

Analisis Klasifikasi Tutupan Lahan Citra Multispektral Sentinel-2 Pada Google Earth Engine

Raden Roro Anna Dyah Retno Manuhoro^{1*}, Lucia Sandra Budiman¹, Wenang Anurogo²

¹Program Studi Sains Informasi Geografi, Fakultas Sains dan Teknologi, Universitas Mahakarya Asia, 55285, Indonesia

²Laboratory for Sustainable Architecture Integrated Design (SAID) and Urbanism Departments of Architecture & Global Smart City, SungKyunKwan University, 2066 Seobu-ro, 16419 Suwon, Republic of Korea

email: annamanuhoro@gmail.com

Info Artikel :

Diterima :

20-07-2026

Disetujui :

26-07-2026

Dipublikasikan :

27-07-2026

ABSTRAK

Tutupan lahan merupakan representasi kondisi fisik permukaan bumi yang terdiri atas unsur alami maupun buatan manusia dan memiliki peran penting dalam berbagai kajian lingkungan, seperti perencanaan tata ruang, pengelolaan sumber daya alam, mitigasi bencana, dan pemantauan perubahan lingkungan. Kota Semarang mengalami perkembangan wilayah yang pesat akibat urbanisasi dan pembangunan infrastruktur yang memicu perubahan tutupan lahan. Penelitian ini bertujuan mengevaluasi hasil klasifikasi tutupan lahan di sebagian wilayah Kota Semarang menggunakan citra satelit Sentinel-2 yang memiliki resolusi spasial 10–60 meter serta informasi multispektral yang mendukung proses klasifikasi. Untuk meningkatkan ketelitian analisis, digunakan transformasi indeks spektral berupa *Normalized Difference Vegetation Index* (NDVI), *Normalized Difference Water Index* (NDWI), dan *Normalized Difference Built-up Index* (NDBI) sebagai variabel tambahan. Pengolahan data dilakukan melalui platform *Google Earth Engine* (GEE) yang berbasis cloud computing sehingga analisis dapat berlangsung secara cepat dan efisien tanpa pengunduhan data dalam jumlah besar. Klasifikasi tutupan lahan dilakukan menggunakan metode K-Means dan *Random Forest*, kemudian dievaluasi dengan *Overall Accuracy* dan Kappa Coefficient untuk menilai tingkat akurasi hasil klasifikasi. Hasil penelitian diharapkan dapat memberikan informasi yang andal mengenai kondisi tutupan lahan serta menjadi dasar dalam perencanaan tata ruang dan pengelolaan lingkungan perkotaan yang berkelanjutan.

Kata Kunci: Tutupan Lahan, Klasifikasi Multispektral, Citra Sentinel-2, Google Earth Engine (GEE), Kota Semarang

ABSTRACT

Land cover represents the physical condition of the Earth's surface, comprising both natural and human-made elements, and plays an essential role in environmental studies such as spatial planning, natural resource management, disaster mitigation, and environmental change monitoring. Semarang City has experienced rapid spatial development driven by urbanization and infrastructure expansion, which has led to significant land cover changes. This study aims to evaluate land cover classification results in part of Semarang City using Sentinel-2 satellite imagery, which provides spatial resolution of 10–60 meters and multispectral information suitable for classification purposes. To improve classification accuracy, spectral index transformations including the Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), and Normalized Difference Built-up Index (NDBI) were applied as additional variables. Data processing was conducted through the Google Earth Engine (GEE) platform, which enables fast and efficient cloud-based computation without the need to download large datasets. Land cover classification was performed using two multispectral classification methods, namely K-Means and Random Forest, and was evaluated using Overall Accuracy and the Kappa Coefficient to assess the reliability of the classification results. The findings are expected to provide reliable spatial information on land cover conditions and serve as a basis for urban spatial planning and sustainable environmental management.

Keywords : Land cover, Multispectral Classification, Sentinel-2 Imagery, Google Earth Engine (GEE), Semarang City



©2026 Penulis. Ini adalah artikel akses terbuka di bawah lisensi Creative Commons Attribution Non Commercial 4.0 International License.
(<https://creativecommons.org/licenses/by-nc/4.0/>)

PENDAHULUAN

Tutupan lahan (*land cover*) merupakan representasi kondisi fisik permukaan bumi yang terdiri atas unsur alami maupun buatan manusia. Informasi mengenai tutupan lahan memiliki peran yang sangat penting dalam berbagai aspek pengelolaan lingkungan, antara lain perencanaan tata ruang, pengelolaan sumber daya alam, pemantauan perubahan lingkungan, mitigasi bencana, serta evaluasi pembangunan berkelanjutan.

Kemajuan teknologi penginderaan jauh memberikan peluang yang besar dalam penyediaan informasi tutupan lahan secara cepat, konsisten, dan mencakup wilayah yang luas. Salah satu sumber data yang banyak dimanfaatkan dalam penelitian tutupan lahan adalah citra satelit Sentinel-2 yang dikembangkan oleh *European Space Agency* (ESA) melalui program *Copernicus*. Sentinel-2 memiliki sensor multispektral dengan 13 saluran spektral dan resolusi spasial 10, 20, dan 60 meter sehingga mampu merekam karakteristik objek permukaan bumi secara lebih rinci. Selain memiliki kualitas radiometrik yang baik, data Sentinel-2 tersedia secara bebas (*open access*) sehingga menjadi salah satu sumber data utama dalam berbagai penelitian pemetaan dan klasifikasi tutupan lahan (European Space Agency, 2015).

Klasifikasi citra multispektral merupakan proses pengelompokan piksel berdasarkan karakteristik respons spektralnya sehingga setiap piksel dapat diidentifikasi ke dalam kelas tutupan lahan tertentu. Menurut Campbell dan Wynne (2011), klasifikasi multispektral umumnya dilakukan menggunakan dua pendekatan utama, yaitu klasifikasi terawasi (*supervised classification*) dan tidak terawasi (*unsupervised*). Seiring berkembangnya metode pengolahan citra digital, berbagai algoritma berbasis seperti *Random Forest*, semakin banyak digunakan karena mampu meningkatkan akurasi klasifikasi dibandingkan metode statistik konvensional.

Transformasi indeks spektral berupa *Normalized Difference Vegetation Index* (NDVI), *Normalized Difference Water Index* (NDWI) dan *Normalized Difference Built-up Index* (NDBI) merupakan parameter yang digunakan dalam analisis tutupan lahan untuk menggambarkan karakteristik vegetasi dan kawasan terbangun secara spasial. NDVI berperan dalam memperjelas kelas vegetasi, sedangkan NDBI memperkuat identifikasi kawasan terbangun. Kombinasi kedua indeks tersebut telah banyak diterapkan untuk meningkatkan akurasi klasifikasi tutupan lahan, terutama pada wilayah yang memiliki heterogenitas objek tutupan lahan yang tinggi seperti kawasan perkotaan (Rodi Sipondak, 2026). Selain vegetasi dan bangunan, identifikasi badan air juga dapat ditingkatkan menggunakan *Normalized Difference Water Index* (NDWI). McFeeters (1996) mengembangkan NDWI dengan memanfaatkan kombinasi saluran/band *Green* dan *Near Infrared* (NIR) untuk meningkatkan kontras antara badan air dengan objek di sekitarnya. Kombinasi ketiga indeks ini meningkatkan akurasi pemisahan objek secara signifikan dengan memanfaatkan respons spektral spesifik dari masing-masing tutupan permukaan bumi.

Perkembangan teknologi komputasi awan (*cloud computing*) turut mengubah paradigma pengolahan data penginderaan jauh. Google Earth Engine (GEE) merupakan platform komputasi geospasial berbasis cloud yang menyediakan akses terhadap arsip data satelit dalam skala global. Selain menyediakan data, GEE juga dilengkapi dengan berbagai fungsi pemrosesan citra, analisis spasial, serta algoritma klasifikasi yang dapat dijalankan secara daring tanpa memerlukan komputer dengan spesifikasi tinggi. Kemampuan tersebut menjadikan GEE sebagai salah satu platform yang paling banyak dimanfaatkan dalam penelitian penginderaan jauh karena mampu mempercepat proses analisis data berskala besar secara efisien (Gorelick et al., 2017).

Berbagai penelitian terdahulu telah menunjukkan bahwa kombinasi citra Sentinel-2 dan Google Earth Engine mampu menghasilkan klasifikasi tutupan lahan dengan tingkat akurasi yang tinggi. Gorelick et al. (2017) menunjukkan bahwa pemanfaatan GEE secara signifikan meningkatkan efisiensi pengolahan data penginderaan jauh dalam skala regional hingga global. Selanjutnya, penelitian Phiri et al. (2020) menunjukkan bahwa Sentinel-2 memiliki kemampuan yang baik dalam membedakan berbagai kelas tutupan lahan karena didukung oleh resolusi spasial yang relatif tinggi serta informasi spektral yang lengkap. Penelitian lainnya juga melaporkan bahwa algoritma *Random Forest* pada platform GEE mampu menghasilkan nilai *Overall Accuracy* di atas 85% pada berbagai kondisi wilayah penelitian. Meskipun demikian, sebagian besar penelitian lebih berfokus pada pengembangan algoritma klasifikasi atau pemetaan perubahan tutupan lahan, sedangkan kajian yang secara khusus mengevaluasi kualitas hasil klasifikasi berdasarkan parameter akurasi pada wilayah perkotaan di Indonesia masih relatif terbatas.

Kota Semarang merupakan salah satu kota metropolitan di Indonesia yang mengalami perkembangan wilayah sangat pesat sebagai akibat pertumbuhan penduduk, urbanisasi, pembangunan infrastruktur, kawasan industri, serta ekspansi permukiman. Dinamika tersebut menyebabkan perubahan tutupan lahan yang cukup signifikan, terutama berupa konversi lahan pertanian, vegetasi, dan ruang terbuka menjadi kawasan terbangun. Perubahan tersebut berimplikasi terhadap meningkatnya kerentanan banjir, berkurangnya ruang terbuka hijau, menurunnya kapasitas infiltrasi tanah, serta

degradasi kualitas lingkungan perkotaan. Oleh karena itu, penyediaan informasi tutupan lahan yang akurat menjadi sangat penting sebagai dasar dalam perencanaan tata ruang, pengelolaan lingkungan, dan pengambilan kebijakan pembangunan berkelanjutan.

Penelitian ini diharapkan mampu menghasilkan informasi mengenai kualitas dan tingkat akurasi klasifikasi tutupan lahan berdasarkan citra Sentinel-2 serta memberikan gambaran mengenai efektivitas Google Earth Engine sebagai platform pengolahan data penginderaan jauh. Hasil penelitian diharapkan dapat menjadi referensi ilmiah dalam pengembangan metode klasifikasi tutupan lahan sekaligus menyediakan informasi spasial yang akurat untuk mendukung perencanaan tata ruang, pengelolaan lingkungan, dan pembangunan wilayah yang berkelanjutan.

METODE PENELITIAN

Penelitian ini menggunakan data citra satelit Sentinel-2 *Multispektral Instrument* (MSI) pada sebagian wilayah Kota Semarang. Pengolahan data dilakukan berbasis *cloud computing* pada *platform Google Earth Engine* (GEE), sehingga seluruh tahapan analisis dilaksanakan melalui *script* yang terintegrasi tanpa memerlukan pengunduhan data citra dalam jumlah besar. Tahapan penelitian diawali dengan pengambilan data citra Sentinel-2 yang disesuaikan dengan area kajian dilengkapi dengan saluran/band yang akan digunakan. Selanjutnya data citra tersebut dilakukan pemilihan dengan meminimalkan keberadaan awan (*cloud masking*) agar area tutupan lahan yang akan diolah dapat teridentifikasi lebih optimal.

Transformasi indeks spektral yang digunakan yaitu NDVI (*Normalized Difference Vegetation Index*), NDWI (*Normalized Difference Water Index*), dan NDBI (*Normalized Difference Built-up Index*). Transformasi NDVI merupakan kombinasi matematis antara saluran/band merah/*Red* dan band NIR (*Near-Infrared*) sebagai indikator keberadaan dan kondisi vegetasi. Sedangkan NDWI menggunakan saluran/band hijau dan NIR untuk membedakan air dari vegetasi atau tanah. Band SWIR (*Short-wave Infrared*) memiliki reflektansi tinggi pada objek bangunan, jalan dan struktur perkotaan yang kemudian dipadukan dengan band NIR pada transformasi NDBI untuk membedakan area terbangun dari objek lain. Ketiga transformasi indeks spektral tersebut dikombinasikan dengan band multispektral Sentinel-2 sebagai variabel masukan dalam proses klasifikasi tutupan lahan.

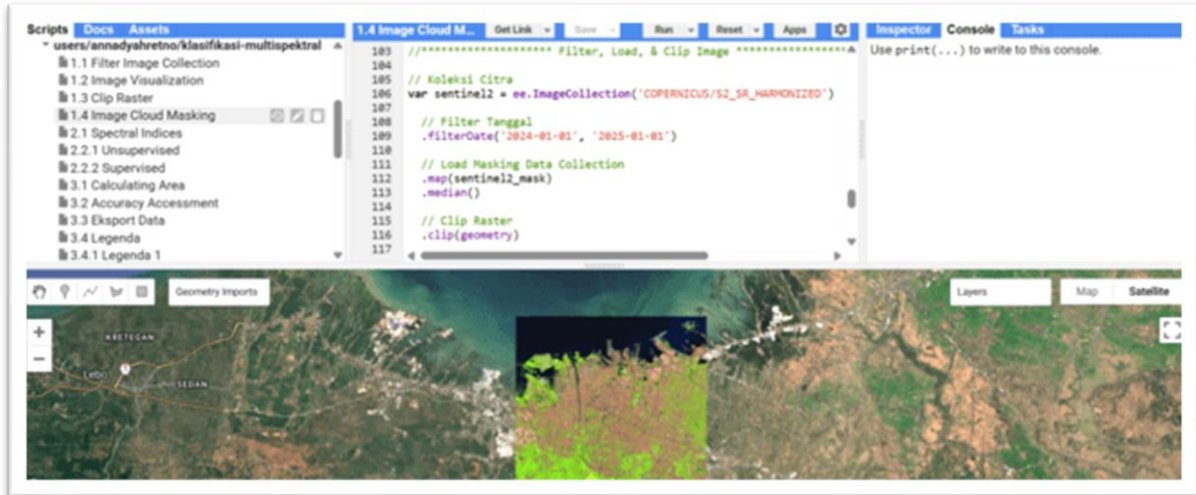
Klasifikasi tutupan lahan dilakukan dengan menggunakan dua pendekatan yaitu klasifikasi *unsupervised* berupa *K-means* dan klasifikasi *supervised* berupa *random forest*. Metode K-Means digunakan untuk mengelompokkan piksel berdasarkan kemiripan karakteristik spektral tanpa data pelatihan, sedangkan Random Forest memanfaatkan sampel pelatihan (*training samples*) yang mewakili masing-masing kelas tutupan lahan untuk membangun model klasifikasi. Hasil klasifikasi selanjutnya dievaluasi melalui uji akurasi menggunakan *Confusion Matrix* dengan membandingkan hasil klasifikasi terhadap data referensi. Tingkat keakuratan klasifikasi dinilai menggunakan parameter *Overall Accuracy* dan *Kappa Coefficient* sehingga diperoleh informasi mengenai tingkat keandalan metode klasifikasi dalam merepresentasikan kondisi tutupan lahan pada wilayah penelitian.

HASIL DAN PEMBAHASAN

Tahap awal pengolahan data dilakukan dengan memperoleh citra satelit Sentinel-2 *Surface Reflectance (COPERNICUS/S2_SR_HARMONIZED)* melalui *platform Google Earth Engine* (GEE). Data citra dipilih berdasarkan batas *Area Of Interest* (AOI) yang mencakup sebagian wilayah Kota Semarang, kemudian difilter berdasarkan rentang waktu perekaman dan persentase tutupan awan yang rendah untuk memperoleh citra dengan kualitas terbaik. Proses penyaringan ini bertujuan mengurangi pengaruh awan dan bayangannya yang dapat mengganggu proses identifikasi objek permukaan bumi serta menurunkan akurasi hasil klasifikasi.

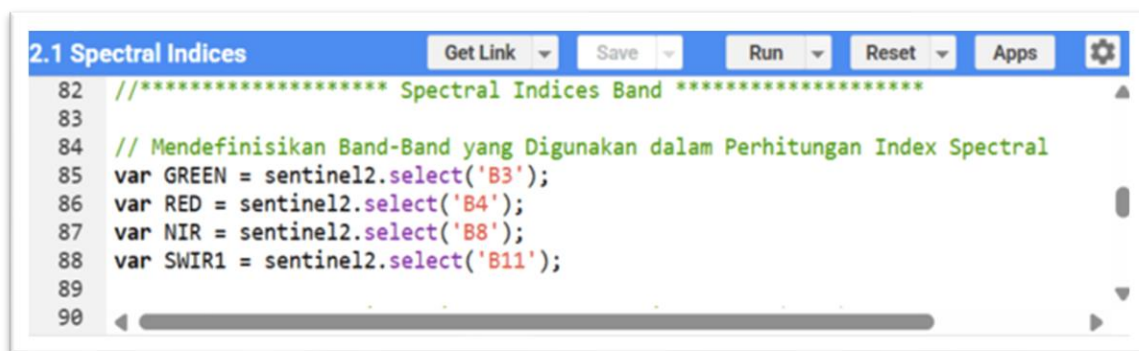
Selanjutnya dilakukan proses *cloud masking* menggunakan band kualitas (*Quality Assessment/QA*) yang tersedia pada data Sentinel-2. Proses ini berfungsi untuk menghilangkan piksel yang teridentifikasi sebagai awan maupun bayangan awan sehingga hanya piksel yang merepresentasikan kondisi permukaan bumi yang digunakan dalam analisis. Setelah proses *masking*, beberapa citra yang telah memenuhi kriteria digabungkan (*mosaic*) menjadi satu citra komposit. Teknik *mosaic* digunakan untuk mengatasi keterbatasan cakupan spasial maupun adanya bagian citra yang masih tertutup awan pada waktu perekaman tertentu, sehingga menghasilkan citra yang lebih utuh dan memiliki kualitas visual yang lebih baik. Pada *Google Earth Engine*, seluruh proses tersebut dilakukan melalui penyusunan *script JavaScript* yang meliputi pemanggilan koleksi data

COPERNICUS/S2_SR_HARMONIZED, penyaringan berdasarkan area kajian (*filterBounds*), penyaringan waktu perekaman (*filterDate*), penerapan fungsi *cloud masking*, pembentukan citra komposit menggunakan metode (*median*), serta pemotongan citra sesuai batas wilayah penelitian (*clip*). Hasil dari tahapan ini berupa citra komposit Sentinel-2 yang telah bebas dari gangguan awan dan siap digunakan sebagai data masukan untuk proses transformasi indeks spektral (NDVI, NDWI, dan NDBI) serta klasifikasi tutupan lahan.



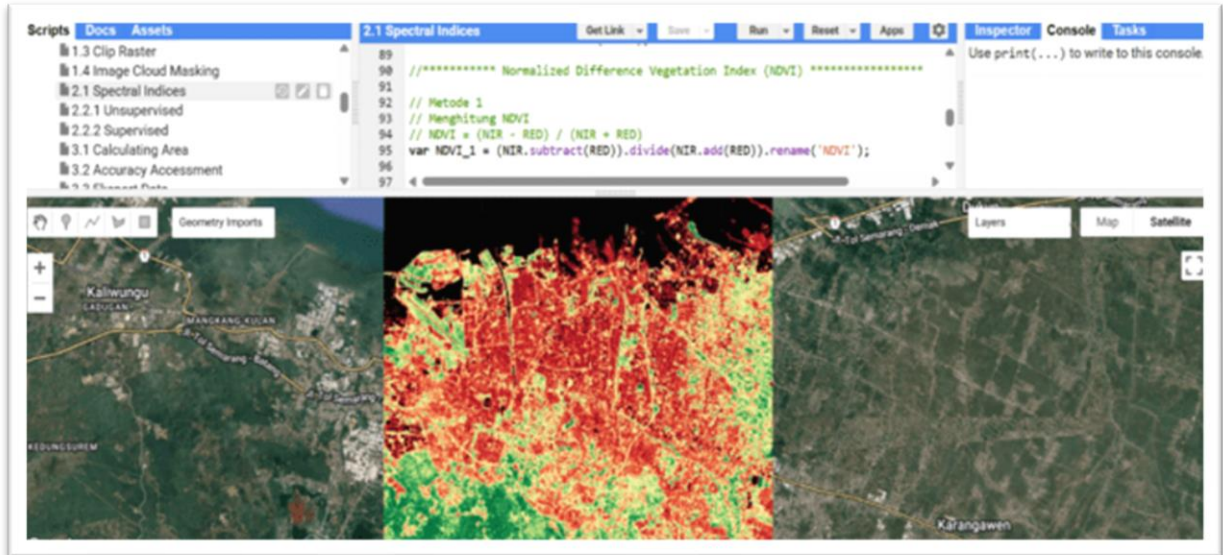
Gambar 1 Data Liputan citra hasil Mosaic Citra dan Cloud Masking beserta data Script

Indeks spektral dalam klasifikasi multispektral tutupan lahan adalah sebagai penambah informasi spektral untuk meningkatkan akurasi dan kemampuan algoritma dalam membedakan kelas tutupan lahan. Hal ini diperlukan karena terkadang nilai spektral dari obyek-obyek permukaan bumi pada band tertentu terdapat kemiripan, sehingga transformasi indeks spektral ini dapat lebih mendefinisikan perbedaan kelas tutupan lahan. Sebelum dilakukan tahapan transformasi indeks spektral, terlebih dahulu dilakukan pendefinisian band-band yang akan dilakukan perhitungan, yaitu meliputi band hijau (*Green*), merah (*Red*), inframerah dekat (*Near-Infrared/NIR*), dan inframerah gelombang pendek (*Short-Wave Infrared/SWIR*).



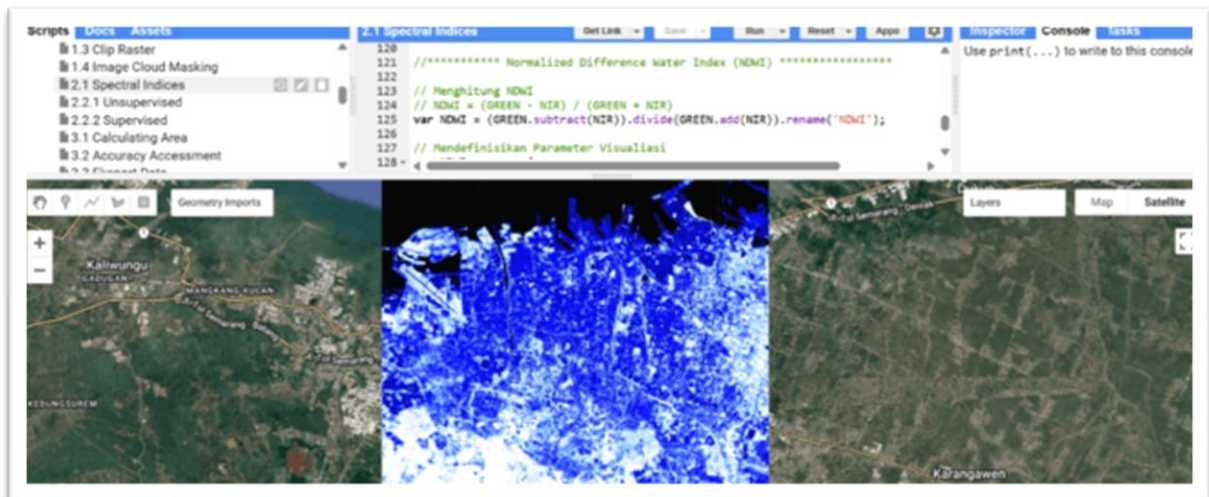
Gambar 2 Data Script Pendefinisian Band-Band untuk Transformasi Indeks Spektral

Transformasi spektral NDVI untuk mendefinisikan tingkat kerapatan vegetasi atau tingkat kehijauan vegetasi, sehingga digunakan band merah dan band inframerah dekat agar dapat membedakan obyek dengan vegetasi sehat dan yang tidak sehat atau obyek non vegetasi. Obyek dengan warna hijau memiliki nilai NDVI yang lebih tinggi daripada obyek dengan warna merah dan hitam atau gelap, seperti yang ditunjukkan pada Gambar 3.



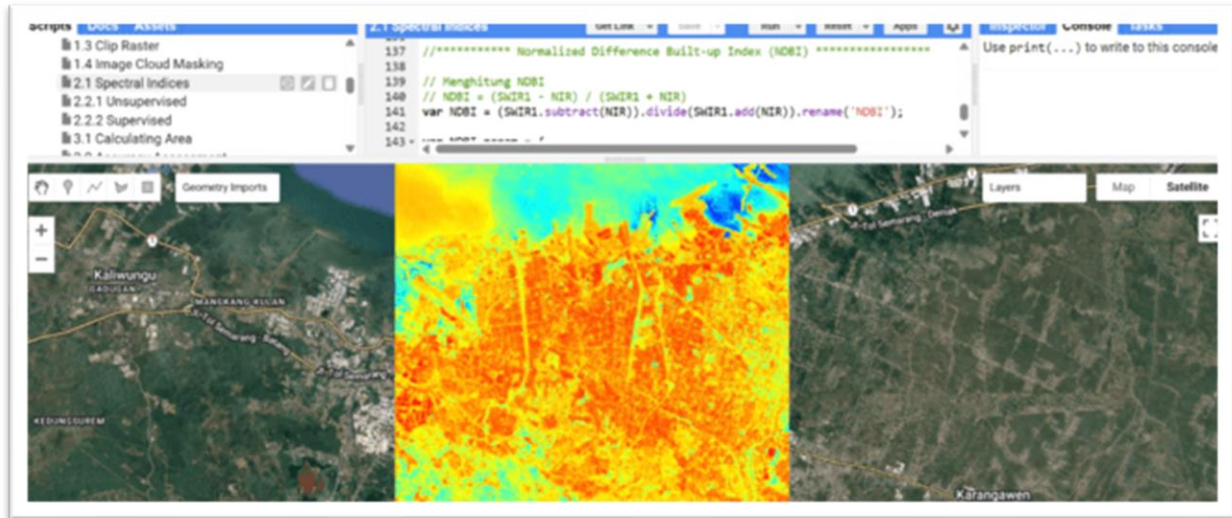
Gambar 3 Hasil Transformasi NDVI beserta Data Script

Transformasi spektral NDWI untuk menekankan keberadaan badan air serta memisahkan dari obyek vegetasi dan tanah atau lahan terbangun. Jika nilai NDWI tinggi maka obyek tersebut memiliki kandungan air tinggi, sedangkan nilai NDWI rendah atau negatif maka obyek tersebut adalah lahan terbangun.



Gambar 4. Hasil Transformasi NDWI beserta Data Script

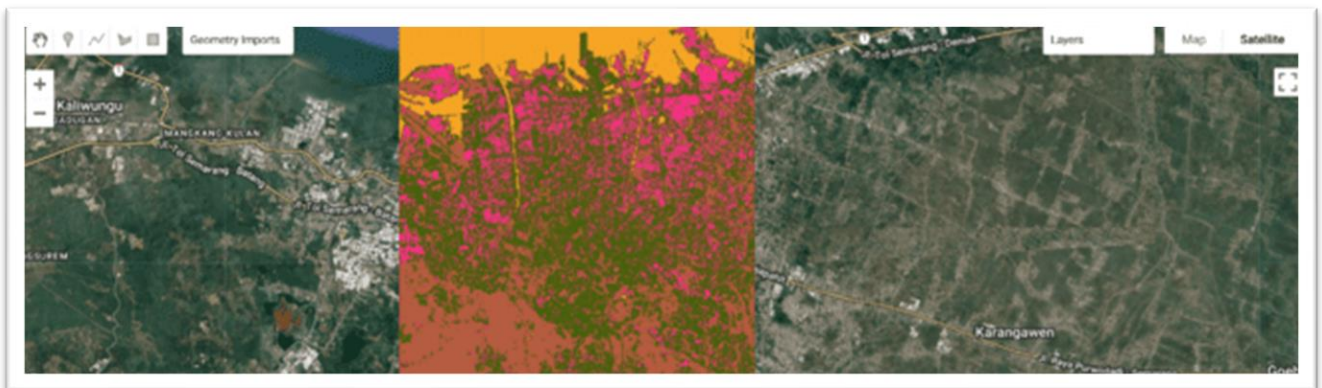
Transformasi spektral NDBI untuk mengidentifikasi area terbangun, seperti permukiman dan infrastruktur dengan menggunakan band SWIR. Jika nilai NDBI tinggi maka merupakan obyek lahan terbangun, sebaliknya jika nilai NDBI rendah maka merupakan obyek vegetasi atau air.



Gambar 5. Hasil Transformasi NDBI beserta Data Script

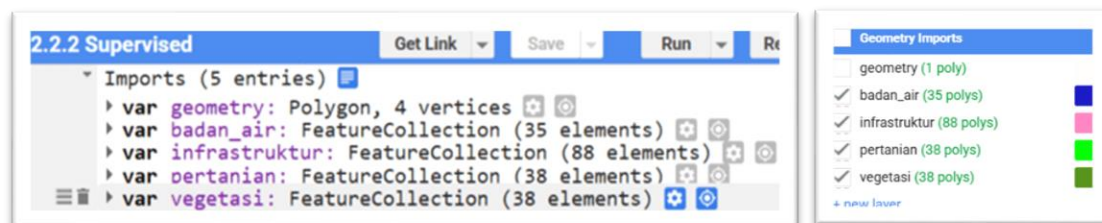
Klasifikasi *unsupervised* pada penelitian ini dilakukan menggunakan algoritma K-Means yang tersedia pada platform Google Earth Engine (GEE). Algoritma K-Means merupakan salah satu metode klasifikasi tidak terawasi (*unsupervised classification*) yang mengelompokkan piksel citra ke dalam sejumlah kelompok (*cluster*) berdasarkan tingkat kemiripan karakteristik spektralnya tanpa menggunakan data latih (*training data*). Pada metode ini, proses pengelompokan dilakukan secara otomatis oleh sistem melalui perhitungan jarak antar nilai spektral setiap piksel, sehingga piksel yang memiliki karakteristik spektral serupa akan tergabung ke dalam kelompok yang sama. Oleh karena itu, peneliti tidak menentukan kelas objek pada tahap awal, melainkan hanya menentukan jumlah *cluster* yang akan dibentuk. Dalam penelitian ini, jumlah *cluster* ditetapkan sebanyak empat kelas yang disesuaikan dengan karakteristik tutupan lahan pada sebagian wilayah Kota Semarang. Setelah proses klasifikasi selesai, setiap *cluster* yang dihasilkan diinterpretasikan berdasarkan karakteristik visual citra, nilai spektral, serta kondisi tutupan lahan yang terdapat pada wilayah penelitian. Tahap interpretasi ini diperlukan karena hasil klasifikasi *unsupervised* hanya menghasilkan kelompok piksel berdasarkan kemiripan spektral tanpa memberikan informasi mengenai jenis tutupan lahannya.

Penerapan algoritma K-Means pada Google Earth Engine memberikan kemudahan dalam proses klasifikasi karena seluruh tahapan komputasi dilakukan secara otomatis pada sistem *cloud computing*. Meskipun algoritma K-Means mampu mengelompokkan piksel secara otomatis berdasarkan kemiripan spektral, hasil klasifikasi masih memiliki beberapa keterbatasan. Objek yang memiliki karakteristik spektral yang hampir sama berpotensi tergabung dalam satu kelompok meskipun secara nyata merupakan kelas tutupan lahan yang berbeda. Sebagai contoh, lahan pertanian pada kondisi tertentu dapat memiliki karakteristik spektral yang mendekati vegetasi atau lahan terbuka, sehingga memungkinkan terjadinya kesalahan klasifikasi. Oleh karena itu, hasil klasifikasi *unsupervised* pada penelitian ini digunakan sebagai gambaran awal distribusi tutupan lahan dan selanjutnya dibandingkan dengan hasil klasifikasi *supervised* menggunakan algoritma *Random Forest* yang memanfaatkan data latih sehingga diharapkan menghasilkan tingkat akurasi yang lebih tinggi.



Gambar 6. Hasil Klasifikasi Tutupan Lahan Algoritma *K-Means*

Selanjutnya, metode klasifikasi *supervised* yang digunakan adalah *Random Forest*, dimana metode klasifikasi ini menggunakan sekumpulan *decision tree* untuk memetakan tutupan lahan. Berbeda dengan klasifikasi *unsupervised*, metode ini membutuhkan data latih (*training sample*) yang mewakili setiap kelas tutupan lahan. Pada tahap ini, data latih dibuat dalam bentuk poligon (*FeatureCollection*) yang didigitasi secara manual berdasarkan interpretasi visual citra Sentinel-2 dan kondisi objek pada wilayah penelitian. Empat kelas tutupan lahan yang digunakan sebagai data pelatihan terdiri atas badan air, infrastruktur, pertanian, dan vegetasi. Jumlah sampel pada masing-masing kelas dibuat relatif seimbang agar model memperoleh informasi spektral yang representatif dan tidak mengalami bias terhadap kelas tertentu. Sebagaimana ditunjukkan pada Gambar 7, jumlah sampel yang digunakan terdiri atas 35 poligon untuk kelas badan air, 88 poligon untuk kelas infrastruktur, 38 poligon untuk kelas pertanian, dan 38 poligon untuk kelas vegetasi.



Gambar 7. Data Script Pengambilan Sampel Data Training

Pada Gambar 7, diketahui bahwa data training sample telah diambil pada setiap kelas tutupan lahannya, yaitu badan air, infrastruktur, pertanian, dan vegetasi, dengan sebaran secara merata pada area liputan citra. Penyebaran sampel secara merata bertujuan agar variasi karakteristik spektral setiap kelas dapat terwakili dengan baik, sehingga model klasifikasi mampu mengenali kondisi objek pada berbagai lokasi di wilayah penelitian. Pengambilan sampel data training berbentuk area (poligon), sehingga jumlah sampel yang diambil berbeda dengan jumlah pikselnya, karena setiap sampel merepresentasikan sekumpulan piksel yang memiliki karakteristik tutupan lahannya. Dengan demikian, jumlah poligon yang digunakan sebagai sampel tidak sama dengan jumlah piksel yang diekstraksi pada proses pelatihan, karena setiap poligon terdiri atas sejumlah piksel sesuai dengan luas area yang didigitasi dan resolusi spasial citra Sentinel-2. Pendekatan ini memberikan representasi data yang lebih baik dibandingkan penggunaan sampel berbentuk titik, karena mampu menangkap variasi nilai spektral dalam satu kelas tutupan lahan.

Setelah proses pengambilan data *training sample* selesai dilakukan, tahapan berikutnya adalah penyusunan *script* klasifikasi *supervised* menggunakan algoritma *Random Forest*. Sebagaimana ditunjukkan pada Gambar 8, proses klasifikasi diawali dengan menambahkan hasil transformasi indeks spektral *Normalized Difference Vegetation Index* (NDVI), *Normalized Difference Water Index* (NDWI), dan *Normalized Difference Built-up Index* (NDBI) ke dalam citra Sentinel-2 menggunakan fungsi *addBands()*. Penambahan indeks spektral tersebut bertujuan memperkaya informasi spektral yang dimiliki setiap piksel sehingga model klasifikasi tidak hanya memanfaatkan band asli Sentinel-2,

tetapi juga karakteristik biofisik yang direpresentasikan oleh masing-masing indeks.

Pada tahap berikutnya seluruh data *training sample* dari empat kelas tutupan lahan, yaitu badan air, infrastruktur, pertanian, dan vegetasi, digabungkan menjadi satu *FeatureCollection* menggunakan fungsi *merge()*. Penggabungan tersebut menghasilkan satu kumpulan data latih yang akan digunakan sebagai referensi dalam proses pembentukan model klasifikasi. Selanjutnya, menentukan variabel masukan (*input features*) berupa band-band Sentinel-2 yang digunakan dalam penelitian, yaitu B2 (Blue), B3 (Green), B4 (Red), B5, B6, B7 (Red Edge), B8 (Near Infrared), B8A (Narrow Near Infrared), B11 dan B12 (Short Wave Infrared), serta tiga indeks spektral NDVI, NDWI, dan NDBI. Kombinasi band multispektral dan indeks spektral tersebut diharapkan mampu meningkatkan kemampuan model dalam membedakan karakteristik spektral setiap kelas tutupan lahan.

```

2.2.2 Supervised
158 //***** Menambahkan Spectral Indices *****
159
160 var sentinel2_with_indices = sentinel2
161   .addBands(NDVI_1.rename('NDVI'))
162   .addBands(NDWI.rename('NDWI'))
163   .addBands(NDBI.rename('NDBI'));
164 print(sentinel2);
165 print(sentinel2_with_indices);
166
167 //***** Supervised Classification *****
168
169 // Menggabungkan Training Samples
170 var roi_ts = badan_air.merge(infrastruktur).merge(pertanian).merge(vegetasi);
171
172 // Band Sample
173 var bands = ['B2', 'B3', 'B4', 'B5', 'B6', 'B7', 'B8', 'B8A', 'B11', 'B12', 'ND
174
175 // Ekstraksi Nilai Band dengan Training Samples
176 var signature = sentinel2_with_indices.select(bands).sampleRegions({
177   collection: roi_ts,
178   properties: ['class'],
179   scale: 20 // Sesuaikan dengan Resolusi Spasial Citra Satelit
180 });
181
182 // Instalasi Algoritma Klasifikasi Supervised
183 var randomforest = ee.Classifier.smileRandomForest(100, 4).train({
184   features: signature,
185   classProperty: 'class',
186   inputProperties: bands
187 });

```

Gambar 8. Data Script Klasifikasi Supervised

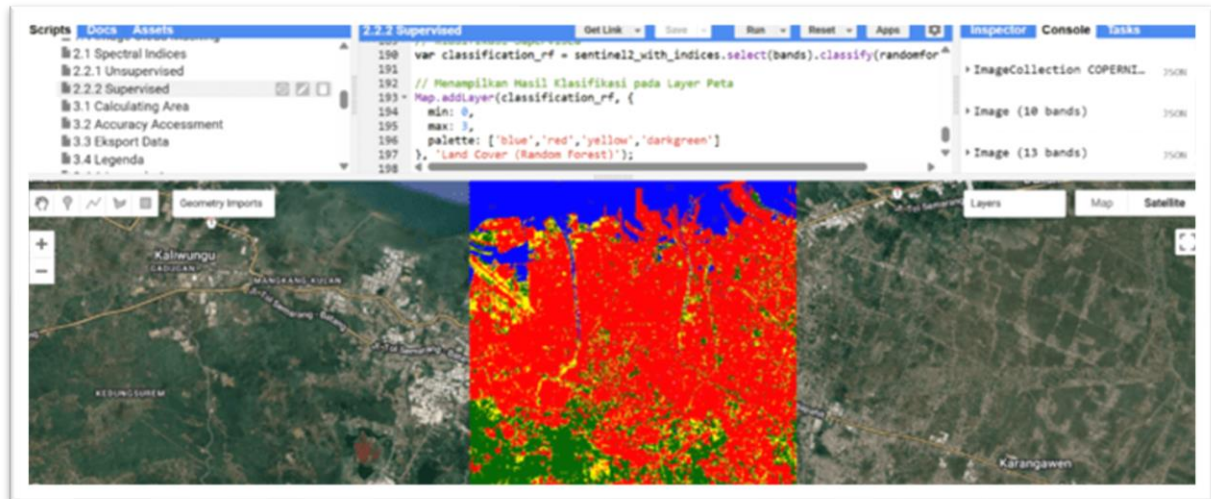
Proses berikutnya adalah ekstraksi nilai spektral dari setiap sampel menggunakan fungsi *sampleRegions()*. Pada tahap ini, setiap poligon data latih dikonversi menjadi sekumpulan piksel yang memiliki informasi nilai reflektansi pada seluruh band dan indeks spektral yang digunakan. Data hasil ekstraksi tersebut kemudian menjadi dataset pelatihan (*training dataset*) yang digunakan untuk membangun model Random Forest. Penggunaan data pelatihan dalam bentuk poligon memungkinkan model memperoleh variasi karakteristik spektral yang lebih representatif dibandingkan apabila hanya menggunakan sampel berbentuk titik, sehingga meningkatkan kemampuan model dalam mengenali setiap kelas tutupan lahan.

Model klasifikasi kemudian dibangun menggunakan fungsi *ee.Classifier.smileRandomForest()*. Pada penelitian ini, algoritma Random Forest diterapkan dengan 100 pohon keputusan (*decision trees*), sedangkan jumlah variabel yang dipilih secara acak pada setiap percabangan ditetapkan sebanyak empat variabel (*variables per split = 4*). Penentuan jumlah pohon yang relatif banyak bertujuan meningkatkan stabilitas model dan mengurangi variasi hasil prediksi, sedangkan pemilihan sebagian variabel pada setiap percabangan membantu menghasilkan pohon keputusan yang beragam sehingga mampu meningkatkan kemampuan generalisasi model terhadap data yang diklasifikasikan. Setelah proses pelatihan (*training*) selesai, model diterapkan pada seluruh citra Sentinel-2 untuk menghasilkan peta klasifikasi tutupan lahan.

Penggunaan kombinasi antara band multispektral Sentinel-2 dan transformasi indeks spektral memberikan kontribusi yang penting terhadap kualitas hasil klasifikasi. NDVI meningkatkan kemampuan model dalam mengenali area vegetasi melalui informasi tingkat kehijauan, NDWI memperjelas identifikasi badan air berdasarkan kandungan air pada permukaan, sedangkan NDBI memperkuat deteksi kawasan terbangun melalui karakteristik reflektansi pada saluran *Short Wave Infrared* (SWIR). Dengan memanfaatkan ketiga indeks tersebut secara bersamaan, separabilitas antar

kelas menjadi lebih baik sehingga potensi kesalahan klasifikasi akibat kemiripan nilai spektral antar objek dapat dikurangi.

Selain itu, implementasi algoritma Random Forest pada Google Earth Engine memberikan keuntungan dari sisi efisiensi komputasi. Seluruh proses ekstraksi data, pembentukan model, hingga klasifikasi dilakukan pada infrastruktur *cloud computing* sehingga mampu mengolah data citra Sentinel-2 dalam waktu yang relatif singkat tanpa bergantung pada spesifikasi komputer pengguna. Pendekatan ini menjadikan proses klasifikasi lebih efisien, mudah direproduksi melalui *script*, serta



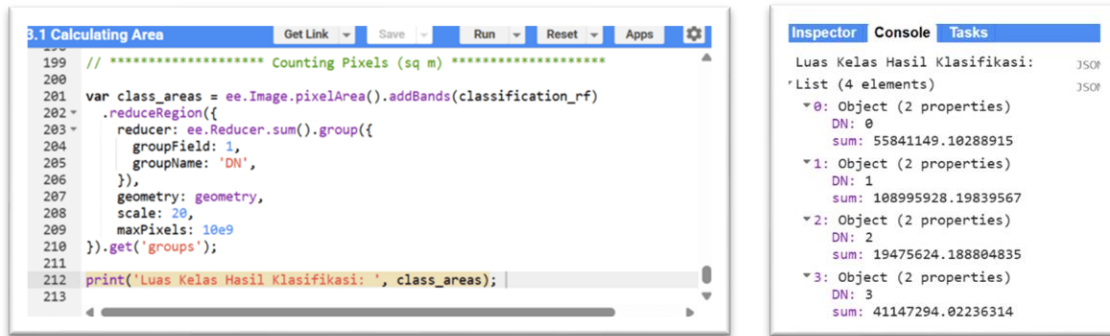
Gambar 9. Hasil Klasifikasi Tutupan Lahan *Random Forest*

Secara umum, hasil klasifikasi memperlihatkan bahwa penggunaan kombinasi band multispektral Sentinel-2 dengan transformasi indeks spektral NDVI, NDWI, dan NDBI mampu meningkatkan kemampuan model dalam membedakan karakteristik setiap kelas tutupan lahan. Penambahan ketiga indeks tersebut memberikan informasi biofisik yang tidak secara langsung diperoleh dari band asli Sentinel-2, sehingga meningkatkan separabilitas antar kelas yang memiliki karakteristik spektral yang berdekatan. Kondisi ini sangat penting terutama dalam membedakan kawasan terbangun dengan lahan terbuka, maupun vegetasi dengan lahan pertanian yang sering memiliki nilai reflektansi yang hampir serupa.

Walaupun secara visual hasil klasifikasi menunjukkan pemisahan kelas yang cukup baik, masih terlihat beberapa piksel yang tersebar secara tidak beraturan. Fenomena ini merupakan karakteristik umum pada klasifikasi berbasis piksel (*pixel-based classification*), terutama pada wilayah yang memiliki heterogenitas tutupan lahan tinggi dan batas antarobjek yang tidak tegas. Oleh karena itu, diperlukan uji akurasi menggunakan *Confusion Matrix*, *Overall Accuracy*, dan *Kappa Coefficient* untuk mengevaluasi tingkat ketelitian hasil klasifikasi secara kuantitatif. Evaluasi tersebut menjadi dasar dalam menentukan tingkat keandalan algoritma Random Forest dalam menghasilkan informasi tutupan lahan pada wilayah penelitian.

Uji akurasi hasil klasifikasi dilakukan dengan tahapan awal yaitu *Calculating Area*, yang merupakan bagian dari proses evaluasi hasil klasifikasi sebelum uji akurasi dilaksanakan. Perhitungan luas area bertujuan untuk mengetahui distribusi spasial setiap kelas tutupan lahan yang dihasilkan dari klasifikasi Random Forest, sekaligus menjadi dasar dalam menentukan jumlah sampel uji akurasi secara proporsional. Dengan mengetahui luas masing-masing kelas, proses pengambilan sampel validasi dapat dilakukan secara representatif sehingga setiap kelas tutupan lahan memiliki peluang yang seimbang untuk dievaluasi sesuai proporsi luas wilayahnya.

Sebagaimana ditunjukkan pada Gambar 10, proses perhitungan luas dilakukan menggunakan *script* pada *Google Earth Engine* dengan menghitung jumlah piksel pada setiap kelas hasil klasifikasi kemudian mengonversinya ke dalam satuan luas (m^2) berdasarkan resolusi spasial citra Sentinel-2. Setiap kelas hasil klasifikasi direpresentasikan dalam bentuk identitas kelas (class ID), yaitu ID 0 untuk kelas badan air, ID 1 untuk kelas infrastruktur, ID 2 untuk kelas pertanian, dan ID 3 untuk kelas vegetasi. Penggunaan identitas numerik tersebut memudahkan proses komputasi pada *Google Earth Engine* dalam melakukan perhitungan statistik setiap kelas tutupan lahan.



Gambar 10. Data Script dan Hasil Perhitungan Luas Area

Hasil perhitungan luas area memberikan informasi mengenai sebaran masing-masing kelas pada wilayah penelitian. Informasi ini tidak hanya digunakan sebagai data kuantitatif mengenai kondisi tutupan lahan, tetapi juga menjadi dasar dalam penyusunan desain pengambilan sampel uji akurasi. Kelas dengan luas yang lebih besar akan memiliki jumlah titik validasi yang lebih banyak dibandingkan kelas dengan luas yang lebih kecil, sehingga proses evaluasi dapat mencerminkan kondisi sebenarnya di lapangan. Pendekatan tersebut bertujuan mengurangi potensi bias dalam penilaian akurasi yang dapat terjadi apabila seluruh kelas diberikan jumlah sampel yang sama tanpa mempertimbangkan proporsi luas wilayahnya.

Tahap *Accuracy Assessment* dilakukan untuk mengevaluasi tingkat ketelitian hasil klasifikasi tutupan lahan yang diperoleh dari algoritma Random Forest. Evaluasi ini bertujuan mengetahui sejauh mana hasil klasifikasi mampu merepresentasikan kondisi tutupan lahan yang sebenarnya. Uji akurasi merupakan tahapan yang sangat penting dalam penelitian klasifikasi citra penginderaan jauh karena hasil klasifikasi yang baik tidak hanya ditunjukkan oleh tampilan visual peta, tetapi juga harus didukung oleh nilai akurasi yang tinggi berdasarkan data referensi.

Proses uji akurasi diawali dengan menentukan titik sampel validasi (*validation samples*) yang berbeda lokasi dengan data training yang digunakan pada proses pembentukan model Random Forest. Pemisahan antara data pelatihan dan data validasi bertujuan menghindari bias evaluasi (*overfitting*), sehingga nilai akurasi yang diperoleh benar-benar menggambarkan kemampuan model dalam mengklasifikasikan data yang belum pernah digunakan pada proses pelatihan.

Jumlah titik sampel uji akurasi ditentukan menggunakan metode *Stratified Random Sampling* yang mengacu pada rekomendasi Olofsson et al. (2014). Metode ini membagi sampel berdasarkan kelas tutupan lahan, kemudian menentukan jumlah titik pada setiap kelas secara proporsional sesuai luas area hasil klasifikasi. Pendekatan ini banyak digunakan dalam evaluasi akurasi klasifikasi citra satelit karena mampu menghasilkan estimasi akurasi yang lebih representatif dibandingkan pengambilan sampel secara acak sederhana.

Sebagaimana disajikan pada Gambar 11, hasil perhitungan menunjukkan bahwa jumlah titik uji akurasi yang digunakan terdiri atas 217 titik untuk kelas badan air, 308 titik untuk kelas infrastruktur, 159 titik untuk kelas pertanian, dan 187 titik untuk kelas vegetasi. Perbedaan jumlah titik pada setiap kelas menunjukkan bahwa distribusi sampel disesuaikan dengan proporsi luas masing-masing kelas tutupan lahan, sehingga seluruh kelas memperoleh representasi yang memadai dalam proses evaluasi.

Id	Land Cover Class	sq m	Pixels	Wi	Si	Wi * Si	Stratified Samples
0	badan_air	56247891.98	140619.73	0.249481	0.3	0.075	
1	infrastruktur	103063782.6	257659.4565	0.457127	0.2	0.091	
2	pertanian	25790796.34	64476.99086	0.114392	0.5	0.057	(Olofsson, et al., 2014)
3	vegetasi	40357524.59	100893.8115	0.179001	0.4	0.072	Wi proporsi wilayah kelas i yang dipetakan
	Total		563649.9888	1		0.295	Si standar deviasi strata i
	N					870.638	So standar deviasi So=0.01

Id	Land Cover Class	Wi * N	Weighted	Equal	Mean
0	badan_air		217	218	217
1	infrastruktur		398	218	308
2	pertanian		100	218	159
3	vegetasi		156	218	187
	Total		871		871

Gambar 11. Penentuan Jumlah Titik Sampel Uji Akurasi

Selanjutnya, titik-titik validasi tersebut didistribusikan secara acak pada masing-masing kelas tutupan lahan sebagaimana diperlihatkan pada Gambar 12. Penyebaran titik secara acak bertujuan agar sampel mampu mewakili variasi karakteristik objek pada seluruh wilayah penelitian serta meminimalkan subjektivitas dalam pemilihan lokasi sampel. Setiap titik validasi kemudian dibandingkan dengan kondisi tutupan lahan berdasarkan interpretasi citra resolusi tinggi maupun data referensi yang digunakan dalam penelitian.

3.2 Accuracy Accessme...
Get Link
Save
Run
Reset
Apps

Imports (9 entries)

```

var geometry: Polygon, 4 vertices
var badan_air: FeatureCollection (63 elements)
var infrastruktur: FeatureCollection (100 elements)
var pertanian: FeatureCollection (86 elements)
var vegetasi: FeatureCollection (57 elements)
var aa_badan_air: FeatureCollection (217 elements)
var aa_infrastruktur: FeatureCollection (308 elements)
var aa_pertanian: FeatureCollection (159 elements)
var aa_vegetasi: FeatureCollection (187 elements)

```

☒ Geometry Imports

☐ geometry (1 poly)

☒ badan_air (63 polys)

☒ infrastruktur (100 polys)

☒ pertanian (86 polys)

☒ vegetasi (57 polys)

☒ aa_badan_air (217 pts)

☒ aa_infrastruktur (308 pts)

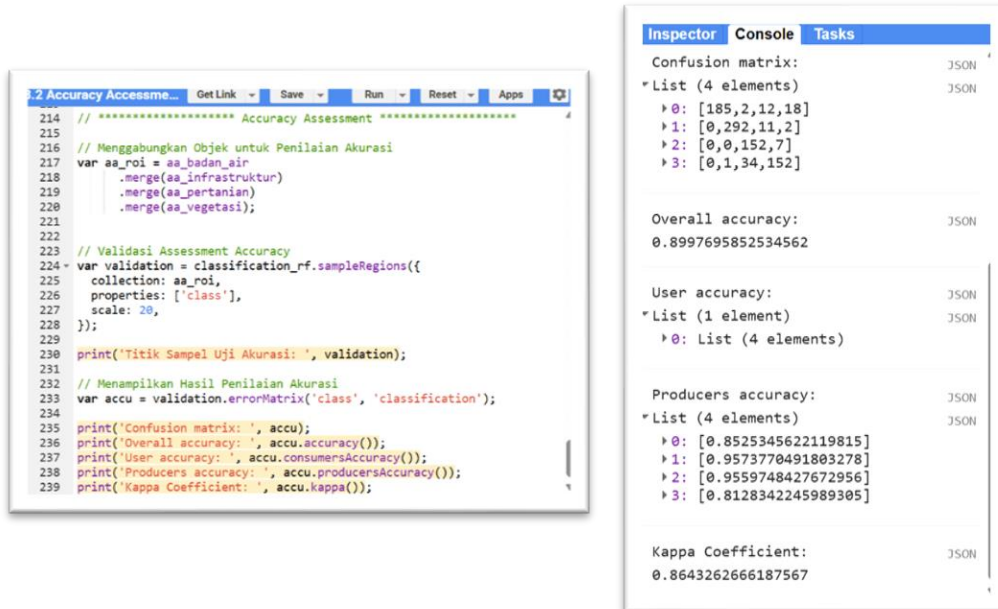
☒ aa_pertanian (159 pts)

☒ aa_vegetasi (187 pts)

+ new layer

Gambar 12. Pengambilan Titik Sampel Uji Akurasi

Setelah seluruh titik validasi diperoleh, dilakukan proses running script pada Google Earth Engine untuk menghasilkan parameter uji akurasi. Proses ini menghasilkan Confusion Matrix, Overall Accuracy, dan Kappa Coefficient sebagaimana ditampilkan pada Gambar 13. Confusion Matrix digunakan untuk membandingkan kelas hasil klasifikasi dengan kelas referensi pada setiap titik validasi, sehingga dapat diketahui jumlah piksel yang terklasifikasi benar maupun yang mengalami kesalahan klasifikasi. Selain menghasilkan jumlah klasifikasi yang benar dan salah pada setiap kelas, *Confusion Matrix* juga menjadi dasar dalam menghitung berbagai parameter akurasi lainnya.



Gambar 13. Data Script dan Hasil Accuracy Assesment

Parameter *Overall Accuracy* (OA) menunjukkan tingkat ketepatan klasifikasi secara keseluruhan, yaitu persentase jumlah piksel yang berhasil diklasifikasikan dengan benar terhadap seluruh sampel uji. Sementara itu, *Kappa Coefficient* (κ) digunakan untuk mengukur tingkat kesesuaian antara hasil klasifikasi dan data referensi dengan mempertimbangkan kemungkinan kesesuaian yang terjadi secara kebetulan (*chance agreement*). Oleh karena itu, nilai Kappa memberikan ukuran reliabilitas klasifikasi yang lebih komprehensif dibandingkan hanya menggunakan *Overall Accuracy*.

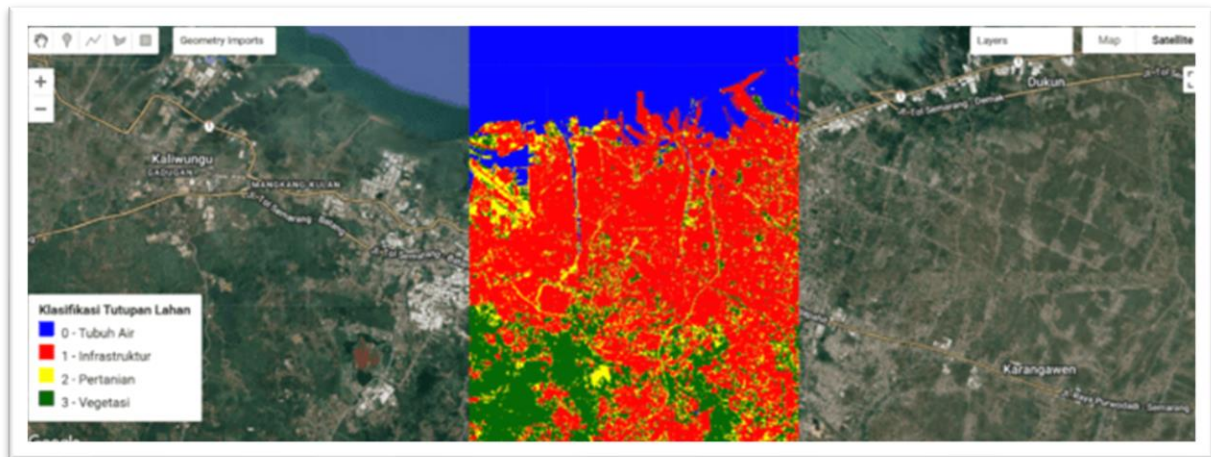
Hasil evaluasi menunjukkan bahwa klasifikasi tutupan lahan menggunakan algoritma Random Forest menghasilkan nilai *Overall Accuracy* sebesar 0,89 (89%) dan Kappa Coefficient sebesar 0,86. Nilai *Overall Accuracy* sebesar 89% menunjukkan bahwa sebanyak 89% dari seluruh sampel uji berhasil diklasifikasikan sesuai dengan kelas tutupan lahan yang sebenarnya. Menurut pedoman evaluasi klasifikasi penginderaan jauh, nilai *Overall Accuracy* di atas 85% menunjukkan bahwa hasil klasifikasi memiliki tingkat ketelitian yang baik dan layak digunakan sebagai sumber informasi spasial.

Sementara itu, nilai *Kappa Coefficient* sebesar 0,86 menunjukkan tingkat kesepakatan yang sangat tinggi antara hasil klasifikasi dan data referensi. Nilai tersebut mengindikasikan bahwa model *Random Forest* mampu mengidentifikasi karakteristik spektral setiap kelas tutupan lahan secara konsisten serta memiliki tingkat kesalahan klasifikasi yang relatif rendah. Tingginya nilai Kappa juga menunjukkan bahwa penggunaan kombinasi band multispektral Sentinel-2 dengan transformasi indeks spektral NDVI, NDWI, dan NDBI memberikan kontribusi positif dalam meningkatkan kemampuan model membedakan kelas badan air, infrastruktur, pertanian, dan vegetasi.

Confusion Matrix		Data				Row Total	UA (%)	CE (%)
		badan air	infrastruktur	pertanian	vegetasi			
	badan air	185	2	12	18	217	85	15
	infrastruktur	0	292	11	2	305	96	4
	pertanian	0	0	152	7	159	96	4
	vegetasi	0	1	34	152	187	81	19
Accuracy								
Column Totals		185	295	209	179	868		
PA (%)		100	99	73	85	OA (%)	90	
OE (%)		0	1	27	15	Kappa	0.864326267	

Gambar 14. Hasil Confusion Matrix

Mengacu pada hasil klasifikasi citra Sentinel-2, wilayah kajian di sebagian Kota Semarang yang tersaji pada Gambar 15, hasil klasifikasi didominasi oleh kelas tutupan lahan infrastruktur atau lahan terbangun (permukiman, jalan, kawasan industri, dan fasilitas umum). Sebaran infrastruktur terutama terkonsentrasi di wilayah tengah hingga utara, mengikuti pola perkembangan kota pesisir yang cenderung mengarah ke kawasan industri, pelabuhan, serta jalur transportasi utama.



Gambar 15. Hasil Klasifikasi Multispektral Tutupan Lahan di Sebagian Kota Semarang

Hasil klasifikasi tutupan lahan menunjukkan bahwa kawasan infrastruktur atau lahan terbangun mendominasi sebagian wilayah kajian di Kota Semarang. Informasi spasial tersebut memberikan gambaran mengenai distribusi kawasan terbangun, vegetasi, pertanian, dan badan air yang dapat dimanfaatkan sebagai dasar dalam penyusunan kebijakan penataan ruang. Dalam rangka mewujudkan pembangunan yang berkelanjutan, arah perencanaan wilayah perlu mempertimbangkan keseimbangan antara pengembangan kawasan terbangun dan pelestarian fungsi lingkungan. Upaya yang dapat dilakukan antara lain melalui pengendalian pemanfaatan ruang untuk mencegah konversi lahan produktif dan ruang terbuka hijau yang tidak sesuai dengan rencana tata ruang, peningkatan kualitas serta kuantitas ruang terbuka hijau sebagai kawasan resapan air dan penyangga ekologis, serta pengelolaan kawasan pesisir secara terpadu mengingat wilayah utara Kota Semarang memiliki karakteristik pesisir yang rentan terhadap banjir rob, abrasi, dan kenaikan muka air laut. Selain itu, informasi hasil klasifikasi tutupan lahan dapat digunakan sebagai salah satu data dasar dalam penyusunan rencana tata ruang berbasis mitigasi bencana, khususnya untuk mengidentifikasi kawasan yang memiliki tingkat kerentanan terhadap banjir, genangan, maupun degradasi lingkungan. Dengan demikian, peta hasil klasifikasi tutupan lahan yang dihasilkan melalui pengolahan citra Sentinel-2 berbasis Google Earth Engine tidak hanya memberikan informasi mengenai kondisi tutupan lahan eksisting, tetapi juga menjadi sumber data spasial yang penting dalam mendukung pengambilan keputusan, pengelolaan lingkungan, dan perencanaan pembangunan wilayah yang lebih efektif, adaptif, dan berkelanjutan.

KESIMPULAN

Citra Sentinel-2 yang digunakan melalui pemanfaatan band multispektral yang dikombinasikan dengan transformasi indeks spektral NDVI, NDWI, dan NDBI, berhasil membedakan tutupan lahan utama yang terdiri dari kelas badan air, infrastruktur, pertanian, dan vegetasi. Keberadaan platform *Google Earth Engine* (GEE), mampu mengintegrasikan proses pengambilan data Sentinel-2, prapengolahan citra, perhitungan indeks spektral, klasifikasi Random Forest, hingga evaluasi akurasi dalam satu platform berbasis cloud computing sehingga proses analisis menjadi lebih cepat dan efisien. Nilai Overall Accuracy sebesar 89% dan Kappa Coefficient sebesar 0,86 menunjukkan bahwa model Random Forest memiliki tingkat ketelitian dan konsistensi yang tinggi dalam mengklasifikasikan tutupan lahan. Hasil dari klasifikasi multispectral melalui platform GEE ini dapat digunakan untuk mendukung perencanaan tata ruang dan pembangunan berkelanjutan di Kota Semarang, khususnya dalam pengendalian pemanfaatan ruang, penguatan ruang terbuka hijau, serta pengelolaan kawasan rawan banjir rob di wilayah pesisir.

DAFTAR PUSTAKA

- Campbell, J. B., & Wynne, R. H. 2011. *Introduction To Remote Sensing*. Guilford Press.
- Gorelick, Noel, Matt Hancher, Mike Dixon, Simon Ilyushchenko, David Thau, and Rebecca Moore. 2017. "Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone." *Remote Sensing of Environment* 202: 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Kaplan, G., & Avdan, U. 2017. Object-based water body extraction model using Sentinel-2 satellite imagery. *European Journal of Remote Sensing*, 50(1), 137-143.
- Lillesand, and Kiefer. 1979. *Remote Sensing and Image Interpretation*. New York: John Willey and sons. 316p
- Lindgren, D.T. 1985. *Land Use Planning and Remote Sensing*. Martinus Nijhoff Publishers, Dordrecht.
- McFeeters, S. K. (1996). *The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features*. *International Journal of Remote Sensing*, 17(7), 1425–1432.
- Noi, P. T., & Kappas, M. (2018). *Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery*. *Sensors*, 18(1), 18.
- Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V. R., Murayama, Y., & Ranagalage, M. (2020). *Sentinel-2 Data for Land Cover/Use Mapping: A Review*. *Remote Sensing*, 12(14), 2291.
- Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). *Monitoring vegetation systems in the Great Plains with ERTS*. NASA Special Publication 351.
- Sipondak, R. (2026). *Evaluasi perubahan tutupan lahan menggunakan indeks NDVI dan NDBI pada kawasan perkotaan Bacan*. *Jurnal Penelitian Nusantara*, 2(2), 138–148.
- Sutanto. 1994. *Penginderaan Jauh jilid 1*. Yogyakarta: Gadjah Mada University Press.
- Tucker, C. J. (1979). *Red and photographic infrared linear combinations for monitoring vegetation*. *Remote Sensing of Environment*, 8(2), 127–150.
- Xu, Rudong & Liu, Jin & Xu, Jianhui. (2018). Extraction of High-Precision Urban Impervious Surfaces from Sentinel-2 Multispectral Imagery via Modified Linear Spectral Mixture Analysis. *Sensors*. 18. 2873. 10.3390/s18092873.
- Zha, Y., Gao, J., & Ni, S. (2003). *Use of Normalized Difference Built-up Index in Automatically Mapping Urban Areas from TM Imagery*. *International Journal of Remote Sensing*, 24(3), 583–594.
- European Union/Copernicus/Google. (2021). Sentinel-2 MSI: MultiSpectral Instrument, Level-2A [Dataset]. Google Earth Engine Data Catalog. https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR_HARMONIZED.