

## Klasifikasi Bunga Menggunakan Naïve Bayes Berdasarkan Fitur Warna Dan Texture

Rafie

Teknik Informatika – STMIK Indonesia Banjarmasin  
Jl. Pangeran Hidayatullah, Sungai Jingah, Banjarmasin, Kalimantan Selatan  
Email : rafiekom@gmail.com

### **ABSTRAK**

Abstrak - Pemrosesan gambar sangat berbeperan pernting dalam mengekstraksi informasi yang berguna dari gambar. Klasifikasi bunga diperlukan untuk mengatasi masalah klasifikasi bunga secara manual serta mempersingkat waktu dalam identifikasi bunga, dalam kasus klasifikasi bunga, pemrosesan gambar adalah langkah penting untuk identifikasi spesies tanaman yang dibantu komputer. Klasifikasi citra bunga di dasarkan pada fitur tingkat rendah seperti warna dan tekstur untuk mendefinisikan dan menggambarkan konten gambar. Penelitian ini bertujuan untuk mengklasifikasi bunga berdasarkan jenis bunga menggunakan Teknik pemrosesan citra. Ektraksi fitur yang digunakan awalnya adalah Hue, Saturation, Value untuk mendapatkan citra warna. Sedangkan untuk mendapatkan citra texture adalah GLCM (Gray Level Co-occurrence Matrik) yaiut Contrast, Correlation, Energy dan Homogenity, kemudian setelah percobaan fitur warna di tambah 3 lagi yaitu Red, Green Blue. Data latih bunga yang digunakan berjumlah 200 gambar yang terdiri dari 4 class bunga (kansas, marguerite, roses dan tulips), masing-masing bunga mewakili dari 50 gambar. Hasil percobaan pertama menunjukkan akurasi menggunakan Naïve Bayes distribusi normal sebesar 66 %, setelah beberapa kali percobaan hingga mendapatkan hasil akurasi tertinggi sebesar 77%, hasil ini diperoleh dengan menerapkan 6 fitur warna dan 4 fitur texture dan menggunakan Naïve Bayes distribusi kernel.

**Kata kunci:** Klasifikasi Bunga, Pengolahan Citra, Naïve Bayes, Ektraksi Fitur.

### **ABSTRACT**

Abstract – The Image Processing is very important in extracting useful information from images. classification of flower is necessary to overcome the problem of classification of flower manually and shorten the time in the identification of flower, in the case of flower classification, image processing is an important step for the identification of computer assisted plant species. The Flower image classification is based on low-level features such as colors and textures to define and describe image content. This study aims to classify flowers by type of flowers using image processing techniques. At first, The feature extraction used Hue, Saturation, Value to get the color image. whereas to get the image of texture is GLCM (Gray Level Co-occurrence Matri) is Contrast, Correlation, Energy and Homogenity, then after the experiment color feature added 3 more colours that is Red, Green and Blue. The data exercise used 200 drawings consisting of 4 classes of flowers (kansas, marguerite, roses and tulips), each representing 50 flowers. The first experimental results show the accuracy of using Naïve Bayes normal distribution of 66%, after several attempts to get the highest accuracy result of 77%, this result is obtained by applying 6 color features and 4 texture features and using Naïve Bayes kernel distribution.

**Keywords:** Classification Flower, Image Processing, Naïve Bayes, Feature Extraction

### **I. Pendahuluan**

Pemrosesan gambar berperan penting dalam mengekstraksi informasi yang berguna dari gambar. Klasifikasi bunga diperlukan untuk mengatasi masalah klasifikasi bunga secara manual serta mempersingkat waktu dalam identifikasi bunga, dalam kasus klasifikasi bunga, pemrosesan gambar adalah langkah penting untuk identifikasi spesies tanaman yang dibantu komputer. Namun, pemrosesan gambar dan proses menerjemahkan gambar ke distribusi statistik fitur low

level bukanlah tugas yang mudah. Tugas-tugas ini rumit karena data gambar yang diperoleh sering bintik-bintik, noise, dan objek target dipengaruhi oleh pencahayaan, intensitas atau iluminasi. Memproses gambar juga tergantung pada jenis peralatan yang menghasilkan gambar dan karakteristiknya yang diekstrak (Hong, 2003:589-592)

Warna adalah bagian penting dalam klasifikasi bunga karena berguna untuk manusia memberikan informasi tambahan untuk segmentasi dan pengenalan kepada

mesin. Di sisi lain, tekstur memberikan informasi tentang distribusi *gray-level* dari satu set piksel yang terhubung, yang terjadi berulang kali di wilayah gambar. Tekstur biasanya dikodekan oleh sejumlah deskripsi, yang diwakili oleh serangkaian ukuran statistik seperti *grey-level co-occurrence matrix (GLCM)*.

*Naïve Bayes* pada penelitian ini digunakan untuk klasifikasi bunga, karena *Naïve Bayes* dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. *Naïve Bayes* didasarkan pada teorema Bayes yang memiliki kemampuan klasifikasi serupa dengan *Decision tree* dan *Neural Network*. *Naïve Bayes* terbukti memiliki akurasi dan kecepatan yang tinggi saat di aplikasi ke dalam database dengan data yang besar (Kusrini, 2009:185).

## 2. Kajian Literatur

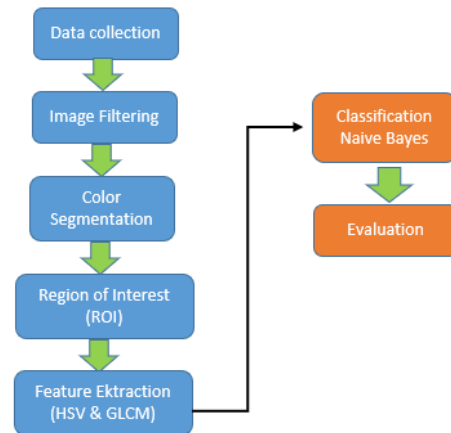
Beberapa penelitian tentang Teknik pemrosesan citra telah banyak di terapkan untuk melakukan klasifikasi terhadap bunga. Salah satu penelitian tentang klasifikasi bunga menggunakan *Neural Network (NN)* berdasarkan fitur warna dan texture, semua gambar dikategorikan ke dalam 7 fitur yaitu *HSV, Contrast, Correlation, Energy* dan *Homogeneity*. Hasil percobaan awal yang terdiri dari 19 kelas bunga dan masing-masing terdiri dari 10 gambar hasil akurasi antara 20 hingga 30% saja. Sedangkan akurasi tertinggi dengan menggabungkan bunga dari warna hanya sebesar 60% (Fadzillah, 2014:81-86). Sedangkan penelitian lain dengan menggunakan *Neural Network (NN)* dan klasifikasi bunga juga. Tetapi jumlah neuron di lapisan tersembunyi adalah 30 dengan ekstraksi *texture* dan warna untuk klasifikasi bunga juga, menghasilkan akurasi tertinggi 95% dengan 5 kelas bunga dan masing-masing terdiri dari 40 gambar sehingga total 200 gambar. (Riddhi, 2017:113-118).

Penelitian Yuita (2014:357-368) menggunakan klasifikasi *k-Nearest Neighbor (kNN)* dengan jarak kosinus, menghasilkan akurasi terbaik sebesar 73,63 %. Dengan Fitur warna diekstraksi dengan menghapus saluran L dalam ruang warna  $L * a * b$ , dan hanya mengambil saluran  $a$  dan  $b$ , karena mengabaikan kondisi pencahayaan yang berbeda dalam gambar bunga. Fitur tekstur diekstrak oleh *Segmentation-based Fractal Texture Analysis*. Sedangkan pada penelitian Fuzy (2017:99-108) dengan kasus klasifikasi berbeda yaitu untuk klasifikasi belimbing dengan menggunakan *Naïve* berdasarkan fitur warna RGB, menghasilkan akurasi sebesar 80%. Dalam penelitian ini menggunakan citra belimbing berjumlah 120 yang terdiri dari data latih berjumlah 90 dan data uji berjumlah 30.

## 3. Metode

Metode penelitian yang di lakukan meliputi pengumpulan data citra bunga, penyaringan gambar,

segmentasi gambar, ekstraksi fitur warna RGB dan ekstraksi fitur texture GLCM, model klasifikasi dengan *Naïve Bayes*, dan evaluasi model dengan melihat hasil klasifikasi menggunakan *confusion matrix*. Tahap penelitian dapat dilihat pada Gambar 1



Gambar 1. Metodologi

### a. Pengumpulan Data

Data yang digunakan adalah citra bunga berjumlah 200 gambar, yang terdiri dari 4 class (Kansas, Marguerity, Roses dan Tulips), masing – masing class terdiri dari 50 gambar yang berbeda-beda. Data gambar di peroleh dari *dataset flower photo* yang kemudian dipilih beberapa gambar sebagai data *training* dalam penelitian ini.

### b. Penyaringan Gambar

Penyaringan gambar pada dasarnya adalah sebuah metode untuk meredakan atau menghilangkan noise pada citra digital atau image. Jenis filter bermacam-macam dan fungsi serta efeknya juga berbeda-beda pula. Filter citra dibagi menjadi dua, yaitu filter linear dan filter non-linear. Filter spasial *linear* adalah filter yang bekerja dengan cara korelasi atau konvolusi. Filter spasial *non-linear* atau biasanya disebut juga dengan filter statistik berdasar urutan adalah filter yang respon nya didasarkan pada urutan atau rangking piksel yang ada dalam citra yang dicakup oleh area filter dengan menggantikan nilai dari piksel yang berada di tengah digantikan dengan nilai hasil pengurutan atau perangkingan tersebut. Filter non-linear memiliki lebih banyak keunggulan dibandingkan dengan filter linear pada ukuran jendela filter yang sama.

Penyaringan gambar yang digunakan dalam penelitian ini adalah filter *median*, karena kemampuannya dalam mengurangi derau yang diakibatkan oleh derau acak misalnya jenis *salt and pepper noise* atau bisa disebut sebagai derau *impulse* (Isma, 2014:123-132). Dibandingkan dengan jenis filter spasial (ruang) non-linear lainnya, filter median merupakan filter yang paling cocok untuk kasus tersebut. Sehingga filter ini dinobatkan menjadi filter yang paling ampuh dalam mengolah citra berderau sejenis.

c. Segmentasi Warna

Segmentasi citra mengambil keuntungan dari perbedaan warna antar daerah dan latar belakang gambar dihapus oleh segmentasi warna. Daerah bunga tersegmentasi dari latar belakang untuk memperoleh objek atau *region of interest (ROI)* sebelum tugas pemrosesan berikutnya dapat dilanjutkan (Jia, 2000:147-156).

d. *Region of Interest (ROI)*

Menghilangkan derau dilakukan dengan menghapus setiap lubang kosong atau wilayah apa pun yang mengelilingi objek yang diinginkan. Tugas ini dapat dilakukan dengan menggunakan fungsi morfologi Matlab seperti imopen untuk menghilangkan *noise*, imclose untuk menghaluskan batas objek dan imclean untuk membersihkan objek untuk hasil segmentasi yang lebih baik (Jia, 2000:147-156).

e. Ekstraksi Fitur

Ekstraksi fitur menangkap karakteristik penting dari pola. Untuk penelitian ini, dua fitur ekstraksi dipertimbangkan, yaitu warna dan tekstur. Semua gambar bunga ditangkap di RGB (*Red, Green, Blue*). Karena ruang warna RGB mudah dipengaruhi oleh intensitas dan pencahayaan dari matahari atau lampu kamera, hal ini mengarah pada ketidakseimbangan persepsi ketidaksamaan warna. Salah satu solusi untuk masalah ini adalah mengonversi gambar ke dalam pemandangan format *Hue, Saturation, Value (HSV)* (Maria, 2006:1447-1454). Format HSV mengabaikan intensitas atau pencahayaan yang disebabkan oleh matahari atau pencahayaan. Mengadopsi dari Fadzillah (2010:33-38), rumus konversi RGB ke HSV digambarkan dalam Gambar 2.

Name	Mathematical Formula
Hue	$H = \cos^{-1} \left\{ \frac{\frac{1}{2}[(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right\}$
Saturation	$S = 1 - \frac{3}{R+G+B} [\min(R, G, B)]$
Value	$V = \frac{1}{3} (R + G + B)$

Gambar 2. Formula Konversi RGB ke HSV

Tekstur gambar dihitung berdasarkan GLCM untuk mendapatkan perkiraan properti gambar di permukaan objek dengan mengukur intensitas piksel di wilayah permukaan yang dipilih. Misalnya, wilayah kelopak dan

simpang bunga bunga dianggap sebagai daerah yang menarik (Fuzy, 2017:99-108).

Biasanya GLCM dihitung pada empat sudut yang berbeda, 0, 45, 90 dan 135 derajat. Empat belas (14) fitur dapat dihitung dari masing-masing GLCM tetapi untuk studi ini, hanya empat fitur yang diekstraksi dari gambar. Fitur seperti Contrast, Correlation, Energy dan Homogeneity, gambar digunakan untuk perhitungan tekstur dan formula konversi pada Gambar 3

Feature	Formula
Contrast	$\sum_{i,j}  i-j ^2 p(i,j)$
Correlation	$\sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)P(i,j)}{\sigma_i \sigma_j}$
Energy	$\sum_{i,j} p(i,j)^2$
Homogeneity	$\sum_{i,j} \frac{p(i,j)}{1+ i-j }$

Gambar 3. Formula Fitur GLCM

f. Klasifikasi *Naïve Bayes*

*Naïve Bayes* merupakan salah satu metode *machine learning* yang menggunakan perhitungan probabilitas. Algoritma ini memanfaatkan metode probabilitas dan statistik sederhana dengan asumsi bahwa antar satu kelas dengan kelas yang lain tidak saling tergantung (*independen*) Dasar dari *Naïve Bayes* yang dipakai dalam pemrograman adalah persamaan (1) Bayes (Fuzy, 2017) :

$$P(Y|X) = \frac{P(Y) \prod_{i=1}^q P(X_i|Y)}{P(X)}$$

.....(1)

P(Y|X) : Probabilitas data dengan vektor X pada kelas Y

P(Y) : Probabilitas awal kelas Y (prior probability)

$\prod_{i=1}^q P(X_i|Y)$ : Probabilitas independen kelas Y dari semua fitur dalam vektor X

P(X) : Probabilitas dari X

g. Evaluasi

Tahapan ini merupakan tahapan untuk menganalisis dan mengevaluasi model yang diperoleh dari masing model yang digunakan. Proses perhitungan akurasi hasil klasifikasi menggunakan rumus *Confusion matrix*. *Confusion matrix* merupakan sebuah tabel yang terdiri atas banyaknya baris data uji yang diprediksi benar dan tidak benar oleh model klasifikasi.

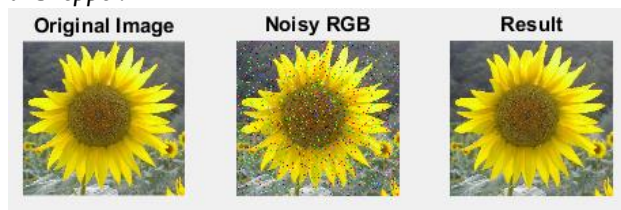
4. Percobaan dan Hasil

Percobaan dalam penelitian ini dilakukan dalam beberapa tahapan mulai dari penyaringan gambar kemudian dilanjutkan ke segmentasi warna dan deteksi wilayah kemudian di lanjutkan ke ekstraksi fitur untuk

mendapatkan ciri. Setelah mendapatkan ciri maka akan dibuat model klasifikasi dengan *Naïve bayes*.

4.1. Hasil Percobaan Penyaringan Gambar

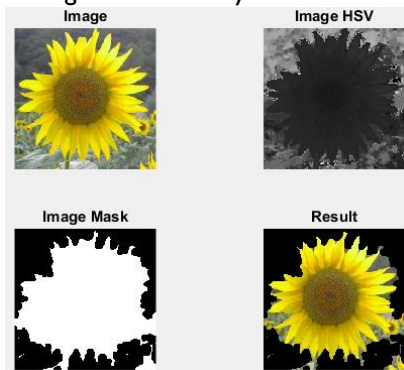
Percobaan tahap pertama ini adalah melakukan penyaringan gambar untuk menghilangkan *Noise*, dalam percobaan ini menggunakan *salt and repper*. Gambar 4 menunjukkan hasil percobaan dengan filter *median salt and repper*.



Gambar 4 Hasil Penyaringan gambar *salt and repper*

4.2. Hasil Percobaan Segmentasi Gambar

Percobaan dalam tahap kedua ini adalah melanjutkan tahap pertama, setelah di dapatkan hasil penyaringan gambar akan dilanjutkan ke tahap kedua yaitu proses segmentasi warna gambar. Segmentasi yang digunakan adalah *threshold* dan kemudian dilanjutkan untuk mendeteksi wilayah. Pada percobaan ini adalah mendeteksi wilayah gambar dari latar belakang akan di hapus. Gambar 5 menunjukkan hasil percobaan segmentasi dengan deteksi wilayah.



Gambar 5. Hasil Segmentasi Warna dan ROI

4.3. Hasil Percobaan Ekstraksi Fitur

Percobaan dalam tahap ini adalah melanjutkan proses dari tahap sebelumnya setelah di dapatkan hasil segmentasi maka gambar akan dilakukan ekstraksi fitur warna dan texture. Hasil percobaan dalam ekstraksi fitur ini adalah dengan mengkonversi warna *Red, Green, Blue* ke *Hue, Saturation, Value*. Kemudian dinormalisasi untuk mendapatkan representasi nilai yang tepat untuk gambar. Tabel. 1 menunjukkan contoh hasil percobaan konversi dengan fitur warna ternormalisasi.

Tabel 1. Contoh Gambar Normalisasi dengan Fitur Warna

Image Color	Numerical Value
-------------	-----------------

H	S	V	H	S	V
			23.06	125	75.49
			37.92	17.23	44.04
			146.34	203.17	32.53

Sedangkan hasil percobaan dalam ekstraksi texture dengan GLCM dalam penelitian ini ada empat fitur yang diekstraksi dari gambar. Fitur seperti Kontras, Korelasi, Energi dan Homogenitas. Hasil percobaan ekstraksi texture dengan GLCM, ada pada tabel 2.

Tabel 2. Contoh Gambar Normalisasi dengan Fitur Warna HSV dan Texture

Numerical Value Color			Numerical Value Texture			
H	S	V	Contrast	Correlation	Energy	Homogeneity
23.66	125.07	75.49	0.97	0.91	0.17	0.18
37.92	17.23	44.04	1.15	0.89	0.53	0.93
146.34	203.17	32.53	0.17	0.78	0.34	0.92

4.4. Hasil Percobaan Klasifikasi *Naïve Bayes*

Proses *training* bertujuan untuk membangun model klasifikasi. Pada percobaan ini menggunakan dataset hasil ekstraksi total data 200 dengan masing-masing class sebanyak 50 dengan probabilitas dari X (*H,S,V,Contrast, Correlation, Energy, Homogeneity*) dan Class (Kansas, Marguerite, Roses, Tulips).

Percobaan pertama dari dataset yaitu pembuatan model klasifikasi dengan menerapkan distribusi semuanya Normal. Maka hasil akurasi yang di dapatkan dalam percobaan pertama adalah sebesar 66 %, dengan *Confusion Matrix* yang dapat dilihat pada tabel 3.

Tabel 3. *Confusion Matrix* Klasifikasi Bunga Distribusi Normal

Kelas Sebenarnya	Kelas Prediksi				Total
	Kansas	Marguerite	Roses	Tulips	
<b>Kansas</b>	44	5	1	0	50
<b>Marguerite</b>	8	35	3	4	50
<b>Roses</b>	8	5	28	9	50
<b>Tulips</b>	3	13	9	25	50

Pada percobaan kedua proses *training Naïve Bayes* ini bertujuan untuk mendapatkan hasil akurasi yang lebih tinggi dari percobaan pertama, percobaan ini menerapkan *Naivebayes kernel*. Maka hasil akurasi yang di dapatkan

pada percobaan kedua ini lebih tinggi dari pada percobaan pertama yaitu sebesar 73 %, dengan *Confusion Matrix*. pada tabel 4.

Tabel 4. *Confusion Matrix* Klasifikasi Bunga Kernel

Kelas Sebenarnya	Kelas Prediksi				Total
	Kansas	Marguerite	Roses	Tulips	
<b>Kansas</b>	47	1	2	0	50
<b>Marguerite</b>	6	36	3	5	50
<b>Roses</b>	8	3	32	7	50
<b>Tulips</b>	4	4	11	31	50

Pada percobaan ketiga ini bertujuan untuk mendapatkan hasil akurasi yang lebih tinggi dari percobaan kedua, percobaan ini menambahkan 3 fitur warna (RGB) untuk ekstraksi fiturnya jadi total 10 fitur yang di ekstraksi yaitu *Red, Gree, Blue, Hue, Saturation, Value, Contrast, Correlation, Energy, Homogeneity*. Maka hasil akurasi yang di dapatkan pada percobaan ketiga ini lebih tinggi dari pada percobaan kedua yaitu sebesar 77 %, dengan *Confusion Matrix* yang dapat dilihat pada tabel 5.

Tabel 5. *Confusion Matrix* Klasifikasi Bunga 10 Fitur dan Distribusi Kernel

Kelas Sebenarnya	Kelas Prediksi				Total
	Kansas	Marguerite	Roses	Tulips	
<b>Kansas</b>	47	0	2	1	50
<b>Marguerite</b>	4	41	2	3	50
<b>Roses</b>	4	5	35	6	50
<b>Tulips</b>	4	5	10	31	50

Berdasarkan tabel 5 diketahui metode *Naive Bayes* untuk klasifikasi bunga masih terjadi kesalahan klasifikasi. Kesalahan ini terjadi paling banyak antara bunga roses dan tulips, pada jenis bunga ini terjadi kesalahan karena disebabkan pada dataset jenis bunga ini memiliki fitur warna yang hampir sama, sehingga dengan hanya 10 fitur ekstraksi masih belum maksimal perlu adanya ekstraksi fitur bentuk bunga.

## 5. Kesimpulan

Ekstraksi fitur warna dan fitur texture GLCM dapat digunakan untuk ekstraksi ciri pada citra. Hasil ekstraksi ciri

digunakan untuk dataset *Naive Bayes* untuk mengenal pola citra dan mengklasifikasi jenis bunga. Hasil percobaan pertama klasifikasi *Naive Bayes* 7 fitur tingkat akurasinya sebesar 66 %, percobaan kedua dengan distribusi kernel akurasinya 73%. Sedangkan dengan menerapkan 10 fitur dengan distribusi kernel menunjukkan tingkat akurasi lebih besar yaitu 77 %.

Pada penelitian selanjutnya juga bisa menggunakan metode *decision tree, support vector machine* atau metode klasifikasi lainnya, sehingga dapat dicari algoritma yang paling tinggi akurasinya.

## Daftar Pustaka

- Fadzillah S, Hawa M.E dan Abdul N.Z, 2014, “*Flower Image Classification Modeling Using Neural Network*”, International Conference on Computer, Control, Informatics and Its Applications, pp-81-86.
- Fadzillah S, Salahuddin M.A dan Shahrul. A.M.Y, 2010, “*Digital Image Classification for Malaysian Blooming Flower*, ” Second International Conference on Computational Intelligence, Modelling and Simulation (CIMSIM), Bali, pp. 33-38
- Fuzy Y.M dan Kana S.S, 2017, “*Klasifikasi Belimbing Menggunakan Naive Bayes Berdasarkan Fitur Warna RGB*”, IJCCS, Vol.11, No.1, pp. 99-108.
- Hong A, Zheru. C, Chen G dan Wang Z, 2003, “*Region-of-interest based flower images retrieval*,” Proc. IEEE International Conference of Acoustics, Speech, and Signal Processing, pp. 589-592.
- Isma I., Sharif M., Mudassar R, dan Musssarat Y. (2014). “*Salt and Pepper Noise Removal Filter for 8-Bit Images Based on Local and Global Occurrences of Grey Levels as Selection Indicator*”, Nepal Journal of Science and Technology Vol. 15, No.2, 123-132.
- Jia. L, James. Z.W dan Gio. W, 2000, “*Integrated Region Matching for Image Retrieval*,” In ACM Multimedia, pp. 147-156.
- Kusrini & Emha T.L, 2009, “*Algoritma Data Mining*”, Penerbit Andi, STMIK AMIKOM. Yogyakarta, Indonesia.
- Maria. E, Nilsback dan Andrew. Z, 2006 “*A visual vocabulary for flower classification*,” In CVPR, Vol. 2, pp. 1447-1454.
- Riddhi H.S, Narendra, dan Zankhana, 2017, “*Flower Classification Using Texture and Color Features*”, ICRISSET2017. International Conference on Research and Innovations in Science, Engineering & Technology. Vol.2, pp.113-118.
- Yuita A.S dan Nanik S, 2014, “*Flower Classification using Combined a\* b\* Color and Fractal-based Texture Feature*”, International Journal of Hybrid Information Technology Vol.7, pp.357-368.