

Klasifikasi Tingkat Kematangan Buah Sawit Berdasarkan Ekstraksi Fitur RGB dan GLCM Menggunakan Algoritma K-NN

Ragil Kurniawan dan Nurahman

Universitas Darwan Ali

Jl. Batu Berlian No.10, Mentawa Baru Hulu, Kec. Mentawa Baru Ketapang,

Kabupaten Kotawaringin Timur, Kalimantan Tengah

Email : ragilk032@gmail.com, nurrahman.ikhitar@gmail.com

Abstrak

Tanaman kelapa sawit merupakan tanaman perkebunan tropis yang memiliki persebaran luas di Indonesia dan merupakan sumber utama produksi minyak kelapa sawit. Dengan adanya produksi buah kelapa sawit berskala besar dari perusahaan kelapa sawit serta perkebunan-perkebunan milik masyarakat lokal, dituntut hadirnya inovasi yang dapat memberi manfaat dan meningkatkan produktivitas perkebunan kelapa sawit itu sendiri. Dengan memperhatikan tingkat kematangan dari buah sawit yang sangat penting dalam penentu hasil dan kualitas panen, dilakukan penelitian yang dapat mengklasifikasi tingkat kematangan buah sawit berdasarkan warna dengan memanfaatkan citra yang diekstraksi menjadi dataset. Penelitian ini menerapkan metode klasifikasi algoritma K-NN terhadap data citra buah kelapa sawit yang diolah dengan bahasa pemrograman python oleh software google colab menggunakan metode ekstraksi fitur RGB dan GLCM. Hasil yang diperoleh berupa dataset yang menyajikan susunan pola numerik dari fitur GLCM dan rata-rata fitur RGB. Klasifikasi dilakukan dengan menggunakan algoritma KNN dan diperoleh hasil akurasi dari nilai $k=9,7,5$, dan 3 senilai 56.00% dan nilai $k=1$ senilai 68.00%.

Kata kunci :Kelapa sawit, Citra, Ekstraksi fitur, RGB, GLCM, Google colab, Klasifikasi, K-NN.

Pendahuluan

Tanaman kelapa sawit (*elaeis guinensis* jack) merupakan tumbuhan tropis dan tergolong ke dalam famili Palmae. Tanaman kelapa sawit berasal dari Nigeria, Afrika Barat. Meski demikian, kelapa sawit dapat tumbuh subur di luar daerah asalnya, termasuk di Indonesia. Menurut [1] tanaman kelapa sawit mulai berbunga dan membentuk buah setelah berusia 2-3 tahun. Buah akan menjadi masak sekitar 5-6 bulan setelah penyerbukan. Proses pematangan buah kelapa sawit dapat diindikasikan melalui perubahan warna kulit buahnya. Buah akan menjadi merah jingga ketika matang. Pada saat buah matang, kandungan minyak pada buah telah maksimal dan Jika terlalu matang, buah kelapa sawit akan jatuh dari tangkai tandannya.

Terdapat beberapa tingkatan dalam kematangan buah tanaman sawit yang akan dipanen. Tiap jenis dan varietas dalam tanaman kelapa sawit memiliki kriteria yang berbeda dalam penentuan tingkat kematangannya. Berdasarkan warna kulit buahnya buah sawit dibagi menjadi tiga jenis va-

rietas [2] yaitu nigrescens, virescens, dan albenscens. Varietas nigrescens memiliki ciri pada waktu muda, buah berwarna ungu kehitaman dan setelah masak berubah menjadi jingga kehitam-hitaman. Tingkat kematangan tersebut akan sangat berpengaruh terhadap kualitas hasil panen, termasuk kualitas minyak sawit yang dihasilkan. Untuk dapat mengetahui tingkat kematangan buah sawit hingga saat ini masih banyak dilakukan dengan metode konvensional, yaitu dengan melihat perubahan warna dan jumlah brondolan yang jatuh. Saat buah dipanen, buah akan dikumpulkan terlebih dahulu dan dipilah sebelum diangkat oleh truk ke pabrik. Buah yang telah matang dan lewat matang akan diangkat oleh truk, sedangkan buah mentah tidak akan diangkat karena tergolong tidak layak untuk diolah. Hanya buah yang layak saja akan diangkat oleh truk dan diolah di pabrik, Jika buah mentah tetap terangkut hingga ke pabrik, akan menimbulkan kerugian, dimana buah mentah tidak akan menghasilkan minyak saat diolah di-pabrik [3]. Metode manual yang dilakukan dalam pemanenan buah sawit dengan menggunakan penga-

matan manusia dan melibatkan banyak pekerja akan memakan waktu banyak dan oleh karena dilakukan oleh manusia, kesalahan seperti terangkutnya buah mentah akan biasa terjadi dalam proses pemanenan.

Tanaman kelapa sawit merupakan tanaman perkebunan dengan persebaran perkebunan kelapa sawitnya yang sangatlah luas di Indonesia. Indonesia sendiri merupakan negara dengan tingkat produksi kelapa sawit terbesar di dunia. Produksi buah kelapa sawit pada perkebunan kelapa sawit yang besar tidak hanya dimiliki oleh perusahaan-perusahaan kelapa sawit saja tetapi juga ada pada perkebunan-perkebunan milik masyarakat lokal. Di era pertumbuhan teknologi yang pesat seperti sekarang ini menuntut hadirnya inovasi yang dapat memberi manfaat dan meningkatkan produktivitas perkebunan kelapa sawit itu sendiri. Penerapan sebuah metode yang dapat secara akurat menentukan tingkat kematangan buah akan sangat membantu dalam proses pemanenan buah sawit. Dengan memanfaatkan kamera untuk mengambil citra digital, dari citra tersebut akan didapatkan informasi yang akan diolah menjadi dataset kemudian diklasifikasikan hingga memperoleh hasil identifikasi tingkat kematangan buah sawit yang baku.

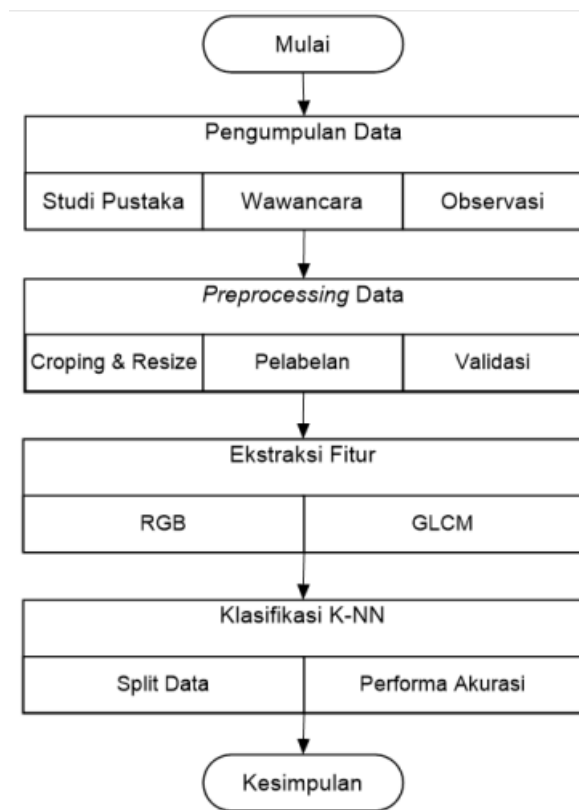
Pengolahan citra digital pada umumnya merupakan proses yang dilakukan untuk mengubah citra analog menjadi digital dan perbaikan kualitas citra. Saat ini pengolahan citra digital telah banyak digunakan untuk berbagai macam kegunaan [4]. Dalam pengolahan citra dilakukan ekstraksi fitur untuk memperoleh dataset yang akan diproses menggunakan metode klasifikasi dan algoritma yang dipilih. Pada penelitian yang dibuat oleh [5] dilakukan ekstraksi fitur warna dengan menghitung rata-rata setiap komponen red, green, dan blue pada data testing dan data training berupa gambar buah jeruk nipis. Kemudian hasil perhitungan dari data testing dan data training tersebut digunakan sebagai inputan pada proses klasifikasi k-nearest neighbor dengan nilai k yang digunakan yakni 1, 3, 5, 7 serta 9. Dari penelitian yang telah dilakukan diperoleh hasil yang menyatakan bahwa dengan menggunakan jarak euclidean $k=7$ dan $k=3$ memiliki nilai persentase 92% dan jarak cityblock distance $k=1$ dan $k=3$ memiliki nilai persentase sebesar 88%. Berdasarkan pernyataan tersebut, fitur warna $k=3$ menunjukkan nilai k terbaik pada klasifikasi tingkat kematangan buah jeruk nipis. Penelitian dengan menggunakan ekstraksi fitur serupa juga dilakukan oleh [6] dengan menjelaskan mengenai pengembangan aplikasi Klasifikasi citra daun Kelapa Sawit yang terkena dampak hama menggunakan metode analisis ekstraksi fitur menggunakan metode Gray Level Co-Occurrence Matrix (GLCM) yang kemudian diklasifikasi menggunakan algoritma k-nearest neighbor (K-NN). Berdasarkan hasil evaluasi dari penelitian yang dilakukan diperoleh nilai akurasi sebesar 83.3%.

Pada penelitian ini dilakukan penerapan metode

klasifikasi algoritma K-NN terhadap data citra yang diolah dengan bahasa pemrograman python oleh software google colab menggunakan metode ekstraksi fitur RGB dan GLCM. Data berupa citra buah kelapa sawit yang diolah menjadi dataset. Data berupa citra diperoleh melalui studi lapangan ke perkebunan tanaman kelapa sawit. Dari studi lapangan dan penelitian yang dilakukan akan didapat hasil klasifikasi dari tingkat kematangan buah kelapa sawit yakni mentah, mengkal, dan masak beserta dengan nilai performa dari klasifikasi tingkat kematangan tersebut.

Metode Penelitian

Terdapat beberapa tahapan dan metode yang digunakan dalam penelitian ini. Tahapan tersebut dimulai dari tahapan pengumpulan data, pengolahan data, tahap ekstraksi fitur RGB dan GLCM dan klasifikasi hasil ekstraksi menggunakan algoritma K-NN. Gambar 1 merupakan diagram tahapan alur penelitian.



Gambar 1: Alur Penelitian

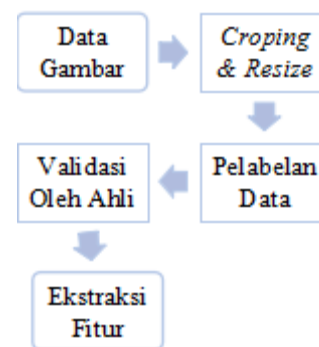
Metode Pengumpulan Data

Tahap pengumpulan data dilakukan dengan melakukan studi kepustakaan, observasi lapangan dan wawancara secara langsung ke pemilik kebun kelapa sawit guna memperoleh informasi yang mendukung penelitian dan data berupa citra yang akan digunakan dalam penelitian ini. Pada penelitian ini studi kepustakaan dilakukan untuk mem-

peroleh referensi dan teori-teori yang mendukung penelitian [7]. Studi kepustakaan dilakukan dengan mengumpulkan informasi dan data, serta dibantu dengan berbagai bahan yang ada di perpustakaan seperti buku referensi, hasil penelitian sebelumnya yang sejenis, artiket, catatan serta berbagai jurnal yang masih berhubungan dengan permasalahan yang ada di dalam penelitian. Untuk sumber data dalam penelitian ini adalah jurnal, dan situs internet yang berhubungan dengan permasalahan yang telah ditentukan. Observasi adalah suatu tindakan terhadap suatu proses atau objek dengan tujuan untuk merasakan dan kemudian memahami pengetahuan tentang suatu fenomena berdasarkan pengetahuan dan gagasan yang telah diketahui sebelumnya [8]. Selain itu, observasi adalah metode atau teknik untuk mengumpulkan data atau informasi dengan mengamati aktivitas yang sedang berlangsung [9]. Dalam penelitian ini, observasi dilakukan pada objek penelitian, yaitu kebun kelapa sawit milik Pak Arifin yang terletak di wilayah Sungai Duren, PT. Sabindo Blok 115, Desa Tabiku, Kecamatan Danau Sembuluh. Observasi dilakukan dengan cara mengamati secara langsung keadaan kebun kelapa sawit tersebut, termasuk kondisi buah kelapa sawit, tingkat kematangan buah, jenis tanaman dan varietas buah kelapa sawit yang ada, serta kegiatan atau proses panen yang dilakukan. Melalui observasi, peneliti dapat memperoleh informasi-informasi yang diperlukan untuk melengkapi data dalam penelitian, memahami lebih lanjut tentang tingkat kematangan buah kelapa sawit di kebun tersebut, dan menggali informasi yang relevan dengan penelitian. Observasi merupakan salah satu metode yang dapat memberikan data yang lebih akurat dan mendalam karena dilakukan secara langsung di lapangan. Informasi yang diperoleh dari observasi ini kemudian dapat digunakan sebagai dasar dalam proses pengolahan citra, ekstraksi fitur, dan klasifikasi tingkat kematangan buah kelapa sawit dalam penelitian ini. Wawancara adalah proses pengumpulan data atau informasi yang terstruktur atau tidak terstruktur dan dapat dilakukan secara langsung atau melalui metode seperti jaringan internet. Tujuan wawancara adalah untuk mendapatkan informasi yang akurat dari sumber yang dapat dipercaya. [9]. Pada penelitian ini wawancara dilakukan dengan pemilik kebun kelapa sawit, yaitu Pak Arifin, yang merupakan narasumber utama dalam pengumpulan data. Dengan melakukan wawancara, peneliti dapat memperoleh data yang lebih komprehensif dan mendalam, serta menggali informasi yang mungkin tidak terdapat dalam sumber lain. Data yang diperoleh dari wawancara ini akan menjadi tambahan penting dalam analisis tingkat kematangan buah kelapa sawit berdasarkan fitur-fitur yang diekstraksi dari citra.

Pre-processing Data

Pada tahapan pre-processing data dilakukan proses resize pada data citra buah kelapa sawit yang pada awalnya berukuran 6000 x 4000 pixel. Proses resize dilakukan untuk memperoleh, nilai citra akan lebih rendah daripada citra aslinya. Ini dimaksudkan agar proses ekstraksi fitur tidak membutuhkan waktu yang lama [10]. Kemudian dilakukan pelabelan dengan memberi label berupa tingkatan kematangan dari masak, mengkal dan mentah. Buah yang masak berwarna kemerahan, mengkal hitam kekuningan dan mentah berwarna ungu kehitaman [1]. Data citra buah kelapa sawit yang telah diberi label kemudian divalidasi melalui penilaian yang dilakukan oleh ahli. Validasi dilakukan oleh Iwan Santosa, M.Si yang merupakan dosen program studi Agribisnis di kampus Universitas Darwan Ali dan telah berpengalaman dibidang tanaman kelapa sawit selama lebih dari 20 tahun. Tahapan pada preprocessing data dapat dilihat pada Gambar 2.



Gambar 2: Preprocessing data

Ekstraksi Fitur

Tahap ekstraksi fitur dilakukan untuk mengidentifikasi dan mengekstraksi fitur-fitur khusus atau karakteristik dari citra tersebut. Fitur-fitur ini dapat mencakup informasi seperti warna, tekstur, bentuk, atau pola yang ada dalam citra. Pada tahap ekstraksi fitur, berbagai teknik dan metode seperti ekstraksi fitur statistik, transformasi domain frekuensi, ekstraksi fitur berbasis bentuk, atau ekstraksi fitur berbasis tekstur dapat digunakan [11]. Dengan menggunakan pengolahan citra sebelum ekstraksi fitur, citra dapat diubah atau disesuaikan agar lebih mudah diekstraksi fiturnya [12]. Hal ini membantu dalam memperoleh fitur-fitur yang lebih relevan dan signifikan untuk analisis atau tugas pengenalan pola yang akan dilakukan selanjutnya, seperti klasifikasi, deteksi objek, atau identifikasi.

RGB (red, green, blue)

Citra berwarna, juga dikenal sebagai citra RGB, terdiri dari warna merah, hijau, dan biru. Perhitungan statistik seperti rerata dan deviasi standar dapat digunakan untuk mengidentifikasi warna [13]. Sebagai ilustrasi, kemampuan ini dapat digunakan untuk identifikasi dan klasifikasi dengan setiap bagian R, G, dan B dihitung.

GLCM (Gray Level Co-Occurrence Matrix)

Ekstraksi fitur GLCM (Gray Level Co-occurrence Matrix) adalah salah satu metode yang umum digunakan dalam pengolahan citra untuk menggambarkan pola tekstur dalam citra. GLCM mengukur hubungan spasial antara dua piksel dalam citra grayscale. Ini adalah metode yang efektif untuk ekstraksi fitur dan analisis tekstur. Metode ini dapat memberikan informasi yang detail tentang citra dari segi tekstur, klasifikasi dan identifikasi. GLCM digambarkan dengan matriks yang berisi frekuensi kemunculan pasangan dua piksel dengan intensitas, jarak, dan orientasi tertentu untuk mendapatkan lima parameter karakteristik: contrast, correlation, energy, homogeneity, dan entropy [14]. Beberapa ciri tekstur pada matriks kookurensi yang bisa diekstraksi yaitu kontras, korelasi, energi dan homogenitas. Berikut penjelasan fitur utama yang digunakan pada penelitian ini yaitu :

1. Energi (energy) Energi merupakan hasil perhitungan yang berkaitan dengan jumlah keberagaman intensitas keabuan dalam citra
2. Korelasi (correlation) Korelasi merupakan representasi dari keterkaitan linear pada derajat citra grayscale. Correlation berkisar dari -1 hingga 1.
3. Kontras (contrast) Kontras merupakan hasil perhitungan yang berkaitan dengan jumlah keberagaman intensitas keabuan dalam citra.
4. Homogenitas (homogeneity) Homogenitas merupakan representasi dari ukuran nilai kesamaan variasi dari intensitas citra. Apabila semua nilai piksel memiliki nilai yang seragam maka homogenitas memiliki nilai maksimum.

Klasifikasi

Klasifikasi dapat didefinisikan sebagai proses pembangunan terhadap suatu model yang mengklasifikasikan sesuatu sesuai dengan karakteristiknya [5]. Klasifikasi adalah cara mengelompokkan objek berdasarkan ciri-ciri yang dimiliki oleh objek klasifikasi. Dalam proses kerja, klasifikasi dapat dilakukan dengan berbagai cara, baik secara manual maupun dengan bantuan teknologi. Untuk klasifikasi yang dilakukan secara manual adalah klasifikasi yang dilakukan oleh manusia tanpa bantuan

algoritma cerdas komputer, dan untuk klasifikasi yang dilakukan dengan bantuan teknologi terdapat beberapa algoritma antara lain nave bayes, support vector machine, pohon keputusan, fuzzy dan artificial. jaringan saraf. [15].

K-Nearest Neighbors (K-NN)

Algoritma K-NN adalah sebuah algoritma untuk mengklasifikasikan objek berdasarkan data latih yang mempunyai jarak paling dekat dengan objek tersebut [16]. Algoritma K-NN tergolong sebagai algoritma supervised. Algoritma K-NN menggunakan klasifikasi ketetanggaan sebagai nilai prediksi dari data training dan data testing.

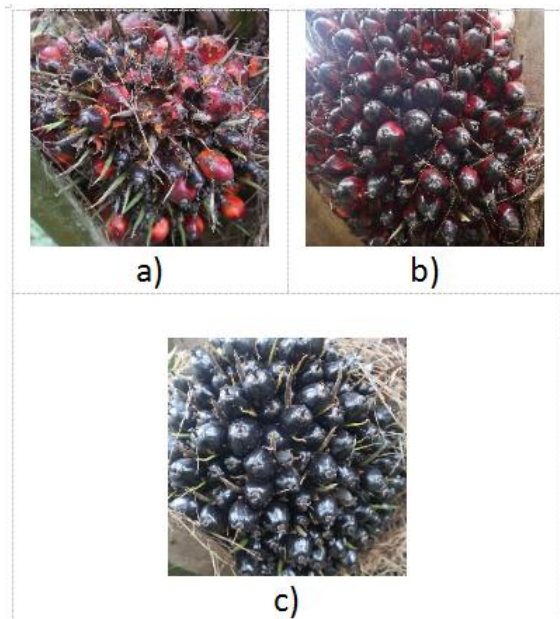
Hasil dan Pembahasan

Preprocessing Data

Pada tahapan pengumpulan data dilakukan Observasi dan wawancara secara langsung di kebun kelapa sawit milik Pak Arifin yang terletak pada wilayah Sungai Duren, PT. Sabindo Blok 115, Desa Tabiku. Kecamatan Danau Sembuluh. Pada tahap ini dilakukan pengumpulan data berupa citra buah kelapa sawit yang diambil dengan ketentuan sebagai berikut:

1. Varietas Lonsum
2. Jenis buah sawit berdasarkan segi warna nigricense
3. Umur tanaman 4-5 tahun
4. Device yang digunakan kamera Canon M100
5. Jarak pengambilan citra 20cm
6. Waktu pengambilan citra pukul 08.00-10.00 dan 14.00-16.30
7. Label atau ketentuan tingkat kematangan buah: Masak, mengkal, mentah

Pengumpulan data citra dilakukan selama 5 hari dan setelah dilakukan tahapan pengumpulan dilakukan kembali pemilahan data citra guna memisahkan citra yang layak dan tidak untuk dilakukan sebagai data penelitian. Citra yang diperoleh dari tahapan sebelumnya sebanyak 100 data citra buah sawit dan akan dilanjutkan untuk tahapan preprocessing data. Tahapan preprocessing data dilakukan dengan menggunakan software Photoshop CS6 hingga diperoleh data citra dengan aspek rasio 1:1 dengan resolusi 1000 x 1000 pixel. Kemudian dilakukan pelabelan tingkat kematangan buah sawit yang divalidasi oleh ahli. Dari tahapan preprocessing data diperoleh total data citra buah sawit sebanyak 100 data dengan 12 data citra buah sawit masak, 13 data citra buah sawit mentah dan 75 data citra buah sawit mentah. Sampel dari data citra buah sawit dapat dilihat dari Gambar 3.



Gambar 3: Gambar data buah sawit: a) buah sawit masak: b) buah sawit mengkal: c) buah sawit mentah

Ekstraksi Fitur RGB dan GLCM Tahapan se-

lanjutnya setelah dilakukan preprocessing data dan diperoleh data yang valid yaitu tahapan ekstraksi fitur RGB dan GLCM. Ekstraksi fitur dilakukan dengan menggunakan software google colab dengan Bahasa pemrograman pemrograman python. Proses ekstraksi fitur dengan google colab dijalankan pada web browser google chrome. Proses ekstraksi fitur dapat dilihat pada Gambar 4 - 8.

Gambar 4 merupakan tampilan source code pada software google colab yang akan menjalankan source code import library. Pada bagian ini, library yang diperlukan diimpor. drive dari google.colab digunakan untuk menghubungkan google colab dengan google drive. greycmatrix dan greycoprops dari skimage.feature digunakan untuk ekstraksi fitur GLCM. cv2 digunakan untuk manipulasi gambar. numpy digunakan untuk operasi numerik. pandas digunakan untuk manipulasi data tabular. os digunakan untuk berinteraksi dengan sistem operasi.

Pada gambar 5 Fungsi extract_rgb_features: Fungsi ini mengambil path gambar sebagai input, membaca gambar menggunakan cv2, mengubah mode warna citra menjadi RGB menggunakan cv2.cvtColor, dan kemudian menghitung rata-rata dan standar deviasi untuk setiap saluran RGB.

```

def extract_rgb(image_path):
    image = cv2.imread(image_path)
    image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)

    # Ekstraksi fitur RGB
    red_mean = np.mean(image[:, :, 0])
    green_mean = np.mean(image[:, :, 1])
    blue_mean = np.mean(image[:, :, 2])

    red_std = np.std(image[:, :, 0])
    green_std = np.std(image[:, :, 1])
    blue_std = np.std(image[:, :, 2])

    # Mengembalikan nilai fitur RGB dalam bentuk array
    return [red_mean, green_mean, blue_mean, red_std, green_std, blue_std]
    
```

Gambar 4: Source code import library

```

def extract_rgb(image_path):
    image = cv2.imread(image_path)
    image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)

    # Ekstraksi fitur RGB
    red_mean = np.mean(image[:, :, 0])
    green_mean = np.mean(image[:, :, 1])
    blue_mean = np.mean(image[:, :, 2])

    red_std = np.std(image[:, :, 0])
    green_std = np.std(image[:, :, 1])
    blue_std = np.std(image[:, :, 2])

    # Mengembalikan nilai fitur RGB dalam bentuk array
    return [red_mean, green_mean, blue_mean, red_std, green_std, blue_std]
    
```

Gambar 5: Source code extract rgb features

```

def extract_glm_features(image_path):
    image = cv2.imread(image_path)
    gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)

    # Hitung matriks GLCM
    glcm = greycomatrix(gray, [1], [0, np.pi/4, np.pi/2, 3*np.pi/4], levels=256, symmetric=True, normed=True)

    # Hitung fitur dari matriks GLCM
    contrast = greycoprops(glcm, prop='contrast')
    homogeneity = greycoprops(glcm, prop='homogeneity')
    energy = greycoprops(glcm, prop='energy')
    correlation = greycoprops(glcm, prop='correlation')

    # Hitung rata-rata dari fitur GLCM
    contrast_mean = np.mean(contrast)
    homogeneity_mean = np.mean(homogeneity)
    energy_mean = np.mean(energy)
    correlation_mean = np.mean(correlation)

    # Mengembalikan nilai fitur GLCM dalam bentuk array
    return [contrast_mean, homogeneity_mean, energy_mean, correlation_mean]
    
```

Gambar 6: Source code extract glm features

```

def create_glm_dataset(folder_path):
    columns = ["Image", "Red mean", "Green mean", "Blue mean", "Red std", "Green std", "Blue std",
               "Contrast_mean", "Homogeneity_mean", "Energy_mean", "Correlation_mean"]
    dataframe = pd.DataFrame(columns=columns)
    for filename in os.listdir(folder_path):
        if filename.endswith(".jpg") or filename.endswith(".png"):
            image_path = os.path.join(folder_path, filename)
            rgb_features = extract_rgb_features(image_path)
            glm_features = extract_glm_features(image_path)
            row = pd.DataFrame({"Image": filename,
                               "Red_mean": [rgb_features[0]],
                               "Green_mean": [rgb_features[1]],
                               "Blue_mean": [rgb_features[2]],
                               "Red_std": [rgb_features[3]],
                               "Green_std": [rgb_features[4]],
                               "Blue_std": [rgb_features[5]],
                               "Contrast_mean": [glm_features[0]],
                               "Homogeneity_mean": [glm_features[1]],
                               "Energy_mean": [glm_features[2]],
                               "Correlation_mean": [glm_features[3]]})
            dataframe = pd.concat([dataframe, row], ignore_index=True)
    return dataframe
    
```

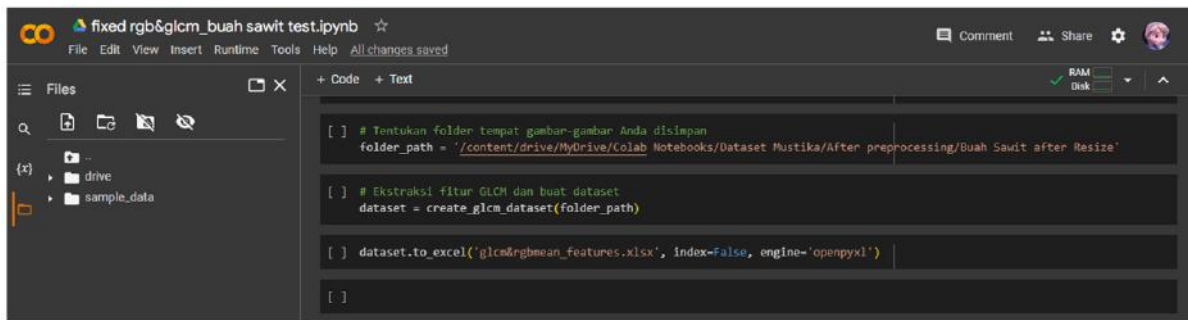
Gambar 7: Source code extract glm dataset

Pada gambar 6 Fungsi `extract_glm_features`: Fungsi ini mengambil path data gambar sebagai input, membaca gambar menggunakan `cv2`, dan mengubahnya menjadi grayscale menggunakan `cv2.cvtColor`. Selanjutnya, fungsi `greycomatrix` dari `skimage.feature` digunakan untuk menghitung matriks GLCM dengan parameter yang ditentukan. Kemudian, fungsi `greycoprops` digunakan untuk menghitung fitur kontras, homogenitas, energi, dan korelasi dari matriks GLCM, dan rata-rata dari setiap fitur GLCM tersebut dihitung.

Pada gambar 7 Fungsi `create_glm_dataset` adalah fungsi yang penting dalam skenario ini. Fungsi ini digunakan untuk membuat dataset yang berisi fitur GLCM dan rata-rata fitur RGB dari data-data dalam folder yang ditentukan. Fungsi ini melakukan iterasi pada setiap file dalam folder tersebut dan melakukan ekstraksi fitur RGB dan

GLCM menggunakan fungsi `extract_rgb_features` dan `extract_glm_features`. Hasil ekstraksi fitur kemudian ditambahkan ke dalam `DataFrame` yang berisi kolom-kolom yang sesuai. Fungsi ini mengembalikan `DataFrame` yang berisi dataset yang telah dibuat. Dengan menggunakan fungsi ini, kita dapat dengan mudah membuat dataset fitur GLCM dan rata-rata fitur RGB dari data-data gambar dalam folder dengan hanya memasukkan path folder yang sesuai sebagai parameter.

Pada bagian kode di Gambar 8, `folder_path` digunakan untuk menentukan lokasi folder tempat gambar-gambar disimpan. Kemudian, fungsi `create_glm_dataset` dipanggil dengan menggunakan `folder_path` sebagai parameter untuk melakukan ekstraksi fitur GLCM dan membuat dataset dari gambar-gambar dalam folder tersebut dan hasilnya disimpan dalam variabel dataset.



Gambar 8: Source code

Selanjutnya, dataset tersebut disimpan dalam format Excel dengan menggunakan fungsi `to_excel`, dengan nama file "glcm&rgbmean_features.xlsx". Dengan menjalankan kode tersebut, fitur GLCM dan rata-rata fitur RGB dari gambar-gambar dalam folder akan diekstraksi dan disimpan dalam file Excel untuk keperluan analisis lebih lanjut. Hasil yang

dipeoleh dari ekstraksi fitur RGB dan GLCM sebelumnya adalah dataset excel yang berisi data dengan fitur-fitur yang telah diekstraksi. Setiap baris mewakili satu data citra dengan proses selanjutnya yaitu tahap klasifikasi menggunakan Algoritma K-NN. Hasil dari ekstraksi fitur yang telah diolah menjadi dataset excel dapat dilihat pada Tabel 1.

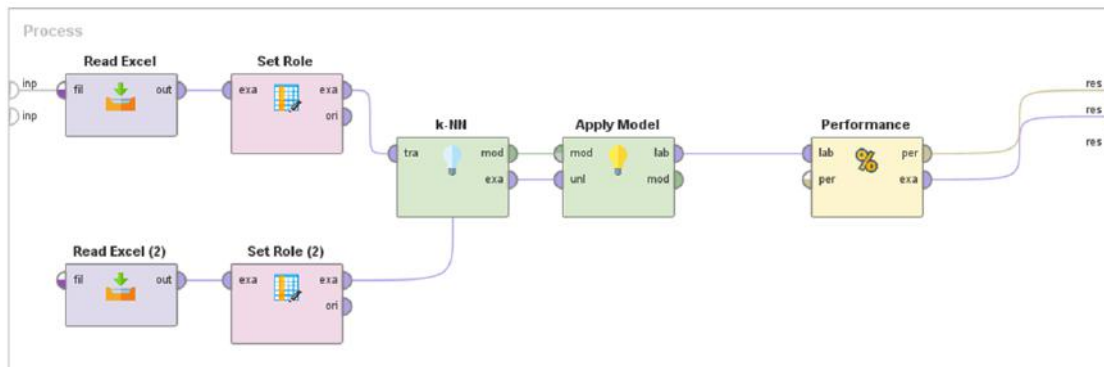
Tabel 1: Dataset citra kelapa sawit setelah ekstraksi fitur RGB dan GLCM

Image	Red mean	Green mean	Blue mean	Red std	Green std	Blue std	Contrast mean	Homogeneity mean	Energy mean	Correlation mean
tbs 1.jpg	107.819267	86.943812	86.373759	59.9122148	54.6867614	56.10915272	526.8242494	0.177215152	0.015329066	0.910462203
tbs 2.jpg	96.250316	76.134412	80.836195	55.7002667	48.4091004848	8.2298177	185.2697704	0.251280649	0.022379318	0.960616663
tbs 3.jpg	103.803132	84.725809	84.484456	58.2288624153	97.48806655	2.1143952	498.195145	0.163950723	0.015442852	0.912922457
tbs 4.jpg	109.018623	81.510823	87.936224	53.3901344650	59.16716950	7.6349175	377.9703368	0.16883796	0.014294586	0.920858818
tbs 5.JPG	113.178101	95.518091	96.994599	60.4379696361	8.618249262	0.06705234	794.3608303	0.149731938	0.013478771	0.890736097
tbs 6.JPG	99.734184	79.325829	76.712409	54.2498109154	0.0602047653	3.6229741	490.8721775	0.206815858	0.01894937	0.912337355
tbs 7.JPG	117.851948	90.850249	94.716904	63.0778239757	7.7062779657	7.71485738	607.848713	0.202522908	0.015138696	0.904872148
tbs 8.JPG	102.179688	83.93548	87.366993	48.9414240151	6.245270152	0.04914852	234.4983886	0.214846888	0.019172107	0.952985337
tbs 9.JPG	128.263608	93.738465	89.773933	79.56269722	75.7046488	73.99309753	777.4863047	0.205541086	0.023129489	0.93133926
tbs 10.JPG	99.517078	70.766131	62.00826	59.7210493957	8.8977295	54.7311867	546.758367	0.228163389	0.021497056	0.915692245
tbs 91.JPG	101.204925	99.321508	104.61061960	65.07915657	9.282960858	8.2029637	360.4877524	0.174772127	0.014854925	0.94718789
tbs 92.JPG	99.305251	97.394167	101.99091469	55.02020367	3.883766567	1.5422606	438.1883506	0.184704356	0.017089984	0.952138644
tbs 93.JPG	97.340326	95.067301	100.01706669	26.98826167	0.434336867	2.6399134	429.7314884	0.183477169	0.01738793	0.952704409
tbs 94.JPG	97.29727	96.919217	102.25592862	88.54409360	6.642362361	4.9961049	350.2745446	0.17777297	0.015347647	0.952916333
tbs 95.JPG	105.19991	102.867409	102.67146269	53.54626866	9.159859366	3.5538532	462.6808262	0.180223298	0.015650453	0.948940864
tbs 96.JPG	91.501545	87.551149	90.826102	66.8749162462	5.847792261	2.4651368	476.1641323	0.191736157	0.019303438	0.940842766
tbs 97.JPG	84.899833	81.475029	84.948085	62.1500660159	7.007671259	4.6846869	524.1898145	0.187604196	0.020102049	0.92752825
tbs 98.JPG	100.87552	98.439273	99.662296	62.8727351561	5.642225262	2.5781663	720.2967922	0.130970843	0.01250134	0.905389477
tbs 99.JPG	117.690791	116.286992	119.03357859	23.25795960	6.906818561	5.3335132	415.8927847	0.144147891	0.011834288	0.942083197
tbs 100.JPG	100.017669	101.520988	105.27674564	63.24805662	5.828814461	4.9121371	306.4829599	0.186700369	0.016595562	0.960914252

Klasifikasi K-NN

Dataset hasil ekstraksi yang berjumlah 100 data diberi label sesuai dengan nama data, kemudian dilakukan pembagian untuk memperoleh data training dan data testing dengan data training berjumlah 75 dan data testing berjumlah 25. Proses klasifikasi dilakukan dengan memanfaatkan software RapidMiner versi 10.1. Pada software RapidMiner

dilakukan pengimputan dataset yang telah dibagi menjadi data training dan data testing dengan menggunakan operator read excel. Dataset yang telah diinput kemudian ditambah operator set roll untuk memberi label atau kelas pada dataset yang telah diinput. Kemudian barulah dataset diproses dengan menggunakan Algoritma K-NN dengan nilai k=1,3,5,7,9. Desain proses klasifikasi K-NN pada software RapidMiner dapat dilihat pada Gambar 9.



Gambar 9: Desain proses klasifikasi K-NN pada RapidMiner

Berikut adalah penjelasan dan urutan proses klasifikasi yang akan dilakukan dalam RapidMiner sesuai dengan Gambar 9:

1. Operator Read Excel: Langkah pertama dalam proses klasifikasi adalah membaca dataset hasil ekstraksi fitur yang disimpan dalam format Excel menggunakan operator "Read Excel" dalam RapidMiner. Operator ini memungkinkan pengguna untuk mengimpor data dari file Excel ke dalam alur kerja RapidMiner.
2. Operator Set Role: Setelah dataset dibaca, operator "Set Role" digunakan untuk mengatur atribut target atau label klasifikasi dari dataset. Operator ini memungkinkan pengguna untuk menentukan atribut yang akan digunakan sebagai target klasifikasi dalam dataset.
3. Operator K-NN: Operator "K-NN" digunakan untuk menerapkan algoritma K-Nearest Neighbors pada dataset. Dalam konfigurasi operator K-NN, pengguna dapat mengatur parameter seperti jumlah tetangga terdekat (K) yang akan digunakan dalam klasifikasi.
4. Operator Apply Model: Setelah model K-NN terlatih, operator "Apply Model" digunakan untuk menerapkan model tersebut pada data yang belum diketahui atau data pengujian. Operator ini akan mengklasifikasikan setiap contoh data pengujian berdasarkan model K-NN yang telah dilatih sebelumnya.
5. Operator Performance: "Performance" digunakan untuk mengevaluasi performa model klasifikasi. Dalam operator ini, berbagai metrik evaluasi seperti akurasi, presisi, recall, dan lainnya dapat dihitung dan ditampilkan sebagai hasil evaluasi klasifikasi.

Dalam rangkaian proses klasifikasi ini, RapidMiner digunakan untuk membaca dataset, mengatur atribut target, melatih model K-NN, menerapkan model pada data pengujian, dan mengevaluasi performa klasifikasi. Proses ini membantu

para peneliti untuk menganalisis dan memahami tingkat kematangan buah kelapa sawit berdasarkan fitur-fitur yang diekstraksi sebelumnya. Dengan menggunakan RapidMiner, proses klasifikasi menjadi lebih terstruktur dan efisien, serta memberikan pemahaman yang lebih dalam tentang klasifikasi tingkat kematangan buah kelapa sawit berdasarkan dataset fitur yang telah diolah. Setelah dilakukan running pada desain proses diperoleh hasil klasifikasi serta performa model dari nilai $k=1,3,5,7,9$.

accuracy: 56.00%

	true masak	true mengkal	true mentah	class precision
pred. masak	0	0	0	0.00%
pred. mengkal	0	0	0	0.00%
pred. mentah	5	6	14	56.00%
class recall	0.00%	0.00%	100.00%	

Gambar 10: Hasil akurasi K-NN dengan nilai $k=9$

Gambar 10 menampilkan hasil akurasi dari metode K-NN dengan nilai $k=9$ yang menyatakan bahwa tingkat akurasi pada metode tersebut adalah 56.00%. Dimana class precision untuk pred. masak adalah 0.00%, pred. mengkal adalah 0.00%, dan pred. mentah adalah 56.00%.

accuracy: 56.00%

	true masak	true mengkal	true mentah	class precision
pred. masak	0	0	0	0.00%
pred. mengkal	0	0	0	0.00%
pred. mentah	5	6	14	56.00%
class recall	0.00%	0.00%	100.00%	

Gambar 11: Hasil akurasi K-NN dengan nilai $k=7$

Gambar 11 menampilkan hasil akurasi dari metode K-NN dengan nilai $k=7$ yang menyatakan bahwa tingkat akurasi pada metode tersebut adalah 56.00%. Dimana class precision untuk pred. masak adalah 0.00%, pred. mengkal adalah 0.00%, dan pred. mentah adalah 56.00%.

accuracy: 56.00%

	true masak	true mengkal	true mentah	class precision
pred. masak	0	0	0	0.00%
pred. mengkal	0	0	0	0.00%
pred. mentah	5	6	14	56.00%
class recall	0.00%	0.00%	100.00%	

Gambar 12: Hasil akurasi K-NN dengan nilai k=5

Gambar 12 menampilkan hasil akurasi dari metode K-NN dengan nilai k=5 yang menyatakan bahwa tingkat akurasi pada metode tersebut adalah 56.00%. Dimana class precision untuk pred. masak adalah 0.00%, pred. mengkal adalah 0.00%, dan pred. mentah adalah 56.00%.

accuracy: 56.00%

	true masak	true mengkal	true mentah	class precision
pred. masak	0	1	0	0.00%
pred. mengkal	0	0	0	0.00%
pred. mentah	5	5	14	58.33%
class recall	0.00%	0.00%	100.00%	

Gambar 13: Hasil akurasi K-NN dengan nilai k=3

Gambar 13 menampilkan hasil akurasi dari metode K-NN dengan nilai k=3 yang menyatakan bahwa tingkat akurasi pada metode tersebut adalah 56.00%. Dimana class precision untuk pred. masak adalah 0.00%, pred. mengkal adalah 0.00%, dan pred. mentah adalah 58.33%.

accuracy: 68.00%

	true masak	true mengkal	true mentah	class precision
pred. masak	2	3	0	40.00%
pred. mengkal	1	1	0	50.00%
pred. mentah	2	2	14	77.78%
class recall	40.00%	16.67%	100.00%	

Gambar 14: Hasil akurasi K-NN dengan nilai k=1

Gambar 14 menampilkan hasil akurasi dari metode K-NN dengan nilai k=1 yang menyatakan bahwa tingkat akurasi pada metode tersebut adalah 68.00%. Dimana class precision untuk pred. masak adalah 40.00%, pred. mengkal adalah 50.00%, dan pred. mentah adalah 77.78%.

Performa akurasi yang diperoleh dari hasil klasifikasi K-NN menggunakan nilai k=9,7,5,3 memiliki nilai yang sama yaitu 56.00%, sedangkan dengan menggunakan nilai k=1 diperoleh hasil akurasi senilai 68.00%. Performa klasifikasi menggunakan K-NN cenderung lebih baik ketika menggunakan nilai k yang lebih kecil. Dalam kasus ini, dengan nilai k=1, algoritma K-NN mampu memberikan akurasi yang lebih tinggi dibandingkan dengan nilai

k yang lebih besar. Namun demikian hasil akurasi yang diperoleh tergolong masih rendah. Dari hasil akurasi yang telah diperoleh dari klasifikasi menggunakan algoritma K-NN dengan dataset yang mencakup ekstraksi fitur RGB dan GLCM tersebut dapat diambil simpulan bahwa masih perlu adanya perbaikan atau evaluasi untuk memperoleh performa akurasi yang lebih baik.

Penutup

Berdasarkan penelitian yang telah dilakukan, dengan menyelesaikan tahapan-tahapan yang ada pada alur penelitian diperoleh hasil berupa dataset hasil ekstraksi fitur yang menyajikan susunan pola numerik dari fitur GLCM dan rata-rata fitur RGB yang diekstraksi dari data citra buah kelapa sawit. Hasil ekstraksi dapat digunakan dalam klasifikasi tingkat kematangan buah sawit yang telah diidentifikasi dengan tiga tingkat kematangan, yaitu masak, mengkal, dan mentah. Klasifikasi dilakukan dengan menggunakan algoritma k-nearest neighbors(K-NN) dan diperoleh pula hasil akurasi dari nilai k=9,7,5, dan 3 senilai 56.00% dan nilai k=1 senilai 68.00%. Dari hasil yang diperoleh tersebut, diketahui bahwa hasil penelitian ini belum cukup untuk dapat memberi keterangan bahwa penggunaan metode dan algoritma yang dipilih telah sesuai untuk diterapkan dalam penelitian lanjutan. Tingkat kematangan buah sawit yang menjadi objek penelitian ini sangatlah penting dalam penentu hasil panen sehingga masih diperlukan adanya perbaikan dan evaluasi untuk memperoleh performa akurasi yang lebih baik. Untuk meningkatkan performa akurasi, penggunaan dataset yang lebih besar dan representatif dapat dipertimbangkan. Evaluasi lebih lanjut dengan menggunakan metrik lain seperti precision, recall, dan F1-score dapat memberikan pandangan yang lebih komprehensif tentang performa model. Selain itu, eksplorasi penggunaan metode validasi yang berbeda juga menjadi langkah yang relevan, seperti validasi silang (cross-validation) untuk menghindari overfitting dan memastikan kehandalan model. Indikator kematangan buah yang berbeda juga bisa dieksplorasi lebih lanjut terutama pada ekstraksi fitur GLCM yang memanfaatkan tekstur. Patut disadari observasi pada penelitian ini, perhatian terutama difokuskan pada ekstraksi fitur berbasis warna dari citra buah kelapa sawit menggunakan komponen RGB dan mengabaikan ekstraksi fitur berbasis tekstur dari citra. Penggunaan variasi tekstur yang lebih kompleks dapat memberikan informasi yang lebih kaya dan akurat tentang tingkat kematangan buah kelapa sawit. Saran lainnya yang mungkin dapat ditambah adalah melakukan tuning parameter lebih mendalam pada algoritma K-NN, seperti menentukan nilai k yang optimal. Selain itu, penggunaan metode klasifikasi lainnya seperti Support Vector Machine (SVM), Decision Trees, atau

Neural Networks dapat dieksplorasi untuk membandingkan performa dan mencari metode yang paling sesuai dengan karakteristik data dan tujuan penelitian.

Ucapan Terimakasih

Ucapan terima kasih kepada pak Arifin selaku pemilik kebun kelapa sawit milik masyarakat yang menjadi objek penelitian dan mejadi tempat memperoleh informasi untuk mendukung penelian ini serta kepada Iwan Santosa, M.Si selaku pakar dalam proses validasi data citra buah kelapa sawit yang merupakan dosen program studi Agribisnis di kampus Universitas Darwan Ali. Ucapan terimakasih yang setulusnya pula disampaikan kepada seluruh pihak yang terlibat secara langsung maupun tidak langsung dalam penelitian ini sehingga penelitian ini dapat diselesaikan dengan sebagaimana semestinya.

Daftar Pustaka

- [1] D. J. Kusuma, "Pelaksanaan Panen Tandan Buah Segar (Tbs) Kelapa Sawit (Elais guineensis Jacq.)", Tugas Akhir Diploma, Jurusan Budidaya Tanaman Perkebunan > Prodi D3 Produksi Tanaman Perkebunan, Politeknik Negeri Lampung, 2021.
- [2] D. Julianto, "Pengaruh Temperatur Terhadap Proses Pemurnian Cpo Pada Crude Oil Tank(Cot) Di Stasiun Klarifikasi Di Pt.Multimas Nabati Asahan Kuala Tanjung", Tugas Akhir Diploma, Universitas Sumatera Utara, Medan, 2012.
- [3] M. Rifqi, M. Akbar dan Y. Fitriasia, "Aplikasi Pendeteksian Kematangan Tandan Buah Segar (TBS) Kelapa Sawit Berdasarkan Komposisi Warna Menggunakan Algoritma K-NN", Politek. Caltex Riau, vol. 6, no. 1, pp. 99–107, doi: <https://doi.org/10.35143/jkt.v6i1.3338>, 2020.
- [4] F. Kosasih, "Estimasi Biomassa Tanaman Nanas Berdasarkan Lebar Daun Menggunakan Algoritma Machine Learning", Skripsi, Fakultas Teknik > Prodi S1-Teknik Elektro, Universitas Lampung, 2023.
- [5] C. Paramita, E. Hari Rachmawanto, C. Atika Sari dan D. R. Ignatius Moses Setiadi, "Klasifikasi Jeruk Nipis Terhadap Tingkat Kematangan Buah Berdasarkan Fitur Warna Menggunakan K-Nearest Neighbor", J. Inform. J. Pengemb. IT, vol. 4, no. 1, pp. 1–6, doi: 10.30591/jpit.v4i1.1267, 2019.
- [6] D. Elvira, "Klasifikasi Citra Daun Kelapa Sawit Yang Terkena Dampak Hama Menggunakan Metode K-Nearest Neighbor (KNN)", Seminar Nasional Informatika Medis (SNIMed), Universitas Islam Riau, 2019.
- [7] D. Restiana and Nurahman, "Analisis Penerapan Metode Cosine Similarity Dalam Pengklasifikasian Surat Menyurat", Tugas Akhir, Sistem Informasi, Universitas Darwan Ali, 2022.
- [8] Putri DCN, Arna Fariza dan Jauari Akhmad Nur Hasim, "Mitigasi Urban Flood Berbasis Mobile Di Surabaya", Jurnal Informatika dan Komputer PENS, Vol. 2, No.2, 2013.
- [9] R. Widia, V. Novianti, Y. Syahidin dan M. Hidayati, "Sistem Informasi Korespondensi Rekam Medis di Rumah Sakit Menggunakan Microsoft Visual Studio", Expert J. Manaj. Sist. Inf. dan Teknol., vol. 11, no. 1, p. 56, doi: 10.36448/expert.v11i1.2013, 2021.
- [10] C. Rahmad, M. Astiningrum dan N. B. Purnomo, "Identifikasi Dan Prediksi Tingkat Kematangan Pisang Candi Dengan Fitur Warna Dan Tekstur Menggunakan Metode K-Nearest Neighbor Cahya", Semin. Inform. Apl. Polinema, pp. 188–193, 2019.
- [11] I. Amalia, "Ekstraksi Fitur Citra Songket Berdasarkan Tekstur Menggunakan Metode Gray Level Co-occurrence Matrix (GLCM)", J. Infomedia, vol. 3, no. 2, pp. 64–68, doi: 10.30811/jim.v3i2.715, 2018.
- [12] D. P. Andini, Y. B. G. Sugiarta dan E. P. S. Zaelani, "Pendeteksian dan Pengenalan Citra Wajah dengan Ekstraksi Fitur Menggunakan Filter Gabor", JTERA (Jurnal Teknol. Rekayasa), vol. 5, no. 2, p. 257, doi: 10.31544/jtera.v5.i2.2020.257-266, 2020.
- [13] A. E. Prasetyo, "Implementasi Scale Invariant Feature Transform- Histogram of Gradients dan Multiclass SVM Untuk Handtracking Pada Kontrol Mouse Pointer", Skripsi, FTIK_Teknik Informatika, Unikom, 2018,
- [14] N. A. Khairudin and Wahyu Saputro, "Klasifikasi Kualitas Mutu Buah Delima Dengan Menggunakan Ekstraksi Gray Level Co-Occurrence Matrix (GlcM) Dan K-Nearest Neighbor (Knn)", J. Inform. Teknol. dan Sains, vol. 4, no. 3, pp. 273–278, doi: 10.51401/jinteks.v4i3.1990, 2022.
- [15] F. A. D. Aji Prasetya Wibawa, Muhammad Guntur Aji Purnama dan Muhammad Fathony Akbar, "Metode-metode Klasifikasi", Pros. Semin. Ilmu Komput. dan Teknol. Inf., vol. 3, no. 1, p. 134, 2018.
- [16] M. P. Vaishnave, K. Suganya Devi, P. Srinivasan, and G. Arutperumjothi, "Detection and classification of groundnut leaf diseases using KNN classifier", IEEE Int. Conf. Syst. Comput. Autom. Netw., pp. 1–5, doi: 10.1109/IC-SCAN.2019.8878733, 2019.