

# Klasifikasi Daun Padi dengan *K-Nearest Neighbor* Berdasarkan Fitur Warna dan *Invariant Moment*

## *Rice Classification with K-Nearest Neighbor based on Color Feature Extraction and Invariant Moment*

Santika Tri Hapsari S, Rahmat Widadi\*, Indah Permatasari

Teknik Telekomunikasi, Teknik Telekomunikasi dan Elektro Institut Teknologi Telkom Purwokerto  
Jalan DI Panjaitan No. 128, Kec. Purwokerto Selatan, Kab. Banyumas, Jawa Tengah Indonesia

\*e-mail: [rahmat@ittelm-pwt.ac.id](mailto:rahmat@ittelm-pwt.ac.id)

(received: 17 Mei 2023, revised: 2 Juli 2023, accepted: 6 Juli 2023)

### Abstrak

Nasi ialah makanan pokok warga Indonesia yang berasal dari tanaman padi. Tanaman padi sering mengalami gagal panen karena terserang penyakit. Tentunya hal tersebut akan berpengaruh pada hasil panen. Oleh karena itu, di era kemajuan teknologi ini dapat menggunakan citra digital yang dapat membantu petani untuk mengklasifikasi penyakit daun padi agar dapat dikendalikan. Salah satu klasifikasi menggunakan *K-Nearest Neighbor* (KNN) yang bersumber pada informasi data pembelajaran dengan jarak yang terdekat. Penelitian menerapkan metode ekstraksi fitur warna dan *invariant moment* agar mendapatkan informasi ciri pembeda suatu objek dari objek yang lain. Data berasal dari *UCI Machine Learning Repository* berjumlah 120 citra yang dibagi menjadi 3 jenis penyakit *bacterial leaf blight*, *brown spot*, dan *leaf smut* dengan setiap kelas memiliki 40 citra. Fitur warna yang digunakan HSV yaitu *Hue*, *Saturation*, dan *Value*. Sedangkan pada *invariant moment* menggunakan tujuh fitur H1 hingga H7 yang dikenalkan oleh Hu. Seleksi fitur dilakukan setelah proses ekstraksi fitur untuk mendapatkan nilai akurasi tertinggi. Selain itu, Variasi jumlah tetangga (*k*) pada KNN juga divariasikan dari  $k=1$  sampai  $k=10$ . Hasil akurasi terbaik didapatkan pada penggunaan fitur yaitu *hue*, *saturation*, *value*, *h2*, *h3*, dan *h7* dan nilai jumlah tetangga pada KNN  $k=1$  dengan akurasi sebesar 81,66%.

**Kata kunci:** Daun Padi, Fitur Warna, *Invariant Moment*, *K-fold*, KNN.

### Abstract

Rice is the staple food of Indonesians which comes from rice plants. Rice plants often experience crop failure due to disease. Of course this will affect the yield. Therefore, in this era of technological advances, digital images can be used to help farmers classify rice leaf diseases so they can be controlled. One of the classifications uses *K-Nearest Neighbor* (KNN) which is sourced from learning data information with the closest distance. Research requires color feature extraction and *invariant moment* methods in order to obtain information on the distinguishing characteristics of an object from other objects. Data comes from the *UCI Machine Learning Repository* totaling 120 images which are divided into 3 types of bacterial disease leaf blight, brown spot, and leaf smut with each class having 40 images. The color features used by HSV are *Hue*, *Saturation*, and *Value*. Meanwhile, the *invariant moment* uses the seven features H1 to H7 introduced by Hu. Feature selection is carried out after the feature extraction process to get the highest accuracy value. In addition, variations in the number of neighbors (*k*) in KNN are also varied from  $k=1$  to  $k=10$ . The best accuracy results are obtained from the use of features, namely *hue*, *saturation*, *value*, *h2*, *h3*, and *h7* and the value of the number of neighbors in KNN  $k=1$  with an accuracy 81.66%.

**Keywords:** Color Feature, *Invariant moment*, *K-fold*, KNN, Rice Leaf.

## 1 Pendahuluan

Indonesia mempunyai kemampuan basis energi hayati khusus posisi yang banyak dengan bermacam tipe tumbuhan pangan semacam padi, umbi, kacang, sayur, buah, serta basis pangan hewani.

<http://sistemasi.ftik.unisi.ac.id>

Secara khusus beras sebagai sumber makanan pokok dan bahan nasi. Nasi ialah santapan utama masyarakat Indonesia yang berperan buat penuh kepentingan karbohidrat [1]. Berdasarkan Berita Resmi Statik No. 21/03/Th. XXV, 1 Maret 2022 mengenai Besar Panen serta Produksi gabah di Indonesia 2021 (Nilai Konsisten) bahwa produksi gabah dalam 2021 ialah sebesar 54, 42 juta ton gabah kering giling (GKG) hadapi penyusutan sebesar 233,91 ribu ton ataupun 0,43 persen dibanding produksi gabah di 2020 yang sebesar 54,65 juta ton GKG. Tidak hanya itu, produksi beras dalam 2021 untuk konsumsi pangan masyarakat menggapai 31,36 juta ton, hadapi penyusutan sebesar 140,73 ribu ton ataupun 0,45 persen dibanding produksi beras di 2020 yang sebesar 31, 50 juta ton [2]. Salah satu faktor masalah yang mengakibatkan penurunan jumlah produksi beras seperti hama dan penyakit padi. Seperti di Kota Padang mengalami penurunan padi pada tahun 2020 sebanyak 6.529,13 ton atau 18,03 % yang diakibatkan oleh serbuan wereng serta penyakit padi. Perihal ini menyebabkan terjadinya kerugian petani dalam produksi padi tersebut [3].

Di era kemajuan pesat teknologi saat ini membawa berbagai dampak perubahan dalam mempermudah manusia dalam menemukan solusi dari masalah yang terdapat. Salah satu teknologi yang bertumbuh dikala ini merupakan citra digital. Citra digital sudah sering digunakan dalam mengenali proses masalah dalam keseharian [4]. Oleh karena itu, citra digital bisa membantu para petani dalam mengidentifikasi penyakit dalam daun padi supaya penyakit daun padi bisa teratasi.

Salah satu cara klasifikasi merupakan memakai *K- Nearest Neighbor* (KNN). KNN ialah cara pengelompokan kepada subjek bersumber pada data pembelajaran dengan jarak yang terdekat. Cara KNN diaplikasikan guna menanggulangi kasus pengenalan yang diukur dengan cara kualitatif ataupun kuantitatif. Tentukan nilai  $k$  terlebih dahulu sebelum mencari data dengan jarak terdekat dengan tetangga [5]. Beberapa objek maupun data yang digunakan terdapat keberagaman yang sangat kompleks. Oleh sebab itu, diperlukan suatu metode ekstraksi fitur mendapatkan informasi atau karakteristik pembeda yang bisa membedakan sesuatu subjek dari subjek yang lain [6].

Sebagian riset lebih dahulu telah sempat memakai cara klasifikasi yang berbeda antara lain dengan memakai cara *Convolution Neural Network* (CNN) dengan membuahkan ketepatan penataran pembibitan menggapai 100%, ketepatan validasi sebesar 83.3%, serta angka ketepatan dalam *confusion matrix* menghasilkan sebesar 92% [7]. Dalam riset lain memakai cara ekstraksi fitur GLCM interval 4 sudut menghasilkan dari 30 data 80% data valid dengan durasi 16 s [8]. Riset klasifikasi penyakit padi yang lain mengenakan citra daun dengan model terlatih Resnet101 dengan hasil akurasi sebesar 100% serta loss sebesar 5,61% [9]. Selanjutnya, terdapat penelitian menggunakan metode klasifikasi *deep learning Efficientnet B3* dengan *transfer learning* dengan hasil pengujian tercapai akurasi 79,53% dan *loss* 0,012 [10]. Penelitian ini bertujuan untuk mengidentifikasi dan menganalisis performa dari klasifikasi penyakit pada daun padi dengan menggunakan KNN berdasarkan fitur warna dan fitur invariant moment.

## 2 Tinjauan Literatur

Penelitian [7] mengenai deteksi penyakit daun padi dengan algoritma *Convolutional Neural Network* (CNN) dilakukan oleh beberapa peneliti Rizal Amegia Saputra et al tahun 2021 melakukan klasifikasi penyakit padi dengan *MobileNetV1* menggunakan fitur. Dataset berjumlah 120 yang terdiri dari 3 penyakit daun padi yaitu *bacterial leaf blight*, *brown spot*, *leaf smut*. Berdasarkan hasil latihan dan uji citra dataset dengan ukuran yang sama yaitu 224x224. Hasil pelatihan data memiliki akurasi 1.0 dan validasi sebesar 0,8333 dengan menggunakan *confusion matrix* mendapatkan tingkat akurasi 92%.

Penelitian Jani Kusanti et al 2018 melakukan klasifikasi penyakit daun padi berdasarkan 4 variasi sudut pada GLCM. Pada penelitian ini dibuatlah sesuatu program aplikasi yang bisa mengenali hama penyakit padi. Oleh sebab itu dibuatlah sesuatu program yang bisa mengetahui hama penyakit padi, melalui metode *backpropagation* dengan 4 variasi ciri sudut GLCM selaku sistem klasifikasi penyakit dalam daun tumbuhan padi. Melalui metode *backpropagation* pada pengolahan citra, dengan klasifikasi dari nilai ciri setiap citra penyakit daun padi yang digunakan. Sistem ini memiliki akurasi mencapai 80% dari 30 data dengan waktu 16 detik dalam melakukan identifikasi penyakit pada citra daun tanaman padi [8].

Penelitian [10] Endang Anggiratih et al 2021 melakukan pengidentifikasi penyakit tanaman padi menggunakan *deep learning Efficientnet B3* dengan *transfer learning*. Pada penelitian ini menggunakan metode *Deep learning* khususnya CNN dengan arsitektur Beberapa penyakit padi yang dapat dianalisa yaitu *brown spot* (bercak coklat) dan *bacterial leaf* (hawar daun). Peneliti menganalisa penyakit tersebut

<http://sistemasi.ftik.unisi.ac.id>

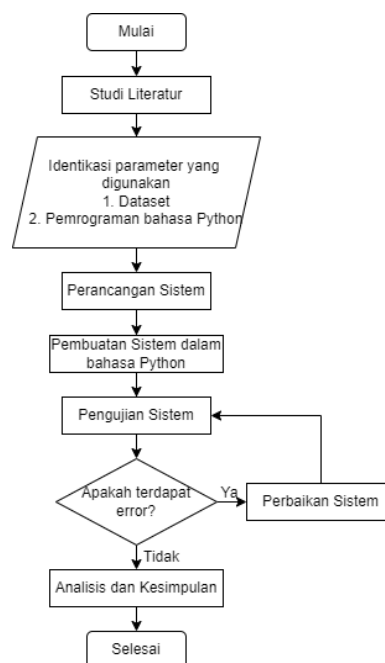
pada padi dengan menggunakan *Deep Learning* tepatnya *Convolutional Neural Network* menggunakan arsitektur *EfficientNet B3* yang dilatih menggunakan data latih (dataset) berupa citra. Akurasi sebesar 79,53% menggunakan *EfficientNet B3* yang mana hasil lebih baik daripada menggunakan *Mobilenet V3*.

Penelitian [11] Zulkifli tahun 2021 melakukan pengidentifikasian penyakit tanaman padi menggunakan *Artificial Intelligence* dan sistem pakar yang mana menggunakan pemeriksaan pakar yang optimal. Salah satu metode dalam pengidentifikasian ini adalah *Expert System Development Life Cycle (ESDLC)* yang dibagi menjadi dua yaitu evaluasi, pemetaan wawasan ahli, konsep, serta pengetesan. Ada pula hasil yang diperoleh pada riset ini merupakan terdapatnya suatu aplikasi pendeteksi penyakit tumbuhan padi berlandaskan kecerdasan buatan dengan memakai cara sistem ahli yang bisa dipakai oleh petani buat mendiagnosis penyakit tumbuhan padi dengan cara mandiri tanpa mesti bertemu langsung dengan pakarnya.

Penelitian [12] Mohtar Khoiruddin *et al* tahun 2022 mengenai klasifikasi daun padi. Metode yang digunakan adalah metode *Convolution Neural Network (CNN)*. Dataset yang digunakan memiliki 3 jenis penyakit yaitu *bacterial leaf blight*, *rice blast*, dan *rice tungro virus* yang berjumlah 6000. Dataset yang digunakan akan dibagi menjadi 0.8 data training, 0.1 data validasi, dan sisanya yaitu 0.1 akan menjadi data testing. Penelitian ini menggunakan nilai epoch yang berbeda yaitu sebesar 25,50,75, dan 100 dengan hasil terbaik menggunakan epoch 100 yang menghasilkan tingkat akurasi sebesar 98%.

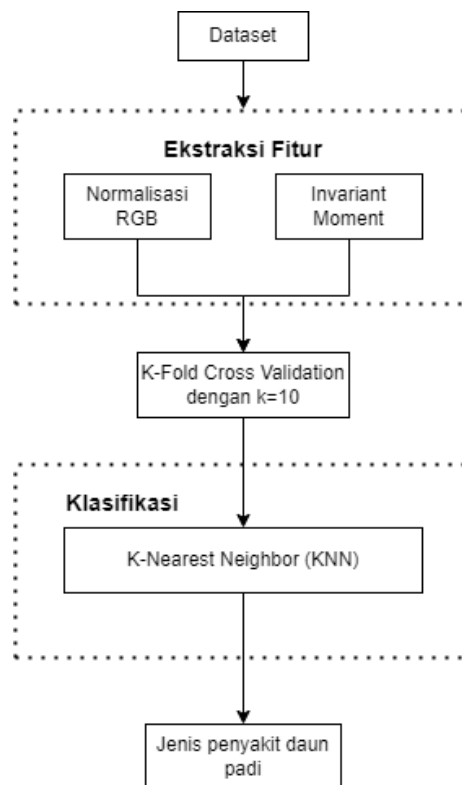
### 3 Metode Penelitian

Pembuatan sistem ini akan dilakukan dalam beberapa langkah. Langkah-langkah pada riset ini, secara umum dapat dilihat dari diagram alur Gambar 1 berikut:



**Gambar 1. Tahapan Penelitian**

Gambar 1 mengenai tahapan penelitian, dalam perancangan sistem terdapat gambaran arsitektur umum dari sistem tersebut seperti Gambar 2 berikut:



**Gambar 2. Arsitektur Umum sistem**

Gambar 2. menjelaskan arsitektur sistem yang terdiri dari tiga bagian yaitu input, proses dan output. Input yang digunakan berupa dataset citra penyakit daun padi yang sudah diklasifikasi menjadi 3 kelas dengan masing-masing kelas memiliki 40 citra. Selanjutnya dilakukan proses ekstraksi fitur warna dan *invariant moment*. Setelah mendapatkan nilai fitur, dataset akan dilakukan proses *k-fold cross validation* untuk membagi data menjadi data *training* dan data testing. Kemudian klasifikasi dilakukan menggunakan metode KKN untuk mengklasifikasikan penyakit daun padi. Penjelasan detail terkait arsitektur umum sistem pada penelitian ini sebagai berikut:

### 3.1. Dataset

Tahap pertama pada penelitian ini membutuhkan citra berupa dataset sebagai inputan. Dataset dalam penelitian ini berjumlah 120 citra yang dibagi menjadi 3 macam penyakit padi seperti pada Tabel 1 di bawah ini.

**Tabel 1. Dataset**

No.	Klasifikasi Daun	Jumlah Dataset
1.	<i>Bacterial Leaf Blight</i>	40
2.	<i>Brown Spot</i>	40
3.	<i>Leaf Smut</i>	40
Jumlah		120

Tabel 1 menjelaskan dataset yang terdiri dari 3 kelas seperti pada yang ada ditabel dengan masing-masing data 40 citra dengan total data yang akan digunakan dalam penelitian ini adalah 120. Dataset tersebut didapatkan dari [13].

### 3.2. Ekstraksi Fitur

Pada tahap berikutnya yaitu pencarian fitur ciri berdasarkan karakteristik dan intrinsik yang signifikan pada citra atau gambar yang sudah diinputkan. Setelah mendapatkan nilai fitur yang sesuai karakteristik dan intrinsik, maka nilai fitur akan diteruskan pada tahap proses klasifikasi. Fitur sendiri ialah ciri khas yang dapat membedakan dari objek satu dengan yang lainnya. Ekstraksi fitur dibedakan

<http://sistemasi.ftik.unisi.ac.id>

menjadi 3 yaitu berdasarkan tekstur, bentuk, serta warna. Penelitian ini berdasarkan fitur warna yaitu pengambilan nilai menggunakan warna HSV yang nantinya mendapatkan nilai *hue*, *saturation*, dan *value*. Nilai *hue* ini mempresentasikan warna dari merah ke hijau, *saturation* dari merah ke merah muda, serta untuk *value* sendiri mempresentasikan dari warna hitam ke putih [14].

$$r = \frac{R}{(R + G + B)}, g = \frac{G}{(R + G + B)}, b = \frac{B}{(R + G + B)} \quad (1)$$

$$V = \max(r, g, b) \quad (2)$$

$$S = \begin{cases} 0 & , \text{jika } V = 0 \\ 1 - \frac{\min(r, g, b)}{V} & , \text{jika } V > 0 \end{cases} \quad (3)$$

$$H = \begin{cases} 0 & , \text{jika } S = 0 \\ \frac{60 * (g - b)}{S * V} & , \text{jika } V = r \\ 60 * \left[ 2 + \frac{b - r}{S * V} \right] & , \text{jika } V = g \\ 60 * \left[ 2 + \frac{r - g}{S * V} \right] & , \text{jika } V = b \end{cases} \quad (4)$$

$$H = H + 360 \text{ jika } h < 0 \quad (5)$$

Keterangan:

r,g,b = Derajat keabuan dari sebuah piksel *red*, *green* dan *blue*

R,G,B = Citra warna *red*, *green* dan *blue* (0-255)

H = *hue* mempresentasikan warna dari merah ke hijau (0-255)

S = *Saturation* mempresentasikan merah ke merah muda (0-255)

V = *Value* mempresentasikan dari warna hitam ke putih (0-255)

Pada persamaan 1 sampai persamaan 5 digunakan untuk mencari nilai fitur warna yang diperlukan pada penelitian ini. Persamaan 1 diperlukan dalam menormalisasi suatu citra agar mendapatkan nilai derajat keabuan sebuah piksel citra *red*, *green*, dan *blue* dalam bentuk citra RGB. Setelah citra dinormalisasi menjadi citra RGB selanjutnya akan dikonversi menjadi citra HSV menghasilkan tiga fitur. Selanjutnya persamaan 2 untuk mencari nilai *value* yang mana nilai ini adalah dengan mencari nilai tertinggi dari r,g,b yang sudah dinormalisasi pada persamaan 1. Persamaan 3 untuk mencari nilai *saturation* yang mana terdapat 2 kondisi, jika *value* 0 atau nilai *value* > 0, maka akan menggunakan rumus yang berbeda. Kemudian persamaan 4 digunakan dalam menghitung nilai *hue* yang mana terdapat beberapa kondisi yang dipengaruhi oleh persamaan 1 hingga 3. Persamaan 5 digunakan jika pada persamaan 4 nilai H kurang dari 0° sehingga nilai akhir H akan ditambah dengan 360°.

Selain menggunakan fitur warna juga menggunakan fitur bentuk yaitu invariant moment yang nantinya akan menghasilkan 7 fitur H. Hu memberitahukan 7 *invariant* yang diserahkan pada persamaan [15].

$$H_1 = \eta_{20} + \eta_{02} \quad (6)$$

$$H_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (7)$$

$$H_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (8)$$

$$H_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (9)$$

$$H_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (10)$$

$$H_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (11)$$

$$H_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} + 3\eta_{12})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (12)$$

Persamaan 6 menggambarkan  $H_1$  sebagai momen inersia yang berhubungan dengan pusat citra dan memiliki sifat refleksi simetri dimana tidak berubah ketika dilakukan

pencerminan. Kemudian Persamaan 7 hingga persamaan 11 yaitu  $H2 - H6$ , juga memiliki sifat refleksi simetri yang serupa dengan  $H1$ . Persamaan 12 menyatakan bahwa  $H7$  memiliki sifat antisimetri yang berarti akan berubah ketika dilakukan pencerminan pada citra. Ketujuh fitur ini, seperti yang dijelaskan dalam persamaan diatas digunakan untuk menjelaskan objek dengan menghitung luas area objek tersebut. Fitur ini didasarkan pada momen pusat yang dinormalisasi. Dengan menggunakan momen yang dihasilkan, fitur-fitur ini dapat digunakan untuk penskalaan, rotasi gambar dan translasi [16].

### 3.3. K-Fold Cross Validation

*K-fold cross validation* adalah pembagian data *training* dan *testing* dalam himpunan  $k$  dengan jumlah yang sama dengan tujuan memperkecil kesalahan prediksi saat mengevaluasi kinerja sistem. Sistem akan dilatih dan diuji sebanyak  $k$  yang ditentukan [17].

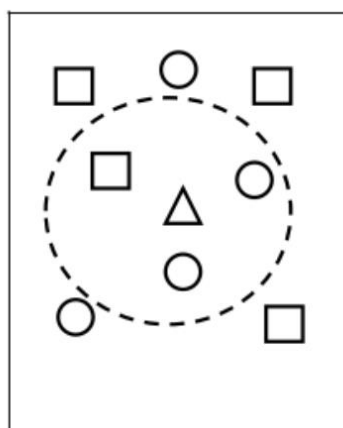
*K-Fold* ini akan digunakan saat proses klasifikasi mencapai tahap *training* dan *testing*. Penelitian menggunakan  $K=10$  yang berarti data masukkan *training* dan *testing* yang dibagi menjadi 12 bagian secara acak. Pada data yang ke-1 akan menjadi *testing* sedangkan pada data ke-2 hingga ke-10 akan menjadi *training*. Proses itu akan diulang sampai bagian data ke-12.

### 3.4. K-Nearest Neighbor

*K-Nearest-Neighbor* (KNN) merupakan algoritma yang digunakan untuk mengklasifikasikan objek untuk data baru. Dalam proses melakukan pengklasifikasian pada objek data baru akan dilatih berdasarkan atribut dan data sampel latih. Data pembelajaran diilustrasikan sebagai atribut digital dengan dimensi- $n$ .

Dengan  $k$  menyatakan jumlah tetangga terdekat pada KNN. Setiap data dihitung jaraknya menggunakan pengukuran jarak Euclidian[18]. Sejumlah  $k$  tetangga terdekat pada data dikelompokkan. Setelah mencari sejumlah  $k$  objek data yang paling dekat jaraknya dengan data uji yang diberikan, kemudian memilih kelas dengan jumlah voting terbanyak [19].

Metode KNN memiliki beberapa kelebihan yaitu KNN relatif tidak sensitif terhadap error dalam dataset, yang berarti dapat mengatasi ketidaktepatan dalam data latih. Kelebihan lainnya adalah KNN dapat digunakan untuk memproses dataset berukuran besar [20]. Dengan demikian KNN merupakan algoritma klasifikasi yang efektif dan fleksibel dalam mengatasi berbagai situasi dalam pengklasifikasian data. Berikut ilustrasi dari *nearest neighbor* dalam bentuk lingkaran adalah 2 data terdekat sehingga bentuk segitiga akan terprediksi pada kelas bentuk lingkaran [20].



**Gambar 3. Ilustrasi Nearest Neighbor**[20]

Pada Gambar 3 menyatakan apabila nilai  $k$  sama dengan tiga maka tiga tetangga terdekat yang akan diperhitungkan jaraknya. Dimana bentuk lingkaran dan persegi merupakan 3 data terdekat. Bentuk lingkaran dan persegi termasuk kedalam 2 kelas yang berbeda sehingga

bentuk segitiga akan diprediksi berdasarkan jarak terdekat dari 2 data tersebut. Oleh karena itu, bentuk lingkaran yang akan menjadi kelas bentuk segitiga [20].

Algoritma KNN menggunakan pendekatan klasifikasi berdasarkan ketetanggaan untuk memprediksi nilai sampel uji yang baru. Dalam algoritma ini, Jarak *euclidean* digunakan data numerik yang didefinisikan sebagai berikut [15]:

$$d_i = \sqrt{\sum_{i=0}^p (x_{2i} - x_{1i})^2} \quad (13)$$

Keterangan:

- $x_{2i}$  = Data *Testing*
- $x_{1i}$  = Data *Training*
- $p$  = Jumlah dimensi data

Jarak *euclidean* dapat dikalkulasikan berdasarkan fitur yang terdapat pada citra. Fitur-fitur citra ini meliputi fitur warna, tekstur dan bentuk. Proses perhitungannya melibatkan *encoding* citra untuk menghasilkan 128 dimensional *face encodings*. Selanjutnya, jarak terdekat dihitung antara data yang ada dalam database menggunakan *euclidean distance* matrik [21]. Langkah-langkah dalam menghitung metode *K-Nearest Neighbor* sebagai berikut ini:

1. Menentukan parameter jumlah tetangga terdekat (k)
2. Menghitung jarak *euclidean* objek terhadap data training yang diberikan
3. Mengurutkan hasil poin secara *ascending* yaitu mengurutkan berdasarkan nilai tertinggi ke terendah
4. Mengumpulkan kategori Y (klasifikasi *nearest neighbor* berdasarkan nilai k)

### 3.5. Confusion Matrix

Data hasil dari *k-fold* kemudian akan dilakukan ke proses *confusion matrix* untuk melihat tingkat akurasi sebagai evaluasi dari sistem. *Confusion matrix* adalah suatu caya yang dipakai dalam mengukur akurasi dari sistem penelitian ini. Pemodelan algoritma yang telah dibuat akan dievaluasi untuk melihat seberapa bagus kinerja sistem melalui nilai akurasi yang didapatkan. Tingkat akurasi ini akan dihasilkan menggunakan data *testing*, data testing lalu akan di prediksi melalui data *training* yang sudah ada menggunakan *confusion matrix* [19]. Dikarenakan pada dataset menggunakan tiga kelas, sehingga penelitian menggunakan *multiple confusion matrix* 3 x 3 sebagai berikut:

**Tabel 2. Confusion Matrix 3x3**

		Prediksi		
		A	B	C
Aktual	A	$P_{AA}$	$P_{BA}$	$P_{CA}$
	B	$P_{AB}$	$P_{BB}$	$P_{CB}$
	C	$P_{AC}$	$P_{BC}$	$P_{CC}$

Sehingga, dalam penelitian ini rumus akurasi yang digunakan dengan menggunakan *Confusion Matrix* 3 x 3 berdasarkan tabel 2 sebagai berikut [22]:

$$accuracy = \frac{P_{AA} + P_{BB} + P_{CC}}{P_{AA} + P_{BA} + P_{CA} + P_{AB} + P_{BB} + P_{CB} + P_{AC} + P_{BC} + P_{CC}} \quad (14)$$

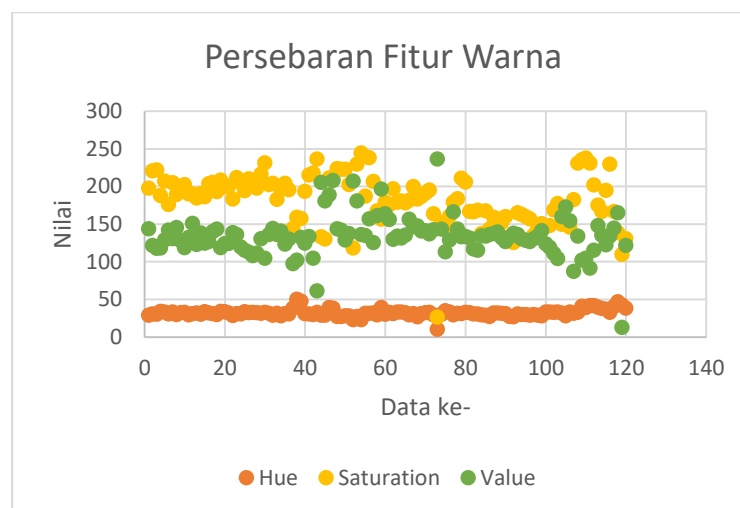
Dengan keterangan sebagai berikut:

- a.  $P_{AA}$  : Dimana ketika prediksi A diklasifikasi dengan benar sebagai A.
- b.  $P_{AB}$  : Dimana ketika prediksi A diklasifikasi dengan salah sebagai B.
- c.  $P_{CA}$  : Dimana ketika prediksi A diklasifikasi dengan salah sebagai C
- d.  $P_{AB}$  : Dimana ketika prediksi B diklasifikasi dengan salah sebagai A.
- e.  $P_{BB}$  : Dimana ketika prediksi B diklasifikasi dengan benar sebagai B.

- f.  $P_{CB}$  : Dimana ketika prediksi B diklasifikasi dengan salah sebagai C.
- g.  $P_{AC}$  : Dimana ketika prediksi C diklasifikasi dengan salah sebagai A.
- h.  $P_{BC}$  : Dimana ketika prediksi A diklasifikasi dengan salah sebagai B.
- i.  $P_{CC}$  : Dimana ketika prediksi C diklasifikasi dengan benar sebagai C.

#### 4 Hasil dan Pembahasan

Pada proses untuk mendapatkan nilai fitur warna yang terdiri dari 3 nilai fitur yaitu *Hue*, *Saturation*, dan *Value* dengan input citra. Citra akan di *preprocessing* dengan tujuan citra akan berfokus pada objek dan *background* akan dipotong hingga tersisa objek. Hal ini bertujuan untuk saat ekstraksi fitur nanti akan berfokus pada objek selanjutnya pengambilan nilai fitur yang mana pada penelitian ini menggunakan fitur warna HSV. Hingga mendapatkan 3 nilai dalam citra sebagai fitur atau karakteristik dari citra dan akan mendapatkan persebaran seperti pada gambar 4 di bawah ini.

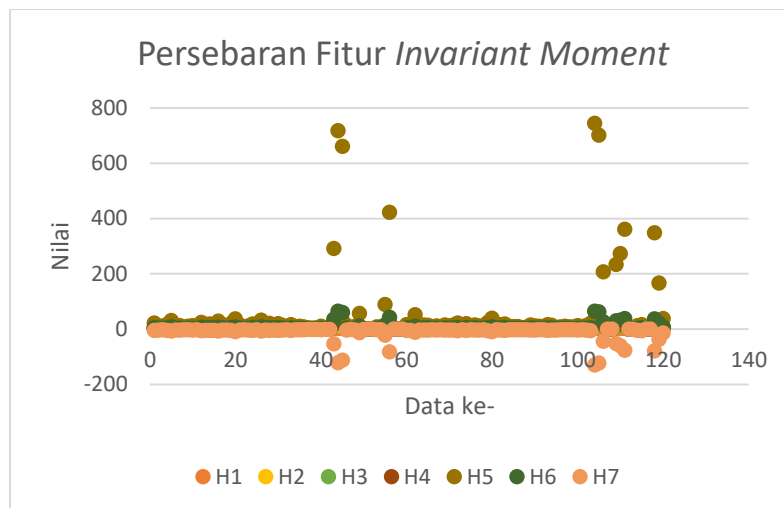


**Gambar 4. Persebaran Fitur Warna**

Gambar 4 menunjukkan penyebaran nilai *hue*, *saturation* dan *value* dalam 120 citra. Citra 1-40 merupakan citra dari kelas penyakit *bacterial leaf blight*, dimana penyebarannya terlihat pada gambar dengan nilai *hue* sekitar 30-40, *saturation* sekitar 150-225, dan *value* 100-150. Berdasarkan hasil tersebut dapat dikatakan bahwa kelas ini baik karena nilai *hue*, *saturation* dan *value* memiliki rentang masing-masing sehingga karakteristik dari penyakit *bacterial leaf blight* dapat terlihat. Kemudian citra 41-80 merupakan kelas penyakit *brown spot* dimana nilai *hue* dapat dilihat stabil dengan rentang nilai 25-40. Sedangkan rentang nilai *saturation* dan *value* diperoleh 120-250 dan 105-210. Pada gambar terlihat bahwa nilai *saturation* dan *value* dari kelas *brown spot* memiliki karakteristik yang tumpang tindih yang berbeda dengan kelas *bacterial leaf blight* yang memiliki jarak antar fitur *saturation* dan *value*. Meskipun demikian kelas *brown spot* memiliki beberapa nilai yang berbeda secara signifikan sehingga kualitas karakteristik pada kelas *brown spot* berkurang. Fitur *saturation* pada kelas *bacterial leaf blight* terlihat pada nilai 41-60, Sedangkan pada kelas *leaf smut* persebaran *saturation* terlihat pada data di data 61-80. Hal ini kurang baik karena nilai *saturation* dan *value* dapat mempengaruhi dalam melakukan pengklasifikasian. Citra 81-120 menunjukkan citra dari kelas penyakit *leaf smut* dengan nilai *hue* 26-50, *saturation* 125-240 dan *value* 85-170. Pada Gambar 3 menunjukkan pada data 81-100 memiliki nilai *hue*, *saturation*, dan *value* yang stabil dan karakteristik antar fitur dapat dikatakan baik. Akan tetapi pada data 101-120 nilai *saturation* dan *value* banyak mengalami *outlier*. Kemudian pada kelas penyakit *leaf smut* memiliki karakteristik dengan nilai *saturation* dan *value* yang lebih rapat dibandingkan dengan kelas *bacterial leaf* dan *brown spot*. Penyebaran ini dapat mempengaruhi pada hasil klasifikasi menggunakan KNN.

Penelitian ini selain menggunakan fitur warna, juga menggunakan fitur bentuk yang mana fitur bentuk yang digunakan adalah *invariant moment*. *Invariant moment* memiliki 7 fitur yaitu H1, H2, H3, H4, H5, H6, dan H7. Tiap citra akan memiliki 7 fitur yang akan menjadi karakteristik dari citra tersebut. Sehingga akan menghasilkan penyebaran seperti pada Gambar 5.





**Gambar 5. Persebaran Invariant Moment**

Pada Gambar 5 Menunjukkan penyebaran dari tiga kelas penyakit. Pada gambar diatas nilai  $H1 - H7$  pada kelas *bacterial leaf blight*, *brown spot* dan *leaf smut* memiliki nilai yang hampir sama. Oleh karena itu, dalam pengklasifikasian penyakit fitur  $H1 - H7$  memiliki pengaruh yang kurang signifikan karena sulit untuk melihat karakteristik khusus dari fitur-fitur tersebut dalam setiap kelas. Citra 1-40 merupakan kelas *bacterial leaf blight*, dimana nilai  $H1 - H7$  terlihat baik dan memiliki sedikit *outlier*. Namun, pada kelas *brown spot* khususnya citra 41-60, terdapat banyak *outlier* yang dapat mengganggu proses klasifikasi karena sulitnya dalam membaca karakteristik setiap fitur. Namun, pada citra 61-80 penyebarannya terlihat baik. Citra 81-120 merupakan kelas *leaf smut* dimana citra 80-100 menunjukkan nilai yang baik dengan sedikit *outlier*, tetapi pada citra 101-120 terdapat nilai fitur yang ekstrem dan jauh dari nilai umumnya.

Data dari analisis ini diambil ketika fitur warna dan *invariant moment* yang sudah didapatkan dan dilakukan proses *selection feature* dengan mengurutkan fitur menggunakan *feature importance* yang berpengaruh dalam klasifikasi. Proses klasifikasi KNN menggunakan *K-fold Cross Validasi* dengan  $k=10$  yang akan diuji menggunakan variasi fitur yang sudah diurutkan. Penelitian ini akan menggunakan notasi pada setiap jenis penyakit ke dalam nilai agar mempermudah pemahaman pada saat analisis.

**Tabel 3. Inisialisai Nilai Setiap Jenis Penyakit Daun Padi**

No.	Klasifikasi Jeni Daun	Nilai
1	<i>Bacterial leaf blight</i>	0
2	<i>Brown spot</i>	1
3	<i>Leaf smut</i>	2

Tabel 3 menunjukkan kelas *bacterial leaf blight* dinotasikan 0, *brown spot* dinotasikan 1, dan *leaf smut* akan dinotasikan 2. Pada saat pengujian menggunakan variasi fitur, setiap variasi fitur akan diuji menggunakan KNN  $k=1$  hingga  $k=10$  dengan hasil *output* berupa akurasi menggunakan *confusion matrix* dengan *k-fold* yang digunakan  $k=10$  sehingga setiap *confusion matrix* memiliki 12 data uji sebanyak 10 kali sehingga total adalah 120 data yang dibagi 40 data tiap kelas. Berikut adalah *confusion matrix* dari tingkat akurasi tertinggi tiap variasi fitur:

**Tabel 4. Confusion Matrix 1 Fitur Terbaik dengan  $k=10$**

		Prediksi		
		0	1	2
Aktual	0	28	9	3
	1	14	16	10
	2	2	8	30

Tabel 4 menunjukkan *confusion matrix* dengan menggunakan fitur *saturation* dalam pengklasifikasian saat aktual 0 mendapatkan nilai 'True' saat prediksi 28 dan salah prediksi berjumlah

12, aktual 1 mendapatkan nilai 'True' =14 dan salah prediksi=24, sedangkan aktual 2 mendapatkan nilai 'True'= 30 data dan salah prediksi=10. Berikut perhitungan dari Tabel 4.

$$\begin{aligned}
 accuracy_{3 \times 3} &= \frac{P_{00} + P_{11} + P_{22}}{P_{00} + P_{10} + P_{20} + P_{01} + P_{11} + P_{21} + P_{02} + P_{12} + P_{22}} \\
 &= \frac{28 + 9 + 3 + 14 + 16 + 10 + 2 + 8 + 30}{74} \\
 &= \frac{120}{120} \\
 &= 0,6166
 \end{aligned}$$

Data tersebut mendapatkan hasil akurasi sebesar 0,6166 atau 61,66%.

**Tabel 5. Confusion Matrix 2 Fitur Terbaik dengan k=9**

		Prediksi		
		0	1	2
Aktual	0	33	4	3
	1	15	18	7
	2	8	3	29

Tabel 5 menunjukkan *confusion matrix* dengan menggunakan fitur *value* dan *saturation* dalam pengklasifikasian saat aktual 0 mendapatkan nilai 'True' saat prediksi 33 dan salah prediksi berjumlah 7, aktual 1 mendapatkan nilai 'True' =18 dan salah prediksi=22, sedangkan aktual 2 mendapatkan nilai 'True'= 29 data dan salah prediksi=11. Berikut perhitungan dari Tabel 5.

$$\begin{aligned}
 accuracy_{3 \times 3} &= \frac{P_{00} + P_{11} + P_{22}}{P_{00} + P_{10} + P_{20} + P_{01} + P_{11} + P_{21} + P_{02} + P_{12} + P_{22}} \\
 &= \frac{33 + 4 + 3 + 15 + 18 + 7 + 8 + 3 + 29}{80} \\
 &= \frac{120}{120} \\
 &= 0,6666
 \end{aligned}$$

Data tersebut mendapatkan hasil akurasi sebesar 0,6666 atau 66,66%.

**Tabel 6. Confusion Matrix 3 Fitur Terbaik dengan k=5**

		Prediksi		
		0	1	2
Aktual	0	32	5	3
	1	14	19	7
	2	5	1	34

Tabel 6 menunjukkan *confusion matrix* dengan menggunakan fitur *hue*, *value*, dan *saturation* dalam pengklasifikasian saat aktual 0 mendapatkan nilai 'True' saat prediksi 32 dan salah prediksi berjumlah 8, aktual 1 mendapatkan nilai 'True' =19 dan 'False'=21, sedangkan aktual 2 mendapatkan nilai 'True'= 34 data dan 'False'=11. Berikut perhitungan dari Tabel 6.

$$\begin{aligned}
 accuracy_{3 \times 3} &= \frac{P_{00} + P_{11} + P_{22}}{P_{00} + P_{10} + P_{20} + P_{01} + P_{11} + P_{21} + P_{02} + P_{12} + P_{22}} \\
 &= \frac{32 + 5 + 3 + 14 + 19 + 7 + 5 + 1 + 34}{85} \\
 &= \frac{120}{120} \\
 &= 0,7083
 \end{aligned}$$

Data tersebut mendapatkan hasil akurasi sebesar 0,7083 atau 70,83%

**Tabel 7. Confusion Matrix 4 Fitur Terbaik dengan k=1**

		Prediksi		
		0	1	2
Aktual	0	32	8	0
	1	8	25	7
	2	1	3	36

Tabel 7 menunjukkan *confusion matrix* dengan menggunakan fitur *hue*, *value*, *saturation*, dan *h7* dalam pengklasifikasian saat aktual 0 mendapatkan nilai 'True' saat prediksi 32 dan 'False'= 8, aktual 1 mendapatkan nilai 'True' =25 dan 'False'=25, sedangkan aktual 2 mendapatkan nilai 'True'= 36 data dan 'False'=4. Berikut perhitungan dari Tabel 7.

$$\begin{aligned}
 accuracy_{3 \times 3} &= \frac{P_{00} + P_{11} + P_{22}}{P_{00} + P_{10} + P_{20} + P_{01} + P_{11} + P_{21} + P_{02} + P_{12} + P_{22}} \\
 &= \frac{32 + 8 + 0 + 8 + 25 + 7 + 1 + 3 + 36}{93} \\
 &= \frac{120}{120} \\
 &= 0,775
 \end{aligned}$$

Data tersebut mendapatkan hasil akurasi sebesar 0,775 atau 77,5%.

**Tabel 8. Confusion Matrix 5 Fitur Terbaik dengan k=1**

		Prediksi		
		0	1	2
Aktual	0	34	6	0
	1	7	27	6
	2	1	3	36

Tabel 8 menunjukkan *confusion matrix* dengan menggunakan fitur *hue*, *value*, *saturation*, *h2*, dan *h7* dalam pengklasifikasian saat aktual 0 mendapatkan nilai 'True' saat prediksi 34 'False'= 6, aktual 1 mendapatkan nilai 'True' =27 dan 'False'=13, sedangkan aktual 2 mendapatkan nilai 'True'= 36 data dan 'False'=4. Berikut perhitungan dari Tabel 8.

$$\begin{aligned}
 accuracy_{3 \times 3} &= \frac{P_{00} + P_{11} + P_{22}}{P_{00} + P_{10} + P_{20} + P_{01} + P_{11} + P_{21} + P_{02} + P_{12} + P_{22}} \\
 &= \frac{34 + 6 + 0 + 7 + 27 + 6 + 1 + 3 + 36}{97} \\
 &= \frac{120}{120} \\
 &= 0,8083
 \end{aligned}$$

Data tersebut mendapatkan hasil akurasi sebesar 0,8083 atau 80,83%.

**Tabel 9. Confusion Matrix 6 Fitur Terbaik dengan k=1**

		Prediksi		
		0	1	2
Aktual	0	34	6	0
	1	7	28	5
	2	1	3	36

Tabel 9 menunjukkan *confusion matrix* dengan menggunakan fitur *hue*, *value*, *saturation*, *h2*, *h3*, dan *h7* dalam pengklasifikasian saat aktual 0 mendapatkan nilai 'True' saat prediksi 34 dan 'False' berjumlah 6, aktual 1 mendapatkan nilai 'True' =28 dan 'False' =12, sedangkan aktual 2 mendapatkan nilai 'True'= 36 data dan 'False'=4. Berikut perhitungan dari Tabel 9.

$$\begin{aligned}
 accuracy_{3 \times 3} &= \frac{P_{00} + P_{11} + P_{22}}{P_{00} + P_{10} + P_{20} + P_{01} + P_{11} + P_{21} + P_{02} + P_{12} + P_{22}} \\
 &= \frac{34 + 6 + 0 + 7 + 28 + 5 + 1 + 3 + 36}{98} \\
 &= \frac{120}{120} \\
 &= 0,8166
 \end{aligned}$$

Data tersebut mendapatkan hasil akurasi sebesar 0,8166 atau 81,66%.

**Tabel 10. Confusion Matrix 7 Fitur Terbaik dengan k=1**

		Prediksi		
		0	1	2
Aktual	0	33	7	0
	1	7	29	4
	2	2	4	34

Tabel 10 menunjukkan *confusion matrix* dengan menggunakan fitur *hue, value, saturation, h1, h2, h3 dan h7* dalam pengklasifikasian saat aktual 0 mendapatkan nilai 'True' saat prediksi 33 dan 'False' berjumlah 7, aktual 1 mendapatkan nilai 'True' =29 dan 'False'=11, sedangkan aktual 2 mendapatkan nilai 'True'= 34 data dan 'False'=6. Berikut perhitungan dari Tabel 10.

$$\begin{aligned}
 accuracy_{3 \times 3} &= \frac{P_{00} + P_{11} + P_{22}}{P_{00} + P_{10} + P_{20} + P_{01} + P_{11} + P_{21} + P_{02} + P_{12} + P_{22}} \\
 &= \frac{33 + 29 + 34}{33 + 7 + 0 + 7 + 29 + 4 + 2 + 4 + 34} \\
 &= \frac{96}{120} \\
 &= 0,8
 \end{aligned}$$

Data tersebut mendapatkan hasil akurasi sebesar 0,8 atau 80%.

**Tabel 11. Confusion Matrix 8 Fitur Terbaik dengan k=1**

		Prediksi		
		0	1	2
Aktual	0	33	7	0
	1	8	26	6
	2	1	6	33

Tabel 11 menunjukkan *confusion matrix* dengan menggunakan fitur *hue, value, saturation, h1, h2, h3, h6 dan h7* dalam pengklasifikasian saat aktual 0 mendapatkan nilai 'True' saat prediksi 33 dan 'False' berjumlah 7, aktual 1 mendapatkan nilai 'True' =26 dan salah prediksi=14, sedangkan aktual 2 mendapatkan nilai 'True'= 33 data dan 'False'=7. Berikut perhitungan dari Tabel 11.

$$\begin{aligned}
 accuracy_{3 \times 3} &= \frac{P_{00} + P_{11} + P_{22}}{P_{00} + P_{10} + P_{20} + P_{01} + P_{11} + P_{21} + P_{02} + P_{12} + P_{22}} \\
 &= \frac{33 + 26 + 33}{33 + 7 + 0 + 8 + 26 + 6 + 1 + 6 + 33} \\
 &= \frac{92}{120} \\
 &= 0,7666
 \end{aligned}$$

Data tersebut mendapatkan hasil akurasi sebesar 0,7666 atau 76,66%.

**Tabel 12. Confusion Matrix 9 Fitur Terbaik dengan k=1**

		Prediksi		
		0	1	2
Aktual	0	33	7	0
	1	10	24	6
	2	1	8	31

Tabel 12 menunjukkan *confusion matrix* dengan menggunakan fitur *hue, value, saturation, h1, h2, h3, h4, h6 dan h7* dalam pengklasifikasian saat aktual 0 mendapatkan nilai 'True' saat prediksi 33 dan 'False' berjumlah 7, aktual 1 mendapatkan nilai 'True' =24 dan 'False'=16, sedangkan aktual 2 mendapatkan nilai 'True'= 31 data dan 'False'=9. Berikut perhitungan dari Tabel 12.

$$\begin{aligned}
 accuracy_{3 \times 3} &= \frac{P_{00} + P_{11} + P_{22}}{P_{00} + P_{10} + P_{20} + P_{01} + P_{11} + P_{21} + P_{02} + P_{12} + P_{22}} \\
 &= \frac{33 + 24 + 31}{33 + 7 + 0 + 10 + 24 + 6 + 1 + 8 + 31} \\
 &= \frac{88}{120} \\
 &= 0,7333
 \end{aligned}$$

Data tersebut mendapatkan hasil akurasi sebesar 0,7333 atau 73,33%.

**Tabel 13. Confusion Matrix 10 Fitur Terbaik dengan k=1**

		Prediksi		
		0	1	2
Aktual	0	33	7	0
	1	9	23	8
	2	2	6	32

Tabel 13 menunjukkan *confusion matrix* dengan menggunakan fitur *hue*, *value*, *saturation*, h1, h2, h3, h4, h5, h6 dan h7 dalam pengklasifikasian saat aktual 0 mendapatkan nilai 'True' saat prediksi 33 dan 'False' berjumlah 7, aktual 1 mendapatkan nilai 'True' =23 dan 'False'=17, sedangkan aktual 2 mendapatkan nilai 'True'= 32 data dan 'False'=8. Berikut perhitungan dari Tabel 13.

$$\begin{aligned}
 accuracy_{3 \times 3} &= \frac{P_{00} + P_{11} + P_{22}}{P_{00} + P_{10} + P_{20} + P_{01} + P_{11} + P_{21} + P_{02} + P_{12} + P_{22}} \\
 &= \frac{33 + 23 + 32}{33 + 7 + 0 + 9 + 23 + 8 + 2 + 6 + 32} \\
 &= \frac{88}{120} \\
 &= 0,7333
 \end{aligned}$$

Data tersebut mendapatkan hasil akurasi sebesar 0,7333 atau 73,33%.

Tabel 4 hingga Tabel 13 memperlihatkan kesalahan prediksi atau bernilai 'False' pada kelas 0 dengan prediksi 1, kelas 1 dengan prediksi kesalahan pada kelas 1, sedangkan kelas 2 sering terjadi salah prediksi di 1. Kesimpulan yang bisa didapatkan bahwa kelas 0 atau *bacterial leaf blight* memiliki kemiripan dengan kelas *brown spot*. Kelas *brown spot* memiliki kemiripan dengan *bacterial leaf blight*, sedangkan kelas *leaf smut* memiliki kemiripan ke brown smut. Hal ini dikarenakan nilai fitur memiliki kemiripan. Selain itu, kelas 1 dapat dilihat di tabel memiliki prediksi yang kecil dibandingkan dengan kelas lainnya. Berikut adalah tabel yang memperlihatkan hasil dari *Confusion matrix* dengan variasi fitur menggunakan k=1 hingga 10.

**Tabel 14. Hasil Confusion Matrix Dengan Variasi Fitur**

k-	Jumlah Fitur Terbaik									
	1	2	3	4	5	6	7	8	9	10
1	46,66%	62,49%	<b>70,83%</b>	<b>77,5%</b>	<b>80,83%</b>	<b>81,66%</b>	<b>80%</b>	<b>76,66%</b>	<b>73,335</b>	<b>73,33%</b>
2	41,66%	63,33%	68,33%	68,33%	69,16%	70,83%	69,16%	67,5%	70%	67,5%
3	56,66%	64,16%	69,16%	67,49%	67,5%	70,83%	71,66%	73,33%	70,83%	70,83%
4	55,83%	64,16%	67,5%	60,83%	61,66%	68,33%	67,5%	68,33%	68,33%	69,16%
5	57,5%	60,83%	<b>70,83%</b>	63,33%	66,66%	70,83%	71,66%	68,33%	66,66%	67,5%
6	58,33%	63,33%	69,16%	62,5%	61,66%	65,83%	68,33%	66,66%	65%	67,5%
7	60,83%	58,33%	69,16%	66,66%	67,5%	71,66%	70,83%	65,83%	69,16%	68,33%
8	60%	65,83%	66,66%	62,5%	64,16%	68,33%	64,16%	63,33%	64,16%	65%
9	60,83%	<b>66,66%</b>	69,16%	65%	64,16%	67,5%	65%	65%	66,66%	67,5%
10	<b>61,66%</b>	65,83%	67,5%	63,33%	65%	67,5%	64,16%	65%	68,33%	65,83%

Hasil klasifikasi menggunakan *confusion matrix* dengan variasi fitur ditunjukkan akurasi pada Tabel 14. Hasil *Confusion matrix* menggunakan variasi fiturnya yang menghasilkan akurasi yang berbeda nilai k yang digunakan. Penelitian akan mencari nilai akurasi terbaik berdasarkan Tabel 14. Berdasarkan beberapa variasi fitur yang digunakan dalam pengklasifikasian penyakit daun padi menggunakan KNN terlihat bahwa untuk tingkat akurasi tertinggi berada pada k=1 dengan menggunakan 6 fitur terbaik. Hasil akurasi tiap variasi fitur memiliki perbedaan. Tiap nilai akurasi tergantung pada fitur mana yang digunakan. Tingkat akurasi terbaik pada penelitian ini yaitu sebesar 0,8166 atau 81,66%. Dari hasil tersebut tingkat akurasi dapat dikatakan baik.

## 5 Kesimpulan

Hasil klasifikasi 3 jenis penyakit daun padi dengan fitur warna (*Hue*, *Saturation*, *Value*) dan *invariant moment* (H1-H7) serta KNN sebagai pengklasifikasi. Berdasarkan hasil penelitian menunjukkan fitur terbaik yang dapat digunakan yaitu *hue*, *saturation*, *value*, h2, h3, dan h7. Sedangkan jumlah tetangga (k) pada KNN yang terbaik didapatkan nilai k=1 dengan nilai akurasi

sebesar 81,66%. Saran untuk penelitian selanjutnya yaitu penggunaan beberapa jenis ekstraksi fitur dan pengklasifikasi lainnya sehingga diharapkan dapat meningkatkan hasil akurasi.

## Referensi

- [1] B. Pujiasmanto, Sutarno, N. Suharyana, and Riyatun, *Padi Hitam: Manfaat, Resep Makanan Beras Hitam, dan Riset Padi Hitam yang Diradiasi Sinar Gamma*. Medan: Yayasan Kita Menulis, 2021.
- [2] Badan Pusat Statistik, “Luas Panen dan Produksi Padi di Indonesia 2021 (Angka Tetap),” *Berita Resmi Statistik*, 2021. [Online]. Available: <https://www.bps.go.id/pressrelease/2022/03/01/1909/produksi-padi-tahun-2021-turun-0-43-persen--angka-tetap-.html>
- [3] N. Selvia, “Diserang Hama dan Penyakit, Produksi Padi makin Menurun,” *Padek*, Padang, Feb. 02, 2021. Accessed: May 27, 2022. [Online]. Available: <https://padek.jawapos.com/sumbar/padang/02/02/2021/diserang-hama-dan-penyakit-produksi-padi-makin-menurun/>
- [4] D. Maulitasari and R. Passarella, *Teori dan Sejarah Citra Forensik*. Palembang: UPT Penerbit & Percetakan Universitas Sriwijaya, 2020.
- [5] A. Purnamawati, W. Nugroho, D. Putri, and W. F. Hidayat, “Deteksi Penyakit Daun pada Tanaman Padi Menggunakan Algoritma Decision Tree, Random Forest, Naïve Bayes, SVM dan KNN,” *InfoTekJar J. Nas. Inform. dan Teknol. Jar.*, vol. 5, no. 1, pp. 212–215, 2020, [Online]. Available: <https://doi.org/10.30743/infotekjar.v5i1.2934>
- [6] F. D. Marleny, *Mengenal Pengolahan Citra Digital menggunakan Python*. Purwokerto: CV. Pena Persada, 2021. [Online]. Available: <https://dosen.itats.ac.id/farida/2020/02/10/mengenal-pengolahan-citra-digital-menggunakan-python-yuuukkk/>
- [7] R. A. Saputra, S. Wasiyanti, A. Supriyatna, and D. F. Saefudin, “Penerapan Algoritma Convolutional Neural Network Dan Arsitektur MobileNet Pada Aplikasi Deteksi Penyakit Daun Padi,” *Swabumi*, vol. 9, no. 2, pp. 184–188, 2021, doi: 10.31294/swabumi.v9i2.11678.
- [8] J. Kusanti, K. Penyakit, D. Padi, and A. Haris, “Klasifikasi Penyakit Daun Padi Berdasarkan Hasil Ekstraksi Fitur GLCM Interval 4 Sudut,” *J. Inform. J. Pengemb. IT*, vol. 03, no. 01, pp. 1–6, 2018, doi: <http://dx.doi.org/10.30591/jpit.v3i1.669>.
- [9] U. N. Oktaviana, R. Hendrawan, A. D. K. Annas, and G. W. Wicaksono, “Klasifikasi Penyakit Padi berdasarkan Citra Daun Menggunakan Model Terlatih Resnet101,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 1216–1222, 2021, doi: 10.29207/resti.v5i2.3607.
- [10] E. Anggiratih, S. Siswanti, S. K. Octaviani, and A. Sari, “Klasifikasi Penyakit Tanaman Padi Menggunakan Model Deep Learning Efficientnet B3 dengan Transfer Learning,” *J. Ilm. SINUS*, vol. 19, no. 1, pp. 75–83, 2021, doi: 10.30646/sinus.v19i1.526.
- [11] Z. Zulkifli, “Sistem Pendeteksi Penyakit Tanaman Padi Berbasis Artificial Intelligence,” *J. Tika*, vol. 6, no. 03, pp. 260–269, 2021, doi: 10.51179/tika.v6i03.813.
- [12] M. Khoiruddin, A. Junaidi, and W. A. Saputra, “Klasifikasi Penyakit Daun Padi Menggunakan Convolutional Neural Network,” *J. Dinda Data Sci. Inf. Technol. Data Anal.*, vol. 2, no. 1, pp. 37–45, 2022, doi: 10.20895/dinda.v2i1.341.
- [13] H. B. Prajapati, J. P. Shah, and V. K. Dabhi, “Detection and Classification of Rice Plant Diseases,” *Intell. Decis. Technol.*, vol. 11, no. 3, pp. 357–373, 2017, doi: 10.3233/IDT-170301.
- [14] M. Fahmi Wibawa, M. A. Rahman, and A. W. Widodo, “Penerapan Ruang Warna HSV dan Ekstraksi Fitur Tekstur Local Binary Pattern untuk Tingkat Kematangan Sangrai Biji Kopi,” vol. 5, no. 7, pp. 2819–2825, 2021, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [15] R. R. Muhima and Farida, “Image Retrieval Batik Klasik Parang Rusak Menggunakan Ekstraksi Fitur Geometric Invariant,” *J. Ilm. NERO*, vol. 4, no. 1, pp. 15–21, 2018, doi: <http://dx.doi.org/10.21107/nero.v4i1.107>.
- [16] S. A. Nurhusni, R. I. Adam, and C. Carudin, “Klasifikasi Kadar Kolesterol Menggunakan Ekstraksi Ciri Moment Invariant dan Algoritma K-Nearest Neighbor (KNN),” *J. Appl. Informatics Comput.*, vol. 5, no. 2, pp. 169–175, 2021, doi: 10.30871/jaic.v5i2.3273.
- [17] L. Mardiana, D. Kusnandar, and N. Satyahadewi, “Analisis Diskriminan Dengan K Fold Cross Validation Untuk Klasifikasi Kualitas Air Di Kota Pontianak,” vol. 11, no. 1, pp. 97–102, 2022, doi: <http://dx.doi.org/10.26418/bbimst.v11i1.51608>.

- [18] Napitupulu, “Klasifikasi Penyakit Pada Daun Tomat dengan Algoritma K-Nearest Neighbor (K-NN) Berdasarkan Ekstraksi Fitur Warna dan Invariant Moment,” *J. Pembang. Wil. Kota*, vol. 1, no. 3, pp. 82–91, 2017.
- [19] D. S. Wahyuni, “Penerapan Algoritma K-Nearest Neighbor Untuk Prediksi Harga Cabai Rawit Di Yogyakarta,” Sanata Dharma University Yogyakarta, 2020.
- [20] H. Leidiyana, “Penerapan Metode K-Nearest Neighbor Pada Penentuan Grade Dealer,” *J. Pengetah. Dan Teknol. Komput.*, vol. 2, no. 2, pp. 108–112, 2017.
- [21] F. Kusumah, Nurjaidin, and M. Ardhiansyah, *ANALISIS SISTEM PENDETEKSI WAJAH PADA GAMBAR DENGAN METODE K-NEAREST NEIGHBOR*. Tangerang: Pascal Books, 2021. Accessed: Jan. 24, 2023. [Online]. Available: [https://www.google.co.id/books/edition/ANALISIS\\_SISTEM\\_PENDETEKSI\\_WAJAH\\_PADA\\_GAMBAR\\_DENGAN\\_METODE\\_K-NEAREST\\_NEIGHBOR/GA/y1xpEAAAQBAJ?hl=id&gbpv=1&dq=kNN+jarak&pg=PA29&printsec=frontcover](https://www.google.co.id/books/edition/ANALISIS_SISTEM_PENDETEKSI_WAJAH_PADA_GAMBAR_DENGAN_METODE_K-NEAREST_NEIGHBOR/GA/y1xpEAAAQBAJ?hl=id&gbpv=1&dq=kNN+jarak&pg=PA29&printsec=frontcover)
- [22] X. Deng, Q. Liu, Y. Deng, and S. Mahadevan, “An Improved Method to Construct Basic Probability Assignment Based on The Confusion Matrix for Classification Problem,” *Inf. Sci. (Ny)*, vol. 340–341, pp. 250–261, 2016, doi: 10.1016/j.ins.2016.01.033.