JUSIFOR: Jurnal Sistem Informasi dan Informatika
Vol. 2, No. 2, Desember 2023, hlmn. 84-93
E-ISSN: 2830-3393| P-ISSN: 2830-2443
DOI: 10.33379/jusifor.v2i2.3728

Perbandingan Performa Klasifikasi Citra Ikan Menggunakan Metode K-Nearest Neighbor (K-NN) Dan Convolutional Neural Network (CNN)

Nizar Abdurrahman[1*], Basuki Rahmat[2], Andreas Nugroho Sihananto[3]

[1,2,3] Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jawa Timur, Indonesia

Info Artikel

Riwayat Artikel:
Diterima : **18-12-2023**
Direvisi : **21-12-2023**
Disetujui : **23-12-2023**

Kata Kunci:

klasifikasi,
k-nearest neighbor,
convolutional neural network

ABSTRAK

Perairan daratan Indonesia memiliki keanekaragaman jenis ikan, mencakup variasi ikan air tawar, payau, dan laut. Meskipun Indonesia memiliki lebih dari 2.000 spesies ikan, hanya sekitar 25% yang dapat dibudidayakan. Beberapa jenis ikan seperti lele, mujair, bandeng, patin, dan gurami menjadi pilihan utama dalam budidaya ikan konsumsi. Dalam konteks ini, penelitian menggunakan teknologi citra digital untuk membandingkan dua metode klasifikasi, yaitu k-NN dan CNN, dengan fokus pada citra ikan. Metode k-NN dan CNN diuji dengan menggunakan dua skenario pembagian *dataset*, yaitu 80:20 dan 90:10. Hasil eksperimen menunjukkan bahwa CNN menghasilkan akurasi yang lebih tinggi dibandingkan k-NN pada kedua skenario. Akurasi CNN mencapai 88% pada rasio data 80:20 dan 80% pada rasio 90:10, sedangkan k-NN mencapai 66% dan 72% pada skenario yang sama. Meskipun performa metode dapat dipengaruhi oleh berbagai faktor, hasil penelitian ini memberikan indikasi bahwa CNN lebih efektif dalam mengklasifikasikan citra ikan. Keunggulan ini dapat memberikan kontribusi penting dalam pemilihan metode klasifikasi untuk pengolahan citra ikan.

***Keywords*:**

classification,
k-nearest neighbor,
convolutional neural network

ABSTRACT

Indonesia's inland waters have a diversity of fish types, including a variety of freshwater, brackish and marine fish. Even though Indonesia has more than 2,000 fish species, only around 25% can be cultivated. Several types of fish such as catfish, tilapia, milkfish, catfish and gourami are the main choices for cultivating consumption fish. In this context, research uses digital image technology to compare two classification methods, namely k-NN and CNN, with a focus on fish images. The k-NN and CNN methods were tested using two dataset division scenarios, namely 80:20 and 90:10. Experimental results show that CNN produces higher accuracy than k-NN in both scenarios. CNN accuracy reached 88% at a data ratio of 80:20 and 80% at a ratio of 90:10, while k-NN reached 66% and 72% in the same scenario. Although the performance of the method can be influenced by various factors, the results of this study provide an indication that CNN is more effective in classifying fish images. This advantage can provide an important contribution in selecting classification methods for fish image processing.

Penulis Korespondensi:
Nizar Abdurrahman,
Informatika,
Universitas Pembangunan Nasional "Veteran" Jawa Timur
Email: 19081010025@student.upnjatim.ac.id

This is an open access article under the CC BY-SA license

1. PENDAHULUAN

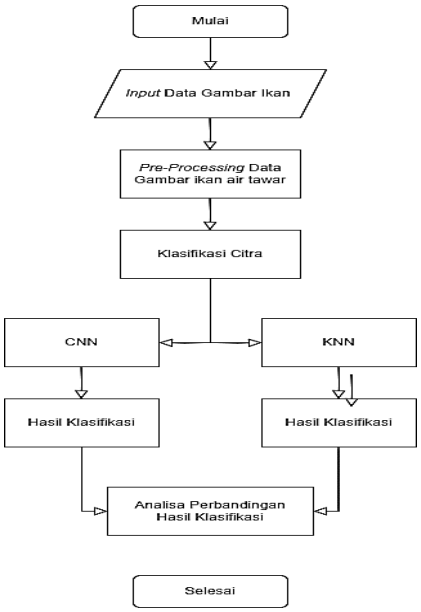
Perairan daratan Indonesia memiliki keanekaragaman jenis ikan, mencakup variasi jenis ikan air tawar, air payau, dan laut. Meskipun terdapat banyak jenis ikan, tidak semuanya cocok untuk dikonsumsi atau dibudidayakan oleh masyarakat Indonesia. Meski demikian, Indonesia dikenal sebagai negara terkaya pertama dalam bidang perikanan, dengan lebih dari 2.000 spesies ikan yang terdapat di perairannya, baik laut maupun perairan tawar seperti danau, sungai, dan rawa. Sayangnya, hanya sekitar 25% dari jumlah spesies tersebut yang dapat dibudidayakan, suatu kenyataan yang disayangkan karena potensi keberagaman hayati Indonesia, khususnya ikan air tawar dan payau, dapat dimanfaatkan secara optimal jika masyarakat lebih mengetahui jenis ikan yang dapat dibudidayakan[1].

Indonesia memiliki berbagai jenis ikan yang menjadi pilihan utama dalam budidaya ikan. Walaupun terdapat banyak jenis ikan, hanya beberapa yang sering dipilih sebagai ikan budidaya, termasuk lele, mujair, bandeng, patin, gurami, dan sebagainya. Ikan-ikan ini umumnya dibudidayakan sebagai konsumsi karena mudah ditemukan di sekitar masyarakat. Meski begitu, sulit bagi generasi muda untuk membedakan jenis ikan, seperti lele, patin, mujair, gurami, dan lainnya, hanya dengan melihatnya secara langsung. Oleh karena itu, penelitian dilakukan menggunakan teknologi citra digital. Dalam era modern ini, berbagai jenis ikan dapat diklasifikasikan melalui pengolahan citra digital, dimungkinkan oleh perkembangan teknologi fotografi digital dan kemajuan ilmu komputer dalam bidang visi komputer [1].

Penelitian ini menggunakan teknologi untuk membantu analisis jenis ikan air tawar dengan mengklasifikasikannya berdasarkan warna menggunakan metode RGB, Citra RGB merupakan citra digital yang memiliki kombinasi warna *red, green* dan *blue* [2]. Dua algoritma klasifikasi, k-NN dan CNN, diterapkan dalam penelitian ini, dengan *Confusion matrix* digunakan untuk membandingkan hasil klasifikasi dan menghitung akurasi. Tujuan penelitian adalah membandingkan performa klasifikasi citra ikan antara k-NN dan CNN. Sebelumnya, k-NN terbukti lebih efektif daripada metode Naïve Bayes, dan penelitian ini mengevaluasi kembali k-NN, membandingkannya dengan CNN. Harapannya, penelitian ini dapat menentukan metode klasifikasi terbaik untuk citra ikan dengan melibatkan jenis-jenis tertentu, hasilnya akan dibandingkan menggunakan *Confusion matrix*.

2. METODE PENELITIAN

3.1. Perancangan Penelitian



Gambar 1. Perancangan Penelitian

Perencanaan penelitian ini melibatkan rangkaian tindakan yang diperlukan agar proses penelitian dapat dilaksanakan secara efektif. Tujuannya adalah agar penulis dapat mengidentifikasi, menjawab, dan mencapai

tujuan pada penelitian ini. Adapun langkah-langkah penelitian tersebut dijelaskan dengan detail pada gambar berikut. Penyusunan rencana penelitian untuk membandingkan kinerja klasifikasi citra ikan menggunakan metode k-Nearest Neighbor (k-NN) dan Convolutional Neural Network (CNN). Langkah awal melibatkan pemasukan data gambar ikan, dengan menggunakan jenis ikan yang umumnya dapat dijumpai di sekitar kita. Tahap berikutnya adalah pemrosesan awal data gambar ikan, yang mencakup normalisasi ukuran data dan pemberian label pada data tersebut. Setelah itu, tahap ketiga melibatkan klasifikasi citra ikan menggunakan kedua metode, yaitu KNN dan CNN. Langkah terakhir melibatkan analisis perbandingan hasil klasifikasi dari kedua metode dan melakukan perbandingan dengan menggunakan *confusion matrix* untuk mengukur akurasi.

3.2. *Dataset*

Dataset adalah sebuah kumpulan informasi yang terstruktur. Sebuah *dataset* terdiri dari kumpulan *file* dalam sistem komputer [3]. Ikan merupakan sekelompok hewan vertebrata yang mendiami lingkungan air, baik itu air tawar maupun laut. Mereka memiliki kerangka tubuh yang terdiri dari tulang belakang atau tulang rawan, memperoleh oksigen melalui insang, dan menggunakan sirip untuk bergerak. Kulit ikan umumnya dilapisi sisik, dan mereka umumnya berkembang biak dengan meletakkan telur. Variasi bentuk dan ukuran ikan sangat beragam, serta memiliki peran penting dalam menjaga keseimbangan ekosistem perairan. Selain itu, ikan memiliki nilai ekonomi yang signifikan sebagai sumber daya pangan dan menjadi fokus penelitian dalam bidang ilmu perikanan dan biologi laut. Kehadiran ikan juga menjadi sumber protein yang paling umum dan mudah diakses. Oleh karena itu, konsumsi ikan dianggap dapat meningkatkan kecerdasan otak pada anak-anak [4].Contoh *file* gambar *dataset* ikan sebagai berikut.

Sumber data untuk penelitian ini merupakan kumpulan gambar ikan konsumsi. Dataset terdiri dari lima jenis ikan konsumsi, yaitu ikan bandeng, ikan mujair, ikan lele, ikan gurami, dan ikan patin. Setiap jenis ikan memiliki 50 data, sehingga totalnya terdapat 250 *dataset*, dengan masing-masing jenis ikan terdiri dari 40 data untuk pelatihan dan 10 data untuk pengujian.

3.3. Pre-Processing

Pre-processing citra merupakan serangkaian langkah atau teknik yang digunakan untuk menyiapkan citra digital sebelum dilakukan analisis atau pengolahan lebih lanjut. Tujuan dari *pre-processing* antara lain untuk menyeragamkan *dataset* agar lebih mudah diolah [5]. Proses *pre-processing* mencakup beberapa langkah, seperti pemotongan (*cropping*) untuk menghilangkan bagian citra yang tidak relevan atau untuk mengisolasi area tertentu, hal ini dimaksudkan agar ukuran citra yang akan diolah dalam sistem mempunyai ukuran yang sama [6]. Serta perubahan ukuran (*resizing*) untuk menyesuaikan ukuran citra dengan kebutuhan analisis.

Selain itu, terdapat tahap pembuatan data *training* dan testing, di mana data dibagi dengan perbandingan tertentu, yaitu 80%:20%, dan 90%:10%. Jumlah citra pada setiap kategori tetap 50, dan pembagian data dilakukan sesuai dengan persentase split yang telah ditentukan. Sebagai contoh, pada split data 80%:20%, 40 citra digunakan sebagai data *training* dan 10 citra lainnya sebagai data testing.

3.4. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) merupakan suatu jenis jaringan saraf yang umumnya digunakan dalam pengolahan data citra. CNN digunakan khususnya untuk melakukan klasifikasi pada data yang sudah memiliki label, dengan menerapkan metode *supervised learning*. Dalam *supervised learning*, target atau hasil yang diharapkan dari input yang diterima oleh jaringan sudah diketahui sebelumnya. Dari segi teknis, CNN adalah suatu arsitektur yang dapat di-latih dan terdiri dari beberapa tahap. *Input* yang diberikan ke CNN berupa objek dalam bentuk citra. Proses yang menjelaskan citra dan mengubahnya menjadi fitur yang dapat dipahami oleh jaringan merupakan ciri khas yang membedakan CNN dari jenis jaringan syaraf lainnya [7].

Kontribusi utama CNN terletak pada penggunaan operasi konvolusi pada lapisan-lapisan dalam struktur jaringan. Operasi konvolusi ini memungkinkan jaringan untuk secara efisien mengekstraksi fitur-fitur lokal dari citra dengan menggunakan jendela konvolusi yang bergerak di seluruh citra. Fitur-fitur yang diekstraksi ini kemudian dipelajari oleh lapisan-lapisan berikutnya dalam jaringan untuk menghasilkan representasi yang semakin abstrak dan kompleks.

CNN telah sukses digunakan dalam berbagai aplikasi pengolahan citra dan pengenalan pola, termasuk klasifikasi citra, deteksi objek, segmentasi citra, dan lain-lain. Keunggulan utamanya terletak pada

kemampuannya untuk secara otomatis mengekstraksi fitur-fitur yang relevan dari data citra tanpa memerlukan desain fitur secara manual, menjadikannya sangat sesuai untuk tugas-tugas pengolahan citra yang kompleks.

3.5. K-Nearest Neighbor (k-NN)

Metode k-Nearest Neighbor (KNN) adalah suatu algoritma yang digunakan untuk klasifikasi objek berdasarkan data latihan yang memiliki jarak terdekat dengan objek tersebut. Proses pengelompokan data baru dilakukan dengan mempertimbangkan jaraknya terhadap sejumlah tetangga terdekat [8].

K-NN merupakan metode berbasis algoritma terpandu yang mengklasifikasikan instansi baru berdasarkan mayoritas kategori dari tetangga-tetangga terdekatnya. Tujuan utama algoritma ini adalah menggunakan atribut dan sampel latihan untuk mengklasifikasikan objek yang baru. Langkah-langkah penyelesaian masalah dengan menggunakan metode K-NN dapat diuraikan sebagai berikut:

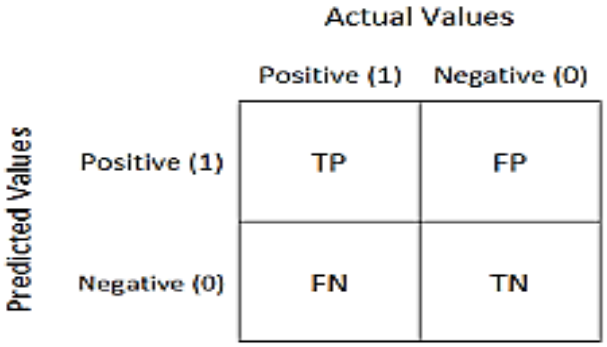
1. Menentukan nilai parameter k (jumlah tetangga terdekat).
2. Menghitung jarak antara setiap sampel data dengan data yang diuji.
3. Mengurutkan data berdasarkan jarak dari yang terkecil hingga terbesar.
4. Mengamati jumlah keputusan terbanyak untuk k data yang diambil.
5. Jika terdapat dua atau lebih kelas yang menjadi tetangga terdekat dari data uji x, maka kondisi seimbang (konflik) terjadi dan diterapkan strategi pemecahan konflik.
6. Untuk setiap kelas yang terlibat dalam konflik, menentukan jarak antara x dan kelas tersebut berdasarkan E tetangga terdekat yang ditemukan pada kelas tersebut.
7. Jika terdapat pola latihan ke-m dari kelas tertentu yang terlibat dalam konflik, maka jarak antara x dan kelas tersebut dihitung menggunakan Persamaan 1 [9] :

$$d_i = \frac{1}{E}\sum_j^n = 1\,|(Xj - Yj)| \quad (1)$$

Algoritma KNN sangat sesuai digunakan untuk masalah klasifikasi dengan dataset yang relatif sederhana. Namun, perlu dicatat bahwa algoritma ini dapat memerlukan komputasi yang intensif ketika jumlah data sangat besar, karena memerlukan perhitungan jarak untuk setiap titik data. Selain itu, KNN juga dapat diperluas untuk masalah regresi, di mana nilai kontinu diprediksi berdasarkan tetangga terdekat. Terdapat variasi dari algoritma KNN, seperti KNN dengan bobot jarak dan KNN dengan kernel, yang memberikan fleksibilitas tambahan dalam mengklasifikasikan dan memprediksi data.

3.6. *Confusion matrix*

Salah satu cara untuk menilai kinerja suatu model klasifikasi adalah melalui pengukuran akurasi model tersebut. Ada beberapa istilah dasar yang menjadi dasar dalam menghitung nilai akurasi, seperti true positive (TP), true negative (TN), false positive (FP), dan false negative (FN). Istilah-istilah tersebut sering dijelaskan dalam bentuk matriks yang dikenal sebagai *confusion matrix*, seperti yang terlihat pada contoh berikut [7].



Gambar 3. *Confusion matrix*

Akurasi dalam konteks klasifikasi merujuk pada persentase kebenaran dalam pengelompokan data setelah diuji pada hasil klasifikasi. Penghitungan akurasi dengan menggunakan *confusion matrix* dapat dijelaskan pada Persamaan 2 (Wulandari, Yasin, & Widiharih, 2020).

$$akurasi = \frac{TP + TN}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN} + \mathrm{TN}} \quad (2)$$

Confusion matrix membantu dalam mengevaluasi performa model dengan menyajikan informasi tentang sejauh mana kemampuan model untuk mengidentifikasi dan mengklasifikasikan data. Dari elemen-elemen yang terdapat di dalamnya, kita dapat menghitung beberapa metrik evaluasi krusial, termasuk akurasi, presisi, *recall* (sensitivitas), F1-score, dan lainnya.

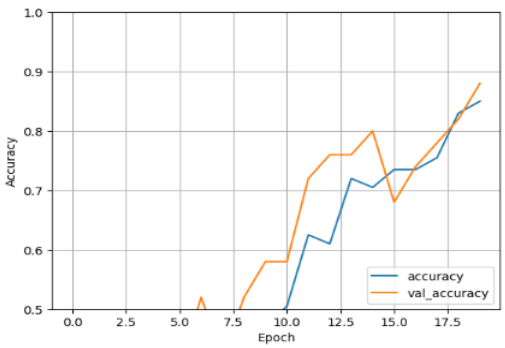
Akurasi, presisi, *recall*, dan F1-score adalah metode-metode pengujian dalam evaluasi performa sistem. Akurasi mengukur tingkat kesamaan antara nilai prediksi dan nilai sebenarnya, mencerminkan jumlah data yang diklasifikasikan secara benar. Presisi membandingkan jumlah informasi relevan yang diperoleh oleh sistem dengan total informasi yang diambil, memperhitungkan baik informasi relevan maupun tidak relevan. *Recall*, sebaliknya, membandingkan jumlah informasi relevan yang diperoleh dengan total informasi relevan yang tersedia, fokus pada rasio benar positif dari seluruh data yang benar positif. F1-score, sebagai ukuran statistik, menggabungkan presisi dan recall untuk menciptakan keseimbangan, berguna ketika distribusi kelas tidak seimbang.

3. HASIL DAN ANALISIS

3.1. Hasil Klasifikasi CNN

Peneliti dalam klasifikasi CNN pada citra ikan menggunakan pengujian dengan pembatasan *epoch* sampai 20 [10], dan melakukan pengujian pada *dataset* dengan rasio 80%:20% dan 90%:10%. Grafik pengujian epoch, Tabel hasil pengujian pada masing-masing rasio, dan *confusion matrix* masing-masing pengujian dapat dilihat sebagai berikut.

1. Pengujian Rasio 80%:20%



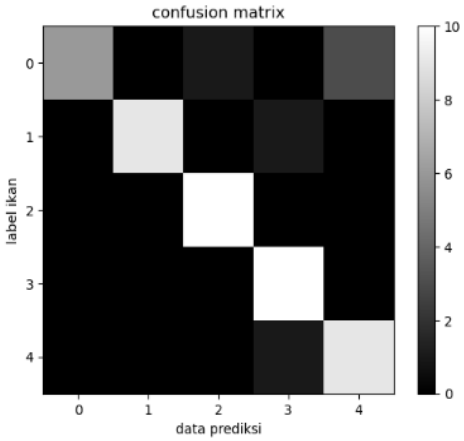
Gambar 4. Grafik Hasil Pelatihan CNN

Grafik di atas merupakan hasil pelatihan algoritma CNN dengan membatasi pelatihan hanya sampai 20 *epoch* yang di mana bertujuan untuk memudahkan evaluasi model CNN. Detail hasil pengujian klasifikasi algoritma CNN dapat dilihat pada Tabel 1. Dari Tabel 1 terlihat bahwa nilai akurasi pengujian rata-rata adalah 88%. Rincian akurasi untuk setiap jenis ikan adalah sebagai berikut: ikan bandeng 60%, ikan gurami 90%, ikan lele 100%, ikan mujair 100%, dan ikan patin 90%. Hasil tersebut menunjukkan bahwa klasifikasi untuk ikan lele dan ikan mujair memiliki tingkat akurasi yang tinggi, yakni mencapai 100%. Untuk selanjutnya untuk hasil classifikasi report dan *confusion matrix*nya dapat dilihat pada Gambar 5.

Tabel 1. Hasil pengujian ikan menggunakan CNN

No	Jenis Ikan	Klasifikasi Benar	Klasifikasi Salah	Akurasi
1	**Bandeng**	6	4	60%
2	**Gurami**	9	1	90%
3	**Lele**	10	0	100%
4	**Mujair**	10	0	100%
5	**Patin**	9	1	90%
	Rata-Rata			**88%**

Gambar 5 menunjukkan hasil *confusion matrix* dari proses klasifikasi ikan dengan menggunakan CNN pada pengujian rasio 80%:20%. Selanjutnya, kami akan menampilkan laporan klasifikasinya pada Gambar 5(a).

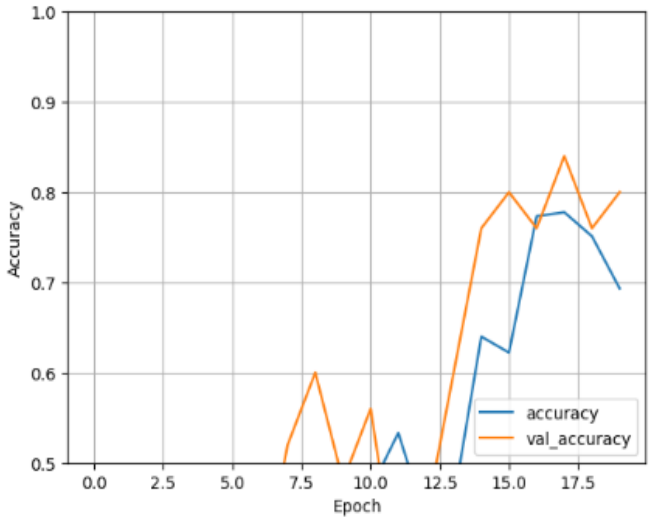


	precision	recall	f1-score	support
0	1.00	0.60	0.75	10
1	1.00	0.90	0.95	10
2	0.91	1.00	0.95	10
3	0.83	1.00	0.91	10
4	0.75	0.90	0.82	10
accuracy			0.88	50
macro avg	0.90	0.88	0.88	50
weighted avg	0.90	0.88	0.88	50

(a) **(b)**

Gambar 5. (a) *Confusion Matrix* dan (b) *Classification Report*

Dari Gambar 5(b), terlihat bahwa *confusion matrix* di atas mencapai nilai akurasi terbaik sebesar 0,88. Pada gambar tersebut, terlihat variasi hasil *confusion matrix* karena masih terdapat beberapa kesalahan dalam proses klasifikasi. Selanjutnya, pada *classification report*, terdapat nilai presisi, *recall*, dan f1-score yang bervariasi.



Gambar 6. Grafik Hasil Pelatihan CNN

2. Pengujian Rasio 90%:10%

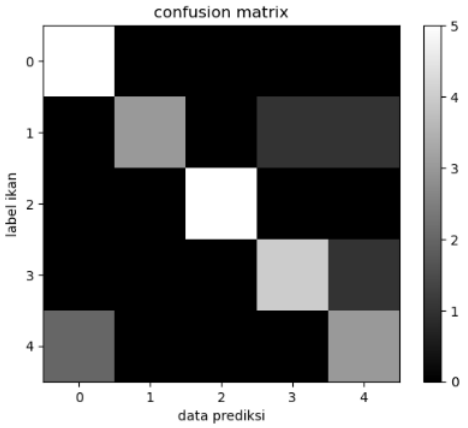
Grafik di atas merupakan hasil pelatihan algoritma CNN dengan membatasi pelatihan hanya sampai 20 *epoch* yang di mana bertujuan untuk memudahkan evaluasi model CNN. Detail hasil pengujian klasifikasi algoritma CNN dapat dilihat pada Tabel 2.

Tabel 2. Hasil pengujian ikan menggunakan CNN

No	Jenis Ikan	Klasifikasi Benar	Klasifikasi Salah	Akurasi
1	**Bandeng**	5	0	100%
2	**Gurami**	3	2	60%
3	**Lele**	5	0	100%
4	**Mujair**	4	1	80%
5	**Patin**	3	2	60%
	Rata-Rata			**80%**

Dari Tabel 2, dapat diamati bahwa hasil akurasi rata-rata dari pengujian mencapai 80%. Dengan akurasi khusus untuk ikan bandeng sebesar 100%, ikan gurami 60%, ikan lele 100%, ikan mujair 80%, dan ikan patin 60%. Dari data di atas, terlihat bahwa klasifikasi ikan bandeng dan ikan lele mencapai akurasi yang tinggi,

yakni 100%. Selanjutnya, kami memaparkan *confusion matrix* dan *classification report* pada Gambar 7(a) dan y(b). Gambar 7(a) menunjukkan hasil *confusion matrix* dari proses klasifikasi ikan dengan menggunakan CNN pada pengujian rasio 90%:10%. Selanjutnya, kami akan menampilkan laporan klasifikasinya pada Gambar 7(b). Dari Gambar 7(b), terlihat bahwa *confusion matrix* di atas mencapai nilai akurasi terbaik sebesar 0,80. Pada gambar tersebut, terlihat variasi hasil *confusion matrix* karena masih terdapat beberapa kesalahan dalam proses klasifikasi. Selanjutnya, pada *classification report*, terdapat nilai presisi, *recall*, dan F1-*score* yang bervariasi.



(a)

	precision	recall	f1-score	support
0	0.71	1.00	0.83	5
1	1.00	0.60	0.75	5
2	1.00	1.00	1.00	5
3	0.80	0.80	0.80	5
4	0.60	0.60	0.60	5
accuracy			0.80	25
macro avg	0.82	0.80	0.80	25
weighted avg	0.82	0.80	0.80	25

(b)

Gambar 7. (a) *Confusion Matrix* dan (b) *Classification Report*

3.2. Hasil Klasifikasi k-NN

Peneliti dalam klasifikasi k-NN pada citra ikan menggunakan pengujian dengan menggunkan nilai k (tetangga terdekat) dari k=1 sampai k=10 dan melakukan pengujian pada dataset dengan rasio 80%:20% dan 90%:10%., Tabel hasil pengujian pada masing-masing rasio, dan *confusion matrix* masing-masing pengujian dapat dilihat sebagai berikut.

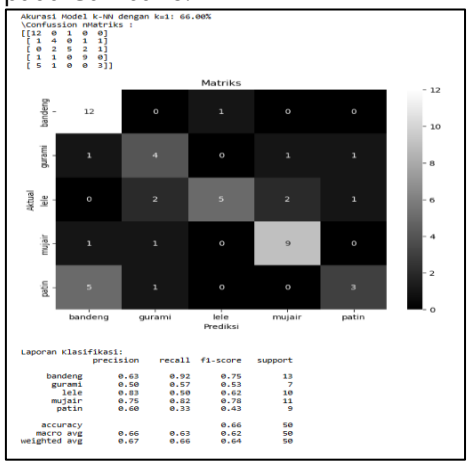
1. Pengujian Rasio 80%:20%

Pada pengujian klasifikasi menggunakan algoritma k-Nearest Neighbors (KNN) dengan rasio data 80:20, langkah utama yang dilakukan adalah membagi data menjadi dua bagian, yaitu data training sebanyak 80% dan data testing sebanyak 20%. Data *training* digunakan untuk melatih model KNN, sementara data testing digunakan untuk menguji performa model yang telah dilatih. Rasio 80:20 merupakan pilihan yang umum digunakan untuk memastikan model memiliki sejumlah data yang memadai untuk pembelajaran tanpa mengorbankan data yang cukup untuk pengujian.

Tabel 3. Hasil Klasifikasi Algoritma k-NN

Nilai K	Rasio	Akurasi
1	80:20	66,00%
2	80:20	58,00%
3	80:20	52,00%
4	80:20	48,00%
5	80:20	46,00%
6	80:20	54,00%
7	80:20	48,00%
8	80:20	42,00%
9	80:20	42,00%
10	80:20	46,00%

Dari Tabel 3, terlihat hasil dari proses klasifikasi untuk setiap nilai K, mulai dari K = 1 hingga K = 10, pada pembagian dataset dengan rasio 80:20. Terdapat nilai akurasi tertinggi yang tercatat pada K = 1, dengan mencapai tingkat akurasi sebesar 66,00%. Oleh karena itu, *confusion matrix* dan laporan klasifikasi pada skenario 4 dibentuk berdasarkan hasil klasifikasi dengan nilai K = 1, karena menghasilkan tingkat akurasi yang lebih tinggi daripada nilai K lainnya. Informasi lebih rinci mengenai *confusion matrix* dan laporan klasifikasi dapat ditemukan pada Gambar 8.



Gambar 8. *Confusion Matrix* dan *Classification Report*

Gambar 8 menunjukkan hasil *confusion matrix* dan laporan klasifikasi, di mana model dievaluasi dengan menggunakan akurasi untuk nilai k = 1 karena memberikan kinerja yang lebih baik dibandingkan dengan nilai k yang lainnya. Model ini mencapai akurasi sebesar 0,66, dan laporan klasifikasi menyajikan nilai-nilai seperti *precision, recall*, F1-score, dan *support*.

2. Pengujian Rasio 90%:10%

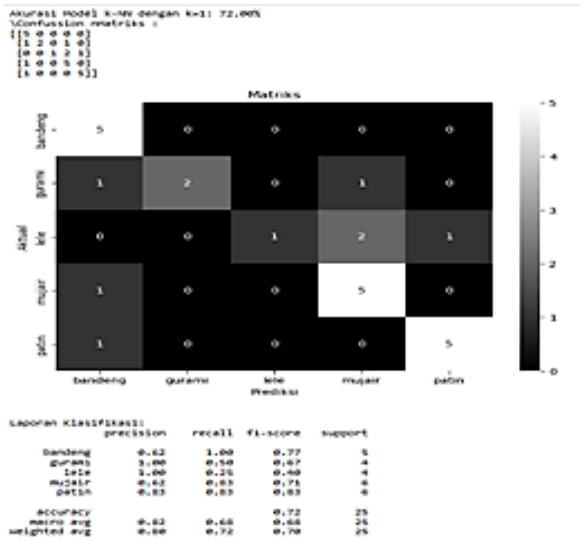
Pada pengujian klasifikasi menggunakan algoritma k-Nearest Neighbors (KNN) dengan rasio data 90:10, langkah utama yang dilakukan adalah membagi data menjadi dua bagian, yaitu data training sebanyak 90% dan data testing sebanyak 10%. Data training digunakan untuk melatih model KNN, sementara data testing digunakan untuk menguji performa model yang telah dilatih. Rasio 90:10 merupakan pilihan yang umum digunakan untuk memastikan model memiliki sejumlah data yang memadai untuk pembelajaran tanpa mengorbankan data yang cukup untuk pengujian.

Tabel 4. Hasil Klasifikasi Algoritma k-NN

Nilai K	Rasio	Akurasi
1	90:10	72,00%
2	90:10	56,00%
3	90:10	56,00%
4	90:10	56,00%
5	90:10	40,00%
6	90:10	52,00%
7	90:10	36,00%
8	90:10	48,00%
9	90:10	40,00%
10	90:10	44,00%

Tabel 4 memperlihatkan hasil klasifikasi untuk setiap nilai K dari 1 hingga 10 pada pembagian dataset dengan rasio 90:10. Terlihat bahwa akurasi tertinggi diperoleh pada K = 1, mencapai 72,00%. Oleh karena itu, *confusion matrix* dan laporan klasifikasi pada skenario 5 merupakan hasil klasifikasi dengan nilai k=1,

karena memberikan hasil yang lebih unggul dibandingkan dengan nilai k lainnya. Informasi terkait *confusion matrix* dan laporan klasifikasi dapat diakses melalui ilustrasi di bawah.



Gambar 9. *Confusion Matrix* dan *Classification Report*

Gambar 9 di atas menunjukkan *output* dari *confusion matrix* dan laporan klasifikasi dengan memanfaatkan model akurasi k = 1, yang menunjukkan performa yang lebih baik dibandingkan dengan nilai k yang lain. Model ini berhasil mencapai akurasi sebesar 0,72 dan menyajikan informasi lengkap dalam laporan klasifikasi, termasuk nilai *precision*, *recall*, F1-*score*, dan *support*.

4. KESIMPULAN

Dalam penelitian ini, perbandingan klasifikasi ikan menggunakan algoritma CNN dan k-NN telah dilakukan melalui serangkaian pengujian dengan pembagian *dataset* 80:20 dan 90:10. Hasil eksperimen menunjukkan bahwa algoritma CNN menghasilkan akurasi sebesar 88% pada rasio data 80:20 dan 80% pada rasio 90:10. Sebaliknya, algoritma k-NN memperoleh akurasi 66% pada rasio 80:20 dan 72% pada rasio 90:10. Dari hasil tersebut, dapat disimpulkan bahwa algoritma CNN memiliki kinerja yang lebih baik dalam mengklasifikasikan citra ikan dibandingkan dengan algoritma k-NN. Akurasi yang lebih tinggi pada kedua skenario pembagian data menunjukkan bahwa CNN mampu mengenali dan mengklasifikasikan jenis ikan dengan lebih akurat. Penting untuk dicatat bahwa performa algoritma dapat dipengaruhi oleh berbagai faktor, termasuk ukuran *dataset*, parameter pengujian, dan karakteristik citra yang digunakan. Meskipun demikian, hasil eksperimen ini memberikan indikasi bahwa CNN dapat menjadi pilihan yang lebih baik dalam konteks klasifikasi citra ikan.

REFERENSI

[1] S. Fauzi, P. Eosina, and G. F. Laxmi, "Implementasi Convolutional Neural Network Untuk Identifikasi Ikan Air Tawar," 2019.

[2] F. Y. Manik and K. S. Saragih, "Klasifikasi Belimbing Menggunakan Naïve Bayes Berdasarkan Fitur Warna RGB," *IJCCS Indones. J. Comput. Cybern. Syst.*, vol. 11, no. 1, p. 99, Jan. 2017, doi: 10.22146/ijccs.17838.

[3] D. Gunawan, "Metode Klasifikasi pada Data Preprocessing Data," no. 1, 2016.

[4] A. Pariyandani, D. A. Larasati, and E. P. Wanti, "Klasifikasi Citra Ikan Berformalin Menggunakan Metode k-NN dan GLCM," vol. 2, 2019.

[5] A. Amalia, M. S. Lydia, S. D. Fadilla, and M. Huda, "Perbandingan Metode Klaster dan Preprocessing Untuk Dokumen Berbahasa Indonesia," *J. Rekayasa Elektr.*, vol. 14, no. 1, pp. 35–42, Apr. 2018, doi: 10.17529/jre.v14i1.9027.

[6] G. F. Laxmi, P. Eosina, and F. Fatimah, "Analisis Perbandingan Metode Prewitt Dan Canny Untuk Identifikasi Ikan Air Tawar," 2017.

[7] I. Wulandari, H. Yasin, and T. Widiharih, "Klasifikasi Citra Digital Bumbu Dan Rempah Dengan Algoritma Convolutional Neural Network (CNN)," *J. Gaussian*, vol. 9, no. 3, pp. 273–282, Aug. 2020, doi: 10.14710/j.gauss.v9i3.27416.

[8] I. E. Hasym and I. Susilawati, "Klasifikasi Jenis Ikan Cupang Menggunakan Algoritma Principal Component Analysis (PCA) Dan K-Nearest Neighbors (KNN)," *KONSTELASI Konvergensi Teknol. Dan Sist. Inf.*, vol. 1, no. 1, pp. 168–179, Apr. 2021, doi: 10.24002/konstelasi.v1i1.4242.

[9] R. R. Waliyansyah and C. Fitriyah, "Perbandingan Akurasi Klasifikasi Citra Kayu Jati Menggunakan Metode Naive Bayes dan k-Nearest Neighbor (k-NN)," *J. Edukasi Dan Penelit. Inform. JEPIN*, vol. 5, no. 2, p. 157, Aug. 2019, doi: 10.26418/jp.v5i2.32473.

[10] B. Rahmat and B. Nugroho, *Pemograman Deep Learning Dengan Python*. Sidoarjo: Indomedia Pustaka, 2021.