

# Analisis Kerusakan Wilayah Akibat Bencana Berbasis Citra Satelit Menggunakan Algoritma ViT-S16 dan BiT-M

Akmal Muzakki Bakir<sup>1</sup>, Muhammad Zaki Rabbani<sup>2</sup>, Muhammad Sya'bani Falif<sup>3</sup>, Gamma Kosala<sup>4</sup>

<sup>1</sup>Fakultas Informatika Universitas Telkom, Bandung, 40257, email: akmalmuzakki@student.telkomuniversity.ac.id

<sup>2</sup>Fakultas Informatika Universitas Telkom, Bandung, 40257, email: zakirah@student.telkomuniversity.ac.id

<sup>3</sup>Fakultas Informatika Universitas Telkom, Bandung, 40257, email: msfalif@student.telkomuniversity.ac.id

<sup>4</sup>Fakultas Informatika Universitas Telkom, Bandung, 40257, email: gammakosala@telkomuniversity.ac.id

Corresponding Author: Akmal Muzakki Bakir

**INTISARI** — Indonesia terletak di kawasan Sirkum Pasifik, yang rawan bencana alam seperti gempa bumi, tsunami, dan gunung meletus, yang menyebabkan kerusakan signifikan. Penelitian ini mengembangkan sistem analisis tingkat kerusakan wilayah akibat bencana berbasis citra satelit. Metodologi yang digunakan menggabungkan pengolahan citra satelit, pembelajaran mesin, dan analisis spasial, dengan harapan dapat membantu percepatan proses penanggulangan bencana melalui informasi yang akurat dan tepat waktu. Penelitian ini mengimplementasikan model Big Transfer Medium (BiT-M) dan Vision Transformers (ViT-S16) untuk klasifikasi area rusak dalam citra satelit. Model BiT-M, khususnya BiT-M R101x3, menunjukkan kinerja unggul dengan akurasi klasifikasi tinggi sebesar 83,22%. Teknik transfer learning dengan model yang telah dilatih sebelumnya, seperti BiT-M R50x1, BiT-M R50x3, BiT-M R101x1, dan BiT-M R101x3, terbukti efektif dan efisien untuk tugas ini. Penelitian ini menggunakan dataset citra satelit sebelum dan sesudah bencana untuk fokus pada klasifikasi daerah dengan dan tanpa kerusakan. Saran penelitian selanjutnya termasuk pengembangan model hibrida, optimisasi model spesifik, dan aplikasi praktis dalam pemantauan bencana serta mitigasi risiko. Penelitian ini berkontribusi pada kemajuan metode klasifikasi citra satelit untuk pemantauan area kerusakan dan pengurangan risiko bencana.

**KATA KUNCI** — Bencana Alam, Citra Satelit, Analisis Kerusakan Wilayah, *Big Transfer Medium* (BiT-M), *Transfer Learning*, *Vision Transformers* (ViT-S16)

## I. PENDAHULUAN

### A. LATAR BELAKANG

Indonesia setiap tahun mengalami berbagai bencana alam seperti banjir, tsunami, gempa bumi, tanah longsor, dan gunung meletus, yang menyebabkan banyak korban jiwa dan kerugian materi. Pada Januari 2024, Indonesia mencatat 38 bencana, terutama banjir dan cuaca ekstrem, menurut BNPB. Bencana ini menyebabkan 49.000 rumah terdampak dan 58.000 orang mengungsi dalam satu pekan. Namun, BNPB menyoroti pentingnya waspada terhadap tanah longsor [1].

Secara geografis, Indonesia terletak di kawasan Sirkum Pasifik atau Cincin Api Pasifik, yang membuatnya rawan gempa bumi dan letusan gunung berapi. Selain itu, iklim tropis dengan curah hujan tinggi meningkatkan risiko bencana alam [2].

Beberapa lembaga pemerintah dan nonpemerintah, seperti BNPB, BPBD, dan PMI, bertugas menangani penanggulangan bencana alam. BNPB menggunakan pemetaan berbasis citra satelit untuk mengidentifikasi wilayah terdampak dan kerusakan bangunan. Namun, proses observasi dan pelabelan kerusakan masih manual, memakan waktu dan tenaga [3].

Tim Barudak Bojongsoang mengusulkan sistem analisis kerusakan berbasis citra satelit dengan teknik *representation learning*. Sistem ini mengidentifikasi dan memprediksi tingkat kerusakan dengan dua kategori: tanpa kerusakan dan terdeteksi kerusakan. Proses deteksi terkomputerisasi ini memungkinkan analisis yang cepat dan akurat, membantu pihak berwenang mengambil tindakan cepat.

### B. RUMUSAN MASALAH

- Bagaimana cara mengembangkan sistem analisis tingkat kerusakan wilayah akibat bencana berbasis citra satelit yang akurat?
- Model apa yang dapat digunakan untuk memprediksi tingkat kerusakan wilayah berbasis citra satelit?
- Bagaimana model prediksi kerusakan berbasis citra satelit menghasilkan klasifikasi kerusakan yang akurat?
- Bagaimana sistem analisis kerusakan berbasis citra satelit dapat membantu lembaga penanggulangan bencana?

### C. TUJUAN, MANFAAT, DAN POTENSI KEBERLANJUTAN

- Mengembangkan sistem analisis kerusakan wilayah berbasis citra satelit yang bermanfaat bagi pemerintah, organisasi non-pemerintah, dan lembaga penelitian untuk mempercepat penanggulangan bencana serta menyediakan informasi untuk

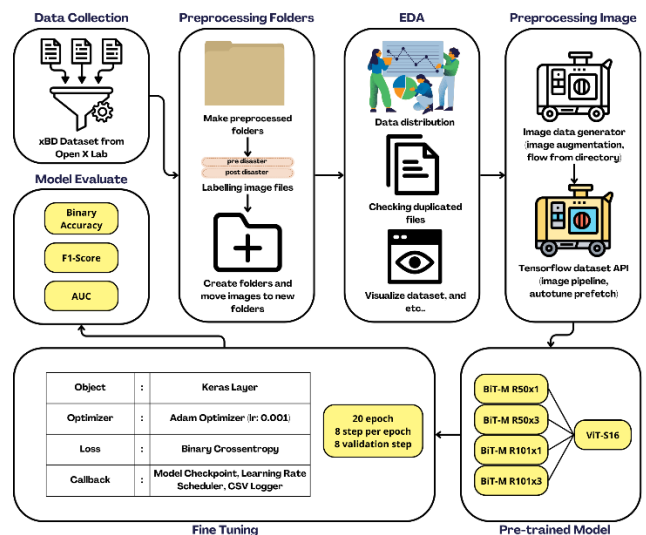
perencanaan distribusi bantuan dan pemulihan akses. Sistem ini dapat terus ditingkatkan untuk akurasi, efisiensi, dan kemudahan penggunaan.

- Membantu meningkatkan ketahanan masyarakat terhadap bencana dengan menyediakan informasi akurat dan tepat waktu tentang kerusakan wilayah. Informasi ini digunakan oleh lembaga pemerintah dan organisasi terkait untuk perencanaan distribusi bantuan, pemulihan, evakuasi, dan pembangunan kembali.
- Mendukung pencapaian *Sustainable Development Goals* PBB, khususnya tujuan nomor 11, 13, dan 16, melalui sistem analisis kerusakan berbasis citra satelit [4].

### D. BATASAN PENELITIAN

- Penelitian ini hanya menggunakan citra satelit publik yang telah dikumpulkan dan diverifikasi oleh ahli pada wilayah tertentu dari [Open X Lab](#).
- Metode yang dipaparkan melakukan prediksi terhadap 2 tingkat kerusakan, yaitu tidak ada kerusakan, dan terdeteksi adanya kerusakan.

### E. ALUR PENELITIAN



Gambar 1. Diagram alur penelitian

Data penelitian dikumpulkan dari [Open X Lab](#), termasuk dataset xBD berbasis citra satelit, dengan mengikuti protokol yang ada. Setelah itu, dilakukan folder preprocessing untuk merapikan data unduhan dan image preprocessing untuk membersihkan serta menyiapkan data untuk pelatihan model. Data augmentation, seperti rotasi gambar dan penambahan noise, juga dilakukan untuk memperbanyak data dan meningkatkan ketahanan model. Data kemudian dilatih menggunakan pretrained model melalui *fine tuning* yang sesuai dengan jenis data dan tujuan penelitian. Proses pelatihan bertujuan membangun model yang dapat memprediksi, mengklasifikasikan, atau menghasilkan keluaran yang diinginkan. Evaluasi model dilakukan dengan data *validation*, menggunakan metrik F1-Score, Binary Accuracy, dan AUC Score untuk mengukur kinerja dan akurasi model.

## II. KAJIAN PUSTAKA

### A. PENELITIAN TERKAIT

Chen, M., & Liu, Y. [5] melakukan studi komparatif model deep learning untuk memperkirakan kerusakan bangunan menggunakan dataset xBD yang berisi ribuan gambar area terdampak bencana. Langkah pertama melibatkan segmentasi bangunan dengan tiga model deep learning (U-Net, FPN, PAN) dan empat encoder (ResNet 50, VGG16, ResNext50 32x4D, DenseNet 121) untuk menemukan kombinasi terbaik. Setelah segmentasi, fitur diekstraksi dan dimasukkan ke dalam jaringan syaraf tiruan untuk mengklasifikasikan kerusakan bangunan.

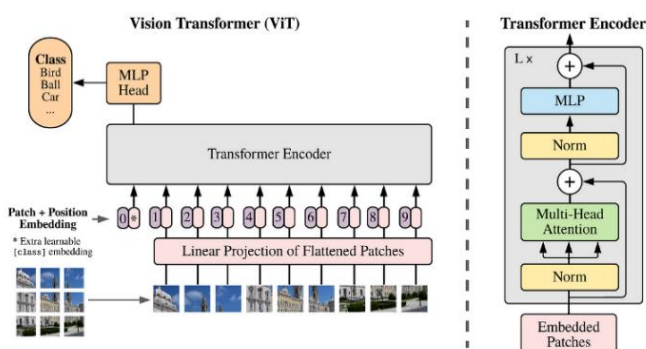
J. Z. Xu, W. Lu. et al. [6] memperkenalkan Big Transfer (BiT) sebagai metode pra-pelatihan baru untuk tugas-tugas computer vision. Dengan melatih dataset besar, BiT menghasilkan representasi visual yang dapat ditransfer ke berbagai tugas seperti klasifikasi gambar, deteksi objek, dan segmentasi. Penelitian ini menunjukkan kinerja unggul BiT pada lebih dari 20 dataset, melampaui metode yang sudah ada, dan menjanjikan model yang kuat dengan penggunaan data dan sumber daya yang lebih efisien.

Dosovitskiy et al. [7] memperkenalkan arsitektur Transformer untuk tugas *computer vision* tanpa menggunakan convolutional neural network (CNN). Alih-alih menggabungkan mekanisme perhatian dengan CNN, mereka mengusulkan model Transformer murni yang memproses urutan patch gambar. Vision Transformer (ViT) ini efektif dalam klasifikasi gambar, menunjukkan kinerja unggul pada ImageNet, CIFAR-100, dan VTAB ketika dilatih sebelumnya pada dataset besar. Selain itu, ViT mencapai hasil ini dengan sumber daya komputasi lebih sedikit dibandingkan dengan jaringan konvolusi canggih.

### B. DASAR TEORI

#### 1) ViT-S16

Vision Transformers (ViT-S16) adalah model Transformer yang dilatih untuk klasifikasi gambar. Model ini menggunakan arsitektur Transformer yang dimodifikasi untuk memproses gambar. ViT-S16 telah terbukti mencapai hasil yang sebanding dengan model convolutional neural network (CNN) terbaik pada beberapa dataset klasifikasi gambar [8].



Gambar 2. Arsitektur Vision Transformers [9]

Arsitektur Transformer sendiri adalah model deep learning yang pertama kali diusulkan pada tahun 2017 untuk menyelesaikan tugas

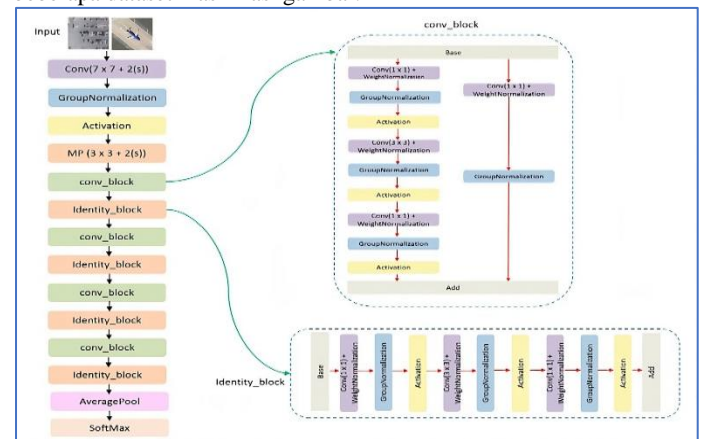
pengolahan kata. Model ini berbeda dari model tradisional seperti Recurrent Neural Networks (RNN) karena tidak menggunakan koneksi berulang [8].

ViT-S16 bekerja dengan cara memecah gambar menjadi patch-patch kecil. Patch-patch ini kemudian diubah menjadi embedding dan diproses oleh Transformer encoder. Transformer encoder menggunakan attention mechanism untuk mempelajari hubungan antar-patch dan menghasilkan representasi global gambar. Representasi ini kemudian digunakan untuk klasifikasi gambar [8].

Dalam penelitian ini model ViT-S akan digabungkan dengan model Big Transfer (BiT-M) agar model yang digabungkan unggul dalam akurasi maupun unggul dalam efisiensi. Selain itu, dengan menggabungkan model dapat menghasilkan performa klasifikasi gambar yang lebih baik daripada model individual.

#### 2) BiT-M

Big Transfer Medium (BiT-M) adalah model klasifikasi gambar yang dilatih pada dataset besar dengan jutaan gambar. Model ini menggunakan arsitektur hybrid yang menggabungkan convolutional neural network (CNN) dan Transformer [10]. BiT-M telah terbukti mencapai hasil yang sebanding dengan model CNN terbaik pada beberapa dataset klasifikasi gambar.



Gambar 3. Arsitektur Big Transfer [11]

BiT-M bekerja dengan cara memecah gambar menjadi patch-patch kecil. Patch-patch ini kemudian diubah menjadi embedding dan diproses oleh CNN dan Transformer. CNN mengekstrak fitur lokal dari gambar, sedangkan Transformer mempelajari hubungan antar patch. Hasil dari CNN dan transformer kemudian digabungkan untuk menghasilkan klasifikasi gambar [10].

Perbedaan utama antara BiT-M dan Vision Transformers adalah arsitekturnya. BiT-M menggunakan arsitektur hybrid yang menggabungkan CNN dan Transformer, sedangkan Vision Transformers menggunakan arsitektur Transformer murni. BiT-M umumnya lebih akurat daripada Vision Transformers, tetapi Vision Transformers lebih efisien.

## III. EKSPERIMEN

### A. SOFTWARE DAN ENVIRONMENT

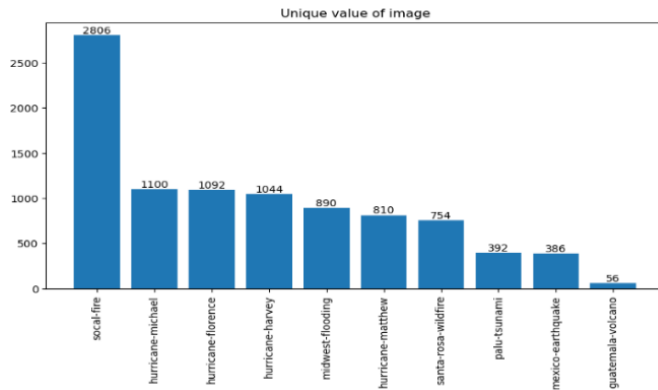
Selama penelitian, kami menggunakan perangkat lokal dengan CPU 1 GHz dan memori 268 MB, serta backend Keras berbasis TensorFlow. IDE yang digunakan adalah Jupyter Notebook dengan Anaconda dan Python 3.9.12. Spesifikasi perangkat keras meliputi Intel Core i5-1035G1, NVIDIA GeForce MX350, RAM 8 GB, dan sistem operasi Windows.

Modul Python yang digunakan mencakup OpenXLab v0.034, Keras v2.15, TensorFlow v2.15, Keras\_Preprocessing v1.1.2, TensorFlow\_Hub v0.16.1, Matplotlib, Pandas, NumPy, serta shutil dan os untuk pra-pemrosesan gambar.

### B. DATASET

#### 1) TENTANG DATASET

Untuk penelitian yang lancar, diperlukan dataset citra satelit yang relevan dan berkualitas tinggi. Dataset xBD dari Open X Lab sangat cocok untuk evaluasi kerusakan wilayah akibat bencana. Dengan lebih dari 22.000 gambar satelit dari 19 kejadian bencana berbeda (sebelum dan sesudah gempa bumi, banjir, kebakaran hutan, letusan gunung berapi, dan kecelakaan mobil), dataset ini menyediakan sumber daya berharga untuk memahami dampak bencana.



Gambar 4. Distribusi kejadian bencana

Meskipun penelitian ini hanya menggunakan 9.330 sampel gambar dari dataset serta dari 10 kejadian yang berbeda, keberagaman kejadian bencana yang tercakup memungkinkan analisis yang komprehensif. Ini membantu pengembangan model prediksi yang lebih akurat untuk menilai tingkat kerusakan dan dampak bencana di berbagai wilayah.

## 2) PROSES PENGUMPULAN DATA

Kami menggunakan Python dan library openxlab untuk melakukan crawling data dengan *access key id* dan *secret key id* dari Open X Lab. Proses dimulai dengan validasi ke library openxlab, lalu pengunduhan tiga file dataset: ``test_images_labels_target.tar.gz``, ``train_images_labels_target.tar.gz``, dan ``hold_images_labels_target.tar.gz``, total berukuran 13.06 GB. Setelah pengunduhan selesai, dataset diekstrak ke folder ``test``, ``train``, dan ``hold``.

Hasil akhir pengumpulan data menyimpulkan bahwa dataset yang berhasil diambil terdiri dari total 9.330 data citra. Dari jumlah tersebut, terdapat 5.598 data yang telah dialokasikan sebagai train data, 1.866 data sebagai test data, dan 1.866 data sebagai validation data.

## 3) PROSES PENGOLAHAN DATA

Setelah data di-*crawling* dan diekstrak, setiap folder (test, train, dan val) memiliki dua sub-folder: labels, dan images. Sub-folder labels berisi file JSON dengan label untuk setiap gambar sesuai nama file, tetapi tidak semua gambar memiliki label. Kami memformat file JSON berdasarkan akhiran label pada nama file (misalnya, ``post_disaster`` dan ``pre_disaster``) dan menyimpan data yang diformat dalam folder ``preprocessed`` di setiap folder train, val, dan test.

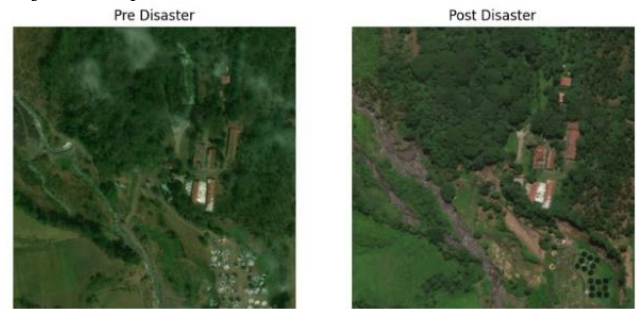
Setelah pemformatan, terdapat 4.665 data dengan label ``post_disaster`` dan 4.665 data dengan label ``pre_disaster``. Data train memiliki 2.799 data untuk masing-masing label, dan data test serta val masing-masing memiliki 933 data per label. Distribusi label yang seimbang ini menguntungkan untuk pengembangan model machine learning.

Kami menerapkan teknik augmentasi data pada dataset gambar untuk meningkatkan variasi, termasuk *rescale data* dengan skala 1/255, *horizontal flip*, *vertical flip*, pengaturan *fill mode* ke ``nearest``, dan *zoom range* sebesar 0,1.

Kemudian, kami membuat generator gambar menggunakan metode ``flow from directory`` untuk menghasilkan batch data dari direktori yang ditentukan, dengan gambar berukuran 224x224 pixel dan label biner (`post_disaster: 0, pre_disaster: 1`) dalam *batch size* 32.

Dan terakhir, kami juga membuat *image pipeline* menggunakan Tensorflow Dataset API untuk mengkonversi generator data ke format yang sesuai dengan API, memungkinkan penggunaan dalam pelatihan

model Tensorflow, dan melakukan prefetching untuk meningkatkan kinerja selama pelatihan.



Gambar 5. Sampel foto sebelum dan sesudah bencana

Gambar 5 adalah sampel gambar sebelum dan setelah terjadi bencana untuk di lokasi yang sama. Gambar bagian kanan menunjukkan kondisi sesudah bencana, sementara gambar bagian kiri menampilkan kondisi sebelum bencana.

## C. IMPLEMENTASI

### A. PENERAPAN TRANSFER LEARNING DAN PRETRAINED MODEL

Kami memanfaatkan transfer learning dengan model BiT-M (Big Transfer) dan ViT-S16 (Vision Transformers). BiT-M dirancang untuk dataset besar dan kompleks, dengan variasi dan versi yang berbeda, sedangkan ViT-S16 adalah pendekatan baru menggunakan transformers untuk pengolahan citra. Model-model pre-trained yang kami gunakan dapat dilihat pada Tabel 1.

TABEL 1.  
PERBANDINGAN MODEL

Model	Arsitektur	Deskripsi
BiT-M R50x1	ResNet-50 + ViT-S16	Model ini didasarkan pada arsitektur ResNet-50 + ViT-S16 dan telah dilatih dengan menggunakan dataset yang luas untuk mempelajari representasi gambar yang abstrak.
BiT-M R50x3	ResNet-50 + ViT-S16	Model ini juga menggunakan arsitektur ResNet-50 + ViT-S16, namun memiliki faktor skala yang tiga kali lebih besar dari BiT-MR50x1.
BiT-M R101x1	ResNet-101 + ViT-S16	Model ini dibangun berdasarkan arsitektur ResNet-101 + ViT-S16, yang lebih dalam dan kompleks daripada ResNet-50, dengan faktor skala 1x.
BiT-M R101x3	ResNet-101 + ViT-S16	Model ini adalah versi lebih besar dari BiT-MR101x1 dengan faktor skala yang lebih besar tiga kali lipatnya.

Implementasi model Big Transfer dalam deep learning melibatkan pengaturan susunan layer yang terdiri dari berbagai parameter seperti *output shape*, jumlah parameter, koneksi antar layer, dan spesifikasi parameter. Arsitektur dan rincian parameter model BiT dan ViT yang diimplementasikan dapat disusun secara rinci pada tabel 2 dan 3 berikut:

TABEL 2.  
ARSITEKTUR PEMBANGUNAN MODEL

Layer	Connected to
Input Layer (224, 224, 3)	-
Keras Layer (BiT-M)	Input Layer
Keras Layer (ViT-S16)	Input Layer
Flatten (BiT-M)	Keras Layer (BiT-M)
Flatten (ViT-S16)	Keras Layer (ViT-S16)
Dropout 0.4 (BiT-M)	Keras Layer (BiT-M)
Dropout 0.4 (ViT-S16)	Keras Layer (ViT-S16)
Concatenate Layer	Dropout (BiT-M), Dropout (ViT-S16)



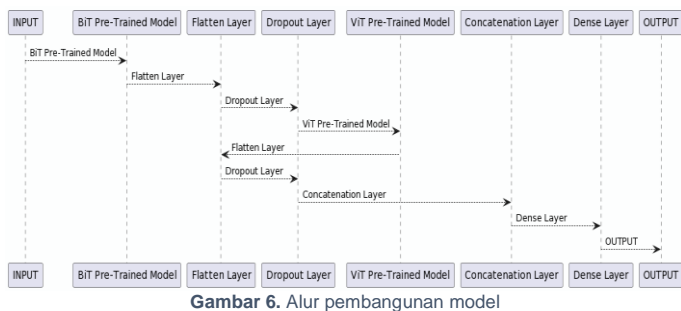
Dense Sigmoid (1)	Concatenate Layer
-------------------	-------------------

TABEL 3.  
RINCIAN PARAMETER DARI BiT-M

Model	Trainable Params	Non-Trainable Params	Size
BiT-M R50x1	3,049	45,551,016	173,77 MB
BiT-M R50x3	7,145	233,224,744	889,71 MB
BiT-M R101x1	3,049	64,543,144	246,22 MB
BiT-M R101x3	7,145	403,840,552	1,50 GB

## B. FINE TUNING

Pada tahapan ini, kami mengimplementasikan pembangunan model, kompilasi, dan pelatihan model menggunakan arsitektur BiT dan ViT yang telah dipilih sebelumnya.



Gambar 6. Alur pembangunan model

Gambar 6 menunjukkan alur pembangunan model dengan lapisan-lapisan yang berbeda. Kami menggunakan dua arsitektur, BiT (Big Transfer) dan ViT (Vision Transformers), dengan lapisan Keras Layer yang telah di-pre-trained. Hasil keluaran dari kedua model di-flatten dan dropout layer diterapkan untuk mengurangi overfitting. Kemudian, hasil keluaran dari kedua model di-concatenate dan diberikan lapisan Dense dengan aktivasi sigmoid untuk klasifikasi biner.

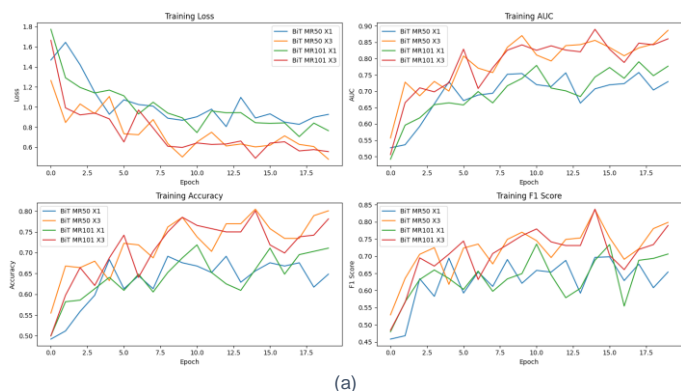
Selanjutnya, model di-compile dengan Adam optimizer dengan *learning rate* 0.001, nilai beta 1: 0.9, nilai beta 2: 0.999, dan nilai epsilon:  $1e-7$ ; *binary crossentropy* sebagai fungsi *loss*; dan metrik evaluasi AUC, *binary accuracy*, dan *f1-score*. Binary accuracy digunakan untuk mengukur seberapa baik model dalam mengklasifikasikan sampel-sampel data ke dalam dua kategori yang saling eksklusif.

Dalam pengaturan *callback*, kami menggunakan *checkpoint* untuk menyimpan model terbaik, *learning rate scheduler* untuk menyesuaikan *learning rate* dengan eksponensial decay -0.1 setiap epoch ke-5, dan *csv logger* untuk mencatat metrik pelatihan.

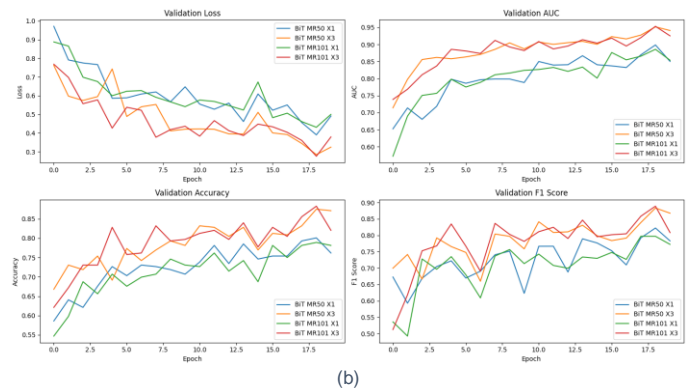
Model dilatih menggunakan *data generator* untuk pelatihan, validasi, dan pengujian. Pelatihan dilakukan sebanyak 20 epoch dengan *step* pelatihan dan validasi sebanyak 8 *step*. Selama pelatihan, digunakan tiga *callback*: *learning rate scheduler*, *checkpoint*, dan *csv logger*. Setelah pelatihan selesai, model dievaluasi pada data uji dan disimpan dalam format H5 untuk penggunaan di masa mendatang.

## IV. ANALISIS HASIL EKSPERIMEN

### A. IMPLEMENTASI HASIL FINE TUNING



(a)



(b)

Gambar 7. Grafik metrik evaluasi, (a) data training, (b) data validation

Gambar 7(a) dan 7(b) menampilkan tiga grafik evaluasi metrik (AUC, akurasi, dan F1-score), serta satu grafik untuk *loss binary crossentropy*. Kedua grafik tersebut menampilkan performa beberapa model machine learning selama 20 epoch, dengan empat grafik yang menunjukkan metrik performa: Training Loss, Training AUC, Training Accuracy, dan Training F1 Score. Keempat model yang dibandingkan adalah BiT M-R50 x1, BiT M-R50 x3, BiT M-R101 x1, dan BiT M-R101 x3. Pada grafik Training Loss, semua model awalnya memiliki loss tinggi namun mengalami penurunan seiring bertambahnya epoch. Model BiT M-R50 x3 memiliki loss awal tertinggi tetapi mengalami perbaikan signifikan. Grafik Training AUC menunjukkan peningkatan AUC seiring bertambahnya epoch, dengan model BiT M-R101 x3 secara konsisten memiliki performa lebih baik. Terjadi fluktuasi dalam akurasi pada grafik Training Accuracy, namun model BiT M-R101 x3 menunjukkan performa lebih baik secara keseluruhan. Pada grafik Training F1 Score, F1 score untuk semua model berfluktuasi tetapi secara umum mengalami peningkatan. Tidak ada model yang secara konsisten unggul, masing-masing model memiliki periode di mana performanya lebih baik daripada yang lain.

Gambar 7(a) dan 7(b) menunjukkan bahwa metrik evaluasi data validasi rata-rata selalu lebih tinggi daripada metrik training. Ini mungkin karena data validasi bersifat independen dan tidak pernah dilihat oleh model selama pelatihan, memberikan gambaran lebih objektif tentang kinerja model. Selain itu, proses regularisasi seperti dropout atau L2 membuat model lebih konservatif pada data training, menghasilkan skor lebih rendah dibandingkan dengan data validasi yang independen.

## B. ANALISIS METRIK EVALUASI

### 1) HASIL EVALUASI

TABEL 4.  
PERBANDINGAN METRIK EVALUASI PADA DATA TESTING

MODEL	LOSS	BINARY ACCURACY	AUC	F1
BIT-M R50X1	0.517452	0.764202	0.848121	0.724654
BIT-M 50X3	0.397575	0.818327	0.909107	0.813505
BIT-M R101X1	0.484615	0.767416	0.856195	0.753117
BIT-M R101X3	<b>0.391256</b>	<b>0.832262</b>	<b>0.917372</b>	<b>0.832258</b>

Tabel 4 menunjukkan hasil evaluasi model-model untuk klasifikasi kerusakan wilayah akibat bencana alam dan kecelakaan menggunakan citra satelit. Empat model dievaluasi: BiT MR50 X1, BiT MR50 X3, BiT MR101 X1, dan BiT MR101 X3. Model BiT MR101 X3 memiliki performa terbaik dengan test loss rendah, test AUC tinggi, akurasi tinggi, dan F1 score baik, menunjukkan kemampuannya dalam klasifikasi yang tepat. Sebaliknya, model BiT MR50 X1 menunjukkan performa terendah dengan test loss tinggi dan skor evaluasi rendah, kemungkinan karena arsitekturnya yang lebih sederhana atau kapasitas representasi yang kurang. Perbedaan performa antara model-model ini

dapat disebabkan oleh arsitektur, ukuran dataset, teknik pelatihan, dan hyperparameter yang digunakan. Evaluasi yang cermat dan pemilihan model yang tepat sangat penting untuk hasil optimal dalam klasifikasi citra satelit.

## 2) HASIL PREDIKSI SEJUMLAH DATASET



**Gambar 8.** Hasil prediksi model BiT-M R101x3, (a) label pre-disaster, (b) label post-disaster

Gambar 8(a) dan 8(b) menyajikan hasil prediksi dari sejumlah dataset citra menggunakan pre-trained model BiT-M R101x3 (dikarenakan BiT-M R101x3 lebih unggul dari ketiga model lainnya). Terlihat sekilas bahwa model ternyata cenderung lebih banyak menghasilkan prediksi benar pada kasus post disaster dibanding pre disaster. Ini mungkin disebabkan oleh karakteristik data yang mendasar. Jika gambar-gambar dalam dataset memiliki perbedaan yang signifikan antara kondisi sebelum dan setelah bencana, misalnya, perubahan tata guna lahan, kehancuran struktur, atau perubahan topografi yang lebih jelas pada gambar-gambar pasca bencana, model cenderung mempelajari pola-pola ini dengan lebih baik. Selain itu, label yang diberikan mungkin lebih mudah diidentifikasi pada gambar-gambar pasca bencana, karena tanda-tanda kehancuran atau dampak bencana bisa lebih jelas terlihat daripada gambar-gambar sebelum bencana yang mungkin memiliki variasi yang lebih luas dalam keadaan normal. Ini dapat menyebabkan model menjadi lebih terampil dalam mengenali kelas pasca bencana. Untuk mengatasi perbedaan ini, penyesuaian strategis dalam pengumpulan data, pemrosesan gambar, atau pendekatan pelatihan mungkin diperlukan untuk meningkatkan keseimbangan dan keadilan dalam prediksi antara kelas pre dan post disaster.

## V. KESIMPULAN

### A. PERBANDINGAN DENGAN PENELITIAN SEBELUMNYA

Penelitian ini mengimplementasikan model Vision Transformers (ViT-S16) dan Big Transfer Medium (BiT-M) untuk klasifikasi kerusakan wilayah pada citra satelit. Berbeda dengan studi sebelumnya yang menggunakan segmentasi citra, penelitian ini langsung melakukan klasifikasi, meningkatkan efisiensi dan kecepatan analisis. Hasilnya menunjukkan kemampuan yang lebih baik dalam memisahkan kelas kompleks dan klasifikasi dengan akurasi tinggi.

Pendekatan transfer learning terbukti efektif untuk klasifikasi citra satelit.

## B. KESIMPULAN PENELITIAN

Penelitian ini berhasil mengimplementasikan model BiT-M untuk klasifikasi kerusakan wilayah pada citra satelit. Model BiT-M R101x3 menunjukkan performa terbaik, dengan akurasi tinggi dalam membedakan kelas kompleks. Penggunaan transfer learning dengan model pre-trained seperti BiT-M menawarkan solusi efektif dan efisien untuk klasifikasi citra satelit.

## C. SARAN PENELITIAN SELANJUTNYA

- Penelitian Hybrid Model: Menggabungkan deep learning dengan metode konvensional atau analisis data spasial untuk meningkatkan performa klasifikasi.
- Pengembangan Model Spesifik : Mengembangkan model khusus untuk klasifikasi kerusakan wilayah pada citra satelit dengan arsitektur lebih kompleks atau teknik regularisasi canggih.
- Penerapan di Bidang Praktis : Mengimplementasikan model dalam aplikasi nyata untuk mendukung pemantauan dan mitigasi bencana alam.

## REFERENSI

- [1] iNews.id, "BNPB Sebut Ada 38 Laporan Bencana di Indonesia dalam Sepekan Terakhir," Nov. 16, 2023. [Online]. Available: <https://www.inews.id/multimedia/video/bnpb-sebut-ada-38-laporan-bencana-di-indonesia-dalam-sepekan-terakhir>. [Accessed: May 18, 2024].
- [2] A. W. Finaka, A. Fauzi, and S. Agam, "Indonesia di Lingkaran Api Pasifik," Indonesia Baik. [Online]. Available: <https://indonesiabaik.id/infografis/indonesia-di-lingkaran-api-pasifik>. [Accessed: May 18, 2024].
- [3] C. D. Pratama and S. Gischa, "Lembaga-Lembaga yang Berperan dalam Penanggulangan Bencana Alam," Kompas.com, Dec. 2, 2020. [Online]. Available: <https://www.kompas.com/skola/read/2020/12/02/155257469/lembaga-lembaga-yang-berperan-dalam-penanggulangan-bencana-alam?page=all>. [Accessed: May 18, 2024].
- [4] INSTIKI, "Mengenal Sustainable Development Goals (SDGs) atau Tujuan Pembangunan Berkelanjutan," INSTIKI, May 2, 2023. [Online]. Available: <https://instiki.ac.id/2023/05/02/mengenal-sustainable-development-goals-sdgs-atau-tujuan-pembangunan-berkelanjutan/>. [Accessed: May 18, 2024].
- [5] Chen, M., & Liu, Y. "Building Damage Detection from Satellite Imagery." Stanford University. [Online]. Available: [https://web.stanford.edu/~markcx/sample-project/IISE\\_Building\\_Damage\\_Detection\\_from\\_Satellite\\_Imagery.pdf](https://web.stanford.edu/~markcx/sample-project/IISE_Building_Damage_Detection_from_Satellite_Imagery.pdf). [Accessed: May 18, 2024].
- [6] J. Z. Xu, W. Lu, Z. Li, P. Khaitan, and V. Zaytseva, "Building Damage Detection in Satellite Imagery Using Convolutional Neural Networks," ArXiv, 2019, vol. abs/1910.06444. [Online]. Available: <https://api.semanticscholar.org/CorpusID:204575963>.
- [7] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," arXiv, 2021. [Online]. Available: <https://doi.org/10.48550/arXiv.2010.11929>.
- [8] A. Berroukham, K. Housni, and M. Lahraichi, "Vision Transformers: A Review of Architecture, Applications, and Future Directions," in IEEE International Conference on Computational Intelligence and Sustainable Technologies (CiSt), Dec. 2023, pp. 205-210. doi: 10.1109/CiSt56084.2023.10410015.
- [9] S.-H. Tsang, "Review: Vision Transformer (ViT)," Feb. 4, 2022. [Online]. Available: <https://sh-tsang.medium.com/review-vision-transformer-vit-406568603de0>. [Accessed: April 12, 2024].
- [10] A. Kolesnikov, L. Beyer, X. Zhai, J. Puigcerver, J. Yung, S. Gelly, and N. Houlsby, "Big Transfer (BiT): General Visual Representation Learning," arXiv, 2020. [Online]. Available: <https://doi.org/10.48550/arXiv.1912.11370>.
- [11] M. M. Al Rahhal, Y. Bazi, T. Abdullah, and M. L. Mekhlafi, "Deep Unsupervised Embedding for Remote Sensing Image Retrieval Using Textual Cues," Applied Sciences, vol. 10, no. 24, Dec. 2020, doi: 10.3390/app10248931.