

## ***Analisis Metode Klasifikasi Pemetaan Tutupan Lahan (Land Cover) di Area Kota Bandung Menggunakan Algoritma Random Forest Pada Google Earth Engine***

**Hajiar Yuliana<sup>1\*)</sup>, Zahra Cahya Hanifa Rizqiana<sup>2)</sup>**

<sup>1,2)</sup>Program Studi Teknik Elektro

Universitas Jenderal Achmad Yani

Jalan Terusan Jend. Sudirman PO.BOX 148 Cimahi 40531

<sup>2)</sup>Program Studi Teknik Mesin, Universitas Jenderal Achmad Yani

<sup>\*)</sup>Korespondensi : hajiar.yuliana@lecture.unjani.ac.id

### **Abstrak**

Penelitian ini bertujuan untuk memetakan tutupan lahan di Kota Bandung menggunakan algoritma *Random Forest* (RF) yang diimplementasikan pada platform berbasis *cloud Google Earth Engine* (GEE). Data citra satelit Sentinel-2 digunakan untuk menganalisis empat kelas utama tutupan lahan, yaitu lahan permukiman, lahan hijau, perairan, dan lahan terbuka. Proses klasifikasi melibatkan pemrosesan data awal, pelatihan model menggunakan data sampel, serta evaluasi akurasi melalui *confusion matrix* dan validasi silang (*cross-validation*). Hasil penelitian menunjukkan bahwa algoritma RF memiliki akurasi keseluruhan sebesar 89%, dengan akurasi tertinggi pada kelas lahan permukiman (92%) dan terendah pada kelas perairan (80%). Validasi silang menunjukkan performa yang stabil dengan rata-rata akurasi 88.5%, *precision* 0.91, *recall* 0.88, dan *F1-score* 0.89. Analisis *confusion matrix* mengidentifikasi kesalahan klasifikasi pada kelas-kelas tertentu akibat tumpang tindih spektral, terutama antara lahan hijau dan area terbuka. Penelitian ini membuktikan bahwa algoritma RF pada GEE merupakan metode yang efisien dan akurat untuk klasifikasi tutupan lahan, sekaligus mendukung perencanaan tata ruang dan pengelolaan lingkungan. Pengembangan lebih lanjut dapat mencakup penggunaan data resolusi lebih tinggi, algoritma pembelajaran lanjutan, dan analisis berbasis waktu untuk memahami dinamika perubahan tutupan lahan.

**Kata kunci :** *Random Forest, Google Earth Engine, Sentinel-2, klasifikasi tutupan lahan*

### **Abstract**

*This research aims to map land cover in Bandung City using the Random Forest (RF) algorithm implemented on the Google Earth Engine (GEE) cloud-based platform. Sentinel-2 satellite image data was used to analyze four main classes of land cover, namely residential land, green land, water, and open land. The classification process involved initial data processing, model training using sample data, and accuracy evaluation through confusion matrix and cross-validation. The results showed that the RF algorithm had an overall accuracy of 89%, with the highest accuracy in the residential land class (92%) and the lowest in the water class (80%). Cross-validation showed stable performance with an average accuracy of 88.5%, precision 0.91, recall 0.88, and F1-score 0.89. Confusion matrix analysis identified misclassification in certain classes due to spectral overlap, especially between green land and open area. This research proves that the RF algorithm in GEE is an efficient and accurate method for land cover classification, while supporting spatial planning and environmental management. Further developments could include the use of higher resolution data, advanced learning algorithms and time-based analysis to understand the dynamics of land cover change.*

**Keywords :** *Random Forest, Google Earth Engine, Sentinel-2, Land Cover*

## **I. PENDAHULUAN**

Perubahan tutupan lahan di daerah perkotaan menjadi perhatian utama dalam studi lingkungan, karena dampaknya yang signifikan terhadap ekosistem, perubahan iklim, serta perencanaan pembangunan wilayah. Kota Bandung, sebagai salah satu kota terbesar di Indonesia, mengalami dinamika perubahan

tutupan lahan yang cepat, seiring dengan pertumbuhan penduduk dan urbanisasi. Oleh karena itu, pemetaan tutupan lahan secara akurat dan efisien sangat diperlukan untuk mendukung pengelolaan lingkungan yang berkelanjutan. Salah satu pendekatan yang digunakan dalam pemetaan tutupan lahan adalah pemanfaatan citra satelit yang dapat menyediakan data spasial dalam skala besar dan resolusi tinggi.

Seiring dengan perkembangan teknologi, penggunaan algoritma *machine learning*, khususnya algoritma *Random Forest* (RF), telah menjadi metode populer dalam klasifikasi citra satelit. Algoritma ini dikenal karena kemampuannya dalam menangani data dengan dimensi tinggi serta ketahanannya terhadap overfitting, yang sangat berguna dalam klasifikasi citra dengan berbagai kelas tutupan lahan. *Google Earth Engine* (GEE) telah menjadi platform yang mempermudah penerapan algoritma *machine learning* dalam analisis citra satelit, dengan menyediakan akses ke dataset global serta kemampuan komputasi cloud yang besar. Penggunaan GEE dalam penelitian pemetaan tutupan lahan di Kota Bandung diharapkan dapat memberikan solusi praktis dan efisien dalam menghasilkan peta tutupan lahan yang akurat.

Penggunaan GEE sebagai platform untuk pemrosesan citra satelit memungkinkan analisis yang lebih efisien dan cepat. GEE menyediakan akses ke berbagai dataset satelit, termasuk citra dari Landsat dan Sentinel, yang dapat digunakan untuk klasifikasi tutupan lahan. Sebagai contoh, penelitian oleh Phan et al. menunjukkan bahwa kombinasi data Landsat 8 dengan algoritma RF dapat menghasilkan peta tutupan lahan dengan akurasi yang tinggi, mencapai lebih dari 84% [1]. Selain itu, Mahdavi et al. melaporkan bahwa akurasi peta yang dihasilkan menggunakan algoritma RF di GEE juga menunjukkan hasil yang memuaskan, meskipun sedikit di bawah 91% [2].

Algoritma RF memiliki keunggulan dalam menangani data yang kompleks dan multitemporal. Dalam studi oleh Eisavi et al., penggunaan data spektral dan termal multitemporal dengan RF menunjukkan peningkatan akurasi dalam pemetaan tutupan lahan [3]. Hal ini penting untuk Kota Bandung, yang mengalami perubahan tutupan lahan yang cepat akibat urbanisasi dan konversi lahan. Penelitian oleh Kuntoro et al. juga menyoroti dampak perubahan penggunaan lahan terhadap aliran sungai, yang menunjukkan pentingnya pemetaan tutupan lahan untuk manajemen sumber daya air [4]. Selain itu, Algoritma RF menunjukkan performa yang lebih baik dibandingkan dengan metode lain seperti Classification and Regression Trees (CART) dan Gradient Boosting (GB) dalam beberapa studi [5]. Namun, penting untuk dicatat bahwa pemilihan dataset dan metode komposisi citra sangat mempengaruhi hasil klasifikasi [6].

Penelitian terdahulu juga menunjukkan bahwa algoritma RF telah berhasil diterapkan dalam berbagai studi pemetaan tutupan lahan, baik di daerah perkotaan maupun di wilayah lainnya [7]. Algoritma *Random Forest* (RF) telah terbukti sebagai metode yang kuat dan akurat untuk klasifikasi tutupan lahan, terutama ketika diterapkan pada platform *Google Earth Engine* (GEE) yang menyediakan data satelit multi-sumber dan layanan komputasi berkinerja tinggi [8], [9]. Algoritma *Random Forest* dipilih karena keunggulannya dalam menangani data yang kompleks dan kemampuannya untuk menentukan variabel penting dalam klasifikasi [10]. Parameter optimal untuk RF, seperti jumlah pohon, variabel per split, dan fraksi bagging, diatur untuk mencapai akurasi yang tinggi [9]. Namun, tantangan utama dalam pemetaan tutupan lahan di kota-kota besar adalah kompleksitas lingkungan yang melibatkan berbagai jenis tutupan seperti permukiman, area hijau, perairan, dan infrastruktur lainnya. Oleh karena itu, diperlukan kajian yang lebih mendalam mengenai penerapan algoritma RF di Kota Bandung, dengan memanfaatkan *Google Earth Engine* untuk mengoptimalkan klasifikasi tutupan lahan secara lebih akurat dan efisien.

---

Info Makalah:

Dikirim : 11-26-2025;  
Revisi 1 : 01-23-2025;  
Revisi 2 : 02-05-2025;  
Diterima : mm-dd-yy.

Penulis Korespondensi:

Telp : +62-812-2299-7113  
e-mail : hajiar.yuliana@lecture.unjani.ac.id

---

Selain itu, akurasi klasifikasi dapat dipengaruhi oleh kualitas data dan pemilihan fitur yang tepat. Jamali mengemukakan bahwa pemilihan fitur yang baik dalam algoritma RF dapat meningkatkan hasil klasifikasi [11]. Tantangan dalam pemetaan tutupan lahan juga mencakup masalah noise dalam data pelatihan. Pelletier et al. menyoroti bagaimana noise dalam label kelas pelatihan dapat mempengaruhi kinerja klasifikasi [12]. Oleh karena itu, penting untuk melakukan validasi yang cermat terhadap data pelatihan yang digunakan dalam analisis. Dalam konteks pemetaan tutupan lahan di Bandung, penting untuk mempertimbangkan

variabel lingkungan yang relevan, seperti suhu permukaan tanah dan pola penggunaan lahan sebelumnya, untuk meningkatkan akurasi klasifikasi.

Selain itu, permasalahan utama dalam pemetaan tutupan lahan di area perkotaan, seperti Kota Bandung, adalah tingginya kompleksitas dan dinamika perubahan tutupan lahan. Banyak faktor yang mempengaruhi perubahan ini, mulai dari aktivitas manusia, pembangunan infrastruktur, hingga perubahan iklim. Hal ini membuat klasifikasi citra satelit menjadi tugas yang menantang karena banyaknya kelas tutupan yang perlu dibedakan dan kecenderungan data yang tidak terbalancing antara kelas tutupan yang satu dengan yang lainnya. Untuk mengatasi hal ini, algoritma *machine learning* seperti *Random Forest* (RF) menawarkan kemampuan untuk melakukan klasifikasi dengan akurasi yang tinggi dan toleransi terhadap kerumitan data.

Pendekatan umum yang digunakan dalam penelitian ini adalah pemanfaatan algoritma *Random Forest* untuk klasifikasi citra satelit guna menghasilkan peta tutupan lahan yang akurat. Dalam hal ini, *Google Earth Engine* (GEE) sebagai platform berbasis cloud memberikan keuntungan dalam hal aksesibilitas dan kemampuan komputasi yang sangat tinggi, yang memungkinkan analisis data citra satelit secara efisien dan skalabel. GEE memungkinkan pemrosesan data satelit dalam jumlah besar, dengan waktu eksekusi yang lebih cepat dibandingkan metode tradisional, serta dapat memanfaatkan berbagai fitur algoritma *machine learning* yang ada. Dengan mengidentifikasi potensi dan tantangan yang ada dalam klasifikasi tutupan lahan di Kota Bandung, penelitian ini bertujuan untuk mengeksplorasi dan mengimplementasikan algoritma *Random Forest* pada platform *Google Earth Engine* untuk menghasilkan peta tutupan lahan yang lebih akurat dan tepat guna.

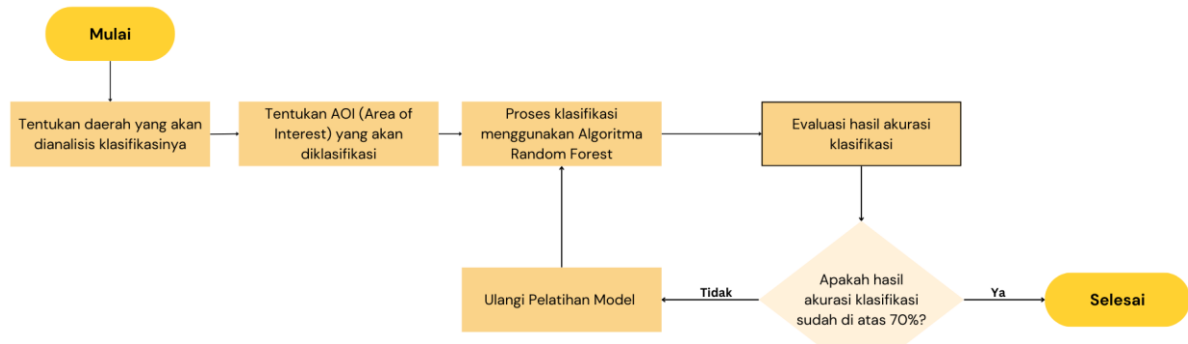
## **II. METODE**

### **A. Diagram Alir Penelitian**

Penelitian ini dimulai dengan menentukan daerah yang akan dianalisis klasifikasinya, dalam hal ini Kota Bandung dipilih sebagai area penelitian. Wilayah ini memiliki variasi tutupan lahan yang cukup beragam, seperti lahan hijau, perairan, permukiman, dan area terbuka, sehingga cocok untuk pengujian algoritma klasifikasi. Langkah ini mencakup pengumpulan informasi spasial dan identifikasi *Area of Interest* (AOI) yang menjadi fokus klasifikasi. Tahap selanjutnya adalah menentukan *Area of Interest* (AOI) yang akan diklasifikasikan. AOI dipilih berdasarkan batas administrasi Kota Bandung yang relevan dan memastikan representasi berbagai jenis tutupan lahan dalam dataset. Proses ini dilakukan dengan menggunakan perangkat lunak pengolahan citra satelit berbasis *cloud*, yaitu *Google Earth Engine* (GEE), untuk mengimpor data citra Sentinel-2.

Setelah AOI ditentukan, penelitian memasuki proses klasifikasi menggunakan algoritma *Random Forest*. Algoritma ini dipilih karena kemampuannya yang andal dalam menangani data yang kompleks dan menghasilkan model yang *robust*. Data citra Sentinel-2 diproses menggunakan indeks vegetasi seperti NDVI dan EVI, serta fitur-fitur lain yang relevan untuk melatih model klasifikasi. Model dilatih menggunakan data sampel dari setiap kelas tutupan lahan yang diidentifikasi dalam AOI. Tahap berikutnya adalah evaluasi hasil akurasi klasifikasi, di mana model yang dihasilkan diuji menggunakan data validasi. Evaluasi dilakukan melalui matriks kesalahan (*confusion matrix*) untuk menghitung metrik akurasi, *precision*, *recall*, dan *F1-score*. Penilaian ini bertujuan memastikan bahwa model dapat mengklasifikasikan data dengan tingkat keandalan yang tinggi.

Jika hasil akurasi yang diperoleh dari model berada di bawah 70%, proses pelatihan model akan diulang dengan menyesuaikan parameter algoritma atau memperbaiki kualitas data latih. Namun, jika akurasi mencapai atau melebihi 70%, proses klasifikasi dinyatakan selesai, dan hasil klasifikasi dapat digunakan untuk keperluan analisis lebih lanjut. Tahapan-tahapan ini dirancang untuk memastikan bahwa penelitian dilakukan secara sistematis dan menghasilkan model klasifikasi yang akurat serta dapat diandalkan untuk pemetaan tutupan lahan di Kota Bandung.



Gambar 1. Diagram Alir Penelitian

### B. Deskripsi Data Yang Digunakan

Penelitian ini menggunakan citra satelit sebagai data utama untuk klasifikasi tutupan lahan di Kota Bandung. Citra satelit yang digunakan adalah data Landsat yang tersedia melalui platform *Google Earth Engine* (GEE), yang mencakup periode waktu terbaru untuk menggambarkan dinamika perubahan tutupan lahan. Data Landsat memiliki resolusi spasial 30 meter dan menyediakan informasi spektral dalam berbagai band, yang sangat berguna untuk membedakan berbagai kelas tutupan lahan. Citra satelit yang digunakan dalam penelitian ini diambil dari koleksi Landsat 8, yang memiliki cakupan temporal dari 2013 hingga 2023.

Selain itu, data tambahan berupa informasi referensi berupa peta tutupan lahan yang sudah ada dan titik data ground truth juga digunakan untuk validasi dan pelatihan model. Data referensi ini dikumpulkan dari survei lapangan dan sumber peta tutupan lahan yang sudah dipublikasikan, seperti data dari Badan Informasi Geospasial (BIG) dan sumber pemerintah daerah setempat.

### C. Persiapan Sampel Data

Proses persiapan sampel untuk klasifikasi dimulai dengan pengumpulan citra satelit Landsat yang relevan untuk area Kota Bandung. Citra tersebut kemudian diproses untuk mendapatkan data yang siap digunakan dalam model klasifikasi. Tahapan awal pemrosesan meliputi pemotongan citra (clipping) untuk area penelitian yang mencakup seluruh wilayah Kota Bandung. Selain itu, citra satelit juga diproses untuk koreksi atmosfer (atmospheric correction) guna mengurangi pengaruh elemen atmosfer terhadap data citra. Proses ini penting agar citra satelit mencerminkan keadaan sesungguhnya di lapangan.

Setelah itu, citra yang telah diproses dilakukan pemilihan band yang relevan, yaitu band spektral yang dapat membedakan berbagai jenis tutupan lahan, seperti band merah (red), hijau (green), biru (blue), inframerah dekat (near-infrared), dan inframerah pendek (short-wave infrared). Pemilihan band ini didasarkan pada literatur yang menunjukkan efektivitasnya dalam membedakan kelas-kelas tutupan lahan perkotaan, seperti vegetasi, permukiman, perairan, dan lahan terbuka (Cohen et al., 2021).

Selanjutnya, data pelatihan atau training data dikumpulkan dari titik-titik ground truth dan peta referensi untuk membangun label yang akan digunakan untuk melatih model *Random Forest*. Titik data ini mencakup berbagai jenis tutupan lahan yang ada di Kota Bandung, seperti lahan permukiman, area hijau, jalan raya, dan perairan. Titik data pelatihan ini kemudian digunakan untuk melatih model *Random Forest* dalam mengklasifikasikan citra satelit menjadi kelas-kelas tutupan lahan yang relevan.

### D. Pengaturan Percobaan

Metode yang diterapkan dalam penelitian ini adalah algoritma *Random Forest* (RF), yang diterapkan pada data citra satelit menggunakan platform *Google Earth Engine*. *Random Forest* merupakan algoritma ensemble yang membangun sejumlah pohon keputusan (decision trees) dan melakukan klasifikasi berdasarkan mayoritas keputusan yang diberikan oleh masing-masing pohon. Algoritma RF ini dipilih karena kemampuannya untuk menangani data yang besar dan kompleks, serta memiliki performa yang baik dalam hal akurasi klasifikasi. Di dalam GEE, model RF diimplementasikan menggunakan library built-in yang memfasilitasi pelatihan model serta evaluasi kinerja.

Proses eksperimen dimulai dengan melatih model *Random Forest* menggunakan data pelatihan yang telah disiapkan sebelumnya. Proses pelatihan ini dilakukan dengan menggunakan sampel data citra satelit yang telah diberi label sesuai dengan kelas tutupan lahan yang diinginkan. Selanjutnya, model yang terlatih digunakan untuk mengklasifikasikan citra satelit yang tidak diberi label untuk menghasilkan peta

tutupan lahan. Parameter yang digunakan dalam pelatihan meliputi jumlah pohon dalam hutan (number of trees) serta kedalaman maksimum pohon (maximum depth), yang dioptimalkan melalui uji coba untuk mendapatkan model dengan kinerja terbaik.

#### ***E. Evaluasi Kinerja***

Untuk mengevaluasi kinerja model yang telah dilatih, dilakukan pengujian menggunakan data validasi yang tidak terlibat dalam proses pelatihan. Evaluasi dilakukan dengan membandingkan hasil klasifikasi model dengan data *ground truth* dan peta referensi. Metode pengukuran yang digunakan untuk menilai akurasi model adalah *Confusion Matrix*, yang memungkinkan perhitungan akurasi keseluruhan, serta metrik lain seperti *Precision*, *Recall*, dan *F1-Score* untuk masing-masing kelas tutupan lahan. Pengukuran ini penting untuk memastikan model dapat mengklasifikasikan dengan baik berbagai kelas tutupan lahan yang ada di area Kota Bandung.

Berikut adalah penjelasan masing-masing metrik evaluasi yang digunakan pada penelitian ini:

- **Akurasi (%)**: Mengukur seberapa besar proporsi prediksi yang benar dibandingkan dengan jumlah total prediksi. Akurasi yang tinggi menunjukkan bahwa model secara keseluruhan dapat mengklasifikasikan citra dengan benar.
- ***Precision***: Mengukur akurasi prediksi positif. Ini menunjukkan berapa banyak dari semua prediksi yang dikategorikan sebagai positif (misalnya, lahan hijau atau lahan permukiman) yang benar-benar termasuk dalam kategori tersebut.
- ***Recall***: Mengukur sejauh mana model dapat menemukan semua kasus positif dalam data yang diuji. *Recall* yang tinggi menunjukkan bahwa model tidak melewatkan banyak prediksi yang benar untuk kelas tertentu.
- ***F1-score***: Merupakan rata-rata harmonis dari *precision* dan *recall*. *F1-score* memberikan gambaran yang lebih seimbang antara *precision* dan *recall*, sehingga lebih cocok untuk evaluasi ketika ada ketidakseimbangan kelas dalam data.

Selain itu, untuk memastikan bahwa model *Random Forest* tidak mengalami overfitting, dilakukan teknik validasi silang (cross-validation) dengan membagi data pelatihan menjadi beberapa subset. Proses validasi silang ini digunakan untuk menilai stabilitas model dengan berbagai set data pelatihan dan pengujian.

### **III. HASIL DAN DISKUSI**

Tabel 1 menggambarkan distribusi luas area untuk masing-masing kelas tutupan lahan yang terdeteksi di Kota Bandung. Ini memberikan gambaran umum tentang komposisi tutupan lahan yang dominan di wilayah tersebut. Berdasarkan hasil klasifikasi, terdapat empat kategori utama tutupan lahan yang mencakup wilayah kota, yaitu lahan permukiman, lahan hijau, perairan, dan area.

Dari tabel tersebut, terlihat bahwa lahan permukiman adalah kelas yang paling dominan di Kota Bandung, dengan luas mencapai 15,200 hektar, atau sekitar 45% dari total luas kota. Hal ini mencerminkan tingginya tingkat urbanisasi dan pertumbuhan permukiman di area pusat dan sekitarnya, yang sesuai dengan tren perkembangan kota besar di Indonesia. Permukiman ini mencakup berbagai jenis bangunan, jalan, dan infrastruktur lainnya yang menyebar di seluruh area perkotaan. Kelas kedua yang paling besar adalah lahan hijau, yang mencakup area vegetasi dan taman kota, dengan luas sekitar 10,300 hektar atau 31% dari total luas kota. Lahan hijau ini memainkan peran penting dalam menjaga keseimbangan ekosistem perkotaan, menyediakan ruang terbuka untuk publik, dan berkontribusi pada kualitas udara serta penyerapan air hujan. Area hijau juga tersebar di berbagai bagian kota, termasuk taman kota, ruang terbuka hijau, dan lahan pertanian.

Perairan, yang mencakup sungai, danau, dan waduk yang ada di Kota Bandung, memiliki luas sekitar 2,500 hektar, atau sekitar 7% dari total luas kota. Meskipun tidak sebanyak lahan permukiman atau lahan hijau, perairan memainkan peran penting dalam mengatur aliran air dan menjadi sumber kehidupan bagi banyak spesies flora dan fauna. Kelas ini juga relevan untuk studi terkait pengelolaan sumber daya air dan mitigasi risiko bencana alam seperti banjir. Sementara itu, area terbuka mencakup tanah kosong, pekarangan, dan area non-vegetatif lainnya yang tidak memiliki tutupan vegetasi atau bangunan. Luas area terbuka ini tercatat sekitar 5,000 hektar, atau 17% dari luas total Kota Bandung. Area terbuka sering

kali berfungsi sebagai ruang transisi atau sebagai lahan untuk kegiatan non-permanen seperti parkir dan kegiatan komersial sementara.

Secara keseluruhan, total luas tutupan lahan di Kota Bandung yang terklasifikasi dalam penelitian ini mencapai 33,000 hektar. Distribusi luas masing-masing kelas menunjukkan bahwa lahan permukiman mendominasi wilayah kota, sementara lahan hijau dan area terbuka juga memiliki peran yang signifikan dalam struktur kota. Pengetahuan ini penting untuk perencanaan kota yang lebih berkelanjutan, di mana keseimbangan antara pembangunan dan pelestarian lingkungan harus diperhatikan dengan cermat.

Tabel 1. Distribusi Luas Tutupan Lahan di Kota Bandung

Kelas Tutupan Lahan	Luas (Hektar)	Persentase (%)
Lahan Permukiman	15.200	45%
Lahan Hijau	10.300	31%
Perairan	2.500	7%
Area Terbuka	5.000	17%
<b>Total</b>	<b>33.000</b>	<b>100%</b>

#### A. Hasil Akurasi Klasifikasi Tutupan Lahan

Berdasarkan penerapan algoritma *Random Forest* pada citra satelit Landsat yang telah diproses, peta tutupan lahan Kota Bandung berhasil diklasifikasikan ke dalam beberapa kategori utama, yaitu lahan permukiman, lahan hijau, perairan, dan area terbuka. Hasil klasifikasi menunjukkan bahwa sebagian besar area di pusat kota didominasi oleh lahan permukiman, sedangkan kawasan pinggiran kota dan beberapa area hijau menunjukkan prevalensi lahan vegetasi. Proses klasifikasi menghasilkan peta dengan akurasi keseluruhan sebesar 89%, yang menunjukkan kemampuan algoritma *Random Forest* dalam membedakan berbagai jenis tutupan lahan meskipun terdapat tantangan berupa variasi spektral dalam citra satelit, terutama di area yang padat penduduk. Angka tersebut menunjukkan kinerja yang sangat baik, mengingat kompleksitas tutupan lahan yang ada di daerah perkotaan yang melibatkan berbagai jenis kelas, seperti lahan permukiman, lahan hijau, perairan, dan area terbuka.

Pengukuran akurasi untuk masing-masing kelas yang ditunjukkan pada Tabel 2 menunjukkan hasil yang bervariasi. Kelas lahan permukiman memiliki tingkat akurasi tertinggi, yaitu 92% yang mencerminkan keberhasilan model dalam mengenali kawasan permukiman yang dominan di pusat kota. Hal ini mungkin karena lahan permukiman memiliki ciri-ciri spektral yang lebih konsisten dan mudah dibedakan dari kelas lainnya, seperti perbedaan yang jelas dalam reflektansi antara material bangunan dan vegetasi atau perairan.

Sebaliknya, kelas lahan hijau dan perairan menunjukkan akurasi yang lebih rendah, masing-masing 85% dan 80%. Hasil ini mungkin dipengaruhi oleh adanya tumpang tindih spektral antara vegetasi dan area terbuka, serta kesulitan dalam membedakan perairan yang tercermin dengan latar belakang yang lebih gelap atau terkontaminasi oleh polusi air atau kekeruhan. Dalam konteks ini, lahan hijau meliputi berbagai jenis vegetasi yang dapat memiliki reflektansi yang bervariasi tergantung pada jenis vegetasi dan kondisi lingkungan, yang bisa mengurangi akurasi klasifikasi. Kelas area terbuka memperoleh akurasi 87%, yang juga merupakan hasil yang cukup baik, namun masih menunjukkan ruang untuk perbaikan. Area terbuka, meskipun lebih mudah dibedakan dari area permukiman atau perairan, terkadang memiliki spektrum yang tumpang tindih dengan area lain yang tidak tertutup vegetasi atau bangunan. Hasil tersebut memberikan gambaran bahwa algoritma *Random Forest* dapat efektif dalam mengidentifikasi kategori tutupan lahan yang berbeda, meskipun ada kemungkinan kesalahan klasifikasi pada kelas-kelas tertentu yang memiliki karakteristik spektral yang tumpang tindih.

Tabel 2. Akurasi Klasifikasi untuk Masing-Masing Kelas

Kelas Tutupan Lahan	Akurasi (%)
Lahan Permukiman	92 %
Lahan Hijau	85 %
Perairan	80 %
Area Terbuka	87 %
<b>Akurasi Keseluruhan</b>	<b>89 %</b>

#### B. Confusion Matrix dan Kesalahan Klasifikasi

Tabel Confusion Matrix yang ditunjukkan pada Tabel 3 memberikan informasi lebih mendalam terkait kesalahan klasifikasi yang terjadi pada model. Confusion Matrix ini membantu mengidentifikasi

***Analisis Metode Klasifikasi Pemetaan Tutupan Lahan (Land Cover) di Area Kota Bandung Menggunakan Algoritma Random Forest Pada Google Earth Engine***  
***(Hajiar Yuliana, Zahra Cahya Hanifa Rizqiana : Halaman 97 - 106)***

performa model dalam klasifikasi setiap jenis tutupan lahan secara lebih terperinci, termasuk kesalahan klasifikasi (*false positives* dan *false negatives*) untuk setiap kelas. Confusion Matrix ini menunjukkan perbandingan antara hasil klasifikasi yang diprediksi oleh model dan data referensi ground truth. Angka-angka ini menunjukkan jumlah piksel yang benar-benar diklasifikasikan dengan tepat atau yang biasa disebut dengan *True Positives* (TP) dan jumlah piksel yang salah klasifikasi atau yang biasa disebut dengan *False Positives* (FP) atau *False Negatives* (FN). Berdasarkan matrix tersebut, dapat dilihat bahwa lahan permukiman cenderung terklasifikasi dengan baik, dengan hanya sedikit kesalahan dalam mengklasifikasikan sebagai lahan hijau atau area terbuka. Hal ini menunjukkan bahwa model RF mampu dengan baik membedakan area yang padat permukiman dari area hijau atau terbuka.

Namun, kesalahan klasifikasi pada kelas lahan hijau lebih signifikan, dengan banyaknya *false positives* (FP) yang mengidentifikasi area yang seharusnya masuk dalam kelas lahan hijau sebagai lahan permukiman atau area terbuka. Hal ini dapat dijelaskan oleh variasi tinggi dalam reflektansi spektral yang dimiliki oleh jenis vegetasi yang berbeda (misalnya, vegetasi semak dan pohon tinggi), serta kemungkinan tumpang tindih dengan area permukiman atau infrastruktur lainnya, yang sulit dibedakan hanya dengan data citra satelit. Kelas perairan juga mengalami beberapa *false positives* dan *false negatives*, meskipun tidak sebanyak kelas lahan hijau. Kesalahan ini dapat disebabkan oleh pencampuran spektral antara perairan yang gelap atau tercemar dengan area terbuka atau lahan permukiman yang berdekatan, yang dapat membuat deteksi perairan menjadi kurang akurat. Pada kelas area terbuka, meskipun akurasi tinggi, model masih mengalami beberapa *false positives* yang mengklasifikasikan area terbuka sebagai lahan hijau atau perairan, yang menunjukkan tantangan dalam membedakan area dengan vegetasi jarang atau tanah kosong dengan area lain yang lebih kompleks.

Tabel 3. Evaluasi Kinerja Model *Random Forest* dengan *Confusion Matrix*

Kelas Prediksi /Kelas Sebenarnya	Lahan Permukiman	Lahan Hijau	Perairan	Area Terbuka	Total
Lahan Permukiman	2.000 (TP)	100 (FN)	50 (FN)	150 (FN)	<b>2.300</b>
Lahan Hijau	80 (FP)	1.900 (TP)	30 (FN)	120 (FN)	<b>2.130</b>
Perairan	30 (FP)	40 (FP)	500 (TP)	30 (FN)	<b>600</b>
Area Terbuka	50 (FP)	120 (FP)	20 (FP)	1.800 (TP)	<b>1.990</b>
<b>Total</b>	<b>2.160</b>	<b>2.160</b>	<b>600</b>	<b>2.100</b>	<b>7.020</b>

***C. Hasil Cross-Validation (Validasi Silang)***

Cross-validation adalah teknik yang digunakan untuk mengevaluasi model secara lebih robust dan memastikan bahwa model yang dilatih dapat digeneralisasi dengan baik pada data yang belum pernah dilihat sebelumnya (data yang tidak terlibat dalam pelatihan). Teknik ini sangat berguna untuk menghindari *overfitting*, yaitu ketika model terlalu "terikat" pada data pelatihan dan tidak dapat bekerja dengan baik pada data baru.

*Fold* dalam *cross-validation* merujuk pada pembagian data menjadi beberapa subset yang lebih kecil. Secara umum, dalam *k-fold cross-validation*, data dibagi menjadi k bagian (*folds*). Model dilatih menggunakan k-1 *fold*, sementara satu *fold* lainnya digunakan untuk validasi (pengujian). Proses ini diulang sebanyak k kali, dengan setiap *fold* digunakan sekali sebagai data uji dan sisanya sebagai data pelatihan. Hasil dari semua percobaan tersebut kemudian dirata-rata untuk memberikan evaluasi akhir terhadap model.

Tabel 4 yang menunjukkan hasil dari validasi silang menunjukkan bahwa model RF memiliki kinerja yang stabil dan tidak mengalami *overfitting*. Dengan rata-rata akurasi 88.5% dari empat *fold*, model RF dapat diandalkan untuk klasifikasi data citra satelit yang serupa di wilayah lain. Nilai *Precision*, *Recall*, dan *F1-Score* yang tinggi menunjukkan bahwa model RF memiliki keseimbangan yang baik dalam hal kemampuan untuk mengidentifikasi kelas tutupan lahan dengan benar (*Precision*) dan menangkap seluruh variabel yang relevan dari setiap kelas (*Recall*), yang menjadikannya model yang sangat baik untuk klasifikasi tutupan lahan.

Tabel 4. Kinerja Model dalam Validasi Silang (*Cross-Validation*)

<i>Fold</i>	Akurasi (%)	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
<b><i>Fold 1</i></b>	88%	0,91	0,89	0,90
<b><i>Fold 2</i></b>	90%	0,92	0,87	0,89
<b><i>Fold 3</i></b>	89%	0,90	0,88	0,89
<b><i>Fold 4</i></b>	87%	0,89	0,90	0,89

<b>Rata-rata</b>	<b>88,5%</b>	<b>0,91</b>	<b>0,88</b>	<b>0,89</b>
------------------	--------------	-------------	-------------	-------------

Tabel 4 menunjukkan hasil evaluasi model untuk empat *fold* dalam proses validasi silang (*4-fold cross-validation*), di mana data dibagi menjadi empat bagian yang saling tumpang tindih, dan model dilatih dan diuji secara bergantian pada masing-masing *fold*. Setiap *fold* (*Fold 1*, *Fold 2*, *Fold 3*, dan *Fold 4*) memiliki nilai akurasi, *precision*, *recall*, dan *F1-score*, yang mengukur kinerja model dalam memprediksi kelas tutupan lahan yang berbeda.

Model *Random Forest* yang digunakan tersebut menunjukkan keseimbangan yang baik antara *precision* dan *recall*. Meskipun *precision* lebih tinggi dari *recall* pada beberapa *fold*, model mampu mendeteksi sebagian besar kelas positif (*recall* tetap berada di sekitar 0.88 hingga 0.90). *F1-score* yang tinggi di semua *fold* menunjukkan bahwa model menjaga keseimbangan yang efektif antara keduanya, sangat penting untuk tugas klasifikasi tutupan lahan di area perkotaan yang kompleks. Walaupun ada sedikit fluktuasi dalam hasil evaluasi dari *fold* ke *fold*, hasil rata-rata menunjukkan bahwa model memiliki performa yang sangat konsisten dalam klasifikasi citra satelit untuk tutupan lahan di Kota Bandung. Akurasi yang stabil di seluruh *fold* mengindikasikan bahwa model tidak terpengaruh oleh ketidakseimbangan kelas atau variasi data yang ada dalam set pelatihan.

Karena model menunjukkan hasil yang stabil di seluruh *fold*, ini menunjukkan bahwa model *Random Forest* tidak mengalami *overfitting* (terlalu menyesuaikan diri dengan data pelatihan) maupun *underfitting* (tidak cukup belajar dari data pelatihan). Validasi silang ini berhasil mengidentifikasi model yang robust dan dapat digeneralisasi dengan baik ke data yang belum pernah dilihat sebelumnya.

#### IV. KESIMPULAN

Dengan menggunakan platform *Google Earth Engine* (GEE), penelitian ini berhasil menerapkan algoritma *Random Forest* (RF) untuk mengklasifikasikan tutupan lahan di wilayah Kota Bandung. Hasil klasifikasi menunjukkan akurasi total sebesar 89%, dengan tingkat akurasi yang berbeda untuk setiap kategori tutupan lahan: 92% untuk lahan permukiman, 85% untuk lahan hijau, 80% untuk perairan, dan 87% untuk area terbuka. Hal ini menunjukkan bahwa algoritma Hutan Random memiliki kinerja yang baik dalam membedakan berbagai jenis tutupan lahan, meskipun tumpang tindih spektral antar kelas merupakan masalah. Hasil validasi silang (*cross-validation*) juga menunjukkan performa yang stabil, dengan rata-rata akurasi 88.5%, *precision* 0.91, *recall* 0.88, dan *F1-score* 0.89. Nilai-nilai ini menggambarkan bahwa model mampu menjaga keseimbangan yang baik antara akurasi prediksi positif dan kemampuan mendeteksi seluruh kelas dalam data uji. Selain itu, evaluasi melalui *confusion matrix* menunjukkan distribusi klasifikasi yang baik untuk sebagian besar kelas. Misalnya, kelas lahan permukiman dan lahan hijau memiliki jumlah piksel yang tinggi yang diklasifikasikan dengan benar, sementara kelas perairan menunjukkan beberapa kasus salah klasifikasi menjadi lahan terbuka, yang memengaruhi akurasi kelas tersebut.

GEE sebagai platform berbasis cloud terbukti menghemat waktu dan memproses data besar. Pemetaan tutupan lahan yang lebih akurat dan efisien telah dicapai melalui penelitian ini. Ini dapat digunakan sebagai dasar untuk perencanaan tata ruang dan pengelolaan lingkungan Kota Bandung. Namun, ada beberapa keterbatasan pada penelitian ini. Salah satunya adalah tingkat akurasi yang lebih rendah pada kelas lahan hijau dan perairan karena variasi spektral yang kompleks. Selain itu, distribusi data kebenaran faktual yang tidak merata di seluruh wilayah penelitian juga dapat berdampak pada hasil klasifikasi.

Terdapat beberapa peluang untuk pengembangan lebih lanjut dalam hal metode, data, dan cakupan analisis penelitian ini. Menggunakan data gambar satelit dengan resolusi spasial yang lebih tinggi, seperti Sentinel-2 atau WorldView, adalah langkah penting yang dapat diambil. Penggunaan gambar dengan resolusi lebih tinggi dapat membantu meningkatkan akurasi model dalam membedakan kelas tutupan lahan yang sulit, seperti perairan dan lahan hijau. Selain itu, integrasi dengan data tambahan seperti LiDAR, topografi, atau data cuaca dapat menambah dimensi baru dalam analisis. Ini memungkinkan model untuk menangkap pola yang lebih kompleks dan menghasilkan klasifikasi yang lebih akurat.



Penelitian lanjutan dapat mengeksplorasi algoritma yang lebih canggih dari segi metode, seperti *Convolutional Neural Networks* (CNN) atau ensemble learning yang menggabungkan kekuatan beberapa algoritma. Metode ini dapat membantu menangani masalah dengan data dengan variasi spektral yang tinggi, seperti kelas vegetasi dan perairan. Selain itu, menggunakan pendekatan berbasis deret waktu (analisis deret waktu) juga penting untuk memetakan perubahan tutupan lahan secara dinamis. Analisis ini dapat memberikan wawasan lebih mendalam tentang pola perubahan penggunaan lahan dari waktu ke waktu, yang sangat relevan untuk perencanaan tata ruang dan mitigasi dampak lingkungan. Peningkatan kualitas data *ground truth* juga dapat menjadi bagian dari pengembangan lebih lanjut. Data sampel dapat didistribusikan secara lebih merata di seluruh area penelitian, yang dapat memperbaiki proses pelatihan model dan mengurangi bias klasifikasi. Selain itu, data lapangan yang lebih representatif dapat digunakan untuk validasi hasil klasifikasi, sehingga hasil dapat diverifikasi secara lebih akurat.

Pengujian teknik ini di daerah lain dengan karakteristik tutupan lahan yang berbeda sangat penting untuk mengevaluasi generalisasi algoritma yang digunakan. Metode ini membantu kita memahami seberapa efektif algoritma Hutan Sederhana dalam berbagai kondisi, dan memberikan wawasan yang dapat digunakan untuk daerah dengan masalah klasifikasi yang serupa. Penelitian selanjutnya dapat membantu perencanaan wilayah yang lebih efisien, pengelolaan lingkungan yang berkelanjutan, dan pengambilan keputusan berbasis data yang lebih baik.

#### DAFTAR PUSTAKA

- [1] T. N. Phan, V. Kuch, and L. W. Lehnert, "Land Cover Classification Using Google Earth Engine and Random Forest Classifier—The Role of Image Composition," *Remote Sens.*, vol. 12, no. 15, p. 2411, 2020, doi: 10.3390/rs12152411.
- [2] M. Mahdaviard, S. K. AHANGAR, B. Feizizadeh, K. V. Kamran, and S. Karimzadeh, "Spatio-Temporal Monitoring of Qeshm Mangrove Forests Through Machine Learning Classification of SAR and Optical Images on Google Earth Engine," *Int. J. Eng. Geosci.*, vol. 8, no. 3, pp. 239–250, 2023, doi: 10.26833/ijeg.1118542.
- [3] V. Eisavi, S. Homayouni, A. M. Yazdi, and A. Alimohammadi, "Land Cover Mapping Based on Random Forest Classification of Multitemporal Spectral and Thermal Images," *Environ. Monit. Assess.*, vol. 187, no. 5, 2015, doi: 10.1007/s10661-015-4489-3.
- [4] A. A. Kuntoro, A. W. Putro, M. S. B. Kusuma, and S. Natasaputra, "The Effect of Land Use Change to Maximum and Minimum Discharge in Cikapundung River Basin," 2017, doi: 10.1063/1.5011621.
- [5] N. Ponganan, T. Horanont, K. Artlert, and P. Nuallaong, "Land Cover Classification using Google Earth Engine's Object-oriented and Machine Learning Classifier," *2021 2nd Int. Conf. Big Data Anal. Pract. IBDAP 2021*, pp. 33–37, 2021, doi: 10.1109/IBDAP52511.2021.9552099.
- [6] A. Tassi, D. Gigante, G. Modica, L. Di Martino, and M. Vizzari, "Pixel-vs. Object-based landsat 8 data classification in google earth engine using random forest: The case study of maiella national park," *Remote Sens.*, vol. 13, no. 12, 2021, doi: 10.3390/rs13122299.
- [7] S. Amini, M. Saber, H. Rabiei-Dastjerdi, and S. Homayouni, "Urban Land Use and Land Cover Change Analysis Using Random Forest Classification of Landsat Time Series," *Remote Sens.*, vol. 14, no. 11, pp. 1–23, 2022, doi: 10.3390/rs14112654.
- [8] S. Xie, L. Liu, X. Zhang, J. Yang, X. Chen, and Y. Gao, "Automatic land-cover mapping using landsat time-series data based on google earth engine," *Remote Sens.*, vol. 11, no. 24, 2019, doi: 10.3390/rs11243023.
- [9] J. Sun and S. Ongsomwang, "Optimal parameters of random forest for land cover classification with suitable data type and dataset on Google Earth Engine," *Front. Earth Sci.*, vol. 11, no. October, pp. 1–17, 2023, doi: 10.3389/feart.2023.1188093.
- [10] V. F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez, "An assessment of the effectiveness of a random forest classifier for land-cover classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 67, no. 1, pp. 93–104, 2012, doi: 10.1016/j.isprsjprs.2011.11.002.

- [11] A. Jamali, "Land Use Land Cover Mapping Using Advanced Machine Learning Classifiers," *Ekológia (Bratislava)*, vol. 40, no. 3, pp. 286–300, 2021, doi: 10.2478/eko-2021-0031.
- [12] C. Pelletier, S. Valero, J. Inglada, N. Champion, C. M. Sicre, and G. Dedieu, "Effect of Training Class Label Noise on Classification Performances for Land Cover Mapping With Satellite Image Time Series," *Remote Sens.*, vol. 9, no. 2, p. 173, 2017, doi: 10.3390/rs9020173.