RUANG: JURNAL ARSITEKTUR
Vol. 20. No. 1, March 2026, pp 69-85
DOI: 10.22487/ruang.v20i1.333
e ISSN 3031-4371

Penerapan GeoAI Berbasis Mask R-CNN untuk Deteksi Kendaraan pada Citra Orthophoto Kawasan Perkotaan

Dita Septyana[a,1], Budi Andresi [b,2], Nadine Sandra Agustina[c,3]

[a] **Program Studi Sains Informasi Geografi, Fakultas Sains, Teknologi Teknik dan Matematika, Universitas Mahakarya Asia, Palu, Indonesia**
[b] **Program Studi S1 Perencanaan Wilayah dan Kota, Jurusan Arsitektur, Fakultas Teknik, Universitas Tadulako, Palu, Indonesia**
[c] **Prodi S1 Agroteknologi, Jurusan Budi Daya Pertanian, Fakultas Pertanian, Universitas Tadulako, Palu, Indonesia**

[1] **dita.septyanageo11@gmail.com;** [2]**budiandresy@gmail.com*;** [3]**ninakabi88@gmail.com**
***dita.septyanageo11@gmail.com**

Submitted: February 12, 2026 | Revised: March 03, 2026 | Accepted: March 04, 2026

ABTRACT

GeoAI technology, which integrates artificial intelligence with spatial analysis, offers a novel approach to extracting urban object information from high-resolution imagery. This study applies Mask R-CNN with a ResNet-50 backbone architecture to detect vehicle objects in orthophoto imagery derived from the processing of 100 UAV photographs over an urban area in Switzerland. A total of 80 vehicle objects were annotated and partitioned into training (70%), validation (15%), and testing (15%) datasets. Model evaluation was conducted using a multi-threshold Intersection over Union (IoU) approach at values of ≥0.5, ≥0.75, and ≥0.95, and analyzed through a confusion matrix alongside Precision, Recall, F1-score, and Mean Average Precision (mAP) metrics. The results demonstrate that the model achieved Precision and Recall scores of 1.00 at IoU ≥0.5; however, performance declined at stricter thresholds, with an aggregate mAP of 0.56, indicating moderate overall performance. These findings suggest that the model is effective for macro-spatial analytical needs such as vehicle count estimation and distribution mapping, yet remains insufficiently stable for applications requiring high geometric precision. Conceptually, this study underscores the importance of multi-threshold evaluation in the application of deep learning for urban spatial analysis, while demonstrating the potential of GeoAI integration in data-driven urban planning.

Keywords: GeoAI, Deep Learning, Object Detection

This is an Open-Access article distributed under the CC-BY-SA license

PENDAHULUAN

Perkembangan kawasan perkotaan yang berlangsung secara dinamis menyebabkan perubahan struktur ruang, intensitas pemanfaatan lahan, serta peningkatan kepadatan aktivitas manusia. Kompleksitas tersebut menuntut pendekatan analisis spasial yang mampu memetakan dan memantau elemen-elemen fisik kota secara cepat, akurat, dan berkelanjutan[1]. Dalam konteks perencanaan wilayah dan kota, ketersediaan data spasial yang presisi menjadi landasan penting dalam pengambilan keputusan, baik untuk pengelolaan transportasi, pengendalian pemanfaatan ruang, maupun evaluasi dinamika kawasan perkotaan [3], [5].

Perkembangan teknologi penginderaan jauh dan kecerdasan buatan telah membuka paradigma baru dalam pemetaan otomatis objek-objek perkotaan. Pendekatan GeoAI (Geospatial Artificial Intelligence) mengintegrasikan analisis spasial dengan algoritma kecerdasan buatan untuk mengolah data geospasial dalam skala besar secara sistematis dan adaptif [2], [6]. GeoAI tidak hanya mempercepat proses ekstraksi informasi spasial, tetapi juga mampu

menangkap pola heterogenitas ruang yang sulit diidentifikasi melalui pendekatan konvensional [6], [16].

Dalam kerangka tersebut, teknik Deep Learning, khususnya Convolutional Neural Networks (CNN), telah banyak digunakan dalam deteksi dan segmentasi objek pada citra satelit maupun citra udara resolusi tinggi [7], [14], [15]. Studi Hoeser dan Kuenzer [7] menunjukkan bahwa pendekatan deep learning secara signifikan meningkatkan akurasi deteksi objek pada data penginderaan jauh dibandingkan metode berbasis fitur manual. Demikian pula, Zhu et al. [14] menegaskan bahwa segmentasi berbasis CNN memungkinkan klasifikasi piksel yang lebih konsisten pada lingkungan perkotaan yang kompleks.

Sejumlah penelitian terdahulu telah mengembangkan model deep learning untuk mendeteksi objek perkotaan seperti bangunan, jaringan jalan, dan kendaraan [3], [4], [11]. Mask R-CNN secara khusus dikenal efektif dalam menghasilkan segmentasi berbasis piksel yang lebih detail dibandingkan metode deteksi berbasis bounding box konvensional [7][2] . Namun demikian, sebagian besar penelitian tersebut berfokus pada capaian akurasi umum dengan threshold evaluasi Intersection over Union (IoU) $\geq$ 0.5, yang secara metodologis masih memberikan toleransi spasial yang relatif longgar. Evaluasi semacam ini belum sepenuhnya merepresentasikan kebutuhan ketelitian geometris dalam praktik perencanaan wilayah dan kota.

Dalam konteks tata ruang, ketelitian spasial memiliki implikasi langsung terhadap analisis kapasitas ruang parkir, estimasi kepadatan kendaraan, identifikasi pola distribusi aktivitas, serta evaluasi intensitas pemanfaatan lahan. Beberapa kajian mutakhir dalam urban analytics juga menekankan pentingnya presisi spasial dalam mendukung kebijakan berbasis data (data-driven urban planning) dan smart city governance[3], [4]. Oleh karena itu, diperlukan evaluasi model deteksi objek yang tidak hanya mempertimbangkan akurasi klasifikasi, tetapi juga menguji sensitivitas model terhadap peningkatan standar presisi spasial.

Berdasarkan celah penelitian tersebut, studi ini bertujuan untuk menerapkan pendekatan GeoAI berbasis Mask R-CNN dengan arsitektur ResNet-50 dalam mendeteksi objek kendaraan pada kawasan perkotaan serta mengevaluasi performa model pada variasi threshold IoU ($\geq$0.5, $\geq$0.75, dan $\geq$0.95). Evaluasi multi-threshold ini dilakukan untuk mengidentifikasi batasan dan potensi penerapan model dalam konteks kebutuhan analisis spasial yang berbeda. [1]

Kontribusi ilmiah penelitian ini terletak pada dua aspek utama. Pertama, penelitian ini menempatkan deteksi objek berbasis deep learning dalam kerangka analisis spasial perkotaan, sehingga hasil model tidak hanya dipahami sebagai capaian teknis klasifikasi citra, tetapi sebagai instrumen pendukung analisis struktur ruang dan dinamika kawasan berbasis data geospasial [2], [6]. Kedua, penelitian ini melakukan evaluasi komprehensif terhadap sensitivitas model pada berbagai tingkat threshold IoU untuk mengidentifikasi trade-off antara akurasi deteksi dan ketelitian spasial, suatu aspek yang masih jarang dibahas secara sistematis dalam studi sebelumnya [7], [16].

Temuan mengenai penurunan performa signifikan pada threshold IoU tinggi memberikan refleksi metodologis bahwa penerapan deep learning dalam studi perkotaan perlu mempertimbangkan skala analisis, kepadatan objek, serta kebutuhan presisi spasial secara kontekstual. Dengan demikian, penelitian ini tidak hanya menawarkan pendekatan teknis GeoAI, tetapi juga memberikan kontribusi konseptual terhadap pemanfaatan kecerdasan buatan dalam praktik perencanaan wilayah berbasis data spasial.

METODE

GeoAI (Geospatial Artificial Intelligence)

GeoAI merupakan pendekatan yang mengintegrasikan kecerdasan buatan (AI) dengan data geospasial untuk menganalisis dan memproses informasi yang berhubungan dengan ruang dan lokasi. Teknologi ini menggabungkan algoritma pembelajaran mesin dan deep learning untuk mengekstraksi wawasan dari data geospasial yang besar dan kompleks, seperti citra satelit, data sensor, dan peta digital. GeoAI tidak hanya mempercepat analisis data spasial, tetapi juga mampu mengidentifikasi pola atau hubungan yang mungkin tidak terdeteksi oleh metode konvensional. Menurut Mai et al. (2025) dalam International Journal of Applied Earth Observation and

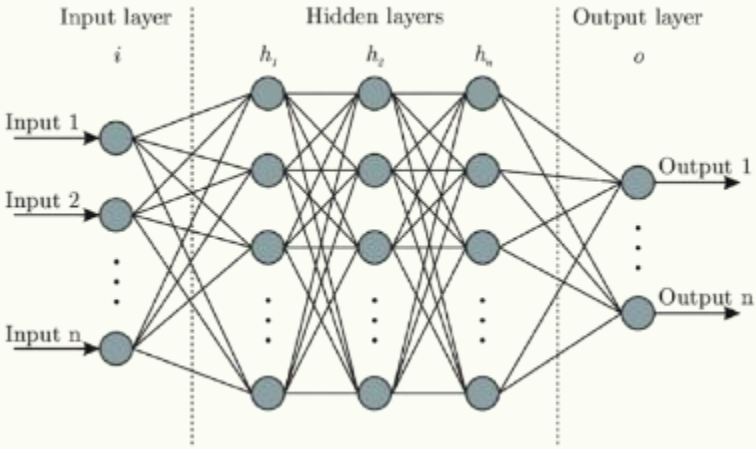
Geoinformation, metode ini menawarkan potensi besar dalam berbagai bidang, mulai dari pemantauan lingkungan, perencanaan kota, hingga prediksi perubahan iklim[5]. Dengan penerapan teknik pembelajaran mesin dan representasi spasial yang lebih canggih, GeoAI dapat menangani heterogenitas data dan kompleksitas spasial, membuatnya menjadi metode yang penting dalam pengambilan keputusan berbasis lokasi dan ruang.

Dalam konteks deteksi objek perkotaan, GeoAI menunjukkan kontribusi yang signifikan dengan memanfaatkan algoritma deep learning, khususnya Convolutional Neural Networks (CNNs), untuk menganalisis citra satelit atau foto udara dan mendeteksi objek-objek seperti bangunan, jalan, atau infrastruktur lainnya. Model deteksi objek ini telah dirancang untuk mengenali fitur spasial yang kompleks dalam citra yang dapat mengandung berbagai jenis objek dengan skala dan bentuk yang bervariasi. Dalam penelitian oleh Li et al. (2025), GeoAI terbukti sangat efektif dalam meningkatkan akurasi deteksi objek perkotaan, terutama dengan menggunakan pendekatan berbasis multi-skala dan deep neural network[6]. Dengan kemampuan untuk mengidentifikasi objek secara otomatis dan akurat, GeoAI memungkinkan pemetaan perkotaan yang lebih efisien dan membantu dalam pengelolaan infrastruktur serta perencanaan kota yang lebih baik. Selain itu, pendekatan ini dapat menangani masalah yang sering dihadapi dalam deteksi objek kecil yang sulit terdeteksi, yang semakin mempertegas relevansi GeoAI dalam bidang ini[7].

Deep Learning untuk Deteksi Objek

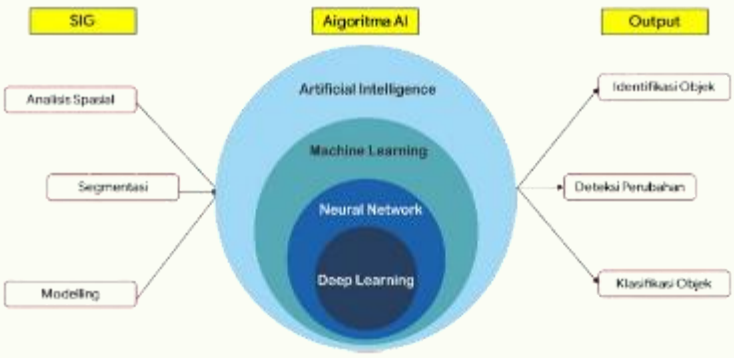
Deep learning telah menjadi pendekatan yang dominan dalam deteksi objek, terutama pada data visual seperti citra satelit dan foto udara. Dalam konteks deteksi objek, Convolutional Neural Networks (CNNs) memainkan peran kunci karena kemampuannya untuk mengekstraksi fitur penting dari gambar secara otomatis[2]. CNNs bekerja dengan cara memproses citra melalui serangkaian lapisan konvolusi yang mampu mengenali pola-pola lokal, seperti tepi, sudut, dan tekstur, yang sangat berguna dalam mengidentifikasi objek spesifik seperti bangunan, jalan, atau kendaraan. Keunggulan dari deep learning dalam deteksi objek adalah kemampuannya untuk belajar langsung dari data besar tanpa membutuhkan fitur yang diprogram secara manual, yang sebelumnya menjadi tantangan dalam metode konvensional. Selain itu, deep learning dapat menangani variasi dalam kondisi gambar, seperti perubahan pencahayaan, sudut pandang, atau gangguan latar belakang, yang biasanya menyulitkan metode deteksi objek tradisional[6]. Penelitian yang dilakukan oleh Li et al. (2025) dalam Journal of Geophysical Research: Machine Learning and Computation menunjukkan bagaimana teknik deep learning, khususnya CNNs, telah berhasil meningkatkan akurasi deteksi objek dalam citra geospasial[6].

Neural network terdiri dari tiga jenis lapisan utama yakni Input Layer: Lapisan pertama yang menerima input data. Setiap neuron di lapisan ini mewakili satu fitur dalam dataset, misalnya, setiap pixel dalam gambar, Hidden Layers: Lapisan-lapisan yang berada di antara input dan output. Di sinilah komputasi utama terjadi. Dalam deep learning, terdapat banyak hidden layers, sehingga jaringan ini dikenal sebagai Deep Neural Niiouetwork (DNN), dan Output Layer: Lapisan terakhir yang memberikan output akhir, bisa berupa probabilitas, label klasifikasi, atau nilai tertentu tergantung pada tipe tugas[8], [9], [10], [11]



Gambar 1. Struktur Neural Network dalam Deep Learning[12].

Seperti halnya Struktur Neural Network, Input data Sistem Informasi Geospasial (SIG) mencakup berbagai kegiatan seperti analisis spasial, segmentasi, dan modelling, yang bertujuan untuk memproses data geospasial guna menemukan pola atau informasi penting. Algoritma AI terbagi menjadi tiga tingkatan: Artificial Intelligence (AI) memberikan kerangka umum untuk analisis berbasis algoritma, Machine Learning (ML) memungkinkan model untuk belajar dari data geospasial dan melakukan prediksi atau klasifikasi, dan Deep Learning (DL), sebagai subkategori dari Machine Learning, memanfaatkan deep neural network untuk menganalisis data geospasial yang lebih kompleks, seperti citra satelit atau data sensor, dengan akurasi tinggi, yang semuanya bekerja secara sinergis untuk mempercepat analisis dan pengambilan keputusan berbasis lokasi[5], [6]. Output yang dihasilkan dari penerapan GeoAI pada sisi kanan gambar meliputi identifikasi objek, deteksi perubahan, dan klasifikasi objek, yang sangat berguna dalam deteksi otomatis objek perkotaan, pemantauan perubahan lingkungan, dan analisis dinamika penggunaan lahan. Dengan demikian, penerapan Deep Learning dalam GeoAI memungkinkan pengenalan dan klasifikasi fitur spasial secara lebih efisien dan akurat.



Gambar 2. Ilustrasi Konsep GeoAI[13]

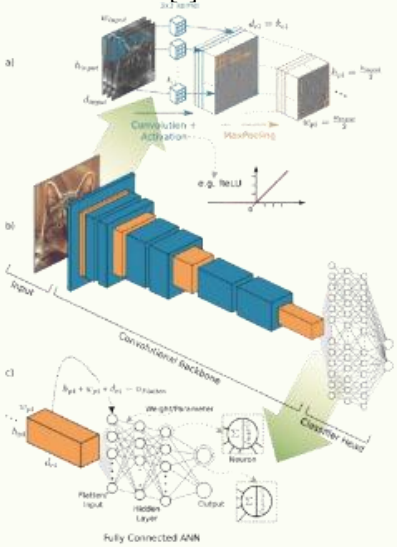
Segmentasi Citra Satelit

Segmentasi citra satelit dalam konteks GeoAI (Geospatial Artificial Intelligence) dan Deep Learning merujuk pada proses pembagian citra penginderaan jauh menjadi unit-unit homogen pada tingkat piksel atau objek untuk mengidentifikasi dan memetakan fitur permukaan bumi secara otomatis. Menurut Zhu et al. (2023) dalam ISPRS Journal of Photogrammetry and Remote Sensing, segmentasi berbasis deep learning, khususnya menggunakan Convolutional Neural Networks (CNNs) yang memungkinkan model untuk mempelajari representasi spasial dan spektral yang kompleks dari citra satelit, sehingga setiap piksel dapat diklasifikasikan ke dalam kelas tertentu seperti bangunan, vegetasi, badan air, atau jalan[14], [15]. Dalam kerangka GeoAI, segmentasi tidak hanya dipandang sebagai tugas pemrosesan citra, tetapi sebagai bagian dari

analisis spasial cerdas yang mempertimbangkan konteks geografis, skala, dan heterogenitas data geospasial.

Skema arsitektur Convolutional Neural Network (CNN) yang digunakan untuk pengenalan gambar, yang terdiri dari tiga tahap utama. Tahap pertama (konvolusi dan aktivasi), citra input diproses dengan kernel 3x3 untuk melakukan operasi konvolusi, yang menghasilkan feature maps yang menggambarkan ciri-ciri penting seperti tepi dan tekstur. Setelah konvolusi, fungsi aktivasi seperti ReLU diterapkan untuk memperkenalkan non-linearitas ke dalam model. Tahap ini diikuti dengan operasi MaxPooling, yang digunakan untuk mengurangi dimensi feature map dengan memilih nilai maksimum dari wilayah kecil, sehingga mengurangi kompleksitas komputasi sambil mempertahankan informasi spasial penting[7]. Tahap kedua (backbone konvolusional), citra diproses melalui beberapa lapisan konvolusi dan pooling, memungkinkan model untuk mengekstraksi fitur yang semakin kompleks dari citra. Proses ini memperdalam representasi citra dan memungkinkan model untuk mengenali pola yang lebih abstrak[7].

Tahap terakhir adalah fully connected ANN, di mana fitur yang diekstraksi dari lapisan konvolusi diratakan menjadi vektor satu dimensi dan dimasukkan ke dalam jaringan syaraf tiruan yang sepenuhnya terhubung. Setiap neuron dalam jaringan terhubung dengan neuron lainnya, dan bobot dihitung untuk menghasilkan prediksi kelas. Di sini, model menghasilkan output dalam bentuk probabilitas untuk setiap kelas, dengan confidence score yang menunjukkan tingkat keyakinan model terhadap klasifikasi tersebut[7].



Gambar 3. Gambaran detail arsitektur Convolutional Neural Network (CNN) untuk segmentasi citra[7]

Deskripsi Dataset

Dataset yang digunakan dalam penelitian ini berupa foto udara resolusi tinggi yang diperoleh dari Open Photogrammetry Format (OPF) Pix4D pada urban area di Negara Switzerland, yang disediakan secara terbuka untuk berbagai keperluan. Dataset ini merupakan hasil akuisisi menggunakan Unmanned Aerial Vehicle (UAV) eBee X (senseFly) yang menghasilkan ortofoto dengan detail spasial tinggi sehingga memungkinkan identifikasi objek kendaraan secara visual. Foto udara yang digunakan merupakan hasil pengolahan 100 foto yang telah diproses menggunakan metode fotogrametri menjadi satu citra orthophoto mosaik yang telah terkoreksi secara geometris dan radiometris.

1. Spesifikasi Dataset

   Spesifikasi dataset foto udara yang digunakan sebagai berikut:

Tabel 1 Spesifikasi Dataset

Location	**Switzerland**
Average Ground Sampling Distance (GSD)	2.41 cm / 0.95 in

Area covered	0.0214 km2 / 2.14 ha / 0.008 sq. mi. / 5.3 acres
Output coordinate system	WGS 84 / UTM zone 32N - EPSG 32632, WGS 84 ellipsoid
Image acquisition	
UAV	eBee X (senseFly)
Image acquisition plan	1 flight, single grid flight plan
Camera	senseFly AeriaX
Images	
Number of images	100
Image size	6000 x 4000 pixels
Image geolocation coordinate system	WGS84 - EPSG 4326, WGS 84 ellipsoid
Ground control points (GCPs)	
Number of GCPs	3
GCPs coordinate system	WGS84 - EPSG 4326, WGS 84 ellipsoid

2. Karakteristik Dataset

Foto Udara yang digunakan memiliki karakteristik:

- Variasi orientasi kendaraan
- Kondisi pencahayaan alami saat akuisisi UAV
- Kendaraan yang dipilih sebagai sampel memenuhi kriteria berikut:
- Terlihat jelas pada citra (tidak tertutup lebih dari 50%)
- Ukuran objek cukup representatif pada resolusi citra
- Tidak terdistorsi secara signifikan oleh sudut pengambilan gambar

3. Proses Labeling

Tahapan labeling meliputi:

- Identifikasi visual objek mobil pada citra
- Pembuatan polygon mask mengikuti batas geometris kendaraan
- Validasi ulang hasil masking untuk memastikan konsistensi

Pendekatan masking dipilih karena memberikan akurasi spasial yang lebih tinggi dibandingkan metode bounding box konvensional, terutama untuk evaluasi berbasis Intersection over Union (IoU)[1].

4. Pembagian Dataset

Orthophoto dibagi menjadi beberapa tile/crop berukuran seragam untuk kebutuhan training model.

- Deliniasi 80 objek mobil dilakukan pada seluruh area orthophoto.
- Dataset dibagi berdasarkan distribusi objek mobil menjadi:
    - 70% objek untuk training
    - 15% objek untuk validation
    - 15% objek untuk testing

Pembagian ini dilakukan secara acak dengan memastikan bahwa objek pada data testing tidak digunakan dalam proses training untuk menghindari data leakage. Pendekatan berbasis distribusi objek ini dipilih karena unit analisis utama dalam penelitian adalah objek kendaraan, bukan citra individual.

Metode Evaluasi Model

Evaluasi performa model dilakukan pada subset data testing yang terdiri dari 15% dari total 80 objek yang diberikan label, yang tidak digunakan dalam proses training maupun validation untuk mencegah data leakage.Evaluasi performa deteksi objek dilakukan menggunakan:

Vol. 20. No. 1, March 2026, pp 69-85

1. Penilaian kinerja deteksi objek dilakukan menggunakan pendekatan multi-threshold Intersection over Union (IoU) dengan nilai ambang ≥0.5, ≥0.75, dan ≥0.95. Threshold 0.5 dipilih karena merupakan standar umum dalam evaluasi object detection, sedangkan threshold 0.75 dan 0.95 digunakan untuk menguji sensitivitas model terhadap peningkatan presisi geometris segmentasi.
2. Confusion Matrix, yang digunakan untuk menganalisis distribusi:
   - True Positive (TP): objek mobil yang terdeteksi dengan benar dan memenuhi threshold IoU
   - False Positive (FP): prediksi mobil yang tidak sesuai dengan ground truth atau tidak memenuhi threshold IoU
   - False Negative (FN): objek mobil pada ground truth yang tidak berhasil terdeteksi
3. Precision, Recall, dan F1-Score untuk menilai kemampuan klasifikasi model dalam mengidentifikasi objek kendaraan secara benar.
   Berdasarkan nilai TP, FP, dan FN, metrik evaluasi dihitung menggunakan rumus[3]:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

4. Mean Average Precision (mAP), yang dihitung sebagai rata-rata nilai Average Precision pada beberapa threshold IoU untuk memberikan gambaran performa model secara agregat dan lebih robust.

Pendekatan multi-metrik ini digunakan untuk memastikan bahwa evaluasi model tidak hanya bergantung pada satu parameter tunggal, tetapi mencerminkan stabilitas dan konsistensi performa model terhadap variasi standar presisi spasial[1].

HASIL DAN PEMBAHASAN

Pembuatan Datasets (Label)

Pembuatan datasets (label) dalam pemrosesan GeoAI dan Deep Learning untuk deteksi objek merupakan tahap krusial yang memengaruhi kualitas dan akurasi model. Proses labeling memerlukan ketelitian tinggi karena setiap objek dalam citra harus diberi label yang tepat, seperti koordinat bounding box untuk objek yang relevan, yang memungkinkan model untuk belajar mengenali pola dan fitur visual pada gambar[16]. Penerapan Deep Learning pada penelitian ini yakni pada objek mobil.

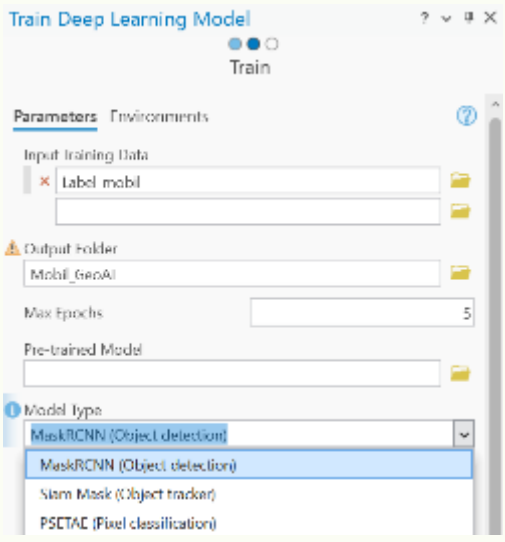
Metode Deep Learning yang digunakan yakni Region-based Convolutional Neural Network (RCNN) Masks. RCNN Masks bertugas untuk mendeteksi dan memetakan objek, seperti kendaraan, dalam citra geospasial secara lebih mendalam dan presisi. Dalam konteks ini, "RCNN Masks" mengacu pada format data yang dihasilkan oleh model, yang tidak hanya mengidentifikasi keberadaan objek (seperti mobil) tetapi juga menyediakan masker piksel yang menggambarkan bentuk dan lokasi objek dalam citra.

RUANG: JURNAL ARSITEKTUR
Vol. 20. No. 1, March 2026, pp 69-85

Gambar 4. Metode Deep Learning RCNN Masks

Training Model

Algoritma Region-based Convolutional Neural Networks (RCNN), yang dilatih dengan menggunakan data citra satelit untuk mendeteksi objek mobil. Model RCNN melalui proses forward pass untuk menghitung output berdasarkan input data, yang kemudian dibandingkan dengan label yang sudah ditentukan untuk menghitung kesalahan (loss)[7]. Dengan menggunakan teknik ini, training model dalam GeoAI dapat menghasilkan model deteksi objek yang sangat akurat.



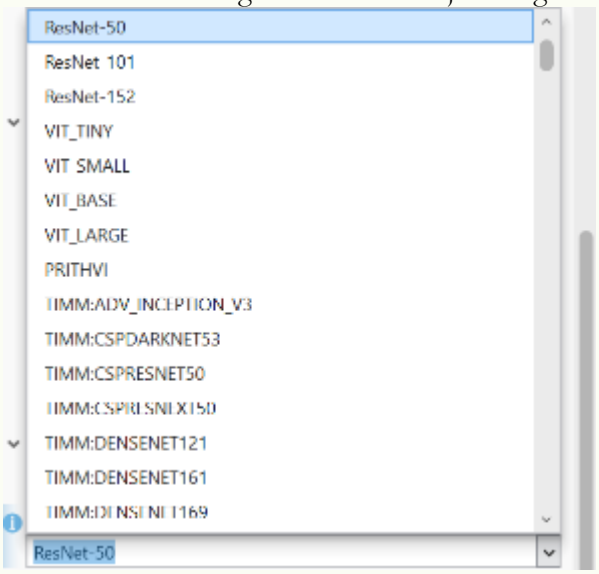
Gambar 5. Train deep Learning Model Mask RCNN (Object Detection)

Training di lakukan pada mobil, dengan melakukan masking pada mobil-mobil yang dijadikan sampel.

RUANG: JURNAL ARSITEKTUR
Vol. 20. No. 1, March 2026, pp 69-85

Gambar 6. Masking Training Model.

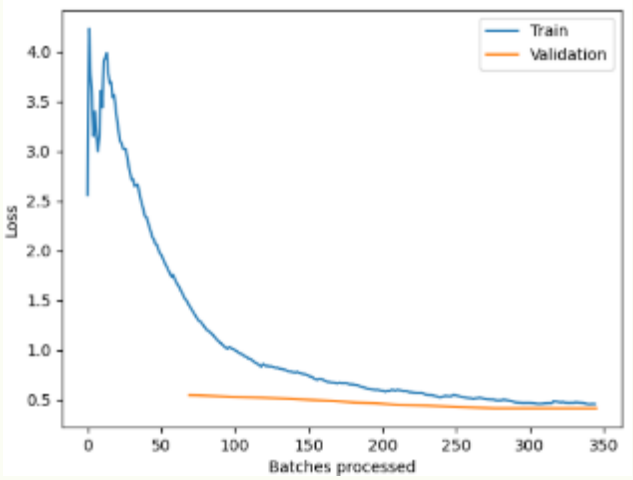
Aritektur Neural Network yang digunakan yakni ResNet-50 (Residual Network 50), ResNet-50 dapat digunakan sebagai model dasar untuk melakukan ekstraksi fitur dari gambar, memungkinkan model mendeteksi dan mengklasifikasikan objek dengan akurasi tinggi.



Gambar 7. Residual Network 50/ResNet-50

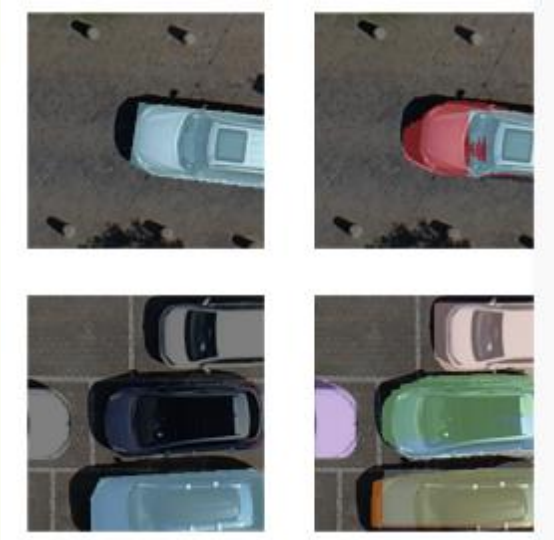
Grafik dari proses training menunjukkan bahwa Nilai training loss (garis biru) menunjukkan penurunan yang tajam, mengindikasikan bahwa model belajar dengan baik dan mengurangi kesalahan pada data pelatihan. Sementara itu, validation loss (garis oranye) stabil pada level yang rendah, menunjukkan bahwa model tidak hanya menghafal data pelatihan tetapi juga dapat menggeneralisasi dengan baik pada data yang belum pernah dilihat sebelumnya. Dengan demikian, model ini tidak mengalami overfitting dan menunjukkan kinerja yang baik, baik pada data pelatihan maupun data validasi.

RUANG: JURNAL ARSITEKTUR
Vol. 20. No. 1, March 2026, pp 69-85



Gambar 8. Grafik Hasil Training Model

Nilai Average Precision Score untuk kategori "Mobil" adalah 0.7111, yang menunjukkan bahwa model memiliki presisi yang cukup baik dalam mendeteksi dan mengklasifikasikan objek mobil.



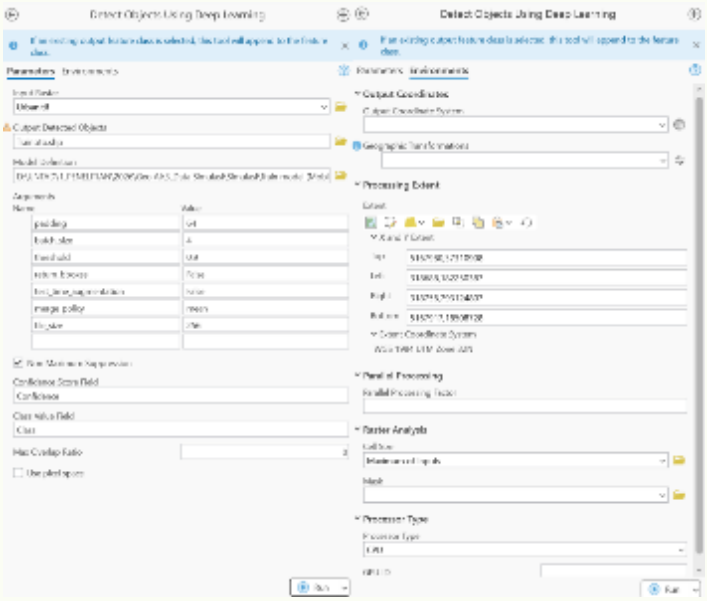
Gambar 9. Hasil Masking Training Model

Gambar 10. Nilai Rata-rata tingkat Presisi Training Model

Deteksi Objek

Deteksi objek dalam pemrosesan GeoAI dan Deep Learning adalah tahap kritis yang memungkinkan sistem untuk mengidentifikasi dan mengklasifikasikan objek-objek tertentu dalam citra satelit. Proses ini menggunakan Convolutional Neural Networks (CNNs), yang diterapkan untuk mengekstraksi fitur spasial dari citra dan kemudian melakukan klasifikasi terhadap objek mobil pada penelitian ini.

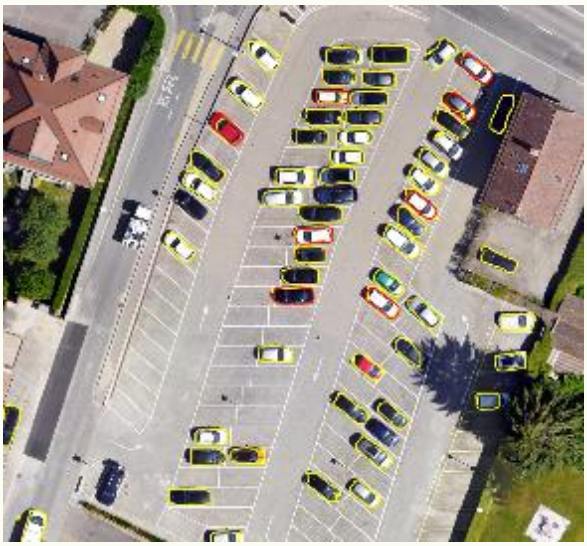
RUANG: JURNAL ARSITEKTUR
Vol. 20. No. 1, March 2026, pp 69-85



Gambar 11. Parameters dan Enviroments pada Tahap Deteksi Objek

Sampel training (garis merah) yang digunakan oleh model untuk belajar mengenali mobil, sementara garis kuning menunjukkan hasil deteksi model pada citra yang sebenarnya, Hasil deteksi menunjukkan bahwa RCNN bekerja dengan baik dalam mendeteksi mobil. Garis kuning yang menggambarkan deteksi objek mencakup sebagian besar mobil yang ada di gambar, dengan bounding box yang cukup tepat untuk memisahkan mobil satu dengan yang lainnya. Hal ini menunjukkan bahwa model berhasil mendeteksi objek dengan akurasi tinggi, meskipun terdapat beberapa kemungkinan kesalahan dalam mendeteksi mobil yang tumpang tindih atau terhalang. Selain itu, model juga menunjukkan kemampuan untuk mempertahankan konsistensi deteksi pada objek dengan variasi posisi dan orientasi.

Namun masih terdapat beberapa kesalahan deteksi, yang ditandai dengan beberapa area yang tidak terdeteksi atau tumpang tindih antara objek. Ini menunjukkan bahwa meskipun RCNN sudah cukup baik, peningkatan model bisa diperlukan untuk memperbaiki deteksi pada objek yang lebih kompleks, seperti kendaraan yang saling menempel atau bagian mobil yang sebagian terhalang.



Gambar 12. Hasil Deteksi Objek

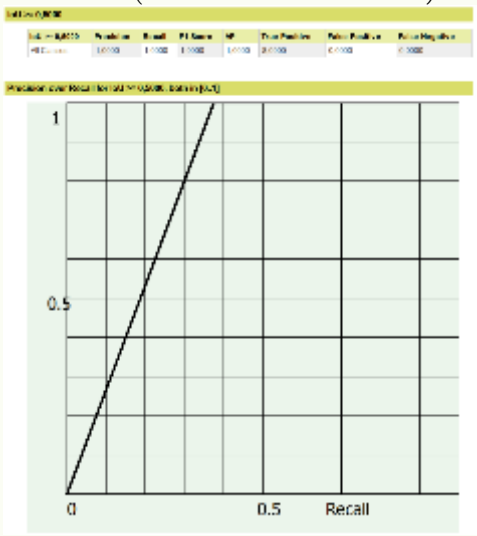
RUANG: JURNAL ARSITEKTUR
Vol. 20. No. 1, March 2026, pp 69-85

Evaluasi Model (berdasarkan IoU)

Evaluasi model dalam pemrosesan GeoAI dan Deep Learning untuk deteksi objek adalah tahap yang sangat penting untuk menilai seberapa baik model dalam mengenali objek-objek yang ada dalam citra satelit secara kuantitatif. Evaluasi ini menggunakan beberpa sampel mobil yang di deliniasi secara detail untuk mengukur ketelitian hasil penerapan RCNN. Pada Gambar di bawah ini Garis Merah merupakan masking untuk evaluasi sedangkan garis kuning adalah hasil deteksi objek menggunakan RCNN.

Gambar 13. Hasil Deteksi RCNN dan uji Akurasi

Hasil evaluasi model berdasarkan IoU (Intersection over Union) >= 0.5

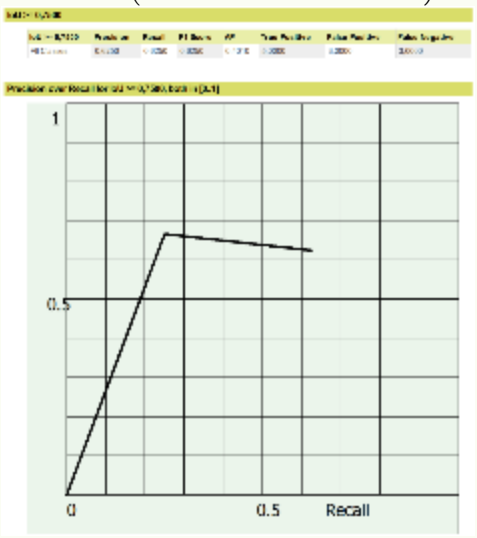


Gambar 14. IoU (Intersection over Union) >= 0.5

IoU >= 0.5, model menunjukkan kinerja sempurna dengan Precision, Recall, F1 Score, dan Average Precision (AP) masing-masing sebesar 1.0000, yang menunjukkan deteksi objek yang sangat akurat dan komprehensif tanpa kesalahan baik False Positive maupun False Negative. Grafik terkait menunjukkan hubungan yang sempurna antara Precision dan Recall, dengan peningkatan yang konsisten seiring dengan meningkatnya Recall.

RUANG: JURNAL ARSITEKTUR
Vol. 20. No. 1, March 2026, pp 69-85

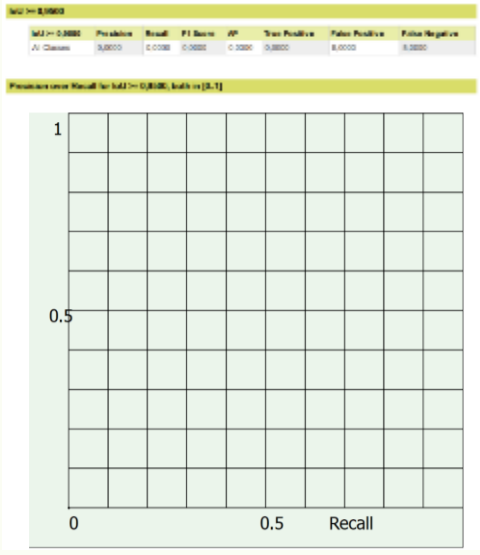
Hasil evaluasi model berdasarkan IoU (Intersection over Union) >= 0.75



Gambar 15. IoU (Intersection over Union) >= 0.75

IoU >= 0.75, model mengalami penurunan signifikan dalam akurasi, dengan Precision dan Recall masing-masing hanya mencapai 0.6250, yang menunjukkan adanya trade-off antara keakuratan deteksi dan jumlah objek yang terdeteksi. Average Precision (AP) turun menjadi 0.4010, dan meskipun True Positive berjumlah 5,000, kesalahan deteksi muncul dengan 3,000 False Positive dan 3,000 False Negative. Grafik yang terkait menunjukkan peningkatan yang lebih kecil dalam Precision dan Recall, menandakan bahwa model mulai kesulitan mencapai keseimbangan yang baik pada nilai IoU yang lebih ketat.

Hasil uji akurasi pada pengolahan GeoAI yang menggunakan IoU (Intersection over Union) >= 0.9500



Gambar 16. IoU (Intersection over Union) >= 0.95

IoU >= 0.95, model gagal dalam deteksi objek, dengan semua nilai evaluasi (Precision, Recall, F1 Score, dan AP) bernilai 0.0000. True Positive tercatat 0, sementara kesalahan False Positive dan False Negative mencapai 8,000 masing-masing. Grafik terkait menunjukkan tidak ada peningkatan dalam Precision maupun Recall, menandakan bahwa threshold IoU yang sangat tinggi membuat model tidak mampu mendeteksi objek dengan benar, bahkan mengklasifikasikan semua objek sebagai kesalahan.

Secara keseluruhan, kinerja model sangat tergantung pada nilai threshold IoU yang digunakan. Pada IoU >= 0.5, model memberikan hasil yang sangat baik, namun seiring peningkatan nilai IoU, performa menurun tajam, yang menunjukkan bahwa model membutuhkan penyesuaian lebih lanjut dalam hal sensitivitas untuk mendeteksi objek pada tingkat presisi yang lebih tinggi.

Evaluasi Robust dan Validasi Model

1. Evaluasi Berbasis Confusion Matrix

Hasil Confusion Matrix

IoU ≥ 0.5

TP = 12, FP = 0, FN = 0

Model berhasil mendeteksi seluruh objek mobil pada data testing tanpa kesalahan klasifikasi. Hal ini menunjukkan kemampuan model dalam mengenali dan menglokalisasi objek pada tingkat overlap moderat.

IoU ≥ 0.75

TP = 8, FP = 4, FN = 4

Pada threshold yang lebih ketat, model mulai mengalami penurunan performa. Empat objek tidak memenuhi standar overlap minimum 75%, dan terdapat empat prediksi yang tidak sesuai secara spasial dengan ground truth. Penurunan ini menunjukkan adanya deviasi batas segmentasi pada beberapa kendaraan, khususnya pada area tepi objek dan bayangan.

IoU ≥ 0.95

TP = 0, FP = 12, FN = 12

Tidak terdapat prediksi yang memenuhi standar overlap sebesar 95%. Hal ini menunjukkan bahwa meskipun objek berhasil dikenali pada threshold lebih rendah, presisi segmentasi pada tingkat pixel-perfect belum tercapai.

2. Perhitungan Precision, Recall, dan F1-Score

Hasil perhitungan menunjukkan:

IoU	Precision	Recall	F1-score
≥0.5	1.00	1.00	1.00
≥0.75	0.67	0.67	0.67
≥0.95	0.00	0.00	0.00

Hasil ini menunjukkan bahwa model memiliki performa sangat baik pada threshold moderat, namun sensitivitas meningkat secara signifikan pada threshold yang lebih ketat.

3. Mean Average Precision (mAP)

Mean Average Precision (mAP) dihitung sebagai rata-rata nilai Average Precision pada tiga threshold IoU (0.5, 0.75, dan 0.95). Hasil perhitungan menunjukkan nilai mAP sebesar 0.56, yang mengindikasikan bahwa model memiliki performa agregat yang moderat. Meskipun deteksi objek sangat akurat pada IoU ≥0.5, penurunan signifikan pada threshold yang lebih ketat mempengaruhi stabilitas performa keseluruhan.

6. Keterbatasan Validasi

Meskipun evaluasi telah dilakukan secara komprehensif, dataset berasal dari satu orthophoto sehingga potensi spatial autocorrelation masih ada. Oleh karena itu, generalisasi model terhadap kawasan perkotaan lain dengan karakteristik berbeda perlu diuji pada penelitian lanjutan.

Implikasi terhadap Perencanaan Kota

Hasil evaluasi menunjukkan bahwa model Mask R-CNN memiliki performa sangat tinggi pada threshold IoU ≥0.5 (Precision dan Recall = 1.00), namun mengalami degradasi signifikan pada threshold yang lebih ketat (IoU ≥0.75 dan ≥0.95). Temuan ini memiliki implikasi langsung terhadap skala dan jenis analisis dalam praktik perencanaan wilayah dan kota.

Pada tingkat IoU ≥0.5, model mampu mengidentifikasi keberadaan dan lokasi kendaraan secara konsisten. Dalam konteks perencanaan kota, tingkat presisi ini sudah memadai untuk

kebutuhan analisis makro seperti Estimasi jumlah kendaraan pada suatu area parkir, Analisis kepadatan kendaraan di koridor jalan, Identifikasi pola distribusi kendaraan dalam kawasan komersial, Evaluasi kapasitas ruang parkir eksisting

Dengan akurasi deteksi yang stabil pada threshold moderat, GeoAI berpotensi menjadi instrumen pendukung dalam penyusunan kebijakan manajemen parkir dan pengendalian lalu lintas berbasis data spasial. Namun demikian, penurunan performa pada IoU ≥0.75 menunjukkan bahwa model belum mampu mempertahankan presisi geometris pada tingkat detail tinggi. Implikasi dari temuan ini adalah bahwa model belum sepenuhnya reliabel untuk analisis mikro-spasial, seperti Pengukuran dimensi kendaraan secara presisi, Analisis jarak antar kendaraan untuk evaluasi standar keselamatan, Delineasi batas parkir individual dengan toleransi geometris ketat. Hasil penelitian ini menunjukkan bahwa penerapan GeoAI dalam konteks perencanaan kota perlu disesuaikan dengan kebutuhan skala analisis. Untuk kebutuhan estimasi kuantitatif dan pemetaan distribusi kendaraan, model sudah cukup memadai. Namun untuk kebutuhan presisi geometris tinggi, diperlukan pengembangan lanjutan, baik melalui peningkatan resolusi data maupun optimasi arsitektur model.

Penggunaan pendekatan ini berpotensi mendukung paradigma perencanaan berbasis data (*data-driven urban planning*), di mana informasi spasial dapat diperoleh secara otomatis dan periodik tanpa survei manual yang memakan waktu dan biaya tinggi. Integrasi deteksi objek berbasis GeoAI dengan Sistem Informasi Geografis (SIG) memungkinkan pembaruan basis data spasial secara dinamis, sehingga mendukung proses monitoring kawasan perkotaan secara berkelanjutan. Namun demikian, karena dataset berasal dari satu orthophoto dengan karakteristik spasial homogen, generalisasi hasil terhadap kawasan dengan morfologi perkotaan yang berbeda masih memerlukan pengujian lebih lanjut. Oleh karena itu, penerapan operasional dalam konteks perencanaan kota sebaiknya dilakukan secara bertahap dengan validasi tambahan pada dataset lintas lokasi.

KESIMPULAN

Penelitian ini menunjukkan bahwa penerapan GeoAI berbasis Mask R-CNN dengan arsitektur ResNet-50 mampu mendeteksi objek kendaraan pada citra orthophoto resolusi tinggi dengan tingkat akurasi yang sangat baik pada threshold IoU ≥0.5. Pada tingkat ini, model mencapai Precision, Recall, dan F1-score sebesar 1.00, yang menunjukkan kemampuan deteksi dan lokalisasi objek secara konsisten pada toleransi spasial moderat.

Namun demikian, evaluasi multi-threshold mengungkapkan bahwa performa model mengalami penurunan signifikan pada IoU ≥0.75 dan gagal memenuhi presisi geometris ekstrem pada IoU ≥0.95. Nilai Mean Average Precision (mAP) sebesar 0.56 menunjukkan bahwa secara agregat model memiliki performa moderat, dengan stabilitas yang cukup pada deteksi objek secara umum, tetapi belum konsisten dalam mempertahankan ketelitian batas segmentasi pada tingkat presisi spasial tinggi. Temuan ini menegaskan adanya trade-off antara kemampuan identifikasi objek dan ketelitian geometris mask hasil segmentasi.

Dalam konteks perencanaan wilayah dan kota, hasil penelitian ini memiliki implikasi yang bersifat kontekstual. Performa tinggi pada IoU ≥0.5 menunjukkan bahwa model telah memadai untuk kebutuhan analisis makro-spasial, seperti estimasi jumlah kendaraan, analisis kepadatan parkir, serta pemetaan distribusi aktivitas kendaraan dalam suatu kawasan. Pendekatan ini berpotensi mendukung perencanaan berbasis data (data-driven urban planning) dengan menyediakan mekanisme ekstraksi informasi spasial secara otomatis dan efisien dari citra udara resolusi tinggi.

Namun demikian, keterbatasan pada threshold IoU yang lebih ketat menunjukkan bahwa model belum sepenuhnya reliabel untuk analisis mikro-spasial yang membutuhkan presisi geometris tinggi, seperti pengukuran dimensi kendaraan atau evaluasi jarak antar kendaraan secara detail. Oleh karena itu, penerapan operasional dalam perencanaan kota perlu disesuaikan dengan skala dan tujuan analisis yang diinginkan.

Secara metodologis, penelitian ini menegaskan pentingnya evaluasi multi-threshold dalam studi deteksi objek berbasis deep learning, khususnya ketika digunakan dalam analisis spasial perkotaan. Ketergantungan pada satu threshold IoU (misalnya ≥0.5) berpotensi memberikan gambaran performa yang terlalu optimistik tanpa mempertimbangkan sensitivitas terhadap presisi geometris.

Meskipun dataset yang digunakan berasal dari satu orthophoto dengan karakteristik spasial yang relatif homogen, hasil penelitian ini memberikan dasar konseptual bagi integrasi GeoAI dalam sistem informasi geografis untuk mendukung monitoring kawasan perkotaan secara berkala. Penelitian lanjutan dengan dataset lintas lokasi dan morfologi perkotaan yang lebih beragam diperlukan untuk meningkatkan generalisasi dan robustnes model dalam konteks implementasi kebijakan perencanaan wilayah berbasis data spasial.

PENGAKUAN

Terima kasih yang sebesar-besarnya kepada rekan-rekan, dan semua pihak yang telah berkontribusi dalam penelitian ini. Terima kasih atas dukungan yang tak henti-hentinya, diskusi yang membangun, dan selalu siap memberikan bantuan.

DEKLARASI PENULIS

Pernyataan Penulis	:	Para penulis memberikan kontribusi substansial terhadap konsepsi dan desain penelitian. Para penulis bertanggung jawab atas analisis data, interpretasi, dan diskusi hasil. Para penulis membaca dan menyetujui manuskrip akhir.
Pernyataan Pendanaan	:	Tidak satu pun dari para penulis menerima dana atau hibah dari lembaga atau badan pendanaan mana pun untuk penelitian ini.
Konflik Kepentingan	:	Penulis menyatakan tidak ada konflik kepentingan.
Informasi Tambahan	:	Tidak ada informasi tambahan yang tersedia untuk makalah ini

DAFTAR PUSTAKA

[1] K. He, G. Gkioxari, P. Dollar, and R. Girshick, *Mask R-CNN*. 2017. doi: 10.1109/ICCV.2017.322.
[2] Michael Batty, "Artificial intelligence and smart cities," *Environ. Plan. B Urban Anal. City Sci.*, vol. 45, no. 1, pp. 3–6, Jan. 2018, doi: 10.1177/2399808317751169.
[3] R. Kitchin, "The ethics of smart cities and urban science," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2083, p. 20160115, Dec. 2016, doi: 10.1098/rsta.2016.0115.
[4] G. Mai *et al.*, "Towards the next generation of Geospatial Artificial Intelligence," *International Journal of Applied Earth Observation and Geoinformation*, vol. 136, p. 104368, Feb. 2025, doi: 10.1016/j.jag.2025.104368.
[5] W. Li, "Artificial Intelligence in Earth Science: A GeoAI Perspective," *Journal of Geophysical Research: Machine Learning and Computation*, vol. 2, no. 3, Sep. 2025, doi: 10.1029/2025JH000691.
[6] T. Hoeser and C. Kuenzer, "Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review-Part I: Evolution and Recent Trends," *Remote Sens. (Basel).*, vol. 12, no. 10, p. 1667, May 2020, doi: 10.3390/rs12101667.
[7] P. N. Zivich and A. I. Naimi, "A primer on neural networks," *Am. J. Epidemiol.*, vol. 194, no. 6, pp. 1473–1475, Jun. 2025, doi: 10.1093/aje/kwae380.
[8] A. Masood and K. Ahmad, "A review on emerging artificial intelligence (AI) techniques for air pollution forecasting: Fundamentals, application and performance," *J. Clean. Prod.*, vol. 322, p. 129072, Nov. 2021, doi: 10.1016/j.jclepro.2021.129072.
[9] M. Madhiarasan and M. Louzazni, "Analysis of Artificial Neural Network: Architecture, Types, and Forecasting Applications," *Journal of Electrical and Computer Engineering*, vol. 2022, pp. 1–23, Apr. 2022, doi: 10.1155/2022/5416722.
[10] Y. , L. X. Zhang and J. Wang, "Real time object detection for urban management: Methods and applications.," *Journal of Urban Computing and Geospatial Intelligence*, 2023.
[11] J. Heaton, "Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning," *Genet. Program. Evolvable Mach.*, vol. 19, no. 1–2, pp. 305–307, Jun. 2018, doi: 10.1007/s10710-017-9314-z.

RUANG: JURNAL ARSITEKTUR
Vol. 20. No. 1, March 2026, pp 69-85

[12] A. G. Prasetyo and N. F. Anshafa, *GeoAI & Deep Learning untuk Urban Mapping: Deteksi Otomatis Objek Perkotaan dengan ArcGIS Pro*. 2025.
[13] L. , W. Z. Zhu and Q. Zhang, "Satellite Image Segmentation Using Deep Learning Techniques for Geospatial Analysis," *ISPRS Journal of Photogrammetry and Remote Sensing*, 2023.
[14] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2018, doi: 10.1109/TPAMI.2017.2699184.
[15] M. Everingham, L. Van Gool, C. Williams, J. Winn, and A. Zisserman, "The Pascal Visual Object Classes (VOC) challenge," *Int. J. Comput. Vis.*, vol. 88, pp. 303–338, Jun. 2010, doi: 10.1007/s11263-009-0275-4.
[16] X. Zhang, Y. Zhou, and J. Luo, "Deep learning for processing and analysis of remote sensing big data: a technical review," *Big Earth Data*, vol. 6, no. 4, pp. 527–560, Oct. 2022, doi: 10.1080/20964471.2021.1964879.