

1. Pendahuluan

1.1 Latar Belakang

Berkembang pesatnya teknologi tidak bisa dipungkiri dapat menyebabkan perubahan aktivitas di kehidupan sehari-hari. Salah satu contoh yang bisa diambil adalah tulisan tangan. Saat ini, media seperti komputer kerap kali digunakan untuk menggantikan tulisan tangan, namun bukan berarti tulisan tangan sudah ditinggalkan. Pada *Accounting Information System Essential Concept and Application 4th Edition*, Joseph W Wilkinson (1999) menjabarkan bahwa pemakaian komputer untuk pemrosesan informasi merupakan suatu keharusan karena teknologi komputer akan memberikan beberapa keuntungan, yaitu diantaranya pemrosesan yang menjadi lebih cepat dan sistem penyimpanan data akan menjadi lebih ringkas dan lebih mudah ketika dibutuhkan. Selain itu, juga dapat dikatakan bahwa dengan menggunakan media komputer akan menghasilkan tulisan yang bersifat umum dibandingkan tulisan tangan yang sifatnya beragam. Setelah itu, muncul masalah baru di mana tidak jarang tulisan yang sudah dicetak diharapkan dapat di-digitalisasi. Hal inilah yang menjadi landasan untuk dibangunnya suatu sistem yang dapat mengenali huruf komputer yang sudah dicetak.

Hingga saat ini sebenarnya sudah banyak dilakukan penelitian terhadap pengenalan karakter, mulai dari huruf latin, huruf hijaiyah, aksara india, hingga aksara-aksara di daerah-daerah Indonesia dan hal ini juga menghasilkan beragam hasil pula. Dalam pengenalan karakter, pemilihan metode untuk mengekstraksi ciri dan mengklasifikasi karakter memegang peranan yang sangat penting dalam menghasilkan akurasi pengenalan. Kombinasi kedua metode yang digunakan akan mempengaruhi akurasi yang dihasilkan sistem tersebut. Misalnya saja, pada penelitian yang dilakukan oleh Syilvia Vaulin [10] yang menggunakan metode *Chain Code* dan *Sequence Alignment* dalam pengenalan huruf komputer dinyatakan bahwa penerapan *Sequence Alignment* sebagai metode klasifikasinya mampu memberikan hasil yang lebih baik dibandingkan dengan penerapan KNN pada penelitian sebelumnya.

Pada tugas akhir ini dibangun suatu sistem pengenalan huruf komputer dengan menggunakan metode *Invariant Moment* yang dikombinasikan dengan *Image Partition* sebagai metode ekstraksi ciri dan *K-Nearest Neighbor* sebagai metode klasifikasi. *Invariant Moment* merupakan pengembangan dari metode *Centralized Moment* yang mampu mengatasi permasalahan transformasi, seperti *Scaling* dan *Rotation*. Metode yang dikembangkan oleh Hu ini sudah beberapa kali digunakan sebagai metode untuk ekstraksi ciri. *K-Nearest Neighbor* (KNN) sendiri merupakan salah satu metode *Supervised Learning* yang kerap digunakan sebagai metode klasifikasi terhadap pengenalan pola. Penelitian ini sudah pernah dilakukan oleh Syilvia Vaulin, dkk [10] yang menggunakan metode *Chain Code* dan KNN sebagai metode ekstraksi ciri dan klasifikasinya. Adapun akurasi rata-rata yang didapatkan sebesar 97.43% untuk jenis huruf *Tahoma*, 97.33% untuk huruf jenis *Verdana*, dan 78% untuk *Arial*. Pada penelitian lainnya, Hanifa Vidya Rizanti, dkk [5] memperoleh akurasi sistem terbaik sebesar 89,74% yang menggunakan *Invariant Moment* sebagai metode ekstraksi ciri yang dikombinasikan dengan *Case Based Reasoning* (CBR) dan *K-Nearest Neighbor* (KNN) sebagai metode

klasifikasi. Dengan mempertimbangkan penggunaan *Image Partition* yang dikombinasikan dengan *Invariant Moment* diharapkan nantinya dapat menghasilkan akurasi pengenalan huruf komputer yang lebih besar dari 90%.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang ada, maka dapat dirumuskan masalah sebagai berikut.

1. Bagaimana cara mengekstraksi ciri huruf komputer dengan menggunakan *Invariant Moment* ?
2. Bagaimana cara menerapkan *K-Nearest Neighbor* untuk mengenali huruf komputer ?
3. Berapa akurasi yang dihasilkan dari sistem?
4. Bagaimana pengaruh penggunaan *Image Partition* pada proses *Feature Extraction* terhadap akurasi yang dihasilkan?

Adapun yang menjadi batasan masalah adalah sebagai berikut.

1. Huruf komputer yang akan dikenali yaitu huruf latin ([a..z] dan [A..Z])
2. Ukuran huruf komputer 12, 16, 20, 24
3. Jenis *font* yang digunakan yaitu *Arial*, *Calibri*, dan *Times New Roman* sebagai data latih dan *font Tahoma*, *Verdana*, dan *Courier New* digunakan sebagai data uji
4. Sistem akan dibangun dengan menggunakan Matlab R2012a
5. Dalam satu citra masukan hanya terdapat satu jenis huruf komputer.

1.3 Tujuan

Adapun tujuan yang ingin dicapai adalah.

1. Membangun sistem pengenalan karakter komputer dengan menggunakan metode *invariant moment* untuk mengekstraksi ciri dan *K-Nearest Neighbor* sebagai metode klasifikasi
2. Menerapkan *K-Nearest Neighbor* sebagai metode klasifikasi pengenalan huruf komputer
3. Menganalisa performansi sistem pengenalan huruf komputer yang dibuat
4. Menganalisa pengaruh penerapan *Image Partition* dalam proses ekstraksi ciri terhadap akurasi sistem
5. Menganalisa faktor-faktor lainnya yang mempengaruhi performansi sistem

1.4 Hipotesa

Pada rencana tugas akhir ini akan digunakan metode *Invariant Moment* sebagai metode ekstraksi ciri dan metode *K-Nearest Neighbor* sebagai metode klasifikasinya. *Invariant Moment* merupakan pengembangan dari metode

Centralized Moments yang hanya konstan pada perubahan posisi namun tidak halnya dengan transformasi (*Translation Invariant*). *Invariant Moment* dikembangkan untuk mengakomodasi perubahan skala dan rotasi pada suatu objek. Metode ini dikembangkan pada tahun 1962 oleh Hu dan dikenal sebagai *Normalized Central Moments*. Metode ini sudah banyak diaplikasikan untuk mengekstraksi ciri atau pola suatu objek. Sebagai contoh, pada penelitian yang dilakukan oleh R.J. Ramtaka [4] *Invariant Moment* digunakan sebagai metode ekstraksi ciri untuk pengenalan tulisan tangan Devanagari Vowels dengan *Fuzzy Gaussian Membership Function* sebagai metode klasifikasinya. Nantinya *Invariant Moment* ini dikombinasikan dengan teknik *Image Partition* untuk memaksimalkan akurasi yang diperoleh. Hasil yang didapatkan cukup baik dengan tingkat akurasi sebesar 94,56%.

Sedangkan *K-Nearest Neighbor* merupakan algoritma *supervised learning* yang melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Adapun pengambilan keputusannya berdasarkan mayoritas dari kategori pada KNN yang mana berarti *class* yang paling banyak muncullah yang akan menjadi *class* hasil klasifikasi [8]. Algoritma ini dianggap memiliki keunggulan yang relatif tidak terpengaruh dari *error* dari data dan juga dapat digunakan dengan kumpulan data dengan jumlah yang besar. Algoritma ini juga sudah banyak diterapkan sebagai metode klasifikasi dalam pengenalan pola. Berdasarkan penjelasan di atas maka diharapkan sistem yang akan dibangun nantinya dengan *Invariant Moment* dan KNN dapat menghasilkan akurasi di atas 90%. Ditambah lagi dengan penerapan proses dilasi/erosi pada tahapan *preprocessing* dengan berbagai macam bentuk diharapkan dapat lebih meningkatkan akurasi sistem pengenalan huruf komputer yang dirancang.

1.5 Metodologi

Adapun metodologi yang digunakan adalah.

1. Studi Literatur

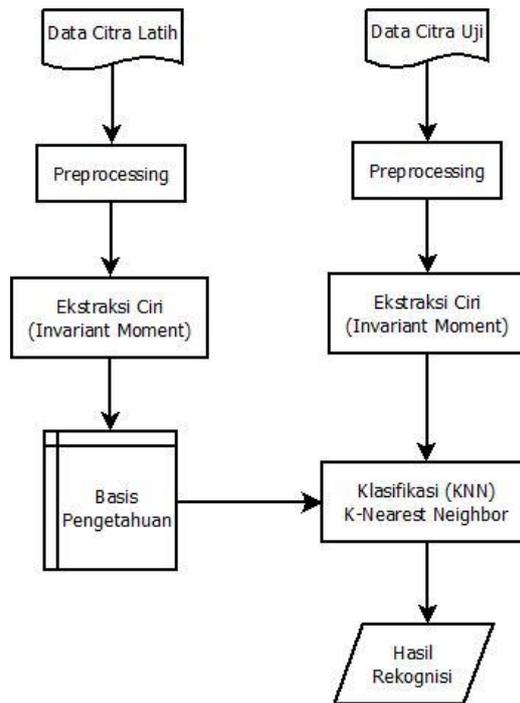
Pada tahap ini dilakukan pengidentifikasian masalah dan penentuan metode yang akan digunakan untuk menyelesaikan masalah berdasarkan penelitian-penelitian sebelumnya dan juga teori yang sudah ada.

2. Pengumpulan Data

Di tahap ini akan dilakukan pengumpulan data berupa huruf komputer untuk data uji dan data latih.

3. Perancangan Model

Pada tahap ini akan dirancang model sistem pengenalan huruf komputer. Berikut gambaran awal diagram blok yang akan dibangun



Gambar 1-1 Rancangan Diagram Blok Sistem

4. Implementasi

Pengimplementasian model yang sudah dirancang sebelumnya.

5. Pengujian dan Analisa Hasil

Pada tahap ini akan dilakukan pengujian terhadap hasil implementasi dan menganalisa akurasi serta faktor-faktor yang mempengaruhinya.

6. Pembuatan Laporan

Di tahap ini akan dilakukan penyusunan laporan dari sistem yang telah dibuat sesuai dengan aturan-aturan yang telah ditetapkan institusi.