

***DECISION SUPPORT SYSTEM UNTUK MENGETAHUI JALAN
BERLUBANG MENGGUNAKAN AUGMENTASI CITRA
BERBASIS WATERSHED DAN PCA CLUSTERING***

SKRIPSI

Oleh :
INTAN TIARA DEWI
NIM. 210605110084



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

***DECISION SUPPORT SYSTEM UNTUK MENGETAHUI JALAN
BERLUBANG MENGGUNAKAN AUGMENTASI CITRA
BERBASIS WATERSHED DAN PCA CLUSTERING***

SKRIPSI

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

Oleh :
INTAN TIARA DEWI
NIM. 210605110084

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

HALAMAN PERSETUJUAN

***DECISION SUPPORT SYSTEM UNTUK MENGETAHUI JALAN
BERLUBANG MENGGUNAKAN AUGMENTASI CITRA
BERBASIS WATERSHED DAN PCA CLUSTERING***

SKRIPSI

Oleh :
INTAN TIARA DEWI
NIM. 210605110084

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 22 April 2025

Pembimbing I,



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

Pembimbing II



Dr. Ir. Fresy Nugroho, MT, IPM
NIP. 19710722 201101 1 001

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



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

HALAMAN PENGESAHAN

DECISION SUPPORT SYSTEM UNTUK MENGETAHUI JALAN BERLUBANG MENGGUNAKAN AUGMENTASI CITRA BERBASIS *WATERSHED* DAN *PCA CLUSTERING*

SKRIPSI

Oleh :
INTAN TIARA DEWI
NIM. 210605110084

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal: 16 Mei 2025

Susunan Dewan Penguji

Ketua Penguji	: <u>Suprivono, M.Kom</u> NIP. 19841010 201903 1 012
Anggota Penguji I	: <u>Nur Fitriyah Ayu Tunjung Sari, M.Cs</u> NIP. 19911226 202012 2 001
Anggota Penguji II	: <u>Dr. Agung Teguh Wibowo Almais, M.T</u> NIP. 19860301 202321 1 016
Anggota Penguji III	: <u>Dr. Ir. Fresy Nugroho, MT, IPM</u> NIP. 19710722 201101 1 001



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




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

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Intan Tiara Dewi
NIM : 210605110084
Fakultas / Program Studi : Sains dan Teknologi / Teknik Informatika
Judul Skripsi : *Decision Support System* untuk Mengetahui Jalan
Berlubang Menggunakan Augmentasi Citra
Berbasis *Watershed* dan *PCA Clustering*

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, 16 Mei 2025

Yang membuat pernyataan,



Intan Tiara Dewi

NIM.210605110084

MOTTO

”Setetes keringat orang tuaku yang keluar, ada seribu langkahku untuk maju”

HALAMAN PERSEMBAHAN

Segala puji dan syukur penulis panjatkan ke hadirat Allah Subhanahu wa ta'ala atas limpahan rahmat, hidayah, serta kekuatan-Nya, sehingga penulis dapat menyelesaikan penyusunan skripsi ini.

Dengan penuh hormat dan cinta, penulis dedikasikan karya ini kepada:

Ayah dan Ibu tercinta, Bapak Supri Handoyo dan Ibu Duwi Novitaningrum, atas cinta, doa, serta pengorbanan yang tiada henti dan selalu mengalir tanpa batas.

Keluarga tersayang,
yang senantiasa memberikan dukungan dan semangat di setiap perjalanan.

Dosen pembimbing serta rekan-rekan seperjuangan,
atas ilmu, arahan, dan dukungan yang sangat berarti sepanjang proses ini.

Untuk diri sendiri,
terima kasih telah bertahan dan terus maju meski menghadapi berbagai rintangan.

Semoga karya ini dapat menjadi langkah awal yang membawa manfaat bagi banyak pihak.

KATA PENGANTAR

Assalamu'alaikum warahmatullahi wabarakatuh

Segala puji bagi Allah Subhanahu wa ta'ala, atas limpahan rahmat dan petunjuk-Nya, yang telah memberikan kekuatan serta kemudahan kepada penulis dalam menyelesaikan skripsi berjudul "*Decision Support System* untuk mengetahui jalan berlubang menggunakan Augmentasi Citra berbasis *Watershed* dan *PCA Clustering*" dengan baik dan lancar. Shalawat serta salam senantiasa tercurah kepada Nabi Muhammad SAW, yang telah membawa umat manusia dari zaman kegelapan menuju cahaya Islam yang penuh berkah dan petunjuk.

Skripsi ini adalah buah dari perjalanan panjang dalam menyelesaikan tugas akhir sebagai bagian dari persyaratan kelulusan di Program Studi Teknik Informatika, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Selama proses penyusunannya, penulis mendapat banyak dukungan dari berbagai pihak. Sebagai wujud penghargaan, penulis mengucapkan terima kasih 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 sabar dan tulus membimbing penulis sepanjang proses penyusunan

skripsi ini. Bimbingan, arahan, serta dukungan beliau tidak hanya bersifat ilmiah, tetapi juga penuh inspirasi dan semangat yang membangkitkan motivasi penulis untuk terus berusaha.

5. Bapak Dr. Ir. Fresy Nugroho, M.T, IPM., selaku Dosen Pembimbing II, atas bimbingan dan pendapat yang diberikan kepada penulis dalam proses penulisan skripsi dan integrasi keilmuan Islam.
6. Bapak Supriyono, M. Kom., selaku Dosen Penguji I, dan Ibu Nur Fitriyah Ayu Tunjung Sari, M.C S selaku Dosen Penguji II, atas segala masukan dan saran berharga yang telah diberikan dalam penyempurnaan skripsi ini.
7. Seluruh dosen, staf administrasi, serta seluruh jajaran di Program Studi Teknik Informatika atas ilmu, dukungan, dan arahnya.
8. Saya ucapkan terima kasih kepada Bapak dan Ibu saya, Bapak Supri Handoyo dan Ibu Duwi Novitaningrum, dua sosok luar biasa yang memiliki peran besar dalam perjalanan hidup saya. Mereka adalah orang-orang yang selalu berusaha sekuat tenaga agar saya, anak pertama mereka, dapat menempuh pendidikan hingga ke tahap ini. Untuk bapak, terima kasih atas setiap tetes keringat dan kerja keras yang telah bapak korbankan demi memenuhi kebutuhan dan pendidikan saya, serta atas keteladanan yang bapak tunjukkan dalam membentuk saya menjadi perempuan yang tangguh. Untuk ibu, terima kasih atas segala motivasi, nasihat, doa, dan harapan yang selalu mengiringi setiap langkah dan usaha saya dalam menempuh pendidikan. Terima kasih atas cinta kasih yang tak pernah pudar oleh waktu, atas kesabaran dan pengorbanan yang tanpa henti mengiringi perjalanan hidup saya. Ibu adalah

sumber kekuatan, inspirasi, dan cahaya yang senantiasa menerangi setiap jalan yang saya lalui. Akhir kata, saya mengucapkan terima kasih atas semua yang telah Bapak dan Ibu berikan, yang nilainya tak terhingga.

9. Ucapan terima kasih juga saya sampaikan kepada kedua adik saya tersayang, Thalita Salsabilla Putri dan Ahmad Fadhil. Terima kasih telah menjadi bagian dari perjalanan hidup saya, menghadirkan tawa, semangat, dan warna dalam setiap hari-hari saya.
10. Penulis juga berterima kasih kepada teman-teman seperjuangan, khususnya keluarga Teknik Informatika Angkatan 21 "Aster", penghuni Lab Database aka Bimbingan Pak Agung (Ima, Najah, Aulia, Putri, Gigih & Roi), keluarga kos aka Pembela Palestina (Raniah, Nenden, Nuor, Nopa, Dau & Lala), teman-teman Disney Lampuz (Nurul, Jemblem, Nopia & Epa), teman-teman Minggu Momen (Mida, Charles & Irham), teman-teman S3 aka Serba Serbi Sambat (Kimal & Jati), teman-teman Ok Gas No Nyd (Kaik & Maji), Kamar 54 Ma'had Fatimah Az-Zahra "54 Asik" dan tim PKL Telkom yang telah memberikan semangat, kenangan, dan kebersamaan yang berarti sepanjang perjalanan ini. Semoga kebaikan dan dukungan kalian semua mendapat balasan yang terbaik dari Allah Subhanahu wa ta'ala. Penulis menyadari bahwa karya ini masih jauh dari sempurna, sehingga segala kritik dan saran yang membangun sangat diharapkan demi penyempurnaan di masa mendatang.

11. Kepada semua pihak dan rekan-rekan penulis yang tidak dapat disebutkan satu per satu, yang telah memberikan dukungan serta bantuan secara tidak langsung dalam penyusunan skripsi ini, penulis menyampaikan terima kasih.

Sebagai akhir dari tulisan ini, penulis menyadari bahwa skripsi ini masih belum sempurna dan masih terdapat berbagai kekurangan. Meskipun demikian, penulis tetap berharap karya ini dapat memberikan manfaat, baik bagi pembaca maupun bagi penulis sendiri.

Malang, 29 April 2025

Penulis

DAFTAR ISI

HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN.....	iv
PERNYATAAN KEASLIAN TULISAN	v
MOTTO	vi
HALAMAN PERSEMBAHAN	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	xii
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xiv
ABSTRAK	xv
ABSTRACT	xvi
الملخص	xvii
BAB I PENDAHULUAN.....	1
BAB II STUDI PUSTAKA.....	8
2.1 Penelitian Terdahulu	8
2.2 Kerusakan Jalan	10
2.3 Metode.....	12
2.3.1 Decision Support System	12
2.3.2 Augmentasi Data Image space Transformation.....	13
2.3.3 Watershed.....	14
2.3.4 PCA Clustering.....	14
BAB III METODE PENELITIAN	15
3.1 Desain Sistem.....	15
3.1.1 Akuisisi dan Pra-Pemrosesan Data.....	16
3.1.2 Segmentasi Citra dengan Metode Watershed	17
3.1.3 Ekstraksi Fitur dan Penyimpanan Data	17
3.1.4 Analisis Data dengan <i>Principal Component Analysis</i> (PCA).....	17
3.1.5 Hasil Akhir dan Interpretasi	18
3.2 <i>Data Collection</i>	18
3.3 Implementasi Metode <i>Watershed</i> dan <i>Principal Component Analysis</i>	19
3.4 Evaluasi Model.....	25
BAB IV HASIL DAN PEMBAHASAN	27
4.1 Langkah Pengujian.....	27
4.2 Hasil Pengujian	27
4.2.1 Proses Watershed.....	28
4.2.2 Proses PCA.....	33
4.3 Augmentasi <i>Clustering</i>	48
4.4 Pembahasan.....	51
BAB IV KESIMPULAN DAN SARAN	54
5.1 Kesimpulan	54
5.2 Saran.....	54

DAFTAR GAMBAR

Gambar 3. 1 Desain Sistem.....	16
Gambar 4. 1 Hasil pra-pemrosesan citra	29
Gambar 4. 2 Hasil citra <i>threshold</i>	30
Gambar 4. 3 Hasil citra <i>Sure background</i>	31
Gambar 4. 4 Hasil citra <i>distance tranform</i>	31
Gambar 4. 5 Hasil citra <i>Sure Foreground</i>	32
Gambar 4. 6 Hasil citra <i>Unknown region</i>	32
Gambar 4. 7 Hasil normalisasi data	33
Gambar 4. 8 Visualisasi Grafik 2D	37
Gambar 4. 9 Visualisasi Grafik 3D	38
Gambar 4. 10 Hasil <i>Clustering 3 Cluster</i> Grafik 2D	39
Gambar 4. 11 Hasil <i>Clustering 3 Cluster</i> Grafik 3D	39
Gambar 4. 12 Hasil <i>Clustering 4 Cluster</i> Grafik 2D	40
Gambar 4. 13 Hasil <i>Clustering 4 Cluster</i> Grafik 3D.....	40
Gambar 4. 14 Hasil <i>Clustering 5 Cluster</i> Grafik 2D	41
Gambar 4. 15 Hasil <i>Clustering 5 Cluster</i> Grafik 3D	41

DAFTAR TABEL

Table 2. 1 Penelitian Terdahulu	8
Tabel 4. 1 Nilai Rasio dan Nilai Eigen	34
Tabel 4. 2 Nilai Rentang (n) tingkat kerusakan jalan.....	39
Tabel 4. 3 Rentang Nilai (n) dalam 4 <i>Cluster</i>	41
Tabel 4. 4 Rentang Nilai (n) dalam 5 <i>Cluster</i>	42
Tabel 4. 5 Nilai <i>Silhouette Score</i> Tiap Data dari Grafik 2D	44
Tabel 4. 6 Nilai <i>Silhouette Score</i> Tiap Data dari Grafik 3D	45
Tabel 4. 7 Hasil Augmentasi <i>Watershed</i>	49

ABSTRAK

Dewi, Intan Tiara. 2025. ***Decision Support System untuk Mengetahui Jalan Berlubang Menggunakan Augmentasi Citra Berbasis Watershed dan PCA Clustering.*** Skripsi. Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Agung Teguh Wibowo Almais, M.T (II) Dr. Fresy Nugroho, M.T

Kata Kunci: Augmentasi Citra, *Clustering*, *Decision Support System*, Jalan Berlubang, *Principal Component Analysis*, *Silhouette Score*, *Watershed*.

Jalan berlubang merupakan salah satu bentuk kerusakan infrastruktur yang dapat membahayakan keselamatan pengguna jalan serta menghambat kelancaran lalu lintas. Proses identifikasi kerusakan jalan secara manual membutuhkan waktu dan sumber daya yang besar serta rentan terhadap kesalahan manusia. Penelitian ini bertujuan untuk mengembangkan sistem pendukung keputusan berbasis pengolahan citra digital untuk mendeteksi dan mengelompokkan tingkat kerusakan jalan berlubang secara otomatis dan efisien. Metode yang digunakan adalah augmentasi citra menggunakan teknik *image space transformation*, *segmentasi* dengan algoritma *Watershed*, serta pengelompokan hasil *segmentasi* menggunakan *Principal Component Analysis* (PCA). PCA digunakan untuk mereduksi dimensi data citra dan mengelompokkan tingkat kerusakan menjadi tiga kategori: ringan, sedang, dan berat. Evaluasi dilakukan menggunakan metode *Silhouette Score* untuk menilai kualitas *clustering*. Hasil pengujian menunjukkan bahwa kombinasi metode *Watershed* dan PCA mampu menghasilkan klasifikasi kerusakan jalan, dengan konfigurasi optimal pada pengelompokan tiga *cluster*. Sistem ini diharapkan dapat menjadi solusi efektif dalam membantu pihak berwenang melakukan penilaian dan perbaikan jalan secara tepat sasaran.

ABSTRACT

Dewi, Intan Tiara. 2025. ***Decision Support System untuk Mengetahui Jalan Berlubang Menggunakan Augmentasi Citra Berbasis Watershed dan PCA Clustering.*** 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) Dr. Fresy Nugroho, M.T

Keywords: Clustering, Decision Support System, Image Augmentation, Potholes, Principal Component Analysis (PCA), Silhouette Score, Watershed.

Potholes are one of the most common types of road damage that can endanger road users and disrupt traffic flow. Manual identification of road damage is time-consuming, labor-intensive, and prone to human error. This study aims to develop a Decision Support System (DSS) based on digital image processing to automatically and efficiently detect and classify the severity of pothole damage. The proposed method involves image augmentation using image space transformation techniques, segmentation using the Watershed algorithm, and clustering using Principal Component Analysis (PCA). PCA is applied to reduce the dimensionality of image data and categorize road damage into three levels: minor, moderate, and severe. Evaluation is conducted using the Silhouette Score to assess the clustering performance. The results show that the combination of Watershed and PCA methods can effectively classify road damage, with optimal clustering achieved using three clusters. This system is expected to serve as an effective solution to assist authorities in making accurate and timely road maintenance decisions.

الملخص

ديوي، إنتان تيارا. 2025. نظام دعم اتخاذ القرار لتحديد الحفر باستخدام تعزيز الصور المستندة إلى مستجمعات المياه وتجميع أطروحة. برنامج دراسة هندسة المعلوماتية، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية. PCA. الحكومية مالانج.المشرف: (أولاً) الدكتور أغونغ تيغوه ويوو أليس، ماجستير في العلوم والتكنولوجيا (ثانياً) الدكتور فريسي نوغروهو، ماجستير في العلوم والتكنولوجيا

الكلمات المفتاحية: تكبير الصور، التجميع، نظام دعم اتخاذ القرار، الحفر، تحليل المكونات الرئيسية، درجة الصورة الظلية، مستجمعات المياه

الحفر هي شكل من أشكال أضرار البنية التحتية التي يمكن أن تعرض سلامة مستخدمي الطرق للخطر وتعيق التدفق السلس لحركة المرور. وتتطلب عملية تحديد أضرار الطرق يدوياً قدرًا كبيراً من الوقت والموارد وهي عرضة للخطأ البشري. يهدف هذا البحث إلى تطوير نظام لدعم اتخاذ القرار يعتمد على معالجة الصور الرقمية لاكتشاف وتصنيف مستوى الأضرار التي لحقت بالحفر بشكل آلي وفعال. تتمثل الطرق المستخدمة في زيادة الصور باستخدام تقنية تحويل فضاء الصورة، والتجزئة باستخدام خوارزمية مستجمعات لتقليل أبعاد بيانات الصورة وتصنيف PCA يُستخدم. (PCA) المياه، وتجميع نتائج التجزئة باستخدام تحليل المكونات الرئيسية مستوى التلف إلى ثلاث فئات: خفيفة ومتوسطة وثقيلة. تم إجراء التقييم باستخدام طريقة "نقاط الصورة الظلية" لتقييم جودة التجميع. أظهرت نتائج الاختبار أن الجمع بين طريقتي مستجمعات المياه والتحليل متعدد الكلور قادر على إنتاج تصنيف الأضرار التي لحقت بالطرق، مع التكوين الأمثل في ثلاث مجموعات عنقودية. من المتوقع أن يكون هذا النظام حلاً فعالاً في مساعدة السلطات على إجراء تقييم وإصلاح الطرق المستهدفة

BAB I

PENDAHULUAN

1.1 Latar Belakang

Jalan raya merupakan infrastruktur transportasi darat yang memiliki peran vital dalam kehidupan manusia, terutama dalam mendukung pertumbuhan ekonomi serta aspek sosial budaya yang mendukung pembangunan nasional. Agar mobilitas masyarakat dapat berjalan lancar dan pelayanan yang diberikan sesuai dengan kapasitas yang dibutuhkan, kondisi jalan harus tetap terjaga. Ketika jalan mengalami kerusakan, aktivitas masyarakat akan terganggu, yang berpotensi menyebabkan kecelakaan (Tenggara dkk., 2021). Oleh karena itu, seiring meningkatnya kebutuhan akan kualitas pelayanan jalan, diperlukan peningkatan mutu jalan dan sarana penunjang lainnya, termasuk menyediakan jalan yang aman dan nyaman.

Dalam hal ini, salah satu sektor yang penting untuk diperhatikan adalah infrastruktur jalan, terutama dalam penanganan kerusakan jalan berlubang. Jalan yang rusak, seperti jalan berlubang, dapat menyebabkan berbagai masalah bagi pengguna jalan, termasuk meningkatkan risiko kecelakaan dan memperlambat arus lalu lintas. Oleh karena itu, penting untuk mengidentifikasi tingkat kerusakan pada jalan guna menentukan prioritas perbaikan yang dibutuhkan. Tingkat kerusakan dapat bervariasi, mulai dari kerusakan ringan yang memerlukan perbaikan kecil hingga kerusakan parah yang membutuhkan perbaikan besar bahkan rekonstruksi ulang.

Di Asia, kecelakaan yang disebabkan oleh kondisi jalan yang buruk, termasuk jalan berlubang, menyumbang sekitar 10% dari total kecelakaan tahunan. Hal ini menunjukkan betapa pentingnya pemantauan dan perbaikan jalan untuk mengurangi risiko kecelakaan (Al Faritzie dkk., 2022). Pengukuran tingkat kerusakan jalan seringkali dilakukan secara manual. Proses pendeteksian dan pencatatan secara manual oleh tenaga kerja sepenuhnya yang mana bisa memakan waktu dua minggu untuk jalan sepanjang 1 km, belum lagi tingkat keakuratan yang rendah, hal ini tentu tidak efektif dalam penentuannya. Terdapat faktor yang menjadi penyebab mengapa penentuan secara manual terbilang tidak efektif, yang pertama memerlukan banyak tenaga dan waktu. Butuhnya banyak orang dikarenakan kerusakan yang terjadi berdampak tidak hanya pada satu arah jalan, hal ini tentu saja membuat waktu penentuan menjadi semakin panjang. Akibatnya bisa menghambat respon dari tim pengecekan. Berikutnya yang kedua yaitu kebergantungan kepada ahli dalam penentuan tingkat kerusakan jalan menjadi salah satu faktor. Dibutuhkan sistem keputusan agar penentuan tingkat kerusakan menjadi konsisten, namun keterbatasan sang ahli menjadi penghambat dalam penentuannya (Fatikasari, 2021).

Untuk mengatasi keterbatasan ini, pengembangan sistem pendukung keputusan diperlukan untuk secara otomatis mendeteksi jalan berlubang menggunakan Augmentasi citra berbasis *Watershed* dan *PCA Clustering* (Sri Rahayu & Halim, 2020a). Penggunaan Augmentasi citra dan *Clustering* dengan metode *Watershed* dan *PCA* yang telah dirancang dalam pendekatan komputasi untuk mengidentifikasi kondisi jalan yang berlubang. Dengan memanfaatkan data

citra jalan, sistem ini memiliki potensi untuk mendeteksi jalan berlubang secara efektif dan efisien. Dengan demikian, diharapkan dampak negatif dari deteksi manual jalan berlubang dapat dihindari, sehingga proses perbaikan jalan dapat dilakukan dengan lebih cepat dan akurat (Sasmito dkk., 2023a).

Penelitian model *Deep Learning* untuk penentuan tingkat kerusakan jalan memiliki potensi yang besar untuk memberikan jalan dalam pengembangan model segmentasi yang lebih kompleks dan selaras dengan konsep islam, bahwa penggunaan model *Deep Learning* berarti sama dengan melakukan pekerjaan yang akurat.

Ayat Al-Qur'an yang dapat dijadikan acuan dalam bekerja secara akurat dan tepat terdapat pada Surat Al-Ma'idah ayat 16 yang berbunyi :

يَهْدِي بِهِ اللَّهُ مَنِ اتَّبَعَ رِضْوَانَهُ ۚ سُبُلَ السَّلَامِ وَيُخْرِجُهُم مِّنَ الظُّلُمَاتِ إِلَى النُّورِ بِإِذْنِهِ ۖ وَيَهْدِيهِمْ إِلَى صِرَاطٍ مُسْتَقِيمٍ

”Dengan Kitab itulah Allah Subhanahu wa ta’ala memberi petunjuk untuk orang yang mengikuti keridaan-Nya menuju jalan keselamatan, Allah Subhanahu wa ta’ala mengeluarkan orang itu dari gelap gulita kepada cahaya dengan izin-Nya, dan menunjukkan ke jalan yang lurus” (QS. Al – Ma’idah: 16).

Penafsiran ayat ini dari Kementrian Agama RI berbunyi ”Ayat ini menjelaskan bahwa Al-Qur'an berfungsi sebagai petunjuk bagi orang-orang yang mencari keridaan Allah Subhanahu wa ta’ala, mengarahkan mereka ke jalan keselamatan di dunia dan akhirat, serta membebaskan mereka dari kegelapan menuju cahaya dan menunjukkan jalan yang benar. Dalam ayat ini terdapat tiga tuntunan penting: mematuhi ajaran Al-Qur'an akan membawa manusia kepada keselamatan dan kebahagiaan, mengikuti ajaran Al-Qur'an akan membebaskan manusia dari segala kesesatan yang diakibatkan oleh tahayul dan khurafat, serta Al-

Qur'an akan mengantarkan manusia pada tujuan akhir agama, yaitu kebahagiaan di dunia dan di akhirat.”

Tafsir QS. Al-Ma'idah (Arah, 2018) ayat 16 diartikan sebagai bimbingan menuju agama Islam yang benar, menjauhkan manusia dari kesesatan, bahaya, dan kekufuran, menuju cahaya kebenaran dan keamanan di dunia dan akhirat. Dalam konteks jalan berlubang, model yang berbasis teknologi ini dapat dianggap sebagai alat yang membantu manusia untuk menemukan dan memperbaiki jalan yang rusak, sehingga menciptakan lingkungan yang lebih aman bagi semua pengguna jalan. Dalam konteks penelitian, ayat ini dapat diintegrasikan dengan menggunakan metode *Watershed* dan PCA yang akurat dan teliti dalam menentukan tingkat kerusakan jalan. Dengan demikian, penelitian dapat menghasilkan keputusan yang akurat dan dapat dipertanggungjawabkan.

Penelitian sebelumnya menunjukkan bahwa metode *Watershed* dapat digunakan untuk segmentasi citra jalan berlubang dengan nilai yang memuaskan. Dalam penelitian ini menunjukkan bahwa algoritma baru yang berbasis *Watershed* dapat mencapai akurasi deteksi hingga 95,34%, jauh lebih baik dibandingkan metode konvensional lainnya yang sering mengalami *over-segmentasi*. Dengan menggabungkan metode *Watershed* dengan Augmentasi, *oversegmentasi* dapat diminimalkan dengan mengelompokkan *region* yang memiliki kesamaan ciri, sehingga menghasilkan segmentasi citra yang lebih efisien. Proses penggabungan ini dikenal sebagai multiresolusi *Watershed*.

Berikutnya, penelitian mengenai pelabelan data kerusakan jalan berlubang menggunakan metode PCA pernah dilakukan dalam permasalahan deteksi dan

klasifikasi kerusakan pasca bencana alam (Almais dkk., 2023). Dalam penelitian ini, dari total 216 data kerusakan yang dianalisis, PCA berhasil mengelompokkan data dengan baik ke dalam tiga kategori: kerusakan ringan, sedang, dan berat, berdasarkan varian sebesar 85,17% untuk PC1 dan 9,36% untuk PC2. Hasil validasi menunjukkan jika penggunaan PCA dalam proses pelabelan data kerusakan memberikan hasil yang sangat akurat, dengan distribusi data yang sesuai dengan target asli. Hal ini menunjukkan jika metode PCA sangat efektif untuk digunakan dalam pengelompokan dan pelabelan data kerusakan jalan berlubang berbasis Augmentasi citra *Watershed*.

Dengan dasar masalah yang telah diuraikan, alasan pemilihan metode *Watershed* dan PCA didasarkan pada penelitian terdahulu. Studi Almais et al. (Almais dkk., 2023) menunjukkan keberhasilan PCA dalam mengelompokkan data kerusakan jalan berlubang berdasarkan data numerik. Hal ini menginspirasi penggunaan PCA dalam analisis berbasis citra untuk optimalisasi pengelompokan data visual. Pemilihan metode *Watershed* dalam segmentasi citra medis, khususnya dalam mengekstraksi *Region of Interest* (ROI) pada citra radiograf. Keberhasilan *Watershed* dalam segmentasi objek dari citra menegaskan potensinya untuk isolasi area penting pada citra dengan akurasi tinggi. Melalui penelitian ini, penulis bertujuan menerapkan kombinasi metode *Watershed* dan PCA dalam segmentasi citra jalan berlubang untuk menentukan tingkat kerusakan secara lebih akurat dan efisien.

1.2 Rumusan Masalah

Dalam penelitian ini, permasalahan yang akan dikaji adalah bagaimana mengevaluasi hasil pengelompokan yang diperoleh melalui *Principal Component Analysis* (PCA) dalam mengklasifikasikan tingkat kerusakan jalan berlubang.

1.3 Batasan Masalah

- a. Penelitian ini akan berfokus pada analisis kerusakan jalan berlubang tanpa memperhitungkan faktor sosial lainnya, dengan menggunakan data berupa citra atau visual.
- b. Data yang digunakan menggunakan data jalan rusak dari *Kaggle* (<https://www.Kaggle.com/datasets/trolololo888/potholes-and-road-damage-with-annotations>).
- c. Menggunakan model Augmentasi data *image space transformation* (*delation, reflection, translation, rotation, cropping, dan Noise injection*).
- d. Menggunakan ekstraksi fitur dari *Unknown region* dan menggunakan ukuran gambar 32x32 piksel.

1.4 Tujuan Penelitian

Melakukan evaluasi terhadap hasil pengelompokan yang diperoleh melalui metode PCA dalam mengklasifikasikan tingkat kerusakan jalan berlubang.

1.5 Manfaat Penelitian

- a. Penelitian ini memberikan manfaat untuk pihak berwenang dengan menyediakan model yang dapat digunakan untuk melakukan perawatan jalan secara lebih efektif berdasarkan data.
- b. Model yang dikembangkan dapat membantu mengidentifikasi kerusakan jalan, seperti jalan berlubang secara akurat dan cepat, sehingga dapat meningkatkan keselamatan pengguna jalan.

BAB II

STUDI PUSTAKA

2.1 Penelitian Terdahulu

Penelitian ini berfokus pada penerapan teknik Augmentasi citra untuk meningkatkan akurasi deteksi jalan berlubang dengan memanfaatkan metode segmentasi berbasis *Watershed*. Penelitian sebelumnya menunjukkan bahwa Augmentasi citra dapat meningkatkan kualitas dataset yang digunakan dalam pelatihan model deteksi (Miftahuddin dkk., 2020a). Dengan menambahkan variasi pada citra, seperti rotasi, pemotongan, dan perubahan pencahayaan, model dapat belajar dari data yang lebih beragam. Hal ini penting untuk meningkatkan generalisasi model dalam mendeteksi jalan berlubang di berbagai kondisi.

Metode *Watershed* merupakan teknik Augmentasi citra yang efektif untuk memisahkan objek dalam citra berdasarkan intensitas piksel (Junod, 2022). Penelitian menunjukkan bahwa metode ini dapat digunakan untuk mendeteksi tepi jalan berlubang dengan akurasi yang tinggi. Dengan menerapkan metode ini, citra yang diambil dari jalan dapat diproses untuk mengidentifikasi area yang mengalami kerusakan.

Table 2. 1 Penelitian Terdahulu

No	Referensi	Topik	Metode	Subjek
1	(Wijaya, 2021)	Deteksi Jalan Berlubang Menggunakan <i>CNN</i>	<i>YOLO (You Only Look Once)</i>	Akurasi keseluruhan sebesar 88% dan <i>kappa accuracy</i> 86%.
2	(Sri Rahayu & Halim, 2020b)	Sistem Temu Kembali Citra Lubang Jalan Aspal Berdasarkan Tingkat Kerusakan	<i>Gray Level Co-occurrence Matrix (GLCM)</i>	Mampu mengidentifikasi lubang jalan berdasarkan tingkat kerusakan.

No	Referensi	Topik	Metode	Subjek
3	(Almais dkk., 2024)	Penilaian Kerusakan Bangunan Pasca Bencana	PCA, GLCM (<i>Gray-Level Co-occurrence Matrix</i>)	Mengidentifikasi kerusakan struktural bangunan pasca bencana menggunakan integrasi analisis GLCM-PCA
4	(Prasasti, Wilis, 2021a)	Segmentasi Citra pada Objek Sebaran Api	<i>Watershed Transform</i>	Segmentasi citra untuk mendeteksi sebaran api menggunakan metode <i>Watershed transform</i> . Dengan kombinasi <i>threshold</i> , <i>hsv</i> , <i>grayscale</i> dan <i>morphology</i> .
5	(Sasmito dkk., 2023b)	Deteksi Kerusakan Jalan Menggunakan Pengolahan Citra <i>Deep Learning</i> di Kota Semarang.	<i>Deep Learning (CNN)</i>	Mempercepat proses identifikasi kerusakan jalan dengan akurasi yang baik.
6	(Rekayasa Sistem Komputer & MIPA Universitas Tanjungpura Jalan Hadari Nawawi Pontianak Telp, 2023)	Segmentasi Citra Menggunakan <i>Watershed</i> dan <i>Intensity Filtering</i>	<i>Watershed + Highpass & Lowpass Filtering</i>	Mengurangi segmentasi berlebihan, hasil segmentasi lebih baik dibandingkan tanpa <i>pre-processing</i>
7	(Enzellina & Suhaedi, 2022a)	Penentuan Faktor Dominan	<i>Principal Component Analysis (PCA)</i>	Menentukan nilai faktor dominan menggunakan metode <i>Principal Component Analysis</i>
8	(Prayogi dkk., 2023)	Deteksi Tumor Otak pada Citra MRI	<i>Watershed dan Thresholding</i>	Mendeteksi Tumor Otak pada Citra MRI Menggunakan Metode <i>Watershed</i> dan <i>Thresholding</i>
9	(Indriyani dkk., 2021)	<i>Improved Watershed Algorithm for Image Segmentation</i>	<i>Improved Watershed + Gaussian Filter</i>	Akurasi segmentasi meningkat hingga 83,34%, namun masih ada over-segmentation pada citra tertentu
10	(Di dkk., 2021)	Reduksi Dimensi untuk Optimasi <i>Clustering</i>	PCA, <i>K-Means</i>	Menerapkan PCA untuk reduksi data produksi pertanian

No	Referensi	Topik	Metode	Subjek
				di Kabupaten Bojogenoro
11	Penelitian saat ini	Pengelompokan Tingkat Kerusakan Jalan	Augmentasi <i>Watershed</i> dan PCA	Mengelompokkan tingkat kerusakan jalan menggunakan Augmentasi <i>Watershed</i> dan PCA

2.2 Kerusakan Jalan

Jalan raya merupakan jalur yang berfungsi untuk memfasilitasi pergerakan lalu lintas dari satu lokasi ke lokasi lainnya. Sebagai sarana transportasi, lalu lintas di jalan raya harus berjalan lancar dan memenuhi persyaratan teknis serta ekonomis yang sesuai dengan fungsi, *volume*, dan karakteristik lalu lintas.

Secara umum, kerusakan jalan dapat dikategorikan menjadi dua jenis, yaitu kerusakan struktural dan fungsional. Kerusakan struktural terjadi ketika ada kegagalan pada lapisan perkerasan atau salah satu komponen perkerasan, yang menyebabkan jalan tidak lagi mampu menahan beban lalu lintas (T Hidayat, 2020). Sementara itu, kerusakan fungsional adalah jenis kerusakan yang mengganggu keselamatan dan kenyamanan pengguna jalan. Berikut adalah beberapa jenis kerusakan jalan (Uspessy & Tenriajeng, 2022):

a. Kerusakan Ringan

Kerusakan ini tidak terlalu mengganggu fungsi jalan dan biasanya hanya membutuhkan perbaikan minimal.

- Retak Halus (*Hairline Cracking*): Keretakan kecil yang belum menyebar atau terlalu dalam. Retak ini biasanya tidak menyebabkan gangguan serius bagi lalu lintas.

- Pengelupasan Permukaan (*Raveling*): Permukaan jalan mengalami pengikisan ringan, di mana agregat pada lapisan aspal mulai terlepas, namun belum mengakibatkan lubang.
- Aus Permukaan: Kehilangan lapisan atas jalan, tetapi masih dalam batas yang tidak terlalu mengganggu keselamatan berkendara.

b. Kerusakan Sedang

Kerusakan yang mulai mengganggu kenyamanan dan keamanan, serta membutuhkan perbaikan lebih intensif.

- Retak Kulit Buaya (*Aligator Cracking*): Keretakan berbentuk jaring, tetapi belum menyebar secara luar. Kondisi ini mulai mempengaruhi daya dukung jalan.
- Ambblas Ringan (*Minor Depression*): Penurunan permukaan jalan yang tidak terlalu dalam (kurang dari 2 cm), yang dapat menyebabkan genangan air saat hujan.
- Gelombang Jalan (*Corrugation*): Permukaan jalan bergelombang akibat pengereman atau masalah stabilitas, yang mulai mengganggu nyaman berkendara.

c. Kerusakan Berat

- Lubang Besar (*Large Potholes*): Lubang dengan kedalaman signifikan yang mampu menampung air, biasanya akibat pelapukan lapisan aspal dan agregat yang terlepas.
- Ambblas Parah (*Severe Depression*): Penurunan jalan yang lebih dari 2 cm, dengan risiko merusak kendaraan dan menyebabkan genangan air.
- Retak Kulit Buaya Parah (*Severe Alligator Cracking*): Retakan yang luas dan dalam yang menyebabkan rusaknya struktur lapisan jalan, sehingga jalan tidak mampu lagi menahan beban lalu lintas dengan baik.

2.3 Metode

2.3.1 Decision Support System

DSS merupakan suatu sistem berbasis komputer yang digunakan untuk membantu proses pengambilan keputusan dalam situasi semi-terstruktur atau tidak terstruktur. DSS dirancang untuk mendukung penilaian, evaluasi, dan pemilihan alternatif yang dilakukan oleh pengambil keputusan dengan menyediakan informasi, model, dan alat analisis yang relevan.

Implementasi DSS dalam penelitian ini menggunakan teknologi pengolahan citra dan teknik machine learning seperti PCA dan clustering untuk mendukung keputusan berdasarkan data visual yang diperoleh dari citra permukaan jalan. DSS akan memberikan rekomendasi status kerusakan jalan berdasarkan hasil klasifikasi tingkat kerusakan (ringan, sedang, berat), sehingga mempermudah proses evaluasi dan penanganan secara sistematis.

2.3.2 Augmentasi Data *Image space Transformation*

Augmentasi data dalam konteks *transformasi* ruang citra (*image space transformation*) dapat dilakukan untuk meningkatkan kualitas model pembelajaran mesin (*Machine Learning*) dan meningkatkan kemampuan deteksi jalan berlubang. Transformasi geometris seperti rotasi, penggeseran, skala, dan refleksi dapat digunakan untuk memperluas *dataset* dengan mengubah orientasi, posisi, ukuran, dan simetri citra. Misalnya, "transformasi geometris seperti rotasi, penggeseran, skala, dan refleksi dapat digunakan untuk memperluas *dataset* dengan mengubah orientasi, posisi, ukuran, dan simetri citra". Selain itu, transformasi *non-linear* seperti transformasi *spline* tipis (*thin plate splines*) dapat digunakan untuk menggambarkan perubahan *non-linear* pada citra. Misalnya, "transformasi *non-linear* seperti transformasi *spline* tipis dapat digunakan untuk menggambarkan perubahan *non-linear* pada citra" (Sayed, 2023).

Penggunaan *masking* dan *filtering* juga efektif untuk menghilangkan *Noise* dan memperjelas detail permukaan jalan. Misalnya, "manipulasi ruang warna lainnya adalah dengan memisahkan Matriks warna RGB secara individual. Transformasi lain melibatkan batasan nilai piksel pada rentang nilai minimum atau maksimum tertentu".

Dengan melakukan Augmentasi data transformasi ruang citra, kita dapat meningkatkan kualitas model pembelajaran mesin dan meningkatkan kemampuan deteksi jalan berlubang. Oleh karena itu, Augmentasi data transformasi ruang citra merupakan langkah penting dalam pengembangan model deteksi jalan berlubang yang akurat.

2.3.3 Watershed

Operasi *Watershed* adalah metode yang digunakan untuk segmentasi citra dengan membagi citra dalam wilayah-wilayah berbeda berdasarkan skala keabuan atau warna. Teknik ini memetakan citra sebagai sebuah relief topografi dan menggambarkan perilaku air di lanskap. Seperti air hujan yang mengalir mengikuti permukaan tanah dan berkumpul di lembah, setiap lembah memiliki wilayah tempat air mengalir. Ketika air dari dua lembah bertemu, sebuah dam dibangun untuk mencegah pencampuran (Prasasti, Wilis, 2021b). Aliran air berhenti ketika puncak dam terlihat, dan batas ini menjadi garis yang memisahkan segmen citra.

2.3.4 PCA Clustering

PCA merupakan metode statistik yang banyak digunakan untuk mereduksi dimensi data, menghilangkan noise, serta memilih fitur penting dalam analisis data dengan dimensi tinggi. Pada penelitian ini, PCA digunakan untuk mengelompokkan data hasil Augmentasi citra dalam mendeteksi jalan berlubang. Metode *Clustering* yang berbasis PCA memungkinkan pengelompokan data tanpa pengawasan (*unsupervised*), sehingga data dapat dikelompokkan berdasarkan pola yang diidentifikasi dari (Almais dkk., 2023).

BAB III

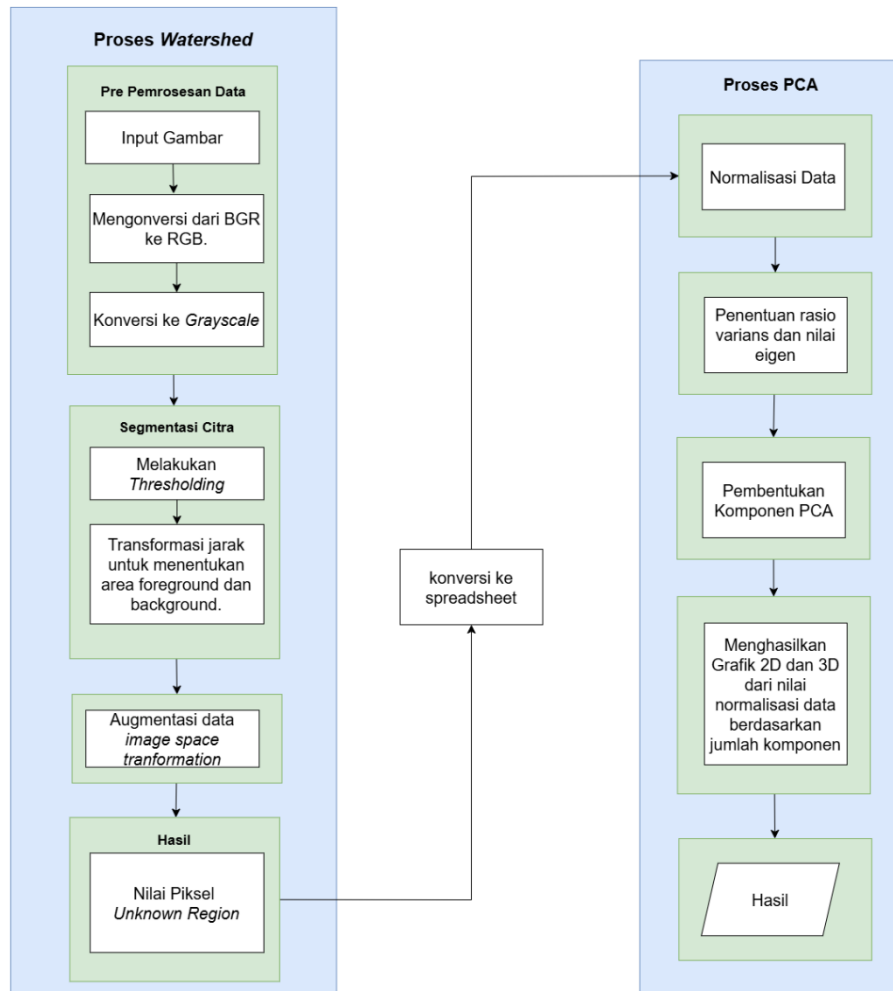
METODE PENELITIAN

3.1 Desain Sistem

Penelitian ini menggunakan metode *Watershed* untuk segmentasi citra yang melibatkan beberapa tahapan, mulai dari *pre-processing* data seperti transformasi ruang citra dan konversi warna, hingga penerapan teknik *thresholding* untuk memisahkan area *foreground* dan *background*. Proses ini menghasilkan nilai *pixel* dari transformasi jarak yang kemudian dikonversi ke dalam format *spreadsheet* sebagai data *input* untuk tahapan selanjutnya.

Metode PCA diterapkan untuk melakukan normalisasi data dan menentukan yang memiliki varian tertinggi. Tujuan dari proses ini adalah mereduksi dimensi data tanpa kehilangan informasi yang signifikan, dengan cara mengidentifikasi rasio varians dan nilai eigen dari data yang telah melalui tahap normalisasi. PCA juga menghasilkan visualisasi data dalam bentuk grafik 2D dan 3D, berdasarkan jumlah komponen yang dipilih.

Pengujian kinerja dilakukan dengan mengevaluasi parameter seperti sensitivitas, spesifisitas, dan akurasi untuk menentukan efektivitas metode dalam mengekstraksi fitur dan mereduksi dimensi citra secara optimal. Interaksi antara proses *Watershed* dan PCA diharapkan dapat meningkatkan kualitas analisis citra dalam penelitian ini, menghasilkan keluaran yang mendukung pengambilan keputusan berdasarkan pola dan fitur yang terkandung dalam citra.



Gambar 3. 1 Desain Sistem

Berikut merupakan langkah-langkah perancangan sistem yang disusun berdasarkan diagram alur:

3.1.1 Akuisisi dan Pra-Pemrosesan Data

Akuisisi dan pra-pemrosesan data dimulai dengan menginput gambar sebagai data awal. Setelah itu, dilakukan Augmentasi data menggunakan *image space transformation* untuk meningkatkan variasi dan kualitas gambar. Selanjutnya, gambar dikonversi dari format warna BGR ke RGB agar sesuai dengan standar

pemrosesan yang digunakan. Setelah konversi warna, gambar diubah menjadi *grayscale* untuk menyederhanakan informasi yang akan digunakan dalam proses segmentasi.

3.1.2 Segmentasi Citra dengan Metode Watershed

Proses segmentasi citra menggunakan metode *Watershed* diawali dengan penerapan *thresholding* untuk memisahkan objek utama dari latar belakang. Setelah itu, dilakukan transformasi jarak guna menentukan area *foreground* dan *background* sehingga batas objek dapat dikenali dengan lebih akurat.

3.1.3 Ekstraksi Fitur dan Penyimpanan Data

Setelah segmentasi selesai, fitur-fitur penting dari citra diekstrak untuk keperluan analisis lebih lanjut. Nilai piksel dari area yang tidak diketahui atau *Unknown region* disimpan sebagai hasil segmentasi. Data yang diperoleh kemudian dikonversi ke dalam format *spreadsheet* agar dapat digunakan dalam tahap analisis berikutnya.

3.1.4 Analisis Data dengan Principal Component Analysis (PCA)

Pada tahap analisis data dengan metode PCA (*Principal Component Analysis*), langkah pertama yang dilakukan adalah normalisasi data agar memiliki distribusi yang lebih seimbang. Setelah itu, rasio varian dan nilai *eigen* dihitung untuk menentukan jumlah yang optimal. Selanjutnya, dibentuk melalui proses reduksi dimensi menggunakan PCA, sehingga hanya fitur yang paling berpengaruh yang dipertahankan. Hasil dari PCA divisualisasikan dalam bentuk grafik 2D dan 3D berdasarkan jumlah komponen yang dipilih.

3.1.5 Hasil Akhir dan Interpretasi

Tahap akhir dari sistem ini adalah interpretasi hasil yang telah diperoleh. Data yang telah diolah dan divisualisasikan digunakan untuk analisis lebih lanjut, baik dalam bentuk laporan maupun sebagai dasar pengambilan keputusan berdasarkan pola yang ditemukan dalam data.

3.2 *Data Collection*

Dalam penelitian ini data yang digunakan didapatkan melalui tahapan proses Augmentasi gambar yang memanfaatkan metode *Watershed* serta *Clustering* berbasis PCA. Sumber data ini berasal dari *Kaggle* (<https://www.Kaggle.com/datasets/trolololo888/potholes-and-road-damage-with-annotations>) dan kriteria yang ditetapkan dalam penelitian bertujuan untuk mendukung evaluasi kondisi jalan berlubang, dengan fokus utama pada segmentasi gambar yang akurat.

Dalam penelitian ini, sebanyak 100 data gambar digunakan untuk analisis segmentasi dan *Clustering*, di mana seluruh data yang digunakan merupakan data sekunder yang diperoleh dari *platform Kaggle*. Tidak ada data yang dikumpulkan langsung dari lapangan, sehingga penelitian ini bergantung pada dokumentasi *visual* yang telah tersedia. Rentang ukuran piksel dari data yang digunakan bervariasi, mulai dari 256×256 piksel hingga 1920×1080 piksel, yang mencerminkan berbagai kondisi pencahayaan dan sudut pengambilan gambar.

Data sekunder ini memungkinkan penelitian untuk mengeksplorasi berbagai jenis kerusakan jalan, termasuk lubang kecil, retakan, serta deformasi aspal yang

lebih luas. Sumber data dari *Kaggle* memberikan akses ke anotasi yang telah disediakan, sehingga proses segmentasi dapat dilakukan dengan lebih terstruktur. Selain itu, penggunaan data ini juga membantu dalam meningkatkan generalisasi model yang dikembangkan, karena mencakup beragam kondisi lingkungan dan karakteristik permukaan jalan dari berbagai lokasi.

Setelah pengumpulan data selesai, salah satu langkah krusial dalam proses pengumpulan data adalah pemeriksaan menyeluruh terhadap kualitas gambar yang dihasilkan. Hal ini dilakukan untuk memastikan bahwa gambar memiliki resolusi dan tingkat ketajaman yang cukup, yang sangat penting bagi analisis selanjutnya. Jika terdapat gambar yang buram, kabur, atau kurang jelas, gambar tersebut akan disaring atau dilakukan perbaikan, bila memungkinkan. Kualitas gambar yang tidak memadai dapat memberikan dampak signifikan terhadap keakuratan hasil segmentasi kerusakan pada jalan. Oleh karena itu, penting untuk memastikan bahwa hanya gambar berkualitas tinggi yang digunakan dalam analisis, mengingat bahwa kualitas visual yang buruk dapat mempengaruhi hasil dari tahap segmentasi, yang merupakan bagian esensial dalam proses evaluasi kondisi jalan berlubang.

3.3 Implementasi Metode *Watershed* dan *Principal Component Analysis*

Tahapan-tahapan dalam Augmentasi citra dengan teknik serta proses *Clustering* berbasis PCA yang terdiri dari beberapa langkah dijelaskan sebagai berikut:

a. Konversi ke *Grayscale*

Proses konversi citra menjadi *grayscale* dilakukan untuk mempermudah analisis dengan mengubah citra berwarna (RGB) menjadi citra dengan satu tingkat kecerahan. Gambar sampel berukuran asli 1080x810 piksel diubah ukurannya menjadi 200x200 piksel. Ukuran yang lebih kecil dipilih untuk mempermudah pemrosesan lebih lanjut pada tahap segmentasi dan analisis citra. Citra *Grayscale* memiliki kedalaman piksel 8 bit, yang menghasilkan 256 tingkat keabuan (Syahfitri dkk., 2023). *Grayscale* dapat diperoleh dari rumus berikut :

$$I_{gray} = 0.299R + 0.587G + 0.114B \quad (3.1)$$

Keterangan:

I_{gray} : nilai piksel *black and white*

R : nilai piksel *red*

G : nilai piksel *green*

B : nilai piksel *blue*

Bobot dalam persamaan ini ditentukan karena mata manusia memiliki sensitivitas lebih tinggi terhadap warna hijau, sehingga kanal hijau diberi bobot yang lebih besar. Akibatnya, citra yang dihasilkan memiliki spektrum kecerahan yang berkisar dari hitam dengan intensitas 0 hingga putih dengan intensitas 255.

b. *Thresholding*

Thresholding diterapkan untuk mengonversi citra *grayscale* menjadi citra biner yang hanya memiliki dua nilai piksel, yaitu 0 (hitam) dan 1 (putih). Gambar 3.3 memberikan contoh citra setelah proses *thresholding*. Tahap awal dari proses ini melibatkan penentuan nilai maksimum dan minimum dari intensitas piksel citra yang akan diproses (Pradani Ayu Widya Purnama & Dr. Ir. Sumijan, 2021).

Menentukan nilai ambang T sebagai kriteria untuk mengubah sebuah piksel menjadi hitam atau putih. Nilai ambang T dihitung dengan rumus :

$$T = \frac{f_{maks} + f_{min}}{2} \quad (3.2)$$

Keterangan:

f_{max} : nilai maksimum dari intensitas piksel

f_{min} : nilai minimum dari intensitas piksel

c. Menghilangkan Noise

Noise atau gangguan dalam citra merupakan ketidaksempurnaan yang muncul dalam bentuk bintik-bintik, sehingga mengurangi kualitas gambar (Yasir et al., 2023). Gangguan ini dapat disebabkan oleh berbagai faktor, seperti kesalahan dalam proses pemindaian, pengambilan gambar yang kurang optimal, atau kotoran yang melekat pada objek yang difoto. Keberadaan *Noise* dapat membuat citra tampak buram atau kurang jelas.

Untuk mengatasi masalah ini, diterapkan berbagai metode pereduksian *Noise*, salah satunya adalah *morfologi opening*. Menurut (Yasir dkk., 2023), *morfologi opening* merupakan teknik dalam pemrosesan citra digital yang bertujuan mengurangi gangguan *Noise* pada citra biner (hitam-putih). Teknik ini terdiri dari dua tahap utama, yaitu dilasi dan erosi.

d. *Sure background* dan *Sure Foreground*

Setelah proses *thresholding* dilakukan pada citra *grayscale*, dihasilkan citra biner dengan dua nilai piksel, yaitu 0 *background* dan 1 untuk *foreground*. Kemudian area *foreground* didefinisikan sebagai piksel dengan nilai 1. Setelah itu menghitung jarak untuk piksel dengan rumus dasar jarak *Euclidean* (Miftahuddin

dkk., 2020b) yang digunakan untuk menghitung *distance transform* antara dua titik

$P_1 = (x_1, y_1)$ dan $P_2 = (x_2, y_2)$ adalah :

$$d = (P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (3.3)$$

Keterangan:

d : jarak
 x_1 dan x_2 : koordinat piksel pada sumbu x
 y_1 dan y_2 : koordinat piksel pada sumbu y

e. Mencari *Unknown region*

Unknown region mengacu pada piksel yang tidak termasuk dalam kategori *foreground* maupun *background*. Nilai wilayah ini dapat ditentukan dengan rumus berikut:

$$Unknown\ Region = Sure\ background - Sure\ foreground \quad (3.4)$$

Dengan ketentuan sebagai berikut:

- Jika *Sure background* bernilai 255 dan *Sure Foreground* bernilai 0, maka *Unknown region* = 255.
- Jika *Sure background* bernilai 255 dan *Sure Foreground* bernilai 255, maka *Unknown region* = 0.

f. Normalisasi Data

Normalisasi data ini agar setiap variabel memiliki skala yang seragam, sehingga kontribusi masing-masing variabel dalam perhitungan menjadi seimbang (Yulianto dkk., 2021). Proses normalisasi data umumnya melibatkan dua langkah utama. Pertama, mengurangi rata-rata dari setiap variabel, dan kedua, membaginya dengan standar deviasi dari variabel tersebut. Proses ini dapat dirumuskan sebagai berikut:

$$X_{norm} = \frac{x - \mu}{\sigma} \quad (3.5)$$

Keterangan:

X_{norm} : data yang sudah dinormalisasi
 x : nilai asli
 μ : rata rata
 σ : standar deviasi

Hasil normalisasi ini memastikan bahwa semua variabel memiliki rata-rata nol dan *varian* satu. Dengan demikian, variabel yang memiliki skala besar tidak akan mendominasi variabel dengan skala kecil dalam perhitungan.

g. Menentukan Matriks Kovarian

Tahap ini bertujuan untuk memahami hubungan antar variabel dalam *dataset* yang telah dinormalisasi. Matriks kovarian adalah sebuah Matriks yang menunjukkan sejauh mana dua variabel berubah bersama-sama (kovariani). Rumus Matriks kovarian adalah sebagai berikut (Enzellina & Suhaedi, 2022b) :

$$C = \frac{1}{(n - 1)} \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})^T \quad (3.6)$$

Keterangan:

C : Matriks kovarian
 n : jumlah data
 x_i : data ke- i
 $\bar{x}$: rata-rata

Nilai kovarian memberikan informasi mengenai arah hubungan antar variabel. Jika kovarian bernilai positif, kedua variabel cenderung meningkat bersama. Sebaliknya, jika bernilai negatif, peningkatan satu variabel akan diikuti dengan penurunan variabel lainnya. Kovarian bernilai nol menunjukkan bahwa tidak ada hubungan *linier* yang jelas antara kedua variabel.

h. Menghitung *Eigenvalues* dan *Eigenvectors*

Setelah memperoleh Matriks kovarian, langkah selanjutnya adalah menentukan nilai *eigen* (*Eigenvalues*) dan vektor *eigen* (*Eigenvectors*). Kedua elemen ini sangat penting karena menentukan arah dan besarnya variasi dalam data, yang menjadi dasar pembentukan *Principal Components*. Nilai *eigen* dihitung menggunakan persamaan berikut:

$$\det(C - \lambda I) = 0 \quad (3.7)$$

Keterangan:

C : Matriks kovarian
 λ : *Eigenvalues*
 I : Matriks identitas

Vektor *eigen* adalah vektor *non-nol* yang arah dan skalanya tidak berubah saat transformasi *linear* diterapkan. Dalam PCA, vektor *eigen* menunjukkan arah dari , yaitu arah dengan variabilitas maksimum dalam data. Untuk setiap nilai *eigen* () yang diperoleh dari persamaan karakteristik, vektor *eigen* () dapat ditemukan dengan menyelesaikan:

$$Cv = \lambda v \quad (3.8)$$

Keterangan:

C : Matriks kovarian
 λ : nilai *Eigen*
 I : Matriks identitas

i. Menentukan Jumlah

Menentukan jumlah merupakan langkah krusial dalam PCA, karena berpengaruh terhadap tingkat akurasi *dataset*. dipilih berdasarkan nilai *eigen* dari Matriks kovarian. Sedangkan Vektor *eigen* dengan nilai tertinggi menangkap

variasi terbesar dalam data dan membentuk sistem koordinat baru yang memungkinkan representasi data dalam dimensi lebih rendah. dihitung menggunakan persamaan berikut:

$$PCA\ Components = X_{norm} \times V \quad (3.9)$$

Keterangan:

X_{norm} : data yang telah di normalisasi

v : vector *eigen*

j. Visualisasi dan *Clustering*

Hasil PCA kemudian divisualisasikan dalam grafik 2D dan 3D untuk memperlihatkan distribusi data pada *Principal Component*. Berdasarkan distribusi ini, dilakukan pengelompokan (*clustering*) menjadi tiga *cluster* dengan rentang nilai koordinat PC1 sebagai berikut (Almais dkk., 2020) :

- *Cluster 1*: Data dengan nilai koordinat kurang dari 0, menggambarkan kondisi kerusakan ringan.
- *Cluster 2*: Data dengan nilai koordinat antara 0 hingga kurang dari 2, menunjukkan kondisi kerusakan sedang.
- *Cluster 3*: Data dengan nilai koordinat lebih dari atau sama dengan 2, menunjukkan kondisi kerusakan berat.

3.4 Evaluasi Model

Metode *Silhouette Score* digunakan untuk mengevaluasi hasil *clustering* dari PCA. Metrik ini mengukur seberapa baik objek dalam *cluster* dikelompokkan dibandingkan dengan objek di *cluster* lain (Paembonan & Abduh, 2021). Nilai *Silhouette Score* dihitung menggunakan rumus berikut (Maharani dkk., 2021) :

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (4.0)$$

Keterangan:

a (i) : rata-rata jarak dari titik data i ke semua titik dalam *cluster* yang sama

b (i) : rata-rata jarak dari titik data i ke semua titik di *cluster* terdekat

Silhouette Score berkisar antara -1 hingga 1. Semakin tinggi nilainya, semakin baik kualitas *clustering*. Skala evaluasi hasil *clustering* adalah sebagai berikut:

- 0,71 - 1,00: *Clustering* sangat kuat
- 0,51 - 0,70: *Clustering* baik
- 0,26 - 0,50: *Clustering* lemah
- $\leq 0,25$: *Clustering* buruk

Dalam penelitian ini, PCA digunakan untuk mengelompokkan data ke dalam tiga *cluster*:

- *Cluster* 1: Kerusakan ringan
- *Cluster* 2: Kerusakan sedang
- *Cluster* 3: Kerusakan berat

Setelah pengelompokan, nilai rata-rata *Silhouette Score* dihitung untuk mengevaluasi kualitas *clustering* secara keseluruhan. Semakin tinggi nilai rata-rata, semakin baik pemisahan antar *cluster*, sedangkan nilai rendah menunjukkan adanya tumpang tindih antara *cluster* yang ada.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Langkah Pengujian

Langkah pengujian dalam penelitian ini bertujuan untuk menilai efektivitas metode *Watershed* dan *Principal Component Analysis* (PCA) dalam melakukan pengelompokkan atau *Clustering* tingkat kerusakan jalan. Data yang digunakan berupa citra jalan yang rusak, diperoleh dari *Kaggle*. Citra tersebut menggambarkan tingkat kerusakan jalan, sehingga dapat dijadikan dasar bagi peneliti dalam menganalisis dan mengelompokkan tingkat kerusakan. Evaluasi hasil *Clustering* dilakukan menggunakan *Silhouette Score* untuk menilai akurasi serta kualitas pemisahan antara *Cluster*.

4.2 Hasil Pengujian

Pada bagian hasil pengujian ini, dibahas proses evaluasi tingkat kerusakan jalan yang dilakukan melalui augmentasi citra menggunakan algoritma *Watershed* serta analisis klasterisasi berbasis *Principal Component Analysis* (PCA). Penelitian ini terdiri dari lima tahapan utama. Tahap pertama adalah pra-pemrosesan citra, yang meliputi pengolahan awal untuk meningkatkan kualitas dan ketajaman citra. Tahap kedua adalah augmentasi citra dengan menerapkan algoritma *Watershed* guna memperjelas batas-batas objek pada citra, khususnya pada area yang mengalami kerusakan. Tahap ketiga adalah konversi hasil segmentasi *Watershed* ke dalam bentuk numerik, sehingga data dapat dianalisis secara kuantitatif. Tahap

keempat adalah klasterisasi menggunakan PCA, yang bertujuan untuk mereduksi dimensi data sekaligus mengelompokkan area jalan berdasarkan tingkat kerusakannya. Terakhir, pada tahap kelima dilakukan evaluasi terhadap hasil klasterisasi menggunakan metode Silhouette Score untuk mengukur sejauh mana hasil pengelompokan memiliki kohesi internal yang baik dan keterpisahan antar klaster yang jelas.

4.2.1 Proses Watershed

Algoritma *Watershed* digunakan untuk Augmentasi citra dengan memisahkan objek berdasarkan perbedaan intensitas piksel. Proses ini dimulai dengan tahap pra-pemrosesan citra untuk mempersiapkan data sebelum augmentasi. Berikut adalah tahapan-tahapan yang dilakukan dalam implementasi algoritma *Watershed*.

a. Pra-pemrosesan Citra

Pra-pemrosesan citra merupakan tahap awal dalam segmentasi menggunakan algoritma *Watershed*. Pada tahap ini, citra dibaca menggunakan *OpenCV* dalam format BGR, kemudian dikonversi ke format RGB agar sesuai untuk proses visualisasi dan analisis lebih lanjut. Selanjutnya, ukuran citra diubah menjadi 32x32 piksel menggunakan metode interpolasi *inter_area* untuk mengurangi kompleksitas komputasi serta memastikan kualitas citra tetap optimal. Setelah itu, citra dikonversi ke *grayscale* guna memperoleh satu kanal intensitas yang lebih mudah diolah dalam proses segmentasi. Hasil dari pra-pemrosesan ini kemudian divisualisasikan dalam tiga bentuk, yaitu citra asli, citra yang telah diubah

ukurannya, dan citra *grayscale*, yang masing-masing memiliki peran penting dalam meningkatkan efisiensi serta akurasi pada tahap segmentasi menggunakan algoritma *Watershed*. Hasil dari tahap pra-pemrosesan ini ditampilkan pada Gambar 4.1 yang menunjukkan bagaimana data disiapkan untuk tahap segmentasi.



Gambar 4. 1 Hasil pra-pemrosesan citra

Gambar 4.1 melihatkan hasil dari tahap pra-pemrosesan citra, yang meliputi konversi citra ke format RGB, perubahan citra menjadi *grayscale*, serta pengubahan ukuran citra menjadi 32×32 piksel. Sebelumnya, ukuran citra asli sebelum diproses adalah 1200×800 piksel.

b. Segmentasi Citra

Proses yang bertujuan untuk memisahkan area penting dalam sebuah gambar dari latar belakang yang tidak diperlukan agar analisis selanjutnya dapat dilakukan dengan lebih efisien. Pada penelitian ini, metode *Watershed* dimanfaatkan untuk melakukan segmentasi melalui serangkaian tahapan yang sistematis. Langkah pertama dalam segmentasi adalah *thresholding* menggunakan metode *Otsu's Binarization*, yang secara otomatis menentukan ambang batas optimal berdasarkan distribusi intensitas piksel pada gambar *grayscale*. Proses ini menghasilkan citra biner di mana objek utama direpresentasikan dalam warna putih, sementara latar

belakang dalam warna hitam, seperti yang ditunjukkan pada Gambar 4.2. Setelah tahap *thresholding*, dilakukan proses penghilangan *noise* atau objek kecil yang tidak diinginkan pada latar belakang menggunakan operasi *morfologi opening*. Teknik ini bertujuan untuk membersihkan area *foreground* dari piksel kecil atau *Noise* yang tidak relevan sehingga menghasilkan citra yang lebih bersih dan siap untuk diproses lebih lanjut.



Gambar 4. 2 Hasil citra *threshold*

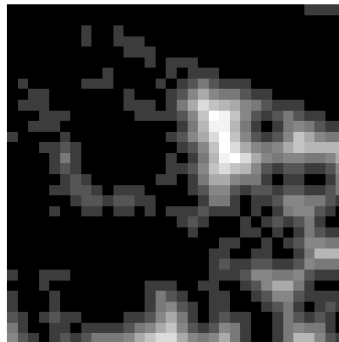
Setelah proses penghilangan *noise* selesai, langkah berikutnya adalah menentukan *Sure background* dan *Sure Foreground*. *Sure background* diperoleh dengan menerapkan operasi dilasi pada hasil *morfologi opening*. Dilasi berfungsi untuk memperluas area latar belakang dengan menambahkan lapisan putih di sekitar objek, sehingga batas antara objek dan latar belakang menjadi lebih jelas. Hasil dari *Sure background* ini dapat dilihat pada Gambar 4.3, yang menampilkan area yang dipastikan sebagai latar belakang.

Sure background, dilated

Gambar 4. 3 Hasil citra *Sure background*

Langkah berikutnya adalah menerapkan *distance transform* untuk mengukur jarak setiap piksel pada area *foreground* terhadap tepi terdekat. Proses ini menghasilkan gradasi intensitas, di mana piksel yang berada lebih jauh dari tepi memiliki nilai intensitas yang lebih tinggi. Hasil dari *distance transform* ditampilkan pada Gambar 4.4, yang menunjukkan distribusi jarak area *foreground* terhadap batas objek.

Distance Transform

Gambar 4. 4 Hasil citra *distance transform*

Setelah memperoleh hasil dari *distance transform*, dilakukan *thresholding* kembali untuk menentukan *Sure Foreground*, yaitu bagian *foreground* yang paling dominan. Langkah ini bertujuan untuk memisahkan objek utama dari latar belakang. Hasil dari *Sure Foreground* ditampilkan pada Gambar 4.5, yang

memperlihatkan area *foreground* yang telah teridentifikasi dengan jelas dan siap untuk tahap segmentasi selanjutnya.



Gambar 4. 5 Hasil citra *Sure Foreground*

Tahap terakhir dalam proses segmentasi adalah menentukan *Unknown region*, yaitu area yang tidak dapat dikategorikan dengan pasti sebagai *foreground* atau *background*. *Unknown region* diperoleh dengan mengurangi *Sure Foreground* dari *Sure background*. Area ini memiliki informasi penting untuk analisis lebih lanjut, terutama dalam proses *Clustering*. Hasil dari *Unknown region* ditampilkan pada Gambar 4.6, yang menunjukkan area yang menjadi perhatian utama pada tahap analisis berikutnya.



Gambar 4. 6 Hasil citra *Unknown region*

c. Konversi Hasil Algoritma *Watershed*

Tahap konversi hasil algoritma *Watershed* dilakukan untuk menyimpan data segmentasi citra ke dalam *file Excel*, sehingga dapat dianalisis dengan lebih mudah.

Proses ini mengekstrak nilai-nilai piksel dari *unknown region* yang telah ditentukan dalam tahap segmentasi dan menyusunnya dalam format tabel. Setiap piksel di *unknown region* direpresentasikan dalam bentuk nilai numerik yang mencerminkan intensitas atau informasi relevan lainnya yang berguna untuk analisis lebih lanjut.

4.2.2 Proses PCA

Langkah akhir adalah penerapan PCA, yang bertujuan untuk mengenali dan mengelompokkan area kerusakan pada bangunan berdasarkan nilai piksel dari *unknown region*.

a. Normalisasi Data

Langkah awal dalam penerapan PCA adalah melakukan normalisasi data agar semua fitur berada dalam skala yang sama. Normalisasi ini menggunakan fungsi *StandardScaler*, yang mengubah setiap fitur memiliki standar deviasi. Hal ini bertujuan untuk mencegah fitur dengan skala besar mendominasi perhitungan PCA. Dalam sistem ini, data yang dinormalisasi diperoleh dari nilai piksel segmentasi menggunakan algoritma *Watershed*, khususnya pada *unknown region*. Hasil normalisasi ini menghasilkan data numerik yang lebih terstruktur, sebagaimana ditunjukkan pada Gambar 4.7.

```
[[ -1.67874412 -0.52915026 -0.43033148 ...  0.82717019  0.8819171
   1.06458129]
 [ -1.57043805 -0.52915026 -0.43033148 ... -1.20894105 -1.13389342
  -0.93933644]
 [ -1.46213197 -0.52915026 -0.43033148 ... -1.20894105 -1.13389342
  -0.93933644]
 ...
 [  1.46213197  1.88982237  2.32379001 ...  0.82717019  0.8819171
 -0.93933644]
 [  1.57043805  1.88982237  2.32379001 ...  0.82717019  0.8819171
 -0.93933644]
 [  1.67874412  1.88982237  2.32379001 ...  0.82717019  0.8819171
   1.06458129]]
```

Gambar 4. 7 Hasil normalisasi data

b. Penentuan Nilai Rasio *Varian* dan Nilai *Eigen*

Hasil normalisasi data pada Gambar 4.7 membentuk sebuah Matriks yang terdiri dari baris dan kolom. Baris dalam Matriks ini merepresentasikan jumlah fitur, sementara kolomnya merupakan parameter yang digunakan. Untuk menentukan *Principal Component* dari Matriks tersebut, digunakan kolom (parameter) yang terdapat pada Gambar 4.7, sehingga jumlah PC yang dihasilkan sama dengan jumlah kolom dalam gambar tersebut.

Langkah berikutnya adalah menghitung rasio *varian* dan nilai *Eigen* dari *Principal Component*. Rasio *varian* menggambarkan proporsi variasi data yang dapat dijelaskan oleh setiap komponen—semakin tinggi nilainya, semakin besar kontribusi komponen tersebut dalam merepresentasikan struktur data. Sementara itu, nilai *Eigen* menunjukkan seberapa kuat suatu komponen dalam mendeskripsikan data, di mana komponen dengan nilai *Eigen* yang lebih besar lebih diprioritaskan. Tabel 4.1 menyajikan presentase nilai rasio varian dan nilai *Eigen* untuk setiap *Principal Component* dalam analisis PCA.

Tabel 4. 1 Nilai Rasio dan Nilai *Eigen*

PC	<i>Explained Variation (%)</i>	<i>Eigen Value (%)</i>
PC1	23.96	8.16
PC2	15.94	5.43
PC3	10.67	3.63
PC4	7.64	2.6
PC5	6.39	2.18
PC6	5.94	2.02
PC7	5.23	1.78
PC8	3.82	1.3
PC9	3.61	1.23
PC10	2.42	0.83
PC11	2.23	0.76
PC12	1.81	0.62
PC13	1.76	0.6
PC14	1.71	0.58
PC15	1.31	0.44
PC16	1.07	0.36
PC17	0.91	0.31

PC18	0.89	0.3
PC19	0.78	0.27
PC20	0.44	0.15

Tabel 4.1 menyajikan persentase nilai rasio *varian* dan nilai *Eigen* untuk setiap *Principal Component* dalam analisis PCA. Dari tabel tersebut, dapat dilihat bahwa PC1 dan PC2 memiliki rasio *varian* tertinggi, masing-masing sebesar 23.96% dan 15.94%, dengan nilai *Eigen* sebesar 8.16% dan 5.43%. Kedua komponen ini berperan paling signifikan dalam analisis karena memberikan kontribusi terbesar dalam menjelaskan variabilitas data.

Penurunan rasio *varian* pada setiap dalam Tabel 4.1 menunjukkan bahwa semakin kecil urutan komponen, semakin sedikit kontribusinya dalam menjelaskan variasi data. Hal ini terjadi karena PCA menyusun berdasarkan seberapa besar variasi data yang dapat dijelaskan, dari yang paling dominan hingga yang paling kecil. Tren penurunan ini berlanjut hingga komponen terakhir, seperti PC19 dan PC20, yang memiliki rasio *varian* sangat rendah, yaitu masing-masing 0.78% dan 0.44%, sehingga kontribusinya terhadap representasi data secara keseluruhan dapat diabaikan.

c. Pembentukan

Dalam sistem ini, pembentukan dilakukan dengan menerapkan *transformasi* PCA pada data yang telah dinormalisasi. Normalisasi bertujuan untuk menyamakan skala semua variabel sehingga tidak ada variabel yang mendominasi analisis. PCA mengolah data dalam bentuk numerik yang merepresentasikan informasi penting,

di mana setiap baris menunjukkan satu objek, sedangkan setiap kolom merepresentasikan variabel yang relevan, seperti intensitas piksel.

PCA mentransformasi data ke dalam bentuk baru yang disebut . Komponen ini bersifat tidak saling berkorelasi dan membantu menyederhanakan pola dalam data. Hasil transformasi ini disusun dalam sebuah tabel (*DataFrame*), dengan setiap kolom merepresentasikan satu . Dalam penelitian ini, PCA menghasilkan 20 , tetapi hanya dua komponen pertama (PC1 dan PC2) yang digunakan dalam tahap *Clustering* karena memiliki kontribusi terbesar terhadap variasi data.

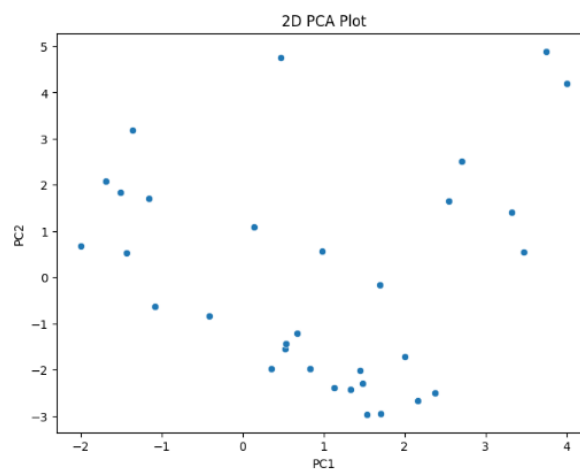
Pemilihan PC1 dan PC2 didasarkan pada nilai *Eigen* yang signifikan, yang menggambarkan seberapa banyak informasi yang dapat dijelaskan oleh setiap komponen secara individual. Penelitian oleh (Almais et al., 2023) mengonfirmasi bahwa dua ini sudah cukup untuk merepresentasikan pola data secara visual dan mendukung proses *Clustering*. Dengan memilih dua , analisis menjadi lebih sederhana tanpa mengurangi efektivitas dalam mengidentifikasi pola penting dalam data. Transformasi PCA ini secara signifikan menyederhanakan data dan mempermudah proses analisis secara keseluruhan.

d. Visualisasi dan *Clustering*

Setelah terbentuk, langkah berikutnya adalah membuat visualisasi grafik 2D dan 3D dari data yang telah dinormalisasi berdasarkan . Tujuan visualisasi ini adalah untuk mengidentifikasi pola *Clustering* dalam data. Pada visualisasi 2D, digunakan *scatter plot* untuk menampilkan distribusi data dalam dua dimensi, di mana setiap titik mewakili satu observasi. Kedekatan antar titik mencerminkan tingkat kemiripan data, sehingga memudahkan dalam mengamati pola *Clustering*.

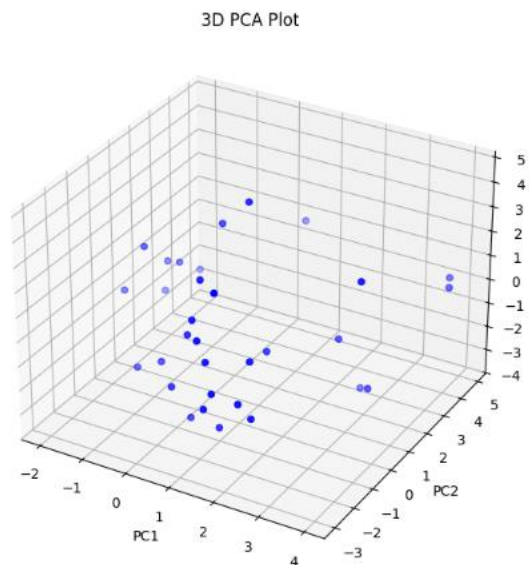
Sementara itu, pada visualisasi 3D, sumbu z ditambahkan dengan menggunakan nilai dari ketiga (PC3), sedangkan sumbu x dan y masing-masing merepresentasikan nilai dari pertama (PC1) dan kedua (PC2). Penambahan dimensi ini memberikan perspektif yang lebih mendalam, memungkinkan identifikasi pola distribusi data dalam ruang tiga dimensi dengan lebih jelas. Dengan mengombinasikan kedua jenis visualisasi ini, pola *Cluster* yang terbentuk dapat diamati secara lebih rinci.

Dalam analisis ini, data divisualisasikan dalam ruang dua dan tiga dimensi berdasarkan . Selanjutnya, proses *Clustering* dilakukan dengan menerapkan 3, 4, dan 5 *Cluster* (Rianti dkk., 2024) untuk mengeksplorasi bagaimana variasi jumlah *Cluster* mempengaruhi hasil pengelompokan data.



Gambar 4. 8 Visualisasi Grafik 2D

Gambar 4.8 menampilkan visualisasi data dalam ruang dua dimensi berdasarkan PC1 dan PC2. Pada grafik tersebut, setiap titik data dipetakan pada sumbu PC1 dan PC2 untuk menunjukkan pola distribusi data, di mana jarak antar titik mencerminkan tingkat kesamaan karakteristik data.



Gambar 4. 9 Visualisasi Grafik 3D

Gambar 4.9 memperlihatkan distribusi data dalam tiga dimensi dengan menggunakan sumbu PC1, PC2, dan PC3. Visualisasi ini memberikan pemahaman yang lebih mendalam mengenai struktur data, terutama dalam membedakan kelompok yang mungkin tidak terlihat pada tampilan dua dimensi. Dengan adanya sumbu z, analisis pola distribusi data menjadi lebih jelas dan menyeluruh.

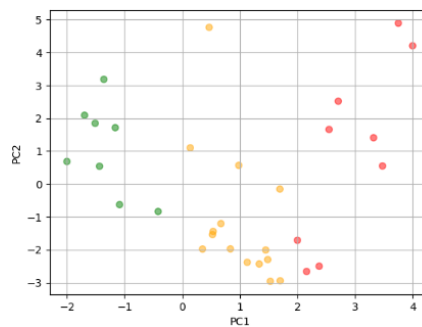
Setelah visualisasi dilakukan, data dikelompokkan (*clustering*) berdasarkan pola distribusinya. Hasil analisis dari Gambar 4.8 dan Gambar 4.9 menunjukkan bahwa data terbagi menjadi 3 *cluster* berdasarkan koordinat titik-titik yang dihasilkan. Metode *Clustering* ini mengacu pada penelitian yang dilakukan oleh (Almais et al., 2023), yang mengelompokkan tingkat kerusakan bangunan ke dalam tiga kategori utama: rusak ringan, rusak sedang, dan rusak berat, sebagaimana tercantum dalam Tabel 4.2. Penentuan label tingkat kerusakan didasarkan pada nilai koordinat PC1 (n), di mana jika $n < 0$, bangunan dikategorikan sebagai rusak ringan,

jika $0 \leq n < 2$, maka masuk dalam kategori rusak sedang, sedangkan jika $n \geq 2$, bangunan diklasifikasikan sebagai rusak berat.

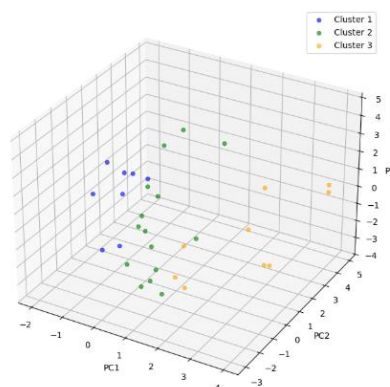
Tabel 4. 2 Nilai Rentang (n) tingkat kerusakan jalan

Rentang Nilai	Tingkat Kerusakan
$n < 0$	Rusak Ringan
$0 \leq n < 2$	Rusak Sedang
$n \geq 2$	Rusak Berat

Tabel 4.2 menjelaskan bahwa setiap *cluster* dipisahkan berdasarkan kedekatannya dalam ruang dua dan tiga dimensi. Penggunaan warna dalam visualisasi membantu membedakan masing-masing *cluster*, sehingga memudahkan dalam mengamati distribusi data pada berbagai kombinasi yang dianalisis.

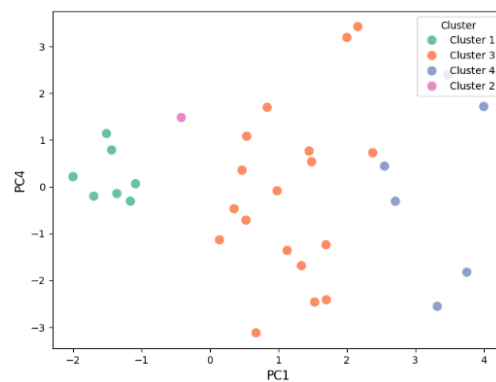


Gambar 4. 10 Hasil *Clustering 3 Cluster* Grafik 2D

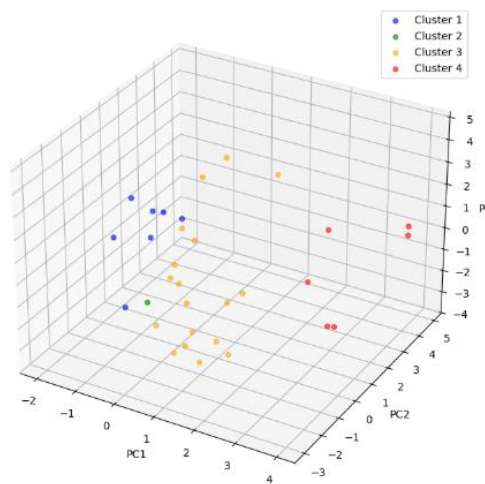


Gambar 4. 11 Hasil *Clustering 3 Cluster* Grafik 3D

Setelah data dikelompokkan ke dalam tiga *cluster*, pengujian dilanjutkan dengan menambahkan jumlah *cluster* menjadi empat dan lima. Penambahan *cluster* ini dilakukan untuk mengeksplorasi apakah pemecahan data ke dalam lebih banyak kelompok dapat memberikan hasil yang lebih detail serta memperjelas perbedaan antar kelompok data.



Gambar 4. 12 Hasil *Clustering 4 Cluster* Grafik 2D



Gambar 4. 13 Hasil *Clustering 4 Cluster* Grafik 3D

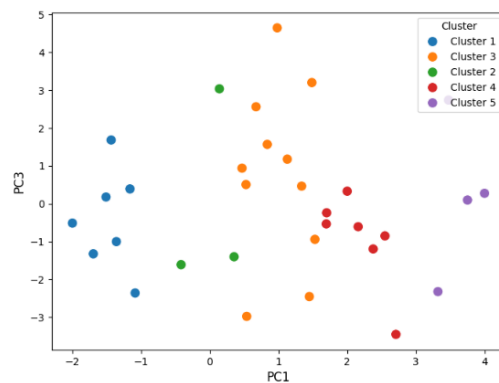
Gambar 4.12 dan 4.13 menampilkan visualisasi empat *cluster* dalam grafik 2D dan 3D. Dengan bertambahnya *cluster*, distribusi data menjadi lebih detail, memisahkan titik yang sebelumnya tergabung dalam kelompok besar. *Cluster* keempat membantu memperjelas batas antar data. Pembagian *cluster* ini didasarkan

pada rentang nilai PC1 yang telah ditentukan yang dijelaskan di Tabel 4.3. Berikut adalah rentang nilai untuk masing-masing *Cluster*:

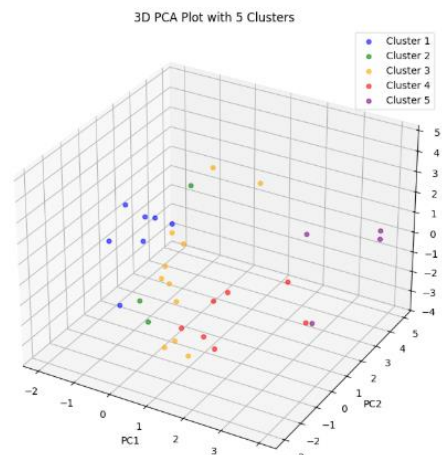
Tabel 4. 3 Rentang Nilai (n) dalam 4 *Cluster*

Cluster	Rentang Nilai	Jenis Kerusakan
<i>Cluster 1</i>	$n < -1$	Rusak Ringan Sekali
<i>Cluster 2</i>	$-1 \leq n < 0$	Rusak Ringan
<i>Cluster 3</i>	$0 \leq n < 2.5$	Rusak Sedang
<i>Cluster 4</i>	$n \geq 2.5$	Rusak Berat

Tabel 4.3 menjelaskan bahwa setiap *cluster* dipisahkan berdasarkan kedekatannya dalam ruang dua dan tiga dimensi. Penggunaan warna dalam visualisasi membantu membedakan masing-masing *cluster*, sehingga memudahkan dalam mengamati distribusi data pada berbagai kombinasi yang dianalisis



Gambar 4. 14 Hasil *Clustering 5 Cluster* Grafik 2D



Gambar 4. 15 Hasil *Clustering 5 Cluster* Grafik 3D

Gambar 4.14 dan Gambar 4.15 menampilkan hasil visualisasi dengan lima *Cluster*. Dengan adanya penambahan satu *Cluster*, data semakin terbagi ke dalam kelompok-kelompok yang lebih kecil. Hal ini memungkinkan pemisahan yang lebih rinci, meskipun beberapa *Cluster* terlihat lebih kecil dan mungkin hanya berisi sedikit titik data.

Peningkatan jumlah *Cluster* menghasilkan lebih banyak variasi dalam pembentukan kelompok, yang dapat membantu mengidentifikasi perbedaan yang lebih halus dalam distribusi data. Namun, hal ini juga berisiko menciptakan *Cluster* yang kurang bermakna atau kurang relevan.

Pembagian titik data ke dalam lima *Cluster* dilakukan berdasarkan nilai PC1, di mana data dikelompokkan sesuai dengan rentang nilai yang lebih spesifik. Rentang nilai untuk masing-masing *Cluster* dapat dilihat pada Tabel 4.4.

Tabel 4. 4 Rentang Nilai (n) dalam 5 *Cluster*

<i>Cluster</i>	Rentang Nilai	Jenis Kerusakan
<i>Cluster 1</i>	$n < -0.8$	Rusak Ringan Sekali
<i>Cluster 2</i>	$-0.8 \leq n < 0.4$	Rusak Ringan
<i>Cluster 3</i>	$0.4 \leq n < 1.6$	Rusak Sedang
<i>Cluster 4</i>	$1.6 \leq n < 2.8$	Rusak Berat
<i>Cluster 5</i>	$n \geq 2.8$	Rusak Berat Sekali

Tabel 4.4 menjelaskan bahwa setiap *cluster* dipisahkan berdasarkan kedekatannya dalam ruang dua dan tiga dimensi. Penggunaan warna dalam visualisasi membantu membedakan masing-masing *cluster*, sehingga memudahkan dalam mengamati distribusi data pada berbagai kombinasi yang dianalisis.

Gambar visualisasi diatas merupakan hasil visualisasi dari proses clustering data menggunakan metode Principal Component Analysis (PCA) dan algoritma pengelompokan (*clustering*), yang ditampilkan dalam ruang dua dimensi antara

pertama (PC1) dan ketiga (PC3). Setiap titik pada grafik merepresentasikan data piksel hasil ekstraksi citra yang telah direduksi dimensinya menggunakan PCA. Titik-titik tersebut dikelompokkan ke dalam lima klaster berbeda yang masing-masing ditandai dengan warna: biru, oranye, hijau, merah, dan ungu. Berdasarkan hasil interpretasi, klaster biru (Cluster 1) diasosiasikan dengan kondisi kerusakan ringan, klaster oranye dan merah (Cluster 2 dan 4) mengindikasikan kerusakan sedang hingga berat, sedangkan klaster ungu (Cluster 5) diduga merupakan kondisi kerusakan berat ekstrem atau data outlier. Klaster hijau (Cluster 3) berada di antara kelompok lain dan dapat mencerminkan transisi antara kerusakan ringan dan sedang. Visualisasi ini memperjelas distribusi tingkat kerusakan dan menunjukkan bahwa pendekatan PCA berhasil mengungkap struktur data yang dapat digunakan untuk klasifikasi kondisi kerusakan secara efektif.

e. Evaluasi Hasil *Clustering*

Dalam penelitian ini, evaluasi hasil *Clustering* dilakukan menggunakan *Silhouette Score* untuk menilai sejauh mana objek dalam satu *Cluster* dikelompokkan dengan baik dibandingkan dengan objek di *Cluster* lainnya. Evaluasi ini diterapkan pada seluruh data, yang terdiri dari 100 gambar yang merepresentasikan kerusakan jalan berlubang. Setiap gambar dikategorikan ke dalam salah satu dari tiga tingkat kerusakan, yaitu ringan, sedang, dan berat.

Silhouette Score dihitung berdasarkan yang digunakan dalam proses *Clustering* dengan *Principal Component Analysis* (PCA). Evaluasi dilakukan pada berbagai skenario *Clustering*, mencakup visualisasi dalam bentuk 2D dan 3D. Untuk visualisasi 2D, kombinasi PC1 dan PC2 digunakan dengan jumlah *Cluster*

yang bervariasi, yakni 3, 4, dan 5 *Cluster*. Sementara itu, visualisasi 3D menggunakan kombinasi PC1, PC2, dan PC3 dengan jumlah *Cluster* yang sama, yaitu 3, 4, dan 5 *Cluster*. Pada Tabel 4.5 merupakan hasil evaluasi *Silhouette Score* untuk visualisasi 2D.

Tabel 4. 5 Nilai *Silhouette Score* Tiap Data dari Grafik 2D

Data Gambar	Nilai <i>Silhouette Score</i>		
	3 <i>Cluster</i>	4 <i>Cluster</i>	5 <i>Cluster</i>
Gambar 1	0.2974	0.1861	0.0736
Gambar 2	0.6204	0.6368	0.6425
Gambar 3	0.4259	0.5512	0.4623
Gambar 4	0.1536	0.1516	0.1211
Gambar 5	0.3575	0.4365	0.2088
Gambar 6	0.2044	0.1478	0.0727
Gambar 7	0.4414	0.3755	0.3302
Gambar 8	0.3628	0.1489	0.1313
Gambar 9	0.3116	0.2726	0.1699
Gambar 10	0.2585	0.1884	0.1780
Gambar 11	0.1668	0.2135	0.1934
Gambar 12	0.3830	0.3032	0.3059
Gambar 13	0.1983	0.1922	0.1015
Gambar 14	0.3704	0.3308	0.2536
Gambar 15	0.3097	0.2835	0.1285
Rata-rata	0.3241	0.2946	0.2249

Tabel 4.5 menampilkan nilai *Silhouette Score* untuk setiap gambar berdasarkan tiga skenario *clustering*, yaitu menggunakan 3, 4, dan 5 *Cluster*. *Silhouette Score* digunakan untuk mengevaluasi kualitas *clustering*, di mana nilai yang lebih tinggi menunjukkan bahwa objek dalam *cluster* lebih terpisah dengan baik. Dari tabel tersebut, dapat diamati bahwa penggunaan 3 *Cluster* menghasilkan nilai *Silhouette Score* yang lebih tinggi dibandingkan dengan 4 dan 5 *Cluster*, yang menunjukkan bahwa pemisahan data lebih optimal dalam konfigurasi ini. Seiring dengan bertambahnya jumlah *Cluster*, nilai *Silhouette Score* cenderung menurun, yang mengindikasikan bahwa penambahan jumlah *cluster* tidak selalu meningkatkan kualitas pengelompokan. Beberapa gambar, seperti Gambar 2

(0.6204), Gambar 3 (0.4259), dan Gambar 14 (0.3704), menunjukkan nilai *Silhouette Score* yang tinggi untuk 3 *Cluster*, menandakan bahwa data dalam gambar-gambar tersebut lebih jelas terkelompok. Sebaliknya, beberapa gambar seperti Gambar 4, Gambar 6, dan Gambar 7 memiliki nilai *Silhouette Score* yang relatif rendah di semua skenario, yang dapat mengindikasikan bahwa objek dalam gambar tersebut lebih sulit untuk dikelompokkan secara optimal. Secara keseluruhan, hasil ini menunjukkan bahwa *clustering* dengan 3 *Cluster* memberikan pemisahan data yang lebih baik dibandingkan dengan 4 atau 5 *Cluster*, meskipun variasi dalam beberapa gambar menunjukkan perlunya analisis lebih lanjut untuk meningkatkan akurasi *clustering*. Tabel 4.6 menyajikan hasil evaluasi *Silhouette Score* pada grafik 3D untuk skenario *clustering* dengan 3, 4, dan 5 *Cluster*:

Tabel 4. 6 Nilai *Silhouette Score* Tiap Data dari Grafik 3D

Data Gambar	Nilai <i>Silhouette Score</i>		
	3 <i>Cluster</i>	4 <i>Cluster</i>	5 <i>Cluster</i>
Gambar 1	0.1944	0.0728	0.0671
Gambar 2	0.6095	0.6180	0.6368
Gambar 3	0.4102	0.4881	0.4209
Gambar 4	0.1619	0.1601	0.1299
Gambar 5	0.3022	0.3618	0.1741
Gambar 6	0.1549	0.0824	0.0881
Gambar 7	0.4509	0.3690	0.4001
Gambar 8	0.3300	0.2021	0.0717
Gambar 9	0.1978	0.1673	0.0363
Gambar 10	0.2039	0.1414	0.1483
Gambar 11	0.1273	0.2741	0.2007
Gambar 12	0.3317	0.2526	0.2946
Gambar 13	0.1506	0.1332	0.0556
Gambar 14	0.3783	0.3387	0.2760
Gambar 15	0.3403	0.3177	0.0198
Rata-rata	0.2896	0.2653	0.2013

Berdasarkan hasil perhitungan *Silhouette Score* pada tabel 4.4, kualitas *clusterisasi* cenderung menurun seiring bertambahnya jumlah *cluster*. Rata-rata

Silhouette Score tertinggi terdapat pada 3 *cluster* dengan nilai 0.2896, diikuti oleh 4 *cluster* (0.2653) dan 5 *cluster* (0.2013). Hal ini menunjukkan bahwa pemisahan data paling optimal terjadi pada 3 *cluster*, sedangkan menambah jumlah *cluster* justru mengurangi kejelasan pemisahan.

Beberapa gambar, seperti Gambar 2 dan Gambar 3, menunjukkan peningkatan nilai *Silhouette Score* saat jumlah *cluster* bertambah, menandakan bahwa data dalam gambar tersebut lebih sesuai jika di *cluster* dengan jumlah lebih banyak. Namun, sebagian besar gambar mengalami penurunan skor, seperti Gambar 5, Gambar 8, dan Gambar 9, yang memiliki penurunan drastis pada 5 *cluster*, menunjukkan bahwa *clusterisasi* menjadi kurang efektif.

Secara keseluruhan, jumlah *cluster* optimal untuk *dataset* ini adalah 3 *cluster*, karena memberikan pemisahan yang lebih baik dibandingkan dengan jumlah *cluster* yang lebih besar. Jika tujuan klasterisasi adalah mendapatkan *cluster* yang lebih terpisah secara jelas, maka 3 *cluster* merupakan pilihan yang lebih baik. Namun, analisis lebih lanjut pada gambar tertentu tetap diperlukan untuk menentukan jumlah *cluster* optimal sesuai dengan karakteristik datanya.

Pada analisis klasterisasi menggunakan *Silhouette Score*, hasil evaluasi menunjukkan bahwa jumlah *cluster* yang optimal adalah 3 *cluster*, baik untuk grafik 2D maupun 3D. Berdasarkan Tabel 4.5, penggunaan 3 *cluster* menghasilkan rata-rata *Silhouette Score* tertinggi sebesar 0.3241 dibandingkan dengan 4 *cluster* (0.2946) dan 5 *cluster* (0.2249). Nilai ini menunjukkan bahwa pemisahan data lebih optimal dengan 3 *cluster* dibandingkan dengan jumlah *cluster* yang lebih besar. Beberapa gambar seperti Gambar 2 (0.6204), Gambar 3 (0.4259), dan Gambar 14

(0.3704) memiliki nilai *Silhouette Score* yang tinggi pada konfigurasi 3 *cluster*, yang menunjukkan bahwa data dalam gambar-gambar tersebut memiliki struktur *cluster* yang lebih jelas. Sebaliknya, beberapa gambar seperti Gambar 4, Gambar 6, dan Gambar 7 memiliki nilai yang lebih rendah, yang mengindikasikan bahwa objek dalam gambar tersebut sulit untuk dikelompokkan secara optimal.

Analisis pada grafik 3D (Tabel 4.6) menunjukkan tren yang serupa. Rata-rata *Silhouette Score* tertinggi diperoleh pada 3 *cluster* (0.2896), diikuti oleh 4 *cluster* (0.2653) dan 5 *cluster* (0.2013). Sebagian besar gambar mengalami penurunan nilai seiring bertambahnya jumlah *cluster*, yang menandakan bahwa peningkatan jumlah *cluster* tidak selalu menghasilkan kualitas klasterisasi yang lebih baik. Namun, beberapa gambar seperti Gambar 2 dan Gambar 3 menunjukkan peningkatan nilai *Silhouette Score* ketika jumlah *cluster* bertambah, yang menandakan bahwa data dalam gambar tersebut lebih sesuai jika dikelompokkan dalam jumlah *cluster* yang lebih banyak. Sebaliknya, beberapa gambar seperti Gambar 5, Gambar 8, dan Gambar 9 mengalami penurunan drastis pada 5 *cluster*, yang menunjukkan bahwa pembagian data menjadi lebih banyak *cluster* tidak selalu efektif.

Secara keseluruhan, hasil analisis menunjukkan bahwa konfigurasi 3 *cluster* memberikan pemisahan data yang lebih jelas dibandingkan dengan 4 atau 5 *cluster*. Pemilihan jumlah *cluster* yang optimal sangat bergantung pada karakteristik data, dan meskipun 3 *cluster* merupakan pilihan terbaik secara umum, beberapa gambar tetap memerlukan analisis lebih lanjut untuk menentukan jumlah *cluster* yang paling sesuai. Oleh karena itu, pendekatan yang fleksibel dan berbasis data sangat dianjurkan dalam menentukan strategi klasterisasi yang optimal.

4.3 Augmentasi *Clustering*

Augmentasi gambar adalah teknik dalam pemrosesan citra yang digunakan untuk memperkaya *dataset* dengan memodifikasi gambar asli agar model pembelajaran mesin lebih tahan terhadap variasi bentuk, posisi, dan pencahayaan. Dalam penelitian ini, digunakan beberapa teknik Augmentasi utama, yaitu rotasi (*rotated*), refleksi (*reflected*), dilasi (*dilated*), dan translasi (*translated*). Teknik ini bertujuan untuk meningkatkan ketahanan model terhadap perbedaan orientasi dan posisi objek dalam gambar tanpa harus menambah jumlah data secara manual.

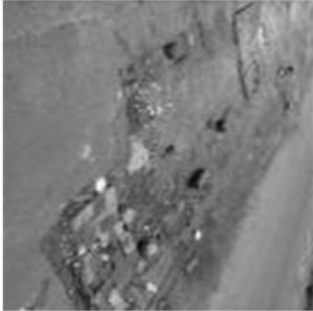
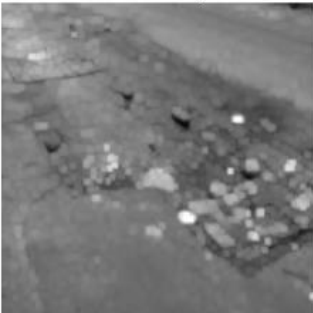
Rotasi dilakukan dengan memutar gambar 90 derajat searah jarum jam untuk memberikan variasi sudut pandang objek. Refleksi dilakukan dengan mencerminkan gambar secara *horizontal*, memungkinkan model mengenali pola yang bersifat simetris. Sementara itu, dilasi digunakan untuk memperbesar area piksel terang dalam gambar menggunakan kernel 3x3, membantu dalam menonjolkan fitur-fitur utama. Terakhir, translasi dilakukan dengan menggeser gambar 5 piksel ke kanan dan ke bawah, sehingga model tidak hanya mengenali objek yang selalu berada di posisi tertentu. Tabel 4.7 menunjukkan hasil dari masing-masing teknik Augmentasi ini.

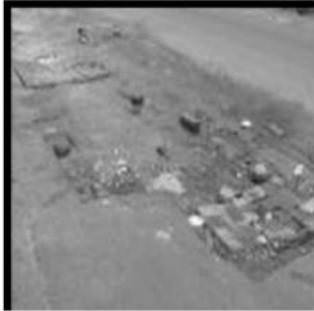
Untuk menguji efektivitas Augmentasi gambar, dilakukan evaluasi menggunakan *Silhouette Score* dengan metode klasterisasi 3 *Cluster*. Hasil uji menunjukkan bahwa Augmentasi gambar membantu meningkatkan kualitas pengelompokan data dengan nilai *Silhouette Score* yang lebih tinggi dibandingkan tanpa augmentasi. Dengan demikian, Augmentasi tidak hanya menambah variasi data tetapi juga meningkatkan pemisahan antar *cluster*, yang berdampak positif

pada kinerja model dalam klasifikasi dan segmentasi gambar. Teknik Augmentasi ini menjadi langkah penting dalam memastikan model pembelajaran mesin dapat berfungsi secara optimal saat diterapkan pada data dunia nyata.

Berikut hasil Tabel 4.7 menyajikan hasil Augmentasi dengan menggunakan evaluasi *Silhouette Score* pada grafik 2D dan 3D untuk skenario *clustering* dengan 3 Cluster:

Tabel 4. 7 Hasil Augmentasi *Watershed*

Gambar	<i>Silhouette Score</i>	
	2D	3D
Rotated Image 	0.1322	0.0824
Reflected Image 	0.1142	0.0470
Dilated Image 	0.2602	0.2184

Translated Image 	0.8478	0.8026
---	--------	--------

Tabel 4.7 yang diberikan membandingkan skor *Silhouette* untuk 2D dan 3D berdasarkan empat jenis perubahan gambar, yaitu *Rotated*, *Reflected*, *Dilated*, dan *Translated Image*. Skor *Silhouette* digunakan untuk menilai seberapa baik *cluster* dalam data tetap terpisah setelah mengalami perubahan.

Pada *Rotated Image*, skor *Silhouette* tercatat sebesar 0.1322 untuk 2D dan 0.0824 untuk 3D, yang menunjukkan bahwa rotasi menyebabkan perubahan distribusi titik data, terutama dalam representasi 3D, sehingga menurunkan kualitas klusterisasi. Sementara itu, *Reflected Image* memiliki skor 0.1142 untuk 2D dan 0.0470 untuk 3D, yang merupakan skor terendah dalam tabel. Hal ini menunjukkan bahwa pencerminan berdampak lebih besar pada klusterisasi, terutama dalam ruang 3D, karena perubahan orientasi yang signifikan.

Pada *Dilated Image*, skor *Silhouette* mencapai 0.2602 dalam 2D dan 0.2184 dalam 3D. Meskipun terjadi perubahan ukuran, hasil ini menunjukkan bahwa klusterisasi masih dapat dipertahankan lebih baik dibandingkan dengan rotasi dan pencerminan. Sementara itu, *Translated Image* memiliki skor tertinggi, yaitu 0.8478 untuk 2D dan 0.8026 untuk 3D, yang menunjukkan bahwa pergeseran posisi tidak banyak memengaruhi pemisahan *cluster*, karena bentuk dan distribusi data tetap relatif tidak berubah.

Dari hasil ini, dapat disimpulkan bahwa *Translated Image* memiliki dampak paling kecil terhadap kualitas klasterisasi, sehingga merupakan metode yang paling stabil. *Dilated Image* masih cukup baik dalam mempertahankan *cluster*, meskipun ada sedikit perubahan pada pemisahan *cluster*. Sebaliknya, *Rotated Image* dan terutama *Reflected Image* memiliki dampak paling besar terhadap kualitas klasterisasi, karena menyebabkan perubahan geometri yang signifikan dan menurunkan skor *Silhouette*, terutama dalam 3D.

4.4 Pembahasan

Penelitian ini bertujuan untuk menganalisis tingkat kerusakan jalan berlubang dengan menggunakan Augmentasi citra melalui algoritma *Watershed* serta teknik *Clustering* berbasis PCA. Proses analisis ini terdiri dari beberapa tahapan utama, yakni pra-pemrosesan citra, Augmentasi dengan algoritma *Watershed*, konversi hasil Augmentasi menjadi data numerik, analisis *Clustering* menggunakan PCA, dan evaluasi model dengan *Silhouette Score*.

Pada tahap awal, citra yang diperoleh melalui proses akuisisi mengalami pra-pemrosesan, yang mencakup konversi ke format RGB dan *grayscale* serta penyesuaian ukuran menjadi 32x32 piksel. Langkah ini bertujuan untuk menyederhanakan pengolahan citra serta mengurangi kompleksitas data tanpa menghilangkan informasi penting. Setelah itu, algoritma *Watershed* diterapkan dalam proses Augmentasi untuk memisahkan objek utama dari latar belakang berdasarkan perbedaan intensitas piksel. Hasil Augmentasi ini menghasilkan area yang belum teridentifikasi, yang kemudian digunakan untuk analisis *Clustering*.

Setelah Augmentasi dilakukan, fitur numerik dari area yang belum teridentifikasi diekstraksi dan dikonversi menjadi data yang lebih mudah dianalisis secara statistik. Data yang diekstraksi kemudian diproses menggunakan PCA guna mereduksi dimensi dan mengidentifikasi yang paling berkontribusi terhadap variasi data. Sebelum analisis PCA dilakukan, data dinormalisasi agar setiap fitur memiliki skala yang seragam, sehingga tidak ada fitur tertentu yang mendominasi analisis.

PCA menghasilkan beberapa , yang selanjutnya digunakan dalam analisis *Clustering*. Dalam penelitian ini, *Clustering* dilakukan dengan dua kombinasi utama: visualisasi 2D menggunakan PC1 dan PC2, serta visualisasi 3D menggunakan PC1, PC2, dan PC3. Kombinasi ini dipilih karena masing-masing komponen memiliki kontribusi signifikan dalam menjelaskan variasi data baik dalam dua maupun tiga dimensi. Data yang telah diklasifikasikan kemudian dianalisis untuk mengidentifikasi pola jalan berlubang.

Proses *Clustering* dilakukan dengan menguji beberapa jumlah *cluster*, yakni 3, 4, dan 5 *cluster*. Hasil visualisasi menunjukkan bahwa pembagian menjadi tiga *cluster* utama—jalan dalam kondisi baik, rusak ringan, dan rusak parah memberikan pemisahan yang paling optimal. Penggunaan kombinasi PC1 dan PC2 dalam visualisasi 2D memberikan pemisahan yang cukup jelas pada 3 *cluster*, sedangkan penambahan jumlah *cluster* menjadi 4 atau 5 menyebabkan penurunan kualitas pemisahan. Begitu pula pada visualisasi 3D menggunakan PC1, PC2, dan PC3, di mana hasil *Clustering* dengan 3 *cluster* tetap memberikan hasil terbaik,

sementara penambahan *cluster* justru menurunkan kejelasan pemisahan pada beberapa gambar.

Pendekatan *Clustering* dalam penelitian ini bertujuan untuk mengelompokkan data berdasarkan kesamaan karakteristik, sejalan dengan bagaimana manusia diciptakan dalam berbagai suku dan bangsa agar dapat saling mengenal dan bekerja sama. Dalam Al-Qur'an, Allah Subhanahu wa ta'ala berfirman:

الَّذِي خَلَقَ سَبْعَ سَمَاوَاتٍ طِبَاقًا مَا تَرَى فِي خَلْقِ الرَّحْمَنِ مِنْ تَفَوتٍ فَارْجِعِ الْبَصَرَ هَلْ تَرَى مِنْ فُتُورٍ

”Yang telah menciptakan tujuh langit berlapis-lapis. Kamu tidak akan melihat pada ciptaan Tuhan Yang Maha Pemurah sesuatu yang tidak seimbang. Maka lihatlah sekali lagi, adakah kamu melihat suatu cacat?” (QS. Al-Mulk: 3).

Menurut tafsir Kementerian Agama RI, ayat ini menegaskan bahwa Allah Subhanahu wa ta'ala menciptakan alam semesta dengan keseimbangan dan tanpa cacat. Jika manusia mengamati ciptaan-Nya dengan saksama, mereka tidak akan menemukan kesalahan atau ketidaksempurnaan dalam aturan yang telah Allah Subhanahu wa ta'ala tetapkan.

Dalam konteks penelitian ini, analisis jalan berlubang menggunakan Augmentasi citra dan PCA *Clustering* bertujuan untuk mengidentifikasi serta memahami pola kerusakan pada jalan. Hal ini selaras dengan konsep dalam ayat tersebut, di mana manusia diingatkan untuk mengamati dan menganalisis lingkungan sekitarnya guna memperbaiki ketidaksempurnaan yang ditemukan, seperti mendeteksi serta menangani kerusakan jalan untuk meningkatkan keselamatan dan kesejahteraan masyarakat.

BAB IV

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini mengembangkan sistem pendukung keputusan untuk mendeteksi jalan berlubang menggunakan Augmentasi citra berbasis algoritma *Watershed* dan *Clustering* berbasis PCA. Hasil penelitian menunjukkan bahwa metode Augmentasi citra dengan algoritma *Watershed* efektif dalam menyoroti area jalan berlubang dengan memisahkan objek utama dari latar belakang berdasarkan perbedaan intensitas piksel. Selain itu, analisis *Clustering* dengan PCA membantu dalam mengurangi dimensi data dan mengidentifikasi pola kerusakan jalan berdasarkan fitur numerik yang diekstraksi dari citra hasil augmentasi. Evaluasi menggunakan *Silhouette Score* menunjukkan bahwa pemilihan tiga *cluster* memberikan pemisahan yang paling optimal dibandingkan dengan empat atau lima *cluster*, baik dalam visualisasi 2D maupun 3D. Dengan hasil yang diperoleh menggunakan 3 *cluster*, metode ini dapat membantu dalam pengelompokan kondisi jalan menjadi beberapa kategori kerusakan, sehingga dapat digunakan sebagai dasar pengambilan keputusan dalam perbaikan dan pemeliharaan jalan.

5.2 Saran

Untuk meningkatkan efektivitas sistem yang dikembangkan, beberapa saran dapat dipertimbangkan dalam penelitian selanjutnya. Pertama, penggunaan citra dengan resolusi lebih tinggi serta penerapan teknik pra-pemrosesan yang lebih

canggih dapat meningkatkan akurasi dalam mendeteksi jalan berlubang. Kedua, kombinasi metode klasifikasi lain, seperti *Deep Learning* atau *Machine Learning*, dapat diterapkan untuk meningkatkan ketepatan dalam mendeteksi dan mengklasifikasikan tingkat kerusakan jalan. Ketiga, perlu dilakukan pengujian pada *dataset* yang lebih luas dan beragam agar sistem dapat bekerja lebih optimal di berbagai kondisi jalan dan lingkungan. Keempat, integrasi dengan sistem pemetaan berbasis GPS dapat memberikan informasi yang lebih komprehensif kepada pihak terkait dalam pemeliharaan jalan. Terakhir, sistem ini dapat dikembangkan lebih lanjut untuk diimplementasikan dalam aplikasi nyata, misalnya melalui penggunaan *drone* atau kamera mobil yang mampu mendeteksi kerusakan jalan secara otomatis dan *real-time*.

DAFTAR PUSTAKA

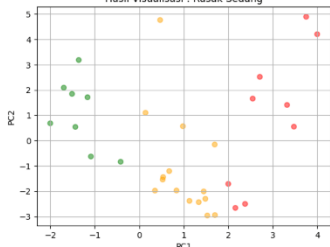
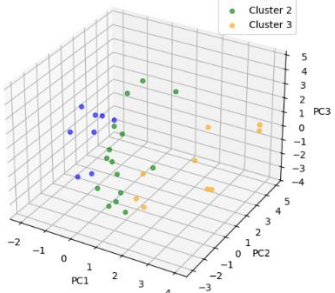
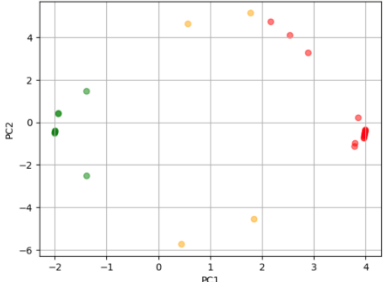
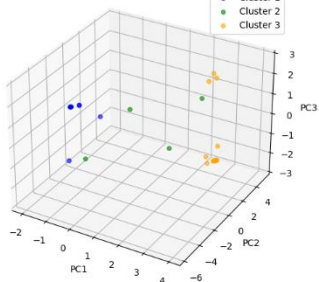
- Al Faritzie, H., Firda, A., & Aprilyanti, S. (2022). Identifikasi Dan Analisis Kerusakan Jalan Pada Ruas Jalan Siaran Sako Kota Palembang. *Bearing : Jurnal Penelitian dan Kajian Teknik Sipil*, 7(4), 223. <https://doi.org/10.32502/jbearing.v7i4.5499>
- Almais, A. T. W., Fatchurrohman, & Holle, K. F. H. (2020). Implementasi fuzzy weighted product penyusunan aksi rehabilitasi rekonstruksi pasca bencana berbasis decision support system dynamic. *Jurnal Eltek*, 18(1), 1. <https://doi.org/10.33795/eltek.v18i1.171>
- Almais, A. T. W., Susilo, A., Naba, A., Sarosa, M., Crysdian, C., Tazi, I., Hariyadi, M. A., Muslim, M. A., Basid, P. M. N. S. A., Arif, Y. M., Purwanto, M. S., Parwatiningsy, D., Supriyono, & Wicaksono, H. (2023). Principal Component Analysis-Based Data Clustering for Labeling of Level Damage Sector in Post-Natural Disasters. *IEEE Access*, 11(July), 74590–74601. <https://doi.org/10.1109/ACCESS.2023.3275852>
- Almais, A. T. W., Susilo, A., Naba, A., Sarosa, M., Juwono, A. M., Crysdian, C., Muslim, M. A., & Wicaksono, H. (2024). Characterization of Structural Building Damage in Post-Disaster using GLCM-PCA Analysis Integration. *IEEE Access*, PP, 1. <https://doi.org/10.1109/ACCESS.2024.3469637>
- Di, P., Bojonegoro, K., Hedyati, D., & Suartana, I. M. (2021). *Penerapan Principal Component Analysis (PCA) Untuk Reduksi Dimensi Pada Proses Clustering Data Produksi*. 05, 49–54.
- Enzellina, G., & Suhaedi, D. (2022a). Penggunaan Metode Principal Component Analysis dalam Menentukan Faktor Dominan. *Jurnal Riset Matematika*, 101–110. <https://doi.org/10.29313/jrm.v2i2.1192>
- Enzellina, G., & Suhaedi, D. (2022b). Penggunaan Metode Principal Component Analysis dalam Menentukan Faktor Dominan. *Jurnal Riset Matematika*, 101–110. <https://doi.org/10.29313/jrm.v2i2.1192>
- Fatikasari, A. D. (2021). Analisis Tingkat Kerusakan Jalan menggunakan metode PCI. *Fatikasari, A. D., & Teknik, F. (2021). Analisis Tingkat Kerusakan Jalan menggunakan metode PCI*. 6(2), 6(2).
- Indriyani, T., Utoyo, M. I., & Rulaningsy, R. (2021). A New Watershed Algorithm for Pothole Image Segmentation. *Studies in Informatics and Control*, 30(3), 131–139. <https://doi.org/10.24846/v30i3y202112>
- Junod, G. (2022). *Makalah IF4073 Interpretasi dan Pengolahan Citra, Semester I Tahun*. 13519096. <https://www.nature.com/articles/srep32412>

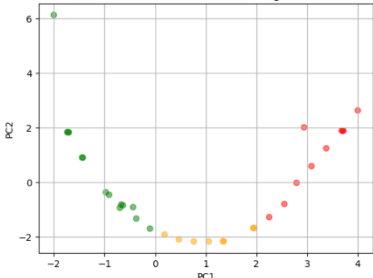
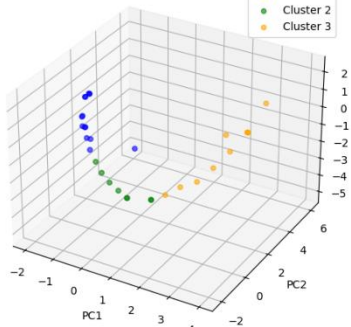
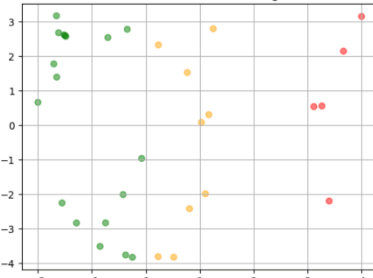
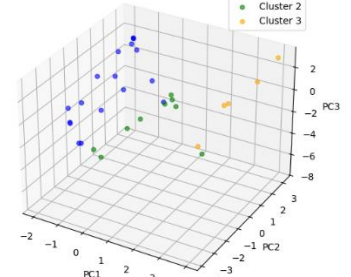
- Maharani, M., Ratnawati, D. E., & Rahayudi, B. (2021). Pengelompokan Kondisi Komputer pada Bank X di Seluruh Indonesia Menggunakan Metode K-Medoids Clustering. ... *Teknologi Informasi dan ...*, 5(5), 1958–1964.
- Miftahuddin, Y., Umaroh, S., & Karim, F. R. (2020a). Perbandingan Metode Perhitungan Jarak Euclidean, Haversine, Dan Manhattan Dalam Penentuan Posisi Karyawan. *Jurnal Tekno Insentif*, 14(2), 69–77. <https://doi.org/10.36787/jti.v14i2.270>
- Miftahuddin, Y., Umaroh, S., & Karim, F. R. (2020b). Perbandingan Metode Perhitungan Jarak Euclidean, Haversine, Dan Manhattan Dalam Penentuan Posisi Karyawan. *Jurnal Tekno Insentif*, 14(2), 69–77. <https://doi.org/10.36787/jti.v14i2.270>
- Paembonan, S., & Abduh, H. (2021). Penerapan Metode Silhouette Coefficient untuk Evaluasi Clustering Obat. *PENA TEKNIK: Jurnal Ilmiah Ilmu-Ilmu Teknik*, 6(2), 48. https://doi.org/10.51557/pt_jiit.v6i2.659
- Pradani Ayu Widya Purnama, M. Kom., & Dr. Ir. Sumijan, M. Sc. (2021). Teori dan Aplikasi Pengolahan Citra Digital Penerapan dalam Bidang Citra Medis. Dalam *Insan Cendekia Mandiri*.
- Prasasti, Wilis, Z. (2021a). Segementasi Citra Menggunakan Metode Watershed Transform dengan Kombinasi Thershold, HSV, Grayscale dan Morphology Untuk Mendeteksi Sebaran API. *Ejournal.Uin-Suska.Ac.Id*, 19(1), 49–54. <http://ejournal.uin-suska.ac.id/index.php/sitekin/article/view/14880>
- Prasasti, Wilis, Z. (2021b). Segementasi Citra Menggunakan Metode Watershed Transform dengan Kombinasi Thershold, HSV, Grayscale dan Morphology Untuk Mendeteksi Sebaran API. *Ejournal.Uin-Suska.Ac.Id*, 19(1), 49–54.
- Prayogi, A., Cendekia Siregar, A., Wahid, R., & Insani, S. (2023). *Deteksi Tumor Otak Menggunakan Metode Watershed dan Thresholding Pada Citra MRI*.
- Rekayasa Sistem Komputer, J., & MIPA Universitas Tanjungpura Jalan Hadari Nawawi Pontianak Telp, F. H. (2023). *Coding : Jurnal Komputer dan Aplikasi Agung Wira Mulia*, [2] Ikhwan Ruslianto, [3] Dwi Marisa Midyanti [1] [2] [3]. 11(01), 11–20.
- Rianti, R., Andarsyah, R., & Awangga, R. M. (2024). Penerapan PCA dan Algoritma Clustering untuk Analisis Mutu Perguruan Tinggi di LLDIKTI Wilayah IV. *Nuansa Informatika*, 18(2), 67–77. <https://doi.org/10.25134/ilkom.v18i2.211>
- Sasmito, B., Setiadji, B. H., & Isnanto, R. (2023a). Deteksi Kerusakan Jalan Menggunakan Pengolahan Citra Deep Learning di Kota Semarang. *Teknik*, 44(1), 7–14. <https://doi.org/10.14710/teknik.v44i1.51908>
- Sasmito, B., Setiadji, B. H., & Isnanto, R. (2023b). Deteksi Kerusakan Jalan Menggunakan Pengolahan Citra Deep Learning di Kota Semarang. *Teknik*, 44(1), 7–14. <https://doi.org/10.14710/teknik.v44i1.51908>

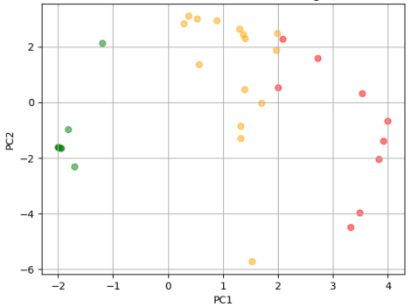
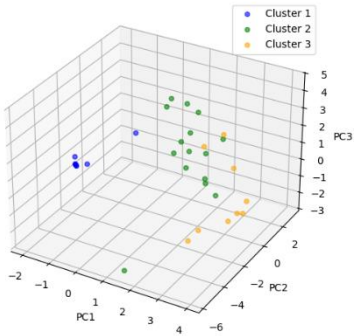
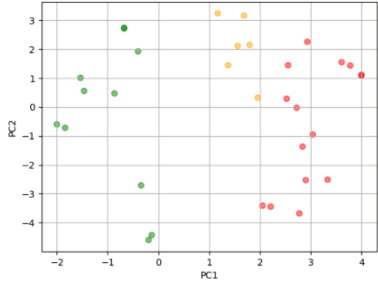
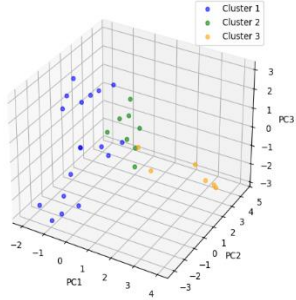
- Sayed, A. H. (2023). Generative Networks. *Inference and Learning from Data*, 2905–2966. <https://doi.org/10.1017/9781009218276.020>
- Sri Rahayu, P., & Halim, J. (2020a). Decision Support System Untuk Kelayakan Pembangunan Jalan Desa Di Kecamatan NamFatikasari, A. D., & Teknik, F. (2021). Analisis Tingkat Kerusakan Jalan menggunakan metode PCI. 6(2).orambe dengan Metode Wighted Aggregated Sum Product Assessment. *Jurnal CyberTech*, x. No.x(x).
- Sri Rahayu, P., & Halim, J. (2020b). Decision Support System Untuk Kelayakan Pembangunan Jalan Desa Di Kecamatan NamFatikasari, A. D., & Teknik, F. (2021). Analisis Tingkat Kerusakan Jalan menggunakan metode PCI. 6(2).orambe dengan Metode Wighted Aggregated Sum Product Assessment. *Jurnal CyberTech*, x. No.x(x). <https://ojs.trigunadharma.ac.id/>
- Syahfitri, A., Ibnutama, K., & Suherdi, D. (2023). Mendeteksi Tingkat Kesegaran Daging Sapi Menggunakan Metode Transformasi Ruang Warna HIS (Hue , Intensity , dan Saturation). 2(November), 923–934.
- Tenggara, M. A. P., Agustin, I. W., & Hariyani, S. (2021). Kinerja Jalan Di Kota Surabaya Berdasarkan Tingkat Pelayanan Jalan. *Planning for Urban Regionand Environment*, 10(3), 119–128.
- Uspessy, M. R., & Tenriajeng, A. T. (2022). Evaluasi Kerusakan Permukaan Jalan Menggunakan Metode Bina Marga dan Prioritas Penanganan Berdasarkan Nilai BCR pada Kelas Jalan Provinsi di Kota Depok. *Rekayasa Sipil*, 16(1), 1–8. <https://doi.org/10.21776/ub.rekayasasipil.2022.016.01.1>
- Wijaya, A. T., Putra, O. V, & Umami, J. (2021). Deteksi Jalan Berlubang Pada Citra Berkabut Menggunakan Convolutional Neural Network Dan Dark Channel Prior. *Prosiding Seminar Sains Dan Teknologi*, 1(1), 131–136. https://publikasiilmiah.unwahas.ac.id/index.php/PROSIDING_SNST_FT/article/view/5035%0Ahttps://publikasiilmiah.unwahas.ac.id/index.php/PROSIDING_SNST_FT/article/viewFile/5035/3777
- Yasir, A., Satria, W., & Yuanda, P. (2023). Digital Image Processing Metode Median Filtering Dan Morfologi Opening Dalam Reduksi Noise Citra. *Warta Dharmawangsa*, 17(4), 1687–1701. <https://doi.org/10.46576/wdw.v17i4.3821>
- Yulianto, S., Apriyadi, R. K., Aprilyanto, A., Winugroho, T., Ponangsera, I. S., & Wilopo, W. (2021). Histori Bencana dan Penanggulangannya di Indonesia Ditinjau Dari Perspektif Keamanan Nasional. *PENDIPA Journal of Science Education*, 5(2), 180–187. <https://doi.org/10.33369/pendipa.5.2.180-187>

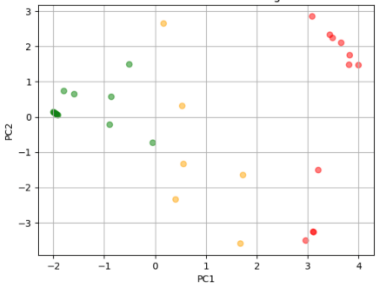
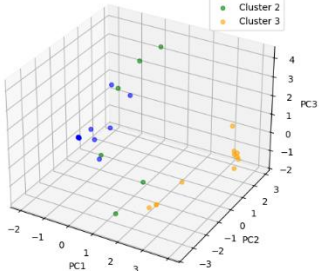
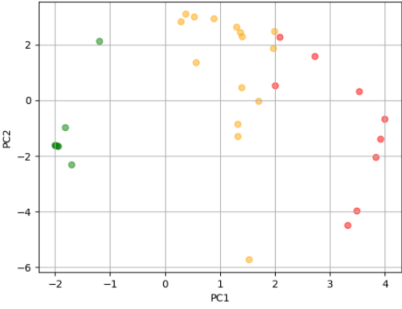
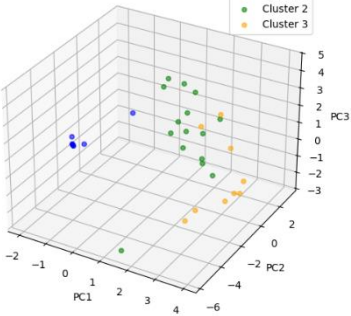
LAMPIRAN

Lampiran 1 Hasil *Clustering* dan *Labeling* Seluruh Data

No.	Gambar	Hasil <i>Watershed</i> - PCA (Kategori Rusak)	Visualisasi
1.		Sedang	Hasil Visualisasi : Rusak Sedang  3D PCA Plot with 3 Clusters 
2.		Berat	Hasil Visualisasi : Rusak Berat  3D PCA Plot with 3 Clusters 

3.		Ringan	Hasil Visualisasi : Rusak Ringan  3D PCA Plot with 3 Clusters 
4.		Ringan	Hasil Visualisasi : Rusak Ringan  3D PCA Plot with 3 Clusters 

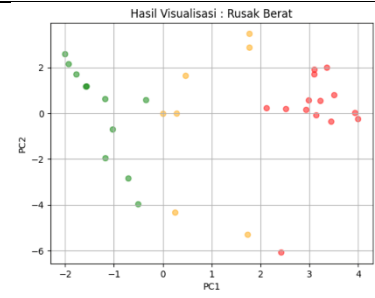
5.		Sedang	Hasil Visualisasi : Rusak Sedang  3D PCA Plot with 3 Clusters 
6.		Berat	Hasil Visualisasi : Rusak Berat  3D PCA Plot with 3 Clusters 

7.		Ringan	Hasil Visualisasi : Rusak Ringan  3D PCA Plot with 3 Clusters 
8.		Sedang	Hasil Visualisasi : Rusak Sedang  3D PCA Plot with 3 Clusters 

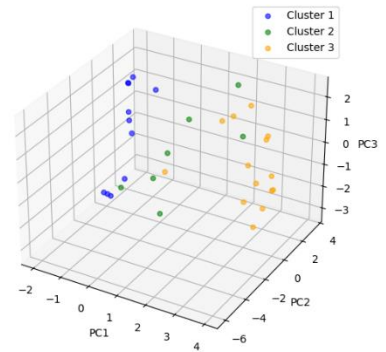
9.



Berat



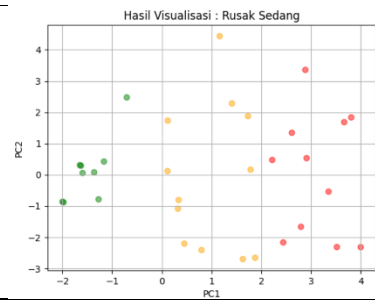
3D PCA Plot with 3 Clusters



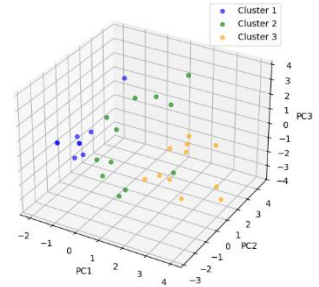
10.

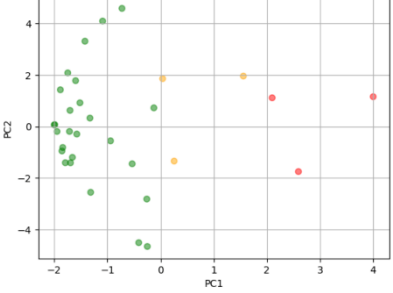
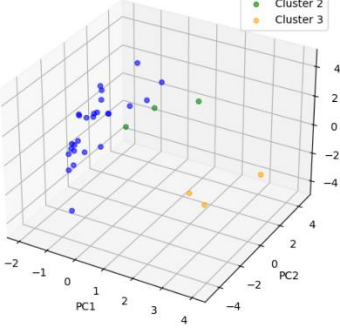
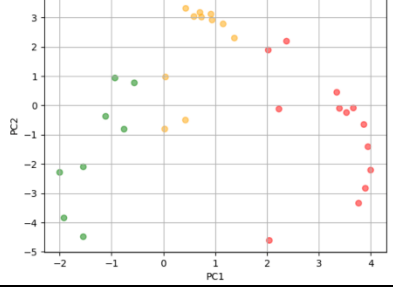
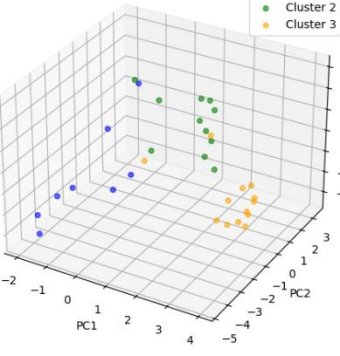


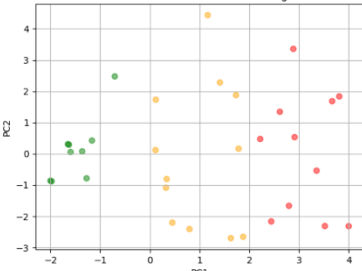
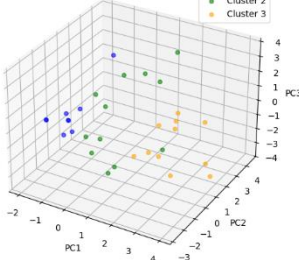
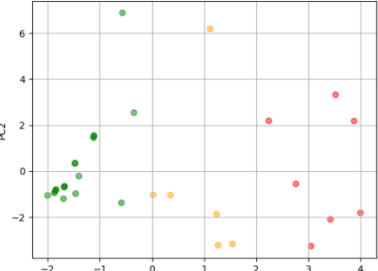
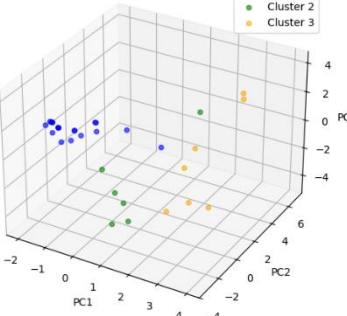
Sedang

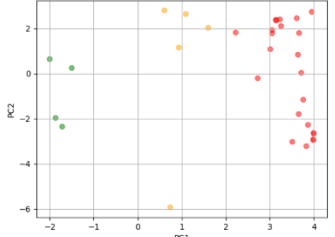
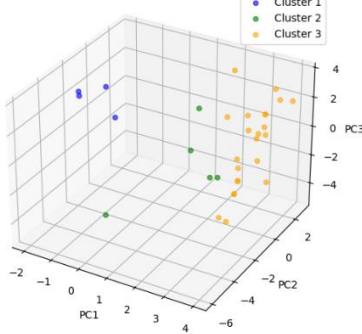
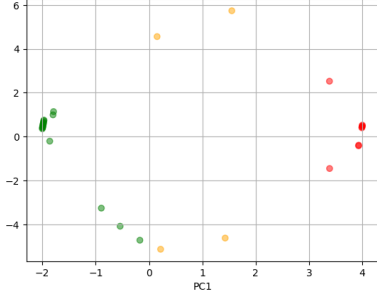
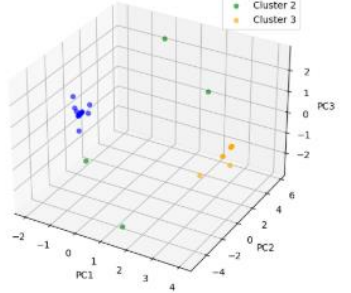


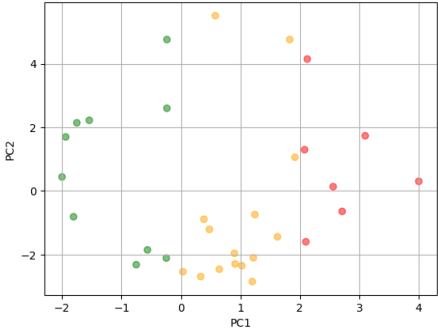
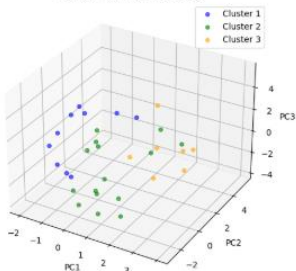
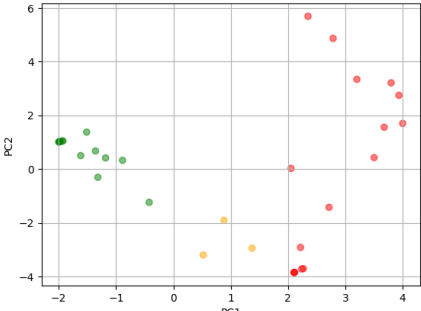
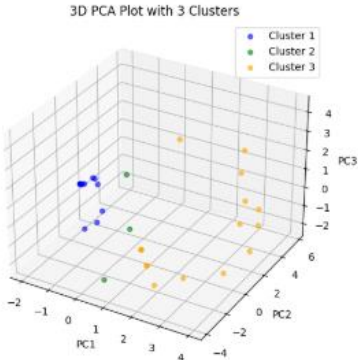
3D PCA Plot with 3 Clusters

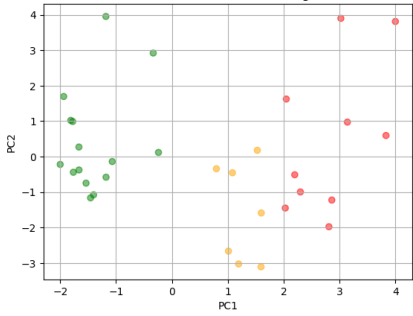
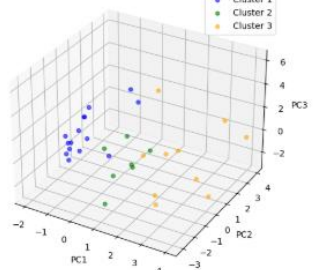
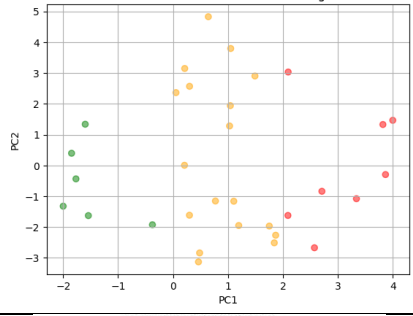


11.		Ringan	Hasil Visualisasi : Rusak Ringan  3D PCA Plot with 3 Clusters 
12.		Berat	Hasil Visualisasi : Rusak Berat  3D PCA Plot with 3 Clusters 

13.		Sedang	Hasil Visualisasi : Rusak Sedang  3D PCA Plot with 3 Clusters 
14.		Ringan	Hasil Visualisasi : Rusak Ringan  3D PCA Plot with 3 Clusters 

15.		Berat	Hasil Visualisasi : Rusak Berat  3D PCA Plot with 3 Clusters 
16.		Ringan	Hasil Visualisasi : Rusak Ringan  3D PCA Plot with 3 Clusters 

17.		Sedang	Hasil Visualisasi : Rusak Sedang  3D PCA Plot with 3 Clusters 
18.		Berat	Hasil Visualisasi : Rusak Berat  3D PCA Plot with 3 Clusters 

19.		Ringan	Hasil Visualisasi : Rusak Ringan  3D PCA Plot with 3 Clusters 
20.		Sedang	Hasil Visualisasi : Rusak Sedang  3D PCA Plot with 3 Clusters 