

Perbandingan Algoritma *Classification and Regression Tree (CART)* dan *Random Forest (RF)* untuk Klasifikasi Penggunaan Lahan pada *Google Earth Engine*

Faiz Al Farikhi ^{a,1}, Retno Widodo Dwi Pramono ^{b,2}

^a Magister Perencanaan Kota dan Wilayah, Universitas Gadjah Mada, Yogyakarta

^b Departemen Teknik Arsitektur dan Perencanaan, Fakultas Teknik, Universitas Gadjah Mada

¹ faiz.al.farikhi@mail.ugm.ac.id

² pramono.wid@ugm.ac.id

Informasi artikel	ABSTRAK
<i>Sejarah artikel</i> Diterima : 31-07-2023 Revisi : 25-09-2023 Dipublikasikan : 25-09-2023	Perubahan penggunaan lahan terus terjadi seiring dengan pertumbuhan jumlah penduduk dan aktivitas ikutannya. Diperlukan analisis yang cepat dan akurat untuk mendukung tatakemelolanya. Salah satu yang dibutuhkan adalah teknik klasifikasi penggunaan lahan yang cepat dan efektif salah satunya dapat dilakukan melalui platform berbasis <i>cloud</i> Google Earth Engine (GEE) dengan memanfaatkan citra multi-spektral dan melibatkan berbagai algoritma pengolahan. Setidaknya terdapat dua metode klasifikasi menggunakan algoritma pohon keputusan (<i>decision tree algorithm</i>) yaitu: <i>Classification and Regression Tree (CART)</i> dan <i>Random Forest (RF)</i> . Perbedaan metode klasifikasi memungkinkan terjadinya perbedaan akurasi dari klasifikasi yang dihasilkan. Belum banyak penelitian yang menjelaskan bagaimana perbedaan tingkat akurasi dan efektifitas kedua metode ini. Penelitian ini bertujuan untuk mengidentifikasi metode klasifikasi yang akurat dan efektif dengan membandingkan nilai akurasi kedua metode pada citra Landsat 8 dan Sentinel 2A. Metode RF menunjukkan nilai akurasi yang lebih tinggi dibandingkan metode CART pada citra Landsat 8 dan Sentinel 2A. Proporsi sampel 60:40 merupakan proporsi yang paling optimal dengan nilai akurasi paling tinggi pada masing-masing metode klasifikasi dan citra yang digunakan. Penggunaan citra Landsat 8 dengan proporsi sampel 60:40 menghasilkan nilai akurasi paling tinggi (OA=93,94% dan Kappa=0,927).
Kata kunci: Klasifikasi Penggunaan Lahan Google Earth Engine Classification and Regression Tree Random Forest Bengawan Solo	

Keywords:	ABSTRACT
LULC Classifier Google Earth Engine Classification and Regression Tree Random Forest Bengawan Solo	Land Use and Land Cover (LULC) change are increasing along with the population growth and their activity-followed. Quick and accurate analyses are needed to support the land use management. One of the effective and accurate platforms to generate LULC classification maps are Google Earth Engine (GEE), a cloud-based geospatial by utilising multi-spectral imagery and various processing algorithms. There are at least two LULC classifier methods using the decision tree algorithm: <i>Classification And Regression Tree (CART)</i> and <i>Random Forest (RF)</i> . Differences classifier method are allowed to produce difference accuracy level. Few studies have explained how the accuracy level and effectiveness of these two methods. This research aims to identify an accurate and effective LULC classifier method by comparing the accuracy values of the two methods on Landsat 8 and Sentinel 2A imagery. The RF algorithm shows higher accuracy result than the CART algorithm on both Landsat 8 and Sentinel 2A images. The most optimal sample proportion is 60:40 with higher accuracy result on both image and classifier algorithm. The Landsat 8 imagery with sample proportion 60:40 shows the highest accuracy result (OA=93,94 and Kappa=0,927).

Pendahuluan

Penggunaan lahan sebagai serangkaian intervensi yang dilakukan oleh manusia terhadap bentuk alami permukaan bumi (Arsyad, 2010) bersifat dinamis seiring dengan dinamika pertumbuhan penduduk serta peningkatan aktivitas ikutannya yang bersifat materil maupun (Vink, 1975). Agar penggunaan lahan dapat optimal dalam memenuhi kebutuhan masyarakat secara berkeadilan dan berkelanjutan, diperlukan tatakelola yang baik didukung dengan kemampuan teknis dan teknologi dalam pengambilan keputusannya. Salah satu aspek penting yang dibutuhkan adalah teknik klasifikasi penggunaan lahan yang akurat, cepat dan efisien yang dapat mengikuti tren perubahan yang sangat dinamis.

Pada perkembangannya, terdapat berbagai macam cara untuk mengidentifikasi penggunaan lahan salah satunya melalui analisis data spasial. Data spasial adalah data yang memiliki atribut spasial yang menunjukkan lokasi, letak dan posisinya di permukaan bumi (Yeung & Hall, 2014). Data spasial memiliki referensi geografis artinya didalamnya memuat posisi dalam sistem koordinat bumi dan dapat ditampilkan dalam skala tertentu. Ketersediaan data spasial saat ini sudah sangat banyak dan bersifat open source sehingga dapat diakses oleh siapapun. Tersedianya data spasial ditambah dengan berkembangnya teknologi pengolahan data spasial menjadikan tingkat kemanfaatan data menjadi lebih besar lagi.

Perkembangan teknologi Geographic Information System (GIS) cukup pesat dan memberikan berbagai kemudahan bagi penggunaannya. Google Earth Engine (GEE) merupakan salah satu platform yang memberikan kemudahan terhadap analisis dan pengolahan data spasial melalui penyediaan data yang besar secara kontinu dengan sistem *cloud* (S. M. Gorelick, 2015). Identifikasi penggunaan lahan dapat dilakukan dengan memanfaatkan platform GEE karena data citra penginderaan jauh dan algoritma klasifikasi tersedia dalam satu platform yang sama. Data spasial dapat diakuisisi dan diolah langsung pada platform GEE sehingga

lebih efisien karena tidak memerlukan banyak aplikasi dan waktu yang dibutuhkan lebih sedikit.

Data spasial berupa citra penginderaan jauh merupakan data yang paling sering digunakan untuk mengidentifikasi tutupan dan penggunaan lahan (Roy et al., 2014). Citra penginderaan jauh merupakan data yang didapatkan melalui perekaman kenampakan permukaan bumi melalui sensor yang terpasang pada wahana luar angkasa. Terdapat berbagai jenis citra penginderaan jauh, perbedaannya terletak pada resolusi spasial, resolusi temporal dan resolusi radiometriknya.

Identifikasi penggunaan lahan dengan memanfaatkan citra penginderaan jauh melibatkan berbagai algoritma salah satunya algoritma pohon keputusan (*decision tree algoritm*). Metode ini mengklasifikasikan populasi ke dalam segmen-segmen menyerupai cabang yang membangun pohon terbalik dengan simpul akar, simpul internal dan simpul daun (Song & Lu, 2015). Penggunaan algoritma pohon keputusan dapat menghasilkan klasifikasi secara cepat dengan data yang besar sehingga identifikasi penggunaan lahan lebih efisien. Algoritma pohon keputusan menggunakan pendekatan statistik non parametrik serta bekerja dengan menyederhanakan hubungan yang kompleks antara variabel input dan variabel target sehingga algoritma pohon keputusan merupakan alat statistik yang kuat.

Terdapat beberapa metode klasifikasi yang menggunakan algoritma pohon keputusan diantaranya adalah Classification and Regression Tree (CART) dan Random Forest (RF) classifier. Perbedaan metode klasifikasi yang digunakan dalam mengidentifikasi penggunaan lahan memungkinkan terjadinya perbedaan akurasi dari klasifikasi yang dihasilkan (Aldiansyah & Saputra, 2022; Yulianto, Nugroho, & Suwarsono, 2021). Sementara itu, perbedaan data spasial juga berpengaruh terhadap akurasi hasil klasifikasi yang ditunjukkan dengan nilai Overall Accuracy (OA) dan Kappa Coefficient (Basheer et al., 2022). Penelitian ini bertujuan untuk mengidentifikasi metode klasifikasi penggunaan lahan yang efektif ditunjukkan dari nilai akurasi paling tinggi.

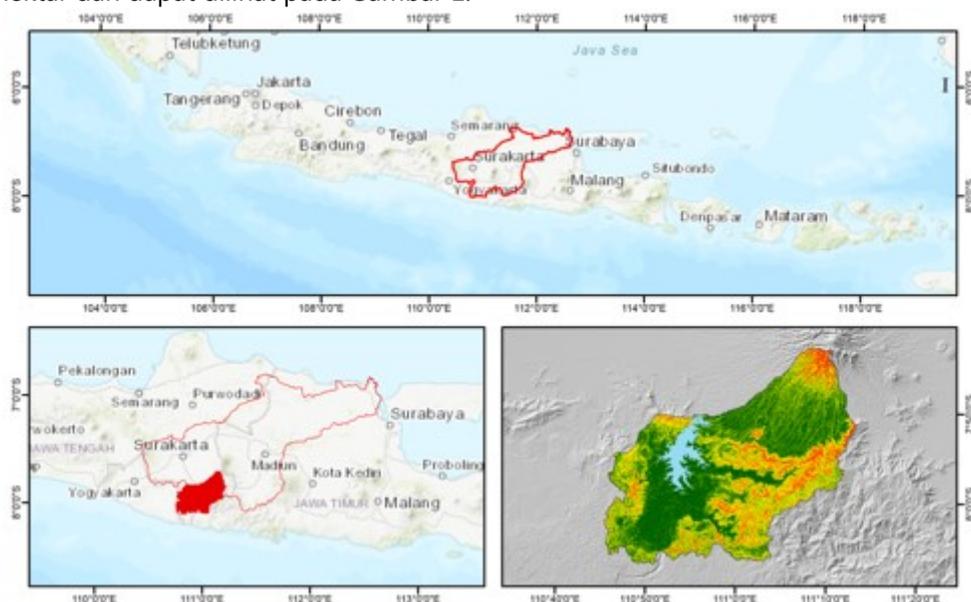
Metode dan data penelitian

Area penelitian

Area penelitian merupakan Daerah Aliran Sungai (DAS) Bengawan Solo bagian hulu yang bermuara di Waduk Gajah Mungkur (WGM) diantaranya adalah Sub DAS Keduang, Sub DAS Tirtomoyo, Sub DAS Temon, Sub DAS Solo Hulu, Sub DAS Alang, Sub DAS Ngunggungahan dan Sub DAS Wuryantoro. DAS Bengawan Solo merupakan salah DAS yang masuk kategori kritis dan diprioritaskan untuk direhabilitasi (RPJMN, 2014). Tujuh Sub DAS tersebut mengalami degradasi lahan yang menyebabkan terjadinya sedimentasi yang cukup parah di WGM (JICA, 2007; Rahman, Harisuseno, & Sisingsih, 2012; Santosa, 2016). Penelitian ini menghasilkan peta penggunaan lahan sehingga dapat dimanfaatkan untuk analisis lanjutan terutama berkaitan dengan optimalisasi guna lahan untuk meminimalisir terjadinya sedimentasi pada WGM. Area penelitian seluas 133 ribu hektar dan dapat dilihat pada Gambar 1.

Data penelitian

Penelitian ini menggunakan dua citra multispektral dengan resolusi spasial 10 meter dan 30 meter yaitu citra Sentinel-2 MSI dan Landsat 8 OLI yang sudah melalui koreksi radiometrik dan geometrik. Citra Sentinel-2 MSI menggunakan data Level-2A dipilih dengan (*filtering*) kriteria cloud cover <10% ("*CLOUDY_PIXEL_PERCENTAGE*", 10) dan menggunakan data nilai tengah (median). Citra Landsat 8 OLI menggunakan data Level 2, Tier 1 Surface Reflectance yang dilakukan masking awan (*Cloud Mask*) dan menggunakan data nilai tengah (median). Band yang digunakan adalah B2, B3, B4, B8 (Sentinel-2) dan B2, B3, B4, B5, B6, B7 (Landsat 8). Data yang digunakan merupakan citra agregat dalam durasi satu tahun (1 Januari – 31 Desember) yang diseleksi berdasarkan rentang nilai tengahnya. Secara lebih detail data yang digunakan dapat dilihat pada Tabel 1.



Gambar 1. Peta area penelitian

Tabel 1. Data penelitian

Nama Data	Periode	Band	Spektrum	Wavelength	Resolusi Spasial	Pemilik Data
Citra Sentinel 2A MSI (SR_Harmonized)	1 Januari 2022 – 31 Desember 2022	B2	Blue	496.6 μm	10m	ESA (European Space Agency)
		B3	Green	560 μm	10m	
		B4	Red	664.5 μm	10m	
		B8	NIR	835.1 μm	10m	
Citra Landsat 8 OLI (T1_L2)	1 Januari 2022 – 31 Desember 2022	B2	Blue	0.452-0.512 μm	30m	USGS (United States Geological Survey)
		B3	Green	0.533-0.590 μm	30m	
		B4	Red	0.636-0.673 μm	30m	
		B5	NIR	0.851-0.879 μm	30m	
		B6	SWIR 1	1.566-1.651 μm	30m	
		B7	SWIR 2	2.107-2.294 μm	30m	

Metode

Pengolahan data dalam penelitian ini menggunakan platform Google Earth Engine (GEE) diantaranya meliputi: *image selection*, *image pre processing*, *data sampling*, *classifier*, dan *accuracy assesment*. GEE merupakan platform berbasis *cloud (internet based)* untuk mengolah data geospasial yang menggabungkan antara basis data spasial dan *machine learning* dalam satu platform sehingga analisis data spasial menjadi lebih efisien. Melalui berbagai kemudahan yang diberikan, saat ini GEE menjadi platform yang sering digunakan dalam pengolahan data spasial salah satunya untuk klasifikasi penggunaan lahan. Pengolahan data pada platform GEE menggunakan bahasa pemrograman JavaScript.

Metode klasifikasi yang digunakan dalam penelitian ini adalah Classification and Regression Tree (CART) dan Random Forest (RF) classifier, keduanya menggunakan algoritma pohon keputusan (*decision tree algorithm*). Keduanya digunakan untuk mengklasifikasikan penggunaan lahan dengan data citra Sentinel-2 dan Landsat 8 untuk membandingkan tingkat akurasi metode pada masing-masing data. Kelas penggunaan lahan dibagi menjadi enam kelas antara lain; air (water), hutan (forest), lahan terbangun (build up area), sawah (paddy), pertanian lahan kering (dryland), perkebunan (plantation). Pembagian kelas penggunaan lahan mengacu pada SNI 7645:2010 dan hasil modifikasi dari beberapa penelitian lain (Ambarwulan dkk., 2023; Basheer dkk., 2022; Yulianto dkk., 2022).

Penelitian ini bertujuan untuk mengidentifikasi algoritma klasifikasi penggunaan lahan yang efektif dengan membandingkan akurasi yang dihasilkan oleh masing-masing algoritma klasifikasi melalui platform Google Earth Engine (GEE). Metode klasifikasi yang dibandingkan dalam penelitian ini adalah metode Classification and Regression Tree (CART) dan Random Forest (RF). Platform GEE merupakan platform pengolahan data spasial berbasis cloud yang menggunakan bahasa pemrograman Javascript. Script yang digunakan dalam penelitian ini adalah "ee.Classifier.smileCart" untuk metode CART dan "ee.Classifier.smileRandomForest" untuk metode RF.

Sampel dan uji akurasi

Total sampel yang digunakan sebanyak 500 titik dengan metode pemilihan titik secara acak

(random sampling). Sampel dibagi menjadi dua kelompok yaitu *training sample* dan *testing sample* dengan proporsi 50-80% (50:50, 60:40, 70:30, dan 80:20). *Training sample* digunakan untuk menjalankan algoritma klasifikasi sementara *testing sample* digunakan untuk menjalankan uji akurasi (*accuracy assesment*) dari hasil klasifikasi yang meliputi Producer's Accuracy (PA), User's Accuracy (UA), Overall Accuracy (OA) dan Kappa Coefficient.

Classification and regression tree (CART)

CART merupakan metode klasifikasi menggunakan logika biner berdasarkan logika pohon keputusan (*decision tree*), hasil klasifikasi didasarkan pada kinerja statistik regresi (Breiman, Friedman, Olshen, & Stone, 1984). Cara kerja algoritma CART adalah memisahkan data melalui simpul-simpul hingga mencapai simpul akhir dengan membentuk hierarki pohon keputusan (Basheer et al., 2022).

Random forest (RF)

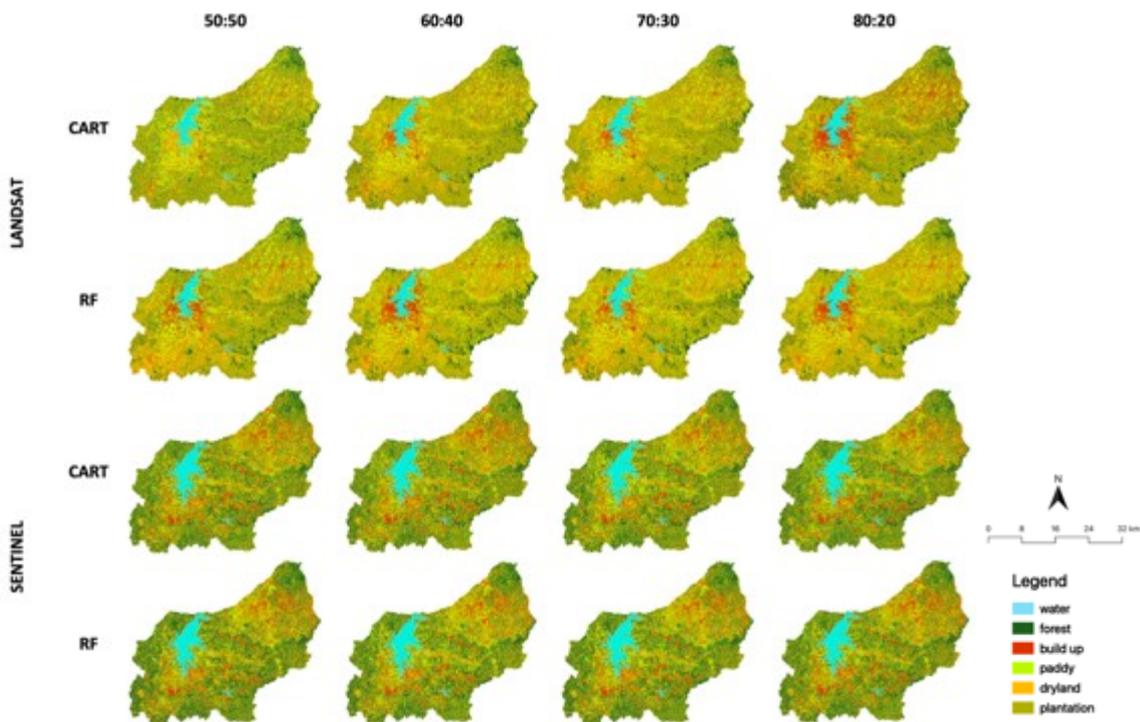
RF classifier merupakan metode klasifikasi yang menggunakan algoritma CART ditambah dengan memilih sampel acak pada setiap simpul yang dihasilkan (Breiman, 2001). Metode klasifikasi RF menghasilkan banyak pohon keputusan dari subset training variable yang dipilih secara acak (Belgiu & Drăgu, 2016). Klasifikasi RF mirip seperti CART namun ditambahkan satu proses tambahan sehingga keputusan diambil dari proses yang lebih kompleks (Basheer et al., 2022).

Hasil dan pembahasan

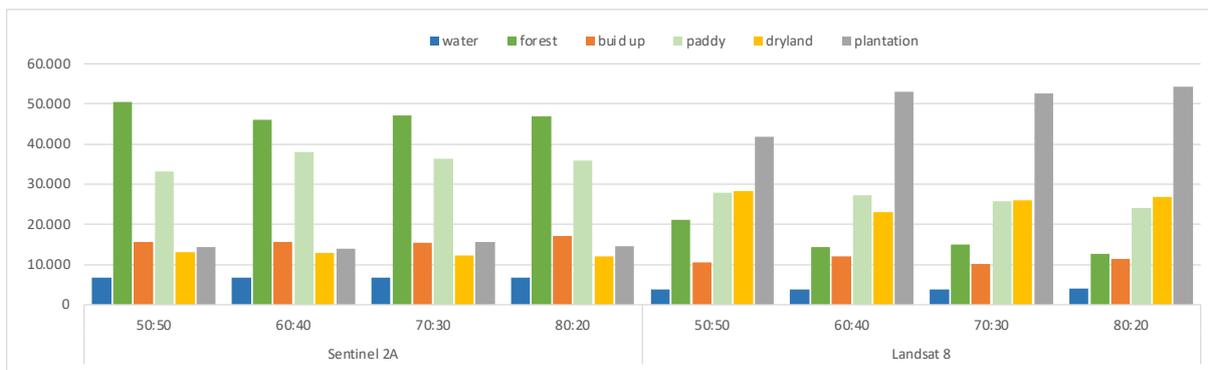
Hasil klasifikasi dari masing-masing metode menunjukkan hasil yang berbeda dan menghasilkan luasan jenis penggunaan lahan yang bervariasi (Gambar 2). Hal ini dimungkinkan terjadi akibat perbedaan citra yang digunakan, citra Sentinel dan Landsat yang digunakan memiliki kenampakan yang cukup berbeda yang disebabkan oleh filter (*ee.filter*) yang diterapkan pada masing-masing citra. Citra Landsat menggunakan filter *median & cloud mask* sementara citra sentinel hanya menggunakan filter *median & awan (no cloud mask)* hal ini karena ketersediaan citra sentinel pada area penelitian cukup terbatas. Keterbatasan data juga mengakibatkan hasil filter median menghasilkan output yang berbeda sehingga periode waktu perekaman yang terpilih berbeda.

Terdapat perbedaan yang cukup signifikan pada luasan masing-masing jenis penggunaan lahan yang dihasilkan citra Landsat dan Sentinel. Hasil klasifikasi citra Landsat lebih banyak menghasilkan kenampakan jenis penggunaan perkebunan (plantation) sementara citra Sentinel lebih banyak kenampakan jenis hutan (forest) (Gambar 3). Perbedaan tersebut dimungkinkan terjadi akibat perbedaan resolusi spasial, citra Sentinel memiliki resolusi spasial lebih detail (10m) sehingga performa klasifikasi lebih baik (Aldiansyah & Saputra, 2022; Immitzer, Vuolo, & Atzberger, 2016). Klasifikasi citra Sentinel menghasilkan luasan jenis penggunaan paddy jauh lebih besar dibandingkan dryland sementara pada citra Landsat jenis penggunaan paddy dan dryland luasanya tidak berbeda jauh. Hal ini

diakibatkan oleh perbedaan masa perekaman citra yang diperoleh dari hasil *filtering* (*ee.filter*), citra Sentinel yang terpilih berada pada periode bulan basah (musim penghujan) sementara citra Landsat yang terpilih berada pada periode bulan kering (musim kemarau). Oleh karena itu hasil klasifikasi Landsat menghasilkan jenis penggunaan dryland lebih banyak dari Sentinel karena pada bulan kering sebagian lahan pertanian padi difungsikan untuk menanam tanaman palawija (Dirgantoro, Suyanto, & Qomariyah, 2015). Tanaman palawija memiliki kerapatan yang berbeda dengan tanaman padi sehingga memungkinkan rentang nilai spektral yang dipantulkan lebih mirip dengan jenis penggunaan dryland.



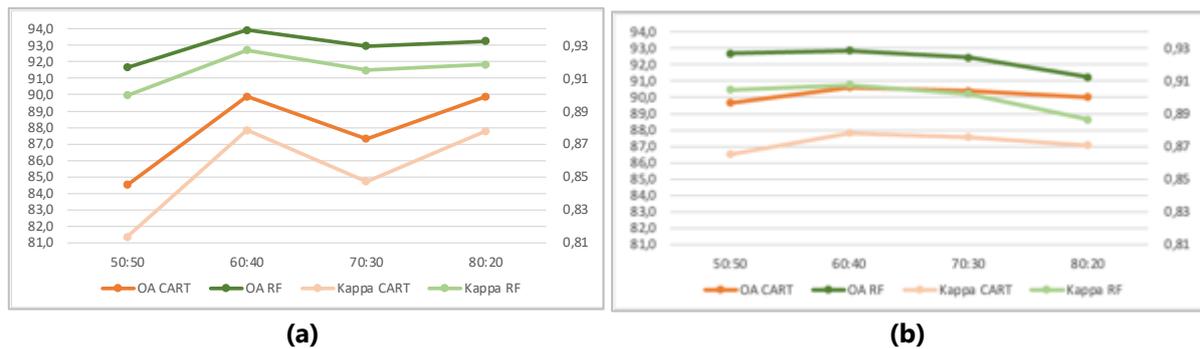
Gambar 2. Peta klasifikasi penggunaan lahan menggunakan citra Landsat 8 dan Sentinel 2A dengan metode CART dan RF dan beberapa proporsi sampel



Gambar 3. Luas area berdasarkan jenis penggunaan lahan hasil klasifikasi metode RF

Tabel 2. Nilai Overall Accuracy (OA) dan Kappa Coefficient (Kappa) berdasarkan besarnya sampel

Citra	Metode	50:50		60:40		70:30		80:20	
		OA (%)	Kappa						
Landsat	CART	84,52	0,814	89,99	0,879	87,32	0,847	89,89	0,878
	RF	91,67	0,900	93,94	0,927	92,96	0,915	93,26	0,918
Sentinel	CART	89,67	0,865	90,61	0,878	90,40	0,876	90,02	0,871
	RF	92,69	0,905	92,86	0,908	92,44	0,902	91,24	0,887



Gambar 4. Nilai OA dan Kappa (a) citra Landsat 8, (b) citra Sentinel 2A

Performa algoritma klasifikasi penggunaan lahan ditunjukkan dari nilai Overall Accuracy (OA), semakin besar nilai OA maka hasil klasifikasi dianggap lebih akurat. Nilai OA dihasilkan dari pengujian silang antara hasil klasifikasi dibandingkan dengan *testing sample* melalui *confusion matrix*. OA merupakan hasil rata-rata yang didapatkan dari Producer's Accuracy (PA) dan User's Accuracy (UA). Akurasi hasil klasifikasi juga dapat dilihat dari Kappa Coefficient, nilai koefisien mendekati 1 maka dianggap lebih akurat dan menunjukkan data raster yang semakin baik (McHugh, 2012). Dalam penelitian ini nilai OA dan Kappa, keduanya digunakan untuk mengetahui sejauh mana akurasi hasil klasifikasi metode CART dan RF pada citra Landsat 8 dan Sentinel 2A.

Pengujian akurasi (*accuracy assesment*) dilakukan pada kedua metode klasifikasi CART dan RF dengan membagi sampel klasifikasi menjadi dua; *training sample* dan *testing sample*. Sampel yang digunakan dalam penelitian ini dibagi kedalam empat proporsi yaitu; 50:50, 60:40, 70:30 dan 80:20. Pembagian sampel bertujuan untuk mengetahui proporsi sampel yang paling efektif untuk menjalankan kedua metode klasifikasi. Hasil perhitungan uji akurasi dari masing-masing metode klasifikasi ditampilkan pada Tabel 2.

Secara umum metode RF memiliki tingkat akurasi yang lebih baik dari CART dengan nilai OA > 90% dan Kappa > 0,90 pada semua skenario pembagian sampel kecuali pada citra Sentinel 80:20 (Kappa=0,87). Metode CART sebagian besar menghasilkan nilai OA < 90% dan Kappa < 0,90, nilai OA > 90% terdapat pada citra Sentinel dengan pembagian sampel 60:40, 70:30 dan 80:20 sementara pada citra Landsat nilai OA < 90%. Beberapa penelitian yang pernah dilakukan juga menunjukkan metode RF memiliki akurasi lebih tinggi dibandingkan dengan metode lainnya (Aldiansyah & Saputra, 2022; Devara & Wijayanto, 2021; Rosyani, Saprudin, & Amalia, 2021).

Melalui skenario pembagian sampel (*training sample & testing sample*) diketahui bahwa nilai OA dan Kappa menunjukkan hasil yang bervariasi, nilainya tidak terlalu jauh namun terdapat perbedaan. Kondisi ini menunjukkan bahwa proporsi sampel berpengaruh terhadap akurasi klasifikasi yang dihasilkan. Secara keseluruhan OA dan Kappa dari masing-masing skenario pembagian sampel memiliki nilai yang linier artinya OA paling tinggi diikuti dengan Kappa tertinggi juga (Gambar 4). Hal ini menunjukkan bahwa keduanya memiliki fungsi yang sama dan dapat digunakan sebagai tolok ukur akurasi hasil klasifikasi penggunaan lahan.

Nilai akurasi paling tinggi pada citra Landsat terdapat pada proporsi sampel 60:40 dengan metode RF (OA=93,94%) sementara paling rendah pada proporsi 50:50 dengan metode CART (OA=84,52). Pada citra Sentinel akurasi paling tinggi terdapat pada proporsi sampel 60:40 dengan metode RF (OA=92,86) dan paling rendah pada proporsi 50:50 dengan metode CART (OA=89,67). Kondisi ini menunjukkan bahwa skenario 60:40 merupakan proporsi pembagian sampel paling efektif pada citra Landsat dan Sentinel, keduanya menghasilkan nilai akurasi yang paling tinggi (OA dan Kappa).

Nilai akurasi yang berbeda dari klasifikasi penggunaan lahan menggunakan citra Landsat dan Sentinel dapat terjadi akibat resolusi spasial yang berbeda, citra Sentinel memiliki resolusi spasial lebih detail (10m) dibandingkan dengan citra Landsat (30m). Perbedaan resolusi spasial berdampak pada tingkat kemudahan interpretasi sampel yang digunakan untuk menjalankan algoritma. Interpretasi sampel lebih mudah dilakukan pada citra Sentinel karena memiliki kerapatan piksel yang lebih rapat dibandingkan dengan citra Landsat. Tingkat kedetilan citra berhubungan dengan tingkat kesulitan pengambilan sampel, ukuran piksel yang tidak terlalu rapat mengakibatkan interpreter kesulitan mendelineasi sampel karena kenampakan yang dihasilkan bias dengan kenampakan di sekitarnya (Kamusoko, 2022). Dalam penelitian ini penulis lebih mudah menginterpretasi penggunaan lahan pada citra Sentinel sehingga pemilihan sampel lebih mudah dilakukan. Berbeda dengan kenampakan yang dihasilkan oleh citra Sentinel, interpretasi penggunaan lahan pada citra Landsat lebih sulit karena perbedaan antar jenis penggunaan lahan tidak terlalu jelas. Hal ini dimungkinkan berkontribusi terhadap lebih rendahnya nilai akurasi yang dihasilkan oleh citra Landsat saat digunakan untuk klasifikasi penggunaan lahan.

Ketersediaan data spasial (citra penginderaan jauh) merupakan salah satu faktor yang cukup krusial dalam analisis penggunaan lahan salah satunya untuk klasifikasi penggunaan lahan. Peta klasifikasi penggunaan lahan dapat digunakan sebagai dasar analisis lanjutan seperti analisis perubahan, analisis dampak dan prediksi penggunaan lahan. Semakin berkembangnya teknologi pengolahan data spasial memunculkan berbagai algoritma yang memudahkan klasifikasi

dan memiliki tingkat akurasi yang cukup baik. Adanya platform pengolahan data spasial seperti GEE sangat membantu dalam pengolahan data, analisis data dan memberikan efisiensi waktu. Sistem *cloud* yang diberikan GEE sangat menyingkat waktu pengolahan data karena menggabungkan platform penyedia data dan platform pengolah data kedalam satu platform sehingga *user* tidak perlu mengunduh data untuk diolah karena bisa diolah langsung pada platform GEE. Berbagai algoritma yang terdapat dalam GEE juga memudahkan pengolahan data.

Cara yang lebih konvensional dalam pengolahan data spasial membutuhkan aplikasi pengolahan yang terinstal dalam komputer yang digunakan. Aplikasi pengolahan data spasial biasanya membutuhkan spesifikasi perangkat keras (*hardware*) yang cukup tinggi untuk mampu menginstal dan menjalankan aplikasi. Data spasial juga harus diunduh ke penyimpanan komputer dan biasanya memerlukan ruang penyimpanan yang cukup besar. Platform GEE didesain sebagai platform berbasis *cloud* yang *compact* dapat melakukan pengolahan data spasial yang mampu menjalankan mekanisme komputasi yang tinggi dan pengolahan data dalam skala yang besar (N. Gorelick et al., 2017). Pengolahan data dengan GEE tidak memerlukan ruang penyimpanan yang besar karena user hanya mengunduh hasil akhir pengolahan. Sistem *cloud* yang dimiliki oleh GEE juga memungkinkan pengolahan data dilakukan dimana saja sehingga lebih fleksibel asalkan perangkat komputer terkoneksi dengan internet.

GEE merupakan platform yang bersifat *open access* artinya bisa diakses oleh siapapun asalkan memiliki izin (*permissions*) salah satunya menggunakan akun *google*. Google akan memberikan API (Application Programming Interface) kepada *user* sehingga dapat membuat *project* dan menuliskan perintah pada Code Editor. Platform GEE bisa memberikan manfaat yang cukup besar kepada peneliti, pemerintah, Non Government Organization (NGO) melalui penyediaan *multi-temporal data* dalam skala besar (*Big Data*) sehingga dapat dijadikan alat untuk melakukan monitoring, perencanaan dan pembuatan kebijakan (Shetty, 2019).

Kesimpulan

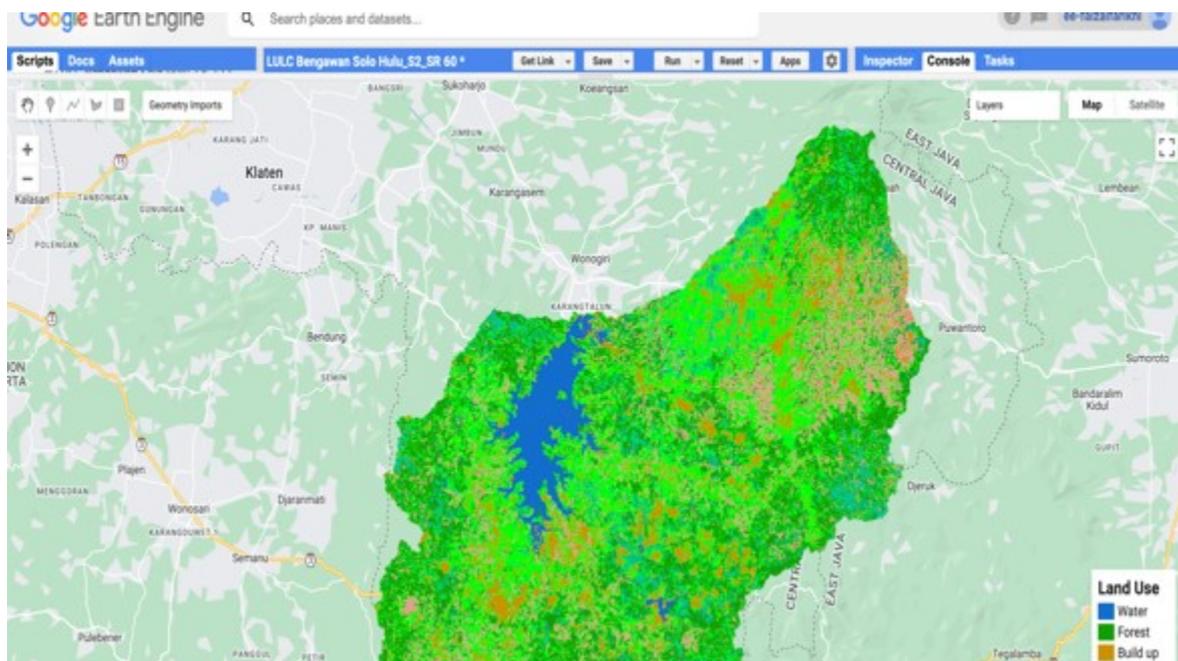
Metode RF merupakan metode klasifikasi yang lebih baik dibandingkan dengan CART ditunjukkan dengan nilai Overall Accuracy (OA) dan Kappa Coeficient (Kappa) yang lebih tinggi. Proporsi sampel yang paling baik untuk menjalankan algoritma klasifikasi penggunaan lahan menggunakan platform Google Earth Engine (GEE) adalah 60:40 dengan hasil akurasi paling tinggi. Citra Landsat 8 menghasilkan rata-rata akurasi lebih tinggi sementara pemilihan sampel lebih mudah menggunakan citra Sentinel 2A karena kenampakan yang dihasilkan lebih mudah diinterpretasi. Penelitian ini dapat dimodifikasi dengan variasi; kelas penggunaan lahan, area penelitian, data yang digunakan, periode waktu, kondisi iklim dan kondisi geografis yang berbeda. Beberapa penelitian menunjukkan nilai akurasi berbeda terutama besarnya proporsi sampel yang menghasilkan akurasi optimal.

Referensi

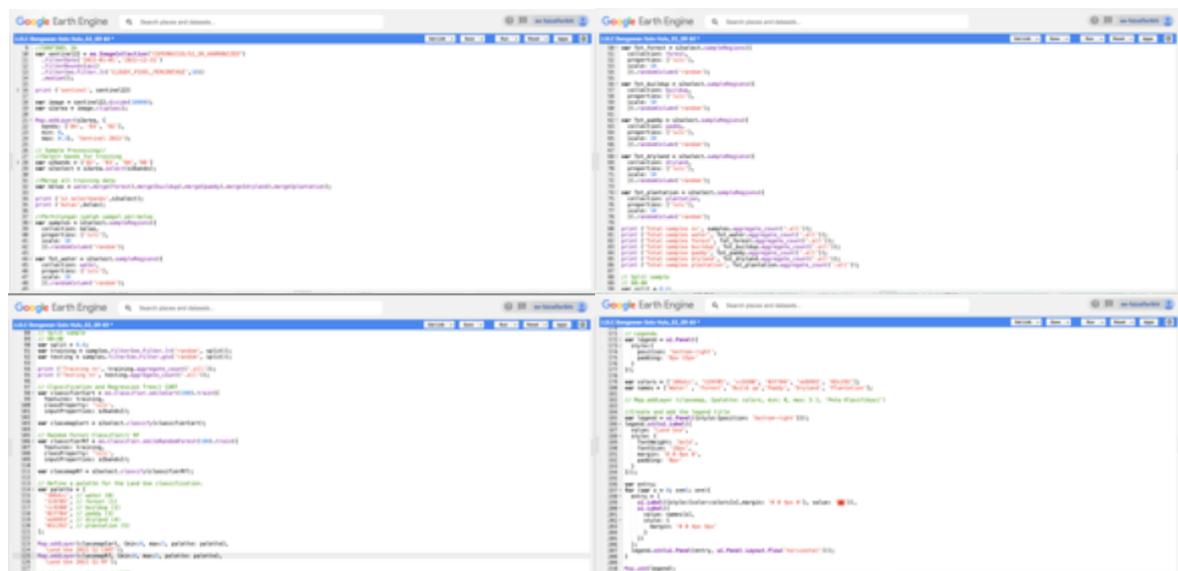
- Aldiansyah, S., & Saputra, R. A. (2022). Comparison of Machine Learning Algorithms for Land Use and Land Cover Analysis Using Google Earth Engine (Case Study: Wanggu Watershed). *Article in International Journal of Remote Sensing and Earth Sciences*, 19(2), 197–210. Retrieved from <https://www.researchgate.net/publication/367379951>
- Ambarwulan, W., Yulianto, F., Widiatmaka, W., Rahadiati, A., Tarigan, S. D., Firmansyah, I., & Hasibuan, M. A. S. (2023). Modelling land use/land cover projection using different scenarios in the Cisadane Watershed, Indonesia: Implication on deforestation and food security. *Egyptian Journal of Remote Sensing and Space Science*, 26(2), 273–283. <https://doi.org/10.1016/j.ejrs.2023.04.002>
- Arsyad, S. (2010). *No Title* (Edisi Kedu). Bogor: IPB Press.
- Basheer, S., Wang, X., Farooque, A. A., Nawaz, R. A., Liu, K., Adekanmbi, T., & Liu, S. (2022). Comparison of Land Use Land Cover Classifiers Using Different Satellite Imagery and Machine Learning Techniques. *Remote Sensing*, 14(19), 1–18. <https://doi.org/10.3390/rs14194978>
- Belgiu, M., & Drăgu, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, 5–32.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification And Regression Tree*. Florida: Chapman & Hall/CRC Press.
- Devara, T., & Wijayanto, A. W. (2021). Machine Learning Applied To Sentinel-2 and Landsat-8 Multispectral and Medium-Resolution Satellite Imagery for the Detection of Rice Production Areas in Nganjuk, East Java, Indonesia. *International Journal of Remote Sensing and Earth Sciences (IJReSES)*, 18(1), 19. <https://doi.org/10.30536/ijjreses.2021.v18.a3538>
- Dirgantoro, N. W. F., Suyanto, & Qomariyah, S. (2015). Simulasi Prediksi Pola Tata Tanam Di Das Tirtomoyo Berdasarkan Neraca Air. *Matriks Teknik Sipil*, 351.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202(2016), 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Gorelick, S. M. (2015). Global change and the groundwater management challenge. *Water Resources Research*, 3031–3051. <https://doi.org/10.1002/2014WR016825>
- Immitzer, M., Vuolo, F., & Atzberger, C. (2016). First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote Sensing*, 8(3). <https://doi.org/10.3390/rs8030166>
- JICA, J. I. C. A. (2007). *Studi Penanganan Sedimentasi Waduk Serbaguna Wonogiri Republik Indonesia*.
- Kamusoko, C. (2022). Land Cover Classification Accuracy Assessment. *Springer Geography*, 80, 105–118. https://doi.org/10.1007/978-981-16-5149-6_6
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemica Medica*, 22(3), 276–282. Retrieved from <https://hrcak.srce.hr/89395>

- Rahman, M. M., Harisuseno, D., & Sisinggih, D. (2012). Studi Penanganan Konservasi Lahan di Sub DAS Keduang DAS Bengawan Solo Kabupaten Wonogiri. *Jurnal Teknik Pengairan*, 3, 250–257.
- Rosyani, P., Saprudin, S., & Amalia, R. (2021). Klasifikasi Citra Menggunakan Metode Random Forest dan Sequential Minimal Optimization (SMO). *Jurnal Sistem Dan Teknologi Informasi (Justin)*, 9(2), 132. <https://doi.org/10.26418/justin.v9i2.44120>
- Roy, D. P., Wulder, M. A., Loveland, T. R., C.E., W., Allen, R. G., Anderson, M. C., ... Zhu, Z. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, 145, 154–172. <https://doi.org/10.1016/j.rse.2014.02.001>
- Santosa, T. J. I. B. (2016). Analysis of Sedimentation in Wonogiri Reservoir. *Jurnal of the Civil Engineering Forum*, 2(1), 27–32.
- Shetty, S. (2019). *Analysis of Machine Learning Classifiers for LULC Classification on Google Earth Engine*. 8(23), 1–22.
- Song, Y. Y., & Lu, Y. (2015). Decision tree methods: applications for classification and prediction. *Shanghai Archives of Psychiatry*, 27(2), 130–135. <https://doi.org/10.11919/j.issn.1002-0829.215044>
- Vink, A. P. A. (1975). *Land Use in Advancing Agriculture*. New York: Springer Verlag.
- Yeung, A. K. W., & Hall, G. B. (2014). Spatial Database Systems: Design, Implementation and Project Management. *The GeoJournal Library*, 87.
- Yulianto, F., Khomarudin, M. R., Hermawan, E., Nugroho, N. P., Chulafak, G. A., Nugroho, G., ... Priyanto, E. (2022). Spatial and temporal distribution of estimated surface runoff caused by land use/land cover changes in the upstream Citarum watershed, West Java, Indonesia. *Journal of Degraded and Mining Lands Management*, 9(2), 3293–3305. <https://doi.org/10.15243/jdmlm.2022.092.3293>
- Yulianto, F., Nugroho, G., & Suwarsono, S. (2021). *Improvement in the Accuracy of the Postclassification of Land Use and Land Cover Using Landsat 8 Data Based on the Majority of Segment-Based Filtering Approach*. 2021.

Lampiran



Lampiran Gambar 1. Tampilan platform GEE hasil klasifikasi metode RF proporsi sampel 60:40 menggunakan citra Sentinel 2A



Lampiran Gambar 2. Tampilan JavaScript pada Code Editor GEE untuk klasifikasi metode RF proporsi sampel 60:40 menggunakan citra Sentinel 2A