

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Sebelumnya

Penelitian dilakukan dengan beberapa tinjauan pustaka serta referensi dan teori yang menjadi dasar untuk melakukan penelitian. Penelitian yang berkaitan dengan Perbandingan Kinerja Pengklasifikasian Citra Bunga Kertas (*Bougainvillea*) Menggunakan *Neural Network* dan *Random Forest* yang dilakukan pada penelitian sebelumnya.

Penelitian sebelumnya yang berjudul “Klasifikasi Jenis Jamur Menggunakan Metode *Neural Network* Dengan *Fitur Inception-V3*” Penelitian ini menggunakan metode *Neural Network* dan *Deep Learning Inception V3* sebagai proses ekstraksi fitur pada citra untuk mengklasifikasikan citra jamur berdasarkan genus dengan aplikasi *Orange Data Mining*. Total dataset yang digunakan sebanyak 2.700, dengan 300 gambar untuk setiap genus. Pengujian menggunakan metode *cross-validation* yang diterapkan pada matriks konfusi untuk mendapatkan nilai presisi, *recall*, *F1-skor*, dan akurasi. Hasil klasifikasi akhir dengan akurasi sebesar 82,5% dan genus Jamur *Boletus* memperoleh hasil terbaik dengan akurasi sebesar 98,9% (Yulianto and Wibisono, 2023).

Penelitian lainnya yang berjudul “Klasifikasi Citra Ikan Menggunakan Algoritma *Convolutional Neural Network* dengan *Arsitektur VGG-16*”. Memanfaatkan arsitektur *VGG-16*, sebuah model CNN yang telah terbukti mampu mengambil fitur-fitur penting dari citra dengan kedalaman yang signifikan. Dataset terdiri dari 1088 citra ikan yang dibagi menjadi empat kelas: *Bangus*, *Glass Perchlet*, *Gold Fish*, dan *Gourami*.

Proses awal melibatkan ekstraksi fitur melalui *image embedding* menggunakan arsitektur *VGG-16*. Penggunaan *VGG-16* memungkinkan ekstraksi fitur yang kuat dan kompleks, dan hasil eksperimen menunjukkan bahwa pendekatan ini berhasil mencapai akurasi *training* sebesar 96,2% (R et al., 2023).

Penelitian yang berjudul “Klasifikasi Citra Daun Anggur Menggunakan *SVM Kernel Linear*”. Penelitian bertujuan meningkatkan perolehan angka akurasi dengan menggunakan bantuan proses ekstraksi fitur, serta membandingkan kinerja beberapa pengklasifikasi yaitu *k-Nearest Neighbor*, *Random Forest*, *Naïve Bayes*, *Neural Network* dan *Support Vector Machine*. Pembagian data latih dan uji dengan perbandingan 60:40. Hasil yang diperoleh mengungguli penelitian sebelumnya yaitu 98.1% pada pengklasifikasi *Support Vector Machine* menggunakan *kernel linear* (Sooai et al., 2023).

Judul penelitian “Klasifikasi Penyakit Kanker Paru Menggunakan Perbandingan Algoritma *Machine Learning*”. Peneliti ini membandingkan tuju algoritma yaitu algoritma *Random Forest*, *Tree*, *Neural Network*, dan *Logistic Regression* dengan nilai akurasinya 90% sementara algoritma *SVM*, *Naive Bayes*, dan *KNN* dengan nilai akurasi di bawah 90% (Idris et al., 2024).

Penelitian lainnya oleh Kristiawan Nugroho dan Sugeng Murdowo yang berjudul “Klasifikasi Jenis Hewan pada Kebun Binatang dengan Menggunakan Metode *Deep Neural Network*”. Dalam penelitian ini menggunakan *Neural Network* untuk memprediksi jenis hewan. Menggunakan *Neural Network* diperoleh hasil akurasi sebesar 98% dengan tingkat presisi sebesar 98.1% dan *Recall* sebesar 98 % (Nugroho and Murdowo-2024).

Tabel 2.1 Perbandingan peneliti terdahulu

No	Peneliti	Judul Penelitian	Metode	Hasil
1.	Okka Hermawan Yulianto dan Setyawan Wibisono(2023)	Klasifikasi Jenis Jamur Menggunakan Metode <i>Neural Network</i> Dengan <i>Fitur Inception-V3</i>	Metode <i>Neural Network</i> dengan <i>Fitur Inception-V3</i>	Hasil klasifikasi akhir dengan akurasi sebesar 82,5% dan genus Jamur <i>Boletus</i> memperoleh hasil terbaik dengan akurasi sebesar 98,9%.
2.	Imam Muslem R, TM Johan, dan Luthfi(2023)	Klasifikasi Citra Ikan Menggunakan Algoritma <i>Convolutional Neural Network</i> dengan <i>Arsitektur VGG-16</i> .	Algoritma <i>Convolutional Neural Network</i> dengan <i>Arsitektur VGG-16</i>	Hasil eksperimen menunjukkan bahwa pendekatan ini berhasil mencapai akurasi training sebesar 96,2%. Selanjutnya, saat proses klasifikasi menggunakan data testing, metode ini menghasilkan akurasi sebesar 99,5%.
3.	Adri Gabriel Sooi1, Paskalis Andrianus Nani, Natalia Magdalena Rafu Mamulak, Corazon Olivia Sianturi, Shine Crossifixio Sianturi, dan Alicia Herlin Mondolang (2023)	Klasifikasi Citra Daun Anggur Menggunakan SVM Kernel Linear	SVM Kernel Linear	Pelatihan data menggunakan ragam pengklasifikasi yang di validasi menggunakan <i>2-fold cross validation</i> . Data yang digunakan adalah sekunder daun anggur yang terdiri dari 7222 citra daun, terbagi dalam empat kelas yang telah tervalidasi dari peneliti terkait. Hasil yang diperoleh mengungguli penelitian sebelumnya yaitu 98.1%
4.	Jatnika Fahmi Idris, Rafid Ramadhani, dan Muhammad Malik Mutoffar(2024)	Klasifikasi Penyakit Kanker Paru Menggunakan Perbandingan Algoritma <i>Machine Learning</i>	Algoritma <i>Machine Learning</i> dengan metode algoritma <i>Random Forest, Tree, Neural Network, Logistic Regression, SVM, Naïve Bayes, dan Knn</i>	Penelitian ini membandingkan 7 algoritma dengan hasil nilai akurasi yang ada, untuk algoritma <i>Random Forest, Tree, Neural Network</i> , dan <i>Logistic Regression</i> memiliki nilai akurasi di atas 90% sementara untuk algoritma SVM, <i>Naïve Bayes</i> , dan kNN memiliki nilai

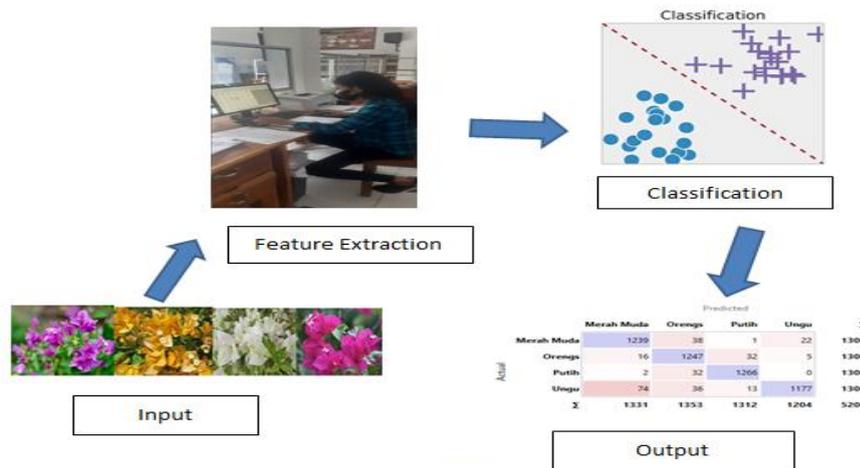
				akurasi di bawah 90%.
5.	Kristiawan Nugroho dan Sugeng Murdowo (2022)	Klasifikasi Jenis Hewan Pada Kebun Binatang dengan Menggunakan Metode <i>Deep Neural Network</i>	Metode <i>Deep Neural Network</i>	<i>Neural Network</i> diperoleh hasil akurasi sebesar 98% dengan tingkat presisi sebesar 98,1% dan Recall sebesar 98%.

Penelitian ini merujuk pada penelitian yang dilakukan oleh Okka Hermawan Yulianto dan Setyawan Wibisono yang berjudul “Klasifikasi Jenis Jamur Menggunakan Metode *Neural Network* Dengan *Fitur Inception-V3*”.dimana penelitiannya menggunakan metode Metode *Neural Network* Dengan *Fitur Inception-V3*. Peneliti mengembangkan penelitian tersebut dengan melakukan Perbandingan Kinerja Pengklasifikasian Citra Bunga Kertas (*Bougainvillea*) Menggunakan *Neural Network* dan *Random Forest* dan citra yang digunakan yaitu citra warna bunga kertas (*bougainvillea*).

2.2 Teori Pendukung

2.2.1 *Machine Learning*

Machine Learning (ML) merupakan cabang dari ilmu kecerdasan buatan dimana suatu pembelajaran bisa dilakukan pada sebuah metode sehingga bisa belajar mandiri dalam menyelesaikan suatu problem. ML melakukan proses pada komputer untuk melakukan tindakan mandiri untuk mencapai tingkat akurasi yang tinggi. Pada ML ini terdapat beberapa metode yang digunakan untuk kelancaran sistem. Proses pada *Machine Learning* dapat digambarkan seperti pada Gambar 2.1 Proses *Machine Learning* sebagai berikut:



Gambar 2.1 Proses *Machine Learning*

Pada gambar 2.1 diatas terlihat bahwa data yang diinput akan dilakukan proses ekstraksi fitur kemudian baru diproses dalam klasifikasi untuk menghasilkan output berupa suatu prediksi apakah data yang diinput sesuai dengan prediksi yang dilakukan. Pada tahapan ini komputer akan diberikan training menggunakan data yang cukup pada proses klasifikasi sehingga nantinya komputer akan bisa memprediksi input suatu objek (Nugroho and Murdowo-2022).

2.2.2 *Orange Data Mining*

Orange Data Mining adalah teknologi pembelajaran mesin sumber terbuka atau perangkat lunak penambangan data. Data mining digunakan untuk membantu dalam penelitian ini menggunakan *Orange*. *Orange* merupakan sebuah *tools open source* untuk pengolahan data mining. *Orange* memiliki *widget* yang berfungsi sebagai unit komputasi untuk membaca, memproses, melakukan visualisasi, melakukan analisis, mengeksplorasi data, dan lain-lain. *Widget* disusun sedemikian rupa sehingga membentuk *workflow* (alur kerja) dan berkomunikasi satu sama lain di lingkungan *Orange*.

Widget data memungkinkan *Orange* untuk memanipulasi data text atau gambar. Dalam menganalisis gambar menggunakan *Orange*, diperlukan sebuah add-ons tambahan yaitu image analytics. *Orange* mengubah data gambar menjadi data yang dapat dibaca oleh komputer sehingga dapat di proses (Idris et al,2024).

2.2.3 Citra

Citra adalah suatu konteks yang merujuk pada lingkungan hidup atau dalam bidang teknologi sendiri citra merujuk pada representasi digital dari objek atau data visual. Citra dibagi menjadi beberapa bagian yaitu citra digital dan citra analog. Citra analog adalah citra bersifat continue seperti gambar pada monitor TV dan foto sinar x, sedangkan pada citra digital adalah citra yang dapat diolah oleh komputer. Nilai suatu pixel memiliki nilai dalam rentang tertentu, dari nilai minimum sampai nilai maksimum (Worung et al, 2020).

2.2.4 Bunga Kertas (*Bougainvillea*)

Bougainvillea adalah tanaman hias yang digunakan untuk menghias taman. Komoditi ini dibudidayakan dalam kehidupan sehari-hari untuk dinikmati keindahannya. Bentuknya berupa pohon kecil yang sukar tumbuh tegak. Keindahannya berasal dari warna bunganya yang berwarna cerah dan menarik perhatian karena tumbuh dengan rimbunnya (Dayatmi et al., 2021).

2.2.5 Ekstraksi Fitur Warna

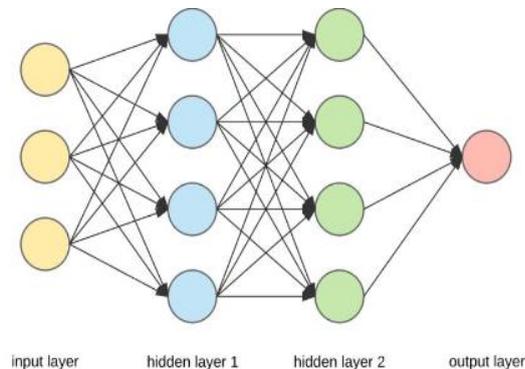
Peneliti menggunakan fitur *squeezeNet(local)* sebagai proses ekstraksi fitur. *SqueezeNet(local)* adalah model jaringan saraf konvolusi yang dirancang untuk memiliki jumlah parameter yang sangat rendah, sehingga

lebih ringan dan cocok untuk perangkat dengan sumber daya terbatas.

Meskipun memiliki jumlah parameter yang lebih sedikit, SqueezeNet masih dapat memberikan kinerja yang baik dalam tugas-tugas seperti klasifikasi gambar (Yulianto et al., 2023).

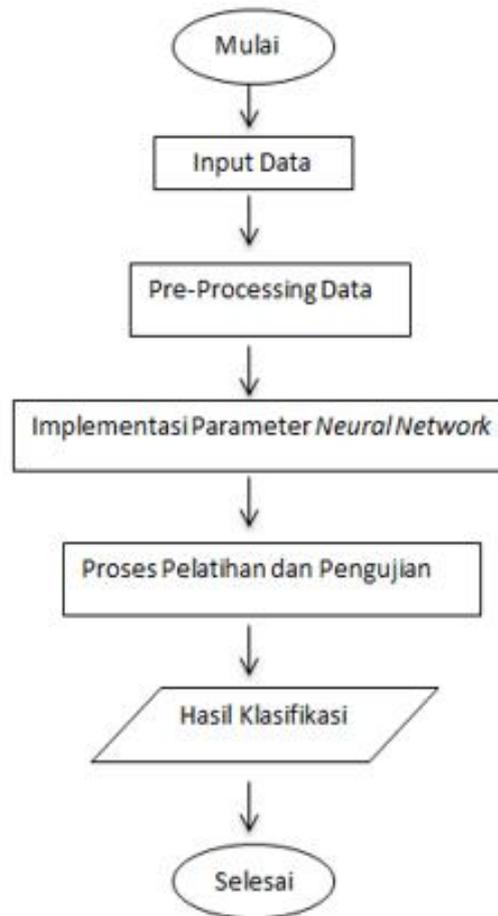
2.2.6 *Neural Network*

Neural Network atau disebut juga Jaringan Syaraf Tiruan (JST) merupakan jaringan pada permodelan komputer yang dibangun untuk meniru cara kerja otak manusia. *Neural Network* bekerja dengan sejumlah neuron dalam sejumlah hidden layer. Peneliti menggunakan propagasi maju (*forward propagation*) dari input layer ke hidden layer pertama, hidden layer kedua, dan output layer (Yulianto et al., 2023).



Gambar 2.2 Input, Hidden dan Output Layer

Pada Gambar 2.2 Input, Hidden dan Output Layer pada penelitian dengan input layer yang digunakan sebanyak 3 neuron, hidden layer pertama 4 neuron, hidden layer kedua 4 neuron dan output layer 1 neuron. Parameter yang mengontrol berapa besar langkah pembelajaran yang diambil algoritma pada setiap interaksi.



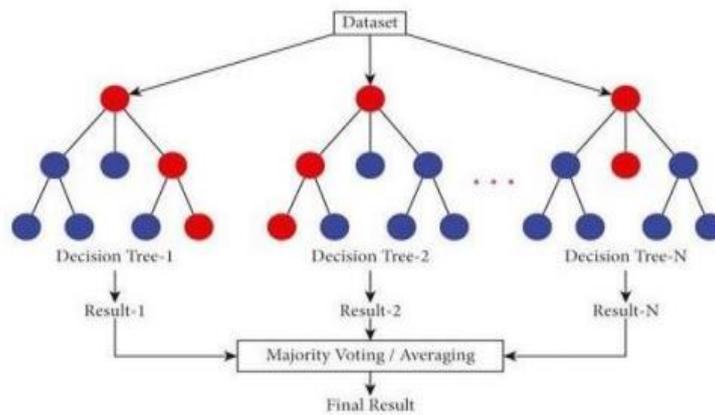
Gambar 2.3 *Flowchart Algoritma Neural Network*

Berikut merupakan tahapan algoritma *Neural Network* dalam pengklasifikasian citra bunga kertas (*bougainvillea*) untuk menentukan kinerja akurasi, presisi, recal, dan f1-skor dalam *orange data mining* (Dayatmi et al., 2021).

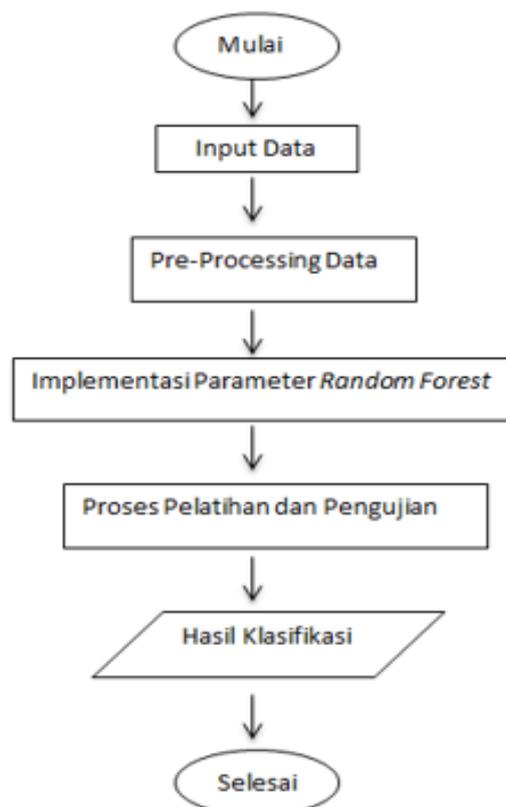
2.2.7 *Random Forest*

Pengklasifikasi *Random Forest* bekerja menggunakan teknik kombinasi atau dikenal berbasis ensemble. Prinsip utamanya adalah penggabungan berbagai pohon keputusan dalam operasinya. Dalam prosesnya, algoritma ini menghasilkan banyak pohon keputusan yang digabung untuk membuat prediksi.

Keuntungan dari penggabungan banyak pohon keputusan adalah akurasi yang tinggi (Sooai et al,2023). *Random Forest* digunakan untuk klasifikasi data dengan cara memilih *vote* terbanyak. Bentuk umum dari model klasifikasi *Random Forest* dapat di lihat pada Gambar 2.4 Metode *Random Forest*.



Gambar 2.4 Metode *Random Forest*

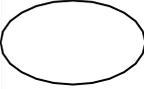
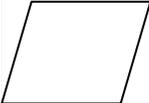


Gambar 2.5 *Flowchart Random Forest.*

2.2.8 Konsep Diagram Alir (*Flowchart*)

Flowchar adalah urutan proses kegiatan yang digambarkan dalam bentuk simbol-simbol atau aliran proses dengan menggunakan anotasi-anotasi. Berikut bentuk simbol-simbol tersebut.

Tabel 2.2 Bentuk-bentuk *symbol flowchart*

Simbol	Pengertian	Keterangan
	Proses	Merupakan kegiatan proses dari operasi program komputer
	Garis aliran (<i>flow line</i>)	Menunjukkan arus data antar simbol/proses
	Conector (<i>On page connector</i>)	Digunakan untuk penghubung dalam satu halaman
	Catatan	Digunakan untuk pencatatan data yang dirkam sebelumnya pada dokumen atau formulir data

2.2.9 *Confusion Matrix*

Confusion matrix digunakan untuk perhitungan akurasi. Hasil evaluasi model klasifikasi didasarkan pada pengujian untuk memperkirakan objek benar dan salah sebuah matrix prediksi yang akan dibandingkan dengan kelas yang asli dari inputan atau berisi informasi nilai actual dan prediksi pada klasifikasi. Tingkat akurasi adalah tingkat keakuratan jaringan yang telah dibuat dalam mengenali masukan citra yang diberikan sehingga menghasilkan keluaran yang benar. *confusion matrix* perhitungan dapat menghasilkan 4

keluaran yaitu *Precision*, *Recall*, *F1-Score* dan *Accuracy* .

Berikut rumus persamaan *Precision*, *Recall*, *F1-Score* dan *Accuracy*.

1. Rumus persamaan Precision

Precision mengukur sejauh mana prediksi positif yang dibuat oleh model adalah benar.

$$\text{Precision} = \frac{TP}{TP+FP} \dots\dots\dots(2.1)$$

2. Rumus persamaan Recall

Recall mengukur sejauh mana model dapat mendeteksi semua instance positif yang seharusnya ada.

$$\text{Recall} = \frac{TP}{TP+FN} \dots\dots\dots(2.2)$$

3. Rumus persamaan F1-Score

F1-Score rata-rata harmonik antara presisi dan recall dapat dihitung dengan rumus.

$$F1 - Score = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \dots\dots\dots(2.3)$$

4. Rumus persamaan Accuracy

Accuracy adalah rasio prediksi yang benar (positif dan negatif) terhadap total jumlah prediksi, dalam konteks klasifikasi mengukur sejauh mana model benar dalam mengklasifikasi data.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \dots\dots\dots(2.4)$$

Keterangan:

1. True Positive (TP)

Klasifikasi yang benar positif menunjukkan bahwa model

dengan benar mengklasifikasikan instance sebagai positif.

2. True Negative (TN)

Klasifikasi yang benar negatif menunjukkan bahwa model dengan benar mengklasifikasikan instance sebagai negatif.

3. False Positive (FP)

False Positive untuk menunjukkan bahwa model secara keliru mengklasifikasikan instance sebagai positif, padahal seharusnya negatif.

4. False Negative (FN)

False Negative untuk menunjukkan bahwa model secara keliru mengklasifikasikan instance sebagai negatif, padahal seharusnya positif .