

# Identifikasi Kejadian Bencana Berdasarkan Gambar Unggahan di Media Sosial dalam Upaya Mitigasi Bencana Menggunakan Inception ResNet V2

WAWU TRI AMBODO<sup>1</sup>, ELLY MATUL IMAH<sup>2</sup>, HARMON PRAYOGI<sup>2,\*</sup>

<sup>1</sup>*Program Studi S1 Sains Data, Universitas Negeri Surabaya, Indonesia*

<sup>2</sup>*Program Studi S1 Kecerdasan Artifisial, Universitas Negeri Surabaya, Indonesia*

## Abstrak

Penanganan bencana memerlukan pemahaman yang tepat terhadap kondisi di tempat kejadian melalui informasi visual yang merepresentasikan dampak dan karakteristik bencana. Data citra bencana memiliki kompleksitas visual yang tinggi dan kemiripan pola antar jenis bencana yang seringkali menyulitkan proses identifikasi secara otomatis. Penelitian ini mengembangkan sistem klasifikasi citra berbasis deep learning menggunakan arsitektur Inception ResNet V2 dengan pendekatan transfer learning dan fine tuning untuk meningkatkan kemampuan ekstraksi fitur visual. Sistem dirancang untuk mengklasifikasikan enam kelas bencana, yaitu Damaged Infrastructure, Fire Disaster, Human Damage, Land Disaster, Non Damage, dan Water Disaster. Model CNN digunakan sebagai model pembanding, sedangkan Inception ResNet V2 digunakan sebagai model utama. Pengujian dilakukan dalam dua skenario, yaitu pelatihan tanpa fine tuning dan pelatihan dengan fine tuning. Hasil evaluasi menunjukkan bahwa model CNN memperoleh akurasi sebesar 88,40%, sedangkan Inception ResNet V2 tanpa fine tuning mencapai akurasi 94,89%. Performa terbaik diperoleh oleh Inception ResNet V2 dengan fine tuning yang mencapai accuracy 95,23%, precision 89,54%, recall 88,88%, dan f1 score 89,17%. Model terbaik yang sudah dibangun digunakan untuk klasifikasi data citra hasil scraping. Hasil menunjukkan bahwa model Inception ResNet V2 mampu mempertahankan performa dengan accuracy 90,26% precision 90,52%, recall 88,64%, dan f1-score 88,91%. Sistem yang dikembangkan memiliki potensi untuk digunakan sebagai pendukung visual kejadian bencana berbasis citra.

**Kata Kunci** *Bencana; CNN; Deep Learning; Inception ResNet V2; Media Sosial*

## Abstract

Disaster management requires an accurate understanding of on site conditions through visual information that represents the impact and characteristics of disasters. Disaster image data exhibit high visual complexity and similarities among different disaster types, which often hinder automated identification. This study develops a deep learning based image classification system using the Inception-ResNet V2 architecture with transfer learning and fine-tuning approaches to enhance visual feature extraction capability. The system is designed to classify six disaster classes, namely Damaged Infrastructure, Fire Disaster, Human Damage, Land Disaster, Non-Damage, and Water Disaster. A CNN model is used as a baseline, while Inception-ResNet V2 serves as the main model. Model evaluation is conducted under two scenarios: training without fine tuning and training with fine tuning. The results show that the CNN model achieves an accuracy of

---

\*Corresponding author. Email: [ellymatul@unesa.ac.id](mailto:ellymatul@unesa.ac.id) or [harmonprayogi@unesa.ac.id](mailto:harmonprayogi@unesa.ac.id).

88.40%, while Inception-ResNet V2 without fine tuning attains an accuracy of 94.89%. The best performance is achieved by the fine-tuned Inception ResNet V2 model, with an accuracy of 95.23%, precision of 89.54%, recall of 88.88%, and an F1-score of 89.17%. The best-performing model is further applied to classify scraped image data. The results indicate that the Inception-ResNet V2 model maintains performance with an accuracy of 90.26%, precision of 90.52%, recall of 88.64%, and an F1-score of 88.91%. The developed system has the potential to be used as a visual support tool for disaster event identification based on image data.

**Keywords** *CNN; Deep Learning; Disaster; Inception ResNet V2; Social Media*

## 1 Pendahuluan

Penggunaan media sosial sebagai sumber informasi bencana mengalami peningkatan signifikan dalam beberapa tahun terakhir karena sifatnya yang *real-time* dan mampu menyediakan informasi langsung dari masyarakat terdampak (Ullah dkk., 2020). Informasi yang dibagikan dapat membantu proses pemantauan, pemetaan, dan penanggulangan bencana secara lebih cepat dibandingkan metode konvensional. Namun, tingginya volume data yang tersebar menyebabkan proses identifikasi informasi yang relevan dan kredibel menjadi sulit dilakukan secara manual. Oleh karena itu, diperlukan metode otomatis yang mampu melakukan analisis dan klasifikasi data media sosial secara efektif untuk mendukung penanganan bencana (Maqbool dkk., 2020).

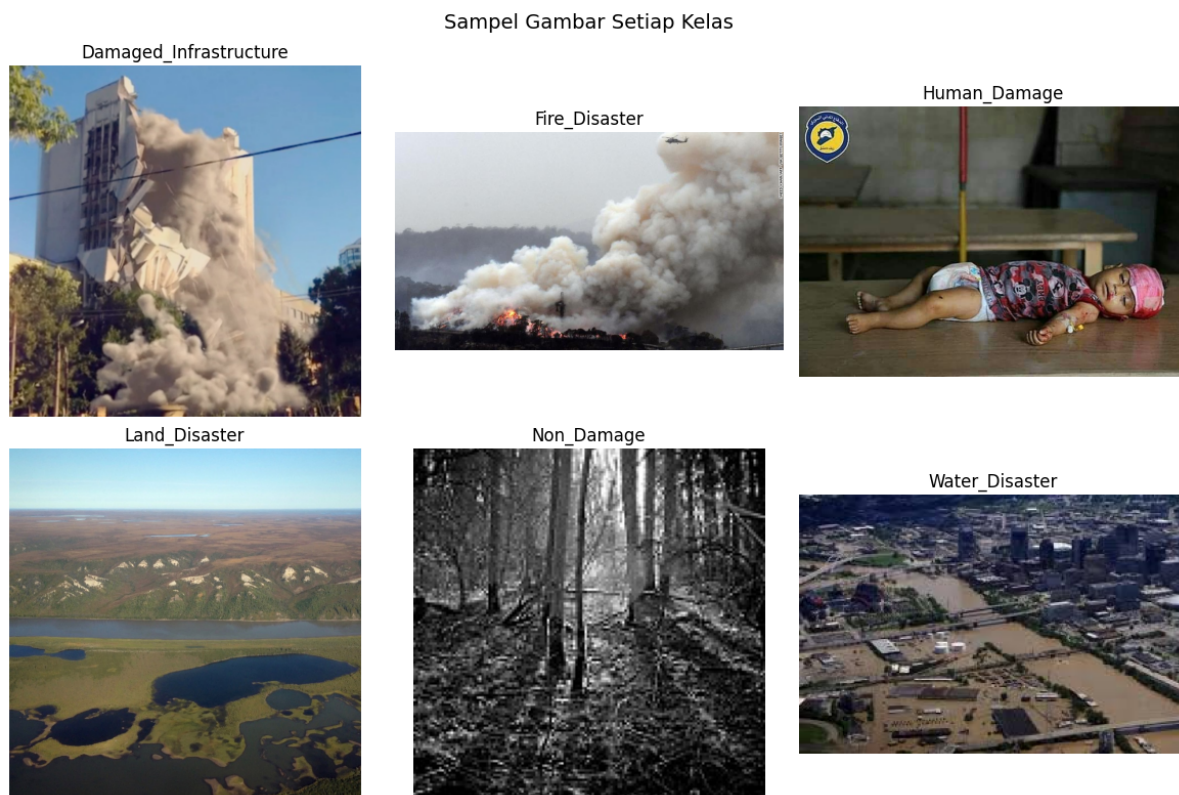
Berbagai penelitian telah dilakukan untuk memanfaatkan data media sosial dalam analisis bencana. Fan dan tim menggunakan *hybrid machine learning pipeline* dan *Named Entity Recognition* (NER) untuk menganalisis serta memetakan peristiwa bencana pada kasus Badai Harvey di Houston, Texas (Fan dkk., 2020). Penelitian Abdullah dan tim menggabungkan *remote sensing* dengan penyebaran informasi pada media sosial untuk mendukung analisis bencana (Abdullah dkk., 2020), sedangkan Puttinaovarat dan tim mengembangkan metode analisis bencana berbasis *remote sensing* dan *mobile Geographic Information System* (GIS) (Puttinaovarat dan Horkaew, 2020). Penelitian-penelitian tersebut menunjukkan bahwa data media sosial memiliki potensi besar dalam mendukung sistem pemantauan bencana, namun tantangan dalam ekstraksi fitur dan peningkatan akurasi klasifikasi masih menjadi perhatian utama. Variasi kondisi visual pada citra bencana juga menyebabkan proses klasifikasi menjadi lebih kompleks sehingga diperlukan model dengan kemampuan generalisasi yang baik Arif dkk. (2020). Perbedaan kualitas citra, sudut pengambilan gambar, kondisi pencahayaan, serta keberadaan objek yang saling tumpang tindih dapat meningkatkan tingkat kesulitan dalam proses ekstraksi fitur dan pengenalan pola visual (Ahuja dkk., 2022).

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan penggunaan arsitektur *Inception ResNet V2* untuk mengklasifikasikan bencana berdasarkan data citra yang diperoleh dari media sosial. Arsitektur ini dipilih karena mampu menggabungkan keunggulan *Inception module* dalam mengekstraksi fitur multiskala dengan kemampuan *Residual Network* dalam menjaga stabilitas aliran gradien sehingga mendukung pelatihan jaringan yang lebih dalam tanpa penurunan performa (Zhuang dkk., 2025). Selain itu, model *Convolutional Neural Network* (CNN) digunakan sebagai *baseline* untuk mengevaluasi peningkatan performa yang dihasilkan oleh *Inception ResNet V2* pada klasifikasi citra bencana berbasis media sosial (Nag dkk., 2020). Pendekatan ini diharapkan mampu menghasilkan performa klasifikasi yang lebih baik pada data citra bencana dengan karakteristik visual yang beragam.

## 2 Metode

### 2.1 Dataset

Dataset yang digunakan berasal dari *Disaster Dataset* yang diperkenalkan oleh Niloy dan tim (Niloy dkk., 2021). Dataset ini berisi citra bencana dan non-bencana yang digunakan untuk kebutuhan klasifikasi citra berbasis *deep learning*. Dataset dikumpulkan dari berbagai sumber dan dirancang untuk mendukung penelitian terkait klasifikasi bencana berbasis citra. Dataset terdiri dari 13.555 citra yang dikelompokkan ke dalam enam kelas, yaitu *Damaged Infrastructure*, *Fire Disaster*, *Human Damage*, *Land Disaster*, *Non Damage*, dan *Water Disaster*. Setiap kelas merepresentasikan kondisi visual yang berbeda, baik akibat kejadian bencana maupun area yang tidak terdampak. Dataset ini sangat sesuai digunakan dalam penelitian deteksi dan klasifikasi bencana berbasis citra. Gambar 1 menunjukkan contoh citra pada setiap kelas dataset yang digunakan dalam penelitian.



Gambar 1: Sampel data pada setiap kelas dataset.

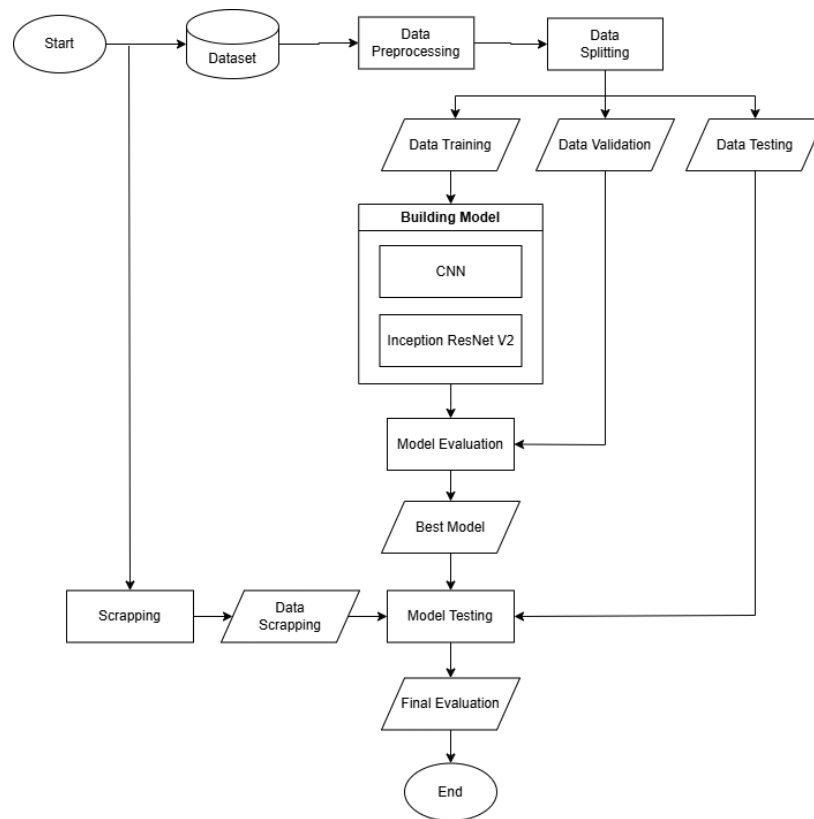
Distribusi data pada setiap kelas ditunjukkan pada Tabel 1. Terlihat bahwa distribusi data antar kelas tidak seimbang. Kelas *Non Damage* memiliki jumlah data terbesar, sedangkan kelas *Human Damage* memiliki jumlah data paling sedikit. Ketidakseimbangan ini perlu diperhatikan dalam proses pelatihan model karena dapat memengaruhi performa klasifikasi, terutama dalam mengenali kelas minoritas. Evaluasi model tidak hanya menggunakan metrik akurasi, tetapi juga precision, recall, specificity, dan f1-score untuk memberikan gambaran performa model yang lebih komprehensif pada setiap kelas.

Tabel 1: Distribusi jumlah data pada setiap kelas.

| Kelas                  | Jumlah Data |
|------------------------|-------------|
| Damaged Infrastructure | 1453        |
| Fire Disaster          | 933         |
| Human Damage           | 240         |
| Land Disaster          | 657         |
| Non Damage             | 9237        |
| Water Disaster         | 1035        |

## 2.2 Alur Penelitian

Penelitian ini dilaksanakan melalui serangkaian tahapan yang sistematis dan terstruktur untuk mencapai tujuan penelitian. Setiap tahapan disusun berdasarkan alur kerja *deep learning* dalam pengolahan citra digital, sehingga proses penelitian dapat berjalan secara terarah dan konsisten. Gambar 2 menunjukkan alur penelitian yang digunakan dalam proses pengembangan model klasifikasi citra bencana.



Gambar 2: Diagram Alur Penelitian.

Penelitian ini terdiri dari beberapa tahap utama yang saling berkaitan dan mendukung proses pengembangan model klasifikasi citra bencana. Setiap tahapan memiliki tujuan dan peran masing-masing dalam memastikan penelitian berjalan secara sistematis dan menghasilkan model dengan kinerja yang optimal. Beberapa tahapan yang dilakukan yaitu:

### 1. Disaster Dataset

Dataset yang digunakan diperoleh dari Kaggle yang berisi citra bencana dan non-bencana yang terbagi ke dalam beberapa kelas, yaitu *Damaged Infrastructure*, *Fire Disaster*, *Human Damage*, *Land Disaster*, *Non Damage*, dan *Water Disaster*. Dataset ini digunakan dalam proses pelatihan dan evaluasi model.

### 2. Preprocessing

Tahap *preprocessing* dilakukan untuk menyesuaikan citra agar sesuai dengan kebutuhan model *deep learning*. Proses ini meliputi penyesuaian ukuran citra, normalisasi nilai piksel, serta augmentasi data untuk meningkatkan variasi data pelatihan dan mengurangi risiko *overfitting*. Penanganan ketidakseimbangan data dilakukan secara mitigatif melalui augmentasi data pada proses pelatihan, seperti rotasi, zoom, dan *horizontal flip*, sehingga variasi citra pada data pelatihan dapat meningkat. Penelitian ini tidak menerapkan *oversampling* atau *undersampling*, namun pemanfaatan arsitektur *Inception ResNet V2* membantu model dalam mengekstraksi fitur visual yang lebih kompleks dan multiskala, sehingga performa klasifikasi pada kelas minoritas tetap dapat dipertahankan dengan baik. Oleh karena itu, evaluasi model tidak hanya menggunakan accuracy, tetapi juga sensitivity, specificity, precision, dan f1-score agar performa model terhadap seluruh kelas dapat dianalisis secara lebih proporsional.

### 3. Pembagian Data

Pembagian data dilakukan dengan membagi dataset ke dalam tiga bagian, yaitu data *training*, data *validation*, dan data *testing* dengan rasio 70:30 acak pada setiap kelas untuk menjaga keseimbangan distribusi data. Data kemudian disimpan dalam direktori terpisah sesuai dengan jenis data.

### 4. Pelatihan Model

Tahap pelatihan dilakukan menggunakan data *training* dengan menerapkan arsitektur CNN dan *Inception ResNet V2*. Data *validation* digunakan untuk memantau kinerja model dan menentukan model terbaik selama proses pelatihan.

### 5. Evaluasi Model

Evaluasi model dilakukan untuk mengukur performa model menggunakan metrik *accuracy*, *recall (sensitivity)*, *precision*, *f1-score*, dan *confusion matrix*. Hasil evaluasi digunakan untuk menentukan model terbaik yang akan digunakan pada tahap pengujian selanjutnya.

### 6. Scraping Data

Proses *scraping* dilakukan untuk memperoleh data tambahan di luar dataset utama. Pengambilan data dilakukan menggunakan *BingImageCrawler* berdasarkan kata kunci pada masing-masing kelas bencana. Pengambilan data tidak dilakukan secara langsung melalui API media sosial seperti Twitter/X karena adanya keterbatasan akses data dan pembatasan API. Oleh karena itu, proses pengumpulan data dilakukan melalui metode alternatif berbasis *web scraping* yang masih memuat unggahan citra bencana dari media sosial.

### 7. Testing Data Scraping

Model terbaik diuji menggunakan data hasil proses *scraping* untuk mengetahui kemampuan generalisasi model terhadap data di luar dataset utama. Data hasil *scraping* kemudian melalui proses seleksi untuk menghapus data duplikat dan citra berkualitas rendah sebelum digunakan sebagai data pengujian tambahan.

## 2.3 Convolutional Neural Network

CNN merupakan jenis jaringan saraf yang menggunakan lapisan konvolusional untuk mengekstraksi fitur dari citra, kemudian menggunakan lapisan yang terhubung sepenuhnya untuk

mengklasifikasikan citra. Arsitektur CNN yang digunakan pada penelitian ini diilustrasikan oleh Gambar 3. Arsitektur ini terdiri dari *input layer*, beberapa *convolutional layer*, *pooling layer*, dan *fully connected layer*. Setiap lapisan memiliki peran dalam mengekstraksi fitur visual dan mengubah representasi citra menjadi informasi yang dapat digunakan untuk proses klasifikasi Gupta dkk. (2022).

CNN mampu menangkap pola lokal seperti tepi, tekstur, dan bentuk melalui operasi konvolusi. Setiap *convolutional layer* terdiri atas sejumlah filter yang melakukan operasi konvolusi terhadap citra masukan untuk menghasilkan *feature map*. Operasi konvolusi pada CNN ditunjukkan pada persamaan 1.

$$h_{ij}^{(k)} = f \left( \sum_m \sum_n w_{mn}^{(k)} \cdot x_{(i+m)(j+n)} + b^{(k)} \right) \quad (1)$$

$h_{ij}^{(k)}$  menyatakan nilai aktivasi pada posisi  $(i, j)$  untuk filter ke- $k$ ,  $w_{mn}^{(k)}$  merupakan bobot kernel,  $x$  adalah input, dan  $f(\cdot)$  fungsi aktivasi *nonlinear* seperti ReLU. *pooling layer* digunakan untuk mempertahankan informasi penting melalui operasi *max pooling* dan *average pooling* (Aamir dkk., 2021). Operasi *max pooling* dan *average pooling* ditunjukkan pada persamaan 2 dan persamaan 3.

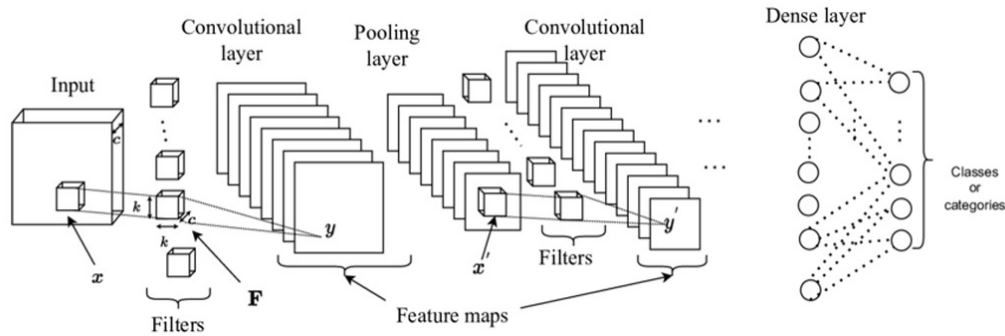
$$p_{ij} = \max_{(m,n) \in R_{ij}} h_{mn} \quad (2)$$

$$p_{ij} = \frac{1}{|R_{ij}|} \sum_{(m,n) \in R_{ij}} h_{mn} \quad (3)$$

Hasil dari proses konvolusi dan *pooling* kemudian diratakan (*flattened*) dan diteruskan ke *fully connected layer* untuk proses klasifikasi. Proses klasifikasi pada *fully connected layer* ditunjukkan pada persamaan 4.

$$y = f(Wx + b) \quad (4)$$

$W$  merupakan bobot,  $x$  vektor fitur hasil *flattening*, dan  $f$  fungsi aktivasi seperti *softmax* untuk menghasilkan probabilitas setiap kelas (Awasthi dkk., 2024). Gambar 3 menunjukkan arsitektur CNN yang digunakan dalam penelitian.



Gambar 3: Arsitektur CNN.



## 2.4 Inception ResNet V2

Inception ResNet V2 merupakan model jaringan konvolusional dalam yang dirancang untuk mengenali pola visual citra dengan kompleksitas tinggi. Arsitektur ini menggabungkan konsep *Inception module* dan *Residual Network* untuk meningkatkan kemampuan ekstraksi fitur sekaligus menjaga stabilitas pelatihan. Koneksi residual membantu mengatasi permasalahan *vanishing gradient*, yaitu kondisi ketika nilai gradien semakin mengecil selama proses *backpropagation* (Alexandre, 2023). Informasi dari layer sebelumnya tetap dipertahankan sehingga proses pembelajaran menjadi lebih stabil dan efisien. Kombinasi kedua konsep tersebut menghasilkan model dengan kemampuan generalisasi yang lebih baik terhadap variasi data citra (Taşci dkk., 2023).

Arsitektur Inception ResNet V2 menerima input citra berukuran  $299 \times 299 \times 3$  yang diproses melalui *Stem block* sebagai tahap awal ekstraksi fitur. Selanjutnya, jaringan terdiri dari beberapa blok utama yaitu *Inception-ResNet-A*, *Reduction-A*, *Inception-ResNet-B*, *Reduction-B*, dan *Inception-ResNet-C*. Setiap blok memanfaatkan kombinasi konvolusi paralel untuk menangkap fitur multiskala, sedangkan blok reduksi berfungsi menurunkan dimensi *feature map* sekaligus meningkatkan jumlah channel sehingga representasi fitur menjadi lebih kaya. Struktur ini memungkinkan model untuk mempelajari pola visual yang kompleks pada citra dengan lebih efektif dibandingkan arsitektur konvensional.

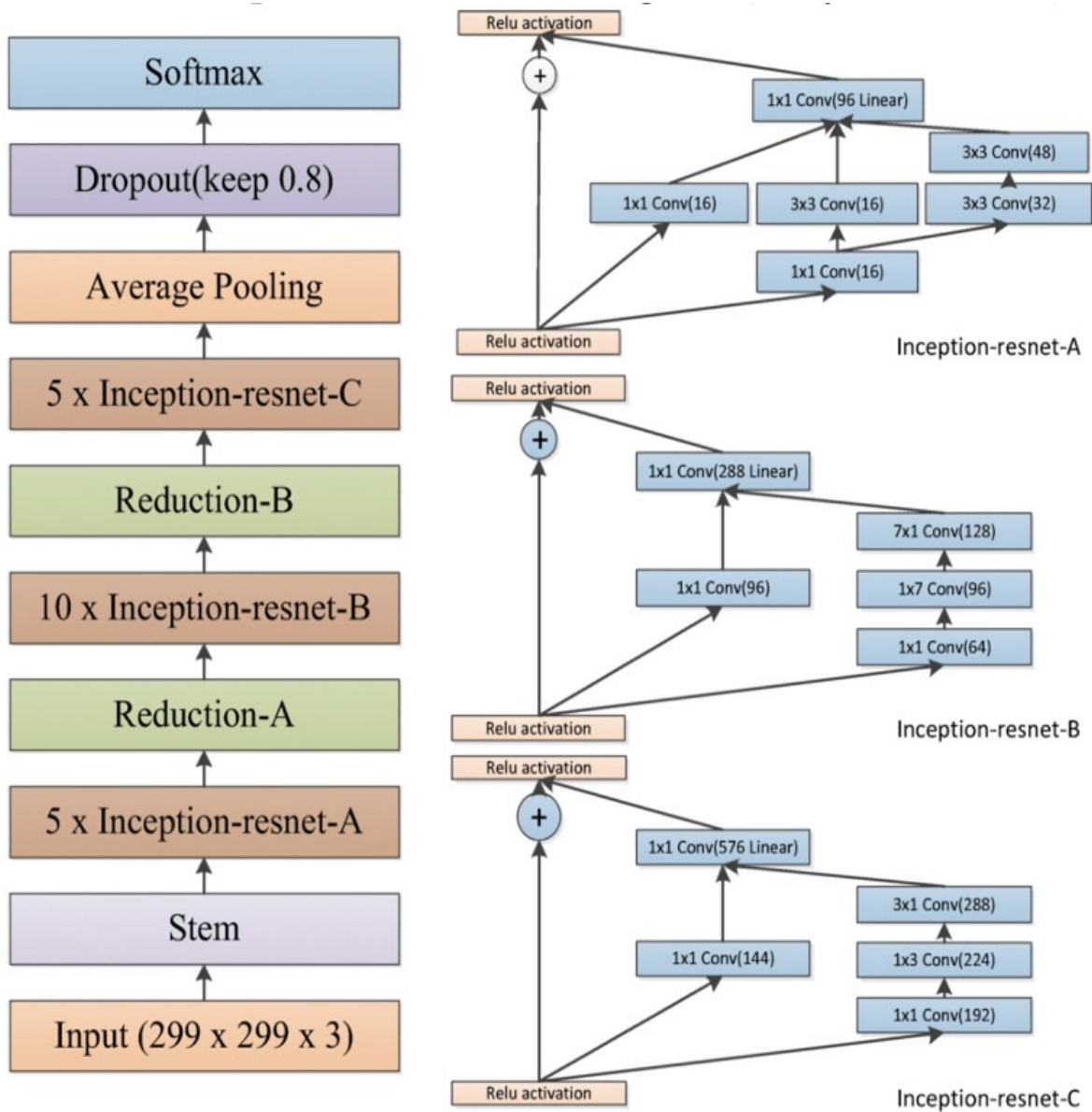
Pada tahap akhir, *feature map* hasil ekstraksi diproses menggunakan *global average pooling* untuk merangkum informasi spasial menjadi vektor fitur global. Selanjutnya diterapkan *dropout* untuk mengurangi risiko *overfitting* dengan cara menonaktifkan sebagian neuron secara acak selama pelatihan. Lapisan terakhir menggunakan fungsi aktivasi *softmax* untuk menghasilkan probabilitas setiap kelas (Froughirad dkk., 2024). Penggunaan *global average pooling* juga membantu mengurangi jumlah parameter pada model sehingga proses pelatihan menjadi lebih efisien tanpa mengurangi kemampuan representasi fitur secara signifikan. Koneksi residual pada setiap blok dapat dinyatakan pada persamaan 5 sebagai:

$$y = f(x) + x \quad (5)$$

dengan  $x$  sebagai input dan  $f(x)$  sebagai transformasi nonlinier. Penjumlahan ini memungkinkan jaringan mempertahankan informasi awal serta menjaga aliran gradien tetap stabil selama proses pelatihan. Proses klasifikasi menggunakan fungsi aktivasi *softmax* yang dirumuskan pada persamaan 6 sebagai:

$$y_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (6)$$

$y_i$  merupakan probabilitas untuk kelas ke- $i$ ,  $z_i$  adalah nilai logit, dan  $K$  jumlah kelas. Fungsi ini memastikan nilai probabilitas berada pada rentang 0 hingga 1 dengan total keseluruhan sebesar 1 (Yahya dkk., 2023). Model *Inception ResNet V2* memiliki keunggulan pada kemampuan ekstraksi fitur multiskala melalui kombinasi *Inception module* dan koneksi residual yang memungkinkan proses pelatihan jaringan dalam berlangsung lebih stabil. Penggunaan arsitektur ini membantu model dalam mengenali pola visual yang kompleks pada citra bencana, termasuk variasi bentuk kerusakan, pencahayaan, tekstur, dan kondisi lingkungan yang berbeda pada setiap kelas. Selain itu, pemanfaatan bobot *pretrained* dari ImageNet memungkinkan proses pelatihan menjadi lebih efisien karena model telah memiliki representasi fitur dasar yang baik sebelum dilakukan penyesuaian pada dataset bencana yang digunakan dalam penelitian (Chen, 2025). Gambar 4 menunjukkan arsitektur *Inception ResNet V2* yang digunakan dalam penelitian.



Gambar 4: Arsitektur Inception ResNet V2.

### 3 Hasil dan Pembahasan

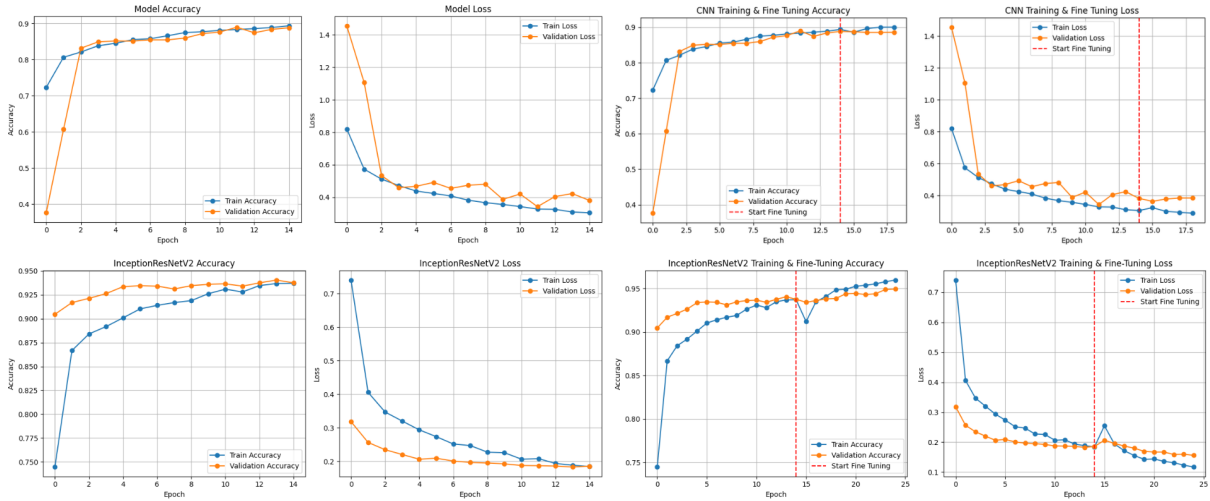
#### 3.1 Kinerja Model

Gambar 5 menunjukkan perkembangan nilai *training accuracy*, *validation accuracy*, *training loss*, dan *validation loss* pada model CNN dan *Inception ResNet V2* sebelum maupun sesudah *fine tuning*. Pada model CNN, *training accuracy* meningkat dari sekitar 72% pada epoch awal menjadi mendekati 89% pada akhir pelatihan. *Validation accuracy* juga mengalami peningkatan signifikan dari sekitar 38% menjadi sekitar 89% yang menunjukkan bahwa model mampu mem-



pelajari pola dasar citra bencana dengan baik. Selain itu, *training loss* menurun dari sekitar 0,81 menjadi 0,31, sedangkan *validation loss* turun dari sekitar 1,46 menjadi 0,38. Selisih yang relatif kecil antara *training accuracy* dan *validation accuracy* menunjukkan bahwa model memiliki kemampuan generalisasi yang cukup baik dan tidak mengalami *overfitting* yang signifikan selama proses pelatihan. Namun setelah *fine tuning* diterapkan, peningkatan performa relatif terbatas karena kurva *accuracy* dan *loss* cenderung stabil di sekitar nilai sebelumnya. Kondisi ini mengindikasikan bahwa CNN telah mencapai kapasitas optimalnya pada tahap pelatihan awal sehingga *fine tuning* tidak memberikan kontribusi yang berarti terhadap peningkatan performa model.

Model *Inception ResNet V2* menunjukkan performa yang lebih tinggi sejak awal pelatihan. *Validation accuracy* telah mencapai lebih dari 90% pada epoch pertama dan terus meningkat hingga mendekati 94%. Nilai *training loss* juga menurun secara konsisten dari sekitar 0,74 menjadi 0,19, sedangkan *validation loss* turun dari sekitar 0,32 menjadi 0,18. Setelah *fine tuning* dilakukan pada epoch ke-15, terjadi peningkatan *training accuracy* dari sekitar 93,8% menjadi 95,6% dan *validation accuracy* dari sekitar 94% menjadi 95%. Selisih yang relatif kecil antara *training accuracy* dan *validation accuracy* menunjukkan bahwa model mampu mempertahankan kemampuan generalisasi yang baik setelah proses *fine tuning*. Meskipun sempat terjadi fluktuasi sesaat setelah proses *fine tuning*, performa model kembali meningkat dan stabil hingga akhir pelatihan. Penurunan nilai *loss* yang konsisten serta tingginya nilai *accuracy* menunjukkan bahwa model mampu memanfaatkan bobot *pretrained* secara efektif untuk mengekstraksi fitur visual yang lebih representatif. Hasil ini menunjukkan bahwa *fine tuning* berhasil meningkatkan kemampuan model dalam menyesuaikan representasi fitur tingkat tinggi dengan karakteristik citra bencana sehingga menghasilkan performa klasifikasi yang lebih baik dibandingkan CNN.



Gambar 5: Grafik *training accuracy*, *validation accuracy*, *training loss*, dan *validation loss* pada model CNN dan *Inception ResNet V2* sebelum dan sesudah *fine tuning*.

### 3.2 Evaluasi Model

Tabel 2 menyajikan perbandingan kinerja model CNN dan Inception ResNet V2 sebelum dan setelah *fine tuning* berdasarkan metrik *accuracy*, *sensitivity*, *specificity*, *precision*, dan *f1-score*. Hasil menunjukkan bahwa Inception ResNet V2 secara konsisten unggul dibandingkan

CNN pada seluruh metrik evaluasi. Model CNN tanpa fine tuning memperoleh accuracy sebesar 0,8840 dengan sensitivity 0,6275 dan f1-score 0,6501. Nilai specificity yang tinggi (0,9742) menunjukkan kemampuan model dalam mengenali kelas Non Damage, namun sensitivity yang relatif rendah mengindikasikan keterbatasan dalam mendeteksi seluruh kelas bencana. Setelah dilakukan fine tuning, performa CNN tidak mengalami peningkatan signifikan, dengan accuracy menurun menjadi 0,8811 dan f1-score menjadi 0,6472.

Inception ResNet V2 tanpa fine tuning telah menunjukkan performa tinggi dengan accuracy 0,9489, sensitivity 0,8833, dan f1-score 0,8883. Nilai precision (0,8942) dan specificity (0,9885) yang tinggi menunjukkan kemampuan model dalam mengklasifikasikan citra secara akurat dan konsisten. Setelah fine tuning, performa model meningkat dengan accuracy mencapai 0,9523, sensitivity 0,8888, dan f1-score 0,8917. Peningkatan ini menunjukkan bahwa fine tuning mampu menyesuaikan representasi fitur tingkat tinggi tanpa mengurangi kemampuan generalisasi model. Inception ResNet V2 dengan fine tuning merupakan model terbaik dalam penelitian ini, karena mampu mencapai keseimbangan optimal antara sensitivity dan precision serta menunjukkan stabilitas yang baik pada seluruh metrik evaluasi.

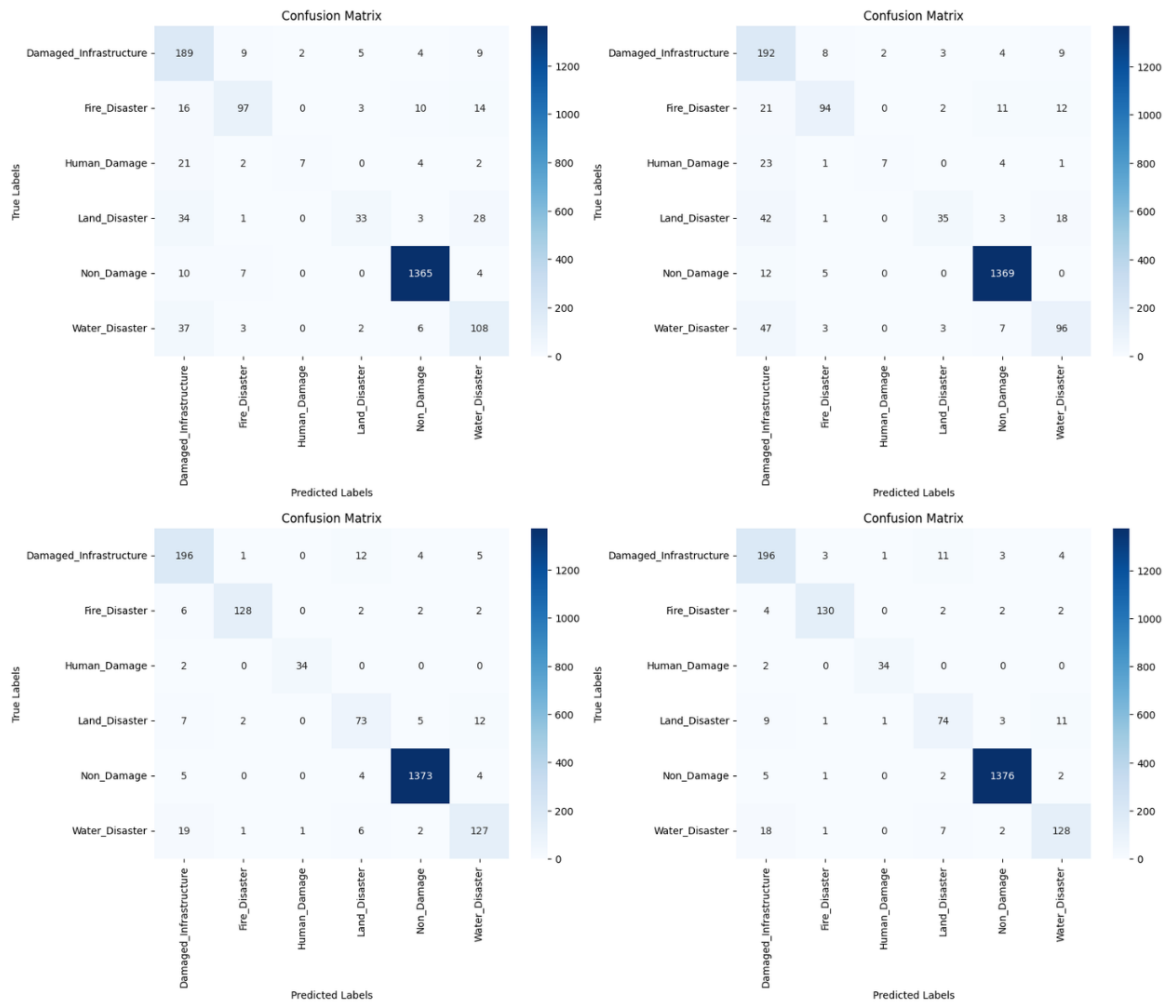
Tabel 2: Perbandingan kinerja model

| Model                                    | Accuracy      | Sensitivity   | Specificity   | Precision     | F1-Score      |
|------------------------------------------|---------------|---------------|---------------|---------------|---------------|
| CNN                                      | 0.8840        | 0.6275        | 0.9742        | 0.7685        | 0.6501        |
| CNN (Fine Tuning)                        | 0.8811        | 0.6172        | 0.9733        | 0.7810        | 0.6472        |
| Inception ResNet V2                      | 0.9489        | 0.8833        | 0.9885        | 0.8942        | 0.8883        |
| <b>Inception ResNet V2 (Fine Tuning)</b> | <b>0.9523</b> | <b>0.8888</b> | <b>0.9897</b> | <b>0.8954</b> | <b>0.8917</b> |

Gambar 6 menampilkan *confusion matrix* dari model CNN sebelum *fine tuning* (kiri atas), CNN sesudah *fine tuning* (kanan atas), *Inception ResNet V2* sebelum *fine tuning* (kiri bawah), dan *Inception ResNet V2* sesudah *fine tuning* (kanan bawah). Nilai pada diagonal utama menunjukkan jumlah prediksi yang benar, sedangkan nilai di luar diagonal menunjukkan kesalahan klasifikasi antar kelas. Semakin besar nilai pada diagonal utama, semakin baik kemampuan model dalam mengenali kelas yang sesuai. Sebaliknya, nilai yang tersebar di luar diagonal menunjukkan adanya kemiripan karakteristik visual antar kelas yang menyebabkan kesalahan prediksi.

Pada model CNN, kelas *Non Damage* memiliki jumlah prediksi benar tertinggi dengan 1.365 citra sebelum *fine tuning* dan meningkat menjadi 1.369 citra setelah *fine tuning*. Peningkatan juga terlihat pada kelas *Damaged Infrastructure* dari 189 menjadi 192 citra, *Human Damage* dari 7 menjadi 8 citra, serta *Land Disaster* dari 33 menjadi 35 citra. Namun, tidak semua kelas mengalami perbaikan performa. Jumlah prediksi benar pada kelas *Fire Disaster* menurun dari 124 menjadi 123 citra dan kelas *Water Disaster* menurun dari 108 menjadi 107 citra setelah *fine tuning*. Hasil ini menunjukkan bahwa proses *fine tuning* pada CNN hanya memberikan peningkatan yang terbatas dan belum mampu meningkatkan performa model secara konsisten pada seluruh kelas.

Model *Inception ResNet V2* menunjukkan distribusi prediksi yang lebih terkonsentrasi pada diagonal utama dibandingkan CNN. Jumlah prediksi benar pada kelas *Damaged Infrastructure* meningkat dari 196 menjadi 197 citra, sedangkan kelas *Fire Disaster* meningkat dari 128 menjadi 130 citra setelah *fine tuning*. Kelas *Non Damage* juga meningkat dari 1.373 menjadi 1.376 citra, sementara kelas *Water Disaster* meningkat dari 127 menjadi 128 citra. Hasil ini menunjukkan bahwa *Inception ResNet V2* mampu mengekstraksi fitur visual dengan lebih baik sehingga menghasilkan performa klasifikasi yang lebih akurat dan stabil dibandingkan CNN.



Gambar 6: *Confusion matrix* pada model CNN dan *Inception ResNet V2* sebelum dan sesudah *fine tuning*.

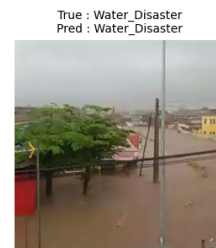
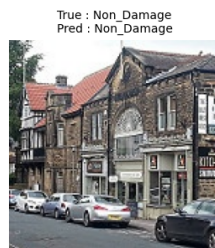
### 3.3 Hasil Prediksi

Gambar 7 menampilkan contoh hasil prediksi benar dan prediksi salah yang dihasilkan oleh model *Inception ResNet V2*. Pada bagian prediksi benar, model mampu mengidentifikasi karakteristik visual utama dari setiap kelas bencana dengan baik. Kelas *Damaged Infrastructure* dikenali melalui indikasi kerusakan fisik pada kendaraan dan bangunan, sedangkan kelas *Fire Disaster* dikenali melalui keberadaan api dan asap yang menjadi objek dominan pada citra. Kelas *Human Damage* berhasil diklasifikasikan dengan benar karena adanya indikasi korban manusia yang terlihat jelas. Sementara itu, kelas *Land Disaster*, *Non Damage*, dan *Water Disaster* dapat dikenali melalui karakteristik visual yang khas seperti tanah retak, lingkungan normal tanpa kerusakan, dan genangan air yang mendominasi area citra.

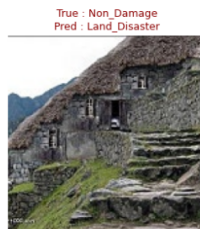
Pada bagian prediksi salah, terlihat bahwa kesalahan klasifikasi umumnya terjadi pada citra dengan karakteristik visual yang mirip antar kelas. Sebagai contoh, citra *Damaged Infrastructure* diprediksi sebagai *Water Disaster* karena adanya refleksi dan genangan air yang lebih dominan.

an dibandingkan objek kerusakan. Citra *Fire Disaster* yang diprediksi sebagai *Water Disaster* menunjukkan bahwa objek api hanya menempati sebagian kecil area citra sehingga fitur yang dipelajari model kurang representatif. Kesalahan juga terjadi pada kelas *Human Damage* yang diprediksi sebagai *Non Damage*, yang mengindikasikan bahwa keberadaan manusia saja belum cukup menjadi pembeda tanpa adanya ciri kerusakan yang jelas. Selain itu, citra *Water Disaster* dan *Land Disaster* menunjukkan kemungkinan terjadinya tumpang tindih karakteristik visual akibat kemiripan tekstur permukaan dan kondisi lingkungan yang diamati dari jarak jauh.

#### Prediksi Benar



#### Prediksi Salah



Gambar 7: Hasil prediksi.

### 3.4 Evaluasi Data Scraping

Data *scraping* diperoleh melalui proses *web scraping* menggunakan *Bing Image* sebagai alternatif sumber data karena keterbatasan akses dan kebijakan privasi yang membatasi pengambilan data secara langsung dari media sosial. Tabel 3 menunjukkan model memperoleh *accuracy* sebesar 0,9026, *precision* sebesar 0,9052, *recall* (*sensitivity*) sebesar 0,8864, *specificity* sebesar 0,9803, serta *f1-score* sebesar 0,8891. Nilai-nilai ini menunjukkan bahwa model tetap mempertahankan performa yang baik ketika dihadapkan pada data dunia nyata dengan karakteristik visual yang lebih beragam. Capaian ini mengindikasikan bahwa model tidak hanya mampu menghafal pola pada data pelatihan, tetapi juga memiliki kemampuan adaptasi yang cukup kuat terhadap variasi gambar dan kondisi lingkungan yang berbeda pada data hasil *scraping*. Selain itu, tingginya nilai *specificity* menunjukkan bahwa model mampu membedakan citra bencana dan non bencana dengan cukup baik meskipun data yang digunakan memiliki kualitas dan kondisi visual yang tidak seragam. Hasil ini menunjukkan bahwa pendekatan berbasis *deep learning* menggunakan *Inception ResNet V2* memiliki potensi untuk diterapkan pada sistem pemantauan bencana berbasis citra secara otomatis.

Tabel 3: Kinerja model pada data *scraping*

| Evaluasi    | Nilai  |
|-------------|--------|
| Accuracy    | 0,9026 |
| Sensitivity | 0,8864 |
| Specificity | 0,9803 |
| Precision   | 0,9052 |
| F1-Score    | 0,8891 |

### 3.5 Penelitian Terdahulu

Penelitian ini membandingkan beberapa penelitian terdahulu yang memiliki fokus serupa pada klasifikasi citra bencana menggunakan pendekatan *deep learning*. Tabel 4 menunjukkan berbagai arsitektur yang telah digunakan, seperti VGG16, ResNet50, EfficientNet-B1, serta kombinasi mBERT dan ResNet50 dengan capaian performa yang bervariasi. Model EfficientNet-B1 yang digunakan oleh Alam dkk memperoleh *accuracy* sebesar 0,82 dan *f1-score* sebesar 0,80, sedangkan model VGG16 yang digunakan oleh Asif dkk dan Zou dkk menghasilkan *accuracy* sebesar 0,83. Wen dkk melaporkan performa yang lebih tinggi menggunakan ResNet50 dengan *accuracy*, *precision*, *recall*, dan *f1-score* masing-masing sebesar 0,87. Sementara itu, kombinasi mBERT dan ResNet50 yang digunakan oleh Islam dkk memperoleh *accuracy* sebesar 0,83.

Dibandingkan dengan penelitian-penelitian tersebut, model *Inception ResNet V2* pada penelitian ini menunjukkan performa terbaik dengan *accuracy* sebesar 0,95, *precision* sebesar 0,89, *recall* sebesar 0,88, dan *f1-score* sebesar 0,89. Hasil ini menunjukkan bahwa arsitektur *Inception ResNet V2* memiliki kemampuan yang baik dalam mengekstraksi fitur visual citra bencana sehingga mampu menghasilkan kinerja klasifikasi yang lebih tinggi dibandingkan beberapa metode yang dilaporkan pada penelitian terdahulu. Namun demikian, masing-masing penelitian menggunakan dataset, skenario pelatihan, *preprocessing*, serta pembagian data yang berbeda sehingga hasil evaluasi tidak dapat dibandingkan secara langsung. Oleh karena itu, perbandingan utama dalam penelitian ini tetap dilakukan pada model CNN dan *Inception ResNet V2* menggunakan dataset dan perlakuan eksperimen yang sama..

Tabel 4: Ringkasan penelitian terdahulu terkait klasifikasi citra bencana

| Peneliti              | Tahun       | Model                      | Accuracy    | Precision   | Recall      | F1-Score    |
|-----------------------|-------------|----------------------------|-------------|-------------|-------------|-------------|
| Alam dkk              | 2023        | EfficientNet-B1            | 0.82        | 0.81        | 0.82        | 0.80        |
| Asif dkk              | 2021        | VGG16                      | 0.83        | 0.80        | 0.82        | -           |
| Zou dkk               | 2021        | VGG16                      | 0.83        | 0.83        | 0.83        | 0.83        |
| Wen dkk               | 2025        | ResNet50                   | 0.87        | 0.87        | 0.87        | 0.87        |
| Islam dkk             | 2025        | mBERT + ResNet50           | 0.83        | 0.83        | 0.83        | 0.83        |
| <b>Penelitian ini</b> | <b>2025</b> | <b>Inception ResNet V2</b> | <b>0.95</b> | <b>0.89</b> | <b>0.88</b> | <b>0.89</b> |

## 4 Kesimpulan

Penelitian ini berhasil membangun model klasifikasi citra bencana menggunakan arsitektur *Inception ResNet V2*. Tahapan penelitian meliputi *preprocessing*, pembagian data, pemanfaatan bobot *pretrained ImageNet* sebagai *feature extractor*, serta penerapan *fine tuning* pada 40 lapisan terakhir *backbone*. Pendekatan tersebut mampu menyesuaikan fitur *pretrained* dengan karakteristik citra bencana secara lebih optimal. Hasil pengujian menunjukkan bahwa *Inception ResNet V2* memberikan performa yang lebih baik dibandingkan CNN. Model mencapai accuracy sebesar 0,9489 pada tahap awal dan meningkat menjadi 0,9523 setelah dilakukan *fine tuning*, dengan precision 0,8954, recall 0,8888, dan f1-score 0,8917. Hasil ini menunjukkan bahwa *Inception ResNet V2* mampu mengekstraksi fitur visual yang kompleks secara lebih efektif dibandingkan CNN.

Model terbaik kemudian diuji menggunakan data citra hasil *scraping* dari media sosial yang merepresentasikan kondisi dunia nyata dengan variasi kualitas citra, sudut pengambilan gambar, dan pencahayaan yang lebih beragam. Hasil pengujian menunjukkan bahwa model tetap mampu mempertahankan performa yang baik dengan accuracy 0,9026, sensitivity 0,8864, specificity 0,9803, precision 0,9052, dan f1-score 0,8891. Tingginya nilai specificity menunjukkan kemampuan model dalam membedakan citra bencana dan non-bencana secara akurat pada data yang lebih beragam. Penelitian ini menunjukkan bahwa integrasi model klasifikasi berbasis *deep learning* dengan citra media sosial berpotensi mendukung sistem pemantauan kejadian bencana secara otomatis dan membantu pengambilan keputusan secara lebih cepat dan tepat.

## Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Program Studi Sains Data Universitas Negeri Surabaya serta seluruh pihak yang telah memberikan dukungan dan masukan selama pelaksanaan penelitian ini.

## Pustaka

- Aamir, M., Ali, T., Irfan, M., Shaf, A., Azam, M. Z., Glowacz, A., Brumercik, F., Glowacz, W., Alqhtani, S., dan Rahman, S. (2021). Natural disasters intensity analysis and classification based on multispectral images using multi-layered deep convolutional neural network. *Sensors*, 21(8):2648.
- Abdullah, T., Bazi, Y., Zuair, M., Mekhalfi, M. L., Rangarajan, L., dan Zuair, M. (2020). Texts:



- Deep bidirectional triplet network for matching text to remote sensing images. *Remote Sensing*, 12(3):405.
- Ahuja, S. N., Michael, M., Selvan, M., dan Mantri, V. (2022). Natural disaster detection in social media and satellite imagery. *ITM Web of Conferences*.
- Alexandre, M. (2023). Resnet: Solving vanishing gradient in deep networks. In *Lecture Notes in Computer Science*, pages 235–247.
- Arif, M. A., Amin, A. A., Ali, A. A., dan Rahman, A. K. M. M. (2020). Visual attention-based comparative study on disaster detection from social media images. *Innovations in Systems and Software Engineering*, 16(3-4):309–319.
- Awasthi, A., Rao, S. B., dan Acharjya, K. (2024). Efficient image feature extraction using convolutional neural networks. In *Proceedings of the International Conference on Advanced Networking and Applications Technologies (ICONAT)*.
- Chen, S. (2025). A comparative study of inception v3 and inceptionresnetv2 for bathroom item classification with and without imagenet pretraining. *Science and Technology of Engineering, Chemistry and Environmental Protection*.
- Fan, C., Wu, F., dan Mostafavi, A. (2020). A hybrid machine learning pipeline for automated mapping of events and locations from social media in disasters. *IEEE Access*, 8:10478–10490.
- Froughirad, F., Eshtivani, R. B., Khajavi, H., dan Rastgoo, A. (2024). Mbception: A new multi-block inception model for enhancing image processing efficiency. *arXiv preprint arXiv:2412.13703*.
- Gupta, J., Pathak, S., dan Kumar, G. (2022). Deep learning (cnn) and transfer learning: A review. In *Journal of Physics: Conference Series*, volume 2273.
- Maqbool, A., Usmani, Z. U. A., Afzal, F., dan Razia, A. (2020). Disaster mitigation in urban pakistan using agent-based modeling with gis. *ISPRS International Journal of Geo-Information*, 9(4):203.
- Nag, S., Pal, T., Basu, S., dan Das Bit, S. (2020). Cnn based approach for post disaster damage assessment. In *International Conference on Distributed Computing and Networking (ICDCN)*.
- Niloy, F. F., Nayem, A. B. S., Sarker, A., Paul, O., Amin, M. A., Ali, A. A., Zaber, M. I., Rahman, A. M., dkk. (2021). A novel disaster image data-set and characteristics analysis using attention model. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 6116–6122. IEEE.
- Puttinaovarat, S. dan Horkaew, P. (2020). Internetworking flood disaster mitigation system based on remote sensing and mobile gis. *Geomatics, Natural Hazards and Risk*, 11(1):1886–1911.
- Taşci, B., Acharya, M. R., Baygin, M., Dogan, S., Tuncer, T., dan Belhaouari, S. B. (2023). Inception and concatenation residual block-based deep learning network for damaged building detection using remote sensing images. *ISPRS Journal of Applied Geomatics*.
- Ullah, W., Ullah, A., Haq, I. U., Muhammad, K., Sajjad, M., dan Baik, S. W. (2020). Cnn features with bi-directional lstm for real-time anomaly detection in surveillance networks. *Multimedia Tools and Applications*.
- Yahya, A. A., Liu, K., Hawbani, A., Wang, Y., dan Hadi, A. N. (2023). A novel image classification method based on residual network, inception, and proposed activation function. *Sensors*, 23(6):2976.
- Zhuang, X., Li, Y., Wang, J., Chen, L., dan Huang, H. (2025). Deep learning-based damage assessment of post-earthquake structures using pretrained cnn models including inception-resnet-v2. *Computers & Geosciences*, 182:104959.