

PEMETAAN TUTUPAN LAHAN MENGGUNAKAN RANDOM FOREST PADA PLATFORM GOOGLE EARTH ENGINE: STUDI KASUS IKN

Hanif Al Fathoni¹⁾

1) Program Studi Informatika Universitas Pembangunan Nasional Veteran Jawa Timur

E-mail: alfatoni922@gmail.com

Abstrak

Transformasi menuju pembangunan berkelanjutan di Ibu Kota Nusantara (IKN) memerlukan informasi spasial yang akurat terkait kondisi tutupan lahan. Penelitian ini bertujuan untuk memetakan tutupan lahan menggunakan algoritma Random Forest (RF) yang diimplementasikan pada platform cloud computing Google Earth Engine (GEE). Data yang digunakan berupa citra Sentinel-2 Level-2A tahun 2024 dengan resolusi spasial 10 meter yang mencakup kawasan inti IKN. Data sampel sebanyak 500 titik dibagi menjadi 70% untuk pelatihan dan 30% untuk validasi, mencakup lima kelas utama tutupan lahan. Proses klasifikasi dilakukan dengan memanfaatkan band spektral utama serta indeks vegetasi (NDVI) dan indeks air (NDWI). Hasil klasifikasi menunjukkan bahwa algoritma RF menghasilkan akurasi keseluruhan (Overall Accuracy) sebesar 81,29% dengan nilai Kappa Coefficient sebesar 0,7641. Peta hasil klasifikasi berhasil membedakan lima kelas utama, yaitu vegetasi alami, vegetasi budidaya, lahan terbangun, lahan terbuka, dan badan air. Penelitian ini membuktikan bahwa integrasi machine learning khususnya algoritma Random Forest dan platform GEE mampu menyediakan informasi spasial yang akurat, efisien, dan dapat direplikasi, sehingga relevan untuk mendukung perencanaan tata ruang, konservasi lingkungan, serta pembangunan berkelanjutan di kawasan IKN.

Kata kunci: Random Forest, Google Earth Engine, Tutupan Lahan, IKN, Machine Learning.

Pendahuluan

Pembangunan berkelanjutan (Sustainable Development) menuntut keseimbangan antara aspek ekonomi, sosial, dan lingkungan. Salah satu indikator penting dalam mendukung kebijakan pembangunan berkelanjutan adalah informasi mengenai tutupan lahan (land cover), yang mencerminkan kondisi fisik permukaan bumi serta dinamika perubahan penggunaan lahan. Sebagaimana dijelaskan oleh Zhu et al. [1], penginderaan jauh berperan penting dalam observasi dan karakterisasi perubahan tutupan lahan untuk mendukung pencapaian Sustainable Development Goals (SDGs). Dalam konteks Indonesia, pembangunan Ibu Kota Nusantara (IKN) sebagai proyek strategis nasional memerlukan ketersediaan data spasial yang akurat dan terkini untuk memantau perubahan lingkungan akibat aktivitas konstruksi yang masif. Rahyojati dan Handiani [2] melaporkan bahwa hingga tahun 2022 kawasan IKN masih didominasi oleh hutan sebesar 91% dari total area, sedangkan area terbangun hanya sekitar 0,4%. Namun, meningkatnya intensitas pembangunan menuntut pemantauan perubahan tutupan lahan secara periodik guna menjaga keseimbangan ekosistem dan mendukung perencanaan tata ruang yang berkelanjutan.

Penelitian ini mengusulkan pemanfaatan platform cloud computing Google Earth Engine (GEE) yang dikombinasikan dengan algoritma machine learning Random Forest (RF). GEE menyediakan akses terhadap katalog data satelit global seperti Sentinel-2, Landsat 8, dan MODIS, serta kemampuan komputasi awan untuk pemrosesan data geospasial berskala besar. Berbagai studi telah menunjukkan keunggulan integrasi GEE dan RF dalam klasifikasi citra satelit dari segi akurasi dan efisiensi, seperti yang ditunjukkan oleh Agaventa et al. [3], Yuliana dan Rizqiana [4], serta Rizaldi et al. [5]. Algoritma RF sendiri dikenal efektif dalam menangani data berdimensi tinggi, tahan terhadap overfitting, dan mampu memproses data multispektral secara paralel di lingkungan komputasi awan seperti GEE [6]. Oleh karena itu, kombinasi GEE dan RF dipilih sebagai pendekatan utama dalam penelitian ini untuk menghasilkan peta tutupan lahan yang akurat dan efisien guna mendukung perencanaan tata ruang, konservasi lingkungan, serta pembangunan berkelanjutan di kawasan IKN.

Studi Pustaka

Klasifikasi Tutupan Lahan

Tutupan lahan (land cover) merupakan representasi kondisi fisik permukaan bumi yang mencakup berbagai elemen seperti vegetasi, badan air, lahan terbangun, dan lahan terbuka. Informasi ini umumnya diperoleh melalui analisis citra satelit menggunakan teknik klasifikasi penggunaan dan tutupan lahan (Land Use/Land Cover atau LULC) [7]. Klasifikasi LULC bertujuan untuk mengelompokkan piksel citra berdasarkan keseragaman nilai spektralnya sehingga dapat menggambarkan distribusi dan perubahan elemen fisik di permukaan bumi. Hasil klasifikasi ini menjadi dasar dalam memahami pola perubahan lingkungan dan dinamika pemanfaatan lahan dari waktu ke waktu.

Algoritma Random Forest

Random Forest merupakan algoritma klasifikasi terawasi (supervised classification) berbasis konsep ensemble learning, yang menggabungkan beberapa model klasifikasi untuk meningkatkan akurasi prediksi [8]. Diperkenalkan oleh Leo Breiman pada tahun 2001, algoritma ini membangun sejumlah pohon keputusan (decision trees) dari subset data yang diambil secara acak, kemudian menentukan hasil akhir melalui mekanisme voting mayoritas untuk klasifikasi atau rata-rata untuk regresi. Pendekatan ini membuat Random Forest lebih akurat, stabil, dan tahan terhadap overfitting, sehingga banyak digunakan dalam pemetaan tutupan lahan berbasis penginderaan jauh yang melibatkan data multispektral berdimensi tinggi.

Google Earth Engine

Google Earth Engine (GEE) merupakan platform komputasi awan yang dirancang untuk pemrosesan dan analisis data penginderaan jauh dalam skala besar. Platform ini menyediakan akses ke berbagai sumber data satelit seperti Landsat, Sentinel, dan MODIS, sehingga memungkinkan pengguna untuk melakukan analisis spasial secara komprehensif. GEE dilengkapi dengan berbagai algoritma klasifikasi seperti CART, Random Forest, Naive Bayes, dan Support Vector Machine (SVM) yang efektif untuk pemetaan tutupan lahan secara terawasi. Kemampuannya dalam mengelola data secara efisien serta fleksibilitas untuk mengembangkan algoritma khusus menjadikan GEE alat yang kuat untuk mendukung penelitian dan pengambilan keputusan berbasis data spasial secara berkelanjutan [9].

Sentinel-2

Satelit Sentinel-2 merupakan bagian dari Program Copernicus yang dikembangkan oleh European Space Agency (ESA) dan diluncurkan pada tahun 2015 untuk mendukung pengamatan Bumi melalui citra multispektral beresolusi tinggi. Satelit ini memiliki 13 pita spektral dengan resolusi spasial 10, 20, hingga 60 meter, yang dirancang untuk menangkap beragam informasi permukaan seperti vegetasi, tanah, badan air, dan infrastruktur buatan manusia. Salah satu keunggulan utama Sentinel-2 adalah ketersediaan data yang terbuka dan gratis, termasuk melalui platform Google Earth Engine (GEE). Dalam penelitian ini, indeks spectral pada citra Sentinel-2 dimanfaatkan untuk memperkuat proses klasifikasi, indeks spectral yang digunakan meliputi Normalized Difference Vegetation Index (NDVI) untuk mengidentifikasi kerapatan vegetasi, Normalized Difference Built-up Index (NDBI) untuk mendeteksi area terbangun, dan Normalized Difference Water Index (NDWI) untuk mengenali badan air. Ketiga indeks tersebut dihitung menggunakan kombinasi pita spektral Band 8 (NIR), Band 4 (Red), Band 11 (SWIR), dan Band 3 (Green), yang membantu meningkatkan akurasi dalam membedakan jenis tutupan lahan berdasarkan karakteristik reflektansi permukaan bumi.

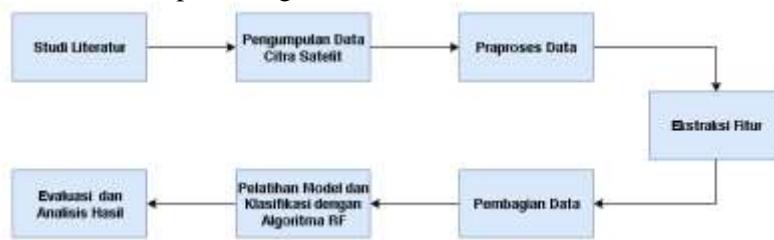
Confusion Matrix dan Evaluasi Akurasi

Confusion Matrix merupakan tabel yang digunakan untuk mengevaluasi kinerja model klasifikasi dengan membandingkan hasil prediksi model terhadap label sebenarnya dari data [10]. Tabel ini menampilkan empat komponen utama, yaitu True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN), yang menggambarkan jumlah prediksi benar maupun salah pada setiap kelas. Berdasarkan nilai-nilai tersebut, dapat dihitung beberapa metrik evaluasi seperti Overall Accuracy, Precision, Recall, dan F1-Score. Overall Accuracy menunjukkan proporsi total prediksi yang benar terhadap seluruh data, sedangkan Precision mengukur tingkat ketepatan model dalam mengklasifikasikan data positif secara benar. Recall menunjukkan kemampuan model dalam mendeteksi semua data positif yang seharusnya dikenali, sementara F1-Score merupakan rata-rata harmonik dari Precision dan Recall yang menggambarkan keseimbangan antara keduanya.

Metodologi Penelitian

Tahapan Penelitian

Penelitian ini memiliki tahapan sebagai berikut.



Gambar 1. Tahapan penelitian

Seluruh proses di atas dijalankan melalui platform komputasi awan Google Colab dengan menggunakan API Python Google Earth Engine.

Studi Literatur

Tahap studi literatur dilakukan untuk mengumpulkan informasi dan dasar teori yang relevan dengan topik penelitian. Kegiatan ini mencakup penelusuran jurnal ilmiah, artikel, dan sumber akademik lain yang membahas penginderaan jauh, tutupan lahan, Google Earth Engine, serta algoritma Random Forest. Tujuannya adalah untuk memahami metode dan pendekatan yang telah digunakan pada penelitian sebelumnya sebagai dasar dalam merancang metodologi penelitian ini.

Pengumpulan Data Citra Satelit

Data utama yang digunakan dalam penelitian ini berupa citra satelit Sentinel-2 Level-2A tahun 2024 yang diakses melalui platform Google Earth Engine (GEE). Sentinel-2 dipilih karena memiliki resolusi spasial tinggi (10–20 meter) serta menyediakan 13 pita spektral yang berguna untuk analisis permukaan bumi. Lokasi penelitian merupakan Kawasan Inti IKN dengan area seluas 56.181

Praproses Data

Tahap praproses dilakukan untuk memastikan kualitas citra yang digunakan dalam klasifikasi. Proses ini meliputi pemotongan citra sesuai area penelitian (cropping), koreksi awan dan bayangan awan menggunakan teknik cloud masking dengan QA Band QA60, serta pembuatan komposit median untuk mengurangi gangguan temporal. Hasil dari tahap ini adalah citra bersih yang siap digunakan untuk analisis lebih lanjut.

Ekstraksi Fitur

Pada tahap ekstraksi fitur, dilakukan perhitungan beberapa indeks spektral seperti Normalized Difference Vegetation Index (NDVI), Normalized Difference Built-up Index (NDBI), dan Normalized Difference Water Index (NDWI). Indeks-indeks ini ditambahkan sebagai fitur tambahan pada dataset karena dapat memperjelas perbedaan antara vegetasi, lahan terbangun, dan badan air. Tahapan ini bertujuan untuk meningkatkan akurasi klasifikasi yang dilakukan oleh algoritma Random Forest.

Pembagian Data

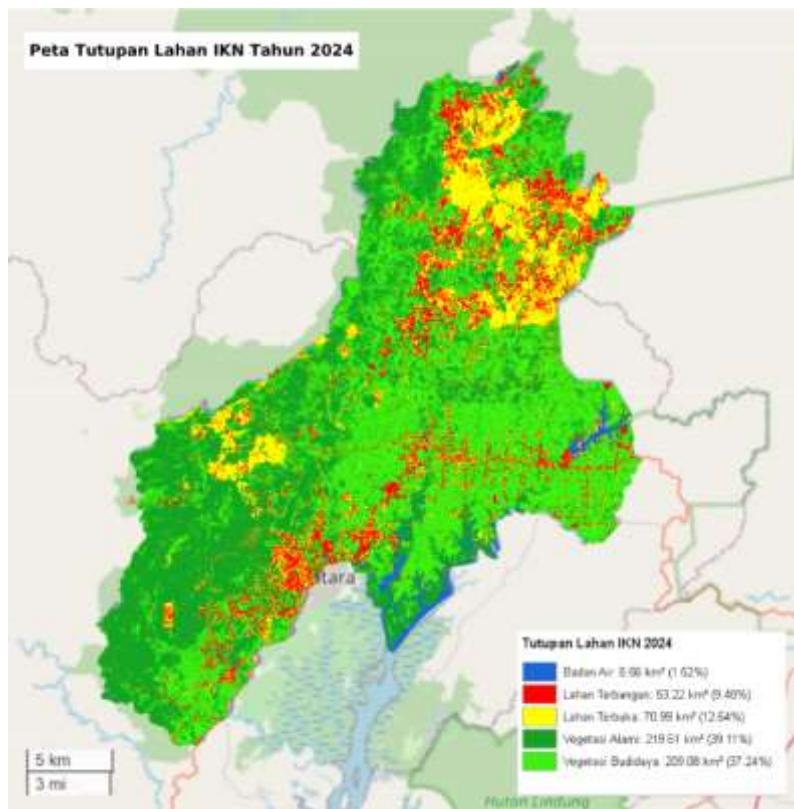
Setelah fitur berhasil diekstraksi, dataset dibagi menjadi dua bagian: data latih (training data) sebesar 70% dan data uji (testing data) sebesar 30%. Data latih digunakan untuk membangun model klasifikasi, sedangkan data uji digunakan untuk mengevaluasi performa model. Pembagian data ini dilakukan secara acak agar hasil klasifikasi tidak bias terhadap distribusi data tertentu.

Evaluasi dan Analisis Hasil

Tahapan akhir penelitian adalah evaluasi dan analisis hasil klasifikasi. Hasil peta tutupan lahan yang dihasilkan oleh algoritma Random Forest dibandingkan dengan data referensi untuk menghitung metrik evaluasi seperti Overall Accuracy, Precision, Recall, dan F1-Score. Selain itu, nilai Kappa Coefficient juga dihitung untuk mengukur tingkat kesesuaian antara hasil klasifikasi dan data sebenarnya. Tahap ini bertujuan untuk menilai sejauh mana metode yang digunakan mampu menghasilkan peta tutupan lahan yang akurat dan dapat diandalkan.

Hasil dan Pembahasan

Peta Tutupan Lahan IKN Tahun 2024



Gambar 2. Hasil peta tutupan lahan IKN tahun 2024

Proses klasifikasi tutupan lahan di kawasan Ibu Kota Nusantara (IKN) tahun 2024 dilakukan menggunakan algoritma Random Forest (RF) pada platform Google Earth Engine (GEE). Berdasarkan hasil klasifikasi, diperoleh lima kelas utama tutupan lahan yaitu vegetasi alami, vegetasi budidaya, lahan terbuka, lahan terbangun, dan badan air. Dari hasil analisis spasial, diketahui bahwa kelas vegetasi alami mendominasi area penelitian dengan luas 219,61 km² (39,11%), diikuti oleh vegetasi budidaya seluas 209,08 km² (37,24%), lahan terbuka sebesar 70,99 km² (12,64%), lahan terbangun sebesar 53,22 km² (9,48%), dan badan air seluas 8,56 km² (1,52%) dari total area penelitian seluas 561,47 km². Dominasi vegetasi alami dan budidaya menunjukkan bahwa wilayah IKN masih didominasi oleh tutupan lahan hijau, meskipun aktivitas pembangunan infrastruktur telah mulai berkembang di beberapa titik tertentu.

Evaluasi Kinerja Random Forest

Evaluasi kinerja model dilakukan menggunakan metrik Overall Accuracy (OA), Kappa Coefficient, Precision, Recall, dan F1-Score, seperti yang ditunjukkan pada diagram pada Tabel 1 berikut.

Tabel 1. Kinerja Model RF

Overall Accuracy	Kappa	Precision	Recall	F1-Score
81,29%	0,7641	81,71%	81,29%	81,32%

Berdasarkan table hasil kinerja di atas, model RF menghasilkan OA sebesar 81,29% dan Kappa Coefficient sebesar 0,7641, yang mengindikasikan tingkat kesesuaian yang baik antara hasil klasifikasi dengan data referensi. Nilai Precision (81,71%), Recall (81,29%), dan F1-Score (81,32%) menunjukkan bahwa model memiliki konsistensi yang baik dalam membedakan tiap kelas tutupan lahan. Secara umum, nilai akurasi di atas 80% menandakan bahwa algoritma Random Forest mampu mengklasifikasikan citra Sentinel-2 dengan performa tinggi untuk area yang luas dan kompleks seperti IKN.

Analisis Akurasi per Kelas

Tabel 2. Confusion Matrix per Kelas

Kelas Aktual (True Label)	Badan Air	Lahan Terbangun	Lahan Terbuka	Vegetasi Alami	Vegetasi Budidaya
Badan Air	100.0% (31)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)
Lahan Terbangun	0.0% (0)	75.0% (15)	25.0% (5)	0.0% (0)	0.0% (0)
Lahan Terbuka	3.4% (1)	27.6% (8)	69.0% (20)	0.0% (0)	0.0% (0)
Vegetasi Alami	0.0% (0)	0.0% (0)	0.0% (0)	77.3% (17)	22.7% (5)
Vegetasi Budidaya	0.0% (0)	0.0% (0)	0.0% (0)	18.9% (7)	81.1% (30)

Berdasarkan hasil confusion matrix, tingkat akurasi tiap kelas menunjukkan variasi yang berbeda. Kelas badan air memiliki nilai akurasi produsen (*producer accuracy*) tertinggi sebesar 100%, menunjukkan kemampuan model dalam mengenali badan air dengan sangat baik. Namun, kelas vegetasi alami (68,97%), vegetasi budidaya (77,27%), dan lahan terbuka (75%) menunjukkan adanya kesalahan klasifikasi yang disebabkan oleh kemiripan spektral antar jenis vegetasi. Kelas lahan terbangun memiliki nilai akurasi yang relatif rendah akibat campuran piksel dengan lahan terbuka dan vegetasi di sekitarnya. Hal ini sejalan dengan kondisi di lapangan, di mana area terbangun di IKN masih tersebar dan belum sepenuhnya berkembang secara homogen.

Kesimpulan

Berdasarkan hasil penelitian pemetaan tutupan lahan menggunakan algoritma Random Forest pada platform Google Earth Engine (GEE) di kawasan Ibu Kota Nusantara (IKN), dapat disimpulkan bahwa model klasifikasi yang digunakan menunjukkan performa yang baik dengan nilai Overall Accuracy sebesar 81,29% dan Kappa Coefficient sebesar 0,7641, yang mengindikasikan tingkat kesesuaian yang kuat antara hasil klasifikasi dengan data referensi. Nilai metrik evaluasi lainnya, yaitu Precision (81,71%), Recall (81,29%), dan F1-Score (81,32%), memperlihatkan konsistensi model dalam mengklasifikasikan tiap kelas tutupan lahan.

Hasil klasifikasi berhasil membedakan lima kelas utama tutupan lahan, yaitu vegetasi alami (39,11%), vegetasi budidaya (37,24%), lahan terbuka (12,64%), lahan terbangun (9,48%), dan badan air (1,52%), yang menunjukkan bahwa wilayah IKN masih didominasi oleh area vegetasi alami dan budidaya. Berdasarkan hasil confusion matrix, kelas badan air memiliki tingkat akurasi tertinggi (producer accuracy 100%), sedangkan lahan terbangun dan lahan terbuka mengalami kesalahan klasifikasi akibat kemiripan spektral dengan area vegetasi di sekitarnya.

Secara keseluruhan, kombinasi Google Earth Engine dan algoritma Random Forest terbukti efektif dalam menghasilkan peta tutupan lahan yang akurat, efisien, dan mudah direplikasi, khususnya untuk wilayah dengan cakupan luas seperti kawasan IKN.

Daftar Pustaka

- [1] Z. Zhu, S. Qiu, dan S. Ye, “Remote sensing of land change: A multifaceted perspective,” *Remote Sensing of Environment*. Elsevier, 2022. [Daring]. Tersedia pada: <https://www.sciencedirect.com/science/article/pii/S0034425722003728>
- [2] A. RAHYOJATI dan D. HANDIANI, “Analisis Tutupan Lahan di Ibu Kota Nusantara (IKN) dengan Memanfaatkan Data Land Cover Sentinel-2,” *Prosiding FTSP Series*. 2024.
- [3] C. D. Agaventa, H. Rumapea, dan ..., “PEMANFAATAN GOOGLE EARTH ENGINE DAN ALGORITMA RANDOM FOREST UNTUK PEMETAAN LAHAN PERKEBUNAN JERUK,” ... *J. Tek. Inform.* ..., 2025, [Daring]. Tersedia pada: <https://ejurnal.methodist.ac.id/index.php/methodika/article/view/4078>
- [4] H. Yuliana dan Z. C. H. Rizqiana, “Analisis Metode Klasifikasi Pemetaan Tutupan Lahan (Land Cover) di Area Kota Bandung Menggunakan Algoritma Random Forest Pada Google Earth Engine,” *Epsil. J. Electr.* ..., 2024, [Daring]. Tersedia pada: <http://epsilon.unjani.ac.id/index.php/epsilon/article/view/119>
- [5] A. Rizaldi, A. Darmawan, H. Kaskoyo, dan ..., “Pemanfaatan Google Earth Engine untuk pemantauan lahan agroforestri dalam skema perhutanan sosial,” *Majalah Geografi* core.ac.uk, 2022. [Daring]. Tersedia pada: <https://core.ac.uk/download/pdf/537346553.pdf>

- [6] D. Apriyanti, I. Layali, M. Gomareuzzaman, dan ..., “MONITORING PERUBAHAN TUTUPAN LAHAN KABUPATEN KLATEN TAHUN 2019 DAN 2023 SELAMA PEMBANGUNAN JALAN TOL YOGYAKARTA–SOLO ...,” ... *J. Geod. Dan ...*, 2025, [Daring]. Tersedia pada: <https://ejournal2.undip.ac.id/index.php/elipsoida/article/view/25698>
- [7] M. W. Naikoo, M. Rihan, dan M. Ishtiaque, “Analyses of land use land cover (LULC) change and built-up expansion in the suburb of a metropolitan city: Spatio-temporal analysis of Delhi NCR using ...,” *Journal of Urban Management*. Elsevier, 2020. [Daring]. Tersedia pada: <https://www.sciencedirect.com/science/article/pii/S2226585620300637>
- [8] C. A. Ramezan, T. A. Warner, A. E. Maxwell, dan B. S. Price, “Effects of training set size on supervised machine-learning land-cover classification of large-area high-resolution remotely sensed data,” *Remote Sensing*. mdpi.com, 2021. [Daring]. Tersedia pada: <https://www.mdpi.com/2072-4292/13/3/368>
- [9] A. Shelestov, M. Lavreniuk, N. Kussul, dan ..., “Exploring Google Earth Engine platform for big data processing: Classification of multi-temporal satellite imagery for crop mapping,” *frontiers in Earth* frontiersin.org, 2017. doi: 10.3389/feart.2017.00017.
- [10] D. Normawati dan S. A. Prayogi, “Implementasi Naïve Bayes classifier dan confusion matrix pada analisis sentimen berbasis teks pada Twitter,” *J-SAKTI J. Sains Komput. Dan ...*, 2021, [Daring]. Tersedia pada: <http://www.tunasbangsa.ac.id/ejurnal/index.php/jsakti/article/view/369>