

**IMPLEMENTASI MASK R-CNN UNTUK MENDETEKSI AREA DI
UNIVERSITAS LAMPUNG DALAM TATA KELOLA LAHAN**

(Skripsi)

Oleh

**MUHAMMAD BAYU TIAR
2117051071**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2026**

**IMPLEMENTASI MASK R-CNN UNTUK MENDETEKSI AREA DI
UNIVERSITAS LAMPUNG DALAM TATA KELOLA LAHAN**

Oleh

MUHAMMAD BAYU TIAR

Skripsi

**Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA ILMU KOMPUTER**

Pada

**Jurusan Ilmu Komputer
Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2026**

ABSTRAK

IMPLEMENTASI MASK R-CNN UNTUK MENDETEKSI AREA DI UNIVERSITAS LAMPUNG DALAM TATA KELOLA LAHAN

Oleh

MUHAMMAD BAYU TIAR

Implementasi teknologi computer vision berbasis deep learning dalam pemetaan area bangunan menjadi salah satu solusi strategis untuk mendukung tata kelola lahan yang lebih efisien. Penelitian ini mengimplementasikan metode Mask Region-Based Convolutional Neural Network (Mask R-CNN) dengan backbone ResNet-50 Feature Pyramid Network (FPN) untuk mendeteksi, men-segmentasi, dan memperkirakan luas bangunan fakultas di Universitas Lampung menggunakan citra udara drone. Dataset yang digunakan terdiri dari 594 citra hasil frame capturing video drone dengan total 5.369 objek bangunan yang dianotasi dalam format COCO JSON. Data dibagi menjadi 70% training, 20% validation, dan 10% testing. Proses pelatihan dilakukan menggunakan transfer learning dengan optimizer AdamW, mekanisme early stopping, serta evaluasi menggunakan COCOeval. Kinerja segmentasi dievaluasi menggunakan metrik mAP, AP50, AP75, dan IoU, sedangkan estimasi luas bangunan dievaluasi menggunakan MAE, MAPE, dan WAPE setelah dilakukan konversi skala dari piksel persegi (px^2) ke meter persegi (m^2) berdasarkan median rasio per kelas. Hasil pengujian menunjukkan bahwa model mencapai mAP sebesar 82,03%, AP50 sebesar 94,52%, AP75 sebesar 89,50%, serta rata-rata IoU sebesar 88,12%. Pada tahap estimasi luas, diperoleh MAE rata-rata sebesar 8,063 m^2 , MAPE sebesar 4,205%, dan WAPE sebesar 3,579%. Hasil tersebut menunjukkan bahwa Mask R-CNN mampu melakukan segmentasi bangunan secara presisi serta memberikan estimasi luas area yang akurat dan konsisten. Dengan demikian, pendekatan ini berpotensi mendukung pemetaan detail kampus dan perencanaan tata kelola lahan berbasis citra drone secara lebih efektif.

Kata Kunci: Mask R-CNN, segmentasi instance, citra drone, deteksi bangunan, estimasi luas area, deep learning, computer vision, COCOeval, tata kelola lahan.

ABSTRACT

IMPLEMENTATION OF MASK R-CNN TO DETECT AREAS AT THE UNIVERSITY OF LAMPUNG IN LAND GOVERNANCE

By

MUHAMMAD BAYU TIAR

The implementation of deep learning-based computer vision technology for building area mapping has become a strategic solution to support more efficient land management. This study implements the Mask Region-Based Convolutional Neural Network (Mask R-CNN) method with a ResNet-50 Feature Pyramid Network (FPN) backbone to detect, segment, and estimate the area of faculty buildings at Universitas Lampung using drone aerial imagery. The dataset consists of 594 images extracted from drone video frames, containing a total of 5,369 annotated building objects in COCO JSON format. The data were divided into 70% training, 20% validation, and 10% testing sets. The training process employed transfer learning with the AdamW optimizer, an early stopping mechanism, and evaluation using COCOeval. Segmentation performance was evaluated using mAP, AP50, AP75, and IoU metrics, while building area estimation was assessed using MAE, MAPE, and WAPE after applying a scale conversion from square pixels (px²) to square meters (m²) based on the median ratio per class. The experimental results indicate that the model achieved an mAP of 82.03%, AP50 of 94.52%, AP75 of 89.50%, and a mean IoU of 88.12%. In the area estimation stage, the model obtained an average MAE of 8.063 m², a MAPE of 4.205%, and a WAPE of 3.579%. These findings demonstrate that Mask R-CNN is capable of performing precise building segmentation while providing accurate and consistent area estimation. Therefore, this approach has strong potential to support detailed campus mapping and drone-based land management planning more effectively.

Keywords: Mask R-CNN, instance segmentation, drone imagery, building detection, area estimation, deep learning, computer vision, COCOeval, land management.

Judul Skripsi : **IMPLEMENTASI MASK R-CNN
UNTUK MENDETEKSI AREA DI
UNIVERSITAS LAMPUNG DALAM
TATA KELOLA LAHAN**

Nama Mahasiswa : **Muhammad Bayu Tiar**

Nomor Pokok Mahasiswa : **2117051053**

Program Studi : **Ilmu Komputer**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**



MENYETUJUI
Komisi Pembimbing

Dr. rer. nat. Akmal Junaidi, M.Sc.
NIP. 19710129 199702 1 001

Ridho Sholehurrahman, M. Mat
NIP. 232111970128101

MENGETAHUI

Ketua Jurusan Ilmu Komputer

Dwi Sakethi, S.Si., M.Kom.
NIP. 19680611 199802 1 001

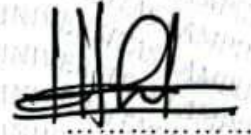
Ketua Program Studi SI Ilmu
Komputer

Tristiyanto, M.I.S., Ph.D.
NIP. 19810414 200501 1 0001

MENGESAHKAN

1. Tim Penguji

Ketua : Dr. rer. nat. Akmal Junaidi, M.Sc.



Sekretaris : Ridho Sholehurrohman, M.Mat



Penguji Utama : Prof. Admi Syarif, Ph.D



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Dr. Eng. Heri Satria, S.Si., M.Si.
NIP. 19711001 200501 1 002



Tanggal Lulus Ujian Skripsi: 28 Januari 2026

PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Muhammad Bayu Tiar

NPM : 2117051071

Dengan ini menyatakan bahwa skripsi saya yang berjudul “Implementasi Mask R-CNN untuk Mendeteksi Area di Universitas Lampung dalam Tata Kelola Lahan” adalah benar hasil karya tulis saya sendiri dan bukan karya orang lain. Seluruh isi dan penulisan yang tertuang dalam skripsi ini telah disusun sesuai dengan kaidah penulisan karya ilmiah yang berlaku di Universitas Lampung. Apabila di kemudian hari terbukti bahwa skripsi ini merupakan hasil penjiplakan atau dibuat oleh pihak lain, maka saya bersedia menerima sanksi sesuai dengan peraturan yang berlaku.

Bandar Lampung, 24 Februari 2026

Penulis,



Muhammad Bayu Tiar

NPM. 2117051071

RIWAYAT HIDUP



Penulis dilahirkan di Natar pada 03 April 2002 sebagai anak pertama dari dua bersaudara dari pasangan Bapak Ivan Bayu Syahputra dan Ibu Sugiarti. Penulis menyelesaikan pendidikan Sekolah Dasar (SD) di SD Negeri 5 Merak Batin Natar pada Tahun 2014, kemudian melanjutkan pendidikan Sekolah Menengah Pertama (SMP) di SMP Yadika Natar dan lulus pada Tahun 2017, serta menyelesaikan pendidikan Sekolah Menengah Atas (SMA) di SMA Negeri 1 Natar pada Tahun 2020. Pada Tahun 2021, penulis terdaftar sebagai mahasiswa Program Studi S1 Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN).

Selama menjadi mahasiswa, penulis melakukan beberapa kegiatan antara lain:

1. Menjadi Anggota Media Informasi Himpunan Mahasiswa Ilmu Komputer (HIMAKOM) Universitas Lampung pada periode 2022–2023.
2. Melaksanakan Kerja Praktik (KP) di Badan Narkotika Nasional Provinsi Lampung pada Tahun 2024.
3. Melaksanakan Kuliah Kerja Nyata (KKN) Periode II Universitas Lampung di Desa Pasir Sakti, Lampung Timur pada Tahun 2024.

MOTTO

“Better late than never but never late is better.”

(Drake)

“Never let success get to your head, and never let failure get to your heart.”

(Drake)

“Stand on your decision, finish what you start, and never quit.”

(Author)

“when writing the story of your life never let anyone else hold the pen.”

(Author)

PERSEMBAHAN

Alhamdulillahirabbil ‘alamin, puji dan syukur penulis panjatkan ke hadirat Allah SWT atas rahmat, karunia, serta pertolongan-Nya yang tiada henti, sehingga penulis dapat melalui setiap proses, tantangan dan rintangan hingga akhirnya dapat menyelesaikan skripsi ini. Segala hikmah, kesehatan dan kekuatan yang diberikan menjadi sumber semangat bagi penulis untuk terus berusaha dan bertahan sampai pada tahap ini.

Skripsi ini saya persembahkan kepada:

Kepada kedua orang tua tercinta Bapak Ivan Bayu Syahputra dan Ibu Sugiarti, serta saudara tercinta, yang senantiasa memberikan doa, kasih sayang, dukungan moral maupun materi serta menjadi motivasi terbesar dalam setiap perjalanan hidup penulis. Segala pengorbanan, kesabaran dan cinta yang tulus menjadi landasan utama bagi penulis untuk terus berusaha dan memberikan yang terbaik. Penulis juga mempersembahkan karya ini kepada seluruh teman-teman Angkatan 2021 Program Studi Ilmu Komputer Universitas Lampung yang telah bersama-sama melalui proses perkuliahan, berbagi pengalaman serta memberikan semangat dan dukungan selama perjalanan akademik.

SANWACANA

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas rahmat, karunia, dan pertolongan-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul *“Implementasi Mask R-CNN untuk Mendeteksi Area di Universitas Lampung dalam Tata Kelola Lahan”*. Skripsi ini dapat terselesaikan berkat doa, dukungan, bimbingan, serta bantuan dari berbagai pihak yang telah memberikan kontribusi secara langsung maupun tidak langsung. Oleh karena itu, pada kesempatan ini penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Kedua orang tua tercinta yang selalu menjadi sumber kekuatan dan inspirasi. Ayah yang tak pernah berhenti mengajarkan arti perjuangan dan kerja keras, serta Ibu yang senantiasa menjadi tempat pulang penuh cinta, doa, dan keteduhan. Tanpa doa dan kasih kalian, langkah ini tidak akan pernah berarti.
2. Bapak Dr. Eng. Heri Satria, S.Si., M.Si. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
3. Bapak Dwi Sakethi, S.Si., M.Kom. selaku Ketua Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
4. Bapak Tristiyato, M.I.S., Ph.D. selaku Ketua Program Studi S1 Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
5. Ibu Yunda Heningtyas, M.Kom. selaku Sekeretaris Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Favorisen R. Lumbanraja S.Kom., M.Si., Ph.D selaku dosen Pembimbing Akademik, yang telah memberikan arahan, motivasi, serta nasihat selama penulis menempuh studi akhir di Jurusan Ilmu Komputer.
7. Bapak Dr. rer. nat. Akmal Junaidi, M.Sc, selaku dosen Pembimbing Utama yang dengan penuh kesabaran telah memberikan arahan, masukan, serta bimbingan hingga skripsi ini dapat terselesaikan dengan baik.

8. Ridho Sholehurrohman, M.Mat, selaku dosen Pembimbing Kedua yang telah memberikan bantuan, masukan, dan saran berharga yang membuat skripsi ini dapat diselesaikan dengan baik.
9. Prof. Admi Syarif, Ph.D selaku dosen Pembahas Utama yang telah meluangkan waktu, memberikan arahan serta memberikan banyak masukan dan motivasi yang bermanfaat dalam kehidupan maupun proses penyusunan skripsi ini.
10. Bapak dan Ibu Dosen Jurusan Ilmu Komputer FMIPA Universitas Lampung, atas segala ilmu, bimbingan serta pengalaman berharga yang telah diberikan kepada penulis selama masa perkuliahan.
11. Ibu Ade Nora Maela selaku staf administrasi Jurusan Ilmu Komputer FMIPA Universitas Lampung yang telah banyak memberikan bantuan dan kemudahan dalam proses administrasi sehingga penulis dapat menyelesaikan skripsi ini dengan lancar.
12. Seluruh staf Jurusan Ilmu Komputer FMIPA Universitas Lampung atas segala bantuan dan kerjasama yang telah memudahkan penulis selama menempuh studi.
13. Adik tersayang, Aura Bayu Agakra, yang selalu menjadi alasan untuk bersemangat sehingga penulis dapat menyelesaikan skripsi ini dengan baik.
14. Rekan-rekan Ilmu Komputer Angkatan 2021, atas persahabatan, perjuangan serta kebersamaan yang telah menjadi bagian berharga dalam perjalanan studi penulis.
15. Untuk teman-teman saya, khususnya khususnya Hanif, Ibe, Nopal El, Elvis yang telah menemani dalam melakukan bimbingan serta memberikan dukungan dalam proses penulisan skripsi dan juga dimasa-masa akhir studi.
16. Seluruh keluarga besar Ucup yang tidak bisa disebutkan satu per satu, atas doa, semangat dan dukungan yang senantiasa mengiringi langkah penulis hingga skripsi ini selesai.

17. Kepada Micelle Yap Aswin atas dukungan, kesabaran, dan ketulusan yang senantiasa diberikan kepada penulis selama proses perkuliahan hingga penyusunan skripsi ini. Kehadiran, doa, serta keyakinan yang diberikan menjadi sumber ketenangan dan kekuatan bagi penulis dalam menghadapi berbagai tantangan akademik, sehingga penelitian ini dapat diselesaikan dengan baik.
18. Kepada Drake, yang melalui karya musik dan lirik-liriknya secara tidak langsung memberikan motivasi dan menemani penulis dalam berbagai sesi pengerjaan skripsi hingga penelitian ini dapat diselesaikan.
19. Teruntuk diri sendiri, yang telah mempertahankan setiap keputusan, menuntaskan setiap proses yang dimulai, dan memilih untuk tidak menyerah dalam menghadapi berbagai rintangan hingga skripsi ini dapat diselesaikan.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat banyak kekurangan, baik dari segi isi maupun penyajian, yang disebabkan oleh keterbatasan pengalaman dan pengetahuan penulis. Oleh karena itu, kritik, saran, dan masukan yang membangun dari para pembaca sangat penulis harapkan, agar dapat menjadi pembelajaran dan bahan perbaikan di masa yang akan datang.

Bandar Lampung, 24 Februari 2026



Muhammad Bayu Tiar

NPM. 2117051071

I. DAFTAR ISI

DAFTAR ISI	i
DAFTAR GAMBAR	iv
DAFTAR TABEL	vi
DAFTAR KODE	vii
I. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
II. TINJAUAN PUSTAKA	5
2.1 Penelitian – Penelitian Terkait	5
2.2 Pengolahan Citra Digital	7
2.3 Video Digital.....	7
2.4 Citra Udara Drone	8
2.5 Skala.....	8
2.6 Deteksi Objek.....	9
2.7 <i>Computer Vision</i>	9
2.8 <i>Artificial Intelligence (AI)</i>	9
2.9 <i>Machine Learning (ML)</i>	10
2.10 <i>Deep Learning</i>	10
2.11 <i>Convolution Neural Network (CNN)</i>	10
2.11.1 <i>Convolution Layer</i>	11
2.11.2 <i>Pooling Layer</i>	11
2.11.3 <i>Fully Connected Layer</i>	12

2.12	<i>Region-Based Convolutional Neural Network (R-CNN)</i>	12
2.13	<i>Region Proposal Network (RPN)</i>	12
2.14	<i>Region of Interest (RoI)</i>	12
2.15	<i>Fully Convolutional Network (FCN)</i>	13
2.16	<i>Mask Region-Based Convolutional Neural Network (Mask R-CNN)</i>	13
2.17	<i>Intersection over Union (IoU)</i>	14
2.18	<i>Average Precision (AP)</i>	16
2.19	<i>Mean Average Precision (mAP)</i>	16
2.20	<i>Common Objects in Context (COCO) JSON</i>	17
2.21	<i>COCOEval</i>	17
2.22	<i>Mean Absolute Error (MAE)</i>	17
2.23	<i>Mean Absolute Percentage Error (MAPE)</i>	18
2.24	<i>Weighted Absolute Percentage Error (WAPE)</i>	18
III.	METODE PENELITIAN	19
3.1	Waktu dan Tempat Penelitian	19
3.2	Perangkat Penelitian	19
3.2.1	Alat Penelitian	19
3.2.2	Bahan Penelitian	20
3.3	Tahapan Penelitian	21
3.3.1	Pengumpulan Data	22
3.3.2	<i>Pre-Processing</i>	22
3.3.3	Pembagian Dataset	23
3.3.4	<i>Training Model Mask R-CNN</i>	23
3.3.5	Evaluasi Metode Mask R-CNN	24
3.3.6	Analisis Hasil	24
IV.	HASIL PENELITIAN DAN PEMBAHASAN	25
4.1	Dataset Penelitian	25
4.2	<i>Pre-processing</i>	27
4.2.1	<i>Frame Capturing</i>	27
4.2.2	<i>Resize</i>	28
4.2.3	Anotasi Data	28

4.3	Pembagian Dataset.....	29
4.4	<i>Training</i> Model	30
4.5	<i>Testing</i> Model	44
4.5.1	Analisis Hasil Deteksi dan Segmentasi.....	44
4.6	Evaluasi Model	47
4.6.1	Evaluasi Hasil Deteksi Objek.....	48
4.6.2	Evaluasi Hasil AP dan mAP	52
4.6.3	Penerapan Faktor Skala ($px^2 \rightarrow m^2$).....	56
4.6.4	Evaluasi Prediksi Luas Area	61
4.7	Analisis Hasil	64
V.	SIMPULAN DAN SARAN	66
5.1	Simpulan	66
5.2	Saran.....	67

DAFTAR PUSTAKA

II. DAFTAR GAMBAR

Gambar 1. Pengolahan Citra Digital (Ahlul Halliwa, 2023)	7
Gambar 2. Hasil Citra Udara Drone (Saihu et al., 2023)	8
Gambar 3. Keterkaitan antara AI, ML, dan Deep Learning	9
Gambar 4. Arsitektur CNN (Tammina, 2019).....	10
Gambar 5. Ilustrasi Convolution Layer (Tammina, 2019)	11
Gambar 6. Ilustrasi Pooling Layer (Tammina, 2019).....	11
Gambar 7. Arsitektur Mask R-CNN (Wulan Dari & Triloka, 2022)	14
Gambar 8. <i>Ground-truth</i> dan <i>Predicted Bounding Box</i> (Rosebrock, 2022).....	14
Gambar 9. Ilustrasi Perhitungan IoU (Rosebrock, 2022)	15
Gambar 10. Flowchart Tahapan Penelitian	21
Gambar 11. Ilustrasi penerbangan drone	26
Gambar 12. Ilustrasi pengambilan drone dari ketinggian	27
Gambar 13. Sampel <i>frame image</i>	28
Gambar 14. Proses anotasi data.....	29
Gambar 15. Ilustrasi pembagian dataset di Roboflow.....	29
Gambar 16. Visualisasi <i>feature map</i> dari <i>backbone</i> ResNet-50.....	32
Gambar 17. Visualisasi <i>multi-scale feature map</i> dari FPN (P2, P3, P4, P5)	33
Gambar 18. Hasil <i>RoIAlign</i> untuk enam RoI pertama.....	34
Gambar 19. Visualisasi keluaran <i>detection head</i>	34
Gambar 20. Grafik <i>training loss</i> dan validasi metrik AP50.....	38
Gambar 21. Grafik <i>training error</i> dan validasi <i>error</i> AP50.....	39
Gambar 22. Grafik <i>training loss</i> dan validasi metrik AP75	40
Gambar 23. Grafik <i>training error</i> dan validasi <i>error</i> AP75.....	41
Gambar 24. Grafik <i>training loss</i> dan validasi metrik mAP	42
Gambar 25. Grafik <i>training error</i> dan validasi <i>error</i> mAP.....	42
Gambar 26. Hasil <i>testing</i> model Mask R-CNN.....	45

Gambar 27. Visualisasi <i>mask ground truth</i> dan <i>mask</i> prediksi	48
Gambar 28. Visualisasi <i>bounding box box prediction</i> dan <i>GT bounding box</i>	49
Gambar 29. Grafik rata-rata nilai IoU per-kelas	51
Gambar 30. Hasil evaluasi AP50 per-kelas	52
Gambar 31. Hasil evaluasi AP75 per-kelas	54
Gambar 32. Hasil evaluasi mAP per-kelas	55
Gambar 33. Hasil segmentasi <i>mask</i>	58
Gambar 34. Distribusi faktor skala ($\text{px}^2 \rightarrow \text{m}^2$) per kelas bangunan.....	59
Gambar 35. Grafik hasil evaluasi kinerja model Mask R-CNN.....	64

III. DAFTAR TABEL

Tabel 1. Penelitian-Penelitian Terkait	5
Tabel 2. Waktu Penelitian	19
Tabel 3. Nilai Hyperparameter Tuning Mask R-CNN	31
Tabel 4. Nilai terbaik performa model Mask R-CNN validasi selama training dan pada pengujian	43
Tabel 5. data luas aktual dan prediksi	55
Tabel 6. Hasil statistik kesalahan prediksi luas berdasarkan metrik MAE, MAPE, dan WAPE.....	55

IV. DAFTAR KODE

Kode Program 1. Struktur utama model Mask R-CNN	35
Kode Program 2. <i>Training loop</i> model Mask R-CNN	36
Kode Program 3. Validasi <i>COCOeval</i> , <i>checkpoint</i> , dan <i>early stopping</i>	37
Kode Program 4. Perhitungan rasio faktor skala ($\text{px}^2 \rightarrow \text{m}^2$).....	51

I. PENDAHULUAN

1.1. Latar Belakang

Universitas Lampung (UNILA) merupakan perguruan tinggi negeri di Provinsi Lampung yang terletak di Kecamatan Rajabasa. UNILA memiliki delapan fakultas yang setiap tahunnya terus melakukan pembangunan gedung maupun pemanfaatan tata kelola lahan. Seiring perkembangan teknologi digital yang semakin pesat, UNILA perlu memanfaatkan teknologi cerdas untuk mendukung pemetaan, analisis rancangan pembangunan, serta perencanaan tata kelola lahan secara lebih efisien. Salah satu pendekatan yang relevan adalah *Computer Vision*, yaitu kemampuan komputer dalam mengenali dan memahami gambar atau video (Ardian, Zakaria, & Susilo, 2016).

Computer Vision merupakan sub-bidang *Artificial Intelligence* (AI) dan *Deep Learning* yang berfokus pada bagaimana komputer dapat melihat dan menafsirkan dunia di sekitarnya (Marpaung, Aulia, & Nabila, 2022). Perkembangan signifikan dalam bidang ini terlihat dari kemajuan pada pengolahan citra digital, di mana gambar dapat diproses untuk mengekstrak informasi penting guna mendukung pengambilan keputusan. Salah satu metode *Deep Learning* yang saat ini banyak digunakan untuk pemrosesan citra adalah Mask R-CNN.

Mask R-CNN (*Mask Region-Based Convolutional Neural Network*) merupakan pengembangan dari Faster R-CNN yang tidak hanya mampu mendeteksi objek, tetapi juga menghasilkan segmentasi *mask* secara piksel demi piksel (Ariya & Lina, 2023). Dengan memanfaatkan *RoIAlign*, Mask R-CNN mampu menjaga keselarasan antara *input* dan *output* sehingga meningkatkan akurasi segmentasi hingga 50% dibandingkan metode sebelumnya (He, Gkioxari, & Girshick, 2017).

Dalam konteks kebutuhan pemetaan dan monitoring di lingkungan kampus, penggunaan *Unmanned Aerial Vehicle* (UAV) atau drone menjadi alternatif yang efektif dibandingkan metode manual maupun citra satelit. Drone dapat menghasilkan citra udara dengan resolusi tinggi secara fleksibel tanpa terhambat kondisi cuaca atau keterbatasan waktu (Saihu, Purba, & Sarkowi, 2023). Integrasi antara teknologi drone dan Mask R-CNN berpotensi memberikan hasil pemetaan yang lebih akurat dan detail, sekaligus mendukung analisis spasial dalam tata kelola lahan.

Beberapa penelitian terdahulu menunjukkan keberhasilan Mask R-CNN dalam meningkatkan akurasi deteksi dan segmentasi objek (He et al., 2017; Chen et al., 2020; Hou & Li, 2024). Namun, penerapannya pada citra drone untuk mendeteksi area kampus, khususnya dalam konteks Universitas Lampung, masih sangat terbatas. Oleh karena itu, penelitian yang berjudul “Implementasi Mask R-CNN untuk Mendeteksi Area di Universitas Lampung” ini berfokus pada implementasi Mask R-CNN untuk mendeteksi area kampus Universitas Lampung dari citra drone. Tujuannya adalah untuk menghasilkan sistem yang tidak hanya mampu mendeteksi dan men-segmentasi bangunan fakultas, tetapi juga memperkirakan luas area bangunan secara kuantitatif, sehingga dapat dimanfaatkan dalam perencanaan pembangunan, tata kelola lahan, dan mendukung inisiatif *Green Metric* kampus.

1.2. Rumusan Masalah

Adapun rumusan masalah dari penelitian ini adalah sebagai berikut:

1. Bagaimana implementasi *Mask R-CNN* untuk mendeteksi area di Universitas Lampung?
2. Bagaimana performansi metode *Mask R-CNN* dalam mendeteksi objek area menggunakan citra udara drone?
3. Bagaimana akurasi model Mask R-CNN dalam memperkirakan luas area bangunan (m^2) dari hasil segmentasi citra drone?

1.3. Batasan Masalah

Adapun batasan masalah dari penelitian ini adalah sebagai berikut:

1. Penelitian dibatasi pada lingkungan Universitas Lampung dan tidak mencakup area di luar kampus.
2. Data citra diperoleh menggunakan drone (UAV) dengan ketinggian tertentu, sehingga kualitas hasil dipengaruhi oleh faktor resolusi kamera, kondisi cuaca, dan sudut pengambilan gambar.
3. Objek yang dideteksi terbatas pada bangunan gedung fakultas (objek statis) menggunakan metode segmentasi instance Mask R-CNN. Objek dinamis seperti kendaraan atau manusia tidak termasuk dalam lingkup penelitian.
4. Penelitian hanya mencakup 3 fakultas (FKIP, FMIPA, dan FK) dengan pertimbangan ketersediaan data citra drone serta keterbatasan waktu penelitian. Oleh karena itu, hasil penelitian belum mewakili seluruh area kampus.
5. Sumber pembanding SDGS UNILA dalam pembangunan berkelanjutan dan *Green Metric* mengenai lingkungan hijau.

1.4. Tujuan

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Mengimplementasikan Mask R-CNN untuk mendeteksi area di lingkungan Universitas Lampung.
2. Mengevaluasi akurasi dan efisiensi Mask R-CNN dalam mendeteksi dan melakukan segmentasi *instance* objek pada citra yang diperoleh dari area di Universitas Lampung.

1.5. Manfaat

Adapun manfaat dari penelitian ini adalah sebagai berikut:

1. Mendukung pengelolaan fasilitas kampus dengan menyediakan informasi spasial berbasis citra drone, sehingga area gedung fakultas dapat terdeteksi dan dipetakan secara lebih akurat.
2. Meningkatkan penerapan teknologi kecerdasan buatan (AI) dalam manajemen lingkungan pendidikan, khususnya melalui deteksi objek dan segmentasi citra untuk keperluan tata kelola lahan dan perencanaan pembangunan kampus.

3. Menjadi acuan bagi penelitian lanjutan yang ingin mengembangkan metode deteksi dan segmentasi berbasis Mask R-CNN maupun model *deep learning* lainnya dalam konteks pengolahan citra udara.
4. Dapat digunakan sebagai bahan perencaann *Green Metric* Universitas Lampung dalam upaya mewujudkan kampus hijau yang berkelanjutan yang berfokus pada kelestarian lingkungan.

II. TINJAUAN PUSTAKA

2.1 Penelitian – Penelitian Terkait

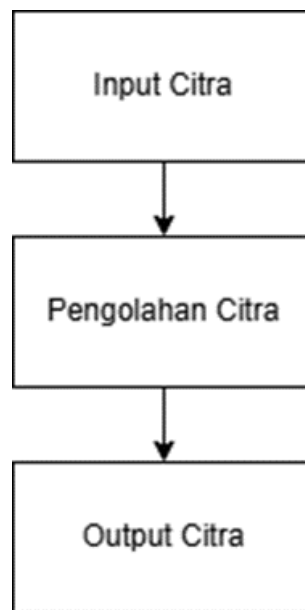
Tabel 1. Penelitian-Penelitian Terkait.

No	Penulis	Judul Paper	Objek	Pokok Bahasan
1.	(He,K., Gkioxari, G., Dollár, P.,& Girshick, R., 2017).	Mask R-CNN	Pendeteksian dan segmentasi <i>instance</i> objek dalam gambar.	Penelitian ini membahas mengenai algoritma Mask R-CNN yang merupakan pengembangan dari algoritma Faster R-CNN. Mask R-CNN memperbaiki kekurangan pada Faster R-CNN dibagian keselarasan segmentasi dengan mengganti metode <i>RoIPooling</i> dengan metode <i>RoIAlign</i> yang mampu meningkatkan akurasi prediksi mask hingga 10%-50%. Mask R-CNN menunjukkan hasil <i>instance segmentation</i> dengan <i>mask AP</i> sebesar 35,7 dan <i>bounding box AP</i> sebesar 38,2.

2.	(Chen, J., Wang, G., Luo, L., Gong, W. & Cheng, Z. 2021).	<i>Building Area Estimation in Drone Aerial Images Based on Mask R-CNN</i>	Bangunan yang teridentifikasi pada citra udara drone	Penelitian ini membahas penggunaan Mask R-CNN untuk memperkirakan luas area bangunan menggunakan citra udara dari drone pada area pedesaan yang rawan bencana. Berdasarkan hasil penelitian ini, Mask R-CNN mencapai <i>F1 Score</i> sebesar 0.875 dan <i>IoU</i> sebesar 0.769.
3.	Hou, T. dan Li, J. (2024)	<i>Application of mask R-CNN for building detection in UAV remote sensing images</i>	bangunan perkotaan dan pedesaan	Penelitian ini membahas penerapan Mask R-CNN untuk meningkatkan akurasi deteksi dan ekstraksi fitur gedung dari citra UAV. Berdasarkan hasil penelitian ini Mask R-CNN menunjukkan kinerja dengan tingkat akurasi klasifikasi mencapai 95% dan akurasi koreksi <i>frame</i> hingga 88,67% serta berhasil mengurangi tingkat kesalahan deteksi dengan <i>False Negative Rate (FNR)</i> serendah 4,5%.

2.2 Pengolahan Citra Digital

Pengolahan citra digital merupakan ilmu yang mempelajari tentang hal-hal yang berkaitan dengan perbaikan kualitas terhadap suatu gambar (meningkatkan kontras, perubahan warna, restorasi citra), transformasi gambar (translasi, rotasi transformasi, skala, geometrik), melakukan pemilihan citra ciri (feature images) yang optimal untuk tujuan analisis, melakukan penyimpanan data yang sebelumnya dilakukan reduksi dan kompresi, transmisi data, dan waktu proses data (Aqdi, 2023). Berikut adalah gambaran sederhana dari proses pengolahan citra digital:



Gambar 1. Pengolahan Citra Digital (Aqdi, 2023).

2.3 Video Digital

Video digital merupakan representasi informasi melalui sekumpulan citra yang digabungkan dan ditampilkan sesuai dengan standar *scanning system*, *frame rate*, dan *frame size* yang digunakan oleh teknologi video tersebut. Video digital tidak hanya digunakan sebagai media penyimpanan atau pengarsipan informasi tetapi juga digunakan sebagai media pertukaran informasi (Wardhani & Madenda, 2016).

2.4 Citra Udara (Drone)

Teknologi drone saat ini berkembang dengan pesat dan telah banyak dimanfaatkan dalam berbagai kegiatan dan bidang keilmuan. Drone merupakan pesawat tanpa awak yang dikendalikan dari jarak jauh oleh *auto pilot*. Penggunaan drone bukan hanya untuk mengambil gambar foto udara dan video, namun juga dapat digunakan dalam proses penginderaan jarak jauh, monitoring tata ruang konstruksi bangunan, dan juga pemetaan lahan (Saihu et al., 2023). Berikut merupakan citra udara yang dihasilkan dengan drone:



Gambar 2. Hasil Citra Udara Drone (Saihu et al., 2023).

2.5 Skala

Konversi luas dari piksel persegi (px^2) ke meter persegi (m^2) dilakukan untuk menerjemahkan hasil segmentasi citra ke ukuran spasial yang sesuai dengan kondisi lapangan. Setiap piksel memiliki representasi ukuran tertentu sesuai resolusi citra, sehingga luas objek dalam piksel perlu dikalibrasi menggunakan faktor skala yang dihitung dari perbandingan antara luas anotasi dan luas aktual. Pendekatan ini banyak digunakan dalam pemetaan dan ekstraksi bangunan berbasis citra, karena hubungan spasial antar-piksel dapat diubah menjadi geometri objek yang dapat diukur secara akurat setelah proses segmentasi (Emelyanov et al., 2024). Adapun perhitungan skala dihitung menggunakan persamaan berikut:

$$Skala = \frac{Luas Aktual (m^2)}{Luas Anotasi (px^2)} \quad (1)$$

2.6 Deteksi Objek

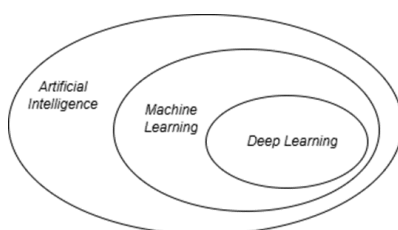
Deteksi objek adalah pendekatan dalam *machine learning* yang berfokus pada identifikasi dan lokalisasi berbagai objek dalam gambar dan video. Setiap objek memiliki fitur atau karakteristik tersendiri yang membantu membedakannya dalam frame foto atau video. Selain mendeteksi fitur, mesin juga melakukan pencocokan fitur atau *features matching* (Marpaung et al., 2022). Deteksi objek merupakan teknologi komputer yang berhubungan dengan citra digital dan pengolahan citra yang digunakan untuk melakukan pendeteksian sebuah objek seperti gedung, manusia, atau benda (Ariya & Lina, 2023).

2.7 Computer Vision

Computer vision merupakan sub-bidang dari *Artificial Intelligence* yang memungkinkan komputer untuk memahami dan menganalisis gambar atau video layaknya cara manusia melihat. *Computer vision* adalah integrasi dari proses akuisisi citra, pengolahan citra, pengenalan dan pembuat keputusan yang digunakan untuk meniru cara kerja sistem visual manusia yang dapat mengenali objek objek tertentu (Putra & Antony, 2018).

2.8 Artificial Intelligence (AI)

Kecerdasan buatan (*Artificial Intelligence* atau AI) merupakan cabang dari ilmu komputer yang berfokus pada pengembangan sistem yang mampu meniru kemampuan berpikir dan bertindak seperti manusia (Nawawi et al., 2024). Melalui AI, komputer dapat belajar dari pengalaman, mengenali pola, mengambil keputusan, serta menyelesaikan tugas-tugas kompleks secara otomatis dan efisien. Teknologi AI mencakup berbagai pendekatan dan subbidang, di antaranya *machine learning* yang memungkinkan sistem belajar dari data tanpa pemrograman eksplisit, serta *deep learning* yang memanfaatkan arsitektur jaringan saraf tiruan berlapis untuk memproses dan mengekstraksi informasi dari data berskala besar (Janiesch et al., 2021).



Gambar 3. Keterkaitan antara AI, ML dan Deep Learning.

2.9 Machine Learning (ML)

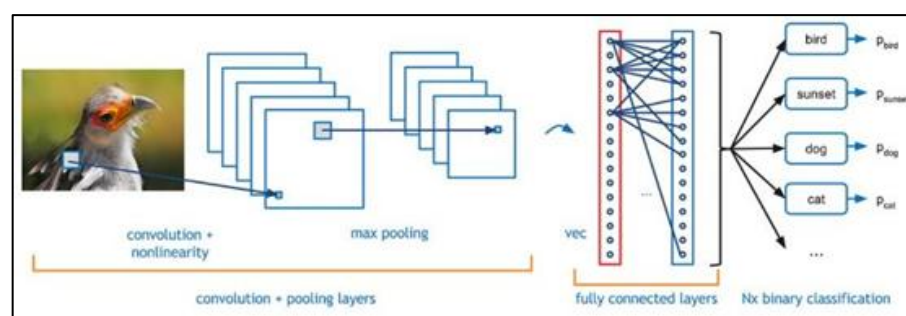
Machine Learning atau pembelajaran mesin, cabang dari AI-kecerdasan buatan, merupakan disiplin ilmu yang mencakup perancangan dan pengembangan algoritma yang digunakan untuk mengembangkan mesin otomatis berdasarkan eksekusi pada algoritma dan kumpulan aturan yang terdefinisi. Oleh karena itu dapat dikategorikan sebagai intruksi yang dijalankan secara otomatis untuk mengubah data menjadi beberapa pola dan diinputkan jauh ke dalam sistem untuk mendeteksi masalah otomatis serta mendapatkan output yang optimal (Raup et al., 2022).

2.10 Deep Learning

Deep Learning mulai banyak digunakan sejak tahun 2006 yang menggunakan mekanisme *deep architecture of learning*. *Learning* atau pembelajaran dalam ini merupakan sebuah prosedur proses estimasi parameter suatu model yang dikembangkan (algoritma) dapat menyelesaikan tugas atau permasalahan tertentu. *Deep Learning* merupakan sebuah teknik berbasis jaringan saraf tiruan telah banyak digunakan sebagai salah satu metode implementasi dalam *Machine Learning* (Diponegoro et al., 2021).

2.11 Convolutional Neural Network (CNN)

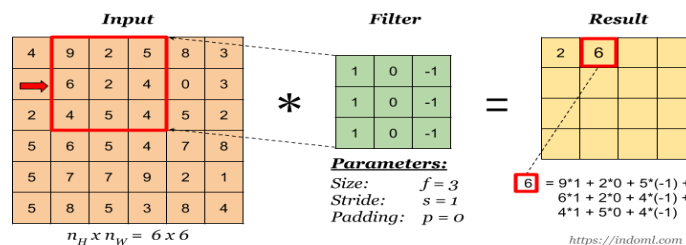
Convolutional Neural Network (CNN) adalah jenis jaringan saraf buatan yang dirancang khusus untuk menganalisis gambar. Salah satu keunggulan CNN adalah kemampuannya memanfaatkan hubungan lokal pada gambar, sehingga bisa menggunakan lebih sedikit parameter karena beberapa di antaranya digunakan bersama. Hal ini membuat CNN lebih efisien dalam penggunaan memori dan komputasi (Tammina, 2019). Metode ini terdiri atas tiga lapisan seperti pada gambar berikut:



Gambar 4. Arsitektur CNN (Tammina, 2019).

2.11.1 Convolution Layer

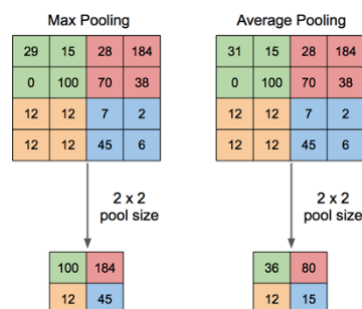
Convolution layer adalah komponen utama dalam jaringan saraf konvolusi (CNN) yang digunakan untuk mengekstraksi fitur penting dari data dalam pemrosesan citra. Proses ini melibatkan penggunaan filter atau kernel yang digeser di seluruh data input untuk menghasilkan peta fitur (*feature map*). Pada setiap posisi, kernel melakukan operasi perkalian elemen dengan data input dan menjumlahkan hasilnya. Operasi ini dikenal sebagai konvolusi, yang berfungsi untuk menangkap pola lokal seperti tepi, tekstur, atau struktur tertentu. Lapisan konvolusi banyak digunakan di berbagai bidang, termasuk pengolahan citra, statistik, dan fisika, karena kemampuannya dalam menyederhanakan data kompleks sambil mempertahankan informasi penting (Tammina, 2019).



Gambar 5. Ilustrasi *Convolution Layer* (Tammina, 2019).

2.11.2 Pooling Layer

pada prinsipnya *pooling layer* terdiri dari filter dengan ukuran tertentu yang akan memperbesar seluruh area *feature map*. Terdapat dua jenis *pooling* yang biasa digunakan yaitu *max pooling* yang merupakan penyatuan maksimum pada setiap *shift* filter akan memilih nilai terbesar dan *average pooling* yang merupakan penyatuan rata-rata akan memilih nilai rata-rata. Tujuan utama dari *pooling layer* adalah untuk memperkecil ukuran *feature map* (*downsampling*) sehingga dapat mempercepat proses komputasi dan mengontrol terjadinya *overfitting* (Prastowo, 2021).



Gambar 6. Ilustrasi *Pooling Layer* (Tammina, 2019)

2.11.3 Fully Connected Layer

Fully Connected Layer (FCL) adalah bagian penting dalam jaringan syaraf tiruan yang berfungsi menghubungkan semua neuron dari satu lapisan ke lapisan berikutnya. Biasanya, lapisan ini berada di bagian akhir jaringan, setelah proses konvolusi dan *pooling* selesai mengambil fitur-fitur penting dari data, misalnya gambar. Peran utama FCL adalah menggabungkan fitur-fitur tersebut menjadi informasi yang dapat digunakan untuk membuat keputusan, seperti mengklasifikasikan atau memprediksi (BJ, Nugroho, Maharani, & Saifullah, 2023).

2.12 Region-Based Convolutional Neural Network (R-CNN)

Region-Based Convolutional Neural Network (R-CNN) merupakan algoritma *Convolutional Neural Network* (CNN) yang dirancang khusus untuk deteksi dan klasifikasi objek dalam gambar. R-CNN bekerja dengan cara memetakan area tertentu dalam gambar yang kemungkinan besar mengandung objek yang ingin dikenali, kemudian melakukan klasifikasi pada area tersebut (Fitriani, 2021).

2.13 Region Proposal Network (RPN)

Region Proposal Network (RPN) merupakan algoritma deteksi objek yang menjadi bagian dari keluarga CNN seperti RCNN, Faster CNN, maupun Mask R-CNN. Dalam prosesnya algoritma ini digunakan untuk pencarian lokasi-lokasi objek dari suatu gambar dengan cepat. Posisi dari objek-objek yang sesuai dengan pencarian akan diberi batas dan ditandai sebagai *Region of Interest* (RoI). Algoritma RPN akan mengambil sub-citra yang sudah ditandai dengan berbagai ukuran objek sebagai input yang diteruskan ke dalam jaringan CNN (BJ, Nugroho, Maharani, & Saifullah, 2023).

2.14 Region of Interest (RoI)

Region of Interest (RoI) adalah teknik yang memungkinkan pemrosesan citra terfokus hanya pada wilayah atau area yang dianggap penting, tanpa harus memproses keseluruhan gambar. Pendekatan ini memberikan beberapa manfaat, seperti mengurangi waktu komputasi, mengoptimalkan penyimpanan data, dan memfasilitasi pengambilan keputusan yang lebih akurat dengan memusatkan perhatian pada bagian citra yang

relevan. RoI bekerja dalam pengkodean area tertentu dari citra digital secara berbeda, sehingga area penting akan memiliki kualitas yang lebih baik dibandingkan area sekitarnya (Pratomo et al., 2020).

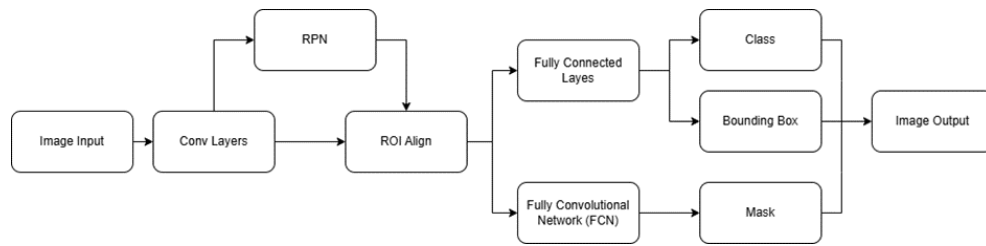
2.15 Fully Convolutional Network (FCN)

Fully Convolutional Neural Network (FCN) merupakan pengembangan dari CNN dimana semua *layer* yang digunakan hanyalah *convolutional layer*. Dibandingkan menggunakan *fully-connected layer* seperti pada CNN dalam prediksi, FCN tetap menggunakan *convolutional layer* untuk mengklasifikasi data ke dalam kelas-kelas. Tujuan utama dari FCN adalah menghasilkan segmentasi semantik, dimana *output* memiliki ukuran yang sama dengan gambar *input* dan menyerupai gambar aslinya. *Output* FCN sendiri melakukan prediksi pada *pixel-pixel* pada citra (Nirmalasari, 2020).

2.16 Mask Region-Based Convolutional Neural Network (Mask R-CNN)

Mask Regional Convolutional Neural Network (Mask R-CNN) adalah jaringan saraf dalam yang bertujuan untuk menyelesaikan masalah segmentasi *instance* dalam pembelajaran mesin atau *computer vision*. *Mask R-CNN* merupakan sebuah *framework* untuk melakukan segmentasi objek yang merupakan perkembangan dari metode sebelumnya, yaitu *Faster R-CNN*. Metode ini menambahkan fitur yang mampu melakukan prediksi terhadap dimana kemungkinan sebuah objek berada dan memberikan mask atau lapisan pada objek. Pada *Mask R-CNN*, diciptakan sebuah layer baru untuk memprediksi *segmentation mask* dari tiap *Region of Interest* (ROI) yang berjalan secara paralel dengan *layer* untuk klasifikasi dan *bounding box regression* yang sudah ada sebelumnya (He et al., 2017).

Mask R-CNN memberikan tiga keluaran untuk setiap objek kandidat, yaitu label kelas, objek pembatas, dan objek topeng (*mask*). Hal tersebut tentunya membuat algoritma deteksi objek dari Mask R-CNN lebih baik karena terdapat variabel tambahan yaitu objek topeng yang tentunya membuat deteksi dari objek yang diinginkan lebih baik. Mask R-CNN juga memiliki beberapa kelebihan diantaranya mudah untuk dilatih, mengungguli semua entri model tunggal yang ada pada setiap tugas, sangat efisien dan hanya menambahkan sedikit tugas diatas *Faster R-CNN*, dan yang terakhir Mask R-CNN mudah di generalisasi ke tugas yang lain (Ayudhawara et al., 2023).

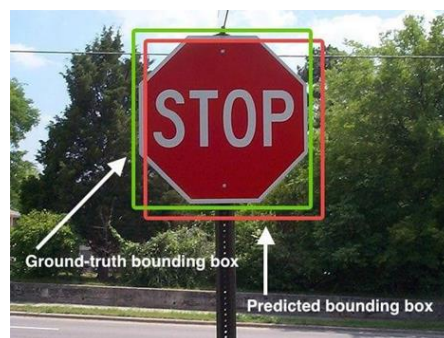


Gambar 7. Arsitektur Mask R-CNN (Wulan Dari & Triloka, 2022).

Berdasarkan gambar 7 dapat dilihat bahwa *instance segmentation* mirip dengan *object detection* di mana masing-masing objek dideteksi satu per satu namun kemudian dikombinasikan dengan *semantic segmentation* sehingga setiap objek dapat diketahui kelas, lokasi serta dapat dibedakan pikselnya. Dalam melakukan deteksi, Mask R-CNN dapat dibagi menjadi tiga bagian, yaitu: *feature extraction network*, *region-proposal network*, dan *instance detection and segmentation networks* (Shi et al., 2019). Mask R-CNN memakai seluruh gambar sebagai input CNN untuk ekstraksi fitur, kemudian *Region Proposal Network* (RPN) mengeluarkan RoI sebagai *region* yang diusulkan dari suatu gambar. Setelah itu Mask R-CNN menggabungkan keluaran ekstraksi fitur gambar dengan *region* yang diusulkan, lalu memperhitungkan kategori, kotak pembatas, dan memberi mask pada objek.

2.17 Intersection over Union (IoU)

Intersection over Union adalah salah satu metrik evaluasi yang umum digunakan dalam bidang *Computer Vision*, khususnya dalam tugas segmentasi gambar dan deteksi objek. IoU dapat digunakan untuk mengukur seberapa baik sebuah model dapat memprediksi lokasi atau segmentasi objek pada gambar dengan membandingkan hasil prediksi model dengan nilai *ground truth* atau nilai yang sebenarnya (Budi et al., 2023).



Gambar 8. *Ground-truth dan Predicted bounding box* (Rosebrock, 2022).

Berdasarkan gambar 8 dapat dilihat bahwa pendeteksi objek telah mendeteksi gambar rambu lalu lintas. Kotak merah adalah predicted *bounding box* sedangkan kotak berwarna hijau adalah ground-truth *bounding box*. Predicted *bounding box* adalah *bounding box* yang diprediksi oleh algoritma. Sedangkan ground-truth *bounding box* adalah *bounding box* yang berasal dari dataset yang sudah diberi label (Darmanto et al., 2023). Dengan kedua *bounding box* tersebut perhitungan IoU dapat dilihat pada persamaan (1) sebagai berikut:

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (2)$$

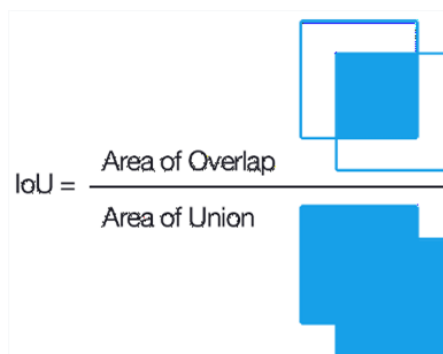
Keterangan:

A = *Bounding box ground thrut.*

B = *Bounding box prediction.*

C = *Bounding box terkecil yang mencakup A dan B.*

Gambar 9 berikut merupakan ilustrasi perhitungan IoU.



Gambar 9. Ilustrasi Perhitungan IoU (Rosebrock, 2022).

Dari gambar 9 dapat diartikan bahwa IoU adalah hasil pembagian dari *area of overlap* dan *area of union*. *Area of overlap* adalah irisan dari *ground- truth bounding box* dan *predicted bounding box*. *Area of union* adalah gabungan dari *ground-truth bounding box* dan *predicted bounding box*. Rentang IoU adalah 0 sampai 1, semakin mendekati 1 semakin baik kesesuaian prediksi dengan *ground truth* (Darmanto et al., 2023).

2.18 Average Precision (AP)

Average Precision adalah metrik yang digunakan untuk mengukur kinerja model deteksi objek berdasarkan *Precision* dan *Recall* (PR). AP membandingkan ketepatan prediksi dari keseluruhan data yang diterima dengan data yang sebenarnya. *Precision* adalah proporsi dari hasil prediksi positif dengan jawaban yang *real positive*. Sedangkan *recall* adalah kebalikan dari *precision*, yaitu proporsi *real positive* yang terprediksi positif (Powers, 2020). Perhitungan *precision* dan *recall* diperoleh dari rumus berikut:

$$\begin{aligned} \textit{precision} &= \frac{\textit{True Positive}}{\textit{True Positive} + \textit{False Positive}} \\ \textit{recall} &= \frac{\textit{True Positive}}{\textit{True Positive} + \textit{False Negative}} \end{aligned} \quad (3)$$

True Positive adalah ketika model berhasil melakukan deteksi objek dan hasil prediksi sesuai dengan label, *False Positive* adalah ketika model berhasil melakukan deteksi objek namun hasil prediksi berbeda dengan label, *False Negative* adalah model tidak dapat mendeteksi objek sama sekali (Powers, 2020). Dari hasil perhitungan kedua nilai tersebut maka akan dilanjutkan untuk menghitung nilai AP (*Average Precision*) dengan cara menghitung rata-rata nilai *precision* dari keseluruhan *recall*. Perhitungan AP dapat dilihat pada rumus berikut:

$$AP = \frac{1}{11} \sum_{\textit{Recall}_i-11} \textit{Precesion}(\textit{Recall}_i) \quad (4)$$

2.19 Mean Average Precision (mAP)

Mean Average Precision (mAP) merupakan metrik utama yang digunakan dalam evaluasi model deteksi objek. mAP dihitung sebagai rata-rata dari nilai AP untuk semua kelas dalam dataset (Pradana, 2024). Perhitungan mAP dilakukan dengan merata-ratakan AP untuk setiap kelas, yang dirumuskan sebagai berikut:

$$mAP = \frac{1}{N} \sum_{i=N}^N AP_i \quad (5)$$

2.20 *Common Objects in Context (COCO) JSON*

COCO JSON adalah format anotasi yang dipakai untuk menyimpan informasi objek dalam bentuk file JSON yang terstruktur, seperti posisi bounding box, kategori objek, dan detail anotasi lainnya. Format ini memudahkan proses pelabelan dan pemrosesan data karena strukturnya jelas serta mudah dibaca oleh model deep learning. Dalam jurnal Gemelli et al. (2024), format COCO JSON dijelaskan sebagai skema anotasi yang umum digunakan dan bahkan dapat diperluas (*augmented*) sesuai kebutuhan. Penggunaan format ini membantu menjaga konsistensi anotasi dan membuat data lebih kompatibel dengan berbagai model deteksi modern (Gemelli et al., 2024).

2.21 *COCOeval*

COCOeval adalah modul evaluasi resmi pada *COCO* API yang digunakan untuk mengukur performa sistem deteksi objek secara lebih ketat dibandingkan standar PASCAL VOC. Evaluasi dilakukan dengan menghitung AP dan mAP pada rentang nilai IoU 0,50–0,95 dengan *interval* 0,05, sehingga penilaian tidak hanya bergantung pada keberhasilan klasifikasi, tetapi juga ketepatan tumpang-tindih prediksi terhadap ground truth. Pendekatan ini membuat *COCOeval* mampu memberikan evaluasi yang lebih komprehensif, stabil, dan representatif terhadap kualitas model deteksi objek (Padilla et al., 2021).

2.22 *Mean Absolute Error (MAE)*

Mean Absolute Error (MAE) adalah sebuah metrik evaluasi yang digunakan untuk menilai seberapa baik model statistik atau model *machine learning* sesuai dengan data empiris atau observasi aktual. MAE dihitung sebagai rata-rata dari nilai absolut kesalahan prediksi (tanpa memperhatikan tanda positif atau negatif). Nilai MAE yang lebih rendah menunjukkan performa model yang lebih baik dalam memodelkan data, karena hal ini mengindikasikan bahwa selisih antara prediksi dan nilai aktual cenderung kecil (Tambunan & Nainggolan, 2023). Nilai MAE dapat ditentukan menggunakan rumus pada persamaan (6) berikut ini:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y' - y_i| \quad (6)$$

Keterangan:

y' : Nilai hasil prediksi
 y : Nilai aktual
 n : Jumlah data

2.23 Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error (MAPE) merupakan indikator statistik yang digunakan untuk menilai tingkat akurasi hasil prediksi pada metode peramalan. Metrik ini menghitung rata-rata persentase kesalahan secara absolut antara nilai prediksi dan nilai aktual (Nurani et al., 2023). Metode MAPE dapat dihitung menggunakan rumus pada persamaan (7) berikut ini.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - Y'_t}{Y_t} \right| \times 100 \quad (7)$$

Keterangan:

Y_t = nilai sebenarnya (aktual) ke - t ($t = 1, \dots, n$),
 Y'_t = nilai hasil perkiraan (prediksi) ke - t ($t = 1, \dots, n$),
 n = banyaknya data yang diuji.

2.24 Weighted Absolute Percentage Error (WAPEE)

Weighted Absolute Percentage Error (WAPEE) merupakan indikator statistik yang digunakan untuk menilai tingkat akurasi hasil prediksi pada metode peramalan. Metrik ini menghitung persentase kesalahan absolut yang ditimbang oleh total nilai aktual; pada data bernilai non-negatif, WAPEE dapat dipandang sebagai MAE yang dinormalisasi oleh rata-rata nilai aktual (Annisa et al., 2015). WAPEE dapat dihitung menggunakan rumus pada persamaan (8) berikut ini.

$$WAPEE = \frac{MAE}{Rata - Rata Luas Aktual} \times 100 \quad (8)$$

III. METODE PENELITIAN

3.1. Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada Desember 2024 sampai November 2025 yang bertempat pada wilayah Bandar Lampung dan Lab. Komputasi Jurusan Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung dengan jadwal penelitian yang akan ditampilkan pada Tabel 2.

Tabel 2. Waktu Penelitian.

NO	Jadwal Kegiatan	BULAN PELAKSANAAN PENELITIAN 2024/2025											
		Des	Jan	Feb	Mar	Apr	Mei	Jun	Jul	Ags	Sep	Okt	Nov
1.	Penulisan Bab 1-3	■	■										
2.	Pengumpulan data			■	■								
3.	Seminar Usul					■							
4.	Pre-Processing Data						■						
5.	Pembagian Data Training dan Data Testing							■					
6.	Training Model Mask R-CNN								■				
7.	Analisis dan evaluasi hasil									■	■		
8.	Penulisan Laporan Bab 4-5											■	
9.	Seminar Hasil Penelitian											■	■

3.2. Perangkat Penelitian

Untuk menunjang penelitian ini dibutuhkan alat dan bahan untuk menyelesaikan beberapa diantaranya:

3.2.1. Alat Penelitian

a. Perangkat Keras

Perangkat keras yang digunakan dalam penelitian ini antara lain:

- Laptop
- Processor : Intel® Core™ i5-12500H 12th Generation

- RAM : 8 GB
- GPU : NVIDIA® GeForce RTX™ 3050 4 GB Graphics
- Storage : 512 GB SSD Hard drive
- System type : 64-bit operating system, x64- based processor

- Drone

- Sensor Kamera : 1/1.3” CMOS
- Resolusi Foto : 12 MP
- Resolusi Video : 4K/30fps HDR
- Baterai : 2453 mAh
- Navigasi : GPS + GLONASS + Galileo
- Waktu Terbang Max : 38 menit

b. Perangkat Lunak

Perangkat lunak yang digunakan dalam penelitian ini antara lain:

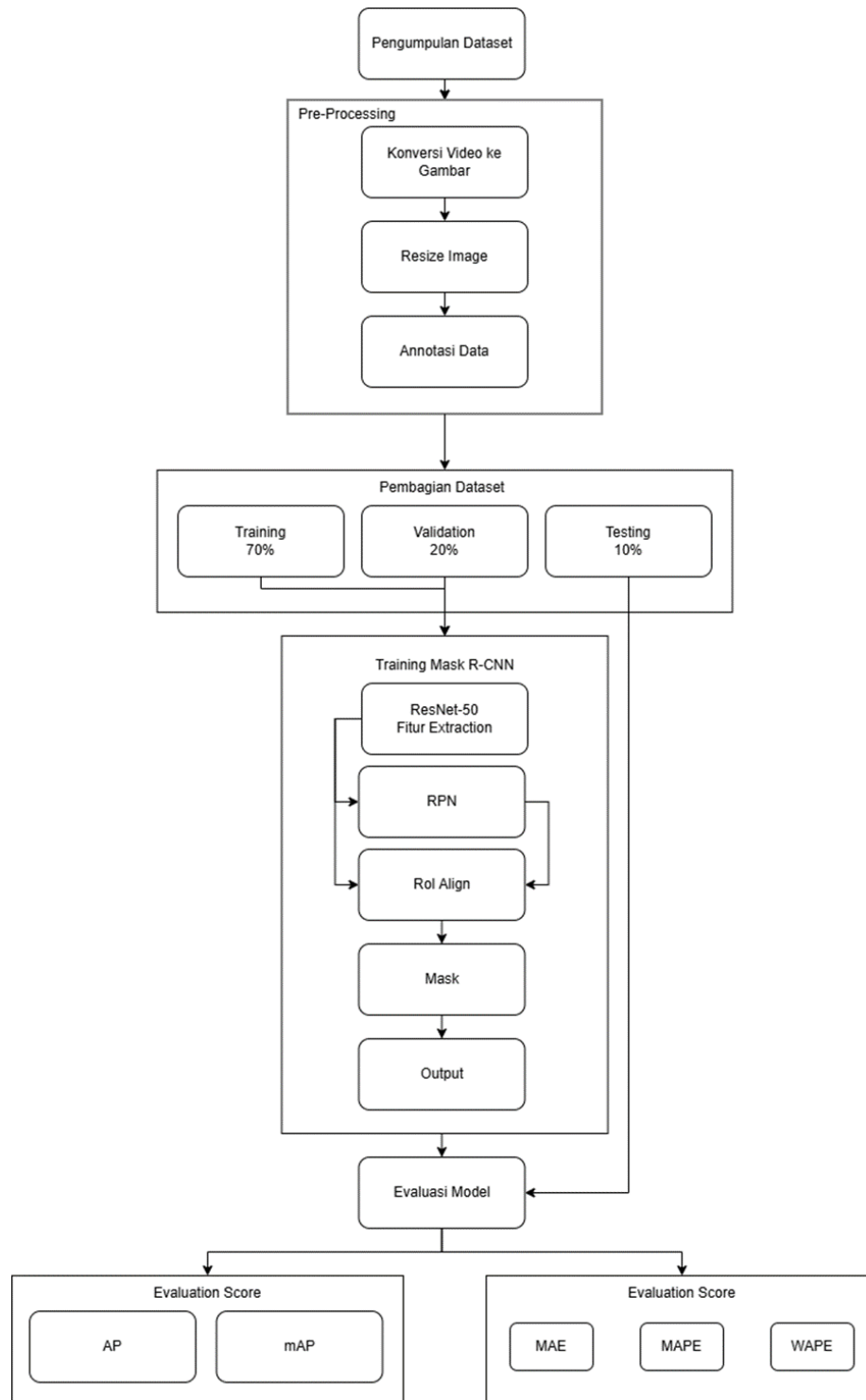
- Sistem operasi Windows 11 Home Single Language 64-bit
- Visual Studio Code versi 1.83.1
- Python versi 3.9.0
- Google Colab
- Google Drive
- Draw.io

3.2.2. Bahan Penelitian

Bahan yang digunakan pada penelitian ini berupa gambar dan video area fakultas Universitas Lampung yang diambil secara langsung menggunakan drone di lingkungan kampus Universitas Lampung.

3.3. Tahapan Penelitian

Desain struktur metodologi penelitian Implementasi Mask R-CNN Untuk Mendeteksi Area di Universitas Lampung Dalam Tata Kelola Lahan yang akan ditunjukkan pada gambar 10 sebagai berikut:



Gambar 10. Flowchart Tahapan Penelitian.

3.3.1. Pengumpulan Data

Dataset yang digunakan pada penelitian ini diperoleh dari hasil rekaman area fakultas di lingkungan Universitas Lampung. Pengambilan data ini dilakukan secara langsung dengan merekam dari udara menggunakan kamera udara (*drone*).

3.3.2. Pre-processing

Pre-processing data dalam penelitian ini merupakan tahap awal dalam pengolahan citra agar siap digunakan untuk melakukan proses *training*. Tahapan ini bertujuan untuk memperbaiki kualitas citra dan mempersiapkan data citra yang digunakan sebagai data input di tahap selanjutnya. *Pre-processing* terdiri dari beberapa tahapan yaitu proses *frame capturing*, *resize* dan anotasi data.

a. *Frame Capturing*

Dalam melakukan deteksi objek diperlukan data citra sebagai data input. Oleh karena itu, pada tahap ini dilakukan *frame capturing* dari video hasil perekaman drone, di mana video diubah menjadi rangkaian citra (*frame-frame*). Hasil *frame capturing* ini kemudian digunakan untuk proses anotasi sehingga citra dapat dipersiapkan sebagai dataset pada tahap selanjutnya.

b. Proses *Resize*

Resize adalah langkah untuk mengubah ukuran citra dan berguna untuk memastikan kualitas serta menghasilkan data yang lebih siap untuk ke proses selanjutnya. *Resize* dilakukan untuk menyesuaikan ukuran citra agar memiliki dimensi yang seragam dan memastikan system hanya memproses area yang relevan. Pada penelitian implementasi Mask R-CNN ini semua citra diubah ukurannya menjadi 854×480 piksel agar keseluruhan citra memiliki rasio yang sama.

c. Proses Anotasi Data

Anotasi adalah pembuatan label pada citra dengan memberikan *bounding box* dan *mask* segmentasi. Untuk melakukan pelatihan model dengan metode *deep learning* berbasis konvolusi, citra perlu dilabeli atau dianotasi. Pada tahap ini, anotasi dilakukan dengan format *COCO JSON* pada setiap karakter gedung di fakultas UNILA.

3.3.3. Pembagian Dataset

Setelah proses anotasi selesai, dataset dibagi menjadi tiga subset menggunakan teknik *simple random sampling* dengan pembagian 70:20:10. Pembagian ini menghasilkan data latih (*training dataset*), data validasi (*validation dataset*), dan data uji (*testing dataset*). Data latih yang berjumlah sekitar 70% digunakan untuk melatih model Mask R-CNN agar dapat mengenali pola dan fitur objek dari citra beranotasi. Data validasi sebesar 20% dimanfaatkan untuk mengevaluasi performa model selama proses pelatihan, sehingga dapat membantu dalam penyesuaian *hyperparameter*, mencegah terjadinya *overfitting*, serta menentukan titik *early stopping* yang optimal. Sementara itu, data uji sebesar 10% berfungsi sebagai data yang sama sekali tidak terlibat dalam proses pelatihan maupun validasi, sehingga mampu memberikan gambaran objektif mengenai performa akhir model terhadap data baru dan mendapatkan evaluasi akhir terhadap performa model.

3.3.4. Training Model Mask R-CNN

Tahapan *training* merupakan proses di mana model mempelajari pola dari dataset yang diberikan untuk mengoptimalkan kemampuannya dalam mengklasifikasi dan mengenali karakter pada citra. Pada tahap training model Mask R-CNN, data latih digunakan untuk melatih model, sementara data validasi digunakan untuk mengevaluasi kinerja model secara berkala selama pelatihan.

Model Mask R-CNN dilatih dengan menggunakan ResNet-50 sebagai *backbone* untuk ekstraksi fitur dari gambar area UNILA. Fitur yang diekstraksi ini kemudian dimasukkan ke dalam *Region Proposal Network* (RPN), yang bertugas mengidentifikasi wilayah potensial yang berisi objek (area UNILA). RPN mengusulkan beberapa wilayah yang diidentifikasi sebagai area fakultas, yang kemudian akan diproses lebih lanjut menggunakan RoI (*Region of Interest*) Align.

Output dari proses ini adalah prediksi koordinat *bounding box* yang menunjukkan lokasi area UNILA dalam gambar, serta menentukan area tersebut merupakan area UNILA atau bukan. Dalam proses *training* dan akurasi prediksi model dipengaruhi oleh *hyperparameter*. *Hyperparameter* ini adalah parameter yang mengendalikan proses pelatihan model dan memiliki pengaruh besar terhadap kinerja serta akurasi prediksi model.

3.3.5. Evaluasi Metode Mask R-CNN

Langkah selanjutnya adalah melakukan evaluasi untuk mengukur performa model dalam mendeteksi area Universitas Lampung serta menilai akurasi hasil segmentasi dan prediksi luas bangunan. Evaluasi segmentasi dilakukan menggunakan metrik IoU, AP, dan mAP, yang masing-masing digunakan untuk mengukur tingkat tumpang tindih antara hasil prediksi dengan *ground truth* serta ketepatan klasifikasi objek pada berbagai ambang batas IoU. Selain itu, untuk menilai kemampuan model dalam memperkirakan luas area hasil segmentasi, digunakan metrik MAE, MAPE, dan WAPE. MAE digunakan untuk menghitung rata-rata selisih absolut antara luas prediksi dan luas aktual dalam satuan meter persegi, MAPE digunakan untuk mengukur kesalahan relatif dalam bentuk persentase terhadap nilai aktual, sedangkan WAPE menunjukkan proporsi kesalahan absolut terhadap total rata-rata luas sebenarnya. Kombinasi metrik tersebut memberikan gambaran yang komprehensif mengenai kinerja model Mask R-CNN, baik dari sisi ketepatan segmentasi spasial maupun akurasi estimasi luas area bangunan yang dihasilkan.

3.3.6. Analisis Hasil

Pada fase ini, kesimpulan dirumuskan berdasarkan hasil evaluasi kinerja model Mask R-CNN dalam melakukan segmentasi dan memprediksi luas area dibandingkan dengan luas aktual. Tujuan utamanya adalah menentukan sejauh mana model mampu memberikan hasil terbaik dalam hal ketepatan segmentasi serta akurasi prediksi luas.

V. SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan penelitian yang telah dilakukan mengenai implementasi Mask R-CNN dalam analisis rancangan pembangunan gedung, perencanaan tata kelola lahan, dan pemetaan area pada fakultas di Universitas Lampung melalui deteksi segmentasi objek serta prediksi luas bangunan dari citra drone, dapat disimpulkan hal-hal berikut:

1. Model Mask R-CNN dengan backbone ResNet-50 FPN digunakan untuk mendeteksi dan men-segmentasi bangunan pada citra drone di lingkungan Universitas Lampung. Dataset dikonversi dalam format COCO, yang berisi anotasi *bounding box* dan *mask polygon* untuk setiap objek bangunan. Melalui algoritma *Region Proposal Network* (RPN), model dapat mengenali area bangunan secara presisi hingga tingkat piksel. Hasil segmentasi kemudian dikonversi menjadi luas area (m^2) yang dihitung dari rasio antara luas anotasi dan luas aktual di lapangan. Kinerja model dievaluasi menggunakan metrik mAP, AP50, dan AP75 untuk menilai akurasi deteksi dan segmentasi, serta MAE, MAPE, dan WAPE untuk mengukur ketepatan estimasi luas area.
2. Berdasarkan hasil pengujian terbaru, model Mask R-CNN dengan *backbone* ResNet-50 FPN menunjukkan kinerja yang baik dalam mendeteksi dan men-segmentasi bangunan pada citra drone di lingkungan Universitas Lampung. Model mencapai mAP sebesar 82,03%, AP50 sebesar 94,52%, dan AP75 sebesar 89,50%, dengan IoU_mean sebesar 88,12%. Nilai-nilai tersebut menunjukkan segmentasi yang akurat dan konsisten pada sebagian besar kelas bangunan; AP50 yang tinggi menegaskan kemampuan model mengenali objek pada variasi bentuk dan ukuran, sementara AP75 dan mAP merefleksikan ketelitian segmentasi pada tingkat tumpang-tindih yang lebih ketat. Pada tahap estimasi luas area, model memperoleh MAE rata-rata $8,063m^2$, MAPE 4,205%, dan WAPE 3,579%, yang berarti kesalahan prediksi luas relatif kecil

terhadap nilai aktual. Secara keseluruhan, metode Mask R-CNN mampu mendeteksi serta men-segmentasi objek bangunan dengan akurat, sekaligus memberikan estimasi luas area yang stabil dan presisi dari citra drone.

Secara keseluruhan, hasil penelitian menunjukkan bahwa metode Mask R-CNN dengan backbone ResNet-50 FPN mampu bekerja dengan baik dalam mendeteksi dan men-segmentasi objek bangunan pada citra drone di lingkungan Universitas Lampung. Model ini tidak hanya mampu menghasilkan segmentasi yang akurat, tetapi juga dapat memberikan estimasi luas area bangunan dengan tingkat kesalahan yang relatif kecil. Kinerja tersebut memperlihatkan bahwa pendekatan berbasis *deep learning* dapat dimanfaatkan secara efektif untuk mendukung proses pemetaan dan analisis tata ruang kampus. Meskipun masih terdapat variasi akurasi antar kelas yang dipengaruhi oleh kondisi citra dan objek bangunan yang terhalang pepohonan, hasil ini menegaskan potensi penggunaan Mask R-CNN sebagai metode yang andal untuk identifikasi serta perhitungan luas objek bangunan dari citra drone.

5.2 Saran

Berikut adalah beberapa saran yang dapat dipertimbangkan untuk pengembangan penelitian di masa mendatang, khususnya dalam penerapan Mask R-CNN untuk deteksi, segmentasi, dan prediksi luas area bangunan dari citra drone:

1. Memperluas dataset dengan jumlah citra yang lebih banyak dan variasi kondisi yang lebih beragam, baik dari segi pencahayaan, ketinggian, sudut pengambilan drone, maupun keragaman bentuk dan ukuran bangunan. Hal ini bertujuan untuk meningkatkan kemampuan generalisasi model agar tetap stabil pada berbagai kondisi lapangan.
2. Mengintegrasikan model ke dalam sistem aplikasi praktis, misalnya berbasis *real-time* pada perangkat *mobile* atau terhubung dengan Sistem Informasi Geografis (SIG). Hal ini akan memperluas pemanfaatan hasil penelitian dalam mendukung pemetaan area kampus serta tata kelola lahan yang lebih efektif.

3. Penelitian selanjutnya dapat mencoba menambahkan variasi data melalui teknik augmentasi citra serta melakukan pengolahan gambar yang lebih baik, seperti pengurangan *noise*, agar model lebih stabil dan akurat pada berbagai kondisi citra drone.
4. Menggunakan drone dengan kestabilan terbang yang lebih baik serta kemampuan pengambilan citra beresolusi tinggi, agar hasil perekaman memiliki detail yang lebih tajam. Selain itu, penggunaan drone dengan durasi terbang lebih lama memungkinkan pengambilan data pada area yang lebih luas dan beragam sudut pandang, sehingga dapat memperkaya variasi dataset dan meningkatkan akurasi hasil segmentasi.
5. Membandingkan Mask R-CNN dengan model *deep learning* terbaru, seperti YOLOv8-Seg, *Segment Anything Model* (SAM), atau kombinasi metode *hybrid*. Perbandingan ini dapat memberikan gambaran komprehensif mengenai model yang paling optimal untuk tugas segmentasi dan estimasi luas area pada konteks pemetaan kampus.

DAFTAR PUSTAKA

- Aqdi, A. H. H. (2023). *Sistem pendeteksi dan penghitung polen hidup dan mati pada tanaman kelapa sawit menggunakan algoritma YOLO V5 berbasis artificial intelligence* [Skripsi, Universitas Lampung]. Universitas Lampung, Bandar Lampung.
- Ardian, R. B., Zakaria, A., & Susilo, G. E. (2016). Study system drainase di Fakultas Teknik Universitas Lampung. *Jurnal Rekayasa Sipil dan Desain*, 4(4), 503–512.
- Ariya, C., & Lina, L. (2023). Perancangan deteksi objek pada rak toko menggunakan metode Mask R-CNN. *Simtek: Jurnal Sistem Informasi dan Teknik Komputer*, 8(2), 295–299. <https://doi.org/10.51876/simtek.v8i2.213>
- Ayudhawara, A., & Satrio, B. (2023). Deteksi ketersediaan tempat parkir menggunakan Mask R-CNN (Studi kasus: Fakultas Teknologi Industri Universitas Islam Sultan Agung). *Seminar Riset Mahasiswa-Computer & Electrical (SERIMA-CE)*, 1, 245–251.
- Bagaskara, R., & Khairunnisa Rizkita, A. (2022). Pendeteksian jumlah bangunan berbasis citra menggunakan metode deep learning. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 6(1), 94–100.
- BJ, B. J., Nugroho, K. T., & Saifullah, D. F. N. M. (2023). Pengaruh jumlah epoch dan step per epoch terhadap performa Mask-RCNN pada deteksi objek tanda tangan. *Journal of Electrical, Electronic, Mechanical, Informatic and Social Applied Science*, 2(1), 7–16. <https://doi.org/10.58991/eemisas.v2i1.20>
- Budi, R., Harianto, R. A., & Setyati, E. (2023). Segmentasi citra area tumpukan sampah dengan memanfaatkan Mask R-CNN. *Journal of Intelligent System and Computation*, 5(1), 58–64. <https://doi.org/10.52985/insyst.v5i1.305>
- Chen, J., Wang, G., & Cheng, Z. (2021). Building area estimation in drone aerial images based on Mask R-CNN. *IEEE Geoscience and Remote Sensing Letters*, 18(5), 891–894. <https://doi.org/10.1109/LGRS.2020.2988326>
- Darmanto, B. F., Setianingsih, C., & Saputra, R. E. (2023). Deteksi pelanggaran pada bahu jalan tol menggunakan algoritma Cascade R-CNN. *E-Proceeding of Engineering*, 10(3), 4488–4497.
- Dewi, T. N. (2024). *Deteksi dan klasifikasi rambu larangan lalu lintas menggunakan metode YOLO V5s di Kota Bandar Lampung* [Skripsi, Universitas Lampung]. Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung.
- Diponegoro, M. H., Kusumawardani, S. S., & Hidayah, I. (2021). Tinjauan pustaka sistematis: Implementasi metode deep learning pada prediksi kinerja murid. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 10(2), 131–138. <https://doi.org/10.22146/jnteti.v10i2.1417>

- Emelyanov, S., Biryukov, A., Sorokin, A., & Emelyanova, E. (2024). *Pixels relationship analysis for extracting building footprints*. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-3, 141–148. <https://doi.org/10.5194/isprs-archives-XLVIII-3-2024-141-2024>
- Gemelli, A., Marinai, S., Pisaneschi, L., & Santoni, F. (2024). *Datasets and annotations for layout analysis of scientific articles*. International Journal on Document Analysis and Recognition, 27, 683–705. <https://doi.org/10.1007/s10032-024-00461-2>
- Grandini, M., Bagli, E., & Visani, G. (2020). Metrics for multi-class classification: An overview. *arXiv preprint*. <https://arxiv.org/abs/2008.05756>
- Harun, A. (2020). Implementasi deep learning menggunakan metode You Only Look Once untuk mendeteksi rokok. *Jurnal Ekonomi*, 18(1), 41–49.
- He, K., Gkioxari, G., & Girshick, R. (2017). Mask R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2980–2988). IEEE. <https://doi.org/10.1109/ICCV.2017.322>
- Hou, T., & Li, J. (2024). Application of Mask R-CNN for building detection in UAV remote sensing images. *Heliyon*, 10(5), e38141. <https://doi.org/10.1016/j.heliyon.2024.e38141>
- Ilahiyah, S., & Nilogiri, A. (2018). Implementasi deep learning pada identifikasi jenis tumbuhan berdasarkan citra daun menggunakan convolutional neural network. *JUSTINDO: Jurnal Sistem dan Teknologi Informasi Indonesia*, 3(1), 49–56.
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31, 685–695. <https://doi.org/10.1007/s12525-021-00475-2>
- JIFTI, M. S. M. (2023). Perancangan UFO drone sebagai media pengindraan citra udara di kampus AKN Pacitan. *Jurnal Ilmiah Teknologi Informasi dan Robotika*, 2(1). <https://doi.org/10.33005/jifti.v2i1.116>
- Kassim, Y. M., Byrne, M. E., & Palaniappan, K. (2020). Small object bird detection in infrared drone videos using Mask R-CNN deep learning. In *IS&T International Symposium on Electronic Imaging Science and Technology*. Society for Imaging Science and Technology. <https://doi.org/10.2352/ISSN.2470-1173.2020.8.IMAWM-085>
- Li, Y., Xu, W., & Li, X. (2021). A novel framework based on Mask R-CNN and histogram thresholding for scalable segmentation of new and old rural buildings. *Remote Sensing*, 13(6), 1070. <https://doi.org/10.3390/rs13061070>
- Marpaung, F., Aulia, F., & Nabila, R. C. (2022). Computer vision dan pengolahan citra digital.
- Nawawi, H. E., Rahmadewi, R., & Nurpulaela, L. (2024). Pendeteksian wajah hasil dari manipulasi AI pada citra digital. *Jurnal Ilmiah Wahana Pendidikan*, 10(2), 402–409.
- Nirmalasari, D. A. (2020). *Pengenalan teks tulisan tangan menggunakan fully convolutional neural network (FCN)* [Skripsi, Institut Teknologi Sepuluh Nopember]. Departemen Teknik Informatika, ITS, Surabaya.

- Padilla, R., Netto, S. L., & da Silva, E. A. B. (2021). *A comparative analysis of object detection metrics with a companion open-source toolkit*. arXiv. <https://doi.org/10.48550/arXiv.2006.07777>
- Powers, D. M. W. (2020). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint*. <https://arxiv.org/abs/2010.16061>
- Prastowo, E. Y. (2021). Pengenalan jenis kayu berdasarkan citra makroskopik menggunakan metode convolutional neural network. *Jurnal Teknik Informatika dan Sistem Informasi*, 7(2). <https://doi.org/10.28932/jutisi.v7i2.3706>
- Pratiwi, H. A., Cahyanti, M., & Lamsani, M. (2021). Implementasi deep learning flower scanner menggunakan metode convolutional neural network. *Sebatik*, 25(1), 124–130. <https://doi.org/10.46984/sebatik.v25i1.1297>
- Pratomo, A. H., Kaswidjanti, W., & Mu'arifah, S. (2020). Implementasi algoritma region of interest (ROI) untuk meningkatkan performa algoritma deteksi dan klasifikasi kendaraan. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 7(2), 155–162. <https://doi.org/10.25126/jtiik.202071718>
- Putra, R. R., & Antony, F. (2018). Sistem computer vision pengenalan pola angka dan operator matematika pada permainan kartu angka berbasis jaringan syaraf tiruan perceptron. *Jurnal Ilmiah Informatika Global*, 9(1). <https://doi.org/10.36982/jiig.v9i1.441>
- Rampersad, H. (2020). Developing. In *Total Performance Scorecard* (pp. 159–183). Routledge. <https://doi.org/10.4324/9780080519340-12>
- Raup, A., Ridwan, W., & Zaqiah, Q. Y. (2022). Deep learning dan penerapannya dalam pembelajaran. *Jiip: Jurnal Ilmiah Ilmu Pendidikan*, 5(9), 3258–3267. <https://doi.org/10.54371/jiip.v5i9.805>
- Rosebrock, A. (2016, November 7). Intersection over union (IoU) for object detection. *PyImageSearch*. <https://www.pyimagesearch.com/2016/11/07/intersection-overunion-iou-for-object-detection/>
- Saihu, E. U. K., Purba, A., & Sarkowi, M. (2023). Pemanfaatan teknologi drone guna pemetaan kesesuaian ruang untuk perencanaan lingkungan perumahan di perkotaan. *Jurnal Profesi Insinyur Universitas Lampung*, 4(1), 13–18. <https://doi.org/10.23960/jpi.v4n1.92>
- Shi, W., van de Zedde, R., & Kootstra, G. (2019). Plant-part segmentation using deep learning and multi-view vision. *Biosystems Engineering*, 187, 81–95. <https://doi.org/10.1016/j.biosystemseng.2019.08.014>
- Tamina, S. (2019). Transfer learning using VGG-16 with deep convolutional neural network for classifying images. *International Journal of Scientific and Research Publications (IJSRP)*, 9(10), 9420. <https://doi.org/10.29322/ijsrp.9.10.2019.p9420>
- Wardhani, I. P., & Madenda, S. (2016). Algoritma ekstraksi video frame berdasarkan analisis histogram warna HCL. *Jurnal Ilmiah KOMPUTASI*, 15(1), 77–84.

- Wang, W., Shi, Y., & Liu, F. (2023). Traditional village building extraction based on improved Mask R-CNN: A case study of Beijing, China. *Remote Sensing*, *15*(10), 2616. <https://doi.org/10.3390/rs15102616>
- Wang, Y., Li, S., & Cai, H. (2022). Improved Mask R-CNN for rural building roof type recognition from UAV high-resolution images: A case study in Hunan Province, China. *Remote Sensing*, *14*(2), 265. <https://doi.org/10.3390/rs14020265>
- Wulan Dari, S., & Triloka, J. (2022). Kajian algoritme Mask R-CNN dan You Only Look Once (YOLO) untuk deteksi penyakit kulit akibat infeksi jamur. *Seminar Nasional Hasil Penelitian dan Pengabdian Masyarakat*, 132–138.
- Zhong, D., He, L., & Lin, Y. (2023). An improved Mask R-CNN: Extraction of door and window instances on village building façade images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *X-1/W1*, 295–301. <https://doi.org/10.5194/isprs-annals-X-1-W1-2023-295-2023>