

BAB II

TINJAUAN PUSTAKA

2.1. Penelitian Terakit

Sejumlah penelitian telah mengaplikasikan algoritma deep learning dalam upaya mendeteksi penyakit pada tanaman kentang. Berbagai metode dan pendekatan yang digunakan dalam penelitian-penelitian tersebut memberikan gambaran mengenai kemajuan dan efektivitas teknik deep learning dalam klasifikasi penyakit tanaman. Tabel 2.1 menyajikan ringkasan penelitian terkait yang mencakup informasi mengenai penulis, judul artikel, tahun publikasi, metode yang digunakan, serta hasil yang diperoleh. Data ini akan dijadikan sebagai referensi dan pembanding dalam pelaksanaan penelitian ini.

Tabel 2. 1. Penenlitian Terkait

Penulis	Judul	Tahun	Metode	Hasil
Chang et al. [7]	Potato Leaf Disease Detection Based on a Lightweight Deep Learning Model.	2024	QueuezeNet 1.1, MNASNet1_3, ShuffleNetV2_ x2_0, EfficientNet_B 1, RegNetY_400 MF	Penelitian ini menggunakan model RegNetY-400MF yang dikombinasikan dengan teknik transfer learning dan augmentasi data. Model berhasil mencapai akurasi tertinggi sebesar 89.87 % dalam mengklasifikasikan tujuh jenis penyakit pada daun kentang.
Bajpai et al. [8]	An Efficient Approach for	2024	ResNet50V2 + Enhanced	Penelitian ini menggabungkan

	Potato Leaf Disease Classification Using Cascaded CNN-Transformers.		Vision Transformer (ViT)	ResNet50V2 dengan Vision Transformer (ViT) yang dimodifikasi untuk meningkatkan kemampuan ekstraksi dan representasi fitur. Hasilnya, model mencapai akurasi sebesar 95% dan rata-rata F1-score sebesar 0.95 dalam mengklasifikasikan penyakit daun kentang,
Chowdhury et al. [9]	Harnessing the Potato leaf disease detection process through proposed Conv2D and resnet50 deep learning models.	2025	ResNet50 dan CNN	Penelitian ini menggunakan arsitektur CNN dan ResNet50 dalam mendeteksi penyakit pada daun kentang. Hasil penelitian menunjukkan bahwa ResNet50 memberikan kinerja yang lebih baik dengan akurasi validasi mencapai 97.24% dan rata-rata F1-score sebesar 0.97.

Kaur et al. [10]	Automating Potato Leaf Disease Detection with Lightning-Fast CNNs: Precision using PyTorch Lightning.	2024	CNN + PyTorch Lightning	Penelitian ini menggunakan arsitektur CNN yang dilatih dengan PyTorch Lightning. Model berhasil mencapai akurasi sebesar 98.6%.
Boukhilfa et al. [11]	DSCSkipNet: An Accuracy-Complexity Trade-off for Effective Potato Disease Identification in Uncontrolled Environments.	2024	DSCSkipNet	Penelitian ini menggunakan arsitektur lightweight CNN yaitu DSCSkipNet yang menggabungkan Depthwise Separable Convolution dan Skip Connections untuk mendeteksi penyakit daun kentang. Hasilnya, model mencapai akurasi rata-rata 80.45%.
Sharma et al. [12]	PotatoLeafGuard: A Novel Hybrid CNN Architecture for Advanced Detection of	2024	Arsitektur hybrid CNN yang menggabungkan CNN dengan model transfer learning	Penelitian ini menggabungkan CNN dan transfer learning menggunakan arsitektur Xception dan VGG16 yang

	Potato Leaf Diseases.		Xception dan VGG16.	dinamakan PotatoLeafGuard. Model tersebut berhasil mencapai akurasi klasifikasi sebesar 94,70%, precision 94%, recall 96%, dan F1-score 97%.
Fathima M & Booba B. [13]	Enhancing Plant Leaf Disease Detection: Integrating MobileNet With Local Binary Pattern and Visualizing Insights With Grad-CAM.	2024	CNN (MobileNet) + fitur tekstur (LBP) + visual interpretasi (Grad-CAM)	Penelitian ini menggunakan pendekatan deteksi penyakit daun tanaman dengan menggabungkan arsitektur MobileNet dan teknik Local Binary Pattern (LBP). Hasil penelitian menunjukkan performa yang sangat baik, dengan akurasi mencapai 98% dan nilai F1-score sebesar 97.9%
Mhala et al. [14]	Enhancing crop productivity with fine-tuned deep	2025	Deep Convolutional Neural Network + fine tuning pada	Model yang dikembangkan mencapai akurasi klasifikasi sebesar 96.53%, dengan

	convolution neural network for Potato leaf disease detection.		model pretrained	precision 96.70%, recall 96.64%, dan F1-score 96.66%.
Firasari et al. [15]	Implementasi Deep Learning pada Deteksi Penyakit Daun Kentang dengan Arsitektur InceptionResN etV2.	2025	InceptionResN etV2	Model menunjukkan performa yang baik dengan akurasi pelatihan sebesar 98,02% dan validasi 95.67%. Namun, masih terdapat tantangan dalam membedakan kelas Late Blight dan Healthy, terlihat dari banyaknya salah klasifikasi pada kelas tersebut.
Rozaqi et al. [16]	Deteksi Penyakit pada Daun Kentang Menggunakan Pengolahan Citra dengan Metode CNN	2021	CNN	Model mencapai akurasi pelatihan 95,27% dan validasi 94,09% setelah 10 epoch. Hasil menunjukkan performa yang baik dan stabil, ditunjukkan oleh tren akurasi yang

				meningkat dan loss yang menurun.
Fuadi et al. [17]	Perbandingan Arsitektur MobileNet dan NASNetMobile untuk Klasifikasi Penyakit pada Citra Daun Kentang.	2022	MobileNet, NASNetMobile (CNN)	NASNetMobile memberikan hasil terbaik dengan akurasi 90.96% NASNetMobile unggul di hampir semua metrik dibanding MobileNet. Oversampling membantu menyeimbangkan dataset.

Penelitian terkait dilakukan oleh Chang dan Lai (2024) yang berfokus pada pengembangan model deep learning untuk klasifikasi penyakit daun kentang. Penelitian ini menggunakan arsitektur RegNetY-400MF karena efisiensinya dari sisi kompleksitas model, serta dilengkapi dengan modul Squeeze-and-Excitation (SE) untuk memperkuat representasi fitur. Dengan memanfaatkan teknik transfer learning dan data augmentation, model berhasil mencapai akurasi tertinggi sebesar 89,87 % dan rata-rata F1-score sebesar 89,83%, menunjukkan kinerja yang seimbang antara akurasi dan efisiensi.

Penelitian lain yang dilakukan oleh Bajpai et al. (2024) menggunakan pendekatan *hybrid* yang mengintegrasikan ResNet50V2 dengan Vision Transformer (ViT). Kombinasi ini dimaksudkan untuk menggabungkan kemampuan CNN dalam mengekstraksi fitur spasial lokal dengan keunggulan Transformer dalam menangkap konteks global. Model tersebut diterapkan untuk klasifikasi tiga kelas utama, yaitu early blight, late blight, dan healthy menghasilkan akurasi pengujian sebesar 95% dan F1-score rata-rata sebesar 0.95, hasil tersebut

menandakan bahwa pendekatan cascaded CNN-Transformer efektif dalam meningkatkan akurasi klasifikasi citra daun kentang.

Selanjutnya, penelitian yang dilakukan oleh Chowdhury dan Das (2025) membandingkan performa model CNN dengan ResNet50 dalam tugas klasifikasi daun kentang berdasarkan tiga kelas penyakit. Data diolah melalui skema pelatihan selama 20 epoch, hasilnya menunjukkan bahwa model ResNet50 menghasilkan akurasi validasi mencapai 97.24%. Adapun penelitian yang dilakukan oleh Kaur tahun 2024 memanfaatkan framework PyTorch Lightning untuk mengembangkan model CNN yang lebih modular, efisien, dan mudah dikembangkan. Penelitian ini diarahkan pada klasifikasi penyakit seperti early blight, late blight, potato scab, blackleg, dan healthy. Dengan dukungan Adam optimizer dan teknik augmentasi data yang sesuai, model CNN yang dibangun mampu mencapai akurasi klasifikasi hingga 98.6%, yang menunjukkan efisiensi pelatihan dan akurasi yang tinggi dalam mendeteksi berbagai jenis penyakit daun kentang.

Penelitian selanjutnya dilakukan oleh Boukhelifa dan Chibani (2024) mengusulkan arsitektur DCSkipNet yang dirancang untuk mendeteksi penyakit daun kentang. Arsitektur ini menggabungkan Depthwise Separable Convolutions (DSC), Skip Connections (SC), serta Pointwise Convolution (PC) untuk menyeimbangkan efisiensi dan akurasi. Melalui skema 4-fold cross-validation, model ini mampu mencapai akurasi rata-rata sebesar 80.45%, dengan macro-F1 score 76%, weighted-F1 score 80%, dan waktu inferensi tercepat sebesar 0.0259 detik.

Penelitian selanjutnya dilakukan oleh Meenakshi Sharma et al. (2024) yang mengembangkan model deep learning bernama PotatoLeafGuard, arsitektur hybrid CNN yang menggabungkan model CNN dasar dengan pendekatan transfer learning menggunakan Xception dan VGG16. Model dilatih menggunakan dataset daun kentang yang diklasifikasikan ke dalam tujuh kelas, yaitu daun sehat serta daun yang terinfeksi oleh jamur, virus, bakteri, hama (pests), nematoda, dan phytophthora. Untuk meningkatkan performa dan mengatasi keterbatasan data, dilakukan augmentasi seperti rotasi, flipping, scaling, dan normalisasi piksel. Hasil

evaluasi menunjukkan bahwa PotatoLeafGuard mampu mencapai akurasi 94.70%, dengan precision 94%, recall 96%, F1-score 97%, dan nilai AUC-ROC sebesar 0.97.

Penelitian yang dilakukan oleh Fathima M. S dan Dr. Booba B (2024) menggunakan arsitektur deep learning yang menggabungkan MobileNetV2 sebagai ekstraktor fitur global dan Local Binary Pattern (LBP) sebagai ekstraktor fitur tekstur lokal untuk mendeteksi penyakit pada daun tanaman. Fitur dari MobileNet dan LBP dikombinasikan pada lapisan fully connected, kemudian diklasifikasikan menggunakan fungsi aktivasi softmax. Selain itu, digunakan Grad-CAM untuk memvisualisasikan bagian citra yang paling berkontribusi terhadap prediksi, sehingga meningkatkan interpretabilitas model. Hasil pengujian menunjukkan bahwa model gabungan ini mampu mencapai akurasi 98%, precision 97%, recall 97.6%, dan F1-score 97.9%, lebih tinggi dibandingkan model lain seperti CNN, VGG16, GoogleNet, dan AlexNet.

Penelitian oleh Pradeep Kumar (2024) fokus pada pengembangan model deep learning untuk mendeteksi penyakit pada daun kentang guna mendukung peningkatan produktivitas pertanian. Dalam penelitian ini, penulis menggunakan arsitektur Deep Convolutional Neural Network (DCNN) yang sudah dilatih sebelumnya, kemudian dilakukan fine-tuning agar model lebih adaptif terhadap karakteristik data daun kentang. Dataset yang digunakan bersumber dari PlantVillage dan mencakup tujuh kelas, termasuk daun sehat dan daun yang terinfeksi berbagai jenis penyakit seperti virus, bakteri, dan jamur. Untuk memperkaya variasi data dan meningkatkan kemampuan generalisasi model, diterapkan beberapa teknik augmentasi seperti rotasi, zoom, dan flipping. Model dievaluasi menggunakan metrik akurasi, precision, recall, dan F1-score. Hasilnya, model mampu mengklasifikasikan citra dengan akurasi hingga 97.8% dan F1-score sebesar 96.3%, menunjukkan bahwa pendekatan ini cukup efektif untuk mendeteksi penyakit daun kentang secara otomatis.

Penelitian lain yang dilakukan oleh Elly Firasari dan F. Lia Dwi Cahyanti (2025) menggunakan arsitektur InceptionResNetV2 untuk mengembangkan sistem

klasifikasi penyakit daun kentang berbasis citra. Model ini dipilih karena mampu dalam menangani fitur multi-skala melalui modul Inception serta stabilitas pelatihan yang ditingkatkan oleh residual connections. Setelah melalui tahap praproses berupa resize, normalisasi piksel, dan augmentasi data, model dilatih menggunakan citra berukuran 299x299 piksel dengan batch size 20 selama 20 epoch. Hasil pelatihan menunjukkan akurasi mencapai 98.02% dan validasi 95.67%. Meskipun performanya tinggi, model masih menunjukkan kelemahan dalam membedakan citra Late Blight dari Healthy, terbukti dari 81 kesalahan klasifikasi pada kelas tersebut.

Penelitian lain dilakukan oleh Abdul Jalil Rozaqi, Andi Sunyoto, dan Rudyanto Arief (2021) yang mengusulkan metode klasifikasi penyakit daun kentang menggunakan algoritma Convolutional Neural Network (CNN). Citra dikonversi menjadi ukuran 150x150 piksel dan dibagi ke dalam data pelatihan dan pengujian dengan rasio 80:20. Arsitektur CNN yang digunakan terdiri dari empat lapisan konvolusi, max pooling, flatten layer, dan dua dense layer. Hasil pelatihan menunjukkan performa yang cukup tinggi, dengan akurasi pelatihan mencapai 95,27% dan akurasi validasi sebesar 94,09% pada epoch ke-10. Grafik akurasi dan loss juga menunjukkan tren konvergen yang stabil, menandakan pelatihan berjalan dengan baik. Namun, distribusi data yang tidak seimbang, di mana jumlah citra daun sehat jauh lebih sedikit dibanding dua kelas lainnya menjadi salah satu kekurangan yang dapat memengaruhi akurasi prediksi pada kelas minoritas tersebut.

Penelitian lain dilakukan oleh Anwar Fuadi dan Aries Suharso (2022) yang membandingkan performa dua arsitektur convolutional neural network, yaitu MobileNet dan NASNetMobile, dalam klasifikasi penyakit daun kentang. Dalam mengatasi ketidakseimbangan data, penulis melakukan oversampling sehingga jumlah citra pada setiap kelas menjadi sama, yakni 1.628 gambar. Data kemudian diproses dengan augmentasi dan dibagi ke dalam lima skema rasio pelatihan-pengujian (mulai dari 90:10 hingga 50:50). Hasil eksperimen menunjukkan bahwa arsitektur NASNetMobile memberikan performa terbaik pada skema 90:10 dengan

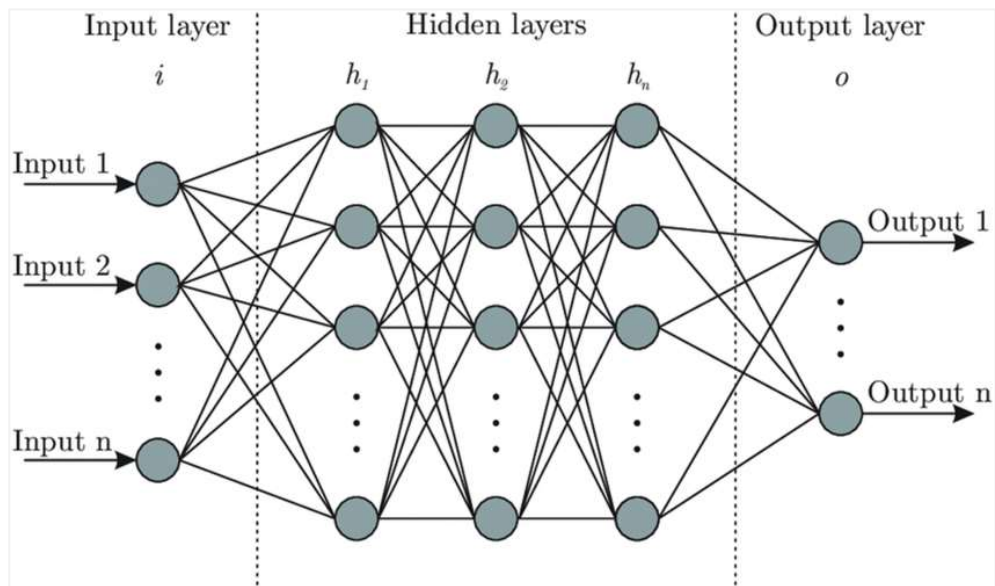
akurasi sebesar 90,96%, precision 90,86%, recall 91,11%, dan F1-score 92,93%. Secara keseluruhan, NASNetMobile unggul dibandingkan MobileNet dalam semua skema pembagian data. Penelitian ini juga menegaskan bahwa pemilihan arsitektur dan teknik penyeimbangan data memainkan peran penting dalam meningkatkan akurasi klasifikasi citra penyakit tanaman.

Temuan-temuan tersebut menjadi pijakan penting dalam pengembangan metode deteksi penyakit daun kentang yang lebih akurat dan efisien pada penelitian ini. Berbagai teknik dan pendekatan yang digunakan dalam penelitian-penelitian sebelumnya memberikan gambaran menyeluruh mengenai strategi yang efektif, sehingga dapat dijadikan acuan dalam merancang sistem klasifikasi yang aplikatif.

2.2. Tinjauan Teori

2.2.1. Deep learning

Deep learning merupakan salah satu cabang dari Machine Learning yang algoritmanya didasarkan pada struktur jaringan saraf manusia. Struktur ini dikenal dengan Artificial Neural Network (ANN) atau dalam bahasa Indonesia disebut Jaringan Saraf Tiruan (JST). Deep learning menggunakan konsep hierarki, konsep ini memungkinkan komputer untuk mempelajari konsep-konsep kompleks dengan menggabungkan elemen-elemen yang lebih sederhana [18]. Jika diilustrasikan sebagai sebuah grafik, konsep-konsep tersebut dibangun secara bertahap diatas konsep lainnya, sehingga membentuk struktur yang dalam dengan berbagai lapisan



Gambar 2.1. Arsitektur Deep Learning [19]

Arsitektur deep learning pada Gambar 2.1 merupakan struktur dari Artificial Neural Network (ANN) yang terdiri dari tiga komponen utama, yaitu lapisan input, lapisan tersembunyi (hidden layers), dan lapisan output. Pada lapisan input, data atau fitur yang akan diproses oleh model dimasukkan ke dalam jaringan. Data ini kemudian diteruskan ke lapisan tersembunyi (hidden layers) yang terdiri dari sejumlah unit atau neuron yang saling terhubung. Hasil dari lapisan tersembunyi diteruskan ke lapisan output yang menghasilkan prediksi atau keputusan akhir dari model, seperti klasifikasi atau regresi. Proses pelatihan model dilakukan dengan cara menyesuaikan bobot yang menghubungkan neuron-neuron di setiap lapisan, sehingga model dapat menghasilkan prediksi yang lebih akurat.

Kemampuan utama dari deep learning terletak pada kemampuannya untuk melakukan feature extraction secara otomatis tanpa memerlukan rekayasa fitur (feature engineering) manual seperti pada metode machine learning konvensional [20]. Dalam konteks pengolahan citra, deep learning mampu mengenali pola visual kompleks seperti tepi objek, tekstur, bentuk, hingga karakteristik spesifik dari gambar secara berlapis. Lapisan awal pada jaringan cenderung mempelajari fitur dasar seperti garis dan sudut, sementara

lapisan yang lebih dalam mempelajari fitur yang lebih abstrak, seperti bentuk spesifik daun atau pola penyakit tertentu. Hal tersebut membuat deep learning cocok digunakan dalam tugas klasifikasi citra daun tanaman, karena mampu menangani kompleksitas visual yang tinggi dan variasi antar kelas dengan lebih efektif. Deep learning memiliki berbagai arsitektur yang dapat disesuaikan dengan jenis permasalahan yang dihadapi. Beberapa pendekatan umum dalam deep learning antara lain:

1. Convolutional Neural Network (CNN)

Pendekatan dengan CNN banyak digunakan dalam pengolahan citra karena kemampuannya dalam mengenali fitur spasial dari gambar secara otomatis [21]. CNN menggunakan layer konvolusi untuk mengekstraksi pola dari citra dan cocok untuk tugas klasifikasi objek, deteksi, maupun segmentasi. Struktur dasar CNN terdiri atas tiga komponen utama, yaitu convolutional layers, pooling layers, dan fully connected layers.

Convolutional layers bertugas mengekstraksi fitur dari input citra. Pooling layers digunakan untuk mereduksi dimensi data sambil mempertahankan informasi penting, sehingga mengurangi kompleksitas komputasi dan menghindari overfitting. Setelah beberapa tahapan konvolusi dan pooling, hasil ekstraksi fitur ini kemudian diteruskan ke fully connected layers untuk proses klasifikasi atau prediksi akhir.

Keunggulan CNN terletak pada kemampuannya untuk belajar secara otomatis dari data visual tanpa memerlukan proses rekayasa fitur (feature engineering) secara manual [22]. Selain itu, CNN juga memiliki kemampuan translational invariance, yaitu kemampuan mengenali objek meskipun posisinya berubah [23].

2. Recurrent Neural Network (RNN)

Recurrent Neural Network (RNN) merupakan salah satu arsitektur deep learning yang dirancang untuk mengolah data yang memiliki

urutan atau keterkaitan antar waktu, seperti teks, audio, atau data sensor. Arsitektur ini bekerja dengan mempertahankan informasi dari input sebelumnya melalui mekanisme *hidden state*, sehingga mampu mengenali pola berdasarkan urutan kejadian.

Setiap unit dalam RNN menerima input tidak hanya dari data saat ini, tetapi juga dari keluaran sebelumnya, menciptakan semacam memori jangka pendek yang memungkinkan jaringan memahami konteks [24]. Karakteristik ini sangat dibutuhkan dalam situasi di mana urutan atau konteks historis memengaruhi keluaran, seperti dalam pemrosesan bahasa alami (*Natural Language Processing*) atau pemantauan kondisi lingkungan secara berkala.

3. Generative Adversarial Network (GAN)

Generative Adversarial Network (GAN) merupakan arsitektur deep learning yang dirancang untuk menghasilkan data baru yang menyerupai data asli. GAN terdiri dari dua jaringan saraf utama, yaitu generator dan discriminator. Generator bertugas menciptakan data sintetis, sementara discriminator berfungsi untuk membedakan antara data asli dan data hasil buatan [25]. Proses pelatihan berlangsung secara berulang, di mana kedua jaringan saling mengasah kemampuannya hingga generator mampu menghasilkan data yang sangat mirip dengan data nyata.

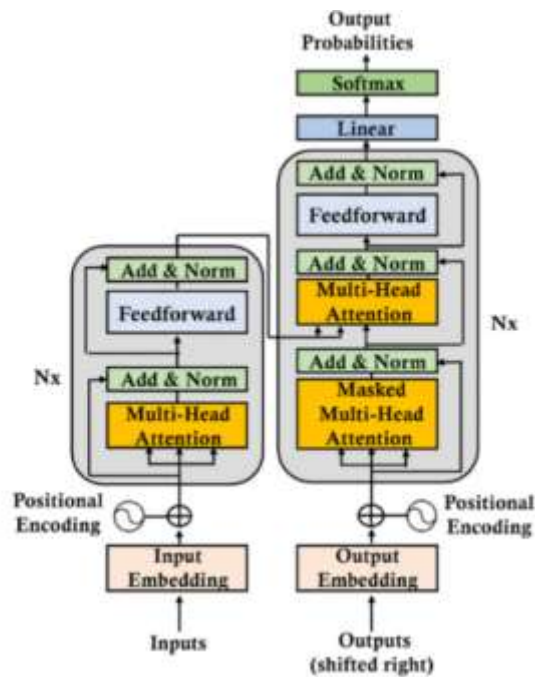
4. Autoencoder

Autoencoder merupakan salah satu arsitektur deep learning yang dirancang untuk mempelajari representasi terkompresi dari data input. Arsitektur ini terdiri dari dua komponen utama, yaitu encoder dan decoder [26]. Encoder bertugas mengubah data input menjadi representasi berdimensi lebih rendah (latent space), sedangkan decoder mencoba merekonstruksi kembali data asli dari representasi tersebut.

Tujuan utama dari proses ini adalah agar representasi yang dihasilkan dapat menangkap fitur-fitur penting dari data secara efisien. Dalam pengolahan citra, autoencoder banyak dimanfaatkan untuk berbagai tugas seperti reduksi dimensi, denoising (penghilangan noise), dan rekonstruksi gambar [27]. Misalnya, dalam tugas penghilangan noise, autoencoder dilatih untuk mengenali pola visual yang bersih dari data citra yang telah terkontaminasi noise, sehingga model dapat menghasilkan versi gambar yang lebih jernih. Pendekatan ini sangat membantu ketika citra hasil akuisisi memiliki kualitas rendah akibat gangguan pencahayaan atau sensor.

2.2.2. Transformer

Transformer merupakan salah satu arsitektur jaringan saraf yang dirancang untuk mengolah data berurutan dengan efisiensi dan akurasi yang tinggi. Arsitektur ini menjadi salah satu inovasi penting dalam bidang Natural Language Processing (NLP) dan kini juga banyak diadaptasi dalam bidang Computer Vision (CV) [28]. Transformer menjadi terobosan besar dalam pengolahan data sekuensial karena mengatasi keterbatasan model sebelumnya seperti RNN dan LSTM, khususnya dalam menangani ketergantungan jarak jauh dalam data. Dengan mekanisme attention yang dapat memproses seluruh input secara paralel, Transformer memungkinkan efisiensi komputasi yang jauh lebih baik, terutama dalam proses pelatihan dengan data dalam jumlah besar.



Gambar 2. 2. Arsitektur Transformer [29]

Arsitektur transformer pada Gambar 2.2 terdiri atas dua komponen utama, yaitu encoder dan decoder, yang masing-masing terdiri dari beberapa lapisan dengan struktur yang seragam. Encoder bertugas mengolah data input dan mengubahnya menjadi representasi vektor yang mengandung informasi kontekstual. Setiap lapisan encoder terdiri dari dua komponen utama, yaitu multi-head attention dan feedforward neural network. Komponen attention berfungsi untuk memungkinkan model memperhatikan berbagai bagian dari input secara paralel, sedangkan feedforward network bertugas memproses representasi yang dihasilkan dari attention. Kedua komponen ini dilengkapi dengan residual connection dan layer normalization guna meningkatkan stabilitas serta efisiensi proses pelatihan.

Di sisi lain, decoder berfungsi menghasilkan output berdasarkan informasi yang diperoleh dari encoder. Struktur pada decoder sedikit lebih kompleks karena terdiri dari tiga komponen utama, yaitu masked multi-head attention, multi-head attention yang mengintegrasikan informasi dari encoder, serta feedforward neural network. Masked attention diterapkan untuk memastikan bahwa token pada posisi tertentu hanya dapat memperhatikan

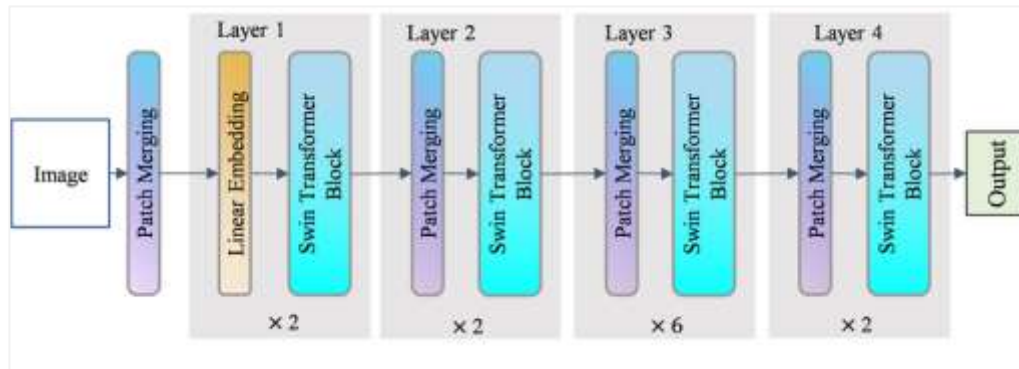
token sebelumnya, sehingga proses prediksi berjalan secara bertahap dan sesuai dengan urutan yang benar.

Sebelum memasuki tahap pemrosesan oleh encoder maupun decoder, data terlebih dahulu diubah ke dalam bentuk vektor melalui proses embedding. Selain itu, positional encoding juga ditambahkan untuk memberikan informasi posisi relatif dari masing-masing token dalam urutan, mengingat arsitektur Transformer tidak memiliki mekanisme pengurutan bawaan seperti pada model berbasis RNN. Setelah seluruh proses selesai, keluaran dari decoder diteruskan ke lapisan linear dan softmax untuk menghasilkan output akhir dalam bentuk token dengan probabilitas tertinggi.

Setelah seluruh proses selesai, keluaran dari decoder diteruskan ke lapisan linear dan softmax untuk menghasilkan output akhir dalam bentuk token dengan probabilitas tertinggi. Kemampuan Transformer dalam menangani paralelisme secara efisien menjadikannya sangat unggul dalam pelatihan skala besar. Hal ini mendorong pengembangan berbagai varian arsitektur Transformer yang lebih spesifik terhadap domain tertentu, seperti BERT, GPT, dan ViT (Vision Transformer), yang masing-masing menyesuaikan struktur dasar Transformer untuk kebutuhan yang berbeda.

2.2.3. Swin Transformer

Swin Transformer (Shifted Window Transformer) merupakan salah satu arsitektur deep learning berbasis Vision Transformer (ViT) yang dirancang khusus untuk menangani tugas-tugas pengolahan citra [30]. Arsitektur ini merupakan solusi atas keterbatasan Vision Transformer dalam hal efisiensi komputasi dan penerapan struktur hierarkis yang penting dalam visi komputer. Berbeda dengan ViT yang menerapkan self-attention secara global, Swin Transformer memperkenalkan pendekatan lokal dengan menggunakan jendela (window) yang saling bergeser antar lapisan, sehingga memungkinkan pertukaran informasi antar area gambar tanpa mengorbankan efisiensi. [31]



Gambar 2. 3. Arsitektur Swin Transformer [32]

Arsitektur Swin Transformer pada Gambar 2.3 dirancang secara hierarkis dan efisien untuk menangani berbagai tugas visi komputer, seperti klasifikasi gambar, segmentasi citra, hingga deteksi objek. Seperti yang ditunjukkan pada Gambar 2.X, proses dimulai dari input citra berdimensi $H \times W \times 3$, yang kemudian dibagi menjadi patch non-overlapping melalui tahap Patch Partition. Setiap patch kemudian di-flatten dan diproyeksikan ke dimensi fitur tertentu menggunakan Linear Embedding, menghasilkan token awal yang merepresentasikan fitur dari masing-masing patch citra.

Setelah itu, arsitektur Swin Transformer dibagi menjadi empat tahap (stage) utama. Stage 1 memproses fitur awal melalui dua blok Swin Transformer yang menerapkan window-based multi-head self-attention (W-MSA) dan shifted window multi-head self-attention (SW-MSA) secara bergantian. Pada tahap ini, resolusi fitur tetap, namun informasi spasial lokal mulai dipelajari. Memasuki Stage 2, dilakukan proses Patch Merging yang menggabungkan patch tetangga (misalnya 2×2) menjadi satu patch dengan dimensi fitur yang lebih besar, sehingga resolusi menurun namun kedalaman informasi meningkat. Tahap ini juga terdiri dari dua blok Swin Transformer. Selanjutnya, Stage 3 memperdalam representasi fitur dengan enam blok Swin Transformer setelah proses patch merging, menghasilkan output dengan resolusi yang lebih rendah namun fitur yang lebih kompleks. Terakhir, Stage 4 melanjutkan pola serupa dengan dua blok Swin Transformer dan menghasilkan fitur akhir dengan resolusi paling rendah namun memiliki konteks global yang paling luas.

Salah satu keunggulan utama dari Swin Transformer adalah kemampuannya dalam menjaga efisiensi komputasi [33]. Dengan hanya melakukan self-attention dalam jendela lokal dan menggeser jendela tersebut antar lapisan, arsitektur ini mampu menangkap hubungan antar area gambar secara efektif tanpa memerlukan komputasi intensif seperti pada Vision Transformer standar.

2.3.4. Knowledge Discovery Database

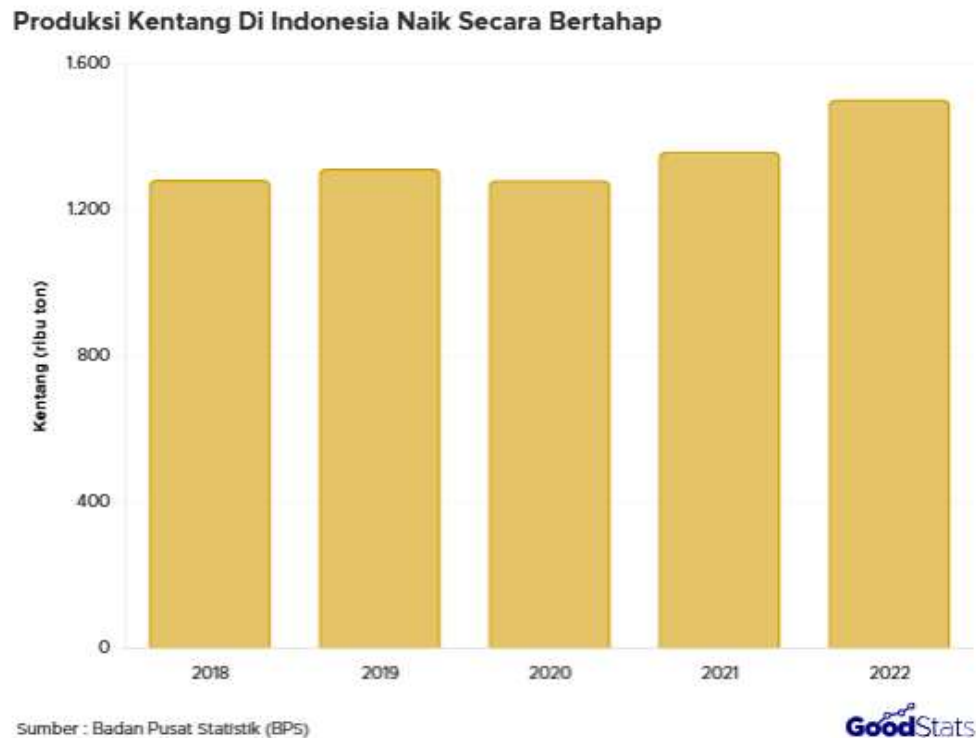
Knowledge Discovery in Database (KDD) adalah suatu proses sistematis untuk menemukan pengetahuan baru yang bermakna, tersembunyi, dan bermanfaat dari kumpulan data yang besar. KDD merupakan rangkaian tahap yang dirancang untuk mengeksplorasi data, mengidentifikasi pola, serta mengekstrak informasi yang dapat digunakan untuk pengambilan keputusan. Proses KDD menjadi landasan dalam bidang data mining, machine learning, dan analitik data karena mampu mengubah data mentah menjadi pengetahuan yang berguna untuk pengambilan keputusan.

2.3.5. Tanaman Kentang

Kentang (*Solanum tuberosum* L.) adalah tanaman umbi-umbian yang termasuk dalam keluarga Solanaceae. Kentang merupakan salah satu komoditas pertanian penting yang dibudidayakan secara luas di seluruh dunia, karena nilai gizi yang tinggi serta kemudahan dalam pengolahan dan konsumsi [34]. Umbi kentang dikenal dengan kaya akan karbohidrat, terutama pati, dan juga mengandung sejumlah vitamin, mineral, serta serat yang penting bagi kesehatan manusia [35]. Tanaman ini biasanya diproduksi melalui pembiakan vegetatif menggunakan umbi sebagai bibit. Selain dimanfaatkan sebagai bahan pangan utama, kentang juga digunakan dalam industri makanan olahan, pembuatan tepung, serta dalam berbagai aplikasi industri lainnya [36].

Sebagai komoditas strategis, kentang menunjukkan tren peningkatan produksi di Indonesia dalam beberapa tahun terakhir. Data dari Badan Pusat Statistik (BPS) menunjukkan bahwa produksi kentang nasional terus

meningkat secara bertahap dari tahun 2018 hingga 2022, sebagaimana ditunjukkan pada Gambar 2.4 berikut.



Gambar 2. 4. Produksi Kentang Di Indonesia

Di banyak negara, termasuk Indonesia, kentang menjadi salah satu tanaman hortikultura unggulan yang memberikan kontribusi signifikan terhadap pendapatan petani dan perekonomian daerah, khususnya di wilayah dataran tinggi. Kentang memiliki siklus tanam yang relatif singkat, sekitar 90–120 hari, sehingga memungkinkan petani untuk melakukan rotasi tanaman secara efisien [37]. Selain itu, nilai jual yang relatif tinggi dibandingkan tanaman pangan lainnya menjadikan kentang sebagai sumber penghasilan utama bagi banyak petani kecil.

Namun, budidaya kentang juga menghadapi berbagai tantangan, di antaranya serangan hama dan penyakit tanaman seperti hawar daun (late blight) yang disebabkan oleh *Phytophthora infestans*. Penyakit ini dapat menyebabkan gagal panen jika tidak ditangani dengan cepat dan tepat. Oleh

karena itu, praktik budidaya yang baik (Good Agricultural Practices), penggunaan bibit unggul yang tahan penyakit, serta penerapan teknologi modern seperti deteksi penyakit berbasis citra dan kecerdasan buatan menjadi penting untuk meningkatkan produktivitas dan ketahanan tanaman. Dalam konteks keberlanjutan, kentang juga memiliki potensi besar sebagai bahan baku bioenergi dan bioplastik karena kandungan patinya yang tinggi. Hal ini menjadikan kentang tidak hanya relevan sebagai komoditas pangan, tetapi juga sebagai bagian dari solusi inovatif untuk mendukung ekonomi sirkular dan industri ramah lingkungan.