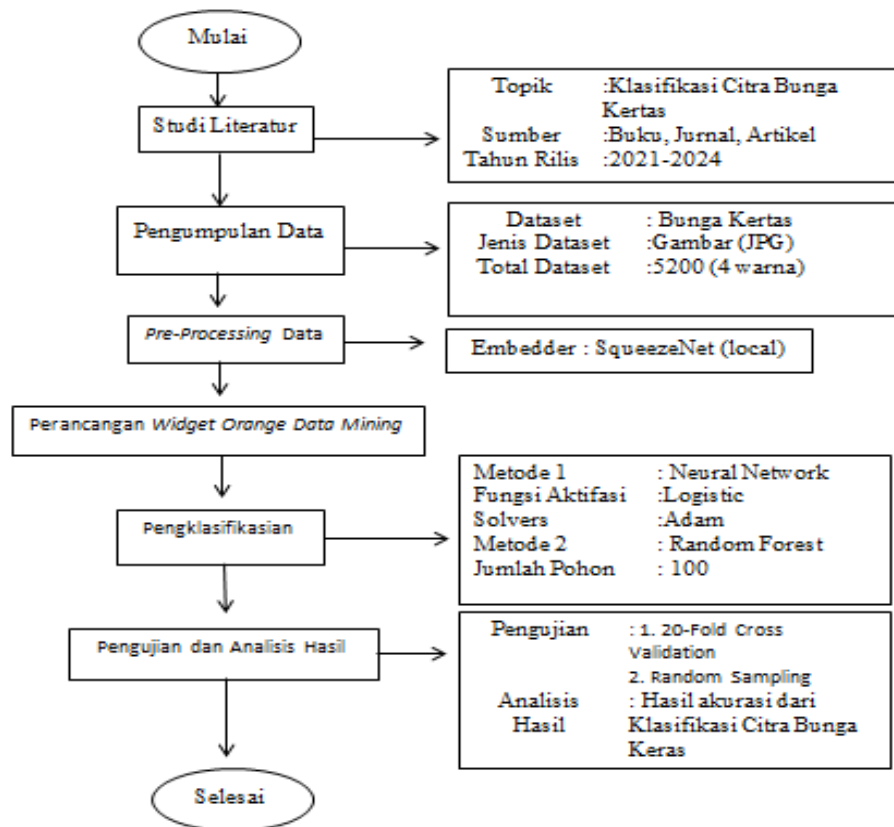


## BAB III

### METODE PENELITIAN

Berikut merupakan gambaran umum tahapan dalam penelitian ini yang terdiri dari studi literatur, pengumpulan data, *pre-processing* data, perancangan *widget orange data mining*, pengklasifikasian, hingga pengujian dan analisis hasil. Tahapan tersebut dapat dilihat pada gambar tahapan metode penelitian berikut:



Gambar 3.1 *Flowchart* Tahapan Penelitian

#### 3.1 Studi Literatur

Tahap pertama yang dilakukan pada penelitian ini adalah studi literatur. Pada tahap ini peneliti melakukan penggalian pengetahuan dan mencari referensi penelitian terdahulu dari berbagai sumber, seperti jurnal, artikel

maupun buku yang berhubungan terhadap penelitian yang dilakukan. Ada beberapa topik penelitian terdahulu yang dicari pada penelitian ini diantaranya pengklasifikasian jenis genus jamur, data mining, algoritma *Neural Network* dan *Random Forest* serta mempelajari bagaimana cara membaca *confusion matrix*.

## 3.2 Proses Pengumpulan Data

### 3.2.1 Pengambilan Data

Pengambilan data citra bunga kertas (*bougainvillea*) di lakukan dengan mencari tempat yang banyak tanaman *bougainvillea* di sekitaran area kota Kupang khususnya Jl Bandara Eltari Kupang. Data yang dikumpulkan berupa data primer. Berupa poin-poin penting pengambilan data:

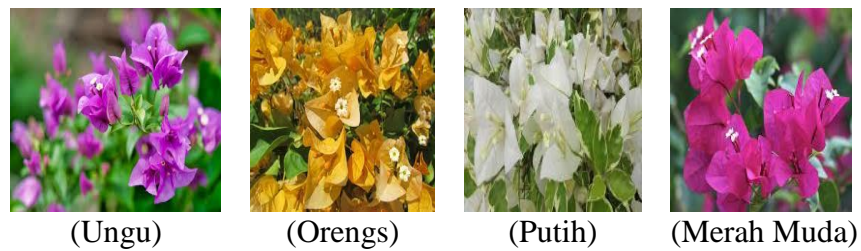
- a. Tanggal pencarian tempat dan mulai pengambilan data bunga kertas (*bougainvillea*) sebanyak 5200 citra (merah muda, *orengs*, putih, dan ungu) pada tanggal minggu, 27 agustus 2023 - Kamis, 31 agustus 2023 dari jam 02:30 PM - 05:00 PM.
- b. Setelah semua data citra didapatkan ,selanjutnya akan di lakukan pre-processing, dalam memudahkan pengolahan citra dengan memotong dan menyamakan semua ukuran citra menjadi 500 x 500 dengan menggunakan aplikasi *I love image*

### 3.2.2 Pengumpulan Data

#### a. Jenis Data

Data yang digunakan dalam penelitian ini merupakan gambar atau citra yang diambil dengan menggunakan *Handphone Samsung M20*. Sifat atau karakteristik dari citra yang digunakan yaitu beragam

warna bunga kertas (*bougainvillea*). Empat jenis warna bunga yang digunakan di antaranya ungu, orengs, putih dan merah muda.



Gambar 3.2 Citra Bunga Kertas (*Bougainvillea*)

Citra yang telah didapat selanjutnya akan melalui *preprocessing* dimana citra tersebut akan dipotong dan semua ukurannya akan disamakan menjadi citra dengan ukuran 500x500 *pixel* dengan bantuan aplikasi *I love Image*.

b. Jumlah Data

Keseluruhan sampel bunga kertas (*bougainvillea*) berjumlah 5200 citra sebanyak 4 warna, dengan masing-masing memiliki 1300 citra. Dari 4 data fitur warna yang masing-masing berjumlah 2 pohon bunga untuk satu warna.

c. Proses Pengambilan Data



Gambar 3.3 Proses Pengambilan Data

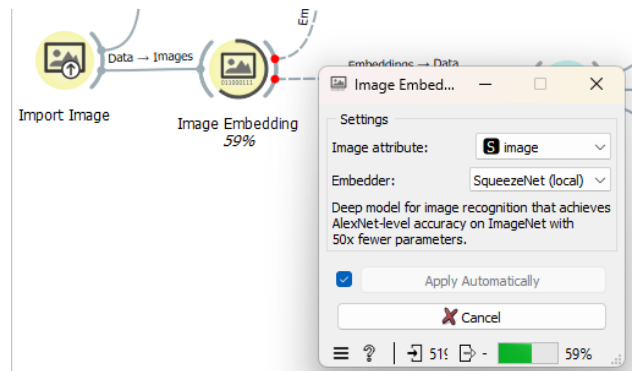
Data primer yang dikumpulkan untuk penelitian ini yaitu *file image* (citra) dengan format *Joint Photographic Expert Group* (JPG/JPEG) yang diperoleh dengan menggunakan kamera digital (*handphone Samsung Galaxy M20*). Objek yang diambil yaitu bunga kertas (*bougainvillea*) dengan jarak 30 cm, Selanjutnya data di bagi menjadi dua kelompok data, yang akan digunakan untuk data latih dan data uji kemudian di klasifikasian dan menghasilkan perbandingan akurasi.

### **3.3 Pre-processing Data**

Setelah data telah dikumpulkan selanjutnya dilakukan tahap pra- proses pada data citra, dimana akan dilakukan proses-proses dasar dari *image processing* seperti *cropping* dan *compressing size* citra. Pada *preprocessing* juga melakukan normalisasi dan ekstraksi fitur . Tahapan ini digunakan untuk mengekstrak bagian-bagian citra, sehingga akan memberikan klasifikasi tersendiri tersendiri terhadap citra dengan bantuan *I love image*. Dataset yang telah siap dimodelkan dibagi menjadi dua, yaitu data *training* dan data *testing* dengan perbandingan 90%:10%.

Pada tahap *pre-processing* data, peneliti menggunakan fitur *squeezeNet(local)* sebagai proses ekstraksi fitur citra bunga. *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. Proses ini diolah pada *widget image embedding Orange*. Hasil pemrosesan ini menghasilkan sebuah data-

data kategori, nama gambar, ukuran gambar, serta ukuran file masing-masing gambar. Berikut tampilan *pre-processing* data pada *Image embedding Orange Data Mining*.

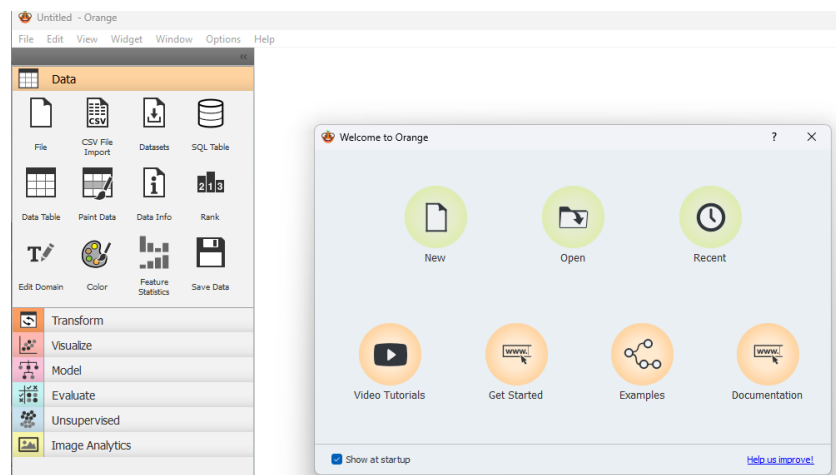


Gambar 3.4 *Widget Image Embedding Orange*

### 3.4 Perancangan *Widget Orange Data Mining*

Proses perancangan sistem klasifikasi citra bunga kertas (*bougainvillea*) dengan metode *Neural Network* dan *Random Forest* dibangun dengan *Tool Orange Data Mining*. Perancangan *widget* dalam *Orange Data Mining* adalah langkah penting dalam menyusun alur kerja (*workflow*) untuk analisis dan pengklasifikasian citra bunga kertas (*bougainvillea*). Langkah-langkah yang dapat dilakukan dalam perancangan *widget* menggunakan *Orange*:

#### 1. Tampilan awal

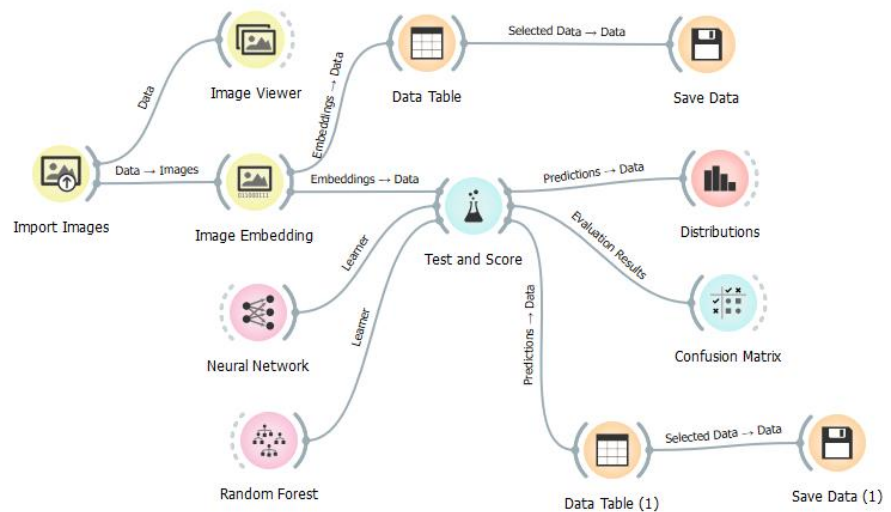


Gambar 3.5 Tampilan *Orange Data Mining*

Menu ini merupakan tampilan awal pada *software Orange Data Mining* yang merupakan tahap pertama tampilan sebelum melakukan klasifikasi citra bunga kertas (*bougainvillea*).

## 2. Tampilan Perancangan Klasifikasi

Dari hasil pembuatan *workflow* menggunakan aplikasi *Orange Data Mining*, menghasilkan sebuah model klasifikasi citra menggunakan metode *Neural Network* dan *Random Forest* dapat dilihat pada Gambar 3.6 *Workflow* Klasifikasi citra bunga kertas (*bougainvillea*).



Gambar 3.6 *Workflow* Klasifikasi citra bunga kertas (*bougainvillea*)

Pada tampilan ini ada beberapa *tools* yang di gunakan yaitu *import image*, *image viewer*, *image embedding*, *data table*, *Neural Network*, *Random Forest*, *test and score*, *distributions*, *confusion matrix* dan *save data* masing-masing mempunyai peran penting dalam pengklasifikasin citra bunga. Tahap yang dilakukan ini untuk kelancaran pemrosesan klasifikasi citra dengan metode *Neural Network* dan *Random Forest*.

### 3.5 Pengklasifikasian

#### 1. Model Klasifikasi *Neural Network*

Proses klasifikasi *Neural Network* Penelitian ini menggunakan propagasi maju (*forward propagation*) dari input layer (5200 neuron) ke hidden layer pertama (100 neuron), hidden layer kedua (1000 neuron), output layer (4 neuron), Fungsi aktifasi logistic, Fungsi aktifasi softmax dan larning rate ( $\alpha=0.0005$ ) dengan rumus:

a. Input layer ke hidden layer pertama

$$z^{(1)} = x * w^{(1)} + b^{(1)} \dots\dots\dots(3.1)$$

$$a^{(1)} = \text{logistic} (z^{(1)})\dots\dots\dots(3.2)$$

b. hidden layer pertama ke hidden layer kedua

$$z^{(2)} = x * w^{(2)} + b^{(2)} \dots\dots\dots(3.3)$$

$$a^{(2)} = \text{logistic} (z^{(2)})\dots\dots\dots(3.4)$$

c. hidden layer kedua ke output layer

$$z^{(3)} = a^{(2)} * w^{(3)} + b^{(3)} \dots\dots\dots(3.5)$$

$$y = \text{softmax} (z^{(3)})\dots\dots\dots(3.6)$$

Ket:

x : matriks input

a : larning rate

W : Bobot (weight) yang menghubungkan setiap neuron pada lapisan sebelumnya dengan neuron pada lapisan berikutnya

B : bias yang ditambahkan pada setiap neuron pada lapisan berikutnya

F : fungsi aktivasi yang digunakan pada setiap neuron (logistic dan softmax)

Y : output yang dihasilkan Model Klasifikasi *Random*

*Forest*

Algoritma ini bekerja dengan cara menggabungkan beberapa pohon keputusan (decision tree) yang dibuat secara acak untuk menghasilkan prediksi yang lebih akurat. secara efisien dapat mengatasi data training dengan jumlah data yang sangat besar. Metode Random Forest menghasilkan satu set pohon acak . Kelas yang dihasilkan berasal dari proses klasifikasi yang dipilih dari kelas yang paling banyak (modus) yang dihasilkan oleh pohon keputusan yang ada. Proses klasifikasi *Random Forest* Penelitian ini menggunakan *Number of trees* (jumlah pohon) 100, *Limit depth of individual trees* (batas kedalaman pohon individu) 4, *Do not split subsets smaller than* (jangan membagi subset yang lebih kecil dari) 1000. Ada beberapa langkah yang bisa digunakan:

- a) membagi data pelatihan menjadi subset dengan ukuran minimal 1000



- b) membuat pohon keputusan. Setiap pohon akan memiliki kedalaman maksimal 4
- c) Pada setiap langkah pembentukan pohon, menggunakan subset acak dari data pelatihan (yang memiliki minimal 1000 sampel) untuk mengurangi *overfitting*.
- d) Setelah pembentukan 100 pohon, peneliti melakukan prediksi dengan masing-masing pohon dan kemudian mengambil rata-rata prediksi (untuk kasus regresi) atau melakukan voting mayoritas (untuk kasus klasifikasi) untuk mendapatkan prediksi akhir

### **3.6 Pengujian dan Analisis Hasil**

Pengujian dan analisis hasil klasifikasi citra bunga kertas (*bougainvillea*) menggunakan metode *Neural Network* dan *Random Forest* untuk mengukur dan memahami seberapa baik model sistem bekerja. Penelitian ini menggunakan dua kali pengujian yaitu pengujian *20-fold cross validation* dan *random sampling*. Pengujian dan analisis hasil di peroleh dari *Confusion matrix*. *Confusion matrix* adalah tabel yang menunjukkan jumlah hasil prediksi yang benar dan salah pada tiap kelas, dapat di lihat pada 2.2.9 *Confusion Matrix* .

### **3.7 Analisis Kebutuhan**

Tahap ini merupakan proses dilaksanakannya analisis kebutuhan dalam perancangan sistem klasifikasi citra bunga kertas (*bougainvillea*) dengan menggunakan algoritma *Neural Network* dan *Random Forest*.

Hal-hal yang dianalisis meliputi kebutuhan perangkat keras (*hardware*), kebutuhan peangkat lunak (*software*) sehingga sistem dapat berjalan dengan baik.

Berikut merupakan beberapa perangkat keras dan perangkat lunak yang dibutuhkan dalam proses pembuatan sistem pengenalan citra.

## 1. Perangkat Keras

Perangkat keras yang digunakan dalam sistem ini meliputi:

### a) PC/Laptop hp

Processor	:11th Gen Intel(R) Core(TM) i3-1115G4 @ 3.00GHz 3.00GHz
Installed RAM	:4,00 GB (3,67 GB usable)
System type	:64-bit operating system, x64-based processor

Dalam perangkat ini berlangsung proses pembuatan sistem.

### b) *Handphone Samsung M20*

Tipe	: PLP IPS
Size	: 6.3 inches, 97.4 cm <sup>2</sup>
Resolusi	: 1080 x 2340 pixels, 19.5:9 rati

Perangkat ini digunakan dalam mengambil citra atau gambar dari objek yang diteliti bunga kertas (*bougainvillea*).

## 2. Perangkat Lunak

Perangkat lunak yang digunakan dalam sistem ini meliputi:

- a) Sistem Operasi *Windows 10*
- b) *Orange Data Mining*
- c) *Excel*
- d) *I love image*