e-ISSN 2621-9883
Elipsoida : Jurnal Geodesi dan Geomatika
Vol 08 No 01, (2025)

EKSTRAKSI OTOMATIS TAPAK BANGUNAN (*BUILDING FOOTPRINT*) PADA ORTOFOTO MENGGUNAKAN *SEGMENT ANYTHING MODEL* (SAM)

Annisa Baroroh, Harintaka*
Departemen Teknik Geodesi, Fakultas Teknik, Universitas Gadjah Mada.
Jl. Grafika Bulaksumur No.2, Mlati, Sleman, DIY-55281
e-mail: harintaka@ugm.ac.id*

(Diterima 22 Maret 2025, Disetujui 19 Juni 2025)

ABSTRAK

Kebutuhan peta dasar skala besar khususnya di Indonesia semakin meningkat seiring berkembangnya waktu. Salah satu fitur geospasial penting adalah bangunan. Penggunaan teknologi UAV mampu menghasilkan produk ortofoto resolusi tinggi yang dapat membantu pemetaan skala besar. Kumpulan data bangunan yang diturunkan dari data ortofoto dapat memberikan informasi untuk pemantauan dan perencanaan suatu daerah, khususnya di daerah yang tidak memiliki peta perencanaan rinci atau data kadaster. Pada umumnya, ekstraksi objek tapak bangunan dilakukan dengan digitasi manual. Namun, perkembangan terbaru menunjukkan teknologi otomatisasi dengan deep learning memiliki keunggulan lebih baik dari sisi kinerja yang lebih singkat. Teknik *deep learning* yang terbaru saat ini adalah SAM (*Segment Anything Model*). SAM merupakan pendekatan baru yang dikembangkan oleh Meta AI untuk segmentasi yang telah dilatih pada dataset sangat besar sehingga tidak memerlukan pelatihan ulang (Kirillov et al., 2023). Penelitian ini memanfaatkan SAM untuk ekstraksi tapak bangunan khususnya wilayah Indonesia. Karakteristik bangunan yang sangat bervariasi menjadi tantangan algoritma SAM dalam mengekstraksi tapak bangunan. Selain SAM, penggunaan metode regularisasi diterapkan untuk memperbaiki bentuk bangunan hasil segmentasi yang tidak teratur dan tegas. Hasil uji akurasi *precision*, *recall*, *f1-score*, dan IoU secara keseluruhan menunjukkan rata-rata nilainya diatas 87 %. Hasil tersebut menunjukkan bahwa SAM mampu melakukan ekstraksi tapak bangunan dengan baik.

Kata kunci : Deep learning, SAM, ortofoto, tapak bangunan, regularisasi

ABSTRACT

The demand for large-scale base maps, particularly in Indonesia, is increasing over time. One of the essential geospatial features is buildings. The use of UAV technology can produce high-resolution orthophoto products that assist in large-scale mapping. Building datasets derived from orthophoto data can provide valuable information for monitoring and planning an area, especially in regions lacking detailed planning maps or cadastral data. Typically, building footprint extraction is done through manual digitization. However, recent advancements indicate that automation technology using deep learning offers better performance in terms of speed. The new deep learning technique is SAM (Segment Anything Model). SAM is a new approach developed by Meta AI for segmentation, trained on a very large dataset, thus eliminating the need for retraining (Kirillov et al., 2023). This research leverages SAM for building footprint extraction, specifically in Indonesian regions. The highly variable characteristics of buildings pose a challenge for SAM's algorithm in extracting building footprints. In addition to SAM, regularization methods are applied to improve the shape of the segmented buildings, making them more structured and defined. The accuracy test results for precision, recall, F1-score, and IoU overall indicate an average value above 87%. These results demonstrate that SAM is capable of effectively extracting building footprints.

Keywords : Deep learning, Segment Anything Model, orthophoto, building footprint, regularization

1. PENDAHULUAN

Teknologi penginderaan jauh masih menjadi pilihan yang banyak digunakan untuk keperluan pemetaan skala besar. Kemampuan akuisisi pada area yang luas tanpa perlu melakukan survei langsung di lapangan merupakan keunggulan teknologi penginderaan jauh (Elachi & Van Zyl, 2021). Penginderaan jauh dapat dimanfaatkan untuk pengumpulan data dan informasi geospasial dalam pembuatan peta dasar skala besar. Peta dasar skala besar mencakup seluruh informasi terkait objek alami maupun buatan yang dapat mendukung pembangunan nasional.

Salah satu teknologi penginderaan jauh yang banyak digunakan adalah UAV (*Unmanned Aerial Vehicle*) atau pesawat udara tanpa awak. UAV memiliki beberapa keunggulan dibandingkan dengan *platform* penginderaan jauh lainnya seperti biaya lebih rendah, penggunaan yang mudah, dan akuisisi data yang fleksibel (Jiang et al., 2020). Teknologi UAV mampu menyediakan sistem akuisisi data spasial resolusi tinggi yang dapat menghasilkan data berupa *point clouds*, DSM (*Digital Surface Model*), dan ortofoto (Yao et al., 2019). Foto udara UAV merupakan data citra resolusi tinggi yang dapat memetakan kenampakan alam dan buatan dengan lebih detail. Dalam konteks pemetaan skala besar, salah satu objek penting yang perlu dipetakan adalah bangunan.

Bangunan bersifat dinamis dan terus bertambah sejalan dengan pertumbuhan penduduk yang begitu pesat, sehingga diperlukan informasi spasial terbaru mengenai bangunan dalam rentang waktu tertentu (Negara & Harintaka, 2021). Ekstraksi tapak bangunan dari data penginderaan sudah banyak berkembang dimulai dari era konvensional hingga otomatis. Pada awalnya, teknik ekstraksi tapak bangunan konvensional yang diperoleh melalui proses digitasi manual dinilai memakan waktu lama, mahal, dan bersifat subjektif (Guo et al., 2021).

Saat ini, teknik otomatis untuk ekstraksi tapak bangunan semakin berkembang pesat. Teknik otomatisasi yang berkembang saat ini menggunakan konsep pembelajaran terkomputerisasi yang dikenal dengan metode pembelajaran mendalam (*deep learning*) (Janiesch et al., 2023). Beberapa tahun terakhir, teknik *deep learning* berbasis algoritma *Convolutional Neural Network* (CNN) menunjukkan hasil yang baik dalam melakukan ekstraksi otomatis tapak bangunan karena mampu menangani sejumlah data besar yang kompleks (Abdollahi et al., 2020; Rastogi et al., 2022; dan Yu et al., 2023). Namun, kinerja CNN dan model *deep learning* lainnya bergantung pada kualitas data pelatihan dan parameter pembelajaran agar mencapai akurasi model yang optimal (Farajzadeh et al., 2023).

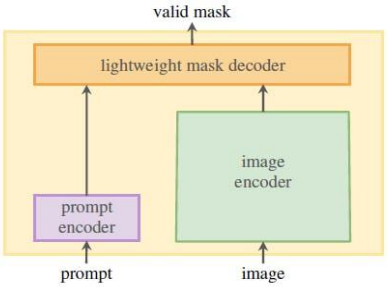
Perkembangan *deep learning* mendorong munculnya model baru yang dapat melakukan segmentasi dengan lebih cepat yang disebut dengan SAM (*Segment Anything Model*). SAM merupakan model dasar untuk segmentasi yang telah dilatih pada dataset besar yang terdiri dari satu miliar *mask* objek dan 11 juta gambar (Kirillov et al., 2023). SAM merupakan pendekatan baru yang dikembangkan oleh tim Meta AI yang menunjukkan kemampuan generalisasi *zero-shot* di berbagai macam dataset citra tanpa memerlukan data pelatihan tambahan (Gui et al., 2024). SAM menunjukkan kemampuan adaptasi yang baik terhadap analisis data penginderaan jauh seperti ekstraksi tapak bangunan (Ji et al., 2023; Osco et al., 2023). SAM untuk segmentasi data geospasial dapat diterapkan pada data penginderaan jauh sehingga mampu menghasilkan data dalam format vektor dan raster (Wu & Osco, 2023).

Ekstraksi tapak bangunan dari data foto udara (ortofoto) menggunakan algoritma SAM mampu menunjukkan generalisasi yang baik, namun memiliki kelemahan yaitu garis poligon yang tidak teratur. Untuk mengatasi kelemahan tersebut, diperlukan proses regularisasi agar garis poligon hasil segmentasi lebih tegas dan teratur. Kombinasi SAM dan regularisasi dapat memberikan hasil tapak bangunan yang lebih representatif. Ekstraksi tapak bangunan dengan SAM dan regularisasi menggunakan data ortofoto memiliki potensi besar untuk mencapai efektifitas dan efisiensi dalam percepatan pemetaan skala besar.

2. *SEGMENT ANYTHING MODEL* (SAM)

Segment Anything Model (SAM) adalah model dasar untuk segmentasi yang dikembangkan oleh tim peneliti Meta AI. SAM menunjukkan kemampuan generalisasi yang baik di berbagai macam dataset gambar karena dapat menghasilkan prediksi yang akurat tanpa proses pelatihan ulang (Osco et al., 2023). Kemampuan SAM tersebut dikenal dengan *zero-shot transfer*, artinya SAM dapat melakukan segmentasi objek dalam gambar/citra tanpa memerlukan data pelatihan tertentu yang berkaitan dengan objek tersebut. Hal ini bisa dicapai dengan menggabungkan antarmuka pengguna tunggal sebagai *prompt* (petunjuk) yang memungkinkan pengguna menentukan objek yang ingin disegmentasi (Ji et al., 2023). Kemampuan SAM didorong oleh visi model dasar yang telah dilatih pada kumpulan data besar SA-1B yang berisi

lebih dari 11 juta gambar dan satu miliar *mask* (Kirillov et al., 2023).



Gambar 1. Komponen algoritma SAM (Kirillov et al., 2023)

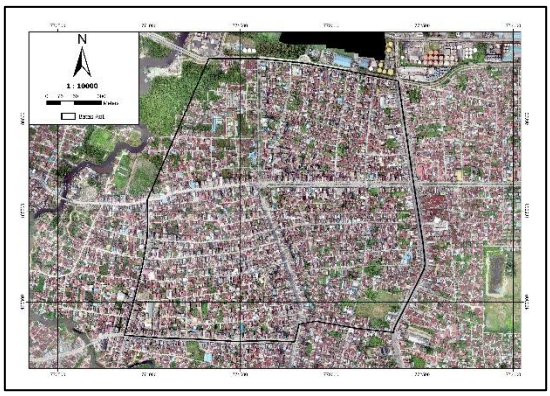
SAM terinspirasi dari konsep NLP (*Natural Language Processing*) di bidang *computer vision*, sehingga mampu melakukan pembelajaran *zero-shot* dan *few-shot* untuk kumpulan data dan tugas baru (Kirillov et al., 2023). SAM terdiri dari tiga komponen utama yang saling berhubungan antara lain *image encoder*, *prompt encoder*, dan *mask decoder*. *Image encoder* (pengkode gambar) bertanggung jawab untuk mengkodekan dan mengekstraksi fitur yang relevan dari gambar/citra *input*. *Image encoder* berbasis *Vision Transformer* (ViT) yang telah dilatih menggunakan metode yang disebut MAE (*Masked Autoencoder*), sehingga mampu memproses input citra resolusi tinggi (Li et al., 2023). *Prompt encoder* (pengkode petunjuk) bertugas mengkodekan *prompt* yang diberikan kepada model. Jenis *prompt* atau petunjuk dikelompokkan menjadi dua yaitu *sparse* (titik, kotak, teks) dan *dense* (*masks*) (Ren et al., 2023). *Mask decoder* adalah komponen yang menghasilkan topeng segmentasi secara cepat berdasarkan gambar dan *prompt* yang dikodekan. *Mask decoder* menggunakan konsep MLP (*Multi-Layer Perceptron*), artinya dapat memetakan keluaran dengan klasifikasi linear dinamis yang menghitung probabilitas *mask* di setiap lokasi dalam gambar.

3. METODOLOGI PENELITIAN

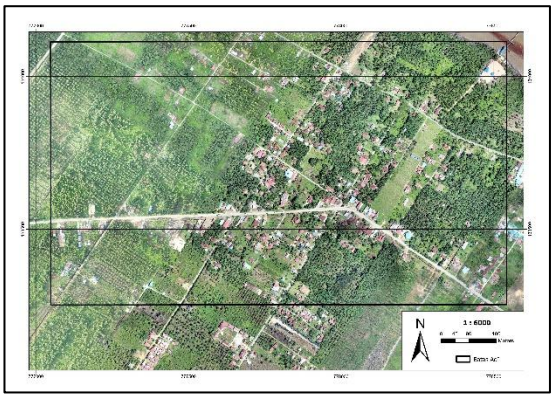
Berikut ini merupakan rincian terkait lokasi penelitian, data yang diperlukan, dan tahapan penelitian.

3.1 Lokasi dan Data

Penelitian dilakukan di Kota Dumai, Provinsi Riau. Secara geografis, Kota Dumai terletak di antara 101°23"37' - 101°8"13' Bujur Timur dan 1°23"23' - 1°24"23' Lintang Utara. Area yang diteliti merupakan permukiman dengan kerapatan bangunan tinggi dan rendah. AOI-1 mewakili area dengan kerapatan bangunan tinggi (padat), sedangkan AOI-2 mewakili area dengan kerapatan bangunan rendah (renggang). AOI-1 dan AOI-2 masing-masing memiliki luas sebesar 204,424 hektar dan 128,629 hektar.



Gambar 2. Lokasi penelitian AOI-1 (Sumber data: Ortofoto UAV dari PT. Hasanah Surveyor Raya, 2023)



Gambar 3. Lokasi penelitian AOI-2 (Sumber data: Ortofoto UAV dari PT. Hasanah Surveyor Raya, 2023)

Penelitian ini menggunakan dua data utama yaitu ortofoto dan model *deep learning* SAM. Data raster ortofoto merupakan hasil akuisisi *drone* pada bulan September-November tahun 2023 dalam format TIFF dengan resolusi spasial sebesar 10 cm. Model *deep learning* SAM diperoleh dari *website* ArcGIS Living Atlas of the World (https://www.arcgis.com/home/item.html?id=9b67b441f29f4ce6810979f5f0667ebe) dalam format DLPK. Model SAM yang disediakan oleh ESRI merupakan paket *segment-geospatial* (samgeo). Seluruh proses pengolahan data dilakukan menggunakan perangkat lunak ArcGIS Pro pada perangkat komputer dengan spesifikasi tinggi.

3.2 Metodologi

Penelitian ini menggunakan dua data utama yaitu *Segment Anything Model* (SAM) dalam format

DLPK dan ortofoto dalam format TIFF. Data raster ortofoto sebelumnya perlu dilakukan *pre-processing* untuk menyesuaikan spesifikasi yang dibutuhkan dalam proses segmentasi dengan SAM. Ortofoto terkoreksi merupakan ortofoto yang sudah dilakukan komposit band RGB, penurunan resolusi radiometrik dari 16-bit ke 8-bit, serta pemotongan (*clip*) wilayah sesuai dengan AOI yang sudah direncanakan.

3.2.1 Ekstraksi Tapak Bangunan dengan SAM dan Regularisasi

Proses segmentasi menggunakan perangkat lunak ArcGIS Pro dilakukan menggunakan *tool Detect Object Using Deep Learning*. Input utama yang diperlukan adalah ortofoto dan model *Segment Anything*. Pengaturan lainnya diperlukan untuk menyesuaikan spesifikasi data dengan perangkat keras yang digunakan. Hal ini dilakukan agar tidak terjadi *error* selama proses pengolahan. Beberapa hal yang perlu diatur antara lain *padding*, *batch size*, *threshold*, dan sebagainya. Pengaturan parameter yang lebih spesifik adalah sebagai berikut: *padding* = 256; *batch_size* = 4; *box_nms_thresh* = 0,7; *points_per_batch* = 64; *stability_score_thresh* = 0,95; *min_mask_region_area* = 0. Proses segmentasi dengan SAM biasanya membutuhkan waktu yang cukup lama karena menyesuaikan spesifikasi perangkat keras yang digunakan. Disarankan untuk menggunakan GPU yang kuat dengan memori GPU minimal 8 GB.

SAM menghasilkan segmentasi berupa seluruh objek yang tampak pada citra masukan dan tidak memberikan label spesifik untuk setiap hasil segmentasi. Penelitian ini berfokus pada objek berupa tapak bangunan sehingga diperlukan proses ekstraksi tapak bangunan untuk memisahkan tapak bangunan dengan objek tersegmentasi lainnya. Hasil segmentasi berupa data vektor poligon, kemudian diintegrasikan dengan data *ground truth* untuk mengekstraksi hanya tapak bangunan. Proses ini memanfaatkan fungsi spasial *overlay*.

Poligon tapak bangunan (*building footprints*) hasil segmentasi menggunakan SAM memiliki bentuk yang tidak tegas dan tidak teratur, sehingga perlu disederhanakan bentuknya melalui proses regularisasi. Proses regularisasi dilakukan secara otomatis di *software* ArcGIS Pro yang berbasis algoritma *polyline compression*.

3.2.2 Evaluasi Hasil

Evaluasi hasil ekstraksi dilakukan dengan dua cara yaitu perhitungan metrik *precision*, *recall*, *f1-score,* dan indeks *Intersection over Union* (IoU). *F1-score* adalah rata-rata harmonis dari *precision* dan *recall*. *Precision* fokus pada keakuratan hasil prediksi, sedangkan *recall* fokus pada kelengkapan hasil prediksi. Matrik evaluasi ini dihitung berdasarkan luas geometrik poligon hasil segmentasi dan data *ground truth*. *Precision*, *recall*, dan *f1-score* dihitung menggunakan Persamaan (1), (2), dan (3) berikut ini (Yang et al., 2018).

$$Precision = \frac{\text{TP}}{(\text{TP+FP})} \quad (1)$$

$$Recall = \frac{\text{TP}}{(\text{TP+FN})} \quad (2)$$

$$f1\ score = 2 \times \frac{precision \times recall}{precision+recall} \quad (3)$$

Indeks IoU menghitung perbandingan antara luasan yang saling tumpang tindih (*intersection*) antara data *ground truth* dengan hasil ekstraksi menggunakan *deep learning* dibandingkan dengan total luasan areanya (*union*). Perhitungan menggunakan indeks IoU dapat dilihat pada Persamaan (4) (Rezatofighi et al., 2019).

$$IoU = \frac{area\ of\ intersection\ (A \cap B)}{area\ of\ union\ (A \cup B)} \quad (4)$$

4. HASIL DAN PEMBAHASAN

4.1. Hasil Segmentasi dengan SAM (*Segment Anything Model*)

Kemampuan SAM dalam melakukan segmentasi memberikan hasil yang baik dengan waktu yang lebih praktis karena tidak memerlukan data pelatihan khusus. Penelitian ini menggunakan SAM dan ortofoto di dua area dengan karakteristik berbeda.

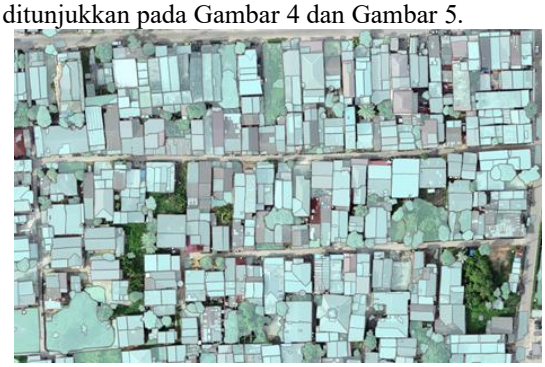
Tabel 1. Rangkuman hasil pengolahan data dengan SAM

Karakteristik kepadatan bangunan	AOI-1 (padat/tinggi)	AOI-2 (renggang/rendah)
Lokasi (kelurahan)	Dumai Kota, Laksamana, Rimba Sekampung, Sukajadi, Bintan	Mundam
GSD	10 cm	10 cm
Luas (Ha)	204,424	128,629
Dimensi [kolom, baris]	[16.504, 15.543]	[14.944, 8.627]
Waktu pengolahan	1 jam 34 menit 55 detik	44 menit 47 detik

Tabel 1 di atas menunjukkan waktu pengolahan data menggunakan SAM dengan deskripsi spesifikasi data raster ortofoto yang digunakan. Berdasarkan waktu pengolahan, SAM

membutuhkan waktu proses lebih lama pada data dengan luas yang lebih besar. Hal ini menunjukkan bahwa semakin besar luas wilayah, maka semakin lama waktu yang diperlukan oleh SAM dalam melakukan segmentasi. Waktu pemrosesan meningkat sebanding dengan luas wilayah yang harus dianalisis karena jumlah piksel yang juga semakin banyak. Area yang lebih besar memerlukan lebih banyak operasi konvolusi dan segmentasi yang perlu dilakukan oleh model. Selain itu, pemrosesan juga dipengaruhi oleh spesifikasi perangkat keras yang digunakan.

Hasil segmentasi dengan SAM secara visual menunjukkan seluruh segmentasi objek yang tampak pada citra. Objek-objek yang tersegmentasi dianggap satu kelas yang sama sehingga hasil SAM masih sangat general. Hasil segmentasi SAM menggunakan ArcGIS Pro mengacu pada konsep keluaran SAM asli yang mendeteksi suatu objek ke dalam beberapa *mask* berdasarkan nilai kepercayaan (*confidence*). Penelitian ini menghasilkan segmentasi berupa *mask* dalam format vektor sehingga hasil segmentasi suatu objek dapat terdiri dari vektor yang saling tumpang tindih. Secara keseluruhan, SAM dapat melakukan segmentasi individu objek secara detail seperti yang ditunjukkan pada Gambar 4 dan Gambar 5.



Gambar 4. Hasil segmentasi dengan algoritma SAM pada AOI-1 (Sumber data: Ortofoto UAV dari PT. Hasanah Surveyor Raya yang diproses oleh penulis)



Gambar 5. Hasil segmentasi dengan algoritma SAM pada AOI-2 (Sumber data: Ortofoto UAV dari PT. Hasanah Surveyor Raya yang diproses oleh penulis)

Berdasarkan visualnya, SAM mampu mendeteksi tapak bangunan dengan baik karena seluruh tapak bangunan dapat tersegmentasi dibandingkan dengan objek lainnya. Hal ini menunjukkan bahwa SAM memiliki potensi baik untuk kedepannya dalam ekstraksi tapak bangunan secara otomatis dan cepat.

4.2. Hasil Ekstraksi dan Regularisasi Tapak Bangunan

Ekstraksi tapak bangunan dilakukan dengan memisahkan antara hasil segmentasi tapak bangunan dengan objek lainnya. Proses pemisahan memanfaatkan fungsi spasial *overlay* yang diintegrasikan dengan data *ground truth*. Penggunaan fungsi spasial *overlay* dapat secara praktis menentukan satu kelas objek yang diinginkan berdasarkan data *ground truth*. Hasil SAM masih berupa data vektor poligon yang saling tumpang tindih sehingga terdapat perbedaan signifikan antara jumlah poligon tapak bangunan hasil segmentasi dengan data *ground truth*. Oleh karena itu perlu dilakukan *editing* untuk mencapai selisih jumlah yang minimal.



(a) Sebelum dilakukan *editing*



(b) Sesudah dilakukan *editing*

Gambar 6. Perbandingan hasil segmentasi tapak bangunan dengan SAM (Sumber data: Ortofoto UAV dari PT. Hasanah Surveyor Raya yang diproses oleh penulis)

Hasil segmentasi tapak bangunan setelah dilakukan editing menjadi lebih rapi dan garis poligon tidak terlihat acak dan saling tumpang tindih. Apabila dilihat berdasarkan segi visual, hasil segmentasi dengan SAM cukup baik dalam merepresentasikan garis tapak bangunan. Secara garis besar, hasil segmentasi telah mewakili representasi garis tapak bangunan utama.

Poligon hasil segmentasi pada umumnya memiliki tepi yang kasar atau bergerigi, yang disebabkan oleh ketidaksempurnaan dalam proses segmentasi. Selain itu, khusus objek bangunan, sudut-sudut pada poligon bisa terlalu tajam atau tidak sesuai dengan bentuk asli bangunan yang biasanya memiliki sudut 90° atau mendekati. Kelemahan hasil segmentasi tersebut diperbaiki melalui proses regularisasi untuk mendapatkan bentuk bangunan yang teratur dan tegas. Proses regularisasi ditunjukkan pada Gambar 7.



(a) (b)

Gambar 7. Regularisasi poligon tapak bangunan. (a) sebelum di regularisasi dan (b) setelah di regularisasi (Sumber data: Ortofoto UAV dari PT. Hasanah Surveyor Raya yang diproses oleh penulis)

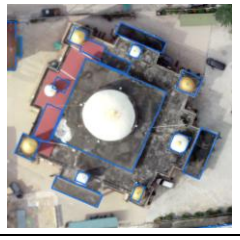
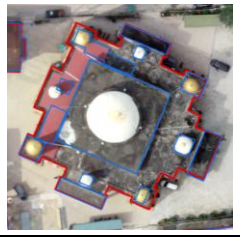
4.3. Evaluasi Hasil

Kehandalan suatu model *deep learning* dapat dinilai berdasarkan segi visual maupun geometrik dari hasil ekstraksi suatu objek. Hasil ekstraksi garis tapak bangunan menggunakan algoritma SAM (*Segment Anything Model*) perlu dilakukan penilaian (*assessment*) untuk mengetahui kehandalan model tersebut. Penilaian dari segi visual maupun geometrik dilakukan dengan membandingkan hasil segmentasi yang sudah dilakukan *editing*, hasil setelah dilakukan regularisasi, dan data *ground truth*.

Secara visual, hasil segmentasi dengan SAM tidak selalu memberikan hasil yang sempurna di setiap wilayah penelitian. Tabel 2 menunjukkan perbandingan hasil ekstraksi dan *ground truth* pada AOI-1 dan AOI-2. Hasil segmentasi dipengaruhi oleh beberapa faktor seperti warna dan tekstur. Ketidaksempurnaan hasil segmentasi karena faktor warna atap yang bervariasi ditunjukkan pada nomor (1). Hasil nomor (2) dan (4) menunjukkan bangunan dengan kombinasi atap seng dan beton serta sedikit kontras warna yang terlalu terang sehingga mempengaruhi hasil segmentasi. Selain itu, bangunan yang memiliki ukuran kecil serta terhalang objek lainnya seperti nomor (3) menjadi kelemahan SAM dalam mengekstraksi tapak bangunan.

Tabel 2 hanya menunjukkan beberapa contoh hasil ekstraksi yang tidak sempurna, sehingga untuk mengetahui kehandalan SAM yang lebih komprehensif diperlukan perhitungan uji akurasi. Hasil SAM perlu dilakukan uji akurasi untuk mengevaluasi seberapa akurat model SAM dalam melakukan segmentasi garis tapak bangunan. Penilaian dilakukan pada hasil segmentasi dan hasil regularisasi untuk memberikan perbandingan yang bervariasi.

Tabel 2. Perbandingan visualisasi antara hasil ekstraksi tapak bangunan dan data *ground truth*

No.	Lokasi	Hasil segmentasi	Hasil regularisasi	Hasil ekstraksi dan data *ground truth*
1.	AOI-1 (kerapatan bangunan tinggi)			
2.				

3.	AOI-2 (kerapatan bangunan rendah)			
4.				

Precision, *recall*, dan *f1-score* adalah matrik evaluasi yang penting dalam analisis hasil segmentasi citra. Metrik-metrik ini digunakan untuk menilai seberapa baik model segmentasi (*Segment Anything Model*) dapat mengekstraksi tapak bangunan pada ortofoto. Selain itu, IoU memberikan ukuran seberapa baik model dapat mengidentifikasi area objek yang tepat dalam citra. Perhitungan dilakukan berdasarkan geometrik luasan poligon tapak bangunan antara hasil segmentasi dan hasil regularisasi dengan data acuan (*ground truth*). Evaluasi hasil berdasarkan nilai *precision*, *recall*, *f1-score* dan IoU ditunjukkan pada Tabel 3.

Tabel 3. Hasil uji akurasi ekstraksi tapak bangunan

	AOI-1 (kerapatan bangunan tinggi)		**AOI-2 (kerapatan bangunan rendah)**	
	Hasil segmentasi	Hasil regularisasi	Hasil segmentasi	Hasil regularisasi
Precision	98.42%	97.78%	97.08%	96.08%
Recall	90.68%	89.59%	92.44%	90.77%
f1- score	94.39%	93.51%	94.71%	93.35%
IoU	89.38%	87.80%	89.94%	87.53%

Berdasarkan hasil uji akurasi, rata-rata nilai akurasi menurun setelah dilakukan regularisasi. Hal ini dapat disebabkan karena adanya perubahan bentuk yang cukup signifikan sehingga terdapat penurunan luas geometrik. Secara keseluruhan, rata-rata nilai akurasi hasil ekstraksi tapak bangunan diatas 87%. Hal ini menandakan bahwa SAM handal dalam melakukan ekstraksi objek dari data citra seperti ortofoto. Kemampuan SAM kedepannya sangat menjanjikan untuk tugas penginderaan jauh lainnya.

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil dan pembahasan yang sudah diperoleh maka dapat disimpulkan bahwa:

1. Penelitian ini menggunakan model *deep learning package* SAM yang telah disediakan oleh Esri. Model ini mampu melakukan segmentasi pada citra atau ortofoto secara otomatis dan tidak memerlukan proses pembuatan model *deep learning*. Hal ini memberikan keuntungan dari segi efektivitas dan efisiensi waktu.
2. SAM menunjukkan kehandalannya dalam melakukan ekstraksi tapak bangunan dengan rata-rata nilai akurasi lebih besar dari 87%. Hal ini menunjukkan bahwa SAM dapat dengan baik digunakan untuk tugas penginderaan jauh seperti segmentasi dan *object detection*.

5.2 Saran

Beberapa saran untuk perkembangan SAM selanjutnya antara lain:

1. Penelitian ini hanya berfokus pada objek tapak bangunan saja, sehingga diperlukan percobaan ekstraksi objek lainnya menggunakan SAM seperti jalan, vegetasi, batas lahan, dan lain sebagainya.
2. SAM merupakan suatu model dasar segmentasi baru yang masih perlu dikembangkan lebih lanjut untuk menyelesaikan permasalahan tugas penginderaan jauh yang lebih kompleks. Integrasi SAM dengan metode *deep learning* lainnya dapat dikembangkan untuk pemanfaatan tugas penginderaan jauh lainnya.

UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Departemen Teknik Geodesi Universitas Gadjah Mada yang telah memberikan bantuan fasilitas perangkat keras dalam pelaksanaan penelitian. Ucapan terimakasih juga disampaikan kepada PT. Hasanah Surveyor Raya yang telah memberikan bantuan data dan dokumen yang dibutuhkan dalam penelitian ini.

DAFTAR PUSTAKA

Abdollahi, A., Pradhan, B., Gite, S., & Alamri, A. (2020). Building Footprint Extraction from High Resolution Aerial Images Using Generative Adversarial Network (GAN) Architecture. *IEEE Access*, *8*, 209517–209527. https://doi.org/10.1109/ACCESS.2020.3038225

Elachi, C., & Van Zyl, J. (2021). Introduction to the Physics and Techniques of Remote Sensing, Third Edition. In *A JOHN WILEY & SONS, INC., PUBLICATION*. https://doi.org/10.1002/9781119523048

Farajzadeh, Z., Saadatseresht, M., Alidoost, F., Engineering, G., & Science, C. (2023). AUTOMATIC BUILDING EXTRACTION FROM UAV-BASED IMAGES AND DSMs USING DEEP LEARNING. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *X-4/W1-202*(February), 171–177.

Gui, B., Bhardwaj, A., & Sam, L. (2024). Evaluating the Efficacy of Segment Anything Model for Delineating Agriculture and Urban Green Spaces in Multiresolution Aerial and Spaceborne Remote Sensing Images. *Remote Sensing*, *16*(2), 1–32. https://doi.org/10.3390/rs16020414

Guo, H., Su, X., Tang, S., Du, B., & Zhang, L. (2021). Scale-Robust Deep-Supervision Network for Mapping Building Footprints from High-Resolution Remote Sensing Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 10091–10100. https://doi.org/10.1109/JSTARS.2021.3109237

Janiesch, C., Zschech, P., & Heinrich, K. (2023). Machine learning and deep learning. *Electronics Markets*, *31*, 71–84. https://doi.org/10.1515/9783110791402-004

Ji, W., Li, J., Bi, Q., Liu, T., Li, W., & Cheng, L. (2023). *Segment Anything Is Not Always Perfect: An Investigation of SAM on Different Real-world Applications*. http://arxiv.org/abs/2304.05750

Jiang, S., Jiang, C., & Jiang, W. (2020). Efficient structure from motion for large-scale UAV images: A review and a comparison of SfM tools. *ISPRS Journal of Photogrammetry and Remote Sensing*, *167*(September 2019), 230–251. https://doi.org/10.1016/j.isprsjprs.2020.04.016

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., & Girshick, R. (2023). Segment Anything. *Arxiv Computer Vision and Pattern Recognition*, 1–30. http://arxiv.org/abs/2304.02643

Li, Y., Wang, D., Yuan, C., Li, H., & Hu, J. (2023). Enhancing Agricultural Image Segmentation with an Agricultural Segment Anything Model Adapter. *Sensors*, *23*(18). https://doi.org/10.3390/s23187884

Negara, T. B., & Harintaka. (2021). Pemodelan Bangunan 3D Menggunakan Footprint Bangunan Hasil Ekstraksi Mask R-CNN dan Dense Point Cloud dari Foto Udara UAV. *Prosiding FIT ISI*, *1*, 248–260.

Osco, L. P., Wu, Q., de Lemos, E. L., Gonçalves, W. N., Ramos, A. P. M., Li, J., & Marcato, J. (2023). The Segment Anything Model (SAM) for remote sensing applications: From zero to one shot. *International Journal of Applied Earth Observation and Geoinformation*, *124*(October), 1–18. https://doi.org/10.1016/j.jag.2023.103540

Rastogi, K., Bodani, P., & Sharma, S. A. (2022). Automatic building footprint extraction from very high-resolution imagery using deep learning techniques. *Geocarto International*, *37*(5), 1501–1513. https://doi.org/10.1080/10106049.2020.1778100

Ren, Y., Yang, X., Wang, Z., Yu, G., Liu, Y., Liu, X., Meng, D., Zhang, Q., & Yu, G. (2023). Segment Anything Model (SAM) Assisted Remote Sensing Supervision for Mariculture—Using Liaoning Province, China as an Example. *Remote Sensing*, *15*(24). https://doi.org/10.3390/rs15245781

Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., & Savarese, S. (2019). Generalized intersection over union: A metric and a loss for bounding box regression. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2019-June*, 658–666. https://doi.org/10.1109/CVPR.2019.00075

Wu, Q., & Osco, L. P. (2023). samgeo: A Python package for segmenting geospatial data with the Segment Anything Model (SAM). *Journal of Open Source Software*, *8*(89), 5663. https://doi.org/10.21105/joss.05663

Yang, H., Wu, P., Yao, X., Wu, Y., Wang, B., & Xu, Y. (2018). Building extraction in very high resolution imagery by dense-attention networks. *Remote Sensing*, *10*(11), 1–16. https://doi.org/10.3390/rs10111768

Yao, H., Qin, R., & Chen, X. (2019). Unmanned aerial vehicle for remote sensing applications - A review. *Remote Sensing*, *11*(12), 1–22. https://doi.org/10.3390/rs11121443

Yu, T., Tang, P., Zhao, B., Bai, S., Gou, P., Liao, J., & Jin, C. (2023). ConvBNet: A Convolutional Network for Building Footprint Extraction. *IEEE Geoscience and Remote Sensing Letters*, *20*, 1–5. https://doi.org/10.1109/LGRS.2023.3250091