



**OPTIMASI DAN EVALUASI MODEL CONVOLUTIONAL NEURAL  
NETWORK DAN VISION TRANSFORMER DALAM MENGATASI  
KETIDAKSEIMBANGAN DATA LONG-TAILED PADA  
KLASIFIKASI CITRA AWAN SATELIT**

**SKRIPSI**

**Oleh:**

**REVATTA MANGGALA NANDIVADHANO**

**20211000036**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS BUDDHI DHARMA  
TANGERANG**

**2025**



**OPTIMASI DAN EVALUASI MODEL CONVOLUTIONAL NEURAL  
NETWORK DAN VISION TRANSFORMER DALAM MENGATASI  
KETIDAKSEIMBANGAN DATA LONG-TAILED PADA  
KLASIFIKASI CITRA AWAN SATELIT**

**SKRIPSI**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program  
Studi Teknik Informatika Fakultas Sains dan Teknologi  
Universitas Buddhi Dharma  
Jenjang Pendidikan Strata 1**

**Oleh:**

**REVATTA MANGGALA NANDIVADHANO**

**20211000036**

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS BUDDHI DHARMA  
TANGERANG**

**2025**

## LEMBAR PERSEMBAHAN

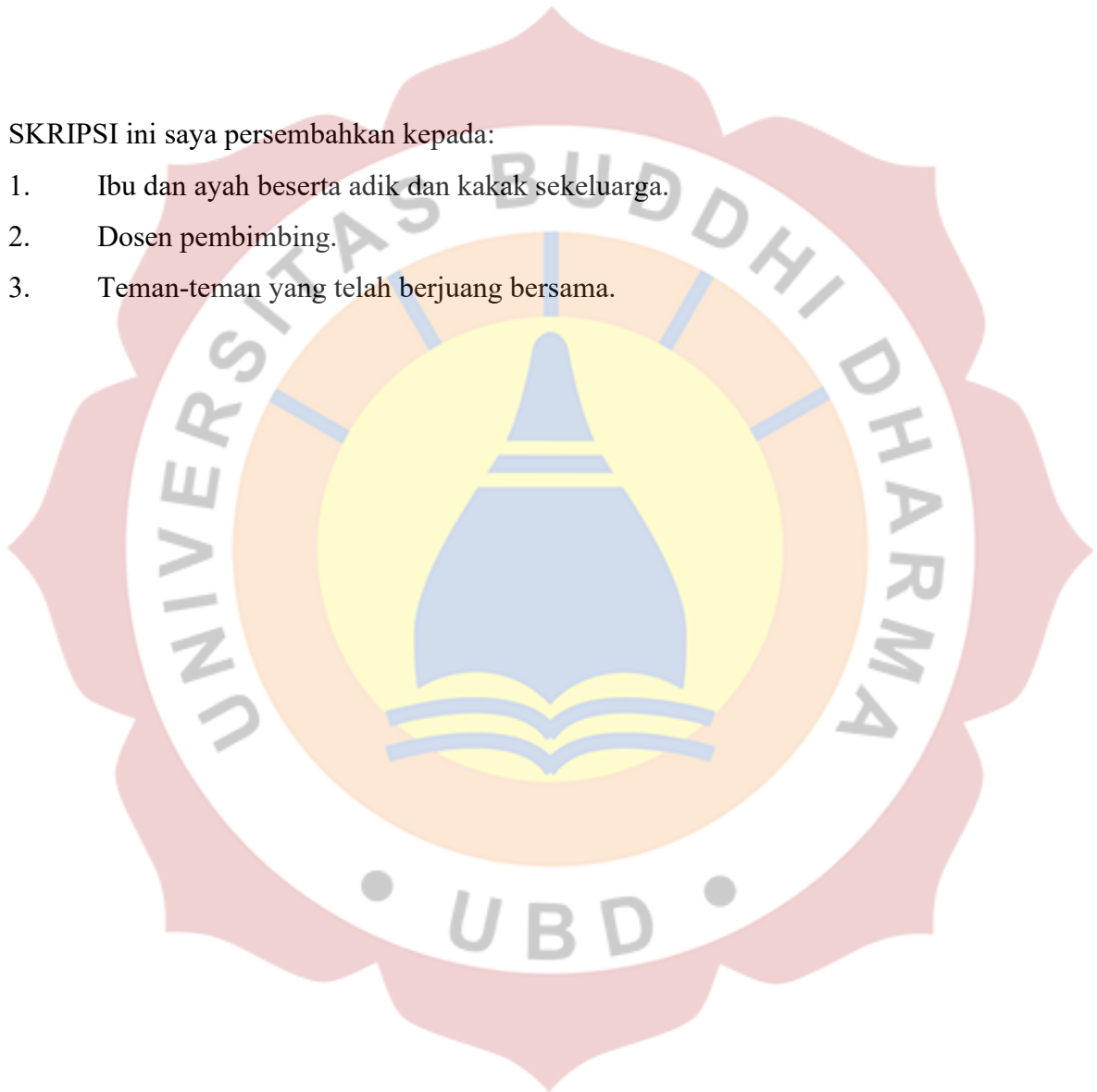
「暁の雨は、おのづから木の葉吹くあらしに似たり」

“Suara hujan pagi adalah kenangan badai di lain hari.”

-Kamo Chōmei (鴨長明)-

SKRIPSI ini saya persembahkan kepada:

1. Ibu dan ayah beserta adik dan kakak sekeluarga.
2. Dosen pembimbing.
3. Teman-teman yang telah berjuang bersama.



# UNIVERSITAS BUDDHI DHARMA

## LEMBAR PERNYATAAN KEASLIAN SKRIPSI

Yang bertanda tangan di bawah ini,

NIM : 20211000036  
Nama : Revatta Manggala Nandivadhano  
Jenjang Studi : Strata 1  
Program Studi : Teknik Informatika  
Peminatan : Data Science

Dengan ini saya menyatakan bahwa:

1. Skripsi ini adalah asli dan belum pernah diajukan untuk mendapat gelar akademik Sarjana atau kelengkapan studi, baik di Universitas Buddhi Dharma maupun di Perguruan Tinggi lainnya.
2. Skripsi ini saya buat sendiri tanpa bantuan dari pihak lain, kecuali arahan dosen pembimbing.
3. Dalam Skripsi ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dan dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan daftar pustaka.
4. Dalam Skripsi ini tidak terdapat pemalsuan (kebohongan), seperti buku, artikel, jurnal, data sekunder, pengolahan data, dan pemalsuan tanda tangan dosen atau Ketua Program Studi Universitas Buddhi Dharma yang dibuktikan dengan keasliannya.
5. Lembar pernyataan ini saya buat dengan sesungguhnya, tanpa paksaan dan apabila dikemudian hari atau pada waktu lainnya terdapat penyimpangan dan ketidak benaran dalam pernyataan ini, saya bersedia menerima sanksi akademik berupa pencabutan gelar akademik yang telah saya peroleh karena Skripsi ini serta sanksi lainnya sesuai dengan peraturan dan norma yang berlaku.

Tangerang, 4 Agustus 2025  
Yang membuat pernyataan,



Revatta Manggala  
Nandivadhano  
20211000036

# UNIVERSITAS BUDDHI DHARMA

## LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH

NIM : 20211000036  
Nama : Revatta Manggala Nandivadhano  
Jenjang Studi : Strata 1  
Program Studi : Teknik Informatika  
Peminatan : Data Science

Dengan ini menyetujui untuk memberikan ijin kepada pihak Universitas Buddhi Dharma, Hak Bebas Royalti Non – Eksklusif (*Non-exclusive Royalty-Free Right*) atas karya ilmiah kami yang berjudul: “OPTIMASI DAN EVALUASI MODEL CONVOLUTIONAL NEURAL NETWORK DAN VISION TRANSFORMER DALAM MENGATASI KETIDAKSEIMBANGAN DATA LONG-TAILED PADA KLASIFIKASI CITRA AWAN SATELIT”.

Dengan Hak Bebas Royalti Non – Eksklusif ini pihak Universitas Buddhi Dharma berhak menyimpan, mengalih-media atau format-kan, mengelolanya dalam pangkalan data (*database*), mendistribusikannya, dan menampilkan atau mempublikasikannya di *internet* atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis atau pencipta karya ilmiah tersebut.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Universitas Buddhi Dharma, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Tangerang, 4 Agustus 2025  
Yang membuat pernyataan,



Revatta Manggala  
Nandivadhano  
20211000036

**UNIVERSITAS BUDDHI DHARMA**  
**LEMBAR PENGESAHAN PEMBIMBING**

**OPTIMASI DAN EVALUASI MODEL CONVOLUTIONAL NEURAL  
NETWORK DAN VISION TRANSFORMER DALAM MENGATASI  
KETIDAKSEIMBANGAN DATA LONG-TAILED PADA  
KLASIFIKASI CITRA AWAN SATELIT**

Dibuat Oleh:

NIM : 20211000036

Nama : Revatta Manggala Nandivadhano

Telah disetujui untuk dipertahankan di hadapan Tim Penguji Ujian

Komprehensif

Program Studi Teknik Informatika

Data Science

Tahun Akademik 2024/2025

Disahkan oleh,

Tangerang, 16 Juni 2025

**Pembimbing**

**Aditiya Hermawan, M.Kom**

NUPTK. 9538766667130223

**UNIVERSITAS BUDDHI DHARMA**  
**LEMBAR PENGESAHAN SKRIPSI**  
**OPTIMASI DAN EVALUASI MODEL CONVOLUTIONAL NEURAL**  
**NETWORK DAN VISION TRANSFORMER DALAM MENGATASI**  
**KETIDAKSEIMBANGAN DATA LONG-TAILED PADA**  
**KLASIFIKASI CITRA AWAN SATELIT**

Dibuat Oleh:

NIM : 20211000036

Nama : Revatta Manggala Nandivadhano

Telah disetujui untuk dipertahankan di hadapan Tim Penguji Ujian

Komprehensif

Program Studi Teknik Informatika

Data Science

Tahun Akademik 2024/2025

Disahkan oleh,

Tangerang, 4 Agustus 2025

Dekan

Ketua Program Studi



**Dr. Yakub, M.Kom., M.M.**

NUPTK. 1836747648130172



**Hartana Wijaya, M.Kom**

NUPTK. 3844759660131192

## LEMBAR PENGESAHAN TIM PENGUJI

Nama : Revatta Manggala Nandivadhano  
NIM : 20211000036  
Fakultas : Sains dan Teknologi  
Judul Skripsi : OPTIMASI DAN EVALUASI MODEL  
CONVOLUTIONAL NEURAL NETWORK DAN VISION  
TRANSFORMER DALAM MENGATASI  
KETIDAKSEIMBANGAN DATA LONG-TAILED PADA  
KLASIFIKASI CITRA AWAN SATELIT

Dinyatakan LULUS setelah mempertahankan di depan Tim Penguji pada hari Senin, 4 Agustus 2025

Nama penguji:  
Ketua Sidang : **Edy, S.T., M.Kom.**  
NUPTK. 7560760661130203  
Penguji I : **Ramona Dyah Safitri, S.Si, M.Si.**  
NUPTK. 8652771672230312  
Penguji II : **Aditiya Hermawan, M.Kom**  
NUPTK. 9538766667130223

Tanda Tangan:



Mengetahui,

Dekan Fakultas Sains dan Teknologi



**Dr. Yakub, M.Kom., M.M.**  
NUPTK. 1836747648130172

## KATA PENGANTAR

Skripsi ini berjudul **“OPTIMASI DAN EVALUASI MODEL CONVOLUTIONAL NEURAL NETWORK DAN VISION TRANSFORMER DALAM MENGATASI KETIDAKSEIMBANGAN DATA LONG-TAILED PADA KLASIFIKASI CITRA AWAN SATELIT”**. Tujuan utama dari pembuatan Skripsi ini adalah sebagai salah satu syarat kelengkapan dalam menyelesaikan program pendidikan Strata 1 Program Studi Teknik Informatika di Universitas Buddhi Dharma. Dalam penyusunan Skripsi ini penulis banyak menerima bantuan dan dorongan baik moril maupun materiil dari berbagai pihak, maka pada kesempatan ini penulis menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Ibu Dr. Limajatini, S.E., M.M., B.K.P, sebagai Rektor Universitas Buddhi Dharma
  2. Bapak Dr. Yakub, S.Kom., M.Kom., M.M, Dekan Fakultas Sains dan Teknologi
  3. Bapak Hartana Wijaya, S.Kom., M.Kom, sebagai Ketua Program Studi Teknik Informatika
  4. Bapak Aditiya Hermawan, S.Kom., M.Kom, sebagai pembimbing yang telah membantu dan memberikan dukungan serta harapan untuk menyelesaikan penulisan Skripsi ini.
  5. Orang tua dan keluarga yang selalu memberikan dukungan baik moril dan materiil.
  6. Teman-teman yang selalu membantu dan memberikan semangat
- Serta semua pihak yang terlalu banyak untuk disebutkan satu-persatu sehingga terwujudnya penulisan ini. Penulis menyadari bahwa penulisan Skripsi ini masih belum sempurna, untuk itu penulis mohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan di masa yang akan datang.

Akhir kata semoga Skripsi ini dapat berguna bagi penulis khususnya dan bagi para pembaca yang berminat pada umumnya.

Tangerang, 13 Agustus 2025

Penulis

## ABSTRAK

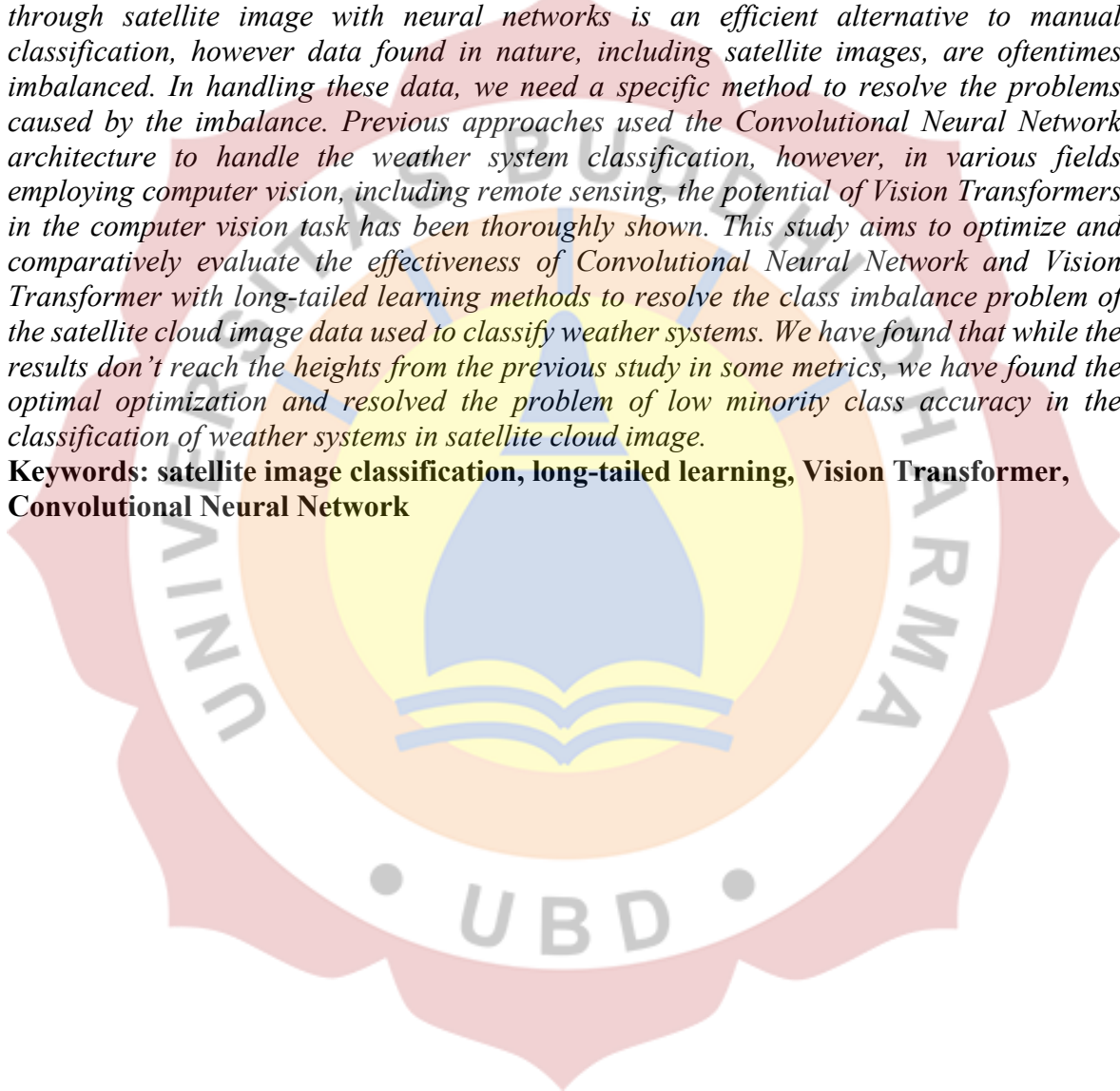
Dengan diperparahnya perubahan iklim, beserta fenomena cuaca ekstrem yang menyusul, pengawasan cuaca yang kontinual semakin diperlukan. Pengawasan aktif dari cuaca dengan memanfaatkan tenaga manusia secara manual merupakan sebuah proses yang inefisien. Klasifikasi sistem cuaca melalui citra satelit menggunakan teknologi jaringan saraf buatan dapat menjadi alternatif yang lebih efisien, akan tetapi data yang ditemukan di alam, seperti citra satelit, seringkali bersifat tidak seimbang. Dalam menangani data-data tersebut, diperlukan metode spesifik yang mengatasi permasalahan yang timbul dari sifat data tersebut. Penelitian-penelitian sebelumnya telah meninjau kelayakan algoritma *Convolutional Neural Network* dalam tugas klasifikasi sistem cuaca, akan tetapi, dalam berbagai bidang yang memanfaatkan penglihatan komputer, termasuk *remote sensing* atau penginderaan jarak jauh, telah menunjukkan bahwa algoritma *Vision Transformer* memiliki potensi. Penelitian ini bertujuan untuk melakukan optimasi dan evaluasi komparatif dari efektifitas *Convolutional Neural Network* dan *Vision Transformer* dengan menerapkan metode pembelajaran data tidak seimbang bersifat *long-tailed* pada tugas klasifikasi sistem cuaca melalui citra awan satelit. Hasil yang ditemukan menunjukkan bahwa walaupun nilai dari beberapa metrik yang diuji tidak mencapai penelitian sebelumnya, penelitian ini telah berhasil dalam melakukan optimasi dan menyelesaikan permasalahan akurasi dari kelas minoritas yang rendah dalam klasifikasi sistem cuaca melalui citra awan satelit.

**Kata kunci:** klasifikasi citra satelit, *long-tailed learning*, *Vision Transformer*, *Convolutional Neural Network*

## **ABSTRACT**

*With the worsening of climate change and subsequent severe weather phenomena, the need to incessantly monitor the weather is at an all time high. Active monitoring of weather by human personnel could be an expensive and inefficient method. Classifying weather systems through satellite image with neural networks is an efficient alternative to manual classification, however data found in nature, including satellite images, are oftentimes imbalanced. In handling these data, we need a specific method to resolve the problems caused by the imbalance. Previous approaches used the Convolutional Neural Network architecture to handle the weather system classification, however, in various fields employing computer vision, including remote sensing, the potential of Vision Transformers in the computer vision task has been thoroughly shown. This study aims to optimize and comparatively evaluate the effectiveness of Convolutional Neural Network and Vision Transformer with long-tailed learning methods to resolve the class imbalance problem of the satellite cloud image data used to classify weather systems. We have found that while the results don't reach the heights from the previous study in some metrics, we have found the optimal optimization and resolved the problem of low minority class accuracy in the classification of weather systems in satellite cloud image.*

**Keywords:** satellite image classification, long-tailed learning, Vision Transformer, Convolutional Neural Network



## DAFTAR ISI

|                                                 |      |
|-------------------------------------------------|------|
| LEMBAR JUDUL LUAR SKRIPSI .....                 | i    |
| LEMBAR JUDUL DALAM SKRIPSI .....                | ii   |
| LEMBAR PERSEMBAHAN .....                        | i    |
| LEMBAR PERNYATAAN KEASLIAN SKRIPSI .....        | ii   |
| LEMBAR PERSETUJUAN PUBLIKASI KARYA ILMIAH ..... | iii  |
| LEMBAR PENGESAHAN PEMBIMBING .....              | iv   |
| LEMBAR PENGESAHAN SKRIPSI .....                 | v    |
| LEMBAR PENGESAHAN TIM PENGUJI .....             | vi   |
| KATA PENGANTAR .....                            | vii  |
| ABSTRAK .....                                   | viii |
| <i>ABSTRACT</i> .....                           | ix   |
| DAFTAR ISI .....                                | x    |
| DAFTAR GAMBAR .....                             | xiv  |
| DAFTAR TABEL .....                              | xvi  |
| DAFTAR LAMPIRAN .....                           | xvii |
| BAB 1 BAB I PENDAHULUAN .....                   | 1    |
| 1.1 Latar Belakang .....                        | 1    |
| 1.2 Identifikasi Masalah .....                  | 4    |
| 1.3 Ruang Lingkup .....                         | 5    |
| 1.4 Tujuan dan Manfaat .....                    | 5    |
| 1.4.1 Tujuan .....                              | 5    |
| 1.4.2 Manfaat .....                             | 5    |
| 1.5 Sistematika Penulisan .....                 | 6    |
| BAB 2 BAB II LANDASAN TEORI .....               | 7    |
| 2.1 Data .....                                  | 7    |

|       |                                                 |    |
|-------|-------------------------------------------------|----|
| 2.2   | <i>Data Mining</i> .....                        | 7  |
| 2.3   | <i>Long-Tailed Data</i> .....                   | 8  |
| 2.4   | <i>CRISP-DM</i> .....                           | 8  |
| 2.5   | <i>Machine Learning</i> .....                   | 10 |
| 2.6   | <i>Deep Learning</i> .....                      | 11 |
| 2.7   | <i>Computer Vision</i> .....                    | 11 |
| 2.8   | <i>Remote Sensing</i> .....                     | 12 |
| 2.9   | <i>Klasifikasi</i> .....                        | 13 |
| 2.10  | <i>Convolutional Neural Network (CNN)</i> ..... | 14 |
| 2.11  | <i>Vision Transformer (ViT)</i> .....           | 16 |
| 2.12  | <i>Python</i> .....                             | 17 |
| 2.13  | <i>Confusion Matrix</i> .....                   | 18 |
| 2.14  | <i>Tinjauan Studi</i> .....                     | 19 |
| 2.15  | <i>Kesimpulan Tinjauan Studi</i> .....          | 28 |
| BAB 3 | <b>BAB III METODE PENELITIAN</b> .....          | 30 |
| 3.1   | <i>Kerangka Pemikiran</i> .....                 | 30 |
| 3.2   | <i>Data Understanding</i> .....                 | 31 |
| 3.3   | <i>Data Splitting</i> .....                     | 36 |
| 3.4   | <i>Data Preprocessing</i> .....                 | 37 |
| 3.4.1 | <i>Data Transformation</i> .....                | 37 |
| 3.5   | <i>Model Training</i> .....                     | 38 |
| 3.5.1 | <i>Arsitektur CLIP</i> .....                    | 38 |
| 3.5.2 | <i>Hyperparameter</i> .....                     | 39 |
| 3.5.3 | <i>Perangkat Keras</i> .....                    | 40 |
| 3.5.4 | <i>Lingkungan Pengembangan</i> .....            | 40 |
| 3.5.5 | <i>Backbone</i> .....                           | 41 |
| 3.5.6 | <i>Metode Evaluasi</i> .....                    | 41 |

|        |                                                                                                                 |    |
|--------|-----------------------------------------------------------------------------------------------------------------|----|
| BAB 4  | BAB IV HASIL DAN PEMBAHASAN.....                                                                                | 43 |
| 4.1    | Evaluasi Model .....                                                                                            | 43 |
| 4.1.1  | Evaluasi Model RemoteCLIP ViT-B-32 .....                                                                        | 43 |
| 4.1.2  | Evaluasi Model GeoRSCLIP ViT-B-32 .....                                                                         | 45 |
| 4.1.3  | Evaluasi Model RemoteCLIP ResNet-50 .....                                                                       | 46 |
| 4.1.4  | Evaluasi Model RemoteCLIP ResNet-50 Dengan CE Loss.....                                                         | 48 |
| 4.1.5  | Evaluasi Model RemoteCLIP ViT-B-32 Dengan CE Loss .....                                                         | 50 |
| 4.1.6  | Evaluasi Model RemoteCLIP ViT-B-32 Dengan CB Loss.....                                                          | 52 |
| 4.1.7  | Evaluasi Model RemoteCLIP ResNet-50 dengan CB Loss .....                                                        | 53 |
| 4.1.8  | Evaluasi Model RemoteCLIP ViT-B-32 Dengan LA Loss<br>(Setengahtr) .....                                         | 55 |
| 4.1.9  | Evaluasi Model RemoteCLIP ViT-B-32 Dengan Focal Loss .....                                                      | 57 |
| 4.1.10 | Evaluasi Model RemoteCLIP ResNet-50 Dengan LADE Loss ....                                                       | 59 |
| 4.1.11 | Evaluasi Model RemoteCLIP ViT-B-32 Tanpa Kelas LabelLess.                                                       | 61 |
| 4.1.12 | Evaluasi Model RemoteCLIP ViT-B-32 Tanpa Kelas LabelLess<br>Dengan CB Loss .....                                | 63 |
| 4.1.13 | Evaluasi Model RemoteCLIP ViT-B-32 Dengan 64 Bottleneck<br>Dimension.....                                       | 65 |
| 4.1.14 | Evaluasi Model RemoteCLIP ViT-B-32 Dengan 128 Bottleneck<br>Dimension.....                                      | 67 |
| 4.1.15 | Evaluasi Model RemoteCLIP ViT-B-32 Dengan 128 Bottleneck<br>Dimension Dan CB Loss .....                         | 69 |
| 4.1.16 | Evaluasi Model RemoteCLIP ViT-B-32 Dengan 128 Bottleneck<br>Dimension Dan CE Loss .....                         | 71 |
| 4.1.17 | Evaluasi Model RemoteCLIP ViT-B-32 Dengan 128 Bottleneck<br>Dimension Dan CB Loss Tanpa Kelas LabelLess.....    | 73 |
| 4.1.18 | Evaluasi Model RemoteCLIP ViT-B-32 Dengan 128 Bottleneck<br>Dimension Dan CB Loss Dilatih Selama 15 Epoch ..... | 75 |
| 4.2    | Komparasi Hasil Evaluasi Model .....                                                                            | 77 |

|                            |                                |    |
|----------------------------|--------------------------------|----|
| 4.3                        | Pembahasan.....                | 78 |
| 4.4                        | Implementasi Web .....         | 81 |
| BAB 5                      | BAB V SIMPULAN DAN SARAN ..... | 84 |
| 5.1                        | Kesimpulan .....               | 84 |
| 5.2                        | Saran .....                    | 85 |
| DAFTAR PUSTAKA.....        |                                | 86 |
| DAFTAR RIWAYAT HIDUP ..... |                                | 94 |



## DAFTAR GAMBAR

|                                                                                   |    |
|-----------------------------------------------------------------------------------|----|
| Gambar 2.1 Siklus CRISP-DM .....                                                  | 9  |
| Gambar 2.2 Ilustrasi Komponen <i>Remote Sensing</i> .....                         | 13 |
| Gambar 2.3 Ilustrasi Proses Konvolusi .....                                       | 15 |
| Gambar 2.4 Ilustrasi LeNet-5 Sebagai Tipe CNN Paling Awal.....                    | 16 |
| Gambar 2.5 Ilustrasi Arsitektur ViT.....                                          | 17 |
| Gambar 3.1 Kerangka Pemikiran .....                                               | 30 |
| Gambar 3.2 Gambar Sebelum <i>Resize</i> .....                                     | 37 |
| Gambar 3.3 Gambar Sesudah <i>Resize</i> .....                                     | 38 |
| Gambar 3.4 Proses <i>Contrastive Pre-training CLIP</i> .....                      | 38 |
| Gambar 3.5 Tampilan <i>Visual Studio Code</i> .....                               | 41 |
| Gambar 3.6 Ilustrasi <i>Confusion Matrix</i> dengan <i>Dataset</i> LSCIDMR-S..... | 42 |
| Gambar 4.1 <i>Confusion Matrix</i> Model 1 .....                                  | 44 |
| Gambar 4.2 <i>Reliability Diagram</i> Model 1.....                                | 44 |
| Gambar 4.3 <i>Confusion Matrix</i> Model 2 .....                                  | 46 |
| Gambar 4.4 <i>Reliability Diagram</i> Model 2.....                                | 46 |
| Gambar 4.5 <i>Confusion Matrix</i> Model 3 .....                                  | 47 |
| Gambar 4.6 <i>Reliability Diagram</i> Model 3.....                                | 48 |
| Gambar 4.7 <i>Confusion Matrix</i> Model 4 .....                                  | 49 |
| Gambar 4.8 <i>Reliability Diagram</i> Model 4.....                                | 50 |
| Gambar 4.9 <i>Confusion Matrix</i> Model 5 .....                                  | 51 |
| Gambar 4.10 <i>Reliability Diagram</i> Model 5.....                               | 51 |
| Gambar 4.11 <i>Confusion Matrix</i> Model 6.....                                  | 52 |
| Gambar 4.12 <i>Reliability Diagram</i> Model 6.....                               | 53 |
| Gambar 4.13 <i>Confusion Matrix</i> Model 7.....                                  | 54 |
| Gambar 4.14 <i>Reliability Diagram</i> Model 7.....                               | 54 |
| Gambar 4.15 <i>Confusion Matrix</i> Model 8.....                                  | 56 |
| Gambar 4.16 <i>Reliability Diagram</i> Model 8.....                               | 56 |
| Gambar 4.17 <i>Confusion Matrix</i> Model 9.....                                  | 58 |
| Gambar 4.18 <i>Reliability Diagram</i> Model 9.....                               | 58 |
| Gambar 4.19 <i>Confusion Matrix</i> Model 10.....                                 | 60 |
| Gambar 4.20 <i>Reliability Diagram</i> Model 10.....                              | 60 |

|                                                      |    |
|------------------------------------------------------|----|
| Gambar 4.21 <i>Confusion Matrix</i> Model 11.....    | 62 |
| Gambar 4.22 <i>Reliability Diagram</i> Model 11..... | 62 |
| Gambar 4.23 <i>Confusion Matrix</i> Model 12.....    | 64 |
| Gambar 4.24 <i>Reliability Diagram</i> Model 11..... | 64 |
| Gambar 4.25 <i>Confusion Matrix</i> Model 13.....    | 66 |
| Gambar 4.26 <i>Reliability Diagram</i> Model 13..... | 66 |
| Gambar 4.27 <i>Confusion Matrix</i> Model 14.....    | 68 |
| Gambar 4.28 <i>Reliability Diagram</i> Model 14..... | 68 |
| Gambar 4.29 <i>Confusion Matrix</i> Model 15.....    | 70 |
| Gambar 4.30 <i>Reliability Diagram</i> Model 15..... | 70 |
| Gambar 4.31 <i>Confusion Matrix</i> Model 16.....    | 72 |
| Gambar 4.32 <i>Reliability Diagram</i> Model 16..... | 72 |
| Gambar 4.33 <i>Confusion Matrix</i> Model 17.....    | 74 |
| Gambar 4.34 <i>Reliability Diagram</i> Model 17..... | 74 |
| Gambar 4.35 <i>Confusion Matrix</i> Model 18.....    | 76 |
| Gambar 4.36 <i>Reliability Diagram</i> Model 18..... | 76 |
| Gambar 4.37 Laman Awal.....                          | 82 |
| Gambar 4.38 Laman Prediksi.....                      | 82 |
| Gambar 4.39 Laman Penjelasan.....                    | 83 |
| Gambar 4.40 Laman Pop-up.....                        | 83 |

## DAFTAR TABEL

|                                                                          |    |
|--------------------------------------------------------------------------|----|
| Tabel 2.1 <i>Confusion Matrix</i> .....                                  | 18 |
| Tabel 3.1 Persebaran Label LSCIDMR-S .....                               | 34 |
| Tabel 3.2 Sampel Data LSCIDMR-S .....                                    | 35 |
| Tabel 3.3 Nama Kelas Dalam <i>Prompt</i> .....                           | 39 |
| Tabel 3.4 Hyperparameter Pada Pelatihan Dan Nilainya .....               | 40 |
| Tabel 3.5 Spesifikasi Perangkat Keras .....                              | 40 |
| Tabel 4.1 Evaluasi Himpunan Tes Model 1 .....                            | 43 |
| Tabel 4.2 Evaluasi Himpunan Tes Model 2 .....                            | 45 |
| Tabel 4.3 Evaluasi Himpunan Tes Model 3 .....                            | 47 |
| Tabel 4.4 Evaluasi Himpunan Tes Model 4 .....                            | 48 |
| Tabel 4.5 Evaluasi Himpunan Tes Model 5 .....                            | 50 |
| Tabel 4.6 Evaluasi Himpunan Tes Model 6 .....                            | 52 |
| Tabel 4.7 Evaluasi Himpunan Tes Model 7 .....                            | 53 |
| Tabel 4.8 Evaluasi Himpunan Tes Model 8 .....                            | 55 |
| Tabel 4.9 Evaluasi Himpunan Tes Model 9 .....                            | 57 |
| Tabel 4.10 Evaluasi Himpunan Tes Model 10 .....                          | 59 |
| Tabel 4.11 Evaluasi Himpunan Tes Model 11 .....                          | 61 |
| Tabel 4.12 Evaluasi Himpunan Tes Model 12 .....                          | 63 |
| Tabel 4.13 Evaluasi Himpunan Tes Model 13 .....                          | 65 |
| Tabel 4.14 Evaluasi Himpunan Tes Model 14 .....                          | 67 |
| Tabel 4.15 Evaluasi Himpunan Tes Model 15 .....                          | 69 |
| Tabel 4.16 Evaluasi Himpunan Tes Model 16 .....                          | 71 |
| Tabel 4.17 Evaluasi Himpunan Tes Model 17 .....                          | 73 |
| Tabel 4.18 Evaluasi Himpunan Tes Model 18 .....                          | 75 |
| Tabel 4.19 Perbandingan Hasil Model Inisial .....                        | 77 |
| Tabel 4.20 Perbandingan Hasil Model Dengan Loss Lain.....                | 77 |
| Tabel 4.21 Perbandingan Hasil Model Dilatih Tanpa Kelas LabelLess.....   | 77 |
| Tabel 4.22 Perbandingan Hasil Model Dilatih Dengan Parameter Lebih ..... | 77 |

## DAFTAR LAMPIRAN

|                       |    |
|-----------------------|----|
| Kartu Bimbingan ..... | 95 |
|-----------------------|----|



# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Cuaca merupakan sebuah aspek penting dalam kehidupan manusia. Kondisi cuaca ekstrem dapat mengakibatkan penghambatan kegiatan fisik manusia pada jangka pendek serta menstimulasi migrasi iklim (Turrisi dkk., 2021). Hal tersebut akan terus terjadi selama dampak dan intensitas dari cuaca ekstrem terus meningkat. Dampak dan intensitas dari cuaca ekstrem meningkat seiring dengan bertambah parahnya perubahan iklim (Clarke dkk., 2022). Perubahan iklim yang disebabkan oleh manusia (antropogenik) yang sekarang sedang terjadi disebabkan oleh emisi gas rumah kaca ke dalam atmosfer (Fawzy dkk., 2020). Emisi tersebut kian meningkat dengan pada tahun 2023 jumlah gas rumah kaca mencapai rekor tertinggi terbaru (UNEP, 2024). Hal tersebut menunjukkan perubahan fenomena cuaca akan terus terjadi dikarenakan kondisi iklim yang semakin berubah. Dengan perubahan iklim yang turut mengubah sifat dari fenomena cuaca, kemampuan untuk mendeteksi fenomena cuaca secara akurat dan tanpa henti semakin diperlukan. Salah satu metode untuk melakukan inferensi sistem cuaca adalah dengan menganalisis bentukan awan dari citra satelit (C. Bai dkk., 2022).

Citra awan satelit merupakan citra satelit yang diambil oleh satelit meteorologi yang berisi informasi tentang awan dan hal lainnya yang terkait. Satelit tersebut memiliki fungsi utamanya untuk melakukan pengawasan cuaca dan iklim bumi dan berperan penting dalam proses analisis cuaca dan peringatan dini (C. Bai dkk., 2022). Penggunaan citra awan satelit untuk melakukan inferensi sistem cuaca dapat menjadi sebuah metode yang efisien. Citra awan telah digunakan sebagai pengganti data deteksi kondisi cuaca yang konvensional (S. Shang dkk., 2023). Citra awan tersebut diambil dari satelit meteorologis, yang telah banyak

dimanfaatkan untuk keperluan observasi dari fenomena seperti awan di atmosfer ataupun kondisi di atas permukaan bumi (Wang dkk., 2022). Kegiatan mengekstraksi informasi dari citra awan satelit tergolong dalam kegiatan dalam *remote sensing*. *Remote sensing*, yang berarti pengindraan jarak jauh, merupakan proses observasi dari jarak jauh (Chuvieco, 2020). Dalam *remote sensing*, metode *machine learning* untuk memproses data *remote sensing* telah diterapkan oleh berbagai peneliti pada tugas-tugas *remote sensing* (Diaz-Gonzalez dkk., 2022; W. Han dkk., 2023; Li dkk., 2023; Marjani dkk., 2024).

Penggunaan *machine learning* dalam bidang meteorologi memiliki potensi untuk memproses inferensi sistem cuaca melalui citra awan satelit. *Machine learning* selain digunakan secara umum pada bidang *remote sensing* juga diterapkan secara spesifik dalam permasalahan *remote sensing* di bagian meteorologis (Tong dkk., 2023). Secara spesifik, metode *deep learning*, yang berupa sub-bidang dari *machine learning*, dapat mempelajari fitur-fitur kompleks dari awan secara adaptif (Wang dkk., 2022). Dua dari algoritma *machine learning* yang digunakan berupa *Convolutional Neural Network* (CNN) dan *Vision Transformer* (ViT). CNN adalah sebuah tipe jaringan saraf tiruan yang memiliki arsitektur yang memungkinkan operasi terhadap informasi spasial dan spektral dari input secara bersamaan (Segal-Rozenhaimer dkk., 2020). CNN merupakan standar de facto dalam tugas penglihatan komputer (Galdran dkk., 2022). CNN juga telah menunjukkan performa yang bagus dalam tugas klasifikasi citra, deteksi target, dan segmentasi semantis (Z. Zhang dkk., 2021). Selain CNN, terdapat algoritma lain yang bernama ViT yang telah banyak diperbincangkan dalam segi perbandingan keunggulan antara ViT dan CNN. ViT merupakan aplikasi dari *Transformer* pada bidang penglihatan komputer, dengan melakukan proses pemotongan citra menjadi potongan-potongan untuk di-input ke dalam sebuah *Transformer* (Dosovitskiy dkk., 2021). ViT telah didemonstrasikan sukses untuk mencapai hasil yang bagus dalam tugas penglihatan komputer, melampaui CNN (Rad, 2024). Mekanisme *self-*

*attention* yang digunakan dalam ViT dapat mempelajari hubungan kompleks antara fitur yang bertetangga serta bertetangga tak langsung (M. Kaselimi dkk., 2023). ViT juga memiliki ketahanan yang lebih baik dibandingkan oleh CNN dalam tipe serangan tertentu (Bai dkk., 2021). Sedangkan peneliti lain menemukan CNN masih memiliki keunggulan dalam tugas klasifikasi tertentu (Galdran dkk., 2022). Dari tinjauan di atas, maka perlu dipertimbangkan perbandingan kedua algoritma tersebut dalam tugas klasifikasi citra awan satelit. Inferensi sistem cuaca dengan klasifikasi citra awan satelit dengan metode *machine learning* memerlukan data dalam skala besar. Sebuah *dataset* citra awan satelit berupa Large-Scale Satellite Cloud Image Database for Meteorological Research (LSCIDMR).

LSCIDMR merupakan *dataset* yang dikembangkan khusus untuk melakukan tugas inferensi sistem cuaca dari citra awan satelit (C. Bai dkk., 2022). Akan tetapi LSCIDMR merupakan sebuah *dataset* yang memiliki permasalahan *inter-class similarity* (kemiripan antar *class* dalam data) dan *class imbalance* (ketidak seimbangan jumlah dari *data point* pada masing-masing *class*) (Yousaf dkk., 2023). Penelitian sebelumnya yang menggunakan LSCIDMR mendapatkan hasil yang memuaskan, namun juga menemukan penggunaan kelas *LableLess* yang sangat tidak seimbang jumlahnya pada *dataset* dapat menghancurkan kualitas dari model (Yousaf dkk., 2023). Penelitian lainnya yang menggunakan LSCIDMR tidak menggunakan *class* tersebut dan mendapatkan hasil yang juga memuaskan (S. Shang dkk., 2023). Akan tetapi, inklusi dari *LableLess* dapat memberikan informasi yang penting mengenai kondisi cuaca yang tidak memiliki fenomena cuaca yang dominan (C. Bai dkk., 2022), sehingga terdapat manfaat untuk melakukan klasifikasi dengan inklusi *Lableless*. Untuk mengatasi permasalahan yang ditimbulkan oleh data yang tidak seimbang, maka diperlukan metode spesifik. Salah satu metode yang dapat digunakan berupa *Long-Tail Learning with Foundation Model: Heavy Fine-Tuning Hurts* (LIFT).

LIFT merupakan sebuah metode untuk melakukan *fine-tuning* pada *foundation model* terhadap data bersifat *long-tailed* yang dirancang oleh (Shi dkk., 2024). LIFT melakukan proses *fine-tuning* secara ringan (dalam jumlah parameter yang dimodifikasi) dan terstruktur (dalam bagian parameter mana yang dimodifikasi). Metode LIFT menggunakan basis *foundation model*, yang mana berupa model yang dilatih pada data yang luas untuk kemudian dapat diadaptasikan sesuai dengan keperluan tugas kehiliran (*downstream tasks*) (Bommasani dkk., 2021). Di mana LIFT menggunakan *foundation model* yang dilatih pada data umum (terutama *Vision-Language Model* CLIP), adaptasi model umum atau *General Vision-Language Model* pada tugas kehiliran dengan domain yang spesifik, seperti tugas dalam *remote sensing*, dapat dilakukan dengan menggunakan model yang dilatih dengan data pada domain spesifik (Z. Zhang dkk., 2024). Dalam *remote sensing*, terdapat beberapa model domain spesifik seperti RemoteCLIP (Liu dkk., 2024) dan GeoRSCLIP (Z. Zhang dkk., 2024), keduanya merupakan model yang didasarkan pada model *Contrastive Image-Language* yang memiliki arsitektur *encoder* ganda untuk gambar dan teks (Radford dkk., 2021).

Berdasarkan latar belakang yang telah diuraikan, penelitian ini akan berfokus pada perbandingan performa CNN dan ViT pada penggunaan *dataset* LSCIDMR untuk melakukan inferensi sistem cuaca.

## 1.2 Identifikasi Masalah

Identifikasi masalah yang didapat dari latar belakang di atas berupa:

- a. Dibutuhkan klasifikasi citra satelit dengan pendekatan *deep learning* untuk memungkinkan inferensi sistem cuaca secara otomatis pada data *long-tailed*.
- b. Perlu diteliti kecakapan dari performa model ViT pada data citra satelit *long-tailed* dibandingkan dengan model CNN.

- c. Perlu dilakukan adaptasi dari kedua model tersebut untuk menangani data citra satelit *long-tailed*.

### 1.3 Ruang Lingkup

Ruang lingkup dari penelitian ini berupa:

- a. Data sekunder berjudul “Large-Scale Satellite Cloud Image Database for Meteorological Research” oleh (C. Bai et al., 2022) yang diambil dari: <https://github.com/Zjut-MultimediaPlus/LSCIDMR>
- b. *Dataset* terdiri atas 104390 potongan gambar yang beresolusi 256\*256 dengan atribut tunggal yang berjumlah 11 kelas.
- c. Pengembangan aplikasi berbasis web untuk menerapkan model terbaik.

### 1.4 Tujuan dan Manfaat

#### 1.4.1 Tujuan

Tujuan yang ingin dicapai oleh penelitian ini berupa:

- a. Melakukan pemodelan CNN dan ViT untuk klasifikasi citra satelit dengan data *long-tailed*.
- b. Melakukan perbandingan performa antara model CNN dan ViT dalam klasifikasi citra satelit.
- c. Melakukan optimasi terhadap CNN dan ViT pada data yang bersifat *long-tailed*.

#### 1.4.2 Manfaat

Manfaat yang dapat diambil dari penelitian ini berupa:

- a. Memudahkan proses inferensi cuaca dari citra satelit.
- b. Menemukan keunggulan antara model ViT dengan CNN pada klasifikasi sistem cuaca dari citra satelit.
- c. Menemukan metode optimasi yang optimal untuk kedua model tersebut untuk data yang tidak seimbang.

## **1.5 Sistematika Penulisan**

Karya tulis ini disusun secara sistematis dalam lima bab utama untuk memudahkan pembaca memahami dan menelaah isi penelitian. Berikut adalah gambaran setiap bab:

### **BAB I PENDAHULUAN**

Bab ini mencakup penjelasan mengenai latar belakang penelitian, identifikasi masalah, ruang lingkup penelitian, tujuan dan manfaat penelitian, serta sistematika penulisan. Bab ini bertujuan memberikan gambaran umum mengenai isi karya tulis.

### **BAB II LANDASAN TEORI**

Bab ini menguraikan teori-teori yang relevan dengan penelitian, meliputi konsep dasar, prinsip umum, dan teori khusus yang berkaitan dengan algoritma atau pendekatan yang digunakan. Selain itu, tinjauan terhadap studi sebelumnya juga disajikan untuk mendukung penelitian.

### **BAB III METODE PENELITIAN**

Bab ini menjelaskan metode yang digunakan dalam penelitian.

### **BAB IV HASIL DAN PEMBAHASAN**

Bab ini memuat hasil penelitian dan evaluasi model yang digunakan.

### **BAB V SIMPULAN DAN SARAN**

Bab terakhir menyajikan kesimpulan dari penelitian yang telah dilakukan, merangkum temuan utama, dan memberikan rekomendasi untuk pengembangan lebih lanjut atau penelitian lanjutan di masa depan.

## BAB II

### LANDASAN TEORI

#### 2.1 Data

Data merupakan bentuk dari pengukuran dan observasi yang biasanya terkuantifikasi, data diperoleh biasanya dalam jangka dilaksanakannya sebuah penelitian. Data merupakan cerminan fakta dari sebuah fenomena, dan dapat dibagi menjadi dua tipe penting, berupa data skor, di mana data dikumpulkan dari kasus-kasus untuk menunjukkan kuantifikasi dari sebuah karakteristik, dan data nominal, di mana data diambil dari pengkategorian kasus ke dalam representasi numerik (Vogt & Johnson, 2015). Dalam bidang ilmu komputer, data didefinisikan sebagai bentuk informasi yang di mana sebuah program komputer beroperasi, dengan karakteristik data sebagai bentuk yang terorganisasi, terstruktur, berulang, dan sering kali terkompresi (Butterfield & Ngondi, 2016).

#### 2.2 Data Mining

*Data Mining* merupakan sebuah proses yang dilakukan untuk mengekstraksi secara otomatis informasi-informasi tersembunyi yang dapat berguna dari kumpulan-kumpulan data, hal ini dilakukan untuk menemukan pola-pola yang baru dan berguna yang selama ini masih tersembunyi (G. Zhou, 2022). Data Mining dalam prosesnya untuk memecahkan permasalahan dan menemukan pengetahuan yang baru memiliki lima tugas utama, yang berupa (Limbong, 2020): Klastering, berguna untuk melakukan pengelompokan dari data yang memiliki atribut atau karakteristik yang mirip; Klasifikasi, berguna untuk melakukan identifikasi atau pembedaan data ke dalam kelompok yang sudah ditentukan; Asosiasi, berguna untuk menunjukkan hubungan antar variabel dalam data, baik secara positif atau negatif; Estimasi, berguna untuk menentukan sebuah parameter yang tidak diketahui; Prediksi, berguna untuk meramalkan tingkat perubahan dari tahapan kejadian kontinu yang belum pernah terjadi. *Data Mining* telah digunakan secara luas dalam berbagai bidang.

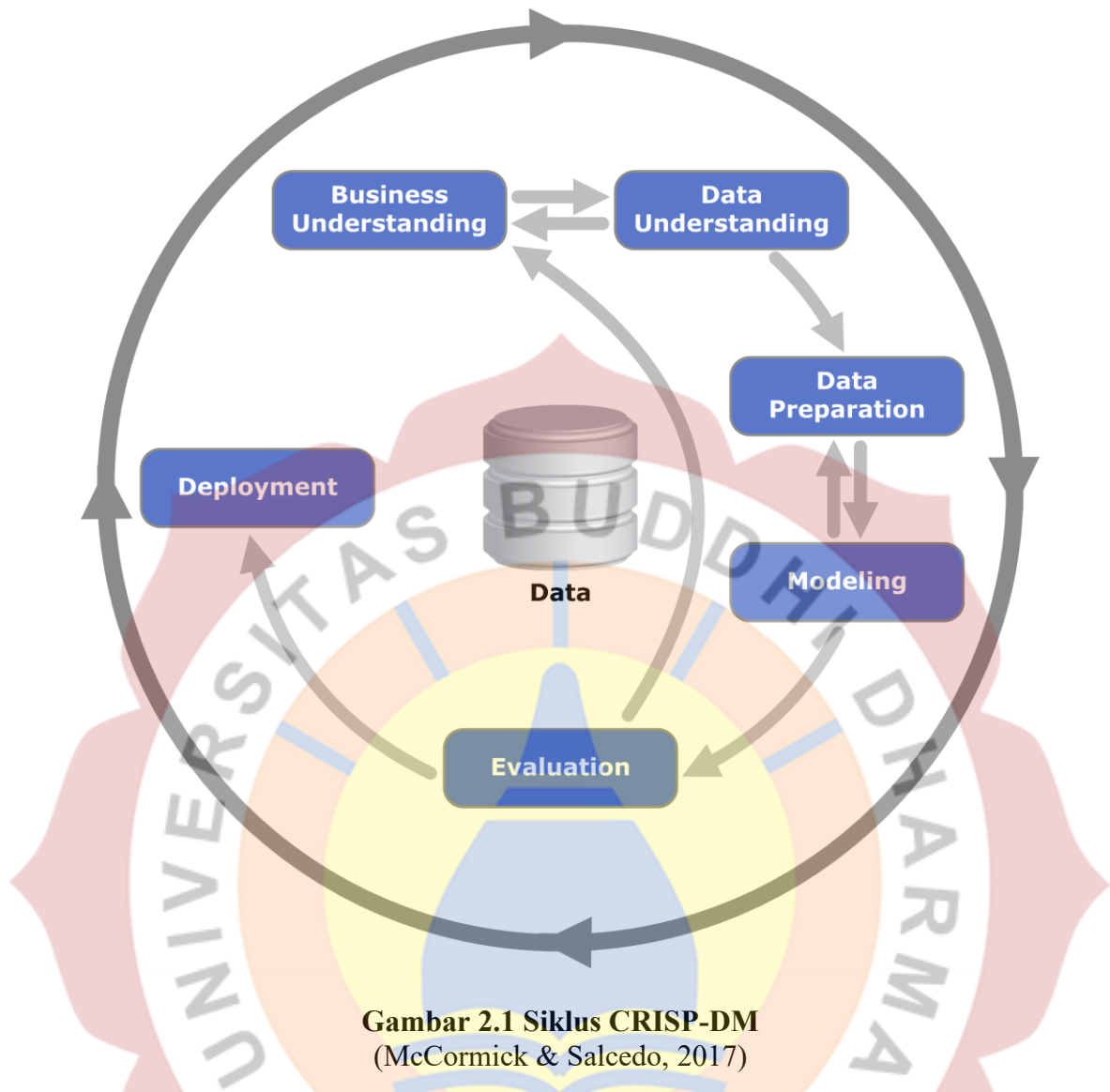
Metode-metode yang digunakan dalam *Data Mining* contohnya berupa (Butterfield & Ngondi, 2016): jaringan saraf buatan, induksi mesin, analisis pengelompokan, dan rangkuman data.

### 2.3 *Long-Tailed Data*

*Long-Tailed Data* merupakan sebuah tipe data di mana dalam data tersebut terdapat kelas-kelas yang memiliki jumlah sampel yang sedikit, serta sebaliknya, memiliki kelas-kelas dengan jumlah sampel yang masif (Kang dkk., 2021; Menon dkk., 2021; Y. Zhang dkk., 2023). Jenis data ini merupakan data yang sering muncul dalam pengaplikasian di dunia natural, yang juga muncul dalam data yang digunakan untuk klasifikasi serta data dalam tugas penglihatan komputer (Kang dkk., 2021; Menon dkk., 2021). Ketidakseimbangan dari komposisi sampel dalam kelas-kelas yang berada di dalam data menyebabkan pemodelan dalam *machine learning* menjadi sebuah tugas yang sulit dilakukan (Y. Zhang dkk., 2023). Dikarenakan ketidakseimbangan kelas di dalam data, model yang dilatih dengan *Long-Tailed Data* akan cenderung kepada kelas *head* (kelas dengan jumlah sampel banyak) dan akan memiliki performa buruk dalam kelas *tail* (kelas dengan jumlah sampel sedikit) (Y. Zhang dkk., 2022).

### 2.4 *CRISP-DM*

CRISP-DM atau *Cross-Industry Standard Process for Data Mining* merupakan sebuah *framework* siklus hidup dari proyek-proyek *data science* (Saltz, 2021).



CRISP-DM terdiri atas enam langkah iteratif yang terdiri atas (Schröer dkk., 2021):

- Business Understanding*, tahapan *Business Understanding* merupakan tahapan yang dilakukan untuk menentukan tujuan dan persyaratan dari proyek, situasi bisnis perlu dipertimbangkan untuk mendapatkan gambaran dari sumber daya yang tersedia dan sumber daya yang diperlukan. Dalam tahap ini tipe dari Data Mining dan kriteria kesuksesan seperti akurasi atau presisi perlu ditentukan.
- Data Understanding*, dalam tahapan ini dilakukan pengumpulan serta peninjauan dari data yang akan digunakan. Dilakukan deskripsi serta pengecekan mutu dari data dengan analisis statistik dan penentuan dari atribut data.

- c. *Data Preparation*, dalam tahapan ini data akan diseleksi dengan mendefinisikan kriteria inklusi dan eksklusi data yang akan digunakan. Data dengan kualitas rendah dapat ditangani dengan membersihkan data tersebut. Tergantung dari pemodelan yang akan dilakukan selanjutnya, dapat juga diperlukan penambahan data baru yang berupa derivasi dari data awal.
- d. *Modelling*, pada tahapan pemodelan dilakukan pemilihan teknik pemodelan yang akan digunakan, pembuatan model percobaan, hingga pembuatan model itu sendiri. Tipe model yang digunakan tergantung dari permasalahan yang telah dirumuskan pada fase *Business Understanding* dan tipe data yang dipahami dari tahapan Data Understanding. Diperlukan alasan yang kuat terhadap pemilihan model yang akan digunakan tersebut.
- e. *Evaluation*, dalam tahapan evaluasi, hasil atau performa dari model akan diperiksa terhadap permasalahan yang dirumuskan pada *Business Understanding*. Dalam tahapan ini dilakukan interpretasi atas hasil dan rumusan atas langkah selanjutnya yang perlu diambil.
- f. *Deployment*, tahapan ini dapat berupa laporan akhir atau penjelasan komponen, panduan pengguna akan dirumuskan pada tahapan ini, yang berisi penjelasan terhadap tahapan Deployment. Tahapan tersebut berupa: perencanaan, pengawasan, dan perawatan.

## 2.5 Machine Learning

Secara luas, *Machine Learning* merupakan sebuah istilah yang merujuk kepada proses yang dilakukan untuk menerapkan sebuah model prediktif pada sekumpulan data, hal ini dilakukan dengan tujuan untuk melakukan imitasi terhadap kemampuan manusia untuk mengenali pola, dengan memanfaatkan komputasi sebagai metode yang lebih objektif (Greener dkk., 2022). *Machine Learning* merupakan sebuah teknik yang meningkatkan

performa dari sebuah sistem dengan cara membuat sistem belajar dari pengalaman lamanya, dalam konteks, pengalaman berarti data dan pembelajaran dilakukan melalui metode komputasi atau algoritma yang bertujuan menghasilkan model yang bisa melakukan prediksi dari hasil pembelajaran tersebut (Z.-H. Zhou, 2021).

## **2.6 Deep Learning**

*Deep Learning* merupakan sebuah sub-bagian dari *Machine Learning* yang terinspirasi secara struktural oleh otak manusia, yang memungkinkan model-model *machine learning* untuk belajar dari data tidak terstruktur dan tidak berlabel (Sharifani & Amini, 2023). *Machine Learning* merupakan konsep yang berdasar atas model komputasi yang bernama jaringan saraf tiruan, dengan model ini dikembangkan pada awalnya untuk meniru mekanisme pembelajaran dan pengolahan informasi di dalam otak manusia (Bishop & Bishop, 2024). Berkembangnya ketersediaan data dan daya komputasi telah memberikan peneliti kesempatan untuk dapat mengembangkan algoritma-algoritma dan model-model yang lebih canggih, yang secara kausal memberikan terobosan dalam berbagai bidang industri (Sharifani & Amini, 2023).

## **2.7 Computer Vision**

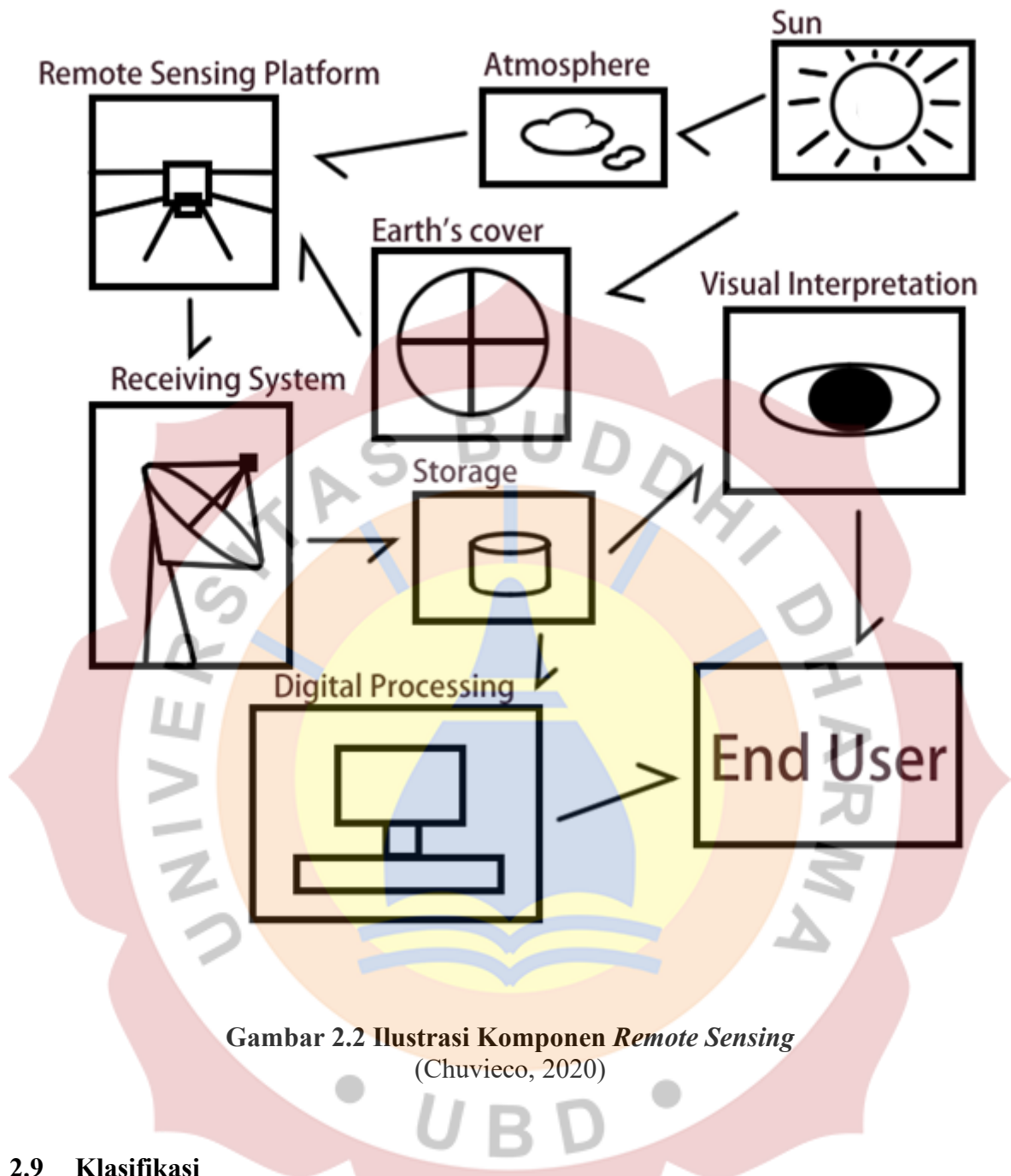
*Computer Vision* merupakan sebuah bidang dalam ilmu komputer dan teknik yang bertujuan untuk membuat komputer yang dapat melihat dunia di luarnya, serta memahami kejadian yang direfleksikan dari penglihatan tersebut (Ikeuchi, 2021). *Computer Vision* dilakukan untuk mencoba mendeskripsikan dunia yang dilihat oleh manusia melalui satu atau lebih citra dan dengan sekaligus melakukan rekonstruksi terhadap properti dari deskripsi tersebut (Szeliski, 2022). *Computer Vision* merupakan salah satu bidang yang sering kali disalah-persepsikan sebagai bidang yang mudah dikarenakan trivialitas dari penggunaan penglihatan oleh manusia (Ikeuchi, 2021), kesalahpahaman ini telah ada

semenjak perkembangan awal dari kecerdasan buatan di mana paham bahwa bagian kognitif kecerdasan jauh lebih sulit dibandingkan bagian perseptual (Szeliski, 2022).

## 2.8 Remote Sensing

*Remote sensing* merupakan istilah yang digunakan untuk mendeskripsikan kegiatan observasi dari Bumi dari jarak jauh, yang secara luas menginklusi kegiatan pengambilan data, penyimpanan, hingga interpretasi di mana data diubah menjadi informasi yang bermanfaat bagi pengguna seperti peneliti, pengolah lahan, petani, dan nelayan (Chuvieco, 2020). Dalam *remote sensing*, tipe data yang digunakan berupa citra dan pencitraan dilakukan oleh platform satelit atau pesawat terbang, dengan platform satelit dibagi kembali menjadi dua, satelit geostasioner yang biasa digunakan untuk analisis cuaca serta iklim dan satelit orbit rendah yang biasa digunakan untuk observasi permukaan daratan serta lautan (Richards & Richards, 2022). *Remote sensing* memiliki enam komponen (Chuvieco, 2020):

- a. Sumber energi, yang memancarkan radiasi elektromagnetik yang berinteraksi dengan sensor.
- b. Permukaan bumi, yang terdiri atas daratan, vegetasi, perairan, bebatuan, hamparan salju, es, dan struktur buatan manusia.
- c. Sensor dan platform, yang merupakan instrumen pengukuran yang merekam energi terpancar dari permukaan bumi.
- d. Sistem penerima, yang berada di permukaan bumi untuk menerima data rekaman dari platform dan sensor.
- e. Analis, yang berupa pelaku analisis yang mengubah data citra menjadi informasi yang bermakna dengan teknik visual dan atau digital.
- f. Komunitas pengguna, yang memanfaatkan informasi yang diekstrak dari data tersebut untuk diterapkan dalam berbagai macam aplikasi.



## 2.9 Klasifikasi

Klasifikasi merupakan sebuah tugas Data Mining yang digunakan untuk memprediksi label atau kelas dari sebuah titik data yang tidak atau belum memiliki label (Zaki dkk., 2020). Secara formal dapat didefinisikan dengan pengklasifikasi sebagai fungsi  $M$  yang memprediksi kelas label  $\hat{y}$  untuk input  $x$  yang didefinisikan dalam persamaan berikut:

$$\hat{y} = M(x) \quad (1)$$

Di mana  $\hat{y} \in \{c_1, c_2, \dots, c_k\}$  sebagai kelas label yang terprediksi.

Keterangan:

$\hat{y}$  = Himpunan kelas label terklasifikasi

$M$  = Model peng-klasifikasi

$x$  = Input

$c_i$  = Kelas label terklasifikasi

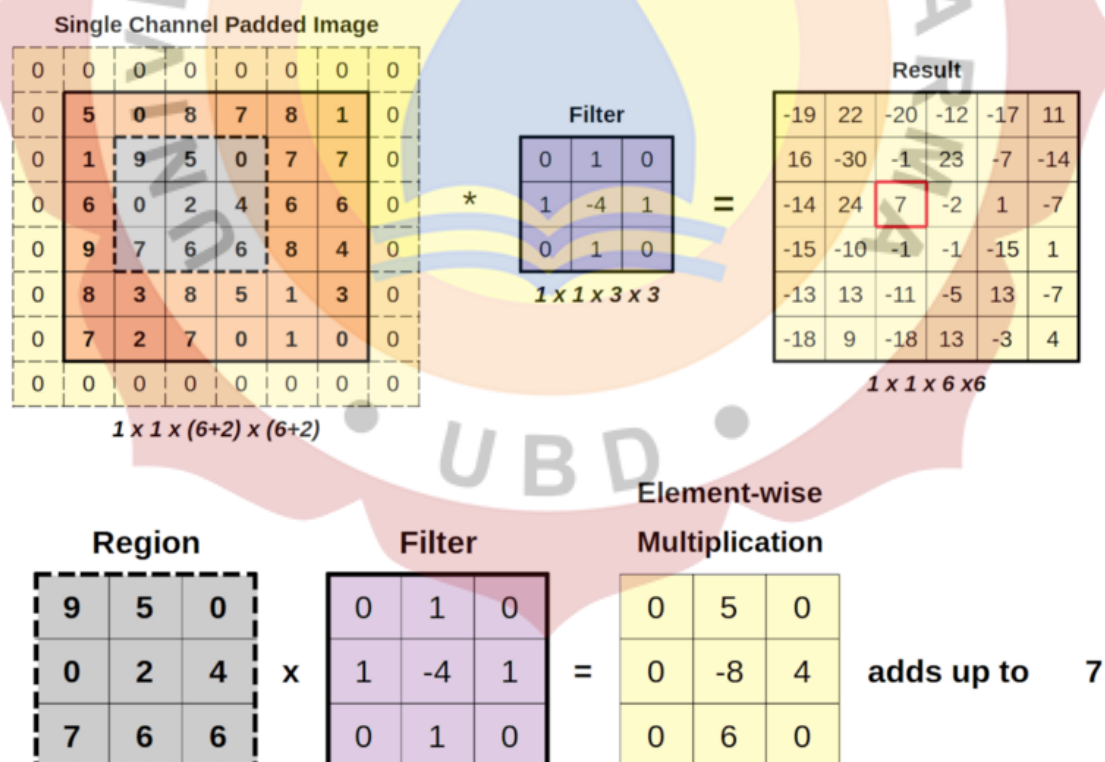
Klasifikasi merupakan salah satu proses yang termasuk dalam *supervised learning* (di mana model diberikan arahan untuk melakukan klasifikasi sesuai dengan data sebelumnya yang memiliki kelas label), di mana proses klasifikasi dimulai dengan tahapan pelatihan dengan data terlabel tersebut untuk mendapat fungsi yang bisa mengkategorisasikan input berdasarkan label dari data yang digunakan dalam pelatihan (J. Han dkk., 2022).

## 2.10 *Convolutional Neural Network (CNN)*

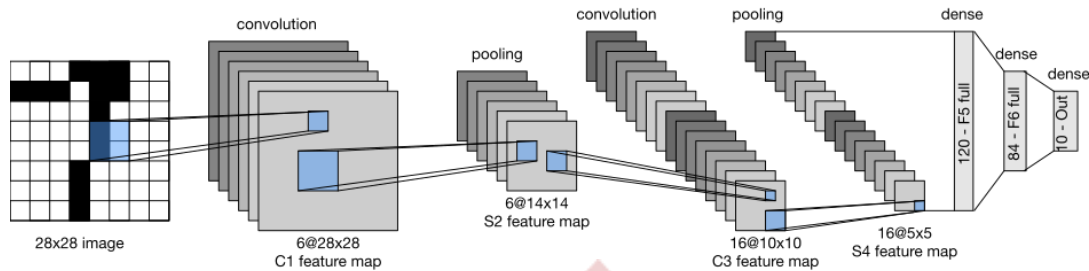
CNN merupakan sebuah tipe jaringan saraf buatan yang digunakan dalam *Computer Vision* untuk klasifikasi gambar dan deteksi objek (Aggarwal, 2018). CNN sama seperti algoritma *machine learning* lainnya di mana *deployment* dilakukan dalam dua tahap, pelatihan dan inferensi, dengan pada tahapan pelatihan diterapkan algoritma *back-propagation* yang bertugas untuk memperbarui parameter dalam CNN untuk memperbaiki performa model, pada tahapan inferensi kelas label data baru ditentukan dalam propagasi *feed-forward* yang berarti mengikuti arah dalam jaringan yang searah (Hassaballah & Awad, 2020). CNN memiliki struktur yang terdiri atas (Krichen, 2023):

- a. *Convolutional layer*, yang merupakan komponen utama CNN yang melakukan operasi konvolusi terhadap input dengan filter-filter yang berupa matriks bergeser yang menghitung *dot product* antara filter dengan input pada setiap iterasinya.

- b. *Pooling layer*, yang digunakan untuk mengurangi dimensi spasial dari *feature map* yang didapat sebagai hasil dari layer konvolusi. Layer ini digunakan untuk mengurangi kompleksitas komputasi dari jaringan.
- c. *Activation layer*, pada layer ini diaplikasikan fungsi aktivasi ke output layer sebelumnya.
- d. *Batch normalisation layer*, layer yang berfungsi menormalisasikan output dari layer sebelumnya.
- e. *Dropout layer*, layer yang menjatuhkan sebagian neuron dari layer sebelumnya, dengan tujuan mencegah *overfitting*.
- f. *Fully connected layer*, layer yang berupa jaringan saraf tiruan tradisional yang menghubungkan setiap neuron dari layer sebelumnya ke setiap neuron pada layer sekarang. Biasa digunakan di akhir untuk mengeluarkan output.



**Gambar 2.3 Ilustrasi Proses Konvolusi**  
(Godoy, 2021)

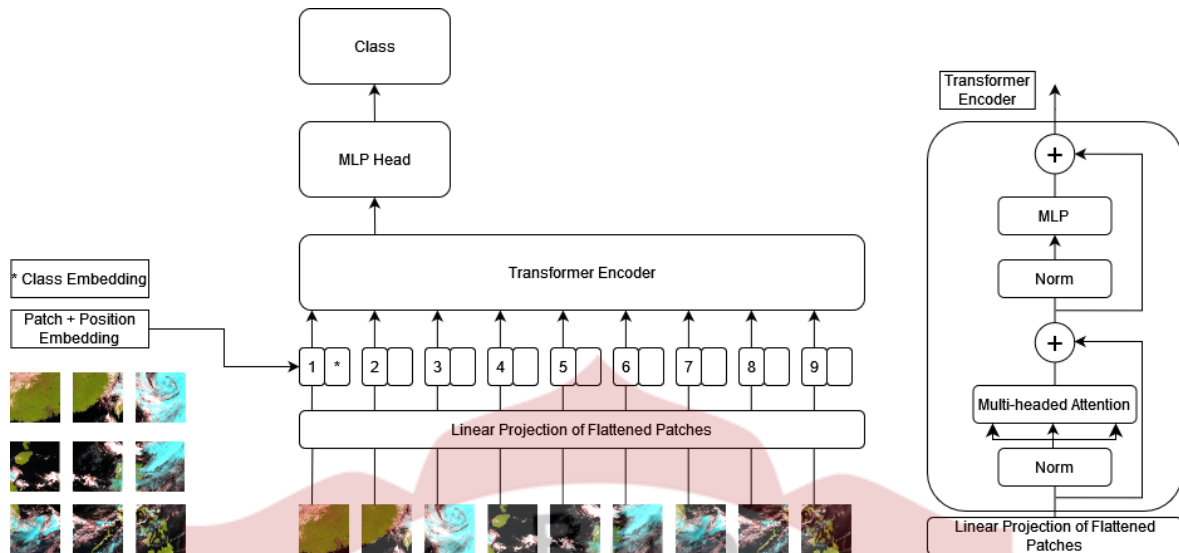


**Gambar 2.4 Ilustrasi LeNet-5 Sebagai Tipe CNN Paling Awal**  
(A. Zhang dkk., 2023)

### 2.11 Vision Transformer (ViT)

ViT merupakan adaptasi dari *Transformer* yang merupakan sebuah arsitektur model *machine learning* de facto standar pada tugas pemrosesan bahasa alami untuk tugas *Computer Vision* (Dosovitskiy dkk., 2021). ViT bekerja dengan cara memecah sebuah gambar menjadi beberapa *patch* yang kemudian dapat diubah menjadi vektor beserta dengan informasi posisional dari *patch* tersebut untuk diolah oleh *Transformer* (Dosovitskiy dkk., 2021). Arsitektur *Transformer*, dan dengan ekstensi *Vision Transformer*, memiliki dua ide utama yang berpengaruh pada perkembangannya, yang berupa (Khan dkk., 2022):

- Self-Attention*, yang berfungsi untuk melakukan estimasi terhadap relevansi dari satu objek kepada objek lainnya. Sebuah layer *Self-Attention* akan mengubah seluruh komponen dari sebuah deret dengan mengagregasikan informasi global dari deret input total.
- Pre-Training*, *Transformer* yang berdasarkan *Self-Attention* biasanya melakukan pelatihan dalam dua tahap, tahapan *Pre-Training* yang dilakukan pada *dataset* dengan skala besar yang biasanya kombinasi dari beberapa *dataset* baik secara *supervised* ataupun *self-supervised*. Kemudian, *weight* yang didapat dari proses tersebut akan diadaptasikan kepada tugas selanjutnya yang menggunakan *dataset* yang berskala sedang.



**Gambar 2.5 Ilustrasi Arsitektur ViT**  
(Dosovitskiy dkk., 2021)

## 2.12 Python

Python merupakan bahasa pemrograman tipe *scripting language* (bahasa pemrograman yang bisa digunakan untuk menulis program kecil dengan singkat) yang merupakan salah satu bahasa terpenting dalam lingkup *data science*, *machine learning*, dan pengembangan umum dalam industri maupun akademik (McKinney, 2022). Python digunakan untuk banyak kegunaan, seperti pengembangan game, aplikasi, penyelesaian masalah bisnis, dan juga digunakan secara ekstensif dalam sektor ilmiah untuk riset akademis dan terapan (Matthes, 2023). Terdapat beberapa alasan terhadap kepopuleran python, yang termasuk (Chityala & Pudipeddi, 2020):

- Tersedia bebas biaya.
- Tersedia dalam semua sistem operasi yang populer seperti Windows, Mac, dan Linux.
- Tidak memerlukan kompilasi dikarenakan Python merupakan bahasa pemrograman terinterpretasi.
- Memberikan kemampuan untuk memprogram dengan lebih cepat.
- Memiliki sintaks yang simpel dan mudah dibaca dan di-*debug*.

- f. Terdapat banyak modul dalam Python yang memudahkan pelaksanaan tugas-tugas seperti komputasi saintifik, visualisasi data, dan lainnya.
- g. Program Python *multi-platform*.
- h. Merupakan bahasa pemrograman dinamis yang memudahkan penggunaan bagi orang-orang yang kurang memiliki pengalaman menulis kode.
- i. Memiliki komunitas dan pengembang yang terdedikasi untuk menjaga kebaruan dari Python.

### 2.13 Confusion Matrix

*Confusion Matrix* merupakan sebuah metode yang digunakan untuk menganalisa performa dari pengklasifikasi, dengan merepresentasikan prediksi yang salah dalam bidang dua dimensi (Heydarian dkk., 2022). Bentuk *Confusion Matrix* yang paling simpel terdiri atas matriks dua kali dua yang merepresentasikan empat kualitas prediksi yang berupa True Positive (TP), False Positive (FP), (True Negative), dan (False Negative) (Krstinić dkk., 2020).

**Tabel 2.1 Confusion Matrix**

| Prediksi\Ground Truth | True | False |
|-----------------------|------|-------|
| True                  | TP   | FP    |
| False                 | FN   | TN    |

*Confusion Matrix* yang terdiri dari  $n$  baris dan  $n$  kolom dapat digunakan untuk menganalisa tingkat benar dan salah dari tiap tipe kelas label yang diprediksi (C. Bai dkk., 2022). Nilai yang didapat dari *Confusion Matrix* akan digunakan untuk mencari nilai *precision*, *recall*, *accuracy*, dan nilai *f1*. Di mana *precision* sebagai rasio TP terhadap prediksi, *recall* sebagai rasio TP terhadap *ground truth*, *accuracy* sebagai rasio TP dan TN terhadap *ground truth*, dan *f1* sebagai persamaan (2). Sebagai variasi dari nilai *f1*, *macro f1* mempertimbangkan semua kelas dengan berat yang sama, dengan menghitung *macro-*

*precision* dan *macro-recall* yang dihitung dari rata-rata dari tiap-tiap kelas (Grandini dkk., 2020). *Macro f1* dihitung dengan persamaan (3).

$$f1 = \frac{2PR}{P + R} \quad (2)$$

Keterangan:

$P$  = *Precision*

$R$  = *Recall*

$$Macro\ f1 = \frac{1}{n} \sum_{x=1}^n f1_x = \frac{1}{n} \sum_{x=1}^n \frac{2P_x R_x}{P_x + R_x} \quad (3)$$

Keterangan:

$n$  = Jumlah kelas label

$P_x$  = *Precision* dari kelas label  $x$

$R_x$  = *Recall* kelas label  $x$

$f1_x$  = *f1* dari kelas label  $x$

## 2.14 Tinjauan Studi

Dalam bidang *remote sensing*, pendekatan dengan *machine learning* terutamanya *deep learning* merupakan salah satu topik utama dalam analisis data yang diperoleh dari sumber seperti satelit ataupun *unmanned aerial vehicles* (UAT). Dengan berkembangnya *deep learning* pada *remote sensing*, kebutuhan untuk data yang ekstensif sesuai dengan berbagai tugas spesifik juga meningkat. Dalam kebutuhan data *remote sensing* untuk proses *deep learning*, banyak dari data *remote sensing* yang memiliki distribusi *long-tailed* (Yang dkk., 2023). Dalam *deep learning*, data yang bersifat *long-tailed* memerlukan penanganan khusus untuk dapat dimanfaatkan secara efektif. Pada bidang *remote sensing* terdapat literatur yang membahas pengaplikasian metode-metode *deep learning* seperti CNN dan ViT.

Dalam bidang data untuk keperluan *deep learning* pada *remote sensing*, penelitian yang berjudul “*LSCIDMR: Large-Scale Satellite Cloud Image Database for Meteorological*

*Research*” yang dilakukan pada tahun 2022 (C. Bai dkk., 2022) bertujuan untuk membuat *database* publik pertama dari citra awan satelit yang diperuntukan bagi penelitian meteorologis dengan nama LSCIDMR. LSCIDMR terdiri atas 104.390 citra beresolusi tinggi dengan 11 kelas (1 kelas label kosong) dalam bentuk anotasi label tunggal (LSCIDMR-S) dan 10 kelas label ganda (LSCIDMR-M). Jumlah label pada LSCIDMR-S mencapai jumlah 40.625 label dan jumlah label pada LSCIDMR-M mencapai jumlah 414.221 label. Dilakukan juga pengujian terhadap kedua bentuk data tersebut dengan beberapa model CNN yang terdiri atas AlexNet, VGGNet-19, ResNet101, dan EfficientNet. Ditemukan bahwa pada LSCIDMR-S dengan 11 kelas dan 10 kelas EfficientNet memiliki hasil yang unggul, sedangkan pada LSCIDMR-M, EfficientNet dan AlexNet memiliki hasil yang unggul.

Dari *dataset* yang telah dijabarkan di atas, terdapat pula penelitian yang telah menerapkan metode-metode optimasi pada metode *deep learning* untuk *dataset* tersebut. Penelitian yang berjudul “*Faster and Lighter Meteorological Satellite Image Classification by a Lightweight Channel-Dilation-Concatenation Net*” yang dilakukan pada tahun 2023 (S. Shang dkk., 2023) merancang sebuah CNN yang bernama Lightweight Channel-Dilation-Concatenation Net (CDC-net) yang didesain untuk melakukan klasifikasi *realtime* dari citra awal satelit. CDC-net menggunakan teknik *depth-wise convolution* dan operasi *FeatureCopy* yang memungkinkan jaringan untuk mengekstraksi fitur dimensi tinggi sembari mengurangi jumlah kedalaman dan parameter dalam jaringan, yang kemudian berdampak pada berkurangnya waktu inferensi. CDC-net juga mengganti layer *pooling* dari CNN yang biasa menggunakan layer *max-pooling* dengan layer *local importance-based pooling* (LIP) yang turut membantu menurunkan waktu inferensi dan meningkatkan akurasi. CDC-net diuji pada *dataset* LSCIDMR-S 10 kelas dan LSCIDMR-M, dengan hasil yang memuaskan. CDC-net mencapai akurasi 3-6% lebih tinggi dengan model ringan lainnya pada LSCIDMR-S dan

juga mencapai hasil terbaik pada LSCIDMR-M sekaligus sebagai model yang melakukan inferensi paling cepat dengan jumlah parameter paling rendah.

Terdapat pula peneliti lain yang melakukan optimasi dalam sisi yang berbeda, di mana akurasi klasifikasi pada LSCIDMR-S diutamakan. Penelitian yang berjudul “*Satellite Imagery-Based Cloud Classification Using Deep Learning*” yang dilakukan pada tahun 2023 (Yousaf dkk., 2023) merancang jaringan saraf buatan bernama SnapResNet yang didesain untuk melakukan klasifikasi secara cepat dan tepat dari citra awan satelit. SnapResNet merupakan jaringan *snapshot-based residual* yang terdiri atas *fully connected layers* (FC-1024), *batch normalisation* (BN), regularisasi L2, *layer dropout*, *layer dense*, dan augmentasi data. Arsitektur yang dimodifikasi didasarkan pada ResNet, modifikasi dapat mengatasi permasalahan *inter-class similarity* pada *dataset* yang digunakan (LSCIDMR-S 10 kelas). Model yang dikembangkan menggapai performa yang jauh lebih baik dibandingkan model lainnya seperti AlexNet, VGG-19, ResNet101, dan EfficientNet dan mencapai akurasi 97.25%. Akan tetapi, sebelumnya dilakukan juga percobaan pada LSCIDMR-S 11 kelas dan hasil yang dicapai tidak memuaskan sehingga tidak di-inklusikan hasilnya dalam penelitian.

Penelitian-penelitian di atas membahas optimasi metode *deep learning* yang didasarkan pada CNN, akan tetapi terdapat pula arsitektur model selain CNN yang dapat diterapkan untuk analisis *remote sensing*, seperti ViT yang merupakan *Transformer* yang diadaptasi untuk tugas *computer vision*. Arsitektur *Transformer* dapat memiliki keunggulan atas CNN. Penelitian yang berjudul “*Are Transformers More Robust Than CNNs?*” yang dilakukan pada tahun 2021 (Bai dkk., 2021) yang bertujuan untuk menemukan apakah *Transformer* lebih tahan terhadap serangan (*robust*) dibandingkan CNN. Sebelumnya, terdapat penelitian yang mendapatkan hasil bahwa *Transformer* memiliki ketahanan yang lebih dibandingkan CNN, akan tetapi peneliti menganggap penelitian yang dilakukan tidak

adil dan belum komprehensif. Pada penelitian ini ditemukan bahwa CNN dapat mencapai ketahanan yang setara dengan *Transformer* pada serangan *adversarial* (serangan sengaja untuk mencapai hasil terburuk) apabila CNN mengadopsi resep pelatihan *Transformer*, menunjukkan ketahanan *Transformer* pada serangan *adversarial* bukan sifat intrinsik *Transformer*. Dalam serangan *out-of-distribution sample*, peneliti menemukan *pre-training* pada *dataset* yang besar tidak diperlukan untuk *Transformer* mengungguli CNN. Keunggulan *Transformer* pada serangan *out-of-distribution sample* dinilai merupakan sifat dari arsitektur *Transformer* setelah diuji menyamakan pengoptimasi, *learning rate scheduler*, dan strategi augmentasi pada CNN dengan *Transformer*.

CNN dan ViT dalam mengatasi permasalahan yang ditimbulkan oleh *dataset* bersifat *long-tailed* memerlukan metode-metode spesifik. Penelitian yang berjudul “*Long-Tail Learning with Foundation Model: Heavy Fine-Tuning Hurts*” yang dilakukan pada tahun 2024 (Shi dkk., 2024) merupakan penelitian yang berfokus pada metode *fine-tuning* atau adaptasi sebuah model untuk tugas kehilangan terutama *foundation model* pada data yang bersifat *long-tailed*. Penelitian ini menunjukkan bahwa pada data *long-tailed*, metode *fine-tuning* secara berat (dalam jumlah *parameter*) dapat menghasilkan penurunan performa pada kelas-kelas *tail*. Penelitian ini kemudian mempersembahkan metode untuk melakukan *fine-tuning* pada *foundation model* secara ringan yang disebut LIFT. Hasil dari eksperimen menunjukkan bahwa performa prediktif model meningkat ditambah dengan berkurangnya waktu pelatihan dan jumlah parameter yang dilatih.

ViT dalam bidang *remote sensing* telah dimanfaatkan dalam tugas pengindraan dengan jarak relatif dekat dengan menggunakan UAV. Penelitian yang berjudul “*Transformer Neural Network for Weed and Crop Classification of High Resolution UAV Images*” yang dilakukan pada tahun 2022 (Reedha dkk., 2022) untuk mengatasi permasalahan pengawasan bibit dan hama nabati pada bidang agrikultur dengan melakukan deteksi hama nabati melalui

metode *deep learning*. Metode *deep learning* akan tetapi menghadapi permasalahan seperti keperluan komputasi yang tinggi, *dataset* dengan label yang berjumlah banyak, dan kesulitan dalam membedakan bibit dan hama nabati dengan atribut yang mirip. Peneliti kemudian menguji performa ViT dalam tugas klasifikasi hama nabati melalui citra yang diambil oleh *Unmanned Aerial Vehicles* atau UAV. Citra yang diambil dijadikan *dataset* untuk dilakukan pelatihan pada ViT. Dikarenakan kebutuhan *dataset* yang besar pada pelatihan ViT, dilakukan augmentasi data dan teknik *transfer learning*. Hasil yang didapat menunjukkan walau dengan *dataset* yang jumlahnya kecil, ViT dapat mencapai hasil yang melampaui model *state-of-the-art* seperti EfficientNet dan ResNet, menyiratkan kemampuan ViT pada kegunaan analisa *Remote Sensing* di bidang agrikultur.

ViT semakin disorot sebagai alternatif dari CNN dalam penginderaan melalui citra satelit. Penelitian yang berjudul “*A Vision Transformer Model for Convolution-Free Multilabel Classification of Satellite Imagery in Deforestation Monitoring*” yang dilakukan pada tahun 2023 (M. Kaselimi dkk., 2023) mencoba mengatasi permasalahan pengawasan deforestasi dengan merumuskan problema klasifikasi multi-label untuk mengawasi penggunaan lahan melalui citra satelit, kerentanan deforestasi maka dinilai dari tipe lahan yang dikembangkan pada area yang berada di sekitar pinggiran hutan. Penelitian merancang sebuah model yang terinspirasi dari ViT yang bernama ForestViT untuk melakukan tugas klasifikasi multi-label. Hasil dari percobaan menemukan bahwa ForestViT bekerja dengan baik pada tugas klasifikasi multi-label, terlebih lagi pada *dataset* yang tidak seimbang. ForestViT melampaui model yang berbasis konvolusi seperti ResNet, VGG, DenseNet, dan MobileNet, dengan kemampuan terbaik dalam mengklasifikasikan kelas minoritas dalam *dataset* tidak seimbang.

Seiring berkembangnya pemanfaatan *deep learning* pada *remote sensing*, tipe model yang digunakan juga semakin beragam, seperti *foundation model*, yang juga digunakan

dalam metode untuk menangani sifat *long-tailed* dari data. Penelitian yang berjudul “*RemoteCLIP: A Vision Language Foundation Model for Remote Sensing*” yang dilakukan pada tahun 2024 (Liu dkk., 2024) yang bertujuan untuk mengatasi kekurangan dari tipe-tipe *foundation model* pada bidang *remote sensing* yang umumnya menggunakan metode-metode yang memerlukan data terannotasi untuk diadaptasi ke tugas kehiliran dengan mendesain *foundation model* bertipe *vision-language* pertama pada bidang *remote sensing*. Para peneliti memecahkan permasalahan kekurangan data untuk *pre-training* dengan mengkonversikan data terannotasi secara heterogen dari berbagai format ke dalam data *image-caption*. RemoteCLIP merupakan model hasil dari adaptasi model CLIP dasar dengan data *remote sensing*. RemoteCLIP memilih tiga arsitektur sebagai *vision encoder*, dari model skala kecil ResNet-50, skala sedang ViT-B-32, dan skala besar ViT-L-14. Hasil dari eksperimen menunjukkan bahwa RemoteCLIP unggul dibandingkan *foundation model* dasar CLIP dalam berbagai tugas dan menetapkan SOTA dalam berbagai *benchmark*.

Pengembangan *foundation model* untuk domain *remote sensing* juga dilakukan oleh peneliti lainnya. Penelitian yang berjudul “*RS5M and GeoRSCLIP: A Large Scale Vision-Language Dataset and A Large Vision-Language Model for Remote Sensing*” yang dilakukan pada tahun 2024 (Z. Zhang dkk., 2024) membahas tantangan dalam adaptasi *General Vision Language Model* (GVLM) pada tugas kehiliran yang bersifat spesifik pada sebuah domain. Penelitian ini memperkenalkan konsep *Domain pre-trained Vision Language Model* (DVLM) yang menjembatani GVLM dengan tugas kehiliran yang spesifik pada domain. Penelitian ini kemudian merancang *image-text dataset* pada domain *remote sensing* sejumlah lima juta gambar dengan deskripsinya dalam Bahasa Inggris menggunakan *caption* dari VLM lain, dataset bernama RS5M. Dari dataset ini dilatih model-model CLIP untuk mendapatkan DVLM dengan berbagai metode *Parameter-Efficient Fine-Tuning* dengan hasil berupa GeoRSCLIP. Hasil dari eksperimen menunjukkan bahwa *fine-tuning*

pada dataset RS5M menunjukkan hasil efektif pada berbagai tugas dibandingkan model basis atau model SOTA.

Penelitian lain yang memanfaatkan CNN dalam *remote sensing* juga menerapkan model dalam tugas klasifikasi yang granular. Penelitian yang berjudul “*CNN Algorithm for Roof Detection and Material Classification in Satellite Images*” yang dilakukan pada tahun 2021 (Kim dkk., 2021) bertujuan untuk merancang algoritma untuk mendeteksi lokasi bangunan dan klasifikasi material atap dari bangunan tersebut dari citra satelit dengan menggunakan CNN. Tujuannya adalah untuk menilai ketahanan dari bangunan tersebut dalam kondisi bencana. Material atap dibedakan menjadi dua, yang berupa atap jadi dan atap tidak jadi. Sebelum identifikasi atap dilakukan, ditentukan dulu lokasi dari bangunan dengan memanfaatkan informasi tiga dimensi dari citra satelit dengan menggunakan peta kedalaman, atau dengan menggunakan informasi dari warna dan noise yang dikontraskan dengan struktur yang mengelilingi bangunan seperti jalanan dan lain sebagainya. CNN yang dirancang terdiri dari 43 layer menunjukkan performa yang unggul dibandingkan dengan CNN GoogleNet.

CNN dalam *remote sensing* tidak terbatas hanya dalam tugas klasifikasi tapi juga segmentasi. Penelitian yang berjudul “*Rich CNN Features for Water-Body Segmentation from Very High Resolution Aerial and Satellite Imagery*” yang dilakukan pada tahun 2021 (Z. Zhang dkk., 2021) bertujuan untuk mengatasi permasalahan segmentasi perairan (*water body*) dari pencitraan *Remote Sensing* dengan resolusi sangat tinggi. Identifikasi sulit dilakukan dikarenakan faktor-faktor seperti vegetasi akuatik, perbedaan warna dari perairan, endapan di sekitar pinggiran perairan, dan bayangan dari tanaman sekitar perairan. Dirumuskan permasalahan untuk diselesaikan dengan modul *multi-feature extraction and combination* (MEC). Modul tersebut terdiri atas tiga submodul yang berfokus pada korelasi *channel* dan spasial pada skala yang berbeda-beda, yang memungkinkan representasi fitur

yang lebih *rich* dengan menangkap hubungan-hubungan yang lokal, besar, dan *inter-channel*. Terlebih lagi, dirancang juga modul *multi-scale prediction fusion* untuk memprediksi kontur perairan dengan lebih baik. Hasil dari pengujian menunjukkan metode yang dirancang mencapai hasil *state-of-the-art* dalam performa segmentasi, dan *robust* dalam skenario citra perairan yang sulit diekstraksi.

Dalam *remote sensing*, jenis citra yang digunakan tidak terbatas pada spektrum kasat mata, alat pengindra seperti satelit dapat memiliki sensor yang jauh melampaui jangkauan tangkapan radiasi elektromagnetik mata manusia. Citra yang ditangkap disebut citra multispektral, dan analisis terhadap citra multispektral juga tidak asing dari pendekatan *deep learning*. Penelitian yang berjudul “*Vision Transformer for Multispectral Satellite Imagery: Advancing Landcover Classification*” yang dilakukan pada tahun 2024 (Rad, 2024) bertujuan untuk mengatasi klasifikasi tutupan lahan (*land cover*) yang krusial untuk memahami perubahan iklim dan dampaknya pada ekosistem dan populasi manusia. Metode yang ada memiliki hambatan dikarenakan keterbatasan data dan kesulitan dalam menangkap hubungan spasial-spektral. Untuk mengatasi permasalahan tersebut, dirancang model ViT yang memanfaatkan mekanisme *self-attention* untuk menangkap dependensi jauh dan global, yang memberikan alternatif menjanjikan dibandingkan metode tradisional berbasis CNN. ViT yang dirancang, secara khusus diperuntukan menangani *dataset* multispektral yang memiliki besar terbatas, dan melakukan identifikasi melalui 44 kelas. ViT dirancang di atas desain arsitektur U-shaped yang digabungkan dengan blok dari *Transformer Sub-Pixel*, dengan hasil sebuah *Transformer* yang dikhususkan untuk analisis citra multispektral. Berdasarkan hasil eksperimen, *Transformer Sub-Pixel* yang dirancang mendapat hasil yang baik dalam akurasi dan efisiensi memodelkan hubungan antara pola spasial dan spektral.

CNN dalam *remote sensing* juga diterapkan dalam otomasi deteksi yang memiliki resiko kerugian atas mis-deteksi tinggi. Penelitian yang berjudul “*Region-Based CNN for*

*Anomaly Detection in PV Power Plants Using Aerial Imagery*” yang dilakukan pada tahun 2022 (Vlaminck dkk., 2022) merancang sistem untuk mengawasi pembangkit listrik tenaga surya melalui pencitraan *drone* dengan CNN. Dikarenakan inspeksi manual secara reguler intensif terhadap sumber daya manusia, biaya, dan memiliki resiko bahaya maka deteksi anomali dilakukan dengan *region-proposal* CNN atau R-CNN. Alasan digunakan R-CNN adalah dikarenakan dalam sebuah panel surya, dapat terdapat lebih dari satu *hotspot* atau anomali. Hasil dari percobaan yang diuji pada *dataset* dari enam buah lokasi pembangkit listrik yang berbeda spesifikasinya menunjukkan nilai *true positive* mencapai 90% dan *false positive* 2-3%, dengan jumlah *false positive* yang paling rendah dibandingkan metode *state-of-the-art* lainnya, yang memungkinkan inspeksi pembangkit listrik yang lebih efisien.

ViT merupakan adaptasi dari *Transformer* untuk keperluan penglihatan komputer, selain ViT vanilla, terdapat juga variasi arsitektur *Transformer* untuk penglihatan komputer, yang juga diterapkan dalam *remote sensing*, secara spesifik sebagai pengenalan awan dari citra satelit. Penelitian yang berjudul “*UATNet: U-Shape Attention-Based Transformer Net for Meteorological Satellite Cloud Recognition*” yang dilakukan pada tahun 2022 (Wang dkk., 2022) yang bertujuan untuk *menangani permasalahan* rekognisi awan dari citra awan satelit yang esensial untuk meningkatkan akurasi dari ramalan cuaca. Ekstraksi tipe awan dari citra satelit dinilai lebih sulit dikarenakan luasnya area tutupan awan yang banyaknya fitur permukaan, mempersulit pengenalan bentuk, besar, tekstur, dan bayangan melalui metode *deep learning* tradisional. Dirancang sebuah *dataset* yang mengandung sembilan tipe awan dan satu langit cerah yang bernama *China region meteorological satellite cloud image dataset* (CRMSCD) bersamaan dengan sebuah model jaringan saraf tiruan yang didasarkan pada arsitektur *Transformer* U-shaped yang mengintegrasikan informasi spasial dan *multi-channel* dengan menggunakan block *self-attention* dari *Swin Transformer* yang dimodifikasi. Blok tersebut memungkinkan pembelajaran hierarkikal dan koneksi jarak jauh

antar fitur dari awan. Dirancang juga *Channel Cross Fusion with Transformer*, untuk memandu *multi-scale channel fusion*, serta modul *Attention-based Squeeze and Excitation*, yang membantu menggabungkan *multi-scale channel information* dengan *decoder features*. Model yang dirancang mencapai hasil pada CRMSCD mendapatkan performa yang mendemonstrasikan efektivitas dari model.

Walau beberapa penelitian menunjukkan keunggulan ViT terhadap CNN, akan tetap terdapat juga kasus-kasus di mana CNN masih memegang posisi sebagai model terbaik untuk sebuah tugas, salah satunya dalam penelitian yang melibatkan estimasi intensitas badai siklon tropis dari citra satelit. Penelitian yang berjudul "*Estimation of Tropical Cyclone Intensity via Deep Learning Techniques from Satellite Cloud Images*" yang dilakukan pada tahun 2023 (Tong dkk., 2023) bertujuan untuk menyelesaikan permasalahan pada estimasi intensitas badai siklon tropis (TC) menggunakan citra awan satelit. Penelitian ini menggunakan Deep CNN atau DCNN dan ViT. Penelitian menerapkan dua strategi, strategi satu tahapan yang langsung menggunakan DCNN atau ViT untuk mengestimasi intensitas TC, dan strategi dua tahapan di mana dilakukan klasifikasi citra awan satelit ke kelompok intensitas dengan DCNN atau ViT dan melakukan transfer learning untuk kemudian diestimasi intensitasnya dengan DCNN atau ViT. Berdasarkan hasil penelitian ditemukan bahwa strategi pertama dengan DCNN mencapai hasil yang terbaik mengalahkan kombinasi strategi dan model lainnya. Akan tetapi ditemukan juga gabungan dari kedua strategi memberikan hasil yang melampaui strategi satu tahapan dengan DCNN, dengan kombinasi yang berupa: strategi satu tahapan dengan DCNN, strategi dua tahapan klasifikasi-estimasi ViT-DCNN.

## **2.15 Kesimpulan Tinjauan Studi**

Dari tinjauan studi di atas didapat bahwa kedua model *deep learning* CNN dan ViT merupakan model yang sering digunakan dalam tugas *computer vision* dalam bidang *remote*

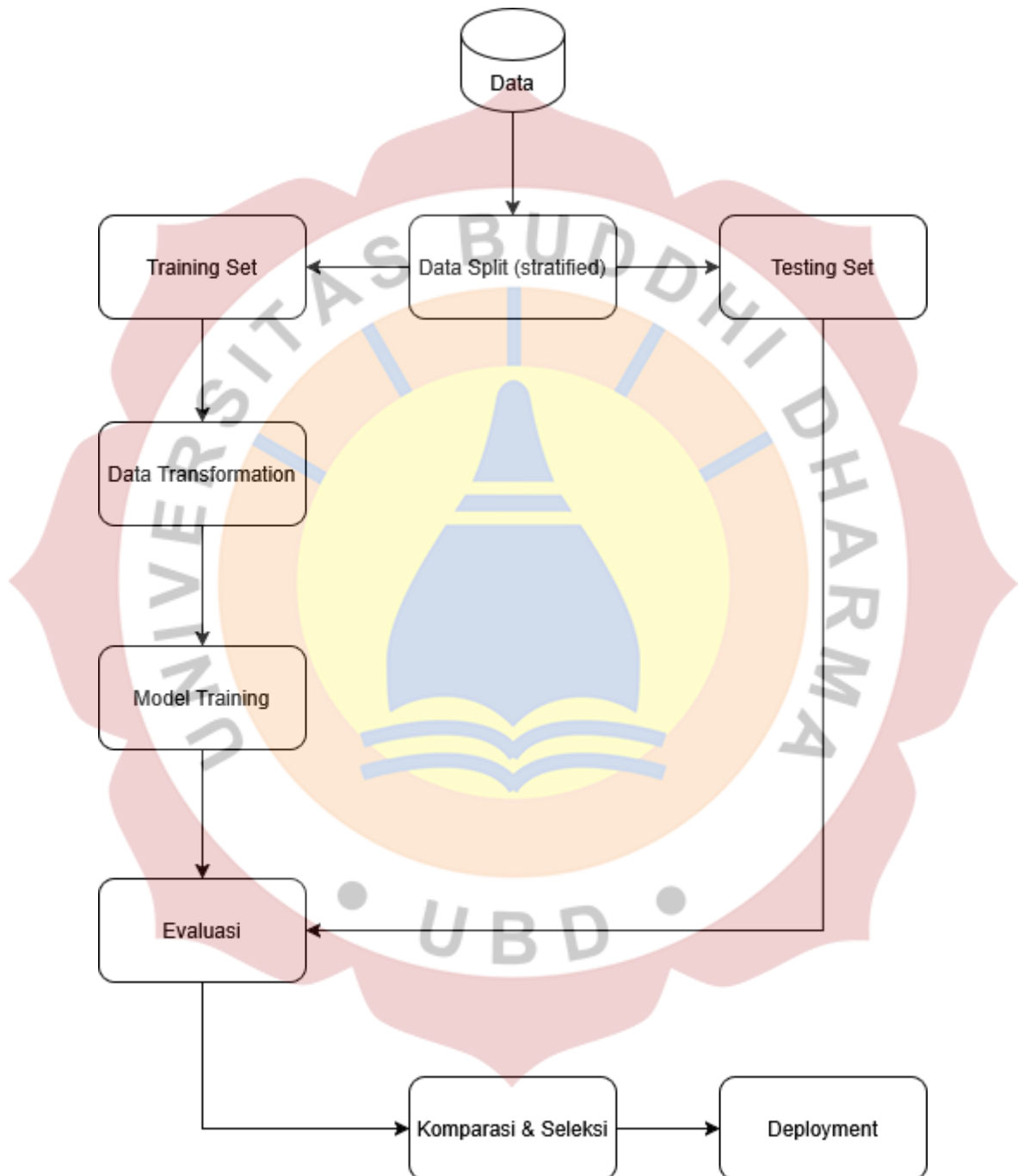
*sensing* baik dalam klasifikasi atau segmentasi, *visible light* ataupun multispektral, citra UAV ataupun satelit. Kedua model dapat mengatasi permasalahan data *long-tailed* yang sering terjadi dalam *remote sensing* dengan memanfaatkan kemampuan *foundation model* dan metode *fine-tuning* yang sesuai. Dari literatur yang telah dibahas, didapat bahwa kedua memiliki kasus di mana salah satu dari keduanya dapat menjadi metode yang lebih unggul. Terutama model ViT yang menunjukkan potensi untuk melampaui performa model CNN pada beberapa kasus dan metrik.

Dari studi yang telah dibahas, upaya untuk menerapkan metode *deep learning* pada klasifikasi sistem cuaca dari citra awan satelit telah dilakukan menggunakan optimasi CNN, akan tetapi terbatas pada bagian dari *dataset* label tunggal yang tidak menginklusi kelas ekstra yang mencakup lebih dari setengah *dataset*. Dengan salah satu penelitian (Yousaf dkk., 2023) mencantumkan performa yang suboptimal dari inklusi kelas tersebut. Oleh karena itu, pada penelitian kali ini akan dilakukan perbandingan performa antara CNN dan ViT, yang merupakan model *deep learning* yang sering digunakan dalam *remote sensing*, pada *dataset* citra awan satelit yang bersifat *long-tailed* dengan memanfaatkan metode LIFT (Shi dkk., 2024) yang berupa *fine-tuning* dari *foundation model* kedua tipe model tersebut untuk mengatasi sifat *long-tailed* dari data citra awan satelit LSCIDMR.

## BAB III

### METODE PENELITIAN

#### 3.1 Kerangka Pemikiran



**Gambar 3.1 Kerangka Pemikiran**

Pada Gambar 3.1 dijabarkan alur kerja dari penelitian, dengan mempertimbangkan ketidakseimbangan dari data, pemisahan data dilakukan secara terstratifikasi untuk

menjamin kedua himpunan data representatif terhadap populasi. Data yang masuk ke dalam himpunan pelatihan akan melalui proses transformasi sederhana untuk disesuaikan ke dalam *input size* dari model. Data yang telah ditransformasi akan masuk ke tahap pelatihan dari model, di mana model akan melakukan *fit* terhadap data dan memulai mempelajari fitur yang terdapat dalam data sesuai, arsitektur CLIP yang digunakan terdiri dari *text encoder* dan *image encoder*, yang sesuai dengan metode LIFT, *text encoder* akan digunakan untuk menginisialisasi *weight* dari *classifier* dan *image encoder* sebagai *backbone* dari model. Setelah pelatihan model selesai, akan dilakukan evaluasi dengan himpunan tes dan hasil dari evaluasi model akan dibandingkan dari model dengan *backbone* dari ViT atau CNN untuk dilihat mana yang berperforma lebih baik dari metrik yang telah ditentukan. Setelah semua proses pelatihan selesai maka model terbaik akan dipilih untuk di-implementasikan.

### 3.2 Data Understanding

Penelitian ini dilakukan menggunakan *dataset* sekunder yang diambil dari “<https://github.com/Zjut-MultimediaPlus/LSCIDMR>”. *Dataset* bernama “*Large-Scale Satellite Cloud Image Database for Meteorological Research-Single-Label Annotation*” (C. Bai dkk., 2022) atau LSCIDMR-S yang terdiri atas 104390 *slice* dan total 11 buah kelas. Data menggunakan citra satelit yang dari satelit geostasioner *Himawari-8* yang diambil dari *Japan Aerospace Exploration Agency (JAXA) P-tree System*. Data dilakukan *pre-processing* oleh pembuat *dataset* dengan proses-proses seperti:

#### a. Channel Selection

Dalam proses ini dilakukan pemilihan *channel* spektrum dari satelit yang akan digunakan untuk mengkonstruksi database. *Channel* yang digunakan berupa *albedo\_3*, *albedo\_4*, dan *albedo\_5*, ketiga channel ini dinilai memiliki warna yang memungkinkan pembedaan antara awan air dan awan es, serta antara permukaan laut, vegetasi, dan gurun pasir.

b. *Time Selection*

Jangka bujur dari satelit berada pada  $80^{\circ}$  timur dan  $160^{\circ}$  barat. Pada jam 00:20 UTC+0, titik subsolar berada di tengah peta komposit. Citra yang ditangkap paling jelas terlihat pada waktu tersebut.

c. *Position Determination*

Posisi citra satelit diambil pada belahan bumi utara, dengan alasan bahwa *dataset* yang diproduksi akan lebih bermanfaat.

d. *Slices Size Determination*

Citra komplit dari satelit dipotong bagian lebih kecil dikarenakan besarnya resolusi asli yang dapat mempersulit pemrosesan kedepannya.

e. *Slicing Naming*

Format dari penamaan citra didasarkan terhadap posisi temporal dan spasial. Dengan bentuk sebagai berikut: *YYMMDD\_ii\_jj.png*. *YYYY* merupakan informasi tahun, *MM* informasi bulan, *DD* informasi tanggal, *ii* informasi urutan pemotongan pada garis lintang, dan *jj* informasi urutan pemotongan pada garis bujur.

Label pada LSCIDMR-S disimpan pada sebuah file csv yang berisikan korespondensi pada tiap potongan citra pada informasi label. Kelas-kelas label dari LSCIDMR-S berupa:

a. *Tropical Cyclone*

*Tropical cyclone* atau badai siklon tropis merupakan pusaran tekanan rendah yang timbul pada permukaan laut tropis atau subtropis. Fitur dari badai siklon tropis berupa sistem awan spiral yang berpusar melawan arah jarum jam. Kriteria inklusi citra pada kelas label ini berupa ada atau tidaknya mata dari badai siklon tropis pada citra.

b. *Extratropical Cyclone*

*Extratropical Cyclone* atau badai siklon ekstratropis merupakan pusaran tekanan rendah berbentuk elips yang terbentuk pada lintang sedang hingga tinggi. Kriteria inklusi citra berupa ada atau tidaknya mata dari badai siklon ekstratropis pada citra.

c. *Frontal Surface*

*Frontal Surface* atau muka fron merupakan antarmuka yang terjadi atas dua masa udara yang berlainan atributnya. Fitur muka fron pada citra dapat dikenali dengan lengkungan siklonis dengan sabuk awan yang membentang dari utara ke selatan. Kriteria inklusi berupa keberadaan muka fron pada citra.

d. *Westerly Jet*

*Westerly jet* atau jet angin antipasat merupakan angin antipasat kuat yang berada pada troposfer di atas sabuk angin antipasat umum. Figur jet angin antipasat dapat dikenali dari pita angin yang tipis dan kencang yang melengkung dari barat ke timur. Kriteria inklusi berupa keberadaan jet angin antipasat pada citra.

e. *Snow*

*Snow* atau salju merupakan akumulasi dari kristal es di permukaan. Figur salju dapat dikenali pada citra dengan warna biru tua, lebih gelap dibandingkan awan es altitud tinggi. Kriteria dari inklusi berupa keberadaan salju dalam citra.

f. *High Ice Cloud*

*High Ice Cloud* atau awan es altitud tinggi merupakan awan yang terdiri secara keseluruhan atau hampir keseluruhan dari kristal es. Awan yang termasuk kategori ini berupa awan sirus, sirokumulus, sirostratus, dan awan berlapis tinggi di altitud tinggi pada musim dingin. Figur awan es altitud tinggi terlihat biru pada citra. Kriteria inklusi berupa  $Area(High\ Ice\ Cloud) > 50\% \& Area(Else) < 20\%$ .

g. *Low Water Cloud*

*Low Water Cloud* atau awan air altitud rendah merupakan awan yang terdiri atas tetesan air. Figur dari awan dikenali dari warna merah jambu pada citra. Kriteria inklusi berupa  $Area(Low\ Water\ Cloud) > 50\% \& Area(Else) < 20\%$ .

h. *Ocean*

*Ocean* atau samudra merupakan istilah yang merujuk kepada perairan terbesar di muka bumi yang mencakup 70% dari permukaan bumi. Figur samudra dikenali dengan warna hitam pada citra. Kriteria inklusi berupa  $Area(Ocean) > 80\% \& Area(Else) < 20\%$ .

i. *Desert*

*Desert* atau gurun merupakan istilah yang merujuk kepada area tandus yang ditutupi oleh pasir. Figur gurun dalam citra dikenali dari warna coklat yang bervariasi. Kriteria inklusi berupa  $Area(Desert) > 50\% \& Area(Else) < 20\%$ .

j. *Vegetation*

*Vegetation* atau vegetasi merupakan istilah yang merujuk pada kumpulan tanaman yang menutupi lahan, termasuk hutan, padang rumput, rawa, dan lainnya. Figur dikenali dengan warna hijau bervariasi. Kriteria inklusi berupa  $Area(Vegetation) > 50\% \& Area(Else) < 20\%$ .

k. *LabelLess*

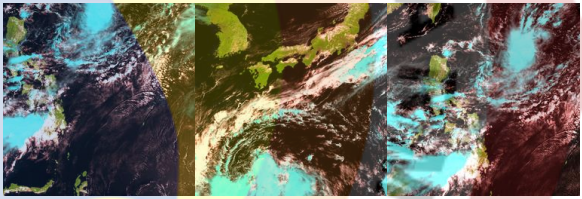
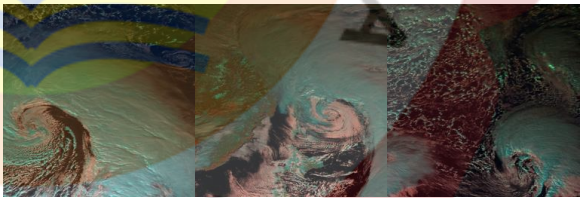
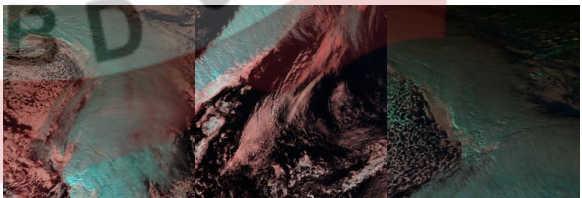
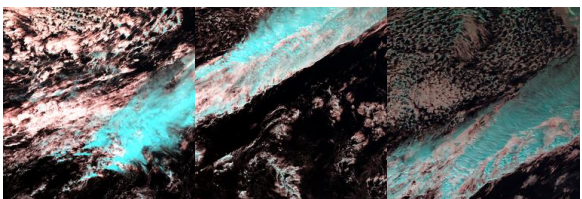
*LabelLess* merupakan sisa dari citra yang tidak memenuhi kriteria inklusi 10 kelas label di atas. Kelas ini dapat berarti tidak terdapat sistem cuaca yang mendominasi di dalam citra.

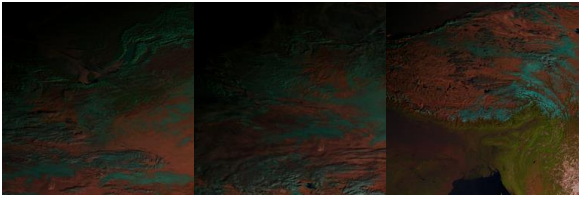
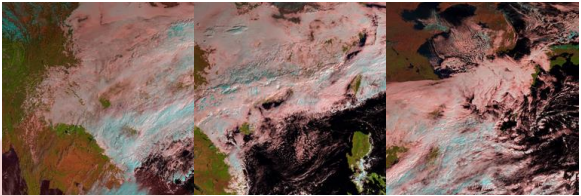
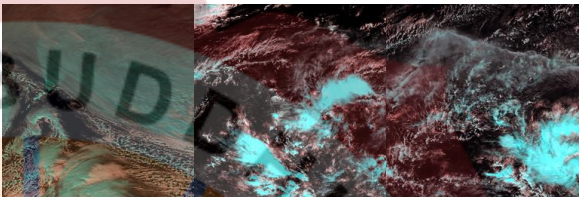
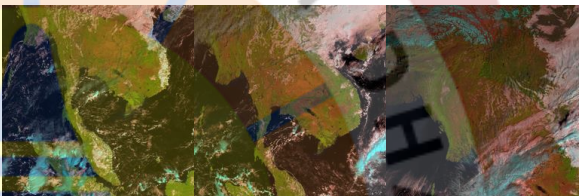
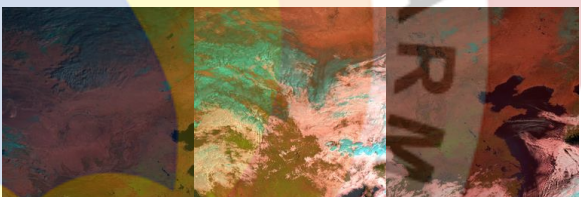
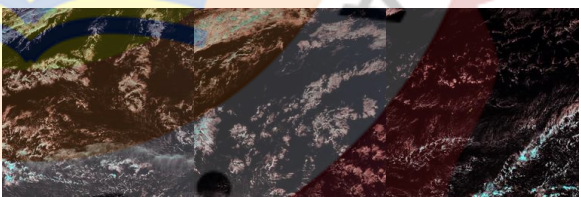
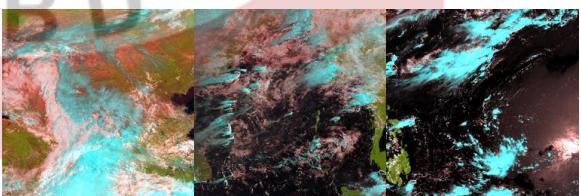
**Tabel 3.1 Persebaran Label LSCIDMR-S**

| Kelas                        | Jumlah | Rasio |
|------------------------------|--------|-------|
| <i>Tropical Cyclone</i>      | 3905   | 3,74% |
| <i>Extratropical Cyclone</i> | 4984   | 4,77% |

|                        |        |        |
|------------------------|--------|--------|
| <i>Frontal Surface</i> | 634    | 0,607% |
| <i>Westerly Jet</i>    | 628    | 0,602% |
| <i>Snow</i>            | 7631   | 7,31%  |
| <i>Low Water Cloud</i> | 1774   | 1,70%  |
| <i>High Ice Cloud</i>  | 5278   | 5,06%  |
| <i>Vegetation</i>      | 7831   | 7,51%  |
| <i>Desert</i>          | 4518   | 4,33%  |
| <i>Ocean</i>           | 4042   | 3,88%  |
| <i>LabelLess</i>       | 63765  | 61,07% |
| Jumlah                 | 104390 | 100%   |

**Tabel 3.2 Sampel Data LSCIDMR-S**

| Kelas                        | Sampel Gambar                                                                        |
|------------------------------|--------------------------------------------------------------------------------------|
| <i>Tropical Cyclone</i>      |  |
| <i>Extratropical Cyclone</i> |  |
| <i>Frontal Surface</i>       |  |
| <i>Westerly Jet</i>          |  |

|                        |                                                                                      |
|------------------------|--------------------------------------------------------------------------------------|
| <i>Snow</i>            |    |
| <i>Low Water Cloud</i> |    |
| <i>High Ice Cloud</i>  |    |
| <i>Vegetation</i>      |   |
| <i>Desert</i>          |  |
| <i>Ocean</i>           |  |
| <i>LabelLess</i>       |  |

### 3.3 Data Splitting

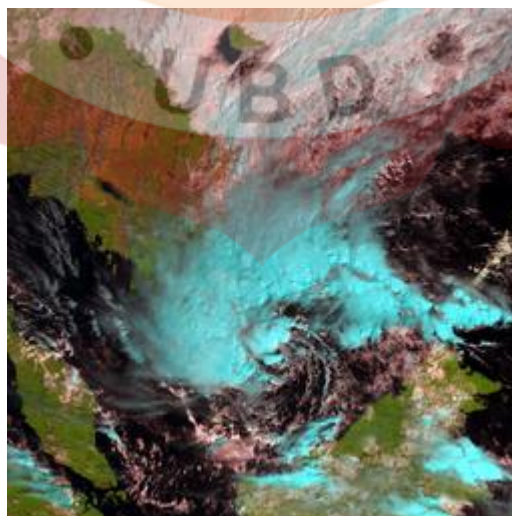
Dalam proses *data splitting*, data akan dipecah menjadi data pelatihan dan data tes dengan rasio 80:20, 80% dari data digunakan untuk pelatihan model dan 20% sisanya digunakan untuk menguji performa model, 80:20 merupakan salah satu dari rasio pemecahan

data yang umum digunakan (Joseph, 2022). Pemecahan data dilakukan secara *stratified*, yang berarti memecah populasi data menjadi beberapa kelompok atau strata (dalam hal ini berdasarkan kelas label) yang kemudian sampel acak diambil dari tiap-tiap kelompok dengan jumlah yang proporsional sesuai dengan jumlah dari kelompok relatif terhadap keseluruhan populasi. *Stratified sampling* memiliki kelebihan dalam sampel yang diambil akan merepresentasikan populasi dengan lebih baik dan meningkatkan validitas eksternal dari studi yang menggunakan *stratified sampling* (Berndt, 2020). Mempertimbangkan sifat *long-tailed* dari *dataset*, maka metode pengambilan sampel untuk pemecahan data secara terstratifikasi berfungsi untuk menjamin representasi terhadap kelas yang memiliki rasio lebih kecil dalam *dataset* untuk termasuk dalam kedua himpunan data baik himpunan pelatihan atau himpunan tes, serta mempertahankan distribusi *long-tail* dari dataset dalam kedua himpunan.

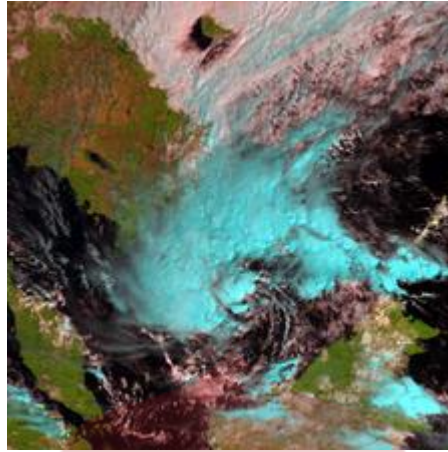
### **3.4 Data Preprocessing**

#### **3.4.1 Data Transformation**

Data pertama akan di-*resize* sesuai dengan besar input yang diterima oleh model, yang berupa 224\*224. Basis gambar di-*resize* dari ukuran 256\*256 dan bukan dari gambar awal 1000\*1000.



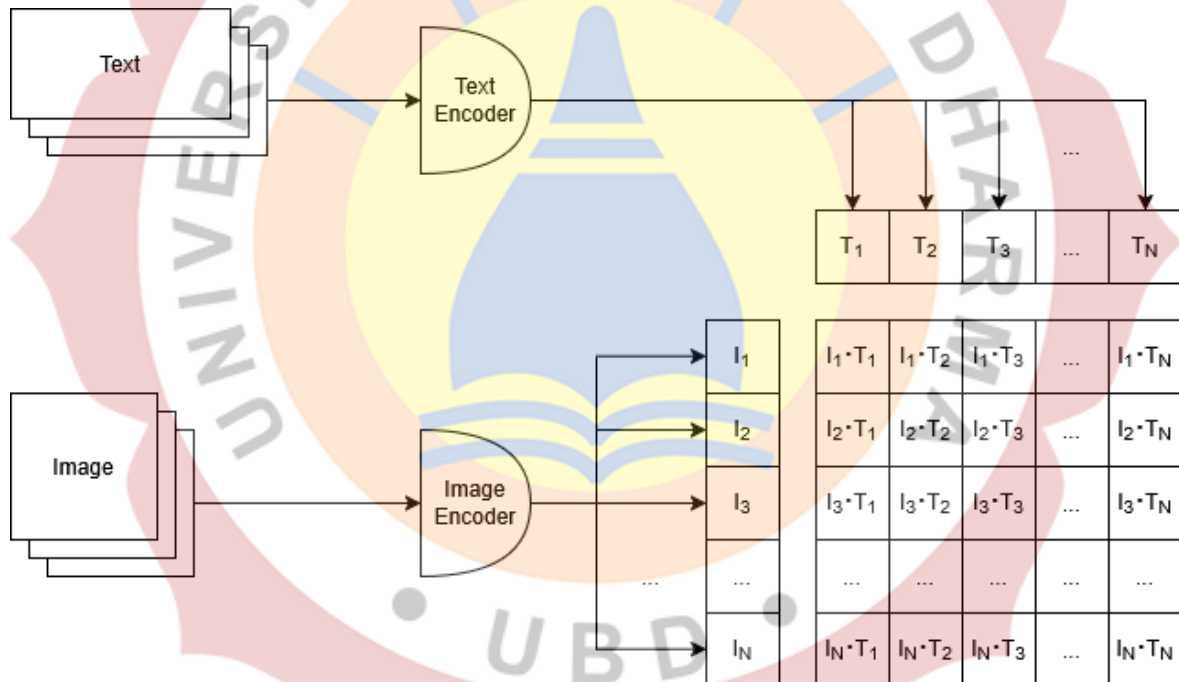
**Gambar 3.2 Gambar Sebelum *Resize***



Gambar 3.3 Gambar Sesudah *Resize*

### 3.5 Model Training

#### 3.5.1 Arsitektur CLIP



Gambar 3.4 Proses *Contrastive Pre-training CLIP*

(Radford dkk., 2021)

CLIP atau *Contrastive Image-Language Pre-training* merupakan sebuah model jaringan saraf buatan yang diberikan *pre-training* pada tugas prediksi *caption* pada gambar dengan metode *contrastive learning* (Radford dkk., 2021). CLIP memiliki arsitektur *dual-*

*encoder*, teks dan gambar, dengan *text encoder* dari CLIP berupa sebuah *Transformer*. Pada metode LIFT yang digunakan pada penelitian kali ini *text encoder* dari CLIP akan digunakan untuk inisialisasi *weight* dari *classifier* dengan mencari fitur semantik dari *prompt*, dan tidak lagi digunakan setelahnya. Pada penelitian kali ini *prompt* didasarkan pada *template* dari eksperimen *zero-shot* dari RemoteCLIP. Dengan prompt berupa “a satellite image showing {class name}.” Beserta *class name* yang dituliskan secara manual sesuai dengan Tabel 3.3. *Image encoder* dari CLIP akan menjadi *backbone* atau bagian dari jaringan saraf buatan yang mempelajari fitur dari data.

**Tabel 3.3 Nama Kelas Dalam *Prompt***

| Label                 | Class Name                                           |
|-----------------------|------------------------------------------------------|
| Desert                | desert                                               |
| Extratropical Cyclone | extratropical cyclone                                |
| Frontal Surface       | frontal surface                                      |
| High Ice Cloud        | high-altitude ice cloud                              |
| Low Water Cloud       | low-altitude water cloud                             |
| Ocean                 | ocean                                                |
| LabelLess             | no dominating weather system or geographical feature |
| Snow                  | snow cover                                           |
| Tropical Cyclone      | tropical cyclone                                     |
| Vegetation            | vegetation                                           |
| Westerly Jet          | westerly jet stream                                  |

### 3.5.2 *Hyperparameter*

*Hyperparameter* pada proses training mengikuti nilai *default* yang diberikan pada *main result* dari LIFT, dengan nilai menurut Tabel 3.4.

**Tabel 3.4 Hyperparameter Pada Pelatihan Dan Nilainya**

| Hyperparameter            | Nilai                                       |
|---------------------------|---------------------------------------------|
| PEFT Method               | Adaptformer (ViT) / Bias Tuning & SSF (CNN) |
| Epoch                     | 10                                          |
| Batch Size                | 128                                         |
| Loss Function             | Logit Adjustment                            |
| Momentum                  | 0.9                                         |
| Learning Rate             | 0.01                                        |
| Input Size                | 224*224                                     |
| Weight Decay              | 0.0005                                      |
| Classifier Initialisation | Text Features                               |

### 3.5.3 Perangkat Keras

Spesifikasi perangkat keras yang digunakan dalam proses pelatihan model dalam penelitian ini dijabarkan pada tabel 3.5

**Tabel 3.5 Spesifikasi Perangkat Keras**

| Perangkat Keras | Spesifikasi                   |
|-----------------|-------------------------------|
| CPU             | AMD Ryzen 5 8600G             |
| GPU             | NVIDIA GeForce GTX 1660 Super |
| RAM             | 16GB DDR5 4.8 GHz             |
| Disk            | 512 GB SSD + 1TB HDD          |

### 3.5.4 Lingkungan Pengembangan

Proses pelatihan model dilaksanakan pada lingkungan di dalam Visual Studio Code dengan bahasa pemrograman Python. Beberapa pustaka Python yang digunakan berupa PyTorch, yang digunakan untuk berbagai kebutuhan *machine learning*, dan terutama paket

torchvision yang berfokus pada penglihatan komputer. Selain itu juga digunakan pustaka seperti OpenCLIP yang berupa implementasi *open source* dari CLIP yang mana RemoteCLIP dan GeoRSCLIP yang digunakan dalam penelitian kali ini tersedia dalam format OpenCLIP.



**Gambar 3.5** Tampilan *Visual Studio Code*

### 3.5.5 *Backbone*

Pada penelitian kali ini, akan digunakan model CLIP dari RemoteCLIP dan GeoRSCLIP (Liu dkk., 2024; Z. Zhang dkk., 2024). RemoteCLIP tersedia dalam ResNet-50, ViT-Base-32, dan ViT-Large-14. GeoRSCLIP tersedia dalam ViT-Base-32, ViT-Large-14, dan ViT-Large-14 dengan  $336 \times 336$  input size. Penelitian kali ini akan berfokus pada ViT-Base-32 dari kedua model CLIP untuk *backbone* berbasis ViT dan ResNet-50 RemoteCLIP untuk CNN. Semua model menerima input dengan besar  $224 \times 224$ .

### 3.5.6 Metode Evaluasi

Pengujian performa model dilakukan dengan menggunakan *Confusion Matrix*. Nilai yang didapat melalui *Confusion Matrix* akan digunakan untuk mencari *precision*, *recall*, *accuracy*, dan *f1* dan *macro f1 score*. Dihitung sesuai dengan persamaan (2) dan (3).

|                       | TC | ETC | FS | WJ | S | LWC | HIC | V | D | O | LL |
|-----------------------|----|-----|----|----|---|-----|-----|---|---|---|----|
| Tropical Cyclone      |    |     |    |    |   |     |     |   |   |   |    |
| Extratropical Cyclone |    |     |    |    |   |     |     |   |   |   |    |
| Frontal Surface       |    |     |    |    |   |     |     |   |   |   |    |
| Westerly Jet          |    |     |    |    |   |     |     |   |   |   |    |
| Snow                  |    |     |    |    |   |     |     |   |   |   |    |
| Low Water Cloud       |    |     |    |    |   |     |     |   |   |   |    |
| High Ice Cloud        |    |     |    |    |   |     |     |   |   |   |    |
| Vegetation            |    |     |    |    |   |     |     |   |   |   |    |
| Desert                |    |     |    |    |   |     |     |   |   |   |    |
| Ocean                 |    |     |    |    |   |     |     |   |   |   |    |
| LabelLess             |    |     |    |    |   |     |     |   |   |   |    |

**Gambar 3.6 Ilustrasi *Confusion Matrix* dengan *Dataset* LSCIDMR-S**