

BAB I

PENDAHULUAN

1.1. Latar Belakang

Kemampuan untuk mengenal atribut pejalan kaki seperti jenis kelamin, umur, gaya berpakaian, dan lain-lain menggunakan citra jarak jauh yang didapat dari kamera keamanan masih menjadi topik yang menarik untuk diteliti. Proses identifikasi atribut pejalan kaki ini dilakukan dengan menggunakan citra pejalan kaki seluruh badan dengan ketiadaan informasi penting seperti wajah, dan beberapa bagian tubuh. Beragamnya gaya berpakaian seseorang dan beberapa faktor lainnya yang memang tidak bisa dikendalikan seperti sudut penglihatan kamera, pencahayaan dan juga banyaknya citra yang berbeda untuk satu atribut yang sama masih menjadi masalah dasar yang dihadapi.

Dengan permasalahan dasar yang dihadapi, dibutuhkan sampel citra pejalan kaki yang cukup beragam dan metode ekstraksi fitur yang cukup kuat untuk melakukan identifikasi atribut pejalan kaki. Pada penelitian ini digunakan metode ekstraksi fitur berdasarkan warna dan tekstur masing-masing menggunakan *Color Histogram* dan *Multi Block Local Binary Pattern* (MB-LBP) dengan metode learning yang digunakan adalah *Adaptive Boosting* (Adaboost).

Warna dipilih karena warna merupakan salah satu komponen citra yang paling mudah dan banyak digunakan dalam merepresentasikan suatu citra. Menggunakan ekstraksi fitur berbasis warna dapat menghemat banyak waktu komputasi. Karena hal tersebut ekstraksi fitur berbasis warna paling banyak dipakai dalam bidang *system retrieval*. Ada banyak metode yang digunakan untuk mengekstraksi fitur berbasis warna seperti *Color Histogram*, *Color Moments*, *Color Correlogram*, *Color Coherence Vector* dan lain lain [1].

Color histogram dipilih karena komputasi yang sederhana dan juga metode ini melihat citra dari perspektif yang berbeda [2]. *Color histogram* merepresentasikan frekuensi distribusi warna pada suatu citra. Kelemahan dari *Color histogram* adalah sifatnya yang tidaklah unik. Kelemahan ini akan sangat berdampak ketika berhadapan dengan data yang jumlahnya sangat besar. Akan tetapi kekurangan dari *color histogram* itu sendiri yang menjadi kelebihan karena dapat mengatasi perubahan pada citra seperti rotasi, sudut penglihatan, dan pergeseran [2].

Untuk fitur tekstur MB-LBP dipilih karena dapat mengkodekan makro struktur pada suatu citra [3]. MB-LBP merupakan pengembangan

dari metode *Local Binary Pattern* (LBP). Jika MB-LBP dapat mengkodekan makro struktur pada suatu citra, maka *LBP* adalah kebalikannya. *LBP* hanya dapat mengkodekan mikro struktur dari suatu citra. MB-LBP dipilih karena selain dapat mengkodekan makro struktur dari suatu citra, MB-LBP memiliki komputasi yang cukup sederhana dan mempunyai toleransi yang cukup tinggi terhadap perubahan pencahayaan.

Selain membutuhkan metode ekstraksi fitur yang cukup kuat, diperlukan juga sebuah metode machine learning yang memiliki akurasi dan kecepatan yang cukup baik. Pada penelitian ini digunakan metode *machine learning Adaptive Boosting* (*Adaboost*). Metode ini pertama kali diperkenalkan oleh Freund dan Schapire pada tahun 1995 [4]. Dinamakan *adaptive* karena algoritma ini dapat menyesuaikan diri dengan data dan metode *classifier* lainnya [5]. Dinamakan *boosting* karena dapat mengurangi kesalahan dari *weak classifier* dan meningkatkan akurasi dari setiap algoritma pembelajaran.

Proses *learning* yang ada pada *adaboost* sangatlah bergantung pada *weak classifier*. *Weak classifier* sendiri adalah *base classifier* yang memiliki performa kurang baik, tetapi masih lebih baik dibandingkan dengan *random guessing* [4].

Base classifier yang digunakan pada penelitian ini adalah *decision stumps*. *Decision stumps* adalah *decision tree* sederhana yang hanya memiliki satu cabang. *Decision Stumps* dipilih karena metode *adaboost* sendiri sudah cukup dengan *base classifier* yang bisa memberikan tebakan yang sedikit lebih baik dari tebakan acak (nilai kebenaran lebih dari $\frac{1}{2}$).

Gabungan dari beberapa *weak classifier* akan menghasilkan *strong classifier*. Kelebihan dari metode ini adalah kita tidak perlu melakukan training terhadap sebuah *weak classifier* hingga menjadi *strong classifier*. Cukup dengan melakukan training terhadap beberapa *weak classifier* dan menggabungkannya menjadi sebuah *strong classifier*.

Sementara untuk *dataset* yang akan digunakan adalah *dataset PEdesTrian Attribute (PETA)*. *Dataset* ini dipilih karena dianggap sebagai *dataset* pejalan kaki terbesar yang pernah dibuat dengan jumlah sampel data mencapai 19000 dengan 61 atribut bernilai tunggal dan 4 atribut bernilai banyak. *Dataset* ini lebih besar dibandingkan dengan yang digunakan oleh Jianqing Zhu, Shengcai Liao, Zhen Lei, Dong Yi, Stan Z.



Li pada penelitian yang berjudul *Pedestrian Attribute Classification in Surveillance: Database and Evaluation* [6].

Permasalahannya adalah bagaimana pengaruh gabungan metode ekstraksi fitur warna dan tekstur terhadap akurasi setiap atribut pejalan kaki yang akan diidentifikasi. Banyaknya variasi gambar dan perbedaan jumlah distribusi positif dan negatif pada satu jenis atribut, kemungkinan akan menyebabkan akurasi yang dihasilkan pada tiap atribut berbeda.

Dengan adanya penelitian ini diharapkan, dapat mengetahui pengaruh gabungan metode ekstraksi fitur warna dan tekstur terhadap akurasi pada setiap atribut pejalan kaki yang akan coba diidentifikasi.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan diatas maka rumusan masalah dari penelitian ini adalah :

1. Bagaimana pengaruh gabungan fitur warna dan tekstur terhadap akurasi yang dihasilkan pada studi kasus identifikasi pejalan kaki.
2. Bagaimana pengaruh jumlah *weak classifier* terhadap akurasi yang dihasilkan pada studi kasus identifikasi pejalan kaki.

1.3 Tujuan Penelitian

Mengacu pada latar belakang serta perumusan masalah diatas maka tujuan dari penelitian ini adalah :

1. Menghitung akurasi dari *adaboost* dengan menggunakan fitur warna dan tekstur untuk identifikasi atribut pejalan kaki.
2. Mengetahui pengaruh jumlah *weak classifier* terhadap akurasi dari identifikasi atribut pejalan kaki.

1.4 Manfaat Penelitian

Berdasarkan latar belakang yang telah dipaparkan diatas didapatkan manfaat yaitu metode yang diusulkan dapat diimplementasikan pada suatu sistem keamanan.

1.5 Batasan Masalah

Batasan masalah pada penelitian ini adalah :

1. *Dataset* yang akan digunakan pada penelitian ini adalah *dataset PEdesTrian Attribute (PETA)* dengan 19000 jumlah citra yang digunakan.

2. Metode ekstraksi fitur berbasis tekstur yang digunakan adalah MB-LBP dengan ukuran deskriptor 3x3.
3. Metode ekstraksi fitur berbasis warna yang digunakan adalah histogram warna dengan ruang warna *hue*.
4. Atribut yang digunakan adalah atribut laki-laki, celana panjang, topi, membawa sesuatu, tas, sepatu.

1.6 Metode Penelitian

Secara umum, penelitian dilakukan dengan beberapa tahapan, yaitu:

a. Studi Pustaka

Tahap ini mengumpulkan seluruh informasi yang dibutuhkan serta mempelajari informasi-informasi tersebut sehingga dapat digunakan untuk membantu dalam menyelesaikan pengerjaan riset tugas akhir ini. Informasi yang dibutuhkan adalah seputar metode pengenalan objek pada citra dengan menggunakan metode Adaboost, *dataset PETA*, *color histogram*, MB-LBP, metode evaluasi *cross-validation* dan *Library Scikit-learn*.

b. Mempersiapkan Data

Pada tahap ini, persiapan dibagi menjadi 2 bentuk :

1. Mendapatkan *dataset* yang digunakan dalam penelitian.
2. Melakukan *preprocessing* label dataset ke dalam bentuk matriks label.

c. Membuat Fungsi Program

Tahap ini membuat setiap fungsi-fungsi yang digunakan. Seperti fungsi ekstraksi fitur, *color histogram* dan MB-LBP, hingga fungsi *cross-validation*. Membuat fungsi ini akan lebih mempermudah dalam pembuatan program untuk penelitian.

d. Uji Coba

Pada tahap ini, dilakukan uji coba terhadap fitur yang telah dibuat pada tahap sebelumnya. Uji coba juga dilakukan pada jumlah *weak classifier* terhadap akurasi dari metode pembelajaran.

e. Analisa dan Evaluasi

Pada tahap ini hasil dari uji coba sistem akan di evaluasi bagaimana akurasi dari proses klasifikasi. Jika hasil yang di dapat kurang dari apa yang diharapkan maka akan dilakukan evaluasi untuk sistem yang dibuat.



f. Dokumentasi

Dokumentasi dilakukan sebagai penyusunan laporan tugas akhir yang berlangsung mulai dari tahap studi pustaka hingga tahap evaluasi.

1.7 Sistematika Penulisan

Adapun sistematika pembahasan tugas akhir ini adalah sebagai berikut :

BAB I PENDAHULUAN

Pada bab ini berisikan latar belakang dari penelitian yang akan dilakukan, rumusan masalah, batasan masalah, tujuan dan manfaat penelitian, metode penelitian dan sistematika penulisan.

BAB II STUDI PUSTAKA

Pada bab ini dijelaskan mengenai materi-materi yang menjadi landasan dalam penyusunan penelitian ini. Materi-materi tersebut mengenai *adaboost*, *decision stumps*, *color histogram*, *MB-LBP*, *scikit-learn*, *dataset PETA*.

BAB III PERANCANGAN SISTEM

Pada bab ini akan menjelaskan rancangan dari seluruh sistem yang akan dibuat, mulai kebutuhan *software* dan *hardware* yang dibutuhkan, rancangan model *Adaboost*, *color histogram*, *MB-LBP* serta alur dari sistem yang akan dikembangkan.

BAB IV IMPLEMENTASI SISTEM

Pada bab ini akan dijelaskan implementasi sistem dari rancangan yang telah dibuat pada bab sebelumnya, dimana sistem akan diuji coba berdasarkan model dan skenario yang telah ditentukan.

BAB V PENUTUP

Pada bab ini berisi kesimpulan dari hasil uji coba serta saran yang dapat digunakan untuk pengembangan selanjutnya.