

**KARAKTERISTIK TINGKAT KERUSAKAN BANGUNAN  
PASCABENCANA ALAM MENGGUNAKAN  
EKSTRAKSI FITUR *GOOGLENET* DAN  
*PRINCIPAL COMPONENT ANALYSIS*  
(PCA)-*CLUSTERING***

**TESIS**

**Oleh:  
HOLIDIYATUL AMANI  
NIM. 230605220016**



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

## **HALAMAN PENGAJUAN**

### **KARAKTERISTIK TINGKAT KERUSAKAN BANGUNAN PASCABENCANA ALAM MENGGUNAKAN EKSTRAKSI FITUR *GOOGLENET* DAN *PRINCIPAL COMPONENT ANALYSIS (PCA)-CLUSTERING***

## **TESIS**

**Diajukan kepada:**

**Universitas Islam Negeri Maulana Malik Ibrahim Malang  
Untuk memenuhi Salah Satu Persyaratan dalam  
Memperoleh Gelar Magister Komputer (M.Kom)**

**Oleh:**

**HOLIDIYATUL AMANI  
NIM. 230605220016**

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

**KARAKTERISTIK TINGKAT KERUSAKAN BANGUNAN PASCABENCANA  
ALAM MENGGUNAKAN EKSTRAKSI FITUR *GOOGLENET* DAN  
*PRINCIPAL COMPONENT ANALYSIS (PCA)-CLUSTERING***

**TESIS**

**Diajukan Kepada:**  
**Fakultas Sains dan Teknologi**  
**Universitas Islam Negeri Maulana Malik Ibrahim Malang**  
**Untuk Memenuhi Salah Satu Persyaratan Dalam**  
**Memperoleh Gelar Magister Komputer (M.Kom)**

**Oleh:**  
**HOLIDIYATUL AMANI**  
**NIM. 230605220016**

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

## HALAMAN PERSETUJUAN

### KARAKTERISTIK TINGKAT KERUSAKAN BANGUNAN PASCABENCANA ALAM MENGGUNAKAN EKSTRAKSI FITUR *GOOGLENET* DAN *PRINCIPAL COMPONENT ANALISYS (PCA)-CLUSTERING*

## TESIS

Oleh:  
**HOLIDIYATUL AMANI**  
**NIM. 230605220016**

Telah diperiksa dan disetujui untuk diuji:  
Tanggal: 6 November 2025

Pembimbing I,



Dr. Agung Teguh Wibowo Almais, M.T  
NIP. 19860301202321 1 016

Pembimbing II,



Dr. Zainal Abidin, M. Kom  
NIP. 19760613 200501 1 004

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



Prof. Dr. Mohammad Faisal, S.Kom., M.T.  
NIP. 19740510 200501 1 007

## HALAMAN PENGESAHAN

### KARAKTERISTIK TINGKAT KERUSAKAN BANGUNAN PASCABENCANA ALAM MENGGUNAKAN EKSTRAKSI FITUR *GOOGLENET* DAN *PRINCIPAL COMPONENT ANALYSIS (PCA)-CLUSTERING*

## TESIS

Oleh:  
**HOLIDIYATUL AMANI**  
**NIM. 230605220016**

Telah Dipertahankan di Depan Dewan Pengaji Tesis  
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan  
Untuk Memperoleh Gelar Magister Komputer (M.Kom)  
Tanggal: 6 November 2025

#### Susunan Dewan Pengaji

Pengaji I : Dr. Ir. Fresy Nugroho, S.T., M.T., IPM., ASEAN Eng  
NIP. 19710722 201101 1 001

Pengaji II : Dr. Fachrul Kurniawan, M. M.T., IPU  
NIP. 19771020 200912 1 001

Pembimbing I : Dr. Agung Teguh Wibowo Almais, M.T  
NIP. 19860301202321 1 016

Pembimbing II : Dr. Zainal Abidin, M. Kom  
NIP. 19760613 200501 1 004

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



Prof. Dr. Ir. Muhammad Faisal, S.Kom., M.T.  
NIP. 19740510 200501 1 007

## HALAMAN PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama : Holidiyatul Amani  
NIM : 230605220016  
Program Studi : Magister Informatika  
Fakultas : Sains dan Teknologi

Menyatakan dengan sebenarnya bahwa Tesis yang saya tulis ini benar-banar merupakan hasil karya saya sendiri, bukan merupakan pengambilalihan 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 Tesis ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 28 Oktober 2025  
Yang Membuat Pernyataan



Holidiyatul Amani  
230605220016

## MOTTO

إِنْهُدْ وَلَا تَكْسَلْ وَلَا تَكُنْ غَافِلًا ﴿فَنَدَمَةُ الْعُقُوبِ لِمَنْ يَتَكَبَّسَ﴾

"Bersungguh-sungguhlah dan jangan bermalas-malasan dan jangan pula lengah,  
karena penyesalan itu bagi orang yang bermalas-malasan."

حَرِبْ وَلَاحِظْ تَكْنْ عَارِفَا

"Cobalah dan perhatikanlah, niscaya kau jadi orang yang tahu."

خَيْرُ النَّاسِ أَحْسَنُهُمْ خُلُقًا وَأَنْفَعُهُمْ لِلنَّاسِ

"Sebaik-baik manusia adalah yang terbaik budi pekertinya dan yang paling bermanfaat  
bagi manusia lainnya."

## **HALAMAN PERSEMBAHAN**

Assalamu'alaikum Wr. Wb.

Kupersembahkan karya Tesis terbaikku ini kepada:

1. Kedua Orang tua saya tercinta, Tahri dan Mei Suryaningsih, serta seluruh keluarga besar, yang telah memberikan dukungan moral, materiel, dan doa yang tiada henti sehingga menjadi kekuatan terbesar bagi penulis.
2. KHR. Azaim Ibrahimy, Selaku pengasuh Ponpes Salafiyah Syafi'iyah, atas segala doa dan bimbingannya.
3. Seluruh Dosen S1 Fakultas SainsTek Universitas Ibrahimy Sukorejo yang telah mendukung untuk terus melanjutkan masa pendidikan saya.
4. Diri sendiri, trimakasih sudah mau berjuang sejauh ini meski jalannya tak semudah yang yang dibayangkan.
5. Saudara dan keluarga besar Muhammad Miftahussurur, Nurul Isya'iyah, Muhammad Abidzar Al Ghifari, Fatimatuz Zahroh, Yuni Faridatur Rohmah, Siska Umaiyyah yang selalu menjadi pendukung dan penyemangat terus-menerus.
6. Sahabat-sahabat dan rekan-rekan seperjuangan Tomy Ivan Sugiharto, Yusril Haza Mahendra, Ahmad Zarkoni, yang telah menjadi teman diskusi dan pemberi semangat dalam suka maupun duka selama menempuh studi.

Wassalamu'alaikum Wr. Wb.

Malang, 28 November 2025



Holidiyatul Amani  
NIM. 230605220016

## KATA PENGANTAR

Assalamu'alaikum Wr. Wb.

Syukur alhamdulillah penulis hantarkan kehadirat Allah SWT yang telah melimpahkan Rahmat dan Hidayah-Nya, sehingga penulis dapat menyelesaikan studi di Program Studi Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang sekaligus menyelesaikan Thesis ini dengan baik

Selanjutnya penulis haturkan ucapan terima kasih seiring do'a dan harapan jazakumullah ahsanal jaza' kepada semua pihak yang telah membantu terselesaiannya Thesis ini. Ucapan terima kasih ini penulis sampaikan kepada:

1. Bapak Dr. H. Agus Mulyono, M.Si., selaku Dekan Fakultas Sains dan Teknologi UIN Maulana Malik Ibrahim Malang.
2. Bapak Prof. Dr. Muhammad Faisal, M.T., selaku Ketua Program Studi Informatika UIN Maulana Malik Ibrahim Malang.
3. Bapak Dr. Agung Teguh Wibowo Almais, M.T., selaku Dosen Pembimbing I, yang telah meluangkan waktu, tenaga, dan pikiran dengan penuh kesabaran dalam memberikan arahan, bimbingan, dan motivasi yang tak ternilai harganya selama proses penelitian ini.
4. Bapak Dr. Zainal Abidin, M. Kom., selaku Dosen Pembimbing II, yang telah memberikan masukan, koreksi, dan wawasan berharga demi kesempurnaan penelitian ini.
5. Bapak Dr. Ir. Fresy Nugroho, S.T., M.T., IPM., ASEAN Eng selaku penguji I sekaligus Dosen Wali dan Bapak Dr. Fachrul Kurniawan, M. MT., IPU selaku penguji II, Fakultas Sains dan Teknologi UIN Maulana Malik Ibrahim Malang, atas ilmu dan pelayanan yang telah diberikan.
6. Segenap sivitas akademika Program Studi Magister Informatika, khususnya para dosen yang telah memberikan ilmu dan bimbingan selama masa studi.
7. Segenap keluarga, yang selalu memberikan semangat dan motivasi hingga saya dapat menyelesaikan tesis ini.
8. Seluruh pihak yang tidak dapat penulis sebutkan satu persatu, yang telah memberikan bantuan baik secara langsung maupun tidak langsung.

Penulis menyadari bahwa dalam penyusunan Tesis ini masih terdapat kekurangan dan penulis berharap semoga Thesis ini bisa memberikan manfaat kepada para pembaca khususnya bagi penulis secara pribadi. Amin Ya Rabbal Alamin.

Wassalamu'alaikum Wr. Wb.

Malang, 28 November 2025



Haudiyatul Amani  
NIM. 230605220016

## DAFTAR ISI

HALAMAN PENGAJUAN .....	i
HALAMAN PERSETUJUAN .....	iii
HALAMAN PENGESAHAN .....	iv
HALAMAN PERNYATAAN.....	v
MOTTO .....	vi
HALAMAN PERSEMBAHAN .....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI .....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL .....	xii
ABSTRAK .....	xiii
ABSTRACT .....	xiv
مستخلص البحث .....	xv
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Pernyataan Masalah.....	4
1.3 Tujuan Penelitian.....	4
1.4 Hipotesis.....	5
1.5 Manfaat Penelitian .....	5
1.6 Batasan Masalah.....	5
1.7 Sistematika Penulisan .....	6
BAB II PUSTAKA TINJAUAN .....	7
2.1 Karakteristik <i>Clustering</i> Berdasarkan Gambar .....	7
2.2 Kerangka Teori.....	13
2.3 Integrasi Al-Qur'an dan Hadis .....	16
2.4 Karakteristik Kerusakan Bangunan .....	18
2.5 Bencana Alam .....	19
BAB III DESAIN DAN INTEGRASI SISTEM PENELITIAN .....	21
3.1 Kerangka Konsep Penelitian .....	21
3.2 Persiapan Data.....	21
3.3 Validasi Data.....	22
3.4 <i>GoogLeNet</i> (Ekstraksi Fitur) .....	23
3.5 <i>Principal Component Analysis (PCA)-Clustering</i> .....	26
3.6 Desain Sistem.....	27
3.7 Skenario Uji Coba .....	35
6.1.1 Validasi Penelitian Terdahulu .....	35
6.1.2 Validasi Ahli Lapangan.....	35
3.8 Instrument Penelitian .....	36
BAB IV IMPLEMENTASI METODE .....	37
4.1 Implementasi <i>GoogLeNet</i> .....	37
4.2 Input <i>Citra</i> .....	37

4.3	Proses Ekstraksi Fitur Menggunakan <i>GoogLeNet</i> .....	38
4.3.1	Pra-pemroses Data .....	39
4.3.2	Pemanggilan Model <i>GoogLeNet</i> .....	39
4.3.3	<i>Proses Forward Pass</i> .....	40
4.3.4	Rekayasa Fitur Tambahan.....	40
4.3.5	Hasil Ekstraksi Fitur.....	41
4.3.6	Penyimpanan Fitur .....	42
4.3.7	Pembentukan dan Normalisasi Vektor Fitur .....	42
4.4	Integrasi Ekstraksi Fitur <i>GoogLeNet</i> dengan PCA- <i>Clustering</i> .....	43
4.4.1	Reduksi Dimensi menggunakan PCA.....	43
4.4.2	Pembentukan Klaster Tingkat Kerusakan .....	46
BAB V	UJI COBA .....	48
5.1	Skenario Uji Coba .....	48
5.2	Input <i>Citra</i> .....	50
5.3	Hasil Ekstraksi Fitur Menggunakan <i>GoogLeNet</i> .....	50
5.4	Visualisasi Fitur Peta <i>GoogLeNet</i> .....	52
5.5	Analisis Fitur Statik Klasik .....	53
5.6	Normalisasi Data.....	55
5.7	Pembentukan N-Komponen PCA .....	56
5.8	<i>Eigen Values</i> dan <i>Eigen Variance Ratio</i> .....	57
5.9	Hasil Visualisasi.....	59
5.10	Hasil Visualisasi Pengelompokan .....	60
BAB VI	PENUTUP .....	64
6.1	Kesimpulan .....	64
6.1.1	Ekstraksi Fitur menggunakan <i>GoogLeNet</i> .....	64
6.1.2	Visualisasi Pengelompokan Menggunakan PCA- <i>Clustering</i> .....	64
6.2	Saran.....	65
	DAFTAR PUSTAKA.....	66

## DAFTAR GAMBAR

Gambar 2.1 Kerangka Teori .....	13
Gambar 3.1 Kerangka Konsep.....	21
Gambar 3.2 Validasi Data .....	22
Gambar 3.3 Arsitektur <i>GoogLeNet (Inception-v3)</i> .....	25
Gambar 3.4 Desain Sistem .....	27
Gambar 3.5 Input <i>Citra</i> dengan Format RGB .....	28
Gambar 3.6 <i>Inception (V3)</i> Module Dengan Ekstraksi Fitur .....	29
Gambar 3.7 Desain Sistem PCA- <i>Clustering</i> .....	32
Gambar 5.1 Kerusakan Bangunan .....	50
Gambar 5.2 Hasil Representasi Intensitas Piksel .....	52
Gambar 5.3 Plot Nilai Eigen dan <i>Rasio Varians</i> untuk Komponen Utama. ....	57
Gambar 5.4 Hasil Visualisasi .....	59

## DAFTAR TABEL

Tabel 2.1 Penelitian Terkait.....	15
Tabel 3.1 Titik Koordinat ( $n$ ) tingkat kerusakan (Almais <i>et al.</i> , 2023).....	34
Tabel 3.2 Instrument Penelitian.....	36
Tabel 5.1 Hasil Ekstraksi Fitur <i>GoogLeNet</i> .....	50
Tabel 5.2 Bentuk Fitur Statik Klasik .....	54
Tabel 5.3 Hasil Normalisasi ( <i>StandardScaler</i> ).....	55
Tabel 5.4 Hasil N-Komponen.....	56
Tabel 5.5 Hasil Analisis Komponen Utama (PCA).....	58
Tabel 5.5 Proyeksi Data Berdasarkan Klaster Hasil PCA- <i>Clustering</i> .....	61
Tabel 5.6 Persentase Hasil Distribusi Label berdasarkan Kluster dari PC1.....	61
Tabel 5.7 Hasil Validasi Eksternal .....	62

## ABSTRAK

Amani, Holidiyatul. 2025. **Karakteristik Tingkat Kerusakan Bangunan Pascabencana Alam Menggunakan Ekstraksi Fitur GoogLeNet dan Principal Component Analysis (PCA)-Clustering.** Tesis. Program Studi Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing (I) Dr. Agung Teguh Wibowo Almais, M.T (II) Dr. Zainal Abidin, M. Kom

Kata kunci: *GoogLeNet*, *PCA-Clustering*, Ekstraksi Fitur, Kerusakan Bangunan, Pascabencana, Analisis *Citra*.

Peningkatan frekuensi bencana dan ketidakakuratan dalam penilaian kerusakan bangunan kerap menghambat proses perencanaan pemulihan, sehingga diperlukan metode yang lebih objektif untuk mendukung kesiapsiagaan dan respons pascabencana. Untuk menjawab kebutuhan tersebut, penelitian ini merancang dan mengimplementasikan sistem identifikasi tingkat kerusakan bangunan pascabencana alam dengan mengintegrasikan arsitektur *deep learning* dan teknik statistik. Metode yang digunakan meliputi ekstraksi fitur menggunakan arsitektur *GoogLeNet*, reduksi dimensi melalui *Principal Component Analysis* (PCA), serta pengelompokan tingkat kerusakan. Proses identifikasi dimulai dengan ekstraksi fitur semantik berdimensi tinggi dari *citra* kerusakan bangunan menggunakan lapisan "mixed2" dari model *GoogLeNet*. Fitur-fitur yang dihasilkan terbukti beragam dan berhasil menangkap perbedaan pola tekstur visual antara area bangunan yang rusak ringan, sedang, dan berat. Selanjutnya, PCA diterapkan pada data yang telah dinormalisasi, dan berhasil mereduksi dimensi data ke dalam dua komponen utama (PC1 dan PC2). Kedua komponen ini secara kolektif mampu menjelaskan 97,77% dari total varian data, menunjukkan efisiensi tinggi dalam merangkum informasi esensial. Berdasarkan ambang batas yang diterapkan pada sumbu PC1, proses klasterisasi berhasil mengelompokkan sampel ke dalam tiga kategori tingkat kerusakan: rusak ringan ( $n < 0$ ), rusak sedang ( $0 \leq n < 2$ ), rusak berat ( $n \geq 2$ ). Hasil distribusi klaster menunjukkan dominasi kategori rusak ringan sebesar 50,0%, diikuti oleh rusak sedang (30,0%), dan rusak berat (20,0%). Konsistensi hasil identifikasi yang diperoleh dari model *GoogLeNet PCA-Clustering* ini terbukti akurat dan dapat diandalkan melalui validasi eksternal dengan penilaian ahli lapangan (BMKG). Penelitian ini menunjukkan bahwa pendekatan hibrida yang terbukti konsisten, akurat, dan dapat digunakan sebagai dasar analisis awal pascabencana.

## ABSTRACT

Amani, Holidiyatul. 2025. **Characteristics of Building Damage Levels After Natural Disasters Using GoogLeNet Feature Extraction and Principal Component Analysis (PCA)-Clustering.** Thesis. Master of Informatics study Program Faculty of Science and Technology State Islamic University Maulana Malik Ibrahim Malang. Advisor (I) Dr. Agung Teguh Wibowo Almais, M.T (II) Dr. Zainal Abidin, M. Kom

Keywords: GoogLeNet, PCA-Clustering, Feature Extraction, Building Damage, Post-Disaster, Image Analysis

The increasing frequency of natural disasters and the inaccuracy of conventional damage-assessment methods often hinder effective recovery planning, highlighting the need for more objective and reliable approaches to support post-disaster preparedness and response. Addressing this issue, this study designs and implements a system for identifying levels of building damage after natural disasters by integrating deep-learning architecture with statistical techniques. The method employs GoogLeNet for semantic feature extraction, Principal Component Analysis (PCA) for dimensionality reduction, and threshold-based clustering for damage categorization. The identification process begins with extracting high-dimensional semantic features from building-damage images using the “mixed2” layer of GoogLeNet, which successfully captures variations in visual texture patterns across light, moderate, and severe damage. PCA, applied to the normalized data, reduces the dimensionality into two principal components (PC1 and PC2), which collectively explain 97.77% of the total variance, demonstrating strong efficiency in summarizing essential information. Using threshold values applied to PC1, the clustering process categorizes samples into three damage levels: light ( $n < 0$ ), moderate ( $0 \leq n < 2$ ), and severe ( $n \geq 2$ ). The resulting distribution shows that light damage dominates at 50.0%, followed by moderate (30.0%) and severe (20.0%). Consistency and accuracy of the identification results are reinforced through external validation with field experts (BMKG). This study demonstrates that the proposed hybrid model yields reliable and consistent outcomes and can serve as a dependable basis for preliminary post-disaster analysis.

## مستخلص البحث

أماني، هوليديا تول. 2025. خصائص مستويات أضرار المباني بعد الكوارث الطبيعية باستخدام استخراج الميزات من **GoogLeNet** وتحليل المكونات الرئيسية (**PCA**) وتجميعها. رسالة ماجستير. برنامج ماجستير علم المعلومات، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانغ .المشرفان (II): د. أغونغ تغوه وييورو ألينس، م.ت (II) د. زين العابدين، م.كوم.

الكلمات المفتاحية: GoogLeNet، تحليل المكونات الرئيسية، التجميع العنودي، استخراج الخصائص، أضرار المباني، ما بعد الكوارث، تحليل الصور.

إنَّ ازدياد وتيرة الكوارث الطبيعية وعدم دقة الأساليب التقليدية في تقييم مستوى الأضرار يُعيقان في الغالب فعالية تحطيط مرحلة التعافي، الأمر الذي يستدعي اعتماد منهجيات أكثر موضوعية وموثوقة لدعم جاهزية الاستجابة بعد الكارثة. وانطلاقاً من هذه الحاجة، يهدف هذا البحث إلى تصميم وتنفيذ نظام لتحديد مستوى ضرر المباني بعد الكوارث الطبيعية من خلال دمج بنية التعلم العميق مع التقنيات الإحصائية. تعتمد المنهجية على استخدام شبكة GoogLeNet لاستخلاص الخصائص الدلالية، وتحليل المركبات الرئيسية (**PCA**) لخفض الأبعاد، إضافةً إلى أسلوب التجميع المعتمد على العيّنات لتصنيف مستويات الضرر. يبدأ نظام التعرف باستخلاص خصائص دلالية عالية الأبعاد من صور أضرار المباني عبر الطبقة mixed2 في غودج GoogLeNet، وقد أثبتت هذه الخصائص قدرتها على تمييز الأنماط البصرية للنسيج بين الضرر الخفيف والمتوسط والشديد. وبعد ذلك، يُطبّق تحليل المركبات الرئيسية على البيانات الموحدة، مما يقلّص الأبعاد إلى مركبين رئيسيين (PC1 و PC2) يفسران معاً نسبة 97.77% من التباين الكلي، وهو ما يدل على كفاءته العالية في تلخيص المعلومات الجلوبية. وبالاستناد إلى قيم العيّنة على المحوّر PC1، تمكن أسلوب التجميع من تصنيف العيّنات إلى ثلاثة مستويات للضرر: ضرر خفيف ( $0 < 0 \leq n$ )، ضرر متوسط ( $0 < 2 \leq n$ )، وضرر شديد ( $2 \geq n$ ). وقد أظهرت نتائج التوزيع سيطرة فئة الضرر الخفيف بنسبة 50.0%， تليها فئة الضرر المتوسط (30.0%)، ثم الضرر الشديد (20.0%). وقد تم تعزيز اتساق ودقة النتائج من خلال التحقق الخارجي بمقارنة تقدّيرات النموذج مع تقييم الخبراء الميدانيين (BMKG). ويرهن هذا البحث على أنَّ النموذج المقترن يتميّز بالاتساق والموثوقية، ويمكن اعتماده كأساس لتحليل أولي بعد الكوارث.

## BAB I

### PENDAHULUAN

#### 1.1 Latar Belakang

Bencana alam didefinisikan sebagai bencana yang disebabkan oleh peristiwa alam (Darmawan *et al.* 2022). Fenomena ini semakin meningkat frekuensinya, baik berupa gempa bumi maupun banjir, sehingga menimbulkan kerugian signifikan dari segi finansial maupun sosial (Murdiaty *et al.*, 2020). Sebagai peristiwa yang umum dialami makhluk hidup di Bumi, bencana alam juga menjadi pengingat bagi umat Muslim untuk percaya bahwa segala sesuatu terjadi atas kehendak Allah SWT. Hal ini tercermin dalam Al-Qur'an, seperti pada surat Al-Hadid ayat 22 dan surah Al-Zalzalah ayat 1-2 menegaskan:

مَا أَصَابَ مِنْ مُصِيبَةٍ فِي الْأَرْضِ وَلَا فِي النُّسُكِ إِلَّا فِي كِتَابٍ أُنْ قَبِيلَ أَنْ تَبْرَأَهُ إِنَّ ذَلِكَ عَلَى اللَّهِ يَسِيرٌ ﴿٢٢﴾

Artinya: "Tidak ada bencana (apa pun) yang menimpa di Bumi dan tidak (juga yang menimpa) dirimu, kecuali telah tertulis dalam Kitab (Lauh Mahfuz) sebelum Kami mewujudkannya. Sesungguhnya hal itu mudah bagi Allah." (Q.S. Al-Hadid: 22)

إِذَا زُلْزِلَتِ الْأَرْضُ زُلْزِلَهَا ﴿١﴾ وَأَخْرَجَتِ الْأَرْضُ أَخْرَاجَهَا ﴿٢﴾

Artinya: "Apabila Bumi diguncangkan dengan guncangan yang dahsyat" "Dan bumi telah mengeluarkan beban-beban berat (yang dikandung) nya." (Q.S. Al-Zalzalah: 1-2)

Ayat-ayat tersebut menegaskan bahwa bencana alam terjadi sesuai ketetapan Allah yang telah ditetapkan sebelumnya, serta menggambarkan guncangan dahsyat ketika bumi mengeluarkan beban-bebannya sebagai peringatan, bukti kekuasaan, dan pengingat bagi manusia agar lebih bersabar dalam menghadapi ujian yang terjadi di Bumi.

Intensitas bencana di Indonesia telah meningkat dalam beberapa tahun terakhir, dipengaruhi faktor alam, non-alam, serta aktivitas manusia. Oleh karena itu, kesiapsiagaan masyarakat menjadi aspek krusial untuk meminimalisir dampaknya (Pokhrel, 2024). Namun dalam konteks penilaian karakteristik tingkat kerusakan tingkat kerusakan bangunan pascabencana masih menggunakan metode konvensional yang masih bergantung pada pengalaman langsung para ahli sering kali memakan waktu dan sumber daya yang besar (Utomo & Marta, 2022).

Tim *Disaster Management Planning and Control Team* (P3B) pun membutuhkan data serta informasi perencanaan pemulihan yang efektif, tetapi banyak masalah di lapangan timbul akibat ketidakakuratan data (Safitri *et al.*, 2022). Ketidaksesuaian ini sering disebabkan oleh perbedaan kategorisasi data awal dari peninjauan lapangan yang beragam, sehingga data bencana yang disampaikan ke Badan Penanggulangan Bencana Daerah (BPBD) Provinsi tidak mencerminkan kondisi aktual. Akibatnya, kriteria persiapan rehabilitasi dan rekonstruksi pasca bencana menjadi tidak konsisten (Fatul Fu *et al.*, 2023).

Untuk mengatasi tantangan ini, diperlukan pendekatan teknologi yang lebih efisien dan akurat, seperti pemanfaatan pembelajaran mendalam (*deep Learning*). Salah satu solusi potensial adalah ekstraksi fitur yang memungkinkan analisis mendalam dengan mengidentifikasi pola kompleks dalam data *citra* (Geng *et al.*, 2025), diikuti oleh pengelompokan menggunakan *Principal Component Analysis* (PCA)-*Clustering*. Metode ini tidak hanya menjanjikan kecepatan, tetapi juga akurasi dalam penilaian kerusakan bangunan, karena *clustering* memfasilitasi identifikasi pola data *citra* yang lebih kompleks (Almais *et al.* 2023). Ekstraksi fitur

sendiri merupakan langkah penting dalam pemrosesan data, yang dikenal sebagai proses pengambilan ciri objek untuk menggambarkan karakteristiknya (Satria & Mushthofa, 2013). Tujuannya adalah mengekstraksi fitur esensial dari data asli, mengurangi fitur melalui algoritma pemilihan tertentu, sehingga meningkatkan akurasi prediksi dan mengurangi biaya komputasi untuk tahap *clustering* selanjutnya (Almais *et al.*, 2024).

Beberapa penelitian terdahulu telah mengeksplorasi pendekatan serupa dengan penelitian (Baral *et al.*, 2024) fokus pada metode klasifikasi berbasis *deep learning* menggunakan algoritma *ResNet-50* dan *GoogLeNet* untuk deteksi kerusakan struktur, di mana *ResNet-50* memanfaatkan *residual learning* guna menghinari *vanishing gradient*, sementara *GoogLeNet* menerapkan *inception modules* untuk ekstraksi fitur multi-skala. Sementara itu (Chen *et al.*, 2023) menggabungkan *GoogLeNet* dengan *transfer learning* untuk klasifikasi tumor payudara pada gambar *Ultrasound*, memilih *GoogLeNet* karena efisiensi komputasinya dan kemampuan ekstraksi fitur multi-skala melalui modul inception.

Meskipun demikian penelitian-penelitian tersebut memiliki keterbatasan, seperti ketergantungan pada dataset tidak seimbang, kurangnya analisis mendalam terhadap fitur yang diekstraksi, serta belum dilakukannya pendekatan pengelompokan tingkat kerusakan tanpa label. Penelitian ini mengusulkan solusi inovatif dengan menggantikan klasifikasi konvensional menggunakan pendekatan PCA-*Clustering* setelah ekstraksi fitur *GoogLeNet*, guna mengidentifikasi karakteristik objek gambar kerusakan bangunan pascabencana secara lebih terstruktur. *GoogLeNet* dipilih untuk mengekstraksi fitur visual seperti retakan,

deformasi, atau kerusakan struktural, karena efisiensinya dalam menghasilkan fitur dengan parameter yang lebih sedikit. Kemudian, PCA-*Clustering* digunakan untuk mereduksi dimensi data dan menghasilkan label berdasarkan skala nilai normalisasi, yang dikelompokkan melalui teknik *Clustering* (Almais *et al.* 2023).

Melalui pendekatan ini, penelitian diharapkan memberikan solusi inovatif untuk meningkatkan efektifitas penilaian kerusakan bangunan pascabencana, dengan memanfaatkan ekstraksi fitur gambar menggunakan *GoogLeNet* dan pengelompokan hasil melalui PCA-*Clustering*. Pada akhirnya, hasil penelitian ini akan menghasilkan informasi tingkat kerusakan yang lebih objektif bagi pemangku kepentingan, termasuk pemerintah dan BPBD, sehingga proses pendistribusian bantuan dapat lebih tepat sasaran berdasarkan kondisi aktual di lapangan.

## 1.2 Pernyataan Masalah

Identifikasi ekstraksi fitur objek gambar sebagai parameter kerusakan bangunan pascabencana alam menggunakan *GoogLeNet*.

Mengukur efektivitas pengelompokan berbasis PCA untuk tingkat kerusakan bangunan pascabencana alam.

## 1.3 Tujuan Penelitian

1. Merancang sistem identifikasi untuk ekstraksi fitur gambar kerusakan bangunan menggunakan arsitektur *GoogLeNet*.
2. Menghasilkan visualisasi pengelompokan berbasis PCA untuk tingkat kerusakan bangunan pascabencana alam.

## 1.4 Hipotesis

Metode *hybrid* yang menggabungkan ekstraksi fitur menggunakan *GoogLeNet* dan reduksi dimensi serta pengelompokan menggunakan *PCA-Clustering* menghasilkan pengelompokan tingkat kerusakan bangunan pascabencana yang lebih tinggi dan perbedaan yang signifikan secara statistik dibandingkan dengan pendekatan tunggal (hanya *GoogLeNet* atau hanya *PCA-Clustering*).

## 1.5 Manfaat Penelitian

1. Meningkatkan pengetahuan tentang penerapan teknologi *deep learning* dan analisis *clustering* dalam kerusakan bangunan setelah bencana.
2. Membantu pengambilan keputusan yang lebih cepat dan akurat terkait rehabilitasi dan rekonstruksi bangunan pascabencana.
3. Memberikan informasi data tentang tingkat kerusakan bangunan di wilayah terdampak, sekaligus meningkatkan kesadaran masyarakat dan pemangku kepentingan untuk mengantisipasi bencana di masa depan.

## 1.6 Batasan Masalah

1. Penelitian ini dibatasi pada penggunaan data gambar kerusakan bangunan yang dihasilkan dari ekstraksi fitur berupa vector fitur.
2. Sumber data gambar diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) Malang.
3. Selain itu, penilaian tingkat kerusakan bangunan tidak mempertimbangkan faktor jenis mutu yang mengacu pada kualitas dan material bangunan, guna menjaga fokus pada aspek teknologi ekstraksi dan pengelompokan fitur.

## **1.7 Sistematika Penulisan**

### **1. BAB I: PENDAHULUAN**

Bab ini membahas tentang Latar Belakang, Rumusan Masalah, Tujuan Penelitian, Batasan Masalah, Manfaat Penelitian serta Sistematika Pembahasan.

### **2. BAB II: STUDI PUSTAKA**

Dalam bab ini akan dijabarkan mengenai studi pustaka yang meliputi penjelasan tentang Karakteristik *Clustering* Berdasarkan Gambar, Penelitian Terkait dan Kerangka Teori.

### **3. BAB III: METODOLOGI PENELITIAN**

Bab ini memuat tentang Prosedur Penelitian, Kerangka Konsep Penelitian, Instrumen Penelitian dan Skenario Uji Coba.

### **4. BAB IV: PEMBAHASAN**

Pada bab ini menjelaskan tentang hasil implementasi karakteristik tingkat kerusakan bangunan menggunakan ekstraksi fitur *GoogLeNet* dan PCA-*Clustering*.

### **5. BAB V: KESIMPULAN**

Bab ini memuat tentang simpulan dari hasil penelitian yang dilakukan dan saran saran untuk pengembangan tesis ini lebih lanjut.

## **BAB II**

### **PUSTAKA TINJAUAN**

#### **2.1 Karakteristik *Clustering* Berdasarkan Gambar**

Penelitian ini bertujuan untuk mengkaji karakteristik pengelompokan data berdasarkan gambar, khususnya dalam konteks penentuan tingkat kerusakan pascabencana alam. Pada kajian literatur ini akan diulas penelitian terdahulu yang relevan mengenai penerapan pengolahan *citra* dan teknik pembelajaran mesin untuk mendeteksi perubahan serta melakukan pengelompokan.

Penelitian (Khajwal *et al.*, 2023) mengusulkan metode klasifikasi kerusakan bangunan pascabencana menggunakan arsitektur *Multi-View CNN* (MV-CNN) yang memanfaatkan gambar dari berbagai sudut pandang untuk meningkatkan akurasi. Permasalahan yang diangkat adalah ketergantungan pada penilaian ahli serta keterbatasan model berbasis gambar tunggal. Model dikembangkan melalui segmentasi bangunan dengan *PSPNet* (berbasis *MobileNet*) dan penggunaan fungsi *loss Earth Mover's Distance* (EMD) untuk menangani sifat ordinal kelas kerusakan FEMA. Eksperimen pada dataset 2.500 gambar pasca-*Hurricane Harvey* (500 bangunan, lima sudut pandang) menghasilkan akurasi 65% untuk lima kelas dan 81% dengan toleransi  $\pm 1$  kelas. Subjek penelitian berfokus pada bangunan residensial, dengan label kerusakan sesuai standar HAZUS-MH FEMA.

Penelitian (Chen *et al.*, 2023) membahas klasifikasi gambar ultrasonografi payudara untuk deteksi kanker menggunakan *GoogLeNet* dengan *transfer learning* dan pra-pemrosesan berbasis model *Total Variation* (TV). Permasalahan utamanya adalah keterbatasan metode konvensional serta biaya tinggi tenaga ahli.

Metode mencakup filter median, pemulihan gambar dengan model TV, ekstraksi fitur multiskala melalui model inception, dan klasifikasi dengan *GoogLeNet*. Hasil menunjukkan akurasi 96,37% dengan loss 0,3492, lebih unggul dibanding *AlexNet* dan *LeNet5*. Dataset terdiri dari 880 *citra* (103 normal, 467 ganas, 210 jinak), dengan input berupa fitur tekstur, bentuk, dan intensitas piksel untuk klasifikasi sehat, jinak, atau ganas.

Penelitian (Almais *et al.*, 2023) menerapkan (PCA) untuk mengelompokkan dan melabeli tingkat kerusakan pascabencana secara *unsupervised*. Permasalahan yang diangkat adalah belum adanya standar numerik dalam penentuan label kerusakan (ringan, sedang, berat) meski aturan pemerintah sudah ada. PCA digunakan untuk reduksi dimensi, dengan PC1 (85,16%) lebih dominan dibanding PC2 (9,35%), dan hasil klasterisasi tiga kelompok menunjukkan kesesuaian dengan label asli:  $n < 0$  (ringan),  $0 \leq n < 2$  (sedang), dan  $n \geq 2$  (berat). Dataset terdiri dari 216 data kerusakan di Jawa Timur meliputi variabel kondisi fisik, struktur, fungsi, dan pendukung bangunan, yang dinormalisasi dengan *StandardScaler*. Hasil menunjukkan PCA efektif tidak hanya untuk reduksi fitur sekaligus pelabelan otomatis, meskipun masih bergantung pada kualitas data survei dan memerlukan validasi lebih lanjut.

Penelitian (Yasi *et al.*, 2024) membahas deteksi banjir melalui klasifikasi gambar menggunakan *deep ensemble learning*. Permasalahan utamanya adalah meningkatnya frekuensi banjir serta keterbatasan metode konvensional akibat variasi pencahayaan, tekstur air, dan latar belakang. Metode yang diusulkan

menggabungkan delapan arsitektur CNN (*MobileNet V2*, *ResNet 50*, *VGG 16*, *DenseNet 201*, *Inception V3*, *EfficientNet B5*, *NasNet Large*, *Xception*) dengan *transfer learning* dan teknik *weighted averaging*. Hasil menunjukkan akurasi validasi 98,51% dan F1-score 97%, dengan *MobileNet V2* mencapai akurasi 97,57% sehingga cocok untuk perangkat ringan. Dataset terdiri dari 671 gambar (340 banjir, 331 non-banjir) dari *FloodNet* dan *Flood Area Segmentation*, dengan input berupa fitur visual seperti tekstur, warna, dan pola air. Penelitian ini relevan untuk sistem deteksi banjir berbasis AI, meskipun tantangannya adalah kebutuhan komputasi tinggi dan kualitas dataset.

Sementara itu (Reis *et al.*, 2024) membahas deteksi retakan jalan pascagempa dengan *deep learning* dan *ensemble learning*. Permasalahan utamanya adalah kebutuhan pemeliharaan infrastruktur jalan pascabencana, di mana metode konvensional kurang akurat pada latar belakang kompleks. Kontribusi utama adalah dataset baru KERCD-2023 berisi gambar jalan retak dan tidak retak dari Turki, serta metode *hybrid* yang mengekstraksi fitur dengan *ResNet152* dan mengklasifikasikannya menggunakan algoritma seperti SVM, KNN, AdaBoost, *Naive Bayes*, dan *ensemble learning* (EL). Hasil menunjukkan kombinasi *ResNet152*, SVM, EL mencapai akurasi tertinggi 98,68%, mengungguli metode lain. Dataset terdiri dari dua kelas (retak dan tidak retak) dengan input berupa fitur visual tekstur dan pola retakan. Temuan ini menjadi dasar penting untuk pengembangan sistem otomatis pemantauan infrastruktur pascabencana.

Penelitian (Almais *et al.*, 2024) membahas integrasi GLCM (*Gray Level Co-*

(*occurrence Matrix*) dan PCA untuk klasifikasi tingkat kerusakan bangunan pascabencana. Permasalahan utamanya adalah kebutuhan metode cepat dan akurat dibanding inspeksi manual yang lambat dan subjektif. Metode melibatkan ekstraksi fitur tekstur (kontras, homogenitas, energi, dll.) melalui GLCM, dilanjutkan reduksi dimensi dan klasterisasi dengan PCA. Hasil menunjukkan jarak piksel 0.5 menghasilkan klaster lebih akurat sesuai validasi surveyor dibanding jarak 1, dengan pemisahan tiga tingkat kerusakan: ringan ( $n < 0$ ), sedang ( $0 \leq n < 2$ ), dan berat ( $n \geq 2$ ). Dataset berupa *citra* bangunan pascabencana dari *Kaggle*, dengan input dinormalisasi sebelum diproses PCA. Keunggulan metode ini adalah efisiensi komputasi dan kemampuan bekerja dengan data terbatas, meski masih memerlukan standar interpretasi koordinat PCA yang jelas.

Penelitian (Wu *et al.*, 2024) membahas klasifikasi tingkat kerusakan rumah pascabanjir menggunakan fungsi gambar *dual-view* dan modul perhatian CBAM. Permasalahan utamanya adalah keterbatasan metode konvensional yang bergantung pada penilaian ahli yang subjektif dan lambat, serta model sebelumnya yang hanya memakai gambar tunggal. Metode yang diusulkan adalah *Dual-View CNN* (DV-CNN) berbasis *ResNet-50* dengan *transfer learning* dan CBAM, yang mencapai akurasi 92,5%, lebih unggul dibanding *ResNet-50/101*. Dataset berisi 5890 gambar *indoor* dan *outdoor* dari 2945 rumah pasca-banjir Zhengzhou 2021 dengan empat tingkat kerusakan (A–D). Input berupa fitur visual *indoor-outdoor* yang dipilih melalui damage filter, lalu divisualisasikan dengan *Grad-CAM* untuk interpretasi. Hasil menunjukkan DV-CNN efektif mendeteksi kerusakan struktural

(retak, runtuh) dan non-struktural (*peeling* dinding). Metode ini menekankan pentingnya integrasi perspektif *indoor-outdoor* untuk penilaian komprehensif, meski perlu adaptasi untuk bencana lain.

Penelitian (Baral *et al.*, 2024) membahas perbandingan kinerja *ResNet-50* dan *GoogLeNet* dalam deteksi kerusakan struktural pascabencana. Permasalahan utamanya adalah keterbatasan inspeksi manual yang lambat, mahal, dan subjektif. Metode menggunakan *transfer learning* dari *ImageNet*, lalu dilatih ulang dengan dataset khusus kerusakan (retak, deformasi, kelainan permukaan) yang diperoleh dari berbagai sumber, termasuk "*Structural-Damage Image Captioning Dataset*". Hasil menunjukkan *GoogLeNet* unggul dengan akurasi 97,5% dibanding *ResNet-50* (97,2%), berkat *modul Inception* yang mengekstrak fitur multi-skala lebih efektif. *GoogLeNet* juga lebih baik pada presisi (97,5%), recall (97,3%), dan F1-score (97,4%). Dataset berupa *citra* 180x180 piksel yang diperluas dengan augmentasi, mencakup bangunan rusak dan tidak rusak. Temuan ini menegaskan relevansi *deep learning* untuk otomatisasi penilaian kerusakan infrastruktur pascabencana.

Penelitian (Shao & Xu, 2025) membahas pengenalan adegan bencana alam dengan *Mamba Multimodal Disaster Recognition Network (Mamba-MDRNet)* yang memanfaatkan model besar terlatih (LLMs) dan mekanisme Mamba. Permasalahan utamanya adalah biaya tinggi dan keterbatasan model visual tradisional dalam menangkap informasi multimodal. Metode melibatkan tiga tahap: ekstraksi fitur (gambar dengan *EfficientNet*, teks dengan *TextCNN*, probabilitas

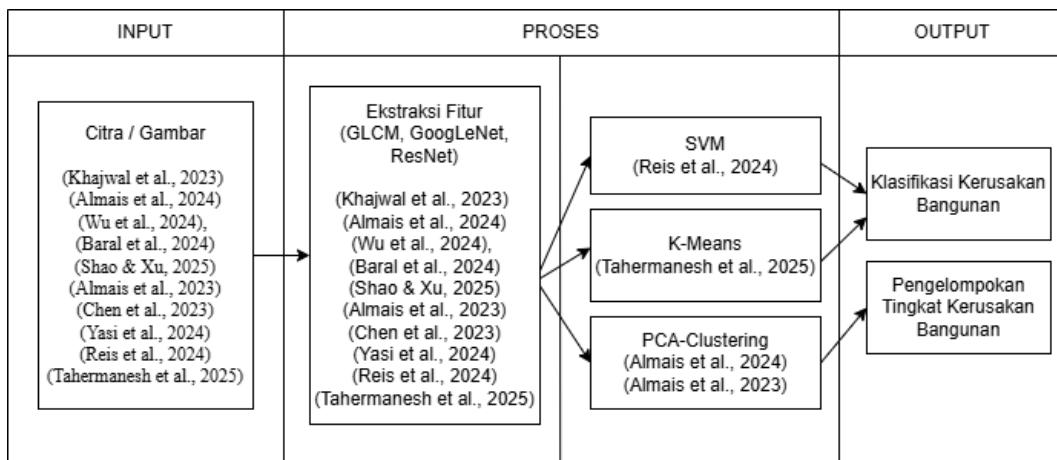
dengan MLP), fusi fitur menggunakan mekanisme Mamba, dan klasifikasi akhir dengan *fully connected layer*. Hasil eksperimen menunjukkan akurasi 97,82%, lebih baik dibanding CNN (91,75%) dan *ResNet18* (97,25%), dengan peningkatan signifikan pada kompleksitas adegan bencana. Dataset terdiri dari 5.277 gambar mencakup 11 jenis bencana (banjir, gempa, badai salju, dll.) dengan input berupa fitur visual, teks, dan distribusi probabilitas dari LLM. Temuan ini menegaskan potensi *Mamba-MDRNet* untuk pemantauan bencana waktunya nyata, terutama pada sistem dengan sumber daya terbatas dan integrasi multimodal.

Penelitian (Tahermanesh *et al.*, 2025) membahas deteksi perubahan pada *citra optik multi-temporal* dengan *clustering* berbasis *deep learning*. Permasalahan utamanya adalah keterbatasan metode tradisional seperti *pixel-based differencing* dan *Change Vector Analysis* dalam menangkap hubungan *spektral-spasial*. Metode yang diusulkan menghadirkan dua model, *3Branch-3DConvNet (3BDCN)* dan *3Branch-3DConvLSTMNet (3BDCLN)*, yang mampu mengekstrak fitur *spektral-spasial* secara lebih komprehensif. Prosesnya melibatkan pembuatan sampel pelatihan semi-otomatis dari *citra* deret waktu dengan *k-means* dan *Elbow method* untuk menentukan jumlah cluster optimal. Hasil eksperimen menunjukkan peningkatan akurasi signifikan dengan koefisien kappa naik 1,5%–8,5% dibanding metode lain. Dataset penelitian berasal dari tiga wilayah (Mazandaran, Danau Urmia, Iran; Danau Poopó, Bolivia) dengan *citra* Landsat dan Sentinel-2, mencakup area urban, vegetasi, air, dan tanah. Variabel input berupa nilai piksel multi-temporal dan fitur spektral-spasial, dengan fokus pemantauan perubahan

tutuhan lahan skala besar.

## 2.2 Kerangka Teori

Berdasarkan kajian di atas, penelitian ini merujuk pada studi sebelumnya yang relevan dengan topik penelitian ini, digunakan sebagai referensi dan bahan perbandingan untuk mendukung pendekatan yang diusulkan. Berikut adalah ringkasan dalam tabel 2.1:



Gambar 2.1 Kerangka Teori

Pemrosesan data citra untuk identifikasi kerusakan bangunan, yang dimulai dari tahap *input*, proses, hingga menghasilkan *output*. Pada tahap *input*, data berupa *citra* atau gambar kerusakan bangunan dikumpulkan dari berbagai sumber penelitian sebelumnya (Khajwal et al., 2023; Almais et al., 2024; Wu et al., 2024; Baral et al., 2024; Shao & Xu, 2025; Chen et al., 2023; Yasi et al., 2024; Reis et al., 2024; Tahermanesh et al., 2025). *Citra* tersebut berfungsi sebagai dasar analisis visual untuk menilai karakteristik kerusakan melalui teknik pembelajaran mesin dan *feature extraction*.

Tahap proses dimulai dengan ekstraksi fitur yang dilakukan menggunakan beberapa pendekatan, seperti GLCM, *GoogLeNet*, dan *ResNet*. Metode-metode ini digunakan untuk menangkap ciri visual penting dari *citra*, termasuk tekstur, pola kerusakan, serta sifat struktural bangunan. Proses ekstraksi fitur merujuk pada penelitian-penelitian terdahulu yang membuktikan efektivitas kombinasi metode statistik dan *deep learning* dalam analisis kerusakan (Khajwal et al., 2023; Almais et al., 2024; Wu et al., 2024; Baral et al., 2024; Shao & Xu, 2025; Chen et al., 2023; Yasi et al., 2024; Reis et al., 2024; Tahermanesh et al., 2025). Selanjutnya, fitur yang telah diperoleh diproses menggunakan tiga pendekatan berbeda, yakni SVM, K-Means, dan PCA-*Clustering*. Setiap metode dipilih berdasarkan tujuan analisis, baik untuk klasifikasi maupun pengelompokan.

Tahap output menghasilkan dua bentuk keluaran utama, yaitu klasifikasi kerusakan bangunan dan pengelompokan tingkat kerusakan bangunan. Metode SVM digunakan untuk melakukan klasifikasi tingkat kerusakan berdasarkan pola yang telah dipelajari (Reis et al., 2024), sedangkan K-Means dan PCA-*Clustering* (Almais et al., 2024; Almais et al., 2023) digunakan untuk mengelompokkan data berdasarkan kemiripan fitur. Hasil ini membantu memudahkan proses evaluasi pascabencana, memberikan gambaran tingkat keparahan kerusakan, serta mendukung perencanaan rehabilitasi dan rekonstruksi bangunan yang lebih tepat sasaran.

Dari beberapa penelitian di atas pada pemanfaatan teknologi kecerdasan buatan, khususnya model *deep learning* seperti *GoogLeNet* dan metode statistik seperti

PCA dan *clustering* (pengelompokan data), untuk mendeteksi dan pengelompokan tingkat kerusakan bangunan pascabencana alam. *GoogLeNet* digunakan untuk mengenali pola visual dari kerusakan dalam *citra*, sementara PCA berfungsi menyederhanakan data besar menjadi fitur penting yang lebih mudah diproses. Metode *clustering* membantu mengelompokkan tingkat kerusakan berdasarkan kemiripan visual. Berdasarkan pada Tabel 2.1:

Tabel 2.1 Penelitian Terkait

Referensi	Topik	Metode	Subjek
(Khajwal <i>et al.</i> , 2023), (Almais <i>et al.</i> , 2024), (Wu <i>et al.</i> , 2024), (Baral <i>et al.</i> , 2024) (Shao & Xu, 2025), (Almais <i>et al.</i> , 2023)	Analisis dan Klasifikasi Kerusakan Bangunan	MV-CNN	Menilai gambar bangunan pasca-badai Harvey
		GLCM, PCA-Clustering	Analisis <i>citra</i> bangunan pasca-bencana
		CNN (ResNet50, GoogLeNet)	Deteksi gambar struktur bangunan pascabencana
		GoogLeNet, ResNet50	Klasifikasi gambar rumah pasca-banjir
		LLM, Mamba-MDRNet	Klasifikasi gambar visual bencana alam
		PCA-Clustering	Mencluster 216 data sektor pasca-bencana
(Chen <i>et al.</i> , 2023)	Klasifikasi kanker payudara	GoogLeNet	Klasifikasi 880 <i>citra</i> US payudara (700 pasien)
(Yasi <i>et al.</i> , 2024)	Klasifikasi dataset area banjir	CNN	Klasifikasi dataset FloodNet & segmentasi area banjir
(Reis <i>et al.</i> , 2024)	Deteksi retakan jalan	ResNet152, SVM	Deteksi gambar jalan pasca-gempa
(Tahermanesh <i>et al.</i> , 2025)	Deteksi perubahan <i>citra</i>	LSTM, 3D-CNN, 3D-ConvLSTM, Clustering	Deteksi <i>citra</i> satelit time-series di 3 lokasi studi
Penelitian Ini	Clustering kerusakan bangunan	GoogLeNet, PCA-Clustering	Mencluster <i>citra</i> bangunan pascabencana alam

Berdasarkan 10 tinjauan literatur pada tabel 2.1, pendekatan yang diusulkan dalam penelitian ini mengintegrasikan elemen dari studi-studi tersebut. Menurut Almais *et al.* (2024), menggunakan GLCM dan PCA-*Clustering* untuk

menganalisis gambar bangunan pascabencana guna mengelompokkan tingkat kerusakan berdasarkan karakteristik tekstur visual (Almais *et al.*, 2024). Sementara itu, Baral dkk. (2024) menerapkan *GoogLeNet*, *ResNet50*, dan *Transfer Learning* untuk mengklasifikasikan gambar rumah pasca banjir, dalam mendekripsi tingkat kerusakan bangunan (Baral *et al.*, 2024). Menurut Chen dkk. (2023), memanfaatkan *GoogLeNet* dan *Transfer Learning* untuk mengklasifikasikan 880 gambar ultrasonografi payudara dari 700 pasien untuk deteksi kanker payudara (Chen *et al.*, 2023). Yasi dkk. (2024) menerapkan *Deep Ensemble CNN* dengan delapan arsitektur untuk klasifikasi dan segmentasi area banjir menggunakan dataset *FloodNet* (Yasi *et al.* 2024). Reis dkk. (2024) menggunakan ResNet152, SVM, dan *Ensemble Learning* untuk mendekripsi kerusakan jalan dari gambar pascagempa Bumi (Reis *et al.*, 2024). Di sisi lain, Tahermanesh dkk. (2025) menggunakan LSTM, 3D-CNN, dan 3D-ConvLSTM berdasarkan Clustering untuk mendekripsi perubahan gambar satelit seri waktu di tiga lokasi studi (Tahermanesh *et al.*, 2025). Perbandingan di atas menunjukkan bahwa penelitian ini melengkapi kekurangan sebelumnya dengan menggabungkan ekstraksi fitur *GoogLeNet* dan PCA-*Clustering* untuk analisis *unsupervised*.

### 2.3 Integrasi Al-Qur'an dan Hadis

Dalam perspektif Islam, bencana alam bukan sekadar fenomena fisik, tetapi juga memiliki dimensi spiritual dan moral yang mengandung pelajaran bagi manusia. Al-Qur'an menjelaskan bahwa kejadian bencana sering kali menjadi ujian

(*ibtilā'*), peringatan (*tadzkiroh*), atau akibat dari perbuatan manusia (*fasād*) di muka bumi. Dalam Surah Ar-Rum ayat 41 disebutkan:

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

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

Selain itu, Al-Qur'an juga menegaskan bahwa jangan berbuat kerusakan di bumi, sebagaimana firman Allah dalam surat Al-Baqarah ayat 11, ayat 205 dan Al-A'raf ayat 56:

وَإِذَا قِيلَ لَهُمْ لَا تُفْسِدُوا فِي الْأَرْضِ قَالُوا إِنَّا هُنَّ مُصْلِحُونَ ﴿١١﴾

Artinya, "Apabila dikatakan kepada mereka, 'Janganlah berbuat kerusakan di bumi,' mereka menjawab, 'Sesungguhnya kami hanyalah orang-orang yang melakukan perbaikan'." (QS. Al-Baqarah:11)

Selanjutnya, ayat dan Hadis pengelompokan manusia menjadi suku/bangsa (*lita'ārafū*) bertujuan untuk saling kenal, namun standar kemuliaan sejati di sisi Tuhan bukanlah suku atau ras, melainkan ketakwaan (*taqwa*). Dalam surat Al-Hujurat Ayat 13 dan Hadis Riwayat Ahmad dan At-Tirmidzi.

يَا أَيُّهَا النَّاسُ إِنَّا خَلَقْنَاكُمْ مِّنْ دُرْجَاتٍ وَجَعَلْنَاكُمْ شُعُوبًا وَقَبَائِلَ لِتَعَارِفُوا إِنَّ أَكْرَمَكُمْ عِنْدَ اللَّهِ أَنْفَكُمْ لَيْسَ اللَّهُ عَلَيْهِ حِلٌّ

Artinya: "Wahai manusia! Sungguh, Kami telah menciptakan kamu dari seorang laki-laki dan seorang perempuan, kemudian Kami jadikan kamu berbangsa-bangsa dan bersuku-suku agar kamu saling mengenal. Sesungguhnya yang paling mulia di antara kamu di sisi Allah ialah orang yang paling bertakwa. Sungguh, Allah Maha Mengetahui, Mahateliti". (Q.S. Al-Hujurat Ayat 13)

يَا أَيُّهَا النَّاسُ، أَلَا إِنَّ رَبَّكُمْ وَاحِدٌ، وَإِنَّ أَبَانِكُمْ وَاحِدٌ، أَلَا لَا فَضْلَ لِعَرَبٍ عَلَى أَغْجَمِيٍّ، ذَلِكَ لِعَجْمَيٍّ عَلَى عَرَبٍ، وَلَا لِأَجْمَرٍ عَلَى أَسْوَدٍ، وَلَا لِأَسْوَدٍ عَلَى أَجْمَرٍ، إِلَّا بِالْتَّغْوِيٍّ.

Artinya: "*Wahai manusia! Ketahuilah, sesungguhnya Tuhan kalian itu satu, dan sesungguhnya bapak kalian itu satu (yakni Adam). Ketahuilah, tidak ada kelebihan bagi orang Arab atas orang non-Arab ('Ajam), dan tidak ada kelebihan bagi orang non-Arab atas orang Arab; tidak ada kelebihan bagi orang berkulit merah atas orang berkulit hitam, dan tidak ada kelebihan bagi orang berkulit hitam atas orang berkulit merah, melainkan dengan takwa.*" (HR. Ahmad dan At-Tirmidzi).

Dengan ayat dan hadis tersebut menegaskan bahwa semua manusia berasal dari satu asal dan diciptakan berbeda-beda bukan untuk saling membedakan derajat, melainkan agar saling mengenal dan bekerja sama. Perbedaan suku, bangsa, atau warna kulit hanyalah bentuk pengelompokan alami, bukan ukuran kemuliaan. Satu-satunya yang membedakan manusia di sisi Allah hanyalah ketakwaan. Dengan demikian, pengelompokan hanya berfungsi untuk mengenali keberagaman, bukan menentukan nilai atau derajat manusia.

#### **2.4 Karakteristik Kerusakan Bangunan**

Karakteristik kerusakan bangunan merujuk pada sifat, pola, dan bentuk kerusakan yang muncul pada suatu struktur bangunan akibat pengaruh eksternal, seperti bencana alam, kegagalan konstruksi, atau faktor lingkungan. Karakteristik ini mencakup kondisi fisik yang dapat diamati pada elemen bangunan, seperti retakan pada dinding, deformasi struktural, keruntuhan sebagian, hingga kerusakan total. Setiap jenis kerusakan memberikan indikasi mengenai tingkat keparahan dan penyebab utama yang memengaruhinya.

Secara umum, karakteristik kerusakan bangunan dapat dikelompokkan berdasarkan tingkat keparahan, meliputi kerusakan ringan, sedang, dan berat.

Kerusakan ringan biasanya berupa keretakan kecil atau kerusakan non-struktural yang tidak memengaruhi kestabilan bangunan. Kerusakan sedang mencakup perubahan bentuk atau kegagalan sebagian komponen struktural yang memerlukan perbaikan signifikan. Sementara itu, kerusakan berat menggambarkan kondisi ketika elemen struktural utama gagal berfungsi sehingga bangunan berpotensi runtuh atau tidak dapat dihuni. Kluster ini digunakan untuk menilai keamanan bangunan dan menentukan strategi rehabilitasi.

Pemahaman mengenai karakteristik kerusakan bangunan sangat penting dalam proses evaluasi pascabencana, perencanaan rekonstruksi, dan pengembangan sistem identifikasi otomatis berbasis *citra*. Dengan mengenali pola visual dan sifat kerusakan secara tepat, para ahli dapat memperkirakan penyebab kerusakan, menghitung risiko lanjutan, serta menetapkan prioritas penanganan. Selain itu, karakterisasi yang akurat membantu lembaga penanggulangan bencana dalam menyusun kebijakan respons yang efektif dan mengoptimalkan alokasi sumber daya.

## 2.5 Bencana Alam

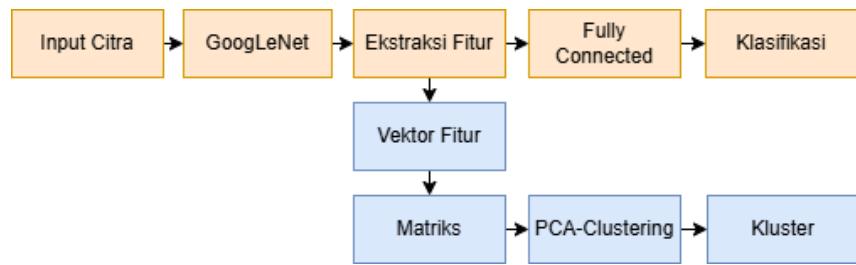
Bencana alam merupakan fenomena yang dapat menyebabkan gangguan berskala luas dan sering kali mengakibatkan kerugian finansial serta sosial yang substansial secara global. Indonesia, sebagai Negara yang rawan bencana, menghadapi tantangan besar dalam mitigasi dampak tersebut (Coronese et al., 2019).

Salah satu aspek krusial pascabencana adalah penilaian kerusakan bangunan. Secara tradisional, proses ini masih sangat bergantung pada metode konvesional yang mengandalkan penilaian subjektif ahli di lapangan. Pendekatan ini memiliki kelemahan signifikan, yakni memakan waktu lama dan membutuhkan sumber daya manusia yang besar. Selain itu, metode konvesional rentan terhadap ketidakakuratan data. Tim *Disaster Management Planning and Control Team* (P3B) sering menghadapi inkonsistensi dalam kategorisasi kerusakan selama survei lapangan, yang menyebabkan perbedaan antara data yang diterima oleh Badan Penanggulangan Bencana Daerah (BPBD) dan kondisi sebenarnya. Ketidaksesuaian data ini pada akhirnya menghambat efektivitas upaya rehabilitasi dan rekonstruksi, sehingga menyoroti kebutuhan mendesak akan pendekatan penilaian yang lebih efisien dan akurat.

## BAB III

### DESAIN DAN INTEGRASI SISTEM PENELITIAN

#### 3.1 Kerangka Konsep Penelitian



Gambar 3.1 Kerangka Konsep

Penelitian sebelumnya ditunjukkan pada warna merah muda, alur analisis dimulai dari Input *citra* yang diproses oleh *GoogLeNet* untuk melakukan ekstraksi fitur, diikuti oleh lapisan *fully connected* dan menghasilkan klasifikasi langsung berupa label kerusakan. Namun, dalam penelitian ini dengan warna biru memiliki alur dimodifikasi setelah ekstraksi fitur oleh *GoogLeNet*. Vektor fitur yang dihasilkan tidak langsung dikelompokkan, melainkan diubah menjadi Matriks, kemudian direkayasa fitur menggunakan GLCM dan diproses lebih lanjut dengan *PCA-Clustering* untuk mengelompokkan tingkat kerusakan secara *unsupervised* berdasarkan karakteristik visual menghasilkan kluster yang lebih interpretatif untuk analisis *spasial* kerusakan bangunan. Pendekatan ini mengatasi keterbatasan klasifikasi tradisional yang bergantung pada data berlabel.

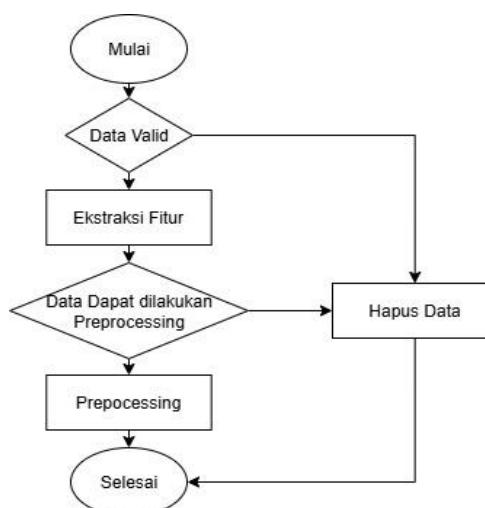
#### 3.2 Persiapan Data

Langkah selanjutnya adalah persiapan data, di mana data yang digunakan berupa *citra* bangunan pascabencana yang diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) Malang. *Citra* tersebut akan melalui proses

ekstraksi fitur menggunakan *GoogLeNet* untuk mengidentifikasi karakteristik visual kerusakan, kemudian dilakukan reduksi dimensi dengan PCA dan pengelompokan tingkat kerusakan menggunakan teknik *clustering*. Pengumpulan data ini memastikan ketersediaan dataset yang representatif dan beragam, sehingga mendukung analisis komprehensif tanpa bergantung pada pengumpulan lapangan yang memakan waktu.

### 3.3 Validasi Data

Setelah pengumpulan data, semua *citra* kerusakan bangunan akan melalui proses validasi untuk memastikan kualitas dan kesesuaianya dengan kriteria analisis, terutama *citra* yang digunakan dalam tahap ekstraksi fitur. Proses ini meliputi pemeriksaan resolusi, sudut pengambilan gambar, serta klarifikasi tingkat kerusakan dengan data lapangan untuk memastikan validitas dataset. Validasi data bertujuan untuk menghindari bias dalam ekstraksi fitur *GoogLeNet* dan pengelompokan PCA. Proses validasi data tersebut dapat dilihat pada gambar 3.2



Gambar 3.2 Validasi Data

Proses validasi data diawali dengan pengumpulan *citra* kerusakan bangunan dari sumber data Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) Malang. Setiap *citra* kemudian diperiksa kelayakannya berdasarkan kriteria seperti kualitas visual, resolusi, kejelasan objek, dan kelengkapan metadata. *Citra* yang memenuhi syarat akan menjalankan ekstraksi fitur awal menggunakan *GoogLeNet* untuk menguji kemampuannya dalam mengidentifikasi karakteristik kerusakan. Tahap ini penting untuk memastikan bahwa data yang digunakan benar-benar valid dan siap diproses lebih lanjut.

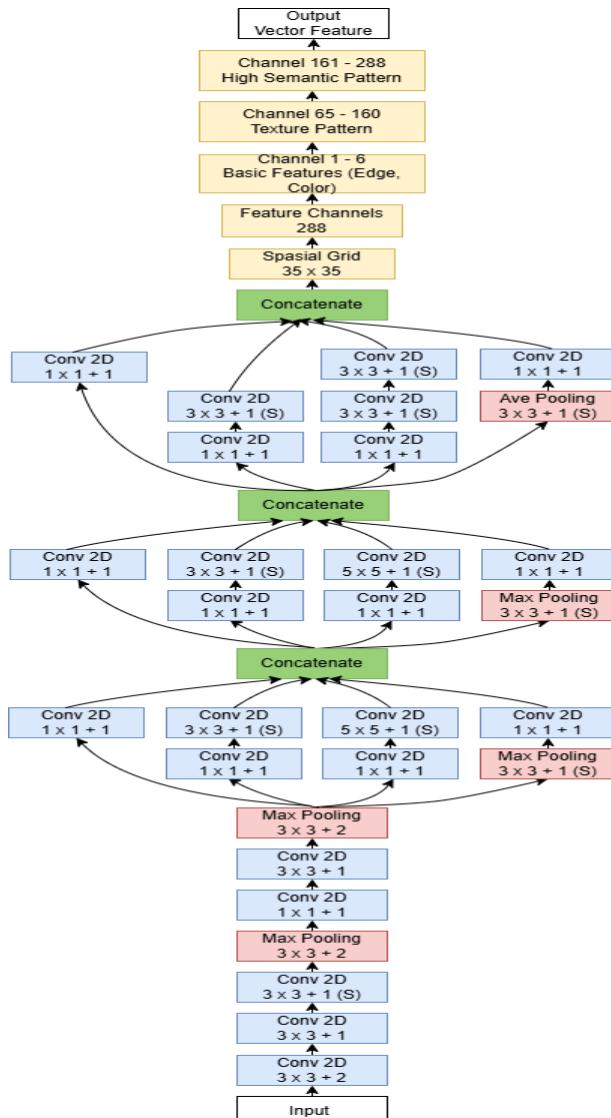
Setelah melalui ekstraksi fitur awal, peneliti menentukan apakah *citra* dapat diproses lebih lanjut atau memerlukan perbaikan. *Citra* yang layak akan memasuki tahap *preprocessing* untuk meningkatkan kualitasnya, sementara *citra* yang tidak memenuhi standar setelah perbaikan akan dihapus dari dataset. Proses penyaringan ini bertujuan untuk menjaga konsistensi dan akurasi data sebelum masuk ke tahap analisis utama.

Pada tahap akhir, *citra* yang telah tervalidasi menjalani *preprocessing* final. Hasilnya adalah dataset *citra* kerusakan bangunan yang siap untuk di ekstraksi fitur mendalam menggunakan *GoogLeNet* dan analisis *Clustering* dengan PCA. Proses Validasi yang ketat ini memastikan bahwa hasil penelitian dapat diandalkan dan menggambarkan kondisi kerusakan bangunan secara akurat, baik untuk keperluan analisis maupun visualisasi.

### 3.4 *GoogLeNet* (Ekstraksi Fitur)

*GoogLeNet* adalah salah satu arsitektur *deep Learning* (pembelajaran mendalam) dalam bentuk *Convolutional Neural Network* (CNN) yang digunakan

untuk ekstraksi fitur semantik. Dalam konteks penilaian kerusakan, *GoogLeNet* yang sudah terlatih (*pre-trained*) digunakan untuk menghasilkan vector fitur berdimensi tinggi dari *citra* bangunan yang rusak. Kemampuan *deep learning* ini memungkinkan analisis mendalam untuk mengidentifikasi pola-pola kompleks dalam *citra* (Anbarasan *et al.*, 2020). Inti proses metode ekstraksi fitur menggunakan *GoogLeNet* yaitu input *citra* (RGB) yang akan diubah ukurannya. *Citra* kemudian melewati serangkaian lapisan operasional konvolusional yang berfungsi untuk mendeteksi serangkaian fitur-fitur tingkat rendah, seperti tepi, sudut, atau tekstur dasar. Inti dari arsitektur *GoogLeNet* adalah *Inception Module*. Modul ini dirancang untuk memproses fitur pada berbagai skala paralel dalam satu blok. Hal ini dicapai dengan menggunakan operasi konvolusional dengan ukuran filter yang berbeda-beda serta *pooling layer*. Untuk mengurangi beban komputasi dan jumlah parameter, modul ini memanfaatkan konvolusi 1x1 sebagai *Bottleneck* sebelum menjalankan konvolusi 3x3 dan 5x5 yang lebih mahal. Strategi ini memungkinkan jaringan menjadi lebih dalam dan luas tanpa mengorbankan performa. Hasil akhir adalah vector fitur berdimensi tinggi yang mana diekstraksi dari *layer* ‘*mixed2*’ yang menangkap representasi hierarkis yaitu pola dasar hingga semantik kompleks dari *citra*. Seperti pada Gambar 3.8:



Gambar 3.3 Arsitektur *GoogLeNet (Inception-v3)*

Modul *Inception* memproses fitur pada beberapa skala secara paralel sehingga memungkinkan ekstraksi pola tekstur (retakan, robekan, deformasi) pada berbagai ukuran. Inti dari arsitektur ini adalah Modul *Inception*, yang mana beberapa operasi konvolusi dengan ukuran filter yang berbeda, pada akhir jaringan digunakan *global average pooling* untuk mereduksi peta fitur menjadi satu vektor fitur per *citra* (vektor tunggal).

### 3.5 Principal Component Analysis (PCA)-Clustering

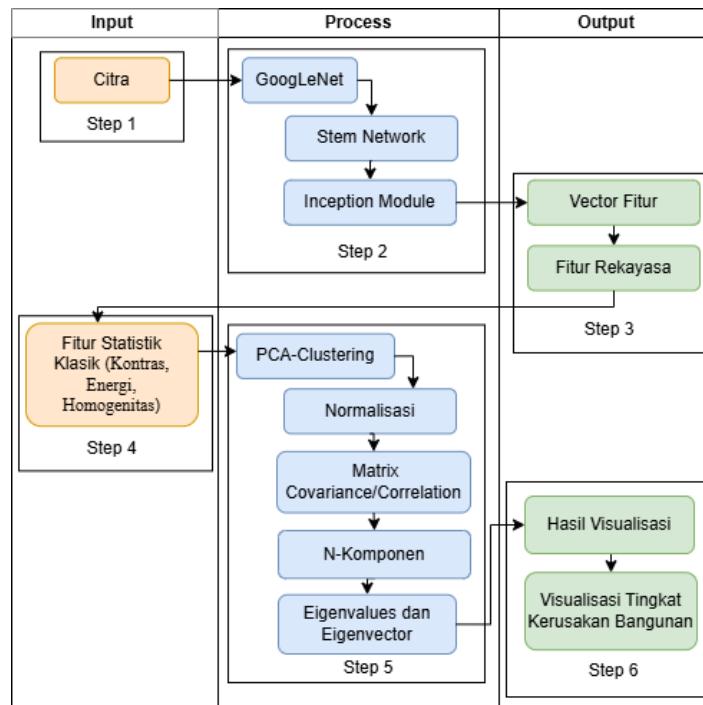
*Principal Component Analysis* (PCA) adalah teknik statistic fundamental yang digunakan untuk pengurangan dimensi secara *unsupervised* (tanpa pengawasan). Tujuannya adalah mentransformasi sekumpulan variable yang mungkin saling berkorelasi (seperti vector fitur berdimensi tinggi dari *GoogLeNet*), menjadi sekumpulan variable baru yang tidak berkorelasi, yang disebut *principal component* (komponen utama). Dengan penelitian ini, PCA tidak hanya digunakan untuk mengurangi dimensi saja tetapi juga sebagai dasar *clustering* (pengelompokan) data. Pendekatan ini memungkinkan identifikasi klaster kerusakan secara objektif dan *data-driven* langsung dari data gambar, tanpa memerlukan pelabelan manual sebelumnya (Almais et al., 2023).

Dengan proses metode PCA-*clustering*, vektor fitur yang di ekstraksi dari *GoogLeNet* dinormalisasi menggunakan *StandartScaler*. Langkah ini penting untuk menstandarkan data dan memastikan semua fitur memiliki skala yang seimbang sehingga tidak ada fitur yang mendominasi analisis PCA. Selanjutnya, dihitung matriks *kovarians/korelasi* dari data yang telah dinormalisasi. Dari matriks kovarians, dilakukan perhitungan *eigenvalues* dan *eigenvectors*. *Eigenvalues* mempresentasikan besaran varians data yang ditangkap oleh setiap *eigenvector* (komponen utama). PCA mengurangi data dengan memproyeksikannya ke komponen utama yang paling signifikan. Dalam studi ini, fokus diberikan pada dua komponen utama pertama (PC1 dan PC2) karena keduanya secara kolektif mampu menangkap 98,07% dari tota varians data, sehingga dianggap representasi yang kuat. Pengelompokan (*Clustering*) dilakukan dengan menerapkan ambang batas

(*threshold*) yang telah ditentukan pada sumbu komponen utama pertama (PC1). Berdasarkan penelitian Almais (2023) dari nilai koordinat PC1 ( $n$ ), data dikategorikan ke dalam tiga tingkat kerusakan: "rusak ringan" ( $n < 0$ ), "rusak sedang" ( $0 < n < 2$ ), dan "rusak berat" ( $n > 2$ ).

### 3.6 Desain Sistem

Untuk mencapai tujuan penelitian ini, sebuah alur telah menggambarkan tahapan analisis pada gambar 3.4:



Gambar 3.4 Desain Sistem

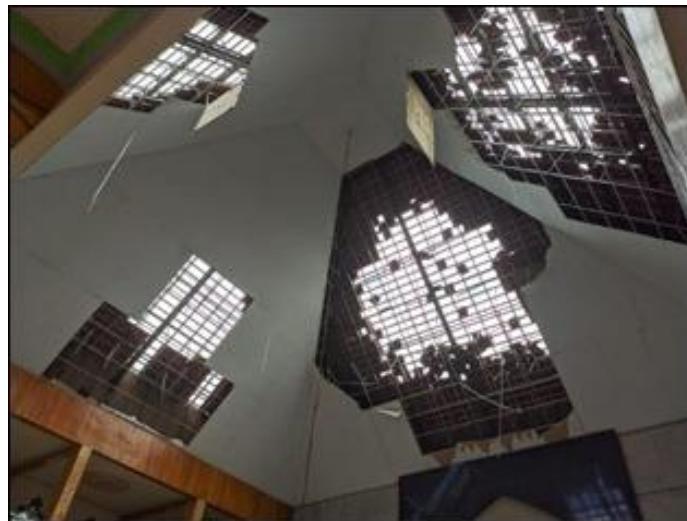
Memperlihatkan *pipeline end-to-end* yang digunakan dalam penelitian ini terlihat setiap langkah modular (*preprocessing, feature extraction, PCA-clustering, dan validasi*) dirancang agar dapat diotomasi dan diskalakan. Pada implementasi, output fitur *GoogLeNet* diambil dari lapisan [*Mixed'2*], kemudian distandarisasi

sebelum masukkan ke PCA. Langkah validasi melibatkan metriks eksternal (penelitian terdahulu). Berikut adalah langkah-langkah menilai tingkat kerusakan bangunan pascabencana alam dalam integrasi ekstraksi fitur *GoogLeNet* dan PCA-*Clustering* secara lebih rinci:

### 3.5.1 Langkah Pertama (*Input Gambar*)

#### 1) Input Gambar

Pada proses ini akan mengumpulkan gambar, seperti pada Gambar 3.5 yang mana berupa gambar bangunan pascabencana alam dari dataset Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) Malang, yang berbentuk dalam format gambar berwarna (RGB) (Cai *et al.*, 2025).



Gambar 3.5 Input *Citra* dengan Format RGB

### 3.5.2 Langkah Kedua Proses *GoogLeNet*

#### 1) *Stem Network*

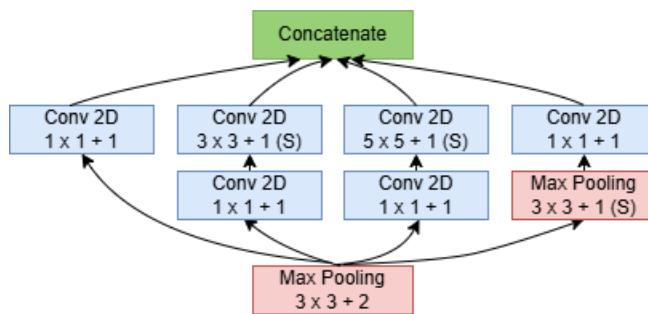
Tahap awal proses *GoogLeNet* yaitu *Resize* (Ubah Ukuran) yang merupakan model pembelajaran mesin yang biasanya memerlukan input dengan ukuran tetap. Oleh karena itu, pada semua gambar mengganti ukurannya ke dimensi standar

menggunakan persamaan 1 (Shen *et al.*, 2024).

$$I((x', y')) = \sum_{i=0}^1 \sum_{j=0}^1 w(i, j) x I(x + i, y + j) \quad (3.1)$$

## 2) Inception Module

Citra akan melewati serangkaian lapisan untuk merubah menjadi representasi fitur (Song *et al.*, 2024). Lapisan ini melakukan operasi konvolusi dengan filter dimana jaringan mulai mendekksi fitur-fitur tingkat rendah seperti tepi, sudut, atau tekstur dasar. Dengan mengindikasikan 10 filter yang berbeda, yang masing-masing belajar mendekksi pola yang berbeda. Lapisan *pooling* (*max pooling*) yang akan mengurangi dimensi *spasial* dan membantu membuat fitur menjadi lebih invariant terhadap translasi kecil dan mengurangi komputasi yang akan melakukan operasi *pooling* secara paralel (Tosta *et al.*, 2019) Pendekatan ini memungkinkan jaringan untuk menangkap fitur dalam berbagai ukuran secara bersamaan dalam satu blok seperti pada gambar 3.6:



Gambar 3.6 Inception (V3) Module Dengan Ekstraksi Fitur

Dengan proses konvolusi  $1 \times 1$  yang berfungsi sebagai *bottleneck* untuk mengurangi jumlah *channel* sebelum operasi konvolusi yang lebih berat, sehingga menurunkan jumlah parameter tanpa kehilangan kemampuan representasi fitur. Melakukan proses tersebut terdapat pada modul *inception* sebelum konvolusi

matriks  $3 \times 3$  dan  $5 \times 5$  (Gopinathan, 2021). Tujuannya untuk meningkatkan kedalaman analisis, memperluas jaringan, dan mengurangi dimensi modul. Secara umum, *GoogLeNet* menggunakan lapisan konvolusi dan *pooling* untuk mengekstraksi data. Modul *inception* dapat menurunkan beban komputasi, meningkatkan kedalaman data, dan pada bagian akhir lapisan penuh terhubung, untuk menyimpan hasil dari proses lapisan *pooling*. Untuk mengurangi jumlah parameter tanpa mengurangi akurasi model, hasil *GoogLeNet* juga menyederhanakan dengan menggunakan metode global *average pooling* yang sederhana (Gliner *et al.*, 2021).

### 3.5.3 Langkah Ketiga Hasil *GoogLeNet*

#### 1) Rekayasa Fitur Melalui Vektor Fitur Dimensi Tinggi

Fitur yang efektif sangat penting pada ekstraksi fitur untuk meningkatkan akurasi penilaian kondisi gambar. Salah satu metode baru untuk mengekstrak fitur dua dimensi, yang dapat menjelaskan kondisi gambar dengan lebih sama (Xie *et al.*, 2024). Output dari jaringan *GoogLeNet* adalah sebuah vektor fitur berdimensi tinggi yang padat dengan informasi penting dari *citra* asli (Song *et al.*, 2024). Dengan menggunakan fitur metode *Gray-Level Co-occurrence Matrix* (GLCM) menganalisis hubungan spasial antar piksel pada *citra* untuk mengkuantifikasi teksturnya. Ini adalah pendekatan rekayasa fitur (*feature engineering*) di mana ciri-ciri spesifik dihitung secara matematis. Fitur Kontras, Energi, dan Homogenitas adalah fitur tekstur statistik klasik yang umumnya diekstraksi menggunakan metode GLCM (Han *et al.*, 2025).

### 3.5.4 Langkah Keempat Input Hasil *GoogLeNet*

### 1) Fitur Statik Klasik (Kontras, Energi, Homogenitas)

Ekstraksi fitur *GoogLeNet* menggunakan fitur *channel* yang menyerupai GLCM (*Gray Level Co-occurrence Matrix*) untuk menangkap informasi tekstur dan pola *spasial* yang tidak hanya bergantung pada warna atau bentuk, tetapi juga hubungan antar piksel di dalam *citra*. *GoogLeNet* menghasilkan banyak *feature map* (*channel*) yang masing-masing mewakili respon terhadap pola visual tertentu seperti tepi, tekstur, atau area kerusakan. Namun, nilai mentah dari *feature map* masih berupa representasi kompleks. Dengan menghitung fitur tekstur GLCM seperti kontras, energi, dan homogenitas dari setiap *channel*, dapat mengubah keluaran *GoogLeNet* menjadi fitur statistik yang lebih bermakna dan mudah diinterpretasikan secara kuantitatif.

Selanjutnya hasil dari *Feature Engineering* menjadi sebuah variabel input untuk proses pengelompokan menggunakan PCA-*Clustering*. Fitur-fitur ini menangkap pola, bentuk, dan tekstur pada level yang lebih tinggi. Proses tersebut melalui *Convert to Spreadsheet* yaitu mengubah data dari format lain ke dalam bentuk *spreadsheet*, seperti *Excel*, biasanya untuk memudahkan analisis data, terutama dalam penelitian yang melibatkan pengumpulan dan pemrosesan data dalam jumlah besar (Brennan 2016). Energi dihitung menggunakan Persamaan 3.2, Kontras dihitung menggunakan Persamaan 3.3, dan Homogenitas dihitung menggunakan Persamaan 3.4:

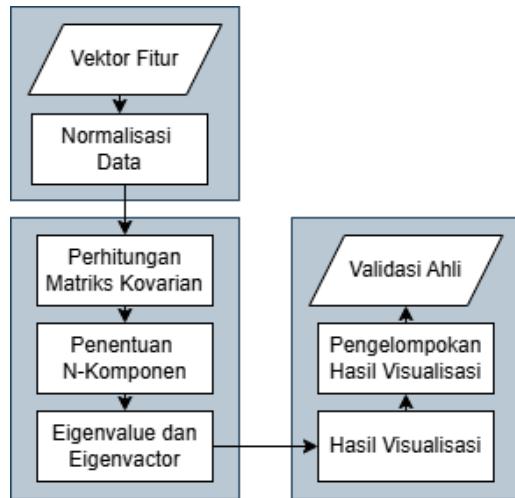
$$\text{Energy} = \sum_{i,j} p(i,j)^2 \quad (3.2)$$

$$\text{Contrast} = \sum_{i,j} |i - j|^2 p(i,j) \quad (3.3)$$

$$Homogeneity = \sum_{i,j} \frac{p(i,j)}{1 + |i - j|} \quad (3.4)$$

### 3.5.5 Langkah Kelima Proses PCA-Clustering

Proses PCA-Clustering mengubah data mentah menjadi informasi terkelompok berdasarkan tingkat kerusakan dengan bantuan PCA, sehingga hasil analisis menjadi lebih terstruktur, efisien. Seperti pada Gambar 3.7 Desain sistem PCA-Clustering



Gambar 3.7 Desain Sistem PCA-Clustering

#### 1) Normalisasi Dataset

PCA adalah teknik statistik dasar yang mengubah sekumpulan variabel yang berpotensi berkorelasi menjadi sekumpulan variabel baru yang tidak berkorelasi yang disebut komponen utama (Jolliffe, 1998). Dalam studi ini, vektor fitur dinormalisasi menggunakan *StandardScaler* Persamaan 3.5 dalam *Python* untuk menstandarkan data input, memastikan semua fitur berada pada skala yang seimbang (Dinc *et al.*, 2014)

$$x_{standart} = \frac{x - mean(x)}{standart deviation(x)} \quad (3.5)$$

Simbol  $\bar{X}$  (*x bar*) mewakili nilai rata-rata dari himpunan  $X$  yang dihitung menggunakan Persamaan 3.6:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad (3.6)$$

## 2) Matriks Kovarians/Korelasi

Proses selanjutnya melibatkan perhitungan matriks *kovarians* (A *et al.*, 2024) menggunakan persamaan 3.7 (dengan  $X$  adalah matriks data yang sudah distandarisasi), hingga memproyeksikan data ke komponen utama yang paling signifikan untuk mendapatkan representasi data dalam dimensi yang lebih rendah (Nolan *et al.*, 2025), salah satunya yang menerapkannya untuk meningkatkan efisiensi sistem temu kembali *citra* (Jahandad *et al.*, 2019).

$$\sum = \frac{1}{N-1} X^T X \quad (3.7)$$

## 3) N-Komponen

Pada tahap ini tidak harus mengambil semua PC. Dalam mereduksi dimensi data, PCA beroperasi dengan mencari sebuah *matriks* proyeksi yang bertujuan untuk meminimalkan nilai *Mean Squared Error* (MSE) atau Kesalahan Kuadrat Rata-Rata antara data hasil proyeksi dan data aslinya (Wang *et al.*, 2018). Karena efisiensinya dalam menangkap varians data. Jadi pemilihan N-Komponen utama berfokus pada PC1, PC2. Pemilihan Komponen yang tepat dengan menggunakan Variabel Ratio, dengan alasan dua Komponen tersebut kuat sebagai analisis.

## 4) *Eigenvalues/Eigenvector*

*Eigenvalues* merepresentasikan skalar yang mengukur seberapa besar variansi data yang dihasilkan oleh setiap *eigenvector* dalam analisis matriks. Sementara itu, rasio variansi merupakan suatu ukuran yang menggambarkan persentase dari

total variansi data yang mampu dijelaskan oleh masing-masing komponen utama tersebut, sehingga dapat menentukan sebaran dan tingkat pentingnya setiap komponen dalam meringkas informasi data (Guilherme Dean Pelegrina, 2024).

### **3.5.6 Langkah Keenam *Output* Visualisasi**

#### 1) Hasil Visualisasi

Data yang telah dikurangi dimensinya divisualisasikan menggunakan diagram pencar. Setiap titik mewakili suatu *instance* data berdasarkan nilai PC1 dan PC2. Titik-titik yang secara visual berdekatan menunjukkan gambar dengan fitur serupa, membentuk kluster. Identifikasi kelompok-kelompok ini merupakan inti dari analisis kluster (Jain *et al.*, 1999).

#### 2) Visualisasi Tingkat Kerusakan Bangunan

Dengan membuat label dari penelitian Almais *et al*, telah melakukan penelitian tentang nilai koordinat (n) pada grafik PCA (Almais *et al.*, 2024). Maka dapat diinterpretasikan hasilnya dalam tabel 3.1 sebagai berikut:

- Jika nilai koordinat PC1 (n) berada dalam rentang ( $n < 0$ ), maka label hasil data PC1 rusak ringan.
- Jika nilai koordinat PC1 (n) berada dalam rentang ( $0 \leq n < 2$ ), maka label hasil data PC2 rusak sedang.
- Jika nilai koordinat PC1 (n) berada dalam rentang ( $n \geq 2$ ), maka label hasil data PC1 rusak Berat.

Tabel 3.1 Titik Koordinat (n) tingkat kerusakan (Almais *et al.*, 2023)

<b>Titik Koordinat (n)</b>	<b>Tingkat Kerusakan</b>
$n < 0$	Rusak Ringan
$0 \leq n < 2$	Rusak Sedang
$n \geq 2$	Rusak Berat

### 3.7 Skenario Uji Coba

Untuk memastikan keakuratan dan keandalan hasil pengelompokan tingkat kerusakan bangunan yang dihasilkan oleh metode *GoogLeNet PCA-Clustering*, dilakukan uji validasi eksternal dengan membandingkan hasil model terhadap data referensi dari penelitian terdahulu serta penilaian langsung oleh ahli lapangan.

#### 6.1.1 Validasi Penelitian Terdahulu

Validasi eksternal dilakukan dengan mengambil sebagian *citra* yang telah digunakan dalam penelitian (Almais et al., 2023), yang sebelumnya telah memiliki kategori kerusakan terverifikasi berdasarkan pengamatan lapangan. Hasil klasterisasi dari model penelitian ini dibandingkan dengan label rujukan tersebut untuk menilai konsistensi pola pengelompokan. Proses ini bertujuan untuk mengidentifikasi sejauh mana sistem berbasis ekstraksi fitur *GoogLeNet* mampu menghasilkan pembagian tingkat kerusakan yang selaras dengan hasil penelitian empiris terdahulu. Tingkat kesesuaian antara hasil model dan data acuan dianalisis secara deskriptif, dengan fokus pada proporsi kesamaan kluster untuk tiap kategori kerusakan (ringan, sedang, dan berat).

#### 6.1.2 Validasi Ahli Lapangan

Selain itu, dilakukan pula validasi oleh seorang ahli lapangan yang memiliki pengalaman dalam bidang penilaian struktur bangunan pascabencana. Ahli tersebut diminta untuk melakukan penilaian visual terhadap sejumlah *citra* hasil pengelompokan *PCA-Clustering*, kemudian menilai apakah kategori kerusakan yang dihasilkan model sesuai dengan kondisi aktual pada *citra*.

Melalui dua bentuk validasi ini, dapat disimpulkan bahwa hasil pengelompokan

metode *GoogLeNet PCA-Clustering* memiliki tingkat konsistensi tinggi terhadap data lapangan dan validasi empiris penelitian sebelumnya, sehingga dapat dipercaya sebagai pendekatan awal dalam sistem analisis otomatis untuk penilaian tingkat kerusakan bangunan pascabencana.

### 3.8 Instrument Penelitian

Instrument penelitian ini menggunakan vektor fitur sebagai variabel independen yang diperoleh dari ekstraksi *citra* kerusakan bangunan, dengan tingkat kerusakan pascabencana alam sebagai variabel *intervening* yang menghubungkan proses analisis. Proses utama melibatkan metode *GoogLeNet* untuk ekstraksi fitur dan *PCA-Clustering* untuk pengelompokan, menghasilkan variabel dependen berupa kluster tingkat kerusakan. Fokus penelitian terletak pada *citra* sebagai variabel independen, yang memengaruhi pembentukan kluster sebagai variabel terikat, sementara tingkat kerusakan berperan sebagai faktor penghubung dalam alur analisis pada Tabel 3.2.

Tabel 3.2 Instrument Penelitian

Variabel Independent	Variabel Intervening	Main Proses	Variabel Dependent
Citra	Ekstraksi Fitur GoogLeNet	PCA-Clustering	Kluster tingkat kerusakan bangunan

## **BAB IV**

### **IMPLEMENTASI METODE**

#### **4.1 Implementasi *GoogLeNet***

Proses implementasi metode ekstraksi fitur menggunakan arsitektur *GoogLeNet* (*Inception v3*) yang diterapkan untuk menganalisis tingkat kerusakan bangunan pascabencana alam. Implementasi dilakukan berdasarkan hasil pengkodean pada lingkungan *python* dengan tahapan implementasi yang mencakup proses input *citra*, pra-pemrosesan data, ekstraksi fitur melalui *GoogLeNet*, pembentukan vektor fitur berdimensi tinggi, normalisasi data, serta persiapan hasil untuk tahap reduksi dimensi dan klasterisasi menggunakan PCA-*Clustering*.

#### **4.2 Input *Citra***

Data input yang digunakan berupa *citra* RGB bengunan pascabencana alam yang dikumpulkan dari sumber data Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) Malang. Setiap *citra* mewakili kondisi bangunan dalam berbagai tingkat kerusakan, meliputi sedang, ringan dan berat.

Sebelum digunakan, semua *citra* diperiksa kelayakannya berdasarkan resolusi, kualitas, visual dan kejelasan objek agar hasil ekstraksi fitur tidak terganggu oleh gangguan noise atau blur. Proses pembacaan *citra* dilakukan menggunakan fungsi cv2.imread() yang menghasilkan representasi numerik dalam bentuk *matriks* intensitas piksel tiga kanal (R,G,B). Untuk menyesuaikan dengan arsitektur *GoogLeNet*, semua *citra* diubah ukurannya. Ukuran ini merupakan standar input bagi pra-latih *GoogLeNet* agar dapat melakukan operasi konvolusi

dengan ukuran filter yang sesuai. Pada proses ini akan menginput gambar yang merupakan ukuran input standar untuk model *InceptionV3*, kemudian mengubah *citra* menjadi *array numerik* dan melakukan pra-pemrosesan (*preprocessing*) agar sesuai dengan format data yang digunakan dalam pelatihan model *ImageNet*. Selanjutnya, model *InceptionV3* dipanggil tanpa lapisan klasifikasi akhir (`include_top=False`) agar berfungsi sebagai *feature extractor* yang hanya menghasilkan representasi fitur dari *citra*, bukan hasil klasifikasi. Pada tahap berikutnya, dipilih salah satu lapisan konvolusional, seperti “*mixed2*”, untuk menentukan tingkat abstraksi fitur yang diambil. Lapisan ini berfungsi mengekstraksi pola visual seperti tepi, bentuk, tekstur, dan struktur bangunan yang relevan terhadap kerusakan. Terakhir, dibuat model turunan baru yang menghasilkan keluaran berupa *feature map* dari lapisan terpilih, sehingga dapat digunakan sebagai input dalam analisis lanjutan, seperti reduksi dimensi dengan PCA atau proses klasterisasi tingkat kerusakan bangunan.

### 4.3 Proses Ekstraksi Fitur Menggunakan *GoogLeNet*

Arsitektur *GoogLeNet* (*Inception-V3*) merupakan model *Convolutional Neural Network* (CNN) yang dirancang untuk menangkap fitur-fitur visual pada berbagai skala melalui kombinasi operasi konvolusi 1x1, 3x3 5x5, dan *max pooling* yang dilakukan secara paralel dalam satu blok yang disebut *Inception Module*.

Ekstraksi dan analisis fitur tekstur dari hasil keluaran model *GoogLeNet* (*InceptionV3*). Proses diawali dengan menghasilkan *feature map* dari *citra* masukan menggunakan model *GoogLeNet* yang telah dilatih sebelumnya, kemudian hasilnya disimpan dalam variabel `features_goolenet`. Selanjutnya, didefinisikan fungsi

`calculate_texture_properties()` untuk menghitung tiga parameter tekstur utama yang menyerupai metode *Gray Level Co-occurrence Matrix* (GLCM), yaitu kontras, energi, dan homogenitas. Nilai kontras dihitung berdasarkan variasi piksel antar area dengan semakin besar menunjukkan perbedaan intensitas yang kuat, energi merepresentasikan keteraturan pola tekstur, dan homogenitas menunjukkan keseragaman area dalam *citra*. Setelah fungsi tersebut dibuat, dilakukan perulangan pada setiap channel hasil ekstraksi *GoogLeNet* untuk menghitung nilai ketiga parameter tersebut pada masing-masing *feature map*. Hasil akhirnya berupa sekumpulan nilai statistik yang merepresentasikan karakteristik tekstur *citra* bangunan, yang kemudian ditampilkan dalam bentuk vektor fitur. Pendekatan ini berfungsi untuk mengubah keluaran kompleks *GoogLeNet* menjadi fitur statistik yang lebih terukur dan mudah dianalisis, sehingga dapat digunakan dalam tahap selanjutnya seperti reduksi dimensi PCA atau klasterisasi tingkat kerusakan bangunan.

#### 4.3.1 Pra-pemroses Data

*Citra* diubah ukurannya dan dinormalisasi ke rentang [0,1] menggunakan transformasi `transforms.Normalize(mean, std)` untuk menyesuaikan distribusi nilai intensitas dengan model pra-latih.

#### 4.3.2 Pemanggilan Model *GoogLeNet*

Model *GoogLeNet* diambil dari pustaka `torchvision.models` dengan parameter `pretrained=True`. Lapisan klasifikasi akhir dihapus untuk mempertahankan lapisan fitur menengah “*mixed2*” yang digunakan sebagai representasi semantic *citra*. Lapisan ini menghasilkan vektor fitur berdimensi tinggi.

### 4.3.3 Proses Forward Pass

Setiap *citra* dijalankan melalui jaringan *GoogLeNet* hingga lapisan *mixed2*, dan hasilnya disimpan sebagai vektor fitur numerik. Proses ini memungkinkan jaringan mendeteksi fitur hierarkis mulai dari tepi, tekstur, hingga pola semantik seperti retakan, deformasi, atau keruntuhan struktur bangunan.

### 4.3.4 Rekayasa Fitur Tambahan

Dengan melakukan rekayasa dan strukturalisasi data fitur tekstur yang diasumsikan telah diekstraksi sebelumnya dari *GoogLeNet* (*texture\_features\_googlenet*). Dengan mengimpor *library pandas*, pertama-tama menghitung jumlah *channel* dan mengubah bentuk (*reshape*) data fitur yang awalnya datar menjadi sebuah *matriks* dua dimensi, di mana setiap baris merepresentasikan satu *channel* dan tiga kolomnya sesuai dengan fitur 'Kontras', 'Energi', dan 'Homogenitas'. Data terstruktur ini kemudian dikonversi menjadi sebuah *Data Frame pandas*, dan sebuah kolom '*Channel*' disisipkan di awal untuk identifikasi sekuensial. Selanjutnya, membuat subset data yang hanya berisi 10 baris teratas (.head(10)) untuk keperluan analisis eksplorasi, selaras dengan metodologi penelitian yang berfokus pada fitur paling diskriminatif. Akhirnya, keseluruhan *DataFrame* (mencakup semua *channel*) diekspor ke dalam sebuah file *Excel* dengan nama *texture\_features\_googlenet.xlsx* untuk penyimpanan, sementara subset 10 baris teratas ditampilkan sebagai *output* untuk verifikasi langsung.

Setiap fitur mendalam (*deep learning*) dari *GoogLeNet*, penelitian ini juga mengekstraksi fitur statistic klasik menggunakan metode *Gray-Level Co-occurrence Matriks* (GLCM). Fitur-fitur yang dihasilkan meliputi Kontras,

Energi, dan Homogenitas, yang berfungsi memperkuat analisis tekstur permukaan bangunan. GLCM dihitung pada saluran fitur terpilih untuk memberikan konteks spasial antar piksel, dan hasilnya digunakan sebagai *input* tambahan untuk proses PCA. Dengan penjelasan for i in range(10): Melakukan iterasi sebanyak 10 kali untuk mengambil channel ke-0 hingga ke-9 dari hasil ekstraksi fitur *GoogLeNet*. Setiap *channel* mewakili filter konvolusi yang menangkap pola tertentu dari *citra* (tepi, tekstur, atau bayangan). *features\_goolenet[0, :, :, i]* untuk mengambil *feature map* dari gambar pertama (*index 0*) pada *channel* ke-i. Hasilnya adalah *matriks* dua dimensi yang merepresentasikan intensitas aktivasi di seluruh area *citra* untuk filter ke-i.

`calculate_texture_properties(channel_data_goolenet)` Fungsi menghitung tiga parameter tekstur utama berdasarkan *Gray-Level Co-occurrence Matrix* (GLCM), yaitu:

- 1) Kontras (*Contrast*): mengukur sejauh mana perbedaan intensitas antara piksel berdekatan. Nilai tinggi menandakan tekstur kasar atau banyak tepi tajam.
- 2) Energi (*Energy*): menunjukkan keteraturan pola. Nilai tinggi berarti tekstur halus atau berulang secara teratur.
- 3) Homogenitas (*Homogeneity*): mengukur keseragaman nilai piksel. Nilai tinggi menunjukkan area yang seragam tanpa variasi tajam.

#### 4.3.5 Hasil Ekstraksi Fitur

Hasil akhir proses ekstraksi fitur berupa 288 vektor fitur *citra* yang masing-masing merepresentasikan satu gambar bangunan pascabencana. Vektor-vektor ini menyimpan informasi visual kompleks yang diperoleh dari lapisan *GoogLeNet*

“mixed2”. Contoh bentuk fitur ditunjukkan bab selanjutnya, sedangkan hasil visualisasi vektor fitur menunjukkan distribusi nilai intensitas yang bervariasi sesuai kondisi kerusakan bangunan.

Proses ini menghasilkan data berukuran besar namun informatif, yang selanjutnya digunakan dalam tahapan reduksi dimensi PCA dan *clustering* untuk mengidentifikasi pola dan tingkat kerusakan bangunan.

#### **4.3.6 Penyimpanan Fitur**

Vektor fitur hasil ekstraksi disimpan dalam format.csv menggunakan pustaka *pandas*. Setiap baris mewakili satu *citra* dengan ribuan kolom berisi nilai fitur hasil konvolusi. Melalui tahapan ini, setiap *citra* bangunan direpresentasikan sebagai vektor numerik yang kaya informasi visual, yang nantinya digunakan untuk proses reduksi dimensi menggunakan PCA.

#### **4.3.7 Pembentukan dan Normalisasi Vektor Fitur**

Hasil ekstraksi fitur *GoogLeNet* berupa vektor fitur berdimensi tinggi perlu distandardkan agar setiap fitur memiliki skala yang seimbang. Normalisasi dilakukan menggunakan metode *StandardScaler* dari pustaka *scikit-learn* dengan persamaan 4.1:

$$X' = \frac{X - X^-}{\sigma} \quad (4.1)$$

Dengan  $X^-$  adalah nilai rata-rata dan  $\sigma$  adalah standar deviasi setiap fitur.

Langkah ini penting untuk mencegah dominasi fitur tertentu terhadap analisis PCA dan menjaga kestabilan hasil komputasi. *Pre-processing* data yang esensial, yakni standardisasi fitur, sebagai persiapan untuk analisis PCA. Pertama, mengidentifikasi dan mengekstrak kolom-kolom numerik (termasuk tipe data

float64 dan int64) dari DataFrame `df_texture_features_top10`. Selanjutnya, sebuah objek *StandardScaler* diinisiasi untuk melakukan standardisasi, sebuah proses yang mentransformasi data agar memiliki rata-rata nol dan standar deviasi satu. Metode `fit_transform` kemudian diterapkan pada data numerik tersebut; langkah ini krusial untuk memastikan fitur-fitur dengan skala yang lebih besar tidak mendominasi atau membiaskan hasil analisis komponen utama. Data yang telah distandardisasi ini dikonversi kembali menjadi *DataFrame* baru (`df_scaled`). Untuk menjaga integritas dan keterbacaan data, kolom '*Channel*' yang berfungsi sebagai identifier diekstrak dari DataFrame `df_texture_features_top10` yang asli. Akhirnya, kolom '*Channel*' ini digabungkan secara horizontal (pd.concat dengan axis=1) dengan *DataFrame* `df_scaled` yang berisi fitur-fitur ternormalisasi, dan 10 baris teratas dari *DataFrame* gabungan tersebut ditampilkan untuk validasi.

#### 4.4 Integrasi Ekstraksi Fitur *GoogLeNet* dengan PCA-*Clustering*

Setelah diperoleh vektor fitur berdimensi tinggi hasil ekstraksi dari *GoogLeNet*, langkah berikutnya adalah mengintegrasikannya dengan metode *Principal Component Analysis* (PCA) untuk mereduksi dimensi serta melakukan pengelompokan tingkat kerusakan bangunan secara *unsupervised*.

Tujuan utama dari integrasi ini adalah mengubah ribuan dimensi hasil *GoogLeNet* menjadi representasi data yang lebih sederhana namun tetap mempertahankan informasi penting, sehingga pola visual antar gambar dapat terlihat dengan jelas dalam ruang dua dimensi.

##### 4.4.1 Reduksi Dimensi menggunakan PCA

Proses reduksi dimensi dilakukan menggunakan algoritma PCA dari pustaka

*scikit-learn*. Setiap vektor fitur yang telah dinormalisasi diubah ke dalam ruang komponen utama. Langkah-langkahnya meliputi:

- 1) Perhitungan Matriks *Kovarians*

Matriks *kovarians* dihitung untuk menilai korelasi antar fitur hasil *GoogLeNet* berdasarkan persamaan 4.2.

$$C = \frac{1}{n-1} (X^T X) \quad (4.2)$$

Di mana  $X$  adalah *matriks* fitur hasil normalisasi.

- 2) Pembentukan Komponen PCA (*n-component*)

PCA membentuk sejumlah komponen utama (*principal components*) yang merupakan representasi baru dari data asli namun dengan dimensi yang lebih rendah. Mengimplementasikan inti dari *Principal Component Analysis* (PCA) untuk reduksi dimensi pada data fitur yang telah distandardisasi (`df_scaled`). Pertama menginisiasi model PCA dengan parameter `n_components=2`, yang secara eksplisit menginstruksikan algoritma untuk mentransformasi data ke dalam ruang dua dimensi. Metode `fit_transform` kemudian dieksekusi pada `df_scaled`, yang berfungsi untuk menghitung dua komponen utama (PC1 dan PC2) dan memproyeksikan data ke sumbu baru tersebut. Membagi data yang dinormalisasi menjadi *Principal Component* (PC) merupakan proses penting dalam analisis PCA karena menentukan representasi baru dari data dengan dimensi yang lebih rendah tanpa kehilangan informasi utama. Jumlah PC yang dipilih secara langsung memengaruhi kualitas hasil analisis, baik dalam hal kemampuan model menjelaskan variansi data maupun keakuratan dalam proses selanjutnya

klasterisasi.

### 3) Perhitungan Nilai Eigen (*Eigenvalues*) dan Vektor Eigen (*Eigenvectors*)

Setiap *eigenvalue* menggambarkan besarnya variansi yang dijelaskan oleh satu komponen utama. Komponen dengan nilai eigen terbesar menampung informasi paling signifikan.

Hasil transformasi ini, yang disimpan dalam *principal components*, kemudian dikonversi menjadi sebuah *DataFrame pandas* (df\_pca) untuk kemudahan analisis dan pelabelan. Selanjutnya mengevaluasi signifikansi dari kedua komponen tersebut dengan menampilkan *rasio varians* yang dijelaskan (*explained\_variance\_ratio\_*) untuk masing-masing komponen, serta menampilkan kembali nilai yang sama di bawah label "*Eigenvalues*". Akhirnya, kode ini mengkalkulasi dan menampilkan total variansi yang dijelaskan oleh kedua komponen tersebut sebagai persentase, yang mengukur efektivitas reduksi dimensi dalam mempertahankan informasi dari dataset asli. Melakukan evaluasi kuantitatif dan visual terhadap hasil *Principal Component Analysis* (PCA). Pertama berfokus pada pelaporan numerik dengan mengekstrak atribut *explained\_variance\_ratio\_* (*ratio varians*) dan *explained\_variance* dari objek PCA. Dengan melakukan iterasi untuk mencetak nilai-nilai ini secara presisi, serta menghitung dan menampilkan rasio variansi kumulatif menggunakan *numpy.cumsum*, menggunakan *matplotlib* untuk membuat visualisasi data-data tersebut. Kode ini menginisiasi sebuah *figure* yang berisi tiga subplot horizontal (1x3): bar chart untuk "*Explained Variance Ratio*", bar chart untuk "*Eigenvalue*", dan bar chart untuk "*Cumulative Explained Variance Ratio*". Dengan menyajikan metrik-metrik ini baik secara tekstual

maupun grafis, kode ini memberikan analisis komprehensif untuk memvalidasi efektivitas reduksi dimensi dan menentukan kontribusi setiap komponen utama.

Berdasarkan penelitian Almais dkk. yang merekomendasikan penggunaan rentang nilai eigen atau *rasio varians* sebagai kriteria seleksi, studi ini menggunakan dua komponen utama (PC) pertama. Komponen-komponen ini menunjukkan *ratio varians* yang stabil dan optimal untuk menganalisis karakteristik kerusakan bangunan. Eksperimen dengan berbagai parameter membuktikan bahwa kedua komponen ini mempertahankan sebagian besar informasi varians dari data asli sambil memungkinkan visualisasi yang efektif dan analisis kluster dalam ruang dimensi rendah.

#### **4.4.2 Pembentukan Klaster Tingkat Kerusakan**

Setelah data direduksi menjadi dua dimensi (PC1 dan PC2), dilakukan proses pengelompokan (*clustering*) untuk menentukan tingkat kerusakan bangunan. Pendekatan pengelompokan dalam penelitian ini menggunakan metode PCA-*Clustering* berbasis ambang nilai pada sumbu PC1.

Menyelesaikan tahapan akhir dalam proses analisis PCA dengan memvisualisasikan hasil proyeksi pada bidang dua dimensi yang dibentuk oleh dua komponen utama, yaitu PC1 dan PC2. Kode ini mengawali proses dengan menggabungkan *DataFrame* df\_pca yang berisi koordinat komponen utama dengan kolom *Channel* dari data awal, memastikan setiap titik data memiliki label identifikasinya. Selanjutnya, menggunakan *library matplotlib.pyplot* untuk menghasilkan *scatter plot* (diagram pencar). Setiap *channel* direpresentasikan sebagai satu titik pada diagram dengan koordinat (PC1, PC2). Titik-titik ini diberi

label dengan nomor *channel* aslinya menggunakan fungsi plt.text untuk memfasilitasi pelacakan visual. Untuk memastikan diagram mudah diinterpretasikan, label sumbu (*xlabel* dan *ylabel*) diatur menjadi '*Principal Component 1*' dan '*Principal Component 2*', judul diagram ditetapkan sebagai '*PCA for Texture Features*', dan *grid horizontal* ditambahkan, yang secara kolektif memungkinkan peneliti untuk mengamati distribusi dan pola pengelompokan *channel* fitur tekstur secara visual.

## BAB V

### UJI COBA

#### 5.1 Skenario Uji Coba

Tahap pertama dimulai dengan pengumpulan data berupa *citra* kerusakan bangunan yang menjadi dasar utama analisis. *Citra* ini dapat berasal dari foto lapangan, dokumentasi pascabencana, atau dataset *citra* publik yang telah digunakan dalam penelitian sebelumnya. Data *citra* tersebut digunakan sebagai *input* awal untuk proses ekstraksi fitur berbasis *deep learning*. Pada tahap ini, kualitas dan representativitas *citra* sangat penting karena akan memengaruhi akurasi proses identifikasi pada tahap berikutnya.

Pada tahap kedua, *citra* yang telah dikumpulkan diproses menggunakan arsitektur *GoogLeNet* yang terdiri dari dua komponen utama, yaitu *Stem Network* dan *Inception Module*, *Stem Network* bertugas melakukan pemrosesan awal, seperti ekstraksi pola dasar dan penyederhanaan struktur visual. Selanjutnya, *Inception Module* mengekstraksi fitur tingkat lanjut dengan memanfaatkan kombinasi filter konvolusi berukuran berbeda untuk menangkap variasi pola tekstur kerusakan bangunan. Hasil dari tahapan ini berupa vektor fitur berukuran tinggi yang merepresentasikan karakteristik visual *citra* secara lebih mendalam.

Tahap ketiga menghasilkan vector fitur atau fitur rekayasa yang diperoleh dari proses ekstraksi *GoogLeNet*. Fitur-fitur ini memuat informasi visual penting seperti pola keretakan, deformasi struktur, dan intensitas kerusakan. Vektor fitur kemudian disiapkan untuk tahap analisis lanjutan guna mengidentifikasi tingkat kerusakan bangunan. Pada tahap ini, vektor fitur menjadi representasi numerik dari *citra* yang

memungkinkan sistem melakukan analisis secara komputasional dan objektif.

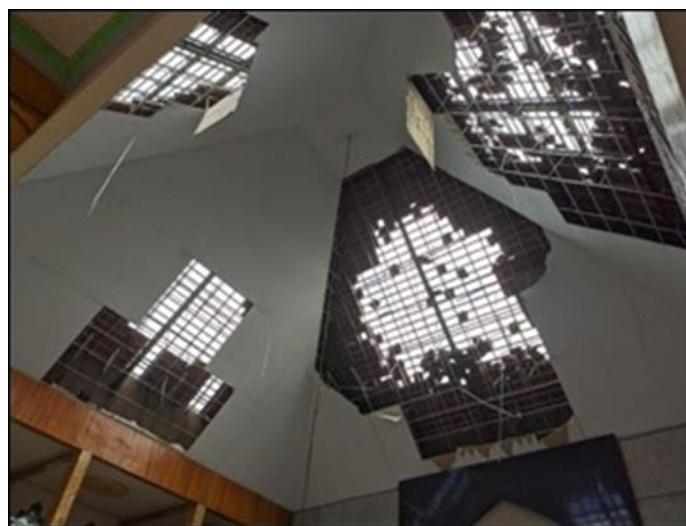
Tahap keempat menambahkan input berupa fitur statistik klasik, seperti kontras, energi, dan homogenitas yang dihitung menggunakan metode tekstur seperti GLCM. Fitur-fitur ini memberikan informasi pelengkap terkait pola distribusi intensitas piksel pada *citra*, sehingga meningkatkan kemampuan sistem dalam membedakan area kerusakan ringan, sedang, maupun berat. Integrasi fitur statistik klasik dan fitur deep learning memperkaya representasi data yang akan dianalisis pada tahap PCA-*Clustering*.

Pada tahap kelima, seluruh fitur yang diperoleh baik dari *GoogLeNet* maupun fitur statistik klasik diproses menggunakan PCA-*Clustering*. Tahapan PCA dimulai dengan normalisasi untuk menyetarakan skala antarvariabel, kemudian dilanjutkan dengan penyusunan matriks kovarian atau korelasi. Setelah itu, PCA menghitung jumlah komponen utama (N-komponen) yang paling berpengaruh berdasarkan nilai *eigen* dan *eigenvector*. Komponen utama ini merangkum sebagian besar variasi data sehingga mempermudah proses *clustering* untuk mengelompokkan tingkat kerusakan bangunan secara lebih efisien dan akurat.

Tahap terakhir menghasilkan *output* berupa visualisasi tingkat kerusakan bangunan yang telah dikelompokkan menggunakan pendekatan PCA-*Clustering*. Visualisasi ini menampilkan penyebaran data berdasarkan komponen utama, sehingga pola perbedaan antar tingkat kerusakan dapat diamati dengan jelas. Hasil visualisasi memudahkan peneliti dan pihak terkait dalam memahami distribusi kerusakan, mengevaluasi keakuratan model, serta mendukung proses pengambilan keputusan pada tahap respons dan pemulihan pascabencana.

## 5.2 Input Citra

Pada proses ini akan menginput gambar yang merupakan ukuran input standar untuk model *InceptionV3*, kemudian mengubah *citra* menjadi *array* numerik dan melakukan pra-pemrosesan (*preprocessing*) agar sesuai dengan format data yang digunakan dalam pelatihan model *ImageNet* seperti pada Gambar 5.1



Gambar 5.1 Kerusakan Bangunan

## 5.3 Hasil Ekstraksi Fitur Menggunakan *GoogLeNet*

Lapisan yang digunakan untuk mengambil keluaran fitur adalah lapisan “*mixed2*”. Lapisan ini menghasilkan peta fitur dengan dimensi (1, 35, 35, 288). Dimensi ini menunjukkan bahwa untuk setiap *citra*, model menghasilkan 288 peta fitur (*channel*). Setiap saluran menggambarkan pola tertentu seperti tepi, tekstur, bentuk, atau area perbedaan intensitas warna yang signifikan pada Tabel 5.1.

Tabel 5.1 Hasil Ekstraksi Fitur *GoogLeNet*

Channel	Kontras	Energi	Homogenitas
Channel 0	Kontras=0.63,	Energi=0.54,	Homogenitas=0.61
Channel 1	Kontras=0.81,	Energi=0.85,	Homogenitas=0.55
Channel 2	Kontras=0.73,	Energi=0.93,	Homogenitas=0.58

<b>Channel</b>	<b>Kontras</b>	<b>Energi</b>	<b>Homogenitas</b>
Channel 3	Kontras=0.26,	Energi=0.08,	Homogenitas=0.79
Channel 4	Kontras=0.53,	Energi=0.41,	Homogenitas=0.65
Channel 5	Kontras=0.49,	Energi=0.29,	Homogenitas=0.67
Channel 6	Kontras=0.57,	Energi=0.44,	Homogenitas=0.64
Channel 7	Kontras=0.80,	Energi=1.04,	Homogenitas=0.56
Channel 8	Kontras=0.53,	Energi=0.33,	Homogenitas=0.65
Channel 9	Kontras=0.63,	Energi=0.64,	Homogenitas=0.61

Hasil perhitungan fitur statistik tekstur pada sepuluh *channel* menunjukkan variasi nilai kontras, energi, dan homogenitas yang menggambarkan perbedaan karakteristik visual pada *citra* kerusakan bangunan *channel* dengan kontras tinggi, seperti *channel* 1 (0.81) dan *channel* 7 (0.80), menandakan adanya perubahan intensitas piksel yang kuat, yang biasanya muncul pada area dengan retakan besar atau kerusakan struktural yang signifikan. Sebaliknya, *channel* dengan kontras rendah seperti *channel* 3 (0.26) menunjukkan tekstur yang lebih seragam dan lebih sedikit variasi intensitas.

Dari sisi energi, beberapa *channel* memiliki nilai tinggi, misalnya *channel* 7 (1.04) dan *channel* 2 (0.93), yang menunjukkan adanya pola tekstur yang lebih teratur dan berulang. Hal ini mengindikasikan bahwa bagian *citra* pada *channel* ini menangkap struktur material yang masih relatif stabil. Sebaliknya, energi rendah pada *channel* 3 (0.08) mengindikasikan tekstur yang lebih acak dan tidak terstruktur, yang sering ditemukan pada *citra* puing atau area yang mengalami kerusakan berat.

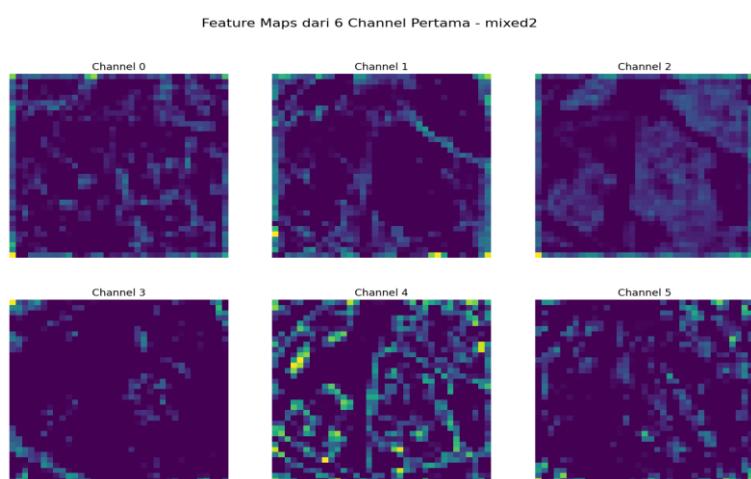
Nilai homogenitas tertinggi terlihat pada *channel* 3 (0.79) dan *channel* 5 (0.67), yang mencerminkan tingkat keseragaman intensitas piksel yang tinggi. Hal ini menunjukkan bahwa *channel* tersebut menangkap area *citra* yang relatif halus dan

tidak memiliki perubahan tekstur yang ekstrem. Sebaliknya, homogenitas yang lebih rendah pada *channel* 1 (0.55) dan *channel* 7 (0.56) mengindikasikan tekstur yang lebih kasar dan kompleks, sesuai dengan karakteristik area yang mengalami kerusakan signifikan.

Secara keseluruhan, variasi nilai kontras, energi, dan homogenitas pada masing-masing *channel* menunjukkan bahwa fitur tekstur yang dihasilkan berhasil menangkap perbedaan pola visual antara area bangunan yang rusak ringan, sedang, dan berat. Perbedaan karakteristik ini kemudian menjadi dasar penting dalam proses ekstraksi fitur dan analisis lebih lanjut menggunakan PCA.

#### 5.4 Visualisasi Fitur Peta *GoogLeNet*

Proses selanjutnya ekstraksi fitur dilakukan dari gambar RGB bangunan pascabencana alam. Setiap *citra* diolah menggunakan model pra-latih *GoogLeNet* dengan memanfaatkan lapisan “*mixed2*” untuk menghasilkan fitur semantik berdimensi tinggi. Konversi ini menghasilkan *matriks* yang mewakili distribusi intensitas piksel untuk setiap gambar, seperti yang ditampilkan pada Gambar 5.2.



Gambar 5.2 Hasil Representasi Intensitas Piksel

Visualisasi *feature map* dari enam *channel* awal hasil ekstraksi fitur menggunakan arsitektur *GoogLeNet*. Setiap *channel* menggambarkan area aktivasi berbeda yang dihasilkan oleh proses konvolusi pada *citra* bangunan, di mana warna biru gelap menunjukkan nilai aktivasi rendah (fitur kurang menonjol), sedangkan warna kuning, hijau terang menandakan nilai aktivasi tinggi atau area yang dianggap penting oleh model. Pada *channel* 0 dan 3, pola aktivasi masih acak dan berstruktur halus, yang menandakan bahwa filter ini mendekripsi fitur dasar seperti tepi dan gradasi cahaya. *Channel* 1 dan 2 mulai memperlihatkan pola yang lebih terstruktur, menyerupai batas dinding atau bidang permukaan bangunan, yang berhubungan dengan area kerusakan sedang. Sementara itu, *channel* 4 menampilkan aktivasi kuat pada beberapa titik dengan pola melingkar dan tajam, menunjukkan sensitivitas tinggi terhadap tekstur retakan atau deformasi yang jelas. *channel* 5 memperlihatkan pola campuran antara area halus dan aktif, mencerminkan transisi dari pola tekstur ringan ke pola dengan kontras lebih tinggi. Secara keseluruhan, keenam *channel* ini menunjukkan bahwa *GoogLeNet* mampu menangkap berbagai karakteristik visual dari *citra* bangunan, mulai dari fitur dasar hingga pola kerusakan yang lebih kompleks, yang penting untuk mendukung proses pengelompokan tingkat kerusakan bangunan.

## 5.5 Analisis Fitur Statik Klasik

Nilai fitur tekstur sampel (kontras, energi, homogenitas) yang diekstraksi dari 10 saluran teratas menggunakan arsitektur *GoogLeNet*. Sepuluh saluran ini mewakili subset dari total 288 vektor fitur dihasilkan melalui pendekatan *Gray-Level Co-occurrence Matrix* (GLCM). Pemilihan 10 saluran teratas dilakukan

untuk ilustratif dan analisis eksplorasi awal, mengingat volume yang besar dari kumpulan fitur lengkap. Distribusi nilai fitur dianalisis untuk mendeteksi nilai penyimpangan (*outlier*) dan mengevaluasi kebutuhan normalisasi akibat perbedaan skala diantara berbagai fitur, pada tabel 5.2 bentuk fitur statik klasik.

**Tabel 5.2 Bentuk Fitur Statik Klasik**

Channel	Contrast	Energy	Homogenitas
0	0.63	0.54	0.61
1	0.81	0.85	0.55
2	0.73	0.92	0.57
3	0.26	0.07	0.79
4	0.52	0.40	0.65
5	0.48	0.28	0.67
6	0.56	0.43	0.63
7	0.80	1.03	0.55
8	0.53	0.32	0.65
9	0.63	0.63	0.61

Nilai kontras bervariasi antara 0.26 hingga 0.81, menunjukkan bahwa beberapa *channel* menangkap pola kasar (misalnya tepi bangunan), sementara lainnya fokus pada area halus. Nilai energi relatif tinggi pada *channel* 1, 2, dan 7 ( $>0.8$ ), yang menandakan bahwa filter-filter ini mendeteksi pola yang berulang atau teratur kemungkinan permukaan dinding atau struktur bangunan yang seragam. Dan Nilai homogenitas berkisar antara 0.55–0.79. Nilai tinggi (misalnya pada *channel* 3 = 0.79) menunjukkan area dengan tekstur seragam, sedangkan nilai rendah (misalnya *Channel* 1 = 0.55) menandakan area dengan variasi intensitas tinggi (seperti retakan atau kerusakan).

Dari hasil tersebut dapat disimpulkan bahwa *GoogLeNet* berhasil mengekstraksi fitur tekstur yang beragam dari *citra* bangunan. Beberapa *channel* fokus pada pola bertekstur halus (homogenitas tinggi), sementara yang lain lebih

sensitif terhadap pola tepi dan kerusakan (kontras tinggi). Kombinasi variasi ini sangat penting dalam analisis kerusakan bangunan karena setiap jenis kerusakan (ringan, sedang, berat) memiliki pola tekstur yang berbeda seperti permukaan retak, dinding pecah, atau area runtuh.

## 5.6 Normalisasi Data

Fitur-fitur telah diskalakan sehingga memiliki rata-rata nol dan simpangan baku satu. Normalisasi ini sangat penting untuk mencegah PCA menjadi bias terhadap fitur-fitur dengan skala yang lebih besar yang ditunjukkan pada tabel 5.3 hasil normalisasi (*StandardScaler*).

Tabel 5.3 Hasil Normalisasi (*StandardScaler*)

Channel	Contrast	Energy	Homogenitas
0	-1.56	0.21	-0.03
1	-1.21	1.37	1.02
2	-0.87	0.85	1.27
3	-0.52	-2.15	-1.63
4	-0.17	-0.44	-0.50
5	0.17	-0.72	-0.90
6	0.52	-0.21	-0.39
7	0.87	1.29	1.65
8	1.21	-0.43	-0.77
9	1.56	0.22	0.29

Vektor fitur berdimensi tinggi yang dihasilkan dari proses ekstraksi fitur *GoogLeNet* berfungsi sebagai masukan utama untuk algoritma PCA. Selama persiapan data, vektor-vektor ini dikonversi ke dalam struktur *spreadsheet* tabel untuk memudahkan analisis kuantitatif. Konversi ini memungkinkan pengelolaan data yang lebih terstruktur dan efisien, terutama saat menangani dataset berskala besar yang mengandung ribuan fitur dari ribuan sampel gambar kerusakan bangunan.

## 5.7 Pembentukan N-Komponen PCA

Proses reduksi dimensi dilakukan menggunakan *Principal Component Analysis* (PCA) dengan dua komponen utama. Penentuan `n_components = 2` bertujuan merangkum informasi terpenting dari fitur hasil ekstraksi *GoogLeNet* ke dalam ruang berdimensi rendah agar pola data lebih mudah diamati dan dianalisis. Data yang telah dinormalisasi kemudian diproyeksikan ke dua komponen utama, sehingga diperoleh nilai PC1 dan PC2 sebagai representasi baru dari setiap *citra* bangunan.

Membagi data yang dinormalisasi menjadi *Principal Component* (PC) merupakan proses penting dalam analisis PCA karena menentukan representasi baru dari data dengan dimensi yang lebih rendah tanpa kehilangan informasi utama. Jumlah PC yang dipilih secara langsung memengaruhi kualitas hasil analisis, baik dalam hal kemampuan model menjelaskan variansi data maupun keakuratan. Seperti pada tabel 5.4 berikut.

Tabel 5.4 Hasil N-Komponen

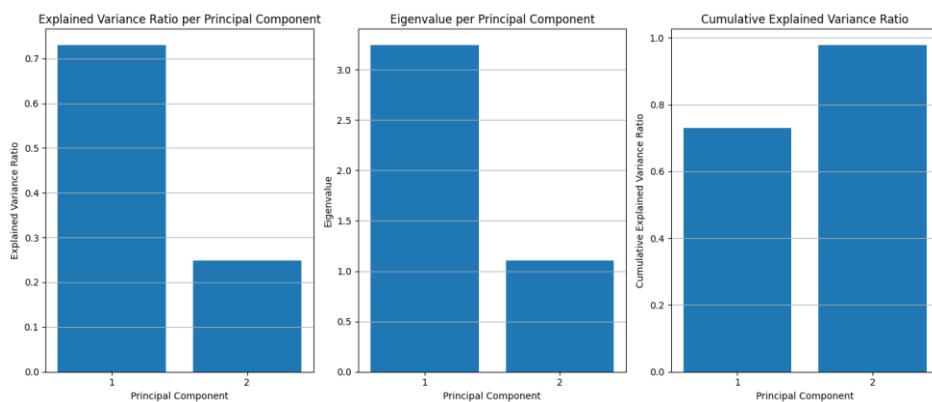
Komponen	Eigenvalues (Nilai Eigen)	Explained Variance Ratio (%)
PC 1	3.2412963	72.93%
PC 2	1.10417824	24.84%
Total	4.34547454	97.77%

Analisis *Principal Component Analysis* (PCA) menunjukkan keberhasilan luar biasa dalam mereduksi dimensi data, karena dua komponen utama (PC 1 dan PC 2) secara kolektif mampu menjelaskan 97.77% dari total varian data. Komponen pertama (PC 1) mendominasi dengan kontribusi 72.93% varian, sementara Komponen kedua (PC 2) menyumbang 24.84%. Kedua komponen tersebut valid secara statistik karena memiliki nilai eigen di atas 1.0, membuktikan bahwa hampir

seluruh informasi penting dari fitur asli telah dikompres menjadi dua variabel baru yang sangat efisien, menjadikannya input optimal untuk tahap pemodelan klasterisasi.

### 5.8 *Eigen Values dan Eigen Variance Ratio*

Hasil transformasi ini, yang disimpan dalam *principal components*, kemudian dikonversi untuk kemudahan analisis dan pelabelan. Selanjutnya mengevaluasi signifikansi dari kedua komponen tersebut dengan menampilkan rasio varians untuk masing-masing komponen, serta menampilkan kembali nilai yang sama di bawah label "*Eigenvalues*". Akhirnya mengkalkulasi dan menampilkan total varians yang dijelaskan oleh kedua komponen tersebut sebagai persentase, yang mengukur efektivitas reduksi dimensi dalam mempertahankan informasi dari dataset asli.



Gambar 5.3 Plot Nilai Eigen dan *Rasio Varians* untuk Komponen Utama.

Kontribusi masing-masing komponen utama (PC) terhadap total variansi data yang dihasilkan dari proses analisis PCA. Grafik pertama menunjukkan rasio variansi yang dijelaskan (*Explained Variance Ratio*), di mana PC1 memiliki kontribusi paling besar sebesar 73,49%, sedangkan PC2 berkontribusi sebesar

24,58% terhadap total variansi. Grafik kedua memperlihatkan nilai eigen (*Eigenvalue*) masing-masing komponen, dengan PC1 bernilai 3,27 dan PC2 bernilai 1,09, yang menggambarkan besarnya informasi atau variansi yang ditangkap oleh masing-masing sumbu utama. Adapun grafik ketiga menunjukkan rasio variansi kumulatif (*Cumulative Explained Variance Ratio*), di mana kombinasi dua komponen utama pertama (PC1 dan PC2) telah mampu menjelaskan 98,07% dari total variansi dataset. Hal ini menandakan bahwa dua komponen utama tersebut sudah cukup mewakili hampir seluruh informasi penting dalam data fitur hasil ekstraksi *GoogLeNet*. Oleh karena efisiensinya dalam merangkum informasi, PC1 dan PC2 dipilih sebagai dasar proyeksi data dalam analisis klasterisasi, karena keduanya mampu menggambarkan pola distribusi *citra* dengan jelas dalam ruang berdimensi rendah, sebagaimana dijelaskan pada Tabel 5.5 Hasil PCA.

Tabel 5.5 Hasil Analisis Komponen Utama (PCA)

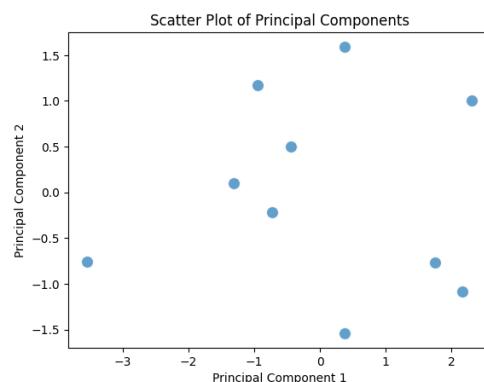
Komponen Utama (PC)	Nilai Eigen	Rasio Variansi	Kumulatif Variansi
PC 1	3,27	0,7349 (73,49%)	0,7349 (73,49%)
PC2	1,09	0,2458 (24,58%)	0,9807 (98,07%)

Hasil analisis *Principal Component Analysis* (PCA) yang digunakan untuk mereduksi dimensi fitur dan mengidentifikasi komponen utama yang paling berpengaruh dalam data. Berdasarkan hasil perhitungan nilai *eigen*, terdapat dua komponen utama (PC1 dan PC2) yang memiliki kontribusi signifikan dalam menjelaskan keragaman data. PC1 memiliki nilai *eigen* sebesar 3,27, sedangkan PC2 memiliki nilai eigen 1,09. Nilai *eigen* yang lebih besar menunjukkan bahwa komponen tersebut menyimpan informasi atau variansi yang lebih penting dari keseluruhan data asli.

Rasio variansi menunjukkan proporsi keragaman data yang dapat dijelaskan oleh masing-masing komponen utama. PC1 mampu menjelaskan 73,49% variansi data, yang berarti komponen ini menjadi faktor dominan dalam menggambarkan struktur informasi data. Sementara itu, PC2 menjelaskan 24,58% variansi tambahan. Dengan demikian, kedua komponen ini secara kolektif mampu mewakili 98,07% variansi total, sebagaimana terlihat pada kolom kumulatif variansi. Angka kumulatif tersebut menunjukkan bahwa sebagian besar informasi penting dalam data telah berhasil dirangkum hanya dalam dua komponen utama.

## 5.9 Hasil Visualisasi

Selanjutnya menampilkan visualisasi dari data yang diproyeksikan ke subruang yang didefinisikan oleh dua komponen utama pertama (PC1 dan PC2). Setiap titik data mewakili vektor fitur yang dihasilkan dari gambar bangunan tunggal dalam dataset seperti pada gambar 5.4 hasil visualisasi.



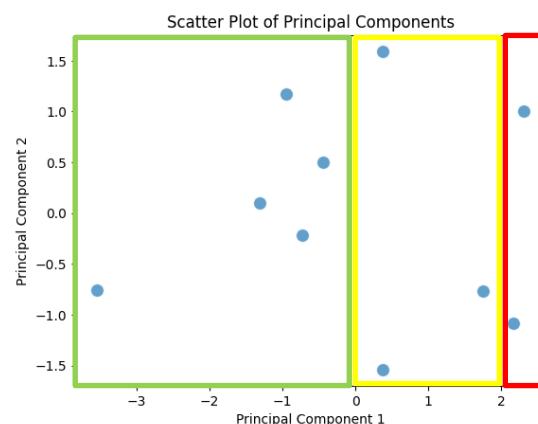
Gambar 5.4 Hasil Visualisasi

Plot sebaran (*scatter plot*) data setelah direduksi ke dalam dua dimensi baru menggunakan *Principal Component Analysis* (PCA). Sumbu horizontal mewakili PC 1, dan sumbu vertikal mewakili PC 2. Setiap titik biru merepresentasikan satu

sampel data. Tujuan utama dari visualisasi ini adalah untuk melihat bagaimana data telah terpisah di ruang dimensi baru. Meskipun titik-titik tersebut belum menunjukkan klaster (pengelompokan) yang jelas dan terpisah secara alami, penyebaran yang luas, terutama di sepanjang sumbu PC 1 (dari sekitar -3.5 hingga +2.5), mengindikasikan bahwa PC 1 berhasil menangkap perbedaan karakteristik yang signifikan antar sampel. Plot ini menjadi dasar visual sebelum dilakukan proses *clustering*, yang akan secara eksplisit menarik garis batas untuk membentuk kelompok-kelompok kerusakan.

### 5.10 Hasil Visualisasi Pengelompokan

Berdasarkan ambang batas yang diterapkan pada sumbu PC1 menurut penelitian Almais dkk, titik data dikategorikan ke dalam tiga kelas tingkat kerusakan ringan ( $n < 0$ ), sedang ( $0 \leq n < 2$ ), dan berat ( $n \geq 2$ ). Seperti pada gambar 5.5 proyeksi data berdasarkan klaster hasil PCA-*Clustering*.



Gambar 5.5 Proyeksi Data Berdasarkan Klaster Hasil PCA-*Clustering*.

Hasil visualisasi menunjukkan bahwa model berhasil memisahkan kelas-kelas yang telah ditentukan dengan baik. Kluster-kluster tersebut membentuk kelompok spasial yang terpisah sesuai dengan label yang diberikan: kluster kerusakan ringan

(hijau) mendominasi wilayah di mana  $PC1 < 0$  kiri, kluster kerusakan sedang (kuning) terkonsentrasi di wilayah tengah ( $0 \leq PC1 < 2$ ), dan kluster kerusakan berat (merah) terletak di sebelah kanan ( $PC1 \geq 2$ ) dijelaskan juga pada tabel 5.5 dalam proyeksi data berdasarkan klaster hasil PCA-*Clustering*.

**Tabel 5.5 Proyeksi Data Berdasarkan Klaster Hasil PCA-*Clustering*.**

Warna	PC1	PC2	Nilai Koordinat (n)	Label
Kuning	0.373353	-1.540810	$0 \leq n < 2$	Rusak Sedang
Merah	2.161966	-1.087124	$n \geq 2$	Rusak Berat
Kuning	1.747507	-0.769260	$0 \leq n < 2$	Rusak sedang
Hijau	-3.535412	-0.754085	$n < 0$	Rusak Ringan
Hijau	-0.724774	-0.215342	$n < 0$	Rusak Ringan
Hijau	0.443805	0.098480	$n < 0$	Rusak Ringan
Hijau	-0.440452	0.500792	$n < 0$	Rusak Ringan
Merah	2.309781	1.006388	$n \geq 2$	Rusak Berat
Hijau	-0.952852	1.168190	$n < 0$	Rusak Ringan
Kuning	0.367692	1.592771	$0 \leq n < 2$	Rusak Sedang

Hasil akhir dari proses PCA-*Clustering*, di mana data sampel yang telah direduksi menjadi skor PC1 dan PC2 berhasil dikelompokkan ke dalam tiga klaster tingkat kerusakan yang bermakna. Proses *clustering* ini mengelompokkan data berdasarkan kriteria Nilai Koordinat (n). Hasil ini menunjukkan keberhasilan algoritma dalam memetakan skor abstrak dari analisis PCA ke dalam label kualitatif yang dapat diinterpretasikan secara praktis, seperti terlihat pada sampel indeks 3 yang memiliki koordinat rendah (-3.535 pada PC1) dan dikluster sebagai rusak ringan. Hasil pengelompokan yang dirangkum dalam Tabel 5.6 menunjukkan persentase distribusi label kerusakan yang tidak merata.

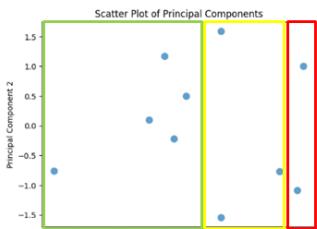
**Tabel 5.6 Persentase Hasil Distribusi Label berdasarkan Kluster dari PC1.**

Label	Persentase
Rusak Ringan	50.0%
Rusak Sedang	30.0%

Rusak Berat	20.0%
-------------	-------

Diketahui bahwa kategori Kerusakan ringan memiliki proporsi terbesar, yaitu sebesar 50,0%, sehingga menjadi kelompok yang paling dominan dalam distribusi tingkat kerusakan bangunan. Dengan ini menunjukkan bahwa sebagian besar bangunan dalam dataset berada dalam kondisi relatif baik, dengan kerusakan bersifat ringan dan tidak mengganggu stabilitas struktur utama. Selanjutnya, kategori kerusakan sedang menempati posisi kedua dengan persentase 30,0%, yang menggambarkan bangunan dengan tingkat kerusakan cukup signifikan, namun masih dapat diperbaiki tanpa perlu rekonstruksi total. Sementara itu, kategori berusakan berat hanya berkontribusi sebesar 20,0%, yang mengindikasikan bahwa jumlah bangunan yang mengalami kerusakan parah atau runtuh sepenuhnya relatif lebih sedikit dibandingkan kategori lainnya. Secara keseluruhan, hasil ini memperlihatkan bahwa area penelitian didominasi oleh bangunan dengan tingkat kerusakan ringan, yang sejalan dengan karakteristik wilayah terdampak bencana dengan intensitas dampak menengah.

Tabel 5.7 Hasil Validasi Eksternal

Gambar	Scatter Plot	Validasi PC1	Validasi Eksternal	Sesuai/Tidak Sesuai
 (Sumber: Gambar dari BMKG)		Rusak Berat	Rusak Ringan	Sesuai

Hasil validasi akhir tingkat kerusakan bangunan yang diperoleh melalui analisis kombinasi *GoogLeNet PCA-Clustering*. Pada kolom pertama ditampilkan *citra*

bangunan dengan kondisi kerusakan ringan, di mana sebagian plafon dan atap mengalami retakan atau runtuh sebagian namun struktur utama masih stabil. Grafik *scatter plot* pada kolom kedua menggambarkan hasil proyeksi data ke dalam ruang dua komponen utama (PC1 dan PC2). Berdasarkan posisi titik pada sumbu PC1, *citra* tersebut terkluster ke dalam kategori “rusak ringan”, yang ditandai dengan area berwarna hijau pada grafik. Hasil ini kemudian divalidasi oleh ahli dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG), yang menyatakan bahwa kondisi visual bangunan pada *citra* tersebut memang sesuai dengan kategori kerusakan ringan. Dengan demikian, hasil identifikasi yang diperoleh dari model *GoogLeNet PCA-Clustering* terbukti konsisten dan sesuai dengan penilaian ahli lapangan, menunjukkan bahwa pendekatan ini mampu memberikan hasil yang akurat dan dapat diandalkan dalam mendeteksi tingkat kerusakan bangunan pascabencana.

## **BAB VI**

### **PENUTUP**

#### **6.1 Kesimpulan**

##### **6.1.1 Ekstraksi Fitur menggunakan *GoogLeNet***

Ekstraksi Fitur Gambar Kerusakan Bangunan Menggunakan Arsitektur *GoogLeNet*. Sistem identifikasi tingkat kerusakan bangunan pascabencana yang dirancang telah berhasil mengimplementasikan metode ekstraksi fitur menggunakan arsitektur *GoogLeNet* yang terintegrasi dengan *Principal Component Analysis* (PCA) dan *Clustering*. Fitur semantik berdimensi tinggi diekstraksi dari *citra* bangunan pascabencana alam dengan memanfaatkan lapisan "*mixed2*" dari model *GoogLeNet*. Hasil ekstraksi fitur, yang diwakili oleh 288 peta fitur (*channel*), berhasil menangkap beragam karakteristik visual, mulai dari fitur dasar seperti tepi dan gradasi cahaya hingga pola kerusakan yang lebih kompleks. Analisis fitur statistik klasik (Kontras, Energi, Homogenitas) pada 10 *channel* teratas menunjukkan keberhasilan *GoogLeNet* dalam mengekstraksi fitur tekstur yang beragam dan sensitif terhadap pola kerusakan (kontras tinggi, homogenitas rendah) dan pola seragam (energi tinggi, homogenitas tinggi).

##### **6.1.2 Visualisasi Pengelompokan Menggunakan PCA-*Clustering***

Visualisasi Pengelompokan Tingkat Kerusakan Bangunan Pascabencana Alam Berbasis PCA. Proses reduksi dimensi menggunakan *Principal Component Analysis* (PCA) dengan dua komponen utama (PC1 dan PC2) menunjukkan efektivitas yang baik. Dua komponen utama ini secara kolektif mampu menjelaskan 98,07% dari total varian data, dengan PC1 menyumbang 73,49% dan PC2

menyumbang 24.58% menjadikan representasi yang efisien dari fitur asli.

Visualisasi data yang telah direduksi menunjukkan bahwa PC1 berhasil menangkap perbedaan karakteristik yang signifikan antar sampel. Visualisasi pengelompokan (*clustering*) berbasis ambang batas PC1 untuk ringan ( $n < 0$ ), sedang ( $0 \leq n < 2$ ), berat ( $n \geq 2$ ), berhasil memisahkan kelas-kelas kerusakan ke dalam kelompok spasial yang jelas dan bermakna, dan rusak ringan memiliki proporsi terbesar, yaitu sebesar 50,0% dengan validasi terhadap penelitian sebelumnya dan ahli lapangan menunjukkan tingkat kesesuaian tinggi.

Hasil validasi eksternal menunjukkan bahwa klasterisasi yang dihasilkan model *GoogLeNet PCA-Clustering* terbukti konsisten dan sesuai dengan penilaian ahli lapangan (BMKG) untuk mengidentifikasi tingkat kerusakan bangunan.

## 6.2 Saran

Memperluas variasi dataset dengan menambahkan *citra* dari berbagai jenis bencana dan tipe bangunan agar model lebih adaptif terhadap kondisi lingkungan yang beragam. Mengeksplorasi arsitektur *deep learning* yang lebih mutakhir seperti *EfficientNet* atau *Vision Transformer* untuk meningkatkan kemampuan ekstraksi fitur secara semantik. Mengintegrasikan data *non-visual*, seperti material bangunan dan intensitas gempa, guna memberikan konteks tambahan pada proses pengelompokan tingkat kerusakan.

## DAFTAR PUSTAKA

- A, T. W., Da-Gang Lu, & Tan, Y.-Q. (2024). Refined dimension-reduction integration method for uncertainty propagation in stochastic systems: Estimation of statistical moments. *Reliability Engineering & System Safety*, 256. [https://doi.org/https://doi.org/10.1016/j.ress.2024.110753](https://doi.org/10.1016/j.ress.2024.110753)
- Almais, A. T. W., Susilo, A., Naba, A., Sarosa, M., Crysdiyan, C., Tazi, I., Hariyadi, M. A., Muslim, M. A., Basid, P. M. N. S. A., Arif, Y. M., Purwanto, M. S., Parwatiningsyah, D., Supriyono, & Wicaksono, H. (2023). Principal Component Analysis-Based Data Clustering for Labeling of Level Damage Sector in Post-Natural Disasters. *IEEE Access*, 11, 74590–74601. <https://doi.org/10.1109/ACCESS.2023.3275852>
- Almais, A. T. W., Susilo, A., Naba, A., Sarosa, M., Juwono, A. M., Crysdiyan, C., Muslim, M. A., & Wicaksono, H. (2024). Characterization of Structural Building Damage in Post-Disaster using GLCM-PCA Analysis Integration. *IEEE Access*, 12(August), 146190–146201. <https://doi.org/10.1109/ACCESS.2024.3469637>
- Anbarasan, M., Muthu, B., Sivaparthipan, C. B., Sundarasekar, R., Kadry, S., Krishnamoorthy, S., R., D. J. S., & Dasel, A. A. (2020). Detection of flood disaster system based on IoT, big data and convolutional deep neural network. *Computer Communications*, 150, 150–157. <https://doi.org/https://doi.org/10.1016/j.comcom.2019.11.022>
- Baral, A., Singh, V., & Lath, A. (2024). Evaluating the Performance of ResNet-50 and GoogleNet for Damage Detection and Classification. *4th International Conference on Sustainable Expert Systems, ICSES 2024 - Proceedings*, 1721–1726. <https://doi.org/10.1109/ICSES63445.2024.10763274>
- Cai, S., Zhao, Y., Lyu, J., Wang, S., Hu, Y., Cheng, M., & Zhang, G. (2025). DDP-DAR: Network intrusion detection based on denoising diffusion probabilistic model and dual-attention residual network. *Neural Networks*, 184(September 2024), 107064. <https://doi.org/10.1016/j.neunet.2024.107064>
- Chen, S.-H., Wu, Y.-L., Pan, C.-Y., Lian, L.-Y., & Su, Q.-C. (2023). Breast ultrasound image classification and physiological assessment based on GoogLeNet. *Journal of Radiation Research and Applied Sciences*, 16(3), 100628. <https://doi.org/10.1016/j.jrras.2023.100628>
- Coronese, M., Lamperti, F., Keller, K., Chiaromonte, F., & Roventini, A. (2019). Evidence for sharp increase in the economic damages of extreme natural disasters. *Proceedings of the National Academy of Sciences of the United States of America*, 116(43), 21450–21455. <https://doi.org/10.1073/pnas.1907826116>
- Darmawan, O. A., Nugroho, F. A., Ardiantoro, N. F., & Wayan, I. (2022). *PELATIHAN MITIGASI BENCANA BAGI SISWA/SISWI MAS BAITUSSALAM MIFTAHUL JANNAH LAMPUNG TENGAH*. 3(2), 272–279.

- Dinc, I., Sigdel, M., Dinc, S., Sigdel, M. S., Pusey, M. L., & Aygun, R. S. (2014). Evaluation of normalization and PCA on the performance of classifiers for protein crystallization images. *Conference Proceedings - IEEE SOUTHEASTCON, March.* <https://doi.org/10.1109/SECON.2014.6950744>
- Fatul Fu, A., Teguh Wibowo Almais, A., & Syauqi, Ia. (2023). Klasifikasi Tingkat Kerusakan Sektor Pasca Bencana Alam Menggunakan Metode MULTIMOORA Berbasis Web. In *Jurnal Informatika Sunan Kalijaga* (Vol. 8, Issue 3).
- Geng, X., Wang, L., Yu, P., Hu, W., Liang, Q., Zhang, X., Chen, C., & Zhang, X. (2025). A method of EEG signal feature extraction based on hybrid DWT and EMD. *Alexandria Engineering Journal*, 113(November 2024), 195–204. <https://doi.org/10.1016/j.aej.2024.11.056>
- Gliner, J. A., Morgan, G. A., Leech, N. L., Gliner, J. A., & Morgan, G. A. (2021). Measurement Reliability and Validity. *Research Methods in Applied Settings*, 319–338. <https://doi.org/10.4324/9781410605337-29>
- Gopinathan, R. T. and S. (2021). Inception Architecture for Brain Image Classification. *Journal of Physics: Conference Series*, 1964(7). <https://doi.org/10.1088/1742-6596/1964/7/072022>
- Guilherme Dean Pelegrina, L. T. D. (2024). A Novel Approach for Fair Principal Component Analysis Based on Eigendecomposition. *IEEE Transactions on Artificial Intelligence*, 5(3), 1195–1206. <https://doi.org/10.1109/TAI.2023.3298291>
- Han, C., Lee, J., Jun, M. B. G., Lee, S. W., & Yun, H. (2025). Visual coating inspection framework via self-labeling and multi-stage deep learning strategies. *Journal of Intelligent Manufacturing*, 36(4), 2461–2478. <https://doi.org/10.1007/s10845-024-02372-9>
- Jahandad, Sam, S. M., Kamardin, K., Amir Sjarif, N. N., & Mohamed, N. (2019). Offline signature verification using deep learning convolutional Neural network (CNN) architectures GoogLeNet inception-v1 and inception-v3. *Procedia Computer Science*, 161, 475–483. <https://doi.org/10.1016/j.procs.2019.11.147>
- Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys*, 31(3), 264–323. <https://doi.org/10.1145/331499.331504>
- Jolliffe, I. T. (1998). Principal components. *Data Handling in Science and Technology*, 20(PART A), 519–556. [https://doi.org/10.1016/S0922-3487\(97\)80047-0](https://doi.org/10.1016/S0922-3487(97)80047-0)
- Khajwal, A. B., Cheng, C. S., & Noshadravan, A. (2023). Post-disaster damage classification based on deep multi-view image fusion. *Computer-Aided Civil and Infrastructure Engineering*, 38(4), 528–544. <https://doi.org/10.1111/mice.12890>
- Murdiaty, M., Angela, A., & Sylvia, C. (2020). Pengelompokkan Data Bencana Alam Berdasarkan Wilayah, Waktu, Jumlah Korban dan Kerusakan Fasilitas

- Dengan Algoritma K-Means. *Jurnal Media Informatika Budidarma*, 4(3), 744. <https://doi.org/10.30865/mib.v4i3.2213>
- Nolan, T. H., Richardson, S., & Ruffieux, H. (2025). Efficient Bayesian functional principal component analysis of irregularly-observed multivariate curves. *Computational Statistics and Data Analysis*, 203(November 2023), 108094. <https://doi.org/10.1016/j.csda.2024.108094>
- Pokhrel, S. (2024). No TitleΕΛΕΝΗ. *Ayaη*, 15(1), 37–48.
- Reis, H. C., Turk, V., Karacur, S., & Kurt, A. M. (2024). Integration of a CNN-based model and ensemble learning for detecting post-earthquake road cracks with deep features. *Structures*, 62(March), 106179. <https://doi.org/10.1016/j.istruc.2024.106179>
- Safitri, A. H., Wibowo Almais, A. T., Syauqi, A., & Melani, R. I. (2022). Pengujian Optimization dan Non-Optimization Query Metode Topsis untuk Menentukan Tingkat Kerusakan Sektor Bencana Alam. *Jurnal ELTIKOM*, 6(1), 89–99. <https://doi.org/10.31961/eltikom.v6i1.532>
- Satria, D., & Mushthofa. (2013). Perbandingan Metode Ekstraksi Ciri Histogram dan PCA untuk Mendeteksi Stoma pada Citra Penampang Daun Freycinetia Comparison of Histogram and PCA as Feature Extraction Methods in Detecting Stoma in Freycinetia Leaf Images Abstrak. *Jurnal Ilmu Komputer Agri-Informatika*, 2(1), 20–28. <https://journal.ipb.ac.id/index.php/jika/article/view/7966>
- Shao, Y., & Xu, L. (2025). Multimodal Natural Disaster Scene Recognition with Integrated Large Model and Mamba. *Applied Sciences (Switzerland)*, 15(3), 1–18. <https://doi.org/10.3390/app15031149>
- Song, Y., Wang, J., Ge, Y., B, L. L., Guo, J., Dong, Q., & Zhifang Liao. (2024). Medical image classification: Knowledge transfer via residual U-Net and vision transformer-based teacher-student model with knowledge distillation. *Jurnal Komunikasi Visual Dan Representasi Gambar*, 102. <https://doi.org/https://doi.org/10.1016/j.jvcir.2024.104212>
- Tahermanesh, S., Mokhtarzade, M., & Beirami, B. A. (2025). Enhancing change detection in multi-temporal optical images using a novel multi-scale deep learning approach based on LSTM. *Advances in Space Research*, 75(10), 7082–7111. <https://doi.org/10.1016/j.asr.2025.02.046>
- Tosta, T. A. A., de Faria, P. R., Neves, L. A., & do Nascimento, M. Z. (2019). Color normalization of faded H&E-stained histological images using spectral matching. *Computers in Biology and Medicine*, 111(June), 103344. <https://doi.org/10.1016/j.combiomed.2019.103344>
- Utomo, D. D., & Marta, F. Y. D. (2022). Dampak Bencana Alam Terhadap Perekonomian Masyarakat di Kabupaten Tanah Datar. *JURNAL TERAPAN PEMERINTAHAN MINANGKABAU*, 2(1), 92–97. <https://doi.org/10.33701/jtpm.v2i1.2395>

- Wang, S., Wang, Y., Chen, Y., Pan, P., Sun, Z., & He, G. (2018). Robust PCA Using Matrix Factorization for Background/Foreground Separation. *IEEE Access*, 6(2), 18945–18953. <https://doi.org/10.1109/ACCESS.2018.2818322>
- Wu, L., Tong, J., Wang, Z., Li, J., Li, M., Li, H., & Feng, Y. (2024). Post-flood disaster damaged houses classification based on dual-view image fusion and Concentration-Based Attention Module. *Sustainable Cities and Society*, 103(February), 105234. <https://doi.org/10.1016/j.scs.2024.105234>
- Xie, J., Wang, W., Xu, C., Fu, M., & Huang, W. (2024). Two-step feature extraction of acoustic emission signals for leakage detection of valves in gas pipelines. *International Journal of Pressure Vessels and Piping*, 212(PB), 105364. <https://doi.org/10.1016/j.ijpvp.2024.105364>
- Yasi, E., Shakib, T. U., Sharmin, N., & Rizu, T. H. (2024). Flood and Non-Flood Image Classification using Deep Ensemble Learning. *Water Resources Management*, 5161–5178. <https://doi.org/10.1007/s11269-024-03906-9>