

***CLUSTERING TINGKAT KERAWANAN DAERAH TANAH LONGSOR
MENGUNAKAN GRAY LEVEL CO-OCCURRENCE MATRIX DAN
PRINCIPAL COMPONENT ANALYSIS***

SKRIPSI

**Oleh :
SALSABILA RAMADANTI NURAINI
NIM. 210605110018**



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025**

***CLUSTERING TINGKAT KERAWANAN DAERAH TANAH LONGSOR
MENGUNAKAN GRAY LEVEL CO-OCCURRENCE MATRIX DAN
PRINCIPAL COMPONENT ANALYSIS***

SKRIPSI

Diajukan kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh :
SALSABILA RAMADANTI NURAINI
NIM. 210605110018

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2025**

HALAMAN PERSETUJUAN

***CLUSTERING TINGKAT KERAWANAN DAERAH TANAH LONGSOR
MENGUNAKAN GRAY LEVEL CO-OCCURRENCE MATRIX DAN
PRINCIPAL COMPONENT ANALYSIS***

SKRIPSI

**Oleh :
SALSABILA RAMADANTI NURAINI
NIM. 210605110018**

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 5 Juni 2025

Pembimbing I,



Dr. Agung Teguh Wibowo Almais, M.T.
NIP. 19860301 202321 1 016

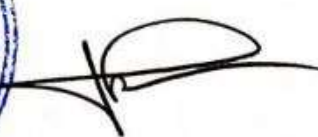
Pembimbing II,



Ashri Shabrina Afrah, M.T.
NIP. 19900430 202012 2 003

Mengetahui,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang




Dr. In Fachrul Kurniawan, M.MT., IPU
NIP. 19771020 200912 1 001

HALAMAN PENGESAHAN

CLUSTERING TINGKAT KERAWANAN DAERAH TANAH LONGSOR MENGUNAKAN *GRAY LEVEL CO-OCCURRENCE MATRIX* DAN *PRINCIPAL COMPONENT ANALYSIS*

SKRIPSI

Oleh :
SALSABILA RAMADANTI NURAINI
NIM. 210605110018

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal: 12 Juni 2025

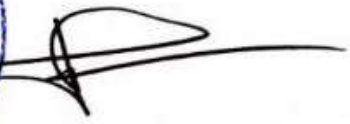
Susunan Dewan Penguji

Ketua Penguji	: <u>Dr. Totok Chamidy, M.Kom</u> NIP. 19691222 200604 1 001
Anggota Penguji I	: <u>Fajar Rohman Hariri, M.Kom</u> NIP. 19890515 201801 1 001
Anggota Penguji II	: <u>Dr. Agung Teguh Wibowo Almais, M.T</u> NIP. 19860301 202321 1 016
Anggota Penguji III	: <u>Ashri Shabrina Afrah, M.T</u> NIP. 19900430 202012 2 003

()
()
()
()

Mengetahui dan Mengesahkan,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang




Dr. Fachrul Kurniawan, M.MT., IPU
NIP. 19771020 200912 1 001

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Salsabila Ramadanti Nuraini
NIM : 210605110018
Fakultas / Program Studi : Sains dan Teknologi / Teknik Informatika
Judul Skripsi : *Clustering Tingkat Kerawanan Daerah Tanah Longsor Menggunakan Gray Level Co-Occurrence Matrix dan Principal Component Analysis*

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 11 Juni 2025
Yang membuat pernyataan,



Salsabila Ramadanti Nuraini
NIM. 210605110018

MOTTO

"Janganlah engkau mengira bahwa engkau dapat meraih ilmu hanya dengan bersantai-santai. Ilmu tidak akan memberikan sedikit pun dari dirinya kepadamu hingga engkau memberikan seluruh dirimu kepadanya. Jika engkau memberikan separuh perhatianmu, maka ia pun akan memberi separuh. Namun jika engkau bermalas-malasan, maka ia akan berpaling seluruhnya."

-Imam Al-Ghazali-

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur dan kerendahan hati kepada Allah SWT,
atas limpahan rahmat, karunia, serta kemudahan-Nya dalam setiap langkah,
sehingga skripsi ini dapat terselesaikan dengan baik sesuai harapan.
Segala ikhtiar dan perjuangan ini penulis persembahkan dengan tulus kepada:

Ayah tercinta, Nurjudi, dan Ibu tercinta, Nanin Yulianik,
yang senantiasa mengalirkan kasih sayang, doa-doa terbaik,
dukungan tanpa lelah, serta teladan keteguhan dan keikhlasan
dalam setiap perjuangan hidup. Terima kasih atas segala pengorbanan
yang tidak dapat terbalas oleh apapun di dunia ini.

Keluarga besar penulis,
yang telah memberikan doa, semangat, dan perhatian yang berarti,
yang meskipun dalam diam tetap menjadi bagian dari kekuatan
untuk terus melangkah dan tidak menyerah dalam menyelesaikan tugas ini.

Dan terakhir, untuk diri penulis sendiri,
Terima kasih telah bertahan sejauh ini.
Semoga skripsi ini menjadi awal dari jalan panjang
yang penuh manfaat dan keberkahan.

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Segala puji dan syukur senantiasa penulis panjatkan ke hadirat Allah subhanahu wa ta'ala atas berkat Rahmat, serta hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “*Clustering Tingkat Kerawanan Daerah Tanah Longsor Menggunakan Gray Level Co-Occurrence Matrix Dan Principal Component Analysis*” dengan baik dan lancar. Salawat serta salam tetap tercurahkan kepada Nabi Muhammad SAW, yang telah membawa kita dari zaman kebodohan menuju zaman kebenaran yakni Islam dan zaman yang penuh dengan ilmu pengetahuan sebagaimana yang di rasakan pada saat ini. Semoga kita semua mendapat syafaatnya di hari akhir kelak.

Penulis mengucapkan rasa terima kasih yang begitu besar kepada seluruh pihak yang memberikan dukungan dan motivasi kepada penulis sehingga dapat menyelesaikan skripsi ini. Ucapan terima kasih ini penulis disampaikan kepada:

1. Prof. Dr. M. Zainuddin, M.A., selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Prof. Dr. Hj. Sri Harini, M.Si., selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Ir. Fachrul Kurniawan, M.MT., IPU., selaku Ketua Program Studi Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Dr. Agung Teguh Wibowo Almais, M.T., selaku Dosen Pembimbing I yang dengan penuh kesabaran dan ketulusan hati memberikan bimbingan,

dan memberi semangat, arahan, masukan, serta membantu penulis dalam mengerjakan dan menyelesaikan skripsi ini.

5. Ashri Shabrina Afrah, M.T., selaku Dosen Pembimbing II yang membimbing dan memberikan arahan kepada penulis dalam proses penulisan skripsi dan penulisan integrasi islam.
6. Dr. Totok Chamidy, M.Kom, selaku Dosen Penguji I dan Bapak Fajar Rohman Hariri, M.Kom., selaku Dosen Penguji II yang telah menguji dan memberikan masukan kepada penulis sehingga penulis dapat menghasilkan skripsi yang baik.
7. Segenap Dosen, Admin, Laboran dan Jajaran Staff Program Studi Teknik Informatika yang telah memberikan ilmu, dukungan, dan bimbingan yang sangat berarti bagi penulis selama menjalani studi di Universitas Islam Maulana Malik Ibrahim Malang.
8. Kedua orang tua tercinta, Bapak Nurjudi dan Ibu Nanin Yulianik yang telah memberi cinta, doa, dan dukungan yang tak terhingga. Tanpa dukungan dan kasih sayang mereka, penulis tidak akan bisa mencapai titik ini. Terima kasih telah menjadi tempat penulis untuk selalu kembali, memberikan doa yang tiada henti, dan memberikan segala yang terbaik demi masa depan penulis.
9. Teman sepembimbingan, Iqlima Rahmafutri Karindah yang telah menemani penulis melalui setiap proses penyusunan skripsi ini. Terima kasih atas kebersamaan yang penuh dengan semangat, motivasi, dan saling mendukung.

10. Teman-teman dekat, Annisa Ayu Permadani, Nurul Aziza Kusuma Dewi, Ferina Putri Soedjono, Vinka Rosalia Fatimah Devi, Jinan Naurah, dan Vivin Octavia Cahyani yang telah menjadi bagian penting dalam perjalanan penulis selama perjalanan kuliah ini. Terima kasih atas kebersamaan, semangat, dan dukungan yang selalu kalian berikan.
11. Teman-teman kos, Salma Chesha dan Sarah Arelia, yang telah memberikan banyak masukan serta membantu penulis dalam memahami berbagai bagian skripsi ini. Terima kasih atas kebersamaan yang penuh dukungan dan semangat selama proses penyusunan skripsi.
12. Seluruh pihak yang telah terlibat secara langsung maupun tidak langsung dalam proses penyusunan skripsi sejauh ini.

Akhir kata, penulis mengakui bahwa penulisan pada skripsi ini masih banyak kekurangan. Penulis berharap semoga skripsi ini diterima sebagai amal ibadah yang tulus dan bermanfaat di sisi Allah Subhanahu Wa Ta'ala. Semoga karya ini menjadi bagian dari kontribusi yang tak terputus dalam rangka memperkuat dan mengembangkan ilmu pengetahuan, serta melaksanakan tugas sebagai hamba Allah yang berkomitmen.

Wassalamualaikum Warahmatullahi Wabarakatuh.

Malang, 5 Juni 2025

Penulis

DAFTAR ISI

HALAMAN PENGAJUAN	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN	iv
PERNYATAAN KEASLIAN TULISAN	v
MOTTO	vi
HALAMAN PERSEMBAHAN.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL.....	xiv
ABSTRAK	xv
ABSTRACT	xvi
الملخص.....	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
BAB II STUDI PUSTAKA	7
2.1 Tanah Longsor	7
2.2 Pengolahan Citra	8
2.3 Ekstraksi Fitur	10
2.4 <i>Gray Level Co-Occurrence Matrix</i>	11
2.5 <i>Clustering</i>	17
2.6 <i>Principal Component Analysis (PCA)</i>	20
2.7 Penelitian Terdahulu	23
BAB III DESAIN DAN IMPLEMENTASI	29
3.1 Desain Sistem.....	29
3.1.1 Input Data.....	30
3.1.2 <i>Pre-processing</i>	31
3.1.3 Proses GLCM (<i>Gray Level Co-Occurrence Matrix</i>)	35
3.1.4 Proses PCA (<i>Principal Component Analysis</i>).....	37
3.2 <i>Data Preparation</i>	38
3.3 Skenario Pengujian	39
BAB IV HASIL DAN PEMBAHASAN.....	42
4.1 Langkah-langkah Uji Coba	42
4.2 Hasil Uji Coba.....	43
4.2.1 Input Data.....	43
4.2.2 <i>Pre-processing</i>	43
4.2.3 Implementasi <i>Gray Level Co-Occurrence Matrix</i>	46

4.2.4 Implementasi <i>Principal Component Analysis</i>	50
4.2.5 Evaluasi Hasil <i>Clustering</i>	71
4.3 Pembahasan.....	86
4.4 Integrasi Islam.....	91
BAB V KESIMPULAN DAN SARAN	95
5.1 Kesimpulan	95
5.2 Saran	96
DAFTAR PUSTAKA	
LAMPIRAN	

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi Arah Pasangan Piksel dari <i>Pixel Of Interest</i>	13
Gambar 2.2 Ilustrasi Perhitungan Matriks GLCM dengan Arah 0° dan Jarak 1... 13	
Gambar 3.1 Desain Sistem.....	30
Gambar 4.1 Dataset Augmentasi Data	44
Gambar 4.2 Hasil Tahapan Pre-processing pada Citra	45
Gambar 4.3 Matriks Nilai dari Hasil Konversi Grayscale Image	46
Gambar 4.4 Hasil Ekstraksi Fitur dari Matriks GLCM.....	50
Gambar 4.5 Nilai Silhouette Score dengan Jarak GLCM 0.5.....	89
Gambar 4.6 Nilai Silhouette Score dengan Jarak GLCM 0.8.....	90

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	24
Tabel 4.1 Matriks GLCM.....	47
Tabel 4.2 Hasil Normalisasi Ekstraksi Fitur GLCM.....	51
Tabel 4.3 Persentase Nilai <i>Variance Ratio</i> dan <i>Eigen Value</i> Rotasi 15°	59
Tabel 4.4 Persentase Nilai <i>Variance Ratio</i> dan <i>Eigen Value Horizontal Flip</i>	59
Tabel 4.5 Persentase Nilai <i>Variance Ratio</i> dan <i>Eigen Value Vertical Flip</i>	60
Tabel 4.6 Eigenvector dari PC1	62
Tabel 4.7 Visualisasi PCA	65
Tabel 4.8 Pembagian Cluster Berdasarkan Rentang Nilai PC1	67
Tabel 4.9 Hasil <i>Clustering</i>	67
Tabel 4.10 Titik Koordinat dan Label Data <i>Clustering</i>	69
Tabel 4.11 Nilai <i>Silhouette Score</i> Rotasi 15° dengan Jarak 0.5	72
Tabel 4.12 Nilai <i>Silhouette Score Horizontal Flip</i> dengan Jarak 0.5.....	73
Tabel 4.13 Nilai <i>Silhouette Score Vertical Flip</i> dengan Jarak 0.5	75
Tabel 4.14 Nilai <i>Silhouette Score</i> Rotasi 15° dengan Jarak 0.8.....	77
Tabel 4.15 Nilai <i>Silhouette Score Horizontal Flip</i> dengan Jarak 0.8.....	78
Tabel 4.16 Nilai <i>Silhouette Score Vertical Flip</i> dengan Jarak 0.8	80
Tabel 4.17 Rentang Nilai PC1 (n) pada 2 Klaster.....	82
Tabel 4.18 Hasil Evaluasi <i>Horizontal Flip Jarak 0.8</i> pada 2 klaster	83
Tabel 4.19 Rentang Nilai PC1 (n) pada 4 Klaster.....	84
Tabel 4.20 Hasil Evaluasi <i>Horizontal Flip Jarak 0.8</i> pada 4 Klaster.....	85

ABSTRAK

Nuraini, Salsabila Ramadanti. 2025. *Clustering Tingkat Kerawanan Daerah Tanah Longsor Menggunakan Gray Level Co-Occurrence Matrix Dan Principal Component Analysis*. Skripsi. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Agung Teguh Wibowo Almais, M.T. (II) Ashri Shabrina Afrah, M.T.

Kata kunci: *Clustering*, Gray Level Co-Occurrence Matrix (GLCM), Principal Component Analysis (PCA), Tekstur Citra, Tanah Longsor.

Identifikasi wilayah rawan longsor berbasis analisis tekstur citra lahan merupakan pendekatan yang potensial dalam mendukung mitigasi risiko bencana. Penelitian ini mengusulkan metode pengelompokan tingkat kerawanan longsor menggunakan ekstraksi fitur tekstur melalui *Gray Level Co-occurrence Matrix* (GLCM) dan reduksi dimensi menggunakan *Principal Component Analysis* (PCA). Fitur yang diekstraksi meliputi kontras, energi, *dissimilarity*, *homogeneity*, korelasi, dan entropi dari citra lahan miring. Penelitian ini mengkaji dua jarak GLCM (0.5 dan 0.8 piksel) serta tiga jenis augmentasi citra (*rotasi*, *horizontal flip*, dan *vertical flip*). Proses klasterisasi dilakukan berdasarkan nilai komponen utama pertama (PC1), yang merepresentasikan dimensi dominan fitur tekstur. Hasil evaluasi menunjukkan bahwa kombinasi jarak GLCM 0.8 dengan augmentasi horizontal flip menghasilkan distribusi klaster yang lebih seimbang dibandingkan jarak 0.5, meskipun dengan nilai Silhouette Score yang sedikit lebih rendah. Rentang nilai PC1 optimal untuk tiga klaster kerawanan adalah: klaster 0 ($-4.3 \leq n < -1.5$), klaster 1 ($-1.5 < n < 1.5$), dan klaster 2 ($1.5 \leq n \leq 4.3$). Temuan ini menunjukkan bahwa pendekatan GLCM-PCA efektif dalam mengidentifikasi pola tekstur yang relevan untuk *clustering* tiga tingkat kerawanan longsor berbasis citra lahan miring

ABSTRACT

Nuraini, Salsabila Ramadanti. 2025. ***Clustering Tingkat Kerawanan Daerah Tanah Longsor Menggunakan Gray Level Co-Occurrence Matrix Dan Principal Component Analysis***. Thesis. Informatics Engineering Study Program Faculty of Science and Technology Maulana Malik Ibrahim State Islamic University Malang. Advisor: (I) Dr. Agung Teguh Wibowo Almais, M.T. (II) Ashri Shabrina Afrah, M.T.

Landslide-prone area identification based on land surface texture analysis is a promising approach to support disaster risk mitigation. This study proposes a method for clustering landslide susceptibility levels using texture feature extraction through the Gray Level Co-occurrence Matrix (GLCM) and dimensionality reduction via Principal Component Analysis (PCA). Extracted features include contrast, energy, dissimilarity, homogeneity, correlation, and entropy from sloped land imagery. The study evaluates two GLCM distances (0.5 and 0.8 pixels) and three types of image augmentation (rotation, horizontal flip, and vertical flip). Clustering is performed based on the first principal component (PC1), which represents the dominant dimension of texture features. Evaluation results show that the combination of a 0.8-pixel GLCM distance with horizontal flip augmentation yields a more balanced cluster distribution compared to a 0.5-pixel distance, although with a slightly lower Silhouette Score. The optimal PC1 ranges for the three susceptibility clusters are: cluster 0 ($-4.3 \leq n < -1.5$), cluster 1 ($-1.5 < n < 1.5$), and cluster 2 ($1.5 \leq n \leq 4.3$). These findings demonstrate that the GLCM-PCA approach is effective in identifying texture patterns relevant to clustering three levels of landslide susceptibility based on sloped land imagery.

Key words: Clustering, Gray Level Co-Occurrence Matrix (GLCM), Image Texture, Landslide, Principal Component Analysis (PCA).

المخلص

نورعبي، سلسبيل راماداني . ٢٠٢٥ . تجميع مستوى قابلية التأثير بالاختيارات الأرضية باستخدام مصفوفة التكرار المشترك للمستوى الرمادي وتحليل المكونات الرئيسية. أطروحة. برنامج دراسة هندسة المعلوماتية في كلية العلوم والتكنولوجيا بجامعة مولانا مالك إبراهيم الإسلامية الحكومية في مالانج. المشرف : (١) الدكتور أغونغ تيغوه ويبيو أليس، ماجستير (٢) أشري شيرينا أفراح، ماجستير.

الكلمات المفتاحية: التجميع، مصفوفة التواجد الرمادي (GLCM)، تحليل المكونات الرئيسية (PCA)، نسيج الصورة، الاختيارات الأرضية

يُعدّ تحديد المناطق المعرضة للاختيارات الأرضية استناداً إلى تحليل نسيج سطح الأرض تحجاً واعداً لدعم جهود التخفيف من مخاطر الكوارث. تقترح هذه الدراسة طريقة لتجميع مستويات القابلية للاختيارات الأرضية من خلال استخراج سمات النسيج باستخدام مصفوفة التكرار الرمادي المشترك (GLCM) وتقليل الأبعاد باستخدام تحليل المكونات الرئيسية (PCA). تشمل السمات المستخرجة كلاً من التباين، والطاقة، والاختلاف، والتجانس، والارتباط، والإنتروبيا، وذلك من صور الأراضي المنحدرة. تقيم الدراسة مسافتين من GLCM (٠,٥ و ٠,٨ بكسل) وثلاثة أنواع من تعزيز الصور (التدوير، والانقلاب الأفقي، والانقلاب الرأسي). يتم تنفيذ عملية التجميع بالاعتماد على المكوّن الرئيسي الأول (١PC)، الذي يُقلّل البُعد المسيطر لسمات النسيج. أظهرت نتائج التقييم أن الجمع بين مسافة GLCM بمقدار ٠,٨ بكسل مع تعزيز الانقلاب الأفقي ينتج توزيعاً أكثر توازناً للمجموعات مقارنةً باستخدام مسافة ٠,٥ بكسل، على الرغم من انخفاض قيمة معامل الصورة الظلية (Silhouette Score) بشكل طفيف. وقد تم تحديد النطاقات المثلى لقيم ١PC لتقسيم ثلاث مجموعات على النحو التالي: المجموعة ٠ (٤,٣- $n \geq ١,٥$)، والمجموعة ١ (١,٥- $n > ١,٥$)، والمجموعة ٢ (١,٥ $n > ١,٥$). تشير هذه النتائج إلى أن منهجية GLCM-PCA فعّالة في التعرف على أنماط النسيج المرتبطة بتصنيف ثلاث مستويات من القابلية للاختيارات الأرضية بالاعتماد على صور الأراضي المنحدرة.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Salah satu bahaya global yang signifikan memengaruhi populasi dan infrastruktur di berbagai belahan dunia yaitu tanah longsor (Emberson et al., 2020). Menurut data dari *Global Landslide Catalog* yang dikembangkan oleh NASA, negara-negara dengan jumlah kejadian tanah longsor tinggi antara lain yaitu India, Nepal, China, Filipina, dan Indonesia (Dandridge et al., 2023). Tingginya frekuensi terjadinya tanah longsor di negara-negara tersebut disebabkan oleh berbagai faktor seperti geologis, hidrometologi, serta aktivitas antropogenik yang mempengaruhi kestabilan tanah dan lereng (Fidan et al., 2024).

Indonesia merupakan negara dengan tingkat kerawanan tanah longsor yang tinggi akibat kondisi vulkanik, curah hujan tinggi, serta deforestasi yang masih aktif. Beberapa wilayah yang paling rentan mengalami tanah longsor meliputi Sumatera, Jawa, dan Sulawesi, dengan tingkat kerawanan paling tinggi adalah Jawa. Provinsi yang paling terdampak bencana yaitu Jawa Barat, Jawa Tengah, dan Jawa Timur (Suwanda Risky et al., 2023).

Salah satu daerah di Pulau Jawa yang memiliki tingkat kerawanan tanah longsor yang tinggi yaitu Kota Batu. Data dari Badan Penanggulangan Bencana Daerah (BPBD) Kota Batu menunjukkan bahwa bencana alam yang mendominasi daerah ini adalah tanah longsor. Pada tahun 2023, tercatat 73 kejadian tanah longsor dari 206 bencana, sementara pada tahun 2024 terdapat 56 kejadian tanah longsor

dari total 122 bencana (Wibowo, 2025).

Wilayah berisiko tinggi umumnya memiliki kemiringan lereng yang curam antara 8-15° (>45%), formasi geologi, tanah laktosol coklat kemerahan, intensitas curah hujan yang tinggi (4000-5000 mm/tahun), serta pertanian lahan kering. Aktivitas penebangan hutan dan pembangunan permukiman di daerah berisiko juga semakin memperparah kondisi tanah yang sudah tidak stabil secara alami (Priyono, 2020).

Bencana alam akibat dari aktivitas manusia seperti deforestasi dan pembangunan di daerah rawan bencana merupakan bentuk ketidakseimbangan yang dilarang dalam agama islam. Hal ini sesuai dalam firman Allah dalam Surah Al-A'raf ayat 56 yang berbunyi sebagai berikut.

﴿٥٦﴾ وَلَا تُفْسِدُوا فِي الْأَرْضِ بَعْدَ إِصْلَاحِهَا وَادْعُوهُ خَوْفًا وَطَمَعًا ۚ إِنَّ رَحْمَتَ اللَّهِ قَرِيبٌ مِّنَ الْمُحْسِنِينَ

“Dan janganlah kamu berbuat kerusakan di muka bumi setelah (Allah) memperbaikinya, dan berdoalah kepada-Nya dengan rasa takut dan penuh harap. Sesungguhnya rahmat Allah sangat dekat kepada orang-orang yang berbuat kebaikan.”(Q.S. Al-A'raf: 56).

Menurut Tafsir Jalalain, ayat Al-Qur'an tersebut menegaskan agar para manusia tidak melakukan kerusakan di muka bumi setelah Allah Swt. menciptakannya dalam keadaan seimbang (al-Mahalli & as-Suyuthi, 2015). Penebangan yang tidak terkendali menghilangkan fungsi alamiah hutan dalam menyerap air dan menahan tanah, sehingga meningkatkan risiko erosi dan longsor. Selain itu, pembangunan di daerah yang secara geologis tidak stabil tanpa perhitungan yang matang dapat mempercepat terjadinya bencana alam. Upaya yang bisa dilakukan salah satunya yaitu dengan menganalisis lingkungan tersebut.

Daerah yang akan dihuni harus diidentifikasi keamanannya agar terhindar dari bencana tanah longsor.

Saat ini, metode identifikasi daerah rawan longsor sering kali masih bergantung pada analisis manual yang memerlukan keahlian khusus, waktu, dan biaya yang tinggi untuk survei lapangan (Pang et al., 2022). Pendekatan konvensional juga masih sering tidak memiliki sistem klasifikasi yang jelas, sehingga menghasilkan hasil yang subjektif dan berpotensi tidak akurat (Cantarino et al., 2023).

Gray Level Co-occurrence Matrix merupakan salah satu metode statistik di dalam pemrosesan gambar yang bisa digunakan untuk proses pengidentifikasian pola tekstur permukaan tanah dari lahan miring (Faisal et al., 2022). Metode ini mengekstraksi beberapa fitur seperti kontras, *dissimilarity*, homogenitas, energi, korelasi dan entropi sehingga membantu dalam membedakan pola gambar berdasarkan tekstur yang ada dan memberikan informasi yang berguna dalam analisis gambar/citra. Sedangkan pengelompokan tingkat kerentanan tanah longsor menjadi kategori ringan, sedang, dan berat menggunakan integrasi dengan metode *PCA clustering*. *PCA* mengurangi dimensi data dengan mentransformasikan fitur ke dalam komponen utama, yang kemudian digunakan pada proses klusterisasi (Almais et al., 2024).

Penelitian yang dilakukan oleh Almais et al., (2024) menggunakan integrasi dari metode GLCM-PCA untuk karakteristik bangunan dan menunjukkan hasil yang baik terutama ketika menggunakan jarak 0,5 piksel. Hasil analisis pada

metode ini dapat mengidentifikasi kerusakan parah dengan akurasi yang sebanding dengan hasil yang diperoleh dari *surveyor* (nilai $PC1 > 2$).

Dengan demikian, penelitian ini menawarkan pendekatan baru dalam penerapan metode GLCM-PCA dalam melakukan *clustering* tingkat kerawanan daerah tanah longsor. Diharapkan metode ini dapat meningkatkan akurasi dalam menganalisis faktor-faktor pemicu longsor, sehingga dapat membantu lembaga terkait dalam perencanaan mitigasi bencana.

1.2 Rumusan Masalah

Permasalahan yang dapat diidentifikasi dari latar belakang di atas adalah sebagai berikut:

1. Bagaimana cara mengidentifikasi daerah rawan longsor secara lebih akurat menggunakan metode GLCM?
2. Sejauh mana metode GLCM-PCA mampu mengelompokkan tingkat kerawanan daerah tanah longsor berdasarkan citra kemiringan lahan secara efektif?

1.3 Batasan Masalah

Terdapat batasan masalah pada penelitian ini, dengan batasan sebagai berikut:

1. Analisis dalam penelitian ini berbasis pada data peta persebaran ancaman bencana tanah longsor dari BPBD Kota Batu.
2. Ruang lingkup penelitian ini hanya mencakup bencana tanah longsor.

3. Data yang digunakan dalam penelitian ini terbatas pada citra lahan miring atau lereng, sehingga faktor penyebab tanah longsor lainnya tidak dianalisis secara langsung.

1.4 Tujuan Penelitian

Penelitian ini memiliki beberapa tujuan, dengan tujuan penelitian sebagai berikut:

1. Menganalisis pola tekstur tanah dengan menerapkan metode GLCM dalam mendeteksi daerah rawan longsor.
2. Menganalisis kualitas integrasi metode GLCM-PCA dalam mengelompokkan daerah rawan longsor berdasarkan citra lahan miring sehingga dapat memberikan hasil penelitian yang lebih objektif, terstruktur, dan efisien.

1.5 Manfaat Penelitian

Terdapat beberapa manfaat pada penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Membantu perkembangan riset ilmiah dalam penggunaan GLCM dan PCA *clustering* sebagai metode untuk mengelompokkan daerah berdasarkan tingkat kerawanan tanah longsor.
2. Menyediakan landasan bagi riset mendatang dalam pengembangan teknologi pemetaan risiko tanah longsor menggunakan metode analisis citra.

3. Memungkinkan masyarakat untuk merencanakan pembangunan dengan mempertimbangkan risiko tanah longsor secara lebih akurat.
4. Membantu masyarakat dalam mengidentifikasi daerah rawan longsor sehingga dapat mengambil langkah pencegahan yang efektif.

BAB II

STUDI PUSTAKA

Bab ini berisi landasan teori dan hasil-hasil penelitian terdahulu yang relevan dengan topik penelitian. Studi pustaka bertujuan untuk memberikan dasar yang konseptual dan teoritis kuat terhadap permasalahan yang dibahas dalam penelitian, serta menunjukkan posisi penelitian yang sedang dilakukan dibandingkan dengan penelitian sebelumnya. Studi ini sangat penting untuk dilakukan agar memastikan bahwa penelitian yang dilakukan tidak berdiri sendiri, tetapi terhubung dan berkontribusi terhadap pengetahuan yang sudah ada sebelumnya.

2.1 Tanah Longsor

Tanah longsor merupakan fenomena geologi yang melibatkan pergerakan material berupa batuan, puing-puing, atau tanah menuruni lereng akibat pengaruh gravitasi (Gakaev, 2021). Proses longsor tanah pada umumnya dimulai ketika air hujan meresap ke dalam pori-pori tanah sehingga menambah bobot keseluruhan tanah serta mengurangi daya rekat antar butiran tanah. Jika kondisi ini terjadi di daerah dengan lereng curam dan tanah pelapukan yang tebal, maka bagian atas tanah akan mengalami pergeseran. Lapisan tanah tersebut dapat dengan mudah bergerak mengikuti kemiringan lereng sehingga memicu terjadinya longsor (Abraham et al., 2020).

Jenis tanah longsor yang paling umum terjadi di Indonesia adalah longsor permukaan dan longsor dangkal yang sering terjadi di lereng curam dengan formasi

geologi kuarter serta jenis tanah tertentu (Fata et al., 2022). Formasi geologi kuarter merupakan lapisan batuan muda yang masih belum stabil sehingga rawan terhadap bencana longsor dan erosi.

Jenis-jenis tanah yang menjadi penyebab tanah longsor yaitu Andisol, Inceptisol, dan Entisol. Andisol terbentuk dari abu vulkanik yang memiliki tingkat porositas tinggi dan daya serap air yang besar sehingga saat kondisi pori-pori tanah terisi penuh oleh air, kestabilannya menurun dan meningkatkan risiko longsor (Budianta et al., 2022). Inceptisol merupakan tanah yang masih dalam tahap perkembangan awal yang umumnya ditemukan pada daerah dengan curah hujan tinggi dan memiliki kohesi bervariasi sehingga rentan terhadap pergerakan tanah (Suryani et al., 2021). Sementara itu, Entisol adalah tanah yang belum berkembang yang sering dijumpai di daerah aliran sungai dan lereng pegunungan sehingga mudah tergerus aliran air hujan (Goyal et al., 2022).

Tanah longsor umumnya diklasifikasikan menjadi beberapa tingkat risiko meliputi tingkat rendah, sedang, dan tinggi (Susetyo et al., 2022). Tingkat kerawanan rendah terdapat pada daerah datar dengan tanah yang stabil (0° - 15°). Tingkat kerawanan sedang terjadi pada lereng dengan kemiringan sedang (15° - 30°). Sementara itu, tingkat kerawanan tinggi pada lereng terjal ($>30^{\circ}$) dengan struktur tanah lepas.

2.2 Pengolahan Citra

Pengolahan citra merupakan salah satu bidang dalam *signal processing* yang memiliki tujuan untuk mengolah gambar sebagai sinyal dua dimensi sehingga dapat mengekstrak informasi atau meningkatkan kualitas visual (Zheng, 2023).

Bidang ini meliputi berbagai teknik yang memanfaatkan gambar atau citra sebagai masukan untuk menghasilkan parameter yang mempresentasikan karakteristik citra.

Pengolahan citra digunakan untuk meningkatkan informasi visual agar lebih mudah diinterpretasikan oleh manusia (Surve et al., 2022). Pengolahan citra mengonversikan gambar menjadi array piksel, yang kemudian diproses oleh komputer untuk berbagai keperluan analisis dan pemrosesan lebih lanjut.

Dalam pengolahan citra digital, gambar direpresentasikan dalam bentuk array dua dimensi yang terdiri dari elemen-elemen piksel. Setiap piksel dalam citra memiliki nilai intensitas tertentu yang menunjukkan tingkat keabuan atau warna yang berkontribusi terhadap informasi visual dalam citra. Representasi citra dalam komputer dapat dibagi menjadi beberapa format, yaitu sebagai berikut.

1. Citra Biner

Representasi gambar dengan dua warna (hitam dan putih) yang sering digunakan dalam pemrosesan dokumen dan segmentasi objek.

2. Citra *Grayscale*

Gambar yang hanya memiliki tingkat kecerahan tanpa informasi warna, dengan skala 0-255 untuk citra 8-bit. Citra ini umumnya digunakan dalam berbagai aplikasi, seperti pemrosesan citra, pengenalan pola, dan analisis tekstur karena kemampuannya dalam menyajikan informasi visual tanpa kompleksitas warna.

3. Citra Berwarna (RGB)

Gambar yang memiliki tiga saluran warna utama seperti merah, hijau, dan

biru, yang dikombinasikan untuk menghasilkan berbagai warna. Citra memungkinkan representasi warna yang lebih alami dan akurat dan digunakan secara luas dalam beberapa aplikasi seperti Fotografi digital, tampilan layar, dan pengolahan citra medis untuk analisis visual yang lebih baik.

4. Citra Multispektral dan Hiperspektral

Gambar ini memiliki lebih dari tiga kanal warna, seperti inframerah dekat, inframerah gelombang pendek, dan ultraviolet yang umumnya digunakan dalam teknik pemantauan dan pengambilan data dari kejauhan menggunakan sensor pada satelit, pesawat, atau drone untuk menganalisis berbagai fenomena lingkungan dan citra permukaan bumi, serta analisis citra satelit (Zangana et al., 2024).

Penelitian ini membutuhkan citra *grayscale* untuk digunakan dalam metode GLCM yang perlu menghitung hubungan spasial antara piksel dengan tingkat keabuan tertentu untuk menganalisis pola tekstur citra lahan miring.

2.3 Ekstraksi Fitur

Ekstraksi fitur merupakan proses penting dalam sistem pengolahan data, khususnya dalam bidang seperti pengolahan citra, pengenalan pola, dan pembelajaran mesin (Suhaidi et al., 2021). Proses ini memiliki tujuan utama yaitu mengubah data mentah menjadi representasi numerik yang lebih terstruktur dan informatif, sehingga dapat digunakan dalam proses analisis di dalam penelitian (Lamer et al., 2022).

Ekstraksi fitur adalah proses penting dalam pembelajaran mesin dan analisis citra yang mengubah data masukan berdimensi tinggi menjadi vektor fitur berdimensi lebih rendah untuk keperluan klasifikasi, klastering, dan pengenalan. Proses ini memainkan peran penting dalam berbagai aplikasi, termasuk analisis teks, klasifikasi dan klastering citra, dan pemrosesan sinyal.

Fitur yang diekstrak harus bersifat diskriminatif, yang berarti harus mampu membedakan antar kelas atau kategori yang berbeda dalam data. Selain itu, fitur yang baik juga sebaiknya bersifat invarian terhadap perubahan skala, rotasi, pencahayaan, atau gangguan lain yang mungkin terjadi pada data mentah (Chakroun et al., 2022).

Secara umum, ekstraksi fitur dapat dilakukan secara manual maupun secara otomatis. Dalam metode manual, fitur ditentukan berdasarkan rumus atau perhitungan tertentu yang telah terbukti efektif pada jenis data tertentu. Sementara itu, dalam pendekatan otomatis, sistem belajar sendiri fitur yang paling relevan dari data melalui proses pelatihan model (Nogales & Benalcázar, 2023).

Ekstraksi fitur yang tepat sangat berpengaruh terhadap kinerja sistem secara keseluruhan. Semakin representatif fitur yang diperoleh, semakin baik pula akurasi dan efektivitas model dalam mengenali pola atau membuat keputusan berdasarkan data tersebut (Antoniuk et al., 2023).

2.4 Gray Level Co-Occurrence Matrix

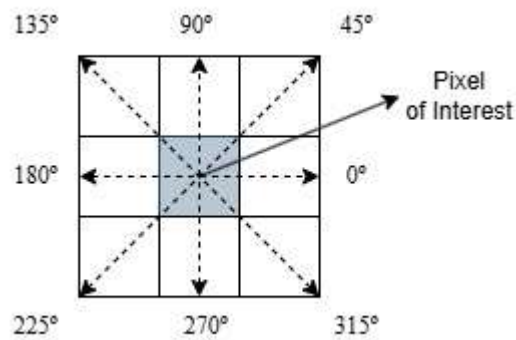
Gray Level Co-Occurrence Matrix merupakan metode yang digunakan dalam analisis tekstur dan ekstraksi fitur dalam pemrosesan citra. GLCM mempunyai beberapa tahapan utama dalam menganalisis tekstur suatu citra. Citra

akan dianalisis, dibaca, dan dikonversi menjadi format *grayscale* untuk menyederhanakan analisis tekstur. Setelah itu, dilakukan penentuan jarak dan orientasi sudut piksel utama (*pixel of interest*) dengan piksel tetangganya agar bisa menghasilkan matriks *co-occurrence*.

Matriks *co-occurrence* atau matriks GLCM merupakan representasi dua dimensi dari frekuensi kemunculan bersama pasangan nilai keabuan dalam sebuah citra yang digunakan untuk menganalisis pola tekstur dalam suatu citra secara statistik, sehingga karakteristik tekstur tersebut dapat dianalisis. Dua dimensi pada matriks ini mengacu pada sumbu baris i yang mewakili nilai keabuan piksel pertama dalam pasangan yang diamati dan sumbu kolom j yang mewakili nilai keabuan piksel tetangga dari piksel pertama berdasarkan pada jarak dan arah tertentu.

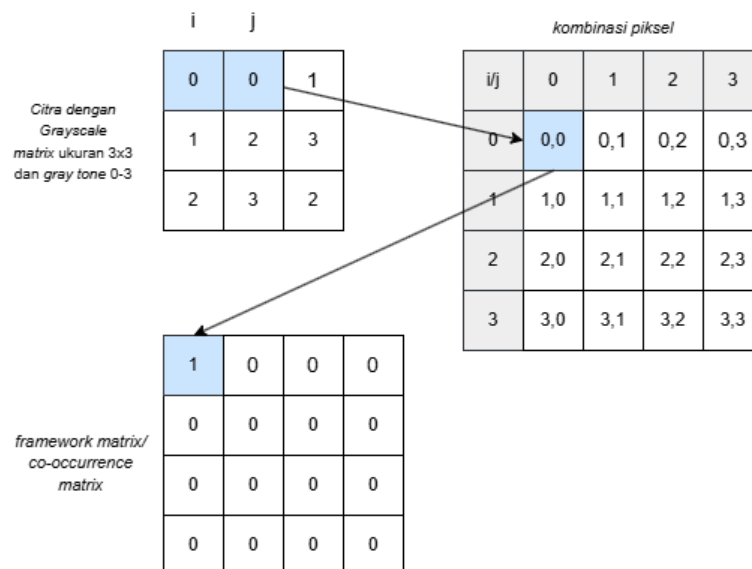
Jarak antar piksel adalah parameter yang menentukan seberapa jauh pasangan piksel dalam arah tertentu. Jarak dinyatakan dalam satuan piksel dan merupakan jarak garis lurus antara piksel pertama dan piksel tetangganya. Lalu, arah merupakan parameter yang menentukan sudut orientasi pasangan piksel yang dihitung hubungannya.

Sudut-sudut arah pasangan piksel yang umum digunakan, yaitu 0° (horizontal ke kanan), 45° (diagonal kanan atas), 90° (vertikal ke atas), 135° (diagonal kiri atas), 180° (horizontal ke kiri), 225° (diagonal kiri bawah), 270° (vertikal ke bawah), dan 315° (diagonal kanan bawah).



Gambar 2.1 Ilustrasi Arah Pasangan Piksel dari *Pixel Of Interest*

Perhitungan matriks GLCM dilakukan berdasarkan parameter jarak dan arah sudut. Matriks ini menghitung frekuensi kemunculan pasangan piksel dengan nilai intensitas *gray level* berdasarkan jarak dan sudut yang telah ditentukan. Secara umum, setiap kali ditemukan kombinasi pasangan nilai intensitas piksel dalam citra, maka nilai dalam matriks pada posisi yang sesuai akan ditambah satu.



Gambar 2.2 Ilustrasi Perhitungan Matriks GLCM dengan Arah 0° dan Jarak 1 pada Citra *Grayscale* 3x3

Gambar 2.2 merupakan contoh perhitungan matriks GLCM pada citra berukuran 3x3 dengan level keabuan dari 0 hingga 3. Parameter arah sebesar 0° dan jarak 1 piksel digunakan dalam ilustrasi tersebut. Proses awal yang ditunjukkan melibatkan pasangan piksel i dan j pertama horizontal (0,0) dengan nilai 1 karena hanya satu kali muncul untuk dua nilai piksel tersebut. Lalu, pasangan piksel i bernilai 0 dan piksel j bernilai 1 juga hanya satu kali muncul. Perhitungan berlanjut hingga semua nilai piksel pada citra *grayscale matrix* terdefinisi. Hasil matriks tersebut akan di-*transpose* dan dinormalisasi agar menghasilkan representasi tekstur dengan lebih lengkap dan stabil.

Matriks GLCM bergantung pada jumlah level keabuan sehingga ukuran dan bentuk matriks bisa menjadi sangat besar. Oleh karena itu, matriks ini perlu diekstraksi menjadi sejumlah nilai fitur ringkas yang merepresentasikan karakteristik tekstur secara efisien. Fitur-fitur yang umum diekstraksi yaitu kontras, *dissimilarity*, homogenitas, energi, korelasi, dan entropi (Saputra et al., 2020). Fitur-fitur tersebut digunakan dalam analisis dan pengelompokan data.

1. Kontras

Fitur ini mengukur variasi intensitas atau perbedaan lokal dalam citra. Nilai kontras tinggi menunjukkan perbedaan intensitas yang besar antara piksel yang berdekatan, menandakan tekstur yang kasar. Sebaliknya, jika nilai kontras rendah menunjukkan perbedaan intensitas yang kecil dan menandakan tekstur yang halus. Berikut merupakan rumus fitur kontras.

$$\text{Contrast} = \sum_{i,j} (i - j)^2 P(i, j) \quad (2.1)$$

Keterangan:

- a) i = indeks baris dalam matriks GLCM
- b) j = indeks kolom dalam matriks GLCM
- c) P = frekuensi pasangan piksel i dan j muncul pada jarak dan sudut tertentu

2. Ketidakteraturan (*Dissimilarity*)

Dissimilarity mengukur sejauh mana intensitas piksel dalam suatu citra berbeda satu sama lain. Jika nilainya tinggi artinya terdapat variasi intensitas yang lebih besar, sementara nilai yang rendah mengindikasikan bahwa piksel memiliki intensitas yang lebih mirip.

$$\text{Dissimilarity} = \sum_{i,j} |i - j| P(i, j) \quad (2.2)$$

Nilai absolut dari selisih antara indeks baris i dan indeks baris j dikalikan dengan probabilitas kemunculan pasangan piksel tersebut sehingga menghasilkan nilai dissimilarity.

3. Homogenitas

Homogenitas atau yang juga dikenal sebagai *Inverse Difference Moment* (IDM), merupakan fitur yang menggambarkan tingkat keseragaman suatu tekstur citra. Semakin tinggi nilainya, maka semakin kecil perbedaan intensitas antara piksel yang berdekatan. Hal ini menunjukkan bahwa tekstur lebih seragam dan memiliki perubahan nilai keabuan yang lebih halus.

$$\text{Homogeneity} = \sum_{i,j} \frac{P(i,j)}{1+(i-j)^2} \quad (2.3)$$

4. Energi

Fitur energi atau Angular Second Moment (ASM), merupakan ukuran keseragaman pola tekstur dalam citra. Tekstur yang memiliki nilai energi

tinggi biasanya lebih homogen dan berpola, sedangkan nilai rendah menunjukkan ketidakteraturan pola dalam citra.

$$\text{Energy} = \sum_{i,j} P(i,j)^2 \quad (2.4)$$

5. Korelasi

Korelasi mengukur sejauh mana tingkat keselarasan antara intensitas piksel yang berdekatan. Jika korelasi tinggi, maka perubahan nilai piksel dalam citra memiliki pola yang seragam, sedangkan korelasi rendah menunjukkan hubungan yang tidak menentu antar piksel.

$$\text{Correlation} = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)P(i,j)}{\sigma_i\sigma_j} \quad (2.5)$$

Keterangan :

- a) μ_i = rata-rata tingkat keabuan untuk baris i
- b) μ_j = rata-rata tingkat keabuan untuk baris j
- c) σ_i = standar deviasi tingkat keabuan untuk baris i
- d) σ_j = standar deviasi tingkat keabuan untuk baris j

6. Entropi

Entropi merupakan ukuran tingkat kompleksitas atau kerandoman tekstur dalam citra berdasarkan matriks *co-occurrence*. Nilai dari fitur ini berbanding terbalik dengan homogenitas yang mengukur tingkat keseragaman tekstur suatu citra. Fitur ini memberikan informasi penting mengenai karakteristik kekasaran dan kehalusan tekstur. Semakin tinggi nilai entropi, maka semakin kompleks atau acak tekstur permukaan citra tersebut. Ketika nilainya mendekati 1, maka citra tersebut memiliki tingkat kekasaran yang tinggi, sedangkan nilai entropi rendah menandakan tekstur citra tersebut lebih seragam dan halus.

$$Entropy = \sum_{i,j} P(i,j) \cdot \log(p(i,j)) \quad (2.6)$$

Dalam penggunaannya, GLCM memiliki beberapa kelemahan yang perlu diperhatikan. Salah satu tantangan utama yaitu ketergantungannya pada parameter yang dapat memengaruhi akurasi hasil analisis, seperti pemilihan jarak dan sudut (Salunkhe et al., 2021). Selain itu, metode ini dapat menghasilkan matriks dengan dimensi yang sangat besar, terutama ketika digunakan pada citra beresolusi tinggi, sehingga membutuhkan sumber daya komputasi yang besar. GLCM juga sensitif terhadap *noise* (bintik-bintik) dalam citra karena menyebabkan pasangan piksel *graylevel* menjadi acak sehingga mengganggu perhitungan fitur tekstur (Wang et al., 2022).

Untuk mengatasi keterbatasan ini, beberapa pendekatan dapat dilakukan, seperti mengombinasikannya dengan metode lain, seperti teknik pengurangan dimensi atau algoritma *machine learning*, agar meningkatkan akurasi dan efisiensi dalam pengolahan citra (Almais et al., 2024).

2.5 *Clustering*

Clustering merupakan salah satu teknik *machine learning* tanpa pengawasan (*unsupervised*) yang mengungkap pola dan struktur data berdasarkan kemiripan karakteristik antar data dengan mengelompokkan data tidak berlabel (tanpa membutuhkan klasifikasi awal) yang serupa ke dalam beberapa *cluster* atau kelompok. Objek-objek dalam satu *cluster* memiliki tingkat kemiripan pola atau nilai fitur yang tinggi, sedangkan objek antar *cluster* memiliki tingkat perbedaan yang signifikan (Feng et al., 2022). Terdapat beberapa jenis metode *clustering* berdasarkan cara kerjanya, yaitu sebagai berikut.

1. *Partitioning*

Partitioning clustering merupakan metode untuk membagi data ke dalam sejumlah *cluster* yang telah ditentukan sebelumnya. Tiap-tiap objek data hanya dapat menjadi anggota dalam satu klaster. Salah satu algoritma dengan jenis metode ini adalah K-Means. Keunggulan dari metode ini adalah kesederhanaan dan efisiensinya saat digunakan pada data dengan jumlah dimensi yang tidak terlalu banyak. Namun, pada metode ini harus menentukan jumlah klaster di awal (Urumov et al., 2022).

2. *Hierarchical*

Data dibentuk seperti *decision tree* atau dendogram yang setiap tingkat menyatukan atau memecah *cluster* berdasarkan kemiripan. Proses *clustering* dalam metode ini dilakukan melalui dua pendekatan, yaitu *bottom-up (agglomerative)* yang setiap data dianggap klaster tersendiri dan secara bertahap digabung hingga terbentuk satu klaster besar berdasarkan kemiripan terdekat serta pendekatan *top-down (divisive)* yang seluruh data dianggap satu klaster besar lalu dipisah menjadi beberapa sub-klaster secara bertahap. Metode ini tidak efisien jika diterapkan pada data besar, namun keunggulan dari metode ini yaitu tidak memerlukan penentuan jumlah klaster sejak awal (Eric U. & Michael O., 2024).

3. *Density-based*

Data dikelompokkan berdasarkan dari kepadatan titik data. *Cluster* terbentuk di daerah dengan kepadatan titik data yang tinggi, dan daerah yang jarang terdapat titik data dianggap *noise*. Salah satu algoritma pada metode

jenis ini adalah DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*). Algoritma ini memiliki keunggulan dalam hal mendeteksi klaster dengan bentuk yang arbitrer (tidak harus bulat atau oval) serta memiliki kemampuan untuk menangani *noise* secara eksplisit. Metode ini tidak mengharuskan untuk menentukan jumlah klaster terlebih dahulu, namun menentukan parameter kepadatan minimal dan jarak maksimum antar titik (Rupanka Bhuyan & Samarjeet Borah, 2013).

4. *Fuzzy*

Metode ini melakukan pengelompokan yang mengizinkan satu data boleh memiliki beberapa klaster tetapi dengan tingkat kedekatan data tersebut terhadap masing-masing klaster bernilai 0 sampai 1. Oleh karena itu, metode ini dinamakan pengelompokan kabur. Metode ini cocok untuk data yang tidak memiliki batas tegas antar klaster, seperti dalam pengelompokan citra kabur atau analisis medis. Metode ini memiliki kelebihan pada fleksibilitasnya untuk menangani data yang tidak jelas kategorinya, namun diperlukan proses komputasi yang lebih kompleks dan waktu yang lebih lama dari metode partisi konvensional (Bian et al., 2021).

Dalam penelitian ini, proses pengelompokan data dilakukan bukan menggunakan algoritma *clustering* tradisional, melainkan dengan memanfaatkan hasil proses reduksi dimensi data dari metode PCA. Hasil dari proses tersebut digunakan untuk mengidentifikasi pola atau kelompok data secara visual berdasarkan sebaran nilai komponen utama (*principal components*). Pendekatan tersebut dinamakan sebagai *clustering* berbasis visualisasi PCA.

2.6 *Principal Component Analysis (PCA)*

Principal Component Analysis (PCA) merupakan metode statistik yang umumnya banyak digunakan untuk mengurangi dimensi data dengan tetap mempertahankan informasi yang paling penting serta dapat digunakan untuk analisis data (Saccenti, 2024). Metode ini mampu mengidentifikasi fitur utama, memvisualisasikan hubungan atau relasi, dan menyederhanakan sekumpulan data yang kompleks (Mohammed et al., 2023).

Walaupun pada dasarnya metode ini merupakan teknik reduksi dimensi, beberapa penelitian menunjukkan bahwa metode ini dapat digunakan secara aktif dalam proses *clustering* dan bahkan dapat digunakan langsung sebagai teknik *unsupervised clustering*. PCA mengubah data berdimensi tinggi menjadi representasi dimensi rendah seperti 2D atau 3D dengan tetap mempertahankan informasi yang penting dalam data. Transformasi ini sangat membantu dalam menyingkap pola tersembunyi yang tidak terlihat dalam dimensi asli sehingga mempermudah proses *clustering* pada data kompleks.

Dalam *clustering*, hasil transformasi metode ini berupa *Principal Components* yang digunakan untuk memvisualisasikan dan mengidentifikasi kluster-kluster data yang serupa. Data dikelompokkan berdasarkan rentang nilai dari komponen utamanya. Dengan demikian, PCA juga dapat digunakan sebagai alat yang kuat untuk data *clustering*. Cara kerja PCA melalui beberapa tahapan yaitu sebagai berikut.

1. Normalisasi Data

Data yang akan dianalisis biasanya dilakukan proses normalisasi agar

memastikan setiap variabelnya seimbang. Dengan menerapkan normalisasi, PCA dapat secara adil menentukan kontribusi setiap fitur dalam pembentukan komponen utama. Normalisasi juga mencegah fitur dengan rentang nilai yang lebih besar mendominasi hasil analisis sehingga dapat menyebabkan kesalahan dalam tahapan selanjutnya.

Normalisasi data dapat dilakukan dengan menggunakan rumus Standard Scale (SS).

$$x_{norm} = \frac{x - \mu}{\sigma} \quad (2.7)$$

Keterangan:

- a) x_{norm} = nilai data setelah normalisasi.
- b) x = nilai asli data.
- c) μ = *mean* (rata-rata) dari data.
- d) σ = standar deviasi dari data.

Pada rumus tersebut, untuk menghitung nilai rata-rata dari data menggunakan rumus berikut:

$$\mu = \frac{\sum_{i=1}^n X_i}{n} \quad (2.8)$$

Keterangan:

- a) μ = rata-rata dari setiap data
- b) X_i = nilai individu setiap data
- c) n = jumlah sampel

2. Kalkulasi *Variance Ratio* dan *Eigen Value*

Perhitungan *Variance Ratio* dilakukan untuk menunjukkan jumlah informasi atau variasi data yang dapat dijelaskan masing-masing komponen.

$$s^2 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{(n-1)} \quad (2.9)$$

Keterangan:

- a) s^2 = varians sampel

Eigen Value dan *Eigen Vector* dihitung untuk menentukan arah dan besarnya variasi data. *Eigen Vector* berfungsi sebagai penunjuk arah *principal component* dalam ruang fitur, sedangkan *Eigen Value* menunjukkan frekuensi jumlah varians yang dijelaskan oleh masing-masing *principal component* (PC). Untuk menghitung *Eigen Value*, diperlukan persamaan 2.10. Sedangkan untuk menentukan *Eigen Vector* diperlukan persamaan 2.11.

$$\det (C - \lambda I) = 0 \quad (2.10)$$

$$Cv = \lambda v \quad (2.11)$$

Keterangan:

- a) C = matriks kovarians
- b) λ = nilai eigen
- c) v = *Eigen Vector*
- d) I = matriks identitas /matriks persegi berukuran sama dengan C

Sebelum menghitung *Eigen Value* dan *Eigen Vector* tersebut, dibutuhkan perhitungan matriks kovarians terlebih dahulu untuk menunjukkan seberapa jauh dua variabel berubah bersamaan. Matriks ini menggambarkan korelasi antara variabel yang menjadi dasar dalam menentukan pola distribusi data.

$$C = \frac{1}{(n-1)} \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})^T \quad (2.12)$$

Keterangan:

- a) n = jumlah data
- b) x_i = data ke-i
- c) $\bar{x}$ = rata-rata

3. Pemilihan *Principal Components* (PC)

Principal component dipilih berdasarkan *Eigen Value* yang lebih besar sehingga menunjukkan bahwa komponen tersebut menjelaskan lebih

banyak varians dalam data. Biasanya jumlah komponen utama dipilih berdasarkan proporsi varians yang ingin dipertahankan.

4. Proyeksi Data ke Ruang PCA

Data kemudian diproyeksikan ke dalam ruang yang baru dengan dimensi lebih kecil sesuai dengan komponen utama yang dipilih. Proses ini menghasilkan representasi yang lebih sederhana tanpa kehilangan informasi penting.

5. Visualisasi Hasil PCA

Hasil PCA dapat divisualisasikan dalam bentuk grafik dimensional untuk melihat distribusi data dalam ruang komponen utama.

6. Clustering Data

Setelah dilakukan visualisasi data, hasil *clustering* akan diperoleh setelah dilakukan identifikasi label berdasarkan rentang nilai dari komponen utamanya.

2.7 Penelitian Terdahulu

Bencana tanah longsor memang merupakan salah satu bencana alam yang sering terjadi di wilayah dengan topografi curam dan kondisi geologi tertentu. Bencana ini sering menimbulkan kerugian dan bahaya bagi masyarakat dan lingkungan. Berbagai penelitian telah dilakukan untuk mengurangi dampak bencana ini dengan mendeteksi dan memetakan kewarisan tanah longsor menggunakan berbagai metode dan pendekatan. Berikut ini merupakan tinjauan beberapa penelitian terkait.

Tabel 2.1 Penelitian Terdahulu

No	Referensi	Topik	Metode	Subjek	Hasil
1.	(Shahabi et al., 2021)	Deteksi Tanah Longsor	PCA, <i>Unsupervised Deep Learning</i> , dan CAE	Membuat peta tanah longsor menggunakan PCA untuk merekonstruksi citra asli dan menormalkan varians data.	Metode CAE berhasil mengekstraksi fitur abstrak dari data MNF, slope, dan NDVI tanpa data berlabel, menghasilkan peningkatan nilai precision, recall, F1-score, dan mIoU dalam deteksi longsor di tiga lokasi studi.
2.	(Xu et al., 2022)	<i>Landform Classification Using Image Texture</i>	GLCM dan CNN	Meningkatkan akurasi dalam klasifikasi bentuk lahan dengan menganalisis tekstur menggunakan GLCM, kemudian melakukan klasifikasi menggunakan CNN.	Integrasi GLCM dengan CNN meningkatkan akurasi klasifikasi bentuk lahan sebesar 6% dan koefisien kappa sebesar 0,4 dibanding metode tradisional.
3.	(Christaki et al., 2022)	<i>Building Change Detection</i>	ANN dan GLCM	Mengevaluasi keefektifan ANN dalam mendeteksi perubahan bangunan menggunakan metode GLCM untuk analisis tekstur citra.	GLCM yang dikombinasikan dengan ANN mencapai akurasi 92% dalam mendeteksi perubahan bangunan permukiman.
4.	(Thangakani et al., 2022)	<i>Malignant and Benign Lesions Detection</i>	CNN, HOG, dan GLCM	Mengembangkan teknik non-invasif berbasis <i>computer vision</i> untuk membedakan jenis lesi menggunakan kombinasi fitur HOG dan	HOG dan GLCM mampu mengklasifikasikan lesi kulit dengan akurasi rata-rata 98,32%, mengungguli metode sebelumnya.

No	Referensi	Topik	Metode	Subjek	Hasil
				GLCM yang diterapkan pada CNN.	
5.	(Tavus & Kocaman, 2023)	<i>Flood Mapping</i>	GLCM dan <i>Random Forest</i>	Menyelidiki pemetaan banjir menggunakan data satelit yang menunjukkan peningkatan akurasi klasifikasi dengan fitur GLCM.	GLCM dan Random Forest meningkatkan akurasi klasifikasi tutupan lahan dari 93% menjadi 99% untuk pemetaan banjir.
6.	(Kim & Yoon, 2023)	<i>Landslide Vulnerability Mapping</i>	PCA dan SMOTE	Menyederhanakan analisis dalam pemetaan kerentanan longsor menggunakan PCA.	Penggunaan PCA pada pemetaan kerentanan longsor menurunkan akurasi karena kurang merepresentasikan karakteristik data asli.
7.	(Kab et al., 2023)	<i>Landslide Susceptibility Assessment</i>	PCA	Memetakan zona kerentanan tanah longsor menggunakan PCA untuk menghasilkan peta kerentanan yang divalidasi indeks kepadatan longsor.	PCA berhasil mengatasi multikolinearitas dalam data, dengan model menghasilkan nilai AUC 0,782 yang menunjukkan presisi yang baik dalam pemetaan kerentanan longsor.
8.	(Yu et al., 2024)	<i>Landslide Detection</i>	GLCM dan <i>Deep Learning</i>	Meningkatkan akurasi deteksi tanah longsor dengan mengintegrasikan fitur tekstur GLCM ke dalam <i>Deep Learning</i> .	GLCM dan deep learning meningkatkan akurasi deteksi longsor dari citra RGB, dengan kenaikan rata-rata 32,6% pada IoU, 19,4% pada F1-score, dan 23,6% pada sensitivitas

No	Referensi	Topik	Metode	Subjek	Hasil
					dibandingkan RGB saja.
9.	(Almais et al., 2024)	Karakterisasi Kerusakan Bangunan Pasca-Bencana	GLCM-PCA	Melakukan penilaian kerusakan structural pada bangunan setelah bencana alam menggunakan integrasi GLCM dan PCA.	GLCM-PCA efektif untuk identifikasi bangunan pasca-bencana, dengan jarak piksel 0,5 menghasilkan informasi lebih akurat dibanding jarak 1, sesuai hasil survei lapangan.

Penelitian yang dilakukan oleh Shahabi et al., (2021) telah berhasil mengembangkan model *unsupervised deep learning* berbasis *Convolutional Auto-Encoder* (CAE) untuk mendeteksi longsor tanpa data berlabel. Citra dianalisis menggunakan transformasi *Minimum Noise Fraction* (MNF) dan PCA untuk meningkatkan akurasi. Hasil penelitian ini menunjukkan bahwa penggunaan fitur dalam proses klasterisasi meningkatkan hasil akurasi dari *confussion matrix* dalam deteksi longsor dibandingkan metode berbasis fitur konvensional.

Penelitian GLCM yang melakukan integrasi dengan *Convolutional Neural Network* (CNN) terbukti efektif dalam meningkatkan akurasi klasifikasi bentuk lahan. Metode ini memberikan peningkatan sekitar 6% dalam akurasi keseluruhan dan 0.4 dalam koefisien kappa dibandingkan metode tradisional (Xu et al., 2022).

Penelitian dari Christaki et al., (2022) menerapkan *Artificial Neural Network* pada fitur GLCM untuk mendeteksi perubahan dalam bangunan permukiman. Perubahan fitur tekstur GLCM yang dikombinasikan dengan ANN dapat memberikan hasil akurasi yang memuaskan yaitu mencapai 92%.

Thangakani et al., (2022) mengusulkan metode deteksi lesi kulit berbasis *computer vision* dengan menggabungkan *Histograms of Oriented Gradients* (HOG) dan GLCM yang mampu mengklasifikasikan lesi sebagai jinak atau ganas dengan akurasi rata-rata 98,32% pada dataset 4.100 gambar. Hasil ini menjadikan gabungan metode HOG-GLCM lebih unggul dari metode sebelumnya.

Tavus & Kocaman, (2023) melakukan pendekatan klasifikasi tutupan lahan untuk pemetaan luas banjir menggunakan Random Forest dan GLCM yang meningkatkan akurasi klasifikasi dari 93% menjadi 99%. Fitur tekstur dari GLCM secara efektif melakukan pemisahan vegetasi yang tergenang dan area banjir lainnya sehingga meningkatkan kualitas peta banjir. Sedangkan penelitian Kim & Yoon, (2023) menunjukkan bahwa dalam membangun *landslide susceptibility map*, data yang diolah menggunakan PCA memiliki keterbatasan dalam merepresentasikan karakteristik asli data sehingga menurunkan akurasi data.

Kab et al., (2023) mengidentifikasi zona kerentanan longsor yang menggunakan PCA untuk menghilangkan multikolinearitas (kondisi ketika terjadi korelasi tinggi antar 2 atau lebih variabel bebas) dalam dataset. Validasi model menghasilkan nilai AUC (*Area Under the Curve*) sebesar 0.782 yang menunjukkan presisi model yang baik.

Penelitian dari Yu et al., (2024) menunjukkan bahwa integrasi fitur tekstur GLCM dengan model *deep learning* secara signifikan meningkatkan akurasi deteksi longsor dari citra RGB penginderaan jauh. Hasil akurasi pada penelitian ini mengalami peningkatan rata-rata 32.6% pada IoU, 19.4% pada F1-score dan 23.6% pada sensitivitas dibandingkan dengan penggunaan data RGB saja.

Almais et al., (2024) melakukan penelitian dalam menentukan karakteristik bangunan pasca-bencana menggunakan integrasi metode GLCM-PCA. Hasil penelitian menunjukkan pemrosesan citra dengan GLCM pada jarak piksel 0.5 lebih akurat dalam menghasilkan informasi daripada dengan jarak piksel 1, sebagaimana dikonfirmasi oleh survei lapangan. Oleh karena itu, metode ini dapat digunakan untuk melakukan identifikasi menggunakan citra.

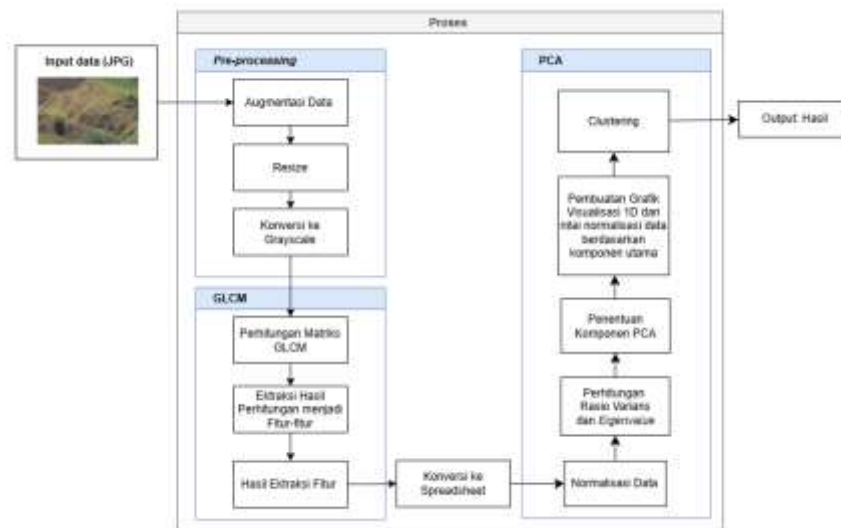
BAB III

DESAIN DAN IMPLEMENTASI

Pada bab ini berisi mengenai tahapan-tahapan dalam proses desain dan implementasi sistem yang digunakan dalam penelitian. Bab ini berisi desain sistem yang memuat perancangan alur sistem mulai dari proses input data, tahap *pre-processing*, hingga penerapan metode GLCM dan PCA untuk ekstraksi dan reduksi fitur citra. Bab ini juga berisi *data preparation* dan skenario pengujian. *Data preparation* membahas secara rinci mengenai persiapan dan pengolahan data yang diperlukan sebelum dilakukan pengujian atau uji coba. Skenario pengujian merupakan bagian yang menjelaskan rancangan skenario yang digunakan untuk mengevaluasi kerja sistem, baik dari segi akurasi dalam pemisahan data, kejelasan visual kluster yang terbentuk, maupun efektivitas metode yang digunakan dalam mengelompokkan tingkat kerawanan lahan. Pengujian ini merupakan langkah penting untuk memastikan bahwa sistem dapat berfungsi dengan baik dan memberikan hasil yang relevan dengan tujuan penelitian. Bab ini menjadi dasar teknik dari keseluruhan proses penelitian yang dilakukan.

3.1 Desain Sistem

Penelitian ini menggunakan *Gray Level Co-Occurrence Matrix* (GLCM) berperan dalam mengidentifikasi karakteristik tekstur citra dan dipadukan dengan metode *Principal Component Analysis* (PCA) untuk mereduksi dimensi data dengan tetap mempertahankan informasi penting. Berikut merupakan tahapan dalam metode penelitian ini.



Gambar 3.1 Desain Sistem

Gambar 3.1 menunjukkan desain sistem yang menggambarkan tahapan-tahapan utama dalam proses pengolahan citra untuk melakukan *clustering* tingkat kerawanan tanah longsor menggunakan metode GLCM dan PCA. Berikut merupakan penjelasan tahapan-tahapan dalam penelitian ini.

3.1.1 Input Data

Input data merupakan tahapan awal dari keseluruhan sistem yang digunakan sebagai sumber utama untuk analisis. Tahapan ini dilakukan dengan memasukkan citra digital berupa gambar lahan miring yang memiliki potensi rawan longsor. Citra yang digunakan menggunakan format JPG. Format ini merupakan salah satu format kompresi gambar standar yang tetap menjaga kualitas visual yang memadai untuk analisis tekstur.

Sebelum diproses dalam sistem, dilakukan pemotongan secara manual pada citra untuk menghilangkan objek tidak relevan yang tidak berkaitan langsung

dengan permukaan lahan sehingga area yang dianalisis benar-benar fokus pada permukaan lahan miring sebagai objek penelitian.

3.1.2 *Pre-processing*

Tahapan ini bertujuan untuk menyiapkan data agar sesuai dengan kebutuhan sebelum dilakukan penelitian. *Pre-processing* pada penelitian ini adalah dengan melakukan proses augmentasi, *resize*, serta pengkonversian ke *grayscale*.

a. **Augmentasi Data**

Salah satu langkah penting dalam *pre-processing* adalah augmentasi citra. Augmentasi citra merupakan tahapan penambahan variasi pada dataset dengan memodifikasi citra tanpa mengubah maknanya. Augmentasi pada citra digunakan untuk membantu memperkaya variasi pola tekstur yang dapat ditangkap oleh metode ekstraksi fitur (Kebaili et al., 2023). Augmentasi citra dilakukan terhadap seluruh data citra asli untuk meningkatkan keragaman data. Augmentasi dilakukan dengan memproses semua citra asli dengan tiga jenis augmentasi. Tiga jenis augmentasi yang digunakan dalam penelitian ini meliputi rotasi 15° (memutar citra ke arah sudut 15°), *horizontal flip* (membalikkan citra ke arah horizontal), dan *vertical flip* (membalikkan citra keatas atau bawah).

1) Augmentasi Rotasi 15°

Augmentasi ini dilakukan dengan menentukan titik pusat dari citra yang akan diputar. Titik pusat dihitung berdasarkan ukuran dimensi gambar dengan rumus dibawah ini.

$$(x_c, y_c) = \left(\frac{w}{2}, \frac{h}{2}\right) \quad (3.1)$$

Keterangan:

- a) w = *weight* atau lebar citra
- b) h = *height* atau tinggi citra

Tahapan selanjutnya adalah menghitung posisi baru setiap piksel setelah dilakukan rotasi. Posisi baru dari suatu piksel dapat dihitung dengan menerapkan transformasi rotasi 2D berikut ini.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \cdot \begin{bmatrix} x - x_c \\ y - y_c \end{bmatrix} + \begin{bmatrix} x_c \\ y_c \end{bmatrix} \quad (3.2)$$

Keterangan:

- a) (x, y) = koordinat piksel asli dalam citra
- b) (x', y') = koordinat piksel setelah rotasi
- c) (x_c, y_c) = koordinat titik pusat citra
- d) θ = sudut rotasi dalam radian
- e) $\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ = matriks rotasi untuk memutar koordinat terhadap sumbu pusat

Sudut rotasi dalam radian pada penelitian ini adalah -15° atau bisa juga dengan $-\frac{\pi}{12}$ karena rotasinya berlawanan dengan arah jarum jam.

Perhitungan posisi baru piksel umumnya menghasilkan bilangan pecahan, maka dilakukan proses interpolasi bilinier untuk menentukan nilai intensitas piksel pada posisi (x', y') . Interpolasi ini penting agar citra hasil rotasi tetap halus dan tidak mengalami distorsi tajam.

Sebagai konsekuensi dari rotasi citra yang bukan merupakan kelipatan 90 derajat, akan muncul area kosong di sekitar citra, khususnya pada bagian sudut. Area ini secara otomatis diisi dengan piksel berwarna hitam (bernilai nol) karena tidak ada nilai intensitas asli pada posisi tersebut. Hal ini terlihat pada hasil rotasi, dimana latar belakang hitam ini tidak berdampak

signifikan terhadap proses ekstraksi fitur karena analisis difokuskan pada area utama citra (Wangiyana et al., 2022).

2) *Augmentasi Horizontal Flip*

Augmentasi ini bekerja dengan membalik posisi setiap piksel pada citra secara horizontal. Dalam proses ini, piksel pada posisi paling kiri ditukar dengan piksel pada posisi paling kanan hingga seterusnya. Secara sistematis, operasi ini dapat dijelaskan sebagai berikut.

$$x' = w - 1 - x \quad (3.3)$$

Posisi piksel sepanjang sumbu horizontal (x) ditukar secara simetris, sementara posisi vertikalnya tetap. Hasil dari augmentasi ini adalah citra baru yang merupakan refleksi cermin dari citra asli terhadap sumbu vertikal (Awaluddin et al., 2023).

3) *Augmentasi Vertical Flip*

Teknik augmentasi ini membalik posisi piksel dari atas ke bawah, sehingga bagian atas citra berpindah ke bawah dan sebaliknya. Augmentasi ini dilakukan dengan menukar posisi piksel sepanjang sumbu vertikal (y) tanpa mengubah posisi horizontalnya. Transformasi ini dinyatakan secara matematis dengan rumus sebagai berikut.

$$y' = h - 1 - y \quad (3.4)$$

Operasi tersebut mencerminkan piksel terhadap garis horizontal di tengah citra (sumbu x tengah), dengan $h - 1$ digunakan karena indeks piksel dimulai dari nol. Dengan operasi ini, citra seolah dicerminkan terhadap

sumbu horizontal di tengah sehingga menghasilkan orientasi baru yang memperkaya data (Nguyen, 2023).

b. Resize

Dataset yang diperlukan untuk penelitian harus dipastikan keseragamannya, terutama dalam keseragaman ukuran sehingga proses mengubah ukuran harus dilakukan. Proses ini juga bertujuan untuk menstandarkan dimensi gambar serta meningkatkan performa model. Penelitian dari Agustin et al., (2022) menunjukkan bahwa gambar yang telah di-*resize* memiliki akurasi lebih besar dibandingkan dengan gambar yang belum diproses dalam penelitiannya yaitu deteksi COVID-19. Untuk itu, kita melakukan proses pengubahan ukuran semua citra dalam dataset menjadi ukuran yang ideal untuk penelitian. Proses *resize* dilakukan secara hati-hati pada citra. Setiap gambar dianalisis terlebih dahulu untuk menentukan proporsi yang tepat dan area penting dari tekstur permukaan lahan tidak terdistorsi secara signifikan saat diubah ukurannya. Hal ini digunakan untuk memastikan bahwa informasi visual pada citra tetap utuh, serta distribusi data antar gambar menjadi lebih merata dan menghindari dominasi area tertentu yang menjadikan penumpukan pola tidak seimbang pada hasil *clustering* nantinya.

c. Konversi ke Grayscale

Pada metode GLCM, gambar perlu dikonversi ke dalam format *grayscale* karena metode ini hanya dapat menganalisis hubungan antar tingkat keabuan dalam suatu citra. GLCM tidak dirancang untuk memproses gambar berwarna, sehingga diperlukan konversi ke *grayscale* (Iqbal et al., 2021). Konversi citra ke *grayscale* yaitu proses mengubah gambar yang memiliki tiga saluran warna (*Red*, *Green*,

Blue) menjadi satu saluran yaitu intensitas keabuan. Proses mengonversikan citra ke *grayscale* citra menggunakan rumus berikut ini.

$$Gray = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B \quad (3.5)$$

Keterangan:

- a) R = nilai piksel dengan komponen warna merah
- b) G = nilai piksel dengan komponen warna hijau
- c) B = nilai piksel dengan komponen warna biru

Koefisien 0.2989, 0.5870, dan 0,1140 digunakan sebagai bobot yang menunjukkan sensitivitas mata manusia dalam menentukan tingkat kecerahan terhadap masing-masing warna.

3.1.3 Proses GLCM (*Gray Level Co-Occurrence Matrix*)

Proses ekstraksi fitur tekstur dari citra menggunakan GLCM memiliki tahapan berikut ini.

a. Perhitungan Matriks GLCM

Pada tahapan ini, dilakukan perhitungan berdasarkan jarak dan sudut yang sudah ditentukan untuk mendapatkan hubungan spasial antar piksel dalam citra berdasarkan nilai *grayscale* dan parameter jarak dan arah sudut. Dari matriks ini, Kalkulasi dilakukan dengan menghitung berapa kali pasangan piksel dengan intensitas tertentu muncul dalam citra. Berikut merupakan langkah-langkah untuk pembuatan matriks GLCM.

1. Pembuatan *Framework Matrix*: *Framework Matrix* dihitung dari *quantization level* berdasarkan jumlah *gray tone*. Pada citra dengan ukuran 8-bit akan memiliki *quantization level* 256 karena nilai *gray tone* diantara 0-255.

2. Pembuatan *Co-Occurrence Matrix*: pembuatan matriks ini yaitu mengisi *framework matrix* dengan mengombinasikan piksel berpasangan dari kiri-atas hingga kanan bawah.
3. Pembuatan *Symmetric Matrix*: perhitungannya dengan menjumlahkan GLCM matriks dengan hasil transposenya.
4. Normalisasi dengan *Matrix Normalization*: normalisasi dilakukan dengan formula 3.6 berikut.

$$glcm_{norm} = \frac{glcm_{value}}{\sum_i^N glcm_{value}} \quad (3.6)$$

Keterangan:

- a) $glcm_{value}$ = nilai elemen tertentu dalam matriks
- b) $\sum_i^N glcm_{value}$ = jumlah seluruh nilai elemen

Hasil dari perhitungan tersebut baru bisa dilakukan ekstraksi menjadi beberapa fitur GLCM (Faisal et al., 2022).

b. Ekstraksi Hasil Menjadi Beberapa Fitur

Setelah mendapatkan matriks GLCM, dilakukan ekstraksi fitur tekstur berdasarkan nilai-nilai matriks. Fitur-fitur yang digunakan, yaitu kontras, *dissimilarity*, *homogeneity*, *energy*, korelasi, dan *entropy*.

c. Hasil Ekstraksi Fitur GLCM

Matriks GLCM yang sudah diekstraksi fitur tekstur menghasilkan data berupa nilai numerik yang menggambarkan karakteristik tekstur citra. Nilai-nilai ini digunakan untuk analisis lebih lanjut (Dwaich & Abdulbaqi, 2021). Hasil ekstraksi GLCM dikonversi menjadi *spreadsheet* untuk memudahkan dalam penelitian menggunakan PCA.

3.1.4 Proses PCA (*Principal Component Analysis*)

Setelah hasil fitur GLCM disimpan dalam *spreadsheet*, dilakukan analisis data menggunakan PCA untuk menentukan tingkat kerawanan tanah longsor pada citra lahan miring tersebut.

a. Normalisasi Data

Nilai-nilai numerik dari hasil ekstraksi fitur GLCM perlu untuk dinormalisasi (Polyakova & Krylov, 2022). Normalisasi data ini menggunakan *Standard Scale* sampai setiap fitur dalam dataset menjadi *z-score* dengan rata-rata 0 dan standar deviasi 1 untuk menghindari dominasi fitur dengan skala besar terhadap fitur lainnya.

b. Perhitungan *Variance Ratio* dan *Eigen Value*

Setelah data dinormalisasi, langkah selanjutnya dalam proses PCA adalah melakukan perhitungan *Variance Ratio* dan *Eigen Value* dari tiap komponen. *Variance Ratio* dikalkulasikan untuk menentukan jumlah komponen utama yang optimal, yaitu dengan memilih komponen yang memiliki kontribusi varians terbesar sampai batas tertentu sehingga dimensi data dapat dikurangi tanpa kehilangan informasi yang signifikan. Sedangkan *Eigen Value* menunjukkan seberapa besar variansi yang dapat dijelaskan oleh masing-masing komponen utama (Badri & Sari, 2021).

c. Pembentukan Komponen PCA

PCA membentuk nilai *principal component* dengan mempertahankan sejumlah kecil komponen yang menjelaskan sebagian besar varians dalam data dari nilai eigen yang sudah diketahui (Karakuzulu et al., 2023).

d. Pembuatan Grafik 1D Berdasarkan Komponen Utama

Setelah penentuan komponen utama, data divisualisasikan dalam bentuk grafik 1D untuk memudahkan dalam mengidentifikasi pola *clustering*. Pembuatan grafik ini digunakan agar struktur internal data terlihat.

e. Clustering

Tahapan ini memiliki tujuan agar dapat mengelompokkan data dari grafik berdasarkan kemiripan pola distribusi data setiap titiknya (Troccoli et al., 2022). Setiap titik dari hasil visualisasi grafik tersebut mempresentasikan nilai hasil ekstraksi fitur pada citra lahan miring yang dikelompokkan menjadi tiga klaster yang sesuai dengan tiga tingkat kerawanan longsor.

f. Hasil dan Evaluasi

Hasil *clustering* yang terbentuk tidak dapat langsung dianggap optimal tanpa melalui proses evaluasi. Evaluasi perlu dilakukan untuk menilai kualitas dan validitas pembentukan klaster serta menilai sejauh mana hasil *clustering* ini mencerminkan pemisahan data yang bermakna (Hendrastuty, 2024). Evaluasi juga menilai kesesuaian hasil klasterisasi dengan tujuan penelitian, yaitu membagi data menjadi tiga klaster untuk merepresentasikan tiga tingkat kerawanan longsor.

3.2 Data Preparation

Data adalah komponen yang penting dalam sebuah penelitian, karena kualitas dan relevansi data sangat memengaruhi hasil akhir penelitian (Condon et al., 2022). Data untuk penelitian ini adalah 50 citra lahan miring yang diperoleh secara langsung oleh peneliti melalui proses pengambilan gambar di beberapa

wilayah Kota Batu. Pengambilan data didasarkan pada peta ancaman tanah longsor tahun 2024 yang telah tersedia dari BPBD Kota Batu, sehingga lokasi-lokasi yang dipilih mempresentasikan tingkat kerawanan yang bervariasi.

Citra tersebut mewakili tempat-tempat yang memiliki karakteristik berbeda. Hal ini dapat memudahkan dalam kegiatan analisis karena memberikan keragaman tekstur permukaan tanah. Citra yang digunakan merupakan citra berwarna (RGB) dengan format *Joint Photographic Experts Group* (JPG). Resolusi dan sudut pengambilan pada data citra tersebut bervariasi tergantung pada kondisi lapangan saat pengambilan gambar.

3.3 Skenario Pengujian

Skenario pengujian dalam penelitian ini bertujuan untuk mengevaluasi keakuratan dalam metode ekstraksi fitur yang digunakan, yaitu kombinasi GLCM-PCA, serta validasi hasil menggunakan *Silhouette Score*. Pengujian dilakukan dengan menggunakan hasil dari penelitian *clustering* tingkat kerawanan tanah longsor di daerah Kota Batu. Hasil dari skenario pengujian ini akan digunakan untuk menentukan sejauh mana metode ekstraksi fitur GLCM-PCA dapat digunakan dalam membedakan karakteristik lahan yang berpotensi mengalami longsor.

Silhouette Score merupakan metrik yang digunakan untuk mengukur seberapa baik suatu data dikelompokkan dalam klaster tertentu (Hasan, 2024). Berikut ini merupakan tahapan dalam menghitung *Silhouette Score*.

1. Menghitung rata-rata jarak dalam klaster yang sama: nilai ini menggambarkan kepadatan klaster, semakin kecil nilainya, semakin dekat

titik dengan klaster lainnya.

2. Menghitung rata-rata jarak pada klaster terdekat: perhitungan ini menunjukkan sejauh mana titik data lebih dekat ke klaster lain dibandingkan dengan klaster asalnya.
3. Menghitung *Silhouette Score* pada semua titik: jika hasil dari nilai ini menunjukkan angka positif, maka menunjukkan bahwa titik berada pada klaster yang sesuai. Jika menghasilkan nilai negatif, maka menunjukkan titik tidak sesuai dengan klasternya.

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (3.7)$$

Keterangan:

- a) $a(i)$ = Mean jarak dari titik data i ke semua titik pada klaster yang sama
- b) $b(i)$ = Mean jarak dari titik data i ke semua titik pada klaster terdekat

4. Menghitung *Silhouette Score* rata-rata: untuk menghitung nilai ini menggunakan formula 3.8.

$$S_{avg} = \frac{1}{N} \sum_{i=1}^N S(i) \quad (3.8)$$

Nilai tersebut dianalisis untuk mengukur seberapa baik objek tersebut cocok dengan *cluster*-nya sendiri. Hasil nilai ini berkisar diantara -1 hingga 1. Jika nilai mendekati 1, maka menunjukkan klasterisasi yang sangat baik. Jika nilai semakin menurun dari angka 1, maka klasterisasi dinilai semakin buruk (Hasan, 2024).

Pengujian dilakukan dengan mengevaluasi kualitas hasil *clustering* berdasarkan jumlah klaster yang telah ditentukan, yaitu sebanyak tiga klaster pada masing-masing data augmentasi dan kedua parameter jarak GLCM. Meskipun sudah ditentukan, diperlukan pengujian lebih lanjut agar memastikan bahwa

pembagian ini sesuai dengan struktur data. Hasil terbaik dipilih berdasarkan nilai *Silhouette Score* tertinggi sebagai representasi paling optimal. Selain itu, dilakukan evaluasi tambahan dengan membandingkan hasil klasterisasi menggunakan dua klaster dan empat klaster untuk menilai sejauh mana jumlah klaster memengaruhi kualitas pemisahan data. Evaluasi ini bertujuan untuk memastikan bahwa pembagian yang dihasilkan mampu merepresentasikan pola tekstur secara optimal.

BAB IV

HASIL DAN PEMBAHASAN

Hasil implementasi dan pengujian sistem disajikan secara sistematis berdasarkan tahapan metode yang telah dijelaskan sebelumnya. Data yang dihasilkan ditampilkan dalam bentuk tabel, grafik, maupun visualisasi kluster untuk menunjukkan kinerja metode yang digunakan. Pembahasan dilakukan dengan menganalisis hubungan antara hasil yang diperoleh, teori yang mendasari, serta temuan dari penelitian terdahulu sehingga diperoleh pemahaman yang lebih komprehensif mengenai pendekatan yang diterapkan dalam penelitian.

4.1 Langkah-langkah Uji Coba

Pengujian ini dilakukan sebagai bentuk untuk mengevaluasi kualitas dari metode *Gray Level Co-Occurrence Matrix* dan *Principal Component Analysis* dalam melakukan penelitian untuk mengelompokkan atau melakukan *clustering* tingkat kerawanan tanah longsor menggunakan data citra lahan miring. Data pada penelitian ini diperoleh dari proses pengambilan gambar secara langsung oleh peneliti di berbagai wilayah Kota Batu. Data tersebut mempresentasikan berbagai karakteristik lahan untuk berbagai tingkat potensi longsor sehingga dapat menjadi acuan dalam analisis pada penelitian ini. Evaluasi pada hasil penelitian ini dilakukan menggunakan *Silhouette Score* pada masing-masing kluster yang dihasilkan agar menghasilkan kualitas yang baik serta pemisahan kluster yang optimal. Selain itu, pengujian dilakukan dengan membandingkan penggunaan parameter jarak GLCM sebesar 0.5 dan 0.8 untuk mengetahui pengaruhnya

terhadap kualitas fitur dan pengelompokkan. Penelitian ini juga melakukan perbandingan beberapa teknik augmentasi, yaitu rotasi 15° , *horizontal flip*, dan *vertical flip* untuk menentukan metode augmentasi yang paling optimal ketika dikombinasikan dengan jarak GLCM yang menghasilkan performa terbaik.

4.2 Hasil Uji Coba

Pada bagian ini, dilakukan pengujian terhadap alur kerja sistem yang telah dirancang untuk mengelompokkan citra menggunakan metode GLCM dan PCA. Proses penelitian ini terdiri dari beberapa tahapan utama yang saling terintegrasi, mulai dari tahapan pra-pemrosesan citra sampai tahap reduksi dimensi dan clustering.

4.2.1 Input Data

Citra yang dimasukkan berasal dari dokumentasi lapangan dalam format RGB. Jumlah citra asli sebelum dilakukan proses augmentasi berjumlah 50 gambar. Setiap gambar sudah divalidasi oleh *surveyor*. Citra digunakan sebagai input data untuk dibaca pada program menggunakan fungsi *imread()* dari OpenCV agar dapat membaca file citra dari *path* yang ditentukan.

4.2.2 Pre-processing

Tahapan ini berisi proses yang dilakukan pada data untuk menyiapkannya sebelum dilakukan penelitian inti.

a. Augmentasi Data

Tahapan pertama dalam proses ini yaitu augmentasi data. Proses ini dilakukan dengan tiga jenis augmentasi yang tidak mempengaruhi tekstur pada

citra. Augmentasi yang dilakukan adalah rotasi 15° , *horizontal flip*, dan *vertical flip* pada semua citra. Jumlah data yang digunakan adalah 150 data augmentasi dari 50 data citra asli.



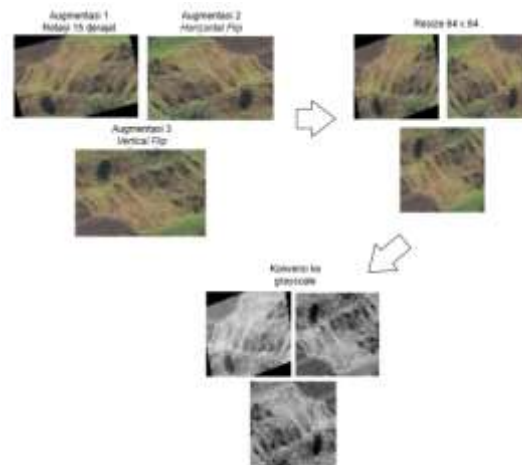
Gambar 4.1 Dataset Augmentasi Data

b. *Resize*

Tahapan selanjutnya adalah melakukan *resize*. Proses *resize* dilakukan dengan mengubah gambar untuk memiliki format ukuran yang sama, yaitu 64x64. Dengan tahapan ini, proses perhitungan GLCM dapat dilakukan secara efisien dan akurat, karena citra memiliki format yang standar dan sesuai dengan analisis hubungan spasial antar piksel. Ukuran ini juga digunakan untuk meningkatkan efisiensi komputasi dan tetap menjaga detail tekstur yang ada pada citra.

c. *Konversi ke Grayscale*

Gambar kemudian dikonversi ke skala keabuan untuk menyiapkannya dalam perhitungan metode GLCM. Konversi ini digunakan untuk menyederhanakan informasi visual dan menghilangkan informasi warna yang tidak diperlukan dalam analisis tekstur, sekaligus mempercepat proses ekstraksi fitur dengan mengandalkan nilai intensitas piksel dalam gambar.



Gambar 4.2 Hasil Tahapan Pre-processing pada Citra

Gambar 4.2 menampilkan hasil dari tahapan *pre-processing* yang mengubah ukuran citra dari ukuran asli 261x174 piksel menjadi ukuran 64x64 piksel dan dikonversi menjadi format *grayscale* untuk menyiapkannya dalam perhitungan matriks *co-occurrence* dalam metode GLCM. Konversi ini bertujuan agar menghilangkan komponen warna RGB, sehingga pada setiap piksel hanya memiliki satu nilai intensitas keabuan. Hal ini perlu dilakukan karena metode GLCM tidak memerlukan informasi warna, melainkan hanya melakukan proses dari nilai intensitas piksel dalam skala keabuan.

Citra *grayscale* memiliki rentang intensitas piksel dari 0 hingga 255. Nilai intensitas 0 mewakili warna hitam dan 255 mewakili warna putih. Nilai-nilai intensitas ini disusun dalam bentuk matriks dua dimensi yang menunjukkan distribusi intensitas keabuan dari seluruh piksel dalam citra. Matriks ini menjadi dasar dalam pembentukan matriks GLCM.



Gambar 4.3 Matriks Nilai dari Hasil Konversi *Grayscale Image*

4.2.3 Implementasi *Gray Level Co-Occurrence Matrix*

Metode GLCM pada penelitian ini digunakan sebagai teknik utama dalam ekstraksi fitur tekstur permukaan lahan miring. GLCM mampu menggambarkan karakteristik visual seperti tingkat kekasaran, kontras, homogenitas, energi, korelasi dan entropi dari permukaan lahan dengan menganalisis keterkaitan nilai intensitas piksel dalam citra *grayscale* sesuai pada parameter jarak dan sudut tertentu. Berikut merupakan tahapan-tahapan dalam proses implementasi GLCM.

a. Perhitungan Matriks GLCM

Nilai matriks pada gambar yang telah dikonversi ke dalam format *grayscale* akan dilakukan perhitungan untuk menentukan intensitas kemunculan pasangan piksel berdasarkan jarak dan arah piksel pasangannya terhadap piksel utama. Perhitungan ini menggunakan fungsi ***graycomatrix()*** dari pustaka *scikit-image*. Hasil dari perhitungan ini berupa matriks dua dimensi yang disebut matriks GLCM. Matriks ini berbentuk persegi, di mana setiap baris merepresentasikan nilai intensitas piksel utama, dan setiap kolom merepresentasikan nilai intensitas piksel pasangannya. Semua gambar yang digunakan memiliki ukuran 256x256 piksel, sehingga perhitungan GLCM dilakukan dengan ukuran matriks 256x256. Nilai pada posisi tertentu dalam matriks, misalnya (i, j), menunjukkan frekuensi kemunculan pasangan piksel dengan nilai intensitas i dan j dalam gambar, sesuai

intensitas piksel untuk melihat tingkat kekasaran tekstur. Energi menjumlahkan kuadrat dari setiap elemen dalam matriks GLCM untuk mengukur tingkat konsistensi pola dalam tekstur. *Homogeneity* memberikan nilai tinggi ketika banyak pasangan piksel yang nilai intensitasnya mirip sehingga menandakan tekstur yang seragam. *Dissimilarity* menghitung selisih nilai pasangan piksel namun menggunakan selisih absolut sehingga lebih sensitif dalam perubahan kecil antar piksel. Korelasi mengukur kesamaan perubahan dari hubungan linier antara nilai intensitas piksel pada pasangan piksel untuk melihat tingkat keteraturan pola tekstur. Lalu entropi menghitung tingkat ketidakteraturan tekstur dari tersebarnya nilai GLCM secara merata dan tidak ada dominasi nilai tertentu.

c. Hasil GLCM fitur dan Konversi ke *Spreadsheet*

Setelah dilakukan ekstraksi fitur, hasil per hitungan dari masing-masing kombinasi jarak dan sudut disimpan pada *file Comma Separated Values (CSV)* yang berbeda. Tahapan ini menggunakan fungsi **df.to_csv()** dari library **pandas** untuk mengubah tampilan *dataframe* dari hasil ekstraksi fitur ke dalam *file CSV*. Penyimpanan pada format ini dilakukan agar memudahkan dalam analisis data lanjutan. Setiap *file CSV* berisi nilai-nilai dari enam fitur, yaitu kontras, *dissimilarity*, *homogeneity*, energi, korelasi, dan entropi yang disusun berdasarkan variasi arah sudut dan nilai jarak GLCM seperti yang ditunjukkan pada Gambar 4.4.

--- Metriks Fitur untuk Distance = 0.5 ---						
	Contrast	Energy	Homogeneity	Dissimilarity	Correlation	Entropy
Sudut:						
0	415.574206	0.003375	0.113990	13.317877	0.001168	11.345458
45	0.000000	0.014355	1.000000	0.000000	1.000000	6.552094
90	336.888952	0.003266	0.223375	11.757448	0.303949	11.252979
135	0.000000	0.014355	1.000000	0.000000	1.000000	6.552094
180	415.574206	0.003375	0.113990	13.317877	0.001168	11.345458
225	0.000000	0.014355	1.000000	0.000000	1.000000	6.552094
270	336.888952	0.003266	0.223375	11.757448	0.303949	11.252979
315	0.000000	0.014355	1.000000	0.000000	1.000000	6.552094
--- Metriks Fitur untuk Distance = 0.8 ---						
	Contrast	Energy	Homogeneity	Dissimilarity	Correlation	Entropy
Sudut:						
0	415.574206	0.003375	0.113990	13.317877	0.001168	11.345458
45	621.231544	0.003368	0.105852	16.564676	0.516574	11.534352
90	336.888952	0.003266	0.223375	11.757448	0.303949	11.252979
135	437.009443	0.002457	0.207277	12.284732	0.078944	11.404467
180	415.574206	0.003375	0.113990	13.317877	0.001168	11.345458
225	621.231544	0.003368	0.105852	16.564676	0.516574	11.534352
270	336.888952	0.003266	0.223375	11.757448	0.303949	11.252979
315	437.009443	0.002457	0.207277	12.284732	0.078944	11.404467

Augmentasi 1
Rotasi 15 derajat

--- Metriks Fitur untuk Distance = 0.5 ---						
	Contrast	Energy	Homogeneity	Dissimilarity	Correlation	Entropy
Sudut:						
0	262.376288	0.000391	0.172296	11.021677	0.778616	11.623658
45	0.000000	0.012015	1.000000	0.000000	1.000000	6.544132
90	256.826298	0.000391	0.177704	11.587382	0.776975	11.614431
135	0.000000	0.012015	1.000000	0.000000	1.000000	6.544132
180	262.376288	0.000391	0.172296	11.021677	0.778616	11.623658
225	0.000000	0.012015	1.000000	0.000000	1.000000	6.544132
270	256.826298	0.000391	0.177704	11.587382	0.776975	11.614431
315	0.000000	0.012015	1.000000	0.000000	1.000000	6.544132
--- Metriks Fitur untuk Distance = 0.8 ---						
	Contrast	Energy	Homogeneity	Dissimilarity	Correlation	Entropy
Sudut:						
0	262.376288	0.000391	0.172296	11.021677	0.778616	11.623658
45	287.346687	0.000379	0.168573	12.463845	0.746934	11.654023
90	256.026290	0.000391	0.177704	11.587382	0.776975	11.614431
135	384.337113	0.000391	0.151796	14.426384	0.663762	11.739344
180	262.376288	0.000391	0.172296	11.021677	0.778616	11.623658
225	287.346687	0.000379	0.168573	12.463845	0.746934	11.654023
270	256.026290	0.000391	0.177704	11.587382	0.776975	11.614431
315	384.337113	0.000391	0.151796	14.426384	0.663762	11.739344

Augmentasi 2
Horizontal Flip

--- Metriks Fitur untuk Distance = 0.5 ---						
	Contrast	Energy	Homogeneity	Dissimilarity	Correlation	Entropy
Sudut:						
0	268.825770	0.000381	0.175556	11.758929	0.773724	11.642998
45	0.000000	0.011959	1.000000	0.000000	1.000000	6.548789
90	237.575387	0.000391	0.179888	11.624008	0.776696	11.614373
135	0.000000	0.011959	1.000000	0.000000	1.000000	6.548789
180	268.825770	0.000381	0.175556	11.758929	0.773724	11.642998
225	0.000000	0.011959	1.000000	0.000000	1.000000	6.548789
270	237.575387	0.000391	0.179888	11.624008	0.776696	11.614373
315	0.000000	0.011959	1.000000	0.000000	1.000000	6.548789
--- Metriks Fitur untuk Distance = 0.8 ---						
	Contrast	Energy	Homogeneity	Dissimilarity	Correlation	Entropy
Sudut:						
0	268.825770	0.000381	0.175556	11.758929	0.773724	11.642998
45	284.611743	0.000369	0.166786	12.454523	0.752444	11.682228
90	257.575387	0.000391	0.179888	11.624008	0.776696	11.614373
135	389.833762	0.000342	0.151496	14.523345	0.661205	11.774589
180	268.825770	0.000381	0.175556	11.758929	0.773724	11.642998
225	284.611743	0.000369	0.166786	12.454523	0.752444	11.682228
270	257.575387	0.000391	0.179888	11.624008	0.776696	11.614373
315	389.833762	0.000342	0.151496	14.523345	0.661205	11.774589

Augmentasi 3
Vertical Flip

Gambar 4.4 Hasil Ekstraksi Fitur dari Matriks GLCM

Gambar 4.4 menunjukkan hasil dari ekstraksi fitur tekstur pada masing-masing citra hasil augmentasi. Pada ketiga jenis augmentasi ini, fitur tekstur dihitung berdasarkan dua jarak piksel, yaitu 0.5 dan 0.8, serta delapan arah sudut. Masing-masing augmentasi citra, menghasilkan nilai-nilai fitur tekstur GLCM yang berbeda yang merepresentasikan karakteristik tekstur permukaan citra yang berbeda karena pengaruh teknik augmentasinya.

4.2.4 Implementasi *Principal Component Analysis*

PCA merupakan metode reduksi dimensi untuk mengurangi jumlah variabel pada dataset namun tetap mempertahankan sebanyak mungkin informasi penting yang terkandung dalam data asli. PCA digunakan untuk mengurangi fitur yang

banyak menjadi beberapa komponen utama. Hal ini memudahkan dalam visualisasi dan interpretasi data. Berikut merupakan tahapan dalam metode PCA untuk visualisasi *clustering*.

a. Normalisasi Data

Pada tahapan ini, data hasil ekstraksi fitur GLCM dalam file csv dibaca dilakukan normalisasi menggunakan fungsi ***StandardScaler()*** sehingga memiliki rata-rata 0 dan standar deviasi 1. Tahapan ini dilakukan agar menyeimbangkan data sehingga tidak ada data yang mendominasi. Proses ini memastikan semua fitur berkontribusi secara adil dalam tahapan selanjutnya, yaitu penentuan *principal component*.

Tabel 4.2 Hasil Normalisasi Ekstraksi Fitur GLCM

No	Jenis Augmentasi	Nama Fitur	Sudut (°)	Jarak	Nilai Fitur
1.	Rotasi 15°	Kontras	0	0.5	1,196976393
			45		-0,989142303
			90		0,781308213
			135		-0,989142303
			180		1,196976393
			225		-0,989142303
			270		0,781308213
			315		-0,989142303
			0	0.8	-0,353443525
			45		1,616877266
			90		-1,112681095
			135		-0,150752646
			180		-0,353443525
			225		1,616877266
			270		-1,112681095
			315		-0,150752646
		Energi	0	0.5	-0,990059074
			45		0,999975416
			90		-1,009891758
			135		0,999975416
			180		-0,990059074
			225		0,999975416

No	Jenis Augmentasi	Nama Fitur	Sudut (°)	Jarak	Nilai Fitur
			270		-1,009891758
			315		0,999975416
			0	0.8	1,113854501
			45		-1,090792963
			90		0,874255025
			135		-0,897316563
			180		1,113854501
			225		-1,090792963
			270		0,874255025
			315		-0,897316563
		<i>Homogeneity</i>	0	0.5	-1,014535677
			45		0,99994679
			90		-0,985357902
			135		0,99994679
			180		-1,014535677
			225		0,99994679
			270		-0,985357902
			315		0,99994679
			0	0.8	0,40757381
			45		-1,547264764
			90		1,198497847
			135		-0,058806893
			180		0,40757381
			225		-1,547264764
			270		1,198497847
			315		-0,058806893
		<i>Dissimilarity</i>	0	0.5	1,120823426
			45		-0,996103654
			90		0,871383881
			135		-0,996103654
			180		1,120823426
			225		-0,996103654
			270		0,871383881
			315		-0,996103654
			0	0.8	-0,246654815
			45		1,606456565
			90		-1,145694985
			135		-0,214106765
			180		-0,246654815
			225		1,606456565
			270		-1,145694985
			315		-0,214106765

No	Jenis Augmentasi	Nama Fitur	Sudut (°)	Jarak	Nilai Fitur
		Korelasi	0	0.5	-1,198633719
			45		0,988946758
			90		-0,779259796
			135		0,988946758
			180		-1,198633719
			225		0,988946758
			270		-0,779259796
			315		0,988946758
			0	0.8	0,405408384
			45		-1,607531991
			90		1,115325527
			135		0,08679808
			180		0,405408384
			225		-1,607531991
			270		1,115325527
			315		0,08679808
		Entropi	0	0.5	1,019547843
			45		-0,99990352
			90		0,980259198
			135		-0,99990352
			180		1,019547843
			225		-0,99990352
			270		0,980259198
			315		-0,99990352
			0	0.8	-0,380666246
			45		1,469896994
			90		-1,286668009
			135		0,197437261
			180		-0,380666246
			225		1,469896994
			270		-1,286668009
			315		0,197437261
2.	Horizontal Flip	Kontras	0	0.5	1,024322069
			45		-0,999850272
			90		0,975378474
			135		-0,999850272
			180		1,024322069
			225		-0,999850272
			270		0,975378474
			315		-0,999850272
			0	0.8	-0,682876638
			45		-0,197644964

No	Jenis Augmentasi	Nama Fitur	Sudut (°)	Jarak	Nilai Fitur
			90		-0,80612534
			135		1,686646941
			180		-0,682876638
			225		-0,197644964
			270		-0,80612534
			315		1,686646941
		Energi	0	0.5	-0,999970895
			45		1
			90		-1,000029104
			135		1
			180		-0,999970895
			225		1
			270		-1,000029104
			315		1
			0	0.8	0,806788885
			45		0,058586452
			90		0,786168781
			135		-1,651544117
			180		0,806788885
			225		0,058586452
			270		0,786168781
			315		-1,651544117
		Homogeneity	0	0.5	-1,006544704
			45		0,999989256
			90		-0,993433809
			135		0,999989256
			180		-1,006544704
			225		0,999989256
			270		-0,993433809
			315		0,999989256
			0	0.8	0,485891254
			45		0,101283078
			90		1,044557927
			135		-1,631732259
			180		0,485891254
			225		0,101283078
			270		1,044557927
			315		-1,631732259
		Dissimilarity	0	0.5	1,019922133
			45		-0,999899771
			90		0,97987741
			135		-0,999899771

No	Jenis Augmentasi	Nama Fitur	Sudut (°)	Jarak	Nilai Fitur
			180		1,019922133
			225		-0,999899771
			270		0,97987741
			315		-0,999899771
			0	0.8	-0,674764322
			45		-0,099396839
			90		-0,884758762
			135		1,658919923
			180		-0,674764322
			225		-0,099396839
			270		-0,884758762
			315		1,658919923
		Korelasi	0	0.5	-1,027908639
			45		0,999802492
			90		-0,971696344
			135		0,999802492
			180		-1,027908639
			225		0,999802492
			270		-0,971696344
			315		0,999802492
			0	0.8	0,674753407
			45		0,195775085
			90		0,815229662
			135		-1,685758154
			180		0,674753407
			225		0,195775085
			270		0,815229662
			315		-1,685758154
		Entropi	0	0.5	1,00181579
			45		-0,999999175
			90		0,99818256
			135		-0,999999175
			180		1,00181579
			225		-0,999999175
			270		0,99818256
			315		-0,999999175
			0	0.8	-0,694343957
			45		-0,077943777
			90		-0,881451704
			135		1,653739438
			180		-0,694343957
			225		-0,077943777

No	Jenis Augmentasi	Nama Fitur	Sudut (°)	Jarak	Nilai Fitur
3.	<i>Vertical Flip</i>		270		-0,881451704
			315		1,653739438
		Kontras	0	0.5	1,024322069
			45		-0,999850272
			90		0,975378474
			135		-0,999850272
			180		1,024322069
			225		-0,999850272
			270		0,975378474
			315		-0,999850272
			0	0.8	-0,682876638
			45		-0,197644964
			90		-0,80612534
			135		1,686646941
			180		-0,682876638
			225		-0,197644964
			270		-0,80612534
			315		1,686646941
		Energi	0	0.5	-0,999970895
			45		1
			90		-1,000029104
			135		1
			180		-0,999970895
			225		1
			270		-1,000029104
			315		1
			0	0.8	0,806788885
			45		0,058586452
			90		0,786168781
			135		-1,651544117
			180		0,806788885
			225		0,058586452
			270		0,786168781
			315		-1,651544117
		<i>Homogeneity</i>	0	0.5	-1,006544704
			45		0,999989256
			90		-0,993433809
			135		0,999989256
			180		-1,006544704
			225		0,999989256
			270		-0,993433809
			315		0,999989256

No	Jenis Augmentasi	Nama Fitur	Sudut (°)	Jarak	Nilai Fitur
			0	0.8	0,485891254
			45		0,101283078
			90		1,044557927
			135		-1,631732259
			180		0,485891254
			225		0,101283078
			270		1,044557927
			315		-1,631732259
		<i>Dissimilarity</i>	0	0.5	1,019922133
			45		-0,999899771
			90		0,97987741
			135		-0,999899771
			180		1,019922133
			225		-0,999899771
			270		0,97987741
			315		-0,999899771
			0	0.8	-0,674764322
			45		-0,099396839
			90		-0,884758762
			135		1,658919923
			180		-0,674764322
			225		-0,099396839
			270		-0,884758762
			315		1,658919923
		Korelasi	0	0.5	-1,027908639
			45		0,999802492
			90		-0,971696344
			135		0,999802492
			180		-1,027908639
			225		0,999802492
			270		-0,971696344
			315		0,999802492
			0	0.8	0,674753407
			45		0,195775085
			90		0,815229662
			135		-1,685758154
			180		0,674753407
			225		0,195775085
			270		0,815229662
			315		-1,685758154
		Entropi	0	0.5	1,00181579
			45		-0,999999175

No	Jenis Augmentasi	Nama Fitur	Sudut (°)	Jarak	Nilai Fitur
			90		0,99818256
			135		-0,999999175
			180		1,00181579
			225		-0,999999175
			270		0,99818256
			315		-0,999999175
			0	0.8	-0,694343957
			45		-0,077943777
			90		-0,881451704
			135		1,653739438
			180		-0,694343957
			225		-0,077943777
			270		-0,881451704
			315		1,653739438

Tabel 4.2 menunjukkan hasil dari data ekstraksi fitur GLCM pada masing-masing jenis augmentasi setelah dilakukan normalisasi untuk menyeimbangkan data.

b. Perhitungan Nilai *Variance Ratio* dan *Eigen Value*

Setelah dilakukan normalisasi pada data, dilanjutkan dengan menghitung nilai *Variance Ratio* dan *Eigen Value* yang digunakan untuk menentukan jumlah komponen utama yang digunakan. Perhitungan rasio varians dilakukan dengan perbandingan total keseluruhan *Eigen Value*, yang menunjukkan seberapa besar kontribusi informasi dari masing-masing komponen. Semakin tinggi nilai *Eigen Value* dan rasio varians pada suatu komponen, maka semakin penting komponen tersebut untuk analisis.

Komponen-komponen yang digunakan di dalam PCA adalah jumlah fitur dari hasil ekstraksi GLCM sehingga dari hasil uji coba menunjukkan bahwa jarak dari parameter metode tersebut mempengaruhi nilai rasio varians dan *Eigen Value*

masing-masing komponen. Jika menggunakan jarak yang berbeda, maka nilai keduanya juga berbeda.

Tabel 4.3 Persentase Nilai *Variance Ratio* dan *Eigen Value* Rotasi 15°

Jarak GLCM	Principal Component	Nilai Rasio Varians (%)	Eigen Value (%)
0.5	PC1	99.58	6.83
	PC2	0.42	0.03
	PC3	0.00	0.00
	PC4	0.00	0.00
	PC5	0.00	0.00
	PC6	0.00	0.00
0.8	PC1	93.51	6.41
	PC2	6.23	0.43
	PC3	0.26	0.02
	PC4	0.00	0.00
	PC5	0.00	0.00
	PC6	0.00	0.00

Tabel 4.3 memperlihatkan persentasi dari nilai rasio varians dan *Eigen Value* pada masing-masing komponen utama berdasarkan dua jarak GLCM yaitu 0.5 dan 0.8. Pada jarak 0.5, komponen pertama (PC1) memiliki *variens ratio* sebesar 99.85% dan *Eigen Value* sebesar 6.83 yang menunjukkan bahwa hampir seluruh informasi dalam data dapat dijelaskan **oleh** komponen ini. Komponen kedua hanya memiliki 0.42% varians, sedangkan komponen lainnya tidak memberikan kontribusi berarti atau memiliki nilai 0%.

Pada jarak GLCM 08, hasil yang diperoleh yaitu PC1 memiliki nilai *Variance Ratio* sebesar 93.51% dan *Eigen Value* sebesar 6.41 sehingga masih mendominasi. Terdapat sedikit kontribusi dari PC2 dan PC3 dengan nilai *Variance Ratio* masing-masing yaitu sebesar 3.21% dan 0.26%.

Tabel 4.4 Persentase Nilai *Variance Ratio* dan *Eigen Value* Horizontal Flip

Jarak GLCM	Principal Component	Nilai Rasio Varians (%)	Eigen Value (%)
0.5	PC1	99.99	6.86
	PC2	0.01	0.03
	PC3	0.00	0.00
	PC4	0.00	0.00

Jarak GLCM	Principal Component	Nilai Rasio Varians (%)	Eigen Value (%)
0.8	PC5	0.00	0.00
	PC6	0.00	0.00
	PC1	99.51	6.82
	PC2	6.39	0.03
	PC3	0.10	0.01
	PC4	0.00	0.00
	PC5	0.00	0.00
	PC6	0.00	0.00

Pada Tabel 4.4, terlihat bahwa PC1 mendominasi nilai rasio varians dan nilai eigen pada semua parameter jarak GLCM yang digunakan. Pada jarak 0.5, Nilai rasio PC1 sebesar 99.99% dengan nilai eigennya 6.86%. Komponen kedua hanya memiliki nilai rasio varians 0.01% dengan nilai eigen 0.03%, dan sisanya tidak memberikan kontribusi yang signifikan. Sementara itu, pada jarak 0.8 nilai *Variance Ratio* pada PC1 sedikit turun menjadi 99.51% dengan nilai eigen 6.82%. Komponen PC2 menyumbang sekitar 0.39% dan PC3 sebesar 0.10%. Berdasarkan hasil ini, proses augmentasi *horizontal flip* tidak secara signifikan mengubah distribusi informasi dalam data, dan informasi masih terpusat pada PC1.

Tabel 4.5 Persentase Nilai *Variance Ratio* dan *Eigen Value Vertical Flip*

Jarak GLCM	Principal Component	Nilai Rasio Varians (%)	Eigen Value (%)
0.5	PC1	100	6.86
	PC2	0.00	0.00
	PC3	0.00	0.00
	PC4	0.00	0.00
	PC5	0.00	0.00
	PC6	0.00	0.00
0.8	PC1	98.56	6.76
	PC2	1.35	0.09
	PC3	0.09	0.01
	PC4	0.00	0.00
	PC5	0.00	0.00
	PC6	0.00	0.00

Tabel 4.5 memperlihatkan nilai rasio varians dan nilai eigen pada masing-masing komponen pada augmentasi *vertical flip*. Pada jarak 0.5, nilai *variance ratio*

sebesar 100% terkonsentrasi sepenuhnya pada PC1 dengan nilai eigen sebesar 6.86%, sedangkan PC2 dan komponen lainnya tidak menyumbang varians data sama sekali. Hal ini menunjukkan bahwa dalam data hasil augmentasi ini dapat direpresentasikan secara penuh oleh satu komponen saja.

Sementara itu, terjadi distribusi varians yang sedikit lebih merata pada hasil persebaran informasi pada jarak GLCM 0.8. PC1 masih mendominasi dengan nilai 98.56% dan nilai eigen 6.76%. PC2 memiliki nilai rasio varians 1.35% dan PC3 bernilai 0.09%.

Berdasarkan hasil dari nilai varians rasio dan nilai eigen pada masing-masing augmentasi, dapat dilihat bahwa ketiga jenis augmentasi tersebut tidak mengubah secara drastis struktur varians data, terutama pada jarak GLCM 0.5 yang menunjukkan dominasi absolut PC1. Pada jarak 0.8 terjadi sedikit penyebaran informasi pada komponen lainnya yang menunjukkan adanya sedikit perubahan pola tekstur sebagai pengaruh dari augmentasi.

c. **Pembentukan Komponen Utama / *Principal Component***

Tahapan ini merupakan langkah penting yang menentukan kualitas dan efektivitas analisis. Komponen yang dipilih harus cukup untuk merepresentasikan sebagian besar variasi data namun tetap mengurangi kompleksitas. Komponen utama dipilih berdasarkan nilai rasio varians dan nilai eigennya karena kedua nilai tersebut menjelaskan seberapa banyak informasi yang ada di dalam komponen.

Berdasarkan hasil analisis perhitungan nilai eigen dan rasio varians pada masing-masing jenis augmentasi, kedua jarak GLCM menghasilkan PC1 yang memiliki nilai rasio varians dan nilai eigen yang besar. Hal ini menunjukkan bahwa

PC1 menghasilkan penyebaran data yang lebih terpisah dan terstruktur dibandingkan komponen lain sehingga dapat dijadikan indikator utama dalam proses visualisasi dan analisis kluster. Oleh karena itu, penelitian ini menggunakan PC1 sebagai komponen utama untuk melakukan analisis *clustering* data menggunakan visualisasi grafik. Selain itu, berdasarkan Elbow Method, komponen-komponen utama yang dipilih adalah yang memiliki nilai eigen lebih dari satu (Zhuang et al., 2022). Hal ini menegaskan bahwa PC1 memang merupakan komponen dengan kontribusi signifikan terhadap varians total data, sehingga pemilihan komponen utama tersebut sudah tepat untuk mewakili karakteristik utama data dalam analisis *clustering*.

Selain mempertimbangkan dari nilai eigen dan rasio varians, proses interpretasi komponen utama mengacu pada *Eigen Vector*. *Eigen Vector* menunjukkan kontribusi relatif dari masing-masing fitur asli yang diproyeksikan ke dalam satu dimensi, yaitu PC1. Masing-masing fitur memiliki nilai *Eigen Vector* yang merepresentasikan arah (positif atau negatif) dan besar pengaruhnya terhadap PC1. Nilai-nilai ini memungkinkan peneliti untuk memahami karakteristik utama yang diwakili oleh PC1. PC1 lebih dipengaruhi oleh fitur-fitur yang mencerminkan kekasaran tekstur atau oleh fitur-fitur yang menggambarkan homogenitas.

Tabel 4.6 *Eigen Vector* dari PC1

Jenis Augmentasi	Jarak GLCM	Nilai <i>Eigen Vector</i> PC1					
		Kontras	Energi	Homogeneity	Dissimilarity	Korelasi	Entropi
Rotasi	0.5	0.407778	-0.40798	-0.40844	0.409011	-0.40773	0.408529
	0.8	0.417595	-0.35202	-0.42173	0.413845	-0.41971	0.420012
Horizontal Flip	0.5	0.408249	-0.40824	-0.40825	0.408257	-0.40823	0.408247
	0.8	0.408719	-0.40773	-0.40588	0.409222	-0.40879	0.409122
Vertical Flip	0.5	0.40825	-0.40824	-0.40824	0.408249	-0.40824	0.40825
	0.8	0.407226	-0.40658	-0.40860	0.410282	-0.40749	0.409289

Tabel 4.6 menunjukkan nilai *Eigen Vector* dari komponen utama, yaitu PC1. Nilai-nilai eigen menunjukkan arah vektor dari masing-masing tekstur terhadap PC1. Setiap nilai menggambarkan kontribusi arah fitur dalam pembentukan PC1 melalui kombinasi linier dari seluruh fitur. Arah dari bobot positif dan negatif pada nilai fitur tersebut dijadikan acuan dalam interpretasi karakteristik data.

Fitur dengan nilai *Eigen Vector* positif menunjukkan keberadaan nilai fitur yang tinggi akan menggeser posisi titik data ke arah positif pada sumbu PC1. Sebaliknya, fitur dengan nilai *Eigen Vector* negatif menunjukkan bahwa nilai fitur yang tinggi akan menggeser ke arah negatif pada PC1.

Dalam penelitian ini, beberapa fitur seperti energi, *homogeneity*, dan korelasi cenderung memiliki nilai *Eigen Vector* yang negatif, yang umumnya merepresentasikan tekstur yang halus, teratur, dan homogen. Jika suatu titik data hasil proyeksi PCA memiliki nilai PC1 yang cenderung negatif, hal ini mengindikasikan bahwa citra pada arah sudut tertentu memiliki ciri-ciri tekstur yang cenderung lebih seragam dan tidak kasar.

Fitur yang cenderung bernilai positif seperti kontras, *dissimilarity*, dan entropi menunjukkan arah berlawanan, yaitu mengarah pada tekstur yang kontras, tidak homogen dan beragam, sehingga dikaitkan dengan permukaan lahan yang lebih tidak stabil. Oleh karena itu, saat titik-titik data hasil PCA diklaster, kelompok dengan titik-titik yang berada di sisi negatif PC1 dapat diinterpretasikan sebagai area dengan ciri tekstur lebih halus dan seragam, yang dalam konteks ini dapat dihubungkan dengan tingkat kerawanan longsor yang lebih rendah. Sebaliknya, kelompok yang berada di sisi positif PC1 menunjukkan arah fitur yang berkaitan

dengan keragaman dan ketidakteraturan tekstur, yang dapat dipresentasikan sebagai tingkat kerawanan yang lebih tinggi.

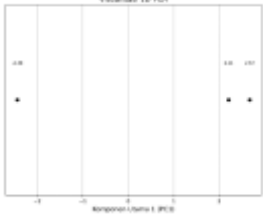
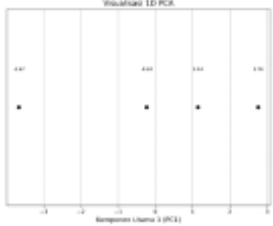
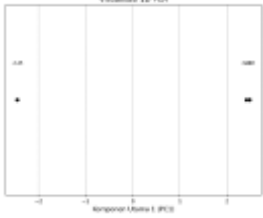
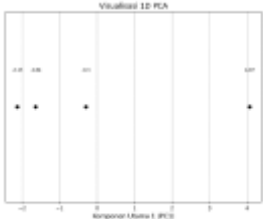
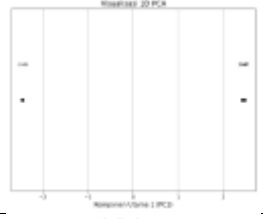
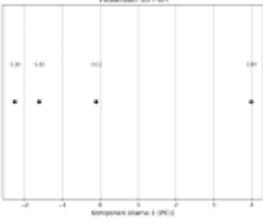
Dengan memahami komposisi PC1 melalui *Eigen Vector*, interpretasi terhadap nilai PC1 lebih bermakna. Nilai PC1 yang tinggi atau rendah tidak hanya mencerminkan proyeksi numerik, tetapi juga memberikan gambaran tentang bagaimana kombinasi fitur-fitur GLCM membentuk representasi tekstur dari citra. Oleh karena itu, analisis *Eigen Vector* berperan penting dalam menjelaskan dimensi baru hasil reduksi, serta mendukung dalam proses analisis.

d. Visualisasi Grafik

Untuk mempermudah dalam proses *clustering*, komponen utama yang telah ditentukan divisualisasikan dalam bentuk grafik 1D dari data komponen utama yang telah dinormalisasi sehingga akan membentuk pola *clustering*. Grafik 1D menampilkan data normalisasi dari PC1. Visualisasi ini digunakan untuk membantu dalam mengamati persebaran data serta melihat kemungkinan terbentuknya klaster yang memiliki karakteristik serupa berdasarkan komponen utama tersebut sehingga proses identifikasi pola dan pemisahan antar klaster menjadi mudah dianalisis.

Visualisasi ini menggunakan tampilan plot. Setiap titik plot mempresentasikan satu observasi. Titik-titik plot yang berdekatan merepresentasikan kemiripan data sehingga memudahkan dalam pengamatan pola *clustering*.

Tabel 4.7 Visualisasi PCA

Jenis Augmentasi	Jarak	Visualisasi Grafik
Rotasi 15°	0.5	
	0.8	
Horizontal Flip	0.5	
	0.8	
Vertical Flip	0.5	
	0.8	

Tabel 4.7 memperlihatkan hasil visualisasi satu dimensi dari ketiga citra augmentasi. Visualisasi tersebut menghasilkan delapan titik yang berasal dari hasil ekstraksi fitur GLCM. Pada visualisasi, delapan titik tersebut ditampilkan dari nilai ekstraksi fitur dari PC1 yang menjadi sumbu x. PC1 dipilih sebagai sumbu x karena komponen ini memiliki nilai eigen tertinggi sehingga mengandung lebih banyak informasi daripada komponen lain. Pada gambar visualisasi 1D yang memvisualisasikan hasil dari GLCM-PCA dengan parameter jarak GLCM 0.5, hanya terlihat tiga titik saja. Namun, visualisasi ini seharusnya memiliki delapan titik. Hal itu dikarenakan beberapa ekstraksi fitur memiliki nilai yang sama, sehingga titik-titik tersebut saling bertumpukan. Namun, titik-titik tersebut masih bisa dilihat dalam visualisasi grafik tiga dimensinya. Pada visualisasi grafik dengan jarak 0.8 di semua augmentasi, distribusi data yang dihasilkan lebih tersebar, walaupun masih ada beberapa titik yang tidak terlihat karena memiliki nilai yang sama.

e. *Clustering*

Penentuan jumlah klaster dilakukan sesuai dari data resmi BPBD Kota Batu yang mengelompokkan tingkat kerawanan daerah tanah longsor menjadi tiga kategori utama. Berdasarkan data tersebut, peneliti menetapkan bahwa proses *clustering* akan diarahkan agar menghasilkan tiga klaster. Penentuan rentang nilai untuk setiap klaster dilakukan berdasarkan nilai PC1 karena komponen ini memiliki kontribusi terbesar dalam menjelaskan variasi data, sebagaimana ditunjukkan dalam Tabel 4.3 sampai Tabel 4.5. Oleh karena itu, distribusi data pada dimensi PC1 dianggap paling representatif untuk menggambarkan perbedaan karakteristik

antar kelompok data. Berikut merupakan penentuan rentang nilai untuk ketiga klaster setelah dilakukan visualisasi grafik sesuai pola distribusi data.

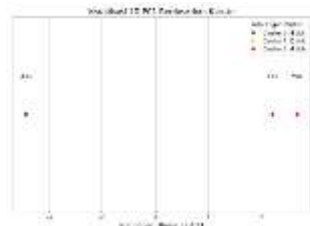
Tabel 4.8 Pembagian Cluster Berdasarkan Rentang Nilai PC1

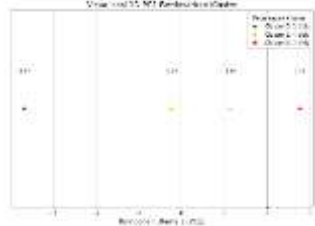
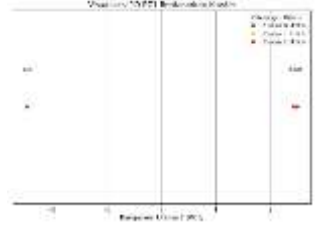
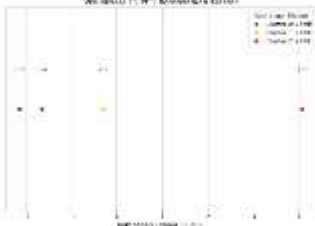
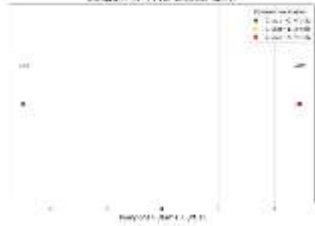
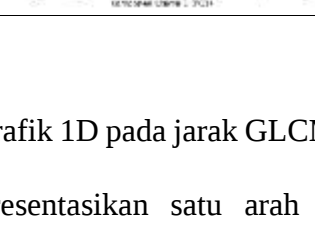
Jarak GLCM	Klaster	Rentang Nilai
0.5	Klaster 0	$-2.8 \leq n < -2$
	Klaster 1	$-1.9 < n < 2$
	Klaster 2	$2 \leq n \leq 2.9$
0.8	Klaster 0	$-4.3 \leq n < -1.5$
	Klaster 1	$-1.5 \leq n < 1.5$
	Klaster 2	$1.5 \leq n \leq 4.3$

Tabel 4.8 menunjukkan pembagian klaster secara manual berdasarkan pola sebaran data yang dihasilkan dari rentang nilai dari komponen utama. Rentang ini ditentukan berdasarkan pengamatan terhadap pola distribusi titik pada seluruh data gambar yang telah divisualisasikan ke ruang satu dimensi melalui PCA. Pola sebaran tersebut dianalisis untuk menentukan titik batas yang memisahkan kelompok data yang cenderung membentuk klaster yang berbeda. Terdapat perbedaan batas rentang nilai PC1 untuk pembagian klaster ini menggunakan jarak 0.5 dan 0.8 sehingga menunjukkan bahwa adanya pengaruh parameter jarak GLCM tersebut dalam struktur distribusi data.

Berikut merupakan visualisasi grafik 1D dengan membagi tiga klaster sesuai penentuan rentang nilai yang telah ditentukan.

Tabel 4.9 Hasil *Clustering*

Jenis Augmentasi	Jarak	Visualisasi Grafik
Rotasi 15°	0.5	

Jenis Augmentasi	Jarak	Visualisasi Grafik
Horizontal Flip	0.8	
	0.5	
	0.8	
	0.5	
Vertical Flip	0.8	
	0.5	

Tabel 4.9 menampilkan visualisasi grafik 1D pada jarak GLCM 0.5 dan 0,8. Masing-masing titik pada grafik merepresentasikan satu arah sudut dalam perhitungan GLCM, dimana satu citra menghasilkan delapan titik karena menggunakan delapan sudut. Oleh karena itu, titik-titik yang ditampilkan

merupakan hasil transformasi nilai fitur tekstur ke dalam sumbu PC1 yang telah diringkas oleh PCA. Titik tersebut dikelompokkan untuk melihat pola kemiripan karakteristik tekstur berdasarkan sudut arah GLCM.

Dalam beberapa grafik visualisasi tersebut, sebaran titik data yang terbentuk masing-masing diberi warna yang berbeda untuk merepresentasikan tiap klaster yang terbentuk. Klaster 0 ditandai dengan warna hijau, klaster 1 dengan warna kuning, serta klaster 2 dengan warna merah. Perbedaan warna dilakukan untuk mempermudah identifikasi visual terhadap pola pemisahan antar klaster.

Tabel 4.10 Titik Koordinat dan Label Data *Clustering*

Jenis Augmentasi	Jarak GLCM	Sudut GLCM	Warna	PC1
Rotasi 15°	0.5	0	Merah	2.670085
		45	Hijau	-2.438889
		90	Merah	2.207692
		135	Hijau	-2.438889
		180	Merah	2.670085
		225	Hijau	-2.438889
		270	Merah	2.207692
	0.8	315	Hijau	-2.438889
		0	Kuning	1.143704
		45	Hijau	-3.668614
		90	Merah	2.760525
		135	Hijau	-0.235614
		180	Kuning	1.143704
		225	Hijau	-3.668614
		270	Merah	2.760525
		315	Hijau	-0.235614
Horizontal Flip	0.5	0	Merah	2.482347
		45	Hijau	-2.449302
		90	Merah	2.416257
		135	Hijau	-2.449302
		180	Merah	2.482347
		225	Hijau	-2.449302
		270	Merah	2.416257
		315	Hijau	-2.449302
	0.8	0	Hijau	-1.641313
		45	Kuning	-0.298374
		90	Hijau	-2.129944

Jenis Augmentasi	Jarak GLCM	Sudut GLCM	Warna	PC1
		135	Merah	4.069633
		180	Hijau	-1.641313
		225	Kuning	-0.298374
		270	Hijau	-2.129944
		315	Merah	4.069633
Vertical Flip	0.5	0	Merah	2.467695
		45	Hijau	-2.449444
		90	Merah	2.431193
		135	Hijau	-2.449444
		180	Merah	2.467695
		225	Hijau	-2.449444
		270	Merah	2.431193
		315	Hijau	-2.449444
	0.8	0	Hijau	-1.615769
		45	Kuning	-0.111856
		90	Hijau	-2.261876
		135	Merah	3.989502
		180	Hijau	-1.615769
		225	Kuning	-0.111856
		270	Hijau	-2.261876
		315	Merah	3.989502

Tabel 4.10 menyajikan informasi mengenai posisi nilai komponen utama dalam pengelompokan klaster yang terbentuk. Posisi masing-masing titik pada sumbu PC1 memiliki makna interpretatif berdasarkan nilai *Eigen Vector* yang telah dijelaskan sebelumnya pada Tabel 4.6. Titik-titik yang berada di klaster 0 atau sebelah kiri (rentang negatif PC1) menunjukkan bahwa arah proyeksi data lebih banyak dipengaruhi oleh fitur-fitur seperti *homogeneity*, energi, dan korelasi sehingga mengindikasikan karakteristik tekstur yang lebih selaras yang dapat dihubungkan dengan indikasi tingkat kerawanan longsor yang lebih rendah.

Sementara itu, titik-titik yang berada di klaster 1 (sekitar nol pada PC1) menunjukkan nilai proyeksi yang berada di antara fitur-fitur positif dan negatif,

yang dapat dikaitkan sebagai tekstur dengan karakteristik sedang, yaitu tidak terlalu homogen maupun terlalu bervariasi.

Titik-titik yang berada di klaster 2 (rentang positif PC1) dipengaruhi oleh fitur-fitur, seperti kontras, *dissimilarity*, dan entropi yang merepresentasikan tekstur tidak seragam dan kasar sehingga dikaitkan dengan indikasi tingkat kerawanan longsor yang lebih tinggi, karena permukaan lahan yang tidak seragam dan cenderung memiliki kestabilan yang lebih rendah.

Dengan demikian, hasil *clustering* dari titik-titik tersebut merepresentasikan tingkat kerawanan longsor berdasarkan sudut-sudut arah tekstur dan parameter jarak piksel dari citra. Setiap sudut mewakili cara berbeda dalam melihat arah hubungan antar piksel dalam GLCM, sehingga memungkinkan analisis tekstur dari berbagai orientasi. Klaster yang mengandung lebih banyak titik dari arah sudut tertentu menunjukkan arah tekstur mana yang paling sering muncul pada karakteristik tertentu, sehingga dapat digunakan untuk memperkuat interpretasi spasial terhadap kondisi kerawanan lahan.

4.2.5 Evaluasi Hasil *Clustering*

Evaluasi terhadap hasil *clustering* dilakukan untuk menilai sejauh mana kualitas pengelompokan data untuk mengidentifikasi tingkat kerawanan tanah longsor. Untuk menguji kualitas tiap klaster, dilakukan pengujian menggunakan Silhouette Score. Evaluasi hasil *clustering* dilakukan secara individual terhadap masing-masing klaster yang terbentuk dan secara keseluruhan. Evaluasi dilakukan pada semua hasil *clustering* dari data augmentasi yang masing-masing terdiri dari 50 data dengan menggunakan jarak GLCM 0.5 dan 0.8.

Tabel 4.11 Nilai Silhouette Score Rotasi 15° dengan Jarak 0.5

Data Gambar	Jumlah Titik			Nilai Silhouette Score			Rata-rata
	Klaster 0	Klaster 1	Klaster 2	Klaster 0	Klaster 1	Klaster 2	
Gambar 1	4	0	4	1	-	0.9367	0.9683
Gambar 2	4	0	4	1	-	0.9160	0.9850
Gambar 3	4	0	4	1	-	0.9994	0.9997
Gambar 4	4	0	4	0.9526	-	1	0.9763
Gambar 5	4	0	4	0.9968	-	1	0.9984
Gambar 6	4	0	4	1	-	0.9737	0.9869
Gambar 7	4	0	4	0.9894	-	1	0.9947
Gambar 8	4	0	4	1	-	0.9727	0.9864
Gambar 9	4	0	4	1	-	0.9672	0.9836
Gambar 10	4	0	4	0.9718	-	1	0.9859
Gambar 11	4	0	4	1	-	0.8896	0.9448
Gambar 12	4	0	4	1	-	0.8965	0.9483
Gambar 13	4	0	4	0.9446	-	1	0.9723
Gambar 14	4	0	4	1	-	0.9616	0.9808
Gambar 15	4	0	4	1	-	0.8568	0.9284
Gambar 16	4	0	4	0.9822	-	1	0.9911
Gambar 17	4	0	4	1	-	0.8941	0.9471
Gambar 18	4	0	4	1	-	0.9732	0.9866
Gambar 19	4	0	4	0.9968	-	1	0.9984
Gambar 20	4	0	4	0.9983	-	1	0.9991
Gambar 21	4	0	4	1	-	0.9944	0.9972
Gambar 22	4	0	4	1	-	0.9260	0.9630
Gambar 23	4	0	4	1	-	0.9539	0.9769
Gambar 24	4	0	4	1	-	0.9403	0.9702
Gambar 25	4	0	4	0.9959	-	1	0.9980
Gambar 26	4	0	4	1	-	0.9886	0.9943
Gambar 27	4	0	4	1	-	0.9862	0.9931
Gambar 28	4	0	4	1	-	0.9490	0.9745
Gambar 29	4	0	4	0.8849	-	1	0.9425
Gambar 30	4	2	2	1	-	0.8515	0.9257
Gambar 31	4	0	4	1	-	0.9524	0.9762
Gambar 32	4	0	4	0.9280	-	1	0.9640
Gambar 33	4	0	4	1	-	0.9275	0.9638
Gambar 34	4	0	4	1	-	0.9086	0.9543
Gambar 35	4	0	4	1	-	0.9396	0.9698
Gambar 36	4	0	4	1	-	0.9235	0.9618
Gambar 37	4	0	4	1	-	0.9958	0.9979
Gambar 38	4	0	4	1	-	0.9745	0.9872
Gambar 39	4	0	4	1	-	0.9940	0.9970
Gambar 40	4	0	4	1	-	0.9703	0.9851
Gambar 41	4	0	4	1	-	0.9438	0.9719
Gambar 42	4	0	4	1	-	0.9854	0.9927
Gambar 43	4	0	4	1	-	0.9715	0.9857
Gambar 44	4	0	4	1	-	0.9607	0.9803
Gambar 45	4	0	4	1	-	0.9709	0.9855
Gambar 46	4	0	4	1	-	0.9794	0.9897
Gambar 47	4	0	4	1	-	0.9194	0.9597
Gambar 48	4	0	4	1	-	0.9488	0.9744
Gambar 49	4	0	4	1	-	0.9338	0.9669

Data Gambar	Jumlah Titik			Nilai Silhouette Score			Rata-rata
	Klaster 0	Klaster 1	Klaster 2	Klaster 0	Klaster 1	Klaster 2	
Gambar 50	4	0	4	1	-	0.9579	0.9790
Rata-rata				0.9328	-	0.9596	0.9691

Tabel 4.11 menampilkan distribusi jumlah titik data pergambar dalam masing-masing klaster dan nilai Silhouette Score dari setiap klaster untuk semua gambar augmentasi rotasi 15°. Nilai Silhouette Score dihitung untuk mengukur seberapa baik suatu titik data cocok dengan klaster yang menjadi anggotanya. Nilai Silhouette Score berada pada rentang angka -1 hingga 1. Nilai yang mendekati 1 menunjukkan bahwa data dikelompokkan dengan baik dan optimal. Berdasarkan tabel tersebut, diketahui bahwa sebagian besar gambar memiliki nilai Silhouette Score yang tinggi. Nilai rata-rata Silhouette Score dari semua gambar mencapai 0.9691 yang menunjukkan bahwa hasil dari *clustering* dengan augmentasi rotasi 15° dan jarak GLCM 0.5 memiliki kualitas yang sangat baik dengan nilai rata-rata klaster 0 sebesar 0.9328 dan klaster 2 sebesar 0.9596. Namun, klaster 1 tidak memiliki nilai silhouette sama sekali, yang berarti tidak ada data yang teralokasi ke dalam klaster tersebut. Hal ini menunjukkan bahwa distribusi data kurang menyebar secara merata ke seluruh klaster, sehingga struktur klasterisasi yang dihasilkan masih belum optimal.

Tabel 4.12 Nilai *Silhouette Score Horizontal Flip* dengan Jarak 0.5

Data Gambar	Jumlah Titik			Nilai Silhouette Score			Rata-rata
	Klaster 0	Klaster 1	Klaster 2	Klaster 0	Klaster 1	Klaster 2	
Gambar 1	4	0	4	1	-	0.9910	0.9955
Gambar 2	4	0	4	1	-	0.8815	0.9408
Gambar 3	4	0	4	1	-	0.9582	0.9791
Gambar 4	4	0	4	1	-	0.8953	0.9476
Gambar 5	4	0	4	0.9999	-	1	0.9999
Gambar 6	4	0	4	1	-	0.9851	0.9925
Gambar 7	4	0	4	1	-	0.8905	0.9452

Data Gambar	Jumlah Titik			Nilai Silhouette Score			Rata-rata
	Klaster 0	Klaster 1	Klaster 2	Klaster 0	Klaster 1	Klaster 2	
Gambar 8	4	0	4	1	-	0.9537	0.9769
Gambar 9	4	0	4	1	-	0.8666	0.9333
Gambar 10	4	0	4	1	-	0.9408	0.9704
Gambar 11	4	2	2	1	1	1	1
Gambar 12	4	2	2	1	1	1	1
Gambar 13	4	2	2	1	1	1	1
Gambar 14	4	0	4	1	-	0.9929	0.9964
Gambar 15	4	2	2	1	1	1	1
Gambar 16	4	0	4	0.9969	-	1	0.9984
Gambar 17	4	0	4	1	-	0.9337	0.9668
Gambar 18	4	0	4	1	-	0.9652	0.9826
Gambar 19	4	0	4	1	-	0.9551	0.9775
Gambar 20	4	0	4	1	-	0.9184	0.9592
Gambar 21	4	0	4	1	-	0.9944	0.9972
Gambar 22	4	0	4	1	-	0.9426	0.9713
Gambar 23	4	0	4	1	-	0.9537	0.9768
Gambar 24	4	2	2	1	1	1	1
Gambar 25	4	2	2	1	1	1	1
Gambar 26	4	0	4	1	-	0.9504	0.9752
Gambar 27	4	0	4	1	-	0.9821	0.9910
Gambar 28	4	0	4	1	-	0.9190	0.9595
Gambar 29	4	2	2	1	1	1	1
Gambar 30	4	2	2	1	1	1	1
Gambar 31	4	0	4	1	-	0.9520	0.9760
Gambar 32	4	2	2	1	1	1	1
Gambar 33	4	2	2	1	1	1	1
Gambar 34	4	2	2	1	1	1	1
Gambar 35	4	2	2	1	1	1	1
Gambar 36	4	0	4	1	-	0.9156	0.9578
Gambar 37	4	0	4	1	-	0.9569	0.9784
Gambar 38	4	0	4	1	-	0.9640	0.9820
Gambar 39	4	0	4	1	-	0.9924	0.9962
Gambar 40	4	0	4	1	-	0.9580	0.9790
Gambar 41	4	0	4	1	-	0.9477	0.9738
Gambar 42	4	0	4	1	-	0.9952	0.9976
Gambar 43	4	0	4	1	-	0.9607	0.9803
Gambar 44	4	0	4	1	-	0.9685	0.9842
Gambar 45	4	0	4	1	-	0.9636	0.9818
Gambar 46	4	0	4	1	-	0.9309	0.9654
Gambar 47	4	2	2	1	1	1	1
Gambar 48	4	0	4	1	-	0.9306	0.9653
Gambar 49	4	0	4	1	-	0.9227	0.9613
Gambar 50	4	0	4	0.9269	-	1	0.9635
Rata-rata				0.9985	0.26	0.9421	0.9674

Berdasarkan pada Tabel 4.12, nilai Silhouette Score pada semua hasil *clustering* pada augmentasi *horizontal flip* menggunakan jarak 0.5 memiliki nilai

yang tinggi dengan rata-rata 0.9674. Nilai Silhouette Score terendah dengan nilai 0.9333 pada gambar 9 yang masih menunjukkan hasil yang sangat baik, namun memiliki kemungkinan adanya ambiguitas dalam pemisahan antar klaster. Rata-rata nilai Silhouette Score pada masing-masing klaster juga masih mendekati optimal dengan nilai tertinggi pada klaster 0 dengan nilai 0.9980, klaster 2 sebesar 0.9388, dan klaster 1 dengan nilai 0.26 yang lebih tinggi daripada hasil evaluasi dari augmentasi rotasi 15°. Nilai klaster 1 ini masih tergolong rendah, yang mengindikasikan kualitas pemisahan yang belum optimal untuk klaster tersebut.

Tabel 4.13 Nilai Silhouette Score *Vertical Flip* dengan Jarak 0.5

Data Gambar	Jumlah Titik			Nilai Silhouette Score			Rata-rata
	Klaster 0	Klaster 1	Klaster 2	Klaster 0	Klaster 1	Klaster 2	
Gambar 1	4	0	4	1	-	0.9950	0.9975
Gambar 2	4	2	2	1	1	1	1
Gambar 3	4	0	4	1	-	0.9575	0.9787
Gambar 4	4	0	4	1	-	0.8972	0.9486
Gambar 5	4	0	4	1	-	0.9972	0.9986
Gambar 6	4	0	4	1	-	0.9840	0.9920
Gambar 7	4	0	4	1	-	0.8908	0.9454
Gambar 8	4	0	4	1	-	0.9534	0.9767
Gambar 9	4	2	2	1	1	1	1
Gambar 10	4	0	4	1	-	0.9412	0.9706
Gambar 11	4	2	2	1	1	1	1
Gambar 12	4	2	2	1	1	1	1
Gambar 13	4	2	2	1	1	1	1
Gambar 14	4	0	4	1	-	0.9907	0.9954
Gambar 15	4	2	2	1	1	1	1
Gambar 16	4	0	4	1	-	0.9984	0.9992
Gambar 17	4	0	4	1	-	0.9335	0.9667
Gambar 18	4	0	4	1	-	0.9671	0.9835
Gambar 19	4	0	4	1	-	0.9651	0.9780
Gambar 20	4	0	4	1	-	0.9093	0.9547
Gambar 21	4	0	4	1	-	0.9909	0.9955
Gambar 22	4	0	4	1	-	0.9419	0.9710
Gambar 23	4	0	4	1	-	0.9530	0.9765
Gambar 24	4	2	2	1	1	1	1
Gambar 25	4	2	2	1	1	1	1
Gambar 26	4	0	4	1	-	0.9478	0.9739
Gambar 27	4	0	4	1	-	0.9773	0.9886
Gambar 28	4	0	4	1	-	0.9178	0.9589
Gambar 29	4	2	2	1	1	1	1
Gambar 30	4	2	2	1	1	1	1
Gambar 31	4	0	4	1	-	0.9536	0.9768

Data Gambar	Jumlah Titik			Nilai Silhouette Score			Rata-rata
	Klaster 0	Klaster 1	Klaster 2	Klaster 0	Klaster 1	Klaster 2	
Gambar 32	4	2	2	1	1	1	1
Gambar 33	4	2	2	1	1	1	1
Gambar 34	4	2	2	1	1	1	1
Gambar 35	4	2	2	1	1	1	1
Gambar 36	4	0	4	1	-	0.9136	0.9568
Gambar 37	4	0	4	1	-	0.9586	0.9793
Gambar 38	4	0	4	1	-	0.9626	0.9813
Gambar 39	4	0	4	0.9982	-	1	0.9966
Gambar 40	4	0	4	1	-	0.9597	0.9798
Gambar 41	4	0	4	1	-	0.9479	0.9739
Gambar 42	4	0	4	1	-	0.9948	0.9974
Gambar 43	4	0	4	1	-	0.9614	0.9807
Gambar 44	4	0	4	1	-	0.9663	0.9832
Gambar 45	4	0	4	1	-	0.9645	0.9823
Gambar 46	4	0	4	1	-	0.9319	0.9659
Gambar 47	4	2	2	1	1	1	1
Gambar 48	4	0	4	1	-	0.9310	0.9655
Gambar 49	4	0	4	1	-	0.9241	0.9620
Gambar 50	4	0	4	1	-	0.9245	0.9622
Rata-rata				0.9999	0.30	0.9608	0.9727

Pada Tabel 4.13, hasil evaluasi Silhouette Score pada data citra dengan augmentasi *vertical flip* dan jarak GLCM 0.5 memiliki rata-rata nilai sebesar 0.9727 dengan nilai terendah pada Gambar 7 dengan nilai rata-ratanya sebesar 0.9454. Hal ini lebih tinggi dari nilai evaluasi pada citra dengan augmentasi horizontal flip, namun masih memiliki nilai yang tinggi dan mendekati nilai optimal. Meskipun begitu, perlu dilihat bahwa persebaran data antar klaster masih belum merata. Kondisi ini mengindikasikan bahwa meskipun kualitas pemisahan antar klaster secara umum tinggi, distribusi data dalam ruang fitur belum sepenuhnya seimbang. Ketidakseimbangan ini dapat mempengaruhi interpretasi hasil clustering dan menunjukkan adanya dominasi klaster tertentu dalam struktur data.

Evaluasi selanjutnya dilakukan dengan augmentasi menggunakan jarak GLCM 0.8 piksel pada ketiga jenis augmentasi untuk perbandingan konsistensi

hasil *clustering* terhadap perubahan parameter jarak GLCM.

Tabel 4.14 Nilai Silhouette Score *Rotasi 15°* dengan Jarak 0.8

Data Gambar	Jumlah Titik			Nilai Silhouette Score			Rata-rata
	Klaster 0	Klaster 1	Klaster 2	Klaster 0	Klaster 1	Klaster 2	
Gambar 1	2	4	2	1	0.5622	1	0.7811
Gambar 2	2	2	4	1	1	0.9531	0.9766
Gambar 3	2	2	4	1	1	0.7714	0.8857
Gambar 4	2	4	2	1	0.3142	1	0.6571
Gambar 5	4	2	2	0.6546	1	1	0.8273
Gambar 6	4	2	2	0.7465	1	1	0.8732
Gambar 7	2	2	4	1	1	0.7181	0.8590
Gambar 8	2	2	4	1	1	0.7745	0.8872
Gambar 9	2	4	2	1	0.8588	1	0.9294
Gambar 10	2	4	2	1	0.6639	1	0.8320
Gambar 11	2	2	4	1	1	0.7398	0.8699
Gambar 12	2	4	2	1	0.9148	1	0.9574
Gambar 13	2	4	2	1	0.9350	1	0.9675
Gambar 14	2	4	2	1	0.8624	1	0.9312
Gambar 15	2	4	2	1	0.7395	1	0.8697
Gambar 16	4	2	2	0.5986	1	1	0.7993
Gambar 17	2	4	2	1	0.5472	1	0.7736
Gambar 18	4	2	2	0.9638	1	1	0.9819
Gambar 19	4	2	2	0.8455	1	1	0.9227
Gambar 20	4	2	2	0.5813	1	1	0.7906
Gambar 21	2	2	4	1	1	0.9591	0.9795
Gambar 22	4	2	2	0.6304	1	1	0.8152
Gambar 23	2	4	2	1	0.8485	1	0.9243
Gambar 24	2	4	2	1	0.9821	1	0.9910
Gambar 25	2	4	2	1	0.8824	1	0.9412
Gambar 26	4	2	2	0.9772	1	1	0.9886
Gambar 27	4	2	2	0.8146	1	1	0.9073
Gambar 28	4	2	2	0.9737	1	1	0.9869
Gambar 29	2	4	2	1	0.8534	1	0.9267
Gambar 30	2	4	2	1	0.6114	1	0.8057
Gambar 31	2	4	2	1	0.7843	1	0.8922
Gambar 32	2	4	2	1	0.6464	1	0.8232
Gambar 33	4	2	2	0.8091	1	1	0.9045
Gambar 34	2	4	2	1	0.7161	1	0.8580
Gambar 35	2	4	2	1	0.4821	1	0.7411
Gambar 36	2	4	2	1	0.8604	1	0.9302
Gambar 37	2	2	4	1	1	0.9639	0.9820
Gambar 38	2	4	2	1	0.1449	1	0.5725
Gambar 39	4	0	4	0.9825	-	0.8509	0.9167
Gambar 40	4	2	2	0.5814	1	1	0.7907
Gambar 41	2	4	2	1	0.1994	1	0.5997
Gambar 42	4	2	2	0.8665	1	1	0.9333
Gambar 43	2	4	2	1	0.2146	1	0.6073
Gambar 44	4	2	2	0.6695	1	1	0.8347
Gambar 45	4	0	4	0.9423	-	0.8015	0.8719
Gambar 46	2	4	2	1	0.4414	1	0.7207
Gambar 47	2	4	2	1	0.9741	1	0.9870

Data Gambar	Jumlah Titik			Nilai Silhouette Score			Rata-rata
	Klaster 0	Klaster 1	Klaster 2	Klaster 0	Klaster 1	Klaster 2	
Gambar 48	4	0	4	0.9709	-	0.7682	0.8696
Gambar 49	2	4	2	1	0.4013	1	0.7007
Gambar 50	2	4	2	1	0.4539	1	0.7269
Rata-rata				0.9322	0.7579	0.9653	0.8443

Berdasarkan pada Tabel 4.14, diketahui bahwa nilai Silhouette Score rata-rata yang diperoleh dari seluruh data citra pada augmentasi rotasi 15 menggunakan jarak GLCM 0.8 mencapai 0.8443. Nilai ini mengindikasikan bahwa hasil *clustering* secara umum cukup baik, meskipun tidak setinggi hasil yang diperoleh pada parameter jarak GLCM 0.5. Klaster 2 menunjukkan performa klaster paling tinggi dengan nilai rata-rata sebesar 0.9653, disusul dengan klaster 0 dengan nilai 0.9322, sedangkan klaster 1 memiliki nilai terendah dengan skor 0.7579. Hal ini menunjukkan adanya perbedaan kualitas pemisahan antar klaster pada beberapa citra, khususnya pada klaster 1 sehingga memiliki kedekatan yang kurang optimal terhadap anggota klaster yang sama. Selain itu, terlihat pada Gambar 39, Gambar 45, dan Gambar 48 tidak memiliki nilai Silhouette Score yang menandakan bahwa tidak adanya titik-titik data pada gambar-gambar tersebut sehingga distribusi datanya masih belum cukup tersebar. Sebaliknya, Gambar 24 menunjukkan nilai Silhouette Score yang sangat tinggi dengan nilai 0.9910 yang menunjukkan kohesi dan pemisahan klaster yang sudah sangat baik.

Tabel 4.15 Nilai Silhouette Score *Horizontal Flip* dengan Jarak 0.8

Data Gambar	Jumlah Titik			Nilai Silhouette Score			Rata-rata
	Klaster 0	Klaster 1	Klaster 2	Klaster 0	Klaster 1	Klaster 2	
Gambar 1	4	2	2	0.7898	1	1	0.8949
Gambar 2	6	0	2	0.9865	-	1	0.9899
Gambar 3	2	4	2	1	0.8844	1	0.9422
Gambar 4	4	0	4	0.7680	-	0.9052	0.8366
Gambar 5	4	2	2	0.9412	1	1	0.9706

Data Gambar	Jumlah Titik			Nilai Silhouette Score			Rata-rata
	Klaster 0	Klaster 1	Klaster 2	Klaster 0	Klaster 1	Klaster 2	
Gambar 6	2	4	2	1	0.5414	1	0.7707
Gambar 7	2	4	2	1	0.6028	1	0.8014
Gambar 8	2	4	2	1	0.6228	1	0.8114
Gambar 9	2	2	4	1	1	0.7652	0.8826
Gambar 10	4	2	2	0.6423	1	1	0.9211
Gambar 11	2	4	2	1	0.9664	1	0.9832
Gambar 12	2	2	4	1	1	0.9976	0.9988
Gambar 13	2	4	2	1	0.4804	1	0.7402
Gambar 14	4	2	2	0.8147	1	1	0.9073
Gambar 15	2	4	2	1	0.8686	1	0.9343
Gambar 16	2	4	2	1	0.1109	1	0.5555
Gambar 17	2	4	2	1	0.8954	1	0.9477
Gambar 18	2	4	2	1	0.8286	1	0.9143
Gambar 19	2	4	2	1	0.7633	1	0.8817
Gambar 20	2	4	2	1	0.4226	1	0.7113
Gambar 21	4	2	2	0.8038	1	1	0.9019
Gambar 22	2	4	2	1	0.9581	1	0.9791
Gambar 23	2	4	2	1	0.5356	1	0.7678
Gambar 24	2	4	2	1	0.0067	1	0.5034
Gambar 25	2	4	2	1	0.3611	1	0.6806
Gambar 26	2	4	2	1	0.9737	1	0.9866
Gambar 27	2	4	2	1	0.6853	1	0.8427
Gambar 28	2	2	4	1	1	0.9143	0.9571
Gambar 29	2	4	2	1	0.5651	1	0.7826
Gambar 30	2	4	2	1	0.8297	1	0.9148
Gambar 31	2	4	2	1	0.9720	1	0.9860
Gambar 32	2	4	2	1	0.7103	1	0.8552
Gambar 33	2	4	2	1	0.6956	1	0.8478
Gambar 34	2	4	2	1	0.9170	1	0.9585
Gambar 35	2	2	4	1	1	0.8905	0.9452
Gambar 36	2	4	2	1	0.8632	1	0.9316
Gambar 37	4	2	2	0.9545	1	1	0.9772
Gambar 38	2	4	2	1	0.1756	1	0.5878
Gambar 39	2	2	4	1	1	0.8339	0.9170
Gambar 40	4	2	2	0.9931	1	1	0.9966
Gambar 41	2	2	4	1	1	0.6852	0.8426
Gambar 42	4	2	2	0.9095	1	1	0.9547
Gambar 43	2	2	4	1	1	0.7775	0.8888
Gambar 44	4	2	2	0.8326	1	1	0.9163
Gambar 45	2	4	2	1	0.8275	1	0.9137
Gambar 46	2	4	2	1	0.8312	1	0.9156
Gambar 47	2	4	2	1	0.9128	1	0.9564
Gambar 48	2	4	2	1	0.6455	1	0.8228
Gambar 49	2	4	2	1	0.6084	1	0.8042
Gambar 50	2	2	4	1	1	0.7655	0.8828
Rata-rata				0.9536	0.8439	0.9707	0.9044

Hasil perhitungan nilai Silhouette Score pada citra augmentasi horizontal flip di Tabel 4.15 menunjukkan pemisahan klaster yang baik karena memiliki rata-rata nilai keseluruhan sebesar 0.9044. Klaster 0 memiliki rata-rata tertinggi sebesar 0.9536, disusul dengan klaster 2 sebesar 0.9707, dan klaster 1 sebesar 0.8439. Nilai ini lebih baik dibandingkan dengan menggunakan augmentasi rotasi 15°, yang ditunjukkan dengan berkurangnya jumlah nilai Silhouette Score yang tidak tersedia menjadi hanya dua. Hal ini mengindikasikan bahwa pemisahan antar klaster serta kedekatan titik-titik data dalam masing-masing klaster telah mendekati kondisi yang optimal.

Tabel 4.16 Nilai Silhouette Score *Vertical Flip* dengan Jarak 0.8

Data Gambar	Jumlah Titik			Nilai Silhouette Score			Rata-rata
	Klaster 0	Klaster 1	Klaster 2	Klaster 0	Klaster 1	Klaster 2	
Gambar 1	4	2	2	0.7566	1	1	0.8783
Gambar 2	0	6	2	-	0.9889	1	0.9917
Gambar 3	2	4	2	1	0.6354	1	0.8177
Gambar 4	4	0	4	0.8070	-	0.8259	0.8164
Gambar 5	2	2	4	1	1	0.9088	0.9544
Gambar 6	2	4	2	1	-0.1289	1	0.4356
Gambar 7	2	4	2	1	0.4695	1	0.7348
Gambar 8	2	4	2	1	0.5534	1	0.7767
Gambar 9	4	2	2	0.8011	1	1	0.9005
Gambar 10	2	4	2	1	0.9508	1	0.9754
Gambar 11	2	4	2	1	0.8206	1	0.9103
Gambar 12	2	4	2	1	0.2056	1	0.6028
Gambar 13	2	4	2	1	0.3974	1	0.6987
Gambar 14	2	4	2	1	0.7183	1	0.8592
Gambar 15	2	4	2	1	0.8593	1	0.9297
Gambar 16	2	2	4	1	1	0.7125	0.8562
Gambar 17	2	4	2	1	0.7164	1	0.8582
Gambar 18	2	4	2	1	0.8609	1	0.9305
Gambar 19	2	4	2	1	0.7812	1	0.8906
Gambar 20	2	2	4	1	1	0.7198	0.8599
Gambar 21	4	2	2	0.8888	1	1	0.9444
Gambar 22	2	4	2	1	0.9929	1	0.9965
Gambar 23	2	4	2	1	0.9086	1	0.9543
Gambar 24	4	2	2	0.1593	1	1	0.5797
Gambar 25	2	4	2	1	0.4160	1	0.7080
Gambar 26	2	4	2	1	0.9652	1	0.9826
Gambar 27	2	4	2	1	0.8131	1	0.9065
Gambar 28	2	4	2	1	0.7420	1	0.8710

Data Gambar	Jumlah Titik			Nilai Silhouette Score			Rata-rata
	Klaster 0	Klaster 1	Klaster 2	Klaster 0	Klaster 1	Klaster 2	
Gambar 29	2	4	2	1	0.6180	1	0.8090
Gambar 30	2	4	2	1	0.7783	1	0.8892
Gambar 31	2	4	2	1	0.8904	1	0.9452
Gambar 32	2	4	2	1	0.9625	1	0.9812
Gambar 33	2	4	2	1	0.6934	1	0.8467
Gambar 34	2	4	2	1	0.8893	1	0.9446
Gambar 35	2	2	4	1	1	0.8532	0.9266
Gambar 36	2	4	2	1	0.9765	1	0.9883
Gambar 37	4	2	2	0.7725	1	1	0.8862
Gambar 38	2	2	4	1	1	0.8831	0.9415
Gambar 39	2	2	4	1	1	0.6931	0.8466
Gambar 40	2	2	4	1	1	0.9397	0.9698
Gambar 41	2	2	4	1	1	0.6898	0.8449
Gambar 42	2	2	4	1	1	0.7853	0.8927
Gambar 43	4	0	4	0.8526	-	0.9651	0.9089
Gambar 44	4	2	2	0.7432	1	1	0.8716
Gambar 45	2	4	2	1	0.6665	1	0.8332
Gambar 46	2	4	2	1	0.5694	1	0.7847
Gambar 47	2	4	2	1	0.9425	1	0.9713
Gambar 48	2	4	2	1	0.4896	1	0.7448
Gambar 49	2	4	2	1	0.5591	1	0.7795
Gambar 50	2	2	4	1	1	0.7817	0.8908
Rata-rata				0.9396	0.7481	0.9551	0.8608

Dengan menggunakan augmentasi *vertical flip* dengan jarak GLCM 0.8, nilai rata-rata Silhouette Score pada semua data adalah sebesar 0.8608. Klaster 2 menunjukkan performa paling optimal dengan nilai rata-rata tertinggi yaitu 0.9551, diikuti oleh klaster 0 sebesar 0.9396, dan klaster 1 sebesar 0.7481. Nilai-nilai ini menunjukkan bahwa hasil clustering masih tergolong baik, walaupun tidak sebaik jika menggunakan augmentasi *horizontal flip*.

Secara keseluruhan, pemilihan parameter jarak GLCM 0.8 menghasilkan nilai evaluasi dengan Silhouette Score menurun jika dibandingkan dengan jarak GLCM 0.5. Meskipun tidak setinggi sebelumnya, nilai tersebut masih tergolong baik dan mencerminkan pemisahan antar klaster yang cukup jelas. Selain itu,, penggunaan jarak GLCM 0.8 juga menunjukkan bahwa distribusi titik data antar

klaster lebih menyebar. Hal ini menunjukkan bahwa meskipun kualitas evaluasi sedikit menurun, penyebaran data dalam ruang fitur menjadi lebih merata, sehingga dapat meningkatkan representasi masing-masing klaster secara keseluruhan.

Dari hasil evaluasi menunjukkan bahwa penggunaan jarak GLCM 0.8 yang dipadukan dengan augmentasi *horizontal flip* memberikan hasil *clustering* yang paling optimal, karena meskipun nilai rata-rata Silhouette Score sedikit menurun dibandingkan jarak 0.5, kombinasi ini menghasilkan distribusi data yang lebih merata dan representasi klaster yang lebih baik secara keseluruhan dibandingkan augmentasi lainnya.

Untuk meninjau hasil *clustering* yang lebih optimal, dilakukan evaluasi terhadap jumlah klaster yang digunakan dalam proses pembentukan klaster dengan melakukan membandingkan beberapa variasi jumlah klaster untuk mengidentifikasi struktur kelompok yang paling representatif terhadap pola tekstur citra. Evaluasi ini difokuskan pada data dengan augmentasi horizontal flip yang menggunakan jarak GLCM 0.8 untuk menentukan jumlah klaster paling optimal dalam memperkuat kualitas pemisahan dan interpretasi antar klaster.

Tabel 4.17 Rentang Nilai PC1 (n) pada 2 Klaster

Klaster	Rentang Nilai
Klaster 0	$-4.3 \leq n < 0$
Klaster 1	$0 \leq n \leq 4.3$

Tabel 4.17 menunjukkan konfigurasi rentang nilai dua klaster yang digunakan dalam evaluasi klasterisasi. Pembagian ini terdiri dari klaster 0 yang memiliki nilai diantara -4.3 hingga 0 dan klaster 1 untuk nilai antara 0 hingga 4.3. Berdasarkan rentang nilai tersebut, dilakukan proses *clustering* sehingga menghasilkan nilai *Silhouette Score* dibawah ini.

Tabel 4.18 Hasil Evaluasi *Horizontal Flip Jarak 0.8* pada 2 klaster

Data Gambar	Jumlah Titik		Nilai Silhouette Score		Rata-rata
	Klaster 0	Klaster 1	Klaster 0	Klaster 1	
Gambar 1	6	2	0.8104	1	0.8578
Gambar 2	6	2	0.9864	1	0.9899
Gambar 3	6	2	0.8940	1	0.9206
Gambar 4	4	4	0.7680	0.9052	0.8366
Gambar 5	6	2	0.8247	1	0.8686
Gambar 6	4	4	0.5053	0.6201	0.5627
Gambar 7	4	4	0.1918	0.7120	0.4519
Gambar 8	4	4	0.7059	0.1749	0.4404
Gambar 9	2	6	1	0.7841	0.8381
Gambar 10	6	2	0.9498	1	0.9624
Gambar 11	6	2	0.6634	1	0.7476
Gambar 12	2	6	1	0.8978	0.9234
Gambar 13	2	6	1	0.9284	0.9463
Gambar 14	6	2	0.8890	1	0.9168
Gambar 15	2	6	1	0.8438	0.8829
Gambar 16	4	4	0.6084	0.7680	0.6882
Gambar 17	2	6	1	0.7836	0.8377
Gambar 18	6	2	0.6756	1	0.7568
Gambar 19	2	6	1	0.7438	0.8079
Gambar 20	4	4	0.3581	0.7568	0.5575
Gambar 21	4	4	0.8536	0.6489	0.7513
Gambar 22	2	6	1	0.7146	0.786
Gambar 23	4	4	0.5357	0.6008	0.5683
Gambar 24	2	6	1	0.9706	0.978
Gambar 25	4	4	0.5843	0.6797	0.6321
Gambar 26	2	6	1	0.6375	0.7282
Gambar 27	2	6	1	0.8702	0.9027
Gambar 28	2	6	1	0.7702	0.8277
Gambar 29	4	4	0.6937	0.3403	0.517
Gambar 30	2	6	1	0.7057	0.7793
Gambar 31	6	2	0.7210	1	0.7908
Gambar 32	2	6	1	0.8601	0.8951
Gambar 33	4	4	0.2012	0.6460	0.4237
Gambar 34	6	2	0.7824	1	0.8368
Gambar 35	2	6	1	0.8506	0.888
Gambar 36	2	6	1	0.7459	0.8094
Gambar 37	4	4	0.9717	0.4109	0.6913
Gambar 38	2	6	1	0.8990	0.9243
Gambar 39	2	6	1	0.8032	0.8525
Gambar 40	6	2	0.8827	1	0.9121
Gambar 41	4	4	0.2728	0.8300	0.5514
Gambar 42	4	4	0.9412	0.4702	0.7057
Gambar 43	4	4	0.4102	0.8640	0.6371
Gambar 44	4	4	0.8707	0.6862	0.7785
Gambar 45	6	2	0.7049	1	0.7787
Gambar 46	2	6	1	0.7959	0.847
Gambar 47	2	6	1	0.7435	0.8077
Gambar 48	4	4	0.4275	0.5843	0.506
Gambar 49	4	4	0.5396	0.5278	0.5337
Gambar 50	4	4	0.3717	0.8612	0.6165

Data Gambar	Jumlah Titik		Nilai Silhouette Score		Rata-rata
	Klaster 0	Klaster 1	Klaster 0	Klaster 1	
	Rata-rata		0.8778	0.8771	0.8249

Dari hasil yang dipaparkan oleh Tabel 4.18, dapat dilihat bahwa hasil evaluasi memperoleh rata-rata nilai *Silhouette Score* gabungan dari kedua klaster sebesar 0.8249, dengan klaster 0 memiliki rata-rata 0.8778 dan klaster 1 dengan rata-rata 0.8771. Nilai ini menunjukkan bahwa hasil pemisahan klaster tergolong cukup baik, karena mendekati nilai maksimum yaitu 1, yang berarti mayoritas data berada cukup jauh dari batas antar klaster dan memiliki kemiripan tekstur yang tinggi dengan anggota klasternya sendiri. Beberapa gambar seperti Gambar 2,3, dan 10 memiliki nilai mendekati 1, mencerminkan konsistensi klaster yang tinggi, sedangkan gambar seperti Gambar 8 dan 29 menunjukkan skor yang lebih rendah akibat posisi data yang lebih dekat dengan batas antar klaster.

Setelah dilakukan evaluasi dengan dua klaster, tahap selanjutnya adalah melakukan pembagian data menjadi empat klaster untuk mengevaluasi kualitas pemisahan data yang lebih kompleks. Rentang nilai PC1 yang digunakan untuk pembentukan empat klaster disesuaikan secara merata berdasarkan distribusi nilai PC1, sebagaimana ditunjukkan pada Tabel 4.19

Tabel 4.19 Rentang Nilai PC1 (n) pada 4 Klaster

Klaster	Rentang Nilai
Klaster 0 (K-0)	$-4.3 \leq n < -2$
Klaster 1 (K-1)	$-2 \leq n < 0$
Klaster 2 (K-2)	$0 \leq n < 2$
Klaster 3 (K-3)	$2 \leq n \leq 4.3$

Evaluasi terhadap hasil *clustering* empat klaster menggunakan nilai *Silhouette Score* untuk masing-masing gambar dan klaster dapat dilihat pada Tabel

4.20 di bawah ini.

Tabel 4.20 Hasil Evaluasi *Horizontal Flip* Jarak 0.8 pada 4 Klaster

Data Gambar	Jumlah Titik				Nilai Silhouette Score				Rata-rata
	K-0	K-1	K-2	K-3	K-0	K-1	K-2	K-3	
Gambar 1	2	4	0	2	1	0.0284	1	-	0.5142
Gambar 2	0	6	0	2	-	0.9864	-	1	0.9899
Gambar 3	2	4	0	2	1	0.8844	-	1	0.9422
Gambar 4	2	2	2	2	1	1	1	1	1
Gambar 5	0	6	0	2	-	0.8247	-	1	0.8686
Gambar 6	2	2	2	2	1	1	1	1	1
Gambar 7	2	2	2	2	1	1	1	1	1
Gambar 8	2	2	2	2	1	1	1	1	1
Gambar 9	2	-	4	2	1	-	0.0733	1	0.5367
Gambar 10	0	6	0	2	-	0.9498	-	1	0.9624
Gambar 11	2	4	0	2	1	0.9663	-	1	0.9832
Gambar 12	2	0	6	0	1	-	0.8978	-	0.9234
Gambar 13	2	0	6	0	1	-	0.9284	-	0.9463
Gambar 14	0	6	0	2	-	0.8890	-	1	0.9168
Gambar 15	2	0	4	2	1	-	0.8685	1	0.9343
Gambar 16	2	2	2	2	1	1	1	1	1
Gambar 17	2	0	4	2	1	-	0.8953	1	0.9477
Gambar 18	2	4	0	2	1	0.8285	-	1	0.9143
Gambar 19	2	0	4	2	1	-	0.7633	1	0.8817
Gambar 20	2	2	2	2	1	1	1	1	1
Gambar 21	2	2	2	2	1	1	1	1	1
Gambar 22	2	0	4	2	1	-	0.9581	1	0.9791
Gambar 23	2	2	2	2	1	1	1	1	1
Gambar 24	2	0	6	0	1	-	0.9706	-	0.978
Gambar 25	2	2	2	2	1	1	1	1	1
Gambar 26	2	0	4	0	1	-	0.9732	1	0.9866
Gambar 27	2	0	4	2	1	-	0.6853	1	0.8427
Gambar 28	2	0	6	0	1	-	0.7702	-	0.8277
Gambar 29	2	2	2	2	1	1	1	1	1
Gambar 30	2	0	4	2	1	-	0.8296	1	0.9148
Gambar 31	2	4	0	2	1	0.9720	-	1	0.986
Gambar 32	2	0	4	2	1	-	0.7103	1	0.8552
Gambar 33	2	2	2	2	1	1	1	1	1
Gambar 34	2	4	0	2	1	0.9170	-	1	0.9585
Gambar 35	2	0	6	0	1	-	0.8506	-	0.888
Gambar 36	2	0	4	2	1	-	0.8632	1	0.9316
Gambar 37	4	0	2	2	0.9544	-	1	1	0.9772
Gambar 38	2	0	6	0	1	-	0.8990	-	0.9243
Gambar 39	2	0	4	2	1	-	-0.0506	1	0.4747
Gambar 40	0	6	0	2	-	0.8827	-	1	0.9121
Gambar 41	2	2	2	2	1	1	1	1	1
Gambar 42	2	2	2	2	1	1	1	1	1
Gambar 43	2	2	2	2	1	1	1	1	1
Gambar 44	2	2	2	2	1	1	1	1	1
Gambar 45	2	4	0	2	1	0.8274	-	1	0.9137
Gambar 46	2	0	4	2	1	-	0.8312	1	0.9156
Gambar 47	2	0	4	2	1	-	0.9127	1	0.9564
Gambar 48	2	2	2	2	1	1	1	1	1

Data Gambar	Jumlah Titik				Nilai Silhouette Score				Rata-rata
	K-0	K-1	K-2	K-3	K-0	K-1	K-2	K-3	
Gambar 49	2	2	2	2	1	1	1	1	1
Gambar 50	2	2	2	2	1	1	1	1	1
Rata-rata					0.9391	0.5757	0.6618	0.9000	0.9586

Hasil tersebut menunjukkan bahwa nilai rata-rata keseluruhan adalah 0.9586 sehingga lebih tinggi dibandingkan hasil rata-rata saat menggunakan tiga kluster. Hal ini berarti pemisahan secara umum pada empat kluster memberikan struktur kelompok yang lebih baik. Sementara itu, saat dilihat secara individu per-kluster, nilai Silhouette Score pada masing-masing kluster lebih tinggi ketika menggunakan tiga kluster. Hal ini disebabkan oleh adanya beberapa kluster tidak terisi titik data dalam beberapa gambar, sehingga skor Silhouette tidak dapat dihitung pada bagian tersebut. Sementara itu, saat menggunakan tiga kluster, hampir semua gambar memiliki distribusi data yang merata ke dalam setiap kluster, sehingga hampir semua gambar memiliki distribusi data yang merata ke dalam setiap kluster. Berdasarkan hal tersebut, meskipun empat kluster secara teoritis menghasilkan rata-rata pemisahan yang lebih baik, implementasinya justru cenderung menghasilkan pembagian yang tidak merata. Banyaknya kluster kosong menunjukkan bahwa model pemisahan kurang efektif dalam menangkap distribusi alami data. Oleh karena itu, penggunaan tiga kluster dapat dianggap lebih representatif dan stabil karena pembagiannya lebih seimbang dan menghasilkan kualitas kluster individu yang lebih baik.

4.3 Pembahasan

Penelitian ini dilakukan dengan menganalisis *clustering* tingkat kerawanan daerah tanah longsor menggunakan metode GLCM dan PCA dengan data penelitian

berjumlah 50 data citra lahan miring. Terdapat 2 tahapan utama dalam penelitian ini, yaitu proses GLCM dan PCA. Pada proses GLCM, citra yang diambil secara langsung di beberapa daerah Kota Batu dilakukan pemotongan untuk menghilangkan objek tidak relevan dan memfokuskan citra pada area lahan miring. Citra yang sudah dilakukan pemotongan, lalu dilakukan proses augmentasi menjadi tiga jenis augmentasi, yaitu rotasi 15°, *horizontal flip*, dan *vertical flip*. Citra hasil dari ketiga jenis augmentasi tersebut diubah ukurannya menjadi ukuran 64x64. Pengubahan ukuran dilakukan untuk mempertimbangkan efisiensi komputasi dan tetap menjaga detail tekstur pada citra sehingga rangkaian proses analisis dapat bekerja secara optimal. Selanjutnya, citra dikonversi menjadi format *grayscale*. Pengkonversian ini dilakukan karena metode GLCM melakukan perhitungan hubungan spasial antar piksel berdasarkan nilai intensitas keabuan. Oleh karena itu, input citra harus berupa citra *grayscale* untuk mendapat informasi tekstur secara akurat. Nilai-nilai intensitas keabuan pada citra *grayscale* digunakan untuk menghitung hubungan antar piksel berdasarkan jarak dan sudut tertentu sehingga menghasilkan fitur statistik seperti *contrast*, *dissimilarity*, *homogeneity*, entropi, *energy*, dan *correlation*. Hasilnya kemudian disimpan dalam format *spreadsheet* untuk mempermudah dalam proses analisis *clustering* menggunakan PCA. Dalam proses ekstraksi ini, digunakan dua nilai parameter jarak GLCM, yaitu 0.5 dan 0.8. Oleh karena itu, proses ini menghasilkan dua set data fitur tekstur yang berbeda untuk analisis perbandingan guna mengevaluasi pengaruh variasi jarak terhadap distribusi nilai fitur, hasil reduksi dan *clustering* dengan PCA, serta kualitas pemisahan klaster yang dihasilkan.

Tahapan selanjutnya adalah proses PCA. Langkah pertama pada proses ini adalah normalisasi data fitur GLCM. Normalisasi data fitur dilakukan untuk menyeimbangkan data agar tidak ada fitur yang mendominasi dalam proses pembentukan komponen utama. Dengan normalisasi, semua fitur memiliki skala yang sama yaitu memiliki rata-rata bernilai nol dan standar deviasi bernilai satu sehingga PCA dapat mengidentifikasi variasi berdasarkan pola hubungan antar fitur, bukan besarnya nilai fitur. Fitur-fitur yang dihasilkan dari ekstraksi metode GLCM, seperti kontras, korelasi, energi, *homogeneity*, *dissimilarity*, dan entropi, diperlakukan sebagai dimensi dalam ruang fitur. Dimensi-dimensi tersebut merupakan komponen yang dijadikan dasar dalam perhitungan PCA dalam membentuk komponen utama.

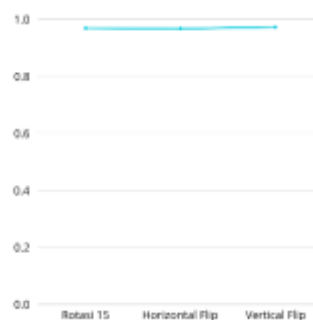
Setelah data dinormalisasi, dilakukan perhitungan nilai eigen dan rasio varians dari tiap-tiap komponen fitur untuk pembentukan komponen utama (PC) sehingga komponen utama yang terbentuk adalah PC1 dan PC2. Ketiga komponen ini digunakan dalam visualisasi grafik dengan 2 dimensi dan 3 dimensi untuk menampilkan pola sebaran data sehingga proses *clustering* bisa dilakukan.

Proses pengelompokkan dilakukan dengan melihat distribusi data yang dihasilkan dari grafik berdasarkan rentang nilai dari PC1 yang memiliki nilai rasio varians terbesar. Pengelompokan (*clustering*) dari grafik ini menghasilkan tiga kluster utama pada masing-masing *clustering* dengan penggunaan kedua jarak GLCM, yaitu kluster 0, kluster 1, dan kluster 2.

Pada jarak GLCM 0.5, kluster 0 diisi oleh nilai PC1 yang kurang dari -2, kluster 1 dengan nilai diantara -2 hingga 2, dan kluster 2 dengan nilai lebih dari 2.

Sementara itu, untuk jarak GLCM 0.8, dilakukan penyesuaian rentang dengan cakupan yang lebih sempit, yaitu klaster 0 dengan nilai kurang dari -1.5, klaster 1 dengan nilai diantara -1.5 hingga 1.5, dan klaster 2 dengan nilai lebih dari 1.5.

Dalam penelitian ini, dilakukan perbandingan hasil *clustering* ini berdasarkan dua parameter jarak GLCM 0.5 dan 0.8 piksel pada masing-masing citra hasil dari ketiga augmentasi. Berdasarkan pada Tabel 4.11 sampai Tabel 4.16, dapat dilihat bahwa evaluasi nilai Silhouette Score pada hasil *clustering* yang menggunakan jarak GLCM 0.5 memiliki nilai yang tinggi sehingga bisa dikatakan optimal pada pembagian klasternya. Nilai rata-rata pada hasil *clustering* dengan augmentasi rotasi 15° yaitu 0.9691, pada augmentasi *horizontal flip* sebesar 0.9674, dan pada augmentasi *vertical flip* sebesar 0.9727.

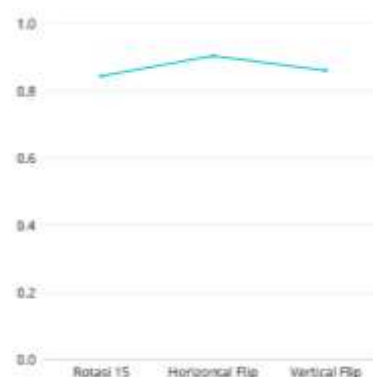


Gambar 4.5 Nilai Silhouette Score dengan Jarak GLCM 0.5

Namun, penggunaan parameter jarak GLCM 0.5 menghasilkan persebaran data yang tidak merata karena banyak titik data yang hanya terkonsentrasi pada klaster 0 dan klaster 2. Sebagian besar klaster 1 tidak terisi sama sekali. Hal ini menunjukkan bahwa struktur klaster tidak sejalan dengan tujuan awal yang membagi ke dalam tiga klaster. Penyebab kondisi ini adalah sebaran titik koordinat yang dihasilkan dari komponen utama tidak merata. Berdasarkan pengamatan visual terhadap data, sebagian besar titik memiliki nilai koordinat yang identik atau

sangat berdekatan, sehingga menyebabkan penumpukan titik pada wilayah tertentu. Akibatnya, proses clustering cenderung mengelompokkan data ke dalam dua klaster dominan saja, sedangkan satu klaster lainnya yaitu klaster 1 menjadi kosong atau tidak memiliki titik data karena tidak ditemukan karakteristik yang cukup berbeda untuk membentuk kelompok sendiri.

Penelitian *clustering* yang menggunakan jarak 0.8 menunjukkan hasil yang optimal, baik dari sisi evaluasi menggunakan nilai Silhouette Score maupun dari sisi distribusi data antar klaster. Hal ini dapat dilihat pada Tabel 4.14 yang menunjukkan bahwa rata-rata nilai Silhouette Score yang dihasilkan menggunakan augmentasi rotasi 15° adalah 0.8443, yang dinilai masih optimal. Pada Tabel 4.15, nilai Silhouette Score yang dihasilkan dari augmentasi *horizontal flip* mengalami kenaikan yaitu sebesar 0.9044 yang berarti pemisahan klaster tergolong baik dan optimal. Sementara itu, hasil evaluasi pada data augmentasi *vertical flip* justru semakin menurunkan performa, dengan nilai sebesar 0.8608 yang mengindikasikan bahwa kualitas pemisahan klaster mulai menurun dibandingkan dengan jenis augmentasi *horizontal flip*.



Gambar 4.6 Nilai Silhouette Score dengan Jarak GLCM 0.8

Walaupun nilai evaluasinya lebih rendah dari penggunaan jarak GLCM 0.5, jarak 0.8 tetap menunjukkan keunggulan dalam hal distribusi data yang lebih menyebar dan merata ke dalam tiga klaster, terutama menggunakan augmentasi *horizontal flip* sehingga sesuai dengan tujuan awal yaitu membagi ke dalam tiga klaster. Dengan demikian, kombinasi metode GLCM dan PCA dalam penelitian ini mampu menghasilkan pemisahan klaster secara visual dan objektif berdasarkan karakteristik tekstur citra lahan miring.

Berdasarkan hasil tersebut, dilakukan pembuatan rentang nilai kedua sebagai upaya evaluatif untuk menguji kesesuaian batas nilai PC1 dalam menghasilkan pembagian klaster yang lebih optimal pada augmentasi *horizontal flip* dan penggunaan jarak GLCM 0.8. Pembagian klaster tersebut meliputi, klaster 1 dengan nilai kurang dari -2, klaster 1 bernilai diantara -2 hingga 1, dan klaster 2 dengan nilai lebih dari 1.

Saat dilakukan evaluasi menggunakan beberapa jumlah klaster, yaitu dua klaster dan empat klaster, hasil yang diperoleh menunjukkan adanya penurunan kualitas pemisahan klaster. Oleh karena itu, rentang nilai pada tiga klaster masih memberikan hasil yang lebih stabil dan optimal dalam membagi data secara seimbang, serta lebih sesuai dengan karakteristik nilai PC1 pada data tekstur citra lahan miring.

4.4 Integrasi Islam

Selain berkontribusi dalam bidang teknologi dan sains, penelitian ini dapat ditinjau dari perspektif nilai-nilai islam, khususnya dalam hal tanggung jawab manusia pada lingkungan alam dan upaya dalam pencegahan bencana alam.

Integrasi ilmu pengetahuan dengan nilai islam menunjukkan bahwa keduanya tidak ada pertentangan. Dalam islam, ilmu pengetahuan dan teknologi berperan sebagai alat untuk mengaktualisasikan peran manusia sebagai pemelihara bumi. Manusia yang disebut sebagai khalifah di bumi diberikan amanah untuk memelihara, menjaga, dan tidak merusak lingkungan alam yang sudah Allah Swt. ciptakan. Allah bersabda dalam Surah *Ar-Rum* ayat 41 yang berbunyi:

ظَهَرَ الْفَسَادُ فِي الْبَرِّ وَالْبَحْرِ بِمَا كَسَبَتْ أَيْدِي النَّاسِ لِيُذِيقَهُمْ بَعْضَ الَّذِي عَمِلُوا لَعَلَّهُمْ يَرْجِعُونَ ﴿٤١﴾

“Telah tampak kerusakan di darat dan di laut disebabkan perbuatan tangan manusia. (Melalui hal itu) Allah membuat mereka merasakan sebagian dari (akibat) perbuatan mereka agar mereka kembali (ke jalan yang benar).” (Q.S. Ar-Rum: 41).

Ayat tersebut senantiasa mengingatkan pada manusia bahwa kerusakan alam yang terjadi, termasuk bencana alam seperti tanah longsor, dapat terjadi karena kelalaian manusia dalam mengelola alam. Oleh karena itu, penelitian ini dirancang untuk mendeteksi tingkat kerawanan tanah longsor pada lahan miring sebagai bagian dari upaya pencegahan bencana untuk mengelola alam yang sesuai dengan prinsip agama islam (H. Aminudin & Syuhada, 2021).

Segala bentuk usaha pencegahan kerusakan dan menjaga keselamatan jiwa serta lingkungan merupakan bagian dari perintah agama dalam islam. Dengan adanya sistem *clustering* berbasis data dan teknologi ini, diharapkan masyarakat dan pemerintah dapat lebih cepat dalam mengambil keputusan yang tepat dalam mencegah kerusakan. Allah Swt. berfirman dalam Surah Al-Baqarah ayat 205 yang berbunyi:

وَإِذَا تَوَلَّى سَعَى فِي الْأَرْضِ لِيُفْسِدَ فِيهَا وَيُهْلِكَ الْحَرْثَ وَالنَّسْلَ وَاللَّهُ لَا يُحِبُّ الْفُسَادَ ﴿٢٠٥﴾

“Dan apabila ia berpaling dari kamu, ia berjalan di muka bumi untuk mengadakan kerusakan padanya, dan merusak tanam-tanaman dan binatang ternak, dan Allah tidak menyukai kerusakan.” (Q.S Al-Baqoroh: 205).

Ayat tersebut dijelaskan dalam Tafsir Jalalain sebagai larangan berbuat kerusakan di bumi. Pembangunan yang tidak memperhatikan aspek kerawanan bencana dapat dikategorikan sebagai bentuk kerusakan tersebut. Maka, penelitian ini merupakan bentuk nyata dari pelaksanaan nilai-nilai islam untuk menjaga bumi dan tidak membuat kerusakan.

Sejalan dengan hal tersebut, Rasulullah Saw. juga mendorong umatnya untuk berperan aktif dalam menjaga kelestarian lingkungan melalui tindakan yang membawa manfaat, salah satunya yaitu menanam pohon. Dalam mengurangi risiko bencana tanah longsor, peran penanaman pohon sangatlah penting. Dalam sebuah hadis yang diriwayatkan dari Jabir ra, Rasulullah bersabda bahwa setiap muslim menanam pohon dan dimakan hasilnya oleh manusia, hewan, atau burung, maka pahala sedekahnya akan terus mengalir hingga hari kiamat.

وفي رواية لا يَغْرِسُ الْمُسْلِمُ غَرْسًا فَيَأْكُلُ مِنْهُ إِنْسَانٌ وَلَا دَابَّةٌ وَلَا طَيْرٌ إِلَّا كَانَ لَهُ صَدَقَةٌ إِلَى يَوْمِ الْقِيَامَةِ

“Dari sahabat Jabir ra, Rasulullah saw bersabda, ‘Tiada seorang muslim yang menanam pohon, lalu buahnya dimakan oleh seseorang, hewan ternak, atau burung, kecuali itu akan bernilai sedekah sampai hari Kiamat.’ (H.R. Muslim).

Hadis ini mengajarkan bahwa upaya menjaga lingkungan, seperti penghijauan dan penanaman pohon, bukan hanya berdampak ekologis tetapi juga spiritual. Maka dari itu, korelasi penelitian tentang kerawanan tanah longsor yang mengedepankan upaya pelestarian lingkungan merupakan wujud dari ajaran Islam

yang melarang kerusakan dan menganjurkan pelestarian bumi. Diharapkan dengan adanya upaya mitigasi bencana dapat dilakukan dengan penuh kesadaran baik secara ilmiah maupun religius, sehingga terwujud harmoni antara manusia, alam, dan nilai-nilai agama.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini berhasil melakukan pemisahan klaster berdasarkan karakteristik tekstur citra lahan miring melalui tahapan ekstraksi fitur menggunakan metode GLCM dan pengelompokan berbasis visualisasi komponen utama dari PCA. Fitur tekstur GLCM seperti kontras, energi, *dissimilarity*, *homogeneity*, korelasi, dan entropi terbukti mampu merepresentasikan karakteristik permukaan lahan yang relevan untuk membedakan tingkat kerawanan longsor. Hal ini menunjukkan bahwa metode GLCM efektif dalam mengidentifikasi pola tekstur yang berkaitan dengan kerawanan daerah rawan longsor. Proses visualisasi dimensi menggunakan PCA memungkinkan pemetaan fitur ke dalam ruang satu dimensi, yaitu PC1. Komponen ini merupakan komponen utama yang digunakan untuk membentuk tiga klaster. Penelitian ini mengkaji dua parameter jarak GLCM (0.5 dan 0.8 piksel) serta tiga jenis augmentasi citra (rotasi, *horizontal flip*, dan *vertical flip*). Evaluasi hasil clustering menunjukkan bahwa penggunaan jarak GLCM 0.5 memberikan nilai Silhouette Score yang tinggi dan stabil, namun cenderung menghasilkan dua klaster dominan dan satu klaster kosong. Sebaliknya, pada penggunaan jarak GLCM 0.8, meskipun memiliki Silhouette Score lebih rendah, menghasilkan distribusi data yang lebih merata ke dalam tiga klaster, terutama saat dikombinasikan dengan augmentasi *horizontal flip*. Kombinasi parameter tersebut terbukti paling sesuai dengan tujuan penelitian, yaitu mengelompokkan data

menjadi tiga tingkat kerawanan berdasarkan kemiripan pola tekstur. Rentang nilai klaster terbaik berdasarkan rentang nilai PC1 adalah klaster 0 ($-4.3 \leq n < -1.5$), klaster 1 ($-1.5 < n < 1.5$), dan klaster 2 ($1.5 \leq n \leq 4.3$). Skema ini menghasilkan pemisahan klaster yang seimbang dan representatif terhadap distribusi nilai fitur serta sudah optimal berdasarkan nilai *Silhouette Score*-nya. Dengan demikian, metode GLCM-PCA terbukti efektif dalam mengelompokkan tingkat kerawanan longsor berdasarkan citra lahan miring, serta memberikan gambaran awal yang dapat digunakan untuk analisis lebih lanjut dalam mitigasi risiko bencana.

5.2 Saran

Berdasarkan penelitian yang telah dilakukan, terdapat beberapa hal yang dapat disarankan untuk penelitian selanjutnya:

1. Metode GLCM yang digunakan terbukti cukup efektif dalam mengekstraksi fitur tekstur, namun metode ini sangat sensitif terhadap kualitas gambar. Gambar yang mengandung gangguan seperti *noise*, blur, dan resolusi yang rendah dapat mempengaruhi dalam hasil ekstraksi fitur. Oleh karena itu, sangat disarankan agar gambar yang digunakan dalam penelitian selanjutnya memiliki kualitas yang baik.
2. Penelitian ini memiliki potensi untuk dikembangkan lebih lanjut agar memperoleh hasil yang lebih akurat dan stabil. Diharapkan pada penelitian selanjutnya dapat mengkaji dan mengembangkan metode yang lebih optimal dalam menghadapi tantangan tersebut.

DAFTAR PUSTAKA

- Abraham, M. T., Satyam, N., Pradhan, B., & Alamri, A. M. (2020). Forecasting of Landslides Using Rainfall Severity and Soil Wetness: A Probabilistic Approach for Darjeeling Himalayas. *Water*, 12(3), 804. <https://doi.org/10.3390/w12030804>
- Agustin, R. W., Noviandini, F., Darmawan, B. M., & Endarko. (2022). *Evaluation of image resize and thresholding binary on Covid-19 detection using convolutional neural network*. 050001. <https://doi.org/10.1063/5.0103183>
- al-Mahalli, J., & as-Suyuthi, J. (2015). *Tafsir Jalalain* (Cet. 2). Surabaya Pustaka eLBA.
- Almais, A. T. W., Susilo, A., Naba, A., Sarosa, M., Juwono, A. M., Crysdiyan, C., Muslim, M. A., & Wicaksono, H. (2024). Characterization of Structural Building Damage in Post-Disaster Using GLCM-PCA Analysis Integration. *IEEE Access*, 12, 146190–146201. <https://doi.org/10.1109/ACCESS.2024.3469637>
- Antoniuk, I., Kurek, J., Krupa, A., Wieczorek, G., Bukowski, M., Kruk, M., & Jegorowa, A. (2023). Advanced Feature Extraction Methods from Images of Drillings in Melamine Faced Chipboard for Automatic Diagnosis of Drill Wear. *Sensors*, 23(3), 1109. <https://doi.org/10.3390/s23031109>
- Awaluddin, B.-A., Chao, C.-T., & Chiou, J.-S. (2023). Investigating Effective Geometric Transformation for Image Augmentation to Improve Static Hand Gestures with a Pre-Trained Convolutional Neural Network. *Mathematics*, 11(23), 4783. <https://doi.org/10.3390/math11234783>
- Badri, F., & Sari, S. U. R. (2021). Penerapan Metode Principal Component Analysis (PCA) Untuk Identifikasi Faktor-Faktor yang Mempengaruhi Sikap Mahasiswa Memilih Melanjutkan Studi ke Kota Malang. *Building of Informatics, Technology and Science (BITS)*, 3(3), 426–431. <https://doi.org/10.47065/bits.v3i3.1139>
- Bian, Z., Chung, F.-L., & Wang, S. (2021). Fuzzy Density Peaks Clustering. *IEEE Transactions on Fuzzy Systems*, 29(7), 1725–1738. <https://doi.org/10.1109/TFUZZ.2020.2985004>
- Budianta, W., Ohta, H., & Takemura, J. (2022). The Effect of Clay-Soil on Landslide: Case Study from Central Java, Indonesia. *IOP Conference Series: Earth and Environmental Science*, 1091(1), 012012. <https://doi.org/10.1088/1755-1315/1091/1/012012>

- Cantarino, I., Carrion, M. A., Martínez-Ibáñez, V., & Gielen, E. (2023). Improving Landslide Susceptibility Assessment through Frequency Ratio and Classification Methods—Case Study of Valencia Region (Spain). *Applied Sciences*, 13(8), 5146. <https://doi.org/10.3390/app13085146>
- Chakroun, M., Bouhamed, S. A., Kallel, I. K., Solaiman, B., & Derbel, H. (2022). Feature selection based on discriminative power under uncertainty for computer vision applications. *ELCVIA Electronic Letters on Computer Vision and Image Analysis*, 21(1), 111–120. <https://doi.org/10.5565/rev/elcvia.1361>
- Christaki, M., Vasilakos, C., Papadopoulou, E.-E., Tataris, G., Siarkos, I., & Soulakellis, N. (2022). Building Change Detection Based on a Gray-Level Co-Occurrence Matrix and Artificial Neural Networks. *Drones*, 6(12), 414. <https://doi.org/10.3390/drones6120414>
- Condon, P., Simpson, J., & Emanuel, M. (2022). Research data integrity: A cornerstone of rigorous and reproducible research. *IASSIST Quarterly*, 46(3). <https://doi.org/10.29173/iq1033>
- Dandridge, C., Stanley, T. A., Kirschbaum, D. B., & Lakshmi, V. (2023). Spatial and Temporal Analysis of Global Landslide Reporting Using a Decade of the Global Landslide Catalog. *Sustainability*, 15(4), 3323. <https://doi.org/10.3390/su15043323>
- Dwaich, H. A., & Abdulbaqi, H. A. (2021). Signature Texture Features Extraction Using GLCM Approach in Android Studio. *Journal of Physics: Conference Series*, 1804(1), 012043. <https://doi.org/10.1088/1742-6596/1804/1/012043>
- Embersson, R., Kirschbaum, D., & Stanley, T. (2020). *New Global Characterization of Landslide Exposure*. <https://doi.org/10.5194/nhess-2019-434>
- Eric U., O., & Michael O., O. (2024). Overview of Agglomerative Hierarchical Clustering Methods. *British Journal of Computer, Networking and Information Technology*, 7(2), 14–23. <https://doi.org/10.52589/BJCNIT-CV9POOGW>
- Faisal, E., Nugroho, A., Basuki, R. S., & Suharnawi, S. (2022). GLCM Based Locally Feature Extraction On Natural Image. *Journal of Applied Intelligent System*, 7(2), 128–134. <https://doi.org/10.33633/jais.v7i2.6569>
- Fata, Y. A., Hendrayanto, H., & Febrianti, N. (2022). Landslide frequency and its relationship with urban development in landform above groundwater basin area of Bogor, Indonesia. *Journal of Degraded and Mining Lands Management*, 9(3), 3561–3572. <https://doi.org/10.15243/jdmlm.2022.093.3561>

- Feng, G., Lin, J., & Wang, K. (2022). Researches Advanced in Clustering Algorithms. *Highlights in Science, Engineering and Technology*, 16, 168–177. <https://doi.org/10.54097/hset.v16i.2498>
- Fidan, S., Tanyaş, H., Akbaş, A., Lombardo, L., Petley, D. N., & Görüm, T. (2024). Understanding fatal landslides at global scales: A summary of topographic, climatic, and anthropogenic perspectives. *Natural Hazards*, 120(7), 6437–6455. <https://doi.org/10.1007/s11069-024-06487-3>
- Gakaev, R. (2021). Exogenous relief-forming processes and phenomena on the territory of the Chechen Republic. *SHS Web of Conferences*, 128, 03001. <https://doi.org/10.1051/shsconf/202112803001>
- Goyal, D., Joshi, V., Gupta, N., & Cabral-Pinto, M. M. S. (2022). Soil Quality Assessment in a Landslide Chronosequence of Indian Himalayan Region. *Land*, 11(10), 1819. <https://doi.org/10.3390/land11101819>
- H. Aminudin, & Syuhada, H. (2021). *Al-Qur'an Hadis Madrasah Aliyah Kelas XII*. Bumi Aksara.
- Hasan, Y. (2024). PENGUKURAN SILHOUETTE SCORE DAN DAVIES-BOULDIN INDEX PADA HASIL CLUSTER K-MEANS DAN DBSCAN. *Jurnal Informatika Dan Teknik Elektro Terapan*, 12(3S1). <https://doi.org/10.23960/jitet.v12i3S1.5001>
- Hendrastuty, N. (2024). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa. *Jurnal Ilmiah Informatika Dan Ilmu Komputer (JIMA-ILKOM)*, 3(1), 46–56. <https://doi.org/10.58602/jima-ilkom.v3i1.26>
- Iqbal, N., Mumtaz, R., Shafi, U., & Zaidi, S. M. H. (2021). Gray level co-occurrence matrix (GLCM) texture based crop classification using low altitude remote sensing platforms. *PeerJ Computer Science*, 7, e536. <https://doi.org/10.7717/peerj-cs.536>
- Kab, A., Djerbal, L., & Bahar, R. (2023). Implementation of PCA multicollinearity method to landslide susceptibility assessment: The study case of Kabylia region. *Arabian Journal of Geosciences*, 16(4), 291. <https://doi.org/10.1007/s12517-023-11374-5>
- Karakuzulu, C., Gümüş, İ. H., Güldal, S., & Yavaş, M. (2023). Determining The Number of Principal Components with Schur's Theorem in Principal Component Analysis. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 12(2), 299–306. <https://doi.org/10.17798/bitlisfen.1144360>

- Kebaili, A., Lapuyade-Lahorgue, J., & Ruan, S. (2023). Deep Learning Approaches for Data Augmentation in Medical Imaging: A Review. *Journal of Imaging*, 9(4), 81. <https://doi.org/10.3390/jimaging9040081>
- Kim, S., & Yoon, H.-K. (2023). Application of classification coupled with PCA and SMOTE, for obtaining safety factor of landslide based on HRA. *Bulletin of Engineering Geology and the Environment*, 82(10), 381. <https://doi.org/10.1007/s10064-023-03403-0>
- Lamer, A., Fruchart, M., Paris, N., Popoff, B., Payen, A., Balcaen, T., Gacquer, W., Bouzillé, G., Cuggia, M., Doutreligne, M., & Chazard, E. (2022). Standardized Description of the Feature Extraction Process to Transform Raw Data Into Meaningful Information for Enhancing Data Reuse: Consensus Study. *JMIR Medical Informatics*, 10(10), e38936. <https://doi.org/10.2196/38936>
- Mohammed, M. A., Akawee, M. M., Saleh, Z. H., Hasan, R. A., Ali, A. H., & Sutikno, T. (2023). The effectiveness of big data classification control based on principal component analysis. *Bulletin of Electrical Engineering and Informatics*, 12(1), 427–434. <https://doi.org/10.11591/eei.v12i1.4405>
- Nguyen, K. (2023). DATA AUGMENTATION ANALYSIS OF VEHICLE DETECTION IN AERIAL IMAGES. *Journal of Computer Science and Cybernetics*, 291–312. <https://doi.org/10.15625/1813-9663/18259>
- Nogales, R. E., & Benalcázar, M. E. (2023). Hand Gesture Recognition Using Automatic Feature Extraction and Deep Learning Algorithms with Memory. *Big Data and Cognitive Computing*, 7(2), 102. <https://doi.org/10.3390/bdcc7020102>
- Pang, D., Liu, G., He, J., Li, W., & Fu, R. (2022). Automatic Remote Sensing Identification of Co-Seismic Landslides Using Deep Learning Methods. *Forests*, 13(8), 1213. <https://doi.org/10.3390/f13081213>
- Polyakova, M. V., & Krylov, V. N. (2022). Data normalization methods to improve the quality of classification in the breast cancer diagnostic system. *Applied Aspects of Information Technology*, 5(1), 55–63. <https://doi.org/10.15276/aait.05.2022.5>
- Priyono, K. D. (2020). RISK ANALYSIS OF LANDSLIDE IMPACTS ON SETTLEMENTS IN KARANGANYAR, CENTRAL JAVA, INDONESIA. *International Journal of GEOMATE*, 19(73). <https://doi.org/10.21660/2020.73.34128>
- Rupanka Bhuyan & Samarjeet Borah. (2013). A Survey of Some Density Based Clustering Techniques. <https://doi.org/10.13140/2.1.4554.6887>

- Saccenti, E. (2024). A gentle introduction to principal component analysis using tea-pots, dinosaurs, and pizza. *Teaching Statistics*, 46(1), 38–52. <https://doi.org/10.1111/test.12363>
- Salunkhe, S., Bachute, M., Gite, S., Vyas, N., Khanna, S., Modi, K., Katpatal, C., & Kotecha, K. (2021). Classification of Alzheimer's Disease Patients Using Texture Analysis and Machine Learning. *Applied System Innovation*, 4(3), 49. <https://doi.org/10.3390/asi4030049>
- Saputra, R. A., Suharyanto, Wasiyanti, S., Saefudin, D. F., Supriyatna, A., & Wibowo, A. (2020). Rice Leaf Disease Image Classifications Using KNN Based On GLCM Feature Extraction. *Journal of Physics: Conference Series*, 1641, 012080. <https://doi.org/10.1088/1742-6596/1641/1/012080>
- Shahabi, H., Rahimzad, M., Tavakkoli Piralilou, S., Ghorbanzadeh, O., Homayouni, S., Blaschke, T., Lim, S., & Ghamisi, P. (2021). Unsupervised Deep Learning for Landslide Detection from Multispectral Sentinel-2 Imagery. *Remote Sensing*, 13(22), Article 22. <https://doi.org/10.3390/rs13224698>
- Suhaidi, M., Abdul Kadir, R., & Tiun, S. (2021). A REVIEW OF FEATURE EXTRACTION METHODS ON MACHINE LEARNING. *Journal of Information System and Technology Management*, 6(22), 51–59. <https://doi.org/10.35631/JISTM.622005>
- Surve, Miss. R. V., Chavan, Miss. N. A., Kamble, Miss. P. S., Bidkar, Miss. G. S., & Chakote, Prof. M. S. (2022). Green Path Project Develops Using Image Processing. *International Journal for Research in Applied Science and Engineering Technology*, 10(5), 833–838. <https://doi.org/10.22214/ijraset.2022.42243>
- Suryani, I., Nontji, M., & Juita, N. (2021). Morphological characteristics and classification of inceptisol in Mamuju regency, West Sulawesi. *IOP Conference Series: Earth and Environmental Science*, 807(4), 042043. <https://doi.org/10.1088/1755-1315/807/4/042043>
- Susetyo, J. A., Kurnianto, F. A., Nurdin, E. A., & Pangastuti, E. I. (2022). Landslide Disaster Mapping in Silo District, Jember Regency. *IOP Conference Series: Earth and Environmental Science*, 975(1), 012011. <https://doi.org/10.1088/1755-1315/975/1/012011>
- Suwanda Risky, Syafriandi, Dony Permana, & Dina Fitria. (2023). Self Organizing Maps Method for Grouping Provinces in Indonesia Based on the Landslide Impact. *UNP Journal of Statistics and Data Science*, 1(3), 180–187. <https://doi.org/10.24036/ujsds/vol1-iss3/15>

- Tavus, B., & Kocaman, S. (2023). FLOOD MAPPING IN MOUNTAINOUS AREAS USING SENTINEL-1 & 2 DATA AND GLCM FEATURES. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-1/W2-2023, 1575–1580. <https://doi.org/10.5194/isprs-archives-XLVIII-1-W2-2023-1575-2023>
- Thangakani, S. A., Sornam, M., & Kavitha, M. (2022, January 2). *Deep Convolutional Neural Network for Skin Lesion Detection with HOG and GLCM features*. | EBSCOhost. <https://openurl.ebsco.com/contentitem/gcd:157171364?sid=ebsco:plink:crawler&id=ebsco:gcd:157171364>
- Troccoli, E. B., Cerqueira, A. G., Lemos, J. B., & Holz, M. (2022). K-means clustering using principal component analysis to automate label organization in multi-attribute seismic facies analysis. *Journal of Applied Geophysics*, 198, 104555. <https://doi.org/10.1016/j.jappgeo.2022.104555>
- Urumov, G., Chountas, P., & School of Computer Science and Engineering, University of Westminster. (2022). Clustering stock price volatility using intuitionistic fuzzy sets. *Notes on Intuitionistic Fuzzy Sets*, 28(3), 343–352. <https://doi.org/10.7546/nifs.2022.28.3.343-352>
- Wang, G., Bo, F., Chen, X., Lu, W., Hu, S., & Fang, J. (2022). A Collaborative Despeckling Method for SAR Images Based on Texture Classification. *Remote Sensing*, 14(6), 1465. <https://doi.org/10.3390/rs14061465>
- Wangiyana, S., Samczyński, P., & Gromek, A. (2022). Data Augmentation for Building Footprint Segmentation in SAR Images: An Empirical Study. *Remote Sensing*, 14(9), 2012. <https://doi.org/10.3390/rs14092012>
- Wibowo, A. (2025, January 6). malang-post.com—Tanah Longsor Dominasi Kejadian Bencana Alam di Kota Batu. *Malang-Post.Com*. <https://malang-post.com/2025/01/06/tanah-longsor-dominasi-kejadian-bencana-alam-di-kota-batu/>
- Xu, Y., Zhu, H., Hu, C., Liu, H., & Cheng, Y. (2022). Deep learning of DEM image texture for landform classification in the Shandong area, China. *Frontiers of Earth Science*, 16(2), 352–367. <https://doi.org/10.1007/s11707-021-0884-y>
- Yu, J., Wen, H., Huang, J., & Liu, B. O. (2024). *Deep Learning-Based Landslides Detection Improving by Gray Level Co-Occurrence Matrix of Rgb Optical Images* (SSRN Scholarly Paper 4870257). Social Science Research Network. <https://doi.org/10.2139/ssrn.4870257>
- Zangana, H. M., Mohammed, A. K., Sallow, Z. B., & Mustafa, F. M. (2024). Exploring Image Representation and Color Spaces in Computer Vision: A

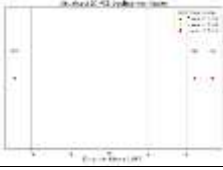
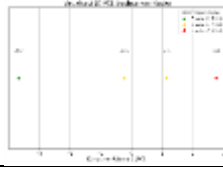
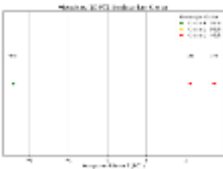
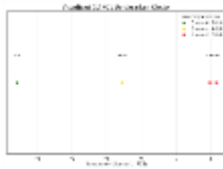
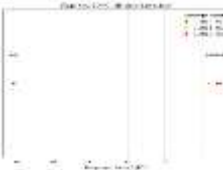
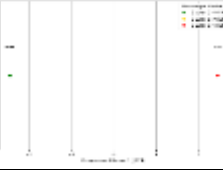
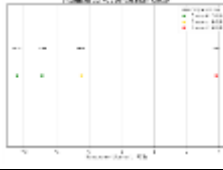
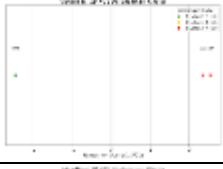
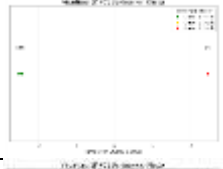
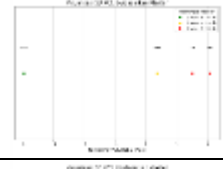
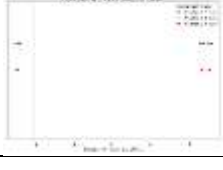
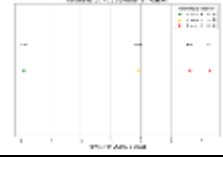
Comprehensive Review. *Indonesian Journal of Computer Science*, 13(3).
<https://doi.org/10.33022/ijcs.v13i3.3998>

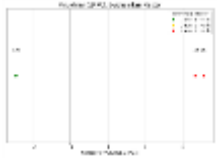
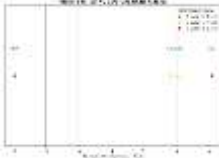
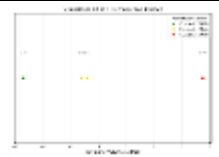
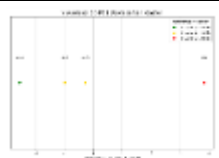
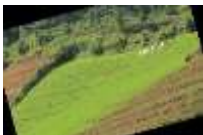
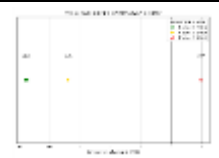
Zheng, W. (2023). Current Technologies and Applications of Digital Image Processing. *Journal of Biomedical and Sustainable Healthcare Applications*, 13–23. <https://doi.org/10.53759/0088/JBSHA202303002>

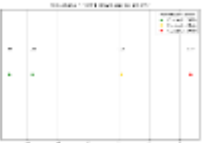
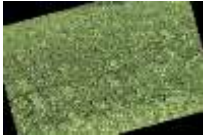
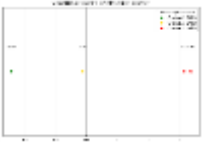
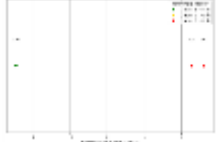
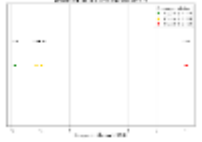
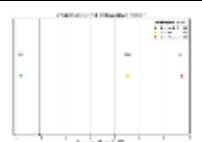
Zhuang, H., Wang, H., & Ji, Z. (2022). findPC: An R package to automatically select the number of principal components in single-cell analysis. *Bioinformatics*, 38(10), 2949–2951.
<https://doi.org/10.1093/bioinformatics/btac235>

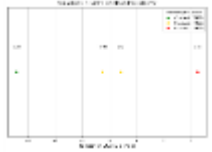
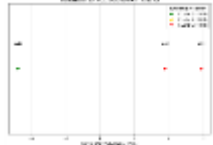
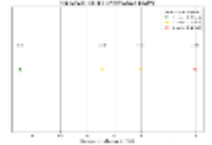
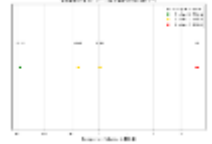
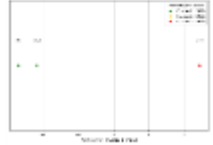
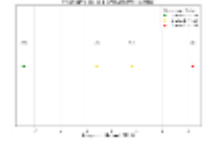
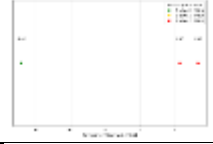
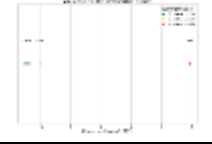
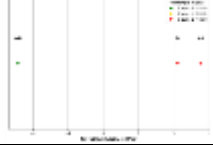
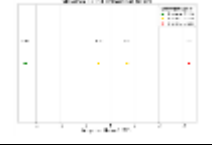
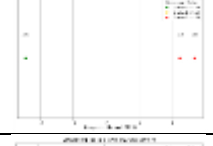
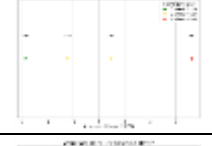
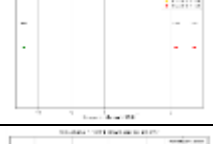
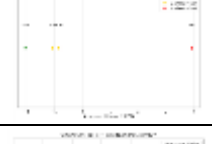
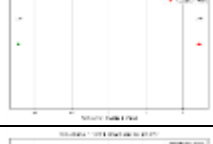
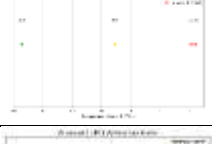
LAMPIRAN

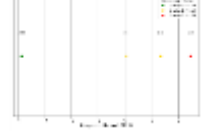
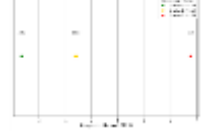
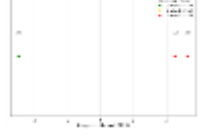
Lampiran 1 Hasil *Clustering* pada Augmentasi Rotasi 15°

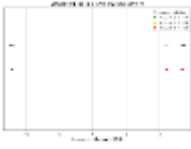
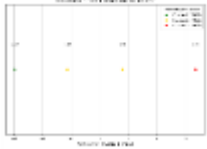
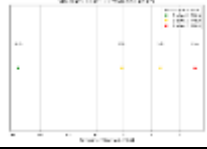
No	Gambar	Visualisasi 1D	
		Jarak GLCM 0.5	Jarak GLCM 0.8
1.			
2.			
3.			
4.			
5.			
6.			
7.			
8.			

9.			
10.			
11.			
12.			
13.			
14.			
15.			
16.			
17.			
18.			

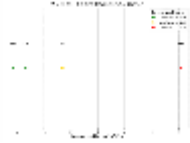
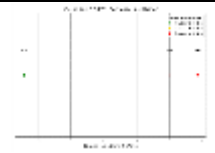
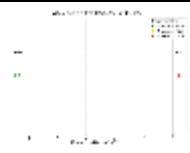
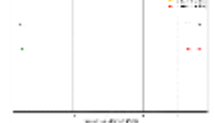
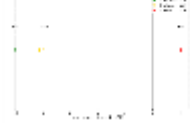
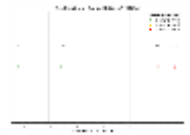
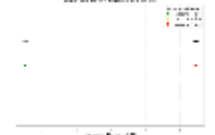
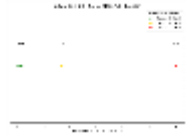
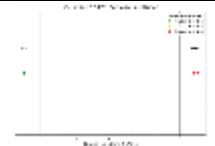
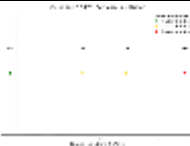
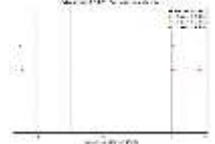
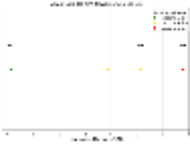
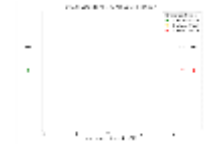
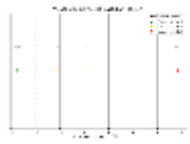
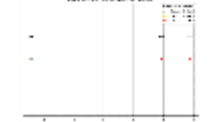
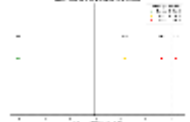
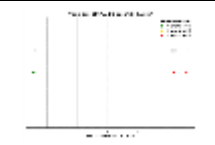
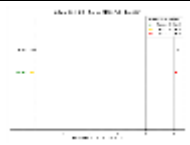
19.			
20.			
21.			
22.			
23.			
24.			
25.			
26.			
27.			
28.			

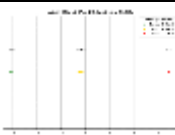
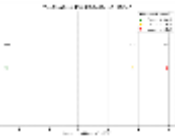
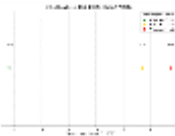
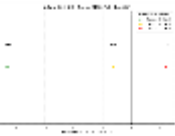
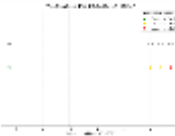
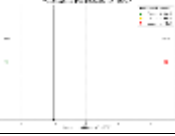
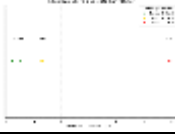
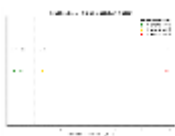
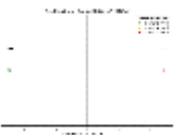
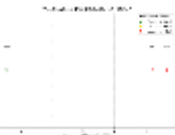
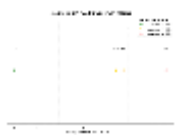
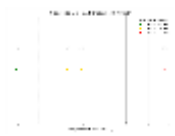
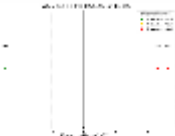
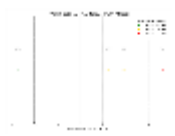
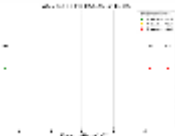
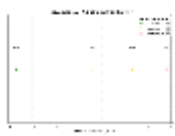
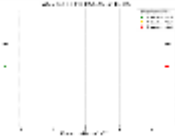
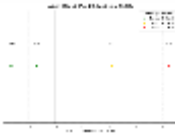
29.			
30.			
31.			
32.			
33.			
34.			
35.			
36.			
37.			
38.			

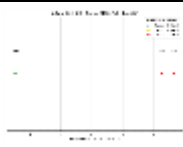
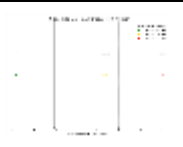
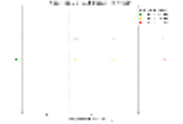
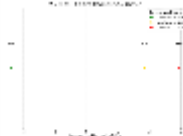
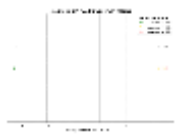
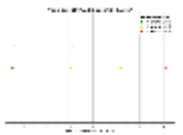
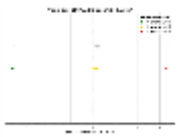
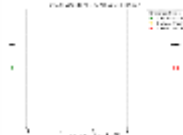
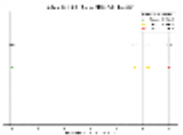
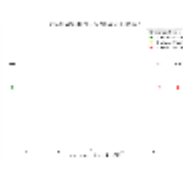
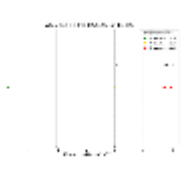
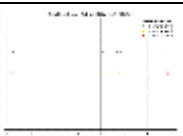
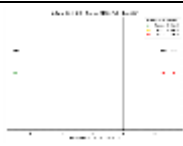
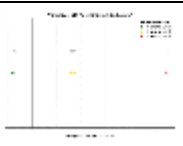
39.			
40.			
41.			
42.			
43.			
44.			
45.			
46.			
47.			
48.			

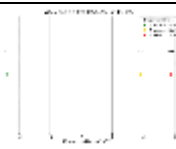
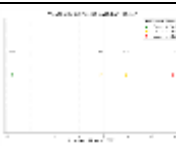
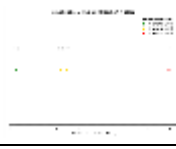
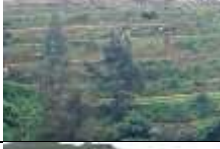
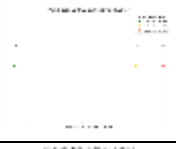
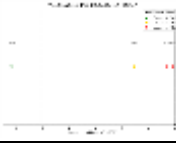
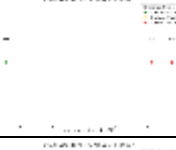
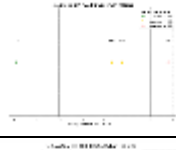
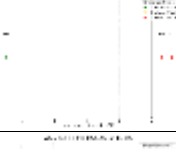
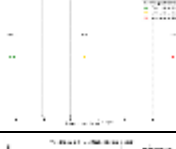
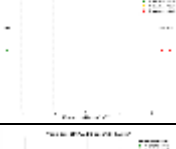
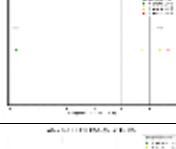
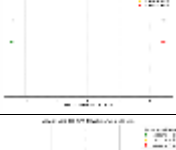
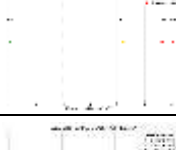
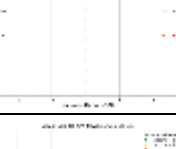
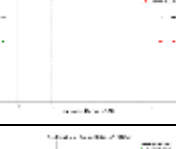
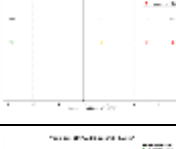
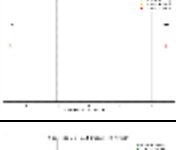
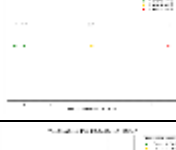
49.			
50.			

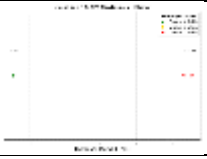
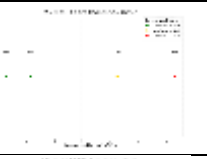
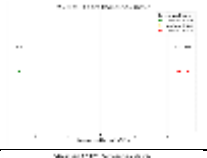
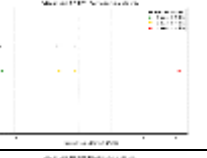
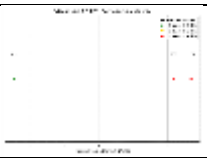
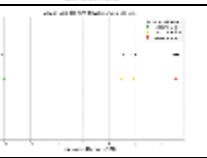
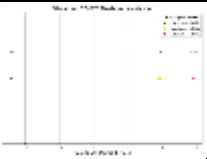
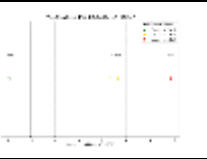
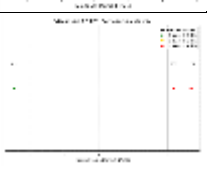
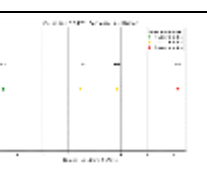
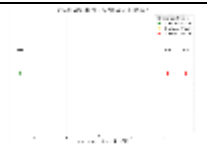
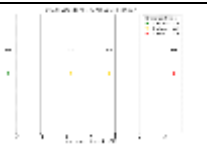
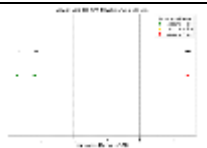
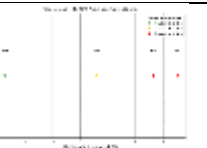
Lampiran 2 Hasil *Clustering* pada Augmentasi *Horizontal Flip*

No	Gambar	Visualisasi 1D	
		Jarak GLCM 0.5	Jarak GLCM 0.8
1.			
2.			
3.			
4.			
5.			
6.			
7.			
8.			
9.			
10.			

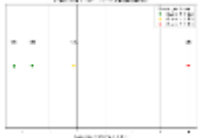
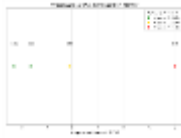
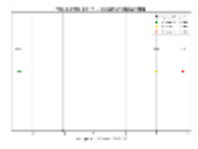
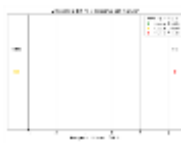
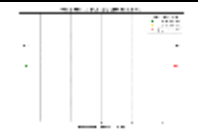
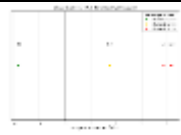
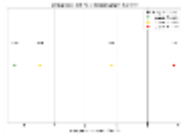
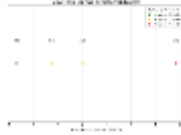
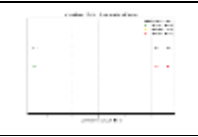
11.			
12.			
13.			
14.			
15.			
16.			
17.			
18.			
19.			
20.			
21.			

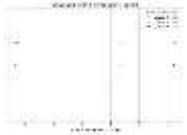
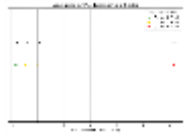
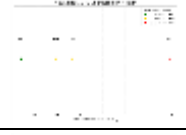
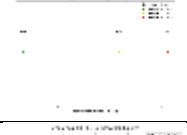
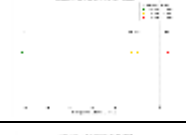
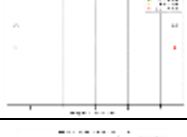
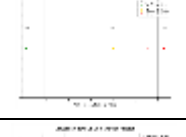
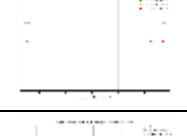
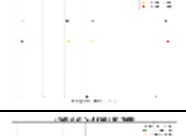
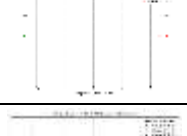
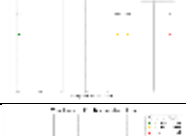
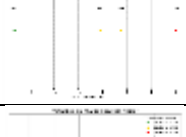
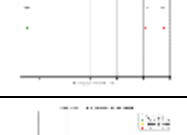
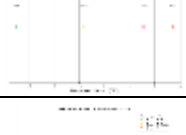
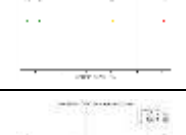
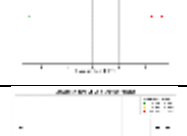
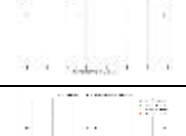
22.			
23.			
24.			
25.			
26.			
27.			
28.			
29.			
30.			
31.			
32.			

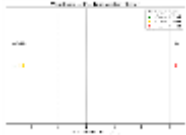
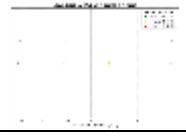
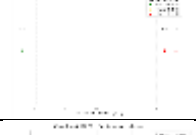
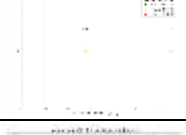
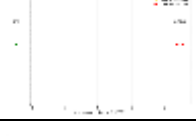
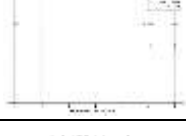
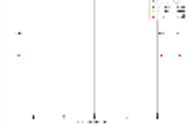
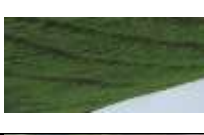
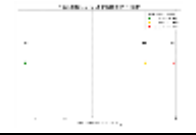
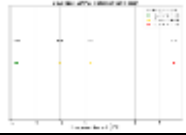
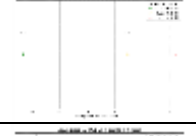
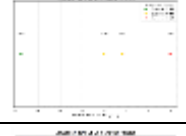
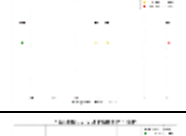
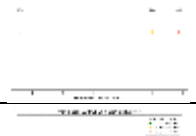
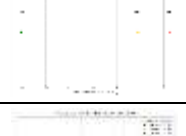
33.			
34.			
35.			
36.			
37.			
38.			
39.			
40.			
41.			
42.			
43.			

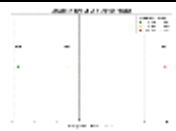
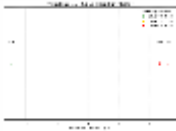
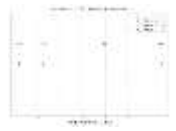
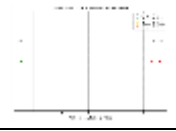
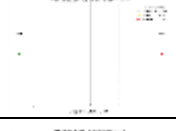
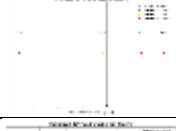
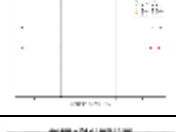
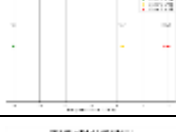
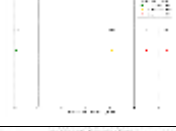
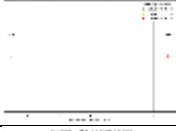
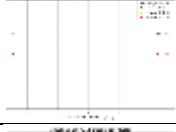
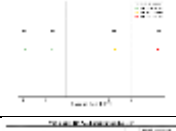
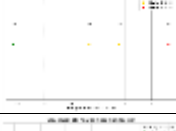
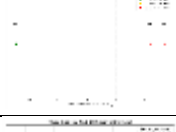
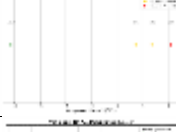
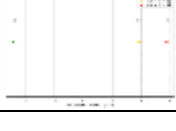
44.			
45.			
46.			
47.			
48.			
49.			
50.			

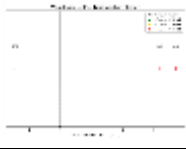
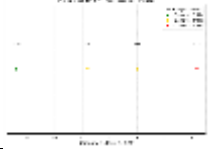
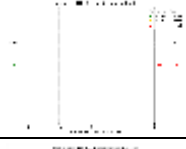
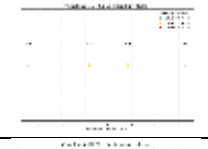
Lampiran 3 Hasil *Clustering* pada Augmentasi *Vertical Flip*

No	Gambar	Visualisasi 1D	
		Jarak GLCM 0.5	Jarak GLCM 0.8
1.			
2.			
3.			
4.			
5.			
6.			
7.			
8.			
9.			
10.			
11.			

12.			
13.			
14.			
15.			
16.			
17.			
18.			
19.			
20.			
21.			
22.			
23.			

24.			
25.			
26.			
27.			
28.			
29.			
30.			
31.			
32.			
33.			
34.			
35.			

36.			
37.			
38.			
39.			
40.			
41.			
42.			
43.			
44.			
45.			
46.			
47.			

48.			
49.			
50.		