

Analisis Keandalan Ekstraksi Garis Tepi Bangunan dari Data Foto Udara Menggunakan Pendekatan *Deep Learning* Berbasis *Mask R-CNN*

Analysis of the Reliability of Building Footprint Extraction Using a Deep Learning Approach Based on Mask R-CNN from Aerial Photos

Agri kristal*¹, Harintaka²

¹Magister Teknik Geomatika, Departemen Teknik Geodesi, Fakultas Teknik, Universitas Gadjah Mada, Yogyakarta

²Departemen Teknik Geodesi, Fakultas Teknik, Universitas Gadjah Mada, Yogyakarta

*Korespondensi penulis: agrikristalbig@gmail.com

Diterima: 16112021; Diperbaiki: 31012022; Disetujui: 04042022; Dipublikasi: 10042022

Abstrak: Kebutuhan peta dasar skala besar khususnya skala 1:5.000 terus meningkat dari waktu ke waktu. Pada umumnya ekstraksi fitur unsur Peta Rupa Bumi Indonesia (RBI) salah satunya adalah bangunan dilakukan dengan digitasi atau *stereoplotting* unsur secara manual baik dari data citra satelit maupun data foto udara. Namun hal itu memiliki kelemahan yaitu membutuhkan waktu yang lama tergantung pada kepadatan dan jumlah bangunan pada area yang akan dipetakan. Di sisi lain pemerintah Indonesia menjadikan percepatan penyelenggaraan Peta RBI skala 1:5000 menjadi salah satu prioritas utama dalam kegiatan kebijakan satu peta. Deteksi dan ekstraksi garis tepi bangunan secara otomatis menggunakan teknologi *computer vision* dari citra optis telah populer beberapa tahun terakhir. Salah satu teknologi yang dikembangkan adalah dengan pendekatan *deep learning*. Ekstraksi garis bangunan dengan pendekatan *deep learning* memiliki kelemahan yaitu poligon garis tepi bangunan yang dihasilkan tidak teratur. Penelitian ini bertujuan untuk mengetahui akurasi hasil regularisasi poligon dari ekstraksi garis tepi bangunan secara otomatis menggunakan metode *deep learning* berbasis *Mask Region-base Convolutional Neural Networks (Mask R-CNN)* dari data foto udara. Hasil penelitian menunjukkan pada area bangunan dengan kepadatan tinggi dan bentuk yang teratur (AoI 1) memiliki nilai indeks *Intersection over Union (IoU)* sebesar 87,8% sedangkan pada area dengan kepadatan tinggi dengan bentuk atap bangunan yang tidak teratur (AoI 2) memiliki nilai indeks *IoU* sebesar 82,6%. Kemudian juga dilakukan perhitungan akurasi posisi pada 25 sampel titik sudut bangunan dengan hasil CE90 pada AoI 1 sebesar 1,183 m dan pada AoI 2 sebesar 1,303 m. Secara geometri data garis tepi bangunan hasil regularisasi tersebut sudah dapat digunakan sebagai unsur bangunan Peta RBI skala 1:5.000 karena dari hasil perhitungan uji ketelitian geometrik horisontal (CE90) menunjukkan bahwa data tersebut masuk dalam ketelitian Peta RBI 1:5.000 kelas satu.

Copyright © 2022 Geoid. All rights reserved.

Abstract: The need of large-scale base map, especially in 1:5,000, is increasing in Indonesia. Furthermore, as the Government of Indonesia has declared 1:5,000 RBI mapping acceleration as one of main priorities of One Map Policy implementation, the need of large-scale topographic map production is also rising. Generally, topographic map feature extraction, including building extraction, is conducted through digitization or manually through feature stereoplotting either from satellite imagery or aerial photography. However, this method is usually time-consuming especially for high building density area mapping. Detection and extraction of building footprint automatically using computer vision of optical imagery have been favoured in recent years due to the time effective process. One of the technologies that have been developed is deep learning approach. However, the building line resulted from deep learning has disadvantage, i.e., irregular building footprint. This study attempts to assess the accuracy of polygon regularization resulted from automatically extracted building footprint using Mask Region-base Convolutional Neural Networks (Mask R-CNN) from aerial photography. The study finds that in high building density area with regular roof shape (AoI 1), the intersection over union (IoU) index is 87.8%. Whereas in high building density area with irregular roof shape (AoI 2) has the IoU index of 82.6%. This study also assesses the positional accuracy of 25 building corner point samples and resulting CE90 of 1.183 m and 1.303 m in AoI 1 and AoI 2 respectively. The geometric horizontal accuracy is classified as the class 1 in accordance with 1:5,000 RBI map accuracy standard. Therefore, this study concludes that geometrically, the building line resulted from the regularization is appropriate as feature in 1:5,000 RBI.

Kata kunci: Ekstraksi garis tepi bangunan, *Mask R-CNN*, Foto udara, Indeks *IoU*, Ketelitian Peta RBI

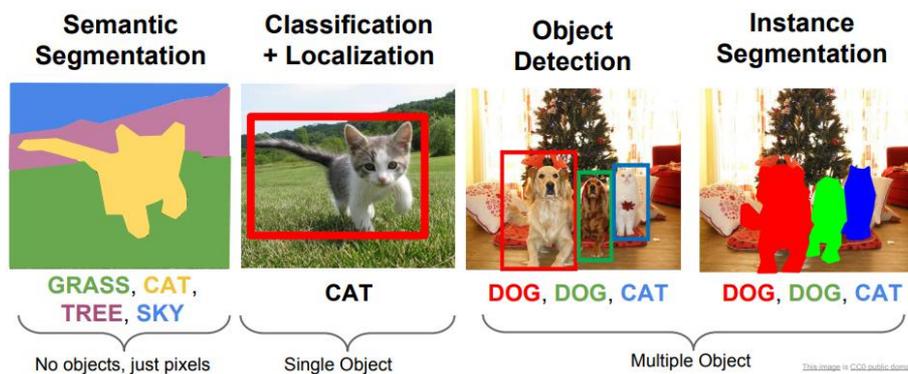
Cara untuk sitasi: Kristal, A., Harintaka. (2022). Analisis Keandalan Ekstraksi Garis Tepi Bangunan dari Data Foto Udara Menggunakan Pendekatan *Deep Learning* Berbasis *Mask R-CNN*. *Geoid*, 17(2), 273-285.

Pendahuluan

Kebutuhan peta dasar skala besar khususnya skala 1:5.000 terus meningkat dari waktu ke waktu. Kebutuhan peta tersebut terkait untuk penyusunan Rencana Detil Tata Ruang (RDTR), pemetaan batas desa, perencanaan pengembangan kota baru, reformasi agraria, dan lain-lain (Abidin, 2018). Pasal 12 Undang-Undang No. 4 Tahun 2011 tentang Informasi Geospasial menyebutkan bahwa peta dasar terdiri dari unsur garis pantai, hipsografi, perairan, nama rupabumi, batas wilayah, transportasi dan utilitas, bangunan dan fasilitas umum, dan penutup lahan. Salah satu elemen dalam unsur penutup lahan jenis kawasan permukiman dan unsur bangunan dan fasilitas umum adalah bangunan.

Pada umumnya ekstraksi fitur unsur peta dasar, salah satunya adalah bangunan, dilakukan dengan dijitasi atau *stereoplotting* unsur secara manual baik dari data citra satelit maupun data foto udara. Hasil dijitasi atau *stereoplotting* secara manual bergantung pada pengalaman dan pengetahuan operator dalam mengenali unsur peta dasar pada citra atau foto udara. Namun hasil interpretasi ini memberikan hasil deliniasi objek yang sesuai dengan kenampakan pada citra atau foto udara secara jelas sesuai dengan interpretasi dari masing-masing operator (Juniati, 2018). Ekstraksi fitur Peta RBI, salah satunya bangunan, yang dilakukan secara manual membutuhkan waktu yang lama serta sumber daya manusia yang banyak (Widyaningrum, 2021). Di sisi lain Pemerintah Indonesia menjadikan percepatan penyelenggaraan Peta RBI skala 1:5000 menjadi salah satu salah satu prioritas utama dalam kegiatan kebijakan satu peta.

Sejak akhir tahun 1990an dan awal tahun 2000an telah dikembangkan ekstraksi fitur bangunan secara otomatis. Ahmadi dkk., (2010); Alshehhi dkk., (2017); Mayer (1999); Qintao Hu dkk., (2021); Rottensteiner dkk., (2012); Zhao dkk., (2018) telah melakukan ekstraksi garis bangunan secara otomatis dari citra optis. Deteksi dan segmentasi fitur bangunan secara otomatis menggunakan teknologi *computer vision* dari citra optis dan data *point cloud* LiDAR telah populer beberapa tahun terakhir (Griffiths & Boehm, 2019). Salah satu teknologi yang dikembangkan adalah dengan pendekatan *deep learning*. *Deep learning* merupakan bidang penelitian yang efektif dalam teknik *machine learning* dan klasifikasi pola/pattern. *Instance segmentation* merupakan salah satu bidang *computer vision* yang menggabungkan antara deteksi objek dan segmentasi semantik (Gambar 1), sehingga dapat dikatakan bahwa *instance segmentation* adalah mengenai pemberian label, prediksi lokasi dan *segmentation mask* berbasis piksel terhadap setiap objek di dalam sebuah citra optis (Viana, dkk., 2020).



Gambar 1. Perbedaan Deteksi Objek, Segmentasi Semantik dan *Instance Segmentation* (Sumber : Li, Johnson, & Yeung, 2017)

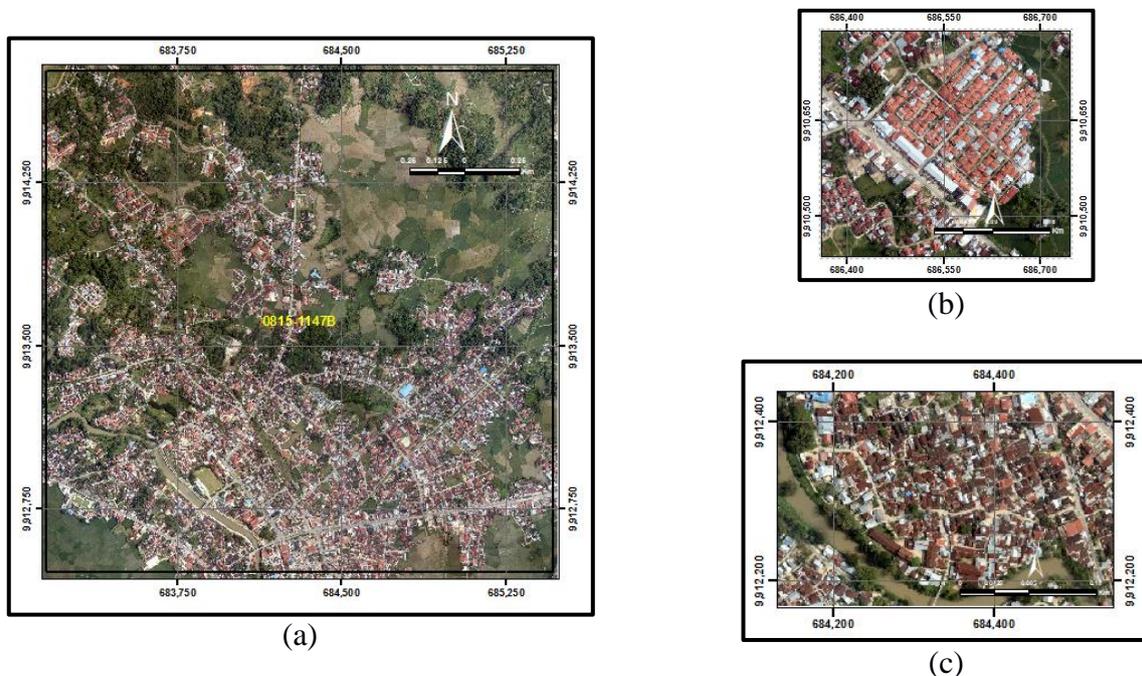
Keunggulan dari model *Mask R-CNN* yaitu output yang dihasilkan berupa label individu objek, *bounding box* dari setiap individu objek dan *segmentation mask* setiap individu objek dalam waktu yang singkat (He, dkk., 2017). Ekstraksi garis bangunan dengan pendekatan *deep learning* berbasis *Mask R-CNN* memiliki kelemahan

yaitu garis poligon bangunan yang dihasilkan tidak teratur. Dalam penelitian (Zhao dkk., 2018) menyatakan bahwa poligon bangunan yang dihasilkan dari *instance segmentation* memiliki bentuk yang tidak beraturan dan jauh berbeda dari batas tapak bangunan yang sebenarnya, sehingga tidak bisa digunakan secara langsung untuk kebutuhan teknis dan kartografi. Untuk mengatasi masalah tersebut maka harus dilakukan regularisasi terhadap poligon hasil dari *instance segmentation* agar poligon bangunan yang dihasilkan lebih teratur, salah satunya dengan menggunakan algoritma *polyline compression* (Gribov, 2015).

Ekstraksi garis bangunan secara otomatis menggunakan pendekatan *deep learning* berbasis *Mask R-CNN* merupakan salah satu solusi dalam mendapatkan garis bangunan yang lebih cepat jika dibandingkan dengan dijitasi manual. Penelitian ini mengkaji tentang kualitas regularisasi hasil ekstraksi fitur bangunan menggunakan pendekatan *deep learning berbasis Mask R-CNN* dari data foto udara. Hasil ekstraksi fitur bangunan ini kemudian dilakukan regularisasi agar poligon bangunan dari hasil *instance segmentation* dapat lebih teratur sehingga dapat digunakan untuk keperluan kartografi dan teknis. Evaluasi terhadap hasil regularisasi dilakukan dengan cara perhitungan indeks IoU dan uji akurasi CE90.

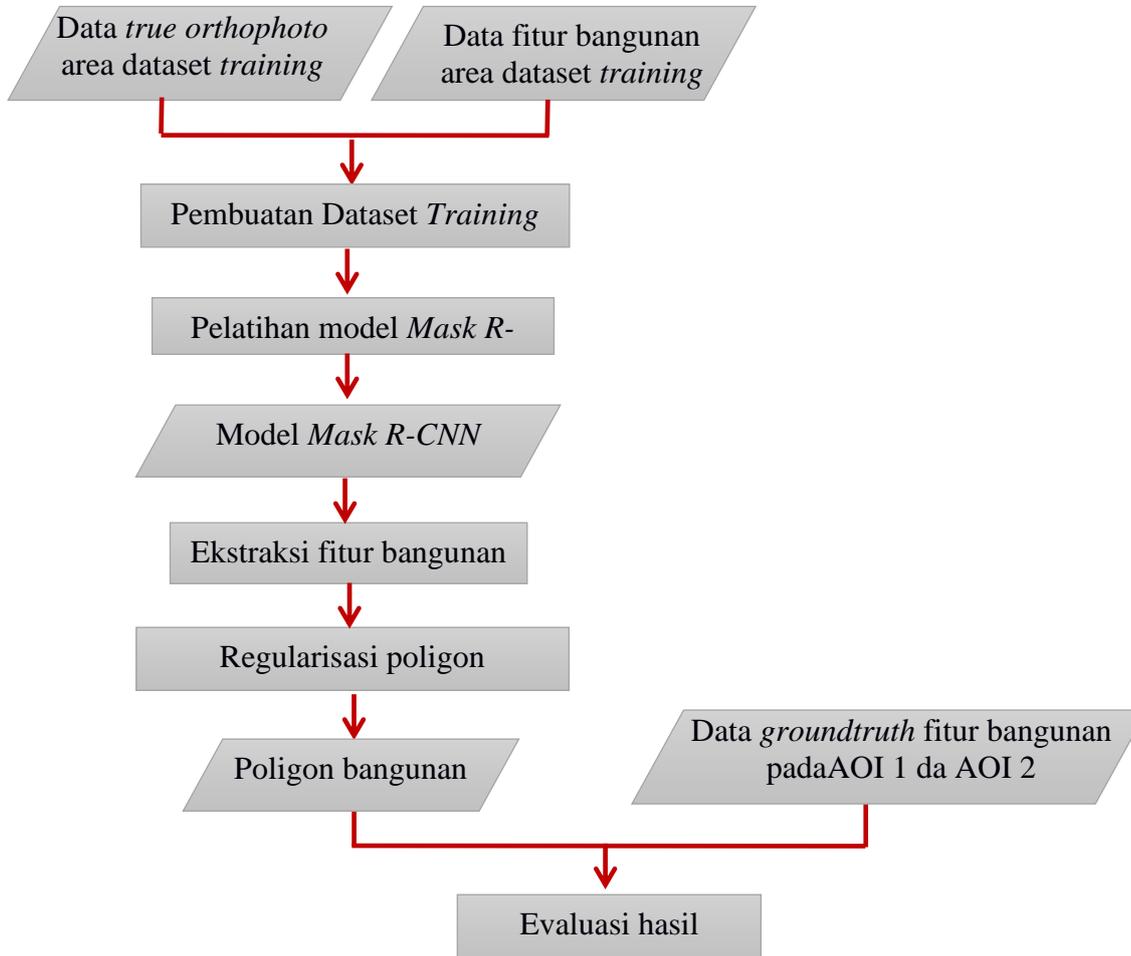
Data dan Metode

Data yang digunakan dalam penelitian ini adalah data *true orthophoto* dengan *Ground Sample Distance (GSD)* 15 cm, data vektor bangunan peta RBI skala 1: 5.000 merupakan data fitur bangunan yang merupakan hasil pengamatan stereoplotting foto udara digital pada tahun 2020. Data fitur bangunan Peta RBI tersebut digunakan sebagai *dataset training* dan juga digunakan sebagai data *groundtruth* untuk evaluasi hasil ekstraksi fitur bangunan. Data foto udara diakuisi pada tahun 2020 dan pembuatan Peta RBI dilakukan pada tahun yang sama yaitu 2020. Lokasi penelitian terletak di sebagian Kota Solok, Provinsi Sumatera Barat. *Dataset training* yang digunakan merupakan data foto udara dan fitur bangunan peta RBI skala 1:5.000 Nomor Lembar Peta (NLP) 0815-1147B. Evaluasi regularisasi hasil ekstraksi bangunan dilakukan terhadap 2 lokasi dengan karakteristik bangunan yang berbeda. *Area of Interest (AoI)* 1 merupakan kawasan permukiman dengan kerapatan bangunan yang tinggi dan sebagian besar bentuk atap bangunan teratur, sedangkan AoI 2 merupakan kawasan permukiman dengan kerapatan bangunan yang tinggi dan sebagian besar bentuk atap bangunan tidak teratur. Gambaran lokasi penelitian dapat dilihat pada Gambar 2.



Gambar 2. Area penelitian (a) *dataset training*, (b) AoI 1, (c) AoI 2 (Sumber: Badan Informasi Geospasial, 2020)

Tahapan penelitian meliputi persiapan, pengumpulan data, pemotongan data sesuai lokasi penelitian, pembuatan dataset *training*, pelatihan model *deep learning*, ekstraksi garis tepi bangunan pada AoI 1 dan AoI 2, regularisasi garis bangunan hasil ekstraksi, dan evaluasi kualitas hasil regularisasi (Gambar 3). Tahap pembuatan dataset *training*, pelatihan model *deep learning*, ekstraksi garis tepi bangunan, dan regularisasi garis tepi bangunan hasil ekstraksi dilakukan menggunakan perangkat lunak *ArcGIS Pro* versi 2.7.

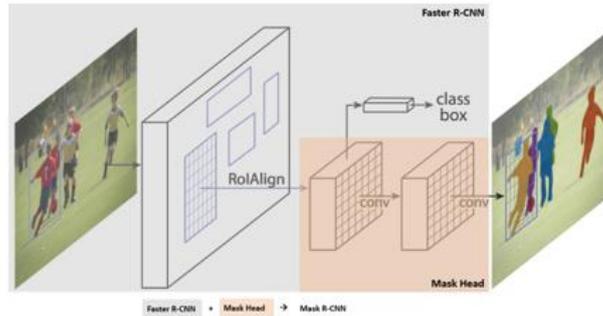


Gambar 3. Diagram Alir Penelitian

Dataset training merupakan kumpulan data yang terdiri dari foto udara dan data vektor bangunan yang akan digunakan sebagai *input* proses pelatihan model *deep learning*. Pembuatan *dataset training* adalah proses mengkonversi data vektor atau data raster berlabel menjadi set data pelatihan *deep learning*. Dalam penelitian ini *dataset training* yang digunakan adalah data vektor bangunan dan data *true orthophoto* pada NLP 0815-1147B. Ukuran *tile* yang digunakan pada penelitian ini adalah 512x512 piksel. Nilai *stride* yang digunakan adalah 256x256 piksel, maka *chip* yang dihasilkan memiliki tampilan 50%. Pelatihan model *deep learning* bertujuan untuk mendapatkan model jaringan terlatih menggunakan dataset *training* tahap sebelumnya dengan pendekatan *deep learning* berbasis *mask R-CNN*.

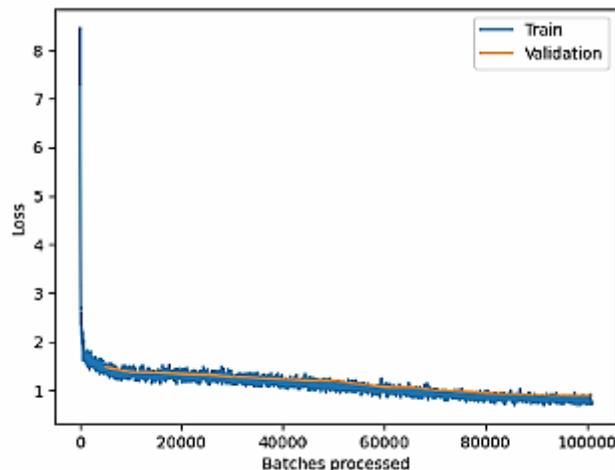
Mask R-CNN merupakan penyempurnaan dari model *Faster R-CNN* dimana output yang dihasilkan berupa label individu objek, *bounding box* dari setiap individu objek dan *segmentation mask* setiap individu objek (He, dkk, 2017). Arsitektur *network* dari model *Mask R-CNN* dibedakan menjadi dua garis besar (Gambar 4), yaitu:

1. *The convolutional backbone architecture* yang digunakan untuk ekstraksi fitur di seluruh image/citra.
2. *The network head* untuk pengenalan *bounding box* (klasifikasi dan regresi) dan prediksi *segmentation mask* pada setiap *Region of Interest (RoI)*.



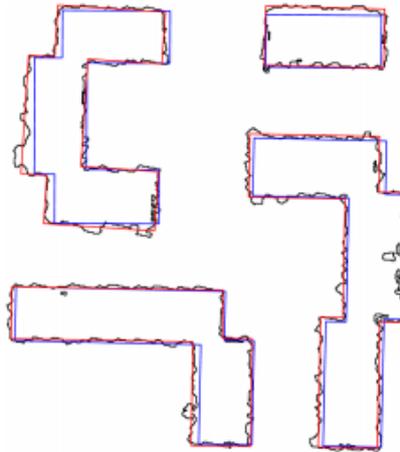
Gambar 4. Framework Model Mask R-CNN (Sumber : He, dkk., 2020)

Model *backbone* yang digunakan dalam penelitian ini adalah Resnet-50. Dalam penelitian ini digunakan 25 *epoch* untuk melatih jaringan, yang artinya dataset akan diproses *forward* dan *backward* melalui model jaringan sebanyak 25 kali dengan memberlakukan *early stopping* untuk menghindari *overfitting* pada model yang dihasilkan. Nilai validasi yang digunakan adalah 10%, yang artinya 90% dari data sampel digunakan untuk proses *training* dan 10% dari data sampel digunakan untuk memvalidasi hasil *training*. Pada *epoch* 25, nilai *training loss* yang dicapai adalah 1,987 dan nilai *validation loss* sebesar 1,157, sedangkan nilai presisi rata-rata adalah 81,3%. Nilai kesalahan internal pada model selama proses pelatihan ditunjukkan oleh nilai *training loss*. *Validation loss* merupakan nilai kesalahan model *deep learning* pada saat melakukan proses terhadap data yang tidak dikenal atau data baru ketika melakukan pelatihan model. Grafik *validation loss* dan *training loss* proses dapat dilihat pada Gambar 5. Dari Gambar 5 dapat dilihat bahwa semakin tinggi nilai *epoch* maka semakin rendah nilai *training loss* dan *validation loss* dalam pelatihan model *deep learning* yang dilakukan.



Gambar 5. Grafik *Training Loss* dan *Validation Loss* pada Proses Pelatihan Jaringan

Proses ekstraksi fitur bangunan merupakan proses untuk melakukan *instance segmentation* pada data *true orthophoto* dengan cara menjalankan model *deep learning* terlatih pada input raster untuk menghasilkan kelas fitur yang berisi objek bangunan. Parameter *threshold* yang digunakan dalam penelitian ini adalah 0,5; 0,6; 0,7; 0,8 dan 0,9. *Threshold* merupakan nilai ambang batas dari nilai *confidence* setiap garis bangunan hasil ekstraksi. Garis bangunan hasil ekstraksi dengan nilai *confidence* dibawah *threshold* secara otomatis akan dihapuskan oleh model. Poligon hasil ekstraksi menunjukkan garis yang tidak teratur sehingga diperlukan proses regularisasi poligon bangunan agar garis tepi bangunan lebih tegas dan lebih teratur. Menurut Gribov, (2015) aplikasi penyederhanaan *polyline* juga dapat dilakukan menggunakan algoritma *polyline compression*. Algoritma ini digunakan untuk memfilter kesalahan acak dari sebuah *polyline*. Cara kerja algoritma ini yaitu dengan membagi *polyline* sumber menjadi beberapa *cluster* dan setiap *cluster* tersebut didekati dengan garis lurus. Hasil yang diperoleh berupa garis tepi bangunan dengan garis yang lebih teratur dan lebih mendekati garis tepi bangunan aslinya (Gambar 6).



Gambar 6. Perbandingan Hasil Ekstraksi (Garis Hitam), Hasil Regularisasi (Garis Merah), dan Data *Ground Truth* (Garis Biru) (Sumber : Gribov, 2015)

Evaluasi kualitas hasil regularisasi dilakukan dengan dua cara yaitu indeks IoU dan akurasi posisi titik sudut bangunan terhadap AoI 1 dan AoI 2. Indeks *Intersection over Union (IoU)* yang juga dikenal sebagai *indeks Jaccard*, merupakan salah satu metode evaluasi paling populer yang digunakan dalam menghitung akurasi dalam hal deteksi objek (Rezatofighi, dkk. 2019). Indeks IoU merupakan perbandingan antara area yang saling tumpang tindih (*intersection*) antara data ground truth dan data hasil prediksi dibandingkan dengan total areanya (*union*). Formula untuk menghitung indeks dapat dilihat pada persamaan 1 (Carvalho, dkk. 2020).

$$IoU(A, B) = \frac{\text{Area of intersection } (A \cap B)}{\text{Area of union } (A \cup B)} \quad (1)$$

Dimana:

A: hasil prediksi melalui model pendekatan *deep learning*

B: *ground truth* (data referensi)

Peraturan Kepala Badan Informasi Geospasial (Perka BIG) No 6 tahun 2018 menyatakan bahwa ketelitian peta merupakan nilai yang menggambarkan tingkat kesesuaian posisi dan atribut sebuah objek di peta dengan posisi dan atribut obyek yang sama yang sebenarnya atau di lapangan. Salah satu ketelitian peta dasar adalah ketelitian geometri yang merupakan nilai yang menggambarkan ketidakpastian koordinat (vertikal dan horisontal) posisi objek yang sama pada peta terhadap posisi yang sebenarnya di lapangan.

Komponen ketelitian horisontal atau *Circular Error 90% (CE90)* didefinisikan sebagai radius lingkaran yang menunjukkan bahwa 90% kesalahan atau perbedaan posisi horisontal objek di peta dengan posisi horisontal objek yang dianggap sebenarnya tidak lebih dari 90%. Nilai RMSE dan CE90 dapat diperoleh menggunakan persamaan 2 dan 3 (Susetyo dkk., 2021 dan Badan Informasi Geospasial, 2018). Ketentuan standar ketelitian geometri Peta RBI yang dihasilkan tercantum pada Perka BIG No 6 tahun 2018.

$$RMSEr = \sqrt{\frac{(X_e - X_g)^2 + (Y_e - Y_g)^2}{n}} \quad (2)$$

$$CE90 = 1,5175 \times RMSEr \quad (3)$$

Dimana:

RMSEr = *Root Mean Square Error* posisi horisontal pada titik uji

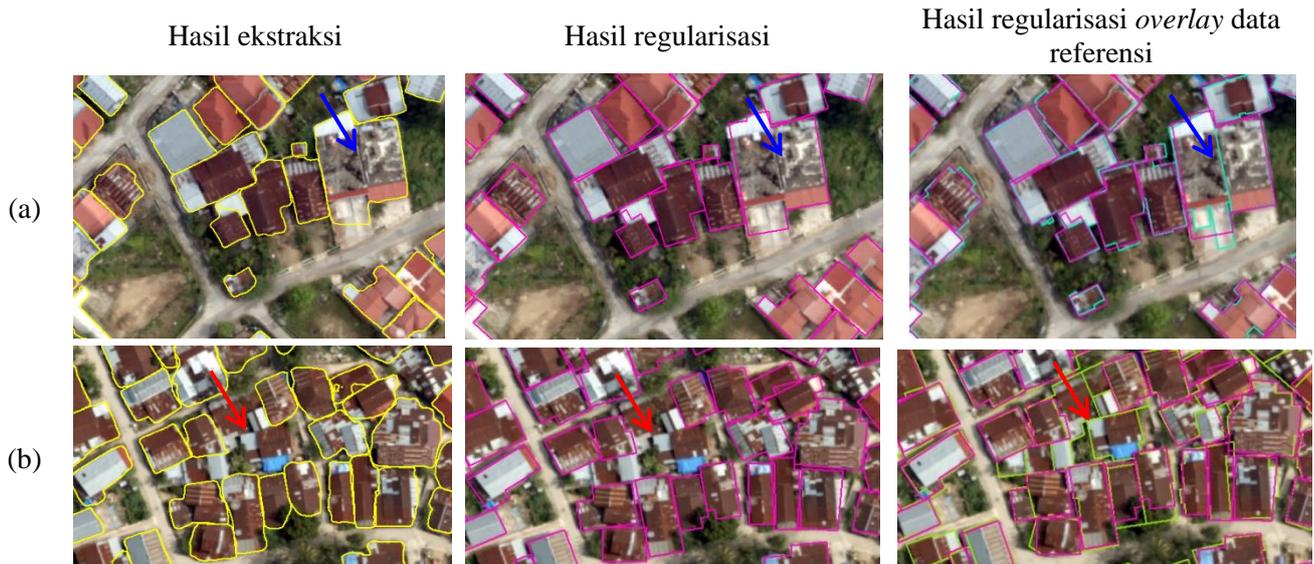
(X_e, Y_e) = koordinat titik pusat bangunan hasil ekstraksi

(X_g, Y_g) = koordinat titik pusat bangunan data referensi

n = jumlah titik uji

Hasil dan Pembahasan

Hasil ekstraksi garis bangunan dapat dikatakan sudah cukup baik secara visual. Sebagian besar bangunan dapat terdeteksi oleh model *deep learning* yang telah dilatih. Beberapa objek bangunan tidak terdeteksi pada wilayah yang memiliki bangunan dengan kerapatan yang sangat tinggi dan bentuk atap yang tidak teratur. Ekstraksi garis bangunan memiliki beberapa kelemahan seperti pada Gambar 7 (a) dan (b). Pada Gambar 7 (a) beberapa objek bangunan yang seharusnya terdeteksi menjadi beberapa poligon masih terdeteksi sebagai satu objek bangunan yang ditunjukkan oleh panah warna biru. Selain itu pada objek bangunan dengan kerapatan yang sangat tinggi terdapat bangunan yang tidak terdeteksi oleh model seperti pada Gambar 7(b) yang ditunjukkan oleh panah warna merah.



Gambar 7. Perbandingan Hasil Ekstraksi, Hasil Ekstraksi Garis Tepi Bangunan dan Data Referensi
(Sumber Citra Asli: Badan Informasi Geospasial, 2020)

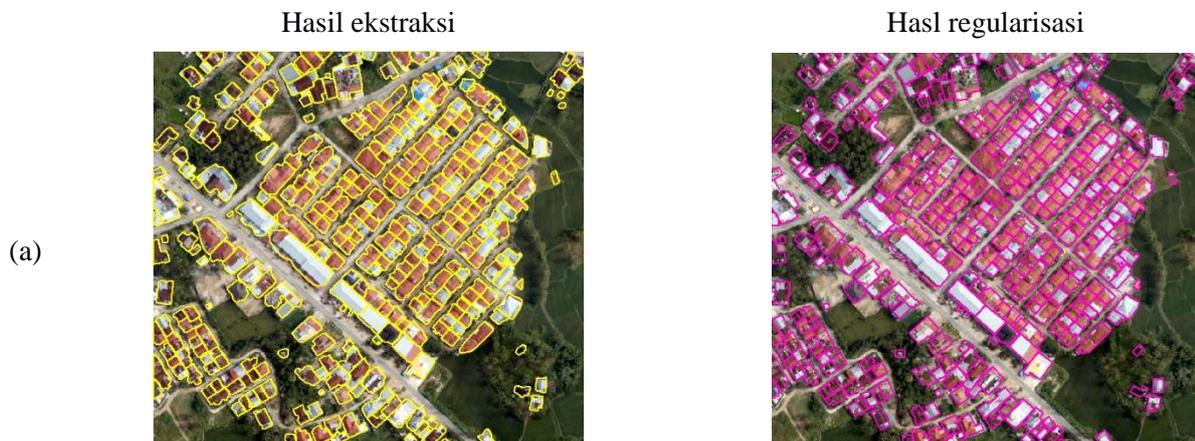
Gambar 8 menunjukkan perbedaan hasil ekstraksi bangunan pada setiap nilai *threshold*. Hasil ekstraksi menunjukkan bahwa semakin kecil nilai *threshold* maka semakin banyak bangunan yang terdeteksi oleh model. Gambar 8 (a) merupakan hasil ekstraksi dengan nilai *threshold* 0,6, dimana hasil ekstraksi menunjukkan bahwa terdapat bangunan yang tidak terdeteksi pada nilai *threshold* 0,7; 0,8 dan 0,9 namun terdeteksi oleh model dengan nilai *threshold* 0,6 yang ditunjukkan pada kotak merah. Hasil ekstraksi bangunan pada nilai *threshold* 0,6 dan 0,7 terdapat objek bukan bangunan yang terdeteksi sebagai bangunan yang ditunjukkan oleh tanda panah biru.





Gambar 8. Hasil Ekstraksi Pada Nilai *Threshold* (a) 0,6; (b) 0,7; (c) 0,8; (d) 0,9
(Sumber Citra Asli: Badan Informasi Geospasial, 2020)

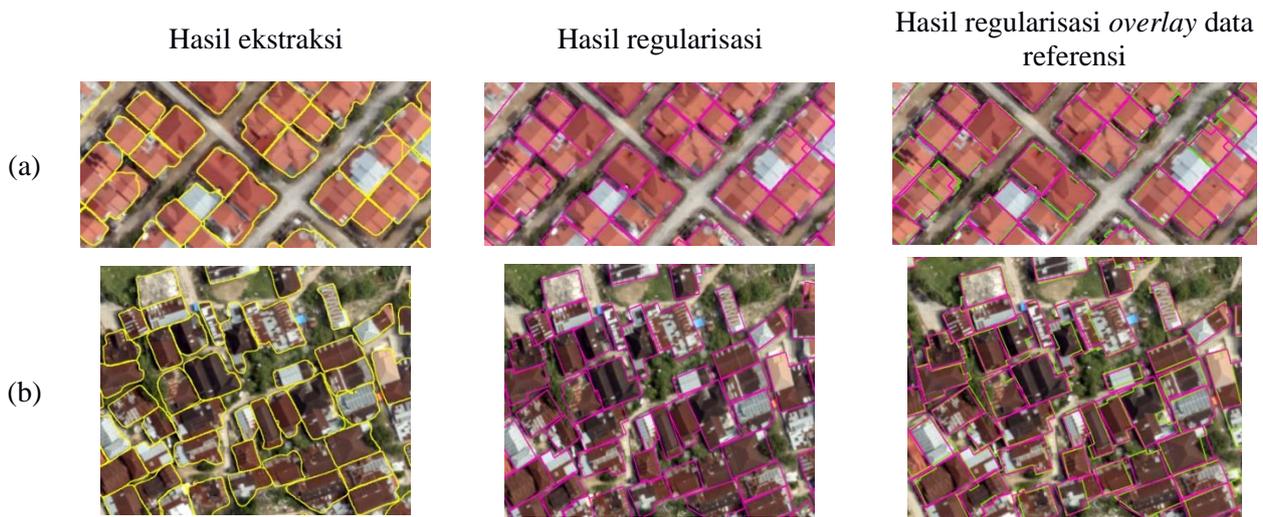
Hasil ekstraksi dari beberapa nilai *threshold* menunjukkan bahwa hasil ekstraksi dengan nilai *threshold* 0,8 menunjukkan hasil yang paling baik. Gambar 9. menunjukkan perbandingan hasil ekstraksi garis bangunan dengan nilai *threshold* yang digunakan 0,8 dan hasil regularisasi garis tepi bangunan pada AoI 1 dan AoI 2. Dapat dilihat pada Gambar 9 (b) dan (d) bahwa hasil regularisasi garis bangunan lebih teratur dan lebih tegas dibandingkan dengan Gambar 9 (a) dan (c) yang merupakan hasil ekstraksi garis bangunan yang belum dilakukan proses regularisasi.





Gambar 9. Perbandingan Hasil Ekstraksi dan Hasil Regularisasi Garis Tepi Bangunan pada AoI 1 (a) dan AoI 2 (b)
(Sumber Citra Asli: Badan Informasi Geospasial, 2020)

Model yang digunakan dapat mengekstraksi garis tepi bangunan pada area yang memiliki bentuk atap bangunan yang teratur dengan baik, sehingga hasil ekstraksi dan regularisasi nya cenderung mendekati garis bangunan pada data referensi (Gambar 10. a). Sedangkan untuk area dengan kerapatan bangunan tinggi dan bentuk yang tidak teratur, hasil ekstraksi garis bangunan dan regularisasi nya jauh berbeda bila dibandingkan dengan data referensi (Gambar 10. c). Dari Gambar 10., dapat dikatakan bahwa model *deep learning* kurang maksimal jika digunakan pada area dengan bangunan yang saling menempel dan bentuk atap yang tidak teratur.



Gambar 10. Perbandingan Hasil Ekstraksi, Hasil Regularisasi Garis Tepi Bangunan dan Data Referensi
(Sumber Citra Asli: Badan Informasi Geospasial, 2020)

Untuk atap bangunan dengan bentuk sederhana dan tidak terlalu kompleks menunjukkan hasil cukup baik seperti ditunjukkan pada Gambar 11. (a) dan (b). Sedangkan untuk bangunan dengan bentuk atap kompleks menunjukkan hasil ekstraksi dan regularisasi yang kurang baik seperti ditunjukkan pada Gambar 11. (c). Walaupun hasil ekstraksi menunjukkan hasil yang kurang maksimal pada bangunan dengan bentuk yang kompleks, namun dapat dilihat hasil ekstraksi garis tepi bangunan dapat mengikuti garis batas bangunan pada *image* dengan detail.





Gambar 11. Perbandingan Hasil Ekstraksi, Hasil Ekstraksi Garis Tepi Bangunan dan Data Referensi (Sumber Citra Asli: Badan Informasi Geospasial, 2020)

Selain evaluasi dari segi visual, evaluasi regularisasi garis tepi bangunan hasil ekstraksi juga dilakukan dengan cara menghitung indeks IoU. Pengujian ini dilakukan dengan melakukan perbandingan antara area yang saling tumpang tindih (*intersection*) antara data *ground truth* dan data hasil ekstraksi dibandingkan dengan total areanya (*union*). Semakin besar nilai indeks IoU maka hasil ekstraksi bangunan semakin akurat. Perhitungan indeks IoU dilakukan di setiap AOI untuk mengetahui kualitas hasil ekstraksi pada area dengan karakteristik bangunan setiap AOI. Tabel 1 menunjukkan hasil perhitungan indeks IoU pada setiap AOI.

Tabel 1. Hasil Perhitungan Indeks IoU pada Setiap AoI

	AoI 1	AoI 2
<i>Intersect</i> (m ²)	48922.015	52029.810
<i>Union</i> (m ²)	55671.932	62963.371
IoU	87.8%	82.6%

Dari Tabel 1. menunjukkan bahwa nilai indeks IoU pada AoI 1 (area yang memiliki kerapatan bangunan yang tinggi dan bentuk atap bangunan teratur) lebih tinggi 5,2% dibandingkan dengan AoI 2 (area yang memiliki kerapatan bangunan yang tinggi dan bentuk atap bangunan tidak teratur). Hal ini menunjukkan bahwa model *deep learning* pada penelitian ini dapat mendeteksi bangunan lebih baik pada area dengan bentuk atap bangunan yang teratur dibandingkan pada area dengan bentuk bangunan yang tidak teratur.



Gambar 12. Sebaran Titik Uji (Sumber Citra Asli: Badan Informasi Geospasial, 2020)

Metode evaluasi selanjutnya adalah menghitung ketelitian geometri horisontal regularisasi garis tepi bangunan hasil ekstraksi. Titik uji yang digunakan untuk menghitung ketelitian geometri horisontal adalah titik sudut bangunan dan merupakan objek yang sama antara data hasil regularisasi dengan data *groundtruth*. Titik sudut bangunan yang digunakan pada setiap AoI berjumlah 25 titik dengan persebaran yang merata pada setiap AoI. Dalam hal ini data *groundtruth* atau data yang dianggap benar adalah data vektor bangunan Peta RBI hasil *stereoplotting* manual. Sebaran titik uji harus merata pada seluruh AoI (Gambar 12.).

Tabel 2. Hasil perhitungan ketelitian geometri horisontal pada AoI 1 dan AoI 2

AoI	CE 90 (m)
AoI 1	1,183
AoI 2	1,303

Hasil uji akurasi titik pusat bangunan pada setiap AoI ini dibandingkan dengan standar ketelitian geometri horisontal peta RBI dalam Perka BIG nomor 6 tahun 2018 tentang Ketelitian Peta Dasar RBI. Nilai CE90 didefinisikan sebagai radius lingkaran yang menunjukkan 90% kesalahan atau perbedaan posisi horisontal objek dengan posisi yang dianggap benar, tidak lebih dari radius tersebut. Dalam penelitian ini menggunakan Peta RBI skala 1:5.000 sehingga standar ketelitian peta dasar yang digunakan adalah kelas 1 (CE90 = 1,5 m), kelas 2 (CE90 = 3 m) dan kelas 3 (CE90 = 4,5 m) seperti yang tercantum pada Perka BIG No 6 tahun 2018. Dalam hal ini data titik sudut bangunan dari hasil *stereoplotting* dianggap sebagai titik referensi dibandingkan dengan posisi titik sudut bangunan hasil regularisasi. Dari Tabel 2. dapat dilihat bahwa nilai CE90 setiap AoI mempunyai nilai kurang dari 1,5 m sehingga setiap AOI masuk ke dalam kelas 1 ketelitian Peta RBI.

Evaluasi hasil penelitian juga dilakukan dalam hal aspek kecepatan proses ekstraksi dan regularisasi garis tepi bangunan. Evaluasi ini dilakukan dengan membandingkan waktu pemrosesan untuk melakukan ekstraksi garis tepi bangunan secara otomatis dibandingkan dengan proses dijitasi secara manual. Kapasitas pengerjaan digitasi interaktif pada foto udara dalam 1 NLP skala 1 : 5.000 berukuran 5,38 km², memerlukan waktu 9 hari kerja pada area dengan kerapatan bangunan yang sangat tinggi (Badan Informasi Geospasial, 2017)). Jadi untuk area dengan luas 0,245 km² dengan kerapatan bangunan yang sangat tinggi memerlukan waktu 9 jam dalam melakukan dijitasi manual. Pada perhitungan waktu proses ekstraksi garis tepi bangunan secara otomatis, proses pelatihan jaringan tidak dimasukkan dan dianggap data model jaringan telah siap digunakan untuk ekstraksi. Berdasarkan model yang diperoleh, waktu yang diperlukan untuk mendapatkan hasil regularisasi garis tepi bangunan pada AoI 1 dan AoI 2 adalah 23 menit 27 detik (Tabel 3). Waktu ini dapat dicapai menggunakan spesifikasi perangkat keras PC dengan spesifikasi *processor AMD Ryzen Threadripper 2950X-Core Processor 3.5 GHz, RAM 32 GB, GPU Nvidia RTX 2080 Ti* dengan memori 11GB.

Tabel 3. Waktu yang dibutuhkan pada setiap tahapan penelitian

Tahap	Luas Area (km ²)	Waktu
Ekstraksi garis tepi bangunan pada AoI 1	0,134	7 menit 5 detik
Ekstraksi garis tepi bangunan pada AoI 2	0,111	14 menit 59 detik
Regularisasi hasil ekstraksi bangunan pada AoI 1	0,134	39 detik
Regularisasi hasil ekstraksi bangunan pada AoI 2	0,111	44 detik

Kesimpulan

Ekstraksi garis tepi bangunan menggunakan pendekatan *deep learning* berbasis *mask R-CNN* memiliki nilai indeks IoU sebesar 87,8% pada area dengan kerapatan bangunan yang tinggi dengan bentuk atap bangunan teratur, sedangkan pada area dengan kerapatan bangunan yang tinggi dengan bentuk atap bangunan tidak teratur mencapai 82,6%. Perhitungan CE90 menunjukkan bahwa data vektor bangunan hasil regularisasi pada AoI 1 sebesar 1,183 m dan pada AoI 2 sebesar 1,303 m. Selain itu proses ekstraksi garis bangunan secara otomatis menggunakan pendekatan *deep learning* berbasis *mask R-CNN* ini lebih cepat bila dibandingkan dengan dijitasi manual, walaupun tidak seakurat hasil dijitasi manual. Hal ini ditunjukkan pada beberapa objek bangunan yang seharusnya terdeteksi menjadi beberapa poligon masih terdeteksi sebagai satu objek bangunan dan pada objek bangunan dengan kerapatan yang sangat tinggi terdapat bangunan yang tidak terdeteksi. Penelitian lebih lanjut mengenai penentuan parameter pada model *deep learning* yang digunakan perlu

dilakukan untuk mengetahui nilai parameter yang tepat agar dapat meningkatkan kualitas hasil ekstraksi garis tepi bangunan.

Ucapan Terimakasih

Ucapan terimakasih disampaikan kepada Kementerian Riset dan Teknologi yang telah memberikan bantuan program beasiswa saintek dan dana penelitian. Ucapan terimakasih juga disampaikan kepada Pusat Pemetaan Rupabumi dan Toponim, Badan Informasi Geospasial yang telah memberikan bantuan data dan dokumen yang dibutuhkan dalam penelitian ini.

Daftar Pustaka

- Abidin, H. Z. (2018). Kebijakan Satu Peta dan Percepatan Peta Dasar Skala Besar. *Seminar Nasional Geomatika 2018*, (Bogor). 5 September. <http://dx.doi.org/10.13140/RG.2.2.35097.06240>
- Ahmadi, S., Zoej, M. J. V., Ebadi, H., Moghaddam, H. A., & Mohammadzadeh, A. (2010). Automatic urban building boundary extraction from high resolution aerial images using an innovative model of active contours. *International Journal of Applied Earth Observation and Geoinformation*, 12(3), 150–157. <https://doi.org/10.1016/j.jag.2010.02.001>
- Alshehhi, R., Marpu, P. R., Woon, W. L., & Mura, M. D. (2017). Simultaneous extraction of roads and buildings in remote sensing imagery with convolutional neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 130, 139–149. <https://doi.org/10.1016/j.isprsjprs.2017.05.002>
- Badan Informasi Geospasial. (2018). Peraturan Badan Informasi Geospasial Nomor 6 Tahun 2018 Tentang Perubahan Atas Peraturan Kepala Badan Informasi Geospasial Nomor 15 Tahun 2014 Tentang Pedoman Teknis Ketelitian Peta Dasar. *Cibinong, Bogor*, 2014–2016. Retrieved from <https://jdih.big.go.id/lihatdoc?id=27330968>
- Badan Informasi Geospasial. (2017). *Analisis Biaya Keluaran Kegiatan Penyelenggaraan Informasi Geospasial*. (2017).
- Gribov, A. (2015). Searching for a Compressed Polyline with a Minimum Number of Vertices (Discrete Solution). *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11009 LNCS, 54–68. https://doi.org/10.1007/978-3-030-02284-6_5
- Griffiths, D., & Boehm, J. (2019). Improving public data for building segmentation from Convolutional Neural Networks (CNNs) for fused airborne lidar and image data using active contours. *ISPRS Journal of Photogrammetry and Remote Sensing*, 154, 70–83. <https://doi.org/10.1016/j.isprsjprs.2019.05.013>
- He, K., Girshick, R., Gkioxari, G., & Doll, P. (2020). Mask R-CNN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2), 386–397. <https://doi.org/10.1109/TPAMI.2018.2844175>
- He, K., Gkioxari, G., Doll, P., Girshick, R., & Ai, F. (2017). Mask R-CNN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2), 386–397. <https://doi.org/https://arxiv.org/pdf/1703.06870.pdf>
- Juniati, E. H. (2018). *2D Semantic Labeling Penutup Lahan di Area Urban dengan Analisis Berbasis Objek Dari Foto Udara dan LiDAR*. (July 2018). <https://doi.org/10.13140/RG.2.2.24118.32320>
- Li, F., Johnson, J., & Yeung, S. (2017). *Detection and Segmentation*. CS231n: Convolutional Neural Networks for Visual Recognition, Stanford University.
- Mayer, H. (1999). Automatic object extraction from aerial imagery - a survey focusing on buildings. *Computer Vision and Image Understanding*, 74(2), 138–149. <https://doi.org/10.1006/cviu.1999.0750>
- Qintao Hu, Liangli Zhen, Yao Mao, Xi Zhou, G. Z. (2021). *Automated building extraction using satellite remote sensing imagery*. <https://doi.org/https://www.sciencedirect.com/science/article/pii/S092658052031089X>
- Rottensteiner, F., Sohn, G., Jung, J., Gerke, M., Baillard, C., Benitez, S., & Breitkopf, U. (2012). THE ISPRS BENCHMARK on URBAN OBJECT CLASSIFICATION and 3D BUILDING RECONSTRUCTION. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1(September), 293–298. <https://doi.org/10.5194/isprsannals-I-3-293-2012>
- Susetyo, D. B., Rizaldy, A., Hariyono, M. I., Purwono, N., Hidayat, F., Windiastuti, R., ... Hartanto, P. (2021). A Simple But Effective Approach of Building Footprint Extraction in Topographic Mapping Acceleration. *Indonesian Journal on Geoscience*, 8(3), 329–343. <https://doi.org/10.17014/ijog.8.3.329-343>
- Viana, F. X., Araujo, G. M., Pinto, M. F., Colares, J., & haddad, D. B. (2020). Aerial Image Instance Segmentation Through Synthetic Data Using Deep Learning. *Learning and Nonlinear Models*, 18(1), 35–46.

<https://doi.org/10.21528/Inlm-vol18-no1-art3>

Widyaningrum, E. (2021). Automatic Object Extraction from Airborne Laser Scanning Point Clouds for Digital Base Map Production. In *TU Delft University*. <https://doi.org/10.4233/uuid>

Zhao, K., Kang, J., Jung, J., & Sohn, G. (2018). Building extraction from satellite images using mask R-CNN with building boundary regularization. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2018-June*, 242–246. <https://doi.org/10.1109/CVPRW.2018.00045>



This article is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/)