

## **BAB II**

### **Tinjauan Pustaka**

#### **2.1 Tinjauan Penelitian Terdahulu**

Penelitian terdahulu terkait topik pembahasan yang di angkat sangat penting untuk dijadikan dasar serta sumber data. Hal ini bertujuan untuk mendukung penelitian yang sedang dilakukan. Selain tu untuk mengajukan gagasan atau argument diperlukan data yang relevan dengan penelitian sehingga dapat dikatakan valid. Tujuan lain dilakukannya studi terhadap penelitian terdahulu adalah untuk menggali kekurangan dan kelebihan suatu metode yang akan digunakan. Dengan begitu peneliti dapat menghindari atau mengatasi permasalahan terkait.

Oleh sebab itu sebelum memulai penelitian perlu dilakukan eksplorasi dan studi pustaka baik melalui internet maupun jurnal-jurnal yang relevan dengan topik penggalian data. Berikut penelitian terdahulu yang digunakan di dalam penelitian klasterisasi citra menggunakan metode *K-Means*. Disediakan pula tabel perbandingan antar – penelitian yang dapat dilihat pada Tabel 2.1

Penelitian pertama dilakukan oleh Sri Hadiani dan Dwiza Riana dengan judul “Segmentasi Citra Bemisia Tabaci Menggunakan Metode *K-Means*” [3]. Penelitian ini menggunakan metode *K-Means* untuk mendeteksi Bermisia Tabaci atau kutu daun pada citra daun. Penelitian ini melakukan pengolahan citra untuk mengetahui susunan warna daun sehingga dapat diketahui citra daun yang menguning disebabkan oleh hama kutu pada daun. Akan tetapi penelitian ini

masih memiliki kelemahan, yaitu objek bemisia tabaci belum dipisahkan dari citra daun sehingga klasterisasi citra hanya dapat mengetahui perubahan warna pada daun yang telah dihinggapi oleh hama tersebut.

Penelitian kedua dilakukan oleh Hyukmin Eum, Jaeyun Bae, Changyong Yoon, dan Euntai Kim dengan judul “*Ship Detection Using Edge-Base Segmentation and Histogram of Oriented Gradient with Ship Size Ratio*”[4]. Penelitian ini menggunakan metode *Support Vector Machine* (SVM) untuk mendeteksi tepi dan *histogram of oriented gradient* HOG dengan rasio ukuran kapal. Metode yang diusulkan dapat mencegah kecelakaan laut dengan mendeteksi kapal dalam jarak dekat. Penggunaan metode SVM merupakan salah satu bentuk penggunaan metode *machine learning* dalam pengenalan citra kapal yang dapat dijadikan referensi dalam penelitian ini. Tetapi pada penelitian terdahulu ini menggunakan tahap *supervised learning* sedangkan pada penelitian ini menggunakan metode *K-Means* yang menggunakan tahap *unsupervised learning*.

Penelitian ketiga dilakukan oleh Sandi Desmanto, Irwan, dan Renni Angreni dengan judul “Penerapan Algoritma *K-Means Clustering* untuk Pengelompokkan Citra Digital dengan Ekstraksi Fitur Warna *RGB*” [5]. Penelitian ini bertujuan menerapkan algoritma *K-Means* dalam pengelompokkan citra digital dengan ekstraksi fitur warna *RGB*. Dalam prosesnya dilakukan proses konversi citra ke matriks pixel, ekstraksi fitur warna *RGB*, dan proses *clustering*. Hasil dari matriks pixel dilakukan proses ekstraksi fitur warna *RGB* untuk mendapatkan

masing-masing nilai red, nilai *green*, nilai *blue* pada tiap citra. Kemudian dilakukan pengambilan nilai rata-rata serta proses *clustering*.

Penelitian keempat dilakukan oleh Nameirakpam Dhanachandra, Khumanthem Manglem and Yambem Jina Chanu dengan judul “*Image Segmentation using K-Means Clustering Algorithm and Subtractive Clustering Algorithm*” [6]. Banyak penelitian telah dilakukan di bidang segmentasi citra menggunakan pengelompokan. Ada berbagai metode dan salah satu metode yang paling populer adalah algoritma pengelompokan *K-Means*. Sebelum menerapkan algoritma *K-Means*, peningkatan peregangan parsial pertama diterapkan pada gambar untuk meningkatkan kualitas gambar. Metode pengelompokan *subtractive* adalah metode pengelompokan data di mana ia menghasilkan *centroid* berdasarkan nilai potensial dari titik data. Jadi *subtractive cluster* digunakan untuk menghasilkan pusat awal dan pusat ini digunakan dalam algoritma *K-Means* untuk segmentasi citra.

**Tabel 2. 1 Penelitian Terdahulu**

No.	Penulis	Judul	Latar Belakang	Hasil
1.	Sri Hadianti dan Dwiza Riana. (2018)	Segmentasi Citra Bemisia Tabaci Menggunakan Metode <i>K-Means</i>	Salah satu hama yang menyerang tanaman adalah Bemisia Tabaci atau kutu yang sering dikenal dengan sebutan Kutu Kebul. Kutu	pengolahan citra dari Bemisia Tabaci atau kutu kebul menggunakan metode <i>clustering K-</i>

			<p>Kebul menyerang bagian daun dari tanaman, akibat dari Kutu Kebul ini yang menyerang bagian daun ini bisa mengakibatkan daun menjadi menguning dan kering, sehingga jika dibiarkan maka daun-daun dalam tanaman tersebut berguguran dan tanaman tersebut bisa mati. Maka dari itu perlu dilakukan upaya untuk mendeteksi citra daun dari awal daun itu tumbuh, agar bisa diberikan tindakan yang mencegah</p>	<p><i>Means,</i> sebab metode ini lebih cocok diterapkan untuk mengelompokkan objek-objek yang terdapat pada suatu citra dengan mengelompokkannya sesuai jenis warna.</p>
--	--	--	---	---

			datangnya hama kutu kebul	
2.	Hyukmin Eum, Jaeyun Bae, Changyong Yoon, dan Euntai Kim. (2016)	<i>Ship Detection Using Edge-Base Segmentation and Histogram of Oriented Gradient with Ship Size Ratio</i>	kapal atau pengontrol pelabuhan mengamati laut untuk waktu yang lama dan melakukan tugas mencegah tabrakan laut saat bahan berbahaya ditemukan di laut. Namun, sulit bagi seseorang untuk mendeteksi semuanya hambatan di laut dan mengambil tindakan, dan ada dua rintangan untuk dideteksi.	Hasil Pertama adalah latar depan dipisahkan dari latar belakang dan kandidat terdeteksi menggunakan sobel tepi dan operasi morfologi di bagian segmentasi berbasis tepi. Kedua, fitur diekstraksi dengan menggunakan descriptor HOG dengan rasio ukuran kapal Ketiga, mesin

				vector dukungan (SVM) memverifikasi apakah kandidat adalah kapal.
3.	Sandi Desmanto, Irwan, dan Renni Angreni	Penerapan Algoritma <i>K-Means Clustering</i> untuk Pengelompokan Citra Digital dengan Ekstraksi Fitur Warna <i>RGB</i>	Banyaknya citra digital yang tersimpan di media penyimpanan membuat sulit dalam pengelompokan. Oleh karena itu, teknik pengelompokan citra digital diperlukan untuk mempercepat proses pengelompokan, salah satunya dapat berdasarkan fitur warna <i>RGB</i> .	Hasil dari penelitian ini didapatkan bahwa proses <i>KMeans</i> clustering dapat digunakan untuk pengelompokan citra digital dimana penentuan nilai <i>centroid</i> awal sangat berpengaruh terhadap hasil <i>cluster</i> .

			<p>Penelitian ini bertujuan untuk menerapkan algoritma <i>K-Means</i> dalam pengelompokkan citra digital dengan ekstraksi fitur warna <i>RGB</i>.</p>	
4.	<p>Nameirakpa m Dhanachan dra*, Khumanthe m Manglem and Yambem Jina Chanu. (2015)</p>	<p><i>Image Segmentation using K-Means Clustering Algorithm and Subtractive Clustering Algorithm</i></p>	<p>Segmentasi citra saat ini menjadi salah satu alat penting di bidang medis tempat ia digunakan untuk mengekstraksi atau wilayah menarik dari latar belakang. Jadi gambar medis tersegmentasi menggunakan teknik yang berbeda</p>	<p>Menghasilkan pusat awal dan pusat ini digunakan algoritma <i>K-Means</i> untuk segmentasi gambar. Akhirnya filter medial diterapkan ke gambar yang disegmentasi untuk menghapus</p>

			<p>dan <i>output</i> proses digunakan untuk analisis lebih lanjut dalam bidang medis. Tapi gambar medis dalam bentuk mentah mereka diwakili oleh <i>array</i> angka di komputer, dengan angka yang menunjukkan nilai-nilai kuantitas fisik yang relevan yang menunjukkan kontras di antaranya berbagai jenis bagian tubuh.</p>	<p>wilayah yang tidak diinginkan dari gambar.</p>
--	--	--	--	---



## 2.2 Landasan Teori

Pada penelitian ini terdapat beberapa teori dasar yang digunakan sebagai acuan dan referensi terkait pembahasan mengenai Analisis Sentimen. Sumber di dapat dari jurnal yang relevan dengan penelitian ini serta dari studi pustaka melalui internet.

### 2.2.1 Segmentasi Citra

Segmentasi citra adalah salah satu metode yang paling banyak digunakan untuk mengklasifikasikan piksel suatu gambar dengan benar dalam suatu keputusan aplikasi yang berorientasi. Ini membagi citra menjadi sejumlah daerah diskrit sehingga piksel memiliki kesamaan tinggi di masing-masing daerah dan sangat kontras antar daerah. Alat yang berharga di berbagai bidang termasuk perawatan kesehatan, citra pemrosesan, gambar lalu lintas, pengenalan pola dll. Ada berbagai teknik untuk segmentasi citra seperti ambang batas berbasis, berbasis tepi, berbasis kluster, berbasis jaringan saraf [6].

### 2.2.2 Citra

Citra adalah suatu representasi (gambaran), kemiripan, atau imitasi dari suatu objek. Citra terbagi 2 yaitu citra yang bersifat analog dan ada citra yang bersifat digital. Citra analog adalah citra yang bersifat *continue* seperti gambar pada monitor televisi, foto sinar X, dan lain-lain. Sedangkan pada citra digital adalah citra yang dapat diolah oleh komputer. [7]

Citra dapat didefinisikan sebagai fungsi  $f(x,y)$  berukuran  $M$  baris dan  $N$  kolom dengan  $x$  dan  $y$  adalah koordinat spasial, dan amplitude  $f$  di titik koordinat  $(x,y)$  dinamakan intensitas atau tingkat keabuan dari citra pada citra tersebut [7].

### 2.2.3 Klasterisasi

Klasterisasi adalah teknik *Machine Learning* yang melibatkan pengelompokan poin data. Diberikan satu set titik data, bisa menggunakan algoritma pengelompokan untuk mengklasifikasikan setiap titik data ke dalam kelompok tertentu. Secara teori, titik data yang berada dalam kelompok yang sama harus memiliki sifat dan atau fitur yang sama, sedangkan titik data dalam kelompok yang berbeda harus memiliki sifat dan atau fitur yang sangat berbeda. Klasterisasi adalah metode pembelajaran tanpa pengawasan dan merupakan teknik umum untuk analisis data statistik yang digunakan di banyak bidang [8].

Citra akan diubah menjadi matriks RGB yang masing-masing komponen warna memiliki nilai pixel masing-masing vektor RGB kemudian dari nilai piksel tersebut diambil nilai piksel masing-masing R, G, dan B. Nilai ini akan dijadikan sebagai atribut-atribut dalam perhitungan algoritma *K-Means*. Jika dua garis vektor adalah saling berdekatan, warna yang akan ditampilkan sangat berbeda, maka akan diambil jalan tengah dengan menghadirkan suatu warna secara kasar dari warna aslinya. Berikut untuk detilnya dari penjelasan di atas, bagaimana pilihan penampilan warna mempengaruhi hasil proses klasterisasi: [9].

1. Langkah yang pertama adalah menetapkan data set dari algoritma yang akan digunakan (*K-Means*), yaitu dengan melakukan pengambilan nilai acak dari k.
2. Kemudian, penampilan *RGB* dari tiap pixel diciptakan, dan menghasilkan dataset dalam 3-vektor.

3. Algoritma *K-Means* diterapkan pada dataset, menetapkan klusterisasi pusat k. Algoritma *K-Means* akan menghadirkan k warna untuk menggambarkan citra tersebut.
4. Tiap-Tiap piksel citra dikonversi dalam suatu garis vektor *RGB*, dan ditampilkan menggunakan rata-rata dari kelompok warna yang dihasilkan.

#### 2.2.4 Machine Learning

*Machine Learning* merupakan pendekatan dalam “kecerdasan Buatan” banyak digunakan untuk menggantikan atau menirukan perilaku manusia untuk menyelesaikan masalah atau otomatisasi. *Machine Learning* mencoba menirukan bagaimana proses manusia atau makhluk cerdas belajar dan mengeneralisasi. Setidaknya ada dua aplikasi utama dalam *Machine Learning* yaitu, klasifikasi dan prediksi [1].

Klasifikasi adalah metode dalam *Machine Learning* digunakan oleh mesin untuk mengklasifikasikan obyek berdasarkan ciri tertentu sebagaimana manusia mencoba membedakan benda satu dengan yang lain.

Prediksi digunakan oleh mesin untuk menerka keluaran dari suatu data masukan berdasarkan data yang sudah dipelajari dalam training.

#### 2.2.5 Metode Supervised Learning

*Supervised Learning* adalah salah satu metode yang terkait dengan *Machine Learning* yang melibatkan alokasikan data berlabel sehingga pola atau fungsi tertentu dapat disimpulkan dari data. Perlu dicatat bahwa *Supervised Learning* melibatkan pengalokasian objek *input*, vektor, sementara pada saat yang sama mengantisipasi nilai *output* yang paling diinginkan, yang sebagian besar

disebut sebagai sinyal pengawas. Properti garis bawah dari *Supervised Learning* adalah bahwa data input diketahui dan diberi label dengan tepat [10].

### 2.2.6 Metode Unsupervised Learning

*Unsupervised Learning* adalah metode kedua dari algoritma *Machine Learning* di mana kesimpulan diambil dari data *input* yang tidak berlabel. Tujuan *Unsupervised Learning* adalah untuk menentukan pola tersembunyi atau pengelompokan data dari data yang tidak berlabel. Sebagian besar digunakan dalam analisis data eksplorasi. Salah satu karakter yang menentukan *unsupervised learning* adalah bahwa *input* dan *output* tidak diketahui [10].

### 2.2.7 Metode K-Means

Algoritma untuk klusterisasi *K-Means* adalah sebagai berikut : [6]

1. Inisialisasi jumlah kluster  $k$  dan pusat.
2. Untuk setiap piksel gambar, hitung jarak Euclidean  $d$ , antara tengah dan setiap piksel gambar menggunakan relasi yang diberikan dibawah ini.

$$d = \| p(x, y) - C_k \|$$

3. Tetapkan semua piksel ke pusat terdekat berdasarkan jarak  $d$ .
4. Setelah semua piksel ditetapkan, hitung ulang posisi baru pusat menggunakan hubungan yang diberikan di bawah ini.

$$C_k = \frac{1}{k} \sum_{y \in C_k} \sum_{x \in C_k} p(x, y)$$

5. Ulangi proses ini hingga memenuhi toleransi atau nilai kesalahan.
6. Bentuk kembali piksel kluster menjadi gambar.

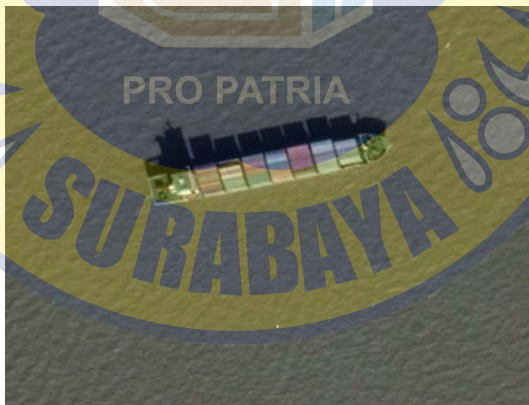
### 2.2.8 Kapal

Transportasi Laut yaitu suatu sistem pemindahan manusia dan Barang yang beroperasi di laut dengan menggunakan Alat sebagai kendaraan dengan bantuan tenaga manusia atau mesin [11].

[12] menurut bahwa tentang alat transportasi laut. Pada jenis-jenis kapal laut yang di bawah ini :

#### 1. Kapal Kargo

Kapal kargo barang dibuat dengan tujuan mengirimkan banyak kargo yang berisi barang dengan berbagai jenis dan macamnya. Barang-barang tersebut akan dikirimkan ke suatu tempat melalui jalur laut. Kapal kargo ini memiliki ukuran besar karena bisa menyimpan barang dan mengantarkannya sekali jalan.



Gambar 2. 1 Gambar Kapal Kargo

#### 2. Kapal Tanker

Kapal Tanker berfungsi untuk menyimpan minyak dan mengantarkannya ke suatu tempat atau tujuan tertentu. Baik mengangkut minyak mentah

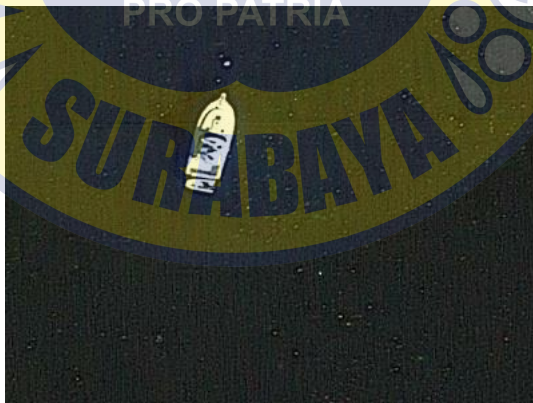
hasil pengeboran di lepas pantai ataupun minyak hasil olahan untuk didistribusikan ke banyak tempat.



Gambar 2. 2 Citra Kapal Tanker

3. Kapal kecepatan tinggi (*high speed craft*)

Kapal kecepatan tinggi sebagai untuk bisa mengantar orang atau barang. Kapasitas penumpangnya tergantung dari ukuran kapal. Namun yang terpenting, kapal tersebut termasuk alat transportasi laut modern.



Gambar 2. 3 Citra Kapal Kecepatan Tinggi