

KLASIFIKASI JENIS IKAN BERBASIS JARINGAN SARAF TIRUAN MENGUNAKAN ALGORITMA PRINCIPAL COMPONENT ANALYSIS (PCA)

Arif Lumute Unihehu¹, Imam Suharjo²

^{1,2}Informatika, Universitas Mercu Buana Yogyakarta, Yogyakarta, Indonesia
¹arifunihehu@gmail.com, ²imam@mercubuana-yogya.ac.id

ABSTRAK

Ikan merupakan hewan berdarah dingin yang banyak di manfaatkan oleh manusia. Ikan adalah kelompok vertebrata poikilotermik yang beraneka ragam dengan jumlah spesies lebih dari 27,000 di seluruh dunia. Banyaknya jumlah spesies ikan menjadi satu masalah dalam membedakan jenis ikan. Tujuan dari penelitian ini adalah membuat sistem Klasifikasi jenis ikan berdasarkan tekstur citra ikan berbasis jaringan syaraf tiruan menggunakan algoritma K-Nearest Neighbors Dan Principal Component Analysis (PCA). Data diambil melalui eksplorasi langsung dan diambil langsung oleh peneliti. Data hanya menggunakan 3 jenis ikan sebagai objek penelitian selanjutnya dilakukan training dan pengujian data uji pada kelas satu, dua dan tiga hanya satu yang tidak dapat dikenali oleh sistem, sedangkan data lainnya dapat dikenali dengan presentase keberhasilan sebesar 93% (Sembilan puluh tiga persen).

Kata Kunci— Principal Component Analysis, KNN, Klasifikasi.Ikan.

ABSTRACT

Fish are cold-blooded animals that are widely used by humans. Fish are a diverse group of poikilothermic vertebrates with more than 27,000 species worldwide. A large number of fish species becomes a problem in distinguishing the types of fish. The purpose of this study was to create a fish type classification system based on the texture of artificial neural network-based fish imagery using K-Nearest Neighbors and Principal Component Analysis (PCA) algorithms. The data was taken through direct exploration and retrieved directly by researchers. The data only uses 3 types of fish as the object of further research conducted training and testing test data in the first, second, and third classes only one can not be recognized by the system, while the other data can be recognized by the percentage of success of 93% (Ninety-three percent).

Keywords— Principal Component Analysis, KNN, Classification, Fish.

1. PENDAHULUAN

Ikan merupakan hewan berdarah dingin yang banyak di manfaatkan oleh manusia. Ikan adalah kelompok vertebrata poikilotermik yang beraneka ragam dengan jumlah spesies lebih dari 27,000 di seluruh dunia. Banyaknya jumlah spesies ikan menjadi satu masalah dalam membedakan jenis ikan, tidak mungkin kita sempat untuk menghafal jenis ikan sebanyak itu, namun paling tidak kita mengetahui ciri – ciri ikan dan klasifikasinya. Permasalahan yang terjadi adalah secara fisik spesies laut atau ikan ini memiliki kemiripan yang hampir sama antara satu jenis dengan jenis yang lainnya, hal tersebut menjadi penyebab untuk manusia dalam mengenali spesies yang akan di di perjual belikan di pasar dagang.

Sebagai upaya mengatasi keterbatasan kemampuan manusia untuk mengenali jenis Ikan dan informasi tentang Ikan maka perlu dibuat suatu pengolahan citra berbasis jaringan saraf tiruan. Pengolahan citra berbasis jaringan saraf tiruan diharapkan dapat membantu mendeteksi jenis spesies atau ikan dari gambar yang dimasukkan ke dalam pengolahan citra, maka berdasarkan uraian permasalahan diatas menjadi acuan penulis menyusun penelitian dengan judul “Klasifikasi Jenis Ikan Berbasis Jaringan Syaraf Tiruan Menggunakan Algoritma K-Nearest Neighbors Dan *Principal Component Analysis* (PCA)”. Penggunaan *Principal Component Analysis* (PCA) dalam melakukan klasifikasi jenis Ikan dikarenakan metode ini dapat mengolah citra wajah seseorang sehingga secara otomatis sistem akan mengenali wajah seseorang atau lainnya melalui ciri – ciri utamanya seperti mata, hidung bibir sebagai identitasnya. Metode ini baik digunakan untuk melakukan klasifikasi terhadap suatu objek. Penggunaan metode ini diharapkan akan membantu nelayan dalam membedakan jenis ikan.

2. TINJAUAN PUSTAKA

2.1. Penelitian Sebelumnya

Penelitian yang dilakukan oleh Sianturi, *dkk.* Dengan judul “Klasifikasi Dokumen Menggunakan Kombinasi Algoritma PCA dan SVM”. Pada penelitian ini Penggunaan PCA dan SVM pada dataset dokumen online menghasilkan waktu komputasi jauh lebih cepat dibandingkan menggunakan SVM. Tetapi SVM menghasilkan akurasi yang lebih baik daripada menggunakan PCA yaitu sebesar 98.95% [1]. Penelitian lain dengan judul “Hasil Ekstraksi Algoritma Principal Component Analysis (PCA) Untuk Pengenalan Wajah Dengan Bahasa Pemograman Java Eclipse IDE”. Pada penelitian ini dengan menggunakan kamera digital sebagai alat melakukan *capture* lalu nilai RGB di normalisasi untuk pengenalan wajah, dengan mendapat akurasi 86.5% dari pengujian sistem. Metode ini baik digunakan untuk melakukan klasifikasi terhadap suatu objek [2]. Penerapan metode *Principal component analysis* juga dilakukan oleh [3] dengan judul penelitian “Penerapan Principal Component Analysis Untuk Peningkatan Kinerja Algoritma Decision Tree Pada Iris Dataset”. Dimana pada penelitian ini klasifikasi digunakan untuk memprediksi

objek yang belum memiliki kelas/label. Penggunaan *principal component analysis* (PCA) dalam optimasi algoritma decision tree, dilakukan saat preprocessing dataset. Berdasarkan hasil perhitungan, akurasi algoritma decision tree setelah dilakukan optimasi dengan PCA terhadap dataset bunga iris sebesar 95.33%. Penelitian lainnya yang dilakukan oleh [4] yang membahas tentang klasifikasi pola primitive menggunakan PCA. sedangkan Penelitian dengan judul “Diagnosa Hama dan Penyakit pada Tanaman Jeruk dengan Menerapkan Jaringan Syaraf Tiruan Learning Vector Quantization” [5]. Pada penelitian ini akan menerapkan sebuah algoritma jaringan syaraf tiruan learning vector quantization untuk diagnosa hama dan penyakit pada tanaman jeruk dengan menerapkan sebuah jaringan LVQ yang memiliki 21 gejala masukan untuk hama dan 34 gejala masukan untuk penyakit pada tanaman jeruk. Sedangkan keluaran terdiri atas 12 kelas hama pada tanaman jeruk dan 17 kelas penyakit pada tanaman jeruk. Penelitian lain yang dilakukan oleh [6] dengan judul “Identifikasi Jenis Daun Tanaman Obat Hipertensi Berdasarkan Citra RGB Menggunakan Jaringan Syaraf Tiruan”. pada penelitian ini dilakukan Pelatihan (*training*) menggunakan JST. Identifikasi menggunakan JST dapat mengidentifikasi daun tanaman obat hipertensi dengan parameter RGB. Bobot paling optimal didapatkan pada epoch 2500 dengan jumlah Neuron 10 pada hidden layer dan nilai MSE (*error*) sebesar 0.0754 pada detik ke 25. Kemampuan pengenalan JST dari masing-masing daun yaitu daun salam 100%, daun kumis kucing 80%, daun seledri 100%, daun alpukat 40%, daun sirsak 60%. Menghasilkan pengenalan rata-rata sebesar 72%.

2.2. Jaringan Saraf Tiruan

Sejak ditemukan pertamakali oleh Mc Culloch dan Pitts, sistem jaringan syaraf tiruan berkembang pesat dan banyak di gunakan oleh banyak aplikasi, jaringan syaraf tiruan (*Artificial Neural Network*) adalah suatu jaringan untuk memodelkan cara kerja sistem syaraf manusia (otak) dalam melaksanakan tugas tertentu. Pemodelan ini didasari oleh kemampuan otak manusia dalam mengorganisasi sel-sel penyusunan (*neuron*), sehingga memiliki kemampuan untuk melaksanakan tugas-tugas tertentu khususnya pengenalan pola dengan efektifitas jaringan tertiggi. [7].

Keunggulan yang utama dari sistem *Artificial Neural Network* (ANN) adalah adanya kemampuan untuk belajar dari contoh yang di berikan atau data training (Sapanta, 2018). Sedangkan untuk kelemahan utamanya dari *Artificial Neural Network* (ANN) yakni di butuhkan pelatihan untuk pengoperasiannya dan di butuhkan waktu yang lama untuk memproses *Artificial Neural Network* (ANN) [8].

2.3. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) Atau di kenal juga dengan nama Hotelling Transform adalah sebuah metode yang menghasilkan transformasi ortogonal yang di

sebut dengan nama eigenimage dimana sebuah citra di presentasikan ke dalam bentuk proyeksi linier searah dengan eigenimage yang bersesuaian dengan nilai eigen terbesar dari matriks kovarian, [9].

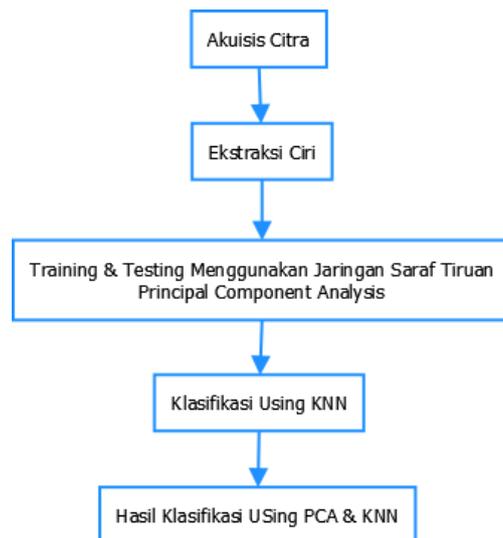
PCA adalah metode statistik yang sudah dipakai dalam berbagai bidang seperti pengenalan pola, penginderaan jauh (remote sensing), kompresi citra dan lain-lain. Teknik ini biasanya dipakai untuk mengenali pola pada data atau sampel berdimensi tinggi dan menyatakan data tersebut dengan cara sedemikian rupa sehingga persamaan dan perbedaan yang ada dalam data tersebut dapat ditentukan [4].

2.4. K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) merupakan metode yang biasa digunakan pada klasifikasi data. Algoritma ini digunakan untuk mengklasifikasikan terhadap objek berdasarkan data pembelajaran yang jarak tetangganya paling dekat atau memiliki nilai selisih yang kecil dengan objek tersebut [10]. Tujuan dari algoritma ini ialah mengklasifikasikan objek baru berdasarkan atribut dan training sample [11]. Prinsip umum dari algoritma ini adalah menemukan k data training untuk menentukan k -nearest neighbor berdasarkan ukuran jarak [12]. Selanjutnya mayoritas dari k tetangga terdekat akan menjadi dasar untuk memutuskan kategori dari sample berikutnya. Selain itu algoritma ini sendiri sering digunakan untuk klasifikasi pada teknik data mining meskipun dapat digunakan untuk estimasi dan prediksi data.

3. METODE YANG DIUSULKAN

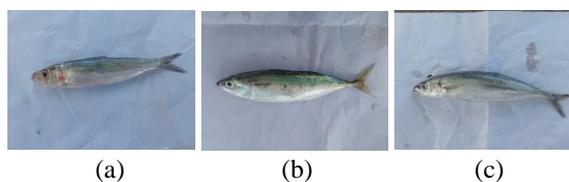
Penelitian ini menggunakan arsitektur jaringan syaraf tiruan dengan mengkombinasikan algoritma dari *principal component analysis* (PCA) dan K-Nearest Neighbor (KNN) untuk melakukan klasifikasi terhadap jenis ikan. Adapun langkah penelitian dimulai dengan akuisisi citra, preprocessing, ekstraksi ciri, pelatihan dan pengujian serta hasil kesimpulan yang didapatkan. Adapun blok diagramnya dapat dilihat pada gambar 1.



Gambar 1. Blok Diagram

3.1 Akuisisi Citra

Akuisisi citra adalah langkah untuk mengambil data yang diperlukan. Dalam penelitian ini akuisisi citra dilakukan dengan menggunakan alat dan kamera smartphone. Data yang digunakan yaitu 3 jenis ikan laut. Proses akuisisi data dilakukan dengan mengambil beberapa ikan dan diletakan pada kertas HVS warna putih dengan pencahayaan yang sama, kemudian dilakukan proses akuisisi data. Data dari hasil akuisisi citra yang digunakan yaitu 3 jenis ikan laut, masing-masing jenis ikan difoto sebanyak 8 citra sehingga total 45 citra, 30 data citra digunakan sebagai data latih dan 15 citra digunakan sebagai data uji.



Gambar 2. (a) Citra Ikan Haring, (b) Citra Ikan Kembung, (c) Citra Ikan Layang

3.2 Preprocessing

Tahapan pre-processing ini meliputi diantaranya adalah cropping yaitu proses pemotongan pada citra yang termasuk dataset lalu dikonversi pada *grayscale*, citra *image* ikan yang telah diakuisisi sekaligus menjadi dataset untuk proses training dan testing pada sistem. Ekstraksi ciri lalu masuk pada training dan uji jaringan syaraf tiruan *Principal Component Analysis* dan KNN.

3.3 Ekstraksi Ciri

Ekstraksi ciri merupakan langkah awal dalam menentukan parameter sebagai interpretasi analisis tekstur citra. Dalam penelitian ini dilakukan dengan empat

ekstraksi yaitu ekstraksi ciri *Hue*, *Saturation*, *Value* dan *Area*. Hasil dari ekstraksi ciri pada dataset akan dijadikan sebagai parameter pada proses training dan testing nantinya.

3.4 Training

Pada tahap ini data citra yang telah diekstraksi ciri akan dimasukan pada tahap selanjutnya yaitu training atau pelatihan, pada proses ini data citra akan ditraining menggunakan algoritma jaringan syaraf tiruan PCA lalu selanjutnya akan masuk pada tahap uji atau testing.

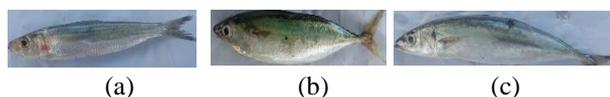
3.5 Testing

Pada tahap pengujian ini data citra yang telah di latih akan di uji dengan menggunakan jaringan saraf tiruan PCA yang sama untuk mendapatkan hasil pengenalan dan klasifikasi sesuai yang diharapkan dan menjadi tujuan dari penelitian ini.

4. HASIL PENELITIAN

4.1 Akuisisi dan Preprocessing Citra

Berikut adalah hasil dari akuisisi serta preprocessing dari objek penelitian ditunjukkan pada gambar 3.



Gambar 3. (a) Citra Cropping Ikan Haring, (b) Citra cropping Ikan Kembung, (c) Citra Cropping Ikan Layang.

4.2 Ekstraksi Ciri

Citra ikan yang terdiri dari tiga jenis citra ikan yang telah melalui proses *preprocessing* kemudian dilakukan proses selanjutnya yaitu pengambilan ciri citra. Ekstraksi ciri citra dilakukan pada citra yang telah melalui *cropping*. Ekstraksi ciri dilakukan dengan menggunakan histogram orde dua atau tingkat dua yaitu matriks *co-occurrence* yakni ekstraksi ciri *hue*, *saturation*, *value* dan *area*. Berikut hasil ekstraksi ciri ditunjukkan pada tabel 1.

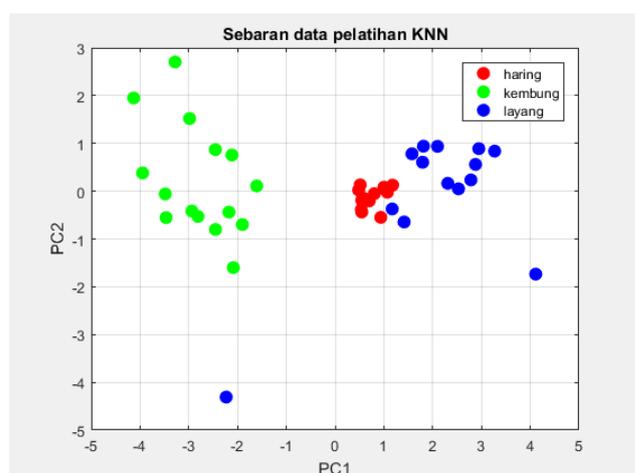
Tabel 1. Ekstraksi Ciri

N	Data uji	Hue	Saturasi	Value	Area
1	Harin g	0.520997231 778420	0.181899951 193295	0.515313011 319439	187 136
2	Harin g	0.521514906 414411	0.180644244 650969	0.522079985 085772	176 907
3	Harin g	0.527769893 187824	0.187226865 219527	0.514619992 092090	186 113
4	Harin g	0.523286093 108283	0.179909998 077061	0.520499742 902620	178 768

5	Harin g	0.520490867 977812	0.185777413 955012	0.516681970 191096	187 641
6	Kemb ung	0.517692069 740028	0.189507990 002789	0.486709122 536951	402 531
7	Kemb ung	0.503556962 432909	0.184218954 667893	0.454764375 601962	175 549
8	Kemb ung	0.509172786 552451	0.183112158 104804	0.469227491 639081	262 443
9	Kemb ung	0.511242017 450049	0.180610787 387618	0.469203320 594324	226 893
10	Kemb ung	0.510338812 527247	0.188768626 135839	0.457742917 122204	245 939
11	Layan g	0.498822299 659204	0.159849740 968884	0.487140138 342529	131 515
12	Layan g	0.553487872 444792	0.196450410 272599	0.537595610 817264	153 797
13	Layan g	0.554851811 312253	0.189093729 698080	0.556001080 616818	153 870
14	Layan g	0.550323417 494378	0.186270621 536707	0.554309915 315632	152 893
15	Layan g	0.540693924 978831	0.184907234 567002	0.541696728 259027	138 800

4.3 Training PCA dan KNN

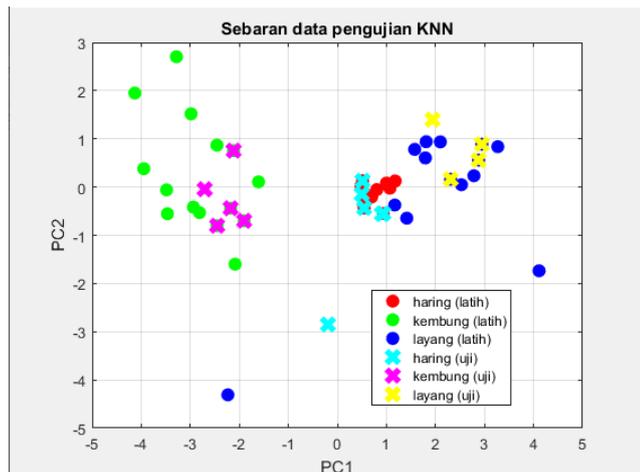
Berdasarkan hasil ekstraksi ciri maka dilakukan pengklasifikasian menggunakan PCA dan KNN untuk mengklasifikasikan ketiga jenis ikan yang diteliti. Adapun hasilnya adalah ditunjukkan melalui gambar berikut.



Gambar 4. Sebaran Data Pelatihan

4.4 Testing

Sebaran data pengujian ditunjukkan melalui gambar 5 berikut.



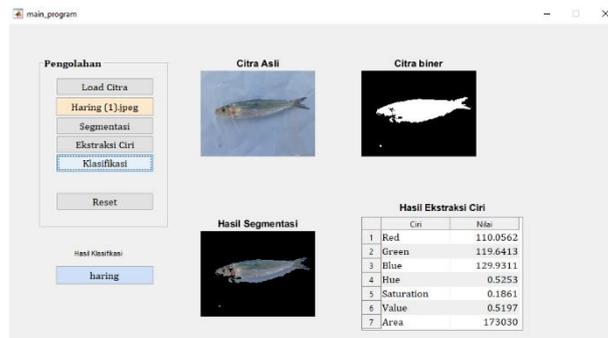
Gambar 5. Sebaran Data Pengujian

Berikut hasil pengujian dengan menggunakan algoritma PCA dan klasifikasi KNN ditunjukkan melalui tabel 2 berikut.

Tabel 2. Hasil Pengujian PCA dan KNN

No	Data uji	Plot Kelas PCA & KNN	Plot Kelas Hasil Uji	Keterangan
1	Ikan Haring	1	1	Benar
2	Ikan Haring	1	1	Benar
3	Ikan Haring	1	1	Benar
4	Ikan Haring	1	1	Benar
5	Ikan Haring	1	1	Benar
6	Kembung	2	2	Benar
7	Kembung	2	2	Benar
8	Kembung	2	2	Benar
9	Kembung	2	2	Benar
10	Kembung	2	2	Benar
11	Layang	3	1	Salah
12	Layang	3	3	Benar
13	Layang	3	3	Benar
14	Layang	3	3	Benar
15	Layang	3	3	Benar

Pada hasil pengujian ini didapatkan satu plot kelas yang salah dalam melakukan klasifikasi yakni data plot kelas ikan haring sehingga, presentase keseluruhan pada proses uji menggunakan PCA dan KNN adalah sebesar 93 % (Sembilan puluh tiga persen). Adapun hasil running sistem klasifikasi yang dibangun ditunjukkan melalui Gambar 6 dibawah.



Gambar 6. Running System Klasifikasi

4.5 Deployment

Ciri yang digunakan adalah ekstraksi ciri orde ke satu dan dua yang digunakan yaitu, nilai *hue*, *saturation*, *value* dan *area*. Pada penelitian ini bobot awal menggunakan ciri dari vektor pelatihan urutan ke-1 dari masing-masing kelas. Kemudian bobot awal diolah dengan nilai maksimum iterasi. Berdasarkan pada data Tabel 2 bahwa data uji pada kelas satu, dua dan tiga hanya satu yang tidak bisa dikenali oleh sistem sedangkan data lainnya dapat dikenali maka, dengan hasil ini memperoleh keberhasilan sebesar 93% (Sembilan puluh tiga persen).

5. KESIMPULAN

Pada penelitian ini telah dilakukan pengujian untuk pengenalan atau klasifikasi ikan laut berbasis teknik pengolahan citra dan jaringan syaraf tiruan principal component analysis. Dari hasil pengujian klasifikasi jenis ikan didapatkan kesimpulan baik dari hasil training data pelatihan maupun sebagai data uji. Pada tahap pelatihan jaringan syaraf tiruan principal component analysis digunakan 45 buah data citra dengan target 3 target, roses pengujian dengan data latih yangtelah disiapkan dan di training, dapat mengenali 93% data citra yang diujikan. Hasil pengujian memperlihatkan bahwa metode Jaringan Syaraf Tiruan menggunakan Algoritma PCA dan KNN dapat menunjukkan akurasi diatas 90%.

UCAPAN TERIMA KASIH

Terima kasih kepada civitas akademik Informatika Universitas Mercu Buana Yogyakarta serta pihak-pihak yang berkontribusi serta berpartisipasi lebih terhadap penelitian ini.

Daftar Pustaka

[1] M. Sianturi, A. Adiwijaya, and S. Faraby, "Klasifikasi Dokumen Menggunakan Kombinasi Algoritma Principal Component Analysis Dan Svm," *eProceedings Eng.*, vol. 4, no. 3, 2017.
 [2] F. Ismawan, "Hasil Ekstraksi Algoritma Principal Component Analysis (PCA) untuk Pengenalan Wajah dengan Bahasa Pemrograman Java Eclipse

- IDE,” *J. Sisfotek Glob.*, vol. 5, no. 1, 2015.
- [3] P. K. Handayani, “Penerapan Principal Component Analysis untuk Peningkatan Kinerja Algoritma Decision Tree pada Iris Dataset,” *Indones. J. Technol. Informatics Sci.*, vol. 1, no. 2, pp. 55–58, 2020.
- [4] I. G. A. Widagda and H. Suyanto, “Klasifikasi Pola Berbentuk Primitif dengan Menggunakan Metode Principal Component Analysis (PCA) The Classification of Primitive-Shaped Patterns by Using Principal Component Analysis Method,” pp. 12–21.
- [5] M. Megawati and R. M. Candra, “Diagnosa Hama Dan Penyakit Pada Tanaman Jeruk Dengan Menerapkan Jaringan Syaraf Tiruan Learning Vector Quantization (Studi Kasus: Badan Penyuluhan Pertanian Kuok),” *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 3, no. 2, pp. 59–62, 2018.
- [6] I. Jamaliah, R. N. Whidhiasih, and M. Maimunah, “Identifikasi Jenis Daun Tanaman Obat Hipertensi Berdasarkan Citra Rgb Menggunakan Jaringan Syaraf Tiruan,” *Penelit. Ilmu Komput. Sist. Embed. dan Log.*, vol. 5, no. 1, pp. 1–11, 2017.
- [7] E. D. Handoyo and L. W. Susanto, “Penerapan Jaringan Syaraf Tiruan metode Propagasi Balik Dalam Pengenalan Tulisan Tangan Huruf Jepang Jenis Hiragana dan Katakana,” *J. Inform.*, vol. 7, no. 1, pp. 39–55, 2011.
- [8] C. Figri and I. Susilawati, “Identifikasi Jenis Tanaman Vinca Berdasarkan Ekstraksi Ciri Citra Bunga Dan Jaringan Syaraf Tiruan Identification Of Vinca Plant Species Based On Feature Extraction Of Flower Images And Artificial Neural Networks,” pp. 54–60, 2020.
- [9] M. L. Aksani, “PERANCANGAN PROGRAM APLIKASI PENGENALAN WAJAH DENGAN MENERAPKAN METODE PRINCIPAL COMPONENT ANALYSIS DAN JARINGAN SYARAF TIRUAN,” *J. Tek.*, vol. 6, no. 1, 2017.
- [10] R. N. Devita, H. W. Herwanto, and A. P. Wibawa, “Perbandingan kinerja metode naive bayes dan k-nearest neighbor untuk klasifikasi artikel berbahasa indonesia,” *JTIK (Jurnal Teknol. Inf. dan Ilmu Komputer) Vol*, vol. 5, 2018.
- [11] M. Lestari, “Penerapan algoritma klasifikasi Nearest Neighbor (K-NN) untuk mendeteksi penyakit jantung,” *Fakt. Exacta*, vol. 7, no. 4, pp. 366–371, 2015.
- [12] S. D. Nugraha, R. R. M. Putri, and R. C. Wihandika, “Penerapan Fuzzy K-Nearest Neighbor (FK-NN) Dalam Menentukan Status Gizi Balita,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput. e-ISSN*, vol. 2548, p. 964X, 2017.