

BAB II

TINJAUAN LITERATUR

2.1 Tinjauan Pustaka

Pendekatan literatur yang diambil dalam penelitian ini diperlukan untuk mengeksplorasi studi-studi terdahulu yang memiliki keterkaitan baik dari segi metode maupun objek penelitian. Beberapa studi terkait menerapkan algoritma Naïve Bayes, metode histogram, dan pendekatan kombinasi pengolahan citra lainnya untuk klasifikasi kualitas air dan membedakan objek visual lainnya. Kelima referensi melibatkan studi sebelumnya yang dilakukan oleh para peneliti guna memahami pendekatan yang digunakan, mengevaluasi kualitas performa metode, dan mengidentifikasi celah penelitian yang perlu diisi oleh penelitian saat ini.

Penelitian yang dilakukan oleh (Sutisna & Yuniar, 2023) berfokus pada pengklasifikasian kualitas air bersih menggunakan algoritma Naïve Bayes. Dataset diperoleh dari dua sumber, yaitu data sekunder melalui wawancara dan observasi, serta data primer yang dikumpulkan dari studi literatur. Penelitian ini memanfaatkan RapidMiner untuk membangun model klasifikasi yang mampu memprediksi kualitas air berdasarkan parameter terukur. Naïve Bayes dipilih karena kemampuannya yang sederhana namun efektif dalam memproses data dengan asumsi independensi antar atribut. Hasil pengujian menunjukkan akurasi sebesar

97,35%, yang membuktikan bahwa metode ini dapat digunakan sebagai alternatif untuk mempercepat penentuan kualitas air dibandingkan pengujian laboratorium konvensional yang memerlukan waktu lebih lama. Temuan ini relevan sebagai acuan penggunaan Naïve Bayes dalam penelitian klasifikasi kejernihan air yang mengandalkan data terukur.

(Suharyanto et al., 2021) meneliti peningkatan kualitas citra bawah air yang mengalami degradasi akibat hamburan dan penyerapan cahaya. Penelitian ini menggunakan pendekatan gabungan metode White Balancing berbasis Gray World, Gaussian Filtering untuk penajaman citra, serta Contrast Limited Adaptive Histogram Equalization (CLAHE) untuk optimasi kontras. Dataset yang digunakan berasal dari Underwater Image Enhancement Benchmark (UIEB) sebanyak 20 citra. Hasil evaluasi menggunakan parameter Peak Signal-to-Noise Ratio (PSNR) dan Root Mean Squared Error (RMSE) menunjukkan bahwa metode kombinasi tersebut mampu memberikan peningkatan kualitas citra yang signifikan dibandingkan penggunaan CLAHE murni. Peningkatan ini penting dalam penelitian berbasis citra, termasuk klasifikasi kejernihan air, karena kualitas citra yang baik akan meningkatkan akurasi proses ekstraksi fitur dan klasifikasi selanjutnya.

(Muchtar et al., 2024) mengembangkan metode klasifikasi kesegaran ikan berbasis analisis warna pada citra area mata ikan menggunakan algoritma Naïve Bayes. Pendekatan ini dimulai dengan segmentasi citra pada ruang warna RGB dan YCbCr, diikuti dengan

ekstraksi fitur entropi yang mengukur tingkat variasi intensitas piksel. Sebanyak 40 dataset citra digunakan untuk pelatihan dan pengujian model. Hasil penelitian menunjukkan akurasi sebesar 97,5%, menandakan bahwa kombinasi analisis warna dan fitur statistik citra sangat efektif untuk membedakan kelas objek. Walaupun objek penelitian adalah ikan, teknik ini sangat relevan untuk diaplikasikan pada citra air, terutama dalam ekstraksi dan analisis fitur warna sebagai dasar klasifikasi tingkat kejernihan.

Penelitian oleh (Sari, 2021) menerapkan algoritma Naïve Bayes untuk mengklasifikasikan kualitas air di wilayah DKI Jakarta menjadi empat kategori, yaitu baik, tercemar ringan, tercemar sedang, dan tercemar berat. Parameter yang digunakan meliputi pH, Dissolved Oxygen (DO), Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), dan beberapa indikator lainnya. Data dikumpulkan melalui observasi, wawancara, dan literatur. Meskipun hasil pengujian menghasilkan akurasi sebesar 50,6%, nilai yang tergolong rendah ini memberikan wawasan penting mengenai tantangan yang dihadapi dalam klasifikasi kualitas air di lingkungan nyata, seperti ketidakseimbangan data dan variabilitas parameter lingkungan. Studi ini memberikan pembelajaran tentang pentingnya kualitas dan jumlah data dalam keberhasilan penerapan Naïve Bayes.

(Istinah ; Mulyana, 2021) merancang sistem klasifikasi kejernihan air berbasis citra digital menggunakan metode Local Binary Pattern

Histogram (LBPH) yang diimplementasikan pada perangkat Android. Penelitian ini memanfaatkan parameter optimal berupa jumlah neighbors = 9, radius = 1, dan threshold BW = 180, dengan jarak pengambilan gambar 20–25 cm. Hasil pengujian menunjukkan akurasi sebesar 95,56% dengan waktu komputasi rata-rata 2,35 detik. Sistem ini mampu mendeteksi kejernihan air secara cepat dan akurat, namun sensitif terhadap jarak pengambilan gambar—citra terlalu dekat menjadi blur, sedangkan terlalu jauh menyebabkan objek tidak terdeteksi. Temuan ini penting karena menunjukkan efektivitas metode histogram dalam klasifikasi berbasis citra, yang relevan untuk penelitian kejernihan air dengan ekstraksi fitur warna.

Dari studi literatur tersebut, penulis mengambil kesimpulan bahwa algoritma Naïve Bayes cukup tepat dan berpotensi tinggi pada klasifikasi kualitas air, baik berdasarkan parameter fisika-kimia maupun citra digital. Teknik histogram dan pengolahan citra juga terbukti meningkatkan akurasi klasifikasi visual. Dengan menambahkan parameter pH dan TDS, pada penelitian ini penulis berusaha untuk pH dan TDS untuk mendukung klasifikasi kelayakan air secara lebih komprehensif dan terintegrasi, di mana pH dan TDS diukur sebagai parameter air.

2.2 Landasan Teori

2.2.1. Kejernihan Air

Kejernihan air merupakan salah satu indikator utama dalam menilai kualitas fisik air secara visual. Air yang jernih menandakan

sedikitnya partikel tersuspensi, sedimen, atau kontaminan lain yang bisa memengaruhi kesehatan manusia maupun ekosistem. Berdasarkan studi (Haryono et al., (2024) kejernihan air dapat dipengaruhi oleh faktor alami seperti endapan lumpur dan organisme air, maupun faktor buatan seperti limbah industri dan domestik. Secara umum, air jernih terlihat bening atau berwarna kebiruan, sementara air keruh akan cenderung berwarna kecoklatan, kekuningan, atau kehijauan. Kejernihan air sangat penting karena air yang keruh sering kali mengandung mikroorganisme patogen, logam berat, atau bahan kimia berbahaya lainnya.

2.2.2. Pengolahan Citra Digital dan Histogram

Pengolahan citra digital adalah teknik pemrosesan citra menggunakan perangkat lunak dan algoritma untuk memperoleh informasi dari sebuah objek visual (Marpaung et al., 2022). Dalam konteks air, citra yang ditangkap melalui kamera diproses untuk mengekstrak nilai warna yang bisa digunakan untuk menilai kejernihan. Salah satu pendekatan yang umum digunakan adalah analisis histogram warna, yaitu grafik distribusi intensitas piksel dalam kanal warna Red, Green, dan Blue (RGB). Setiap warna memiliki rentang nilai dari 0–255 piksel dan histogram menunjukkan berapa banyak piksel yang berada pada tiap tingkat intensitas. Histogram ini memberikan gambaran sebaran dan dominansi warna dalam gambar, sehingga sangat berguna untuk membedakan air jernih dan keruh secara visual (Ardiansyah, 2009).

Ekstraksi fitur dari histogram warna dilakukan dengan mengambil

nilai seperti rata-rata (mean), modus (nilai paling sering muncul), atau deviasi standar dari masing-masing kanal RGB. Fitur-fitur ini menjadi representasi numerik dari karakteristik visual air dan dapat digunakan sebagai input pada algoritma klasifikasi seperti Naïve Bayes. Menurut (Gonzalez & Woods, 2018), analisis histogram memberikan informasi yang cukup akurat untuk membedakan citra berdasarkan warna dominan, terlebih jika digabungkan dengan pengelompokan atau klasifikasi berbasis machine learning. Dengan demikian, pengolahan histogram warna merupakan metode yang relevan dan efektif dalam pendekatan klasifikasi kualitas air berbasis citra.

Pada pengolahan citra digital, kejernihan air dapat dideteksi dengan memanfaatkan warna dan intensitas cahaya pada citra air yang ditangkap oleh kamera. Warna yang dihasilkan oleh air berkaitan erat dengan indeks kejernihannya, karena partikel tersuspensi mempengaruhi reflektansi cahaya dan menghasilkan perbedaan spektrum warna. Oleh karena itu, pengamatan kejernihan air tidak hanya bisa dilakukan secara laboratorium, tetapi juga dapat dianalisis melalui pendekatan pengolahan citra yang lebih sederhana dan efisien untuk lingkungan lapangan atau daerah dengan fasilitas terbatas

2.2.3. Naïve Bayes

- **Pengertian Naïve Bayes**

Naïve Bayes merupakan algoritma klasifikasi berbasis probabilistik yang bekerja berdasarkan Teorema Bayes. Metode ini mengasumsikan

bahwa semua fitur bersifat independen terhadap satu sama lain yang membuatnya sangat sederhana dan cepat dalam melakukan proses klasifikasi. Dalam konteks penelitian ini, Naïve Bayes digunakan untuk mengklasifikasikan tingkat kejernihan air (jernih, agak keruh, keruh) berdasarkan fitur warna (nilai RGB) yang diperoleh dari histogram citra. Meski asumsi independensi fitur sering kali tidak terpenuhi secara sempurna di dunia nyata, penelitian oleh (Han et al., 2012) dan (Zhang et al., 2022) menunjukkan bahwa Naïve Bayes tetap memberikan performa yang kompetitif dalam klasifikasi data numerik, termasuk dalam konteks kualitas air.

Keunggulan utama Naïve Bayes adalah kemampuannya menangani dataset berukuran kecil, kecepatan dalam pelatihan model, dan toleransi terhadap data yang tidak lengkap. Dalam studi oleh (Herwindiati, 2024), Naïve Bayes tetap menghasilkan prediksi yang layak meskipun hanya menggunakan beberapa fitur utama. Hasil klasifikasi ini kemudian dapat diverifikasi atau diperkuat dengan data tambahan seperti nilai pH dan TDS, sehingga hasil akhir menjadi lebih akurat dan terpercaya.

- **Cara Kerja Naïve Bayes**

Cara kerja Naïve Bayes berfokus pada perhitungan peluang setiap kelas terhadap data yang diuji. Secara umum, prosesnya adalah sebagai berikut:

1. Menentukan Kelas

Sistem menentukan kategori yang tersedia, misalnya jernih, agak keruh, dan keruh.

2. Menghitung Probabilitas Awal (Prior Probability)

Probabilitas awal dihitung berdasarkan jumlah data pada masing-masing kelas di data latih.

3. Menghitung Probabilitas Fitur terhadap Kelas (Likelihood)

Setiap fitur, seperti nilai warna RGB, pH, dan TDS, dihitung peluang kemunculannya pada setiap kelas.

4. Menggabungkan Seluruh Probabilitas

Seluruh probabilitas fitur dikalikan dengan probabilitas awal kelas untuk memperoleh nilai akhir.

5. Menentukan Hasil Klasifikasi

Kelas dengan nilai probabilitas terbesar dipilih sebagai hasil klasifikasi.

Dengan mekanisme tersebut, sistem tidak menebak secara acak, melainkan berdasarkan peluang matematis dari data yang telah dipelajari sebelumnya.

- **Probabilitas pada Naïve Bayes**

Probabilitas merupakan konsep utama dalam metode Naïve Bayes. Probabilitas menunjukkan tingkat kemungkinan suatu peristiwa terjadi dengan nilai antara 0 sampai 1. Dalam konteks klasifikasi, probabilitas digunakan untuk menentukan seberapa besar peluang suatu data masuk ke

dalam kelas tertentu. Terdapat tiga jenis probabilitas utama dalam Naïve Bayes:

1. Prior Probability (Probabilitas Awal)

Peluang suatu kelas sebelum melihat data fitur. Nilai ini diperoleh dari perbandingan jumlah data tiap kelas dengan total data.

2. Likelihood (Probabilitas Bersyarat)

Peluang suatu fitur muncul pada kelas tertentu. Contohnya, seberapa sering nilai RGB tertentu muncul pada kelas air jernih.

3. Posterior Probability (Probabilitas Akhir)

Peluang suatu kelas setelah mempertimbangkan seluruh fitur yang dimiliki data uji. Nilai inilah yang digunakan untuk menentukan hasil klasifikasi.

Melalui pendekatan probabilitas ini, Naïve Bayes mampu memberikan keputusan yang bersifat kuantitatif dan terukur.

- **Rumus Naïve Bayes**

Secara matematis, Naïve Bayes dirumuskan sebagai berikut:

$$P(C|X) = \frac{P(X|C) \times P(C)}{P(X)}$$

Keterangan:

- **P(C|X)** : Probabilitas data X termasuk kelas C
- **P(X|C)** : Probabilitas fitur X pada kelas C
- **P(C)** : Probabilitas awal kelas C
- **P(X)** : Probabilitas keseluruhan data

Berdasarkan rumus tersebut, metode Naïve Bayes digunakan untuk menghitung peluang suatu data termasuk ke dalam kelas tertentu dengan mempertimbangkan probabilitas awal kelas dan probabilitas setiap fitur yang diamati. Seluruh nilai probabilitas fitur dikalikan untuk memperoleh nilai akhir masing-masing kelas, kemudian kelas dengan nilai terbesar dipilih sebagai hasil klasifikasi.

2.2.4. Parameter Fisik: pH dan TDS

Parameter fisik seperti pH (*potential of Hydrogen*) dan TDS (*Total Dissolved Solids*) merupakan indikator penting dalam menilai kualitas air secara ilmiah. pH menunjukkan tingkat keasaman atau kebasaan air, air minum yang layak konsumsi umumnya memiliki pH antara 6,5–8,5 (Quality, n.d.). Nilai pH yang terlalu rendah atau tinggi dapat menandakan adanya bahan kimia berbahaya atau pencemaran. Sementara itu, TDS menunjukkan konsentrasi zat padat terlarut dalam air seperti garam, logam berat, dan mineral organik. Air dengan nilai TDS tinggi cenderung memiliki rasa tidak enak, serta bisa menyebabkan endapan atau bahkan gangguan kesehatan jika dikonsumsi secara terus-menerus. Standar kualitas air WHO menyatakan bahwa TDS ideal berada di bawah 500 mg/L untuk air minum.

Dalam kaitannya dengan klasifikasi kejernihan air, nilai pH dan TDS bisa dijadikan sebagai **variabel verifikasi** terhadap hasil klasifikasi berbasis warna. Misalnya, jika citra air diklasifikasikan sebagai "jernih", tetapi nilai TDS menunjukkan angka di atas ambang batas (>1000 ppm),

maka dapat disimpulkan bahwa air tersebut hanya “tampak jernih” namun secara kimiawi kurang layak. Penelitian oleh (Pratami et al., 2025) menyatakan bahwa adanya korelasi antara kejernihan visual dengan parameter fisik: semakin keruh air, semakin tinggi nilai TDS dan seringkali juga terjadi pergeseran nilai pH ke ekstrem. Oleh karena itu, menggabungkan parameter pH dan TDS ke dalam sistem klasifikasi bukan hanya memperkuat hasil, tetapi juga memberikan validasi ilmiah terhadap sistem berbasis visual.