

EKSTRAKSI TAPAK BANGUNAN DARI ORTHOPHOTO MENGUNAKAN MODEL MASK R-CNN (Studi kasus: Kelurahan Ploso, Pacitan, Jawa Timur)

Rahmat Ardiansyah¹, Ir. Achmad Ruchlihadiana R.,M.M.², Levana Apriani, S.T.,M.T.³

¹Mahasiswa Teknik Geodesi Universitas Winaya Mukti, Bandung

²Dosen Pembimbing Satu Teknik Geodesi Universitas Winaya Mukti, Bandung

³Dosen Pembimbing Dua Teknik Geodesi Universitas Winaya Mukti, Bandung

ABSTRACT

The demand for accurate, efficient, and up-to-date spatial data continues to grow in the digital era. In this context, base maps play a crucial role in supporting various activities such as spatial planning, infrastructure development, and disaster management. Traditionally, the process of updating base maps, especially for extracting building footprints from aerial imagery, has been conducted manually—requiring substantial time and resources. However, advances in remote sensing and artificial intelligence now enable automated object extraction through deep learning approaches. This study aims to evaluate the effectiveness of the Mask R-CNN model in extracting building footprints from orthophoto imagery in Ploso Subdistrict, Pacitan District, East Java Province, Indonesia.

A quantitative method was applied using secondary data comprising high-resolution orthophotos (6 cm spatial resolution) and manually digitized building data as ground truth. The Mask R-CNN model was trained under three different training data proportions: 30%, 50%, and 70% of the total study area. The workflow included image preprocessing, object labeling, model training, inference on test data, and post-processing of extracted results. Model performance was evaluated using Precision, Recall, F1 Score, and Intersection over Union (IoU) metrics.

The best results were achieved under the 70% training data scenario, yielding a Precision of 0.6685, Recall of 0.3808, and F1 Score of 0.4852. The model successfully detected 1,117 true positives, with 554 false positives and 1,816 false negatives. These findings indicate that increasing the proportion of training data improves model performance in identifying and extracting building objects from high-resolution aerial imagery. Therefore, the Mask R-CNN method proves to be an effective and efficient solution for automated base map updates.

Keywords: *building extraction, deep learning, Mask R-CNN, aerial imagery, IoU*

ABSTRAK

Kebutuhan terhadap data spasial yang akurat, efisien, dan mutakhir semakin meningkat seiring dengan berkembangnya era digital. Dalam konteks ini, peta dasar menjadi elemen penting dalam mendukung berbagai aktivitas seperti perencanaan tata ruang, pembangunan infrastruktur, dan manajemen bencana. Proses pemutakhiran peta, khususnya ekstraksi objek bangunan dari citra udara, umumnya dilakukan secara manual dan membutuhkan waktu serta sumber daya yang besar. Dengan kemajuan teknologi penginderaan jauh dan kecerdasan buatan, ekstraksi objek kini dapat dilakukan secara otomatis menggunakan pendekatan *deep learning*. Penelitian ini bertujuan untuk mengevaluasi efektivitas model *Mask R-CNN* dalam mengekstraksi bangunan dari *orthophoto* di wilayah Kelurahan Ploso, Kecamatan Pacitan, Kabupaten Pacitan, Provinsi Jawa Timur.

Penelitian ini menggunakan pendekatan kuantitatif dengan data sekunder berupa *orthophoto* resolusi tinggi beresolusi spasial 6 cm serta data hasil digitasi bangunan sebagai acuan validasi. Model *Mask R-CNN* dilatih menggunakan tiga skenario proporsi data pelatihan, yaitu 30%, 50%, dan 70% dari total wilayah studi. Proses dilakukan mulai dari pra-pemrosesan citra, pelabelan objek, pelatihan model, inferensi terhadap data uji, hingga pasca-pemrosesan hasil ekstraksi. Evaluasi kinerja model dilakukan menggunakan metrik *Precision*, *Recall*, *F1 Score*, dan *Intersection over Union* (IoU).

Hasil terbaik diperoleh pada skenario pelatihan 70% dengan nilai *Precision* sebesar 0,6685, *Recall* sebesar 0,3808, dan *F1 Score* sebesar 0,4852. Model mampu menghasilkan *True Positive* sebanyak 1.117 objek, dengan jumlah *False Positive* dan *False Negative* masing-masing sebesar 554 dan 1.816. Hasil ini menunjukkan bahwa semakin besar proporsi data pelatihan, semakin baik pula performa model dalam mengenali dan mengekstraksi objek bangunan dari citra resolusi tinggi. Dengan demikian, metode *Mask R-CNN* terbukti efektif dalam mendukung otomatisasi pemutakhiran peta dasar secara efisien.

Kata kunci: ekstraksi bangunan, *deep learning*, *mask R-CNN*, citra foto udara, IoU

PENDAHULUAN

Latar Belakang

Kemajuan teknologi penginderaan jauh dan kecerdasan buatan mendorong efisiensi dalam pemetaan spasial. Salah satu tantangan utama adalah ekstraksi objek bangunan dari citra udara secara cepat dan akurat. Metode manual melalui digitasi *onscreen* memerlukan waktu dan tenaga yang besar. *Mask R-CNN* hadir sebagai solusi dengan kemampuan segmentasi yang unggul. Wilayah kelurahan Ploso dipilih karena perubahan penggunaan lahannya yang dinamis. Penelitian ini mengevaluasi *Mask R-CNN* dalam mengekstraksi bangunan dari *orthophoto* dan membandingkan hasilnya dengan data digitasi manual.

Kebutuhan terhadap data geospasial yang akurat dan mutakhir sangat penting dalam menunjang

pembangunan nasional, khususnya dalam hal penataan ruang, pemantauan lingkungan, dan mitigasi bencana. Bangunan sebagai salah satu unsur utama penutup lahan memiliki peran penting dalam analisis perubahan tata guna lahan dan perkembangan wilayah. Oleh karena itu, ekstraksi informasi tapak bangunan dari citra penginderaan jauh menjadi hal yang esensial. Metode konvensional seperti digitasi manual memiliki keakuratan tinggi, namun tidak efisien untuk skala besar karena memerlukan waktu dan tenaga yang signifikan.

Seiring berkembangnya teknologi pengolahan citra digital dan kecerdasan buatan, terutama metode *deep learning*, pendekatan otomatis menjadi sangat menjanjikan. *Mask R-CNN* merupakan salah satu algoritma segmentasi citra berbasis *deep learning* yang terbukti efektif dalam mengenali dan

memisahkan objek dalam citra secara presisi, termasuk objek bangunan. Model ini mampu melakukan deteksi objek berbasis bounding box sekaligus segmentasi piksel secara instans, yang membuatnya unggul dibanding metode terdahulu.

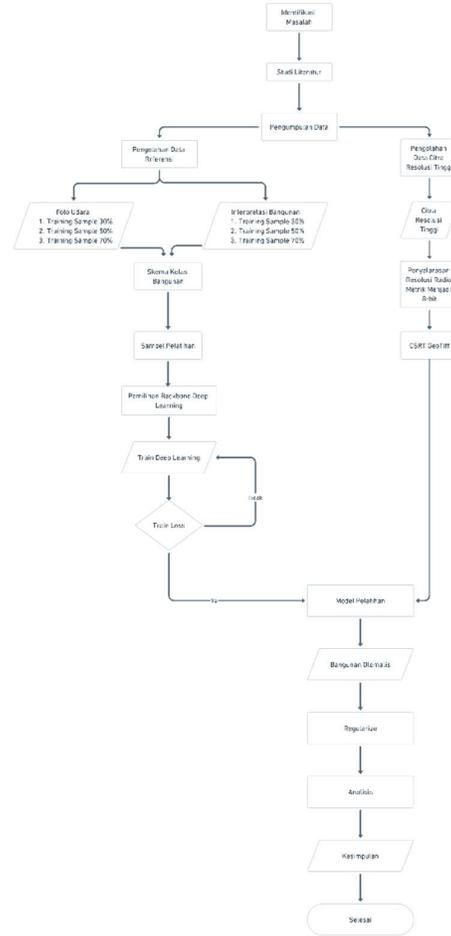
Dalam konteks penelitian ini, Kelurahan Ploso, Kecamatan Pacitan, dipilih sebagai lokasi studi karena ketersediaan data orthophoto resolusi tinggi dan karakteristik penggunaan lahan yang beragam. Penelitian ini bertujuan untuk mengevaluasi kemampuan Mask R-CNN dalam mengekstraksi tapak bangunan secara otomatis dan menilai keakuratannya dengan membandingkan hasilnya terhadap data referensi berupa digitasi manual.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen berbasis *deep learning*. Tujuannya adalah untuk mengevaluasi performa algoritma *Mask R-CNN* dalam mengekstraksi tapak bangunan dari citra *orthophoto* resolusi tinggi. Jenis data yang digunakan dalam penelitian ini meliputi data citra *orthophoto* dengan resolusi spasial 6cm dan data referensi berupa *shapefile* hasil digitasi manual bangunan. Wilayah studi yang dipilih adalah Kelurahan Ploso, Kecamatan Pacitan, yang secara administratif berada di Kabupaetn Pacitan, Provinsi Jawa Timur.

Kerangka Pemikiran

Kerangka pemikiran yang dilaksanakan dalam kegiatan ini dapat dilihat dari diagram di bawah ini:



Gambar 2. Kerangka Pemikiran

Tahap Pengolahan Data

Tahapan penelitian ini terdiri dari beberapa langkah utama, yaitu: Pra-pemrosesan data citra, pembuatan data pelatihan, pelatihan model *Mask R-CNN*, inferensi dan prediksi tapak bangunan, dan evaluasi akurasi hasil deteksi menggunakan metrik evaluasi. Setiap tahapan dirancang untuk memastikan proses ekstraksi dilakukan secara sistematis dan menghasilkan keluaran yang dapat dibandingkan secara kuantitatif.

Pada tahapan pra-pemrosesan, *orthophoto* dikonversi ke format RGB 8-bit dan dipotong menjadi *tile* dengan ukuran tertentu untuk mempermudah pelatihan model. Kemudian dilakukan pelabelan objek bangunan berdasarkan

shapefile referensi hasil digitasi manual. Setiap objek bangunan diberi label untuk memungkinkan model belajar mengenali pola spasial dan visual dari bangunan.

Model *Mask R-CNN* dilatih dengan tiga skenario proporsi data pelatihan yang berbeda, yaitu 30%, 50%, 70% dari total dataset. Hal ini bertujuan untuk mengetahui pengaruh umlah data pelatihan terhadap performa model. Proses pelatihan mengetahui dilakukan menggunakan GPU untuk mempercepat komputasi, dan parameter pelatihan seperti *batch size*, *learning rate*, serta jumlah iterasi ditetapkan agar tetap konsisten pada ketiga skenario untuk menjaga validasi perbandingan.

Setelah pelatihan selesai, model digunakan untuk mendeteksi objek bangunan pada citra uji hasil deteksi dibandingkan dengan data referensi untuk mengevaluasi akurasi model. Metrik evaluasi yang digunakan meliputi *Precision*, *Recall*, *F1 Score*, dan *Intersection Over Union (IoU)*. Metrik ini membantu untuk mengetahui seberapa akurat model dalam mendeteksi dan mengklasifikasikan tapak bangunan, serta sejauhmana hasil segmentasi mendekati data sebenarnya.

Perangkat keras yang digunakan untuk pengolahan *deep learning* dengan algoritma khususnya *R-CNN* pada perangkat lunak pengolah SIG harus mendukung perintah *CUDA*. Perangkat keras yang digunakan untuk mendukung penelitian ini menggunakan kartu grafis Nvidia RTX 4060TI dengan konfigurasi pendukung lain dengan memori sebesar 16gb, processor Intel I7-12700F, RAM 16gb, 512 SSD NVME.

Tahap Analisis

Dalam penelitian tugas akhir ini digunakan metode analisis untuk hasil akhir pengolahan data, metode analisis dalam penelitian ini menggunakan *IoU (Intersection Over Union)*, pengujian dilakukan untuk melihat ketelitian hasil bangunan yang didapatkan oleh metode *deep learning* dengan persamaan matematis *IoU* dengan cara memban-

dingkan data bangunan hasil ekstraksi otomatis dan data referensi. Untuk tujuan menghitung hasil bangunan yang telah dibentuk dengan hasil prediksi dari model bangunan sehingga menghasilkan nilai akurasi.

Hasil akhir dari evaluasi ini akan menghasilkan data metrik dari area yang dievaluasi. Berikut adalah formula umum untuk menghitung *IoU*:

$$IoU = \frac{Area\ Of\ Overlap}{Area\ Of\ Union} \dots \text{(Rumus 1)}$$

Keterangan:

IoU : Perpotongan Antar Data

AOV : Irisan antara objek prediksi

AOU :Luas gabungan objek yang diprediksi dengan *ground truth*

Selanjutnya jika hasil yang terdeteksi berpotongan dengan objek referensi atau *ground truth* maka dianggap sebagai ekstraksi yang berhasil (*true positive*), kemudian jika objek terdeteksi tidak berpotongan dengan data referensi maka dianggap sebagai kesalahan ekstraksi (*false positive*), jika objek referensi tidak berpotongan dengan hasil deteksi maka dianggap juga sebagai kesalahan ekstraksi (*flase negative*). Pernyataan ini dapat ditulis sebagai berikut:

$$Precision = \frac{TP}{TP + FP} \dots \dots \dots \text{(2)}$$

$$Recal = \frac{TP}{TP + FN} \dots \dots \dots \text{(3)}$$

$$F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall} \text{(4)}$$

HASIL PEMBAHASAN

Penelitian ini menghasilkan model ekstraksi tapak bangunan berbasis *Mask R-CNN* yang dilatih pada tiga skenario proporsi data pelatihan, yaitu 30%, 50%, dan 70%. Setiap skenario digunakan untuk mengevaluasi pengaruh jumlah data pelatihan terhadap performa deteksi bangunan dari *orthophoto* resolusi tinggi.

Hasil Pelatihan Model

Pelatihan Model *Mask R-CNN*

Pada penelitian ini, penulis mengimplementasikan *Mask R-CNN* menggunakan *backbone* ResNet-50, model dilatih menggunakan data bangunan referensi hasil dari digitasi manual (*on screen*) dengan tiga skenario yaitu 30%,50%, dan 70%. Model ini dilatih dengan jumlah *epoch* sebanyak 100, dengan ukuran *batch* 4. Pada skenario 30% memakan waktu sebanyak 3 jam 47 menit 46 detik, dan untuk skenario 50% memakan waktu 8 jam 27 menit 25 detik, dan untuk skenario 70% memakan waktu 8 jam 2 menit 53 detik. Dalam pelaksanaan pelatihan ini akan menghasilkan *training log* yang berbeda-beda pada skenario 30% mendapatkan hasil *average_precision* sebesar 0.592196, untuk skenario 50% menghasilkan *average_precision* sebesar 0.5850052, dan pada skenario 70% menghasilkan *average_precision* sebesar 0.67809270. Jumlah data sampel pelatihan dapat dilihat pada Tabel 1.

Tabel 1. Perbandingan Data Pelatihan

Jenis Data	Persentase	Jumlah Data
Skenario 1	30%	600
Skenario 2	50%	1.600
Skenario 3	70%	2.200

Ekstraksi Tapak Bangunan

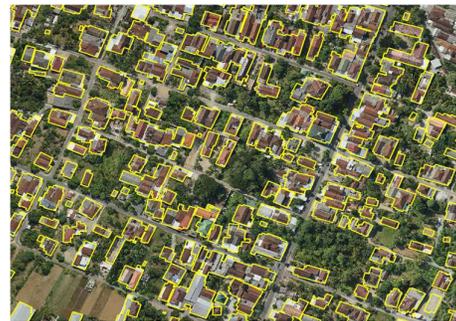
Model yang telah terlatih digunakan untuk mendeteksi tapak bangunan di Kelurahan Ploso. Hasil ekstraksi bangunan dapat dilihat pada gambar 3. Terlihat bahwa model sudah cukup baik dalam mendeteksi bangunan namun masih harus diperbaiki karna masih terdapat kesalahan geometri. Meskipun *Mask R-CNN* dapat melakukan segmentasi tingkat objek, hasil dari ekstraksi tapak bangunan masih perlu dilakukan *post-processing* untuk membuat bangunan lebih teratur.



Gambar 3. Hasil Ekstraksi Bangunan

Pelaksanaan *Post-Processing*

Setelah melakukan pendeteksian objek bangunan diperlukan *post-processing* untuk memperbaiki bentuk geometri dari tapak bangunan yang sudah terdeteksi. Hasil ekstraksi tapak bangunan yang sudah terdeteksi masih kurang akurat sehingga perlu dilakukan regularisasi bangunan untuk membuat bentuk geometri bangunan lebih teratur. Dengan menggunakan regularisasi polygon efektif untuk penyesuaian batas antar bangunan agar menjadi bangunan yang lebih terstruktur. Untuk bangunan yang telah melewati tahapan regularisasi dapat dilihat pada gambar 4.



Gambar 4. Hasil dari Bangunan yang telah melalui Regularisasi

Hasil Evaluasi

Hasil bangunan otomatis akan dilakukan tahap uji dengan data bangunan yang bersumber dari data referensi, ber-dasarkan perpotongan dari hasil bangunan yang terbentuk diantara kedua data tersebut. Metode *Intersection*

Over Union (IoU) adalah metode yang diterapkan untuk uji akurasi terhadap hasil ekstraksi bangunan otomatis dalam penelitian ini penulis menggunakan IoU 50% yang berarti objek dianggap positif jika terdapat interseksi minimal 50% antara polygon bangunan hasil ekstraksi dengan bangunan referensi. Untuk hasil uji akurasi pada tiap-tiap skenario dapat dilihat pada Tabel 2, Tabel 3, dan Tabel 4.

Tabel 2. Hasil Uji Akurasi Skenario 30%

<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>	<i>True Positive</i>	<i>False Positive</i>	<i>False Negative</i>
0.5374	0.2820	0.3699	827	712	2106

Tabel 3. Hasil Uji Akurasi Skenario 50%

<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>	<i>True Positive</i>	<i>False Positive</i>	<i>False Negative</i>
0.6167	0.3648	0.4584	1070	665	1863

Tabel 4. Hasil Uji Akurasi Skenario 70%

<i>Precision</i>	<i>Recall</i>	<i>F1 Score</i>	<i>True Positive</i>	<i>False Positive</i>	<i>False Negative</i>
0.6684	0.3808	0.4852	1117	554	1816

Untuk *precision* adalah nilai pemodelan yang dihasilkan dari jumlah unsur bangunan yang berhasil dibentuk dan terverifikasi dengan data referensi dari hasil digitasi, yang juga dapat disebut dengan *True Positive*, sedangkan *recall* merupakan perhitungan dengan menghitung unsur kesalahan pembentukan bangunan atau dapat disebut *False Negative*, kemudian *F1-Score* adalah bentuk perhitungan yang menyeluruh baik dari jumlah yang berhasil dibentuk maupun kesalahan pembentukan unsur bangunan baik dari *False Negative* dan *False Positive* sehingga mempunyai nilai rata-rata diantara *Precision* dan *Recall*.

Data matriks tersebut merupakan tolak ukur dari sebuah pemodelan dari proses *training sample* untuk mengetahui sejauh mana hasil pemodelan dikatakan bagus sehingga dapat diputuskan untuk penggunaan lebih lanjut atau pengulangan pada proses *training sample*.

KESIMPULAN

Berdasarkan analisis yang telah dilakukan dari kedua hasil ekstraksi bangunan secara otomatis menggunakan Mask R-CNN dengan data referensi hasil digitasi yang berbeda, maka diperoleh kesimpulan sebagai berikut:

1. Secara keseluruhan, proses ekstraksi bangunan yang dilakukan dikatakan baik berdasarkan skor ketelitian yang diukur menggunakan metrik IoU. Pada data Skenario 1, diperoleh nilai *precision* sebesar 0.5374, *recall* 0.2820, dan *F1 score* 0.3699. sedangkan pada pelatihan data Skenario 2, nilai *precision* mencapai 0.6167, *recall* 0.3648, dan *F1 score* 0.4584 dan pada Skenario 3 diperoleh nilai *precision* sebesar 0.668, *recall* 0.3808, dan *F1 score* 0.4852. Meskipun demikian, masih terdapat bangunan yang belum berhasil terdeteksi serta adanya objek lain seperti jalan, lapangan, dan vegetasi yang terkadang teridentifikasi sebagai bangunan. Selain itu, variasi model atap dan warna atap bangunan juga mempengaruhi hasil dari ekstraksi. Oleh karena itu, diperlukan keterlibatan tenaga manusia untuk melakukan kontrol kualitas dan melakukan penyuntingan pada data hasil ekstraksi agar data akhir menjadi lebih akurat dan sesuai dengan standar yang ditetapkan.
2. Setelah dilakukan pelatihan pada kedua data, yaitu Skenario 1 dengan jumlah data 30% dari area penelitian dan Skenario 2 dengan jumlah data 50% dan skenario 3 dengan jumlah data 70% dari area penelitian, dengan parameter yang sama, dapat disimpulkan bahwa penggunaan jumlah data pelatihan yang lebih banyak menghasilkan kualitas ekstraksi bangunan yang lebih baik. Pada pelatihan data Skenario 1, jumlah bangunan yang digunakan untuk pelatihan ± 600 bangunan, sedangkan pada pelatihan data Skenario 2 sebanyak ± 1.600

bangunan dan pada pelatihan data skenario 3 sebanyak \pm 2.200 bangunan. Hasil pengujian akurasi menggunakan IoU menunjukkan bahwa model yang dilatih dengan data lebih banyak (Pelatihan data Skenario 3) mampu mendeteksi objek bangunan dengan jumlah dan ketepatan yang lebih tinggi dibandingkan model yang dilatih dengan data lebih sedikit (Pelatihan data Skenario 1) walaupun hasil yang dihasilkan tidak signifikan antara skenario dan skenario 3 akan tetapi dengan jumlah data yang dilatih semakin banyak maka hasil prediksi objek akan semakin baik. Temuan ini menegaskan pentingnya ketersediaan data pelatihan yang memadai untuk meningkatkan performa model ekstraksi bangunan.

SARAN

Adapun saran yang dapat diterapkan untuk penelitian selanjutnya:

1. Karna dalam melakukan proses Training Data membutuhkan waktu yang lama maka diharuskan menggunakan perangkat keras yang lebih baik untuk mempersingkat waktu, dan juga dengan perangkat keras yang lebih baik akan dapat membantu untuk hasil bacaan yang lebih baik.
2. Menambahkan parameter tambahan di luar parameter bawaan perangkat lunak dapat meningkatkan kualitas hasil pelatihan. Namun, dalam penelitian ini, penulis hanya menggunakan parameter default dari perangkat lunak tanpa melakukan penyesuaian atau penambahan parameter lainnya.

DAFTAR PUSTAKA

- Boesch, G. (2024, January 4). *viso.ai*. Retrieved from What is Intersection over Union (IoU)?: https://viso.ai/computer-vision/intersection-over-union-iou/?utm_source=chatgpt.com
- Chang, K. (2016). *Introduction to Geographic Information System. 8th ed.* McGraw-Hill Education.
- Danoedoro, P. (1996). *Pengolahan Digital Teori dan Aplikasinya dalam Bidang Pengindraan Jauh*. Yogyakarta: Fakultas Geografi Universitas Gadjah Mada.
- Danoedoro, P. (2012). *Pengantar Penginderaan Jauh Digital*. Yogyakarta: Andi.
- Ekadinata, A., Dewi, S., Hadi, D. P., Nugroho, D. K., & Johana. (2009). *Sistem Informasi Geografis, Untuk Pengolahan Bentang Lahan Berbasis Sumber Daya Alam*.
- Emil Nafan, S. M., Fajrul islami, S. M., & Gushelmi, S. M. (2022). *Dasar-Dasar Deep Learning dan Contoh Aplikasinya*. CV. Mitra Cendekia Media.
- Erkamin, M., Mukhlis, I. R., Putra, Adiwarmam, M., Rassarandi, I. F., Rumata, I. N., . . . Hermawan, D. E. (2023). *SISTEM INFORMASI GEOGRAFIS (SIG)*. Yogyakarta: PT. Green Pustaka Indonesia.
- Gomez, C., & White, P. A. (2021). *Advancement in GIS for Disease Surveillance and Analysis*.
- He, K. G. (2017). *Mask R-CNN Proceedings of the IEEE International Conference on Computer*.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *Computer Vision and Pattern recognition*, 770-778.
- Kahlil, Munggaran, M. R., Kurnianggoro, L., & Adhiguna Mahendra, N. Z. (2023). *Computer Berbasis Deep Learning Untuk Aplikasi Pertanian*. Syiah Kuala University Press.
- Kuncoro, M. (2003). *Metode Riset Untuk Bisnis & Ekonomi Bagaimana meneliti & Menulis tesis ?* jakarta: Erlangga.

- Lillesand, T., & Kiefer, F. W. (2004). *Remote Sensing and Image Interpretation. Fifth Edition*. New York: Jhon Wiley and Sons.
- Lillesand, T., Kiefer, R. W., & Chipman, J. W. (2015). *Remote Sensing And Image Interpretaion*. Jphn Wiley and Sons.
- Lillesand, T., Kiefer, R., & Chipman, J. (2015). *Remote Sensing and Image Interpretation*. Jhon Wiley & Sons.
- Minsky, M. (1961). Steps toward artificial intelligence. . *Proceedings of the IRE*, 49 (1), 8-30.
- Nurmanini, S., Darmawahyuni, A., Sapitri, A. I., Rachmatullah, M. N., Firdaus, & Tutuko, B. (2021). *Pengenalan Deep Learning dan Implementasinya*. Palembang: UPT. Penerbit dan Percetakan(Universitas Sriwijaya).
- Ok, A. O. (2013). Automated detection of buildings form single VHR multispectral images using shadow information and graph cuts. . *Photogrammetry and Remote Sensing*, 86.
- Ren, S., He, K., Girshick, R., & Sun, J. (2016). *Fater R-CNN: Towards Real-Time Object Detection With Region Proposal Networks*. *Computer Vision and Pattern Recongnition* .
- Science, D. (2019). *Mask R-CNN*. Retrieved from Data Science: <https://datascience.eu/computer-vision/mask-r-cnn/>
- Sena, S. (2017). *Pengenalan Deep Learning Part 7: Convolutional Neural Network (CNN)*. Retrieved from <https://medium.com/@samuelsena/pengenalan-deep-learning-part-7-convolutional-neural-network-cnn-b003b477dc94>.
- Sugiyono. (2018). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Sutanto. (1991). *Pengindraan Jauh Untuk Survey Sumber Daya Alam*. Gadjah Mada University Press.
- Sutanto. (1994). *Pengindraan jauh Jilid I dan II*. Yogyakarta: Gadjah Mada Unigersity.
- Syahidi, A. A. (2024). *Pengolahan Citra Digital dan Dasar Visi Komputer dengan Python*. Banjarmasin Utara: POLIBAN PRESS.
- Yogi, A. (n.d.). *Transposed Convolution vs Convolutin Layer: Examples*. Retrieved from https://vitalflux.com/transposed-convolution-vs-convolution-layer-examples/?_x_tr_sl=en&_x_tr_tl=id&_x_tr_hl=id&_x_tr_pto=i mgs
- Zhu, X. X. (2021). *Deep Learning in Remote Sensing :A Comprehensive Review and List of Resources*. *IEEE Geoscience and Remote Sensing Magazine*