

PEMBUATAN DATASET BIJI JAGUNG: KLASIFIKASI MENGGUNAKAN K-NEAREST NEIGHBOR

¹Bima aditiansyah, ²T Muhammad Imanullah, ³Dedy Abdullah, ⁴A.R. Walad Mahfuzhi

¹Universitas Muhammadiyah Bengkulu, Indonesia

bimaaditiansyah@gmail.com; muhammad.imanullah@umb.ac.id; dedyabdullah@umb.ac.id; walad@umb.ac.id

Article Info

Article history:

Received, 2025-04-02

Revised, 2025-04-15

Accepted, 2025-05-28

Kata Kunci:

Klasifikasi;
dataset;
k-nearest neighbor;
biji jagung;

ABSTRAK

jagung adalah salah satu tanaman pangan penghasil karbohidrat yang penting di Indonesia, selain padi dan gandum. Jenis jagung biasa dilihat dari bentuknya. Mengklasifikasikan jenis jagung bisa dikatakan cukup sulit karena bentuk biji jagung sangatlah beragam, saat ini kita memasuki zaman modern, dimana zaman yang sudah maju ini klasifikasi jenis jagung bisa menggunakan pengolahan citra digital. Salah satu cara untuk meningkatkan akurasi dalam proses grading benih jagung adalah dengan menggunakan teknologi pengolahan citra digital. Dalam citra digital klasifikasi bentuk sebuah objek bisa menggunakan beragam metode salah satunya seperti K-Nearest Neighbors, Ada beberapa tahapan dalam penelitian ini seperti pembentukan dataset, preprocessing, penerapan algoritma k-nearest neighbor, pengujian, analisis akhir. berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma k-nearest neighbor dalam pengolahan citra digital mampu terbukti efektif dalam mengklasifikasi biji jagung. pada tahap latih dengan nilai K=1 mendapatkan akurasi 100% dan pada tahapan pertama tingkat akurasi yang didapat sebesar 78%, tahapan kedua mendapatkan akurasi sebesar 80,3% dan tahapan terakhir mendapatkan akurasi sebesar 76,5% dengan dijumlahkan ketiganya dan di bagi tiga mendapatkan hasil sebesar 78,26%. meskipun demikian, masih terdapat beberapa keterbatasan dalam penelitian ini, seperti pengaruh pencahayaan, resolusi gambar, dan kondisi fisik biji jagung yang dapat mempengaruhi akurasi deteksi akan lebih bagus jika menggunakan kamera professional. oleh karena itu, pengembangan lebih lanjut diperlukan, seperti integrasi dengan machine learning atau kecerdasan buatan (AI) untuk meningkatkan efisiensi dan akurasi sistem klasifikasi biji jagung dimasa mendatang.

Keywords:

Classification;
Dataset;
k-nearest neighbor;
corn kernels

ABSTRACT

Corn is one of the important carbohydrate-producing food crops in Indonesia, besides rice and wheat. Corn types are usually seen from their shape. Classifying corn types can be said to be quite difficult because the shape of corn kernels is very diverse. Currently we are entering the modern era, where this advanced era, corn type classification can use digital image processing. One way to improve the accuracy in the corn seed grading process is by using digital image processing technology. In digital images, the classification of an object's shape can use various methods, one of which is K-Nearest Neighbors. There are several stages in this research such as dataset formation, preprocessing, Application of the k-nearest neighbor algorithm, testing, final analysis. Based on the research that has been done, it can be concluded that the k-nearest neighbor algorithm in digital image processing is proven to be effective in classifying corn kernels. In the training stage with a value of K = 1, it gets 100% accuracy and in the first stage the accuracy level obtained is 78%, the second stage gets an accuracy of 80.3% and the final stage gets an accuracy of 76.5% by adding the three and dividing by three to get a result of 78.26%. However, there are still several limitations in this study, such as the influence of lighting, image resolution, and the physical condition of corn kernels that can affect detection accuracy. It would be better if using a professional camera. Therefore, further development is needed, such as integration with machine learning or artificial intelligence (AI) to improve the efficiency and accuracy of the corn kernel classification system in the future.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Penulis Korespondensi:

Bima aditiansyah,
Program Studi Teknik Informatika,
Universitas Muhammadiyah Bengkulu,
Email: bimaaditiansyah@gmail.com

1. PENDAHULUAN

Jagung merupakan salah satu tanaman pangan penghasil karbohidrat yang penting di Indonesia, selain padi dan gandum. Kebutuhan jagung saat ini mengalami peningkatan dapat dilihat dari data produksi di tingkat nasional yang dari tahun ke tahun selalu meningkat (Rosiani, Mentari, and Prastyia n.d.) Jagung dapat diolah sebagai bahan konsumsi manusia maupun hewan. tanaman jagung yang banyak di tanam oleh masyarakat di Indonesia adalah tipe mutiara, serta terdapat juga jagung yang bertipe brondong, jagung tipe tepung, jagung gigi kuda, dan jagung manis (Lutfi 2019). adapun beberapa contoh jenis jagung, seperti jagung hibrida, jagung sorgum, jagung bisi, jagung elos, dan masih banyak lainnya. jenis jagung bisa diketahui dari bentuk bijinya. ada biji yang berukuran besar bulat, ada yang berukuran kecil pipih, ada yang berukuran sedang sedikit lonjong, ada yang berbentuk bulat, ada yang memiliki lekukan pada bagian tengah ataupun pada bagian atas biji, dan lain sebagainya.

Mengklasifikasi jenis jagung bisa dikatakan cukup sulit karena dari bentuk biji jagung sangatla beragam bagi orang yang berpengalaman di bidang pertanian jagung, membedakan jenis jagung bukanlah menjadi kendala yang berat. namun bagi masyarakat yang awam membedakan jenis jagung merupakan kendala tersendiri. saat ini kita sudah memasuki zaman Industri, dimana zaman yang sudah maju ini klasifikasi jenis jagung bisa menggunakan pengolahan citra digital (Maulidiansyah and Abdillah 2023). Salah satu cara untuk meningkatkan akurasi dalam proses grading benih jagung adalah dengan menggunakan teknologi pengolahan citra digital. pengolahan citra digital telah menjadi alat yang efektif dalam menganalisis dan memanipulasi citra. dalam konteks ini, pengolahan citra digital dapat digunakan untuk mengklasifikasikan benih jagung (Zahara et al. 2024).

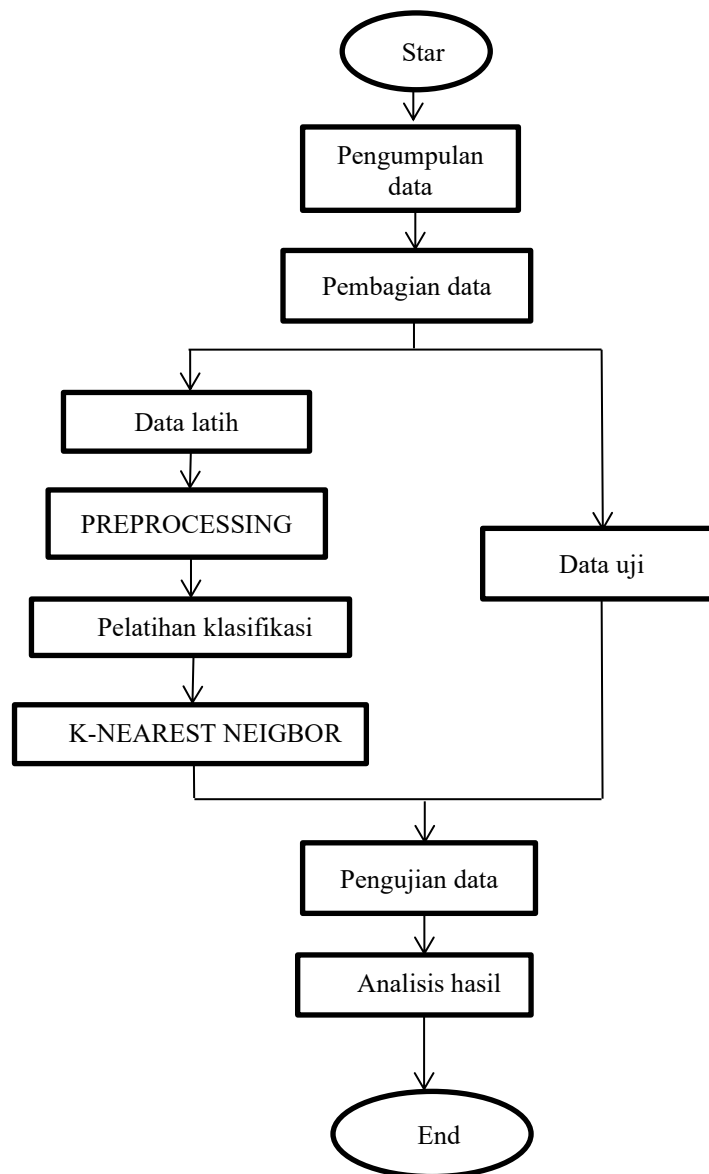
Pengolahan citra digital adalah salah satu teknologi image processing yang memanfaatkan citra suatu gambar untuk keperluan tertentu contohnya seperti di bidang pertanian mengidentifikasi kematangan suatu buah, kualitas sayur dan sebagainya. di Indonesia, sebagian besar masyarakat mengandalkan disektor pertanian (Astiningrum, Syulistyo, and Zakariya n.d.).

Dalam citra digital klasifikasi bentuk sebuah objek bisa menggunakan beragam macam metode. ada beberapa metode yang dapat digunakan dalam klasifikasi bentuk objek seperti K-Nearest Neighbors, Backpropagation, Local Binary Pattern, Super Vector mechindan lain-lain. dalam klasifikasi objek menggunakan beberapa metode tersebut tentunya memiliki hasil akhir yang tidak sama hal ini dapat terjadi karena mengacu pada dataset yang diteliti lebih berfokus pada metode yang mana (Wang 2020). dataset adalah sekumpulan data yang dapat digunakan sebagai bahan percobaan riset. beberapa studi mengumpulkan data mereka sendiri sebagai bahan percobaan (Salsabila, Dewanta, and Nuha 2023). beberapa yang lain mengambil dari tempat lain seperti *Papers With Code*, *data.go.id*, *satudata.go.id*. *dataset* dapat direpresentasikan dalam berbagai bentuk misalnya bentuk tabel dalam basis data, bentuk matriks, bentuk teks, bentuk gambar, bentuk Comma Separated Value (CSV) dan sebagainya (Rachmat and Lukito 2016).

K-Nearest Neighbors atau yang biasa disebut KNN. KNN merupakan sebuah algoritma pengklasifikasian antara suatu data berdasarkan objek yang terdekat. KNN juga merupakan algoritma pengklasifikasian dengan rumus yang cukup sederhana sehingga mudah untuk diaplikasikan (Christopher and Mulyana 2022). algoritma ini dapat digunakan untuk mengklasifikasikan data berdasarkan pola dan kemiripan dengan data (Fauzi et al. 2023). Alasan peneliti melakukan penelitian yang menghasilkan dataset adalah agar bisa bermanfaat bagi orang lain, bisa berkontribusi khususnya bagi sektor pertanian, dapat digunakan sebagai referensi untuk penelitian lanjutan. serta bertujuan untuk memperoleh hasil akurasi yang optimal dalam mengklasifikasikan biji jagung menggunakan K-Nearest Neighbor dengan melakukan pengujian terhadap nilai k.

2. METODE PENELITIAN

Tahapan penelitian yang digunakan dalam penelitian ini dapat dilihat pada gambar 1.



Gambar 1. Alur tahapan penelitian

Gambar 1 menunjukkan alur proses penelitian ini dibagi menjadi beberapa tahap. Pertama, peneliti mengumpulkan data yang dibutuhkan untuk penelitian berupa data biji jagung. Kedua, pembagian data menjadi dua, data latih dan data uji. Ketiga, preprocessing dalam tahapan ini aplikasi yang digunakan adalah matlab R2017. Proses preprocessing dilakukan proses perubahan data dari citra warna ke citra grayscale, biner, imcomplement, morfologi, filling holes, area opening, ekstraksi ciri warna RGB. Keempat, melakukan pelatihan model K-Nearest neighbor pada data latih menggunakan fungsi yang telah disediakan oleh MATLAB. Pada tahap ini, model akan mempelajari pola dan hubungan antara fitur-fitur citra dengan kategori atau label yang sesuai. Setelah model K-Nearest neighbor terlatih, langkah selanjutnya adalah melakukan pengujian pada data uji menggunakan citra-citra yang belum pernah dilihat sebelumnya. Model akan menghitung jarak atau kesamaan antara data baru dengan data latihan untuk menemukan K tetangga terdekatnya. Kelima, pengujian data. Pengujian dilakukan dengan membandingkan hasil ciri latih dan target latih menggunakan K-Nearest neighbor dengan citra RGB dan kelas yang diketahui. Proses ini memastikan apakah K-Nearest neighbor yang diterapkan mampu mendeteksi kelas-kelas data latih yang ditentukan. Keenam, hasil data uji akan dibandingkan dengan tiga tahapan pengujian yang berguna untuk memastikan bahwa model klasifikasi biji jagung dengan algoritma k-nearest neighbor yang dikembangkan tidak hanya

akurat pada data yang digunakan untuk melatih, tetapi juga kokoh, dapat meningkatkan generalisasi, dan siap untuk digunakan dalam skenario dunia nyata.

3. HASIL DAN ANALISIS

preprocessing menggunakan matlab untuk membaca file citra data latih didalam workspace dan menginisialisasi variabel citra, kemudian membuat gambar skala abu-abu (Grayscale) adalah gambar yang hanya memiliki satu warna yaitu abu-abu.



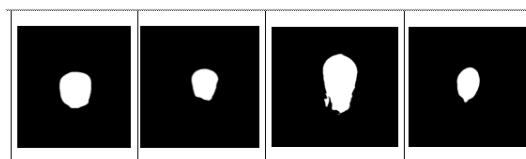
Gambar 1. Menampilkan citra grayscale

mengubah gambar grayscale menjadi gambar biner. Citra biner adalah citra digital yang hanya memiliki dua warna, yaitu hitam dan putih, dan setiap piksel direpresentasikan oleh satu bit data (0 atau 1).



Gambar 2. Menampilkan citra biner

komplemen adalah citra baru yang dihasilkan dengan membalik nilai piksel pada citra biner asli. Jika piksel pada citra asli berwarna hitam (nilai 0), maka pada citra komplemennya akan menjadi putih (nilai 1), dan sebaliknya.



Gambar 3. Menampilkan citra komplemen

filling holes bertujuan untuk mengisi area gelap (lubang) dalam sebuah gambar dengan warna atau nilai yang sesuai dengan sekitarnya, sehingga menghasilkan gambar yang lebih halus dan utuh.



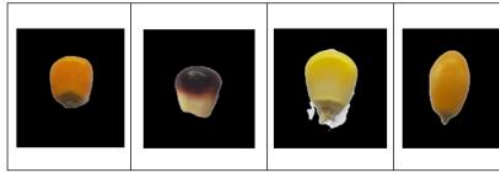
Gambar 4. Menampilkan citra filling holes

area opening adalah operasi morfologi yang menghilangkan objek-objek kecil dan detail yang tidak diinginkan dari suatu citra, sehingga menghasilkan citra yang lebih halus dan bersih dari noise.



Gambar 5. Menampilkan citra area opening

menampilkan citra Red, green, blue, (RGB) dari hasil segmentasi citra. mengisi variabel ciri biji dengan hasil segmentasi dan mengisi variable target biji dengan nama kelas biji.



Gambar 6. Menampilkan hasil segmentasi citra RGB

Berdasar hasil processing citra didapatkan nilai red, green, blue seperti tabel berikut.

Tabel 1. daftar rata-rata nilai RGB data latih pada Jagung Ayam

No.	Red	Green	BLUE
1.	0.653322702240578	0.371027658420862	0.0830732925484273
2.	0.707262385801672	0.474955234210918	0.154785150463603
3.	0.666560910671929	0.446973941222170	0.109152342777319
4.	0.685204510413592	0.465028792112477	0.131885481865708
5.	0.685002771061789	0.479180526072398	0.154477796601115
6.	0.698153838606302	0.471768490436117	0.152984733881042
.			
.			
.			
100.	0.627087792953577	0.369018510090463	0.101948046470878

Tabel 2. daftar rata-rata nilai RGB data latih pada Jagung hitam

No.	Red	Green	Blue
1.	0.396929382000514	0.313670431891514	0.267929845907615
2.	0.395723865328139	0.312536703415162	0.287174683678130
3.	0.409284512340167	0.326930181440009	0.288705483965484
4.	0.458712966724410	0.354451608211870	0.274714324853330
5.	0.410351503858841	0.318853320086926	0.270719462668560
6.	0.417163060910441	0.320564971591938	0.274471838688690
.			
.			
.			
100	0.411329638656576	0.283372211388357	0.234991166997646

Tabel 3. daftar rata-rata nilai RGB data latih pada Jagung manis

No.	Red	Green	Blue
1.	0.754749000399334	0.627518735877955	0.239557253992085
2.	0.742460843688105	0.613885066290801	0.207691563794099
3.	0.741022421925794	0.659174629849695	0.437216261482390
4.	0.698481851037615	0.622847006524109	0.412928811896058
5.	0.713517539906746	0.627476601479916	0.349612635267674
6.	0.728884269273493	0.634209676134078	0.371046556506362
.			
.			
.			
100.	0.646189775070822	0.523295429461588	0.193179739804341

Tabel 4. daftar rata-rata nilai RGB data latih pada Jagung popcorn

No.	Red	Green	Blue
1.	0.625947148333699	0.502714908581126	0.281309603299851
2.	0.703157098876899	0.498123948605997	0.166412894539976
3.	0.693801820913992	0.485768688806735	0.110674601317589
4.	0.694247012384694	0.485989946293789	0.157854143861882

5.	0.703648650522293	0.496087748633608	0.158782433635312
6.	0.699318847929997	0.492734798705071	0.156847448322020
.			
.			
.			
100.	0.691772053354745	0.490062704436021	0.126581572384827

Pada penerapan algoritma KNN , mengambil nilai ($K = 1$) dan mendapatkan ciri latih dan target latih dari 1-100 merupakan biji jagung ayam, 101-200 merupakan biji jagung manis, 201-300 merupakan biji jagung manis, 301-400 merupakan biji jagung popcorn. hasil akurasi pelatihan sebesar = 100% dari target latih dan ciri latih.

```

akurasi_pelatihan =
100
fx >>

```

Gambar 7. Menampilkan hasil akurasi pelatihan

Pada tahapan pengujian pertama dengan data uji bahwa 29 data uji yang diklasifikasikan ke dalam kelas yang salah (tidak sesuai target). Sehingga akurasi sistem dalam mengklasifikasikan biji jagung adalah $(103/132) \times 100\% = 78\%$

```

akurasi_pengujian =
78.0303
fx >>

```

Gambar 8. Menampilkan hasil akurasi pengujian tahap 1

pada tahapan pengujian kedua dengan data uji bahwa 26 data uji yang diklasifikasikan kedalam kelas yang salah (tidak sesuai target). Sehingga akurasi sistem dalam mengklasifikasikan biji jagung adalah $(106/132) \times 100\% = 80,3\%$

```

akurasi_pengujian =
80.3030
fx >>

```

Gambar 9. Menampilkan hasil akurasi pengujian tahap 2

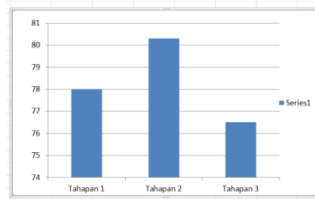
pada tahapan pengujian ketiga dengan data uji bahwa 31 data uji yang diklasifikasikan kedalam kelas yang salah (tidak sesuai target). Sehingga akurasi sistem dalam mengklasifikasikan biji jagung adalah $(106/132) \times 100\% = 76,5\%$

```

akurasi_pengujian =
76.5152
fx >>

```

Gambar 10. Menampilkan hasil akurasi pengujian tahap 3



Gambar 11. Menampilkan diagram hasil akurasi ketiga tahapan

Berdasarkan jumlah akurasi yang didapat menunjukkan bahwa algoritma k-nearest neighbor cukup baik dalam mengklasifikasi biji jagung dengan citra (RGB). kemudian peneliti memasukkan dataset berupa gambar dan file coding ke github dengan judul classification-of-corn-seeds atau bisa menggunakan link <https://github.com/bima-aditiansyah/classification-of-corn-seeds> dan peneliti juga memasukkan diwebsite Kaggle dengan judul corn kernel dataset.

4. KESIMPULAN

dapat disimpulkan bahwa algoritma k-nearest neighbor dalam pengolahan citra digital mampu terbukti efektif dalam mengklasifikasi biji jagung. algoritma ini mampu mengidentifikasi fitur Red Green Blue pada citra dengan tingkat akurasi yang cukup tinggi. pada tahapan pelatihan peneliti melakukan uji coba algoritma dengan nilai K 1, dengan hasil akurasi 100% pada tahapan pengujian peneliti melakukan tiga tahapan agar algoritma k-nearest neighbor benar-benar efisien dalam klasifikasi. pada tahapan pertama hasil penelitian ini menunjukkan bahwa tingkat akurasi yang didapat sebesar 78%, tahapan kedua mendapatkan akurasi sebesar 80,3% dan tahapan terakhir mendapatkan akurasi sebesar 76,5% dengan dijumlahkan ketiganya dan di bagi tiga mendapatkan hasil sebesar 78,26%, yang menunjukkan keandalan k-nearest neighbor ini dalam klasifikasi biji jagung.

REFERENSI

1. Astiningrum, Mungki, Arie Rachmad Syulistyo, and M Alfin Zakariya. "Warna Dan Tekstur Menggunakan HSV DAN GRAY LEVEL RUN LENGTH MATRIX (GLRLM)." : 37–44.
2. Christopher, Arvin, and Teady Matius Surya Mulyana. 2022. "Klasifikasi Tumbuhan Angiospermae Menggunakan Algoritma K-Nearest Neighbor Berdasarkan Pada Bentuk Daun." 07: 1233–43.
3. Fauzi, Mohammad et al. 2023. "Penerapan Algoritma K-Nearest Neighbor Dalam Klasifikasi Penyakit Jantung." 4(4): 850–60.
4. Lutfi, Moch. 2019. "Implementasi Metode K-Nearest Neighbor Dan Bagging Untuk Klasifikasi Mutu Produksi Jagung." 10(2): 130–37.
5. Maulidiansyah, and Isnaini Abdillah. 2023. "Klasifikasi Jenis Jagung Berdasarkan Bentuk Biji Menggunakan Metode You Only Look Once (YOLO) by Maulidiansyah Maulidiansyah." 2023.
6. Rachmat, Antonius, and Yuan Lukito. 2016. "SENTIPOL: Dataset Sentimen Komentar Pada Kampanye PEMILU Presiden Indonesia 2014 Dari Facebook Page." *Konferensi Nasional Teknologi Informasi dan Komunikasi 2016* (December): 218–28. http://publikasi-fti.ukdw.ac.id/Paper_Klasifikasi_Komentar_Spam_pada_Instagram_Berbahasa_Indonesia_Menggunakan_K_NN_Yuan_Lukito.pdf.
7. Rosiani, Ulla Delfana, Mustika Mentari, and Andi Novan Prastya. "Klasifikasi Kualitas Biji Jagung Berdasarkan Deteksi Warna Dan Bentuk Menggunakan Metode."
8. Salsabila, Balqis Hasna, Favian Dewanta, and Hilal Hudan Nuha. 2023. "Pembuatan Dataset Benih Padi Varietas Ciherang Berbasis Sistem Pengolahan Citra." 10(5): 4253–56.
9. Wang, L I N. 2020. "Tinier-YOLO: A Real-Time Object Detection Method for Constrained Environments." *IEEE Access* 8: 1935–44.
10. Zahara, Muhamad, Anugrah Putra, Feri Candra, and Esa Prakasa. 2024. "Klasifikasi Kualitas Varietas Benih Jagung Bima 20 Menggunakan Metode Random Forest." 10(2): 367–85.