

IMPLEMENTASI MASK R-CNN UNTUK APLIKASI SEGMENTASI OBJEK SUATU AREA DARI UDARA

Raihan Aura Ferdiansyah¹, Suharjanto Utomo², Arief Suryadi Satyawan³, Hernawati⁴

^{1,2,3,4}Teknik Informatika, Universitas Nurtanio Bandung

Email: ¹raihanaf321@gmail.com

ABSTRAK

Kemajuan teknologi drone dan kecerdasan buatan telah memungkinkan analisis lingkungan dari udara dengan efisiensi tinggi. Penelitian ini mengimplementasikan Mask R-CNN untuk segmentasi objek dari citra udara, dengan fokus pada identifikasi sawah, bangunan, dan landasan. Dataset citra udara diperoleh dari Google Earth Pro, dan anotasi dilakukan menggunakan VGG Image Annotator (VIA). Model dilatih dengan parameter $IMAGE_MIN_DIM = 512$, $IMAGE_MAX_DIM = 512$, dan $LEARNING_RATE = 0.001$. Hasil evaluasi menunjukkan akurasi segmentasi sebesar 63% dan *mean Average Precision* (mAP) sebesar 0.8182. Meskipun model menunjukkan performa yang baik dalam mendeteksi sawah dan landasan, akurasi untuk bangunan masih rendah (51.02%). Penelitian ini memberikan kontribusi dalam pengembangan teknologi segmentasi objek real-time menggunakan drone, dengan rekomendasi untuk peningkatan dataset dan optimisasi parameter pada penelitian lanjutan.

Kata kunci: Mask R-CNN, Segmentasi Objek, Citra Udara, Drone, Kecerdasan Buatan.

ABSTRACT

Advancements in drone technology and artificial intelligence have enabled highly efficient aerial environmental analysis. This study implements Mask R-CNN for object segmentation from aerial imagery, focusing on the identification of rice fields, buildings, and runways. Aerial image datasets were obtained from Google Earth Pro, and annotations were carried out using the VGG Image Annotator (VIA). The model was trained with the parameters $IMAGE_MIN_DIM = 512$, $IMAGE_MAX_DIM = 512$, and $LEARNING_RATE = 0.001$. Evaluation results show a segmentation accuracy of 63% and a mean Average Precision (mAP) of 0.8182. Although the model demonstrated good performance in detecting rice fields and runways, the accuracy for building detection remained relatively low (51.02%). This study contributes to the development of real-time object segmentation technology using drones, with recommendations for dataset enhancement and parameter optimization in future research.

Keywords: Mask R-CNN, Object Segmentation, Aerial Imagery, Drone, Artificial Intelligence.

1. PENDAHULUAN

Perkembangan teknologi drone dan kecerdasan buatan (AI) dalam satu dekade terakhir telah memberikan dampak besar dalam berbagai bidang, termasuk pemantauan lingkungan, pertanian presisi, serta mitigasi bencana. Drone kini tidak hanya digunakan untuk pengambilan gambar dari udara, tetapi juga dimanfaatkan untuk pengumpulan data spasial secara efisien dan real-time. Salah satu tantangan utama dari data citra udara yang dikumpulkan adalah bagaimana mengekstraksi informasi penting secara otomatis, seperti mengenali objek-objek tertentu dengan cepat dan akurat. Di sinilah segmentasi objek menjadi krusial.

Segmentasi objek dari citra udara memiliki peran yang sangat penting dalam berbagai konteks. Dalam bidang pertanian, misalnya, kemampuan untuk mengidentifikasi dan membedakan sawah dari jenis lahan lainnya dapat membantu petani dan pemerintah dalam menghitung luas tanam, memantau pertumbuhan tanaman, serta mengidentifikasi wilayah yang membutuhkan intervensi. Di sisi lain, deteksi bangunan dan landasan pacu merupakan bagian penting dalam pemetaan perkotaan, pengawasan wilayah, hingga navigasi drone, terutama dalam skenario darurat seperti kebakaran, gempa, atau banjir. Akurasi dalam segmentasi objek dari citra udara menjadi penentu utama dalam efektivitas sistem pemantauan berbasis drone tersebut.

Dalam upaya memenuhi kebutuhan tersebut, Mask R-CNN muncul sebagai salah

satu algoritma deep learning yang menjanjikan. Mask R-CNN memiliki kemampuan segmentasi instance yang presisi, memungkinkan pengenalan bentuk dan batas objek secara rinci dalam sebuah gambar. Model ini telah berhasil digunakan dalam berbagai bidang, mulai dari deteksi bangunan dalam citra satelit, segmentasi buah untuk pertanian, hingga analisis medis seperti deteksi tumor. Namun, penerapan Mask R-CNN pada citra udara masih menghadapi sejumlah tantangan, seperti kompleksitas latar belakang, keberagaman bentuk dan ukuran objek, serta keterbatasan dataset berkualitas tinggi yang sesuai untuk pelatihan model.

Melihat potensi dan tantangan tersebut, penelitian ini dilakukan untuk mengimplementasikan Mask R-CNN dalam proses segmentasi objek dari citra udara, dengan fokus pada tiga kategori objek: sawah, bangunan, dan landasan. Dataset citra udara diperoleh dari Google Earth Pro, yang kemudian dianotasi secara manual menggunakan VGG Image Annotator (VIA) guna memastikan presisi data latih. Tujuan akhirnya adalah mengembangkan sistem segmentasi yang tidak hanya akurat, tetapi juga cukup efisien untuk diterapkan dalam berbagai aplikasi praktis di lapangan.

Dengan pendekatan ini, penelitian ini diharapkan dapat memberikan kontribusi nyata dalam pengembangan sistem pemantauan berbasis drone yang lebih cerdas dan otomatis. Selain itu, hasil penelitian ini juga dapat menjadi acuan dalam pengembangan lebih lanjut di bidang pengolahan citra udara dan penerapan teknologi AI untuk kebutuhan analisis spasial.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan eksperimen terapan dalam bidang visi komputer untuk mengimplementasikan dan mengevaluasi metode Mask R-CNN dalam segmentasi objek dari citra udara. Tujuan utama adalah mengidentifikasi objek sawah, bangunan, dan landasan melalui segmentasi instance berbasis deep learning. Tahapan dalam penelitian ini disusun secara sistematis sebagai berikut:

1. Pengumpulan Data

Tahap awal adalah pengumpulan data berupa citra udara dari Google Earth Pro. Citra yang dikumpulkan mencakup objek-objek target yaitu sawah, bangunan, dan

landasan, dengan berbagai sudut pengambilan gambar.

2. Anotasi Data

Data yang telah dikumpulkan dianotasi menggunakan VGG Image Annotator (VIA). Proses anotasi dilakukan dengan cara memberi label dan membuat bounding box serta mask pada masing-masing objek di citra. Hasil anotasi digunakan sebagai *ground truth* dalam proses pelatihan dan evaluasi model.

3. Pra-pemrosesan Data

Untuk menjaga konsistensi input ke dalam model, semua citra diskalakan ke resolusi tetap dengan parameter:

- IMAGE_MIN_DIM = 512,
- IMAGE_MAX_DIM = 512.

Konversi ukuran dilakukan agar semua gambar memiliki format dan dimensi yang kompatibel dengan arsitektur Mask R-CNN.

4. Implementasi Mask R-CNN

a. Pemilihan Model

Menggunakan arsitektur Mask R-CNN yang sudah terlatih (*pre-trained*) dan tersedia dalam pustaka seperti *TensorFlow*

b. Pelatihan Model

Melakukan pelatihan (*training*) model dengan menggunakan dataset yang telah diproses.

c. Parameter dan Konfigurasi

Menentukan parameter penting seperti learning rate, image per gpu, dan jumlah epoch berdasarkan eksperimen awal dan tinjauan literatur. Mengatur konfigurasi model untuk mendeteksi dan mensegmentasi sawah, bangunan, dan landasan secara spesifik.

5. Evaluasi dan Analisis

a. Metode Evaluasi

Menggunakan metrik evaluasi seperti mean *Average Precision* (mAP), *confusion matrix*, *precision recall*.

b. Analisis Hasil

Analisis hasil segmentasi untuk mengidentifikasi kekuatan dan kelemahan model. Ini termasuk evaluasi visual dari hasil segmentasi serta analisis kuantitatif berdasarkan metrik evaluasi. Melakukan analisis kesalahan untuk memahami kondisi di mana model gagal atau kurang optimal, dan memberikan

rekomendasi untuk perbaikan lebih lanjut.

6. Validasi dan Pengujian
 - a. Validasi model menggunakan data yang tidak termasuk dalam data pelatihan untuk memastikan generalisasi model terhadap data baru.
 - b. Menerapkan model yang telah dilatih dan divalidasi pada citra udara baru untuk menguji kinerja.

3. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk mengimplementasikan model Mask R-CNN dalam proses segmentasi objek dari citra udara, dengan fokus pada tiga kategori objek utama yaitu sawah, bangunan, dan landasan. Dalam bab ini, dipaparkan hasil analisis sistem, perancangan, serta implementasi teknis model, termasuk konfigurasi pelatihan dan evaluasi awal.

3.1 Analisa Sistem Berjalan

Berdasarkan studi literatur dan telaah terhadap sistem segmentasi yang telah ada, diketahui bahwa model seperti U-Net dan Faster R-CNN telah banyak digunakan untuk segmentasi objek dari citra udara. Namun, kedua pendekatan tersebut memiliki keterbatasan dalam segmentasi instance terutama pada objek yang memiliki bentuk dan ukuran bervariasi.

Sebagai alternatif yang lebih tepat, Mask R-CNN dipilih karena kemampuannya dalam menangani segmentasi instance secara lebih presisi. Model ini tidak hanya mampu mendeteksi keberadaan objek, tetapi juga menghasilkan mask piksel yang merepresentasikan bentuk objek secara lebih detail.

3.2 Perancangan Sistem

a. Dataset

Dataset yang digunakan dalam penelitian ini terdiri dari 113 citra udara, yang dibagi menjadi:

- 70 citra untuk pelatihan (*training*),
- 33 citra untuk validasi (*validation*),
- 10 citra untuk pengujian (*testing*).

Semua citra diperoleh dari Google Earth Pro dengan variasi ketinggian pengambilan gambar antara 90 hingga 2000 meter, sehingga mencakup berbagai skala objek dan jenis lingkungan seperti daerah perkotaan, pedesaan, serta area

bandara. Objek yang dianotasi dalam dataset meliputi sawah, bangunan, dan landasan, dengan tujuan memperluas cakupan generalisasi model.

b. Preprocessing Data

Agar citra dapat digunakan secara konsisten dalam proses pelatihan, dilakukan tahap preprocessing, yakni:

1. Resize citra ke dimensi tetap 512x512 piksel.
2. Normalisasi intensitas piksel, untuk menghindari ketidakkonsistenan pencahayaan.
3. Anotasi objek dilakukan menggunakan VGG Image Annotator (VIA), dengan jenis anotasi berupa bounding box dan mask (segmen piksel objek).
4. Anotasi divalidasi secara manual guna menjaga kualitas label.

Langkah ini penting untuk mempersiapkan data agar dapat diterima dan diproses secara optimal oleh arsitektur Mask R-CNN.

c. Pelatihan Model

Model Mask R-CNN dikonfigurasi dengan parameter sebagai berikut:

1. IMAGE_MIN_DIM = 512
2. IMAGE_MAX_DIM = 512
3. LEARNING_RATE = 0.001
4. STEPS_PER_EPOCH = 70
5. NUM_CLASSES = 4 (termasuk latar belakang)

Model dilatih selama 40 epoch dengan waktu pelatihan total sekitar 10 jam. Selama proses pelatihan, digunakan TensorBoard untuk memantau nilai loss, serta visualisasi kurva akurasi dan deteksi yang dihasilkan dari setiap iterasi epoch. Proses pelatihan mengoptimalkan tiga komponen utama dalam Mask R-CNN:

1. Klasifikasi objek,
2. Regresi bounding box (penyesuaian lokasi objek),
3. Segmentasi mask.

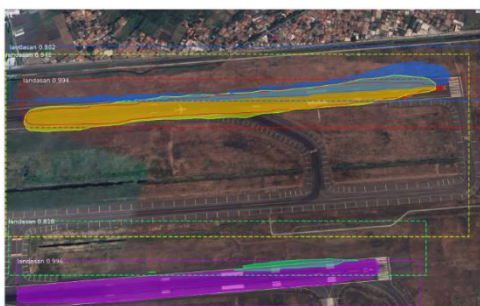
3.3 Implementasi Program

Implementasi dilakukan menggunakan Python 3.6, TensorFlow, dan Jupyter Notebook. Hardware yang digunakan meliputi Intel Core i3 Gen 10, 16GB RAM, dan GPU GTX 1650 4GB. Lingkungan ini dipilih karena kestabilan dan kompatibilitasnya dengan

pustaka *deep learning*, serta dukungan GPU yang mampu mempercepat proses training meskipun dengan spesifikasi menengah. Penggunaan Jupyter Notebook memberikan fleksibilitas tinggi dalam pengujian kode secara interaktif serta mendukung dokumentasi eksperimen yang rapi.



Gambar 2. Hasil Segmentasi Citra Sawah dan Bangunan



Gambar 3. Hasil Segmentasi Citra Landasan

3.4 Pengujian Model

Pengujian dilakukan pada 10 citra uji yang belum pernah dilatih sebelumnya. Model kemudian dievaluasi menggunakan beberapa metrik, yaitu akurasi keseluruhan, *precision*, *recall*, dan *mean Average Precision (mAP)*.

Tabel 1. Hasil Evaluasi Model Mask R-CNN

Kelas	Precision (%)	Recall (%)	Keterangan
Sawah	94.44	73.91	Tinggi, stabil
Bangunan	51.02	86.21	Variatif, kompleks
Landasan	100.00	100.00	Deteksi sempurna
Rata-rata	81.82 (mAP)	86.71	-

Akurasi Keseluruhan : 63%
 mean Average Precision (mAP): 0.8182

Hasil ini menunjukkan bahwa model sangat efektif dalam mendeteksi objek landasan dan cukup baik untuk sawah. Namun, hasil pada objek bangunan menunjukkan akurasi

(precision) yang rendah meskipun recall cukup tinggi. Hal ini kemungkinan disebabkan oleh kompleksitas bentuk bangunan, tumpang tindih objek, serta variasi latar belakang yang tinggi pada lingkungan urban.

Visualisasi hasil segmentasi juga menunjukkan bahwa model mampu membedakan objek dengan baik melalui masking berwarna yang jelas, meskipun beberapa objek kecil atau mirip bangunan gagal dideteksi.

4. KESIMPULAN

Penelitian ini berhasil mengimplementasikan metode Mask R-CNN untuk segmentasi objek dari citra udara, dengan fokus pada identifikasi sawah, bangunan, dan landasan. Model yang dikembangkan menunjukkan performa yang cukup baik, dengan akurasi keseluruhan sebesar 63% dan nilai mean Average Precision (mAP) sebesar 0.8182. Hasil segmentasi paling optimal diperoleh pada objek sawah dan landasan, dengan precision masing-masing 94.44% dan 100%, serta recall sebesar 73.91% dan 100%. Namun, segmentasi pada objek bangunan masih memiliki precision yang relatif rendah (51.02%), yang kemungkinan besar disebabkan oleh variasi bentuk, tekstur, dan latar belakang yang kompleks.

Temuan ini menunjukkan bahwa Mask R-CNN memiliki potensi signifikan dalam segmentasi objek dari citra udara, terutama untuk aplikasi pemantauan pertanian, perencanaan wilayah, dan navigasi drone. Meski demikian, beberapa aspek teknis masih perlu ditingkatkan agar model dapat berfungsi lebih optimal pada objek dengan kompleksitas tinggi.

Sebagai tindak lanjut, terdapat beberapa saran untuk pengembangan penelitian ke depan:

1. Perluasan dan variasi dataset dengan menambahkan lebih banyak citra pelatihan dari berbagai kondisi lingkungan, seperti cuaca berbeda, waktu siang dan malam, serta musim yang bervariasi untuk meningkatkan kemampuan generalisasi model.
2. Penerapan teknik augmentasi data, seperti rotasi, flipping, zooming, dan variasi brightness guna memperkaya data latih dan mencegah overfitting.
3. Eksperimen hyperparameter lebih lanjut, termasuk pengaturan ulang learning rate,

batch size, dan jumlah epoch, agar diperoleh konfigurasi pelatihan terbaik yang meningkatkan akurasi dan efisiensi waktu proses.

Dengan peningkatan di atas, diharapkan performa segmentasi Mask R-CNN terhadap objek citra udara semakin akurat dan aplikatif dalam skenario dunia nyata.

DAFTAR PUSTAKA

1. He, K., Gkioxari, G., Dollar, P., & Girshick, R. (2018). Mask R-CNN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2), 386-397.
2. Han, Q., Yin, Q., Zheng, X., & Chen, Z. (2021). Remote sensing image building detection method based on Mask R-CNN. *Complex & Intelligent Systems*, 7(3), 1845-1855.
3. Wang, S., Sun, G., Zheng, B., & Du, Y. (2021). A Crop Image Segmentation and Extraction Algorithm Based on Mask R-CNN. *Entropy*, 23(10), 10-13.
4. Zhang, Q., Chang, X., & Bian, S. (2019). Vehicle-Damage-Detection Segmentation Algorithm Based on Improved Mask R-CNN. *IEEE Access*, 7, 6997-7001.
5. Chiao, J.-Y., Chen, K.-Y., Liao, K. Y.-K., Hsieh, P.-H., Zhang, G., & Huang, T.-C. (2019). Detection and classification the breast tumors using mask R-CNN on sonograms. *Medicine*, 98(18), e15345.