

Ekstraksi Data Bangunan Dari Data Citra *Unmanned Aerial Vehicle* Menggunakan Metode *Convolutional Neural Networks (CNN)* (Studi Kasus: Desa Campurejo, Kabupaten Gresik) *Building Data Extraction of Unmanned Aerial Vehicle Image Data Using Convolutional Neural Networks (CNN) Method (Case Study: Campurejo Village, Gresik Regency)*

Citra Ayu Sekar Kinasih*, Husnul Hidayat

Departemen Teknik Geomatika, FTSPK-ITS, Kampus ITS Sukolilo, Surabaya, 60111, Indonesia

*Korespondensi penulis: citraayusekark@gmail.com

Diterima: 16082021; Diperbaiki: 02112021; Disetujui: 15112021; Dipublikasi: 10012022

Abstrak: Seiring meningkatnya pembangunan akibat pertumbuhan penduduk, diperlukan suatu pengawasan dalam pemanfaatan lahan secara tepat salah satunya melalui pemetaan sebaran bangunan. Pemetaan sebaran bangunan dapat dilakukan dengan cara menganalisis citra penginderaan jauh yang diambil menggunakan berbagai wahana salah satunya menggunakan wahana *Unmanned Aerial Vehicle (UAV)* yang mampu menyediakan citra resolusi sangat tinggi. Namun, selama ini proses klasifikasi seringkali dilakukan dengan cara digitasi secara manual yang dianggap kurang efektif dan efisien sehingga dibutuhkan cara ekstraksi otomatis. Dalam penelitian ini metode *Convolutional Neural Networks (CNN)* digunakan untuk mengatasi tantangan ekstraksi bangunan menggunakan data citra foto udara resolusi tinggi pada Desa Campurejo, Kabupaten Gresik dengan menggunakan algoritma Mask R-CNN, di mana algoritma ini diharapkan mampu membantu proses klasifikasi secara otomatis dengan menggunakan data masukan (*training data*). Hasil klasifikasi kemudian dilakukan validasi dan uji akurasi sehingga mampu menghasilkan peta sebaran bangunan skala besar yaitu 1:5000. Akurasi hasil klasifikasi bangunan dengan metode Mask R-CNN diuji dengan menggunakan *confusion matrix* yang menghasilkan nilai *precision* 94,78%, *recall* 82,63%, *F1 Score* 88,29% dan *accuracy* 79,03% untuk wilayah 1 dan untuk wilayah 2 menghasilkan nilai *precision* 98,10%, *recall* 78,37%, *F1 Score* 87,13% dan *accuracy* 77,20%. Sementara jumlah bangunan yang dapat dideteksi pada wilayah 1 sebanyak 2102 bangunan dan wilayah 2 sebanyak 247 bangunan. Prosedur ini menunjukkan potensi yang besar untuk memanfaatkan metode *Convolutional Neural Networks (CNN)* dalam melakukan ekstraksi bangunan.

Copyright © 2022 Geoid. All rights reserved.

Abstract: Along with increasing development due to population growth, a proper monitoring of land use is needed, one of which is through mapping the distribution of buildings. Mapping the distribution of buildings can be done by analyzing remote sensing images taken using various vehicles, one of which is the *Unmanned Aerial Vehicle (UAV)* which can provide very high-resolution images. However, up to now the classification process is often done by manual digitization which is considered less effective and efficient so that an automatic extraction method is needed. In this study, the *Convolutional Neural Networks (CNN)* method was used to overcome the challenges of building extraction using high resolution aerial photo image data in Campurejo Village, Gresik Regency using the Mask R-CNN algorithm, where this algorithm is expected to be able to help the classification process automatically by using the Mask R-CNN algorithm. input data (*training data*). Then the results of the classification are validated and tested for accuracy to produce a large-scale building distribution map, namely 1: 5000. The accuracy of the building classification results using the Mask R-CNN method was tested using a *confusion matrix* which resulted in a *precision* value of 94.78%, *recall* 82.63%, *F1 Score* 88.29% and *accuracy* 79.03% for region 1 and for region 2 resulted *precision* value 98.10%, *recall* 78.37%, *F1 Score* 87.13% and *accuracy* 77.20%. While the number of buildings that can be detected in area 1 is 2102 buildings and area 2 is 247 buildings. This procedure shows great potential to utilize the *Convolutional Neural Networks (CNN)* method in extracting buildings.

Kata kunci: Bangunan, *Convolutional Neural Networks*, Mask-RCNN, Foto Udara

Cara untuk sitasi: Kinasih, C.A. & Hidayat, H. (2021). Ekstraksi Data Bangunan Dari Data Citra *Unmanned Aerial Vehicle* Menggunakan Metode *Convolutional Neural Networks (CNN)* (Studi Kasus: Desa Campurejo, Kabupaten Gresik). *Geoid*, 17(1), 81 - 92.

Pendahuluan

Bangunan adalah salah satu variabel masukan utama untuk membentuk model demografis dan sosial ekonomi (Xie dan Zhou 2017). Bangunan Gedung menurut Undang- Undang nomer 28 tahun 2002 adalah wujud fisik hasil pekerjaan konstruksi yang menyatu dengan tempat kedudukannya, sebagian atau seluruhnya berada di atas dan/atau di dalam tanah dan/atau air, yang berfungsi sebagai tempat manusia melakukan kegiatannya, baik untuk hunian atau tempat tinggal, kegiatan keagamaan, kegiatan usaha, kegiatan sosial, budaya, maupun kegiatan khusus. Seiring meningkatnya pembangunan akibat pertumbuhan penduduk menuntut adanya peningkatan tugas pengawasan, peningkatan standar prosedur dan perencanaan dalam pemanfaatan ruang untuk peruntukan lahan secara tepat salah satunya melalui pemetaan sebaran bangunan.

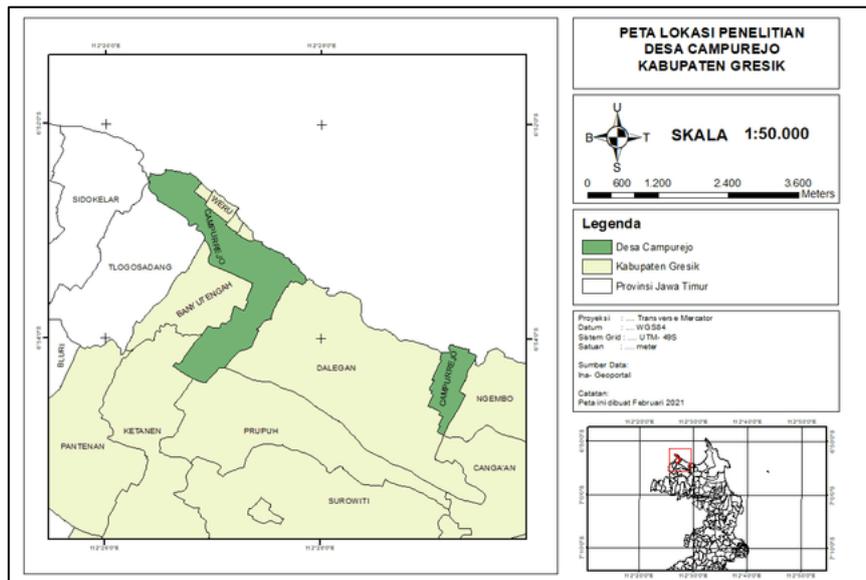
Dalam praktiknya, pemetaan sebaran bangunan dapat dilakukan dengan cara menganalisis citra penginderaan jauh yang diambil menggunakan berbagai wahana salah satunya menggunakan wahana *Unmanned Aerial Vehicle (UAV)* yang mampu menyediakan citra resolusi sangat tinggi yang dapat digunakan untuk pembuatan peta skala besar. Dalam penelitian pemetaan desa dalam skala 1: 2000 menggunakan orthophoto dari data citra UAV *rotary wing* yang dilakukan oleh Hidayat dan Sukojo (2017) dihasilkan hasil peta yang dapat diterima sesuai standar yang berlaku oleh Badan Informasi Geospasial. Namun, selama ini proses pemetaan dan pengambilan informasi seringkali dilakukan dengan cara digitasi secara manual yang dianggap kurang efektif dan efisien karena memakan banyak waktu, tenaga dan biaya serta memiliki hasil yang cenderung bergantung terhadap subjektivitas. Ekstraksi otomatis bangunan dari data penginderaan jauh sangat penting untuk banyak aplikasi, seperti pembaruan peta, pemodelan kota, analisis pertumbuhan kota, pemantauan perubahan, penanggulangan bencana, dll (Khoshelham, K. et al., 2010). Ekstraksi fitur merupakan proses pengukuran, mengolah atau mendeteksi fitur dari gambar sampel. Dua jenis fitur yang paling umum dalam ekstraksi adalah ekstraksi ciri geometris dan ekstraksi ciri warna (Gavali dan Banu, 2019). Dalam tata perkotaan yang kompleks seringkali terjadi kesulitan dalam mengklasifikasikan jenis bangunan.

Algoritma *Deep Learning (DL)* semakin banyak digunakan dalam aplikasi penginderaan jauh (Kussul, N., et al., 2017), salah satunya *Convolutional Neural Networks (CNN)*. CNN adalah arsitektur hierarki yang dapat dilatih pada kumpulan data skala besar untuk melakukan pengenalan dan deteksi objek (Hu, F., et al., 2015). Setiap proses pemahaman gambar yaitu klasifikasi gambar dan objek menggunakan model CNN memerlukan input data pelatihan. Penelitian ini menggunakan algoritma Mask R-CNN yaitu salah satu model yang dikembangkan dari model Faster- RCNN yaitu *region- based convolutional neural networks* yang dimana mampu mendeteksi objek dalam gambar sekaligus menghasilkan *segmentasi instance* dimana bertujuan untuk mendeteksi kelas objek bersama dengan prediksi kotak pembatas dalam gambar.

Oleh karena itu, penulis bertujuan untuk menggunakan potensi CNN untuk membantu melakukan fitur klasifikasi otomatis yang menghindari digitasi secara manual untuk mengatasi tantangan ekstraksi bangunan pada citra resolusi tinggi pada daerah Desa Campurejo, Kabupaten Gresik. Algoritma Mask-RCNN digunakan untuk mengatasi tantangan proses segmentasi bangunan dimana algoritma ini dapat membantu dalam melakukan pendeteksian batas tidak hanya berupa *semantic segmentation* namun juga *instance segmentation*. Wilayah desa Campurejo dipilih karena memiliki tingkat kompleksitas yang tinggi. Dimana wilayah ini memiliki sebuah eksklave yang membagi wilayahnya menjadi dua wilayah. Hal ini membuat desa campurejo memiliki kondisi dan bentang alam yang berbeda- beda sehingga berpengaruh pula terhadap persebaran dan bentuk geometri bangunannya.

Data dan Metode

Lokasi yang dijadikan studi kasus pada penelitian ini adalah Desa Campurejo, Kecamatan Panceng, Kabupaten Gresik. Wilayah Desa Campurejo mempunyai 2 wilayah yang terpisah dengan total luas wilayah ± 677 Ha. Secara administrasi Desa Campurejo terletak pada koordinat $6^{\circ}52'28.75''$ LS - $6^{\circ}54'53.96''$ LS dan $112^{\circ}26'41.2''$ BT- $112^{\circ}29'8.19''$ BT.



Gambar 1. Peta Lokasi Penelitian

Data yang digunakan dalam penelitian ini adalah data batas administrasi (.shp) Desa Campurejo dan data foto udara Desa Campurejo, Kecamatan Panceng, Kabupaten Gresik (.tiff). Data foto udara yang digunakan adalah data citra dengan akuisisi menggunakan *Unmanned Aerial Vehicle* (UAV) pada tahun 2020 yang memiliki resolusi spasial masing- masing 4,17 cm/pix untuk wilayah 1 dan 4,05 cm/pix untuk wilayah 2.

Pengolahan data dilakukan dengan menggunakan perangkat lunak ArcGIS Pro dengan lisensi ITS serta *Deep Learning Framework* agar dapat dilakukan proses deteksi objek bangunan menggunakan *deep learning* pada aplikasi.

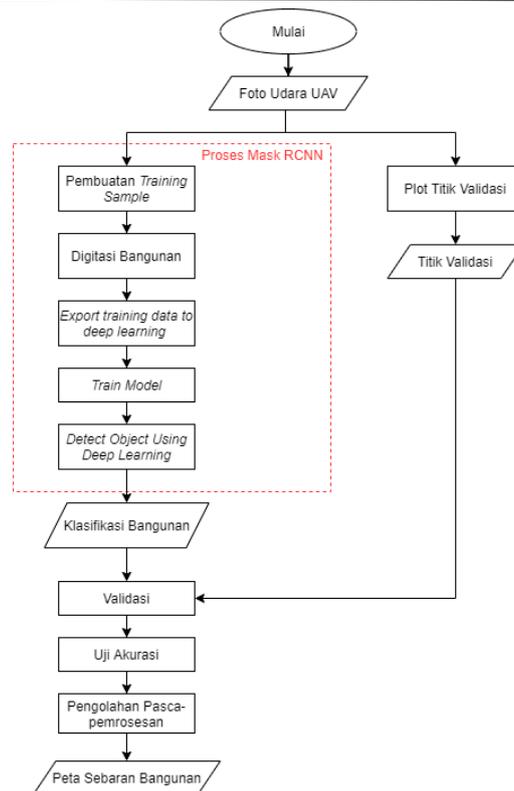
Proses pengolahan menggunakan metode Mask R-CNN untuk ekstraksi bangunan terdiri dari beberapa tahap utama, yaitu (1) Pembuatan *Training Sample*; (2) *Train Model*; (3) Klasifikasi dan Deteksi Objek; (4) Validasi dan Uji Akurasi; (5) Pengolahan Pasca Pemrosesan.

1. Pembuatan *Training Sample*

Pembuatan *training sample* dilakukan dengan cara pembuatan label objek. Dalam penelitian ini dilakukan pelabelan dengan menggunakan satu kelas yaitu bangunan. *Training sample* untuk proses Mask-RCNN dilakukan secara manual dengan digitasi beberapa bangunan yang dijadikan *sample*. Hasil dari label objek kemudian dilakukan proses *export training data to deep learning* agar selanjutnya dapat dilakukan proses *train model*. Dalam penelitian ini digunakan *training sample* bangunan kurang lebih 25% wilayah dari citra dengan total jumlah training sample ditunjukkan pada tabel 1 untuk mengetahui kemampuan model dalam melakukan deteksi objek. *Training sample* yang digunakan harus mampu mewakili keberagaman jenis bangunan yang berada diwilayah studi baik ciri geometris maupun ciri spectral/ warna nya. Semakin beragam jenis *training sample* yang ada, semakin baik model akan mendeteksi bangunan yang ada.

Tabel 1. Jumlah *Training Sample*

Kelas	Metode	Training Sample
Bangunan	Campurejo Wilayah 1	1261
	Campurejo Wilayah 2	262



Gambar 2. Diagram Alir Pengolahan

2. Train Model

Tahap *train model* dilakukan dengan menggunakan gambar dan label yang dihasilkan dari tahap sebelumnya. Dalam tahap ini perlu diatur beberapa parameter untuk membentuk model seperti *Max Epoch*, *Batch Size*, *Model Type*, *Backbone Model* dan *Validation*. Hasil *train model* kemudian akan dievaluasi hasil prediksi dan bentuk grafik pembelajarannya, apabila telah sesuai dapat dilanjutkan tahap berikutnya.

3. Klasifikasi dan Deteksi Bangunan

Hasil model yang telah terbentuk kemudian diuji coba kedalam data citra dengan menggunakan bantuan alat *geoprocessing* “*Detect Object Using Deep Learning*”. Pada tahap ini parameter seperti *padding*, *threshold* dan *batch size* perlu ditentukan. Dalam hal ini sasaran yang dapat dideteksi adalah bangunan- bangunan yang ada di desa Campurejo sehingga dapat diklasifikasi menjadi satu kelas yang sama yaitu bangunan.

4. Validasi dan Uji Akurasi

Dalam tahap validasi dilakukan penentuan titik- titik validasi berupa bangunan dari data foto udara UAV desa Campurejo, yang kemudian dilakukan *overlay* antara *vector* bangunan hasil deteksi dengan titik- titik validasi tersebut. Selain itu, data validasi lainnya adalah berupa digitasi polygon bangunan dimana selanjutnya akan mempertimbangkan Rasio *Intersection over Union (IoU)*. Rasio *Intersection over Union (IoU)* digunakan sebagai ambang batas untuk menentukan apakah hasil yang diprediksi adalah *true positive* atau *false positive*. IoU adalah jumlah tumpang tindih antara kotak pembatas di sekitar objek yang diprediksi dan kotak pembatas di sekitar data referensi tanah. Dalam penelitian ini digunakan ambang batas 0,5 yang artinya bangunan dianggap bertampalan apabila memiliki *overlay* melebihi 50%.

Tahap uji akurasi dilakukan dengan perhitungan *confusion matrix*. Pengujian dilakukan terhadap sampel yang mewakili objek tertentu dalam suatu polygon hasil klasifikasi dengan titik yang sesuai

dilapangan dalam hal ini yaitu titik validasi. Dari nilai *True Negative (TN)*, *False Positive (FP)*, *False Negative (FN)*, dan *True Positive (TP)* dapat diperoleh nilai akurasi, presisi, *F1 Score* dan *recall*.

$$Akurasi = \frac{\sum_{i=1}^l \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i}}{l} * 100\% \quad (1)$$

$$Presisi = \frac{\sum_{i=1}^l TP_i}{\sum_{i=1}^l (FP_i + TP_i)} * 100\% \quad (2)$$

$$Recall = \frac{\sum_{i=1}^l TP_i}{\sum_{i=1}^l (TP_i + FN_i)} * 100\% \quad (3)$$

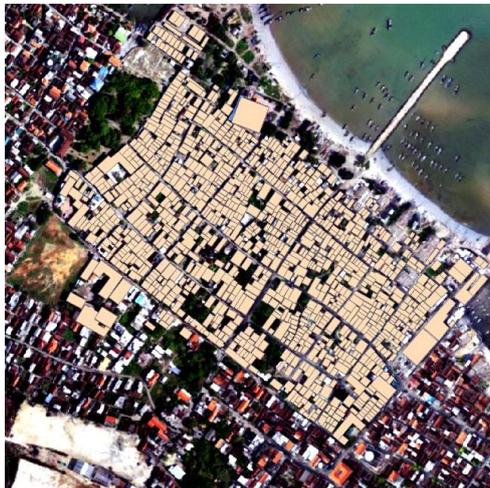
$$F1\ Score = \frac{(Precision * Recall)}{(Precision + Recall) * 2} * 100\% \quad (4)$$

5. Pengolahan Pasca Pemrosesan

Dalam tahap ini dilakukan proses perbaikan hasil klasifikasi seperti *regularize polygon* untuk memperbaiki hasil delineasi bangunan, *clipping* untuk menyeleksi bangunan, *merge* untuk memperbaiki hasil polygon yang bertampalan (*overlay*), dan *layouting* peta.

Hasil dan Pembahasan

1. Hasil *Training Data*



Gambar 3. *Training Data* Dari Hasil Digitasi



Gambar 4. Hasil Proses *Export Training Data*

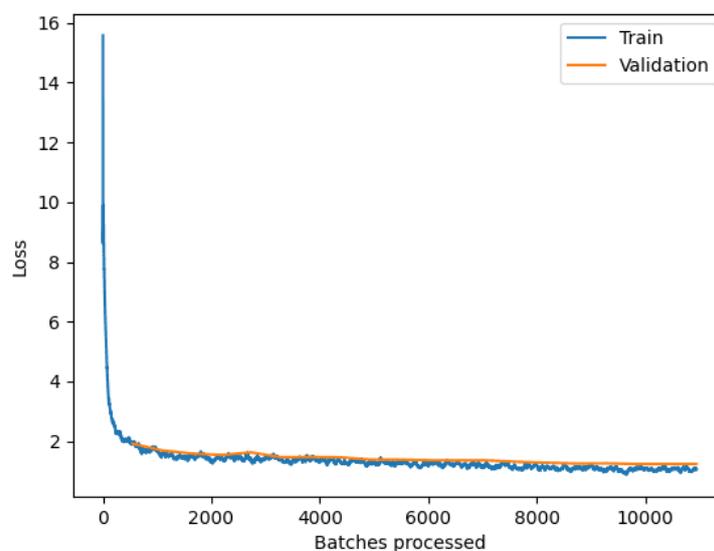
Pembuatan *training sample* menggunakan label objek pada citra sebagai data masukan. Sumber data masukan yang digunakan untuk training data yaitu *training data* dari hasil digitasi yang ditunjukkan pada Gambar 3 untuk proses Mask R-CNN. Pemberian label dikelompokkan menjadi satu kelas yaitu bangunan. Hasil pelabelan kemudian dilakukan proses *export training data for deep learning* sehingga dapat digunakan dalam proses pembentukan *deep learning*. Hasil dari proses ini ditunjukkan pada Gambar 4 yaitu berupa *image chips* yaitu kumpulan chip gambar dari raster yang dipotong sesuai dengan *tile size XY* dan *label chips* yaitu kumpulan label gambar chip dari raster yang telah dipotong.

2. Train Model

Pada pembentukan Model ini algoritma CNN yang digunakan adalah Mask R-CNN. Proses *train model* menggunakan *image chips* dan *label chips* dari proses sebelumnya. Dalam proses *train model* ini digunakan *Max Epoch 20*, *Batch Size 2*, *Model Type MaskRCNN*, *Backbone Model ResNET 50* dan *Validation 20%* dengan data masukan yaitu masing- masing citra. Setiap *train model* akan menunjukkan hasil prediksi seperti Gambar 5 dan hasil kurva pembelajaran pada Gambar 6.



Gambar 5. Hasil Prediksi Model



Gambar 6. Hasil Kurva Pembelajaran Model

Kurva pembelajaran adalah kurva yang menunjukkan plot kinerja pembelajaran model selama beberapa pengalaman atau waktu. Hasil kurva diamati untuk menentukan apakah model termasuk kedalam kategori *underfit*, *overfit* atau *good fit*. Berdasarkan hasil kurva pembelajaran model yang telah dibuat dapat dikatakan bahwa hasil ini termasuk kedalam kondisi model *good fit*. Model dikatakan masuk kedalam kategori *good fit* apabila memiliki ciri plot *training loss* dan *validation loss* menurun ketitik yang stabil dan memiliki celah kerugian yang kecil.

3. Hasil Klasifikasi dan Deteksi Bangunan

Klasifikasi bangunan dilakukan dengan menggunakan deteksi otomatis menggunakan *deep learning* algoritma Mask-RCNN dari hasil *train model*. Dalam proses pendeteksian nilai parameter yang digunakan yaitu *padding* 64, *threshold* 0.9 dan *batch size* 4. Nilai parameter yang digunakan dalam proses pendeteksian didapatkan melalui proses *trial and error* untuk menemukan parameter yang tepat. Klasifikasi ini hanya menghasilkan satu kelas yaitu bangunan. Jumlah bangunan yang mampu dideteksi oleh metode ini yaitu untuk wilayah 1 5.670 *polygon* dan wilayah 2 sebanyak 558 *polygon*. Hasil klasifikasi ini masih perlu dilakukan pengeditan secara manual.



Gambar 7. Hasil Klasifikasi Bangunan

4. Validasi dan Uji Akurasi



Gambar 8. Sebaran Titik Validasi

Validasi sebuah proses untuk menguji ketepatan dari hasil deteksi bangunan dengan keadaan sesungguhnya dilapangan. Hasil validasi ini kemudian digunakan untuk melakukan perhitungan untuk uji akurasi. Data validasi yang digunakan adalah titik- titik letak bangunan dan beberapa *sample* bangunan berupa *polygon* yang diplot dari data citra UAV desa Campurejo. Citra UAV dianggap mampu mewakili keadaan sesungguhnya dilapangan karena memiliki resolusi yang tinggi. Untuk daerah Campurejo wilayah 1 memiliki 4639 titik validasi dan wilayah 2 memiliki 527 titik validasi yang mewakili jumlah bangunan diwilayah tersebut seperti yang ditunjukkan pada Gambar 8.

Sementara itu, banyaknya *sample* bangunan berupa *polygon* yang digunakan yaitu untuk wilayah 1 sebanyak 260 bangunan sementara wilayah 2 sebanyak 50 bangunan yang ditunjukkan pada Gambar 9. *Sample polygon* yang digunakan merupakan hasil digitasi secara manual, bangunan yang dipilih untuk digunakan sebagai validasi yaitu bangunan- bangunan yang tidak dijadikan dalam *training data*. Titik- titik validasi tersebut kemudian digunakan dalam perhitungan uji akurasi.



Gambar 9. Sebaran *Polygon* Validasi

Tahap uji akurasi dilakukan dengan perhitungan *confusion matrix*. Pengujian dilakukan hasil klasifikasi bangunan dengan titik yang sesuai dilapangan dalam hal ini yaitu titik validasi dan *sample polygon* validasi. Kemudian dapat ditentukan nilai *True Positive* (TP), *False Positive* (FP), *False Negative* (FN), dan *True Negative* (TN) sehingga dapat diperoleh nilai akurasi, presisi, *recall* dan *F1 Score*.

Tabel 2. Perhitungan Akurasi

	Titik Validasi (<i>Point</i>)		Titik Validasi (<i>Polygon</i>)	
	Wilayah 1	Wilayah 2	Wilayah 1	Wilayah 2
Jumlah bangunan (terdeteksi)	5670	558	281	56
Jumlah bangunan (validasi)	4639	527	260	50
<i>True Positive</i>	3833	413	167	39

False Positive	211	8	21	6
False Negative	806	114	93	11
Precision	94,78%	98,10%	88,83%	86,67%
Recall	82,63%	78,37%	64,23%	78,00%
F1 Score	88,29%	87,13%	74,55%	82,11%
Accuracy	79,03%	77,20%		

5. Analisa Hasil

Dari hasil pengolahan, dapat dilihat bahwa metode Mask R-CNN mampu membantu proses klasifikasi ekstraksi bangunan secara otomatis dengan memasukan sejumlah *training data*. Hal ini dinilai sangat membantu dalam meningkatkan efisiensi dan efektivitas waktu dibandingkan proses digitasi secara manual dimana CNN mampu melakukan ekstraksi ribuan data bangunan secara otomatis dari *training data* dan hanya perlu divalidasi secara manual.

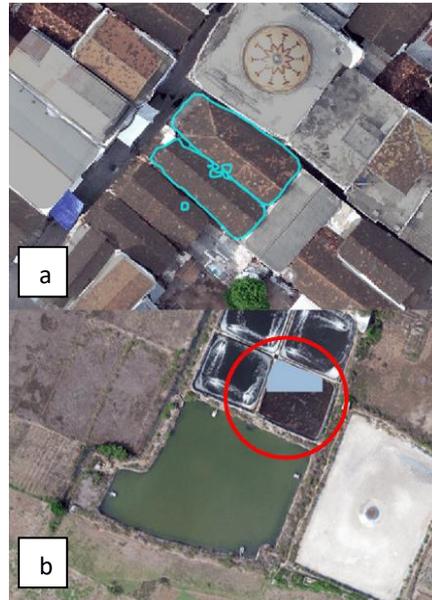
Hasil yang ditunjukkan pada Tabel 2 menggambarkan nilai untuk penilaian akurasi objek untuk kelas bangunan menggunakan dua data validasi berbeda. Nilai akurasi dilakukan pada seluruh citra wilayah desa Campurejo baik wilayah 1 maupun wilayah 2. Penilaian akurasi menghitung antara nilai TP (*True Positive*), TN (*True Negative*), FP (*False Positive*), dan FN (*False Negative*). Nilai pertampalan (*overlay*) ditunjukkan oleh nilai TP atau dengan kata lain model mampu mendeteksi dengan benar bangunan yang ada. Sementara nilai TN tidak memiliki nilai dalam penelitian ini karena nilai TN menunjukkan model mampu mendeteksi non-bangunan sebagai non-bangunan. Hal ini dikarenakan deteksi hanya dilakukan untuk bangunan yang memiliki satu kelas sehingga objek selain itu dianggap sebagai non bangunan. Oleh sebab itu nilai TN tidak diperhitungkan didalam penelitian ini.

Dari hasil uji akurasi klasifikasi bangunan menggunakan titik validasi dapat dilihat bahwa metode Mask R-CNN yang digunakan masih belum memenuhi nilai akurasi klasifikasi bangunan yang disyaratkan oleh Peraturan Kepala Badan Informasi Geospasial Nomor 15 Tahun 2014 yaitu memiliki ketelitian 85%. Uji akurasi menggunakan validasi titik menghasilkan nilai *precision* 94,78%, *recall* 82,63%, *F1 Score* 88,29% dan *accuracy* 79,03% untuk wilayah 1 dan untuk wilayah 2 menghasilkan nilai *precision* 98,10%, *recall* 78,37%, *F1 Score* 87,13% dan *accuracy* 77,20%. Sementara hasil menggunakan uji akurasi menggunakan data validasi berupa data luasan (*polygon*) juga belum memenuhi persyaratan tersebut dengan dapat dilihat menghasilkan nilai *precision* 88,83%, *recall* 64,23% dan *F1 Score* 74,55% untuk wilayah 1 dan untuk wilayah 2 menghasilkan nilai *precision* 86,67%, *recall* 78% dan *F1 Score* 82,11%. Meskipun belum memiliki nilai akurasi yang optimal, namun hasil model memiliki model yang baik dengan dikategorikan sebagai *good fit* dan telah mampu memprediksi dengan baik. Dari nilai hasil *precision*, *recall* dan *F1 Score* juga menunjukkan hasil yang baik.

Dalam melakukan ekstraksi bangunan terdapat dua jenis fitur yang paling umum dalam ekstraksi yaitu ciri geometris dan ciri warna. Keragaman bentuk bangunan, serta padatnya bangunan yang saling menempel satu sama lain di desa Campurejo membuat model sulit untuk mendeteksi bangunan seperti ditunjukkan pada Gambar 10a dimana bangunan berbeda dideteksi sebagai satu bangunan. Selain itu, bangunan yang memiliki bentuk dan ukuran yang tidak biasa seperti gubuk dan bangunan yang memiliki banyak atap juga sulit dideteksi oleh sistem. Dari ciri warna atau spektral, bangunan campurejo memiliki warna yang berbeda-beda bergantung pada pemilihan bahan. Selain itu bentang alam yang berbeda mengakibatkan beberapa bangunan terlihat berbaur dengan vegetasi, tanah dan air disekitarnya seperti yang terjadi pada Gambar 10b dimana tambak memiliki warna yang hampir mirip seperti genteng bangunan sehingga model mendeteksinya sebagai bangunan. Strategi yang dapat dilakukan yaitu dengan memperbanyak jenis ragam warna, ukuran, bentuk pada sampel *training data*.

Sejauh ini, dapat dilihat bahwab metode Mask-RCNN memiliki potensi yang sangat besar dalam melakukan klasifikasi bangunan. Proses perbaikan manual diperlukan untuk memperbaiki dan menganalisa kekurangan dari proses klasifikasi secara otomatis. Oleh sebab itu masih diperlukan peran operator terlatih manusia untuk

melakukannya. Meskipun memiliki hasil yang belum optimal, namun hingga saat ini hasil pendeteksian otomatis yang cepat belum mampu dicapai dengan metode lain. Sehingga dengan digunakannya metode ini diharapkan mampu membantu proses klasifikasi otomatis secara cepat. Proses klasifikasi otomatis ini apabila telah membuahkan hasil yang optimal dapat dimanfaatkan untuk banyak aplikasi dan bidang terutama dalam percepatan penyediaan peta dasar di Indonesia.



Gambar 10. Kesalahan Deteksi

6. Hasil Pengolahan Pasca Pemrosesan

Pengolahan pasca pemrosesan yaitu proses perbaikan hasil klasifikasi yang dilakukan secara manual meliputi proses *regularize polygon* untuk memperbaiki hasil delineasi bangunan, *clipping* untuk menyeleksi bangunan-bangunan yang berada didalam wilayah desa Campurejo, *merge* untuk memperbaiki hasil polygon yang bertampalan (*overlay*), dan terakhir proses *layouting* peta.

Setelah dilakukan pengolahan pasca pemrosesan dihasilkan jumlah *attribute polygon* bangunan yang ditunjukkan oleh Tabel 3. Titik validasi digunakan untuk menghitung selisih jumlah bangunan yang dianggap benar dan jumlah bangunan yang terdeteksi. Dari hasil metode Mask R-CNN untuk wilayah 1 terdapat 2102 bangunan dengan selisih 142 bangunan dengan data validasi dan untuk wilayah 2 terdapat 247 bangunan dengan selisih 57 bangunan dengan data validasi.

Tabel 3. Jumlah Bangunan

	Mask R-CNN	
	Wilayah 1	Wilayah 2
Jumlah bangunan (validasi)	2244	304
Jumlah bangunan (terdeteksi)	2102	247
Selisih	142	57

Terlepas dari pendeteksian otomatis, suatu kontrol kualitas dan analisa lebih lanjut oleh operator masih tetap dibutuhkan untuk menghilangkan dan memperbaiki kesalahan. Dalam penelitian ini peran manusia dalam melakukan perbaikan dan pengeditan secara manual masih banyak diperlukan. Meskipun masih diperlukan pengeditan manual namun ditinjau dari kecepatan ekstraksi bangunan, metode ini dapat dikatakan lebih cepat dibanding digitasi secara manual, terutama untuk menghasilkan peta skala besar pada area yang luas dan bangunan yang relatif banyak. Waktu yang cukup lama dibutuhkan dalam proses pembuatan model (*train model*).



Gambar 11. Peta Hasil Klasifikasi Sebaran Bangunan Desa Campurejo dengan Metode Mask R-CNN

Kesimpulan

Dalam penelitian ini dilakukan penggunaan metode Mask R-CNN untuk melakukan ekstraksi bangunan dari data citra UAV dengan memasukkan *training sample* sebagai data masukan. *Training sample* yang digunakan berasal dari digitasi manual. Setelah dilakukan deteksi bangunan secara otomatis, dilakukan validasi dan uji akurasi. Nilai akurasi yang didapat dari hasil pengolahan dengan metode Mask R-CNN ini masih belum

mampu memenuhi standar ketelitian dari Peraturan Kepala Badan Informasi Geospasial Nomor 15 Tahun 2014 untuk bangunan yaitu 85%. Proses pasca pemrosesan masih diperlukan dengan cara perbaikan secara manual oleh operator terlatih (operator) untuk memperbaiki kesalahan-kesalahan yang terjadi. Meskipun begitu, berdasarkan tingkat efisiensi waktu metode ini memiliki potensi yang besar dalam melakukan pendeteksian bangunan secara otomatis sehingga diharapkan mampu membantu proses percepatan penyediaan peta dasar di Indonesia.

Ucapan Terimakasih

Penulis mengucapkan terima kasih kepada Laboratorium Geospasial Departemen Teknik Geomatika Institut Teknologi Sepuluh Nopember yang telah mendukung dalam penyediaan sarana prasarana untuk mengerjakan penelitian ini.

Daftar Pustaka

- Gavali, P., & Banu, J. S. (2019). Deep convolutional neural network for image classification on CUDA platform. In *Deep Learning and Parallel Computing Environment for Bioengineering Systems* (pp. 99-122). Academic Press.
- Hidayat, H., & Sukojo, B. M. (2017). Analysis of Horizontal Accuracy for Large Scale Rural Mapping Using Rotary Wing UAV Image. In *IOP Conference Series: Earth and Environmental Science* (Vol. 98, No. 1, p. 012052). IOP Publishing.
- Hu, F., Xia, G. S., Hu, J., & Zhang, L. (2015). Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sensing*, 7(11), 14680-14707.
- Khoshelham, K., Nardinocchi, C., Frontoni, E., Mancini, A., & Zingaretti, P. (2010). Performance evaluation of automated approaches to building detection in multi-source aerial data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1), 123-133.
- Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, 14(5), 778-782.
- Pemerintah Indonesia. 2002. *Undang-Undang Republik Indonesia Nomor 28 Tahun 2002 Tentang Bangunan Gedung*.
- Xie, J., & Zhou, J. (2017). Classification of urban building type from high spatial resolution remote sensing imagery using extended MRS and soft BP network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(8), 3515-3528.



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