

Pengenalan Karakter Huruf Hiragana Menerapkan Metode Template Matching Correlation

Theresia Sri Melati Pasaribu^{1*}

¹Fakultas Ilmu Komputer dan Teknologi Informasi, Program Studi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia

Email: ^{1*}teremelati.pasaribu@gmail.com

Abstrak

Bahasa Jepang merupakan salah satu bahasa yang lumayan sulit untuk dipelajari, karena kita harus mengerti keseluruhan hurufnya, tata cara penulisannya dan tata cara bacanya. Karakter huruf Jepang merupakan karakter yang cukup kompleks dan memiliki karakteristik unik dibandingkan dengan karakter latin apabila ditulis dengan tulisan tangan, dimana terdapat keterkaitan antar piksel. Gaya penulisan dari masing-masing jenis tulisan bisa sangat bervariasi untuk setiap orang. Perbedaanannya lebih terasa untuk tulisan Hiragana karena bersifat lebih kursif. Oleh karena itu muncullah ide untuk membuat sebuah perancangan aplikasi yang dapat mengenali pengenalan huruf Jepang atau huruf Hiragana dengan menggunakan metode Template Matching Correlation. Huruf dalam bentuk citra sehingga nilai pixel citra tersebut dilakukan pencocokan dengan teknik metode Template Matching Correlation. Template Matching Correlation merupakan metode pencocokan setiap piksel pada suatu matriks citra digital dengan citra yang menjadi acuan untuk menguji kedekatan sebuah data dengan yang lain. Sehingga pada metode Template Matching Correlation, nantinya akan terdapat data yang dijadikan referensi untuk dibandingkan dengan data yang akan diuji. Semakin besar nilai perhitungan korelasinya, maka kemiripan antara data uji dengan sample template yang bersangkutan semakin besar. Metode korelasi ini memiliki nilai keakuratan yang cukup besar atau tepat.

Kata Kunci : Huruf Hiragana, Pengenalan Pola, Template Matching Correlation

1. PENDAHULUAN

Tulisan bahasa Jepang terdiri dari tiga jenis yaitu Kanji logografi (karakter Tionghoa), syllabic Kana. Kana adalah sistem syllabic, di mana satu pengucapan selalu diwakili oleh sepasang karakter yaitu, Hiragana dan Katakana. Sebaliknya, sebagian besar karakter Kanji yang digunakan dalam bahasa Jepang sehari-hari memiliki dua atau lebih pengucapan yang berbeda. Karena ada jenis karakter Hiragana dan Katakana pada setiap pengucapan sehingga terdapat tantangan dalam mengenali setiap jenis sistem tulisannya. Selain itu, gaya penulisan dari masing-masing jenis tulisan bisa sangat bervariasi untuk setiap orang. Perbedaanannya lebih terasa untuk tulisan Hiragana karena bersifat lebih kursif. Oleh karena itu muncullah ide untuk membuat sebuah perancangan aplikasi yang dapat mengenali pengenalan huruf Jepang atau huruf Hiragana dengan menggunakan metode Template Matching Correlation. Huruf dalam bentuk citra sehingga nilai pixel citra tersebut dilakukan pencocokan dengan teknik metode Template Matching Correlation.

Template Matching Correlation merupakan metode pencocokan setiap piksel pada suatu matriks citra digital dengan citra yang menjadi acuan untuk menguji kedekatan sebuah data dengan yang lain. Sehingga pada metode Template Matching Correlation, nantinya akan terdapat data yang dijadikan referensi untuk dibandingkan dengan data yang akan diuji. Semakin besar nilai perhitungan korelasinya, maka kemiripan antara data uji dengan sample template yang bersangkutan semakin besar. Metode korelasi ini memiliki nilai keakuratan yang cukup besar atau tepat.

Melihat tingkat keberhasilan yang cukup tinggi yang dihasilkan oleh para peneliti sebelumnya maka penulis akan menggunakan metode Template Matching Correlation untuk mengenali pola tulisan tangan huruf Jepang Hiragana dasar.

2. TEORITIS

A. Pengenalan Pola

Pengenalan pola adalah salah satu bidang dalam pembelajaran mesin (*machine learning*) yang menitikberatkan pada metode klasifikasi objek ke dalam kelas-kelas tertentu untuk menyelesaikan masalah tertentu. Pengenalan pola bisa menjadikan mesin secerdas manusia untuk mengenali pola dalam berbagai kondisi lingkungan [2]. Aplikasi pengenalan pola bertujuan untuk melakukan proses pengenalan terhadap suatu objek (misalnya citra) ke dalam salah satu kelas tertentu, berdasarkan pola yang dimilikinya. Di dalam mengenali sebuah pola, komputer memerlukan sebuah algoritma untuk menentukan tingkat kesamaan antara pola uji dan pola yang ada pada database. Beberapa contoh pengenalan pola misalnya pengenalan pola sidik jari, tanda tangan dan lain-lain.

B. Huruf Hiragana

Huruf Hiragana pada awalnya terdiri dari 48 suku kata, namun simbol untuk suku kata “Wi” dan “We” telah dihapuskan oleh Departemen Pendidikan Jepang pada tahun 1946. Sehingga menjadi 46 suku kata. Perubahan ini merupakan sebagian dari perbaikan suku kata. Fungsi huruf *Hiragana* adalah untuk menuliskan kata-kata asli Bahasa Jepang yang bukan serapan bahasa asing dan sebagai *Furigana* yaitu kata yang menerangkan bagaimana membaca sebuah huruf *Kanji*. Huruf *Hiragana* dibagi menjadi dua kelompok, yaitu huruf hiragana dasar dan imbuhan [3].

C. Template Matching Correlation

Template Matching Correlation merupakan metode pencocokan setiap piksel pada suatu matriks citra digital dengan citra yang menjadi acuan untuk menguji kemiripan sebuah data dengan yang lain. Pencocokan citra yang menghasilkan tingkat kemiripan/kesamaan yang tinggi

menentukan suatu citra tersebut dikenali sebagai salah satu dari citra template.

Data inputan dan template gambar, nantinya akan menghasilkan data berupa matriks. Dan akan dicari nilai kesamaan antar dua buah matriks tersebut dengan menghitung nilai korelasinya. Nilai korelasi dua buah matriks didapatkan dengan menggunakan rumus korelasi seperti yang ditunjukkan pada perhitungan 2.1 berikut ini [3] [4].

$$r = \frac{\sum_{k=1}^n (x_{ik} - \bar{x}_i) \cdot (x_{jk} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ik} - \bar{x}_i)^2 \cdot \sum_{k=1}^n (x_{jk} - \bar{x}_j)^2}} \quad (1)$$

Nilai \bar{x}_i dirumuskan dengan persamaan 2. dan \bar{x}_j dirumuskan dengan persamaan 3.

$$\bar{x}_i = 1/n \sum_{k=1}^n x_{ik} \quad (2)$$

$$\bar{x}_j = 1/n \sum_{k=1}^n x_{jk} \quad (3)$$

Keterangan :

r = nilai korelasi antara dua buah matriks (rentang nilai antara -1 dan +1)

x_{ik} = nilai piksel ke-k dalam matriks i

x_{jk} = nilai piksel ke-k dalam matriks j

\bar{x}_i = rata-rata nilai piksel matriks i

\bar{x}_j = rata-rata nilai piksel matriks j

n = jumlah piksel dalam suatu matriks

3. ANALISA

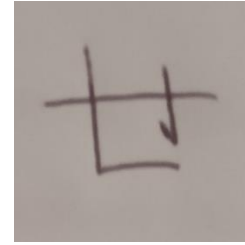
A. Analisa Masalah

Pada analisa masalah ini, penulis akan memberikan pengenalan gambar pola pada huruf-huruf Jepang Hiragana, data atau sampel yang akan digunakan adalah tulisan tangan yang ditulis dikertas putih dengan tinta berwarna hitam. Kemudian data tersebut diinputkan pada sebuah kamera digital dimana terdiri dari beberapa sampel data yang dipilih yaitu huruf 'se' dari Huruf Hiragana dasar. Tulisan tangan yang diperoleh akan dilihat tingkat perbedaan pada bentuk pola penulisan, dan dapat dinilai bobot matriks dari setiap huruf dengan menerapkan citra digital agar lebih mudah dilihat dan dapat mudah dibaca, maupun ditulis. Ketebalan garis, tinggi garis dan kelengkungan garis sangat berkaitan dengan pengenalan pola Huruf Hiragana pada setiap tulisan tangan, apabila tingkat penulisan sangat berbeda maka akan dapat mempengaruhi dalam pengucapan hurufnya dan nilai pengenalan hasil pembobotan pada citra digital. Oleh karena itu, untuk melakukan pengenalan tulisan tangan berbentuk huruf 'se' menggunakan metode *Template Matching Correlation* harus melalui tahapan-tahapan tertentu sehingga dapat menjadi masukan yang baik. Jumlah data yang diambil sebanyak 3 pola tulisan tangan berbentuk Huruf Hiragana dari 3 orang. Maka pada penelitian ini penulis menggunakan metode *Template Matching Correlation* sebagai proses penerapan dari pengenalan pola Huruf Hiragana.

B. Penerapan Template Matching Correlation

Dalam merancang sebuah sistem, khususnya sistem yang berbasisi aplikasi perlu dilakukan analisa. Analisa berguna untuk meminimalisir terjadinya kesalahan pada saat akan dirancangnya sebuah aplikasi. Analisa merupakan

upaya untuk melakukan pemahaman tertentu terhadap sesuatu masalah yang dilakukan dalam pengkajian. Pemahaman tersebut bertujuan mengetahui kekurangan-kekurangan pada metode atau metode yang digunakan.



Gambar 1. Huruf Hiragana Dasar 'se'

Gambar di atas yang akan dianalisa yang berbentuk citra RGB dengan ukuran 2 X 3 dengan menghasilkan nilai piksel yang dapat dilihat pada gambar di bawah ini.

Template Matching Correlation adalah salah satu teknik dalam pengolahan citra digital yang berfungsi untuk mencocokkan tiap-tiap bagian dari suatu acuan dengan citra yang menjadi Template (acuan). Dari matriks citra di atas akan dilakukan perhitungan dengan metode Template Matching Correlation untuk menyesuaikan gambar masukkan dan gambar keluaran. Adapun jumlah perbedaan piksel dinyatakan juga sebagai algoritma pembandingan antara image sketsa terhadap template citra.

Tabel 1. Nilai RGB Citra Resolusi 2 x 3

R = 153	R = 146	R = 156
G = 142	G = 135	G = 145
B = 136	B = 129	B = 139
R = 151	R = 144	R = 153
G = 140	G = 133	G = 142
B = 134	B = 127	B = 136

Nilai citra di atas resolusi 2 x 3 piksel, nilai piksel diambil dengan menggunakan aplikasi matlab. Nilai grayscale citra tersebut atau derajat keabuannya sama yaitu nilai red = green = blue.

Penyelesaian :

$$r = \frac{\sum_{k=1}^n (x_{ik} - \bar{x}_i) \cdot (x_{jk} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ik} - \bar{x}_i)^2 \cdot \sum_{k=1}^n (x_{jk} - \bar{x}_j)^2}}$$

Keterangan :

r = nilai korelasi antara dua buah matriks (rentang nilai antara -1 dan +1)

x_{ik} = nilai piksel ke-k dalam matriks i

x_{jk} = nilai piksel ke-k dalam matriks j

\bar{x}_i = rata-rata nilai piksel matriks i

\bar{x}_j = rata-rata nilai piksel matriks j

n = jumlah piksel dalam suatu matriks

$$\bar{x}_i = \frac{1}{n} \sum_{k=1}^n x_{ik}$$

$$\bar{x}_i = \frac{1}{18} (153 + 142 + 136 + 146 + 135 + 129 + 156 + 145 + 139)$$

$$\bar{x}_i = \frac{1}{18}(1.281) = 71,166$$

$$\bar{x}_i = \frac{1}{18}(151 + 140 + 134 + 144 + 133 + 127 + 153 + 142 + 136)$$

$$\bar{x}_i = \frac{1}{18}(1.260) = 70$$

$$\bar{x}_i = 71,166 + 70 = 141,166$$

$$\bar{x}_i = \frac{141,166}{2} = 70,583$$

Nilai rata-rata matrik i adalah $\bar{x}_i = 70,583$

$$\bar{x}_j = \frac{1}{n} \sum_{k=1}^n x_{ik}$$

$$\bar{x}_j = \frac{1}{18}(153 + 142 + 136 + 151 + 140 + 134)$$

$$\bar{x}_j = \frac{1}{18}(856) = 47,555$$

$$\bar{x}_j = \frac{1}{18}(146 + 135 + 129 + 144 + 133 + 127)$$

$$\bar{x}_j = \frac{1}{18}(814) = 45,222$$

$$\bar{x}_j = \frac{1}{18}(156 + 145 + 139 + 153 + 142 + 136)$$

$$\bar{x}_j = \frac{1}{18}(871) = 48,388$$

$$\bar{x}_j = 47,555 + 45,222 + 48,388$$

$$\bar{x}_i = \frac{141,165}{3} = 47,055$$

Nilai rata-rata matriks i adalah $\bar{x}_j = 47,055$

$$(x_{ik} - \bar{x}_i) = (153 - 70,583) + (142 - 70,583) + (136 - 70,583) + (146 - 70,583) + (135 - 70,583) + (129 - 70,583) + (156 - 70,583) + (145 - 70,583) + (139 - 70,583)$$

$$(x_{ik} - \bar{x}_i) = (82,417) + (71,417) + (65,417) + (75,417) + (64,417) + (58,417) + (85,417) + (74,417) + (68,417)$$

$$(x_{ik} - \bar{x}_i) = 645,753$$

$$(x_{ik} - \bar{x}_i) = (151 - 70,583) + (140 - 70,583) + (134 - 70,583) + (144 - 70,583) + (133 - 70,583) + (127 - 70,583) + (153 - 70,583) + (142 - 70,583) + (136 - 70,583)$$

$$(x_{ik} - \bar{x}_i) = (80,417) + (69,417) + (63,417) + (73,417) + (62,417) + (56,417) + (82,417) + (71,417) + (65,417)$$

$$(x_{ik} - \bar{x}_i) = 624,753$$

$$\sum_{k=1}^n (x_{ik} - \bar{x}_i) = 645,753 + 624,753$$

$$\sum_{k=1}^n (x_{ik} - \bar{x}_i) = 1.270,506$$

Dari perhitungan nilai template pada citra didapat nilai

$$\sum_{k=1}^n (x_{ik} - \bar{x}_i) = 1.270,506$$



Gambar2. Citra Acuan

Tabel 2. Nilai RGB Citra Acuan

R = 220	R = 222	R = 222
G = 220	G = 222	G = 222
B = 220	B = 222	B = 222
R = 217	R = 220	R = 220
G = 217	G = 220	G = 220
B = 217	B = 220	B = 220

Nilai piksel di atas merupakan matriks dari target yang dijadikan acuan untuk mencocokkan dengan template secara keseluruhan. Matrik citra di atas merupakan nilai piksel citra yang akan dicocokkan pada citra acuan dengan menggunakan rumus. $(x_{jk} - \bar{x}_j)$

$$(x_{jk} - \bar{x}_j) = (220 - 47,055) + (220 - 47,055)$$

$$+ (220 - 47,055) + (222 - 47,055)$$

$$+ (222 - 47,055) + (222 - 47,055)$$

$$+ (222 - 47,055) + (222 - 47,055)$$

$$+ (222 - 47,055)$$

$$(x_{jk} - \bar{x}_j) = (172,945) + (172,945) + (172,945)$$

$$+ (174,945) + (174,945) + (174,945)$$

$$+ (174,945) + (174,945) + (174,945)$$

$$(x_{jk} - \bar{x}_j) = 1.565,505$$

$$(x_{jk} - \bar{x}_j) = (217 - 47,055) + (217 - 47,055)$$

$$+ (217 - 47,055) + (220 - 47,055)$$

$$+ (220 - 47,055) + (220 - 47,055)$$

$$+ (220 - 47,055) + (220 - 47,055)$$

$$+ (220 - 47,055)$$

$$(x_{jk} - \bar{x}_j) = (169,945) + (169,945) + (169,945)$$

$$+ (172,945) + (172,945) + (172,945)$$

$$+ (172,945) + (172,945) + (172,945)$$

$$(x_{jk} - \bar{x}_j) = 1.547,505$$

$$(x_{jk} - \bar{x}_j) = (153 - 47,055) + (142 - 47,055)$$

$$+ (136 - 47,055) + (151 - 47,055)$$

$$+ (140 - 47,055) + (134 - 47,055)$$

$$(x_{jk} - \bar{x}_j) = (105,945) + (94,945) + (88,945)$$

$$+ (103,945) + (92,945) + (86,945)$$

$$(x_{jk} - \bar{x}_j) = 573,67$$

$$(x_{jk} - \bar{x}_j) = (146 - 47,055) + (135 - 47,055)$$

$$+ (129 - 47,055) + (144 - 47,055)$$

$$+ (133 - 47,055) + (127 - 47,055)$$

$$(x_{jk} - \bar{x}_j) = (98,945) + (87,945) + (81,945) + (96,945) \\ + (85,945) + (79,945)$$

$$(x_{jk} - \bar{x}_j) = 531,67$$

$$(x_{jk} - \bar{x}_j) = (156 - 47,055) + (145 - 47,055) \\ + (139 - 47,055) + (153 - 47,055) \\ + (142 - 47,055) + (136 - 47,055)$$

$$(x_{jk} - \bar{x}_j) = (108,945) + (97,945) + (91,945) \\ + (105,945) + (94,945) + (88,945)$$

$$(x_{jk} - \bar{x}_j) = 588,67$$

$$(x_{jk} - \bar{x}_j) = 573,67 + 531,67 + 588,67$$

$$(x_{jk} - \bar{x}_j) = 1,694.01$$

$$\sum_{k=1}^n (x_{jk} - \bar{x}_j)^2 = (1.694,01)^2$$

$$\sum_{k=1}^n (x_{jk} - \bar{x}_j)^2 = 2.869.669,88$$

$$r = \frac{\sum_{k=1}^n (x_{ik} - \bar{x}_i) \cdot (x_{jk} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ik} - \bar{x}_i)^2 \cdot \sum_{k=1}^n (x_{jk} - \bar{x}_j)^2}}$$

$$r = \frac{1.270,506 \cdot 1.694,01}{\sqrt{(1.270,506)^2 \cdot (1.694,01)^2}}$$

$$r = \frac{1.270,506 \cdot 1.694,01}{\sqrt{1.614.185,5 \cdot 2.869.669,88}}$$

$$r = \frac{2.152.249,87}{2.152.249,87}$$

$$r = 1$$

Nilai korelasi dari perhitungan nilai piksel dengan menggunakan metode Template Matching Correlation adalah sebesar 1 yang artinya nilai korelasi adalah kecocokan pola sesuai dengan yang diharapkan.

4. IMPLEMENTASI

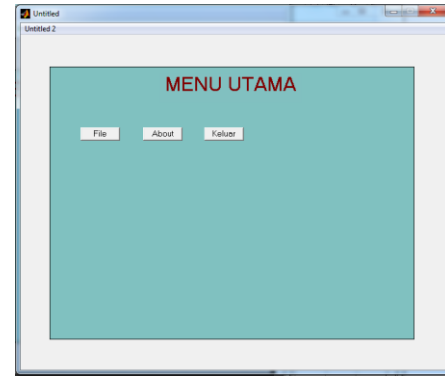
Berikut ini merupakan kebutuhan sistem yaitu Perangkat keras dan perangkat lunak yang digunakan adalah memiliki spesifikasi minimal sebagai berikut :

- Prosesor Intel Pentium IV 1,8 Ghz
- Memory 512 Mb
- Harddisk minimal 80 Gb
- VGA 128 Mb
- Monitor dengan resolusi 800x600
- Keyboard dan Mouse
- Sistem Operasi Windows 7

A. Tampilan Pengujian

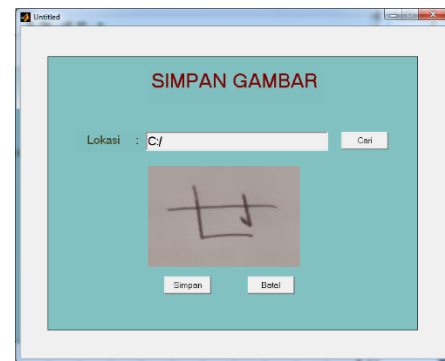
Berikut adalah tampilan program hasil pengujian metode template matching correlation dengan menggunakan aplikasi matlab, dapat dilihat pada gambar berikut ini :

Menu Utama ini terdiri dari 3 toolbar, yaitu File, About, dan Keluar. File berfungsi untuk menampilkan proses dari algoritma Template Matching Correlation, About berisi data singkat tentang penulis, Keluar untuk keluar dari program.



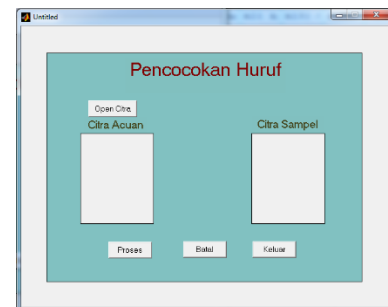
Gambar 3. Tampilan Menu Utama

Tampilan ini berfungsi untuk menyimpan gambar yang akan di jadikan sebagai gambar sampel. Terdiri dari bar link pencarian gambar, tombol cari untuk mencari gambar berdasarkan link yang telah di inputkan, gambar yang sudah di temukan untuk memastikan bahwa gambar yang akan di gunakan tidak salah, tombol simpan untuk menyimpan gambar dan tombol batal untuk membatalkan gambar apabila gambar salah.



Gambar 4. Tampilan simpan gambar

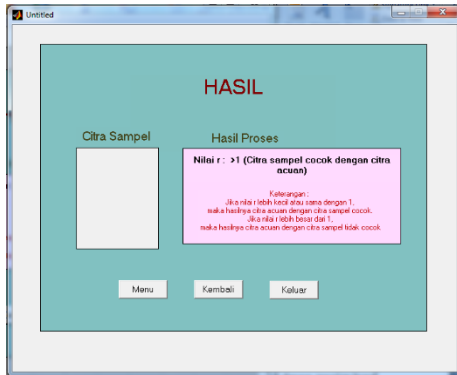
Pada tampilan ini gambar yang telah di inputkan akan di cocokkan pada gambar sampel. Tampilan ini berisi citra acuan, citra sampel, tombol proses di mana pada gambar acuan dan gambar sampel akan di hitung nilai korelasi dari kedua matriks citra tersebut, tombol batal untuk kembali ke menu penyimpanan gambar dan tombol keluar untuk keluar dari program.



Gambar 5. Tampilan Pencocokan Huruf

Tampilan penyimpanan citra akhir/hasil adalah proses terakhir dalam aplikasi ini. Dalam proses ini citra

akhir akan disimpan di lokasi yang diinginkan pengguna. Citra akhir yang disimpan akan disimpan dalam format jpg. Dalam proses ini berguna agar citra akhir yang dijadikan masukan tidak akan hilang sehingga dapat digunakan lagi di kemudian hari.



Gambar 6. Tampilan Hasil

5. KESIMPULAN

Setelah membahas dan menyajikan semua bab sebelumnya, dapat dibuat kesimpulan Template Matching Correlation merupakan metode pencocokan setiap piksel pada suatu matriks citra digital dengan citra yang menjadi acuan untuk menguji kedekatan sebuah data dengan yang lain. Rentang nilai korelasi adalah antara -1 hingga 1. Jika nilai korelasi semakin mendekati 1 maka tingkat kesamaan antara data testing dengan template huruf yang bersangkutan semakin besar. Posisi huruf dan jenis font yang digunakan pada citra acuan sangat berpengaruh terhadap hasil klasifikasi. Proses prediksi ini akan dilakukan dengan mengubah nilai-nilai piksel aslinya dengan menggunakan template matching.

REFERENCES

- [1] Nurdin Usman., Jakarta, Indonesia: Grasindo, 2002.
- [2] Usman Ahmad, Pengolahan citra digital dan teknik pemrogramannya. Yogyakarta, Indonesia: Graha Ilmu, 2005.
- [3] Ardi Sanjaya², Ahmad Bagus Setawan³ Ariska Fitria Anggelina¹, "Pengenalan Pola Tulisan Huruf Jepang (Hiragana) Menggunakan Partisi Citra," Generation Journal, vol. 2, p. 27, Januari 2018.
- [4] Sinar Sinurat², Henry Kristian Siburian³ Atriyanti Ratu Pelita Dongoran¹, "PERANCANGAN [4] [4] APLIKASI UNTUK PENCOCOKKAN IMAGE BUNGA KAMBOJA DENGAN MENGGUNAKAN METODE TEMPLATE MATCHING," Jurnal Riset Komputer (JURIKOM), vol. 3, p. 59, Agustus 2016.
- [5] Wahana Komputer, Ragam Aplikasi Pengolahan Image dengan Matlab. Jakarta, Indonesia: Elex Media Komputindo, 2013.