

**PENERAPAN ARSITEKTUR *VISION TRANSFORMER*  
DALAM PENENTUAN JENIS TANAH PADA CITRA DIGITAL**

**SKRIPSI**

**Sebagai Salah Satu Syarat untuk Memperoleh Gelar  
Sarjana Sains Bidang Studi Matematika**

**Oleh :**

**Ilham Tri Wibowo**

**08011281823024**



**JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS SRIWIJAYA**

**2022**

**LEMBAR PENGESAHAN**

**PENERAPAN ARSITEKTUR *VISION TRANSFORMER*  
DALAM PENENTUAN JENIS TANAH PADA CITRA DIGITAL**

**SKRIPSI**

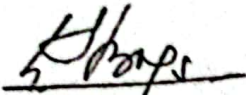
**Sebagai Salah Satu Syarat untuk Memperoleh Gelar  
Sarjana Sains Bidang Studi Matematika**

**Oleh**

**ILHAM TRI WIBOWO**

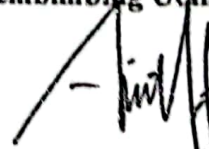
**NIM. 08011281823024**

**Pembimbing Kedua**



**Dr. Bambang Suprihatin, M.Si**  
**NIP. 197101261994121001**

**Indralaya, 4 Agustus 2022  
Pembimbing Utama**



**Anita Desiani, S.Si., M.Kom**  
**NIP. 19771211 2003122002**

**Mengetahui,  
Ketua Jurusan Matematika**



**Drs. Sugandi Yahdin, M.M**  
**NIP. 195807271986031003**

## HALAMAN PERSEMBAHAN

*Kupersembahkan skripsi ini untuk:*

*Yang Maha Kuasa Allah Subhanahu Wa Ta'ala*

*Orang tuaku Tersayang*

*Saudara-saudaraku tercinta*

*Keluarga besarku,*

*Semua guru dan dosenku,*

*Sahabat-sahabatku,*

*Almamaterku*

Motto

*“Whatever you are, be a good one”*

## KATA PENGANTAR

Puji syukur kehadiran Allah Subhanahu Wa Ta'ala yang telah memberikan rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Penerapan Arsitektur *Vision Transfomer* Dalam Penentuan Jenis Tanah Pada Citra Digital” sebagai salah satu syarat memperoleh gelar sarjana sains bidang studi Matematika di Fakultas MIPA Universitas Sriwijaya.

Penulis menyadari bahwa proses pembuatan skripsi ini merupakan proses pembelajaran yang sangat berharga serta tak lepas dari kekurangan dan keterbatasan. Dengan segala hormat dan kerendahan hati, penulis mengucapkan terima kasih dan penghargaan kepada:

1. Orang tuaku tercinta, Mama Suharti yang melahirkanku, dan tidak pernah berhenti berjuang serta memberikan yang terbaik untukku sebagai putranya, Ayah tiriku Bagus Priyo Adhi Wibowo yang sabar atas sikapku, terkhusus Almarhum Ayahku Buimin yang sudah berjuang untuk hidupku. Terima kasih karena tak pernah lelah mendidik, menasehati, membimbing, mendukung dan terus mendo'akan anaknya. Semoga semua yang telah dilakukan menjadi amal ibadah.
2. Bapak Drs Sugandi Yahdin, M.M selaku Ketua Jurusan Matematika FMIPA Universitas Sriwijaya yang telah memberikan arahan dan motivasi kepada penulis selama proses perkuliahan dan Ibu Dr. Dian Cahyawati Sukanda, M.Si selaku Sekretaris Jurusan Matematika FMIPA Universitas Sriwijaya yang telah mengarahkan urusan akademik kepada penulis.

3. Ibu Anita Desiani, M.Kom selaku dosen pembimbing utama yang telah bersedia meluangkan waktu, tenaga, pikiran untuk memberikan bimbingan, pengarahan dan didikan berharga selama proses pembuatan skripsi, kompetisi atau program mahasiswa, dan perjalanan perkuliahan ini.
4. Bapak Dr.Bambang Suprihatin, M.Si selaku dosen pembimbing pendamping yang telah bersedia meluangkan waktu, tenaga, pikiran untuk memberikan bimbingan, pengarahan dan didikan berharga selama proses pembuatan skripsi dan perjalanan perkuliahan ini.
5. Bapak Drs. Ali Amran, M.T dan Ibu Oki Dwipurwani, M.Si selaku dosen pembahas dan penguji yang telah memberikan tanggapan, kritik serta saran yang sangat bermanfaat untuk perbaikan dan penyelesaian skripsi ini.
6. Bapak Dr. Ngudiantoro, M.Si., selaku sekretaris tim pelaksana tugas akhir penulis dan Ibu Dr. Evi Yuliza, M.Si selaku ketua tim pelaksana tugas akhir dan Ibu Dr. Yulia Resti, M.Si selaku dosen pembimbing akademik yang telah membimbing dan mengarahkan urusan akademik selama masa perkuliahan penulis.
7. Seluruh Dosen di Jurusan Matematika FMIPA yang telah memberikan ilmu, nasihat, serta bimbingan selama proses perkuliahan dan seluruh guru yang telah memberikan ilmu yang sangat bermanfaat hingga mengantarkan penulis pada pendidikan ini.
8. Pak Irwansyah selaku admin dan Ibu Hamidah selaku pegawai tata usaha Jurusan Matematika FMIPA yang telah membantu penulis selama perkuliahan.

9. Kakak-kakak dan kakak-kakak iparku tercinta Yanti Widya Ningrum, Dina Kustari, Rajudin dan Riki Aripin serta Keponakan-keponakanku tersayang Zavier Assyauqie Syafiq, Raya Intan Putri Perdana, Firel Attariz Calief dan Alesha Maureena Ayu.
10. Semua sahabat seperjuangan selama masa perkuliahan dan skripsi ini, Ambyar, Sesat Club, ViT Teams, teman-teman komputasi dan teman dari masa SMA ku. Terima kasih sudah menjadi orang-orang baik dan random yang menjadi penyemangat disekeliling penulis yang selalu mendukung, membantu, dan memberikan semangat kepada penulis serta pelajaran luar biasa dari proses pertemanan kita.
11. Kakak-kakak tingkat angkatan 2016 dan 2017 bidang minat komputasi yang telah banyak membantu serta berbagi ilmu selama proses skripsi.
12. Keluarga Matematika 2018, BPH COIN 2019/2020 dan rekan-rekan perlombaan selama perkuliahan.
13. Serta semua pihak yang tidak dapat penulis sebutkan satu per satu, semoga segala kebaikan yang diberikan mendapat balasan terbaik dari Allah.
14. *Thank u for myself.*

Semoga skripsi ini dapat menambah pengetahuan dan bermanfaat bagi mahasiswa/mahasiswi Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya dan pihak yang memerlukan.

Indralaya, Juli 2022

Penulis

# **APPLICATION OF VISION TRANSFORMER ARCHITECTURE IN DETERMINING SOIL TYPE IN DIGITAL IMAGE**

**By:**

**Ilham Tri Wibowo  
08011281823024**

## **ABSTRACT**

Soil is a thin layer of tiny particles formed by the natural process of weathering rocks on the earth's surface. Soil characteristics vary greatly from place to place. Soil classification currently plays an important role in agriculture to determine the right plants for certain soil types. So it is necessary to have a technology that is able to classify soil types in order to minimize planting errors and increase agricultural production. Therefore, this study was conducted to classify soil images in determining soil types. Broadly speaking, soil can be classified into 5 classes, namely Black Soil, Cinder Soil, Laterite Soil, Peat Soil and Yellow Soil. Deep learning is often used to perform various tasks in machine learning such as classification. One architecture that is very well classified is the Vision Transformer (ViT). The advantages of ViT is that it performs better on small datasets and has patching techniques in its architecture. In this study, soil classification was carried out to determine the type of soil using the ViT architecture. The research stages are data collection, data pre-processing, data augmentation, application of ViT, training, testing, and performance evaluation as well as analysis and interpretation. The study used soil image data sourced from Kaggle by obtaining performance evaluations, namely accuracy, sensitivity, specificity, F1-score, and Cohen's kappa respectively 97.17%, 93.20%, 97.62%, 92.99% , and 89.85%. Based on these results, it shows that the ViT Small Variant architecture is able to perform classification tasks to determine the type of soil from the image data used.

**Keyword:** Classification, Soil Type, Vision Transformer.

## **PENERAPAN ARSITEKTUR *VISION TRANSFORMER* DALAM PENENTUAN JENIS TANAH PADA CITRA DIGITAL**

**Oleh:**

**Ilham Tri Wibowo  
08011281823024**

### **ABSTRAK**

Tanah adalah lapisan tipis partikel kecil yang terbentuk oleh proses alami pelapukan batuan di permukaan bumi. Jenis tanah yang berbeda memiliki karakteristik yang sangat berbeda. Klasifikasi tanah saat ini sangat berperan penting dalam bidang pertanian untuk menentukan tanaman yang tepat pada jenis tanah tertentu. Sehingga diperlukan adanya teknologi yang mampu mengklasifikasikan jenis tanah agar dapat meminimalisir kesalahan penanaman serta meningkatkan hasil produksi pertanian. Maka dari itu, penelitian ini dilakukan untuk mengklasifikasikan citra tanah dalam menentukan jenis tanah. Secara garis besar tanah dapat diklasifikasikan menjadi 5 kelas yaitu *Black Soil*, *Cinder Soil*, *Laterite Soil*, *Peat Soil* dan *Yellow Soil*. *Deep learning* sering digunakan untuk melakukan berbagai tugas dalam pembelajaran mesin seperti klasifikasi. Salah satu arsitektur yang sangat baik dalam klasifikasi adalah *Vision Transformer (ViT)*. Kelebihan *ViT* memiliki performa lebih baik pada kumpulan data kecil dan memiliki teknik *patching* di dalam arsitekturnya. Pada penelitian ini, dilakukan klasifikasi tanah untuk menentukan jenis tanah menggunakan arsitektur *ViT*. Tahapan penelitian adalah pengumpulan data, *pre-processing* data, augmentasi data, penerapan *ViT*, *training*, *testing*, dan evaluasi kinerja serta analisis dan interpretasinya. Penelitian menggunakan data citra tanah yang bersumber dari Kaggle dengan memperoleh evaluasi kinerja yaitu akurasi, sensitivitas, spesifisitas, *F1-score*, dan *cohen's kappa* masing-masing sebesar 97,17%, 93,20%, 97,62%, 92,99%, dan 89,85%. Berdasarkan hasil yang tersebut, menunjukkan bahwa arsitektur *ViT Small Variant* mampu melakukan tugas klasifikasi untuk menentukan jenis tanah dari data citra yang digunakan.

Kata Kunci: Klasifikasi, Jenis Tanah, *Vision Transformer*.



## PERNYATAAN

Saya yang bertanda tangan di bawah ini :

Nama : Ilham Tri Wibowo

NIM : 08011281823024

Jurusan : Matematika

Menyatakan dengan ini saya bersungguh-sungguh bahwa skripsi yang berjudul “Penerapan Arsitektur *Vision Transfomer* Dalam Penentuan Jenis Tanah Pada Citra Digital” merupakan karya yang saya susun sendiri dan saya tidak melakukan penjiplakan dari karya manapun serta saya melakukan pengutipan sesuai dengan pedoman keilmuan yang berlaku seperti tertuang dalam Peraturan Menteri Pendidikan Nasional Republik Indonesia Nomor 17 Tahun 2010 terkait Pencegahan dan Penanggulangan Plagiat di Perguruan Tinggi. Apabila dikemudian hari terdapat pelanggaran yang ditemukan dalam skripsi saya ataupun adanya pengaduan dari pihak lain terhadap keaslian skripsi saya, maka saya bersedia menanggung sanksi yang dijatuhkan kepada saya.

Demikianlah pernyataan ini dibuat dengan sungguh-sungguh tanpa adanya paksaan dari pihak manapun.

Indralaya, 4 Agustus 2022

Yang membuat pernyataan



Ilham Tri Wibowo

NIM. 08011281823024

## DAFTAR ISI

<b>HALAMAN JUDUL</b> .....	<b>i</b>
<b>HALAMAN PENGESAHAN</b> .....	Error! Bookmark not defined.
<b>HALAMAN PERSEMBAHAN</b> .....	<b>iii</b>
<b>KATA PENGANTAR</b> .....	<b>iv</b>
<b>ABSTRACT</b> .....	<b>vii</b>
<b>ABSTRAK</b> .....	<b>viii</b>
<b>DAFTAR ISI</b> .....	<b>ix</b>
<b>DAFTAR TABEL</b> .....	<b>xi</b>
<b>DAFTAR GAMBAR</b> .....	<b>xi</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1. Latar Belakang .....	1
1.2. Perumusan Masalah.....	4
1.3. Pembatasan Masalah .....	4
1.4. Tujuan.....	5
1.5. Manfaat.....	5
<b>BAB II TINJAUAN PUSTAKA</b> .....	<b>6</b>
2.1. Citra Digital .....	6
2.2. Klasifikasi Citra.....	8
2.3. Pre-Processing dan Augmentasi .....	9
2.4. Arsitektur Vision Transformer .....	9
2.4.1. Patches.....	10
2.4.2. <i>Position and CLS Embedding</i> .....	11
2.4.3. <i>Position Encoding</i> .....	11
2.4.4. <i>Transformer Encoder</i> .....	12
2.4.5. <i>Self Attention</i> .....	12
2.4.6. <i>Multi Self Attention (MSA)</i> .....	14
2.4.7. <i>Multi Layer Perceptron (MLP)</i> .....	15
2.4.8. <i>Vision Transformer Variants</i> .....	15
2.5. Activation Function : Gaussian Error Linear Unit (GELU).....	15
2.6. Activation Function : Softmax .....	16
2.7. Loss Function : Categorical Cross Entropy.....	16
2.8. <i>Layer Normalization</i> .....	17

2.9.	Optimization Function : Adam (Adaptive Momen Estimation) .....	17
2.10.	Confusion <i>Matrix</i> .....	18
<b>BAB III METODOLOGI PENELITIAN .....</b>		<b>22</b>
3.1.	Tempat .....	22
3.2.	Waktu .....	22
3.3.	Alat .....	22
3.4.	Metode Penelitian .....	22
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>		<b>26</b>
4.1.	Deskripsi Data .....	26
4.2.	<i>Pre-processing</i> dan Augmentasi Data .....	27
4.3.	Implementasi Arsitektur <i>Vision Transformer</i> dalam Klasifikasi Citra ..	29
4.4.	Operasi Manual pada Arsitektur <i>Vision Transformer</i> .....	32
4.5.	Proses Pelatihan Data .....	56
4.6.	Proses Pengujian Data .....	59
4.7.	Evulasi Kinerja Model .....	62
4.8.	Analisis dan Interpretasi Hasil .....	67
<b>BAB V KESIMPULAN DAN SARAN .....</b>		<b>69</b>
5.1.	Kesimpulan .....	69
5.2.	Saran .....	69
<b>DAFTAR PUSTAKA .....</b>		<b>70</b>

## DAFTAR TABEL

Tabel 2. 1 Gambar 5 Jenis Tanah.....	7
Tabel 2. 2 Confusion Matrix Ukuran 5x5.....	19
Tabel 4. 1. Data Sampel Citra Jenis Tanah.....	26
Tabel 4. 2. Nilai bobot hidden layer 1, hidden layer 2 dan output.....	49
Tabel 4. 3 Confusion Matrix.....	59
Tabel 4. 4 Hasil Akurasi.....	63
Tabel 4. 5 Hasil Sensitivitas.....	63
Tabel 4. 6 Hasil Spesifisitas.....	64
Tabel 4. 7 Hasil F1-Score.....	65
Tabel 4. 8 Hasil Cohen's Kappa.....	66
Tabel 4. 9 Hasil Penilaian Evaluasi Kinerja Arsitektur ViT.....	66
Tabel 4. 10 Perbandingan Evaluasi Kinerja Model.....	67

## DAFTAR GAMBAR

Gambar 2. 1 Representasi Citra Digital .....	6
Gambar 2. 2 Ilustrasi Arsitektur ViT.....	9
Gambar 2. 3 Ilustrasi Transformer Encoder.....	12
Gambar 2. 4. Ilustrasi Scaled Dot Product Attention.....	13
Gambar 2. 5 Ilustrasi Multi Head Attention .....	14
Gambar 4. 1 Flip-UP-to-DOWN (FlipUD) dan Flip-Left-to-Right (FlipLR).....	28
Gambar 4. 2 Rotasi Citra pada Augmentasi.....	28
Gambar 4. 3 Komponen Utama pada Arsitektur ViT .....	29
Gambar 4. 4 Patch and Positional Embedding pada Arsitektur ViT.....	30
Gambar 4. 5 Transformer Encoder pada Arsitektur ViT .....	31
Gambar 4. 6 Multi Layer Perceptron Head pada Arsitektur ViT.....	31
Gambar 4. 7 Ilustrasi Patching.....	32
Gambar 4. 8 Ilustrasi Flat Patches .....	33
Gambar 4. 9. Hasil Proses Training Per-Epoch pada Arsitektur ViT .....	57
Gambar 4. 10. Grafik Akurasi Training Arsitektur ViT .....	58
Gambar 4.11. Grafik Loss Training Arsitektur ViT .....	58

# **BAB I**

## **PENDAHULUAN**

### **1.1. Latar Belakang**

Dalam beberapa tahun terakhir, pendekatan klasifikasi berbasis *Deep Learning* (DL) telah memberikan kinerja yang canggih (Alom et al., 2019). Deep learning memiliki manfaat untuk mendefinisikan model data tidak terstruktur seperti foto, suara, video, dan teks (Liang et al., 2017). Deep learning telah terbukti menjadi teknik yang Jenis tanah yang berbeda memiliki karakteristik yang sangat berbeda menjanjikan untuk klasifikasi citra (Y. Jiang et al., 2019). Klasifikasi citra adalah proses mengidentifikasi dan mengelompokkan citra ke dalam kategori tertentu (Lee & Shin, 2020). Saat ini, klasifikasi tanah sangat penting untuk pertanian karena membantu mengidentifikasi tanaman terbaik untuk berbagai jenis tanah, mengurangi kesalahan penanaman dan meningkatkan hasil panen. (Adila et al., 2018).

Secara garis besar tanah dikelompokkan menjadi 5 jenis tanah diantaranya *Black Soil, Cinder Soil, Laterite Soil, Peat Soil* dan *Yellow Soil* (Hatano et al., 2021). Jenis tanah yang berbeda memiliki karakteristik yang sangat berbeda (Rahman et al., 2018). Tanah adalah lapisan tipis partikel kecil yang terbentuk oleh proses alami pelapukan batuan di permukaan bumi (Frisca Panjaitan & M. Madjid B. Damanik, 2015). Tanah terdiri dari mineral, bahan organik, endapan, dan berbagai ukuran partikel (Herwandi et al., 2017). Klasifikasi jenis tanah akan dilakukan dengan memanfaatkan *dataset* yang ada. Salah satu DL yang paling populer adalah *Convolutional Neural Network* (CNN) (Albawi et al., 2017).

CNN telah menjadi metode DL yang efektif dalam pemecahan masalah klasifikasi citra dengan tingkat akurasi yang tinggi (Prajapati et al., 2017). CNN dapat memiliki ratusan hingga jutaan parameter yang masing – masing belajar untuk mendeteksi fitur citra yang berbeda (Salman et al., 2020). Multi-Layer Perceptron (MLP) telah dikembangkan menjadi CNN yang khusus dibuat untuk menangani data dua dimensi atau citra. CNN merupakan pengembangan dari *Multi-Layer Perceptron* (MLP), yang dirancang khusus untuk mengolah data dua dimensi atau data citra (James et al., 2020). CNN memiliki kelebihan dalam mengolah *input* data berukuran  $m \times n$  seperti citra (Rizwan I Haque & Neubert, 2020).

Pada penelitian klasifikasi jenis tanah diantaranya (Hamzah et al., 2021) menerapkan arsitektur CNN menghasilkan accuracy sebesar 97%. (Azizi et al., 2020) menerapkan arsitektur Inception-v4 menghasilkan nilai akurasi 95,83%. Dapat dilihat bahwa dari kedua penelitian tersebut hanya menghitung nilai *accuracy* namun tidak memaparkan nilai sensitivitas, spesifisitas, *F1-Score*, dan *Cohen's kappa*. Arsitektur CNN memiliki kelemahan yaitu memerlukan jumlah data yang besar agar mendapatkan kinerja yang baik dan proses dalam pelatihan model yang cukup lama (Barbero-Gómez et al., 2021). Dikarenakan keterbatasan arsitektur CNN dibutuhkan sebuah teknik memperbanyak data salah satunya adalah *Patching*. *Patching* adalah proses membagi citra asli menjadi potongan atau *patch* yang lebih kecil berguna untuk menambah data yang tersedia untuk pelatihan dan sangat membantu ketika memiliki kumpulan data yang terbatas (Misra et al., 2020). Terdapat beberapa teknik *Patching* yang dilakukan oleh beberapa peneliti diantaranya (Qi et al., 2016) menerapkan metode *patching circular* yaitu

mengambil *patch* melingkar simetris terhadap setiap garis yang melewati pusat. (Misra et al., 2020) metode *patching* yang digunakan *neighbourhood local information* dengan ukuran citra yang ditentukan sebesar 224 x 224 piksel. Diperlukan sebuah arsitektur yang memiliki sebuah teknik *patching* di dalam arsitekturnya seperti *Vision Transformer (ViT)*.

*ViT* merupakan model transformasi murni pada urutan *patch* citra untuk mengklasifikasikan seluruh citra (Han et al., 2022). Citra yang dianalisis dibagi menjadi *patch*, kemudian diubah menjadi urutan dengan meratakan dan menyematkan (Bazi et al., 2021). *ViT* biasanya menggunakan token kelas yang mengumpulkan informasi global untuk memprediksi kelas keluaran (Z. Jiang et al., 2021). Untuk menghasilkan hubungan kontekstual jarak jauh antara piksel dalam citra, *ViT* menggunakan proses *multihead attention* sebagai elemen bangunan dasar (Bazi et al., 2021). Selain itu, *ViT* memiliki performa lebih baik pada kumpulan data kecil (Nijhawan et al., 2022).

Beberapa penelitian telah menunjukkan hasil kinerja yang sangat baik menggunakan arsitektur *ViT* diantaranya (Tanzi et al., 2022) melakukan perbandingan klasifikasi citra antara arsitektur *ViT* dan arsitektur *InceptionV3*. Hasil evaluasi menunjukkan *ViT* mendapatkan hasil yang lebih baik sebesar 25% *accuracy* lebih tinggi, 20% *precision* lebih tinggi, 23% *recall* lebih tinggi, 23% *F1-Score* lebih tinggi. (Huang et al., 2021) menerapkan perbandingan arsitektur CNN dan *ViT* untuk klasifikasi daur ulang sampah dengan *accuracy* 5.09% lebih baik pada *ViT* dalam pengklasifikasi citra. (Z. Jiang et al., 2022) menerapkan arsitektur *ViT* untuk menganalisis citra pada penyakit mata dengan membandingkan hasil



*accuracy* beberapa model seperti *VGG16*, *Resnet50*, *Densenet121*, *Efficientnet*, dan *ViT* berturut-turut adalah 98,51%, 97,32%, 97,02%, 34,16% dan 99,69%. (Gao et al., 2021) melakukan klasifikasi berdasarkan citra CT paru-paru untuk mengidentifikasi penyakit COVID dan NON-COVID dengan membandingkan antara klasifikasi arsitektur ViT dan DenseNet meliputi penilaian 2,9% *accuracy* dan 4% *F1-score* lebih tinggi menggunakan ViT. (Gheflati & Rivaz, 2021) mengusulkan klasifikasi kanker payudara berdasarkan citra payudara *ultrasound* (US) dengan hasil kinerja *accuracy* dan *Area Under the Curve* (AUC) sebesar 86,7% dan 95%.

Berdasarkan keterbatasan dalam arsitektur CNN serta permasalahan pada klasifikasi jenis tanah berdasarkan kebutuhan data klasifikasi jenis tanah. Pada penelitian ini mengusulkan arsitektur *ViT* pada klasifikasi jenis tanah. Penerapan arsitektur *ViT* diharapkan akan meningkatkan hasil kinerja prediksi pada klasifikasi jenis tanah. Hasil evaluasi kinerja arsitektur akan dilihat berdasarkan nilai *accuracy*, sensitivitas, spesifisitas, *F1-Score*, dan *Cohen's kappa*.

## **1.2. Perumusan Masalah**

Permasalahan pada penelitian ini akan membahas bagaimana hasil evaluasi kinerja dari arsitektur ViT dalam melakukan klasifikasi jenis tanah dengan mengukur nilai *accuracy*, sensitivitas, spesifisitas, *F1-Score*, dan *Cohen's kappa*.

## **1.3. Pembatasan Masalah**

Beberapa pembatasan masalah pada penelitian ini :

1. Penelitian ini dilakukan hanya untuk mengklasifikasikan jenis tanah yang memanfaatkan arsitektur *ViT* pada citra lima jenis tanah yang berbeda yaitu

*Black Soil, Cinder Soil, Laterit Soil, Peat Soil, dan Yellow Soil*. Hanya dataset Soil Types yang digunakan dalam penelitian ini.

2. Pembahasan evaluasi dari model arsitektur klasifikasi akan mengukur kinerja berdasarkan *accuracy*, sensitivitas, spesifisitas, *F1-Score*, dan *Cohen's kappa*.

#### **1.4. Tujuan**

Tujuan dari penelitian ini adalah untuk mengetahui hasil evaluasi kinerja *accuracy*, sensitivitas, spesifisitas, *F1-Score*, dan *Cohen's kappa* dari penerapan arsitektur *ViT* untuk klasifikasi jenis tanah.

#### **1.5. Manfaat**

Manfaat dari hasil penelitian ini :

1. Dapat diperoleh model yang dapat mengklasifikasikan jenis tanah yang lebih akurat dengan memanfaatkan arsitektur *ViT*.
2. Dapat digunakan sebagai referensi bagi penelitian yang akan melakukan penelitian pada bidang klasifikasi dan pengembangan *deep learning*.

## DAFTAR PUSTAKA

- Adila, W. N., Regasari, R., & Nurwasito, H. (2018). Sistem Pendukung Keputusan (SPK) Pemilihan Tanaman Pangan Pada Suatu Lahan Berdasarkan Kondisi Tanah Dengan Metode Promethee. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(5), 2118–2126.
- Agus, F., Anda, M., Jamil, A., & Masganti. (2016). Lahan gambut Indonesia : Pembentukan, karakteristik, dan potensi mendukung ketahanan pangan. In *Lahan Gambut Indonesia*.
- Albawi, S., Mohammed, T. A. M., & Alzawi, S. (2017). Understanding of a Convolutional Neural Network. *IEEE*, 1–6. <https://doi.org/10.1109/ICEngTechnol.2017.8308186>
- Alom, M. Z., Yakopcic, C., Hasan, M., Taha, T. M., & Asari, V. K. (2019). Recurrent residual U-Net for medical image segmentation. *Journal of Medical Imaging*, 6(01), 1. <https://doi.org/10.1117/1.jmi.6.1.014006>
- Ansari, A., & Borse, R. Y. (2019). Image Processing and Analysis. *International Journal of Engineering Research and Applications (IJERA)*, 3(4), 335–354. <https://doi.org/10.1016/b978-0-12-816501-0.00014-7>
- Azizi, A., Gilandeh, Y. A., Mesri-Gundoshmian, T., Saleh-Bigdeli, A. A., & Moghaddam, H. A. (2020). Classification of soil aggregates: A novel approach based on deep learning. *Soil and Tillage Research*, 199, 104586. <https://doi.org/10.1016/j.still.2020.104586>
- Ba, J. L., Kiros, J. R., & Hinton, G. E. (2016). Layer Normalization. *ArXiv Preprint ArXiv:1607.06450*. <http://arxiv.org/abs/1607.06450>
- Banerjee, K., Vishak Prasad, C., Gupta, R. R., Vyas, K., Anushree, H., & Mishra, B. (2021). Exploring alternatives to softmax function. *Proceedings of the 2nd International Conference on Deep Learning Theory and Applications, DeLTA 2021*, 81–86. <https://doi.org/10.5220/0010502000810086>
- Barbero-Gómez, J., Gutiérrez, P. A., Vargas, V. M., Vallejo-Casas, J. A., & Hervás-Martínez, C. (2021). An ordinal CNN approach for the assessment of neurological damage in Parkinson's disease patients. *Expert Systems with Applications*, 182, 115271. <https://doi.org/10.1016/j.eswa.2021.115271>
- Bazi, Y., Bashmal, L., Al Rahhal, M. M., Dayil, R. Al, & Ajlan, N. Al. (2021). Vision transformers for remote sensing image classification. *Remote Sensing*, 13(3), 1–20. <https://doi.org/10.3390/rs13030516>
- Cahya, F. N., Hardi, N., Riana, D., & Hadianti, S. (2021). Klasifikasi penyakit mata menggunakan convolutional neural network (cnn). *Jurnal Sistem Informasi*, 10, 618–626.
- Chen, C. F., Fan, Q., & Panda, R. (2021). CrossViT: Cross-Attention Multi-Scale Vision Transformer for Image Classification. *Proceedings of the IEEE*

- International Conference on Computer Vision*, 347–356.  
<https://doi.org/10.1109/ICCV48922.2021.00041>
- Chicco, D., Warrens, M. J., & Jurman, G. (2021). The Matthews Correlation Coefficient (MCC) is More Informative Than Cohen's Kappa and Brier Score in Binary Classification Assessment. *IEEE Access*, 9(Mcc), 78368–78381.  
<https://doi.org/10.1109/ACCESS.2021.3084050>
- Deng, X., Liu, Q., Deng, Y., & Mahadevan, S. (2016). An improved method to construct basic probability assignment based on the confusion matrix for classification problem. *Information Sciences*, 340–341, 250–261.  
<https://doi.org/10.1016/j.ins.2016.01.033>
- Frisca Panjaitan, J., & M. Madjid B. Damanik. (2015). Klasifikasi Tanah Berdasarkan Taksonomi Tanah 2014 di Desa Sembahe Kecamatan Sibolangit. *Jurnal Agroekoteknologi Universitas Sumatra Utara*, 3(4), 1447–1458.
- Gao, X., Qian, Y., & Gao, A. (2021). COVID-VIT: Classification of Covid-19 from CT chest images based on vision transformer models. *ArXiv Preprint ArXiv:2107.01682*, 59.
- Gheflati, B., & Rivaz, H. (2021). Vision Transformer for Classification of Breast Ultrasound Images. *CoRR*, abs/2110.1.  
<https://doi.org/https://doi.org/10.48550/arXiv.2110.14731>
- Hamzah, Diqi, M., & Ronaldo, A. D. (2021). Effective Soil Type Classification Using Convolutional Neural Network. *International Journal of Informatics and Computation*, 3(1), 20–29. <https://doi.org/10.35842/ijicom.v3i1.33>
- Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., Tang, Y., Xiao, A., Xu, C., Xu, Y., Yang, Z., Zhang, Y., & Tao, D. (2022). A Survey on Vision Transformer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–23. <https://doi.org/10.1109/TPAMI.2022.3152247>
- Hatano, R., Shinjo, H., & Takata, Y. (2021). The Soils of Japan. In *Geographical Review* (Vol. 29, Issue 1). <https://doi.org/10.2307/210080>
- Hendrycks, D., & Gimpel, K. (2016). Gaussian Error Linear Units (GELUs). *CoRR*, abs/1606.0, 1–9. <http://arxiv.org/abs/1606.08415>
- Herwandi, H., Marsudi, M., & Aprianto, A. (2017). PENGARUH GRADASI DAN KEPADATAN RELATIF (DR) TERHADAP NILAI PERMEABILITAS TANAH PASIR. *Jurnal Mahasiswa Teknik Sipil Universitas Tanjungpura*, 4, 1–13.
- Huang, K., Lei, H., Jiao, Z., & Zhong, Z. (2021). Recycling waste classification using vision transformer on portable device. *Sustainability (Switzerland)*, 13(21), 1–14. <https://doi.org/10.3390/su132111572>
- James, R. M., Kusriani, & Arief, M. R. (2020). Classification of X-ray COVID-19 Image Using Convolutional Neural Network. *IEEE*, 1–6.

<https://doi.org/10.1109/ICORIS50180.2020.9320828>

- Jiang, Y., Li, Y., & Zhang, H. (2019). Hyperspectral Image Classification Based on 3-D Separable ResNet and Transfer Learning. *IEEE Geoscience and Remote Sensing Letters*, *16*(12), 1949–1953. <https://doi.org/10.1109/LGRS.2019.2913011>
- Jiang, Z., Hou, Q., Yuan, L., Zhou, D., Shi, Y., Jin, X., Wang, A., & Feng, J. (2021). All Tokens Matter: Token Labeling for Training Better Vision Transformers. *Advances in Neural Information Processing Systems*, *34*, 18590–18602. <http://arxiv.org/abs/2104.10858>
- Jiang, Z., Wang, L., Wu, Q., Shao, Y., Shen, M., Jiang, W., & Dai, C. (2022). Computer-aided diagnosis of retinopathy based on vision transformer. *Journal of Innovative Optical Health Sciences*, *15*(2), 1–9. <https://doi.org/10.1142/S1793545822500092>
- Juneja, M., Thakur, N., Thakur, S., Uniyal, A., Wani, A., & Jindal, P. (2020). Gc-net for classification of glaucoma in the retinal fundus image. *Machine Vision and Applications*, *31*(5), 1–18. <https://doi.org/10.1007/s00138-020-01091-4>
- Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. *ArXiv Preprint ArXiv:1412.6980*, 1–15.
- Krstinić, D., Braović, M., Šerić, L., & Božić-Štulić, D. (2020). Multi-label Classifier Performance Evaluation with Confusion Matrix. *Comput Sci Inf Technol*, *10*, 01–14. <https://doi.org/10.5121/csit.2020.100801>
- Lee, I., & Shin, Y. J. (2020). Machine learning for enterprises: Applications, algorithm selection, and challenges. *Business Horizons*, *63*(2), 157–170. <https://doi.org/10.1016/j.bushor.2019.10.005>
- Li, X., Fan, P., Li, Z., Chen, G., Qiu, H., & Hou, G. (2021). Soil Classification Based on Deep Learning Algorithm and Visible Near-Infrared Spectroscopy. *Journal of Spectroscopy*, *2021*. <https://doi.org/10.1155/2021/1508267>
- Liang, H., Sun, X., Sun, Y., & Gao, Y. (2017). Text feature extraction based on deep learning: a review. *Eurasip Journal on Wireless Communications and Networking*, *2017*(1), 1–12. <https://doi.org/10.1186/s13638-017-0993-1>
- Luque, A., Carrasco, A., Martín, A., & de las Heras, A. (2019). The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition*, *91*, 216–231. <https://doi.org/10.1016/j.patcog.2019.02.023>
- Mishra, S., Vanli, O. A., Huffer, F. W., & Jung, S. (2016). Regularized discriminant analysis for multi-sensor decision fusion and damage detection with Lamb waves. *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2016*, *9803*, 728–741. <https://doi.org/10.1117/12.2217959>
- Misra, D., Crispim-Junior, C., & Tougne, L. (2020). Patch-Based CNN Evaluation

- for Bark Classification. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12540, 197–212. [https://doi.org/10.1007/978-3-030-65414-6\\_15](https://doi.org/10.1007/978-3-030-65414-6_15)
- Nijhawan, R., Batra, A., Loyola-González, O., Kumar, M., & Jain, D. K. (2022). Food Classification of Indian Cuisines Using Handcrafted Features and Vision Transformer Network. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4014907>
- Picek, L., Šulc, M., Matas, J., Heilmann-Clausen, J., Jeppesen, T. S., & Lind, E. (2022). Automatic Fungi Recognition: Deep Learning Meets Mycology. *Sensors*, 22(2), 633. <https://doi.org/10.3390/s22020633>
- Prajapati, S. A., Nagaraj, R., & Mitra, S. (2017). Classification of Dental Diseases Using CNN and Transfer Learning. *IEEE*, 70–74. <https://doi.org/10.1109/ISCBI.2017.8053547>
- Qi, X., Zhao, G., Shen, L., Li, Q., & Pietikäinen, M. (2016). LOAD: Local orientation adaptive descriptor for texture and material classification. *Neurocomputing*, 184, 28–35. <https://doi.org/10.1016/j.neucom.2015.07.142>
- Rahmad, F., Suryanto, Y., & Ramli, K. (2020). Performance Comparison of Anti-Spam Technology Using Confusion Matrix Classification. *IOP Conference Series: Materials Science and Engineering*, 879(1), 012076. <https://doi.org/10.1088/1757-899X/879/1/012076>
- Rahman, S. A. Z., Mitra, K. C., & Islam, S. M. M. (2018). Soil Classification Using Machine Learning Methods and Crop Suggestion Based on Soil Series. *International Conference of Computer and Information Technology*, 1, 1–4. <https://doi.org/10.1109/ICCITECHN.2018.8631943>
- Rizwan I Haque, I., & Neubert, J. (2020). Deep learning approaches to biomedical image segmentation. *Informatics in Medicine Unlocked*, 18, 100297. <https://doi.org/10.1016/j.imu.2020.100297>
- Roshini, R., & Alex, J. S. R. (2022). Automatic segmentation of optic cup and optic disc using multiresunet for glaucoma classification from fundus image. In M. Saraswat, H. Sharma, & K. V. Arya (Eds.), *Intelligent Vision in Healthcare* (pp. 33–44). Springer Singapore. [https://doi.org/10.1007/978-981-16-7771-7\\_4](https://doi.org/10.1007/978-981-16-7771-7_4)
- Saifullah, S. (2020). K-Means Clustering for Egg Embryo'S Detection Based-on Statistical Feature Extraction Approach of Candling Eggs Image. *Sinergi*, 25(1), 43. <https://doi.org/10.22441/sinergi.2021.1.006>
- Salman, F. M., Abu-Naser, S. S., Alajrami, E., Abu-Nasser, B. S., & Ashqar, B. A. M. (2020). COVID-19 Detection using Artificial Intelligence. *International Journal of Academic Engineering Research*, 4(3), 18–25. <http://dSPACE.alazhar.edu.ps/xmlui/handle/123456789/587>

- Sianturi, F. S., & Agustina, D. H. (2020). Stabilisasi tanah laterit dengan penambahan kapur terhadap kuat geser tanah. *Sigma Teknika*, 3(1), 33–38.
- Soleymani, M., Bonyani, M., Mahami, H., & Nasirzadeh, F. (2021). Construction material classification on imbalanced datasets for construction monitoring automation using Vision Transformer (ViT) architecture. *CoRR*, abs/2108.0.
- Tanzi, L., Audisio, A., Cirrincione, G., Aprato, A., & Vezzetti, E. (2022). Vision Transformer for femur fracture classification. *Injury*, 53(7), 2625–2634. <https://doi.org/10.1016/j.injury.2022.04.013>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems*, 30. <https://doi.org/10.1109/2943.974352>
- Wang, C., Sun, J., Xu, W., & Chen, X. (2019). Depth learning standard deviation loss function. *Journal of Physics: Conference Series*, 1176(3). <https://doi.org/10.1088/1742-6596/1176/3/032050>
- Whipple, A. V., Cobb, N. S., Gehring, C. A., Mopper, S., Flores-Rentería, L., & Whitham, T. G. (2019). Long-term studies reveal differential responses to climate change for trees under soil- Or herbivore-related stress. *Frontiers in Plant Science*, 10(February), 1–14. <https://doi.org/10.3389/fpls.2019.00132>
- Xu, J., Sun, X., Zhang, Z., Zhao, G., & Lin, J. (2019). Understanding and improving layer normalization. *Advances in Neural Information Processing Systems*, 32, 1–11.
- Yang, H., Zhang, X., Xu, M., Shao, S., Wang, X., Liu, W., Wu, D., Ma, Y., Bao, Y., Zhang, X., & Liu, H. (2020). Hyper-temporal remote sensing data in bare soil period and terrain attributes for digital soil mapping in the Black soil regions of China. *Catena*, 184, 104259. <https://doi.org/10.1016/j.catena.2019.104259>
- Yuan, K., Guo, S., Liu, Z., Zhou, A., Yu, F., & Wu, W. (2021). Incorporating Convolution Designs into Visual Transformers. *Proceedings of the IEEE International Conference on Computer Vision*, 559–568. <https://doi.org/10.1109/ICCV48922.2021.00062>
- Yunus, A., Hasibuan, S., & Syafridiman. (2018). Profil Tanah Dasar Kolam Podsolik Merah Kuning (PMK) dengan Umur Berbeda pada Kolam Budidaya Ikan Patin (*Pangasius sp.*) secara Intensif. *Angewandte Chemie International Edition*, 6(11), 951–952.
- Zhai, X., Kolesnikov, A., Houlsby, N., & Beyer, L. (2021). Scaling Vision Transformers. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 12104–12113. <http://arxiv.org/abs/2106.04560>
- Zhu, X., Jia, Y., Jian, S., Gu, L., & Pu, Z. (2021). ViTT: Vision transformer tracker. *Sensors*, 21(16), 1–16. <https://doi.org/10.3390/s21165608>

