

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terkait

Tabel 1.1 merangkum studi-studi sebelumnya yang mengimplementasikan metode deep learning dalam klasifikasi dan deteksi penyakit pada tanaman kentang. Temuan dari penelitian-penelitian tersebut digunakan sebagai dasar perbandingan serta untuk menyoroti adanya celah penelitian yang akan dikaji.

Tabel 2. 1 Tinjauan Terdahulu

Judul	Tahun	Metode	Hasil
Automating Potato Leaf Disease Detection with Lightning-Fast CNNs: Precision using PyTorch Lightning [14]	2024	Model klasifikasi gambar menggunakan arsitektur CNN untuk mendeteksi penyakit pada tanaman kentang. Pelatihan dilakukan secara efisien dengan PyTorch Lightning, serta optimisasi menggunakan Adam optimizer. Untuk meningkatkan generalisasi, diterapkan data augmentation seperti rotasi, flipping, dan kontras. Dataset dibagi menjadi 70% training, 15% validation, dan 15% testing.	Model CNN yang dikembangkan dengan PyTorch Lightning terbukti efektif dalam mendeteksi penyakit pada daun kentang, dengan capaian akurasi hingga 98,6%. Penggunaan PyTorch Lightning juga mempercepat dan menyederhanakan proses pelatihan model, membuatnya lebih terstruktur dan efisien.
DSCSkipNet: An Accuracy-Complexity Trade-off for Effective Potato Disease Identification in Uncontrolled Environments [15].	2024	DSCSkipNet terdiri dari Feature Extractor dan Classifier. Feature Extractor memanfaatkan enam blok konvolusi gabungan DSC dan NSC, serta Skip Connections untuk mempertahankan fitur penting. Pointwise Convolution, LeakyReLU, Batch Normalization, dan MaxPooling digunakan untuk efisiensi dan	Model mencapai akurasi tertinggi sebesar 83,23%, menunjukkan kestabilannya dalam kondisi dunia nyata. Kelas Healthy memiliki performa terendah karena kemiripan visual dengan kelas penyakit lain, sementara kelas Bacteria mencatat F1-score tertinggi sebesar 95%,

Judul	Tahun	Metode	Hasil
		stabilitas pelatihan. Classifier terdiri dari tiga Fully Connected Layers, dengan output 7 neuron dan Softmax Activation. Model dilatih menggunakan Adam optimizer, Cross-Entropy Loss, dan 4-Fold Cross-Validation untuk mencegah overfitting dan meningkatkan generalisasi.	menandakan kemampuan model yang sangat baik dalam mendeteksi infeksi bakteri. Penggunaan Skip Connections turut berperan dalam meningkatkan stabilitas model, khususnya saat mengekstraksi fitur dari gambar dengan kondisi kompleks.
Potato Leaf Disease Detection Based on a Lightweight Deep Learning Model [5].	2024	Metodologi dalam penelitian ini meliputi pemanfaatan arsitektur RegNetY-400MF yang dilengkapi modul squeeze-and-excitation (SE) untuk meningkatkan kemampuan ekstraksi fitur. Proses pelatihan dilakukan dengan transfer learning dan fine-tuning untuk menyesuaikan model terhadap dataset penyakit daun kentang. Kombinasi batch size 16 dan learning rate 0.001 dipilih sebagai parameter optimal. Selain itu, digunakan teknik augmentasi data seperti random cropping, brightness adjustment, flipping, dan rotasi untuk memperkaya variasi data dan meningkatkan akurasi model.	Model RegNetY_400MF menunjukkan performa paling unggul dibandingkan model lainnya dengan akurasi tertinggi sebesar 89,87% serta standar deviasi terendah 1,28%, menandakan hasil yang konsisten. Model ini mengungguli EfficientNet_B1 (89,66%) dan ShuffleNetV2_x2_0 (88,06%), sehingga dapat disimpulkan bahwa RegNetY_400MF lebih andal dalam mendeteksi penyakit daun kentang pada penelitian ini.
PotatoLeafGuard: A Novel Hybrid CNN Architecture [6].	2024	Penelitian ini mengembangkan PotatoLeafGuard, model Hybrid CNN yang menggabungkan Xception dan VGG16 untuk meningkatkan deteksi penyakit daun kentang. Menggunakan dataset 3.076 gambar dengan	PotatoLeafGuard mencapai akurasi terbaik sebesar 94,70%, mengungguli model CNN standar, Xception, dan VGG16. Kombinasi CNN dengan transfer learning (Xception & VGG16) terbukti

Judul	Tahun	Metode	Hasil
		tujuh kelas penyakit, model ini melalui preprocessing augmentasi data, normalisasi, dan pembagian data (70% pelatihan, 12% validasi, 18% pengujian). Proses pelatihan melibatkan dropout, batch normalization, dan pooling untuk efisiensi komputasi dan mengurangi overfitting.	efektif dalam mengatasi keterbatasan dataset. Namun, tantangan utama yang dihadapi termasuk variasi pencahayaan dan latar belakang, jumlah dataset yang terbatas untuk beberapa kategori penyakit, serta kemiripan gejala antar penyakit.
Enhancing crop productivity with fined-tuned deep convolution neural network for Potato leaf disease detection [16].	2025	Penelitian ini mengembangkan model deep learning berbasis transfer learning untuk deteksi enam penyakit daun kentang, fokus pada mitigasi class imbalance dan peningkatan akurasi melalui fine-tuning CNN. Menggunakan dataset 3.076 gambar, preprocessing mencakup data augmentation dan early stopping. Model dilatih dengan arsitektur CNN seperti DenseNet201, ResNet152V2, dan NasNetMobile, dioptimalkan dengan fine-tuning.	Model DenseNet201 dengan fine-tuning dan augmentasi menghasilkan akurasi terbaik sebesar 81,31%. Class imbalance dapat diatasi melalui augmentasi strategis dan L2 regularization. Tantangan utama yang dihadapi meliputi keterbatasan dataset, terutama untuk kategori nematoda, serta overfitting pada model ResNet152V2 dan NasNetMobile akibat variasi data. Selain itu, variasi pencahayaan dan kondisi daun juga mempengaruhi akurasi deteksi model.
Machine Learning Techniques for Potato (Solanum Tuberosum) Leaf Diseases Detection and	2024	Penelitian ini mengembangkan model machine learning untuk deteksi dan klasifikasi penyakit daun kentang, mengatasi variasi lingkungan dan kemiripan gejala penyakit. Dataset yang digunakan berasal dari PlantVillage, Kaggle, dan data	Model CNN dengan transfer learning mencapai akurasi tertinggi 97%, dengan ekstraksi fitur menggunakan GLCM dan CCM yang meningkatkan klasifikasi. Tantangan utama termasuk variasi warna dan tekstur daun

Judul	Tahun	Metode	Hasil
Classification [17].		lapangan, dengan gambar diambil menggunakan smartphone, kamera digital, dan drone. Preprocessing mencakup ekstraksi fitur tekstur dan warna menggunakan GLCM dan CCM, deteksi pola dengan SIFT, dan reduksi dimensi gambar dengan DWT. Model yang diuji meliputi Random Forest, CNN, SVM, Logistic Regression, KNN, dan ANN, kemudian dibandingkan untuk menentukan model terbaik.	yang mempengaruhi akurasi, kurangnya variasi dalam dataset, serta kesamaan gejala antar beberapa penyakit yang menyebabkan kesalahan klasifikasi. Sebagai rekomendasi, kombinasi CNN dengan Transformer atau Graph Neural Network (GCN) dapat digunakan untuk meningkatkan akurasi deteksi.
Efficient Data Augmentation Methods for Crop Disease Recognition in Sustainable Environmental Systems [7].	2025	Penelitian ini mengusulkan metode augmentasi data untuk meningkatkan kinerja model deep learning dalam mengenali penyakit tanaman, seperti VGGNet, ResNet, DenseNet, EfficientNet, ViT, and DeiT. Dataset baru dibangun dengan 24 kelas penyakit, termasuk sampel buah dan daun, untuk menangani berbagai pola penyakit dan meningkatkan generalisasi model. Augmentasi dilakukan dengan transformasi geometris dan peningkatan ruang warna untuk menguji efektivitas model konvolusional dan transformer.	Hasilnya menunjukkan bahwa teknik augmentasi ini meningkatkan akurasi klasifikasi, dengan skor F1 lebih dari 98%. Analisis peta fitur mengonfirmasi bahwa model efektif dalam menangkap karakteristik penyakit. Penelitian ini menekankan pentingnya augmentasi data untuk mengembangkan solusi deteksi penyakit yang efisien dan berkelanjutan.
Deep learning for mango leaf disease identification: A vision	2024	Penelitian ini menggunakan teknik augmentasi data untuk meningkatkan kinerja model deep learning dalam mendeteksi penyakit tanaman. Data yang	Hasil menunjukkan bahwa meskipun SqueezeNet memiliki ukuran dan jumlah parameter terkecil, akurasi lebih rendah. Model Vision

Judul	Tahun	Metode	Hasil
transformer perspective [10].		diolah melalui transformasi geometris dan perubahan warna. Model yang digunakan meliputi Convolutional Neural Network (CNN), yaitu SqueezeNet, ShuffleNet, dan EfficientNet, serta Vision Transformer (ViT) dengan DeiT dan SwinTransformer.	Transformer (ViT) seperti DeiT mencapai akurasi tertinggi 99.75%, dengan waktu pelatihan yang hampir setara dengan model CNN meskipun memiliki ukuran lebih besar. Model DeiT yang dioptimalkan menawarkan kombinasi kecepatan, akurasi, dan waktu inferensi yang singkat, menjadikannya pilihan terbaik untuk penerapan pada server.
An Efficient Approach for Potato Leaf Disease Classification Using Cascaded CNN-Transformers [18].	2024	Deteksi dini penyakit pada daun kentang dapat ditingkatkan dengan memperbaiki kualitas fitur dan akurasi klasifikasi, sehingga memungkinkan intervensi yang lebih cepat dan mendukung praktik pertanian yang berkelanjutan. Salah satu pendekatan yang dilakukan adalah dengan mengkombinasikan model Vision Transformer (ViT) dan ResNet50V2 yang telah dimodifikasi, yaitu dengan mengganti lapisan proyeksi linier menjadi lapisan konvolusional. Model yang digunakan untuk klasifikasi mencakup tiga kelas, yaitu Early Blight, Late Blight, dan daun yang sehat.	Model ini berhasil mencapai akurasi sebesar 97% dan terbukti sangat efisien serta cepat dalam prosesnya, dengan waktu pemrosesan yang kurang dari 1 detik. Keunggulan ini menjadikannya sangat potensial untuk diterapkan dalam deteksi penyakit tanaman secara cepat dan akurat, memungkinkan tindakan pencegahan atau perbaikan dapat dilakukan dengan lebih segera untuk mendukung pertanian yang lebih optimal.
Harnessing the Potato leaf disease detection	2025	Penelitian ini menganalisis dan membandingkan kinerja model deep learning CNN dan	Hasil analisis menunjukkan bahwa ResNet50 lebih unggul dibandingkan CNN dalam

Judul	Tahun	Metode	Hasil
process through proposed Conv2D and ResNet50 deep learning models [19].		ResNet50 dalam mendeteksi penyakit pada daun kentang, terutama pada early blight dan late blight, menggunakan teknik transfer learning. Model CNN dibangun dari awal dengan arsitektur Convolutional Neural Network standar, sementara ResNet50 diterapkan sebagai model transfer learning yang disesuaikan dengan dataset baru. Data dibagi dengan proporsi 90% untuk pelatihan dan 10% untuk validasi, dengan 3.251 gambar yang dikategorikan dalam tiga kelas, yaitu Early Blight, Late Blight, dan Healthy Leaf, yang bersumber dari Kaggle.	mendeteksi penyakit pada daun kentang, dengan akurasi validasi mencapai 97% berbanding 76% pada CNN. Performa CNN yang lebih rendah disebabkan oleh jumlah lapisan yang terbatas dan ekstraksi fitur yang kurang optimal. ResNet50 lebih efektif berkat penerapan transfer learning dari model yang sudah dilatih sebelumnya.

Penelitian-penelitian sebelumnya telah menunjukkan kemajuan signifikan dalam deteksi penyakit daun kentang menggunakan berbagai pendekatan deep learning, mulai dari CNN konvensional, arsitektur lightweight seperti RegNetY dan ShuffleNet, hingga model hybrid seperti gabungan Xception dan VGG16. Pendekatan berbasis transfer learning, augmentasi data, dan regularisasi juga telah diterapkan untuk meningkatkan akurasi dan mengurangi overfitting. Model CNN terbukti unggul dalam mendeteksi fitur spasial, sedangkan model hybrid dan transfer learning memberikan performa yang lebih baik dalam kondisi data terbatas atau kompleks. Namun, sebagian besar penelitian ini masih menghadapi tantangan dalam generalisasi, terutama pada dataset dunia nyata yang memiliki pencahayaan bervariasi, latar belakang kompleks, dan gejala antar penyakit yang mirip. Selain itu, model transformer seperti ViT dan DeiT mulai dieksplorasi, tetapi belum banyak diterapkan secara khusus untuk klasifikasi penyakit daun kentang dengan jumlah kelas yang lebih banyak dan kompleksitas visual yang tinggi.

Berdasarkan tinjauan tersebut, dapat disimpulkan bahwa masih terdapat gap dalam penerapan Vision Transformer yang efisien seperti DeiT untuk mendeteksi tujuh kelas penyakit daun kentang, khususnya dengan mempertimbangkan keterbatasan dataset, kebutuhan akan inferensi cepat, dan interpretabilitas model. Oleh karena itu, penelitian ini bertujuan untuk mengeksplorasi dan menguji performa model DeiT dalam klasifikasi multi-kelas penyakit daun kentang, serta membandingkannya dengan model baseline lainnya. Pendekatan ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem deteksi penyakit tanaman yang lebih akurat, efisien, dan mudah diimplementasikan dalam praktik pertanian berkelanjutan.

2.2 Tinjauan Teori

2.2.1 Deep Learning

Deep learning merupakan salah satu cabang dari kecerdasan buatan yang menggunakan algoritma berbasis jaringan saraf berlapis untuk mempelajari dan mengenali pola dari data yang sangat besar. Deep learning mengambil konsep dari manusia yang mempelajari berdasarkan pada jaringan syaraf otak, yang memproses informasi secara bertahap. Deep learning memungkinkan sistem untuk secara otomatis belajar dari data yang telah diberi label, sehingga dapat memahami informasi dengan cara yang lebih mendalam. Dengan begitu, komputer dapat mengenali pola-pola kompleks dan membuat keputusan dengan tingkat akurasi yang tinggi [20]. Beragam model deep learning telah dikembangkan untuk memperluas kemampuan teknologi ini melalui pendekatan yang beragam, diantaranya [21]:

1. Stacking Autoencoders

Model ini menggabungkan beberapa autoencoder, yaitu unsur deep learning untuk menyederhanakan data tanpa kehilangan informasi penting. Autoencoders bekerja dengan mengekstrak fitur dari data, yang berguna untuk mengurangi dimensi data atau meningkatkan pemahaman terhadap informasi yang ada.

2. Convolutional Neural Networks (CNNs)

Model ini dirancang untuk secara efisien menganalisis data dalam bentuk gambar dan video. Dengan memanfaatkan lapisan konvolusi, model ini mampu mendeteksi pola dalam data, seperti perbedaan warna, bentuk, atau struktur tertentu. Proses ini memungkinkan sistem untuk secara otomatis mengenali berbagai elemen dalam data, termasuk objek, teks, atau suara, dengan tingkat ketelitian yang sangat tinggi.

3. Deep Learning on Graphs

Pendekatan ini digunakan untuk menganalisis data yang memiliki hubungan atau struktur yang saling terhubung, seperti data tentang orang yang saling berhubungan atau rute-rute jalan. Dengan memanfaatkan Graph Neural Networks (GNNs), deep learning dapat memahami bagaimana elemen-elemen dalam data tersebut saling berhubungan dan membuat prediksi berdasarkan interaksi antar elemen tersebut.

4. Deep Probabilistic Neural Networks

Menggabungkan teknik deep learning dengan model probabilistik, pendekatan ini membantu model dalam mengukur ketidakpastian dan menghasilkan prediksi yang lebih kuat, terutama ketika data yang diproses memiliki elemen ketidakpastian atau gangguan, sehingga dapat menghasilkan prediksi yang lebih akurat dan handal.

5. Deep Fuzzy Neural Networks

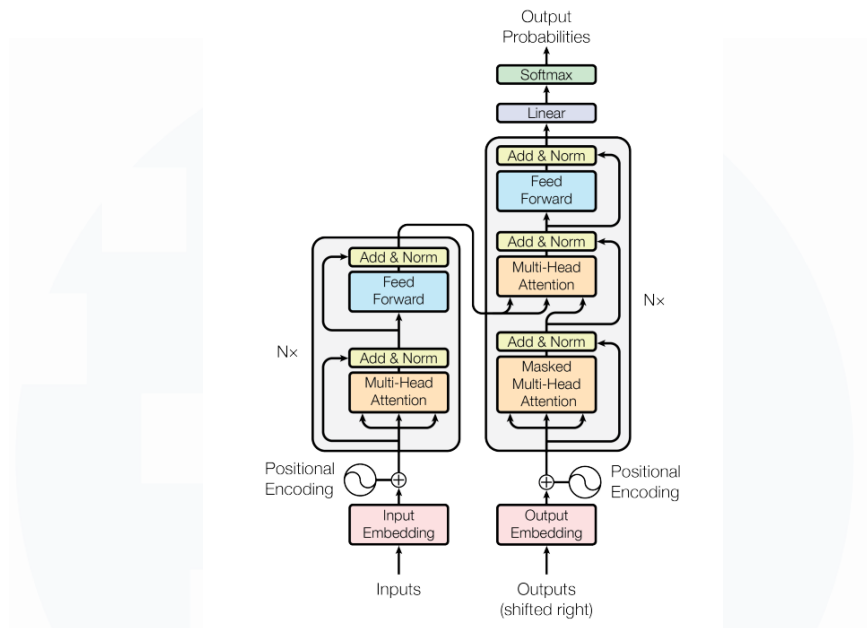
Dengan menggabungkan logika fuzzy ke dalam deep learning, model ini memungkinkan sistem untuk menangani data yang tidak jelas atau tidak pasti. Pendekatan ini sangat berguna untuk membuat prediksi ketika informasi yang tersedia tidak sepenuhnya pasti atau terdefinisi dengan jelas.

6. Generative Adversarial Networks (GANs)

Model ini menggunakan dua jaringan saraf yang bekerja saling berlawanan, yaitu generator dan discriminator, untuk membuat data baru yang sangat mirip dengan data asli. GANs banyak digunakan untuk menghasilkan gambar, suara, atau teks buatan yang terlihat atau terdengar alami.

2.2.2 Transformer

Transformer merupakan arsitektur model yang pertama kali diperkenalkan pada tahun 2017 oleh [22] dengan artikel yang berjudul “Attention is All You Need”. Model ini diformulasikan untuk mengolah data dalam bentuk rangkaian, seperti kalimat atau susunan kata, dengan cara yang lebih efisien dibandingkan pendekatan sebelumnya. Salah satu teknik yang membuat Transformer lebih unggul dalam kecepatan pelatihan yaitu memanfaatkan mekanisme attention sehingga dapat memproses keseluruhan rangkaian elemen secara paralel. Transformer dapat memahami hubungan antar elemen dalam suatu rangkaian terlebih pada jarak jauh, tidak seperti model yang memproses data satu per satu.



Gambar 2. 1 Arsitektur Transformer

Sumber 1 ([23])

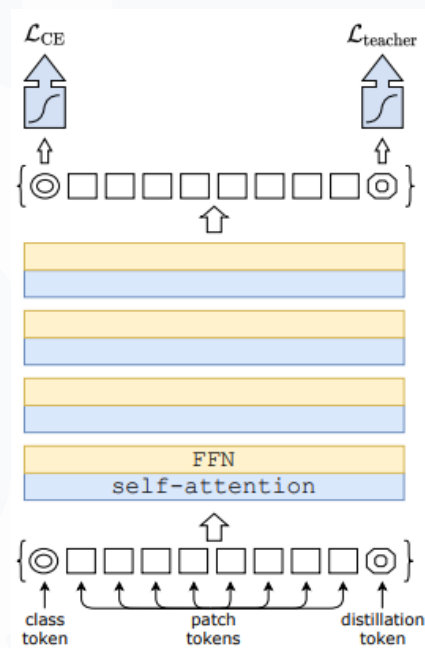
Arsitektur Transformer terdiri atas dua bagian utama, yaitu encoder dan decoder. Encoder berfungsi untuk memetakan input ke dalam representasi internal, sedangkan decoder menggunakan representasi tersebut untuk menghasilkan output. Setiap layer pada bagian encoder memiliki tiga komponen utama. Pertama adalah multi-head-self-attention yang berfungsi untuk memperhatikan berbagai bagian input secara bersamaan dan menghitung relevansi antar elemen dalam sebuah rangkaian. Hasilnya akan diproses oleh komponen kedua yaitu feed-forward neural network, yang merupakan jaringan saraf sederhana yang berfungsi untuk memperkuat representasi data. Dan terakhir terdapat residual connection dan layer normalization untuk menjaga stabilitas proses pelatihan serta mempercepat pencapaian hasil optimal.

Struktur layer pada decoder memiliki kemiripan, namun dilengkapi dengan satu komponen tambahan. Selain masked self-attention yang membatasi informasi agar prediksi hanya mempertimbangkan elemen sebelumnya, terdapat juga encoder-decoder attention yang memungkinkan decoder untuk memperhatikan hasil dari encoder. Komponen ini penting agar output dapat

disesuaikan dengan konteks input secara lebih akurat. Setelah itu, informasi tetap melewati feed-forward neural network serta ditambahkan residual connection dan normalisasi seperti pada encoder, sehingga keseluruhan sistem bekerja secara terstruktur dan efisien

2.2.3 Data-efficient Image Transformer

Data-efficient Image Transformer (DeiT) merupakan pengembangan dari model Vision Transformer (ViT) yang dirancang untuk meningkatkan efisiensi pelatihan dalam pengenalan citra, terutama ketika data yang tersedia terbatas. Sementara ViT memerlukan dataset besar untuk menghasilkan performa optimal, DeiT mengatasi keterbatasan ini dengan menggunakan teknik knowledge distillation. Dalam teknik ini, DeiT belajar dari model besar (teacher model) untuk meniru cara kerjanya, sehingga bisa memahami pola dari data meski jumlahnya tidak banyak.



Gambar 2. 2 Arsitektur Data-efficient Image Transformer (DeiT)

Sumber 2 ([24])

Keunggulan utama DeiT adalah kemampuannya untuk mencapai performa tinggi pada dataset terbatas tanpa perlu arsitektur konvolusional. Selain distilasi pengetahuan, DeiT juga memanfaatkan strategi pelatihan efisien seperti

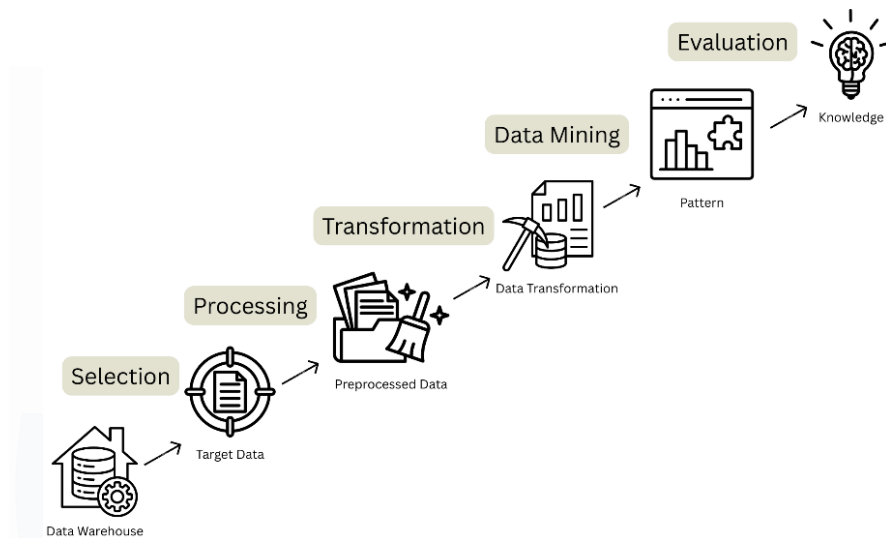
augmentasi data dan tuning hiperparameter. Pendekatan ini membuat DeiT kompetitif dengan model CNN dan ViT dalam tugas-tugas seperti klasifikasi gambar, sambil tetap hemat dalam sumber daya komputasi dan waktu pelatihan.

DeiT pertama kali diperkenalkan oleh Facebook AI Research (FAIR) pada akhir tahun 2020 sebagai jawaban atas keterbatasan ViT dalam skenario dengan data terbatas. Tujuan utamanya adalah membuat Transformer lebih ramah terhadap skala kecil seperti ImageNet tanpa perlu pretraining besar. Dengan memanfaatkan distilasi dan inovasi pelatihan lainnya, DeiT menjadi model yang efisien dan fleksibel untuk berbagai aplikasi visi komputer [24].

2.2.4 Knowledge Discovery in Databases (KDD)

Knowledge Discovery in Databases (KDD) merupakan suatu pendekatan sistematis untuk menemukan pola, informasi, dan pengetahuan tersembunyi dari data dalam jumlah besar dan kompleks. Konsep KDD pertama kali diperkenalkan oleh [25] pada tahun 1996 dan ditekankan kembali dalam penelitian oleh [26] bahwa KDD adalah proses non-trivial untuk mengidentifikasi pola yang valid, baru, berguna, dan dapat dipahami dari data. KDD hadir sebagai respons atas keterbatasan pendekatan statistik konvensional dalam menghadapi kompleksitas big data, khususnya pada data non-numerik seperti teks, gambar, suara, atau sensor. Dalam konteks ini, KDD menjadi penting untuk penelitian berbasis citra seperti deteksi penyakit tanaman kentang yang membutuhkan analisis mendalam terhadap pola visual.

[26] menekankan bahwa KDD bersifat dialektik, menggabungkan pendekatan deduktif dan induktif. Artinya, proses KDD tidak hanya mengekstraksi pola dari data secara otomatis, tetapi juga melibatkan pemahaman domain untuk memandu proses penemuan pengetahuan secara ilmiah dan kontekstual. KDD berbeda dengan eksplorasi data biasa karena memiliki tahapan terstruktur, mulai dari pemahaman domain, pengumpulan dan pembersihan data, transformasi data, pemilihan metode data mining, hingga evaluasi dan integrasi pengetahuan.



Gambar 2. 3 Tahapan Knowledge Discovery in Databases (KDD) [23]

Penjelasan dari tahapan KDD, berdasarkan gambar 2.3 sebagai berikut:

1. Selection

Tahap pertama adalah seleksi data (data selection), di mana data yang relevan dipilih dari berbagai sumber yang tersedia. Tidak semua data cocok untuk digunakan, sehingga pada tahap ini penting untuk mempertimbangkan kualitas, cakupan, dan keterkaitan data terhadap permasalahan yang diteliti.

2. Preprocessing

Tahap kedua adalah praproses data (data preprocessing). Data mentah sering kali mengandung noise, duplikasi, data hilang, atau inkonsistensi. Praproses bertujuan untuk membersihkan dan menyusun ulang data agar siap dianalisis lebih lanjut. Langkah ini mencakup pembersihan data, normalisasi, transformasi format, dan penanganan data yang hilang.

3. Transformation

Tahap keempat adalah reduksi dan transformasi data (data transformation/reduction). Pada tahap ini, data yang telah diproses diubah ke dalam bentuk yang lebih sesuai dan efisien untuk algoritma

penambangan data. Ini bisa berupa reduksi dimensi, ekstraksi fitur penting, atau pengelompokan variabel.

4. Data Mining

Tahap keempat adalah penambangan data (data mining), yaitu proses inti dari KDD di mana algoritma statistik atau machine learning digunakan untuk mengidentifikasi pola dan hubungan dalam data. Metode yang digunakan bisa berupa klasifikasi, klusterisasi, asosiasi, regresi, atau deteksi anomali tergantung pada tujuan analisis.

5. Evaluation

Tahap keenam adalah interpretasi dan evaluasi pola (pattern evaluation & interpretation). Setelah pola ditemukan, perlu dilakukan interpretasi untuk memastikan bahwa hasil tersebut memang valid, bermanfaat, dan dapat dipahami oleh pengguna akhir. Evaluasi ini dapat mencakup metrik akurasi, validasi silang, atau analisis makna dari pola yang ditemukan.