

# **DATA MINING**

UTS Mini Proyek : Clustering Citra Udara



**Disusun Oleh :**

Giasat Ahmad	(140810220018)
Devalco Aghazzan Muslion	(140810220079)
Angga Prasetyo	(140810220086)

PROGRAM STUDI S-1 TEKNIK INFORMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS PADJADJARAN  
JATINANGOR  
2024

# DAFTAR ISI

<b>DAFTAR ISI</b>	<b>1</b>
<b>BAB I</b>	<b>2</b>
<b>PENDAHULUAN</b>	<b>2</b>
1.1 LATAR BELAKANG	2
1.2 TUJUAN	3
1.3 MANFAAT	3
<b>BAB II</b>	<b>4</b>
<b>METODOLOGI</b>	<b>4</b>
2.1 METODOLOGI PENELITIAN	4
2.2 DATA INPUT	4
2.3 PREPROCESSING DATA	4
2.4 CLUSTERING	5
<b>BAB III</b>	<b>7</b>
<b>HASIL DAN PEMBAHASAN</b>	<b>7</b>
3.1 DATA TRAINING	7
3.2 DATA PROCESSING	8
3.3 DEPLOYMENT	10
<b>BAB IV</b>	<b>11</b>
<b>KESIMPULAN DAN SARAN</b>	<b>11</b>
4.1 KESIMPULAN	11
4.2 SARAN	11
<b>LAMPIRAN</b>	<b>12</b>
<b>SCREENSHOT IMPLEMENTASI APLIKASI</b>	<b>12</b>
<b>DAFTAR PUSTAKA</b>	<b>15</b>

# **BAB I**

## **PENDAHULUAN**

### **1.1 LATAR BELAKANG**

Perencanaan tata ruang menjadi semakin penting seiring dengan pertumbuhan populasi dan urbanisasi yang pesat di banyak kota besar. Urbanisasi ini seringkali menimbulkan tantangan bagi pemerintah daerah dalam menjaga keseimbangan antara pembangunan perumahan dan pelestarian lingkungan. Salah satu elemen yang krusial dalam perencanaan ini adalah keberadaan ruang terbuka hijau, yang diwajibkan oleh Undang-Undang Nomor 26 Tahun 2007 untuk mencakup minimal 30% dari luas suatu kota. Keberadaan ruang terbuka hijau tidak hanya penting untuk menjaga keseimbangan ekosistem, tetapi juga sebagai tempat aktivitas masyarakat yang sehat dan berkelanjutan.

Namun, meningkatnya kebutuhan akan lahan untuk hunian seringkali berdampak pada berkurangnya ruang terbuka hijau, yang pada akhirnya mengganggu keseimbangan lingkungan. Dalam konteks ini, penggunaan teknologi seperti citra satelit dapat membantu dalam melakukan analisis tata ruang yang lebih akurat, untuk memetakan area yang sesuai untuk pembangunan hunian tanpa merusak lingkungan.

Proyek ini menggunakan metode K-Means Clustering, yang merupakan salah satu algoritma pengelompokan data, untuk menganalisis citra satelit dalam mengidentifikasi penggunaan lahan dan menentukan kelayakan hunian suatu area. Metode ini memungkinkan pemisahan area berdasarkan karakteristik visual seperti warna, yang dapat dihubungkan dengan kategori lahan seperti bangunan, jalan, dan ruang terbuka hijau. Hasil analisis ini akan berguna bagi pengambil kebijakan dalam menentukan area hunian yang layak, sambil tetap mempertahankan proporsi ruang hijau yang diamanatkan oleh undang-undang.

## **1.2 TUJUAN**

Tujuan dari proyek ini adalah untuk mengembangkan aplikasi sederhana yang menggunakan analisis citra satelit dengan metode K-Means Clustering guna mengidentifikasi penggunaan lahan dan memberikan rekomendasi kelayakan hunian berdasarkan distribusi lahan dan keberadaan ruang terbuka hijau.

## **1.3 MANFAAT**

Dalam pembuatan aplikasi Analisis Citra Satelit Menggunakan K-Means Clustering untuk Menentukan Kelayakan Hunian Suatu Tempat , beberapa manfaat yang dapat diambil adalah sebagai berikut :

1. Kesadaran Lingkungan: Meningkatkan pemahaman masyarakat mengenai pentingnya keberadaan ruang terbuka hijau.
2. Pengambilan Keputusan: Memberikan data analitis yang mendukung pemerintah dalam perencanaan tata ruang.
3. Pemanfaatan Teknologi: Menunjukkan aplikasi teknologi penginderaan jauh dalam perencanaan tata ruang yang berkelanjutan.
4. Konservasi Lahan: Membantu melindungi ruang terbuka hijau dari konversi yang berlebihan untuk keperluan hunian.
5. Kualitas Hidup: Memfasilitasi penyediaan lingkungan hidup yang lebih sehat dan seimbang melalui tata ruang yang baik.

## **BAB II**

### **METODOLOGI**

#### **2.1 METODOLOGI PENELITIAN**

Metodologi penelitian ini terdiri dari beberapa tahapan yang saling berkaitan untuk mencapai tujuan analisis citra satelit menggunakan K-Means Clustering dalam menentukan kelayakan hunian suatu tempat.

Pengumpulan data dilakukan melalui beberapa langkah penting. Pertama, citra satelit diperoleh dari penyedia data seperti Google Earth, Sentinel, atau Landsat. Pemilihan citra harus memperhatikan resolusi dan waktu pengambilan yang relevan dengan analisis untuk memastikan akurasi hasil. Selain itu, data pendukung juga dikumpulkan, meliputi informasi demografi, penggunaan lahan, dan data lingkungan lainnya yang relevan. Pengumpulan data ini bertujuan untuk memperkuat analisis dan memberikan konteks yang lebih luas dalam penentuan kelayakan hunian.

#### **2.2 DATA INPUT**

Proyek ini menggunakan gambar citra satelit sebagai data input. Citra satelit tersebut akan dianalisis untuk menentukan kelayakan hunian berdasarkan karakteristik visual dan pola lahan yang terdeteksi.

#### **2.3 PREPROCESSING DATA**

Proses preprocessing adalah tahap persiapan data mentah agar siap untuk dianalisis. Langkah-langkah preprocessing meliputi:

1. Normalisasi Ukuran

Semua citra satelit diubah ukurannya menjadi 200 x 200 pixel. Hal ini dilakukan untuk memastikan konsistensi dalam analisis, sehingga perbandingan antar citra menjadi lebih akurat.

## 2. Ekstraksi Fitur Warna

Citra dianalisis berdasarkan dua ruang warna: RGB (Red-Green-Blue) dan HSV (Hue-Saturation-Value). Ekstraksi fitur warna ini bertujuan untuk menangkap perbedaan karakteristik visual antara area lahan yang berbeda, seperti ruang terbuka hijau, bangunan, dan jalan.

## 2.4 CLUSTERING

Clustering adalah teknik untuk mengelompokkan data berdasarkan kemiripan fitur. Dalam proyek ini, digunakan metode K-Means Clustering dengan prosedur sebagai berikut:

1. Pemilihan Jumlah Cluster (k): Menentukan jumlah cluster k, yaitu jumlah kategori lahan yang ingin diidentifikasi.
2. Inisialisasi Centroid: Memilih titik awal centroid secara acak untuk setiap cluster.
3. Perhitungan Jarak: Menghitung jarak antara setiap titik data (piksel citra) dengan centroid.
4. Pengelompokan Data: Menempatkan setiap titik data ke dalam cluster berdasarkan jarak terdekatnya dengan centroid.
5. Pembaruan Centroid: Menghitung ulang posisi centroid berdasarkan rata-rata posisi titik dalam setiap cluster.
6. Pengulangan Proses: Proses ini diulang hingga posisi centroid stabil dan tidak ada perubahan signifikan.

Algoritma K-Means kami pilih dalam pengembangan aplikasi ini karena efisiensinya dalam komputasi, sehingga memungkinkan analisis yang cepat dan akurat. Algoritma ini juga

memberikan hasil yang mudah diinterpretasikan dan sesuai untuk memisahkan area berdasarkan variasi warna yang terlihat pada citra satelit, seperti area hijau, bangunan, dan infrastruktur lainnya.

## BAB III

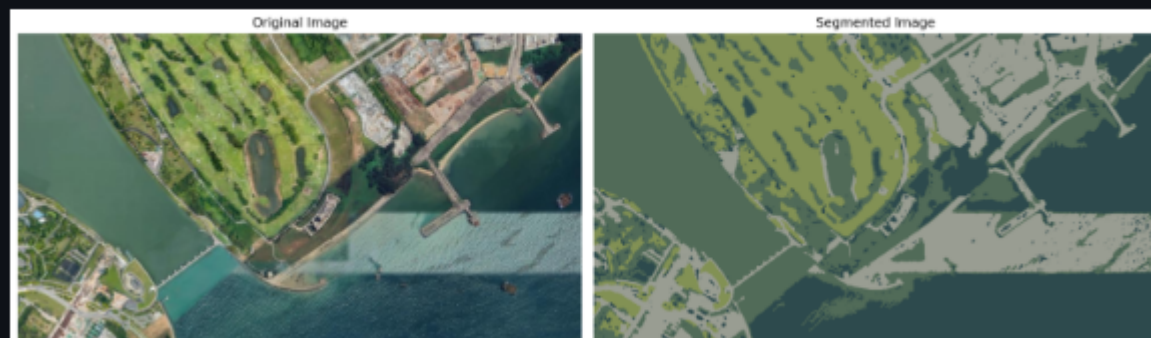
### HASIL DAN PEMBAHASAN

#### 3.1 DATA TRAINING

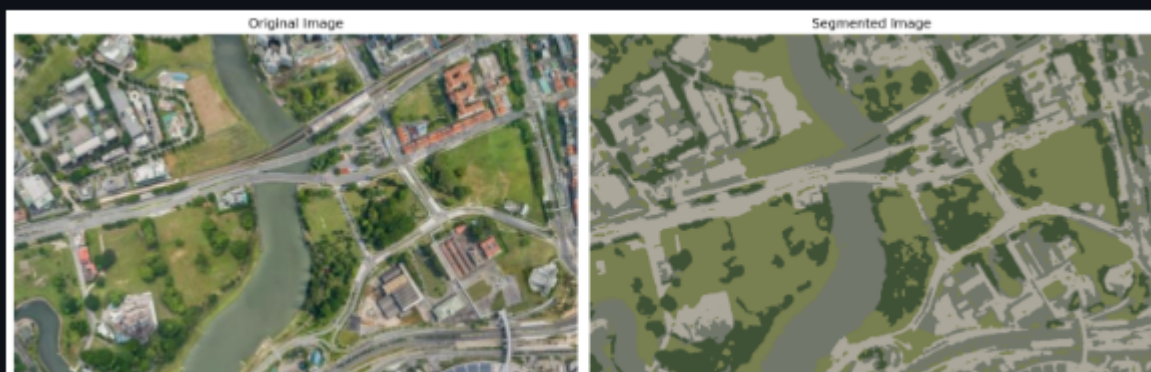
Processing training image 1/3...



Processing training image 2/3...



Processing training image 3/3...



Model training completed and saved.

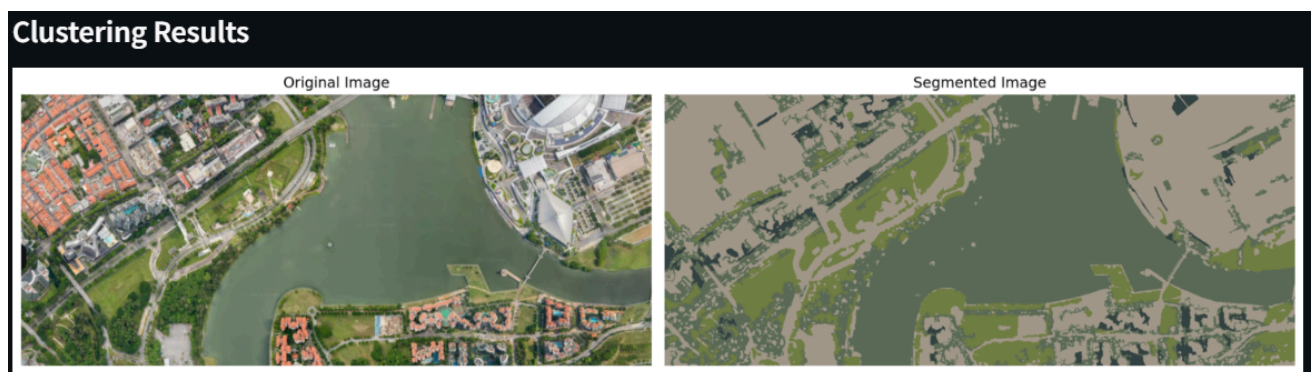


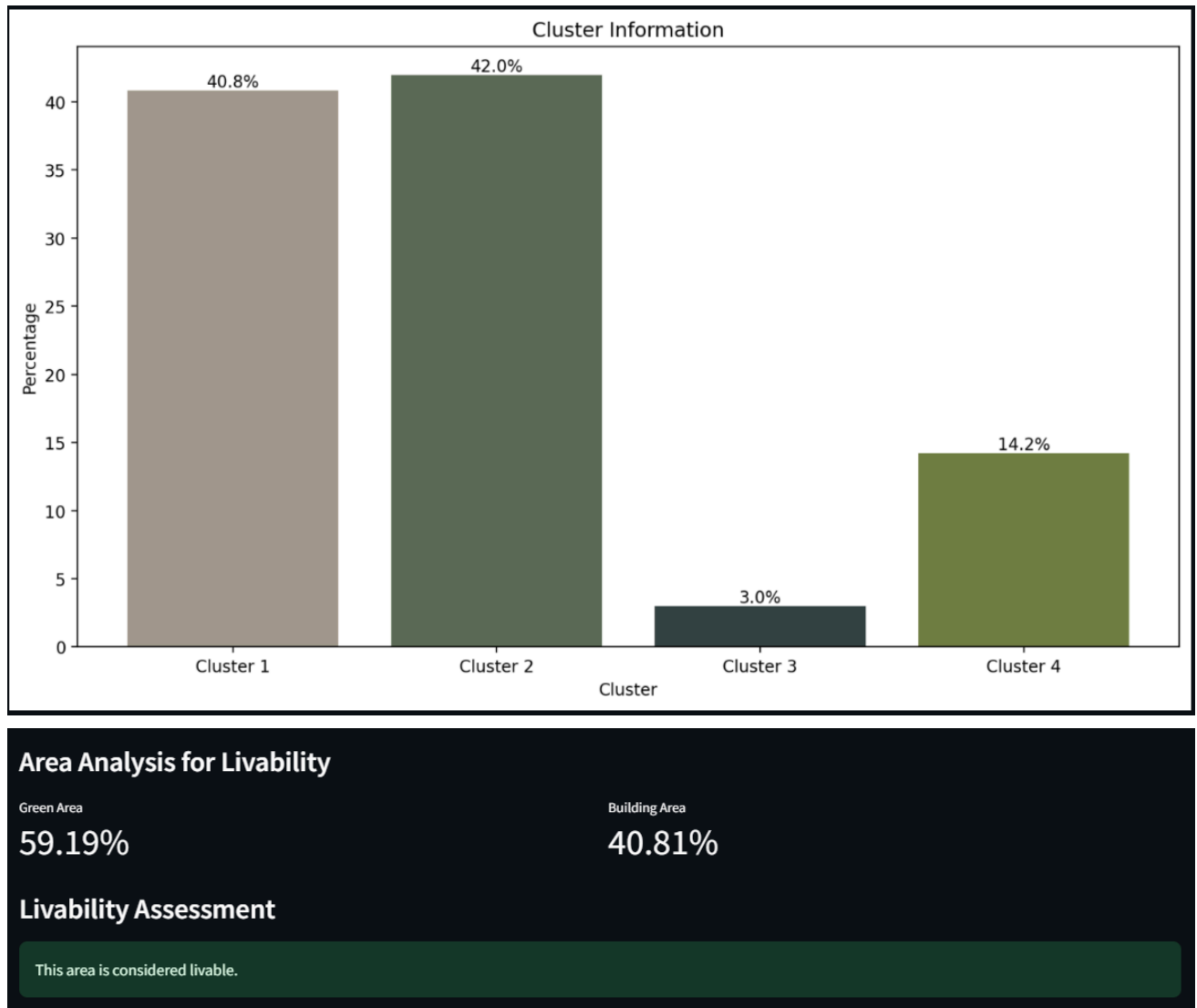
Setelah *User* mengunggah beberapa gambar sebagai data latih melalui Streamlit file uploader. Gambar-gambar ini kemudian dikonversi menjadi array numpy sehingga dapat diproses lebih lanjut dalam tahap ekstraksi fitur. Data latih kemudian disiapkan, seperti mengubah ukuran dan format gambar, Kemudian diekstrak fitur dari gambar yang telah dipreproses, seperti informasi piksel

Setelahnya, model K-Means diinisialisasi dengan jumlah cluster sesuai input pengguna dan dilatih dengan fitur tersebut sehingga menghasilkan citra tersegmentasi. Hasil segmentasi divisualisasikan dengan sebuah grafik yang menunjukkan citra asli dan citra tersegmentasi. Setelah semua data diproses, fitur dari seluruh gambar digabungkan menjadi satu array besar, yang menggabungkan fitur dari setiap gambar untuk membentuk satu set data.

Model K-Means kemudian dilatih ulang menggunakan semua fitur gabungan untuk memastikan model mampu mengenali pola dalam seluruh dataset. Model yang telah dilatih disimpan dalam file untuk digunakan kembali saat pengguna memproses gambar baru, sehingga mempercepat proses tanpa perlu melatih ulang.

### 3.2 DATA PROCESSING





Data proses disiapkan terlebih dahulu dengan cara yang sama dengan data latih, Pemrosesan data dimulai dengan mengonversi citra menjadi tipe data float32 dan menormalisasinya ke rentang 0-1. Langkah ini dilakukan untuk memastikan data terstandarisasi, yang memudahkan model dalam mengolahnya. Selanjutnya, citra diproses menggunakan Gaussian Blur untuk menghaluskan dan mengurangi noise, sehingga fokus analisis bisa pada pola-pola besar dan objek utama dalam gambar.

Setelah itu, citra dikonversi dari ruang warna RGB ke LAB, karena ruang warna LAB lebih akurat dalam menggambarkan kecerahan dan warna, mirip dengan cara mata manusia memandang. Setelah citra diproses, langkah berikutnya adalah ekstraksi fitur.

Proses ekstraksi fitur dimulai dengan mengekstraksi informasi warna dari setiap piksel dalam citra yang telah diubah ke ruang warna LAB. Setiap piksel diubah menjadi data berbentuk array 2D, dengan tiga komponen warna (L, A, B) yang memberikan informasi terkait intensitas dan nuansa warna. Selain itu, informasi spasial juga ditambahkan, yaitu koordinat x dan y dari setiap piksel, yang dinormalisasi agar sesuai dengan skala fitur warna. Dengan menggabungkan informasi warna dan lokasi ini, setiap piksel direpresentasikan dengan lima fitur (L, A, B, x, y).

Fitur yang sudah terkumpul kemudian digunakan oleh model K-Means untuk mengelompokkan piksel ke dalam segmen-segmen berdasarkan kesamaan warna dan lokasi. Area-area kecil yang tidak relevan kemudian dibersihkan untuk memastikan hanya segmen signifikan yang tersisa. Setelah itu, analisis dilakukan untuk mengidentifikasi area hijau (vegetasi) dan bangunan berdasarkan warna rata-rata segmen-segmen tersebut. Persentase area hijau dan bangunan dihitung untuk mengevaluasi karakteristik lingkungan dalam citra (Jika Yang dipilih di awal adalah “Clustering Kelayakan Huni Suatu Area” bukan “Clustering Biasa”)

### 3.3 DEPLOYMENT

Deployment aplikasi ini dilakukan dengan bantuan library Python Streamlit yang memudahkan proses pembuatan dan penyebaran aplikasi web secara cepat dan efisien. Aplikasi ini mengimplementasikan algoritma KMeans clustering secara manual tanpa menggunakan library eksternal untuk proses clustering-nya.

Aplikasi ini dirancang untuk melakukan segmentasi area kota berdasarkan citra udara atau satelit, dengan dua mode operasi: clustering biasa dan analisis kelayakan huni suatu area. Hasil dari deployment bisa dilihat pada link berikut ini [WEB Kelompok 8](#) Berikut adalah fitur-fitur utama yang terdapat pada aplikasi web:

- Pemilihan Mode Clustering

Pengguna dapat memilih antara dua mode: "Clustering Biasa" dan "Clustering Kelayakan Huni Suatu Area". Masing-masing mode menyediakan analisis yang berbeda sesuai kebutuhan pengguna.

- Pelatihan Model KMeans

Pengguna dapat mengunggah beberapa gambar untuk melatih model KMeans. Aplikasi akan memproses setiap gambar, mengekstrak fiturnya, dan melatih model berdasarkan fitur-fitur tersebut. Hasil segmentasi untuk setiap gambar latihan ditampilkan secara visual.

- Upload dan Proses Gambar

Setelah model dilatih, pengguna dapat mengunggah gambar baru untuk dianalisis menggunakan model yang telah dilatih.

- Visualisasi Hasil Clustering

Hasil segmentasi ditampilkan secara visual, membandingkan gambar asli dengan gambar yang telah disegmentasi. Informasi tambahan seperti persentase setiap cluster juga disajikan.

- Analisis Area (untuk mode Kelayakan Huni)

Dalam mode ini, aplikasi melakukan analisis area hijau dan area bangunan berdasarkan hasil clustering, menghitung persentase masing-masing area.

- Penilaian Kelayakan Huni (untuk mode Kelayakan Huni)

Berdasarkan analisis area, aplikasi memberikan penilaian apakah suatu area dianggap layak huni atau tidak, sesuai dengan kriteria yang telah ditentukan.

Aplikasi ini menggunakan implementasi KMeans manual yang memungkinkan pengguna untuk memahami proses clustering secara lebih mendalam tanpa bergantung pada library machine learning eksternal. Dengan fitur-fitur ini, aplikasi menyediakan alat yang fleksibel untuk analisis citra, baik untuk tujuan umum maupun untuk penilaian kelayakan huni area kota berdasarkan citra satelit atau udara.

## **BAB IV**

### **KESIMPULAN DAN SARAN**

#### **4.1 KESIMPULAN**

Proyek ini berhasil mengembangkan aplikasi sederhana untuk menganalisis citra satelit menggunakan metode K-Means Clustering. Hasil analisis memungkinkan identifikasi penggunaan lahan dan kelayakan hunian suatu tempat berdasarkan karakteristik visual dari citra satelit. Dengan metode ini, pemerintah dan perencana kota dapat lebih mudah menentukan area yang layak untuk hunian, sambil mempertahankan keberadaan ruang terbuka hijau sesuai dengan ketentuan yang berlaku.

Metode K-Means Clustering terbukti efektif dalam memisahkan area berdasarkan pola warna yang mewakili kategori lahan seperti bangunan, jalan, dan ruang terbuka hijau. Dengan hasil clustering ini, keputusan yang lebih bijak dapat diambil dalam mengelola tata ruang dan pembangunan kota secara berkelanjutan.

#### **4.2 SARAN**

Untuk pengembangan lebih lanjut, beberapa saran yang dapat kami sebagai pengembang aplikasi ini adalah sebagai berikut:

1. Penggunaan Data dengan Resolusi Tinggi: Meningkatkan akurasi analisis dengan menggunakan citra satelit yang memiliki resolusi lebih tinggi.
2. Integrasi Algoritma Lain: Kombinasi dengan metode pengolahan citra lain seperti Support Vector Machines (SVM) atau Random Forest untuk hasil yang lebih komprehensif.

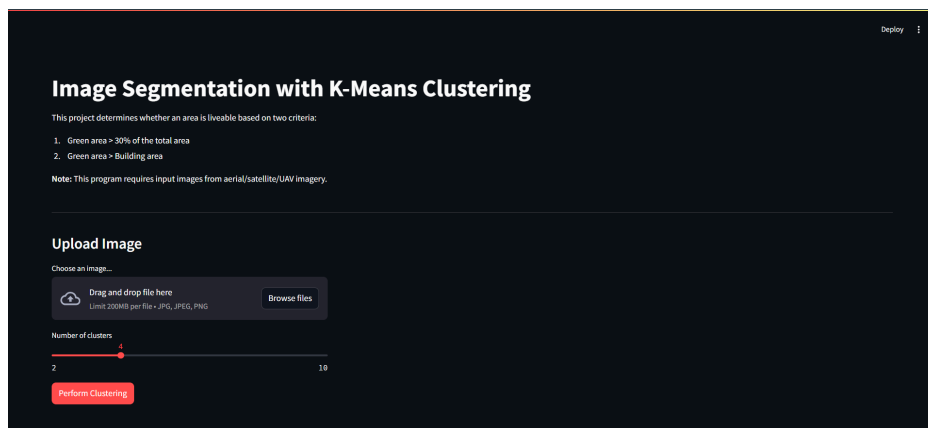
3. Analisis Temporal: Mengembangkan fitur untuk menganalisis perubahan penggunaan lahan dari waktu ke waktu dengan menggunakan data citra satelit dari periode yang berbeda.

Dengan kami berikan saran-saran ini, kami berharap kedepannya bagi para pembaca yang akan membuat aplikasi yang serupa, dapat menjadikan saran ini sebagai pertimbangan agar dapat membuat aplikasi yang lebih efektif dan bermanfaat dalam pengembangannya.

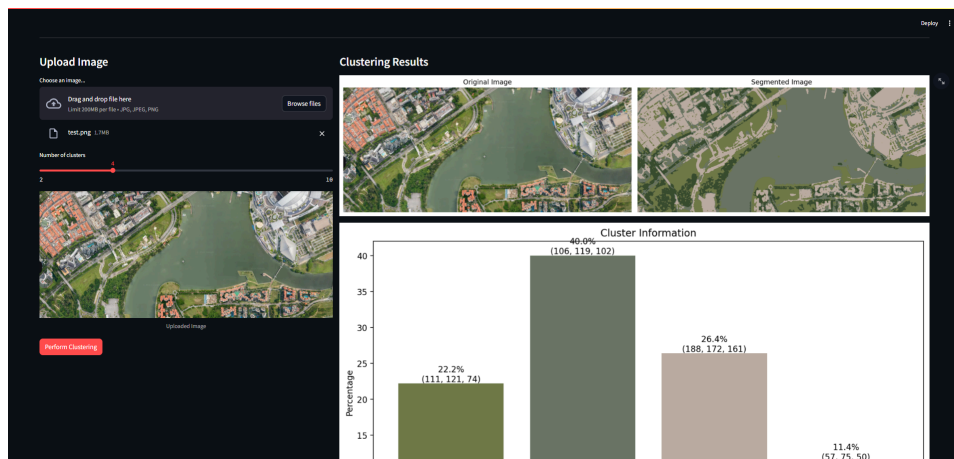
# LAMPIRAN

## SCREENSHOT IMPLEMENTASI APLIKASI

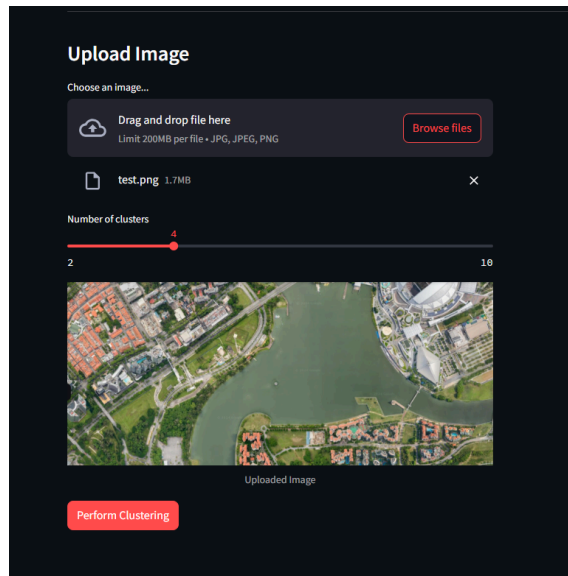
- Tampilan Halaman Utama



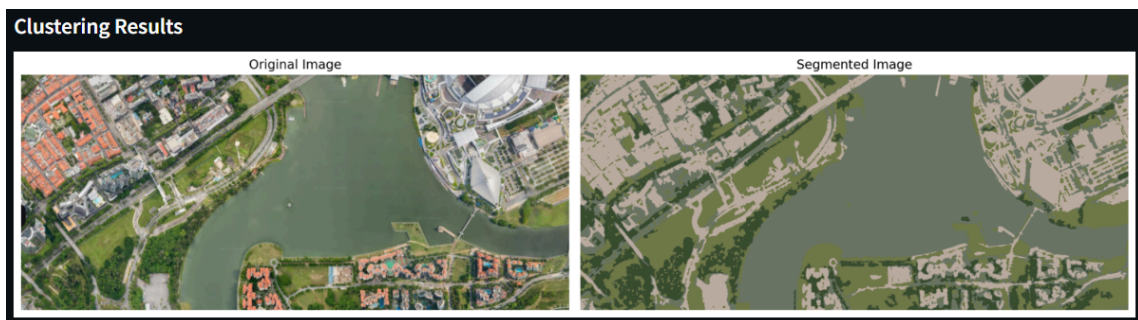
- Tampilan Halaman Utama Dengan Hasil



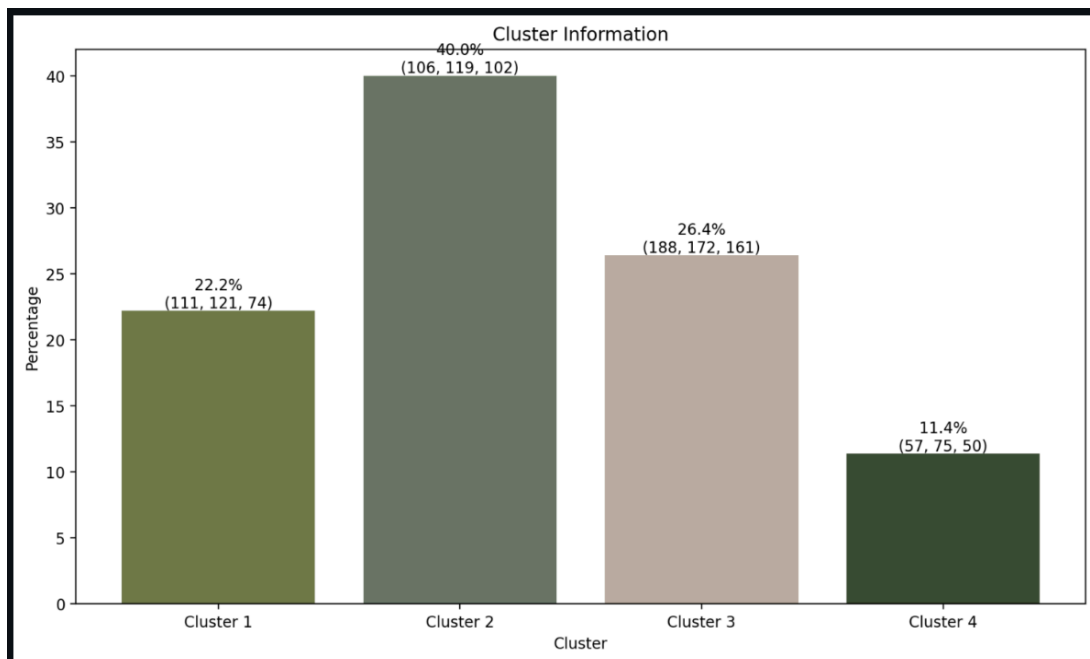
- Tampilan Component Input Gambar



- Tampilan Component Hasil Segmentasi Input Gambar

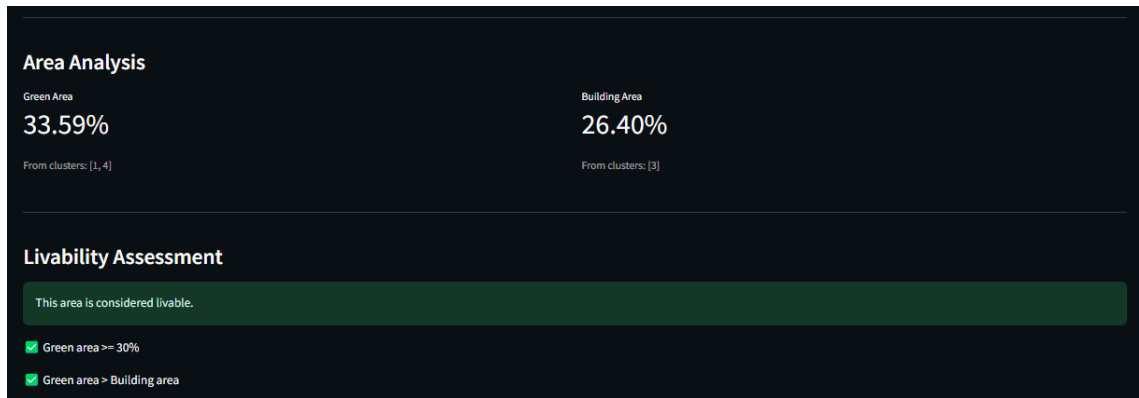


- Tampilan Component Diagram Persentase Cluster



- Tampilan Component Kesimpulan





## DAFTAR PUSTAKA

<https://peraturan.bpk.go.id/Details/39908/uu-no-26-tahun-2007>  
<https://socs.binus.ac.id/2017/03/09/clustering/>