SEGMENTASI CITRA BANGUNAN UNTUK MENGUKUR KERUSAKAN PASCA BENCANA ALAM MENGGUNAKAN *DUAL-ENCODER U-NET WITH RESNET-50 BACKBONE AND ATTENTION MECHANISM*

THESIS

Oleh:
REVALDI RAHMATMULYA
NIM. 230605210006

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

SEGMENTASI CITRA BANGUNAN UNTUK MENGUKUR KERUSAKAN PASCA BENCANA ALAM MENGGUNAKAN *DUAL-ENCODER U-NET WITH RESNET-50 BACKBONE AND ATTENTION MECHANISM*

THESIS

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

Oleh:
REVALDI RAHMATMULYA
NIM. 230605210006

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

SEGMENTASI CITRA BANGUNAN UNTUK MENGUKUR KERUSAKAN PASCA BENCANA ALAM MENGGUNAKAN *DUAL-ENCODER U-NET WITH RESNET-50 BACKBONE AND ATTENTION MECHANISM*

THESIS

Oleh:
REVALDI RAHMATMULYA
NIM. 230605210006

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 27 Mei 2025

Pembimbing I,	Pembimbing II,
Dr. Ir. Mokhamad Amin Hariyadi, M.T	Dr. Agung Teguh Wibowo Almais, M.T
NIP. 19670118 200501 1 001	NIP 19860301 2023321 1 016

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

Dr. Cahyo Crysdian
NIP. 19740424 200901 1 008

SEGMENTASI CITRA BANGUNAN UNTUK MENGUKUR KERUSAKAN PASCA BENCANA ALAM MENGGUNAKAN *DUAL-ENCODER U-NET WITH RESNET-50 BACKBONE AND ATTENTION MECHANISM*

THESIS

Oleh:
REVALDI RAHMATMULYA
NIM. 230605210006

Telah Dipertahankan di Depan Dewan Penguji Thesis
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Magister Komputer (M.Kom)
Tanggal: 27 Mei 2025

Susunan Dewan Penguji

Penguji I	: Dr. Cahyo Crysdian NIP. 19740424 200901 1 008	( )
Penguji II	: Dr. Ririen Kusumawati, S.Si., M.Kom NIP. 19720309 200501 2 002	( )
Pembimbing I	: Dr. Ir. Mokhamad Amin Hariyadi, M.T NIP. 19670018 200501 1 001	( )
Pembimbing II	: Dr. Agung Teguh Wibowo Almais, M.T NIP 19860301 2023321 1 016	( )

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

Dr. Cahyo Crysdian
NIP. 19740424 200901 1 008

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama	: Revaldi Rahmatmulya
NIM	: 230605210006
Fakultas / Program Studi	: Sains dan Teknologi / Magister Informatika
Judul Thesis	: SEGMENTASI CITRA BANGUNAN UNTUK MENGUKUR KERUSAKAN PASCA BENCANA ALAM MENGGUNAKAN *DUAL-ENCODER U-NET WITH RESNET-50 BACKBONE AND ATTENTION MECHANISM*

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

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

Malang,
Yang membuat pernyataan,

Revaldi Rahmatmulya
NIM. 230605210006

MOTTO

"Lies will always be a lies, but sometimes, they can be more noble than the truth"

HALAMAN PERSEMBAHAN

Dengan mengucap alhamdulillahi rabbil alamin, saya persembahkan Thesis ini untuk :

1. Kedua orang tua saya Wisnu Nurhariadi, S.T. dan Wiwik Hernamala yang selalu mendoakan saya dan mendukung saya hingga bisa mencapai di titik ini.
2. Guru – guru saya Gus Edy Nuswantoro, Ustad Heri Wahyu Purnomo, Habib Sholeh, Gus Ulin, Kyai Mukhtar, Kyai Ikhsan, karena telah memberikan ilmu yang manfaat dunia akhirat serta senantiasa mendoakan saya.
3. Almarhum guru – guru saya (Alm) Kyai Hamid, (Alm) Habib Abdullah bil faqih, (Alm) Habib Alwi, (Alm) Habib Salim, (Alm) Habib Abdurrohman, (Alm) Kyai Kholil, (Alm) Kyai Abdul faqih, (Alm) Kyai As'ad, (Alm) Kyai Said, (Alm) Kyai Su'aidi, (Alm) Asrori, (Alm) Kyai Zainullah, (Alm) Kyai Jamal, (Alm) Kyai Muttaqin, (Alm) Kyai Ismail, (Alm) Kyai Munawwir, (Alm) Kyai Arwani, (Alm) Kyai Muntaha, (Alm) Ustad Musa, (Alm) Kyai Sulaiman, (Alm) Kyai Hanan, (Alm) Kyai Abu Amar, (Alm) Kyai Ahmad Shiddiq, (Alm) Kyai Musta'in, (Alm) Mbah Kyai Khusnan, (Alm) Mbah Kyai Thoyyib, (Alm) Mbah Kyai Jazuli, karena berkat barokah sanad keilmuan beliau dan rahmat Allah SWT saya dapat sampai di titik ini.
4. Kedua adik tersayang saya Kyla Anadilareva dan Ahza Revandila.
5. Teman – teman seperjuangan saya Abdurrozzaaq Ashshiddiqi Zuhri, Achmad Fahreza Alif Pahlevi, Aldian Faizzul Anwar, Aji Bagas Prakasa, Fahrendra Khoirul Ihtada, Galan Ramadan Harya Galib karena selalu mendukung, membantu, dan menemani saya baik dalam keadaan susah maupun senang.

KATA PENGANTAR

Assalamualaikum Warahmatullahi Wabarakatuh.

Segala puji hanya milik Allah Subhanahu Wa Ta'ala atas segala nikmat dan kasih sayang-Nya yang telah memudahkan penulis untuk menyelesaikan thesis yang berjudul "Segmentasi Citra Bangunan untuk Mengukur Kerusakan Pasca Bencana Alam Menggunakan Dual-Encoder U-Net With ResNet-50 Backbone and Attention Mechanism". Semoga shalawat dan salam senantiasa terlimpah kepada Nabi Muhammad Sallalahu 'Alaihi wa Sallam. Dan semoga kita semua mendapat syafaatnya di hari kiamat nanti, Aamiin.

Penulis mengucapkan rasa syukur dan terima kasih yang tak terhingga kepada semua pihak-pihak yang selalu memberikan bantuan dan motivasi kepada penulis untuk dapat menyelesaikan Thesis ini. Ucapan ini penulis sampaikan kepada:

1. Prof. Dr. H. M. Zainuddin, M.A. selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Prof. Dr. Hj. Sri Hariani, M.Si. selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Cahyo Crysdian. selaku Ketua Program Studi Magister Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Dr. Ir. Mokhamad Amin Hariyadi, M.T. selaku dosen pembimbing I dan Dr. Agung Teguh Wibowo Almais, M.T. selaku dosen pembimbing II yang telah memberikan bantuan dan arahan kepada penulis, sehingga bisa menuntaskan Thesis ini.

5. Dr. Cahyo Crysdian. selaku dosen penguji I dan Dr Ririen Kusumawati, S.Si., M.Kom. selaku dosen penguji II yang telah menguji serta memberikan masukan sehingga penulis dapat menuntaskan Thesis dengan baik.
6. Segenap Dosen, Admin, Laboran dan Mahasiswa Program Studi Magister Informatika yang telah memberikan banyak dukungan dan bimbingan selama pengerjaan Thesis ini.
7. Papa, Mama, serta adik - adik saya yang selalu memberikan dukungan dan motivasi untuk terus berusaha, dan doa yang tak putus-putusnya selalu disampaikan agar dapat menuntaskan Thesis ini dengan lancar dan baik.

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

Wassalamualaikum Warahmatullahi Wabarakatuh.

Malang, 26 Mei 2025

Penulis

DAFTAR ISI

HALAMAN PERSETUJUAN ..... **iii**
HALAMAN PENGESAHAN ..... **iv**
PERNYATAAN KEASLIAN TULISAN ..... **v**
MOTTO ..... **vi**
HALAMAN PERSEMBAHAN ..... **vii**
KATA PENGANTAR ..... **viii**
DAFTAR ISI ..... **x**
DAFTAR TABEL ..... **xii**
DAFTAR GAMBAR ..... **xiii**
ABSTRAK ..... **xv**
ABSTRACT ..... **xvi**
خلاصة ..... **xvii**
BAB I PENDAHULUAN **1**
1.1 Latar Belakang 1
1.2 Pernyataan Masalah 5
1.3 Batasan Masalah 5
1.4 Tujuan Penelitian 6
1.5 Manfaat Penelitian 6
BAB II STUDI PUSTAKA **7**
2.1 Segmentasi Bangunan 7
2.2 Segmentasi Kerusakan Bangunan 10
2.3 Kerangka Teori 11
BAB III METODOLOGI PENELITIAN **15**
3.1 Kerangka Konsep 15
3.2 Desain Penelitian 16
3.2.1 Data Collection 17
3.2.2 Desain sistem 17
3.2.2.1 Input Data 18
3.2.2.2 Preprocessing 19
3.2.2.3 Patching Data 20
3.2.2.4 Augmentasi Data 21
3.2.2.5 Augmentasi Data 21
3.2.2.6 Unet with ResNet50 Backbone and Attention Mechanism 22
3.2.2.6.1 Encoder 23
3.2.2.6.2 Convolution Layer 24
3.2.2.6.3 Pooling Layer 25
3.2.2.6.4 Bottleneck 25
3.2.2.6.5 Decoder 26
3.2.2.6.6 CBAM (Convolutional Block Attention Module) 27

3.2.2.6.7 Pixelwise 29
3.3 Uji Coba 30
3.3.1 Skenario Pengujian 30
3.3.2 Training Model Convoutional Neural Network 32
3.3.3 *Reassembly Segmented Image* 34
3.3.4 Hasil Awal 36
3.4 Evaluasi 40
3.5 Intrument Penelitian 41
BAB IV U-NET WITH RESNET-50 BACKBONE AND ATTENTION MECHANISM 43
4.1. Model Arsitektur 43
4.2. Uji Coba 44
4.2.1 Skenario 1 44
4.2.2 Skenario 2 47
4.2.3 Skenario 3 50
4.2.4 Skenario 4 53
4.3. Analisis Awal 56
4.3.1 Analisis Pengaruh Loss Function 57
4.3.2 Analisis Pengaruh Optimizer 58
4.3.3 Analisis Kinerja Per Kelas 58
4.4 Kesimpulan Analisis Awal 60
BAB V DEFAULT UNET WITH PRETRAINED ENCODER 62
5.1 Model Arsitektur 62
5.2. Uji Coba 63
5.2.1 Skenario 5 63
5.2.2 Skenario 6 66
5.2.2 Skenario 7 69
5.2.2 Skenario 8 72
5.4. Analisis Awal 75
5.4.1 Analisis Pengaruh Loss Function 75
5.4.2 Analisis Pengaruh Optimizer 76
5.4.3 Analisa Kinerja Per Kelas 77
5.4 Kesimpulan Analisis Awal 79
BAB VI PEMBAHASAN 81
6.1 Perbandingan Kinerja 81
6.2 Perbandingan Kinerja Analisis Confusion Matrix 82
6.3 Analisis Dampak Mekanisme Atensi (CBAM) 84
BAB VII PENUTUP 96
7.1 Kesimpulan 96
7.2 Saran 98
DAFTAR PUSTAKA 101

DAFTAR TABEL

Tabel 2.1. Rangkuman kerangka teori 13
Tabel 3.1 Label Kelas 17
Tabel 3.2 Contoh data citra satelit beserta *ground truth* 19
Tabel 3.3 Skenario pertama 31
Tabel 3.4 Skenario kedua 31
Tabel 3.5 Skenario ketiga 31
Tabel 3.6 Kombinasi Skenario *Proposed Model* 32
Tabel 3.7 Kombinasi Skenario dual encoder U-Net 32
Tabel 3.8 Definisi *Hyperparameter* 34
Tabel 3.9 Instrumen Penelitian 41
Tabel 4.1. Tabel Kelompok Skenario 1 44
Tabel 4.2 Evaluasi Model per Kelas skenario 1 46
Tabel 4.3 Evaluasi Rerata Skenario 1 46
Tabel 4.4 Evaluasi per kelas skenario 2 49
Tabel 4.5 Evaluasi rerata skenario 2 49
Tabel 4.6 Evaluasi model per kelas skenario 3 52
Tabel 4.7 Evaluasi rerata skenario 3 52
Tabel 4.8 Evaluasi model per kelas skenario 4 55
Tabel 4.9 Evaluasi rerata skenario 4 55
Tabel 4.10 Rangkuman hasil pengujian kelompok skenario 1 56
Tabel 5.1 Kombinasi skenario kelompok 2 63
Tabel 5.2 Evaluasi model per kelas skenario 5 65
Tabel 5.3 Evaluasi rerata skenario 5 65
Tabel 5.4 Evaluasi model per kelas skenario 6 68
Tabel 5.5 Evaluasi rerata skenario 6 68
Tabel 5.6 Evaluasi moddel per kelas skenario 7 71
Tabel 5.7 Evaluasi rerata skenario 7 71
Tabel 5.8 Evaluasi model per kelas skenario 8 74
Tabel 5.9 Evaluasi rerata skenario 8 74
Tabel 5.10 Rangkuman hasil pengujian kelompok skenario 2 75
Tabel 6.1 Perbandingan Kinerja Model Usulan (S4) vs Model Pembanding (S6) 81

DAFTAR GAMBAR

Gambar 2.1. Kerangka Teori 12
Gambar 3.1. Kerangka konsep penelitian 15
Gambar 3.2. Desain penelitian 16
Gambar 3.3. Desain sistem 18
Gambar 3.4 Alur prepocessing 20
Gambar 3.5 Patching Image 20
Gambar 3.6 Augmentasi data 21
Gambar 3.7 Model Arsitektur 22
Gambar 3.8. ResNet 50 Backbone Sumber : (Sharma *et al*., 2025) 24
Gambar 3.9. CBAM Module Sumber : (Xu *et al*., 2023) 27
Gambar 3.10 Pixelwise 30
Gambar 3.11 Alur training 33
Gambar 3.12. Accuracy and Loss Training 36
Gambar 3.13. Gambar Hasil Prediksi Model 37
Gambar 3.14. Evaluation Matrics 38
Gambar 3.15. Confusion Matrix 39
Gambar 4.1 Proposed model 43
Gambar 4.2 Grafik Loss Skenario 1 44
Gambar 4.3 Hasil Segmentasi Skenario 1 45
Gambar 4.4 *Confussion Matrix* Skenario 1 46
Gambar 4.5 Grafik loss skenario 2 47
Gambar 4.6 Hasil segmentasi skenario 2 48
Gambar 4.7 *Consfussion Matrix* skenario 2 49
Gambar 4.8 Grafik loss skenario 3 50
Gambar 4.9 Hasil segmentasi skenario 3 51
Gambar 4.10 *Confussion matrix* skenario 3 52
Gambar 4.11 Grafik loss skenario 4 53
Gambar 4.12 Hasil segmentasi skenario 4 54
Gambar 4.13 *Confussion matrix* skenario 4 55
Gambar 5.1 Model pembanding 62
Gambar 5.2 Grafik loss skenario 5 63
Gambar 5.3 Hasil segmentnasi skenario 5 64
Gambar 5.4 *Confussion matrix* skenario 5 65
Gambar 5.5 Grafik loss skenario 6 66
Gambar 5.6 Hasil segmentasi skenario 6 67
Gambar 5.7 *Confussion matrix* skenario 6 68
Gambar 5.8 Grafik loss skenario 7 69
Gambar 5.9 Hasil segmentasi skenario 7 70
Gambar 5.10 *Confussion matrix* skenario 7 71
Gambar 5.11 Grafik loss skenario 8 72
Gambar 5.12 Hasil segmentasi skenario 8 73
Gambar 5.13 *Confussion matrix* skenario 8 74

Gambar 6.1 Komparasi Confussion matrix skenario 4 (kiri) dan confussion matrix skenario 6 (kanan) 82
Gambar 6.2 Hasil segmentasi skenario 4 85
Gambar 6.3 Hasil Segmentasi skenario 6 85

ABSTRAK

Rahmatmulya, Revaldi. 2025. **Segmentasi Citra Bangunan untuk Mengukur Kerusakan Pasca Bencana Alam Menggunakan *Dual-Encoder U-Net With ResNet-50 Backbone and Attention Mechanism***. Thesis. Program Studi Magister Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Ir. Mokhamad Amin Hariyadi, M.T. (II) Dr. Agung Teguh Wibowo Almais, M.T.

Kata kunci: *Convolutional Neural Network, Machine Learning, Segmentasi.*

Penilaian tingkat kerusakan bangunan pasca bencana alam merupakan langkah krusial untuk respons darurat dan perencanaan rekonstruksi. Metode manual yang ada saat ini seringkali lambat, subjektif, dan memakan sumber daya besar. Penelitian ini bertujuan untuk mengembangkan dan mengevaluasi model deep learning berbasis U-Net untuk segmentasi kerusakan bangunan secara otomatis dan akurat menggunakan citra satelit. Model yang diajukan mengintegrasikan arsitektur U-Net dengan backbone ResNet50 pretrained, pendekatan dual encoder dengan input ganda (citra pra dan pasca bencana), serta mekanisme atensi Convolutional Block Attention Module (CBAM) pada jalur decoder. Kinerja model usulan ini dibandingkan secara komprehensif dengan model pembanding (U-Net ResNet50 tanpa CBAM) melalui serangkaian skenario pengujian yang melibatkan variasi loss function (Crossentropy, Dual Focal) dan optimizer (Adam, RMSprop) pada dataset Xview2. Hasil penelitian menunjukkan bahwa model usulan dengan konfigurasi terbaik (Dual Focal Loss + RMSprop) mencapai kinerja superior dengan nilai Overall Accuracy (OA) 0.6009, mean F1-Score (mF1) 0.5145, dan mean Intersection over Union (mIoU) 0.3939. Angka ini menunjukkan peningkatan mIoU sebesar 8.24% dibandingkan model pembanding terbaik. Analisis membuktikan bahwa penambahan CBAM secara signifikan meningkatkan kemampuan model dalam menghasilkan batas segmentasi yang lebih jelas dan akurasi klasifikasi tingkat kerusakan yang lebih baik, meskipun identifikasi 'Minor Damage' tetap menjadi tantangan. Penelitian ini menyimpulkan bahwa model U-Net yang dimodifikasi dengan ResNet50, dual input, dan CBAM merupakan pendekatan yang efektif dan menjanjikan untuk otomatisasi penilaian kerusakan bangunan pasca bencana.

ABSTRACT

Rahmatmulya, Revaldi. 2025. **Post-Disaster Building Damage Segmentation Using Dual-Encoder U-Net With ResNet-50 Backbone and Attention Mechanism.** Thesis. Master Program of Computer Science, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Advisors: (I) Dr. Ir. Mokhamad Amin Hariyadi, M.T. (II) Dr. Agung Teguh Wibowo Almais, M.T

Post-disaster building damage assessment is a crucial step for emergency response and reconstruction planning. Current manual methods are often slow, subjective, and resource-intensive. This research aims to develop and evaluate a U-Net based deep learning model for automatic and accurate building damage segmentation using satellite imagery. The proposed model integrates the U-Net architecture with a pretrained ResNet50 backbone, a dual encoder approach with dual inputs (pre- and post-disaster images), and the Convolutional Block Attention Module (CBAM) attention mechanism in its decoder path. The performance of the proposed model was comprehensively compared against a baseline model (U-Net ResNet50 without CBAM) through a series of test scenarios involving variations of loss functions (Crossentropy, Dual Focal) and optimizers (Adam, RMSprop) on the Xview2 dataset. The results indicate that the proposed model with its best configuration (Dual Focal Loss + RMSprop) achieved superior performance, reaching an Overall Accuracy (OA) of 0.6009, a mean F1-Score (mF1) of 0.5145, and a mean Intersection over Union (mIoU) of 0.3939. This represents an 8.24% increase in mIoU compared to the best baseline model. Analyses from result confirm that the addition of CBAM significantly enhances the model's ability to produce clearer segmentation boundaries and improve damage level classification accuracy, although identifying 'Minor Damage' remains a challenge. This study concludes that the modified U-Net model utilizing ResNet50, dual input, and CBAM is an effective and promising approach for automating post-disaster building damage assessment.

Keywords: *Convolutional Neural Network, Machine Learning, Segmentation*

خلاصة

رحمموليا ، ريفالدي. 2025. **تجزئة صور البناء لتحديد مدى الضرر الناجم عن الكوارث الطبيعية باستخدام شبكة *U-Net* ذات التشفير المزدوج مع العمود الفقري *ResNet-50* وآلية** الانتباه. اطروحه. برنامج دراسة ماجستير المعلوماتية ، كلية العلوم والتكنولوجيا ، جامعة مولانا مالك إبراهيم الإسلامية الحكومية ، ملانج. المشرف: (I) د. محمد أمين حرياديي ، ماجستير (II) د. أغونغ تيغوه وييو أمليس ، ماجستير

الكلمات المفتاحية: *الشبكة العصبية التلافيفية ، التعلم الآلي ، التجزئة.*

يعد تقييم مدى الأضرار الناجمة عن المباني بعد وقوع كارثة طبيعية خطوة حاسمة للاستجابة للطوارئ والتخطيط لإعادة الإعمار. غالبا ما تكون الأساليب اليدوية الموجودة اليوم بطيئة وذاتية وتستهلك موارد كبيرة. تهدف هذه الدراسة إلى تطوير وتقييم نموذج التعلم العميق القائم على U-Net للتجزئة التلقائية والدقيقة لأضرار المباني باستخدام صور الأقمار الصناعية. يدمج النموذج المقترح بنية U-Net مع العمود الفقري ResNet50 المدرب مسبقا ، ونهج التشفير المزدوج مع المدخلات المزدوجة (صور ما قبل الكارثة وبعدها) ، وآلية توهين وحدة انتباه الكتلة التلافيفية (CBAM) على مسار وحدة فك التشفير. تمت مقارنة أداء هذا النموذج المقترح بشكل شامل مع النموذج المقارن (U-Net ResNet50 بدون CBAM) من خلال سلسلة من سيناريوهات الاختبار التي تتضمن اختلافات في وظيفة الخسارة (Crossentropy ، Dual Focal) والمحسن (Adam ، RMSprop) على مجموعة بيانات Xview2. أظهرت النتائج أن النموذج المقترح بأفضل تكوين (الخسارة البؤرية المزدوجة + RMSprop) حقق أداء متفوقا بقيمة الدقة الكلية 0.6009 (OA) ، ومتوسط F1-Score (mF1) 0.5145 ، ومتوسط تقاطع على الاتحاد 0.3939 (mIoU). يظهر هذا الرقم زيادة في mIoU بنسبة 8.24٪ مقارنة بأفضل نموذج مقارنة. يثبت التحليل أن إضافة CBAM يحسن بشكل كبير من قدرة النموذج على إنتاج حدود تجزئة أكثر وضوحا ودقة تصنيف أفضل لمستوى الضرر ، على الرغم من أن تحديد "الضرر الطفيف" لا يزال يمثل تحديا. وخلصت الدراسة إلى أن نموذج U-Net المعدل باستخدام ResNet50 والمدخلات المزدوجة و CBAM هو نهج فعال وواعد لأتمتة تقييم أضرار المباني بعد الكارثة.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kerusakan bangunan akibat bencana alam merupakan masalah serius yang dihadapi manusia. Hal ini disebabkan karena dampak ekonomi yang sangat besar, kebutuhan berbagai bantuan setelah bencana, serta rekonstruksi bangunan kembali setelah bencana alam (Wang *et al*., 2025) .Penentuan Tingkat kerusakan diperlukan diperlukan sebelum melakukan rekonstruksi ulang. Dalam beberapa kasus, kerusakan ringan hanya akan memerlukan sedikit perbaikan dan biaya yang tidak besar, namun kerusakan juga bisa parah dan memerlukan biaya yang tinggi untuk melakukan perbaikan bahkan rekonstruksi ulang (Pramono *et al*., 2022).

Penentuan Tingkat seringkali dilakukan secara manual, hal ini tentu tidak efektif dan memakan banyak waktu. Pendekatan semi otomatis dilakukan dengan memanfaatkan *remote sensing* (Hatic *et al*., 2023). Pendekatan yang lebih efektif dilakukan untuk mengubah dari semi otomatis menjadi otomatis, untuk mengatasi keterbatasan ini, model pengembangan *model deep learning* dibutuhkan untuk secara otomatis bisa menentukan tingkat kerusakan bangunan pasca bencana alam (Zaryabi *et al*., 2022a).

Pengukuran tingkat kerusakan bangunan pasca bencana alam secara seringkali dilakukan dengan cara manual, hal ini tentu tidak efektif dalam penentuannya. Terdapat beberapa faktor yang menjadi penyebab mengapa penentuan secara manual terbilang tidak efektif, yang pertama memerlukan banyak tenaga dan waktu. Butuhnya banyak orang dikarenakan kerusakan yang terjadi

berdampak tidak hanya satu bangunan, hal ini tentu saja membuat waktu penentuan menjadi semakin panjang. Akibatnya bisa menghambat respon dari tim penyelamat (Zaryabi *et al*., 2022b). Berikutnya yang kedua yaitu kebergantungan kepada ahli dalam penentuan tingkat kerusakan bangunan pasca bencana alam juga menjadi salah satu faktor. Dibutuhkan sorang ahli agar penentuan tingkat kerusakan menjadi konsisten, namun keterbatasan sang ahli menjadi penghambat dalam penentuannya (Pahlavani *et al*., 2006).

Untuk mengatasi keterbatasan ini, model pengembangan *model deep learning* dibutuhkan untuk secara otomatis bisa menentukan tingkat kerusakan bangunan pasca bencana alam (Zaryabi *et al*., 2022b). Penulis menggunakan model *deep learning* dengan metode *covolutional neural network* menggunakan arsitektur yang telah penulis rancang dalam pendekatan komputasi untuk menentukan tingkat kerusakan bangunan pasca bencana alam. Dengan memanfaatkan data – data berupa citra bangunan yang rusak akibat bencana alam, model *deep learning* ini memliki potensi untuk menentukan tingkat kerusakan pasca bencana alam yang efektif dan efisien. Dengan demikian, diharapkan dampak – dampak yang terjadi akibat dilakukannya cara manual dalam penentuan tingkat kerusakan bangunan pasca bencana alam tidak sampai terjadi.

Salah satu strategi deep learning yang bisa diterapkan adalah dengan melakukan segmentasi semantik. Segmentasi semantik yang kadang juga disebut sebagai segmentasi citra berbasis konten – adalah masalah dalam bidang *computer vision* di mana tugas utamanya adalah menentukan kelas semantik dari setiap piksel dalam sebuah gambar. Biasanya, masalah ini diselesaikan dengan pendekatan

supervised learning, yaitu dengan mengandalkan dataset gambar yang telah diberi anotasi pada tingkat piksel, dan melatih model machine learning menggunakan data tersebut untuk menjalankan tugas segmentasi (Csurka *et al*., 2023).

Salah satu model machine learning yang populer digunakan adalah U-Net. U-Net yang berbasis *fully convolutional network* merupakan model lanjutan dari *deep learning* yang menjadi salah satu standart struktur *convolutional neural network* dalam perkembangan pendeteksian (Khelifi and Mignotte., 2020). U-Net menggabungkan informasi tingkat tinggi yang dapat digunakan untuk membedakan kategori, serta informasi beresolusi tinggi yang memberikan dasar akurat untuk pemetaan lokasi dan segmentasi, hal ini sesuai untuk mendeteksi Tingkat kerusakan bangunan.

Al-Qur'an dalam surat Ar – Rum ayat 41 berbunyi:

ظَهَرَ ٱلْفَسَادُ فِى ٱلْبَرِّ وَٱلْبَحْرِ بِمَا كَسَبَتْ أَيْدِى ٱلنَّاسِ لِيُذِيقَهُم بَعْضَ ٱلَّذِى عَمِلُوا۟ لَعَلَّهُمْ يَرْجِعُونَ

"*Telah nampak kerusakan di darat dan di laut disebabkan karena perbuatan tangan manusia, supaya Allah merasakan kepada mereka sebahagian dari (akibat) perbuatan mereka, agar mereka kembali (ke jalan yang benar)." (QS. Ar – Ruum: 41)*

Tafsir ayat ini yang berasal dari Kementrian Agama RI berbunyi "Bila pada ayat-ayat sebelumnya Allah menjelaskan sifat buruk orang musyrik mekah yang menuhankan hawa nafsu, melalui ayat ini Allah menegaskan bahwa kerusakan di bumi adalah akibat mempertuhankan hawa nafsu. Telah tampak kerusakan di darat dan di laut, baik kota maupun desa, disebabkan karena perbuatan tangan manusia yang dikendalikan oleh hawa nafsu dan jauh dari tuntunan fitrah. Allah menghendaki agar mereka merasakan sebagian dari akibat perbuatan buruk mereka

agar mereka kembali ke jalan yang benar dengan menjaga kesesuaian perilakunya dengan fitrahnya. Perbuatan buruk manusia akan mendatangkan azab sebagaimana azab yang telah menimpa umat-umat terdahulu. Azab itu juga akan datang kepada umat-umat di masa sekarang maupun yang akan datang sebagai sunatullah jika mereka memiliki karakter yang sama. Karena itu, katakanlah, wahai nabi Muhammad, kepada siapa saja yang meragukan hakikat ini, 'bepergianlah di muka bumi, di mana saja yang bisa kamu jangkau, lalu lihatlah bagaimana kesudahan orang-orang dahulu yang dihancurkan akibat perilaku buruk mereka. Itu semua karena kebanyakan dari mereka adalah orang-orang yang mempersekutukan Allah dan menuhankan hawa nafsu" (Kemenag RI, 2016)

Tafsir QS. Ar-Ruum ayat 41 mengajarkan kepada kita bahwa salah satu penyebab datangnya bencana adalah karena perbuatan manusia itu sendiri. Oleh karena itu, kita tetap harus sabar, tawakal, dan berikhtiar dalam menghadapi bencana yang datang. Salah satu bentuk ikhtiar kita adalah memanfaatkan teknologi yang ada untuk mengatasi dampak yang sudah diakibatkan oleh bencana. Salah satunya yaitu dengan memanfaatkan *U-Net* sebagai alat dalam otomatsi penentuan tingkat kerusakan bangunan bencana alam.

Penelitian oleh Wu *et al* (2021) mengenai penentuan tingkat kerusakan bangunan pasca bencana alam pernah dilakukan dengan menggunakan metode *U-Net with SeResNeXt Backbone*. Dalam penelitian ini, peneliti menggunakan data *remote sensing*. Hasil percobaan menunjukkan bahwa model mampu menghasilkan nilai F1 score dalam persen sebesar 79.2. Hal ini menunjukkan bahwa *U-Net* dengan menggunakan *backbone* dapat memberikan hasil yang memuaskan.

Berikutnya penelitian oleh Weber and Kan (2020) mengenai segmentasi kerusakan bangunan menggunakan metode *covolutional neural network* dengan arsitektur *U-net dual encoder*. Dalam penelitian ini, model *U-Net* mampu menghasilkan nilai *F1 score* dalam persen sebesar 74.1. Dalam penelitian ini dapat diisimpulkan bahwa *U-Net* dengan menggunakan arsitektur *dual encoder* mampu memberikan nilai *F1 score* yang baik.

Dengan dasar masalah yang telah diuraikan, alasan pemilihan U-Net sebagai model otomasi penentuan kerusakan bangunan didasarkan pada dua penelitian yang berhasil meraih nilai F1 Score yang baik. Dalam penelitian ini, *U-Net* akan dimodifikasi sehingga dapat menggunakan *backbone* dengan *pretrained model*, menerapkan *dual encoder* dengan dua input data, dan menerapkan *attention mechanism* pada proses *decoder*. Sebagian dari metode yang diajukan ini adalah gabungan dari dua penelitian yang telah disebutkan di atas.

1.2 Pernyataan Masalah

Bagaimana peningkatan kinerja model U-Net dalam pemanfaatan *pretrained backbone, dual encoder with dual input,* dan *attention mechanism* ?

1.3 Batasan Masalah

1. Data yang digunakan adalah citra satelit dari xBD Dataset.
2. Data citra wilayah yang digunakan adalah dampak dari bencana tornado, tsunami, kebakaran hutan, banjir, dan badai angin.

1.4 Tujuan Penelitian

Mengetahui seberapa bagus peningkatan kinerja model dalam melakukan segmentasi citra bangunan pasca bencana alam.

1.5 Manfaat Penelitian

1. Bermanfaat bagi tim SAR untuk menentukan prioritas evakuasi korban bencana alam berdasarkan keparahan Tingkat kerusakan bangunan suatu wilayah.

2. Bermanfaat bagi pemerintah dalam penentuan anggaran bantuan renovasi bangunan terdampak bencana alam.

BAB II

STUDI PUSTAKA

2.1 Segmentasi Bangunan

Penelitian yang dilakukan oleh Cai *et al* (2024) membahas tentang segmentasi semantik pada citra penginderaan jauh beresolusi tinggi, yang memang memiliki tantangan tersendiri karena kerumitan objek dan latar belakangnya. Dalam studi ini, peneliti mengusulkan sebuah model baru bernama Res50-SimAM-ASPP-Unet, yang merupakan hasil modifikasi dari arsitektur U-Net. Model ini menggabungkan ResNet50 sebagai backbone untuk ekstraksi fitur yang lebih dalam, kemudian menambahkan SimAM, sebuah modul attention ringan yang tidak menambah parameter tetapi mampu menyoroti area penting pada citra. Tak hanya itu, mereka juga menyisipkan ASPP *(Atrous Spatial Pyramid Pooling)* untuk menangkap informasi spasial pada berbagai skala, yang terbukti berguna dalam menangani objek berukuran beragam. Proses pelatihan dilakukan pada dataset ISPRS Vaihingen dan Potsdam, dengan ukuran citra 512×512 piksel, menggunakan *optimizer Adam* dan *cross-entropy loss* selama 100 epoch. Hasilnya cukup mengesankan: model ini mencatatkan mIoU sebesar 83,7% dan OA 90,6% di dataset Vaihingen, serta mIoU 85,3% dan OA 91,4% di dataset Potsdam. *F1-score* masing-masing juga tinggi, yakni 88,1% dan 89,0%. Berdasarkan performa tersebut, peneliti menyimpulkan bahwa kombinasi ResNet50, SimAM, dan ASPP mampu meningkatkan ketepatan segmentasi objek urban secara signifikan, sekaligus menjaga efisiensi dalam proses pelatihan.

Selanjutnya penelitian oleh Yuanjun *et al* (2023), penelitian ini bertujuan untuk melkukan semantic segmentation terhadap remote sensing images. Dataset yang digunakan pada penelitian ini berasal dari WHDLD Dataset, LanCover.ai dataset, dan Massachusetts dataset. Metode yang digunakan pada penelitian ini yaitu CTMU-Unet yaitu modifikasi arsitektur Unet dengan menambahkan modul CTM didalamnya. Hasil dari penelitian ini mampu memberikan akurasi dalam persen 92.06 FWIOU, 84.6 mIoU, 90.74 MPA, dan 91.24 F1 Score pada dataset LandCover.ai dataset. Berikutnya hasil yang memuaskan juga berhasil didapatkan pada dataset Massachusetts-building dataset yaitu dalam persen 88.2 FWIOU, 81.92 mIoU, 88.3 MPA dan 89.73 F1 Score.

Selanjutnya penelitian oleh Liu Z *et al* (2020), penelitian ini bertujuan untuk dapt melakukan segmentasi terhadap bangunan pada satellite imagery. Dataset yang digunakan pada penelitian ini adalah gambar yang berasal dari open source dataset Kaggle. Metode yang diajukan oleh peneliti adalah menggunakan Unet dengan memanfaatkan pre-trained ResNet34. Hasil dari penelitian ini dalam persen mampu mendapatkan 83 mIoU yang sudah dapat melakukan segmentasi pada bangunan dengan baik.

Selanjutnya penelitian oleh Wenmei *et al* (2022), penelitian ini bertujuan dalam melakukan segmentasi pada rentang area genangan air. Penelitian ini menggunakan dataset yang didapatkan dari dua satelit *Sentinel-1A* untuk mendapatkan data gambar wilayah tergenang air. Penelitian ini menggunakan Unet dan Unet-CBAM yang nantinya keduanya akan dikomparasi dengan traditional *SAR image water extraction methods*. Hasil dari penelitian ini menunjukkan bahwa

Unet mampu memberikan hasil yang memuaskan dalam *recall, precision,* dan *F1-parameter*, tetapi Unet tidak baik dalam mendeteksi garis dan rentang genangan air yang kecil. Sedangkan metode yang diajukan yaitu Unet-CBAM mampu memberikan peningkatan hasil baik dalam *recall, precision,* dan *F1-parameter*. Nilai yang dihasilkan oleh Unet-CBAM dalam persen yaitu *recall* 92.4 (naik 0.8%), *precision* 83.4 (naik 0.3%), dan *F1 Score* (87.7) naik 0.6%.

Selanjutnya penelitian oleh Sun *et al* (2023), penelitian ini bertujuan dalam melakukan segmentasi pada *high-resolution remote sensing images*. Dataset yang digunakan dalam penelitian ini adalah gambar *remote sensing* dari ISPRS Vaihingen, Postdam, dan LoveDA datasets. Metode yang digunakan dalam penelitian ini mengajukan metode *semantic segmentation network based on sparse self-attention and feature alihnment* (SAANet). Lebih jelasnya, pada metode ini *self-attention module* (SPAM) akan membagi dan Menyusun ulang fitur map pada position dimension serta melakukan *positions attention operations* (PAM) pada sub region yang sudah dikembalikan dan disusun ulang. Sedangkan *sparse channel self-attention module* (SCAM) akan mengumpulkan, Menyusun ulang fitur map pada channel dimension serta melakukan *channel attention operations* (CAM) pada sub channel yang sudah dikembalikan dan disusun ulang. Pengujian dari SAANNet akan dibandingkan dengan DeepLabv3+, HRNet, EMANet, PSPNet, CCNNet, FCN, dan DANet. Hasil pengujian akan melihat seberapa besar akurasi segmentasi terhadap permukaan, bangunan, low vegetation, pohon, mobil, background. Lalu matrik pengukuran akan menggunakan mIoU, mF1, dan OA. Hasil dari penelitian ini terbagi menjadi 3 sesuai penggunaan dataset. Pertama pada penggunaan dataset

ISPRS Vaihingen SAANet mampu mendapatkan akurasi tertinggi pada semua segmentasi kecuali permukaan, serta mendapatkan akurasi tertinggi diantara semua matrik pengukuran. Kedua pada dataset ISPRS Postdam, kali ini SAANet mampu menghasilkan akurasi tertinggi pada segmentasi maupun matrtik pengukuran. Ketiga pada dataset LoveDA SAANet hanya mampu meraih nilai akurasi tinggi pada segmentasi background. Sedangkan seluruh matrik pengukuran mendapatkan akurasi tertinggi.

2.2 Segmentasi Kerusakan Bangunan

Penelitian oleh Wu *et al* (2021), penelitian ini bertujuan untuk membangun model yang dapat melakukan segmentasi tingkat kerusakan kepada bangunan pada data gambar remote sensing. Dataset yang digunakan pada penelitian ini didapatkan dari Maxar Open Data Program. Metode yang digunakan pada penelitian ini adalah UNet dengan menambahkan attention module serta menggunakan berbagai macam backbone seperti Dual Path Net (DPN), SEResNeXt backbone, Squeeze-and-Excitation Networks Backbone (SENet), RestNet34 Backbone. Hasil dari penelitian ini, dari semua backbone yang diuji, SEResNeXt backbone mendapatkan hasil yang terbaik dengan nilai akurasi dalam persen 79.2 F1 Score.

Penelitian yang dilakukan oleh Weber and Kan (2020), penelitian ini bertujuan untuk dapat melakukan segemntasi terhadap bangunan untuk menentukan kerusakannya. Dataset yang digunakan pada penelitian ini berasal dari xBD Dataset yang berisi gambar remote sensing. Metode yang digunakan dalam penelitian ini adalah menggunakan Unet dengan menambahkan multi-temporal fusion dan

menggunakan Mask R-CNN sebagai backbone. Model pada penelitian ini mampu menghasilkan nilai akurasi dalam persen sebesar 73.8 F1 Score.

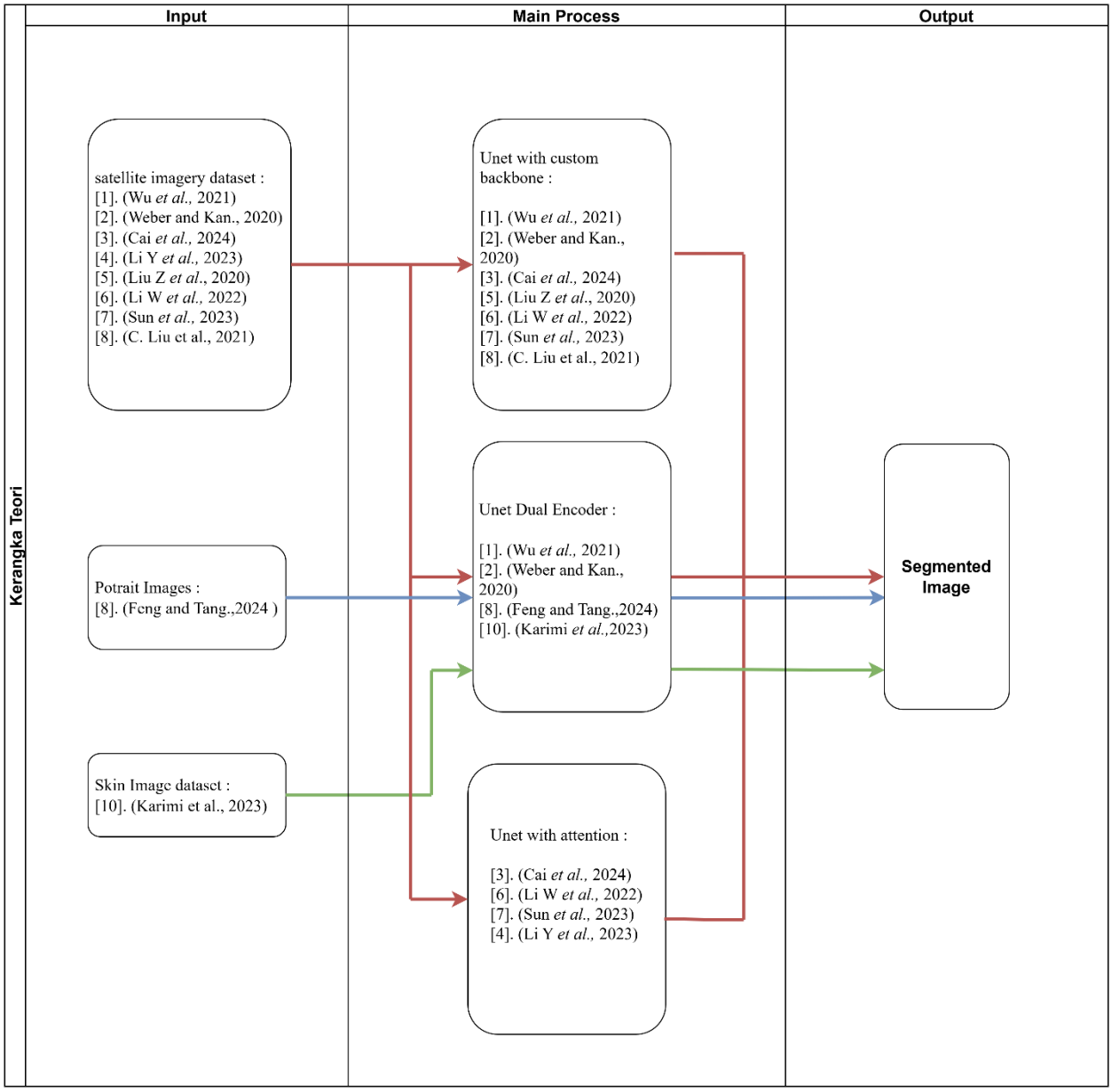
Selanjutnya penlitian oleh Feng and Tang (2024), penelitian ini bertujuan untuk dapat melakukan semantic segmentation manusia pada gambar yang portrait. Dataset yang digunakan pada penelitian ini diambil dari Personalized Guidance Research Based on Cross modal High Order Reasoning Model. Metode yang diajukan untuk penelitian ini adalah Unet dengan dual-modal information complementarity. Hasil dari metode yang diajukan pada penelitian ini mampu menghasilkan akurasi dalam persen nilai PA sebesar 96.58, dan MIoU sebesar 94.48. Hasil ini menjadi yang terbaik diantara metode lainnya sebagai perbandingan yaitu Unet, LinkNet, PortraitNet, Deeplab v3+, dan Trans Unet.

Selanjutnya penelitian oleh Karimi *et al* (2023), penelitian ini bertujuan untuk melakukan segmentasi pada skin lesion pada manusia. Hal ini menjadi tantangan bagi peneliti karena skin lesion memiliki bentuk, warna, ukuran, dan texture yang bervariasi. Dataset yang digunakan pada penelitian ini berasal dari ISIC 2016, ISIC 2017 ISIC 2018, dan PH2. Metode yang digunakan pada penlitian ini menggunakan DEU-Net yaitu dual-encoder unet. Hasil dari penelitian ini mampu menghasilkan model dengan akurasi yang tinggi dalam persen yaitu 96.64 accuracy, 87.48 IoU, dan 92.90 Dice Coefficient.

2.3 Kerangka Teori

Dalam melakukan penelitian tentang penentuan tingkat kerusakan bangunan pasca bencana alam dengan metode *convolutional neural network* perlu mengacu

pada jurnal-jurnal sebelumnya yang telah melakukan sebagai kerangka teori seperti yang terlihat pada Gambar 2.1



Gambar 2.1. Kerangka Teori

Rangkuman kerangka teori pada gambar 2.1 dengan menyajikan evaluasi kinerja model pada penelitian sebelumnya dapat dilihat pada Table 2.1.

Tabel 2.1. Rangkuman kerangka teori

No	Metode	Peneliti	IoU	F1 Score	OA
1	U-Net + Attention mechanism	(W. Li *et al*., 2022)	-	First dataset : 87.7 Second dataset :88.0	-
2	SAANet with ResNet101 Backbone and SPAM + SCAM	(Sun *et al*., 2023)	First dataset : 68.5 Second dataet : 73.79 Third dataset : 50.1	First dataset : 80.22 Second dataet : 83.57 Third dataset : 66.3	First dataset : 86.72 Second dataet : 88.22 Third dataset : 65.45
3	Unet with ResNet50 backbone and SimAM + ASPP	(Cai *et al*., 2024)	81.1	90.45	-
4	SeResNxt + HRNet	(C. Liu *et al*., 2021)	-	0.65	-
5	Unet with Dual-modal compelemntary module + SK-ASPP module	(Feng and Tang, 2024)	95.48	-	-
6	DEU-Unet Dual Encoder	(Karimi *et al*., 2023)	87.48	-	-
7	Unet + CTM Module and ResNet50 Backbone	(Y. Li *et al*., 2023)	First dataset : 84.6 Second dataset : 81.92	First dataset : 91.24 Second dataset : 89.73	-
8	Unet with ResNet34 Backbone	(Z. Liu *et al.,* 2020)	83.0	-	-
9	Unet dual encoder with Attention and ResNet34 Backbone, Senet Backbone, SeRESNeXt backbone, FPN Backbone	(Wu *et al*., 2021)	-	79.2 F1 Score on SeResNeXt backbone	-
10	Unet with Mask R-CNN backbone	(Weber and Kan, 2020)	-	73.8	-

Berdasarkan Table 2.1 U-Net dengan menggunakan backbone dapat dengan baik menentukan tingkat kerusakan bangunan. Tidak hanya itu, U-Net dengan attention mechanism serta U-Net dengan dual encoder juga dapat mengahsilkan nilai matrik

yang baik. Oleh karena itu, penulis memutuskan untuk menggunakan U-Net backbone pretrained model + attention mechanism dalam penelitian ini.

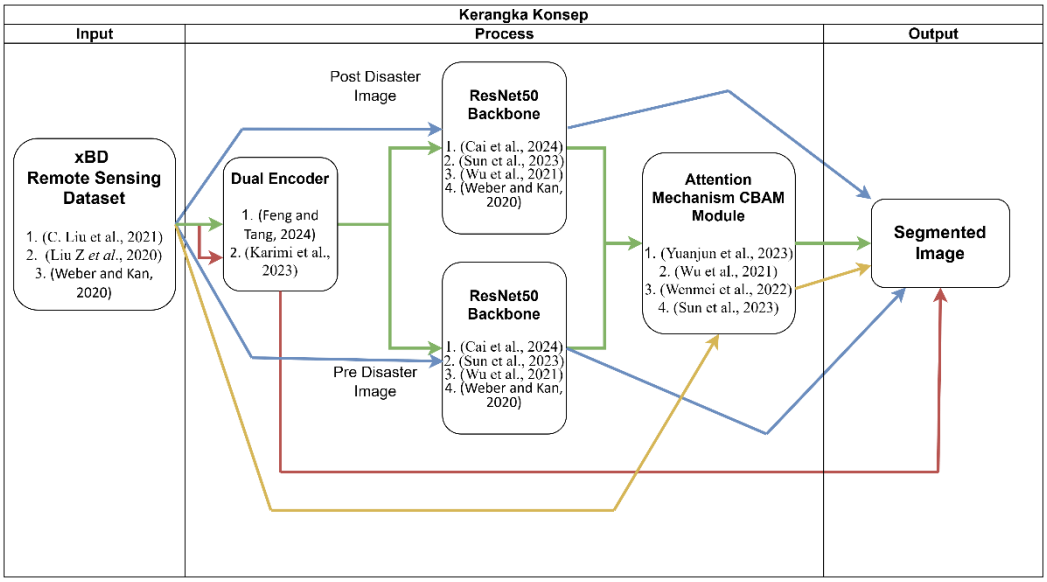
BAB III

METODOLOGI PENELITIAN

Pada bab ini, penulis akan memaparkan metode yang diajukan beserta penjelasan alur proses dari penelitian ini. Penjelasan pada bab ini meliputi kerangka konsep, dan desain penelitian. Penjelasan metode yang diajukan ada dibahas dalam sub bab pada penjelasan desain sistem.

3.1 Kerangka Konsep

Kerangka konsep merupakan gambaran hubungan antara konsep variable yang akan diteliti. Dalam kerangka konsep ini, variable yang ditambahkan pada proposed model akan dipaparkan beserta sumber dari penggunaan variable tersebut. Kerangka konsep tersaji pada Gambar 3.1.



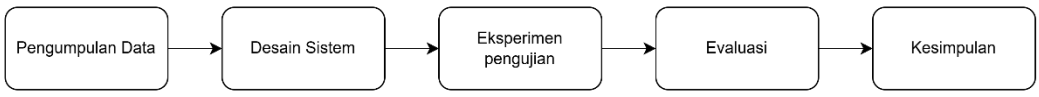
Gambar 3.1. Kerangka konsep penelitian

Gambar 3.1 menjelaskan tentag konsep dari metode yang digunakan pada penelitain ini. Garis warna merah, biru, dan kuning adalah garis proses dari penelitian sebelumnya, sedangkan garis berwarna hijau adalah konsep penelitian

yang diajukan pada penelitian ini. Dimulai dari pengumpulan dataset berupa *remote sensing image* dari xBD *Dataset* kemudian dilakukan *preprocessing image* sebelum digunakan untuk pelatihan. Selanjutnya, setelah melewati *preprocessing,* data akan terbagi menjadi dua yaitu data *pre-disaster image* dan data *post-disaster image*. Model akan memproses dua inputan dalam bersamaan namun dengan jalur yang berbeda, hal ini disebut sebagai *dual encoder*. Setelah kedua jenis data masuk, proses pertama yang dilakukan adalah ekstrasi fitur, pada penelitian ini proses ekstrasi fitur akan menggunakan *backbone* dari ResNet50 yang berupa *pretrained model*. Setelah melakukan ekstrasi fitur, fitur dari kedua jenis data akan disatukan kemudian akan dilakukan mekanisme *attention* dengan menggunakan CBAM module. Terakhir akan muncul output berupa bobot yang sudah optimal.

3.2 Desain Penelitian

Desain penelitian akan membahas alur dari penelitian secara menyeluruh mulai dari pengumpulan data, desain sistem, eksperimen pengujian, evaluasi dan terakhir kesimpulan. Penjelasan pada setiap alur akan ada pada subbab masing masing. Alur penelitian dapat dilihat pada berikut pada Gambar 3.2 :



Gambar 3.2. Desain penelitian

Gambar 3.2 merupakan alur dari penelitian yang akan dilakukan oleh penulis pada penelitian ini. Dimulai dari tahap pemgumpulan data yang didapatkan dari sumber internet, perancangan desain sistem, melakukan eksperimen dengan

scenario pengjuian yang telah di tentukan, mengevaluasi hasil pengujian, dan terakhir memberikan kesimpulan dari penelitian yang telah dilakukan.

3.2.1 Data Collection

Data yang digunakan pada penelitian ini merupakan *high resolution remote sensing image* yang didapatkan dari penyedia *dataset public xBD Dataset* dari *xView2: Assess Building Damage*. Dataset ini akan terdiri dari tiga jenis yaitu gambar dengan kondisi sebelum bencana alam (*pre-disaster*) dan gambar dengan kondisi setelah bencana alam (*post-disaster*) dan *Ground truth* yang berisi label untuk warna tiap kelas. Label kelas tersaji pada Tabel 3.1.

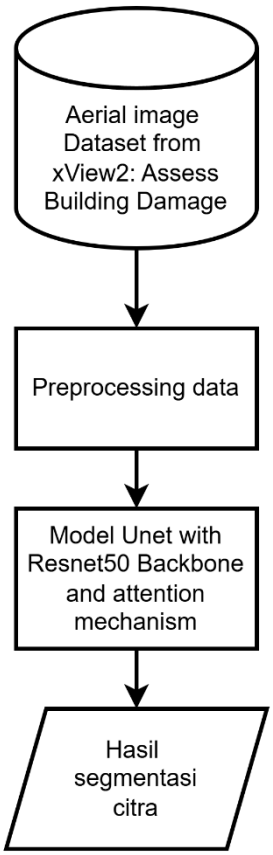
Tabel 3.1 Label Kelas

Kode Kelas	Deskripsi	Warna
0	*Background*	
1	*No damage*	
2	*Minor Damage*	
3	*Major Damage*	
4	*Destroyed*	

Pada table 3.1 terlihat deskripsi kelas beserta warnanya yang terdapat pada ground truth dataset. Total label kelas kerusakan terdapat 5 kelas yaitu *no damage, minor damage, major damage, dan destroyed, background.*

3.2.2 Desain sistem

Pada penelitian ini, penulis sudah menyusun desain sistem yang akan menggambarkan sistem dari penelitian ini. Desain sistem untuk penelitian ini dapat dilihat pada Gambar 3.3.

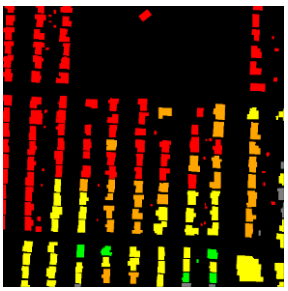
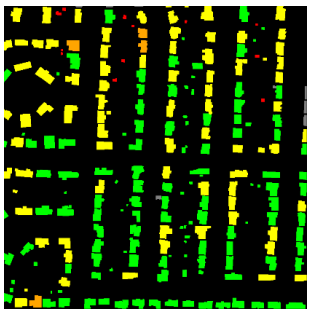


Gambar 3.3. Desain sistem

3.2.2.1 Input Data

Data yang akan digunakan pada penelitian ini berjumlah 332 gambar high resolution remote sensing image. Ukuran yang didapatkan dari dataset berukuran 1024 x 1024. Gambar yang akan diguankan akan berformat *Portable Network Graphics* (*.png). Contoh dataset tersaji pada Tabel 3.2

Tabel 3.2 Contoh data citra satelit beserta *ground truth*

No	Citra satelit daerah pasca bencana alam	*Ground truth*
1.		
2.		

3.2.2.2 Preprocessing

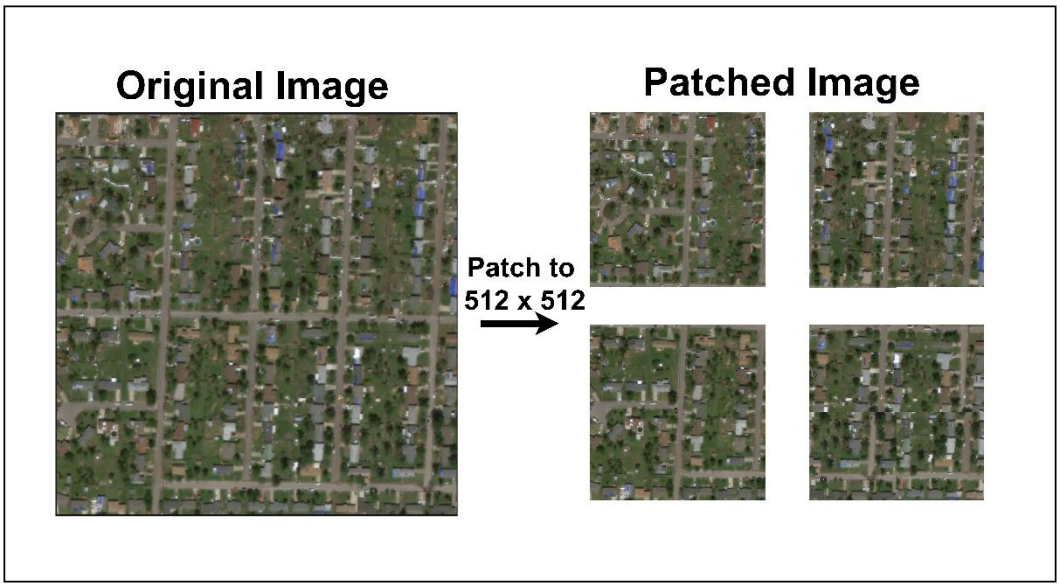
Preprocessing data gambar pada machine learning merupakan suatu rangkaian dan tahapan yang dilakukan pada *raw image* data sebelum digunakan pada *machine learning model*. Tujuannya adalah untuk meningkatkan kualitas, konsistensi, dan relevansi data, sehingga meningkatkan kinerja dan akurasi model (Bala and Behal., 2024).

Proses pertama setelah data dimasukkan adalah melakukan preprocessing. Langkah ini wajib dilakukan suapaya data yang telah dimasukkan (*Raw Data*) dapat langsung digunakan oleh *model machine learning*. Preprocessing yang akan dilakukan dapat dilihat alurnya berikut pada Gambar 3.4 :

Gambar 3.4 Alur prepocessing

3.2.2.3 Patching Data

Langkah pertama yang dilakukan yaitu proses *patching*. Proses ini akan memotong gambar menjadi beberapa bagian atau komponen yang lebih kecil (Chen *et al*., 2021). Pada proses ini nantinya akan memotong gambar yang semula berukuran 1024 x 1024 menjadi 512 x 512 . Sehingga dalam satu gambar berukuran 1024 x 1024 akan berjumlah menjadi 4 gambar. Proses patching data dapat dilihat pada Gambar 3.5



Gambar 3.5 Patching Image

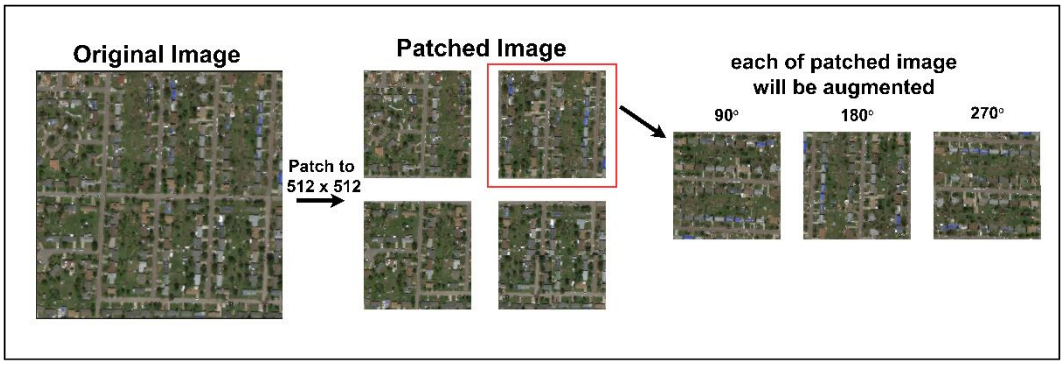
Persamaan pada proses ini ditulis dalam Persamaan 3.1 :

$$N_{\text{patch}} = \left(\frac{W}{P}\right) \times \left(\frac{H}{P}\right) \tag{3.1}$$

Keterangan :
- Npatch adalah jumlah patch yang dihasilkan dari satu citra.
- W dan H masing-masing adalah lebar dan tinggi citra asli dalam piksel.
- P adalah ukuran sisi patch (dalam hal ini 512 piksel).

3.2.2.4 Augmentasi Data

Selanjutnya setiap gambar akan menjalani proses *augmentasi.* Proses ini akan merotasi gambar pada derajat 90°, 180°, 270°. Tujuan dari proses ini agar data yang digunakan akan semakin kaya dan mengurangi resiko *overfitting.* Proses Agumentasi dapat dilihat pada Gambar 3.6



Gambar 3.6 Augmentasi data

Persamaan pada proses ini ditulis dalam Persamaan 3.2 :

$$(x', y') = R_{\theta} \cdot (x, y) \tag{3.2}$$

Keterangan :
- (x, y) adalah koordinat sebelum rotasi
- (x', y') adalah koordinat sesudah rotasi
- R_{θ} adalah matriks rotasi

3.2.2.5 Augmentasi Data

Setelah melakukan augmentasi, akan dilakukan proses normalisasi. Proses ini akan membagi nilai dari setiap *pixel* dengan 255, sehingga nilai tiap *pixel* hanya akan bernilai 0 – 1. Hal ini bertujuan untuk mengurangi beban komputasi saat melakukan proses pelatihan. Persamaan untuk proses ini dapat dilihat pada Persamaan 3.3 :

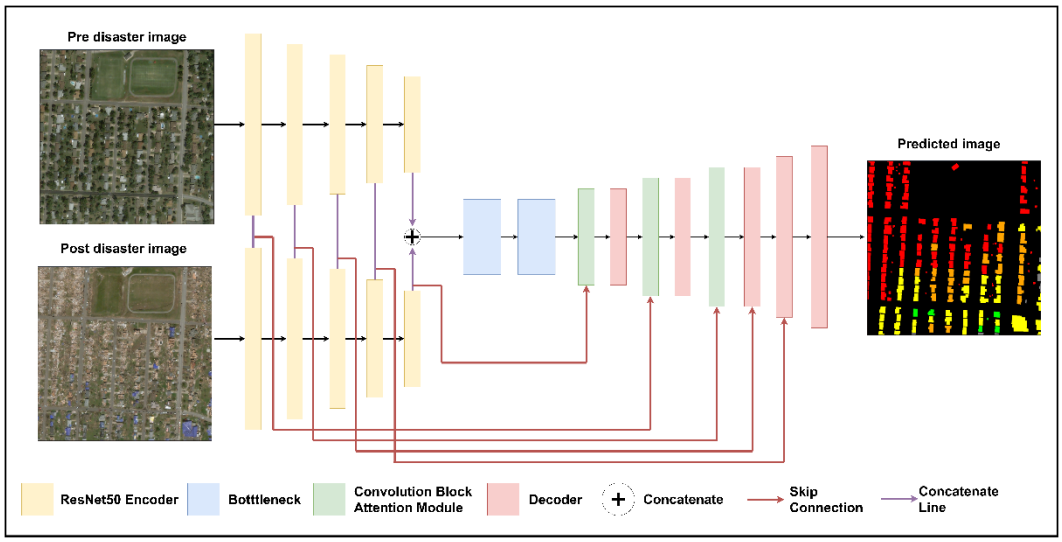
$$pixel\ normalization\ = \frac{original\ pixel}{255} \tag{3.3}$$

Keterangan :

- Pixel normalization adalah pixel yang sudah dinormalisasikan
- Original pixel adalah nilai pixel asli pada gambar

3.2.2.6 Unet with ResNet50 Backbone and Attention Mechanism

Berikutnya adalah model arsitektur yang diajukan pada penelitian ini. Pada penelitian ini akan memodifikasi U-Net yaitu mengubah encoder nya dengan encoder dari ResNet50 lalu menambahkan attention mechanism pada decodernya. Gambaran dari arsitektur dapat dilihat pada Gambar 3.7 yang dimodifikasi dari sumber (Sun *et al*., 2023) :



Gambar 3.7 Model Arsitektur

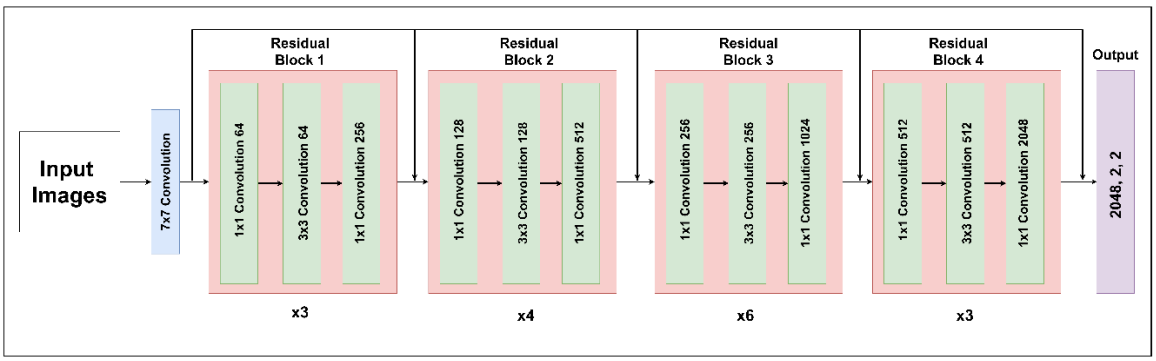
Model yang digunakan dalam penelitian ini merupakan pengembangan dari arsitektur U-Net yang dirancang untuk segmentasi kerusakan bangunan berdasarkan dua kondisi citra, yaitu sebelum (*pre-disaster*) dan sesudah bencana (*post-disaster*). Pada bagian encoder, digunakan dua buah ResNet50 yang telah dilatih sebelumnya (*pretrained*) untuk mengekstraksi fitur dari masing-masing citra secara independen. Fitur dari kedua encoder kemudian digabungkan menggunakan

proses *concatenation*, lalu dikompresi kanalnya dengan konvolusi 1×1 agar sesuai dengan struktur decoder. Proses ini memungkinkan model untuk menggabungkan informasi dari dua kondisi citra secara efektif, menangkap konteks temporal dan spasial dari objek yang diamati.

Untuk meningkatkan fokus model terhadap informasi penting selama proses decoding, digunakan modul *Convolutional Block Attention Module* (CBAM) pada beberapa tahap setelah proses *upsampling*. CBAM terdiri dari dua tahap perhatian, yaitu *channel attention* dan *spatial attention*, yang masing-masing membantu model dalam mengenali fitur penting berdasarkan kekuatan sinyal fitur di tiap kanal dan lokasi spasial dalam peta fitur. Dengan memasukkan CBAM secara bertingkat di beberapa blok decoder, model diharapkan mampu menyaring informasi yang tidak relevan dan meningkatkan akurasi prediksi pada area-area kerusakan. Pendekatan ini memberikan kontribusi terhadap peningkatan performa model segmentasi, terutama dalam konteks citra satelit pasca-bencana.

3.2.2.6.1 Encoder

Dalam artsitektur U-Net, encoder atau bisa dibilang *backbone* merupakan bagian dari jaringan yang bertanggung awab dalam mengekstrasi dan mengkompresi fitur dari input data gambar. Encoder terdiri dari rangkaian konvolusi dan *pooling layers* yang secara progresif mengurangi *spatial dimensions* dari gambar sambil menambahkan jumlah saluran fitur untuk menangkap representasi data input yang semakin abstrak dan kompleks (A. Naji and Majeed Saleh., 2024). Dalam penelitian ini, encoder yang digunakan adalah milik ResNet50. Berikuit adalah gambar alur ResNet50 Encoder pada Gambar 3.8.



Gambar 3.8. ResNet 50 Backbone
Sumber : (Sharma *et al*., 2025)

ResNet-50 digunakan untuk mengekstraksi fitur hierarkis dari citra masukan melalui lima tahapan utama: Conv1, Conv2_x, Conv3_x, Conv4_x, dan Conv5_x. Setiap tahapan terdiri dari beberapa residual block yang memperdalam representasi fitur sambil mengurangi resolusi spasialnya. Proses dimulai dari konvolusi awal dengan kernel 7×7 dan max-pooling, dilanjutkan dengan blok-blok residual yang menggunakan shortcut connection agar pelatihan jaringan lebih stabil dan efisien.

Ketika digunakan sebagai encoder, bagian *fully-connected* (FC) dan softmax pada akhir ResNet-50 dihilangkan. *Output* akhir encoder berupa feature map berdimensi [B, 2048, H/32, W/32], yang merepresentasikan fitur semantik tingkat tinggi dari citra *input*. *Output* ini kemudian dapat digunakan sebagai representasi dasar dalam arsitektur lain seperti decoder U-Net, atau dimodifikasi lebih lanjut, seperti melalui *block attention* yang memperkuat informasi penting dalam citra.

3.2.2.6.2 Convolution Layer

Proses ini yang akan dikerjakan pertama kali pada tiap blok *encoder*, rumus dari konvolusi ialah (Goodfellow *et al*., 2015.) :

$$C(i,j) = \sum_{m}^{1} \sum_{n}^{1} I(i+m, j+n) * K(m,n) \quad (3.4)$$

Keterangan :
(i, j) = indeks piksel.
(m, n) = indeks kernel.
K = kernel.
C = hasil konvolusi.
I = citra input.

3.2.2.6.3 Pooling Layer

Proses Selanjutnya gambar *input* yang sudah dilakukan konvolusi, hasilnya akan dilakukan *pooling*. Pada Resnet50 encoder menggunakan *max pooling*. Berikut adalah rumusnya pada Persamaan 3.5 :

$$Y(i,j) = \max_{(m,n) \in R(i,j)} X(m,n) \quad (3.5)$$

Keterangan :
Y(i,j) = nilai output pooling pada posisi (i,j)
X(m,n) = nilai input feature map di dalam jendela pooling.
R(i,j) = region (area) input yang dicakup oleh jendela pooling di posisi (i,j).
max = operasi maksimum (mengambil nilai tertinggi dalam jendela).

3.2.2.6.4 Bottleneck

Setelah melakukan fitur ekstrasi pada encoder, kedua fitur yaitu fitur yang berasal dari data pre-disaster dan fitur yang berasal dari post-disaster akan digabungkan. Proses ini dinamakan concatenation, yaitu proses menggabungkan dua atau lebih struktur data seperti array, vektor, matriks, atau set fitur ujung ke ujung sepanjang dimensi tertentu (Yan *et al*., 2023).

Selanjutnya setelah penggabungan akan menuju pada proses bottleneck. Pada proses ini, fitur – fitur yang telah didapatkan dari *encoder* dan digabungkan dalam proses concatenation akan dilakukan proses konvolusi seperti biasa namun bedanya adalah dengan tidak melakukan *pooling*. *Feature map* dari blok terakhir encoder akan memasuki blok *bottleneck*, proses yang akan terjadi yaitu akan

melakukan dua kali proses konvolusi. Setelah Fungsi dari bottleneck sendiri ialah supaya mempertahankan informasi – informasi dari gambar tersebut.

3.2.2.6.5 Decoder

Decoder dalam U-Net adalah bagian dari arsitektur yang bertanggung jawab untuk merekonstruksi resolusi spasial dari citra masukan setelah dikompresi oleh encoder. Decoder melakukan operasi upsampling untuk mengembalikan ukuran citra secara bertahap dan menyempurnakan peta fitur untuk menghasilkan keluaran segmentasi yang tepat (Ronneberger *et al*., 2015).

Proses yang terjadi pada decoder adalah transpose convolution. Transponse convolution dalam decoder U-Net merupakan teknik upsampling yang digunakan untuk meningkatkan resolusi spasial peta fitur. Teknik ini juga dikenal sebagai "deconvolution" atau "fractionally strided convolution." Dalam decoder, lapisan transpose convolution membalikkan downsampling yang dilakukan oleh encoder, yang memungkinkan jaringan merekonstruksi ukuran gambar asli dari representasi fitur terkompresi (Ramya *et al*., 2023) .

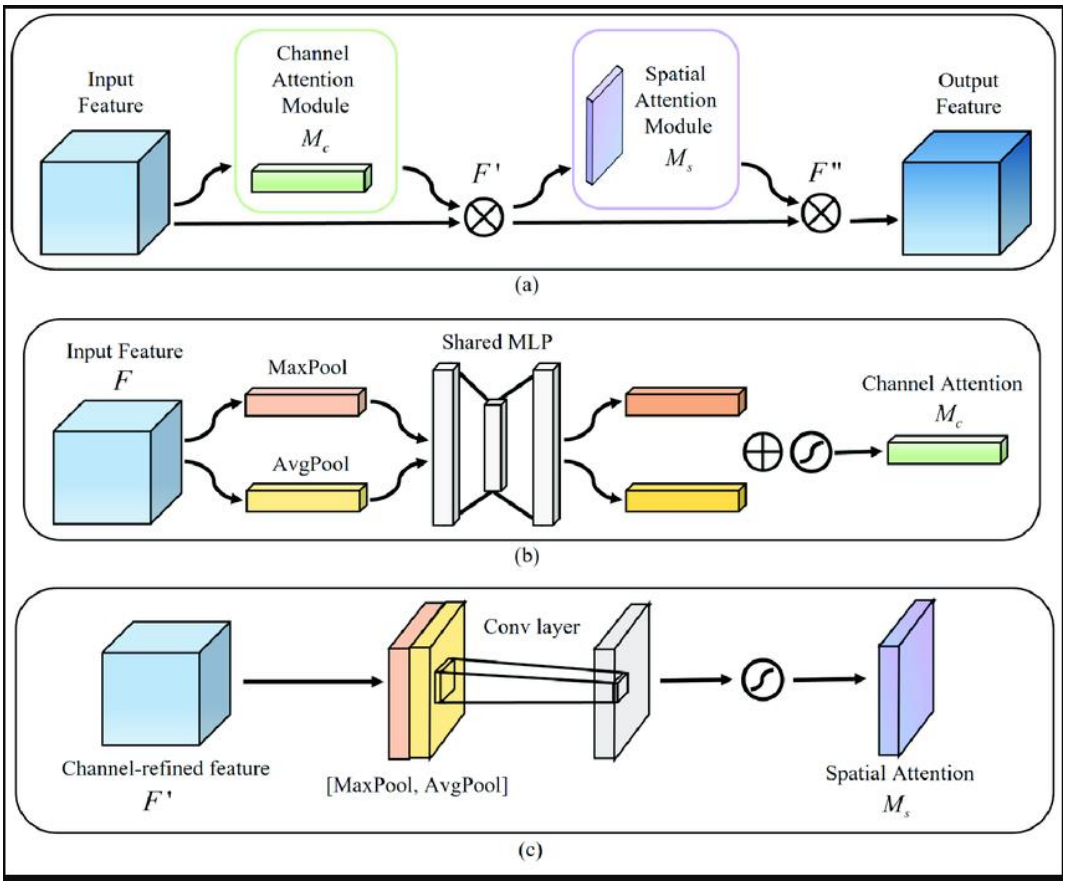
Berikut adalah rumus *transpose convolution* pada decoder pada Persamaan 3.6 :

$$C(i,j) = \sum_{m}^{1} \sum_{n}^{1} I(i-m, j-n) * K(m,n) \qquad (3.6)$$

Keterangan :
(i, j) = indeks piksel.
(m, n) = indeks kernel.
K = kernel.
C = hasil konvolusi.
I = citra input.

3.2.2.6.6 CBAM (Convolutional Block Attention Module)

Convolutional Block Attention Module (CBAM) adalah modul attention ringan dan efisien yang dapat dipasang secara plug-and-play pada berbagai arsitektur CNN. CBAM berfungsi untuk meningkatkan representasi fitur dengan cara memberikan perhatian (attention) pada informasi penting di dimensi channel maupun spasial. Dengan menyesuaikan bobot fitur berdasarkan konteks semantik dan spasial, CBAM membantu jaringan fokus pada bagian gambar yang paling relevan untuk tugas tertentu, seperti klasifikasi, deteksi, atau segmentasi(Woo *et al*., 2018) . Gambar rangkaian proses CBAM dapat dilihat pada Gambar 3.9.



Gambar 3.9. CBAM Module
Sumber : (Xu *et al*., 2023)

Dalam Gambar 3.9 (a) merupakan Gambaran utuh *Convolutional Block Attention Module*. Gambar (b) *Channel Attention Module* merupakan Gambaran lebih spesifik alur dari *block channel attention*. Gambar (c) *Spatial Attention Module* merupakan Gambaran lebih spesifik alur dari *block spatial attention*.

CBAM (*Convolutional Block Attention Module*) bekerja dalam dua tahap berurutan, yaitu *channel attention* dan *spatial attention*, untuk menyaring informasi penting dalam feature map. Pertama, pada tahap channel attention, fitur input diproses melalui *global average pooling* dan *max pooling* secara independen di setiap channel, kemudian dilewatkan ke jaringan *feed-forward* (MLP) untuk menangkap hubungan antar *channel*. *Output* dari kedua *pooling* digabung dan diaktivasi dengan fungsi sigmoid untuk menghasilkan *channel attention map*, yang kemudian dikalikan dengan input awal. Selanjutnya, pada tahap *spatial attention*, hasil dari *channel attention* diproses lagi dengan *average pooling* dan *max pooling* namun kali ini dilakukan di sepanjang dimensi *channel,* menghasilkan dua peta spasial yang digabung dan dilewatkan ke layer konvolusi 2D dengan kernel 7×7. Setelah diaktivasi dengan sigmoid, diperoleh spatial attention map yang juga dikalikan dengan input sebelumnya. Proses ini memungkinkan jaringan fokus pada fitur "apa" yang penting dan "di mana" lokasinya dalam gambar.

Berikut adalah persamaan dari channel attention pada Persamaan 3.7 dan Persamaan 3.8 :

$$M_c(F) = \sigma\left(MLP\big(AvgPool(F)\big) + MLP\big(MaxPool(F)\big)\right) \tag{3.7}$$

$$F' = M_c(F) \otimes F \tag{3.8}$$

Keterangan :
- F: Input feature map

- AvgPool, Maxpool : global average dan max pooling
- MLP: Dua layer FC (biasanya diimplementasikan sebagai 1x1 convolution)
- σ: Sigmoid activation
- Mc(F): Channel attention map
- ⊗: element-wise multiplication

Kemudian persamaan dari spatial attention adalah berikut pada Persamaan 3.9 dan Persamaan 3.10 :

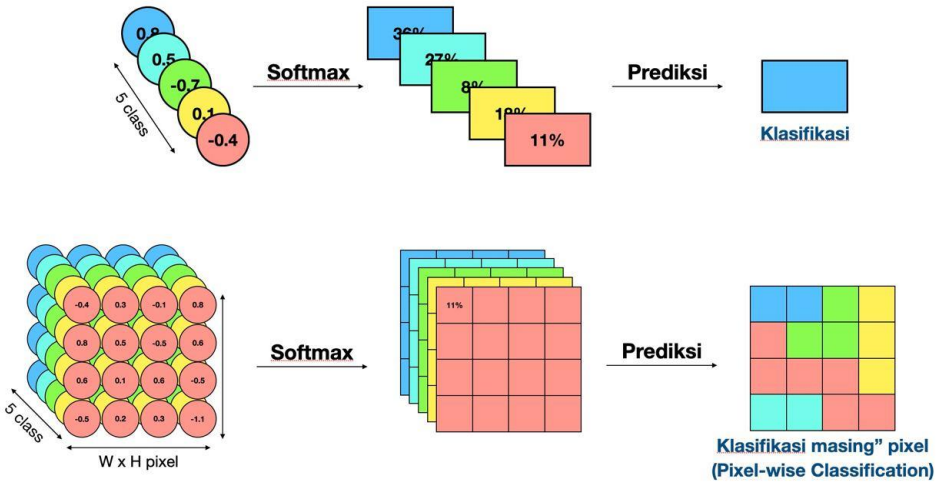
$$M_s(F') = \sigma\big(Conv^{7\times7}([AvgPool(F'); MaxPool(F')])\big) \tag{3.9}$$

$$F'' = M_s(F') \otimes F' \tag{3.10}$$

Keterangan :
- F′: Output dari channel attention
- [·;·]: Concatenation di dimensi channel
- Conv7×7: Convolution layer dengan kernel 7×7
- Ms(F′): Spatial attention map

3.2.2.6.7 Pixelwise

Setelah gambar telah melewati blok *decoder* terakhir, langkah terakhir yaitu melakukan proses yang dinamakan *pixelwise*. Proses ini akan me klasifikasikan per pixel gambar dengan menggunakan aktivasi *softmax* untuk menentukan nilai mana yang paling besar pada setiap kelas yang telah ditentukan. Proses ini akan digabungkan dengan *cross entropy* yaitu nantinya perhitungan *loss function* akan dilakukan per *pixel* untuk dilihat nilainya. Jika ternyata nilai dari *loss function* tidak begitu baik maka akan dilakukan proses *backpropagation* untuk memperbarui nilai bobot pada *filter*. Gambar 3.10 diambil dan dimodifikasi dari (Strohmann *et al*., 2019)



Gambar 3.10 Pixelwise

3.3 Uji Coba

Proses ini akan menjelaskan tentang uji coba pada penelitian ini yang terbagi menjadi skenario pengujian dan hasil awal. Skenario pengujian akan berisi kombinasi skenario yang akan dilakukan pada penelitian ini, sedangkan hasil awal adalah hasil percobaan dengan dataset yang telah ditentukan untuk melihat kerja awal dari sistem yang telah dirancang.

3.3.1 Skenario Pengujian

Pada penelitian ini, penulis akan menjabarkan apa saja skenario percobaan yang akan dilakukan pada penelitian ini. Hal ini dilakukan agar peneliti mengerti pada skenario mana model memiliki bobot paling optimal dan akurasi paling bagus.

Skenario pertama adalah skenario dalam penggunaan *model U-Net*. Dalam arsitektur yang peneliti ajukan, model yang digunakan adalah modifikasi U-Net dengan dual encoder ResNet50 backbone + attention CBAM module.Berikutnya

penggunaan model U-Net dengan hanya menggunakan dual encoder digunakan sebagai pembanding.

Tabel 3.3 Skenario pertama

Kode Skenario	***Model***
1	*Dual Encoder U-Net with ResNet50 Nackbone + CBAM*
2	*Dual Encoder U-Net*

Skenario kedua yaitu peneliti akan mencoba untuk melakukan pelatihan model dengan menggunakan *loss function* yang bervariatif. Skenario pengujian dapat dilihat pada Tabel 3.4.

Tabel 3.4 Skenario kedua

Kode Skenario	***Loss Function***
A	CrossEntropy
B	Dual Focal

Skenario ketiga yaitu peneliti akan mencoba untuk melakukan pelatihan model dengan menggunakan *optimizer* yang bervariatif. Skenario pengujian dapat dilihat pada Tabel 3.5

Tabel 3.5 Skenario ketiga

Kode Skenario	***Optimizer***
I	Adam
II	RMSprop

Tabel 3.6 merupakan kombinasi skenario untuk *proposed model*, dimana untuk semua *loss function* dan *optimizer* akan menggunakan *proposed model*. Nantinya hasil dari kombinasi ini akan di analisis dan dibandingkan

Tabel 3. 6 Kombinasi Skenario *Proposed Model*

Skenario	Kombinasi	Keterangan
1	1-A-I	Crossentropy with adam
2	1-B-I	Dual focal with adam
3	1-A-II	Crossentropy with RMSprop
4	1-B-II	Dual focal with RMSprop

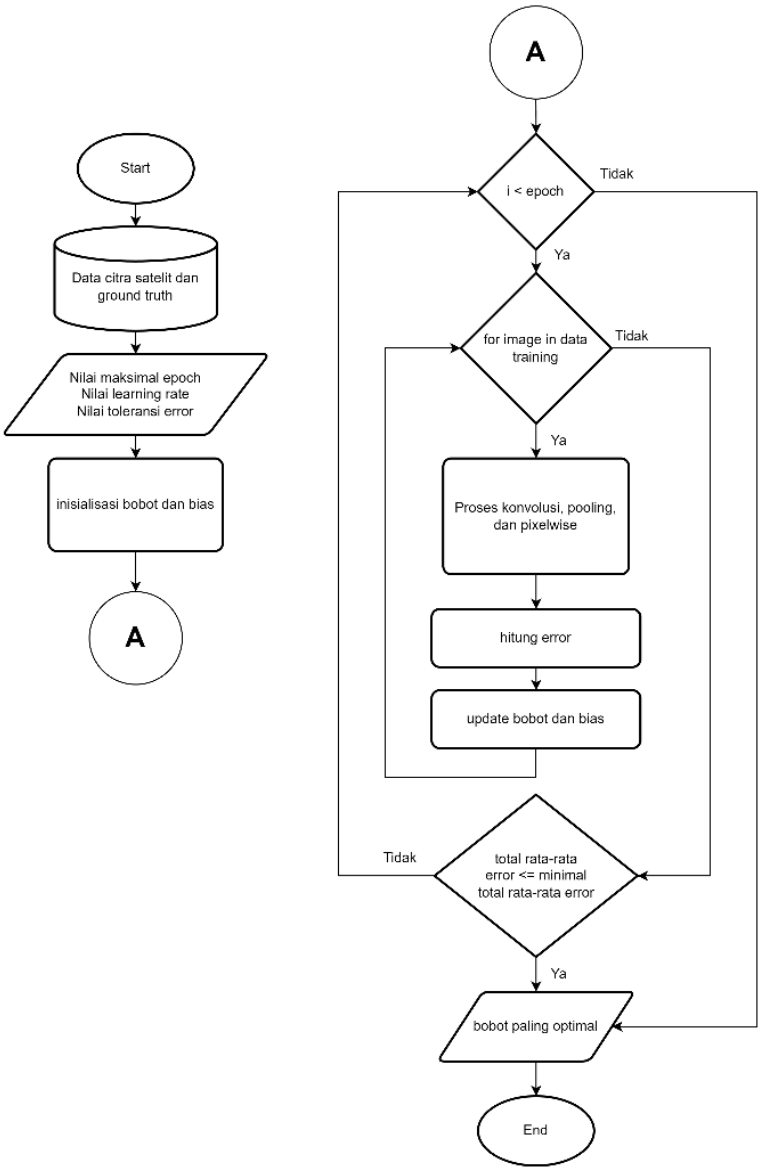
Tabel 3. 7 Kombinasi Skenario dual encoder U-Net

Skenario	Kombinasi	Keterangan
5	2-A-I	Crossentropy with adam
6	2-B-I	Dual focal with adam
7	2-A-II	Crossentropy with RMSprop
8	2-B-II	Dual focal with RMSprop

Tabel 3.7 merupakan kombinasi skenario untuk *U-Net dual encoder*, dimana untuk semua *loss function* dan *optimizer* akan menggunakan *U-Net dual encoder model*. Nantinya hasil dari kombinasi ini akan di analisis dan dibandingkan.

3.3.2 Training Model Convoutional Neural Network

Untuk bisa mendapatkan model yang mampu melakukan segmentasi dengan akurat, maka langkah berkutnya yaitu melakukan proses *training*. Proses training bertujuan agar model *convolutional neural network* memiliki nilai bobot yang paling optimal. Sehingga dengan nilai bobot yang paling optimal mampu melakukan prediksi untuk segmentasi citra yang diberikan. Proses *training* akan menggunakan gambar citra aerial daerah pasca bencana alam dan citra *ground truth* yang telah diinisialisasi di awal. *Flowchart* dari proses *training* diberikan pada Gambar 3.11 yang diambil dan dimodifikasi dari (Ojha *et al*., 2012)



Gambar 3.11 Alur training

Setelah memanggil data inputan, pada proses training perlu dilakukan juga inisialisasi parameter yang akan digunakan dalam pelatihan. Parameter tersebut biasa disebut dengan *hyperparameter*. Diantara *hyperparameter* tersebut ialah *learning rate, epoch, batch size,* dan toleransi *error*. Setelah di inisialisasi, nilai dari *hyperparameter* akan bersifat konstanta atau tidak akan berubah – ubah selama

proses pelatihan. Berikut adalah tabel definisi dari setiap *hyperparameter* di atas pada Tabel 3.8.

Tabel 3.8 Definisi *Hyperparameter*

No	*Hyperparameter*	Definisi
1.	*Learning rate*	*Learning rate* adalah parameter yang menentukan seberapa besar langkah yang diambil selama proses optimasi pada model *machine learning*. Penelitian ini akan menggunakan 0.001 untuk *learning rate*
2.	*Epoch*	*Epoch* adalah satu kali putaran lengkap melalui seluruh dataset pelatihan selama pelatihan model *machine learning*. Penelitian ini akan menggunakan 50 epoch
3.	*Batch size*	*Batch size* adalah jumlah contoh pelatihan yang digunakan dalam satu iterasi pada proses pelatihan model. Penelitian ini akan menggunakan 4 *batch size*
4.	Toleransi *error*	Toleransi *error* mengacu pada tingkat kesalahan yang dapat diterima atau batas kesalahan yang diizinkan selama pelatihan model, sering digunakan untuk menghentikan pelatihan jika tingkat kesalahan telah mencapai batas tertentu.

Setelah melakukan inisialisasi *hyperparameter*, maka selanjutnya proses pelatihan model akan berlangsung. Hasil akhir dari proses pelatihan adalah bobot yang sudah optimal, pada model ini bobot tersebut akan berformat .pth.

3.3.3 *Reassembly Segmented Image*

Penyusunan kembali *segmented images* adalah proses komputasi yang bertujuan untuk mengembalikan *original image* dari bagian-bagiannya. Proses ini menggabungkan teknik dari *image processing*, *pattern recognition*, dan *optimization* untuk mengotomatisasi tugas yang sebelumnya dilakukan secara manual dan memakan waktu lama, dengan aplikasi yang signifikan di bidang sains, forensik, dan restorasi media digital (Sakpal *et al*., 2015).

Pada proses ini, setelah model dilatih maka bobot dari model yang berformat .pth akan di load dan digunakan untuk melakukan *predict image*. *Predict image* akan menginputkan gambar berukuran 1024 x 1024 yang tidak ikut kedalam data latih untuk nantinya dilakukan segmentasi guna mengetahui Tingkat kerusakan bangunannya. Data yang diinputkan harus terbagi menjadi dua jenis yaitu *pre-disaster image* dan *post-disaster image*. Setelah data diinputkan, karena model dilatih dengan menggunakan data berukuran 512 x 512 maka data yang semula berukuran 1024 x 1024 akan di patch dulu menjadi 512 x 512 baru kemudian dilakukan *predict.* Setelah proses *predict* selesai, output dari model adalah gambar tersegmentasi berukuran 512 x 512 oleh karena itu dilakukan proses penggabungan kembali gambar menjadi ukuran semula yaitu 1024 x 1024.

Prosesnya dilakukan dengan menghitung jumlah baris dan kolom *patch* yang diperlukan, lalu menempatkan setiap *patch* ke posisi yang sesuai di dalam gambar berdasarkan indeks baris dan kolomnya. Hasil akhirnya adalah sebuah citra besar yang terdiri dari potongan-potongan gambar kecil yang disusun secara berurutan dari kiri ke kanan dan atas ke bawah. Persamaan proses ini dapat dilihat pada Persamaan 3.11 dan Persamaan 3.12

$$N_{\text{row}} = \frac{H}{P}, \quad N_{\text{col}} = \frac{W}{P} \tag{3.11}$$

$$Ifull(x,y) = Ipatch^{(i,j)(x',y')dengan} x = j \times P + x', y = i \times P + y' \tag{3.12}$$

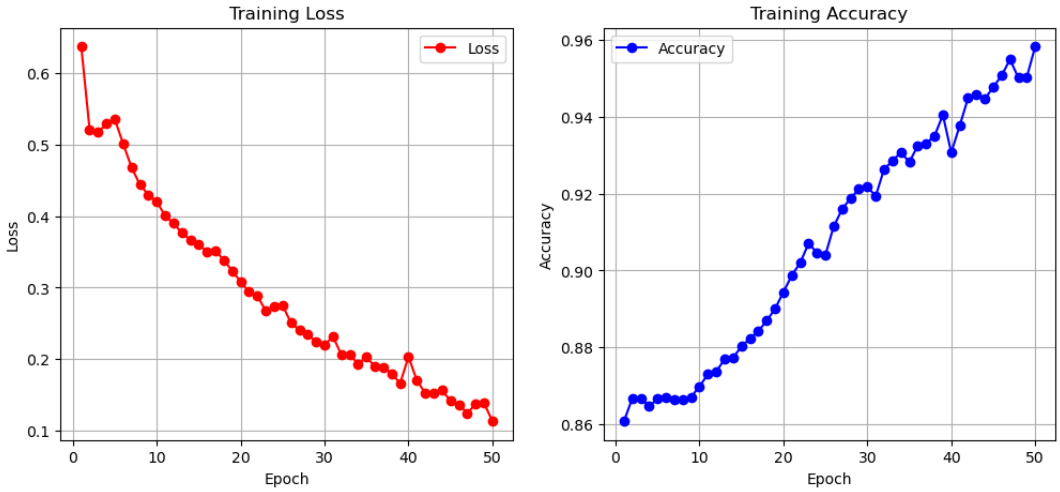
Keterangan :

- N row dan N col adalah jumlah baris dan kolom patch.
- H dan W adalah tinggi dan lebar gambar input (misal 1024).
- P adalah ukuran patch (misal 512).
- I full (x, y) adalah nilai piksel di gambar gabungan pada posisi (x, y).

- I patch (i, j) (x prime, y prime) adalah nilai piksel pada patch ke-i, j pada posisi lokal (x prime, y prime).
- x prime dan y prime adalah koordinat di dalam patch, mulai dari 0 sampai P minus 1.

3.3.4 Hasil Awal

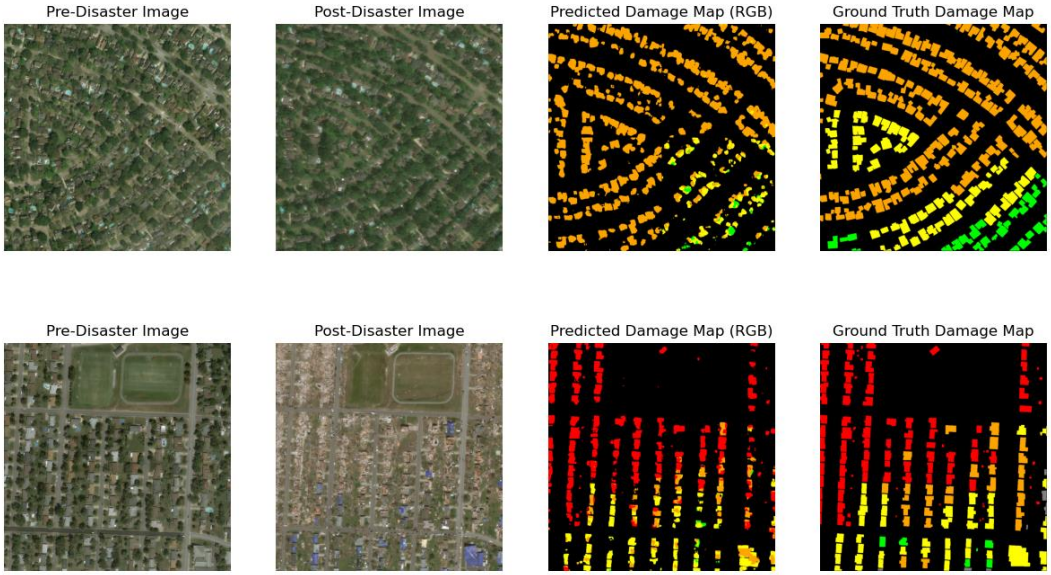
Dataset yang akan digunakan pada penelitian ini berjumlah 15.000 data. Pada sub bab ini akan dilakukan percobaan awal dengan hanya menggunakan 5.000 data. Pada pengujian percobaan ini, model yang ddigunakan adalah proposed model, dengan perbandingan 80:20 untuk train dan test. Learning rate akan di set pada 0.001 , batch size akan di set pada 4, dan epoch train sebanyak 50 kali. Berikut adalah hasil dari percobaan awal pada Gambar 3.12 :



Gambar 3.12. Accuracy and Loss Training

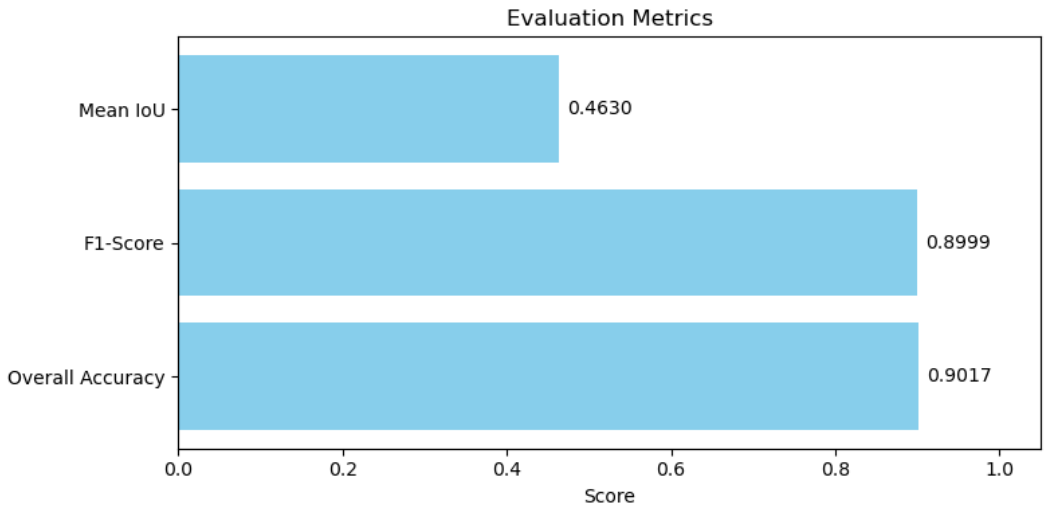
Pada hasil training, proposed model berhasil mendapatkan nilai yang memuaskan yaitu Loss: 0.1134, Accuracy: 0.9583. Dapat dilihat pada gambar 3.10 bahwa model dapat mempelajari dengan baik dataset yang diberikan. Hal ini diindikasikan pada grafik yang konsisten naik pada accuracy dan grafik yang konsisten menurun pada loss.

Berikutnya model yang telah dilatih akan digunakan untuk memprediksi gambar yang sudah dipisahkan saat pembgaian dataset. Berikut adalah salah satu contoh hasil prediksi yang dapat dilihat pada Gambar 3.13.



Gambar 3.13. Gambar Hasil Prediksi Model

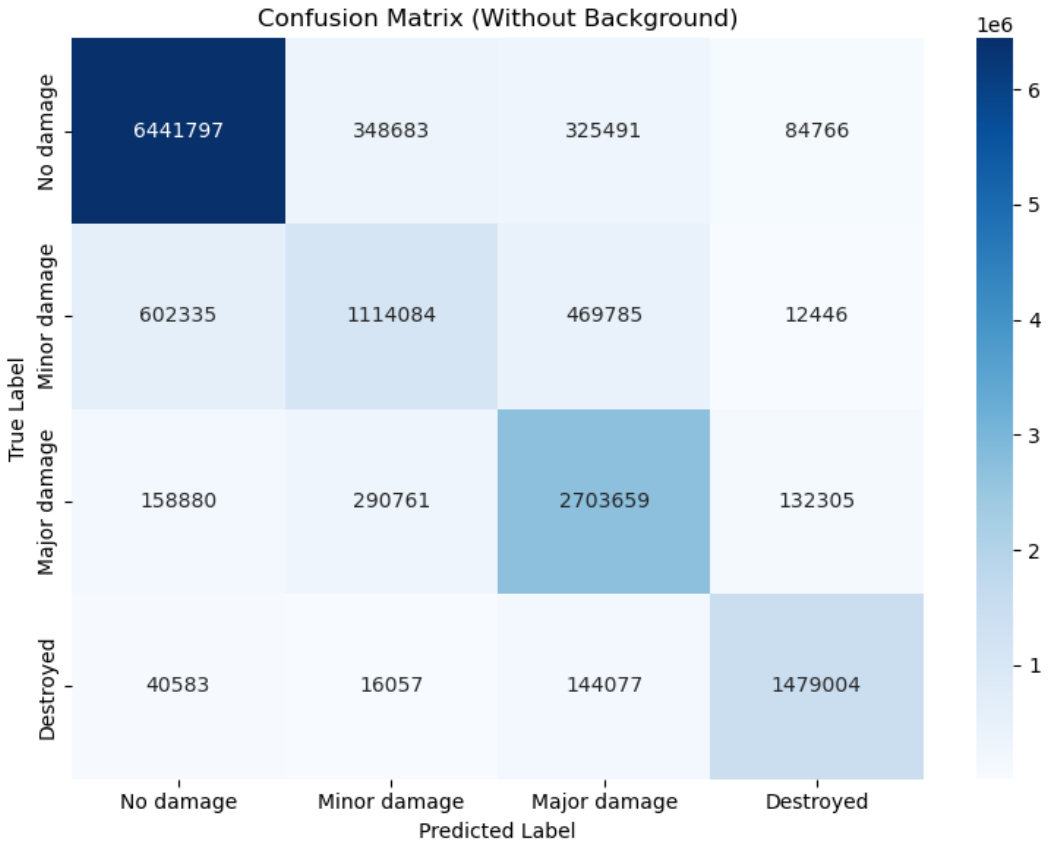
Berikutnya setelah model selesai training dan prediction, model akan melakukan evaluation terhadap hasil prediksinya. Evaluation matrics yang akan digunakan pada model ini adalah mIoU, F1-Score dan OA. Evaluasi ini akan menggunakan 20 persen dari data yang sudah di split di awal tadi. Berikut adalah hasil nya pada Gambar 3.14:



Gambar 3.14. Evaluation Matrics

Berdasarkan hasil evaluasi model segmentasi pada gambar 3.13 yang dilakukan pada data uji, diperoleh nilai Overall Accuracy (OA) sebesar 0.9017, F1-score sebesar 0.8999, dan Mean Intersection over Union (mIoU) sebesar 0.4630. Nilai OA dan F1 yang tinggi menunjukkan bahwa model memiliki performa yang sangat baik secara umum dalam mengklasifikasikan piksel ke dalam kelas-kelas kerusakan yang sesuai. Namun, nilai mIoU yang masih berada di kisaran 0.46 mengindikasikan bahwa model belum sepenuhnya optimal dalam memetakan batas-batas wilayah kerusakan secara presisi, terutama pada kelas-kelas minoritas atau yang bentuknya lebih kompleks. Hal ini cukup umum dalam tugas segmentasi dengan distribusi kelas yang tidak seimbang dan kompleksitas citra tinggi, namun hasil ini tetap mencerminkan performa model yang kuat dan potensial untuk dikembangkan lebih lanjut.

Berikutnya juga akan menampilkan confusion matrix dari hasil pengujian data sementara ini, berikut adalah hasil confusion matrix nya pada Gambar 3.15 :



Gambar 3.15. Confusion Matrix

Hasil confusion matrix pada gambar 3.14 menunjukkan bahwa model memiliki performa sangat baik dalam mengklasifikasikan piksel latar belakang (background) dengan akurasi 96.38%, menandakan deteksi area non-bangunan sangat andal. Untuk kelas “no damage”, model juga menunjukkan akurasi tinggi sebesar 61.78%, meskipun masih ada kebingungan yang cukup signifikan dengan kelas "background" (29.10%) dan “minor damage” (8.58%). Kelas "minor damage" menunjukkan akurasi moderat di angka 43.51%, namun sering tertukar dengan "background" (30.93%) dan “no damage” (20.01%), yang menunjukkan model kesulitan membedakan tingkat kerusakan yang lebih ringan. Sementara itu, kelas “major damage” dan “destroyed” masing-masing memiliki akurasi 37.31% dan 59.46%, tetapi sering diklasifikasikan sebagai tingkat kerusakan yang lebih ringan,

menunjukkan tantangan model dalam mengenali kerusakan berat. Secara keseluruhan, model bekerja baik dalam mengidentifikasi background dan bangunan tanpa kerusakan, namun perlu ditingkatkan dalam membedakan level kerusakan yang lebih tinggi, terutama antara "minor", "major", dan "destroyed".

3.4 Evaluasi

Setelah model memberikan hasil segmentasi citra, langkah selanjutnya yaitu perhitungan evaluasi model terhadap hasil yang sudah diberikan. Evaluasi model pada penelitian ini akan menggunakan *Mean Intersection Over the Union Score, F1 Score, dan OA* atau biasa disingkat MIoU. Nilai yang dihasilkan akan mempunyai rentang 0 – 1 dimana semakin mendekati angka 1 maka bentuk hasil prediksi akan semakin mendekati ground truth (Khan *et al*., 2021). Rumus perhitungan mIoU dapat dilihat pada Persamaan 3.13 :

$$MIoU = \sum_{1}^{n} \frac{TP}{(TP + FP + FN)} * \frac{1}{n} \tag{3.13}$$

Keterangan :
TP = nilai true positive
FP = nilai false positive
FN = nilai false negative

Selanjutnya rumus perhitungan F1-Score dapat dilihat pada Persamaan 3.14 :

$$Dice_c = \frac{1}{C} \sum_{c=1}^{C} \frac{2 \cdot TP_c}{2 \cdot TP_c + FP_c + FN_c} \tag{3.14}$$

Keterangan :
TP = nilai true positive
FP = nilai false positive
FN = nilai false negative
C = jumlah total kelas

Selanjutnya rumus perhitungan OA dapat dilihat pada Persamaan 3.15 :

$$OA = \frac{\sum_{c=1}^{C} TP_c}{\sum_{c=1}^{C}(TP_c + FP_c + FN_c + TN_c)} \quad (3.15)$$

Keterangan :
TP = nilai true positive
FP = nilai false positive
FN = nilai false negative
C = jumlah total kelas

3.5 Intrument Penelitian

Instrumen penelitian merupakan komponen penting dalam mendeskripsikan variabel-variabel yang digunakan selama proses penelitian. Instrumen ini mencakup identifikasi terhadap jenis variabel yang terlibat, baik sebagai penyebab, perantara, maupun hasil dari proses yang dilakukan.

Dalam penelitian ini, variabel-variabel yang digunakan dibagi menjadi tiga kategori, yaitu variabel *independent*, variabel *intervening*, dan variabel *dependent*. Variabel *independent* merupakan komponen input yang digunakan dalam pemodelan segmentasi kerusakan bangunan. Variabel *intervening* menjelaskan proses atau mekanisme internal dalam jaringan yang menghubungkan input dan output. Sementara itu, variabel *dependent* merupakan hasil akhir dari sistem, yaitu citra segmentasi yang menunjukkan tingkat kerusakan bangunan pascabencana.

Rincian dari masing-masing variabel disajikan pada Tabel 3.9

Tabel 3.9 Instrumen Penelitian

Jenis	Variabel	Keterangan
Variabel Independen	- Citra pre-disaster - Citra post-disaster - Arsitektur model (UNet + dual ResNet-50 + CBAM)	Input utama yang digunakan dalam proses pelatihan dan prediksi model segmentasi
Variabel Intervening	- Ekstraksi fitur oleh encoder - Mekanisme attention CBAM - Proses pelatihan model	Proses internal yang menjembatani input (variabel bebas) dengan hasil segmentasi akhir
Variabel Dependent	Citra hasil segmentasi kerusakan bangunan	Output utama dari model segmentasi yang menunjukkan tingkat kerusakan pada masing-masing area

Variabel *independent* dalam penelitian ini mencakup data citra *pre-disaster* dan *post-disaster* yang berasal dari dataset *xView2*, serta arsitektur model segmentasi yang digunakan, yaitu UNet dengan dual encoder berbasis ResNet-50 *pretrained* dan dilengkapi dengan mekanisme *attention* CBAM.

Variabel *intervening* terdiri atas proses-proses internal dalam model *deep learning*, seperti ekstraksi fitur oleh encoder, penguatan informasi spasial dan channel melalui mekanisme CBAM, serta proses pembelajaran model selama pelatihan. Variabel ini berperan sebagai perantara yang menghubungkan input dengan output segmentasi.

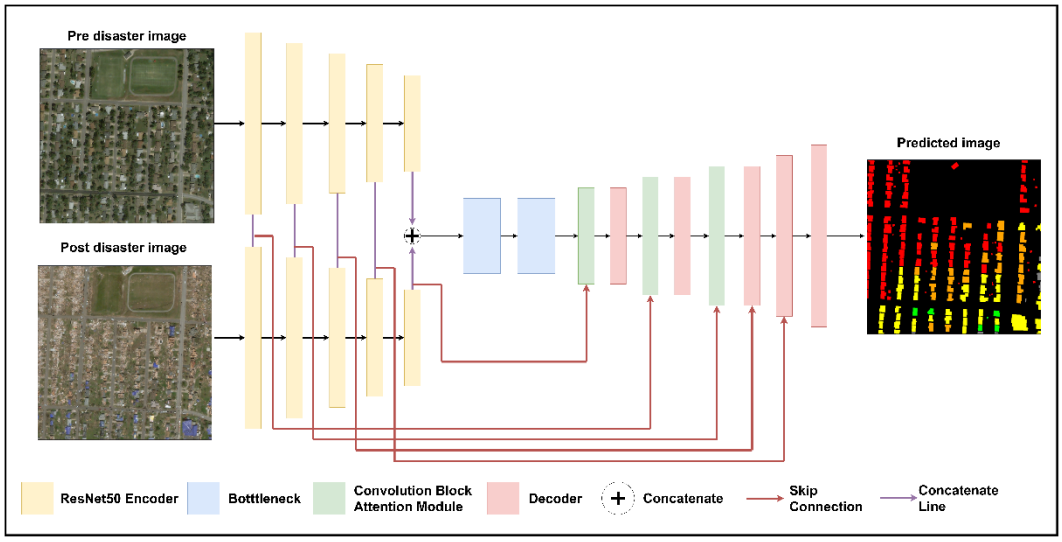
Variabel *dependent* dari penelitian ini adalah citra hasil segmentasi kerusakan bangunan, yaitu citra keluaran dari model yang memetakan tingkat kerusakan pada setiap piksel area terdampak. Hasil segmentasi ini kemudian dievaluasi kualitasnya menggunakan metrik *F1-score*, *mean Intersection over Union* (mIoU), dan *Overall Accuracy* (OA).

BAB IV
U-NET WITH RESNET-50 BACKBONE AND ATTENTION MECHANISM

Pada bab ini akan disajikan hasil pelatihan model berdasarkan skenario pengujian yang telah disusun yang terdiri dari loss dari setiap skenario. Berikutnya, hasil prediksi segmentasi akan disajikan berupa gambar yang telah tersegmentasi. Kemudian hasil evaluasi model akan ditampilkan nilai evaluasi setiap skenario pengujian. Terakhir akan disajikan analisis awal pengujian.

4.1. Model Arsitektur

Model yang digunakan pada bab ini adalah proposed model yang telah penulis susun pada bab 3. Model ini merupakan hasil modifikasi model dasar U-Net dengan adanya penambahan satu encoder pada proses ekstrasi fitur, kemudian digunakan module ResNet50 sebagai *pretrained* encoder, dan terakhir ada penambahan *attention mechanism* pada tahap decoder. Gambar arsitektur model dapat dilihat pada Gambar 4.1.



Gambar 4.1 Proposed model

4.2. Uji Coba

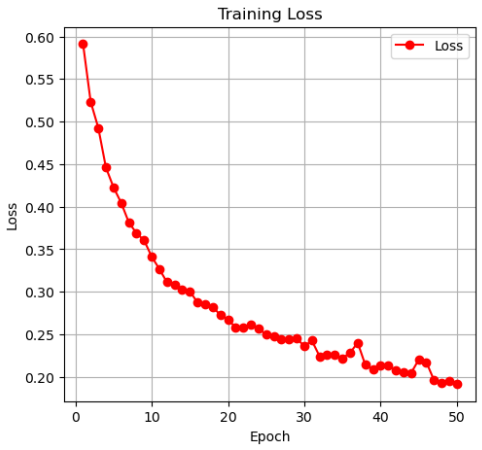
Pada bab ini akan melakukan 4 skenario uji coba dari total 8 skenario uji coba yang sudah dijelaskan pada sub bab 3.4. Skenario yang akan dilakukan dapat dilihat pada Tabel 4.1.

Tabel 4. 1. Tabel Kelompok Skenario 1

Skenario	Kombinasi	Keterangan
1	1-A-I	Crossentropy with adam
2	1-B-I	Dual focal with adam
3	1-A-II	Crossentropy with RMSprop
4	1-B-II	Dual focal with RMSprop

4.2.1 Skenario 1

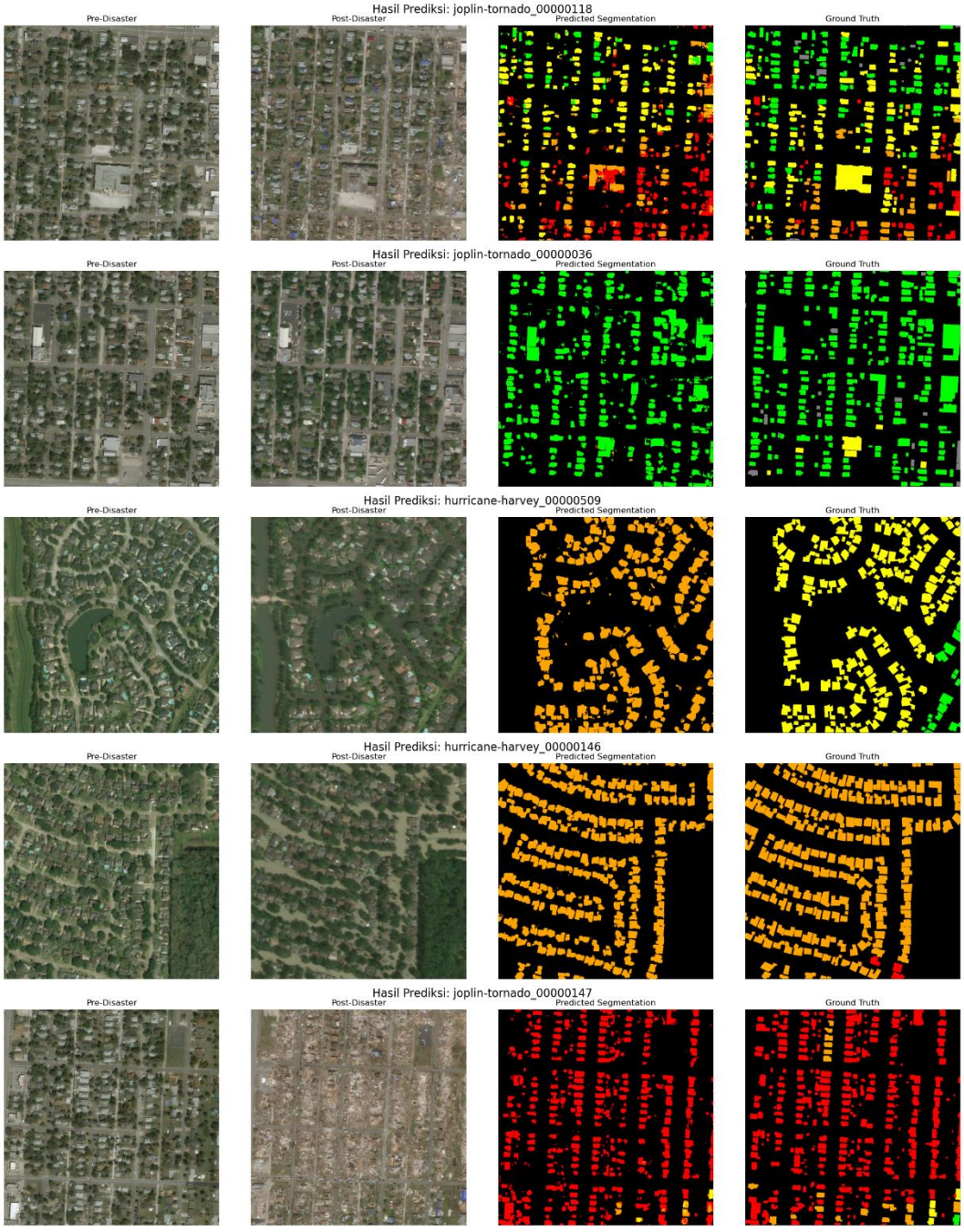
Pada skenario ini telah dilakukan pelatihan model selama 50 *epoch* dengan *learning rate* sebesar 0.001. *Loss function* yang digunakan pada skenario ini menggunakan crossentropy dengan *optimizer* menggunakan adam. Hasil training model pada skenario pertama berhasil mendapatkan nilai sebesar *Loss*: 0.1914. Gambar grafik *loss* dapat dilihat pada Gambar 4.2.



Gambar 4.2 Grafik Loss Skenario 1

Selanjutnya setelah melewati proses pelatihan, model digunakan untuk melakukan prediksi gambar untuk melihat hasil segmentasi tingkat kerusakan

bangunan. Telah diberikan 5 gambar yang diambil dari *data testing* sebagai contoh prediksi hasil segmentasi. Hasil dari segmentasi dapat dilihat pada Gambar 4.3.



Gambar 4.3 Hasil Segmentasi Skenario 1

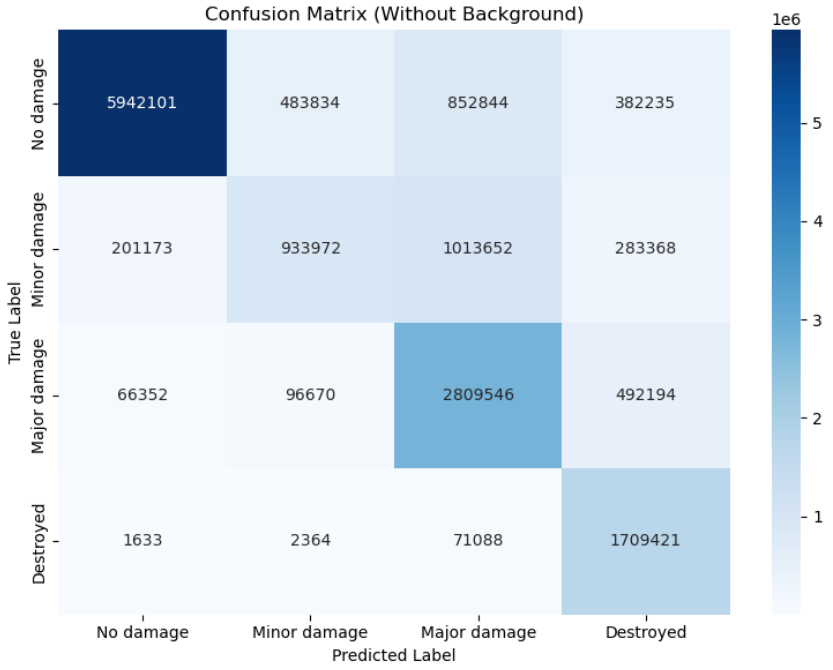
Setelah model melakukan prediksi untuk mendapatkan hasil segmentasi, selanjutnya model akan dievaluasi dengan menggunakan *test dataset* yang telah di *split data*. Jumlah data yang digunakan pada proses *testing* ini sebanyak 1594 data. Evaluasi yang telah dilakukan adalah IoU dan F1 – score yang akan dievaluasi per kelas. Kemudian model juga akan dievaluasi dengan melihat nilai mIoU, mF1-score dan *Overal accuracy*. Nilai IoU dan f1- score perkelas dapat dilihat pada Tabel 4.2 dan nilai mIoU, mF1 dan Overal accuracy dapat dilihat pada Tabel 4.3. Gambar *confussion matrix* dapat dilihat pada Gambar 4.4.

Tabel 4.2 Evaluasi Model per Kelas skenario 1

Number	Class	F1 - Score	IoU
1	No Damage	0.745787	0.594625
2	Minor Damage	0.379268	0.234010
3	Major Damage	0.618636	0.447844
4	Destroyed	0.683455	0.519128

Tabel 4.3 Evaluasi Rerata Skenario 1

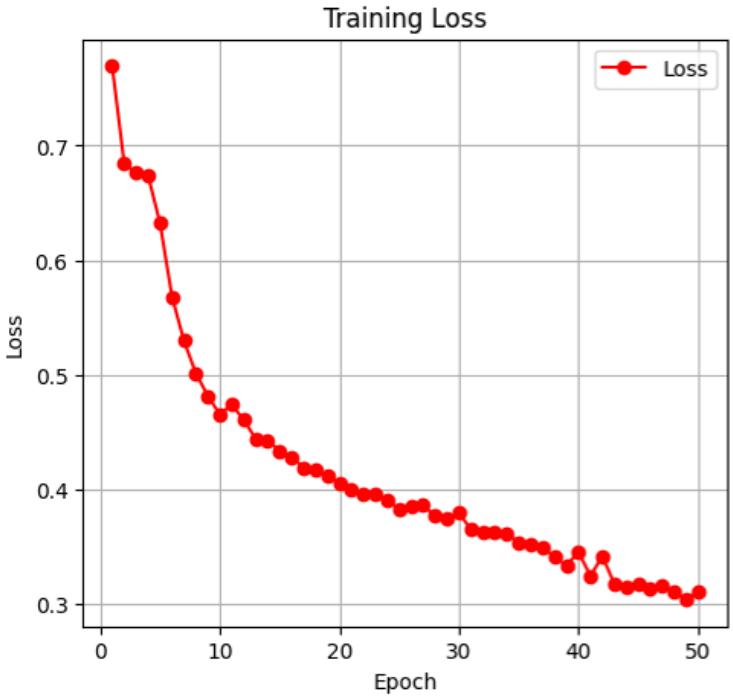
Overall Accuracy (OA)	Mean F1 Score (mF1)	Mean IoU (mIoU)
0.5813	0.4854	0.3591



Gambar 4.4 *Confussion Matrix* Skenario 1

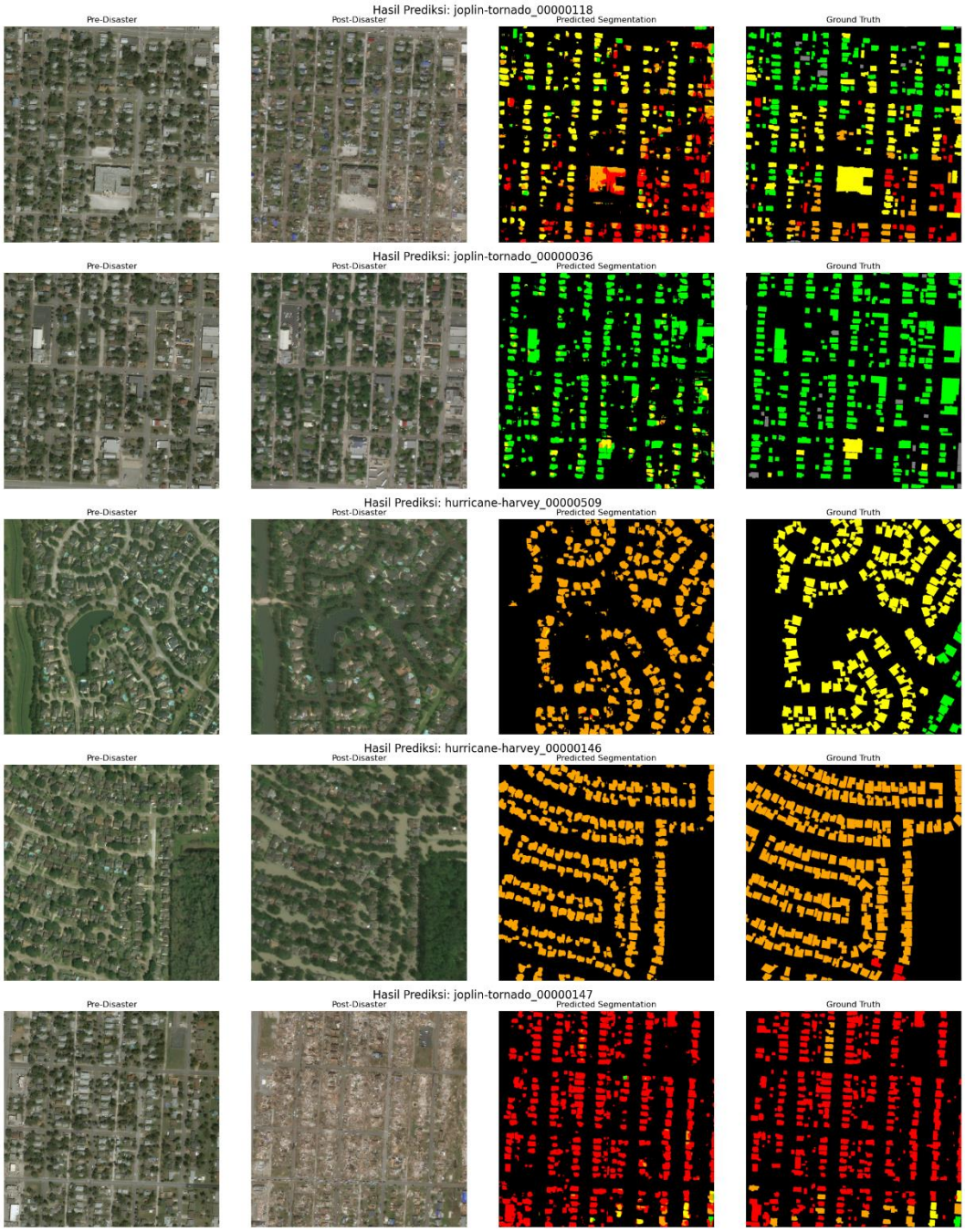
4.2.2 Skenario 2

Pada skenario ini telah dilakukan pelatihan model selama 50 epoch dengan *learning rate* sebesar 0.001. *Loss function* yang digunakan pada skenario ini menggunakan dual focal loss dengan *optimizer* menggunakan adam. Hasil training model pada skenario pertama berhasil mendapatkan nilai sebesar *Loss:* 0.3108,. Gambar grafik *loss* dapat dilihat pada gambar 4.5.



Gambar 4.5 Grafik loss skenario 2

Selanjutnya setelah melewati proses pelatihan, model digunakan untuk melakukan prediksi gambar untuk melihat hasil segmentasi tingkat kerusakan bangunan. Telah diberikan 5 gambar yang diambil dari *data testing* sebagai contoh prediksi hasil segmentasi. Hasil dari segmentasi dapat dilihat pada Gambar 4.6.



Gambar 4.6 Hasil segmentasi skenario 2

Setelah model melakukan prediksi untuk mendapatkan hasil segmentasi, selanjutnya model akan dievaluasi dengan menggunakan *test dataset* yang telah di *split data*. Jumlah data yang digunakan pada proses *testing* ini sebanyak 1594 data. Evaluasi yang telah dilakukan adalah IoU dan F1 – score yang akan dievaluasi per

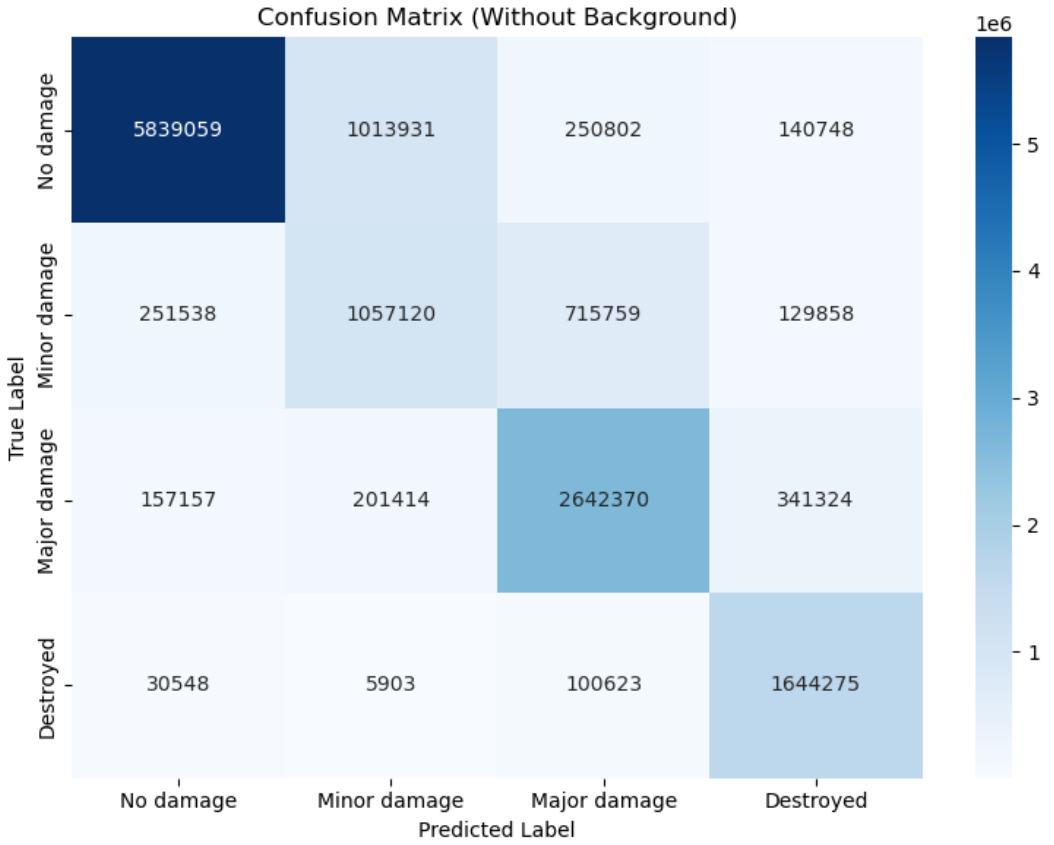
kelas. Kemudian model juga akan dievaluasi dengan melihat nilai mIoU, mF1-score dan *Overal accuracy*. Nilai IoU dan f1- score perkelas dapat dilihat pada tabel 4.2 dan nilai mIoU, mF1 dan Overal accuracy dapat dilihat pada Tabel 4.5. Berikutnya gambar *confussion matrix* dapat dilihat pada Gambar 4.7.

Tabel 4.4 Evaluasi per kelas skenario 2

Number	Class	F1 - Score	IoU
1	No Damage	0.729784	0.574535
2	Minor Damage	0.371789	0.228342
3	Major Damage	0.656859	0.489047
4	Destroyed	0.748882	0.598570

Tabel 4.5 Evaluasi rerata skenario 2

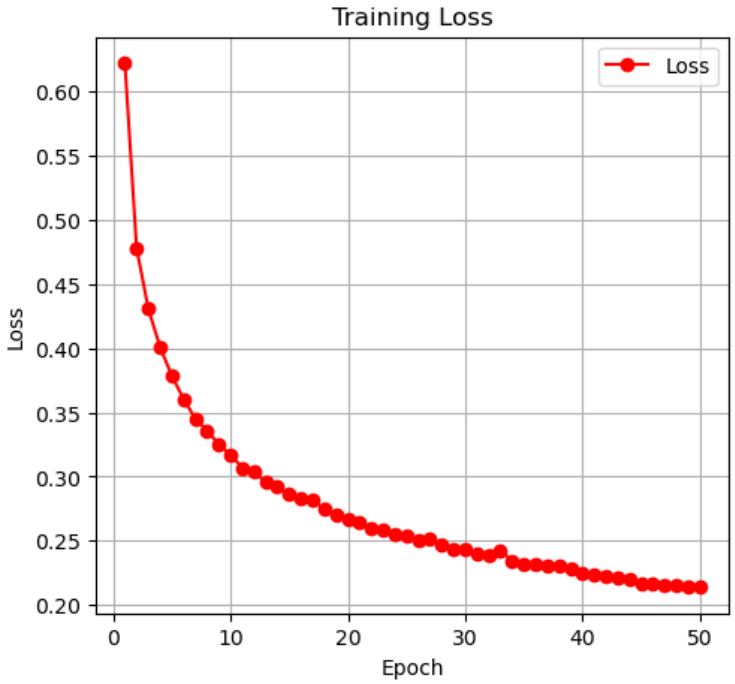
Overall Accuracy (OA)	Mean F1 Score (mF1)	Mean IoU (mIoU)
0.5705	0.5015	0.3781



Gambar 4.7 *Consfussion Matrix* skenario 2

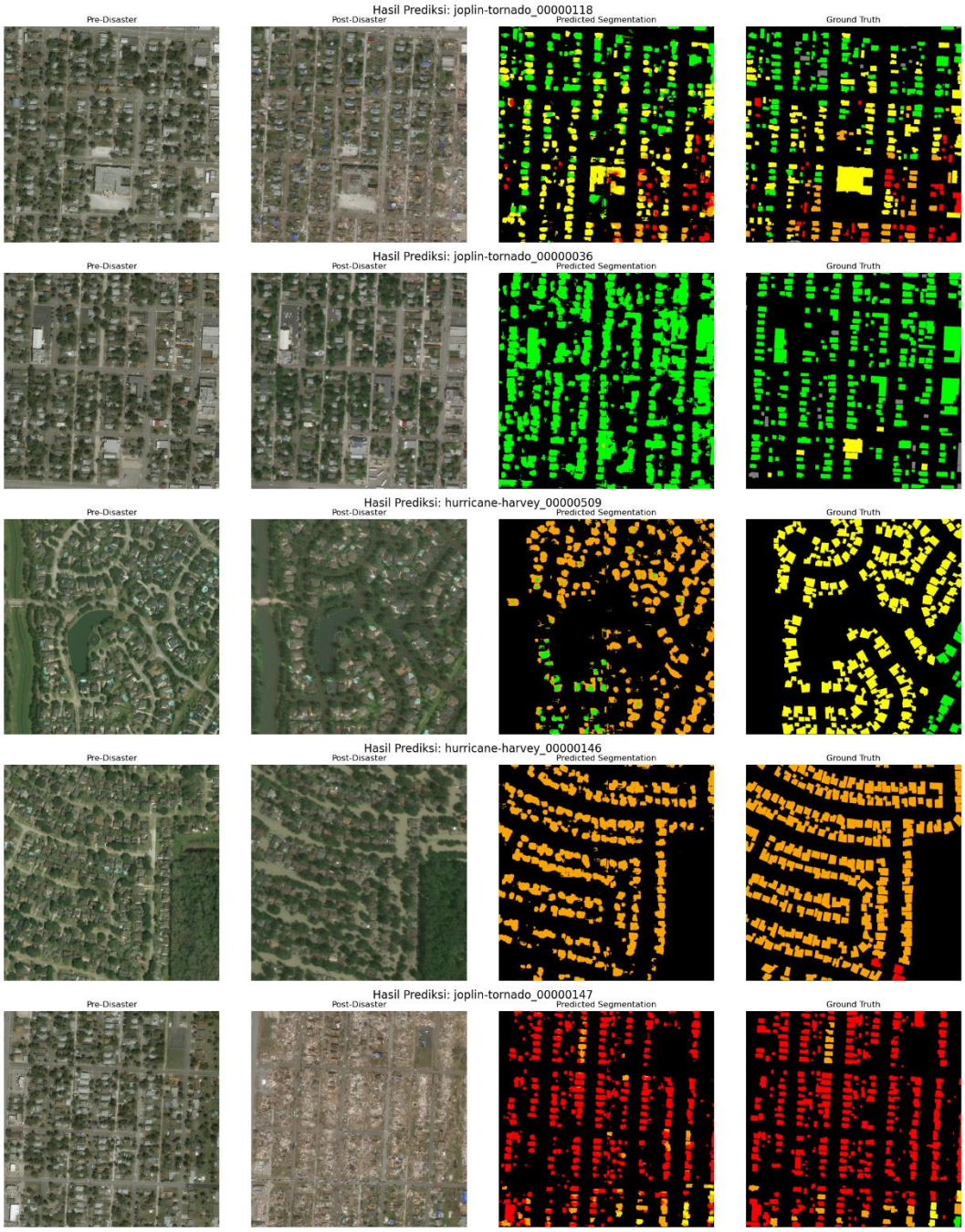
4.2.3 Skenario 3

Pada skenario ini telah dilakukan pelatihan model selama 50 epoch dengan *learning rate* sebesar 0.001. *Loss function* yang digunakan pada skenario ini menggunakan crossentropy dengan *optimizer* menggunakan RMSprop. Hasil training model pada skenario pertama berhasil mendapatkan nilai sebesar *Loss*: 0.2145. Gambar grafik *loss* dapat dilihat pada Gambar 4.4.



Gambar 4. 8 Grafik loss skenario 3

Selanjutnya setelah melewati proses pelatihan, model digunakan untuk melakukan prediksi gambar untuk melihat hasil segmentasi tingkat kerusakan bangunan. Telah diberikan 5 gambar yang diambil dari data *testing* sebagai contoh prediksi hasil segmentasi. Hasil dari segmentasi dapat dilihat pada Gambar 4.9.



Gambar 4.9 Hasil segmentasi skenario 3

Setelah model melakukan prediksi untuk mendapatkan hasil segmentasi, selanjutnya model akan dievaluasi dengan menggunakan *test dataset* yang telah di *split data*. Jumlah data yang digunakan pada proses *testing* ini sebanyak 1594 data.

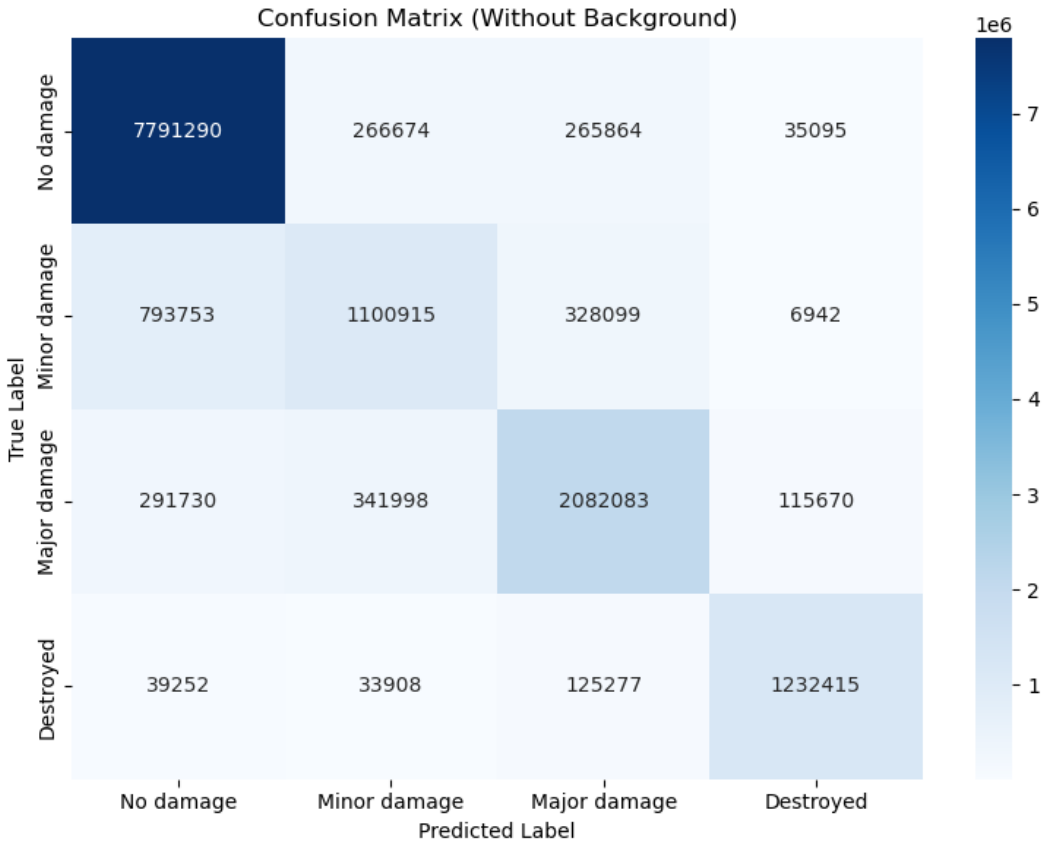
Evaluasi yang telah dilakukan adalah IoU dan F1 – score yang akan dievaluasi per kelas. Kemudian model juga akan dievaluasi dengan melihat nilai mIoU, mF1-score dan *Overal accuracy*. Nilai IoU dan f1- score perkelas dapat dilihat pada Tabel 4.6 dan nilai mIoU, mF1 dan *Overal accuracy* dapat dilihat pada Tabel 4.7. Berikutnya Gambar *confussion matrix* dapat dilihat pada Gambar 4.10

Tabel 4.6 Evaluasi model per kelas skenario 3

Number	Class	F1 - Score	IoU
1	No Damage	0.835980	0.718184
2	Minor Damage	0.427392	0.271772
3	Major Damage	0.583443	0.411874
4	Destroyed	0.699203	0.537519

Tabel 4.7 Evaluasi rerata skenario 3

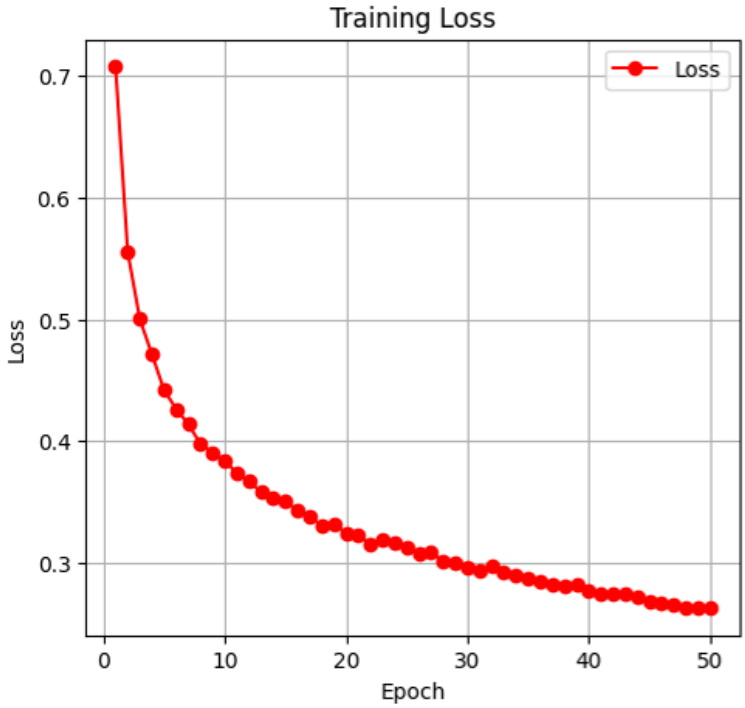
Overall Accuracy (OA)	Mean F1 Score (mF1)	Mean IoU (mIoU)
0.6227	0.5092	0.3879



Gambar 4.10 *Confussion matrix* skenario 3

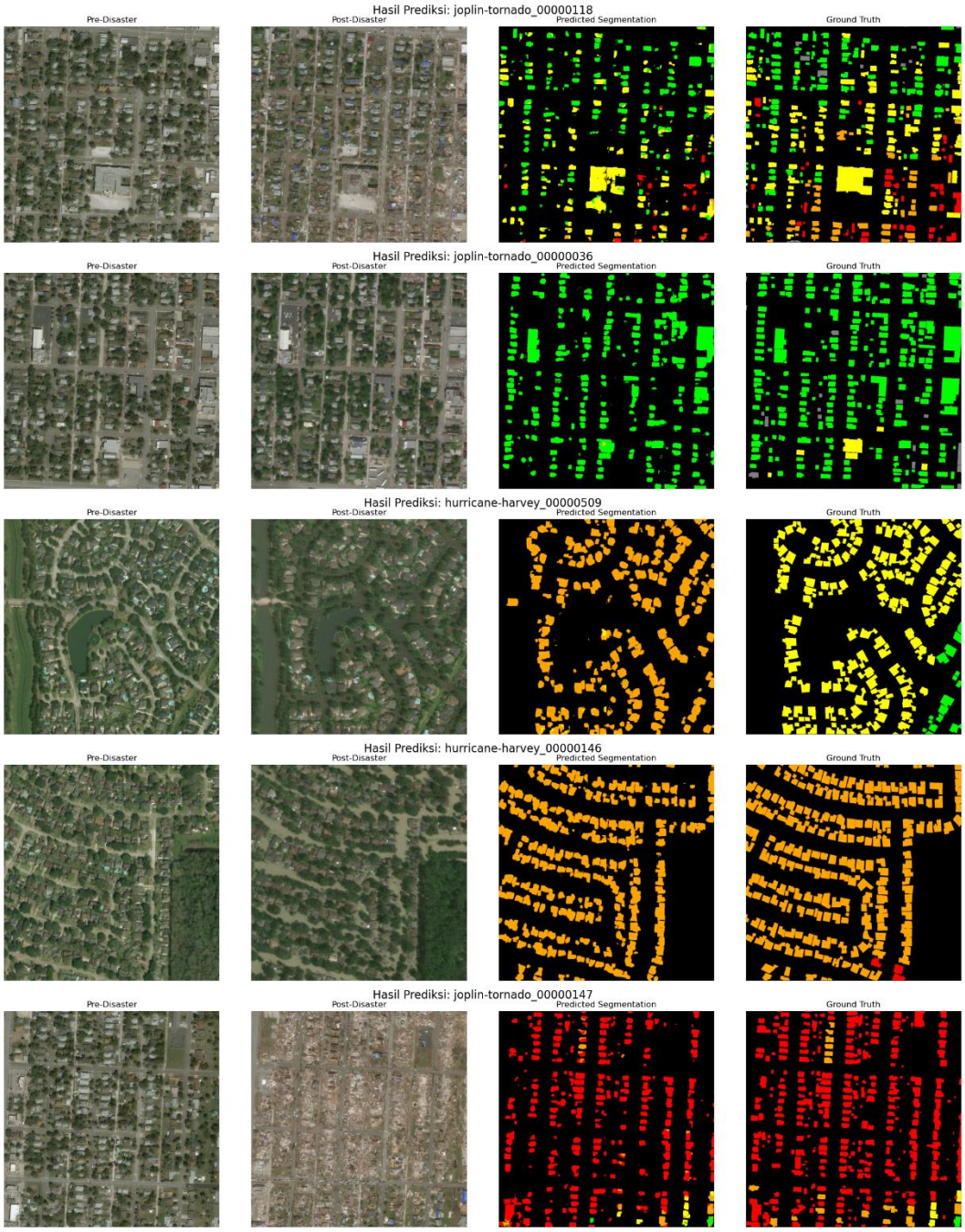
4.2.4 Skenario 4

Pada skenario ini telah dilakukan pelatihan model selama 50 epoch dengan learning rate sebesar 0.001. Loss function yang digunakan pada skenario ini menggunakan dual focal dengan *optimizer* menggunakan RMSprop. Hasil training model pada skenario pertama berhasil mendapatkan nilai sebesar *Loss*: 0.2627. Gambar grafik *loss* dapat dilihat pada Gambar 4.11.



Gambar 4.11 Grafik loss skenario 4

Selanjutnya setelah melewati proses pelatihan, model digunakan untuk melakukan prediksi gambar untuk melihat hasil segmentasi tingkat kerusakan bangunan. Telah diberikan 5 gambar yang diambil dari *data testing* sebagai contoh prediksi hasil segmentasi. Hasil dari segmentasi dapat dilihat pada Gambar 4.12.



Gambar 4.12 Hasil segmentasi skenario 4

Setelah model melakukan prediksi untuk mendapatkan hasil segmentasi, selanjutnya model akan dievaluasi dengan menggunakan *test dataset* yang telah di *split data*. Jumlah data yang digunakan pada proses *testing* ini sebanyak 1594 data. Evaluasi yang telah dilakukan adalah IoU dan F1 – score yang akan dievaluasi per

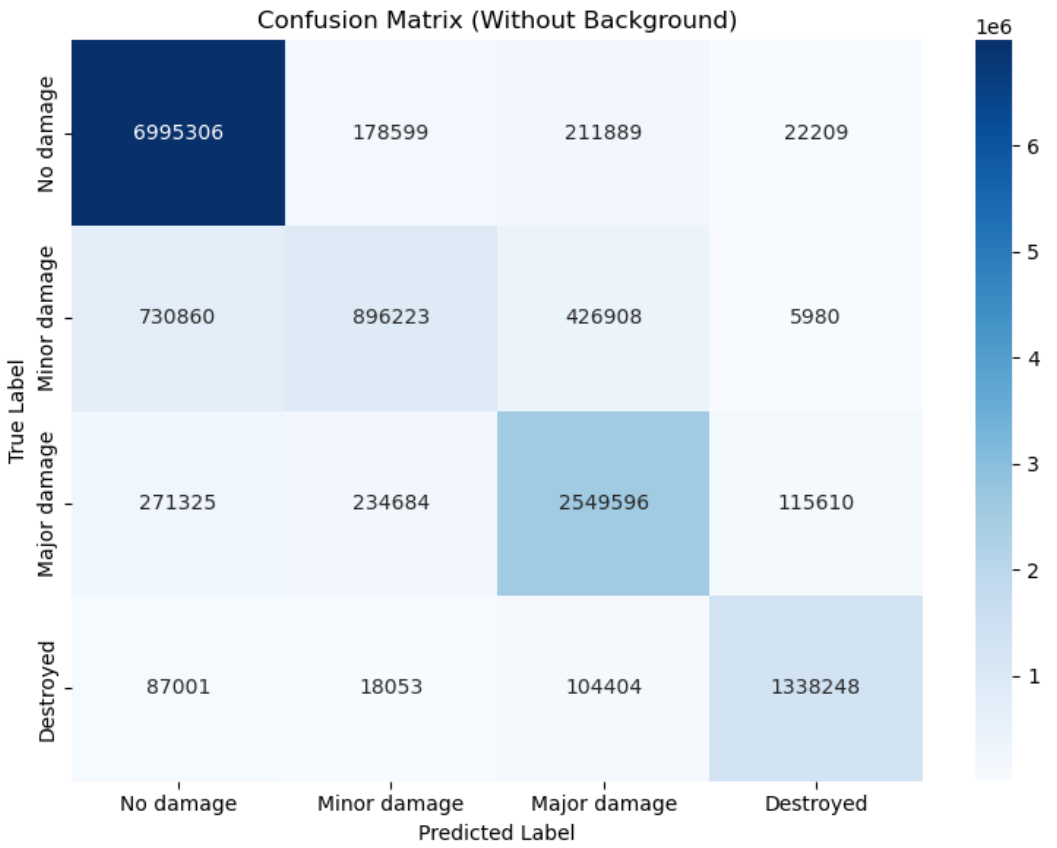
kelas. Kemudian model juga akan dievaluasi dengan melihat nilai mIoU, mF1-score dan *Overal accuracy*. Nilai IoU dan f1- score perkelas dapat dilihat pada Tabel 4.8 dan nilai mIoU, mF1 dan Overal accuracy dapat dilihat pada Tabel 4.9. Berikutnya Gambar *confussion matrix* dapat dilihat pada Gambar 4.13.

Tabel 4.8 Evaluasi model per kelas skenario 4

Number	Class	F1 - Score	IoU
1	No Damage	0.785621	0.646932
2	Minor Damage	0.378484	0.233414
3	Major Damage	0.668421	0.501977
4	Destroyed	0.739952	0.587241

Tabel 4.9 Evaluasi rerata skenario 4

Overall Accuracy (OA)	Mean F1 Score (mF1)	Mean IoU (mIoU)
0.6009	0.5145	0.3939



Gambar 4.13 *Confussion matrix* skenario 4

4.3. Analisis Awal

Setelah menyajikan hasil eksperimen dari keempat skenario pengujian untuk model usulan (U-*Net ResNet50 Pretrained with CBAM*), bagian ini akan melakukan analisis perbandingan untuk mengidentifikasi konfigurasi terbaik serta memahami pengaruh dari *loss function* dan *optimizer* yang berbeda terhadap kinerja model. Untuk memudahkan perbandingan, hasil metrik evaluasi utama dari keempat skenario dirangkum dalam Tabel 4.10.

Tabel 4.10 Rangkuman hasil pengujian kelompok skenario 1

Skenario	Kombinasi	Loss function	Optimizer	OA	mIoU	mF1-Score
1	1 – A – I	Crossentropy	Adam	0.5813	0.4854	0.3591
2	1 – B – I	Dual Focal	Adam	0.5705	0.5015	0.3781
3	1 – A – II	Crossentropy	RMSProp	**0.6227**	0.5092	0.3879
4	1 – B – II	Dual Focal	RMSProp	0.6009	**0.5145**	**0.3939**

Berdasarkan Tabel 4.10, terlihat bahwa Skenario 3 (Crossentropy + RMSprop) menghasilkan nilai *Overall Accuracy* (OA) tertinggi, yaitu 0.6227. Namun, jika dilihat dari metrik mIoU dan mF1-Score, yang seringkali dianggap lebih representatif untuk tugas segmentasi (terutama dengan potensi ketidakseimbangan kelas), Skenario 4 (Dual Focal + RMSprop) menunjukkan kinerja terbaik dengan mIoU 0.3939 dan mF1 0.5145.

Perbedaan ini menunjukkan bahwa Skenario 3 mungkin lebih baik dalam mengklasifikasikan piksel secara keseluruhan, tetapi Skenario 4 sedikit lebih unggul dalam hal keseimbangan antara presisi dan recall serta tumpang tindih antara prediksi dan *ground truth* di seluruh kelas. Kedua skenario yang menggunakan RMSprop (Skenario 3 dan 4) secara konsisten menunjukkan kinerja

yang lebih baik dibandingkan dengan dua skenario yang menggunakan Adam (Skenario 1 dan 2) pada hampir semua metrik utama.

4.3.1 Analisis Pengaruh Loss Function

Untuk menganalisis dampak *loss function*, kita membandingkan pasangan skenario dengan *optimizer* yang sama:

- Adam (Skenario 1 vs 2): Penggunaan Dual Focal Loss (Skenario 2) menghasilkan mIoU (0.5015) dan mF1 (0.3781) yang lebih tinggi dibandingkan Crossentropy (Skenario 1: mIoU 0.4854, mF1 0.3591), meskipun OA-nya sedikit lebih rendah.
- RMSprop (Skenario 3 vs 4): Hasil serupa terlihat di sini, di mana Dual Focal Loss (Skenario 4) memberikan mIoU (0.5145) dan mF1 (0.3939) yang lebih tinggi daripada Crossentropy (Skenario 3: mIoU 0.5092, mF1 0.3879), meskipun OA Skenario 3 lebih tinggi.

Hasil ini mengindikasikan bahwa Dual Focal Loss cenderung memberikan sedikit keunggulan dalam hal mIoU dan mF1 untuk model usulan pada dataset ini. Hal ini mungkin disebabkan oleh kemampuan Dual Focal Loss untuk memberikan bobot lebih pada sampel yang sulit atau kelas yang kurang terwakili, yang relevan dalam konteks penilaian kerusakan bangunan. Meskipun hasil *Overall Accuracy* dari model yang menggunakan Dual Focal sedikit lebih kecil ketimbang model yang menggunakan crossentropy.

4.3.2 Analisis Pengaruh Optimizer

Untuk menganalisis dampak *optimizer*, kita membandingkan pasangan skenario dengan *loss function* yang sama:

- Crossentropy (Skenario 3 vs 1): RMSprop (Skenario 3) secara signifikan mengungguli Adam (Skenario 1) di semua metrik (OA: 0.6227 vs 0.5813, mF1: 0.5092 vs 0.4854, mIoU: 0.3879 vs 0.3591).
- Dual Focal Loss (Skenario 4 vs 2): RMSprop (Skenario 4) juga mengungguli Adam (Skenario 2) di semua metrik (OA: 0.6009 vs 0.5705, mF1: 0.5145 vs 0.5015, mIoU: 0.3939 vs 0.3781).

Hasil ini menunjukkan bahwa RMSprop secara konsisten memberikan kinerja yang lebih baik dibandingkan Adam dalam keempat skenario pengujian untuk model usulan ini. *Optimizer* RMSprop tampaknya lebih efektif dalam menavigasi *loss landscape* untuk mencapai konvergensi yang lebih baik dalam 50 epoch pelatihan yang dilakukan.

4.3.3 Analisis Kinerja Per Kelas

Melihat tabel evaluasi per kelas, beberapa poin penting dapat diamati. Analisis ini diperkuat dengan mengamati *confusion matrix* untuk setiap skenario yaitu pada Gambar 4.4, Gambar 4.7, Gambar 4.10, dan Gambar 4.13. Berdasarkan confussion matrix berikut adalah poin – poin hasil analisis :

- Kelas Minor Damage

Seperti yang ditunjukkan oleh skor F1 dan IoU yang rendah, *confusion matrix* secara visual mengkonfirmasi bahwa *'Minor Damage'* adalah kelas yang paling

menantang. Misalnya, pada Gambar 4.13 confussion matrix skenario 4, terlihat bahwa sejumlah besar piksel *'Minor Damage'* (730,860) salah diklasifikasikan sebagai *'No Damage'*, dan sejumlah besar lainnya (426,908) salah diklasifikasikan sebagai *'Major Damage'*. Pola kebingungan antara *'Minor Damage'* dengan *'No Damage'* dan *'Major Damage'* ini terlihat konsisten di keempat skenario, meskipun Skenario 3 dan 4 (dengan RMSprop) tampak sedikit lebih baik dalam membedakannya dibandingkan Skenario 1 dan 2.

- Kelas No Damage

Pada Gambar 4.10 confussion matrix skenario 3 menunjukkan nilai *True Positive* tertinggi untuk kelas *'No Damage'* (7,791,290), yang menjelaskan tingginya nilai OA pada skenario ini. Namun, Gambar 4.13 *confussion matrix* skenario 4 juga menunjukkan kinerja yang kuat (6,995,306) dan tampak memiliki *false negative* yang lebih sedikit (yaitu, lebih sedikit *'No Damage'* yang salah diklasifikasikan sebagai kelas lain) dibandingkan Skenario 1 dan 2.

- Kelas Major & Destroyed

Skenario 4 pada Gambar 4.13 yaitu *confussion matrix* skenario 4, menunjukkan nilai *True Positive* yang relatif kuat untuk *'Major Damage'* (2,549,596) dan '*Destroyed*' (1,338,248). Jika dibandingkan dengan Skenario 3, Skenario 4 tampak lebih baik dalam membedakan *'Major Damage'* dan *'Destroyed',* di mana Skenario 3 cenderung lebih banyak salah mengklasifikasikan *'Major Damage'* ke kelas lain (terutama *'No Damage'*). Sedangkan untuk skenario 1 memiliki *true positive* terbanyak untuk kelas '*Destroyed*' dan kelas '*Major Damaged'* jika dibandingkan

dengan skenario 2. Hal ini menarik mengingat skenario 1 dan skenario 4 memiliki susunan hyperparameter yang saling berbeda.

- Pola Kesalahan Umum

Secara umum, *confusion matrix* menunjukkan bahwa kesalahan klasifikasi paling sering terjadi antara kelas-kelas yang berdekatan tingkat kerusakannya (misalnya, *'No Damage' <-> 'Minor Damage'*, *'Minor Damage' <-> 'Major Damage'*). Namun, pada Gambar 4.4 confussion matrix skenario 1 menunjukkan kebingungan yang cukup tinggi bahkan antara kelas yang jauh, seperti *'No Damage'* yang salah diklasifikasikan sebagai 'Major Damage' (852,844). Bahkan untuk klasifikasi kelas *minor damage* (933,972) lebih kecil ketimbang terklasifikasi *major damage* (101,362). Hal ini menunjukkan bahwa Skenario 1 memiliki kinerja yang paling lemah.

4.4 Kesimpulan Analisis Awal

Berdasarkan analisis (metrik OA, mIoU, mF1) dan diperkuat oleh *confusion matrix*, Skenario 4 (Dual Focal Loss + RMSprop) menunjukkan kinerja paling seimbang dan menjanjikan secara keseluruhan, terutama jika mIoU dan mF1 dijadikan acuan utama. Meskipun Skenario 3 unggul di OA karena tingginya *True Positive 'No Damage'*, *confusion matrix* Skenario 4 menunjukkan distribusi prediksi yang lebih baik di kelas 'Major' dan 'Destroyed', yang penting untuk penilaian risiko bencana. Oleh karena itu, Skenario 4 dipilih sebagai konfigurasi terbaik untuk model usulan (U-Net ResNet50 CBAM) dan akan digunakan sebagai representasi model usulan dalam bab pembahasan (Bab 6) untuk dibandingkan

dengan model pembanding. Meskipun demikian, perlu dicatat bahwa kelas *'Minor Damage'* tetap menjadi tantangan bagi model usulan ini.

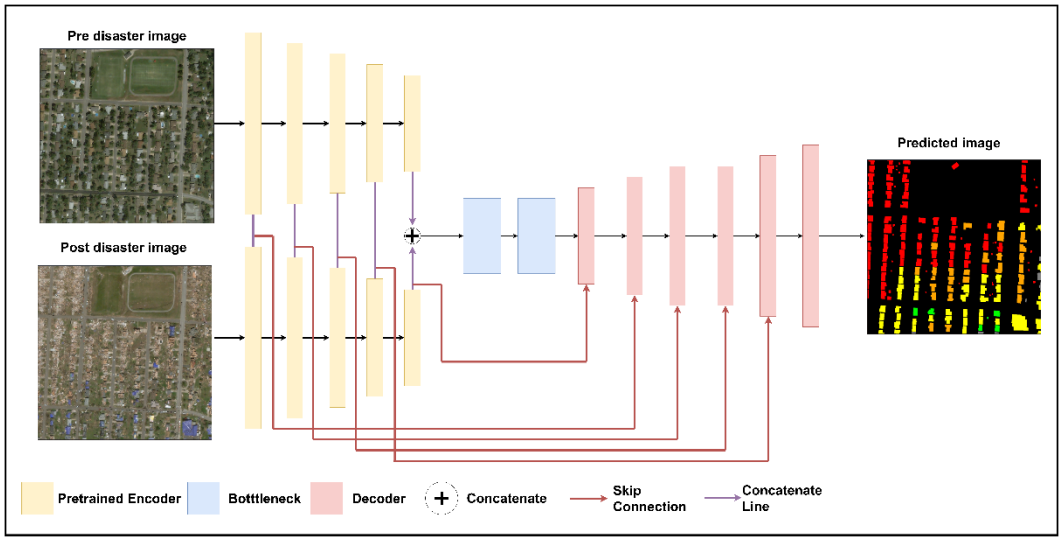
BAB V

DEFAULT UNET WITH PRETRAINED ENCODER

Pada bab ini akan disajikan hasil pelatihan model berdasarkan skenario pengujian yang telah disusun yang terdiri dari loss dari setiap skenario. Berikutnya, hasil prediksi segmentasi akan disajikan berupa gambar yang telah tersegmentasi. Kemudian hasil evaluasi model akan ditampilkan nilai evaluasi setiap skenario pengujian. Terakhir akan disajikan analisis awal pengujian.

5.1 Model Arsitektur

Model yang digunakan pada bab ini adalah model pembanding yang diajukan penulis. Model ini merupakan model *default* unet dengan penambahan penggunaan *pretrained* encoder. Gambar arsitektur model dapat dilihat pada Gambar 5.1.



Gambar 5.1 Model pembanding

5.2. Uji Coba

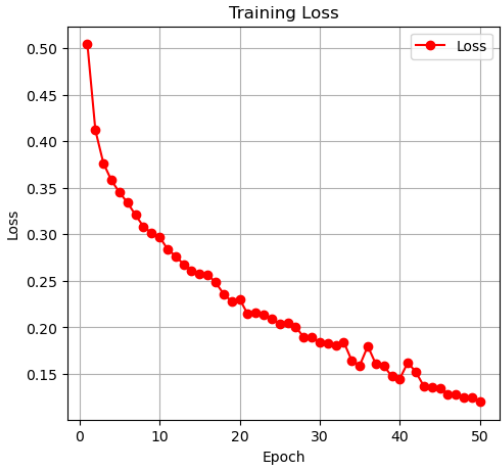
Pada bab ini akan melakukan 4 skenario uji coba dari total 8 skenario uji coba yang sudah dijelaskan pada sub bab 3.5. Skenario yang akan dilakukan dapat dilihat pada Tabel 5.1.

Tabel 5.1 Kombinasi skenario kelompok 2

Skenario	Kombinasi	Keterangan
5	2-A-I	Crossentropy with adam
6	2-B-I	Dual focal with adam
7	2-A-II	Crossentropy with RMSprop
8	2-B-II	Dual focal with RMSprop

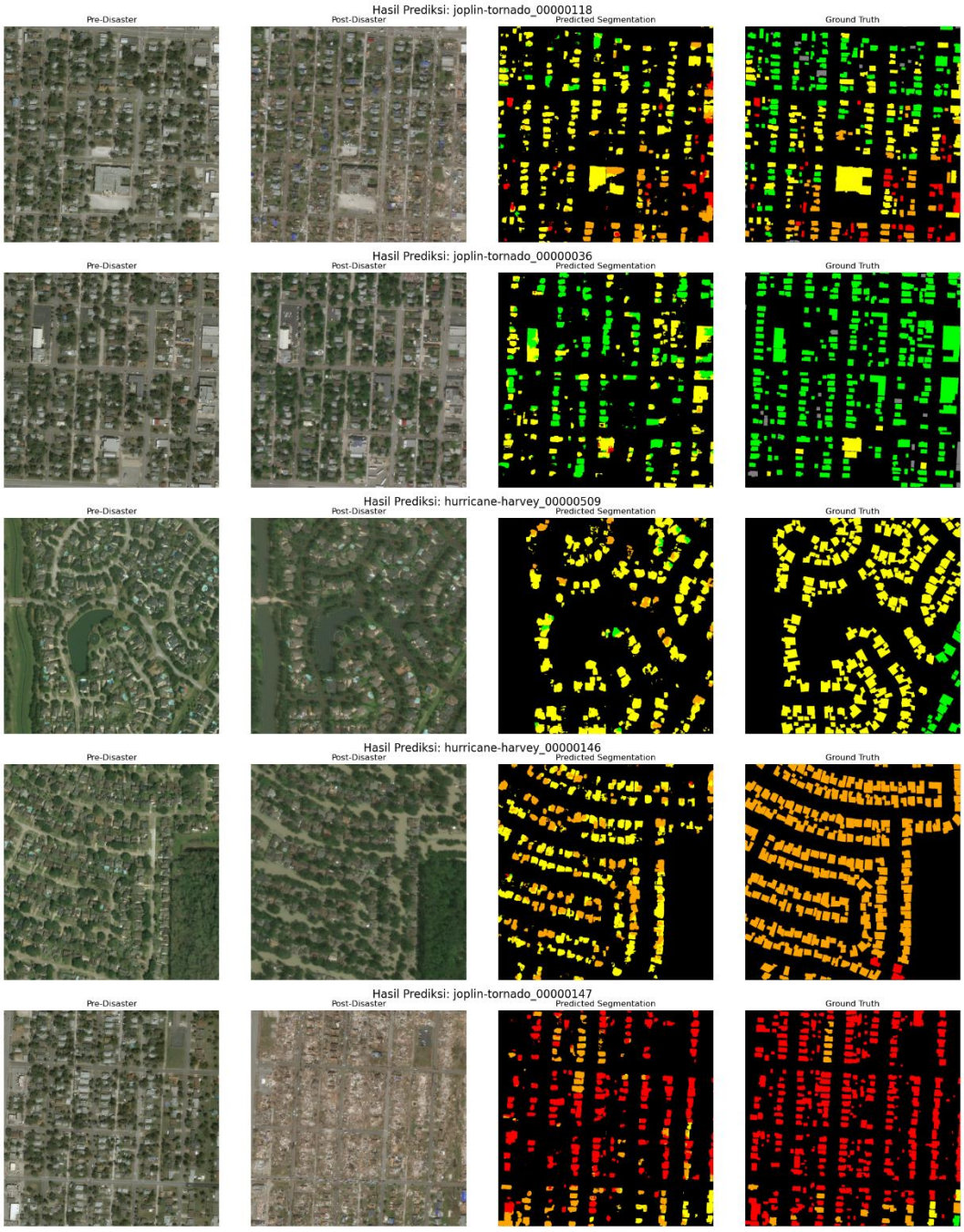
5.2.1 Skenario 5

Pada skenario ini telah dilakukan pelatihan model selama 50 epoch dengan learning rate sebesar 0.001. *Loss function* yang digunakan pada skenario ini menggunakan crossentropy dengan *optimizer* menggunakan adam. Hasil *training* model pada skenario pertama berhasil mendapatkan nilai sebesar *Loss:* 0.1200. Gambar grafik *loss* dapat dilihat pada Gambar 5.2.



Gambar 5.2 Grafik loss skenario 5

Selanjutnya setelah melewati proses pelatihan, model digunakan untuk melakukan prediksi gambar untuk melihat hasil segmentasi tingkat kerusakan bangunan. Telah diberikan 5 gambar yang diambil dari *data testing* sebagai contoh prediksi hasil segmentasi. Hasil dari segmentasi dapat dilihat pada Gambar 5.3.



Gambar 5.3 Hasil segmentnasi skenario 5

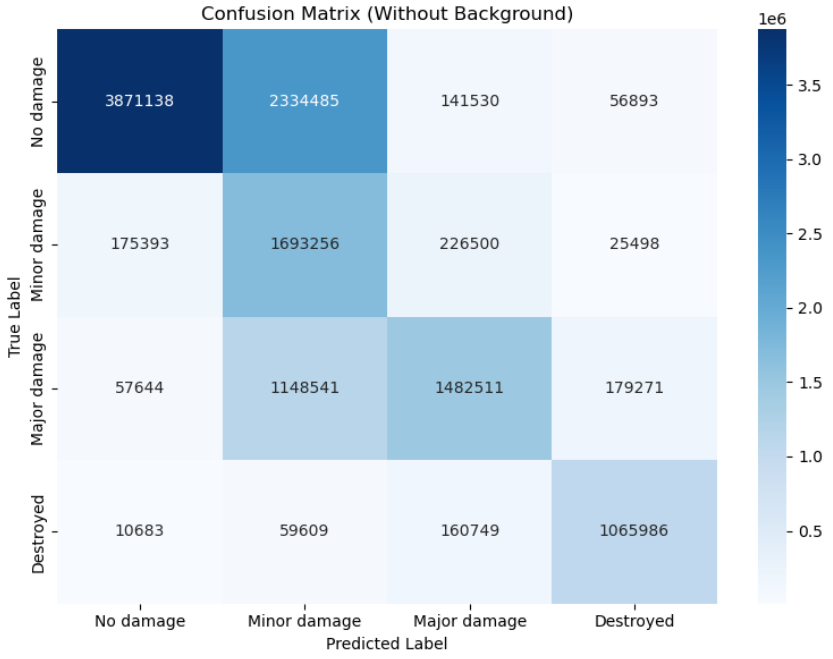
Setelah model melakukan prediksi untuk mendapatkan hasil segmentasi, selanjutnya model akan dievaluasi dengan menggunakan *test dataset* yang telah di split data. Jumlah data yang digunakan pada proses *testing* ini sebanyak 1594 data. Evaluasi yang telah dilakukan adalah IoU dan F1 – score yang akan dievaluasi per kelas. Kemudian model juga akan dievaluasi dengan melihat nilai mIoU, mF1-score dan *Overal accuracy*. Nilai IoU dan f1- score perkelas dapat dilihat pada Tabel 5.2 dan nilai mIoU, mF1 dan Overal accuracy dapat dilihat pada Tabel 5.3. Berikutnya Gambar *confussion matrix* dapat dilihat pada Gambar 5.4.

Tabel 5.2 Evaluasi model per kelas skenario 5

Number	Class	F1 - Score	IoU
1	No Damage	0.559465	0.388373
2	Minor Damage	0.391768	0.243601
3	Major Damage	0.467139	0.304749
4	Destroyed	0.615692	0.444766

Tabel 5.3 Evaluasi rerata skenario 5

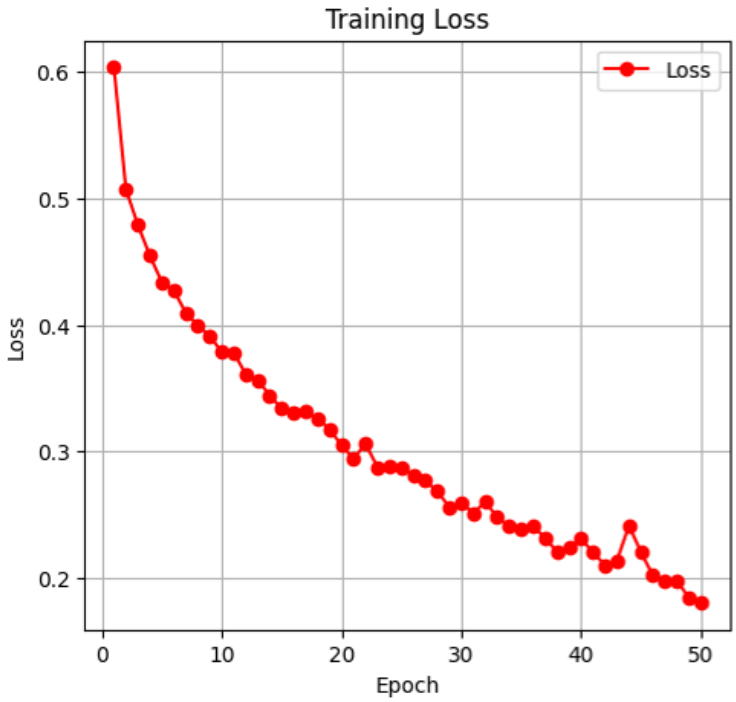
Overall Accuracy (OA)	Mean F1 Score (mF1)	Mean IoU (mIoU)
0.4139	0.4068	0.2763



Gambar 5.4 *Confussion matrix* skenario 5

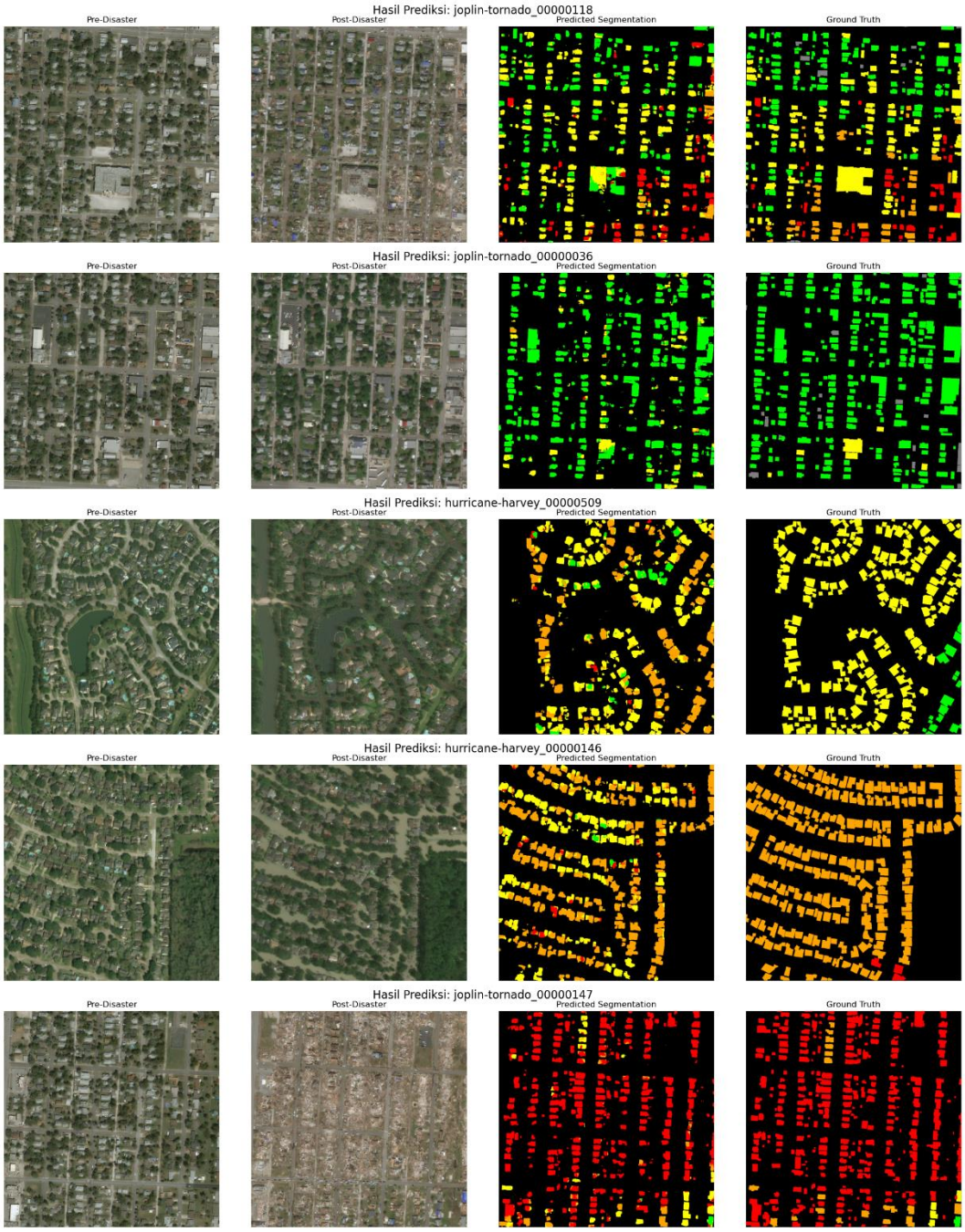
5.2.2 Skenario 6

Pada skenario ini telah dilakukan pelatihan model selama 50 epoch dengan learning rate sebesar 0.001. *Loss function* yang digunakan pada skenario ini menggunakan dual focal loss dengan optimizer menggunakan adam. Hasil *training* model pada skenario pertama berhasil mendapatkan nilai sebesar *Loss*: 0.1805. Gambar grafik *loss* dapat dilihat pada Gambar 5.5.



Gambar 5.5 Grafik loss skenario 6

Selanjutnya setelah melewati proses pelatihan, model digunakan untuk melakukan prediksi gambar untuk melihat hasil segmentasi tingkat kerusakan bangunan. Telah diberikan 5 gambar yang diambil dari *data testing* sebagai contoh prediksi hasil segmentasi. Hasil dari segmentasi dapat dilihat pada Gambar 5.6



Gambar 5.6 Hasil segmentasi skenario 6

Setelah model melakukan prediksi untuk mendapatkan hasil segmentasi, selanjutnya model akan dievaluasi dengan menggunakan *test dataset* yang telah di split data. Jumlah data yang digunakan pada proses *testing* ini sebanyak 1594 data. Evaluasi yang telah dilakukan adalah IoU dan F1 – score yang akan dievaluasi per

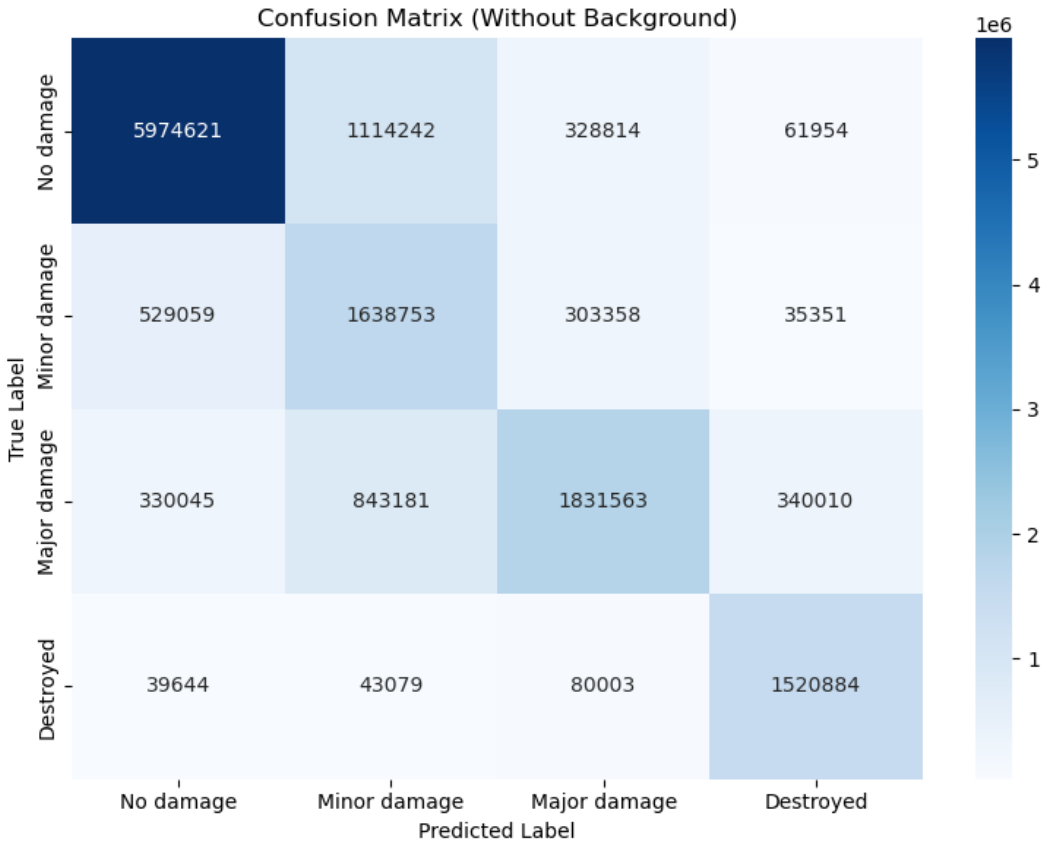
kelas. Kemudian model juga akan dievaluasi dengan melihat nilai mIoU, mF1-score dan *Overal accuracy*. Nilai IoU dan f1- score perkelas dapat dilihat pada Tabel 5.4 dan nilai mIoU, mF1 dan Overal accuracy dapat dilihat pada Tabel 5.5. Berikutnya Gambar *confussion matrix* dapat dilihat pada Gambar 5.7.

Tabel 5.4 Evaluasi model per kelas skenario 6

Number	Class	F1 - Score	IoU
1	No Damage	0.719954	0.562444
2	Minor Damage	0.465056	0.302979
3	Major Damage	0.532458	0.362823
4	Destroyed	0.743114	0.591234

Tabel 5.5 Evaluasi rerata skenario 6

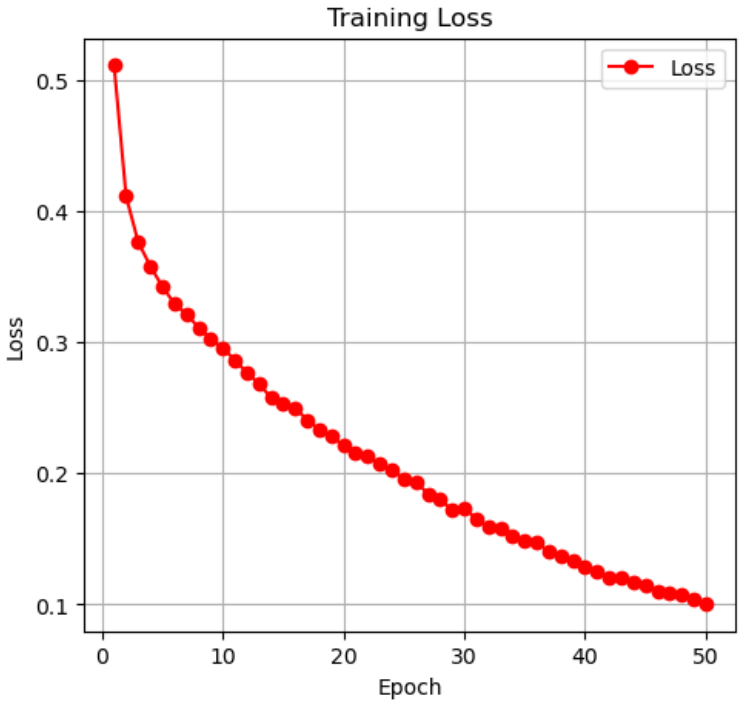
Overall Accuracy (OA)	Mean F1 Score (mF1)	Mean IoU (mIoU)
0.5594	0.4921	0.3639



Gambar 5.7 *Confussion matrix* skenario 6

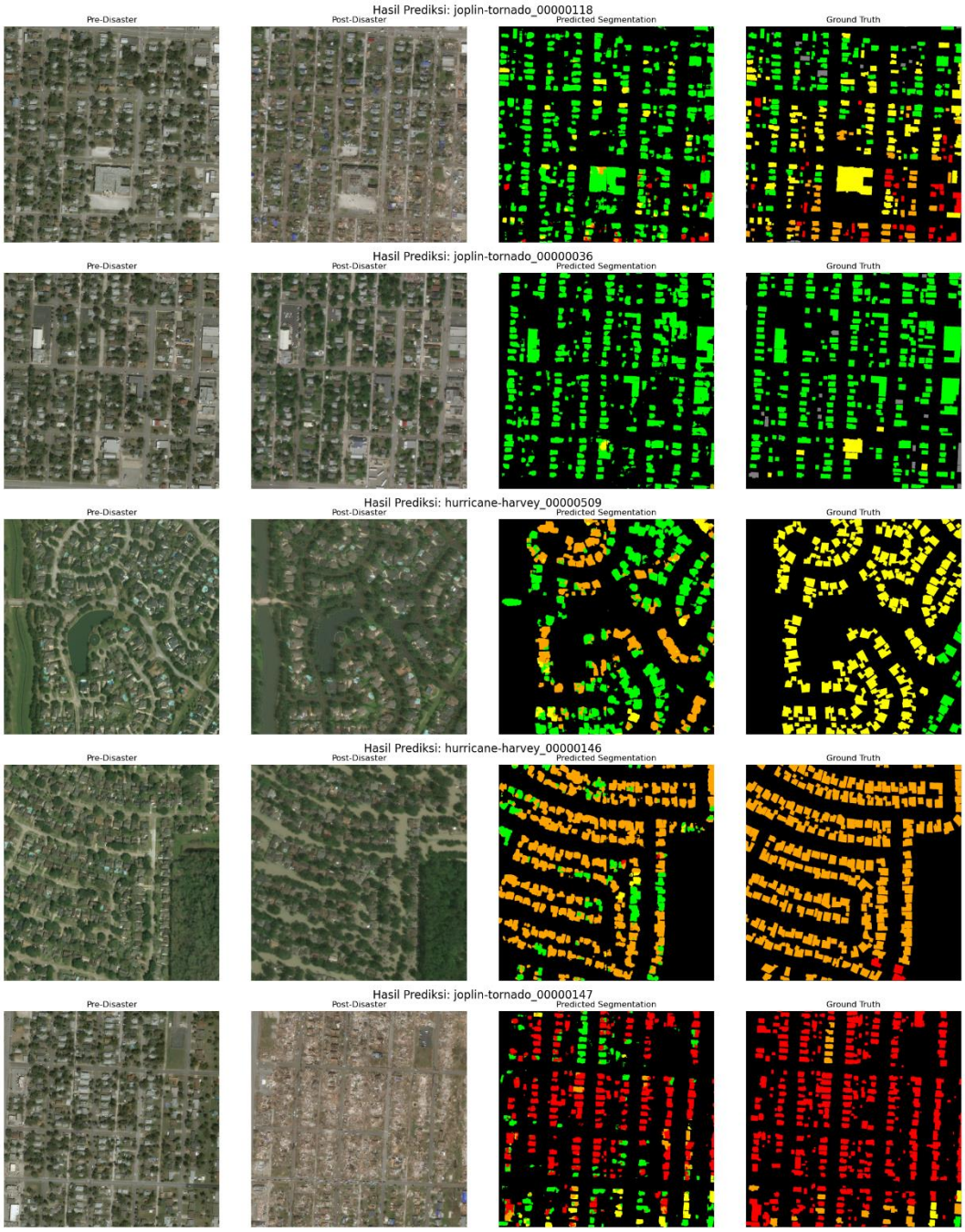
5.2.2 Skenario 7

Pada skenario ini telah dilakukan pelatihan model selama 50 epoch dengan learning rate sebesar 0.001. Loss function yang digunakan pada skenario ini menggunakan dual crossentropy dengan *optimizer* menggunakan RMSprop. Hasil *training* model pada skenario pertama berhasil mendapatkan nilai sebesar *Loss*: 0.1002. Gambar grafik *loss* dapat dilihat pada Gambar 5.8.



Gambar 5.8 Grafik loss skenario 7

Selanjutnya setelah melewati proses pelatihan, model digunakan untuk melakukan prediksi gambar untuk melihat hasil segmentasi tingkat kerusakan bangunan. Telah diberikan 5 gambar yang diambil dari data *testing* sebagai contoh prediksi hasil segmentasi. Hasil dari segmentasi dapat dilihat pada Gambar 5.9.



Gambar 5.9 Hasil segmentasi skenario 7

Setelah model melakukan prediksi untuk mendapatkan hasil segmentasi, selanjutnya model akan dievaluasi dengan menggunakan *test dataset* yang telah di split data. Jumlah data yang digunakan pada proses *testing* ini sebanyak 1594 data. Evaluasi yang telah dilakukan adalah IoU dan F1 – score yang akan dievaluasi per

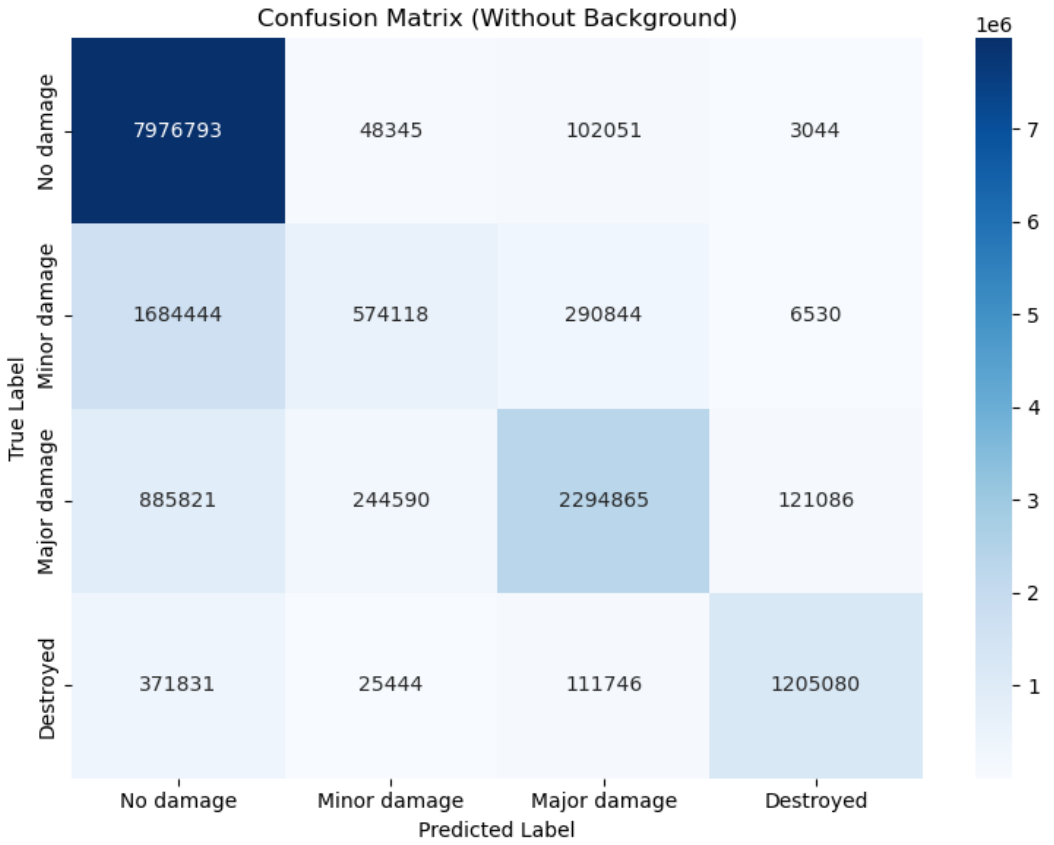
kelas. Kemudian model juga akan dievaluasi dengan melihat nilai mIoU, mF1-score dan *Overal accuracy*. Nilai IoU dan f1- score perkelas dapat dilihat pada Tabel 5.6 dan nilai mIoU, mF1 dan Overal accuracy dapat dilihat pada Tabel 5.7. Berikutnya Gambar *confussion matrix* dapat dilihat pada Gambar 5.10.

Tabel 5.6 Evaluasi moddel per kelas skenario 7

Number	Class	F1 - Score	IoU
1	No Damage	0.772842	0.629782
2	Minor Damage	0.266983	0.154057
3	Major Damage	0.643232	0.474092
4	Destroyed	0.694408	0.531872

Tabel 5.7 Evaluasi rerata skenario 7

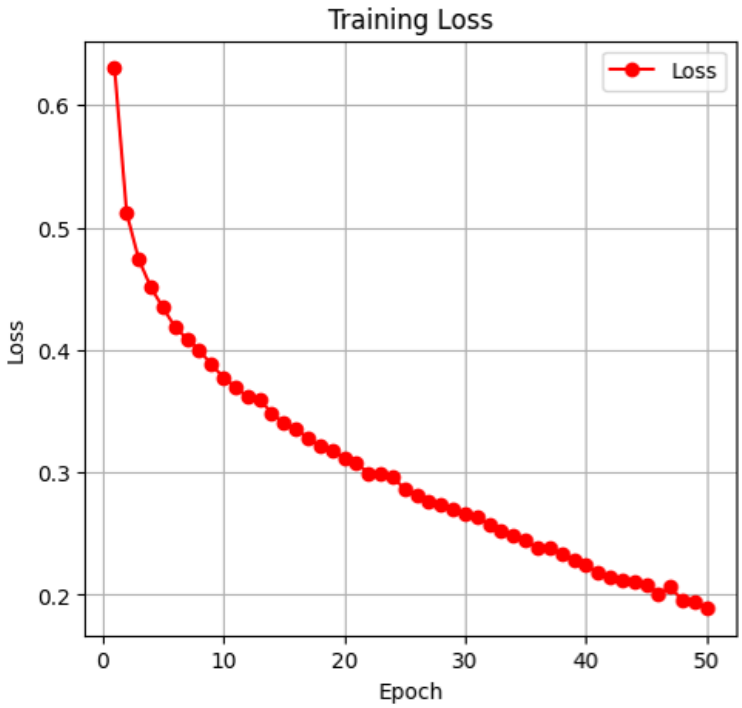
Overall Accuracy (OA)	Mean F1 Score (mF1)	Mean IoU (mIoU)
0.6147	0.4755	0.3580



Gambar 5.10 *Confussion matrix* skenario 7

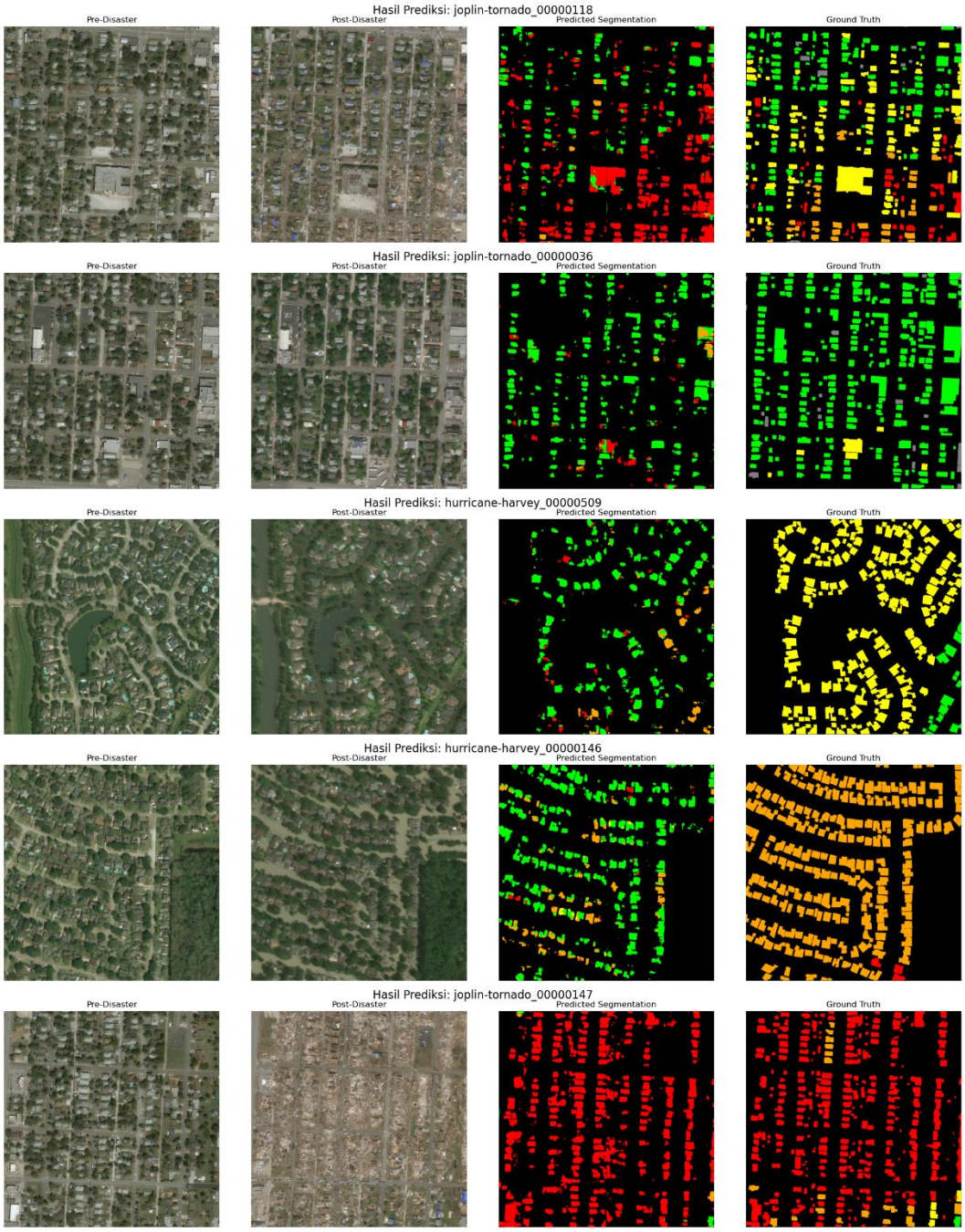
5.2.2 Skenario 8

Pada skenario ini telah dilakukan pelatihan model selama 50 epoch dengan learning rate sebesar 0.001. *Loss function* yang digunakan pada skenario ini menggunakan dual dual focal dengan *optimizer* menggunakan RMSprop. Hasil *training* model pada skenario pertama berhasil mendapatkan nilai sebesar *Loss*: 0.1889. Gambar grafik *loss* dapat dilihat pada Gambar 5.11.



Gambar 5.11 Grafik loss skenario 8

Selanjutnya setelah melewati proses pelatihan, model digunakan untuk melakukan prediksi gambar untuk melihat hasil segmentasi tingkat kerusakan bangunan. Telah diberikan 5 gambar yang diambil dari *data testing* sebagai contoh prediksi hasil segmentasi. Hasil dari segmentasi dapat dilihat pada Gambar 5.12.



Gambar 5.12 Hasil segmentasi skenario 8

Setelah model melakukan prediksi untuk mendapatkan hasil segmentasi, selanjutnya model akan dievaluasi dengan menggunakan *test dataset* yang telah di split data. Jumlah data yang digunakan pada proses *testing* ini sebanyak 1594 data. Evaluasi yang telah dilakukan adalah IoU dan F1 – score yang akan dievaluasi per

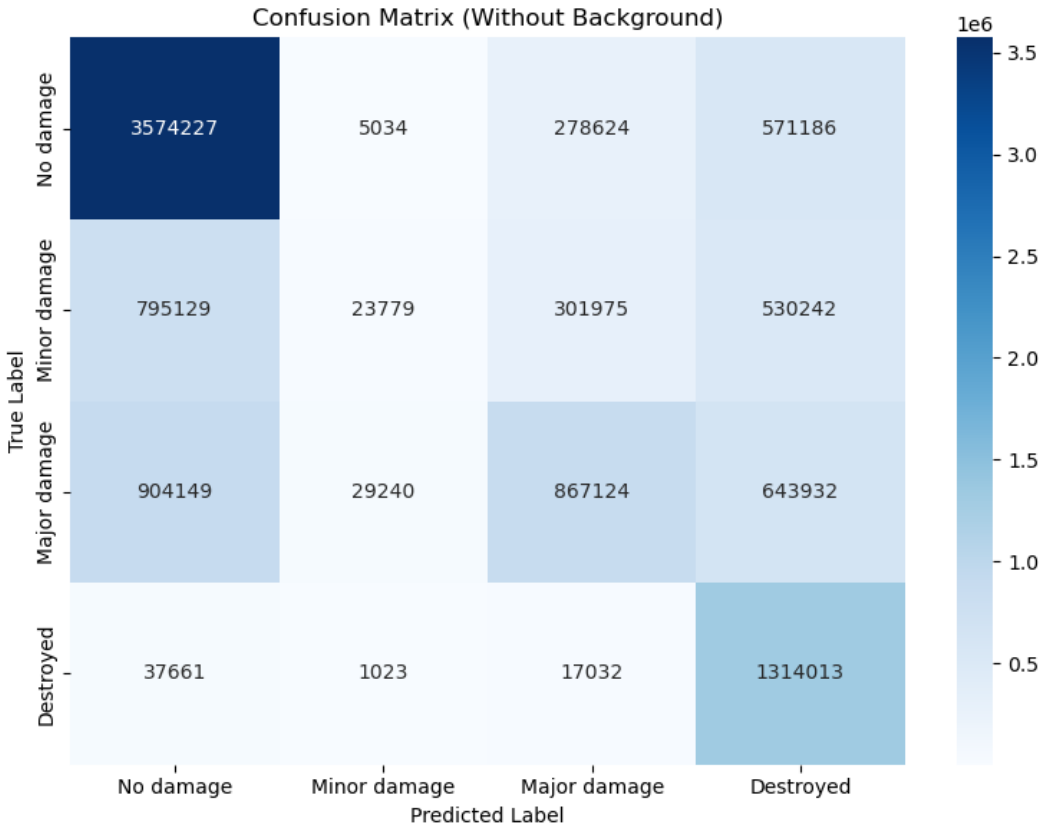
kelas. Kemudian model juga akan dievaluasi dengan melihat nilai mIoU, mF1-score dan *Overal accuracy*. Nilai IoU dan f1- score perkelas dapat dilihat pada Tabel 5.8 dan nilai mIoU, mF1 dan *Overal accuracy* dapat dilihat pada Tabel 5.9. Berikutnya Gambar *confussion matrix* dapat dilihat pada Gambar 5.13.

Tabel 5.8 Evaluasi model per kelas skenario 8

Number	Class	F1 - Score	IoU
1	No Damage	0.475453	0.311865
2	Minor Damage	0.013716	0.006905
3	Major Damage	0.298974	0.175761
4	Destroyed	0.505930	0.338625

Tabel 5.9 Evaluasi rerata skenario 8

Overall Accuracy (OA)	Mean F1 Score (mF1)	Mean IoU (mIoU)
0.2948	0.2588	0.1666



Gambar 5.13 *Confussion matrix* skenario 8

5.4. Analisis Awal

Bagian ini menyajikan analisis mendalam terhadap hasil pengujian model pembanding (U-Net dengan *backbone* ResNet50 *pretrained*, tanpa mekanisme atensi CBAM). Analisis ini bertujuan untuk mengevaluasi kinerja model pembanding dalam empat skenario pengujian yang sama (Skenario 5 hingga 8) seperti yang diterapkan pada model usulan. Tujuannya adalah untuk memahami perilaku model pembanding dan mengidentifikasi konfigurasi terbaiknya, yang nantinya akan digunakan sebagai *baseline* dalam perbandingan dengan model usulan di Bab 6. Ringkasan hasil metrik evaluasi utama dari Skenario 5 hingga 8 disajikan dalam Tabel 5.10.

Tabel 5.10 Rangkuman hasil pengujian kelompok skenario 2

Skenario	Kombinasi	Loss function	Optimizer	OA	mIoU	mF1-Score
5	2 – A – I	Crossentropy	Adam	0.4139	0.4068	0.2763
6	2 – B – I	Dual Focal	Adam	0.5594	**0.4921**	**0.3639**
7	2 – A – II	Crossentropy	RMSProp	**0.6147**	0.4755	0.3580
8	2 – B – II	Dual Focal	RMSProp	0.2948	0.2588	0.1666

5.4.1 Analisis Pengaruh Loss Function

Pengaruh *loss function* pada model pembanding tampaknya sangat bergantung pada *optimizer* yang digunakan:

- Dengan Adam (Skenario 5 vs 6)

Dual Focal Loss (Skenario 6) secara signifikan meningkatkan kinerja dibandingkan Crossentropy (Skenario 5). Terjadi peningkatan mIoU dari 0.4068 untuk skenario 5 dan 0.4921 untuk skenario 6. OA pada skenario 6 pun lebih unggul ketimbang skenario 5 yaitu 0.5594 untuk skenario 6 dan 0.4139 untuk skenario 5. Ini menunjukkan bahwa dengan Adam, Dual Focal Loss sangat efektif pada model ini.

- Dengan RMSprop (Skenario 7 vs 8)

Situasi berbalik drastis. Dual Focal Loss (Skenario 8) justru menurunkan kinerja secara dramatis dibandingkan Crossentropy (Skenario 7). Terjadi penurunan mIoU yaitu 0.4755 untuk skenario 7 dan 0.2588 untuk skenario 8. Pada nilai mF1 juga demikian, nilai 0.3580 untuk skenario 7 dan 0.1666 untuk skenario 8. Namun, OA pada skenario 7 memiliki nilai tertinggi pada keempat skenario yaitu 0.6147. Ini menunjukkan bahwa dengan RMSProp, Dual focal memilliki hasil yang sangat buruk, sedangkan gabungan crossentropy dan RMSProp memberikan nilai OA yang tertinggi serta model dengan crossentropy memiliki nilai yang lebih stabil ketimbang model yang menggunakan dual focal jika dipasangkan dengan RMSProp.

Kesimpulan ini menyoroti bahwa tidak ada *loss function* yang secara universal lebih baik untuk model pembanding; kinerjanya sangat sensitif terhadap pilihan *optimizer*.

5.4.2 Analisis Pengaruh Optimizer

Analisis pengaruh *optimizer* juga menunjukkan hasil yang bergantung pada *loss function*:

- Dengan Crossentropy (Skenario 7 vs 5)

RMSprop (Skenario 7) meningkatkan kinerja secara signifikan dibandingkan Adam (Skenario 5), dengan peningkatan mIoU yaitu 0.4755 untuk skenario 7 sedangkan skenario 5 hanya 0.4068. Begitu pula pada nilai OA yang mana skenario dengan RMSProp (skenario 7) memiliki nilai tertinggi yaitu 0.6147.

- Dengan Dual Focal Loss (Skenario 6 vs 8)

Adam (Skenario 6) justru jauh lebih unggul daripada RMSprop (Skenario 8) dalam hal mIoU dan mF1. Dapat dililhat dari nilai mIoU skenario 6 yaitu 0.4921 yang lebih unggul ketimbang skenario 8 yaitu 0.2588. Begitu pula pada nilai mF1 nya yaitu 0.3639 untuk skenario 6 dan 0.1666 untuk skenario 8.

Hal ini kembali menekankan adanya interaksi yang kuat dan tidak terduga, terutama kegagalan kombinasi Dual Focal Loss dan RMSprop pada model pembanding ini. Mungkin RMSprop dengan Dual Focal Loss mengalami kesulitan konvergensi atau terjebak dalam *local minimum* yang buruk pada arsitektur tanpa CBAM.

5.4.3 Analisa Kinerja Per Kelas

Analisis *confusion matrix* Melihat tabel evaluasi per kelas (Tabel 5.2, 5.4, 5.6, 5.8) dan diperkuat dengan pengamatan *confusion matrix* untuk setiap skenario model pembanding Gambar 5.4, 5.7, 5.10, dan 5.13, beberapa poin penting mengenai kinerja per kelas dapat diidentifikasi. Berikut adalah poin-poin hasil analisis:

- Kelas Minor Damage

Kelas ini terbukti menjadi tantangan yang sangat signifikan bagi model pembanding, bahkan lebih sulit dibandingkan model usulan. Kinerja terburuk terlihat pada Skenario 8 Gambar 5.13, yang hampir gagal total mengidentifikasi kelas ini dengan *True Positive* (TP) hanya 23.779. Skenario 7 Gambar 5.10 juga sangat lemah dengan TP 574.118, namun dengan misklasifikasi masif ke *'No Damage'* (1.684.444). Skenario 6 Gambar 5.7 menunjukkan TP tertinggi untuk

kelas ini (1.638.753), menjadikannya yang terbaik di antara model pembanding, meskipun ia juga menderita *false positive* yang tinggi dari kelas *'No Damage'* (1.114.242) dan *'Major Damage'* (843.181), yang menunjukkan adanya kebingungan. Skenario 5 Gambar 5.4 juga kesulitan, dengan kecenderungan besar untuk salah mengklasifikasikan 'No Damage' sebagai *'Minor Damage'* (2.334.485).

- Kelas No Damage

Skenario 7 Gambar 5.10 menunjukkan dominasi absolut dalam mengidentifikasi kelas *'No Damage'*, dengan TP mencapai 7.976.793. Ini menjelaskan mengapa Skenario 7 memiliki OA tertinggi di antara semua skenario model pembanding. Skenario 6 Gambar 5.7 juga menunjukkan kinerja yang kuat (5.974.621 TP), namun tidak setinggi S7. Sebaliknya, Skenario 5 (3.871.138 TP) dan terutama Skenario 8 (3.574.227 TP) menunjukkan kinerja yang jauh lebih rendah pada kelas ini.

- Kelas Major & Destroyed

Untuk kelas *'Major Damage'*, Skenario 7 Gambar 5.10 secara mengejutkan menunjukkan TP tertinggi (2.294.865), diikuti oleh Skenario 6 (1.831.563). Namun, Skenario 6 menunjukkan TP yang lebih tinggi untuk kelas *'Destroyed'* (1.520.884) dibandingkan Skenario 7 (1.205.080). Skenario 5 dan 8 menunjukkan kinerja yang jauh lebih lemah pada kedua kelas kerusakan parah ini. Ini menunjukkan adanya trade-off antara Skenario 6 dan Skenario 7. Skenario 7 baik di *'No Damage*' dan *'Major damage'*, sementara Sskenario 6 lebih baik di *'Destroyed'* dan *'Minor damage'*, menjadikannya lebih seimbang secara keseluruhan (seperti yang ditunjukkan oleh mIoU/mF1).

- Pola Kesalahan Umum & Anomali Skenario 8

Secara umum, model pembanding juga menunjukkan pola kesalahan antara kelas-kelas yang berdekatan. Namun, yang paling menonjol adalah kegagalan Skenario 8 Gambar 5.13. *Confusion matrix*-nya menunjukkan diagonal yang sangat lemah dan misklasifikasi yang tersebar luas, mengindikasikan bahwa kombinasi Dual Focal Loss dengan RMSprop sangat tidak stabil atau tidak cocok untuk arsitektur U-Net ResNet50 tanpa CBAM. Selain itu, Skenario 5 Gambar 5.4 menunjukkan kecenderungan unik untuk sangat banyak memprediksi kelas *'Minor Damage'*, yang kemungkinan besar disebabkan oleh *false positive* yang tinggi. Skenario 7 Gambar 5.10 menunjukkan bias kuat ke arah *'No Damage'*, yang meningkatkan OA tetapi mengorbankan akurasi kelas lain, terutama 'Minor Damage'. Skenario 6, meskipun tidak sempurna, tampak paling sedikit menunjukkan bias ekstrem atau kegagalan total.

5.4 Kesimpulan Analisis Awal

Analisis terhadap model pembanding menunjukkan perilaku yang berbeda dan terkadang tidak terduga dibandingkan model usulan. Temuan utama adalah:

1. Tidak ada *optimizer* atau *loss function* yang secara konsisten unggul; kinerjanya sangat bergantung pada kombinasi keduanya.
2. Kombinasi Dual Focal Loss dengan RMSprop (Skenario 8) mengalami kegagalan signifikan, menyoroti potensi ketidakstabilan atau sensitivitas model pembanding.

3. Skenario 7 (Crossentropy + RMSprop) memberikan OA tertinggi, tetapi bias terhadap kelas 'No Damage' dan lemah pada 'Minor Damage'.
4. Skenario 6 (Dual Focal + Adam) menunjukkan kinerja paling seimbang dan mIoU/mF1 tertinggi.

Berdasarkan pertimbangan keseimbangan kinerja antar kelas, yang lebih penting untuk evaluasi segmentasi, Skenario 6 (Dual Focal Loss + Adam) dipilih sebagai konfigurasi terbaik untuk model pembanding (U-Net ResNet50). Konfigurasi inilah yang akan digunakan sebagai *baseline* untuk dibandingkan secara langsung dengan model usulan terbaik (Skenario 4) pada Bab 6.

BAB VI

PEMBAHASAN

Bab sebelumnya telah menyajikan hasil pengujian dan analisis awal untuk model usulan (U-Net ResNet50 CBAM) dan model pembanding (U-Net ResNet50). Telah diidentifikasi bahwa Skenario 4 (Dual Focal Loss + RMSprop) merupakan konfigurasi terbaik untuk model usulan, sementara Skenario 6 (Dual Focal Loss + Adam) menunjukkan kinerja terbaik untuk model pembanding. Bab ini bertujuan untuk melakukan pembahasan mendalam dengan membandingkan secara langsung kinerja kedua model terbaik tersebut. Analisis ini akan mencakup perbandingan metrik evaluasi, perbandingan *confusion matrix* dan contoh visual, serta diskusi mengenai dampak mekanisme atensi CBAM.

6.1 Perbandingan Kinerja

Untuk melihat perbandingan langsung antara model usulan (Skenario 4) dan model pembanding (Skenario 6), metrik evaluasi utama disajikan dalam Tabel 6.1.

Tabel 6.1 Perbandingan Kinerja Model Usulan (S4) vs Model Pembanding (S6)

Model	Loss function	Optimizer	OA	mF1	mIoU
Model pembanding (skenario 6)	Dual Focal	Adam	0.5594	0.4921	0.3639
Proposed Model (skenario 4)	Dual Focal	RMSProp	0.6009	0.5145	0.3939
Peningkatan (%)			+7.42%	+4.55%	+8.24%

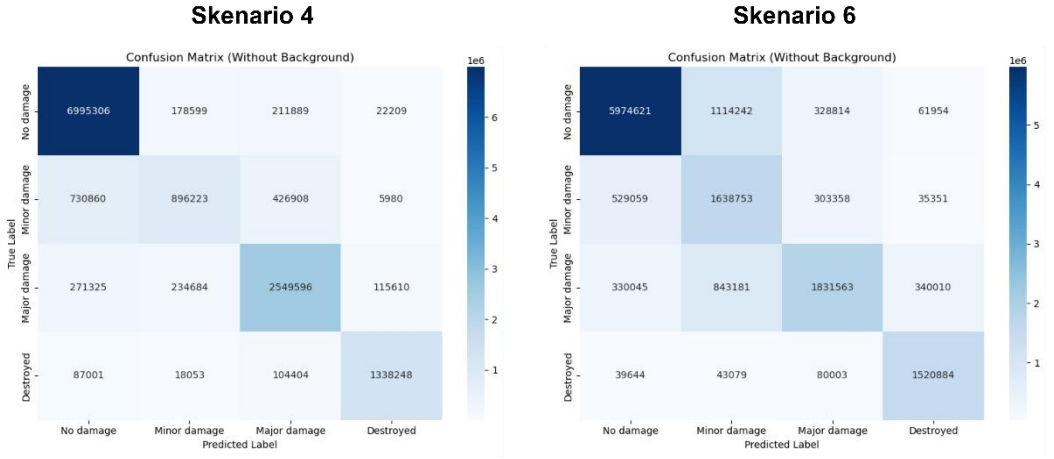
Dari Tabel 6.1, terlihat jelas bahwa model usulan (Skenario 4) secara konsisten mengungguli model pembanding (Skenario 6) di semua metrik evaluasi utama. Model usulan menunjukkan peningkatan *Overall Accuracy* (OA) sebesar 7.42%,

peningkatan *mean F1-Score* (mF1) sebesar 4.55%, dan yang paling signifikan, peningkatan *mean Intersection over Union* (mIoU) sebesar 8.24%.

Peningkatan mIoU, yang merupakan metrik standar utama untuk tugas segmentasi, menunjukkan bahwa penambahan mekanisme atensi CBAM pada model usulan berhasil meningkatkan kemampuan model untuk menghasilkan peta segmentasi yang lebih akurat secara spasial, dengan tumpang tindih yang lebih baik antara prediksi dan *ground truth* dibandingkan model pembanding. Peningkatan OA dan mF1 lebih lanjut memperkuat kesimpulan bahwa model usulan memiliki kinerja yang superior.

6.2 Perbandingan Kinerja Analisis Confusion Matrix

Untuk memahami *bagaimana* model usulan mencapai keunggulan ini, kita perlu membandingkan *confusion matrix* dari kedua model terbaik pada Gambar 6.1.



Gambar 6.1 Komparasi Confussion matrix skenario 4 (kiri) dan confussion matrix skenario 6 (kanan)

Analisis perbandingan *confusion matrix* menunjukkan beberapa poin penting:

1. Akurasi 'No Damage'

Model usulan (Skenario 4) memiliki *True Positive* (TP) 'No Damage' yang lebih tinggi (6.995.306) dibandingkan Skenario 6 (5.974.621). Selain itu, Skenario 4 menunjukkan *False Negative* (kesalahan prediksi *'No Damage'* menjadi kelas lain) yang lebih rendah, terutama ke kelas 'Major Damage' dan 'Destroyed'. Ini menunjukkan Skenario 4 lebih baik dalam mengenali area yang tidak rusak.

2. Akurasi 'Major Damage'

Model usulan (Skenario 4) menunjukkan keunggulan signifikan dalam mengidentifikasi 'Major Damage', dengan TP sebesar 2.549.596 dibandingkan Skenario 6 yang hanya 1.831.563. Skenario 6 cenderung lebih banyak salah mengklasifikasikan 'Major Damage' sebagai 'Minor Damage' (843.181) dibandingkan Skenario 4 (234.684). Ini adalah indikasi kuat bahwa CBAM membantu model mengenali fitur-fitur kerusakan yang lebih parah.

3. Akurasi 'Destroyed'

Meskipun Skenario 6 memiliki TP *'Destroyed'* yang sedikit lebih tinggi (1.520.884 vs 1.338.248), Skenario 4 menunjukkan *False Positive* yang jauh lebih rendah untuk kelas ini (lebih sedikit kelas lain yang salah diklasifikasikan sebagai *'Destroyed'*). Ini menunjukkan Skenario 4 mungkin lebih 'hati-hati' tetapi lebih presisi saat menyatakan sesuatu hancur.

4. Tantangan 'Minor Damage'

Kinerja pada kelas 'Minor Damage' menarik. Skenario 6 menunjukkan TP *'Minor Damage'* yang lebih tinggi (1.638.753) dibandingkan Skenario 4 (896.223). *Namun*, jika dilihat lebih dalam, Skenario 6 mencapai ini dengan tingkat kebingungan yang

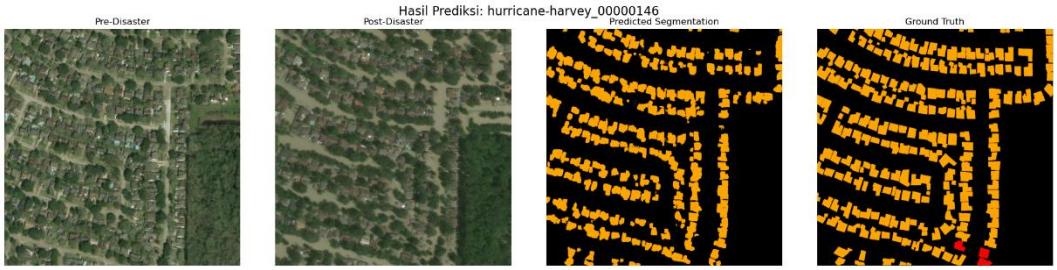
sangat tinggi, terutama dari *'Major Damage'* (843.181) dan 'No Damage' (1.114.242) yang salah diklasifikasikan sebagai *'Minor Damage'*. Sebaliknya, Skenario 4, meskipun TP *'Minor Damage'*-nya lebih rendah, *confusion matrix*-nya tampak lebih 'bersih', dengan lebih sedikit *false positive* yang masuk ke kelas *'Minor Damage'*. Ini mungkin berarti Skenario 4 lebih presisi (meskipun recall-nya lebih rendah) pada kelas ini, sementara Skenario 6 cenderung 'menebak' *'Minor Damage'* lebih sering. Ini menunjukkan bahwa meskipun CBAM tidak sepenuhnya menyelesaikan masalah *'Minor Damage'*, ia mungkin membantu mengurangi ambiguitas klasifikasi secara keseluruhan.

Secara keseluruhan, *confusion matrix* mendukung temuan bahwa model usulan (Skenario 4) lebih baik. Ia menunjukkan kebingungan antar kelas yang lebih rendah (elemen *off-diagonal* yang umumnya lebih kecil) dan akurasi yang lebih baik pada kelas 'No Damage' dan 'Major Damage', meskipun 'Minor Damage' tetap menjadi tantangan.

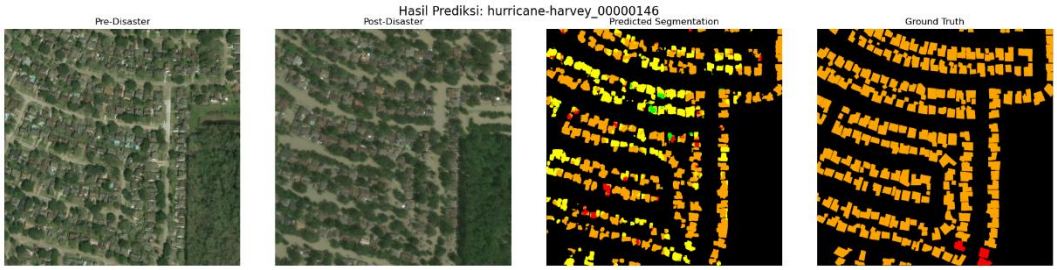
6.3 Analisis Dampak Mekanisme Atensi (CBAM)

Keunggulan model usulan (Sskenario 4) atas model pembanding (Skenario 6), sebagaimana ditunjukkan pada Tabel 6.1, secara signifikan dipengaruhi oleh integrasi modul *Convolutional Block Attention Module* (CBAM). CBAM, melalui mekanisme *channel* dan *spatial attention*, dirancang untuk meningkatkan representasi fitur yang relevan bagi tugas segmentasi. Analisis visual terhadap hasil prediksi semakin memperjelas kontribusi CBAM.

Gambar 6.2 dan Gambar 6.3 menyajikan perbandingan hasil prediksi untuk area yang sama, di mana Gambar 6.2 menunjukkan hasil Skenario 4 (model usulan dengan CBAM) dan Gambar 6.3 menunjukkan hasil Skenario 6 (model pembanding tanpa CBAM).



Gambar 6.2 Hasil segmentasi skenario 4



Gambar 6.3 Hasil Segmentasi skenario 6

Berdasarkan Gambar 6.2 dan Gambar 6.3, terlihat perbedaan signifikan antara kedua hasil segmentasi:

- Kejelasan Batas Segmentasi: Hasil segmentasi dari Skenario 4 (dengan CBAM) menunjukkan batas-batas bangunan yang jauh lebih jelas dan tegas dibandingkan dengan Skenario 6. Batas yang lebih kabur pada Skenario 6 mengindikasikan kesulitan model pembanding dalam membedakan secara presisi antara piksel bangunan dan latar belakang di tepian objek.
- Peningkatan nilai evaluasi secara keseluruhan dapat diamati pada Skenario 4 dibandingkan dengan Skenario 6, yang ditunjukkan oleh peningkatan nilai OA,

mF1, dan mIoU masing-masing sebesar +7.42%, +4.55%, dan +8.24%. Hal ini menunjukkan bahwa proposed model memiliki performa yang lebih baik secara umum dalam melakukan segmentasi kerusakan. Peningkatan ini dipengaruhi oleh penggunaan mekanisme attention CBAM, yang membantu model memfokuskan perhatian pada area-area relevan, sehingga menghasilkan prediksi yang lebih akurat, batas objek yang lebih jelas, dan segmentasi yang lebih konsisten terhadap ground truth. Dengan demikian, kemampuan CBAM untuk meningkatkan pemahaman spasial model berkontribusi tidak hanya pada peningkatan mIoU, tetapi juga pada peningkatan presisi dan akurasi secara keseluruhan.

- Kaitan segmentasi kelas, temuan ini dikonfirmasi secara kuat oleh *confusion matrix*. Mari kita lihat contoh spesifik 'rusak berat' (Major Damage) yang salah diprediksi sebagai 'rusak ringan' (Minor Damage):
    - Pada *confusion matrix* Skenario 6, jumlah piksel 'Major Damage' yang salah diprediksi sebagai 'Minor Damage' sangat tinggi, yaitu 843,181.
    - Sementara itu, pada *confusion matrix* Skenario 4, jumlah kesalahan yang sama jauh lebih rendah, yaitu 234,684.

Ini menunjukkan bahwa model Skenario 6 (tanpa CBAM) hampir 3.6 kali lebih sering melakukan kesalahan spesifik ini dibandingkan Skenario 4 (dengan CBAM). Kemampuan CBAM untuk membantu model mengenali fitur-fitur yang lebih diskriminatif antar kelas kerusakan secara jelas mengurangi tingkat kebingungan ini. Hal ini juga tercermin pada nilai mF1 Skenario 4 yang lebih tinggi

(0.3939 vs 0.3639), yang mengindikasikan keseimbangan presisi dan recall yang lebih baik di seluruh kelas.

Secara keseluruhan, analisis visual memperkuat dan memberikan konteks pada data. Hasil Skenario 4 yang lebih superior bukan hanya sekadar angka yang lebih tinggi, tetapi secara nyata terlihat dalam bentuk segmentasi yang lebih rapi dan klasifikasi yang lebih akurat pada level piksel. Perbedaan visual yang jelas antara Skenario 4 dan Skenario 6, terutama dalam hal kejelasan batas dan pengurangan kesalahan klasifikasi, menjadi bukti kuat bahwa mekanisme atensi CBAM secara efektif meningkatkan kinerja kemampuan model U-Net ResNet50 dalam tugas segmentasi kerusakan bangunan yang kompleks. CBAM membantu model untuk tidak hanya melihat bangunan, tetapi juga untuk memahami di mana batasnya dan membedakan tingkat kerusakannya dengan lebih baik.

Penelitian ini, yang berfokus pada aspek teknis pengembangan model *deep learning* untuk segmentasi kerusakan bangunan, pada hakikatnya tidak terlepas dari nilai-nilai dan pandangan hidup Islam. Sebagai agama yang *rahmatan lil 'alamin* (rahmat bagi seluruh alam), Islam mendorong umatnya untuk memanfaatkan ilmu pengetahuan dan teknologi demi kemaslahatan umat manusia dan alam semesta. Upaya untuk mengembangkan metode penilaian kerusakan pasca bencana yang cepat dan akurat sejalan dengan beberapa prinsip dasar ajaran Islam.

Dalam pandangan Islam, bencana alam dapat dipahami sebagai ujian atau cobaan dari Allah SWT untuk menguji kesabaran, keimanan, dan kepedulian sosial. Sebagaimana firman Allah SWT dalam QS. Al-Baqarah [2]: 155:

وَلَنَبْلُوَنَّكُمْ بِشَيْءٍ مِّنَ الْخَوْفِ وَالْجُوْعِ وَنَقْصٍ مِّنَ الْاَمْوَالِ وَالْاَنْفُسِ وَالثَّمَرٰتِۗ وَبَشِّرِ الصّٰبِرِيْنَ ١٥٥

Artinya :

"Dan Kami pasti akan menguji kamu dengan sedikit ketakutan, kelaparan, kekurangan harta, jiwa, dan buah-buahan. Dan sampaikanlah kabar gembira kepada orang-orang yang sabar."

Tafsir Kemenag RI mengenai ayat ini berbunyi “Kehidupan manusia memang penuh cobaan. Dan kami pasti akan menguji kamu untuk mengetahui kualitas keimanan seseorang dengan sedikit ketakutan, kelaparan, kekurangan harta, jiwa, dan buahbuahan. Bersabarlah dalam menghadapi semua itu. Dan sampaikanlah kabar gembira, wahai nabi Muhammad, kepada orang-orang yang sabar dan tangguh dalam menghadapi cobaan hidup, yakni orang-orang yang apabila ditimpa musibah, apa pun bentuknya, besar maupun kecil, mereka berkata, inna' lilla'hi wa inna' ilaihi ra'ji'un (sesungguhnya kami milik Allah dan kepada-Nyalah kami kembali). Mereka berkata demikian untuk menunjukkan kepasrahan total kepada Allah, bahwa apa saja yang ada di dunia ini adalah milik Allah; pun menunjukkan keimanan mereka akan adanya hari akhir. Mereka itulah yang memperoleh ampunan dan rahmat dari tuhannya, dan mereka itulah orang-orang yang mendapat petunjuk sehingga mengetahui kebenaran” (Kemenag RI, 2016).

Menurut Tafsir Kementerian Agama RI, ayat ini menegaskan bahwa ujian adalah keniscayaan hidup yang bertujuan menguji kualitas iman dan kesabaran. Kesabaran (*sabar*) di sini tidak hanya berarti kepasrahan pasif, melainkan keteguhan hati yang disertai *ikhtiar* (usaha) mencari solusi. Pemahaman ini

memberikan landasan spiritual bahwa dalam menghadapi bencana, kita termotivasi untuk melakukan respons aktif. Pengembangan model segmentasi kerusakan menggunakan AI dalam penelitian ini adalah bentuk *ikhtiar* modern untuk memahami skala bencana, yang menjadi dasar tindakan pertolongan dan rehabilitasi yang efektif, sebagai wujud nyata kepedulian dan kesabaran aktif.

Selain keharusan untuk merespons bencana dengan ikhtiar, Islam juga sangat menekankan pentingnya mencari dan mengembangkan ilmu pengetahuan. Allah SWT berfirman dalam QS. Al-Mujadila [58]: 11:

يٰٓاَيُّهَا الَّذِيْنَ اٰمَنُوْٓا اِذَا قِيْلَ لَكُمْ تَفَسَّحُوْا فِى الْمَجٰلِسِ فَافْسَحُوْا يَفْسَحِ اللّٰهُ لَكُمْۚ وَاِذَا قِيْلَ انْشُزُوْا فَانْشُزُوْا يَرْفَعِ اللّٰهُ الَّذِيْنَ اٰمَنُوْا

مِنْكُمْۙ وَالَّذِيْنَ اُوْتُوا الْعِلْمَ دَرَجٰتٍۗ وَاللّٰهُ بِمَا تَعْمَلُوْنَ خَبِيْرٌ ۝١١

Artinya :

"Wahai orang-orang yang beriman, apabila dikatakan kepadamu "Berilah kelapangan di dalam majelis-majelis," lapangkanlah, niscaya Allah akan memberi kelapangan untukmu. Apabila dikatakan, "Berdirilah," (kamu) berdirilah. Allah niscaya akan mengangkat orang-orang yang beriman di antaramu dan orang-orang yang diberi ilmu beberapa derajat. Allah Mahateliti terhadap apa yang kamu kerjakan."

Tafsir Al – Mishbah mengenai ayat ini berbunyi "Menurut M. Quraish Shihab dalam Tafsir Al-Misbah, ayat ini menunjukkan betapa tinggi penghargaan Islam terhadap ilmu pengetahuan. Peningkatan derajat dijanjikan bagi mereka yang memiliki ilmu dan menggunakannya dengan landasan iman. Ilmu adalah alat untuk

memahami kebesaran Allah melalui ciptaan-Nya dan untuk mengelola kehidupan di bumi dengan lebih baik" (M. Quraish Shihab, *Tafsir Al-Misbah*)

M. Quraish Shihab dalam Tafsir Al-Misbah menjelaskan bahwa ayat ini menegaskan tingginya kedudukan ilmu dalam Islam sebagai alat untuk memahami alam dan menyelesaikan permasalahan. Amanah keilmuan inilah yang coba diwujudkan dalam penelitian ini. Pemanfaatan metode *deep learning* yang canggih seperti U-Net, *pretrained backbone* ResNet50, dan *attention mechanism* CBAM, merupakan implementasi dari pemanfaatan ilmu pengetahuan dan teknologi—sebagai karunia Allah—untuk masalah kemanusiaan. Upaya menghasilkan model yang akurat (mencapai mIoU 0.3939) juga sejalan dengan prinsip *itqan* (profesionalisme) dalam bekerja, karena penilaian yang tepat adalah langkah krusial yang harus dilakukan sebaik mungkin.

Pemanfaatan ilmu dan teknologi ini pada akhirnya diarahkan untuk memenuhi salah satu pilar utama ajaran Islam, yaitu kewajiban tolong-menolong (*ta'awun*), sebagaimana ditegaskan dalam QS. Al-Ma'idah [5]: 2:

يٰٓاَيُّهَا الَّذِيْنَ اٰمَنُوْا لَا تُحِلُّوْا شَعَاۤىِٕرَ اللّٰهِ وَلَا الشَّهْرَ الْحَرَامَ وَلَا الْهَدْيَ وَلَا الْقَلَاۤىِٕدَ وَلَآ اٰۤمِّيْنَ الْبَيْتَ الْحَرَامَ يَبْتَغُوْنَ فَضْلًا مِّنْ رَّبِّهِمْ وَرِضْوَانًاۗ وَاِذَا حَلَلْتُمْ فَاصْطَادُوْاۗ وَلَا يَجْرِمَنَّكُمْ شَنَاٰنُ قَوْمٍ اَنْ صَدُّوْكُمْ عَنِ الْمَسْجِدِ الْحَرَامِ اَنْ تَعْتَدُوْاۘ وَتَعَاوَنُوْا عَلَى الْبِرِّ وَالتَّقْوٰىۖ وَلَا تَعَاوَنُوْا عَلَى الْاِثْمِ وَالْعُدْوَانِۖ وَاتَّقُوا اللّٰهَۗ اِنَّ اللّٰهَ شَدِيْدُ الْعِقَابِ ۝٢

Artinya :

"Wahai orang-orang yang beriman, janganlah kamu melanggar syiar-syiar (kesucian) Allah, jangan (melanggar kehormatan) bulan-bulan haram, jangan

(mengganggu) hadyu (hewan-hewan kurban) dan qalā'id (hewan-hewan kurban yang diberi tanda), dan jangan (pula mengganggu) para pengunjung Baitulharam sedangkan mereka mencari karunia dan rida Tuhannya! Apabila kamu telah bertahalul (menyelesaikan ihram), berburulah (jika mau). Janganlah sekali-kali kebencian(-mu) kepada suatu kaum, karena mereka menghalang-halangimu dari Masjidilharam, mendorongmu berbuat melampaui batas (kepada mereka). Tolong-menolonglah kamu dalam (mengerjakan) kebajikan dan takwa, dan jangan tolong-menolong dalam berbuat dosa dan permusuhan. Bertakwalah kepada Allah, sesungguhnya Allah sangat berat siksaan-Nya."

Tafsir dari Kemenag RI mengenai ayat ini yang berbunyi "Ayat berikut berisi hukum-hukum Allah yang berkaitan dengan tata cara pelaksanaan ibadah haji. Wahai orang-orang yang beriman! janganlah kamu melanggar syiar-syiar kesucian Allah, yakni segala amalan yang dilakukan dalam melaksanakan ibadah haji seperti tata cara melakukan tawaf dan sa'i, serta tempat-tempat mengerjakannya, seperti kakbah, safa, dan marwah, jangan engkau melanggarnya dengan berburu ketika dalam keadaan ihram dan jangan pula melanggar kehormatan bulanbulan haram, yaitu bulan zulkaidah, zulhijah, muharram, dan rajab, janganlah pula engkau melanggar kehormatannya dengan berperang pada bulan itu kecuali untuk membela diri ketika diserang. Jangan pula mengganggu hadyu, yaitu hewan-hewan kurban yang dihadiahkan kepada kakbah untuk mendekatkan diri kepada Allah, hewan-hewan itu disembelih di tanah haram dan dihadiahkan dagingnya kepada fakir miskin, dan qala'id, hewan-hewan kurban yang diberi tanda, dikalungi dengan tali sebagai tanda yang menunjukkan bahwa hewan

itu telah dipersiapkan untuk dikurbankan dan dihadiahkan, dan jangan pula mengganggu orang-orang yang mengunjungi baitulharam, untuk melaksanakan ibadah haji atau umrah, mereka mencari karunia berupa keuntungan duniawi, dan keridaan yang berupa ganjaran dari tuhannya. Akan tetapi, apabila kamu telah menyelesaikan ihram, maka bolehlah kamu berburu apabila kamu mau. Jangan sampai kebencian sebagian kamu kepada suatu kaum karena mereka menghalang-halangimu dari mengunjungi masjidilharam, mendorongmu berbuat melampaui batas kepada mereka dengan cara membunuh mereka atau melakukan kejahatan kepada mereka. Dan tolong-menolonglah kamu dalam mengerjakan kebajikan, melakukan yang diperintahkan Allah, dan takwa, takut kepada larangannya, dan jangan tolong-menolong dalam berbuat dosa, melakukan maksiat dan permusuhan, sebab yang demikian itu melanggar hukum-hukum Allah. Bertakwalah kepada Allah, takut kepada Allah dengan melakukan perintah-Nya dan meninggalkan larangan-Nya, karena sungguh Allah sangat berat siksaan-Nya kepada orang-orang yang tidak taat kepada-Nyapada ayat yang lalu telah dijelaskan beberapa perbuatan yang diharamkan. Ayat ini menguraikan lebih terperinci makanan-makanan yang diharamkan. Ada sepuluh jenis makanan yang diharamkan, semuanya berasal dari hewan. Diharamkan bagimu memakan bangkai, darah yang keluar dari tubuh, sebagaimana tersebut dalam surah alan'a'm 6: 145, daging babi, dan daging hewan yang disembelih bukan atas nama Allah, demikian pula diharamkan daging hewan yang tercekik, yang dipukul, yang jatuh, yang ditanduk, dan yang diterkam binatang buas, kecuali yang sempat kamu sembelih. Hewan yang tercekik, yang dipukul, yang jatuh, yang ditanduk, dan yang diterkam binatang buas adalah halal hukumnya

kalau sempat disembelih sebelum mati. Dan diharamkan pula hewan yang disembelih untuk berhala. Dan diharamkan pula mengundi nasib dengan anak panah. Orang arab jahiliah menggunakan anak panah untuk menentukan apakah mereka akan melakukan suatu perbuatan atau tidak. Mereka mengambil tiga buah anak panah yang belum memakai bulu, masing-masing anak panah itu ditulis dengan kata-kata lakukan, jangan lakukan, dan anak panah yang ketiga tidak ditulis apa-apa. Semua anak panah itu diletakkan dalam sebuah tempat dan disimpan dalam kakbah. Bila mereka hendak melakukan suatu perbuatan, maka mereka meminta agar juru kunci kakbah mengambil salah satu dari tiga anak panah itu. Mereka melakukan perbuatan atau tidak melakukan perbuatan sesuai dengan bunyi kalimat yang tertulis dalam anak panah yang diambilnya. Kalau yang terambil anak panah yang tidak ada tulisannya, maka undian diulangi sekali lagi. Janganlah melakukan yang demikian itu karena itu suatu perbuatan fasik. Pada hari ini, yaitu pada waktu haji wada', haji terakhir yang dilakukan oleh nabi Muhammad, orang-orang kafir telah putus asa untuk mengalahkan agamamu, sebab itu janganlah kamu takut kepada mereka, tetapi takutlah kepada-ku. Pada hari ini telah aku sempurnakan agamamu untukmu, dan telah aku cukupkan nikmat-ku bagimu, dan telah aku ridai islam sebagai agamamu. Tetapi barang siapa dalam keadaan terpaksa, dibolehkan memakan makanan yang diharamkan oleh ayat ini karena lapar, bukan karena ingin berbuat dosa, maka sungguh, Allah maha pengampun, maha penyayang." (Kemenag RI, 2016).

Tafsir Kemenag RI menguraikan ayat ini sebagai perintah untuk saling mendukung dalam segala bentuk kebaikan (*al-birr*), termasuk membantu korban

bencana. Model segmentasi yang diajukan dalam tesis ini secara langsung berfungsi sebagai alat pendukung *ta'awun*. Dengan menyediakan data segmentasi kerusakan yang cepat dan akurat, model ini membantu memastikan bahwa pertolongan (*al-birr*) disalurkan secara tepat sasaran, adil, dan efisien, sehingga semangat tolong-menolong dapat diwujudkan secara lebih efektif.

Lebih jauh lagi, upaya ini tidak berhenti pada respons darurat, tetapi juga menyentuh aspek pembangunan kembali, yang merupakan bagian dari tugas manusia untuk memakmurkan bumi (*i'mar al-ardh*). Allah berfirman dalam QS. Hud [11]: 61:

وَاِلٰى ثَمُوْدَ اَخَاهُمْ صٰلِحًاۗ قَالَ يٰقَوْمِ اعْبُدُوا اللّٰهَ مَا لَكُمْ مِّنْ اِلٰهٍ غَيْرُهٗۗ هُوَ اَنْشَاَكُمْ مِّنَ الْاَرْضِ وَاسْتَعْمَرَكُمْ فِيْهَا فَاسْتَغْفِرُوْهُ ثُمَّ تُوْبُوْٓا اِلَيْهِۗ اِنَّ رَبِّيْ قَرِيْبٌ مُّجِيْبٌ ﴿٦١﴾

Artinya :

" Kepada (kaum) Samud (Kami utus) saudara mereka, Saleh. Dia berkata, "Wahai kaumku, sembahlah Allah! Sekali-kali tidak ada tuhan bagimu selain Dia. Dia telah menciptakanmu dari bumi (tanah) dan menjadikanmu pemakmurnya. Oleh karena itu, mohonlah ampunan kepada-Nya, kemudian bertobatlah kepada-Nya. Sesungguhnya Tuhanku sangat dekat lagi Maha Memperkenankan (doa hamba-Nya)."

Tafsir al – mishbah berbunyi tentang ayat ini "kata *'ista'marakum'* (menjadikanmu pemakmurnya) menyiratkan sebuah tugas dan amanah dari Allah kepada manusia untuk mengelola, membangun, mengembangkan, dan menjaga

kelestarian bumi. Ini mencakup pembangunan fisik (infrastruktur, bangunan) dan pembangunan sosial-ekonomi" (Dr. M. Quraish Shihab, 2001).

Menurut Tafsir Al-Misbah, ayat ini menyiratkan tugas dan amanah manusia untuk mengelola, membangun, dan mengembangkan bumi. Meskipun penelitian ini berfokus pada *penilaian* kerusakan, ia menjadi langkah fundamental pertama dalam proses *i'mar al-ardh* pasca bencana, yaitu rekonstruksi. Dengan menyediakan data awal yang akurat, model segmentasi ini secara tidak langsung berkontribusi pada upaya membangun kembali kehidupan sebagai bagian dari tanggung jawab manusia sebagai *khalifah*.

Dengan demikian, dapat disimpulkan bahwa penelitian teknis ini tidak hanya memiliki signifikansi ilmiah, tetapi juga berakar kuat pada nilai-nilai keislaman. Ia menunjukkan bagaimana teknologi *deep learning* dapat menjadi sarana bagi umat Islam untuk mewujudkan nilai-nilai agamanya: ber-*ikhtiar* dalam menghadapi cobaan, mengamalkan ilmu pengetahuan, mengoptimalkan *ta'awun*, dan menjalankan amanah sebagai *khalifah* di muka bumi.

BAB VII

PENUTUP

7.1 Kesimpulan

Berdasarkan evaluasi yang telah dilakukan, penelitian ini membuktikan bahwa integrasi *attention mechanism* (CBAM) secara signifikan meningkatkan kinerja model U-Net yang memanfaatkan *pretrained backbone* (ResNet50) dan pendekatan *dual encoder with dual input*. Peningkatan kinerja ini divalidasi melalui perbandingan antara model usulan (yang mengintegrasikan ketiga elemen) dengan model pembanding (tanpa CBAM) menggunakan serangkaian skenario pengujian pada dataset Xview2.

Berdasarkan hasil pengujian dan analisis yang telah dipaparkan, kesimpulan utama yang dapat ditarik untuk menjawab pernyataan masalah adalah sebagai berikut:

1. Pemanfaatan *Pretrained Backbone* dan *Dual Encoder* sebagai Dasar:

Kinerja model U-Net yang *hanya* memanfaatkan *pretrained backbone* (ResNet50) dan *dual encoder* (model pembanding Skenario 6) menunjukkan kinerja dasar yang cukup baik, dengan mencapai mIoU sebesar 0.3639 dan mF1 sebesar 0.4921. Hal ini membuktikan bahwa penggunaan *pretrained backbone* efektif dalam mengekstraksi fitur kompleks dari citra satelit, dan pendekatan *dual input* relevan untuk tugas perbandingan citra pra dan pasca bencana. Namun, kinerja ini masih menunjukkan keterbatasan, terutama dalam hal keseimbangan akurasi antar kelas dan tingkat kebingungan yang relatif tinggi.

2. Peningkatan Kinerja Signifikan melalui *Attention Mechanism* (CBAM)

Kinerja model U-Net mengalami peningkatan yang signifikan ketika *attention mechanism* (CBAM) diintegrasikan ke dalam arsitektur (model usulan Skenario 4). Penambahan CBAM terbukti mampu meningkatkan mIoU menjadi 0.3939 (naik 8.24%), mF1 menjadi 0.5145 (naik 4.55%), dan OA menjadi 0.6009 (naik 7.42%) dibandingkan model pembanding. Peningkatan ini menunjukkan bahwa CBAM secara efektif membantu model untuk:

- Meningkatkan fokus pada fitur spasial dan kanal yang paling informatif.
- Memperbaiki lokalisasi batas-batas bangunan.
- Mengurangi misklasifikasi antar kelas kerusakan, terutama untuk kelas *'Major Damage'* dan *'No Damage'*.

3. Kinerja Kombinasi Optimal

Sebagai jawaban langsung atas pernyataan masalah, kinerja model U-Net yang memanfaatkan *ketiga* elemen (pretrained backbone, dual encoder, dan attention mechanism CBAM) adalah yang paling superior dalam penelitian ini. Kombinasi ini (model usulan Skenario 4) menghasilkan model segmentasi kerusakan bangunan yang lebih akurat, seimbang, dan andal dibandingkan model yang tidak menggunakan *attention mechanism*. Kinerja ini (mIoU 0.3939) menunjukkan bahwa pendekatan yang diusulkan adalah strategi yang valid dan efektif untuk tugas segmentasi kerusakan bangunan.

Meskipun demikian, perlu dicatat bahwa kinerja terbaik sekalipun masih menunjukkan tantangan, terutama dalam mengidentifikasi kelas *'Minor Damage'* secara akurat. Hal ini mengindikasikan bahwa meskipun pendekatan yang diusulkan berhasil meningkatkan kinerja secara keseluruhan, masih terdapat ruang untuk penyempurnaan lebih lanjut.

Secara keseluruhan, penelitian ini berhasil menunjukkan dampak positif dan signifikan dari penambahan *attention mechanism* (CBAM) di atas fondasi U-Net dengan *pretrained backbone* dan *dual encoder*, sehingga memberikan jawaban yang jelas mengenai peningkatan kinerja model dalam memanfaatkan ketiga teknologi tersebut untuk segmentasi kerusakan bangunan.

7.2 Saran

Berdasarkan kesimpulan dan keterbatasan yang teridentifikasi selama penelitian, berikut adalah beberapa saran yang dapat menjadi acuan untuk penelitian selanjutnya di bidang ini:

1. Peningkatan Penanganan Kelas 'Minor Damage'

Fokus pada pengembangan metode untuk meningkatkan deteksi kerusakan ringan, misalnya dengan:

- Menggunakan *loss function* yang dirancang khusus untuk menangani kelas minoritas atau kasus sulit.
- Menerapkan teknik *data augmentation* yang lebih canggih untuk memperbanyak sampel 'Minor Damage'.

- Menjelajahi arsitektur atau modul tambahan yang sensitif terhadap fitur-fitur subtil.
- Menggunakan data citra dengan resolusi spasial yang lebih tinggi.

2. Eksplorasi Arsitektur dan Backbone:
   - Mencoba *backbone* yang lebih modern dan efisien (misalnya, EfficientNet, ViT - Vision Transformers).
   - Menjelajahi mekanisme atensi lainnya (misalnya, SE-Net, Triplet Attention) atau kombinasi beberapa mekanisme atensi.
   - Memodifikasi arsitektur U-Net (misalnya, U-Net++, Attention U-Net) untuk melihat potensi peningkatannya.
3. Pengembangan Dataset dan Data:
   - Menggabungkan dataset Xview2 dengan dataset lain untuk meningkatkan variasi dan generalisasi.
   - Menjelajahi penggunaan data multi-modal, seperti data SAR (Synthetic Aperture Radar) yang tidak terpengaruh awan, atau data LiDAR/DEM (Digital Elevation Model) untuk informasi ketinggian, yang dapat membantu membedakan kerusakan.
4. Optimasi Hyperparameter dan Pelatihan:
   - Melakukan pencarian *hyperparameter* yang lebih ekstensif (misalnya, menggunakan *Grid Search* atau *Bayesian Optimization*) untuk menemukan

kombinasi *learning rate*, *batch size*, dan parameter *optimizer* yang lebih optimal.

- Menerapkan teknik pelatihan yang lebih canggih, seperti *learning rate scheduler* yang berbeda atau pelatihan dengan *ensemble models*.

5. Validasi di Area Bencana Berbeda:

Melakukan pengujian dan validasi model pada data dari lokasi bencana yang berbeda (jenis bencana, tipe bangunan, kondisi geografis) untuk menguji ketahanan dan kemampuan generalisasi model.

Dengan menindaklanjuti saran-saran ini, diharapkan penelitian di masa depan dapat menghasilkan model segmentasi kerusakan bangunan yang lebih akurat, andal, dan dapat diimplementasikan secara efektif dalam mendukung upaya tanggap darurat dan mitigasi bencana.

DAFTAR PUSTAKA

A. Naji, S., & Majeed Saleh, N. (2024). Digital Image Forgery Detection And Localization Using The Innovated U-Net. *Iraqi Journal for Computers and Informatics*, *50*(1), 195–207. https://doi.org/10.25195/ijci.v50i1.484

Bala, B., & Behal, S. (2024). A Brief Survey of Data Preprocessing in Machine Learning and Deep Learning Techniques. *2024 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, 1755–1762. https://doi.org/10.1109/I-SMAC61858.2024.10714767

Cai, J., Shi, J., Leau, Y.-B., Meng, S., Zheng, X., & Zhou, J. (2024). Res50-SimAM-ASPP-Unet: A Semantic Segmentation Model for High-resolution Remote Sensing Images. *IEEE Access*, 1–1. https://doi.org/10.1109/ACCESS.2024.3519260

Chen, Z., Zhu, Y., Zhao, C., Hu, G., Zeng, W., Wang, J., & Tang, M. (2021). *DPT: Deformable Patch-based Transformer for Visual Recognition*. https://doi.org/10.1145/3474085.3475467

Csurka, G., Volpi, R., & Chidlovskii, B. (2023). *Semantic Image Segmentation: Two Decades of Research*.

Dr. M. Quraish Shihab. (2001). *Tafsir Al-Mishbah Jilid 06*.

Feng, G., & Tang, C. (2024). Portrait Semantic Segmentation Method Based on Dual Modal Information Complementarity. *Applied Sciences*, *14*(4), 1439. https://doi.org/10.3390/app14041439

Goodfellow, I., Bengio, Y., & Courville, A. (2015). *Deep Learning*.

Hatic, D., Rauhut, M., Polushko, V., Codastefano, M., Rosa, S., McKay, P., Stompanato, F., Beleza, A., & Hagen, H. (2023). Disaster impact analysis in Africa: EDDA - UAV building damage dataset of rural and urban Mozambique. *Remote Sensing*. https://api.semanticscholar.org/CorpusID:264357977

Karimi, A., Faez, K., & Nazari, S. (2023). DEU-Net: Dual-Encoder U-Net for Automated Skin Lesion Segmentation. *IEEE Access*, *11*, 134804–134821. https://doi.org/10.1109/ACCESS.2023.3337528

Kemenag RI. (2016). *Kemenag - Tafsir Ringkas Jilid 1*.

Kemenag RI. (2016). *Tafsir Wajiz I* (1st ed., Vol. 1). Lajnah Pentashihan Mushaf Al-Qur'an.

Khan, M. Z., Gajendran, M. K., Lee, Y., & Khan, M. A. (2021). Deep Neural Architectures for Medical Image Semantic Segmentation: Review. *IEEE Access*, *9*, 83002–83024. https://doi.org/10.1109/ACCESS.2021.3086530

Khelifi, L., & Mignotte, M. (2020). Deep Learning for Change Detection in Remote Sensing Images: Comprehensive Review and Meta-Analysis. *IEEE Access*, *8*, 126385–126400. https://doi.org/10.1109/ACCESS.2020.3008036

Li, W., Wu, J., Chen, H., Wang, Y., Jia, Y., & Gui, G. (2022). UNet Combined With Attention Mechanism Method for Extracting Flood Submerged Range. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 6588–6597. https://doi.org/10.1109/JSTARS.2022.3194375

Li, Y., Zhu, Z., Li, Y., Zhang, J., Li, X., Shang, S., & Zhu, D. (2023). CTMU-Net: An Improved U-Net for Semantic Segmentation of Remote-Sensing Images Based on the Combined Attention Mechanism. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *16*, 10148–10161. https://doi.org/10.1109/JSTARS.2023.3326960

Liu, C., Ge, L., & Sepasgozar, S. M. E. (2021). Post-Disaster Classification of Building Damage Using Transfer Learning. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 2194–2197. https://doi.org/10.1109/IGARSS47720.2021.9554795

Liu, Z., Chen, B., & Zhang, A. (2020). Building segmentation from satellite imagery using U-Net with ResNet encoder. *2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, 1967–1971. https://doi.org/10.1109/ICMCCE51767.2020.00431

Ojha, V. K., Dutta, P., Saha, H., & Ghosh, S. (2012). Detection of Proportion of Different Gas Components Present in Manhole Gas Mixture Using Backpropagation Neural Network. *Technology*, *37*(Icint), 11–15.

Pahlavani, P., Samadzadegan, F., & Delavar, M. R. (2006). *A GIS-Based Approach for Urban Multi-criteria Quasi Optimized Route Guidance by Considering Unspecified Site Satisfaction* (pp. 287–303). https://doi.org/10.1007/11863939_19

Pramono, A., Mahendra, I. W. S., Wijaya, I. B. A., Agustina, I. A., & Herman, R. T. (2022). Diagnosis and Repair of the Cracking House Next to the River. *IOP Conference Series: Earth and Environmental Science*, *998*(1), 012002. https://doi.org/10.1088/1755-1315/998/1/012002

Ramya, S., Kumar, S. P., R, M. M. S., Ram, G. D., Lingaraja, D., & Aravind, T. (2023). Left Ventricle Segmentation in Cardiac MR Images Using Deep Neural Network Models. *2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN)*, 218–223. https://doi.org/10.1109/ICPCSN58827.2023.00041

Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation* (pp. 234–241). https://doi.org/10.1007/978-3-319-24574-4_28

Sharma, N., Gupta, S., Elkamchouchi, D. H., & Bharany, S. (2025). Encoder–Decoder Variant Analysis for Semantic Segmentation of Gastrointestinal Tract Using UW-Madison Dataset. *Bioengineering*, *12*(3), 309. https://doi.org/10.3390/bioengineering12030309

Strohmann, T., Bugelnig, K., Breitbarth, E., Wilde, F., Steffens, T., Germann, H., & Requena, G. (2019). Semantic segmentation of synchrotron tomography of multiphase Al-Si alloys using a convolutional neural network with a pixel-wise weighted loss function. *Scientific Reports*, *9*(1), 1–10. https://doi.org/10.1038/s41598-019-56008-7

Sun, L., Zou, H., Wei, J., Cao, X., He, S., Li, M., & Liu, S. (2023). Semantic Segmentation of High-Resolution Remote Sensing Images Based on Sparse Self-Attention and Feature Alignment. *Remote Sensing*, *15*(6), 1598. https://doi.org/10.3390/rs15061598

Wang, B., Zhao, C., Li, J., Sheng, Q., & Ling, X. (2025). U2DDS-Net: A New Architecture Based on U2Net With Disaster Type for Building Damage Assessment Under Natural Disasters. *The Photogrammetric Record*. https://api.semanticscholar.org/CorpusID:275737080

Weber, E., & Kan, H. (2020). *BUILDING DISASTER DAMAGE ASSESSMENT IN SATELLITE IMAGERY WITH MULTI-TEMPORAL FU-SION*. https://github.com/ethanweber/xview2

Woo, S., Park, J., Lee, J.-Y., & Kweon, I. S. (2018). *CBAM: Convolutional Block Attention Module* (pp. 3–19). https://doi.org/10.1007/978-3-030-01234-2_1

Wu, C., Zhang, F., Xia, J., Xu, Y., Li, G., Xie, J., Du, Z., & Liu, R. (2021). Building Damage Detection Using U-Net with Attention Mechanism from Pre- and Post-Disaster Remote Sensing Datasets. *Remote Sensing*, *13*(5), 905. https://doi.org/10.3390/rs13050905

Xu, Y., Hou, S., Wang, X., Li, D., & Lu, L. (2023). A Medical Image Segmentation Method Based on Improved UNet 3+ Network. *Diagnostics*, *13*(3), 576. https://doi.org/10.3390/diagnostics13030576

Yan, Y., Wang, H., Zhang, J., Song, X., Lou, Y., Dai, T., Wang, L., & Lv, M. (2023). Text Classification Algorithm for Medical Adverse Events Based on Deep Learning. *2023 IEEE International Conference on Medical Artificial Intelligence (MedAI)*, 235–244. https://doi.org/10.1109/MedAI59581.2023.00037

Zaryabi, E. H., Kalantar, B., Moradi, L., Halin, A. A., & Ueda, N. (2022a). MSBDA-Net: Multi-scale Siamese Building Damage Assessment Network. *2022 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, 1–6. https://doi.org/10.1109/CSDE56538.2022.10089353

Zaryabi, E. H., Kalantar, B., Moradi, L., Halin, A. A., & Ueda, N. (2022b). MSBDA-Net: Multi-scale Siamese Building Damage Assessment Network. *2022 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, 1–6. https://doi.org/10.1109/CSDE56538.2022.10089353