

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Plankton merupakan organisme mikroskopis yang hidup di perairan laut dan sungai serta memiliki peran yang penting dalam ekosistem akuatik [1]. Plankton terbagi menjadi dua jenis yaitu *phytoplankton* (plankton tumbuhan) dan *zooplankton* (plankton hewan) [2]. *Phytoplankton* adalah jenis yang bersifat autotrof yang menjadikannya sebagai produsen utama yang memberikan suplai energi bagi organisme lain yang berada di perairan [3]. Sebaliknya, *zooplankton* bersifat heterotrof yang artinya bergantung pada organisme lain untuk memperoleh sumber energi yang dibutuhkan [4]. Keberadaan plankton yang beragam dan melimpah dapat menjadi indikator kondisi kesehatan ekosistem laut dan turut berperan besar dalam siklus karbon global melalui proses fotosintesis dan respirasi [5].

Menurut penelitian yang dilakukan oleh Aditee Mitra [6], menyebutkan bahwa plankton memberikan kontribusi sekitar 50% dari total produksi primer di perairan global, hal tersebut menunjukkan betapa pentingnya plankton dalam mendukung produktivitas ekosistem akuatik. Selain itu, penelitian oleh Satria Cahya [7] menyebutkan bahwa variasi spesies plankton dapat digunakan sebagai bioindikator perubahan lingkungan laut seperti peningkatan suhu dan penurunan kualitas air akibat polusi.

Keanekaragaman spesies plankton di Indonesia sangat tinggi dengan lebih dari 2000 spesies yang telah diidentifikasi secara global. Keberagaman ini mencerminkan adaptasi plankton terhadap berbagai kondisi lingkungan yang berbeda baik itu suhu, salinitas maupun nutrisi yang mempengaruhi kelimpahan plankton [8]. Namun, klasifikasi spesies plankton secara tradisional dilakukan melalui pengamatan visual dibawah mikroskop oleh ahli biologi maritim. Metode ini sangat memakan waktu, tenaga kerja dan keahlian khusus untuk melakukan identifikasi. Tingkat subjektivitas pengamat juga dapat menyebabkan terjadinya variasi dalam hasil klasifikasi yang dapat mengurangi konsistensi dan akurasinya.

Dengan meningkatnya jumlah sampel plankton yang harus dianalisis sehingga kebutuhan akan metode klasifikasi yang lebih efisien dan akurat menjadi semakin mendesak [9].

Kemajuan teknologi dalam bidang pengolahan citra digital dan *machine learning* menawarkan solusi untuk otomasi proses klasifikasi plankton. Algoritma *machine learning*, terutama metode *supervised learning* telah menunjukkan kemampuan yang baik dalam melakukan identifikasi dan klasifikasi objek berdasarkan fitur-fitur yang diekstraksi dari citra [10]. Akan tetapi, penerapan *machine learning* dalam melakukan klasifikasi plankton menghadapi tantangan terkait dengan kompleksitas morfologi spesies plankton yang beragam [9].

Kondisi ini menghadirkan tantangan terhadap pembuatan model *machine learning* karena rentan terhadap *overfitting* dan kesulitan dalam belajar pola antar kelas yang beragam [10]. Selain itu, variasi morfologis antar spesies yang mirip seperti *Prorocentrum* dan *Thalassiosira* dapat menyulitkan algoritma dalam membedakan antar kelas secara akurat.

Untuk mengatasi hal tersebut, penelitian ini mengusulkan pendekatan klasifikasi menggunakan metode *ensemble learning*. Metode *ensemble learning* menggabungkan prediksi dari beberapa model *machine learning* untuk meningkatkan akurasi dalam melakukan klasifikasi terhadap data yang akan diujinya. Pendekatan ini diharapkan dapat meningkatkan performa klasifikasi spesies plankton [11].

Penelitian ini juga akan memperkuat kualitas dataset melalui tahap augmentasi data dan *feature engineering* untuk mengekstrak fitur morfologis dan tekstur yang relevan dari citra plankton. Augmentasi data dapat meningkatkan jumlah data per kelas dengan memodifikasi citra yang ada melalui teknik seperti rotasi, *flipping* dan *cropping*. Ekstraksi fitur ini diharapkan dapat menangkap karakteristik morfologis dan tekstur yang membedakan antar spesies plankton dengan lebih baik [11].

## 1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana cara mengoptimalkan proses segmentasi dan ekstraksi fitur dari citra plankton untuk memperoleh representasi numerik yang akurat?
2. Bagaimana penerapan metode *ensemble learning*, khususnya penggabungan SVM dan *random forest* dapat meningkatkan performa klasifikasi spesies plankton?
3. Bagaimana pengaruh berbagai skema evaluasi model *train-test split*, *cross-validation*, dan *gridsearch* terhadap performa klasifikasi?

## 1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Mengoptimalkan proses ekstraksi fitur dari citra plankton agar diperoleh representasi numerik yang akurat dan relevan.
2. Mengimplementasikan metode *ensemble learning* yang menggabungkan SVM dan *random forest* untuk meningkatkan akurasi dan stabilitas klasifikasi spesies plankton.
3. Mengevaluasi performa model klasifikasi berdasarkan berbagai pendekatan validasi *train-test split*, *cross-validation*, dan *gridsearch*.

## 1.4 Batasan Masalah

Untuk menjaga fokus dan kejelasan dalam penelitian, batasan masalah yang ditetapkan sebagai berikut:

1. Penelitian ini menggunakan dataset citra plankton dari dua sumber yaitu BRIN (Badan Riset dan Inovasi Nasional) dan WHOI (*Woods Hole Oceanographic Institution*) dengan total 2.472 citra melalui proses augmentasi.
2. Fokus klasifikasi dilakukan pada lima genus plankton, yaitu *Ceratium*, *Dinobryon*, *Guinardia delicatula*, *Prorocentrum*, dan *Thalassiosira*, dengan representasi data yang telah diseimbangkan per kelas.

3. Fitur yang diekstraksi dari citra terbatas pada fitur morfologis (*area, perimeter, aspect ratio, circularity, solidity*), *Hu Moments*, dan fitur tekstur berbasis GLCM (*contrast, energy, homogeneity, correlation*). Fitur warna tidak digunakan karena tidak representatif akibat variasi pencahayaan dan transparansi objek.
4. Segmentasi citra dilakukan menggunakan metode U-Net sebagai pendekatan utama, sementara metode *thresholding* dan *K-Means* hanya digunakan dalam percobaan awal dan tidak digunakan dalam hasil akhir.
5. Model klasifikasi yang digunakan terbatas pada *support vector machine* (SVM), *random forest*, dan *voting classifier* sebagai bentuk *ensemble learning*. Tidak dilakukan perbandingan dengan model *deep learning* atau model klasifikasi lainnya.
6. Evaluasi performa model dilakukan dengan beberapa pendekatan, yaitu *train-test split* (80:20 dan 95:5), *gridsearch* untuk pencarian parameter terbaik dan validasi silang (*5-fold* dan *10-fold cross-validation*).

## 1.5 Metodologi Penelitian

Metode penelitian ini disusun dengan tujuan untuk memastikan setiap tahap menghasilkan data dan analisis yang sesuai dengan tujuan penelitian. Tahapan penelitian yang dilakukan adalah sebagai berikut:

### 1. Pengumpulan Data

Data yang digunakan dalam penelitian ini berasal dari dua sumber yaitu BRIN (Badan Riset dan Inovasi Nasional) yang berlokasi di KST Samaun Samdikun dan WHOI (*Woods Hole Oceanographic Institution*). Data dikumpulkan dalam bentuk citra dari lima genus plankton dan diklasifikasikan ke dalam folder sesuai kelasnya.

### 2. Preprocessing Data

Pengolahan citra dilakukan untuk mempersiapkan citra awal agar dapat digunakan pada analisis selanjutnya. Proses ini bertujuan untuk meningkatkan kualitas citra serta memastikan bahwa citra yang digunakan dalam penelitian atau analisis memiliki format yang sesuai dan konsisten.

Langkah pertama dalam pengolahan citra adalah penyesuaian ukuran gambar. Hal ini dilakukan untuk memastikan bahwa semua citra memiliki dimensi yang seragam, sehingga memudahkan pemrosesan. Selain penyesuaian ukuran, proses pengolahan citra juga melibatkan teknik augmentasi data. Augmentasi data dilakukan untuk meningkatkan jumlah dan keragaman data yang tersedia, dengan cara memodifikasi citra yang ada melalui teknik-teknik seperti rotasi, *flipping* dan lainnya.

### **3. Augmentasi Data**

Augmentasi Data merupakan teknik untuk memperbanyak data dengan cara memanipulasi dan memodifikasi data yang sudah ada. Teknik ini sangat penting ketika bekerja dengan dataset kecil karena kurangnya data serta untuk memberikan banyak variasi terhadap data.

### **4. Ekstraksi Fitur**

Ekstraksi fitur bertujuan untuk mengekstrak informasi penting yang terkandung dalam citra, yang dapat digunakan untuk membedakan objek atau pola yang ada. Setiap citra plankton nantinya akan dilakukan ekstraksi fitur yang akan memberikan nilai berupa nilai matematis dan nantinya akan digunakan untuk melakukan pembuatan model *machine learning* tersebut.

### **5. Pelatihan Model Klasifikasi**

Pelatihan model dilakukan dengan memanfaatkan algoritma *support vector machine* (SVM) dan *random forest* untuk melakukan klasifikasi. Nantinya setiap algoritma akan melakukan pelatihannya masing-masing dan akan digabungkan untuk meningkatkan akurasi dari model tersebut. Data yang digunakan dibagi menjadi dua bagian, yaitu data pelatihan dan data pengujian untuk memastikan model dapat dilatih dan dievaluasi dengan baik.

### **6. Pengujian Model**

Pengujian model dilakukan untuk mengevaluasi hasil dari model yang sudah dibuat. Proses ini bertujuan untuk memberikan gambaran mengenai sejauh mana kinerja model dalam mengklasifikasikan plankton.

## **7. Hasil Klasifikasi**

Hasil penelitian ini diharapkan dapat membentuk model klasifikasi yang dapat membedakan spesies plankton berdasarkan dari fitur morfologi dan tekstur. Model ini akan menjadi dasar bagi sistem yang dapat membantu manusia dalam mengidentifikasi plankton.

### **1.6 Sistematika Penulisan**

Sistematika penulisan penelitian ini disusun untuk memberikan pemahaman yang jelas tentang keseluruhan proses penelitian yang dilakukan. Adapun sistematika penulisan penelitian ini adalah sebagai berikut:

#### **Bab I Pendahuluan**

Bab ini menjelaskan mengenai latar belakang yang mendasari dilakukannya penelitian, rumusan masalah yang akan diselesaikan dalam penelitian, tujuan yang ingin dicapai, serta batasan masalah untuk memperjelas cakupan penelitian.

#### **Bab II Tinjauan Pustaka**

Bab ini membahas penelitian-penelitian terdahulu yang relevan dengan topik yang diteliti. Tinjauan Pustaka berfungsi sebagai dasar teori yang penting untuk memahami latar belakang serta menyusun kerangka konseptual dalam penelitian ini.

#### **Bab III Metodologi Penelitian**

Bab ini menjelaskan tahapan-tahapan yang akan dilaksanakan dalam penelitian, dimulai dari pengumpulan data, teknik pengolahan citra, proses ekstraksi fitur serta metode pelatihan dan pengujian dalam model tersebut.

#### **Bab IV Pembahasan**

Bab ini akan menyajikan dimulai dari hasil perancangan hingga hasil dari penelitian yang diperoleh. Hasil penelitian dibahas secara mendetail untuk menjawab rumusan masalah dan membandingkannya dengan temuan dari penelitian terdahulu.

## **Bab V Penutup**

Bab ini menyajikan kesimpulan yang diperoleh dari penelitian yang sudah dilakukan berdasarkan hasil dan pembahasan yang telah diuraikan di bab sebelumnya. Selain itu, bab ini juga mengemukakan implikasi penelitian, kontribusinya terhadap perkembangan ilmu pengetahuan, serta memberikan saran untuk penelitian dimasa mendatang.

