

Klasifikasi Buah dan Sayur menggunakan fitur ekstraksi HOG dan Metode KNN

Firnanda Al-Islama Achyunda Putra ¹, Aditya Galih Sulaksono ², Listanto Tri Utomo³,
Ahmad Rizal Khamdani ⁴

^{1,2,3,4} Fakultas Teknologi Informasi, Sistem Informasi, Universitas Merdeka, Malang, Indonesia
Jalan Terusan Dieng No. 62-64 Klojen, Pisang Candi, Kec. Sukun, Kota Malang, Indonesia Email:
¹firnanda.putra@unmer.ac.id, ²khamdani.ahmad@student.unmer.ac.id

Abstrak

Abstrak – Perkembangan otomatisasi menggunakan algoritma saat ini sangat berkembang pesat. Banyak peneliti yang menggunakan metode *machine learning* untuk mempermudah aktivitas industry supaya lebih efisien. Buah merupakan kebutuhan pokok bagi masyarakat yang tentunya hampir setiap hari di konsumsi. Oleh karena banyak permintaan buah, maka banyak industry yang mengefisiensikan pemilihan kualitas dan jenis buah. Kebutuhan pokok selain buah yaitu sayuran. Sayuran dapat berasal dari kacang-kacangan, daun, atau biji-bijian yang dapat dimasak. Terdapat banyak variasi baik bentuk, warna, serta tekstur dari buah dan sayuran. Terkadang sulit untuk membedakan antara jenis yang memiliki kesamaan dalam hal ini. Oleh karena itu, dibutuhkan suatu sistem untuk membantu mengklasifikasikan buah dan sayuran dengan lebih mudah. Dalam penelitian ini, jenis buah dan sayuran diklasifikasikan berdasarkan hasil ekstraksi dari *Histogram Oriented of Gradient*. Metode yang digunakan dalam penelitian ini adalah *Histogram Oriented of Gradient* (HOG) dan *K-Nearest Neighbor* (K-NN). Proses HOG digunakan untuk ekstraksi fitur tersebut yaitu untuk mendapatkan karakteristik buah dan sayuran, sedangkan K-NN digunakan untuk proses klasifikasi citra. HOG banyak digunakan untuk ekstraksi ciri atau nilai dalam pemrosesan citra digital. Algoritma HOG merupakan algoritma *basic* yang mudah untuk di terapkan dalam berbagai aspek deteksi. Biasanya ekstraksi HOG digunakan untuk mendeteksi manusia, kendaraan maupun objek lainnya. Setelah nilai didapatkan dari nilai Histogram proses selanjutnya adalah mengklasifikasi menggunakan algoritma K-NN. Penelitian ini merupakan penelitian yang menggunakan HOG untuk Ekstraksi Fitur dan Algoritma K-NN untuk klasifikasi. Setiap nilai bobot citra latih dan citra uji akan dibandingkan dengan nilai Euclidean atau biasa disebut jumlah tetangga. Penelitian dengan metode ini memberikan hasil pengujian dengan tingkat akurasi 76,54% untuk buah-buahan, sedangkan hasil pengujian untuk sayuran memberikan nilai akurasi sebesar 71,22%.

Kata kunci : Klasifikasi Buah, HOG, K-NN.

1. Pendahuluan

Pemrosesan citra gambar (*image processing*) merujuk pada teknik dan metode yang digunakan untuk memanipulasi dan menganalisis gambar digital. Gambar digital adalah kumpulan piksel yang mewakili informasi visual. Tujuan dari pemrosesan citra gambar adalah untuk mengekstrak informasi yang berguna dari gambar, meningkatkan kualitas gambar, atau mempersiapkan gambar untuk analisis lebih lanjut.

Pemrosesan citra gambar memiliki banyak aplikasi praktis dalam bidang seperti kedokteran, astronomi, geografi, teknologi informasi, dan lain sebagainya. Dalam kedokteran, contohnya, pemrosesan citra gambar digunakan untuk mendiagnosa penyakit, seperti kanker dan penyakit jantung, dan untuk merencanakan operasi. Dalam astronomi, citra digital dari langit malam dapat diproses untuk mengidentifikasi dan mempelajari objek-objek seperti galaksi dan bintang.

Penelitian yang dilakukan oleh (Nayak et al., 2019) Proyek yang diusulkan mampu mengenali buah berdasarkan fitur seperti bentuk, warna, dan tekstur. Ini meningkatkan pengetahuan orang awam tentang beberapa langka dan buah yang tidak diketahui. Proyek ini terutama berkonsentrasi pada mengurangi usaha manusia dan membuat hidup manusia lebih mudah. Buah pengakuan akan dapat mengurangi arus yang sedang berlangsung masalah. Ini mengurangi kebingungan di antara buah tertentu.

Beberapa teknik pemrosesan citra gambar yang umum digunakan termasuk pengolahan filter, transformasi Fourier, segmentasi citra, deteksi tepi, pengenalan pola, dan rekayasa citra. Pemrosesan citra gambar dapat dilakukan menggunakan berbagai alat seperti software pemrosesan citra, library pemrosesan citra pada bahasa pemrograman tertentu, maupun menggunakan platform seperti Matlab dan Python (F. Liu et al., 2017).

Klasifikasi buah telah banyak dilakukan oleh peneliti sebelumnya seperti (X. Liu et al., 2019) menggunakan ekstraksi hanya dengan dari warna, tetapi hal tersebut dapat digunakan untuk mengklasifikasi sebuah buah. (Al-falluji, 2016) juga melakukan klasifikasi buah menggunakan edge detection atau biasa disebut deteksi tepi. Adapun klasifikasi yang saat ini sedang berkembang yaitu menggunakan CNN. Penelitian yang dilakukan oleh (Momeny et al., 2020) yaitu melakukan klasifikasi buah ceri dilakukan berdasarkan fisiknya penampilan. Sampel dataset ditingkatkan menggunakan metode augmentasi. Gambar berupa buah ceri diklasifikasikan dengan algoritma (CNN). Pada penelitian ini *hybrid pooling* digunakan untuk meningkatkan generalisasi dalam CNN yang dalam CNN dengan metode *hybrid pooling* dibandingkan dengan metode penyatuan lainnya. Fitur dari gambar buah ceri diekstraksi oleh HOG dan LBP dan diklasifikasikan dengan algoritma KNN, JST, Fuzzy dan EDT.

Adapun penelitian lain yang menggunakan CNN adalah (Vijayalakshmi & Peter, 2021) Dalam penelitian ini, peneliti mengklasifikasikan buah pisang. Model penelitian berbasis deep learning untuk otomatisasi identifikasi buah pisang telah dilakukan. Sistem identifikasi ini meliputi pra pemrosesan, fitur ekstraksi, dan klasifikasi. Selama preprocessing, gambar buah diubah ukurannya. Tahap ekstraksi ciri menggunakan algoritma CNN lima lapis untuk mengekstrak fitur gambar buah. Klasifikasi dilakukan dengan menggunakan RF algoritma. Hasil algoritma ini dibandingkan dengan algoritma K-NN. Adapun penelitian yang menggunakan CNN untuk mendeteksi sebuah lemon (Jahanbakhshi et al., 2020) klasifikasi lemon asam dilakukan berdasarkan pada ketidaksempurnaan mereka yang tampak. Kemudian, properti gambar yang diekstraksi diklasifikasikan dengan metode yang diusulkan (HOG dan LBP) dengan algoritma KNN, JST, Fuzzy, SVM dan DT. Akhirnya, gambar diklasifikasikan dengan CNN yang ditingkatkan dan hasilnya dibandingkan dengan hasil yang diperoleh melalui algoritma lain. Kumpulan data meningkat dengan metode augmentasi data. Dari pooling stokastik berdasarkan sparsity digunakan untuk meningkatkan hasil CNN. Convolutional Neural Network (CNN), dengan konfigurasi metode yang diusulkan, dapat mengklasifikasikan lemon asam dengan akurasi 100% di semua ukuran gambar. Klasifikasi yang dapat dilakukan oleh metode cnn juga dapat mendeteksi sebuah penyakit dari buah itu sendiri, Metode klasifikasi hibrida kami menggabungkan model DCNN dengan fitur PHOG untuk mengklasifikasikan buah menggunakan 5G dan teknologi cloud. Gambar diambil di perangkat pintar dan ditransfer ke cloud melalui jaringan 5G (Nasir et al., 2020).

Adapun penelitian yang dilakukan oleh (H. Abd al karim & A. Karim, 2021) yaitu tentang

klasifikasi buah. Selain itu, terdapat banyak kesulitan dalam klasifikasi buah. Dalam metode yang diusulkan, dan untuk mengatasi kesulitan tersebut, Tiga metode extractor \ deskriptor fitur telah digunakan, yaitu (LPB, HOG, dan GLCM) untuk mengekstraksi fitur dan menyimpannya dalam vektor fitur, selain itu classifier pohon keputusan digunakan untuk mengkategorikan buah.

Penelitian yang dilakukan oleh (Al-falluji, 2016) membahas terkait fitur ekstraksi HOG dapat di kombinasikan dengan algoritma ANN. Metode Histogram of Oriented Gradients (HOG) adalah salah satu teknik ekstraksi fitur pada pemrosesan citra untuk pengenalan objek. Teknik ini sangat populer digunakan pada aplikasi pengenalan wajah, kendaraan, dan lain sebagainya. Penelitian yang dilakukan oleh (Rismiyati & Wibawa, 2019; Yohannes et al., 2020) menggunakan ekstraksi ciri HOG dan menggabungkan dengan ELM yang memiliki hasil yang bagus. Cara kerja HOG adalah dengan menghitung distribusi orientasi gradien pada setiap piksel dalam citra. Gradien pada suatu piksel adalah perbedaan intensitas citra di sekitar piksel tersebut, yang menggambarkan arah perubahan intensitas yang signifikan pada piksel tersebut. Kemudian, distribusi orientasi gradien tersebut dihitung dalam blok-blok yang tumpang tindih pada citra, dan dihitung histogram untuk setiap blok. Histogram ini kemudian di-normalisasi untuk menghilangkan efek perbedaan pencahayaan pada citra.

Setelah histogram dihitung dan dinormalisasi, fitur HOG dapat dihasilkan dengan menggabungkan histogram dari blok-blok yang berdekatan pada citra. Fitur ini kemudian dapat digunakan untuk melatih model klasifikasi, seperti Support Vector Machine (SVM), untuk pengenalan objek.

Klasifikasi buah yang menggunakan ekstraksi fitur HOG juga dilakukan oleh (Naik & Patel, 2017). Kontur buah seperti warna, ukuran, bentuk, tekstur dan cacat yang berbeda merupakan atribut buah yang sangat penting untuk klasifikasi dan grading. Dengan kemajuan dalam pemrosesan citra dan ketersediaan perangkat keras dan perangkat lunak berbiaya rendah, pekerjaan klasifikasi dan penilaian buah secara manual telah diganti dengan sistem visi mesin otomatis. Alasan lain dari otomatisasi non-destruktif adalah kemampuannya untuk menghasilkan hasil yang akurat, cepat, objektif, dan efisien dibandingkan pekerjaan manual.

Penelitian yang dilakukan oleh (Behera et al., 2021) menggunakan LBP serta HOG untuk ekstraksi fitur. Pada penelitian ini penulis mencoba untuk mengklasifikasi antara papaya yang masih belum matang, sedikit matang serta masih belum matang. Yang menjadi ciri atau fitur adalah dari warna papaya itu sendiri. Klasifikasi status kematangan buah pepaya dengan dua pendekatan yaitu Machine Learning dan Transfer Learning Approach. Dalam pendekatan pembelajaran mesin LBP, fitur HOG dan

GLCM diklasifikasikan dengan KNN, SVM dan Naïve Bayes dengan fungsi kernel yang berbeda (total sembilan pengklasifikasi) dievaluasi.

Keuntungan dari metode HOG adalah robustness terhadap perbedaan pencahayaan pada citra dan mampu mengekstraksi fitur yang penting untuk pengenalan objek. Namun, teknik ini juga memiliki kekurangan, yaitu kemampuan ekstraksi fitur yang kurang baik pada citra yang sangat kompleks atau mengandung banyak perbedaan intensitas.

Penelitian yang dilakukan oleh (Wah & San, 2018) Klasifikasi K-Nearest Neighbor (KNN) adalah metode klasifikasi pada pembelajaran mesin (machine learning) yang sederhana dan mudah diimplementasikan. Metode KNN digunakan untuk melakukan klasifikasi suatu data ke dalam kelas-kelas yang sudah diketahui berdasarkan kumpulan data latih (training data) yang sudah diberi label kelas.

Cara kerja metode KNN adalah (*Fruits Recognition System Based on Colour , Shape , Principal Component and Region Features*, 2019) dengan menghitung jarak antara data uji (test data) dengan semua data latih yang sudah diberi label kelas. Data latih yang paling dekat dengan data uji akan dijadikan sebagai tetangga terdekat (nearest neighbors). Kemudian, kelas mayoritas dari tetangga terdekat tersebut akan dijadikan sebagai label kelas untuk data uji.

Parameter yang penting pada metode KNN adalah nilai K, yaitu jumlah tetangga terdekat yang akan diambil dalam proses klasifikasi. Nilai K yang terlalu kecil dapat menyebabkan model overfitting, sedangkan nilai K yang terlalu besar dapat menyebabkan model underfitting.

Keuntungan dari metode KNN adalah mudah diimplementasikan dan dapat digunakan pada data dengan jumlah dimensi yang besar (Rismiyati & Wibawa, 2019; Yohannes et al., 2020) Namun, metode ini juga memiliki kekurangan, yaitu sensitif terhadap data yang noise atau data yang memiliki nilai yang ekstrem. Selain itu, proses klasifikasi pada KNN juga cukup lambat pada kumpulan data yang besar karena harus menghitung jarak antara data uji dengan semua data latih. Klasifikasi buah yang menggunakan metode SVM dilakukan oleh (Ghazal et al., 2021; Kuang et al., 2016). Kedua penelitian tersebut menggunakan klasifikasi SVM untuk menentukan kelas buah. Python merupakan salah satu bahasa pemrograman yang sering digunakan untuk melakukan analisis data dan pembuatan model klasifikasi. Beberapa library yang sering digunakan pada Python untuk melakukan klasifikasi adalah Scikit-learn. Scikit-learn adalah library machine learning yang populer pada Python. Library ini menyediakan banyak algoritma klasifikasi, seperti *K-Nearest Neighbor*, *Decision Tree*, *Random Forest*, *Support Vector Machine* (SVM), dan lain-lain. *K-nearest Neighbor* adalah salah satu metode dari

scikit learn yang akan digunakan pada penelitian ini. Dengan menggunakan library dan langkah-langkah dasar tersebut, Python dapat digunakan untuk melakukan klasifikasi data dengan mudah dan efisien.

Berdasarkan data dari BPS (Badan Pusat Statistik), Indonesia memiliki produksi buah buahan yang sangat besar dan bertambah setiap tahunnya. Hal ini dikarenakan negara Indonesia memiliki tanah yang sangat baik dari kontur ataupun tingkat kesuburan tanah. Penelitian terkait pengklasifikasian buah dan sayuran dapat berkontribusi di bidang efisiensi pabrik yang akan melakukan impor ataupun ekspor buah buahan.

Penelitian ini dilakukan karena pengolahan citra atau gambar sangat banyak digunakan di bidang sains dan teknologi. Sebagai contoh, untuk mengklasifikasi jenis buah biasanya dilakukan secara manual, dengan adanya system ini, maka gambar akan di olah dan jika di kombinasikan dengan alat pemisah gambar maka akan sangat berguna untuk pabrik-pabrik atau penjual buah.

Metode dari penelitian ini adalah menggunakan metode algoritma KNN yang dikombinasikan dengan ekstraksi ciri HOG serta memadukannya dengan sekaligus menganalisis waktu komputasi dari metode yang di usulkan. Ekstraksi ciri HOG merupakan ekstraksi yang pada umumnya digunakan untuk pendeteksian objek. Yang membedakan model yang diusulkan dari penelitian yang telah dilakukan adalah kami penulis sekaligus menganalisa hasil waktu komputasi.

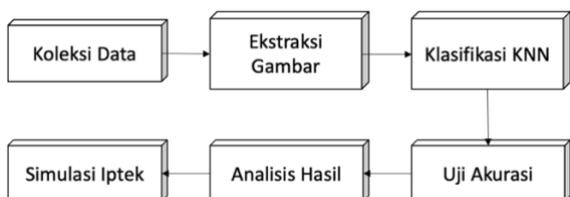
Penelitian ini dilakukan dengan menggunakan metode KNN. Terdapat beberapa model metode yang telah dilakukan oleh penelitian terdahulu seperti metode KNN yang dikombinasikan dengan ekstraksi ciri HOG. Tetapi pada penelitian tersebut tidak menampilkan atau menganalisis waktu komputasi. Penelitian ini dilakukan untuk membahas gap terkait belum adanya analisis waktu komputasi. Penelitian ini berpedoman pada penelitian penelitian terdahulu yang telah dilakukan uji akurasi dan juga dilakukan perbandingan antara salah satu metode dengan metode yang lain.

2. Metode

Kegiatan ini dilakukan dengan berbagai tahapan yang perlu dilakukan secara urut, dikarenakan apabila pekerjaan satu tidak selesai maka tidak akan dapat melanjutkan ke tahapan berikutnya. Penelitian dimulai dari koleksi data, data didapatkan secara online. Peneliti tidak melakukan pengambilan data secara manual. Dikarenakan data untuk dapat dilakukan pengujian harus cukup untuk mendapatkan hasil yang maksimal. Setelah data sudah di dapatkan penulis melakukan proses atau tahapan selanjutnya yaitu ekstraksi gambar. Ekstraksi gambar dilakukan dengan algoritma HOG. Setelah gambar sudah di ekstraksi yang mana akan menghasilkan nilai fitur dari HOG yaitu sebanyak

3700 nilai fitur. Setelah mendapatkan fitur dilanjutkan dengan proses klasifikasi KNN. Metode knn dipilih berdasarkan ekstraksi fitur yang paling sering digunakan untuk mendeteksi atau mengklasifikasi objek. Proses selanjutnya adalah melakukan pengujian akurasi, setelah akurasi didapatkan maka akan dilakukan analisis terhadap hasil.

Gambar 1 menjelaskan bagaimana tahapan penelitian ini dilakukan. Akhir proses dari penelitian ini adalah sebuah analisis hasil yang didapatkan dari ekstraksi ciri HOG serta algoritma klasifikasi KNN.

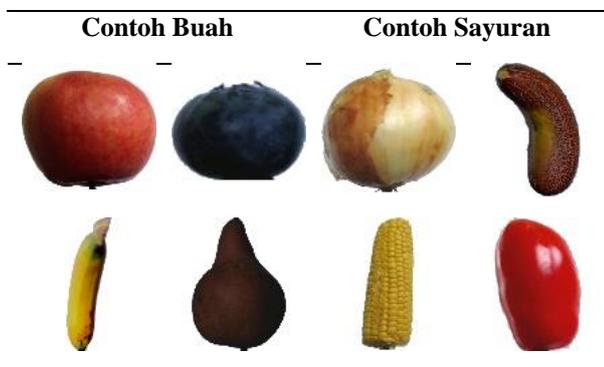


Gambar 1. Block Sistem penelitian

2.1 Koleksi Data

Data diambil dari situs yang berada di internet yaitu <https://www.kaggle.com/code/databeru/fruit-and-vegetable-classification/>. Kaggle merupakan bank dataset yang sangat umum digunakan oleh peneliti. Umumnya peneliti mengambil dataset untuk mencoba atau membuat metode baru. Dataset yang berada di situs tersebut tidak hanya buah dan sayuran saja, tetapi terdapat data yang memiliki karakteristik bermacam macam. Total keseluruhan dataset yang digunakan yaitu sebanyak 90.380 data berupa gambar. Yang mana terdiri dari 131 buah buahan dan sayuran. Tabel 1 adalah contoh gambar dari buah buahan serta sayuran.

Tabel 1. Contoh Data Buah dan Sayuran



2.2 Ekstraksi Gambar menggunakan Histogram Oriented of Gradient

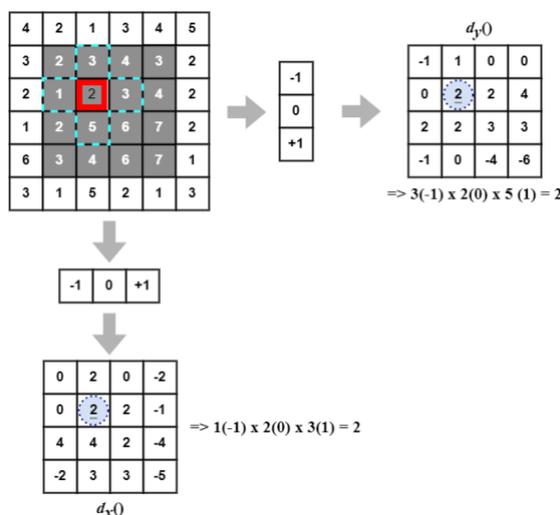
Metode penyelesaian masalah, serta tahapan tahapan dari metode HOG. Dalam naskah, nomor kutipan secara berurutan dalam tanda kurung siku [3], juga tabel angka dan angka secara berurutan

seperti yang ditunjukkan pada tabel 1 dan gambar Ekstraksi Gambar yang digunakan adalah menggunakan metode Histogram oriented of Gradient. Metode ini sering digunakan untuk mendeteksi sebuah objek.

Terdapat beberapa tahapan penting dalam ekstraksi ciri HOG. Yang pertama yaitu dengan mengambil gambar masukan yang diinginkan untuk menghitung fitur HOG. Gambar inputan pada penelitian ini adalah sebuah gambar berupa citra. Kemudian mengubah ukuran gambar dengan rasio 1:2 (lebar:tinggi) yakni 64 x 128 piksel. Ukuran gambar sebaiknya 64 x 128 piksel untuk membagi gambar menjadi 8 x 8 dan 16 x 16 saat mengekstrak fitur.

Proses selanjutnya adalah Menghitung orientasi (arah x dan y) pada gambar. Proses menghitung orientasi ada dua cara yakni menggunakan operator sobel atau *centered derivatice* (turunan terpusat).

Gambar 2 menjelaskan analisis visual dari ekstraksi ciri HOG. Perhitungan tersebut untuk memperoleh nilai gradien X dan nilai gradien Y.

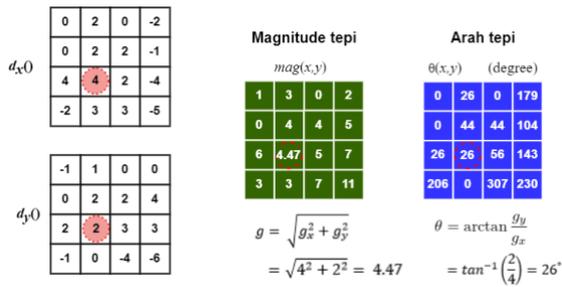


Gambar 2. Penghitungan orientasi arah X dan Y

Langkah selanjutnya yaitu Menghitung magnitude dan gradien, berikut cara perhitungan nilai magnitude dan gradien. Rumus dari cara menghitung gradien adalah

$$Magnitude(\mu) = \sqrt{G_x^2 + G_y^2} \quad Angle(\theta) = |\tan^{-1}(G_y/G_x)| \quad (1)$$

Dari hasil komputasi perhitungan berdasarkan orientasinya, maka selanjutnya *menghitung* nilai magnitude tepi dan arah tepi gradien. Gambar 3 merupakan visual analisis dari perhitungan nilai magnitude tepi.



Gambar 3. Menghitung nilai magnitude tepi

Langkah selanjutnya yaitu Membuat histogram menggunakan gradien dan orientasi. Untuk histogram, pastikan untuk membaginya menjadi 9 bin terpisah, masing-masing sesuai dengan sudut dari 0–160 derajat dengan kelipatan 20. Jika nilai pada tiap cell dari proses arah tepi melebihi atau sama dengan nilai 180 maka dilakukan proses polarity change. Polarity change membuat deskriptor menjadi invarian terhadap perubahan relatif antara foreground dan background. Pada deskriptor HOG juga perlu dipastikan gambar dibagi ke dalam cell sehingga histogram gradien dapat dihitung untuk setiap cell-nya. Sebagai contoh, jika gambar sebelumnya memiliki ukuran 64 x 128 piksel, maka gambar menjadi 8 x 8 cell.

Langkah selanjutnya adalah Menghitung HOG pada 8 x 8 cell (9 x 1). Pada langkah sebelumnya telah dijelaskan bahwa histogram yang dibuat di deskriptor fitur HOG membagi gambar menjadi 8 x 8 cell, dan histogram gradien berorientasi dihitung untuk setiap pikselnya. Hal ini bertujuan agar mendapatkan fitur (atau histogram) untuk patch yang lebih kecil mewakili keseluruhan gambar. Pembagian cell dapat menggunakan 8 x 8 menjadi 16 x 16 atau 32 x 32. Jika kita membagi gambar menjadi 8x8 sel dan menghasilkan histogram, kita akan mendapatkan matriks berukuran 9 x 1 untuk setiap sel. Matriks ini dihasilkan dengan salah satunya menggunakan metode ke-4 yang telah di bahas. Hasil yang didapatkan dari ekstraksi fitur HOG adalah Kita akan memiliki 105 (7x15) blok berukuran 16x16. Masing-masing dari 105 blok ini memiliki vektor 36x1 sebagai fitur. Oleh karena itu, fitur total untuk gambar tersebut adalah 105 x 36x1 = 3780 fitur.

2.3 Klasifikasi K-nearest Neighbor

Klasifikasi K-NN merupakan salah satu klasifikasi yang sangat umum digunakan untuk mengklasifikasi sebuah objek. Telah banyak peneliti sebelumnya yang meneliti terkait deteksi objek menggunakan metode ini. Klasifikasi K-NN beroperasi dengan mengelompokkan suatu objek berdasarkan kesamaannya dengan objek-objek lainnya. Tahap awal dari klasifikasi K-NN adalah menemukan nilai k yang optimal, kemudian menentukan cara menghitung jarak antara objek-objek tersebut. Setelah itu, hasil perhitungan jarak

diurutkan dan digunakan untuk mengklasifikasikan objek menggunakan algoritma K-NN.

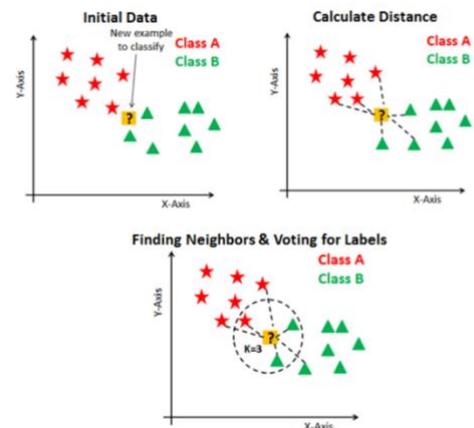
Formula *Euclidean distance* pada 1-dimensional space adalah sebagai berikut.

$$dis(x_1, x_2) = \sqrt{\sum_{i=0}^n (x_{1i} - x_{2i})^2} \tag{2}$$

Jika ada lebih dari satu, kita dapat menjumlahkannya seperti di bawah ini.

$$dis = \sqrt{\sum_{i=0}^n (x_{1i} - x_{2i})^2 + (y_{1i} - y_{2i})^2 + \dots} \tag{3}$$

Gambar 4 menjelaskan terkait bagaimana sebaran dari sebaran data menggunakan algoritma KNN. Terdapat beberapa tahapan untuk mendapatkan hasil klasifikasi algoritma tetangga yang dapat di hitung. K = jumlah tetangga terdekat. Pertama Hitung jarak antara data yang akan dievaluasi dengan semua data pelatihan. Selanjutnya Urutkan jarak yang terbentuk (urut naik) dan tentukan jarak terdekat sampai urutan ke-K. setelah itu Pasangkan kelas (C) yang bersesuaian. Dan yang terakhir adalah Cari jumlah kelas terbanyak dari tetangga terdekat tersebut, dan tetapkan kelas tersebut.



Gambar 4. Sebaran data oleh algoritma KNN

Lebih detail nya metode K-NN wajib menentukan nilai tetangga yaitu nilai K yang terbaik. Nilai K semakin besar belum tentu memiliki hasil yang lebih baik.

2.4 Uji Akurasi

Uji Akurasi digunakan untuk mengetahui bagaimana akurasi yang didapatkan oleh algoritma *Histogram Oriented of Gradient* dengan dikombinasikan menggunakan algoritma klasifikasi *K-Nearest Neighbor*. Uji akurasi didapatkan dari perbandingan kondisi actual yang dibandingkan dengan hasil sistem yang didapatkan.

2.5 Analisis Hasil

Hasil yang di dapatkan akan dianalisis secara objektif. Baik hasil akurasi yang bagus ataupun hasil akurasi yang kurang bagus. Akurasi pengenalan objek dengan menggunakan KNN dan HOG tergantung pada kualitas dataset yang digunakan, serta parameter dan konfigurasi yang dipilih untuk setiap teknik. Oleh karena itu, penting untuk melakukan eksperimen dan penyesuaian yang tepat untuk mendapatkan hasil yang optimal. Penulis menggunakan dataset yang sangat banyak sehingga dengan harapan banyaknya dataset akan memperbanyak training data serta harapannya akan memiliki nilai akurasi yang bagus.

2.6 Kegunaan

Kegunaan dari penelitian yang diusulkan adalah untuk bahan referensi oleh sebuah perusahaan yang akan menerapkan pengklasifikasian buah menggunakan metode yang berbasis citra. Banyak hal yang dapat dianalisis dari pemrosesan citra selain klasifikasi. Fokus pada penelitian ini adalah untuk mempermudah bagi perusahaan yang ingin menerapkan pemrosesan citra untuk membantu dan mempercepat perusahaan dalam memilah buah. jika perusahaan memiliki produk buah yang sangat banyak dan bermacam macam, maka riset ini sangat penting untuk dikembangkan. Yang sebelumnya perusahaan akan memilah secara manual, dengan adanya sistem otomatisasi klasifikasi maka perusahaan akan lebih mudah dan murah dalam mengklasifikasi buah secara otomatis.

3. Hasil dan Pembahasan

Pada sub bab ini penulis akan menjelaskan terkait akurasi dari hasil analisis algoritma HOG di kombinasikan dengan algoritma KNN. Seperti yang dijelaskan pada pendahuluan algoritma HOG memang banyak digunakan oleh peneliti sebelumnya untuk mendeteksi sebuah objek. Pada penelitian ini penulis mencoba untuk menggunakan algoritma HOG yang biasanya digunakan untuk mendeteksi objek, kini digunakan untuk mengklasifikasi sebuah objek.

Untuk memperkaya dari penelitian ini penulis menganalisis penggunaan nilai K yaitu 3, 5 dan 7. Pada gambar 4 telah dicontohkan untuk gambaran dari model atau nilai K yang digunakan. Penulis menganalisis hasil akurasi untuk nilai K 3, 5 dan 7. Tabel 2 menjelaskan hasil dari nilai K yang digunakan terhadap metode yang diusulkan.

Tabel 2. Hasil Pengujian Akurasi Klasifikasi

Data	K=3	K=5	K=7
Buah Buahan	77.32 %	75.7 %	76.6 %
Sayuran	72.4 %	70.19 %	71.09 %

Algoritma klasifikasi KNN juga telah banyak digunakan oleh para peneliti sebelumnya untuk mengklasifikasi sebuah objek. Nilai K yang digunakan pada penelitian ini adalah 3, 5 dan 7. Angka tersebut digunakan melihat dari referensi (Ginantra, 2016; Isman et al., 2021; Ula et al., 2022) yang juga menggunakan nilai K yang sama.

Hasil evaluasi dari model terdiri dari beberapa dua komponen utama yaitu menggunakan nilai K yang berbeda model menghasilkan nilai akurasi yang berbeda juga. Pada Tabel 2 nilai K yang terbaik adalah nilai K=3 untuk objek adalah buah buahan. Sedangkan pada objek berupa sayuran menghasilkan K=3 juga yang memiliki nilai akurasi yang terbaik.

Hal ini dikarenakan nilai K yang Nilai optimal untuk parameter k dalam algoritma K-Nearest Neighbors (KNN) sangat bergantung pada karakteristik data yang digunakan. Secara umum, ketika nilai k ditingkatkan, dampak dari gangguan atau noise dalam klasifikasi akan berkurang, tetapi pada saat yang sama, batas antara berbagai klasifikasi menjadi lebih tidak jelas atau kabur. Pemilihan nilai k yang baik bisa dilakukan melalui proses optimasi parameter, seperti penggunaan metode cross-validation.

Data yang kita gunakan terdiri dari data latih dan data uji. Data latih adalah data yang digunakan untuk membuat sebuah model dari data. Sedangkan data uji biasanya disebut dengan sebutan data testing adalah data yang digunakan untuk menguji algoritma yang digunakan. Proporsi ideal data uji dan data latih yang umum digunakan oleh peneliti sebelumnya adalah 80% untuk data Latih dan 20% untuk data Uji. Pemisahan data latih dan data uji dengan rasio 80:20 adalah salah satu praktek yang sering diterapkan dalam pemodelan statistik dan pembelajaran mesin. Keputusan ini didasarkan pada upaya mencapai keselarasan antara jumlah data yang digunakan untuk pelatihan model (data latih) dan data yang digunakan untuk menguji kinerja model (data uji).

Penggunaan perbandingan data latih tersebut adalah sebagai Upaya untuk menghindari overfitting. Misalnya data yang dibagi, atau rasio yang digunakan adalah 95% data latih dan 5% data uji maka resiko overfitting akan lebih tinggi.

Tabel 3. Hasil Pengujian Akurasi Klasifikasi dan waktu komputasi rata rata

Data	Waktu Komputasi (detik)	Akurasi
Buah Buahan	20:15:43	76,54%
Sayuran	13:23:21	71,22%

Manfaat dari dilakukannya penelitian ini untuk mendukung teknologi yang terkini. Telah banyak teknologi yang digunakan di negara maju untuk mengklasifikasi sebuah objek. Pada penelitian ini peneliti berkontribusi dalam mencari nilai akurasi antara metode ekstraksi ciri HOG dengan klasifikasi KNN.

4. Kesimpulan

Kesimpulan dari penelitian yang telah dilakukan oleh penulis adalah algoritma ekstraksi ciri Histogram oriented of Gradient dapat digunakan untuk mengklasifikasi buah. Jumlah dataset yang digunakan pada penelitian ini adalah 90.380 data berupa gambar. Yang mana terdiri dari 131 buah buahan dan sayuran. Data diambil dari sumberdata terkenal yaitu kaggle.com. Hasil akurasi menunjukkan bahwa HOG cocok untuk ekstraksi ciri buah. Akurasi yang didapatkan adalah 76.54% sedangkan untuk akurasi yang didapatkan oleh algoritma yang penulis usulkan untuk mengklasifikasi sebuah sayuran memiliki akurasi sebanyak 71,22%. Penelitian yang dilakukan pada penelitian ini hanya berupa buah dan sayuran. Dalam riset selanjutnya harapannya ada riset yang dapat memilah atau mengklasifikasi kualitas buah dan sayuran yang memiliki grade A grade B ataupun grade C. Jika memungkinkan tidak hanya buah dan sayuran saja yang perlu di klasifikasi tetapi bisa juga berupa kualitas makanan dan tingkat fresh atau tidaknya dari buah dan sayuran tersebut.

Daftar Pustaka:

Al-falluji, R. A. A. (2016). Color , Shape and Texture based Fruit Recognition System. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 5(7), 2108–2112.

Behera, S. K., Rath, A. K., & Sethy, P. K. (2021). Maturity status classification of papaya fruits based on machine learning and transfer learning approach. *Information Processing in Agriculture*, 8(2), 244–250. <https://doi.org/10.1016/j.inpa.2020.05.003>

Fruits Recognition System Based on Colour , Shape , Principal Component and Region Features. (2019). 6(2), 226–231.

Ghazal, S., Qureshi, W. S., Khan, U. S., Iqbal, J., Rashid, N., & Tiwana, M. I. (2021). Analysis of visual features and classifiers for Fruit classification problem. *Computers and Electronics in Agriculture*, 187(April), 106267. <https://doi.org/10.1016/j.compag.2021.106267>

Ginatra, N. L. W. S. R. (2016). Deteksi Batik Parang Menggunakan Fitur Co-Occurrence Matrix Dan Geometric Moment Invariant Dengan Klasifikasi KNN. *Lontar Komputer : Jurnal Ilmiah Teknologi Informasi*, 7(1), 40. <https://doi.org/10.24843/lkjiti.2016.v07.i01.p05>

H. Abd al karim, M., & A. Karim, A. (2021). Using Texture Feature in Fruit Classification. *Engineering and Technology Journal*, 39(1B), 67–79. <https://doi.org/10.30684/etj.v39i1b.1741>

Isman, Andani Ahmad, & Abdul Latief. (2021). Perbandingan Metode KNN Dan LBPH Pada Klasifikasi Daun Herbal. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(3), 557–564. <https://doi.org/10.29207/resti.v5i3.3006>

Jahanbakhshi, A., Momeny, M., Mahmoudi, M., & Zhang, Y. D. (2020). Classification of sour lemons based on apparent defects using stochastic pooling mechanism in deep convolutional neural networks. *Scientia Horticulturae*, 263(December 2019), 109133. <https://doi.org/10.1016/j.scienta.2019.109133>

Kuang, H., Hang Chan, L. L., Liu, C., & Yan, H. (2016). Fruit classification based on weighted score-level feature fusion. *Journal of Electronic Imaging*, 25(1), 013009. <https://doi.org/10.1117/1.jei.25.1.013009>

Liu, F., Snetkov, L., & Lima, D. (2017). *Summary on fruit identification methods: A literature review*. 119(ESSAEME), 1629–1633. <https://doi.org/10.2991/essaeme-17.2017.338>

Liu, X., Zhao, D., Jia, W., Ji, W., & Sun, Y. (2019). A Detection Method for Apple Fruits Based on Color and Shape Features. *IEEE Access*, 7, 67923–67933. <https://doi.org/10.1109/ACCESS.2019.2918313>

Momeny, M., Jahanbakhshi, A., Jafarnejhad, K., & Zhang, Y. D. (2020). Accurate classification of cherry fruit using deep CNN based on hybrid pooling approach. *Postharvest Biology and Technology*, 166(December 2019), 111204. <https://doi.org/10.1016/j.postharvbio.2020.111204>

Naik, S., & Patel, B. (2017). Machine Vision based Fruit Classification and Grading - A Review. *International Journal of Computer Applications*, 170(9), 22–34. <https://doi.org/10.5120/ijca2017914937>

- Nasir, I. M., Bibi, A., Shah, J. H., Khan, M. A., Sharif, M., Iqbal, K., Nam, Y., & Kadry, S. (2020). Deep learning-based classification of fruit diseases: An application for precision agriculture. *Computers, Materials and Continua*, 66(2), 1949–1962. <https://doi.org/10.32604/cmc.2020.012945>
- Nayak, A., Manjesh, R., & Dhanusha, M. (2019). Fruit Recognition using Image Processing. *International Journal of Engineering Research & Technology (IJERT)*, 7(08), 1–6. www.ijert.org
- Rismiyati, & Wibawa, H. A. (2019). Snake Fruit Classification by Using Histogram of Oriented Gradient Feature and Extreme Learning Machine. *ICICOS 2019 - 3rd International Conference on Informatics and Computational Sciences: Accelerating Informatics and Computational Research for Smarter Society in The Era of Industry 4.0, Proceedings*, 2–6. <https://doi.org/10.1109/ICICoS48119.2019.8982528>
- Ula, M., Zulhusna, R., Putra Fhonna, R., & Pratama, A. (2022). Penerapan Model Klasifikasi K-Nearest Neighbor Dalam Pencarian Kesesuaian Pekerjaan. *Metik Jurnal*, 6(1), 18–23. <https://doi.org/10.47002/metik.v6i1.343>
- Vijayalakshmi, M., & Peter, V. J. (2021). CNN based approach for identifying banana species from fruits. *International Journal of Information Technology (Singapore)*, 13(1), 27–32. <https://doi.org/10.1007/s41870-020-00554-1>
- Wah, T. N., & San, P. E. (2018). Performance Comparison of Rice Detection based on kNN and ANN Techniques. *The 9th International Conference on Science and Engineering, December 2018*, 1–5. https://www.researchgate.net/publication/349683438_Performance_Comparison_of_Rice_Detection_based_on_kNN_and_ANN_Techniques
- Yohannes, Y., Pribadi, M. R., & Chandra, L. (2020). Klasifikasi Jenis Buah dan Sayuran Menggunakan SVM Dengan Fitur Saliency-HOG dan Color Moments. *Elkha*, 12(2), 125. <https://doi.org/10.26418/elkha.v12i2.42160>