



**PENGEMBANGAN ALGORITMA BERBASIS
VISION TRANSFORMER PADA DETEKSI
PENYAKIT TANAMAN KAKAO**

KUALIFIKASI

REZA AL HUSNA

99223134

PROGRAM DOKTOR TEKNOLOGI INFORMASI

UNIVERSITAS GUNADARMA

JUNI 2024

DAFTAR ISI

DAFTAR ISI

BAB 1	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah	5
1.4 Tujuan Penelitian	6
1.5 Kontribusi dan Manfaat Penelitian	6
BAB 2	7
TELAAH PUSTAKA	7
2.1 Pengolahan Citra	7
2.2 Ekstraksi Fitur	10
2.3 Kecerdasan Buatan	13
2.3.1 Deep Learning	13
2.4 Transformer	14
2.4.1 Komponen Utama Transformer	15
2.5 Vision Transformer (ViT)	18
2.5.1 Lapisan-lapisan dalam Vision Transformer (ViT)	19
2.6 Teknik Feature Fusion berbasis Attention	20
2.7 Confusion Matrix	21
2.8 Tanaman Kakao	23

2.9	Penyakit Tanaman Kakao	24
2.10	Perbandingan Tinjauan	27
2.11	Roadmap Penelitian	47
BAB 3	49
METODE PENELITIAN	49
3.1	Tahapan Penelitian.....	49
3.2	Akuisisi Data Penyakit Daun Tanaman Kakao	49
3.3	Pre-Processing	50
3.3.1	Resize Dataset.....	50
3.3.2	Grayscale	50
3.3.3	Augmentasi Dataset	51
3.3.4	Ekstraksi Fitur.....	51
3.3.4.1	Histogram Oriented of Gradients (HoG)	51
3.3.4.2	Local Binary Pattern (LBP)	51
3.3.5	Splitting Data	51
3.3.5.1	Data Training	51
3.3.5.2	Data Testing	52
3.4	Pengembangan dan Pelatihan Model.....	52
3.5	Pengujian dan Evaluasi Model	52
3.6	Jadwal Penelitian	53
Bibliography	55

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Konsep *Artificial Intelligence* (AI) melibatkan pembuatan dan kemajuan sistem komputer yang mampu melakukan tugas-tugas yang umumnya memerlukan kecerdasan manusia, seperti persepsi visual, pengenalan suara, pengambilan keputusan, dan terjemahan antar bahasa (Russel, 2016). Ini mencakup bidang-bidang seperti pengolahan citra, pemrosesan data, *Machine Learning* (ML), *Deep Learning* (DL), dan teknologi visi komputer (Kale, 2019). AI telah menjadi fokus penelitian dan pengembangan teknologi yang menarik dan telah terbukti memberikan dampak signifikan dalam berbagai industri, termasuk dalam industri pertanian (Kumar et al., 2019).

ML merupakan cabang dari AI untuk membantu sistem komputer belajar dari data dan meningkatkan kinerjanya tanpa pemrograman eksplisit (Kamilaris & Prenafeta-Boldú, 2018). Kemajuan terbesar dalam perkembangan AI terdapat dalam bidang DL (Kumar et al., 2019). DL yang merupakan subbidang dari ML, menggunakan jaringan saraf tiruan untuk mengatasi masalah yang sangat kompleks dengan tingkat akurasi yang tinggi. Algoritma DL telah berhasil diterapkan dalam berbagai aplikasi, termasuk pengenalan citra, pemrosesan bahasa alami, dan analisis data besar (LeCun, Bengio, & Hinton, 2015). DL sebagai subbidang dari ML telah membawa revolusi dalam visi komputer dengan memungkinkan pemodelan data yang kompleks dan pengenalan pola yang sangat efisien. Dalam industri pertanian, DL di dalam tugas visi komputer dapat memberikan solusi cerdas untuk melakukan pemantauan, prediksi, maupun otomatisasi dalam berbagai masalah pada tanaman, salah satunya dalam mendeteksi penyakit tanaman.

Selama beberapa dekade terakhir, pertanian telah menjadi sumber pendapatan utama bagi beberapa negara dan memberikan kontribusi signifikan terhadap perekonomian global (Parez et al., 2023). Tanaman Kakao (*Theobroma Cacao*) merupakan sebuah komoditas penting dunia. Indonesia sebagai salah satu negara eksportir utama kakao dalam perdagangan internasional. Indonesia setiap tahun mampu memproduksi biji kakao sekitar 700.000 ton. Namun, produksi kakao Indonesia terus menurun dalam tiga tahun terakhir. Tahun 2022, Indonesia memproduksi kakao mencapai 650.600 ton biji kakao. Produksi tersebut menurun 3,04% dibandingkan produksi tahun 2021 sebanyak 688.200 ton, Sementara produksi pada tahun 2021 turun hampir tiga persen dibanding tahun sebelumnya tercatat 720.660 ton (Badan Pusat Statistik Indonesia, 2023). Kendala dalam menjaga produksi kakao di Indonesia perlu didukung dengan berbagai program untuk mengantisipasi penurunan produksi. Penurunan produksi kakao memiliki dampak meluas (Balai Informasi Standar Instrumen Pertanian, 2023). Beberapa faktor berkontribusi terhadap penurunan produksi ini diantaranya disebabkan oleh banyaknya pohon kakao yang sudah tua, kurang produktif, dan kurangnya pemeliharaan intensif juga menyebabkan sebagian besar tanaman kakao terserang penyakit. Penyakit pada kakao menyebabkan penurunan produksi kakao yang cukup besar yaitu sekitar 50% (Arnawa et al., 2012). Penyakit kakao merupakan salah satu faktor yang berkontribusi terhadap penurunan biji-bijian secara global serta dapat mempengaruhi pertumbuhan, kesehatan, dan produktivitas tanaman, sehingga dapat merusak hasil panen.

Beberapa penyakit yang sering menyerang tanaman kakao adalah penyakit kanker batang, penyakit *Vascular Streak Dieback* (VSD), *Cocoa Swollen Shoot Virus Disease* (CSSVD), *Leaf Blotch*, *Witches Broom Diseases* (WBD), *Antraknosa*, penyakit penggerek buah kakao, penyakit busuk buah, penyakit hawar benang (CABI Plantwise, 2019). Penyakit tersebut merupakan penyakit yang paling merugikan, karena terdapat hampir di seluruh area penanaman kakao. Gejala yang dialami pada tanaman dapat bervariasi tergantung pada jenis penyakit, jenis tanaman yang terinfeksi,

dan kondisi lingkungan. Beberapa gejala pada tanaman kakao yang terkena penyakit diantaranya bercak atau lesi pada daun, daun menguning atau rontok, pertumbuhan terhambat atau kerusakan pada vena, keropeng atau deformitas pada buah, bunga tidak berkembang, batang dan akar mengalami pembusukan, produksi yang menurun, dan kematian pada tanaman (Sari A, 2023). Namun, banyak penyakit tanaman kakao memiliki kesamaan gejala tersebut, sehingga sulit untuk mengidentifikasi jenis penyakit, dan akurasi pengenalan penyakit mungkin rendah pada kondisi alami (Patil dan Kumar, 2022). Karena itu, penelitian ini mengusulkan metode ML dan DL dalam tugas visi komputer untuk melakukan identifikasi dalam deteksi penyakit tanaman kakao yang lebih akurat.

Banyak penelitian yang memanfaatkan ML dan DL untuk tugas visi komputer dalam mengenal penyakit tanaman. Secara umum, penelitian-penelitian sebelumnya menggunakan model seperti *Support Vector Machine* (SVM) dari ML digabungkan dengan metode ekstraksi fitur dan *Convolutional Neural Network* (CNN) dari ML. Pada penelitian (Rodriguez, Alfaro, Paredes, Esenarro, & Hilarion, 2021), digunakan HoG, LBP untuk ekstraksi fitur dan SVM untuk klasifikasi penyakit pada tanaman. Penelitian (Harakannanavar, Rudagi, Puranikmath, Siddiqua, & Pramodhini, 2022), digunakan *Discrete Wavelet Transform* (DWT), *Principal Component Analysis* (PCA), dan *Gray Level Co-occurrence Matrix* (GLCM) untuk ekstraksi fitur lokal dan model SVM, K-Nearest Neighbors (KNN), dan CNN melakukan klasifikasi pada penyakit daun tanaman. Penelitian (Rachmad, Syarief, Rifka, Sonata, Setiawan, & Rochman, 2022), digunakan LBP untuk mengekstraksi fitur dan KNN untuk klasifikasi penyakit daun jagung. Penelitian (Hosny, El-Hady, Samy, Vrochidou, & Papakostas, 2023), digunakan LBP dan CNN untuk mendeteksi penyakit daun tanaman. Selain dari model yang sudah disebutkan diatas, terdapat salah satu inovasi terbaru model DL yang membawa kemajuan signifikan yaitu *Vision Transformer* (ViT), yang memperkenalkan pendekatan baru untuk pengenalan citra. ViT mengadopsi konsep transformer yang awalnya dikembangkan untuk tugas-tugas *Natural Language Processing* (NLP).

Transformer menggunakan *attention mechanism* untuk fokus pada bagian-bagian penting dari input data, sehingga meningkatkan kemampuan model dalam memahami konteks yang kompleks (Vaswani et al., 2017). Dalam bidang pertanian, ViT diterapkan untuk memantau pertumbuhan tanaman dan mendeteksi penyakit tanaman, yang membantu petani dalam mengambil keputusan yang lebih tepat dan cepat.

Berbeda dengan beberapa penelitian diatas, penelitian ini mengadopsi pendekatan *feature fusion* berbasis *attention*. *Feature fusion* berbasis *attention* merupakan metode yang menggabungkan fitur-fitur yang diekstraksi dengan *attention mechanism* untuk meningkatkan kinerja model. *Attention mechanism* memberikan kemampuan pada model untuk fokus pada bagian-bagian penting dari input, sehingga meningkatkan representasi fitur yang lebih relevan (Hu et al., 2018).

Penelitian (Borhani et al., 2022) mengeksplorasi model ViT dalam pengklasifikasian penyakit tanaman. Menggabungkan blok konvolusional dengan blok *attention* transformer membantu model mendapatkan akurasi yang tinggi. Hasil akurasi dari model hybrid yang diusulkan masing-masing dataset mencapai 100% pada dataset daun padi, 91,7% pada dataset karat gandum, dan 98,77% dataset *PlantVillage*. Penelitian (Lye dan Ng, 2023) melakukan klasifikasi penyakit tanaman menggunakan ViT. Penelitian ini berfokus identifikasi pada keseluruhan daun dibandingkan bagian individual atau bagian tertentu. Mereka menggunakan modul *attention mechanism Locality Self Attention* (SLA) untuk membantu fokus dalam identifikasi keseluruhan daun. Hasil akurasi yang diperoleh adalah 89,58%, serta ketahanan model yang memuaskan terhadap perubahan orientasi dan posisi daun. (Tabbakh & Barpanda, 2023) mengeksplorasi pendekatan *feature fusion* yaitu menggabungkan model transfer learning dengan model ViT yang diberi nama TLMViT untuk klasifikasi penyakit daun tanaman. Penelitian ini melakukan dua fase ekstraksi fitur, yaitu ekstraksi fitur awal menggunakan model *pre-trained* dan ekstraksi fitur mendalam dilakukan menggunakan model ViT. Model diklasifikasikan menggunakan MLP dan diuji dengan lima model *pre-trained* dan ViT. Model yang diusulkan pada penelitian ini dapat

mengungguli model transfer learning lainnya dengan peningkatan akurasi sebesar 1,11% dan 1,099%.

Berdasarkan uraian latar belakang dan analisis terhadap penelitian-penelitian sebelumnya, Model transformer mampu mencatat bagian-bagian tertentu yang berbeda secara bersamaan dan memahami keterkaitan bagian tersebut, yang sangat penting untuk pengenalan fitur penyakit (Gao R et al., 2024). Pada penelitian (Jesse, douha, dan Lenka, 2024) mengidentifikasi penyakit tanaman kakao menggunakan ViT masih memiliki keterbatasan model dalam variasi latar belakang yang kompleks serta mengandung noise yang tinggi. Selain itu, penelitian tersebut masih terbatas hanya pada satu kelas penyakit saja. Oleh karena itu, masih terbuka peluang untuk mengembangkan metode menggunakan teknik *feature fusion* berbasis *attention mechanism* untuk mengoptimalkan model dengan menggabungkan ekstraksi fitur dan ViT, serta menambahkan variasi pada dataset yang bertujuan untuk meningkatkan ketahanan dan generalisasi pada model.

1.2 Rumusan Masalah

Berdasarkan yang telah diuraikan pada latar belakang masalah diatas, maka direncanakan perumusan masalah yang mendasari penelitian sebagai berikut.

1. Bagaimana mengembangkan algoritma dan metode untuk identifikasi penyakit daun kakao agar lebih baik menggunakan pendekatan *feature fusion* berbasis *attention mechanism*?
2. Bagaimana membangun prototype sistem deteksi yang tepat dan akurat untuk identifikasi penyakit pada daun kakao agar mempermudah pengguna dalam mengenali daun yang terkena penyakit?

1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Fitur yang dianalisis dalam penelitian ini adalah bentuk daun kakao yang menunjukkan gejala penyakit.
2. Dataset yang digunakan dalam penelitian ini terdiri dari 4 penyakit daun kakao, yaitu: penyakit antraknosa (*Colletotrichum gloeosporioides*), penyakit *vascular streak dieback* (VSD), penyakit *Leaf Blotch* dan penyakit *cocoa swollen shoot virus disease* (CSSVD).

1.4 Tujuan Penelitian

Tujuan penelitian yang ingin dicapai pada penelitian ini adalah:

1. Menghasilkan algoritma dan metode pada identifikasi penyakit daun kakao dengan pendekatan *feature fusion* berbasis *attention mechanism* untuk mendapatkan hasil yang lebih akurat dalam mengenali jenis penyakit daun kakao.
2. Menghasilkan sebuah sistem deteksi yang mampu secara efektif mengenali dan mengidentifikasi penyakit pada daun kakao yang terinfeksi, dengan tujuan untuk meningkatkan akurasi dan keandalan dalam proses diagnosis secara lebih efisien.

1.5 Kontribusi dan Manfaat Penelitian

1. Kontribusi kebaruan ilmu: Pengembangan algoritma yang menerapkan pendekatan *feature fusion* dengan *attention mechanism* untuk menciptakan model Vision Transformer (ViT) yang efektif dalam mendeteksi penyakit pada tanaman kakao.
2. Kontribusi dari sisi teknologi: Menghasilkan sebuah prototype sistem yang dapat digunakan untuk mendeteksi penyakit pada tanaman kakao secara efektif.
3. Kontribusi dari sisi agrikultural: Dapat membantu para petani untuk mengenali gejala penyakit pada tanaman dengan lebih cepat dan efektif, yang pada gilirannya dapat berdampak positif pada pertumbuhan tanaman kakao.

BAB 2

TELAAH PUSTAKA

2.1 Pengolahan Citra

2.1.1 Definisi Citra

Citra adalah representasi dari objek mengandung informasi visual yang dapat diinterpretasikan oleh manusia maupun mesin untuk berbagai tujuan seperti dokumentasi, analisis, atau pemantauan. Informasi dalam citra dapat dibagi menjadi dua jenis: informasi umum dan informasi abstrak. Informasi umum merupakan hasil pengolahan tanpa bantuan perangkat khusus, seperti warna, bentuk, dan tekstur. Sedangkan informasi abstrak merupakan hasil pengolahan dengan bantuan perangkat pengetahuan khusus, seperti penyakit tanaman yang menunjukkan gejala seperti bercak lesi, kerusakan pada vena, pembengkakan serta gejala tertentu lainnya.

Secara matematis, sebuah citra dapat didefinisikan ke dalam sebuah fungsi dua dimensi $f(x, y)$, dimana x dan y merupakan koordinat spasial dan f merupakan nilai intensitas warna pada koordinat (x, y) , f dapat dikatakan piksel dari sebuah citra yang dapat dikodekan dalam 24-bit citra berwarna dimana 3 komponen utamanya yaitu R (Red), G (Green), B (Blue), 8 bit untuk citra *gray-level* atau 1 bit untuk citra biner. (Madenda, 2015).

$$F = [f(x, y)] = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,M) \\ \vdots & \vdots & \dots & \vdots \\ f(N,0) & f(N,1) & \dots & f(N,M) \end{bmatrix} \quad (2.1)$$

2.1.2 Citra Berwarna

Citra berwarna adalah citra yang secara visual memiliki kandungan informasi warna, dimana warna direpresentasikan dalam nilai-nilai piksel yang mengandung komponen *luminance*, *hue* dan *chrominance/saturation*. *Luminance* merupakan ukuran Tingkat kecerahan suatu warna. *Hue* adalah salah satu sifat utama warna yang

direpresentasikan dalam nilai derajat (0° - 360°). Warna-warna dasar seperti warna merah memiliki nilai hue 0° atau 360° , hijau memiliki nilai hue 120° , dan biru memiliki nilai hue 240° . *Chrominance* merepresentasikan tinggi rendahnya kandungan Cahaya putih dalam sebuah warna. Semakin rendah nilai *chrominance* maka warna semakin memucat (memutih), sebaliknya semakin tinggi nilai *chrominance* maka warna semakin mendekati saturasi (*pure color*) (Madenda, 2015).

Citra warna adalah citra yang masing-masing piksel mempunyai 3 komponen warna spesifik, yaitu komponen R (Red), G (Green) dan B (Blue). Warna setiap piksel ditentukan oleh kombinasi dari intensitas warna *red*, *green*, dan *blue* yang disimpan pada bidang warna di Lokasi piksel. Format file grafis menyimpan citra warna sebagai 24 bit, yang berasal dari komponen *red*, *green* dan *blue* masing-masing 8 bit. Hal ini menyebabkan citra warna mempunyai 24 juta kemungkinan warna (Andono et al., 2017).



Gambar 2.1. Citra warna RGB (Madenda, 2015)

2.1.3 Citra Grayscale

Citra *Gray level* (skala keabuan) adalah citra dimana nilai pikselnya hanya diwakilkan oleh nilai *luminance*, yang umumnya dikodekan dalam 8 bit didefinisikan memiliki skala keabuan yang bervariasi dari nilai 0 sampai 255 (2^8-1). Nilai 0 merepresentasikan warna hitam, sedangkan nilai 255 merepresentasikan warna putih.

Diantara nilai tersebut merepresentasikan warna keabuan yang bervariasi dari hitam cerah menuju putih (Madenda, 2015).



Gambar 2.2. Citra Grayscale (Madenda, 2015)

2.1.4 Citra Biner

Citra biner merupakan bagian dari citra *grayscale* yang hanya memiliki dua level keabuan, yaitu 0 untuk warna hitam dan 1 untuk warna putih, sehingga setiap piksel dari citra biner dikodekan hanya menggunakan 1 bit. Citra biner dihitung dengan menggunakan nilai ambang (*threshold*) bila nilai piksel lebih kecil daripada batas ambang maka nilai tersebut diubah menjadi 0 hitam, sedangkan bila lebih besar atau sama dengan nilai ambang maka nilai tersebut diubah menjadi 1 putih (Madenda, 2015).



Gambar 2.3. Citra Biner (Madenda, 2015)

2.2 Ekstraksi Fitur

2.2.1 Histogram of Oriented Gradients

Histogram of Oriented Gradient (HOG) adalah sebuah teknik dalam pengolahan citra yang memanfaatkan distribusi intensitas gradien dari citra untuk mengenali objek. Pendekatan HOG melibatkan perhitungan gradien di berbagai wilayah citra. Setiap citra memiliki karakteristik yang tercermin dalam nilai gradiennya. Proses ini melibatkan pembagian citra ke dalam sel (cell), dimana masing-masing sel membentuk histogram gradien, dan kemudian sel-sel ini digabungkan menjadi blok (block). Blok histogram ini kemudian dinormalisasi. Setiap piksel dalam sel menghasilkan kontribusi pada histogram berdasarkan nilai gradiennya. Histogram ini mencerminkan distribusi derajat keabuan piksel dalam citra. Gabungan dari histogram ini berfungsi sebagai deskriptor yang merepresentasikan objek tersebut.

Pertama untuk setiap piksel dari citra akan dihitung gradiennya dari sumbu x dan y dengan menggunakan persamaan:

$$Sx(x, y) = I(x + 1, y) - I(x - 1, y) \quad (2.2)$$

$$Sy(x, y) = I(x, y + 1) - I(x, y - 1) \quad (2.3)$$

Rumus diatas menjelaskan bahwa gradient didapat dengan mengurangi nilai piksel di depannya dengan nilai piksel di belakangnya sesuai dengan sumbunya. Setelah didapat nilai gradien dari sumbu x dan y untuk setiap pixelnya, proses selanjutnya adalah menghitung besar nilai dan arah gradiennya dengan menggunakan persamaan 2.x

$$M(x, y) = \sqrt{Sx(x, y)^2 + Sy(x, y)^2} \quad (2.4)$$

$$\theta(x, y) = \tan^{-1} \frac{Sy(x, y)}{Sx(x, y)} \quad (2.5)$$

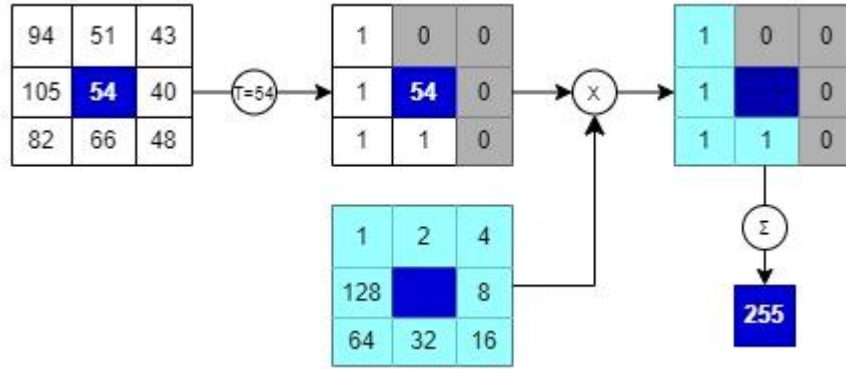
Nilai magnitude gradien dan sudut gradien dianggap sebagai elemen kunci dalam langkah-langkah berikutnya. Sx dan Sy mengindikasikan gradien pada sumbu x dan y secara berturut-turut. Proses berikutnya membagi setiap piksel dalam citra

menjadi beberapa sel, di mana setiap sel dihitung distribusi HoG-nya melalui proses vote. Langkah pertama dalam proses vote dalam HoG adalah menetapkan nilai-nilai bin dengan membagi total sudut gradien ke dalam jumlah bin orientasi yang telah ditetapkan sebelumnya. Kemudian, setiap arah sudut gradien dari setiap piksel dalam sel akan ditempatkan dalam rentang bin orientasi yang telah ditentukan sebelumnya, dan kemudian nilai magnitudenya akan dibagi oleh bin orientasi terkait. Setelah HOG dibuat untuk setiap sel, langkah selanjutnya adalah normalisasi terhadap hasil vote pada setiap bin dalam sel. Normalisasi dilakukan dalam blok yang ditentukan dengan ukuran $m \times n$ sel. Ada empat metode normalisasi yang berbeda: L2-Norm, L2-Hys, L1-sqrt, dan L1-norm.

Persamaan normalisasi dapat dijelaskan sebagai berikut: Proses normalisasi blok dilakukan dengan menggunakan sliding window yang akan bergerak dengan pergeseran sebesar 1x ukuran sel secara vertikal dan horizontal. Proses ini memiliki tumpang tindih untuk beberapa sel yang dinormalisasi, sehingga memberikan informasi yang redundan, namun dapat meningkatkan akurasi. Akhirnya, hasil normalisasi dari setiap blok digabungkan menjadi satu vektor fitur besar.

2.2.2 Local Binary Pattern

Local Binary Patterns (LBP) adalah algoritma ekstraksi fitur yang mengekstrak fitur statistik dan structural fitur tekstur gambar. Metode LBP awalnya diperkenalkan oleh Timo Ojala. Cara kerja operator LBP melibatkan perbandingan nilai keabuan dari piksel-piksel yang berdekatan. LBP adalah metode ekstraksi fitur tekstur yang bersifat *rotation invariant*. Nilai LBP diperoleh dari proses *thresholding*, kemudian nilai tersebut dikalikan dengan bobot biner. Bentuk dasar LBP menggunakan nilai piksel tengah sebagai ambang batas untuk piksel-piksel tetangganya dalam matriks berukuran 3x3. Proses ini menghasilkan pola biner yang menggambarkan karakteristik tekstur (Mujib et al., 2018; Indrawati et al., 2018; Rachmad A et al., 2022).



Gambar 2.4 Tahapan Perhitungan LBP (Prakasa Esa, 2015)

Versi dasar dari operator LBP menggunakan nilai piksel pusat sebagai acuan untuk menentukan *threshold* pada piksel tetangga dalam area 3×3 . Proses penentuan *threshold* akan menghasilkan pola biner yang menggambarkan karakteristik tekstur. Persamaan dasar dari LBP dapat dinyatakan sebagai berikut:

$$LBP(x_c, y_c) = \sum_{n=0}^7 2^n g(I_n - I(x_c, y_c)) \quad (2.6)$$

$LBP(x_c, y_c)$ Merupakan nilai LBP pada piksel pusat (x_c, y_c) . I_n dan $I(x_c, y_c)$ merujuk pada nilai piksel tetangga dan nilai pusat. Indeks n menunjukkan indeks dari piksel tetangga. Fungsi $g(x)$ bernilai nol jika $x < 0$ dan $g(x)$ bernilai 1 jika $x \geq 0$. Gambar 2.4, piksel pusat dengan nilai 54 akan dipilih sebagai *threshold*. Piksel tetangga akan diberi nilai 0 jika nilainya lebih kecil dari *threshold*. Sebaliknya, jika nilai piksel tetangga lebih besar atau sama dengan *threshold*, maka akan diberi nilai 1. Nilai LBP dihitung dengan melakukan perkalian skalar antara matriks biner dan bobot. Akhirnya, hasil perkalian tersebut dijumlahkan untuk merepresentasikan nilai LBP. Nilai LBP dari matriks 3×3 pada gambar 2.4 adalah $2^0 + 2^5 + 2^6 + 2^7 = 1 + 32 + 64 + 128$, yang sama dengan 225 (Perkasa Esa, 2015).

2.3 Kecerdasan Buatan

Artificial Intelligence (AI) diartikan sebagai kemampuan sistem untuk menafsirkan data eksternal dengan benar, untuk belajar dari data tersebut, dan menggunakan pembelajaran tersebut untuk mencapai tujuan dan tugas tertentu melalui adaptasi yang fleksibel. Kecerdasan dibuat dan dimasukkan ke dalam suatu mesin (komputer) agar dapat melakukan pekerjaan seperti yang dilakukan manusia. Penerapan kecerdasan buatan antara lain sistem pakar, logika fuzzy, jaringan saraf tiruan dan robotika (Awangga dan Batubara, 2020).

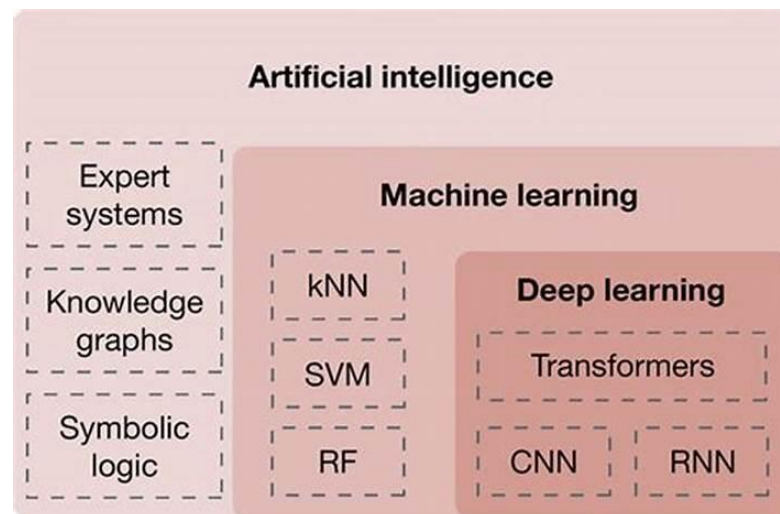
AI dimunculkan oleh seorang professor dari Massachusetts Institute of Technology yang bernama John McCarthy pada tahun 1956 pada Dartmouth Conference. Tujuan utama konferensi tersebut didefinisikan kecerdasan buatan, yaitu mengetahui dan memodelkan proses-proses berpikir manusia dan mendesain mesin agar dapat menirukan kelakuan manusia tersebut (Wirawan, 2017). Konsep dan teknologi kecerdasan buatan disempurnakan oleh Alan Turing dengan meneliti dan menguji coba algoritma AI yang diberi nama “Turing Test” (Jamaluddin dan Sulistyowati, 2022).

AI adalah cabang ilmu komputer yang bertujuan untuk mengembangkan sistem dan mesin yang mampu melakukan tugas yang biasanya memerlukan AI. AI melibatkan penggunaan algoritma dan model matematika untuk memungkinkan komputer dan sistem lainnya untuk belajar dari data, mengenali pola, dan membuat keputusan yang cerdas (Eriana dan afrizal, 2023).

2.3.1 Deep Learning

Deep Learning (DL) merupakan sebuah sub-bidang dari *machine learning* yang menggunakan arsitektur jaringan syaraf tiruan (*Neural Network*) yang terdiri dari lapisan-lapisan sebagai model dan mempelajari representasi data yang kompleks. Pembelajaran mendalam melibatkan rangkaian lapisan yang berurutan, biasanya dalam jumlah puluhan hingga ratusan, yang secara otomatis mempelajari pola dari data yang

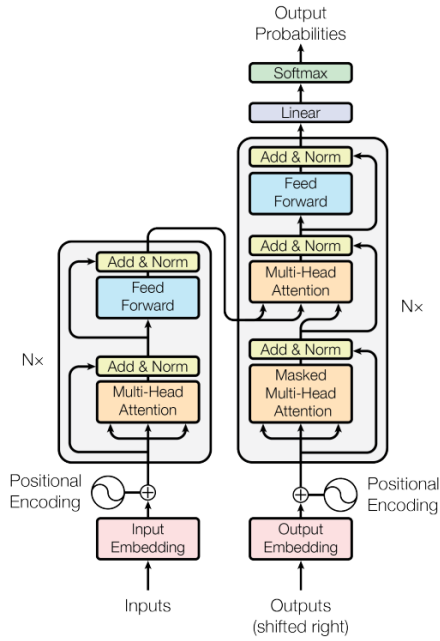
diberikan. Lapisan-lapisan *Neural Network* ini memiliki struktur bertumpuk, di mana setiap lapisan ditempatkan di atas lapisan lainnya (Chollet, 2018). Beberapa jenis algoritma deep learning yang telah banyak digunakan adalah *Convolutional Neural Network* (CNN), *Recurrent Neural Networks* (RNNs), dan *Generative Adversarial Networks* (GANs), Autoencoders, dan Transformer.



Gambar 2.5 Cabang Ilmu Kecerdasan Buatan (Cau et al., 2023)

2.4 Transformer

Transformer adalah arsitektur jaringan saraf yang diperkenalkan pada tahun 2017 oleh para peneliti dari Google Brain dalam makalah berjudul "*Attention Is All You Need*." Transformer telah merevolusi berbagai aplikasi dalam pemrosesan bahasa alami (NLP) dan, baru-baru ini, dalam visi komputer. Transformer merevolusi *Natural Language Processing* (NLP) dan berbagai aplikasi lainnya dengan menghilangkan ketergantungan pada arsitektur RNN dan *Long Short-Term Memory* (LSTM). Arsitektur Transformer dapat dilihat pada Gambar 2x



Gambar 2.6 Model arsitektur Transformer

Sumber: Vaswani et al 2017

2.4.1 Komponen Utama Transformer

Komponen utama transformer meliputi *multi-head self-attention mechanism* dan struktur layer encoder-decoder antara lain:

1. Positional Encoding

Positional encoding menambahkan informasi posisi ke dalam input embedding karena transformer tidak memiliki struktur sekuensial intrinsik.

$$PE_{(Pos, 2i)} = \sin\left(\frac{pos}{10000^{2i/d_{\text{model}}}}\right) \quad (2.7)$$

$$PE_{(Pos, 2i+1)} = \cos\left(\frac{pos}{10000^{2i/d_{\text{model}}}}\right) \quad (2.8)$$

Dimana pos adalah posisi token dalam urutan dan i adalah indeks dimensi

2. Self-Attention Mechanism

Self-attention mechanism memungkinkan model untuk fokus pada bagian-bagian tertentu dari setiap posisi output. Dalam decoder, output sebelumnya dikompresi ke dalam sebuah *Query* (Q dari m dimensi) dan output teksnya dihasilkan oleh pemetaan *Query* dan *Keys* dan *Values* (Weng, 2018).

$$\text{attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (2.9)$$

Dimana Q (Query), K (Key), dan V (Value) adalah matriks yang diperoleh dari input embedding melalui transformasi linier, dan d_k adalah dimensi dari key.

3. Multi-Head Attention

Multi-head attention menggunakan beberapa perhatian parallel (head) yang digabungkan untuk menangkap berbagai jenis informasi.

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O \quad (2.10)$$

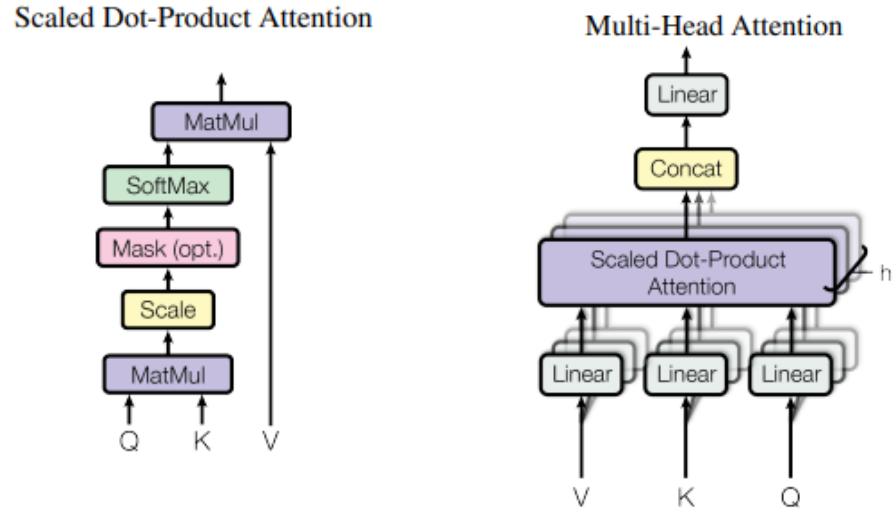
Dimana setiap head dihitung sebagai:

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (2.11)$$

Dimana QW_i^Q, KW_i^K, VW_i^V , dan W^O adalah matrik bobot yang berbeda untuk setiap head.

Transformer mengadopsi *scaled dot-product attention*: outputnya adalah jumlah nilai yang dibobotkan, di mana bobot yang ditetapkan untuk setiap nilai ditentukan oleh *dot-product* dari query dengan semua *keys*: *Multi-head self-attention mechanism* digunakan untuk mempelajari ketergantungan antara token-token dalam urutan data. *Multi-head self-attention mechanism* beroperasi dengan melakukan perhitungan *scaled dot-product attention* secara parallel berulang kali. *Output independent*

attention hanya digabungkan dan diubah secara linier menjadi dimensi yang diharapkan (Weng, 2018).



Gambar 2.7 (Kiri) Scaled Dot-Product Attention. (Kanan) Multi-Head Attention

Sumber: Vaswani et al 2017

4. Feed-Forward Neural Network (FFN)

Setiap posisi dalam urutan diproses melalui jaringan feed-forward yang sama setelah *attention mechanism*.

$$\text{FFN}(x) = \max(0, xW_1 + b_1)W_2 + b_2 \quad (2.12)$$

Dimana W_1, W_2 adalah matriks bobot, dan b_1, b_2 adalah bias

5. Layer Normalization

Layer normalization digunakan untuk menstabilkan dan mempercepat pelatihan dengan menormalkan output dari setiap lapisan sebelum dilanjutkan ke lapisan berikutnya.

$$\text{LayerNorm}(x) = \frac{x - E[x]}{\sqrt{\text{Var}[x] + \epsilon}} \cdot \gamma + \beta \quad (2.13)$$

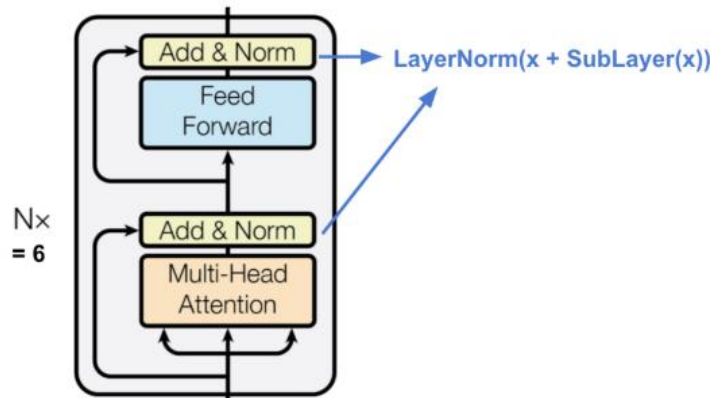
Dimana $E[x]$ adalah rata-rata, $\text{Var}[x]$ adalah variasi, ϵ adalah konstanta kecil untuk stabilitas numerik, dan γ, β adalah parameter yang dapat dilatih.

6. Residual Connections

Residual connections menambahkan input asli ke output dari sub-lapisan (self-attention atau FFN) untuk memastikan aliran gradien yang lebih baik dan stabilitas pelatihan.

$$\text{Output} = x + \text{SubLayer}(x) \quad (2.14)$$

Dimana x adalah input asli dan $\text{SubLayer}(x)$ adalah output dari sub-lapisan.



Gambar 2.8 Encoder Transformer

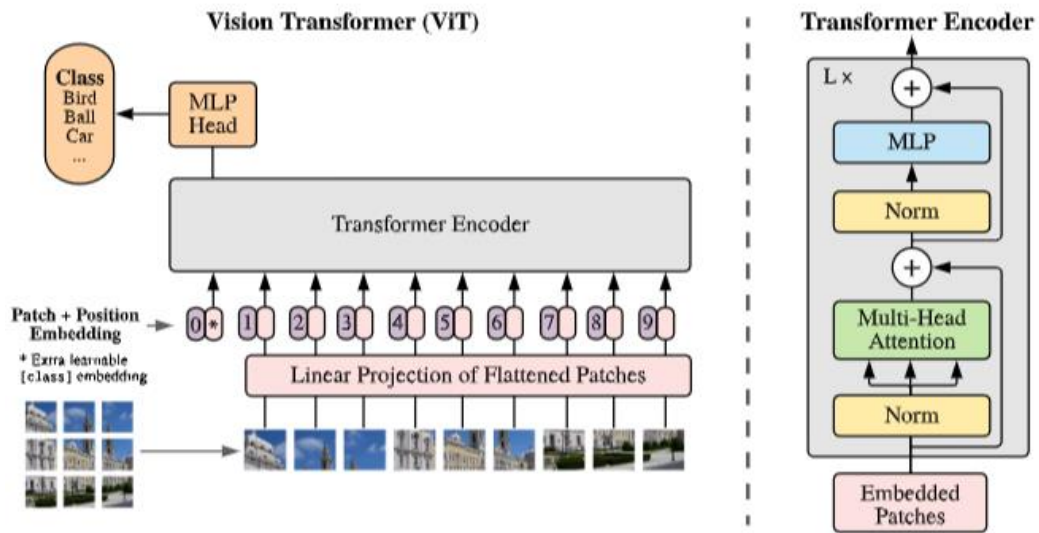
Sumber: Vaswani et al., 2017

2.5 Vision Transformer (ViT)

Vision Transformer (ViT) merupakan model DL yang menggunakan arsitektur Transformer, yang awalnya dikembangkan untuk NLP, untuk tugas-tugas visi komputer seperti pengenalan citra dan klasifikasi citra. ViT menggantikan lapisan konvolusi tradisional dengan *self-attention mechanism* untuk menangkap informasi global dalam citra melalui pembagian citra menjadi patch-patch kecil yang kemudian diolah oleh lapisan Transformer (Dosovitskit et al., 2020). ViT muncul sebagai algoritma canggih yang menggantikan atau menggabungkan teknik tradisional seperti CNN. ViT merupakan jenis *Neural Network* untuk pemrosesan citra dalam visi

komputer. *Backbone* ViT adalah *self-attention mechanism* yang biasanya digunakan dalam NLP (Khalid, Fayez, Awos & Ilamparithi, 2023).

ViT memisahkan citra menjadi token visual, berbeda dengan model berbasis konvolusi, serangkaian lapisan konvolusi dan pengumpulan digunakan untuk mengekstrak fitur dari citra. ViT menggunakan pendekatan berbasis *attention*, yang memungkinkan model untuk mempertimbangkan konteks lokasi manapun dalam urutan input.



Gambar 2.9 Transformer Encoder

Sumber: Dosovitskiy et al., 2021

2.5.1 Lapisan-lapisan dalam Vision Transformer (ViT)

1. Patch Embedding

Citra di input dengan ukuran $H \times W \times C$ dibagi menjadi patch-patch kecil berukuran tetap $P \times P$, dan setiap patch diratakan (flattened) menjadi vector satu dimensi $P^2 \times C$. Vektor diproyeksikan ke ruang embedding menggunakan lapisan linier.

2. Positional Encoding

Positional encoding ditambahkan ke setiap vector embedding untuk memberikan informasi posisi. Positional encoding pada ViT menggunakan rumus yang sama seperti dalam transformer untuk teks.

3. Transformer Encoder

Citra yang telah diubah menjadi serangkaian embedding patch diproses melalui beberapa lapisan encode transformer, terdiri dari:

$$\text{EncoderLayer}(x) = \text{LayerNorm}(x + \text{MultiHeadSelfAttention}(x)) \quad (2.15)$$

$$\text{EncoderLayer}(x) = \text{LayerNorm}(x + \text{FFN}(X)) \quad (2.16)$$

4. Classification Token (CLS Token)

Token khusus (CLS) ditambahkan ke urutan patch sebelum diproses oleh Transformer. Representasi akhir dari token digunakan untuk prediksi klasifikasi.

5. Output Layer

Output dari token (CLS) diberikan ke lapisan dense akhir yang menghasilkan prediksi kelas.

2.6 Teknik Feature Fusion berbasis Attention

Attention-based feature fusion adalah teknik dalam DL yang mengintegrasikan konsep *feature fusion* dengan *attention mechanism* untuk meningkatkan kinerja model dalam berbagai tugas, seperti klasifikasi citra, segmentasi, dan deteksi objek. *Attention mechanism* memberikan kemampuan kepada model untuk memfokuskan perhatian pada bagian-bagian penting dari data input, sehingga menghasilkan representasi fitur yang lebih relevan (Hu et al., 2018).

Dalam proses ini, fitur dari berbagai sumber atau teknik diekstraksi dan digabungkan untuk menciptakan representasi yang lebih kaya. Proses penggabungan ini bisa melibatkan kombinasi fitur dari berbagai lapisan dalam jaringan neural atau dari beragam model. *Attention mechanism* kemudian digunakan untuk memberikan

bobot berbeda pada fitur-fitur yang digabungkan, sesuai dengan relevansinya terhadap tugas yang sedang dihadapi.

Attention mechanism membantu model untuk fokus pada fitur-fitur yang lebih penting dan mengabaikan yang kurang relevan. Beberapa manfaat *Attention-based feature fusion* diantaranya memberikan peningkatan akurasi dan model dapat membuat prediksi lebih akurat, mengurangi kompleksitas sehingga model menjadi lebih efisien serta generalisasi yang lebih baik dengan fokus pada fitur-fitur yang paling signifikan. Modul *feature fusion* bertanggung jawab untuk menggabungkan fitur-fitur yang diekstraksi dari model encoding informasi lokal dan global secara organik, yang bertujuan untuk menghasilkan representasi citra yang lebih representative (Chen Z et al., 2024).

2.7 Confusion Matrix

Confusion Matrix merupakan suatu metode yang dapat digunakan untuk mengukur kinerja suatu model klasifikasi. Confusion matrix mengandung informasi yang membandingkan hasil klasifikasi yang dilakukan oleh sistem untuk diukur keakuratannya (Rahman, dkk., 2017).

Tabel 2.1 *Confusion Matrix*

		Prediksi	
		Positif	Negatif
Aktual	Positif	True Positive	False Negative
	Negatif	False Positive	True Negative

Pada tabel Confusion Matrix diatas TP dan TN menyatakan classifier mengenali dengan benar. Sebaliknya FP dan FN menyatakan bahwa classifier salah mengenali dengan baik.

Perhitungan Akurasi, Presisi dan Recall yang dilakukan oleh confusion matrix berdasarkan Tabel diatas dapat menggunakan persamaan sebagai berikut :

1. Akurasi

Merupakan persentase jumlah data yang dilakukan pada klasifikasi atau prediksi secara benar oleh algoritma. Berikut rumus dari akurasi:

$$\text{Akurasi} = \frac{TP + TN}{TP + FN + FP + TN} \times 100\% \quad (2.17)$$

2. Presisi

Mengukur tingkat keakuratan model saat mendeteksi objek. Precision merupakan perbandingan antara *True Positive* (TP) atau banyaknya objek yang terdeteksi benar dan tepat oleh model dengan keseluruhan pendeteksian objek (TP dan FP), *False Positive* (FP) atau banyaknya model mendeteksi objek yang salah namun tepat seperti persamaan 1 (Carlos santos et al., 2022):

$$\text{Presisi} = \frac{TP}{TP + FP} \times 100\% \quad (2.18)$$

3. Recall

Recall mengukur banyaknya kekeliruan objek yang terdeteksi oleh model dan mengharuskan model untuk terbebas dari kesalahan saat mendeteksi suatu objek. *False Negative* (FN) atau banyaknya model mendeteksi objek yang salah dan tidak tepat. Nilai recall semakin rendah saat semakin banyaknya objek yang terdeteksi salah oleh model dan tidak tepat. Matriks ini membandingkan banyaknya objek yang terdeteksi benar dan tepat dengan seluruh objek yang dideteksi (deteksi objek yang benar dan tepat juga deteksi yang salah dan tidak tepat) seperti persamaan 2 (Nariman Mamdouh, 2021).

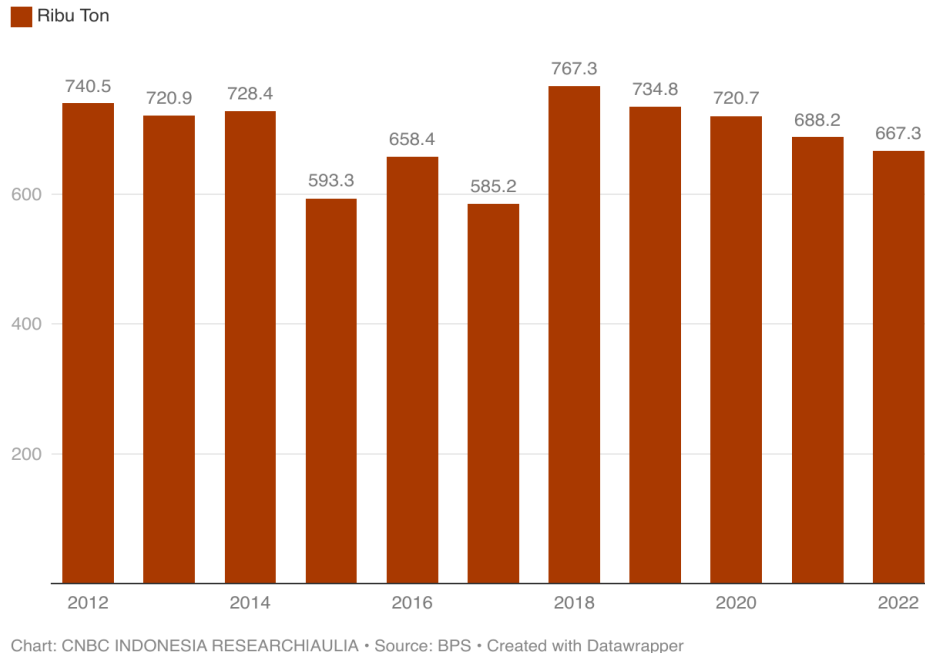
$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% \quad (2.19)$$

2.8 Tanaman Kakao

Tanaman kakao (*Theobroma cacao L.*) berasal dari hutan-hutan tropis di Amerika Tengah dan di Amerika Selatan bagian Utara. Tanaman Kakao atau sering disebut dengan tanaman coklat merupakan tanaman pohon tahunan (*perennial*) dengan tinggi mencapai 5 meter. Di Indonesia tanaman kakao diperkenalkan oleh orang Spanyol pada tahun 1560 di Minahasa dan Sulawesi. Indonesia merupakan Negara sebagai penghasil kakao ketiga terbesar di dunia setelah Pantai Gading dan Ghana. Perkebunan kakao menurut status pengusahaan dalam periode 10 tahun terakhir (2012-2021), sebagian besar dikelola oleh Perkebunan Rakyat (97,57%), 1,01% dikelola Perkebunan Besar Negara (PBN) dan sisanya 1,42% dikelola Perkebunan Besar Swasta (PBS). Berdasarkan data BPS, produksi kakao di Indonesia sebanyak 667.300ton pada tahun 2022. Jumlah tersebut lebih rendah 3,04% dibandingkan pada tahun sebelumnya yang mencapai 688.200 ton. Kondisi ini salah satunya disebabkan oleh banyaknya pohon kakao yang sudah tua serta serangan penyakit dan hama pada tanaman kakao, sehingga tak lagi produktif.

Volume Produksi Kakao di Indonesia

*Berdasarkan data tahun 2012-2022



Gambar 2.10 Produksi Kakao di Indonesia

Sumber : CNBC Indonesia, 2023

Wilayah produksi terbesar kakao meliputi Sulawesi Tengah, Sulawesi Tenggara, Sulawesi Selatan, Sulawesi Barat, Lampung, Aceh, Sumatera Barat, dan Sumatera Utara. Namun, produksi kakao terbesar berasal dari pulau Sulawesi mencapai 75% dari total produksi kakao Indonesia. Pada tahun 2021 produksi kakao di Sulawesi Tengah, yakni mencapai 130.600 ton. Sulawesi Tenggara dan Sulawesi Barat dengan produksi kakao masing-masing 114.800ton dan 107.700 ton.

2.9 Penyakit Tanaman Kakao

Penyakit pada tanaman kakao dan daun kakao dapat disebabkan oleh berbagai faktor, termasuk infeksi oleh jamur, bakteri, virus, atau kondisi lingkungan yang tidak sesuai. Penyakit pada tanaman kakao menimbulkan kerugian yang cukup besar. Setiap tahun kerugian yang ditimbulkan bisa mencapai jutaan rupiah setiap hektar tanaman. berikut beberapa penyakit yang umumnya mempengaruhi tanaman kakao:

a. Antraknosa

Penyakit antraknosa pada kakao merupakan penyakit yang disebabkan oleh jamur *Colletotrichum Gloeosporioides*. Gejala penyakit antraknosa pada tanaman kakao meliputi bercak-bercak pada daun, batang, dan buah. Bercak-bercak ini awalnya berwarna kecil, basah, dan berlendir, namun kemudian dapat membesar dan mengering.



Gambar 2.11 Antraknosa

Sumber: <https://plantwiseplusknowledgebank.org/doi/pdf/10.1079/pwkb.2019780065>

7

b. *Vascular Streak Dieback* (VSD)

Penyakit *Vascular Streak Dieback* (VSD) disebabkan oleh jamur *Basidiomycete* (*Ceratobasidiales*). Gejalanya terdapat lesi panjang pada daun yang mengikuti pembuluh tanaman (vena). Daun kemudian mengering dan gugur.



Gambar 2.12 Vascular Streak Dieback (VSD)

Sumber: <https://plantwiseplusknowledgebank.org/doi/pdf/10.1079/pwkb.2019780065>

7

c. *Cocoa Leaf Blotch* (Penyakit Coklat pada daun)

Penyakit *Cocoa Leaf Blotch* (Penyakit Coklat pada daun) pada tanaman kakao disebabkan oleh jamur *Phytophthora palmivora*. Gejalanya terdapat bercak coklat pada buah dan daun.



Gambar 2.13 *cocoa Pod and leaf blotch*

Sumber: https://apps.lucidcentral.org/pppw_v10/text/web_full/entities/cocoa_black_pod_006.htm

d. *Cocoa Swollen Shoot Virus Disease* (CSSVD)

Penyakit *Cocoa Swollen Shoot Virus Disease* (CSSVD) ini disebabkan oleh virus *Cocoa Swollen Shoot*. Gejala yang ditimbulkan yaitu terdapat pucuk tanaman yang membesar, daun-daun muda menjadi keriting, dan tanaman mengalami hambatan pertumbuhan. Penyakit ini dapat menyebabkan kerugian hasil yang signifikan.



Gambar 2.14 *Cocoa Swollen Shoot Virus Disease* (CSSVD)

Sumber: <https://plantwiseplusknowledgebank.org/doi/pdf/10.1079/pwkb.2019780065>

2.10 Perbandingan Tinjauan

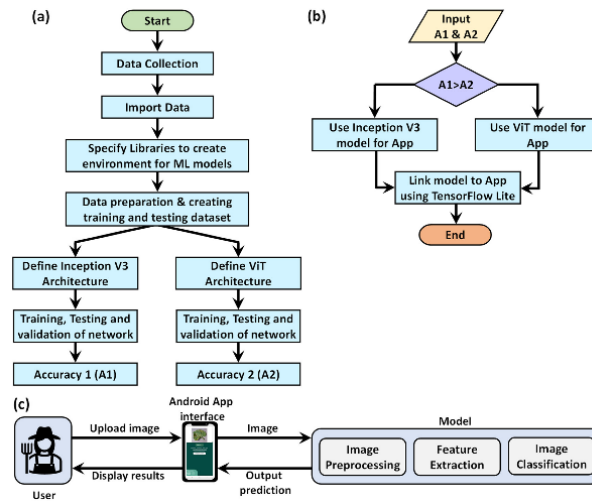
(Jesse, Douha, & Lenka, 2024) Image Classification For Csvgd Detection In Cacao Plants

Penelitian (Jesse et al., 2024) melakukan deteksi penyakit Cacao Swollen Shoot Virus Disease (CSSVD) pada tanaman kakao berbasis DL dengan tiga model yaitu VGG16, ResNet50, dan ViT. berfokus pada pengembangan dan evaluasi model Transformer untuk visi komputer yang relative baru untuk klasifikasi citra yang dapat mendeteksi penyakit CSSVD pada tanaman kakao dengan presisi tinggi. Penelitian ini menggunakan dataset KaraAgroAI Cocoa yang terdiri dari tiga kelas: CSSVD, Antraknosa dan sehat. Model VGG16 dan ResNet50 masing-masing terdiri dari 16 dan 50 lapisan konvolusi dengan optimasi Adam dan ukuran batch 64 dilatih selama 20 epoch. Pada model ViT, citra dipecah menjadi patch berukuran 16x16 piksel kemudian diflatten dan dipetakan ke dimensi tertentu menggunakan proyeksi linear. Embedding posisi ditambahkan untuk mempertahankan informasi posisi, dan dimasukkan kedalam encoder transformer yaitu lapisan multi-head self-attention (MSA) dan blok MLP. Hasil akurasi akhir yang diperoleh penelitian ini menunjukkan model ResNet50 memberikan kinerja terbaik dengan 94%, 92% VGG16, dan 80% ViT (Jesse et al., 2024).

(Barman, Sarma, Rahman, Deka, Lahkar, Sharma, dan Saikia, 2024) ViT-SmartAgri: Vision Transformer and Smartphone-Based Plant Disease Detection for Smart Agriculture

Penelitian mengembangkan model ViT-SmartAgri yang menggabungkan ViT dan Inception V3 untuk deteksi penyakit tanaman pada daun tomat. Penelitian ini berfokus pada ekstraksi dan kombinasi fitur dari berbagai skala dan konteks untuk mendeteksi penyakit pada taun tomat. Citra diubah ke ukuran seragam 64x64 piksel. ViT digunakan untuk ekstraksi fitur dan dimasukkan ke dalam encoder transformer yang menggunakan *self-attention* untuk menangkap hubungan global antar patch. Serta

Inception V3 digunakan untuk ekstraksi fitur lokal melalui berbagai filter ukuran yang berbeda. Mereka menggunakan Pendekatan *Feature fusion* untuk membentuk vector fitur gabungan dimasukkan ke lapisan *fully connected* melalui softmax activation. Model yang diusulkan dapat mengenali pola-pola yang kompleks dan variatif. Hasil akhir akurasi yang diperoleh model tersebut sebesar 95,76% ViT, dan 94,02% Inception V3 (Barman et al., 2024).



Gambar 2.15. Metodologi yang diusulkan (Barman et al., 2024)

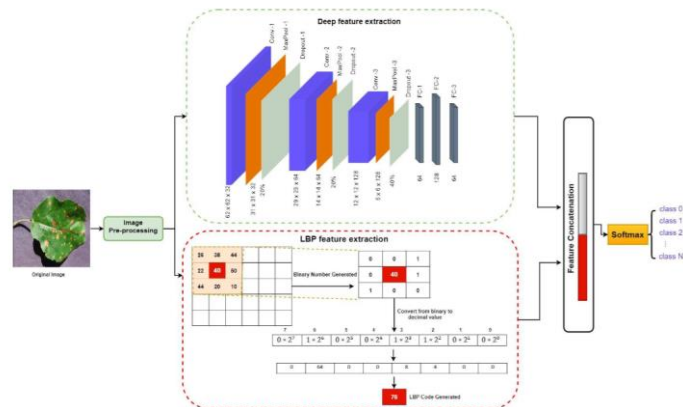
(Gao R, Dong Z, Wang Y, Cui Z, Kamu M, Dong B, Lu Y, Wang X, Yihong, & Yan S, 2024) Intelligent Cotton Pest and Disease Detection: Edge Computing Solutions with Transformer Technology and Knowledge Graphs

Penelitian ini berbasis DL dengan model transformer untuk deteksi hama dan penyakit kapas. Penelitian ini berfokus pada menggabungkan arsitektur transformer dengan teknologi knowledge graph untuk meningkatkan akurasi dan efisiensi deteksi hama dan penyakit kapas. Dalam penelitian ini model transformer terdiri dari komponen inti yaitu *joint-attention mechanism*, normalisasi layer, jaringan feed-forward, dan residual connection. Joint-attention mechanism pada penelitian ini memungkinkan model menganalisis data masukan dari berbagai perspektif secara bersamaan. Dalam model yang diusulkan mampu memiliki akurasi dan ketahanan yang

tinggi dalam menangani latar belakang yang kompleks seperti mampu mengidentifikasi tepi daun kapas dalam situasi tepi daun yang buram atau tercampur dengan warna latar belakang. Hasil akhir dari kinerja model yang diusulkan mencapai Tingkat akurasi sebesar 94%, mAP 95% dan FPS 49,7 (Gao R et al., 2024).

(Hosny, El-Hady, Samy, Vrochidou, & Papakostas, 2023) Multi-Class Classification of Plant Leaf Diseases Using Feature Fusion of Deep Convolutional Neural Network and Local Binary Pattern

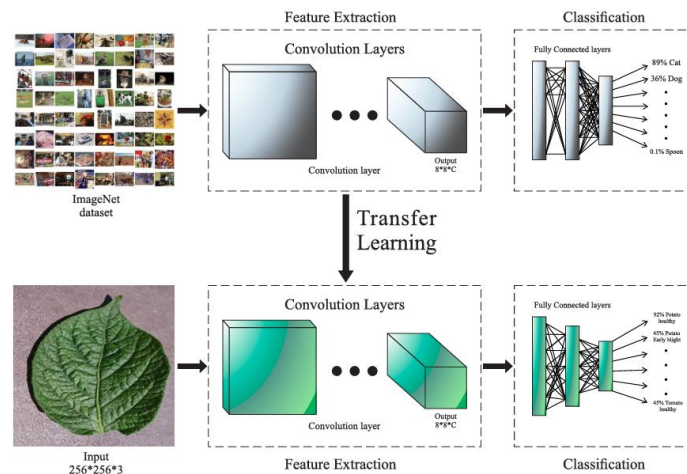
Penelitian ini berfokus pada klasifikasi multi-kelas penyakit daun tanaman menggunakan kombinasi fitur dari jaringan saraf konvolusi (CNN) dan LBP yang bertujuan untuk meningkatkan akurasi klasifikasi dengan menggabungkan ekstraksi fitur global dari CNN dan kemampuan ekstraksi fitur local dari LBP. Citra diubah menjadi 64x64 untuk input ke 3 lapisan konvolusional dengan kernel 3x3, lapisan max-pooling dengan kernel 3x3, dan 4 dense lapisan *fully connected*. Selanjutnya, melakukan ekstraksi fitur menggunakan LBP. Setelah itu, model dilatih dengan konfigurasi parameter tuning yaitu batch size 32, learning rate 0.01, dan epoch 50. Model diujicoba pada 3 jenis tanaman penyakit daun yaitu: pada dataset penyakit daun apel, penyakit daun tomat, dan penyakit daun anggur. Hasil akhir yang diperoleh dari model yang diusulkan masing-masing sebesar 98,80%, 96,50%, dan 98,30%. (Hosny et al., 2023)



Gambar 2.16. Metodologi yang diusulkan (Hosny et al., 2023)

(Tabbakh & Barpanda, 2023) A Deep Features Extraction Model Based on the Transfer Learning Model and Vision Transformer 'TLMViT' for Plant Disease Classification

Pada penelitian mengusulkan pendekatan baru berbasis pembelajaran mendalam yaitu model *feature fusion* dengan menggabungkan model transfer learning dengan ViT yang diberi nama model TLMViT untuk klasifikasi penyakit daun tanaman. Penelitian ini berfokus pada pendekatan ekstraksi fitur mendalam. Penelitian ini melakukan empat tahap yaitu akusisi data, dimana Kumpulan data PlantVillage dan gandum digunakan untuk melatih dan mengevaluasi model. Selanjutnya mereka melakukan augmentasi untuk meningkatkan jumlah sampel pelatihan dan mengatasi masalah overfitting. Tahapan berikutnya dilakukan ekstraksi fitur daun sebanyak dua fase: ekstraksi fitur awal menggunakan model berbasis pre-trained dan ekstraksi fitur mendalam dilakukan menggunakan model ViT. Tahapan terakhir, klasifikasi menggunakan pengklasifikasian MLP. Model yang diusulkan diuji dengan lima model berbasis pre-trained yang diikuti oleh ViT secara individual. Model yang diusulkan dapat mengungguli model transfer learning lainnya dengan peningkatan akurasi sebesar 1,11% dan 1,099%. (Tabbakh & Barpanda, 2023).



Gambar 2.17. Metodologi yang diusulkan (Tabbakh & Barpanda., 2023)

(Lye, & Ng, 2023) Processing Plant Diseases Using Transformer Model

Penelitian (Lye H & Ng K, 2023) melakukan klasifikasi penyakit tanaman menggunakan ViT. Penelitian ini berfokus identifikasi pada keseluruhan daun dibandingkan bagian individual atau bagian tertentu. Pada penelitian ini melakukan beberapa tahapan yaitu tahapan awal, melakukan pra-pemrosesan seperti resize data, implementasi modul *shift patch tokenization*, visualisasi token atau *patches*, dan implementasi *patch encoder*. Tahapan selanjutnya, menambahkan lapisan *encoding* menggunakan modul *Locality Self Attention* (LSA) yang membantu model mempelajari hubungan antar-token, modul MLP. Setelah itu, model ViT dilatih selama 30 epoch dengan *batch size* 256 citra per *batch*. Data dibagi menjadi dua bagian yaitu 90% data latih dan 10% data validasi. Hasil akurasi yang diperoleh adalah 89,58%, serta ketahanan model yang memuaskan terhadap perubahan orientasi dan posisi daun (Lye & Ng, 2023).

(Rachmad, Syarief, Rifka, Sonata, Setiawan, Eka Rochman, 2022) Corn Leaf Disease Classification Using Local Binary Patterns (LBP) Feature Extraction

Penelitian (Rachmad et al, 2022) berfokus untuk mengembangkan model menggunakan ekstraksi fitur LBP yang dikombinasikan dengan K-NN untuk klasifikasi. Selain itu, pengujian dilakukan juga menggunakan beberapa metode klasifikasi seperti Naïve Bayes dan Adaboost. Dataset yang digunakan yaitu citra penyakit daun jagung terdiri dari tiga jenis penyakit. Jumlah data yang digunakan sebanyak 3500 dataset. Tahap awal dilakukan perubahan ukuran citra menjadi 300x300 serta citra RGB dikonversi ke citra Grayscale. Setelah itu, melakukan ekstraksi fitur menggunakan LBP dan diklasifikasikan menggunakan KNN. Selanjutnya, model yang diusulkan dilakukan perbandingan dengan Naïve Bayes dan Adaboost. Hasil akhir akurasi yang diperoleh yaitu model yang diusulkan menunjukkan akurasi lebih tinggi dibandingkan kedua model lainnya sebesar 81,1% (Rachmad et al., 2022)

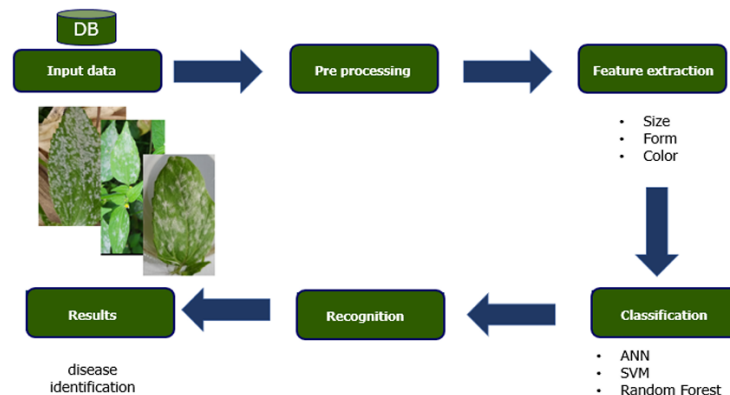
(Borhani, Khoramdel, & Najafi, 2022) A deep learning-based approach for automated plant disease classification using vision transformer

Penelitian ini memperkenalkan model ringan untuk klasifikasi penyakit tanaman secara real-time menggunakan struktur ViT. (Borhani et al., 2022) membandingkan model berbasis CNN dengan kompleksitas yang hampir sama dan juga mengeksplorasi model dengan menggabungkan blok CNN dengan blok transformer. Blok CNN terdiri dari dua lapisan konvolusional dengan kernel 3x3 dan 2 lapisan ReLU serta satu max pooling layer dengan kernel 2x2. Blok kedua yaitu blok transformer dengan masukan lapisan normalisasi diikuti oleh lapisan *multi-head*, dimana empat *multi-head attention* digunakan pada semua blok transformer kemudian keluaran ditambahkan ke masukan blok MLP. Penelitian diujicoba pada tiga kumpulan data yaitu Kumpulan data penyakit daun padi, kumpulan data klasifikasi karat gandum dan Kumpulan data *PlantVillage*. Pengujian pada ketiga Kumpulan data dilakukan dengan tiga resolusi berbeda: 50x50 *pixel*, 100x100 *pixel*, dan 200x200 *pixel*. Setelah dilakukan pengujian pada ketiga Kumpulan data, model berbasis ViT mencapai performa yang lebih akurat dibandingkan model CNN maupun hybrid. Namun, pada penelitian ini blok *attention* yang digunakan lebih lambat dibandingkan blok konvolusional. Menggabungkan blok konvolusional dengan blok *attention* membantu model memprediksi lebih cepat daripada model berbasis ViT dan mendapatkan akurasi lebih tinggi dibandingkan model CNN (Borhani et al., 2022).

(Rodriguez, Alfaro, Paredes, Esenarro, & Hillario, 2021) Machine Learning Techniques in the Detection of Cocoa (*Theobroma cacao L.*) Diseases

Penelitian ini berfokus pada pengembangan sistem deteksi penyakit pada pohon kakao menggunakan ML. penelitian (Rodriguez et al., 2021) menggunakan pendekatan *feature fusion* yang menggabungkan fitur-fitur yang diekstraksi untuk meningkatkan akurasi model deteksi penyakit. Penelitian ini melalui beberapa tahapan awal preprocessing dengan melakukan normalisasi warna yang dikonversi ke skala abu-abu dan ukuran citra disesuaikan seragam menjadi 256x256 piksel. Tahapan berikutnya,

melakukan ekstraksi fitur menggunakan HoG dan LBP untuk mengekstraksi fitur tekstur. *Support Vector Machine* (SVM), *Random Forest* (RF), dan *Artificial Neural Network* (ANN) digunakan untuk klasifikasi model. Hasil akurasi yang diperoleh diperoleh model SVM dengan fitur HoG mencapai akurasi 70%, model SVM dengan fitur LBP mencapai akurasi 75% dan model ANN dengan fitur LBP mencapai akurasi tertinggi sebesar 85% (Rodriguez et al., 2021).

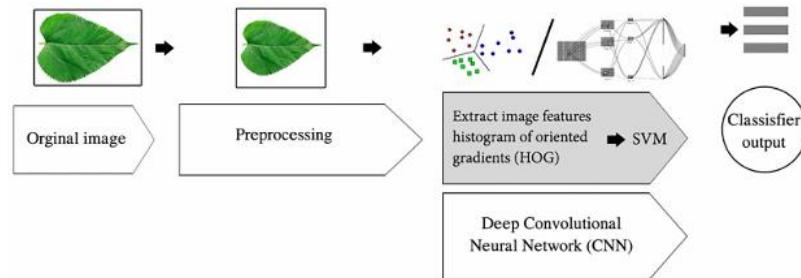


Gambar 2.19. Metodologi yang diusulkan (Rodriguez et al., 2021)

(Bao, Kiet T, Dinh T, Hiep H, 2020) Plant species identification from leaf patterns using histogram of oriented gradients feature space and convolution neural networks

Berfokus pada identifikasi spesies tanaman berdasarkan pola daun menggunakan dua metode yaitu: ekstraksi fitur HoG dan SVM sebagai klasifikasi. Serta metode CNN untuk melakukan ekstraksi fitur otomatis dan melakukan klasifikasi spesies tanaman. (Bao et al., 2020) membandingkan kinerja antara metode tradisional (HoG-SVM) dengan metode DL (CNN) dalam mengidentifikasi spesies tanaman dari gambar daun. HoG-SVM membagi citra menjadi sel-sel kecil (8x8 piksel) dan menghitung histogram arah gradien. Metode CNN ukuran citra menjadi 128x128 piksel, arsitektur CNN terdiri dari beberapa lapisan konvolusi, ReLU, dan Pooling. Pelatihan dilakukan dengan menggunakan dataset augmented untuk meningkatkan jumlah data pelatihan. Dan klasifikasi dilakukan pada lapisan *fully connected* dengan fungsi softmax. Hasil akurasi yang dicapai HoG-SVM pada dataset Swedish sebesar

98% dan 92% dataset Flavia. Akurasi CNN dicapai pada dataset Swedish sebesar 90,22% dan 95,5% dataset Flavia (Bao et al., 2021)



Gambar 2.20. Metodologi yang diusulkan (Bao et al., 2021)

(Aminul et al., 2019) Automatic Plant Detection Using HOG and LBP Features With SVM

Penelitian ini melakukan identifikasi spesies daun pada tanaman menggunakan HoG dan LBP sebagai ekstraksi fitur dan multiclass SVM melakukan klasifikasi citra daun tanaman. Penelitian ini melalui tahapan preprocessing seperti segmentasi citra yaitu citra RGB dikonversi ke citra grayscale, setelah itu, citra grayscale diubah menjadi citra biner menggunakan Teknik thresholding dan tahap normalisasi citra. Tahapan ekstraksi fitur dilakukan dengan HoG dan LBP serta SVM untuk klasifikasi citra daun. HoG-SVM menggunakan 3 ukuran sel (cells) berbeda yaitu 2x2, 4x4, dan 8x8 yang masing-masing mendapatkan akurasi sebesar 77,5%, 81,25% dan 85,31%. Akurasi LBP-SVM yaitu 40,6% sedangkan akurasi yang didapatkan dengan metode *feature fusion* HoG dan LBP dengan SVM mencapai akurasi sebesar 91,25% (Aminul et al., 2019).

Penelitian yang sudah dilakukan peneliti terdahulu terangkum dalam table 2.1 berikut.

Tabel 2.1 Ringkasan Kajian Penelitian

NO.	Peneliti, Tahun	Subjek Penelitian	Metode Penelitian	Kelebihan	Kekurangan
1.	Atuhurra Jesse, N'guessan Yves-Roland Douha, Pabitra Lenka (2024)	Deteksi penyakit CSSVD pada daun kakao	ConvNet (VGG16 ResNet50) dan Vision Transformer	- Model ResNet50 mendapatkan hasil terbaik mencapai presisi 95,39%, recall 93,75%, 94,34% dan akurasi 94%.	- Masih perlu perluasan dataset yang lebih beragam untuk meningkatkan kemampuan generalisasi model dalam mendeteksi penyakit pada kondisi nyata yang lebih luas.
2.	Utpal Bartender, Parismita Sarma, Mirzanur	Deteksi pada 10 penyakit daun tomat	ViT-SmartAgri	- Model ViT menunjukkan performa yang lebih tinggi dibandingkan	- Model memerlukan lebih banyak waktu pelatihan dan lebih banyak data untuk

	Peneliti, Tahun	Subjek Penelitian	Metode Penelitian	Kelebihan	Kekurangan
	Rahman, Vaskar Deka, Swati Lahkar, Vaisali Sharma, dan Manob Jyoti Saikia (2024)			<ul style="list-style-type: none"> - dengan model Inception V3 memperoleh akurasi 95,76% ViT dan 94,02%. - Model yang diusulkan mampu mengenali pola-pola yang kompleks dan variatif 	<ul style="list-style-type: none"> - mencapai performa yang optimal.
3.	Ruicheng Gao, Zhan Cai Dong, Yuqi Wang, Zhuowen Cui, Muyang kamu, Bowen Dong, Yuchun Lu, Xuaner Wang,	Deteksi hama dan penyakit pada daun kapas	Transformer Based dan Knowledge Graph	<ul style="list-style-type: none"> - Hasil akurasi model yang diusulkan telah mencapai 94% - Penggunaan <i>joint-attention mechanism</i> dan <i>joint-head design</i> mampu membuat model menangani 	<ul style="list-style-type: none"> - Diversitas dan skala dataset masih perlu perluasan untuk meningkatkan kemampuan generalisasi model. - Konsep knowledge <i>graph</i> seperti

	Peneliti, Tahun	Subjek Penelitian	Metode Penelitian	Kelebihan	Kekurangan
	Lagu Yihong, Dan Shu Yan (2024)			- latar belakang yang kompleks dan mendeteksi objek kecil seperti serangan hama.	- ConceptNet belum sepenuhnya menangkap hubungan visual yang relevan
4.	Khalid M. Hosny, Walaa M. El-Hady, Farid M. Samy, Eleni Vrochidou, And George A. Papakostas (2023)	Klasifikasi penyakit daun tanaman tomat, apel, dan anggur	Multi-Class Classification of Plant Leaf Diseases Using Feature Fusion of Deep Convolutional Neural Network and Local Binary Pattern	- Pendekatan Feature fusion memberikan hasil yang lebih baik dibandingkan dengan metode konvensional dalam hal akurasi klasifikasi - Pendekatan ini membuktikan efisiensi dalam mengolah dataset -	- Penggunaan dua Teknik ekstraksi fitur dapat meningkatkan kompleksitas komputasi yang membutuhkan sumber daya yang lebih besar - Akurasi model dapat menurun jika citra yang digunakan memiliki kualitas rendah atau

NO.	Peneliti, Tahun	Subjek Penelitian	Metode Penelitian	Kelebihan	Kekurangan
				<ul style="list-style-type: none"> - besar dengan berbagai jenis penyakit daun - Model yang diusulkan dapat menggeneralisasi dengan baik <p>Akurasi yang diperoleh pada tanaman apel, tomat, anggur masing-masing sebesar 98,80%, 96,50%, dan 98,30%.</p>	terdapat noise yang signifikan.
5.	Amer Tabbakh And Soubhagya Sankar	Klasifikasi penyakit daun pada PlantVillage dataset dan	Model hybrid yang disebut TLMViT	<ul style="list-style-type: none"> - Model hybrid yang diusulkan mengungguli model transfer learning 	<ul style="list-style-type: none"> - Model dilatih pada Kumpulan data plantvillage masih mengalami penurunan

NO.	Peneliti, Tahun	Subjek Penelitian	Metode Penelitian	Kelebihan	Kekurangan
	Barpanda (2023)	dataset tanaman gandum	menggabungkan model berbasis transfer learning dan vision transformer	<ul style="list-style-type: none"> - lainnya dengan peningkatan akurasi sebesar 1,11% dan 1,099%. Menggabungkan fitur pre-trained dengan self-attention vision transformer mampu menangkap pola relevan secara efektif. 	<ul style="list-style-type: none"> - performa yang signifikan saat diuji pada citra dengan latar belakang, sudut yang lebih bervariasi.
6.	Hong-Zheng Marcus Lye, dan Kok-Why Ng (2023)	10 penyakit daun pada tanaman tomat	Vision Transformer (ViT)	<ul style="list-style-type: none"> - Model yang diusulkan mencapai akurasi pengujian sebesar 89,58% dengan variasi yang relative kecil. - 	<ul style="list-style-type: none"> - Keterbatasan daya komputasi membatasi ukuran input selama proses pelatihan, yang berdampak pada kualitas dan akurasi model

NO.	Peneliti, Tahun	Subjek Penelitian	Metode Penelitian	Kelebihan	Kekurangan
				<ul style="list-style-type: none"> - Model menunjukkan ketahanan yang baik terhadap perubahan orientasi dan posisi daun dalam citra. - Penambahan fitur seperti Shift Patch Tokenization dan LSA membantu model menjadi lebih robust. 	<ul style="list-style-type: none"> - Beberapa kelas memiliki akirasi atau skor F1 yang lebih rendah karena jumlah citra yang sedikit
7.	Aeri Rachmad, Mohammad Syarief, Silfia Rifka, Fifin Sonata, Wahyudi	3 jenis penyakit daun pada tanaman jagung	Ekstraksi fitur LBP dan metode klasifikasi KNN	<ul style="list-style-type: none"> - Metode KNN menunjukkan hasil AUC sebesar 94,1%, F1-Score 80,9%, precision 81,8%, recall 81,1%. 	<ul style="list-style-type: none"> - Penggunaan LBP belum mengakomodasi semua variasi tekstur yang ada pada daun jagung secara optimal.

NO.	Peneliti, Tahun	Subjek Penelitian	Metode Penelitian	Kelebihan	Kekurangan
	Setiawan, Eka Mala Sari Rochman (2022)				
8.	Yixin Chen, Xiyun Wang, Zhibo Chen, Kang Wang, Ye Sun, Jiarong Jiang, Xuhao Liu (2023)	7 penyakit pada daun camellia oleifera	<i>Attention</i> dan <i>multi- dimensional feature fusion</i> neural network (AMDFNet)	<ul style="list-style-type: none"> - Model yang diusulkan mencapai akurasi 86,78%. - AMDFNet memiliki jumlah parameter yang lebih sedikit, hanya seperempat dari parameter yang digunakan oleh ResNet50, sehingga lebih efisien dalam hal komputasi. 	<ul style="list-style-type: none"> - Penelitian ini masih terbatas pada beberapa jenis penyakit umum pada daun Camellia oleifera, sehingga kemampuan klasifikasi terhadap penyalit yang lebih jarang masih perlu ditingkatkan. - Penggunaan <i>attention mechanism</i> dan <i>feature fusion</i> multidimensional

NO.	Peneliti, Tahun	Subjek Penelitian	Metode Penelitian	Kelebihan	Kekurangan
				- Model yang diusulkan dirancang agar dapat dijalankan di perangkat mobile, memungkinkan penyakit yang cepat dan akurat.	menambah kompleksitas dalam implementasi dan pemeliharaan model.
7.	Aeri Rachmad, Mohammad Syarief, Silfia Rifka, Fifi Sonata, Wahyudi Setiawan, Eka Mala Sari Rochman (2022)	3 jenis penyakit daun pada tanaman jagung	Ekstraksi fitur LBP dan metode klasifikasi KNN	- Metode KNN menunjukkan hasil AUC sebesar 94,1%, F1-Score 80,9%, precision 81,8%, recall 81,1%.	- Penggunaan LBP belum mengakomodasi semua variasi tekstur yang ada pada daun jagung secara optimal.

NO.	Peneliti, Tahun	Subjek Penelitian	Metode Penelitian	Kelebihan	Kekurangan
8.	Yasamin Borhani, Javad Khoramdel dan Esmaeil Najaf (2022)	Klasifikasi penyakit daun padi, gandum dan plantvillage	Vision Transformer (ViT)	- Model berbasis ViT dan Kombinasi CNN-ViT menunjukkan akurasi yang lebih tinggi dibandingkan dengan model klasik	- Penggunaan blok atensi memperlambat kecepatan prediksi dibandingkan dengan model berbasis CNN -
9.	Ciro Rodriguez, Oswaldo Alfaro, Pervis Paredes, Doris Esenarro, Francisco Hilario (2021)	5 jenis penyakit pada tanaman kakao.	Teknik Machine Learning menggunakan algoritma HoG dan LBP, dan SVM	- Algoritma SVM, RF, dan ANN yang digunakan menunjukkan kemampuan yang baik dalam memprediksi kondisi tanaman kakao	- Dataset yang digunakan masih terbatas dan perlu ditingkatkan untuk menghasilkan hasil yang lebih akurat.

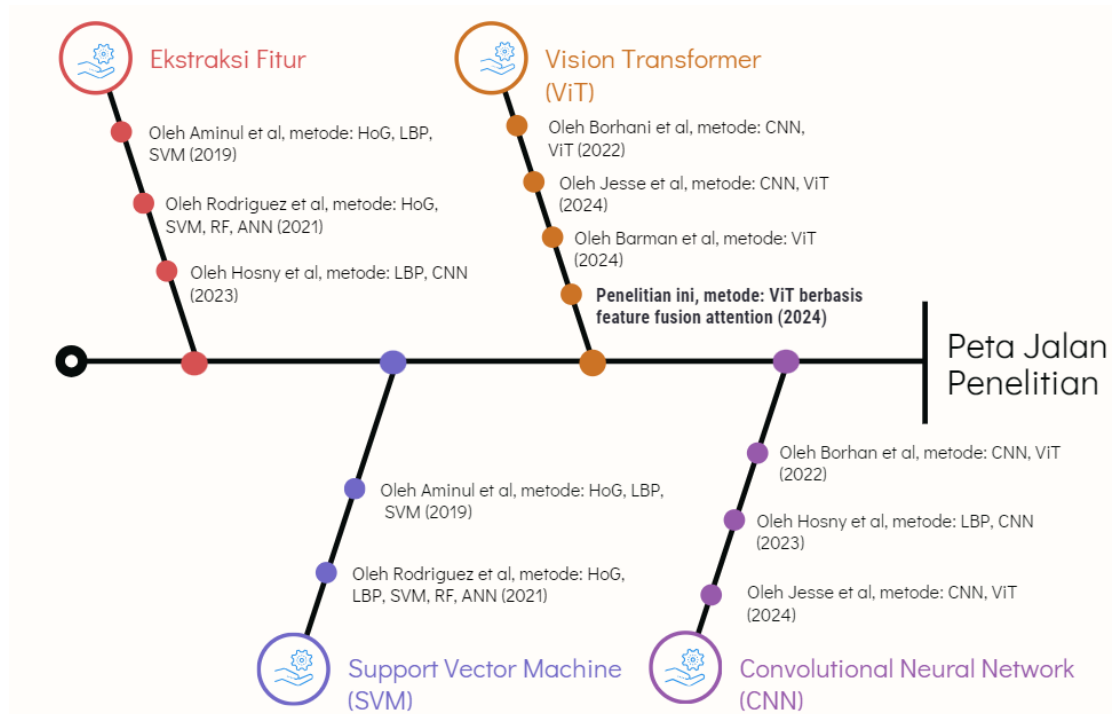
NO.	Peneliti, Tahun	Subjek Penelitian	Metode Penelitian	Kelebihan	Kekurangan
					- Mengekplorasi algoritma ekstraksi fitur lain seperti SIFT dan SURF untuk meningkatkan hasil.
10.	Truong Quoc Bao, Nguyen Thanh Tan Kiet, Truong Quoc Dinh & Huynh Xuan Hiep (2019)	Menggunakan <i>Swedish leaf dataset</i> dengan 15 jenis tanaman dan 75 jenis daun dengan total 1125 citra daun. dan <i>Flavia dataset</i> dengan 32 jenis berbeda yang berjumlah 1907 citra daun.	Membandingkan Histogram of oriented gradients feature space dan convolution neural networks	<ul style="list-style-type: none"> - Model HoG-SVM menunjukkan akurasi yang baik sebesar 98% dan flavia 92%. - Model CNN menunjukkan akurasi lebih tinggi dengan Swedish 98,22% dan flavia 95,5% - CNN menunjukkan performa yang lebih 	<ul style="list-style-type: none"> - Efektivitas model HoG-SVM sangat bergantung pada ukuran dan kualitas citra input serta Panjang vector fitur - Membutuhkan perhitungan fitur manual yang memakan waktu. - CNN membutuhkan banyak data pelatihan -

NO.	Peneliti, Tahun	Subjek Penelitian	Metode Penelitian	Kelebihan	Kekurangan
				- baik Ketika kedua dataset digabungkan	- untuk mencapai performa yang baik Memerlukan komputasi yang lebih tinggi dan sumber daya GPU untuk pelatihan.
11.	Mohammad Aminul Islama, Md. Sayeed Iftekhhar Yousuf, M. M. Billah (2019)	<i>Flavia leaf dataset</i>	Fitur HoG dan LBP menggunakan SVM	- Kombinasi HoG dan LBP dengan SVM menunjukkan peningkatan signifikan dalam akurasi deteksi dibandingkan menggunakan HoG atau LBP saja. - Akurasi tertinggi yang dicapai adalah	- LBP & SVM hanya mencapai akurasi sebesar 40,6%, menunjukkan bahawa LBP kurang efektif jika digunakan sendiri tanpa kombinasi fitur lain.

NO.	Peneliti, Tahun	Subjek Penelitian	Metode Penelitian	Kelebihan	Kekurangan
				- 91,25% dengan menggabungkan fitur HiG dan LBP	

2.11 Roadmap Penelitian

Gambar 2.20 menunjukkan roadmap Penelitian yang terkait dengan topik penelitian yang sudah dilakukan oleh peneliti dalam periode 2019-2024.



Gambar 2.20 Roadmap Penelitian

Pada tahun 2019, dilakukan deteksi pada penyakit daun tanaman menggunakan ekstraksi fitur HoG dan LBP serta diklasifikasikan menggunakan model SVM (Aminul et al., 2019). Pada tahun 2021, sebuah penelitian yang mengembangkan sistem deteksi penyakit pohon kakao dengan pendekatan *feature fusion* menggunakan HoG dan LBP untuk ekstraksi fitur dan SVM, RF, ANN sebagai model klasifikasi (Rodriguez et al., 2021). Pada tahun 2022, melakukan klasifikasi untuk identifikasi penyakit daun jagung menggunakan ekstraksi fitur LBP dan KNN sebagai model klasifikasi penyakit daun jagung (Rachmad et al., 2022). Pada tahun yang sama, terdapat penelitian yang mengeksplorasi model ViT dalam klasifikasi penyakit tanaman yang menggabungkan fitur CNN dengan *attention mechanism* (Borhani et al, 2022). Pada tahun 2023,

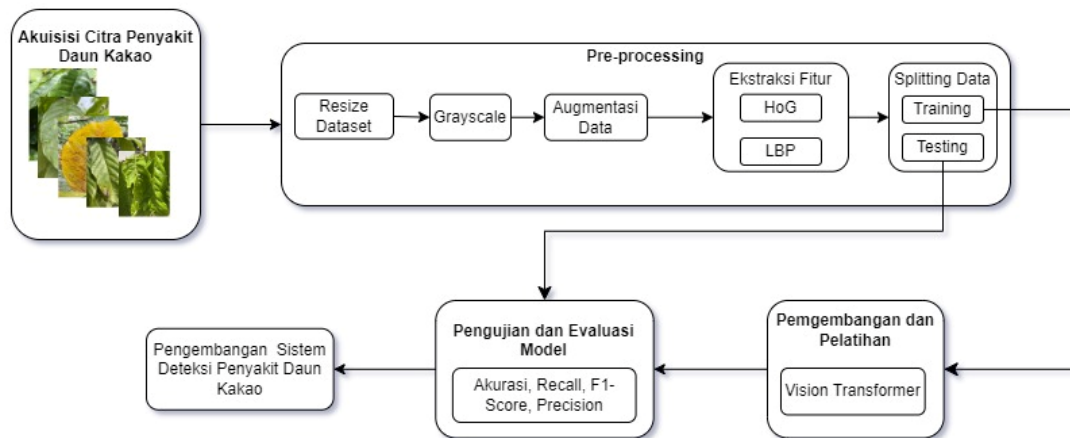
melakukan deteksi digunakan LBP dan CNN untuk mendeteksi penyakit pada tanaman (Hosny et al., 2023). Pada tahun 2024, melakukan klasifikasi citra untuk mendeteksi penyakit CSSVD pada tanaman kakao menggunakan tiga model klasifikasi yaitu VGG16, ResNet50, dan ViT (Jesse et al., 2024). ViT digunakan untuk ekstraksi fitur dan dimasukkan ke dalam encoder transformer menggunakan *self-attention* untuk menangkap hubungan global antar patch untuk deteksi penyakit daun tanaman (Barman et al., 2024). Pada tahun ini 2024, akan diajukan oleh peneliti “PENGEMBANGAN ALGORITMA BERBASIS VISION TRANSFORMER PADA DETEKSI PENYAKIT TANAMAN KAKAO”.

BAB 3

METODE PENELITIAN

3.1 Tahapan Penelitian

Secara garis besar penelitian ini terdiri dari beberapa tahapan, yaitu Akuisisi data, pre-processing data, pengembangan dan pelatihan model, pengujian dan evaluasi model, serta pengembangan system deteksi penyakit daun kakao, ditunjukkan pada Gambar 3.1.



Gambar 3.1 Tahapan Penelitian

3.2 Akuisisi Data Penyakit Daun Tanaman Kakao

Pengumpulan citra Penyakit Daun tanaman kakao dikumpulkan secara langsung oleh peneliti (data primer) dan juga menggunakan data yang dikumpulkan oleh peneliti lain (data sekunder). Terdapat 4 kelas penyakit dan satu kelas daun sehat yang akan digunakan dalam penelitian ini, yaitu: Daun sehat, penyakit *antraknosa* (*Colletotrichum gloeosporioides*), penyakit *vascular streak dieback* (VSD), penyakit *Leaf Blotch* dan penyakit *cocoa swollen shoot virus disease* (CSSVD).



Gambar 3.2 Contoh 4 Jenis Penyakit Daun Tanaman Kakao

Dataset primer akan dilakukan pengambilan foto penyakit daun tanaman kakao yang terdapat pada kebun kakao di daerah Kabupaten Solok, Provinsi Sumatra Barat. Pengambilan akan dilakukan dari jarak 20 cm dari kamera yang bertujuan menangkap detail kecil seperti bercak kecil atau lesi pada daun, perubahan warna serta tekstur permukaan daun. Dataset sekunder menggunakan dataset yang telah digunakan umum oleh para peneliti lain terkait penyakit daun tanaman kakao.

3.3 Pre-Processing

3.3.1 Resize Dataset

Perubahan ukuran citra dilakukan menggunakan metode *nearest neighbor interpolation*. Cara kerja dari metode ini dengan cara mengambil nilai piksel terdekat dari citra asli untuk menentukan nilai piksel baru dalam citra yang akan diubah ukurannya. Citra diubah ukurannya menjadi seragam (224 x 224 piksel). Faktor skala dihitung dengan membandingkan dimensi citra asli, dimana (W, H) dengan dimensi baru (W', H') yang akan diubah. untuk setiap piksel dalam citra baru dengan koordinat (i', j') , hitung koordinat terdekat di citra asli (i, j) . Selanjutnya map nilai piksel yaitu mengambil nilai piksel dari citra asli pada koordinat (i, j) dan menetapkan nilai ke piksel baru di koordinat (i', j') dalam citra yang diubah ukurannya.

3.3.2 Grayscale

Pada tahap ini citra RGB dikonversi ke Grayscale untuk membantu menyederhanakan dan memfokuskan informasi intensitas cahaya yang lebih relevan. Gejala penyakit pada daun tanaman kakao seperti perubahan warna, bintik-bintik atau

nekrosis dapat lebih mudah diidentifikasi melalui variasi intensitas Cahaya. Grayscale dapat mempertahankan informasi penting dengan lebih sederhana.

3.3.3 Augmentasi Dataset

Augmentasi data dilakukan untuk meningkatkan variasi pada dataset yang akan digunakan serta untuk mencegah terjadinya overfitting. Teknik augmentasi yang diterapkan penelitian ini seperti *rotasi, flipping, zooming, dan cropping*.

3.3.4 Ekstraksi Fitur

3.3.4.1 Histogram Oriented of Gradients (HoG)

Ekstraksi fitur HoG digunakan dalam penelitian ini untuk menangkap bentuk dan tekstur. HoG berfokus pada gradien intensitas lokal dan arah tepi yang menggambarkan struktur dan tekstur dari daun yang terkena penyakit pada tanaman kakao, HoG dapat menangani perubahan dalam rotasi dan skala yang memungkinkan pendeteksian penyakit yang konsisten pada pengambilan citra dari sudut atau jarak yang berbeda.

3.3.4.2 Local Binary Pattern (LBP)

Penerapan ekstraksi fitur LBP pada penelitian ini untuk menangkap tekstur lokal dalam citra. LBP membantu dalam menangkap informasi tekstur seperti bercak-bercak, lubang kecil yang terdapat pada daun, perubahan warna daun yang tidak merata dan perubahan permukaan lainnya.

3.3.5 Splitting Data

3.3.5.1 Data Training

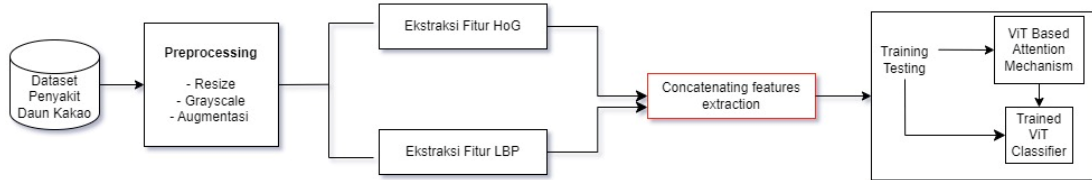
Data training digunakan untuk melatih model untuk mengenali pola maupun karakteristik visual yang membedakan daun sehat dengan daun yang terinfeksi penyakit. Melalui proses pelatihan ini model mengoptimalkan parameter untuk meminimalkan kesalahan dalam memprediksi.

3.3.5.2 Data Testing

Data testing digunakan untuk melakukan pengujian pada model yang telah dilatih sebelumnya untuk mengevaluasi kinerja model.

3.4 Pengembangan dan Pelatihan Model

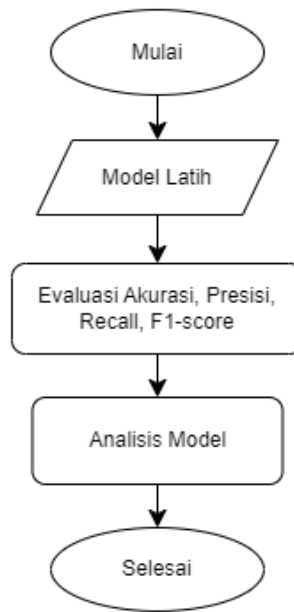
Data citra daun kakao yang telah melalui preprocessing dan ekstraksi fitur, kemudian digunakan untuk pelatihan dan pembuatan model *deep learning* menggunakan pendekatan *feature fusion* berbasis *attention mechanism* yaitu fitur ekstraksi HoG dan LBP digabungkan menggunakan metode *concatenation* ke dalam lapisan *vision transformer* yang menggunakan *attention mechanism*. *Attention mechanism* dalam vision transformer memberikan fokus yang berbeda pada fitur HoG dan LBP. Penggunaan *attention mechanism* dapat meningkatkan akurasi model dengan mengurangi pengaruh noise atau informasi yang tidak relevan dalam gambar.



Gambar 3.3 Pengembangan dan Pelatihan Model

3.5 Pengujian dan Evaluasi Model

Pengujian dan Evaluasi model dilakukan untuk melihat akurasi model saat mengidentifikasi penyakit daun tanaman kakao. Proses evaluasi dimulai dengan pengujian yang terdiri dari data yang belum pernah dilihat oleh model selama melakukan fase pelatihan. Matrik evaluasi digunakan untuk mengukur kinerja model secara menyeluruh. matrik evaluasi yang digunakan seperti akurasi, presisi, recall, dan F1-score.



Gambar 3.4 Pengujian dan Evaluasi Model

3.6 Jadwal Penelitian

Jadwal penelitian merupakan rancangan kegiatan yang dilakukan selama penelitian beserta estimasi waktu tiap kegiatan seperti yang ditunjukkan tabel 3.1

Tabel 3.1 Jadwal Penelitian

Kegiatan	Tahun Pertama (Semester 1)											
	1	2	3	4	5	6	7	8	9	10	11	12
Studi Literatur												
	Tahun Kedua (Semester 2 & 3)											
Pembuatan Proposal												
Pengumpulan Dataset (Sekunder)												
Pengumpulan Dataset (Primer & Sekunder)												

[illegible]

Bibliography

- Altuntaş, Y., & KOCAMAZ, F. (2021). Deep feature extraction for detection of tomato plant diseases and pests based on leaf images. *Celal Bayar University Journal of Science*, 17(2), 145-157.
- Aminul Islam, M., Billah, M., Sayeed Iftekhhar Yousuf, M., & Billah, M. M. (2019). Automatic Plant Detection Using HOG and LBP Features With SVM. *Article in International Journal of Computer*. <http://ijcjournal.org/>
- Barman, U., Sarma, P., Rahman, M., Deka, V., Lahkar, S., Sharma, V., & Saikia, M. J. (2024). ViT-SmartAgri: Vision Transformer and Smartphone-Based Plant Disease Detection for Smart Agriculture. *Agronomy*, 14(2). <https://doi.org/10.3390/agronomy14020327>
- Boukabouya, R. A., Moussaoui, A., & Berrimi, M. (2022, November). Vision Transformer Based Models for Plant Disease Detection and Diagnosis. *In 2022 5th International Symposium on Informatics and its Applications (ISIA)* (pp. 1-6). IEEE.
- Borhani, Y., Khoramdel, J., & Najafi, E. (2022). A deep learning based approach for automated plant disease classification using vision transformer. *Scientific Reports*, 12(1), 11554.
- Cau, R., Pisu, F., Muscogiuri, G., Mannelli, L., Suri, J. S., & Saba, L. (2023). Applications of artificial intelligence-based models in vulnerable carotid plaque. *Vessel Plus*. <https://doi.org/10.20517/2574-1209.2023.78>
- Chen, Y., Wang, X., Chen, Z., Wang, K., Sun, Y., Jiang, J., & Liu, X. (2023). Classification of *Camellia oleifera* Diseases in Complex Environments by Attention and Multi-Dimensional Feature Fusion Neural Network. *Plants*, 12(14), 2701.

- Chen, Z., Wang, G., Lv, T., & Zhang, X. (2024). Using a Hybrid Convolutional Neural Network with a Transformer Model for Tomato Leaf Disease Detection. *Agronomy*, 14(4), 673.
- Chen, Z., Zhou, H., Lin, H., & Bai, D. (2024). TeaViTNet: Tea Disease and Pest Detection Model Based on Fused Multiscale Attention. *Agronomy*, 14(3), 633.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.
- Faisal, M., Leu, J. S., & Darmawan, J. T. (2023). Model Selection of Hybrid Feature Fusion for Coffee Leaf Disease Classification. *IEEE Access*, 11, 62281–62291. <https://doi.org/10.1109/ACCESS.2023.3286935>
- Gao, R., Dong, Z., Wang, Y., Cui, Z., Ye, M., Dong, B., ... & Yan, S. (2024). Intelligent Cotton Pest and Disease Detection: Edge Computing Solutions with Transformer Technology and Knowledge Graphs. *Agriculture*, 14(2), 247.
- Giri, A., Saxena, D. R. R., Saini, P., & Rawte, D. S. (2020). Role of artificial intelligence in advancement of agriculture. *International Journal of Chemical Studies*, 8(2), 375-380.
- Harakannanavar, S. S., Rudagi, J. M., Puranikmath, V. I., Siddiqua, A., & Pramodhini, R. (2022). Plant leaf disease detection using computer vision and machine learning algorithms. *Global Transitions Proceedings*, 3(1), 305–310. <https://doi.org/10.1016/j.gltp.2022.03.016>
- Harvyanti, A. F. M., Baihaki, R. I., Ridlo, Z. R., & Agustin, I. H. (2023, May). Application of Convolutional Neural Network for Identifying Cocoa Leaf Disease. In *1st International Conference on Neural Networks and Machine Learning 2022 (ICONNSMAL 2022)* (pp. 283-304). Atlantis Press.

- Hosny, K. M., El-Hady, W. M., Samy, F. M., Vrochidou, E., & Papakostas, G. A. (2023). Multi-Class Classification of Plant Leaf Diseases Using Feature Fusion of Deep Convolutional Neural Network and Local Binary Pattern. *IEEE Access*, 11, 62307–62317. <https://doi.org/10.1109/ACCESS.2023.3286730>
- Hossain, S., Tanzim Reza, M., Chakrabarty, A., & Jung, Y. J. (2023). Aggregating Different Scales of Attention on Feature Variants for Tomato Leaf Disease Diagnosis from Image Data: A Transformer Driven Study. *Sensors*, 23(7). <https://doi.org/10.3390/s23073751>
- Jackulin, C., & Murugavalli, S. (2022). A comprehensive review on detection of plant disease using machine learning and deep learning approaches. *Measurement: Sensors*, 24, 100441.
- Javaid, M., Haleem, A., Khan, I. H., & Suman, R. (2023). Understanding the potential applications of Artificial Intelligence in Agriculture Sector. *Advanced Agrochem*, 2(1), 15-30.
- Jesse, A., & Lenka, P. (2024). Image Classification for CSSVD Detection in Cacao Plants. *arXiv preprint arXiv:2405.04535*.
- Li, G., Wang, Y., Zhao, Q., Yuan, P., & Chang, B. (2023). PMVT: a lightweight vision transformer for plant disease identification on mobile devices. *Frontiers in Plant Science*, 14, 1256773.
- Li, X., & Li, S. (2022). Transformer help CNN see better: a lightweight hybrid apple disease identification model based on transformers. *Agriculture*, 12(6), 884.
- Li, X., Li, X., Zhang, S., Zhang, G., Zhang, M., & Shang, H. (2023). SLViT: Shuffle-convolution-based lightweight Vision transformer for effective diagnosis of sugarcane leaf diseases. *Journal of King Saud University-Computer and Information Sciences*, 35(6), 101401.

- Lopes, J. F., da Costa, V. G. T., Barbin, D. F., Cruz-Tirado, L. J. P., Baeten, V., & Barbon Junior, S. (2022). Deep computer vision system for cocoa classification. *Multimedia Tools and Applications*, 81(28), 41059-41077.
- Lu, X., Yang, R., Zhou, J., Jiao, J., Liu, F., Liu, Y., ... & Gu, P. (2022). A hybrid model of ghost-convolution enlightened transformer for effective diagnosis of grape leaf disease and pest. *Journal of King Saud University-Computer and Information Sciences*, 34(5), 1755-1767.
- Lu, Y., & Young, S. (2020). A survey of public datasets for computer vision tasks in precision agriculture. *Computers and Electronics in Agriculture*, 178, 105760.
- Lye, H. Z. M., & Ng, K. W. (2023). Processing Plant Diseases Using Transformer Model. *JOIV: International Journal on Informatics Visualization*, 7(4), 2550-2557.
- Mamadou, D., Kacoutchy, J. A., Ballo, A. B., & Kouassi, B. M. (2023). Cocoa Pods Diseases Detection by MobileNet Confluence and Classification Algorithms. *International Journal of Advanced Computer Science and Applications*, 14(9).
- Madenda, S. (2015). *PENGOLAHAN CITRA & VIDEO DIGITAL*. Jakarta: Erlangga.
- Önler, E. (2023). Feature fusion based artificial neural network model for disease detection of bean leaves. *Electronic Research Archive*, 31(5), 2409–2427. <https://doi.org/10.3934/era.2023122>
- Quoc Bao, T., Tan Kiet, N. T., Quoc Dinh, T., & Hiep, H. X. (2020). Plant species identification from leaf patterns using histogram of oriented gradients feature space and convolution neural networks. *Journal of Information and Telecommunication*, 4(2), 140-150.
- Rachmad, A., Syarief, M., Rifka, S., Sonata, F., Setiawan, W., & Rochman, E. M. S. (2022, December). Corn leaf disease classification using local binary patterns

- (LBP) feature extraction. *In Journal of Physics: Conference Series* (Vol. 2406, No. 1, p. 012020). IOP Publishing.
- Rodriguez, C., Alfaro, O., Paredes, P., Esenarro, D., & Hilario, F. (2021). *Machine Learning Techniques in the Detection of Cocoa (Theobroma cacao L.) Diseases* (Vol. 25). <http://annalsofrscb.ro>
- Roy, A. M., & Bhaduri, J. (2021). A deep learning enabled multi-class plant disease detection model based on computer vision. *Ai*, 2(3), 413-428.
- Soh, K. S., Mounq, E. G., Danker, K. J. J., Dargham, J. A., & Farzamnia, A. (2024). Cocoa Diseases Classification using Deep Learning Algorithm. *In ITM Web of Conferences* (Vol. 63, p. 01014). EDP Sciences.
- Sun, C., Zhou, X., Zhang, M., & Qin, A. (2023). SE-VisionTransformer: Hybrid Network for Diagnosing Sugarcane Leaf Diseases Based on Attention Mechanism. *Sensors*, 23(20), 8529.
- Tabbakh, A., & Barpanda, S. S. (2023). A Deep Features extraction model based on the Transfer learning model and vision transformer" TLMViT" for Plant Disease Classification. *IEEE Access*.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- Vera, D. B., Oviedo, B., Casanova, W. C., & Zambrano-Vega, C. (2024). Deep Learning-Based Computational Model for Disease Identification in Cocoa Pods (Theobroma cacao L.). *arXiv preprint arXiv:2401.01247*.
- Yang, B., Li, M., Li, F., Wang, Y., Liang, Q., Zhao, R., ... & Wang, J. (2024). A novel plant type, leaf disease and severity identification framework using CNN and transformer with multi-label method. *Scientific Reports*, 14(1), 11664.

Yong, W. C., Ng, K. W., Haw, S. C., Naveen, P., & Ng, S. B. (2024). Leaf Condition Analysis Using Convolutional Neural Network and Vision Transformer. *International Journal of Computing and Digital Systems*, 16(1), 1-10.

Yong, W. C., Ng, K. W., Haw, S. C., Naveen, P., & Ng, S. B. (2024). Leaf Condition Analysis Using Convolutional Neural Network and Vision Transformer. *International Journal of Computing*