

Klasifikasi Layak Makan Buah Apel Berdasarkan Fitur Warna Menggunakan Algoritma KNN dengan Ekstraksi Fitur LBP dan Validasi Model dengan Cross Validation

Deo Haganta Depari

Program Studi Informatika, UPNVJ

Jl. RS. Fatmawati Raya, Pd. Labu, Kec. Cilandak, Kota Depok, Jawa Barat 12450

Email: deohaganta12@gmail.com

Albestty Islamyati Rafeli

Program Studi Informatika, UPNVJ

Jl. RS. Fatmawati Raya, Pd. Labu, Kec. Cilandak, Kota Depok, Jawa Barat 12450

Email: albestty@gmail.com

David

Program Studi Informatika, UPNVJ

Jl. RS. Fatmawati Raya, Pd. Labu, Kec. Cilandak, Kota Depok, Jawa Barat 12450

Email: david123ersa@gmail.com

**hanya berkontribusi pada ide judul.*

Abstract:

The color and appearance of various fruits are sometimes difficult to determine whether the apples are included in the category of ripe or raw, plus it also needs to be seen whether the fruit is suitable for eating or not, therefore in this journal will be discussed and explained how to determine the feasibility of apples to be eaten using the K-NN (K-Nearest Neighbor) and LBP (Local Binary Pattern) and Model Validation methods with Cross Validation, by sampling apple images from a dataset available on the "kaggle.com" site, the data is then processed in order to facilitate the classification process then the results with the best settings obtained from the trial are cross validation scores 0.75 and 60 percent percentage of 20 apple image data.

Keywords: **Apel, K-NN (K-Nearest Neighbour), LBP (Local Binary Pattern) dan Cross Validation**

Abstrak:

Warna dan tampak buah yang beragam terkadang sulit untuk menentukan apakah buah apel tersebut sudah termasuk kategori matang atau mentah, ditambah juga perlu dilihat apakah

buah tersebut layak dimakan atau tidak, oleh karena itu di dalam jurnal ini akan dibahas dan dijelaskan bagaimana cara mengetahui kelayakan buah apel untuk dimakan dengan menggunakan metode K-NN (*K-Nearest Neighbour*) dan LBP (*Local Binary Pattern*) dan Validasi Model dengan Cross Validation, dengan pengambilan sampel citra buah apel dari dataset yang tersedia di situs “kaggle.com”, data tersebut kemudian diproses agar memudahkan proses klasifikasi kemudian didapatkan hasil dengan pengaturan terbaik yang didapat dari uji coba yaitu skor cross validation 0.75 dan persentase 60 persen dari 20 data citra apel.

Kata Kunci: *Apel, K-NN (K-Nearest Neighbour), LBP (Local Binary Pattern) dan Cross Validation*

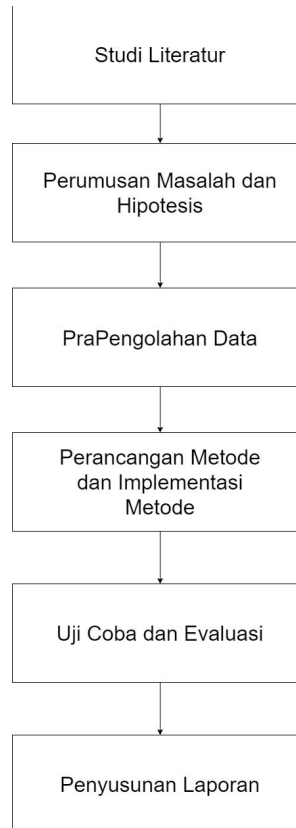
1. Pendahuluan

Penelitian ini berkaitan dengan kondisi pandemic yang disebabkan oleh virus covid 19, pengurangan nya sentuhan atau disebut “*low touch*” di kondisi sekarang merupakan hal yang akan dihadapi kedepannya oleh setiap industri, terutama untuk kebutuhan pangan yaitu makanan, salah satu contohnya adalah buah apel, yang dipetik kemudian dikelompokkan sesuai kematangan dan kelayakan nya, Tentunya hal ini pastinya memerlukan banyak buruh untuk menyentuh dan mengelompokkan apel tersebut.

Hal tersebut merupakan hal yang perlu dihindari dari sekarang dan kedepannya untuk mengurangi penyebaran virus covid 19, maka dari itu perlu adanya sebuah mesin yang otomatis mengelompokkan apel apel tersebut menjadi tiga kategori yaitu: Layak, Kurang Layak dan Tidak layak, Penelitian lebih lanjut dilakukan dengan menggunakan metode Algoritma KNN dengan Ekstraksi Fitur LBP dan Validasi Model dengan *Cross Validation*.

2. Metode penelitian

Secara umum, penelitian ini dilakukan dalam beberapa tahapan yaitu diawali dari studi literatur, perumusan masalah dan hipotesis, perancangan metode dan implementasi serta uji coba dan evaluasi.



Gambar 1. Rancangan Penelitian

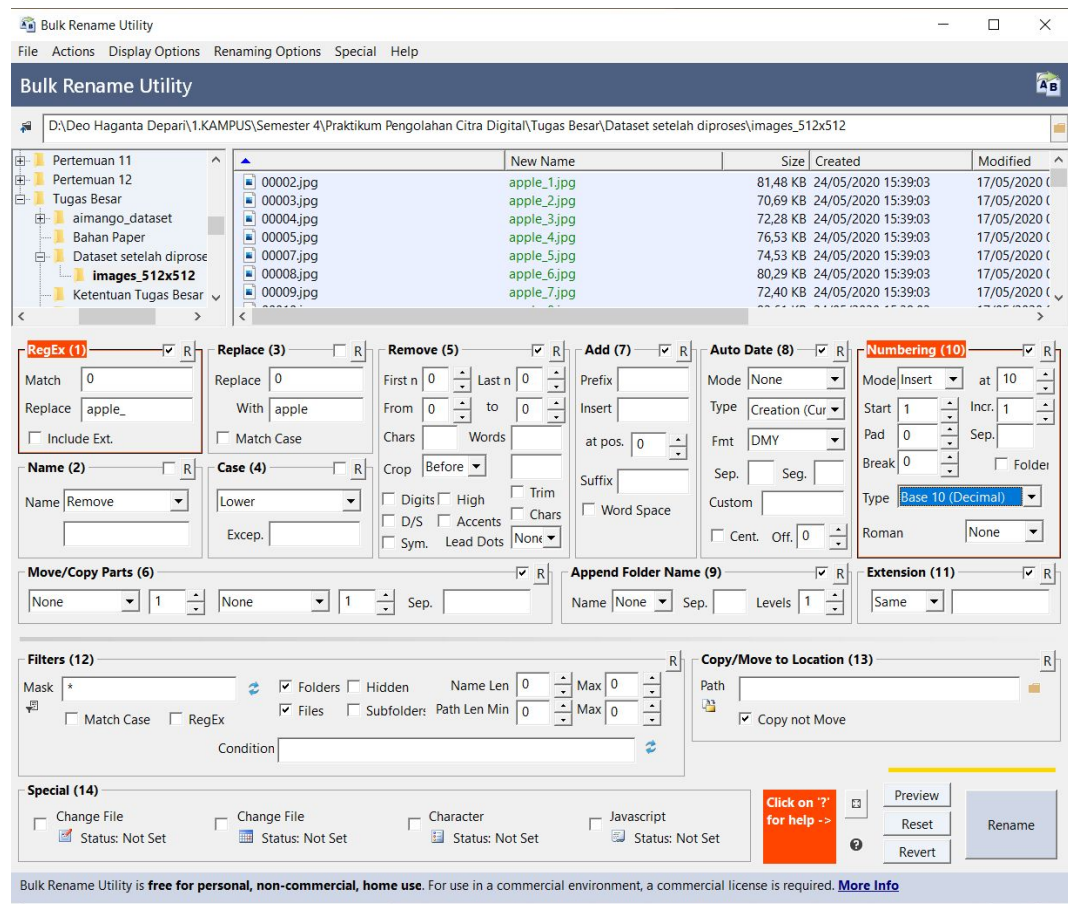
- a. *Studi Literatur*, mempelajari beberapa literatur tentang identifikasi yang berkaitan, mulai dari metode pra proses, metode ekstraksi fitur dan metode pelatihan dan pengenalan
- b. *Perumusan masalah*, merumuskan permasalahan yang akan diteliti dan mencari solusinya, mengungkapkan hipotesis mengenai solusi permasalahan yang diteliti Yang ditemukan ialah untuk menentukan kelayakan makan sebuah apel, maka di awal perlu adanya inspeksi manual yang mengklasifikasi apel tersebut, sehingga kedepannya bisa menggunakan dataset yang telah diklasifikasikan tadi untuk menentukan kelas apakah apel yang baru
- c. *Pra Pengolahan Data*, Data set yang telah diunduh mempunyai beberapa kesalahan penomoran dalam nama citra dan juga dalam pelabelan di file csv, maka dari itu data perlu diolah terlebih dahulu agar saat masuk ke Metode, data mudah untuk di proses

Pertama buat terlebih dahulu folder yang akan menampung dataset yang telah diolah.



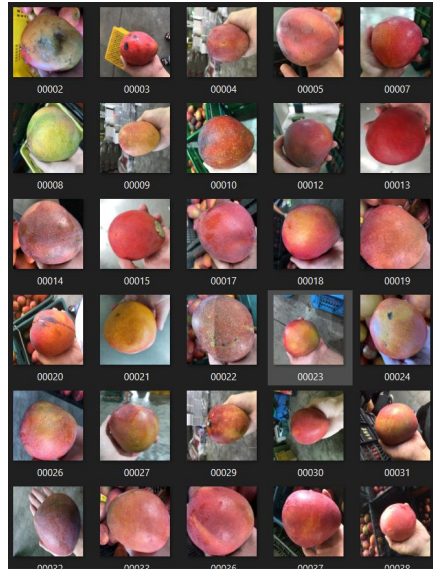
Gambar 2. Folder untuk menampung Dataset yang telah diproses

Dikarenakan dataset mempunyai enam ribu empat ratus citra, maka menggunakan aplikasi “*Bulk Rename Utility*”, akan memudahkan proses penggantian nama citra, sehingga citra yang ada di dataset yang belum berurutan bisa diurutkan.

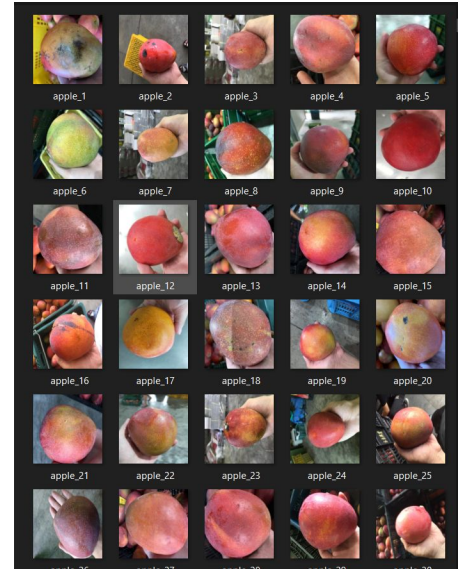


Gambar 3. Merubah Nama Dataset menjadi berurutan

Berikut perubahan nama yang ada di dataset *aimango* yang telah di download dari web *kaggle.com* kemudian semua data citra tersebut diurutkan penamaannya berdasarkan angka setelah itu *background* dari dataset tersebut dihapus lalu di download kembali dengan penambahan nama di belakang nama sebelumnya dengan “*removebg-preview*” dan di hapus satu persatu penambahan nama tersebut.

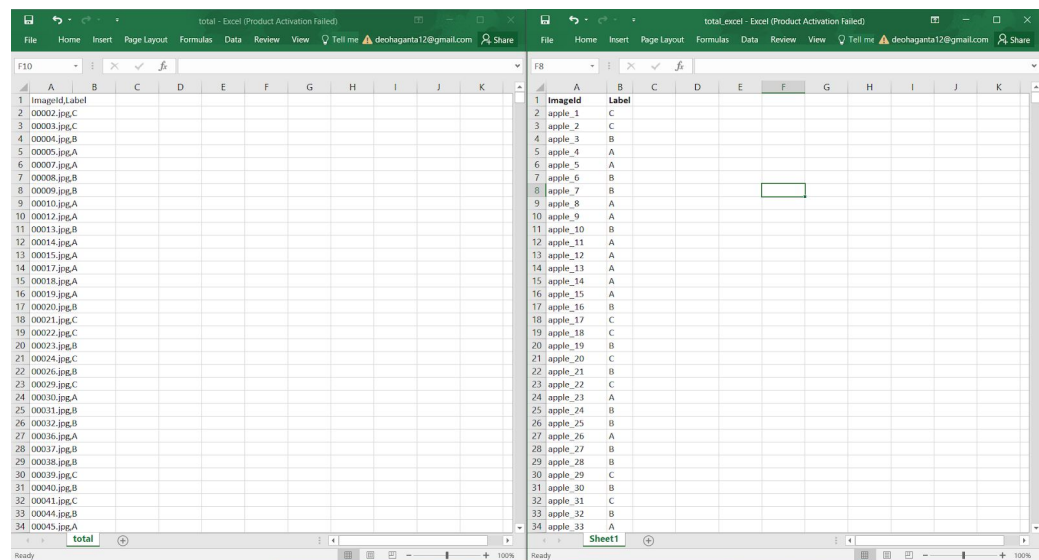


Gambar 4. Citra Sebelum diurutkan



Gambar 5. Citra Setelah diurutkan

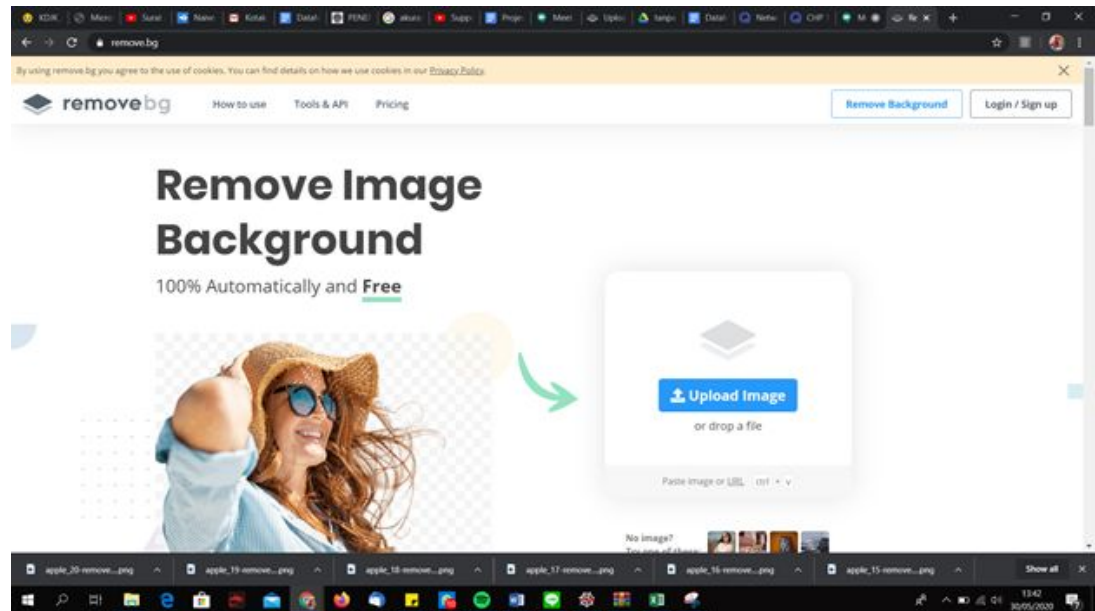
Kemudian perlu juga untuk merubah nama yang ada di file csv, sehingga tidak ada kesalahan label saat masuk implementasi metode nanti



Gambar 6. File CSV sebelum diurutkan dan Excel setelah diurutkan

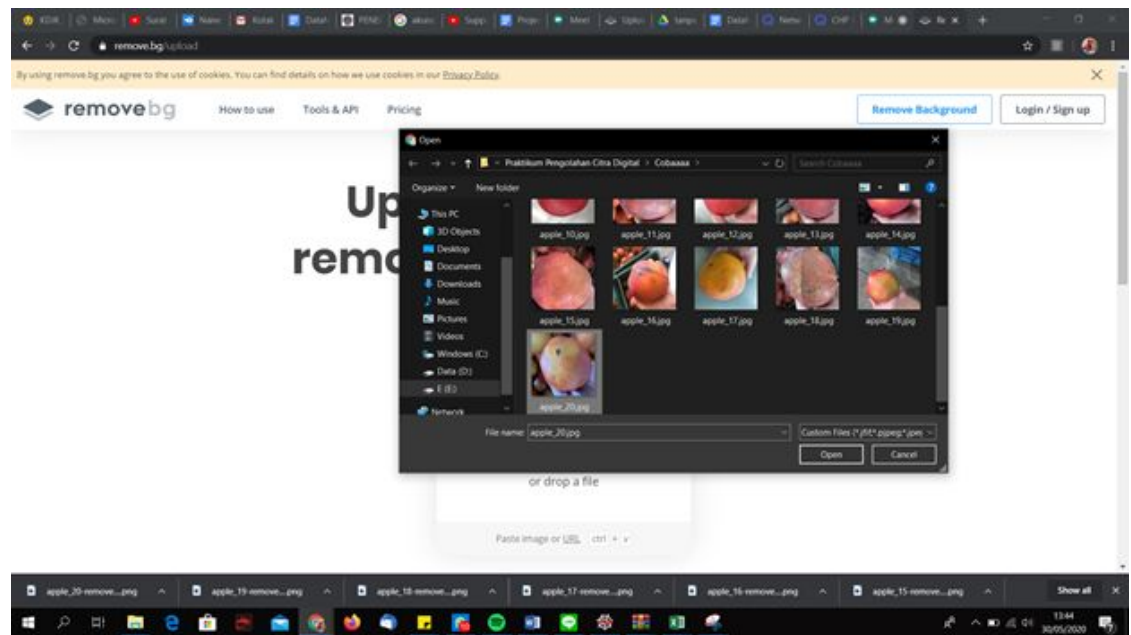
Menghilangkan Background pada Citra

Tulis alamat web, yaitu '*remove.bg*' untuk menghilangkan *background*, kemudian klik *upload file* untuk mengupload citra yang akan dihilangkan *background*nya.



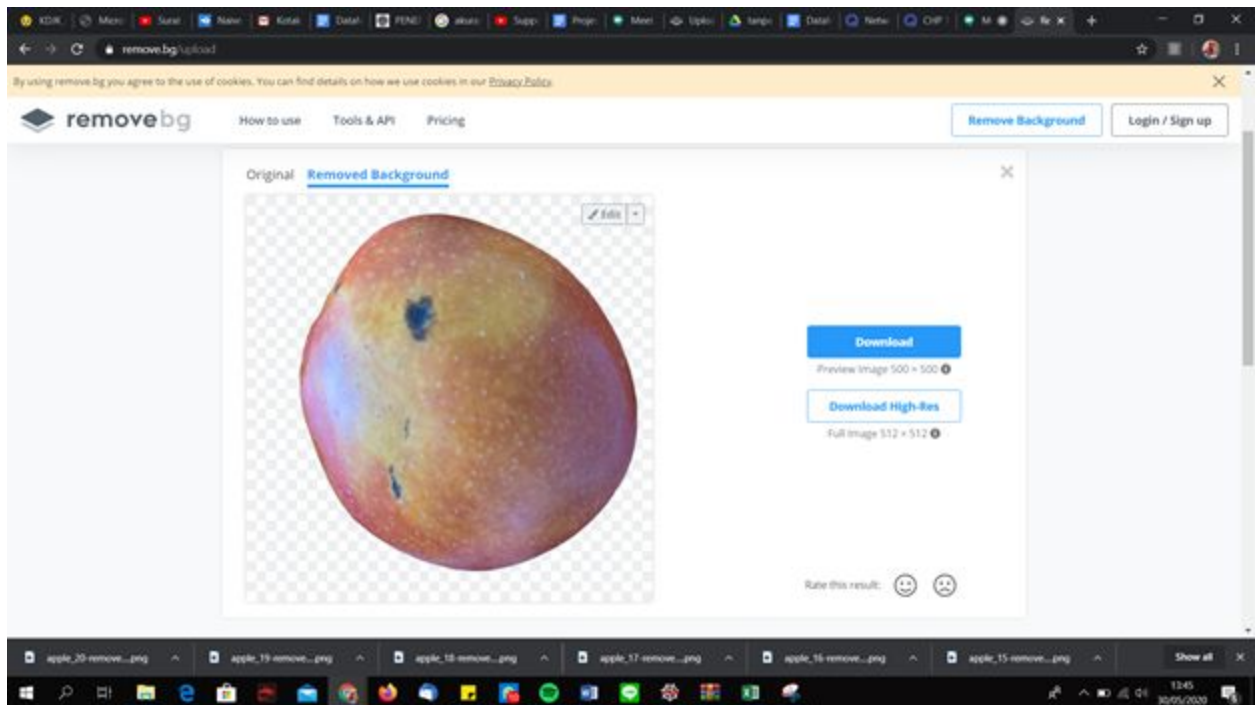
Gambar 7. Tampilan awal pada web *remove.bg*

Pilih gambar yang akan dihilangkan *background*-nya kemudian klik ‘open’.



Gambar 8. Tampilan web setelah mengklik ‘Upload Image’

Pilih ‘download’ untuk mendownload citra yang telah dihilangkan *backgroundnya*.



Gambar 9. Tampilan setelah di pilih gambar yang akan dihilangkan *backgroundnya*

- d. *Perancangan metode dan implementasi*, merancang metode penyelesaian dari masalah yang telah dirumuskan dan mengolah dataset berdasarkan penelitian yang diperoleh dari studi literatur yang berkaitan.

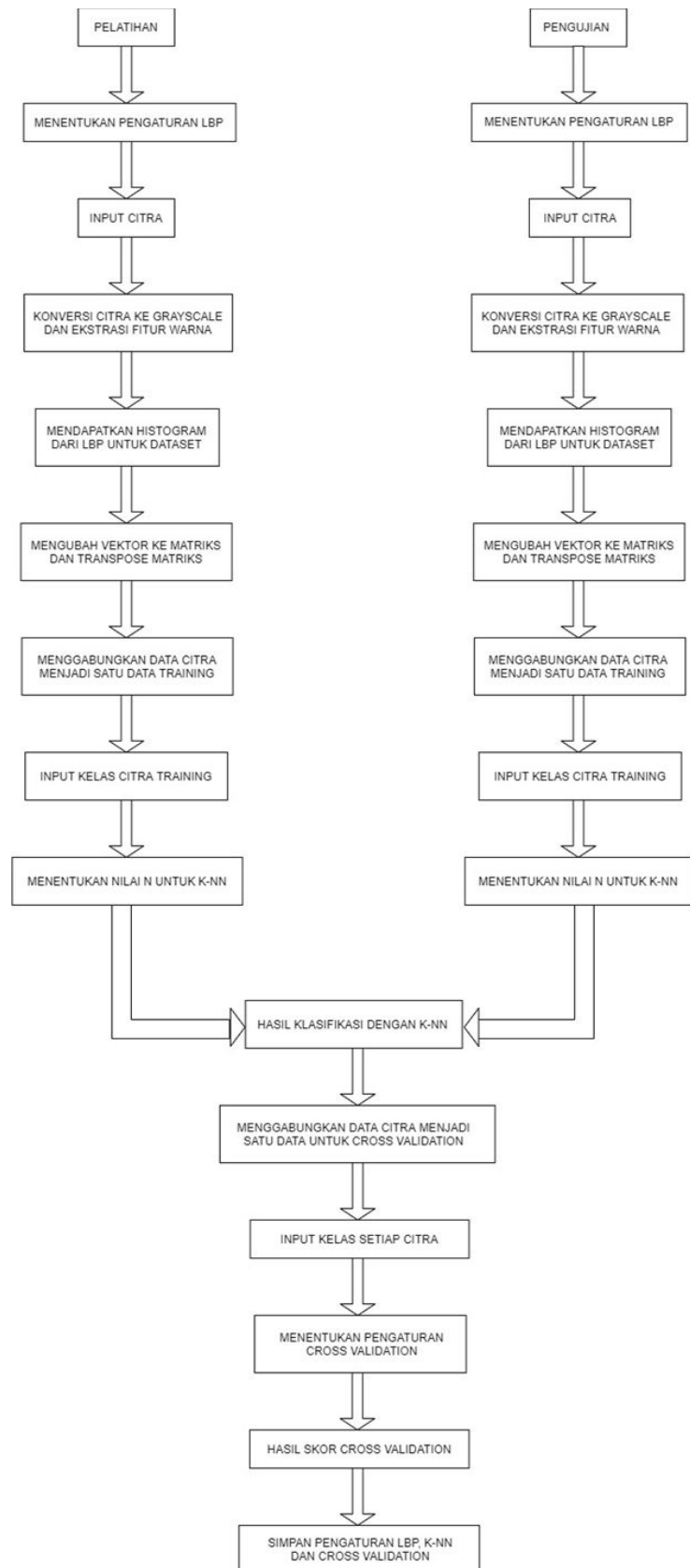
Penelitian ini bertujuan untuk mengklasifikasikan level kematangan buah apel menggunakan metode Algoritma *K-Nearest Neighbors* (K-NN) dengan Ekstraksi Fitur LBP. Citra didapatkan dari situs open source yaitu “kaggle.com”. Rancangan metode dilakukan dengan menambahkan library yang *dibutuhkan*, selanjutnya memasukkan data asli yang berupa citra RGB (Red Green Blue) dan klasifikasi label yang telah didapatkan inspeksi manual di awal, kemudian dilakukan konversi Citra RGB ke *grayscale*, dan melakukan fitur *extraction* dengan *Local Binary Pattern*, setelah itu buat histogram dari Low Binary Pattern untuk dataset lalu mengubah vektor menjadi matriks dan melakukan transpose matriks untuk dataset tersebut.

Setelah itu menggabungkan citra menjadi satu matriks data *training* dan tentukan target melakukan percobaan *K-Nearest Neighbors* (K-NN), penambahan library dilakukan untuk membantu percobaan tersebut.

Pada percobaan tersebut data tidak dipisahkan kedalam data *training* dan data *testing*, tetapi digabungkan pada keseluruhan data dikarenakan dataset tidak dipisah kedalam data *training* dan data *testing* maka *variable response* disesuaikan.

Percobaan tersebut dilakukan dengan ketentuan klasifikasi *K-Nearest Neighbour* (K-NN), selanjutnya menghitung *score* dengan *cross validation* dan menggunakan perhitungan akurasi dengan parameter *cross validation*

- e. *Uji coba dan Evaluasi*, melakukan pengujian dan evaluasi terhadap metode yang telah dirancang dengan menerapkan beberapa skenario, pengujian dilakukan untuk mendapatkan pengaturan *n neighbors*, *radius*, *score*, metode, dan *cross validation* yang terbaik dan akan digunakan untuk menentukan klasifikasi kelas citra, pengaturan tersebut akan dipilih berdasarkan skor *cross validation* tertinggi, lalu banyak nya kelas yang terklasifikasi dengan benar, dari pengaturan terbaik yang didapatkan, kemudian akan digunakan untuk mengklasifikasi setiap empat citra dari dataset dua puluh citra.



Gambar 10. Rancangan pengujian

Dari proses pengujian tersebut maka didapatkan hasil bahwa pengaturan terbaik adalah:

Score Cross Validation= 0.75, *Radius*=25, *n_neighbors*= 3, *CV*= 5, *Method*=*ror*, dan 1 klasifikasi kelas yang benar

- f. *Penyusunan laporan*, penyusunan laporan dilakukan mulai dari awal hingga akhir penelitian

3. Hasil dan Pembahasan

a. Implementasi

Pengambilan citra didapatkan dari situs *open source*, *kaggle.com*, dengan judul dataset “*aimango*”, citra kemudian diproses untuk memudahkan klasifikasi, lalu dengan pengaturan terbaik yang telah didapatkan sebelumnya saat uji coba dan evaluasi, citra dibaca dengan pengaturan tersebut dan konversi ke *grayscale* dan melakukan *feature extraction* dari setiap citra, kemudian di dapatkan histogram dari *Local Binary Pattern* untuk setiap citra, lalu diubah vektor matriks dan *mentranspose* matriks dari setiap citra.

Kemudian dalam penggabungan data citra menjadi satu matriks untuk data *training*, penggabungan data citra menjadi satu matriks untuk data *testing* dan input kelas citra data *training* akan dilakukan sebanyak lima iterasi, dimana setiap iterasi, dari dua puluh data citra, enam belas akan digunakan sebagai *training* dan empat sebagai *testing*, akan ditukar dan input kelas citra akan diubah sesuai dengan citra yang digunakan untuk *training*

b. Hasil Klasifikasi

Berikut hasil klasifikasi pengujian pada setiap iterasi dengan akurasi “*cross validation*” tujuh puluh lima persen dari dua puluh citra apel,

Tabel 1. Hasil Klasifikasi untuk data *testing* citra 1-4

No	Nama	Target	Hasil	Keterangan
1	Apple_1	Tidak Layak	Layak	Salah
2	Apple_2	Tidak Layak	Layak	Salah

3	Apple_3	Kurang Layak	Kurang Layak	Benar
4	Apple_4	Layak	Layak	Benar

Tabel 2. Hasil Klasifikasi untuk data *testing* citra 5-8

No	Nama	Target	Hasil	Keterangan
1	Apple_5	Layak	Layak	Benar
2	Apple_6	Kurang Layak	Kurang Layak	Benar
3	Apple_7	Kurang Layak	Layak	Benar
4	Apple_8	Layak	Layak	Salah

Tabel 3. Hasil Klasifikasi untuk data *testing* citra 9-12

No	Nama	Target	Hasil	Keterangan
1	Apple_9	Layak	Layak	Benar
2	Apple_10	Kurang Layak	Kurang Layak	Benar
3	Apple_11	Layak	Layak	Benar
4	Apple_12	Layak	Kurang Layak	Salah

Tabel 4. Hasil Klasifikasi untuk data *testing* citra 13-16

No	Nama	Target	Hasil	Keterangan
1	Apple_13	Layak	Layak	Benar
2	Apple_14	Layak	Tidak Layak	Salah
3	Apple_15	Layak	Layak	Benar
4	Apple_16	Kurang Layak	Kurang Layak	Benar

Tabel 5. Hasil Klasifikasi untuk data *testing* citra 17-20

No	Nama	Target	Hasil	Keterangan
1	Apple_17	Tidak Layak	Layak	Salah
2	Apple_18	Tidak Layak	Layak	Salah

3	Apple_19	Kurang Layak	Layak	Salah
4	Apple_20	Tidak Layak	Tidak Layak	Benar

4. Kesimpulan

Dataset yang digunakan dalam penelitian ini menggunakan 20 citra yang diambil dari situ *kaggle.com*, dengan judul dataset “aimango”. Dataset dibagi menjadi 4 data *testing* dan 16 data *training*, citra yang telah diproses sehingga mempunyai *background* putih dan pemilih pengaturan *n neighbors*, radius, skor, metode dan *cross validation* memiliki pengaruh terhadap nilai kelayakan buah apel, penelitian ini dilakukan dengan pengujian untuk proses klasifikasi kematangan buah apel dengan menggunakan metode KNN(*K Nearest Neighbor*) dan *feature extraction* dengan LBP (*Low Binary Pattern*), dan Validasi Model dengan *Cross Validation*, setelah dilakukan pengujian dan klasifikasi didapatkan bahwa skor *cross validation* 0.75 dan persentase 60 persen dari 20 data citra apel.

5. Saran

Di penelitian ini akurasi tidak terlalu besar, dikarenakan data citra yang digunakan tidak cukup banyak. Diharapkan di penelitian selanjutnya akan lebih baik jika data percobaan ditambahkan, karena di penelitian ini hanya menggunakan 20 data citra serta setiap citra harus dihilangkan *background*nya satu persatu dan waktu pengerjaan yang cukup sempit (*deadline* cukup dekat).

Daftar Pustaka

- Muchammad Arief. 2019. “Klasifikasi Kematangan Buah Jeruk Berdasarkan Fitur Warna Menggunakan Metode SVM” Program Teknik Informatika, Fakultas Ilmu Komputer, Universitas Nahdlatul Ulama Sidoarjo, *Volume 4*, Nomor 1, Juli 2019
- Xiejialun. (March 2020). “Aimango”(6th Version) [data file and csv class].
<https://www.kaggle.com/xiejialun/aimango>.