

IDENTIFIKASI JAMUR BERACUN DAN TIDAK BERACUN DENGAN K-NEAREST NEIGHBOR

MATA KULIAH MACHINE LEARNING & TEKNIK PENGOLAHAN CITRA

16523228 Arif Sulaksana Putra

16523156 Khoirunnisa Shofwatul Laili

ABSTRAKSI

Teknologi *Machine Learning* dan pengolahan citra digital telah banyak digunakan dalam berbagai penelitian untuk dapat melakukan identifikasi dan klasifikasi. Salah satu jenis metode yang dinilai cepat dengan akurasi yang tinggi adalah KNN atau *K-Nearest Neighbor*. Penelitian ini mencoba menerapkan metode KNN dalam melakukan klasifikasi jenis-jenis jamur yang beracun dan tidak beracun melalui gambar. Hasil yang didapatkan adalah metode ini dapat melakukan klasifikasi dengan tingkat akurasi sebesar 67.78 persen.

BAB 1 PENDAHULUAN

1.1 LATAR BELAKANG

Kata “Jamur” sudah tidak asing lagi bagi kita. Jamur adalah salah satu jenis tumbuhan yang tidak mandiri karena kehidupannya selalu bergantung pada organisme lain sehingga disebut tumbuhan heterotrofik (Wiardani, 2010). Jamur adalah organisme yang selnya berinti, dapat membentuk spora, tidak berklorofil, dan berupa benang-benang tunggal atau benang-benang yang bercabang dengan dinding dari *selulosa* atau *khitin* atau keduanya (Suarnadwipa & Hendra, 2008).

Jika dipandang dari sudut nilai konsumsi, jamur dapat dibagi menjadi dua yaitu jamur konsumsi dan jamur beracun. Dari ribuan jamur yang telah diidentifikasi, hanya terdapat tidak lebih dari 50 jenis jamur yang dapat dikonsumsi. Bahkan beberapa jenis jamur tergolong dalam jamur yang sangat beracun. Jenis jamur yang dapat dikonsumsi antara lain, jamur shiitake (*Lentinusedodes*), jamur kuping (*Auricularia sp.*), jamur tiram (*Pleurotus ostreatus*), jamur kancing (*Agaricus bisporus*), dan lainnya (Agromedia Pustaka, 2002).

Laporan tentang kesalahan konsumsi jamur yang parah dan fatal telah meningkat di seluruh dunia sejak tahun 1950-an (Diaz, 2005). Hal ini disebabkan karena masih minimnya pengetahuan masyarakat mengenai jamur mana yang bisa dikonsumsi dan jamur mana yang beracun. Oleh karena itu perlu sistem klasifikasi jamur yang dapat membedakan jamur mana saja yang bisa dikonsumsi dan yang tidak bisa dikonsumsi.

Teknologi visi komputer dan *machine learning* telah menghadirkan solusi untuk mempermudah manusia dalam proses klasifikasi tanaman hidup berdasarkan ciri-ciri fisiknya. Riska dan Rosadi (2015), menggabungkan metode *Unconstraint Hit or Miss Transformation*

(UHMT) dengan empat *structuring element* (SE) untuk mengklasifikasikan tanaman mangga berdasarkan struktur daun. Metode NBC (Naive Bayes Classification) dan KNN (K-Nearest Neighbor) juga diterapkan untuk klasifikasi tanaman herbal berdasarkan bentuk citra daun oleh Liantoni dan Nugroho (2015). Sedangkan metode terbaik untuk klasifikasi menurut Riska dan Rosadi (2005) adalah *dynamic architecture for artificial neural networks* (DAN2), dan *support vector machines* (SVM).

Klasifikasi jenis-jenis jamur dilakukan dengan mengidentifikasi bentuk citra jamur. Proses klasifikasi tanaman jamur memerlukan ekstraksi fitur dari bentuk jamur tersebut. Ada banyak metode yang bisa digunakan untuk melakukan klasifikasi berdasarkan fitur bentuk yang ada pada tanaman. Pada penelitian ini, digunakan metode *K-Nearest Neighbor* (KNN) untuk melakukan proses pengelompokan jenis-jenis jamur. Metode ini dianggap sebagai metode yang bisa melakukan klasifikasi dengan cepat dan *simple*. Dengan menerapkan metode ini, diharapkan proses identifikasi jenis-jenis jamur beracun dan tidak beracun dapat dilakukan dengan lebih cepat dan mudah.

1.2 RUMUSAN MASALAH

Dari latar belakang di atas, didapatkan rumusan masalah:

1. Bagaimana cara mengidentifikasi jamur yang beracun dan tidak beracun?
2. Bagaimana algoritma *K-Nearest Neighbor* (KNN) yang dapat melakukan klasifikasi jenis-jenis jamur?
3. Bagaimana tingkat akurasi metode *K-Nearest Neighbor* (KNN) yang digunakan untuk klasifikasi jamur?

BAB 2 KAJIAN PUSTAKA

2.1 K-NEAREST NEIGHBOR

Untuk melakukan klasifikasi, dapat menggunakan metode K-Nearest Neighbor. Tahap-tahap penggunaan K-Nearest Neighbor adalah sebagai berikut :

1. Menentukan Nilai k
Tahap pertama adalah menentukan nilai k. Penentuan nilai k dapat ditentukan secara bebas.
2. Menghitung jarak
Tahap kedua yaitu menghitung jarak. Rumus jarak yang digunakan adalah *Euclidean Distance*. Rumus menghitung *Euclidean Distance* dinyatakan sebagai berikut.

$$D(x, y) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2}$$

D adalah jarak antara data latih (x) dan data uji (y) yang akan diklasifikasikan.

3. Menentukan Nilai Nearest Neighbor Value

Tahap ketiga adalah menentukan nilai tetangga terdekat dari tiap-tiap jarak. Hasil yang didapat dari penghitungan ini berupa kategori kelas yang akan dijadikan acuan ketika pengklasifikasian berdasarkan nilai tetangga terdekat.

4. Membuat Klasifikasi Berdasarkan Nearest Neighbor Mayoritas

Tahap terakhir adalah membuat klasifikasi data uji dapat dilakukan berdasarkan kelas mayoritas. Misalnya jika mayoritas nilai tetangga terdekat bernilai kelas “jamur beracun” maka kelas data uji yang belum diketahui termasuk dalam kelas “jamur beracun”.

2.2 COLOR HISTOGRAM

Histogram Warna atau *color histogram* adalah representasi distribusi warna dalam sebuah gambar yang didapatkan dengan menghitung jumlah pixel dari setiap bagian *range* warna, secara tipikal dalam dua dimensi atau tiga dimensi. Dalam pembuatan *histogram*, nilai RGB mempunyai *range* dari 0 sampai 255 akan punya kombinasi warna sebesar 16777216 (didapatkan dari : 255 x 255 x 255) (Kusumaningtyas & Asmara, 2016).

2.3 GRAY LEVEL COCCURENCE MATRICES (GLCM)

GLCM adalah matriks derajat keabuan yang merepresentasikan hubungan suatu derajat keabuan dengan derajat keabuan lain. GLCM merupakan dasar dari teknik tekstur Haralick, GLCM digunakan untuk analisis pasangan piksel yang bersebelahan tergantung dengan sudut yang digunakan (Wibowo, 2016). Adapun 5 persamaan tekstur fitur yang digunakan dalam perancangan sistem ini , yaitu:

a. *Angular Second Moment*

Menunjukkan ukuran sifat homogenitas citra.

$$ASM = \sum_{i=1}^L \sum_{j=1}^L \{GLCM(i,j)\}^2 \quad (1)$$

b. *Contrast*

Kontras adalah perhitungan perbedaan intensitas antara piksel satu dan piksel yang berdekatan diseluruh gambar.

$$Contrast = \sum_{i=1}^L n^2 \{ \sum_{|i-j|} GLCM(i,j) \} \quad (2)$$

c. *Correlation*

Menunjukkan ukuran ketergantungan linear derajat keabuan citra sehingga dapat memberikan petunjuk adanya struktur linear dalam citra.

$$Correlation = \frac{\sum_{i=1}^L \sum_{j=1}^L GLCM(i,j) - \mu_i' \mu_j'}{\sigma_i' \sigma_j'} \quad (3)$$

d. Inverse Different Moment

Menunjukkan kehomogenan citra yang berderajat keabuan sejenis.

$$IDM = \sum_{i=1}^L \sum_{j=1}^L (GLCM(i,j) \log GLCM(i,j)) \quad (4)$$

e. Entropy

Menunjukkan ukuran ketidakaturan bentuk.

$$Entropy = - \sum_{i=1}^L \sum_{j=1}^L (GLCM(i,j) \log GLCM(i,j)) \quad (5)$$

2.4 HU MOMENT

Ekstraksi ciri orde pertama merupakan metode ekstraksi ciri yang didasarkan pada karakteristik nilai *gray level* citra.. Sehingga dapat dihitung beberapa parameter ciri orde pertama, antara nilai *phi* yang terdiri dari tujuh ciri $\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6$, dan ϕ_7 . *Moment* yang dihasilkan dapat digunakan untuk menangani translasi, penskalaan, dan rotasi gambar. Penciptanya Hu, menciptakan tujuh moment invariant yang di dalam bukunya (Kadir dan Adhi, 2012) seperti berikut:

$$\phi_1 = \pi_{20} + \pi_{02} \quad (1)$$

$$\phi_2 = (\pi_{20} - \pi_{02})^2 + (2\pi_{02})^2 \quad (2)$$

$$\phi_3 = (\pi_{30} - 3\pi_{12})^2 + (\pi_{03} - 3\pi_{21})^2 \quad (3)$$

$$\phi_4 = (\pi_{30} - \pi_{12})^2 + (\pi_{03} + \pi_{21})^2 \quad (4)$$

$$\phi_5 = (\pi_{30} - 3\pi_{12})(\pi_{30} - \pi_{12})[(\pi_{30} + \pi_{12})^2 - 3(\pi_{21} + \pi_{03})^2] +$$

$$(\pi_{03} - 3\pi_{12})(\pi_{03} + \pi_{21})[\pi_{03} + \pi_{12})^2 - 3(\pi_{12} + \pi_{30})^2] \quad (5)$$

$$\phi_6 = (\pi_{20} - \pi_{02})[(\pi_{30} + \pi_{12}) - (\pi_{21} + \pi_{03})] +$$

$$4\pi_{11}(\pi_{30} + \pi_{12})(\pi_{03} + \pi_{21}) \quad (6)$$

$$\phi_7 = (3\pi_{21} - \pi_{03})(\pi_{30} + \pi_{12})[(\pi_{30} + \pi_{12})^2 - 3(\pi_{21} + \pi_{03})^2] +$$

$$(\pi_{30} - 3\pi_{12})(\pi_{21} + \pi_{03})[\pi_{03} + \pi_{21})^2 - 3(\pi_{30} + \pi_{12})^2] \quad (7)$$

BAB 3 METODOLOGI PENELITIAN

3.1 AKUISISI CITRA

Proses akuisisi citra adalah proses untuk memperoleh data yang akan diteliti. Penelitian ini menggunakan citra jamur yang didapatkan melalui situs *www.kaggle.com*. Sedangkan jumlah dataset yang digunakan adalah sebanyak 4.341 data citra jamur dalam format JPG. Kemudian, dataset yang didapatkan akan dibagi menjadi data latih dan data uji.

3.2 PRE-PROCESSING CITRA

Pre-processing merupakan tahap yang dilakukan untuk mempersiapkan citra yang masih kasar sehingga dapat diolah lebih lanjut (Wahyuningrum, et. al., 2017). Pre-processing bertujuan untuk memperbaiki kualitas citra sebelum ekstraksi ciri dan diklasifikasi. Citra yang digunakan akan diubah menjadi citra dengan aras keabuan. Proses ini berfungsi untuk mereduksi citra tiga dimensi menjadi satu dimensi saja dengan nilai intensitas yang sama, sehingga dapat mempercepat proses komputasi (Wahyuningrum, et. al., 2017). Kemudian, citra juga bisa diubah menjadi citra biner apabila diperlukan.

3.3 EKSTRAKSI CIRI

Ekstraksi ciri adalah proses untuk mendapatkan ciri khusus pada suatu citra. Proses ekstraksi dilakukan dengan melakukan deteksi tepi untuk mendapatkan bentuk citra, dan mengukur kedalaman warna untuk mendapatkan warna citra. Hasil yang diperoleh akan digunakan untuk menentukan jenis-jenis jamur berdasarkan ciri yang didapatkan untuk kemudian dikelompokkan menjadi jamur yang layak untuk dimakan ataupun jamur yang mengandung racun.

3.4 KLASIFIKASI

Proses klasifikasi bertujuan untuk menentukan kelompok jenis-jenis jamur supaya jamur dapat dikenali sebagai jamur yang layak makan atau beracun. Klasifikasi pada penelitian ini dilakukan dengan metode *K-Nearest Neighbor* (KNN).

3.5 PENGUKURAN TINGKAT AKURASI

Parameter yang diamati untuk mengetahui performansi sistem adalah akurasi dan waktu komputasi (Wahyuningrum, et. al., 2017).

BAB 4 IMPLEMENTASI

4.1 OBJEK PENELITIAN

Objek yang digunakan dalam penelitian ini adalah citra jamur yang didapatkan dari situs *www.kaggle.com*.



Gambar 1. *Objek citra jamur*

4.2 TAHAP PENELITIAN

Tahap-tahap penelitian dijelaskan dalam diagram alir berikut ini.



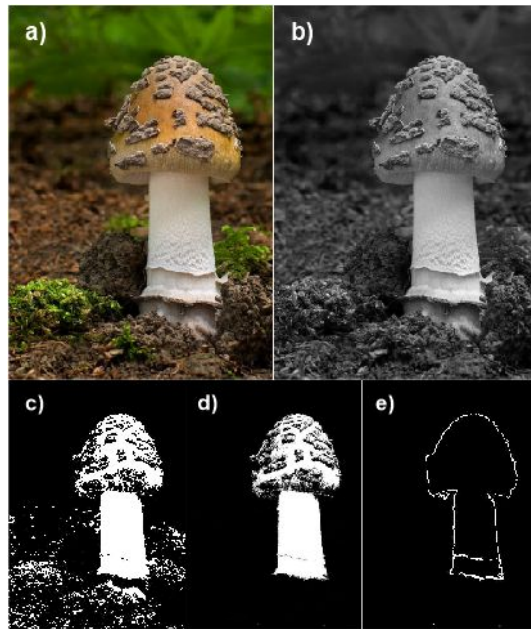
Gambar 2. *Diagram alir tahap-tahap penelitian*

a) Akuisisi Citra

Citra jamur yang akan digunakan berjumlah 4.341 dengan campuran kelas jamur beracun dan tidak beracun. Kemudian, data tersebut dibagi menjadi data latih sebanyak 75% dan data uji sebanyak 25%.

b) *Pre-processing*

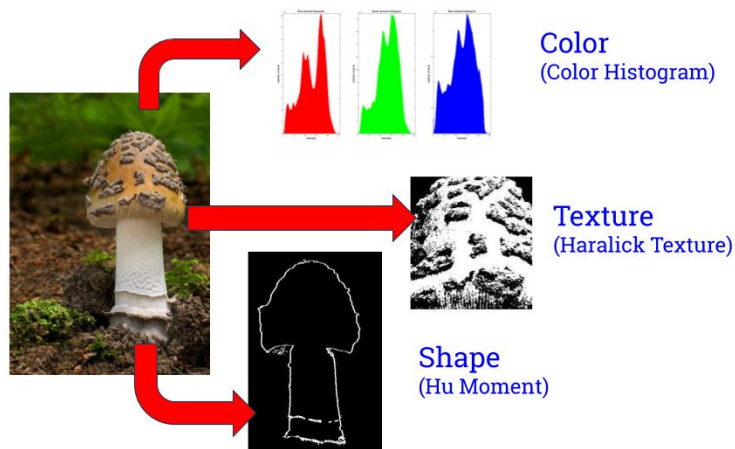
Proses ini dilakukan dalam beberapa tahapan, yaitu perubahan citra menjadi citra aras keabuan, perubahan menjadi citra biner, pemisahan objek, kemudian deteksi tepi.



Gambar 3. *Pre-processing. a) Kondisi citra awal. b) Perubahan menjadi citra aras keabuan. c) Perubahan menjadi citra biner. d) Pemisahan objek / segmentasi. e) Deteksi tepi*

c) Ekstraksi Ciri

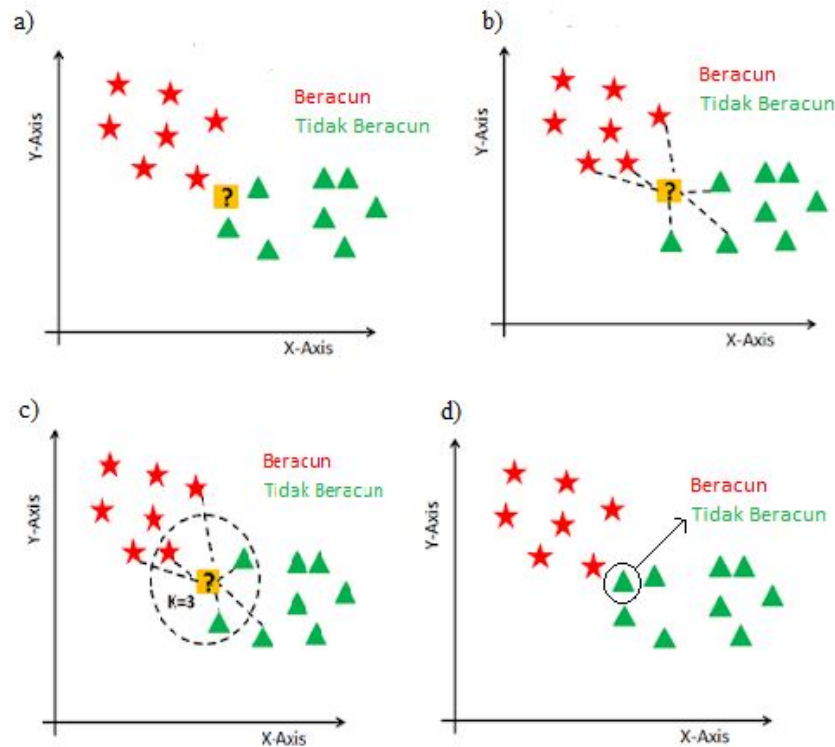
Proses ini bertujuan untuk mendapatkan ciri khusus pada objek. Terdapat tiga ciri yang perlu diekstrak untuk memperoleh klasifikasi jenis jamur beracun dan tidak beracun yaitu warna, bentuk dan tekstur. Kami menggunakan metode *Color Histogram* untuk mengekstrak ciri warna, metode *Haralick Texture* dari *Grey Level Cooccurrence Matrices*(GLCM) untuk mengekstrak ciri texture dan metode *Hu Moment* untuk mengekstrak ciri bentuk.



Gambar 4. *Proses ekstraksi ciri untuk mendapatkan ciri khusus objek*

d) Klasifikasi

Proses klasifikasi jamur beracun dan tidak beracun dilakukan dengan metode *K-Nearest Neighbor*. Metode ini menentukan kelompok jenis jamur dari citra jamur yang baru masuk dengan menghitung jarak terdekat.



Gambar 5. Klasifikasi dengan metode KNN. a) Menginisiasi data baru. b) Menghitung jarak. c) Menentukan Nearest Neighbor Value. d) Membuat Klasifikasi Berdasarkan Nearest Neighbor Mayoritas

e) Pengukuran tingkat akurasi

Proses evaluasi dilakukan untuk mengukur seberapa baik algoritma ini jika diterapkan. Pengukuran tingkat akurasi diterapkan dengan menggunakan confusion matriks.

	<i>Prediction Class</i>			
<i>Actual Class</i>		Yes	No	Total
	Yes	TP	FN	P
	No	FP	TN	N
	Total	P'	N'	P+N

Tabel 1. *Confusion Matrix*

Setelah menerapkan metode KNN dengan nilai k sama dengan 1, maka diperoleh *confusion matrix* sebagai berikut.

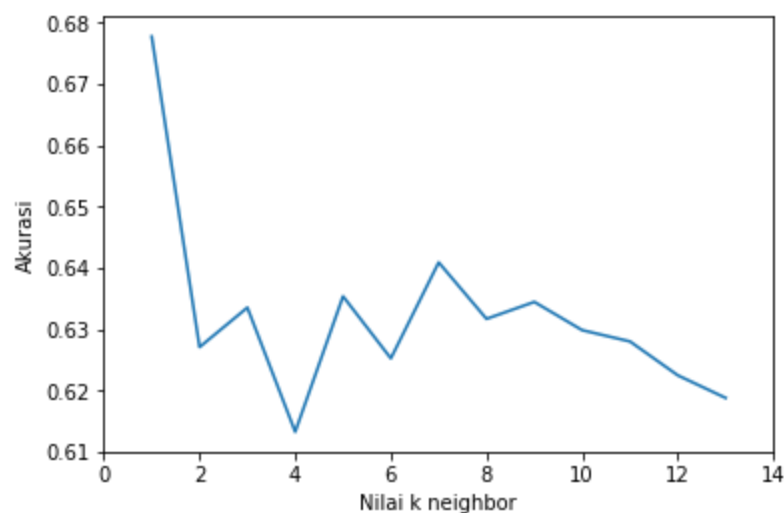
	<i>Prediction Class</i>				
<i>Actual Class</i>		Beracun	Tidak Beracun	Total	Recognition (%)
	Beracun	316	163	479	65.97 %
	Tidak Beracun	187	420	607	69.19 %
	Total	503	583	1089	

Tabel 2. *Confusion matrix* penerapan metode KNN dengan $k = 1$.

	Formula	Proses	Hasil (%)
Akurasi	$TP + TN / P + N$	$316 + 420 / 1086$	67.78 %
Error Rate	$FP + FN / P + N$	$187 + 163 / 1086$	32.23 %
Precision	$TP / TP + FP$	$316 / 503$	62.82 %
Recall	TP / P	$316 / 479$	65.97 %

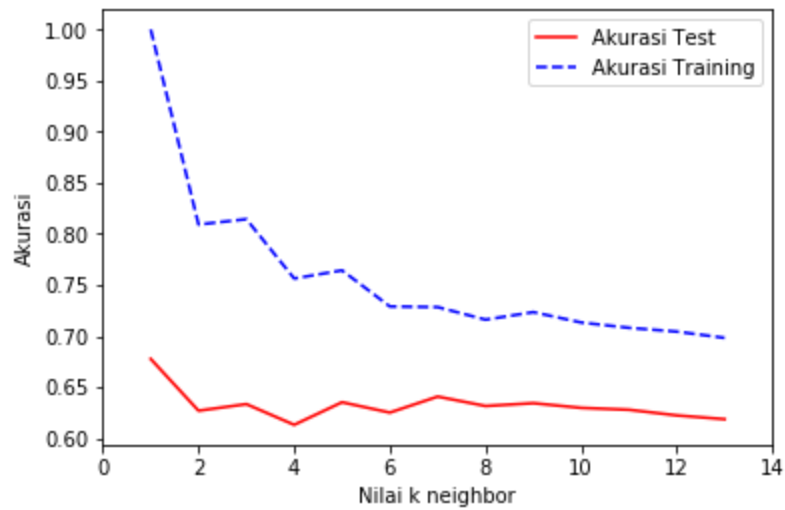
Tabel 3. *Pengukuran tingkat akurasi*

Metode KNN juga diterapkan dengan nilai k antara 1 sampai dengan 13. Hasil akurasi yang diperoleh ditampilkan dalam grafik sebagai berikut.



Gambar 6. *Tingkat akurasi metode KNN dengan nilai k antara 1 sampai dengan 13*

Berikut ini adalah grafik yang menunjukkan tingkat akurasi setelah metode diterapkan pada data latih dan data uji.



Gambar 7. Perbandingan tingkat akurasi antara data latih dan data uji.

KESIMPULAN

Setelah dilakukan penelitian ini, maka kami dapat menyimpulkan bahwa identifikasi jamur beracun dan tidak beracun melalui ekstraksi ciri warna, bentuk, dan tekstur dapat dilakukan dengan metode *K-Nearest Neighbor*. Dari implementasi di atas ternyata tingkat akurasi terbaik yang didapatkan melalui perhitungan *confusion matrix* dengan rentang nilai $k = 1$ hingga 13 adalah KNN dengan nilai $k = 1$ sebesar 67.78%.

REFERENSI

Wiardani, I. (2010). Budidaya Jamur Konsumsi. Penerbit Andi.

Suarnadwipa, N., & Hendra, W. (2008). Pengeringan jamur dengan dehumidifier. *Jurnal Ilmiah Teknik Mesin CAKRAM Vol*, 2(1), 30-33.

Agromedia Pustaka. (2002). Budi Daya Jamur Konsumsi. Agromedia Pustaka. Jakarta.

Diaz, J. H. (2005). Evolving global epidemiology, syndromic classification, general management, and prevention of unknown mushroom poisonings. *Critical care medicine*, 33(2), 419-426.

Liantoni, F., & Nugroho, H. (2015). Klasifikasi Daun Herbal Menggunakan Metode Naïve Bayes Classifier Dan Knearest Neighbor. *Jurnal Simantec*, 5(1).

Riska, S. Y., Cahyani, L., & Rosadi, M. I. (2015). Klasifikasi Jenis Tanaman Mangga Gadung dan Mangga Madu Berdasarkan Tulang Daun. *Jurnal Buana Informatika*, 6(1).

Wahyuningrum, R. D., Hidayat, B., & Darana, S. (2017). Deteksi Kualitas Keju Berdasarkan Segmentasi Warna Dan Tekstur Dengan Metoda Discrete Cosine Transform (dct) Dengan Klasifikasi K-Nearest Neighbor (k-nn) Pada Citra Digital. *eProceedings of Engineering*, 4(2).

Zubair, A., & Muslikh, A. R. (2017, September). IDENTIFIKASI JAMUR MENGGUNAKAN METODE K-NEAREST NEIGHBOR DENGAN EKSTRAKSI CIRI MORFOLOGI. In Seminar Nasional Sistem Informasi (SENASIF) (Vol. 1, No. 1, pp. 965-972).

Kusumaningtyas, S., & Asmara, R. A. (2016). Identifikasi Kematangan Buah Tomat Berdasarkan Warna Menggunakan Metode Jaringan Syaraf Tiruan (JST). *Jurnal Informatika Polinema*, 2(2), 72-72.

Kadir, Abdul., Susanto Adhi (2012), Teori dan Aplikasi Pengolahan Citra, C.V Andi Offset, Yogyakarta