

# 1. Pendahuluan

## 1.1 Latar belakang masalah

Teknologi yang kian berkembang pesat sangat banyak memberikan pengaruh terhadap perkembangan ilmu pengetahuan salah satunya adalah dalam bidang pengenalan pola. Sistem pengenalan pola banyak dikembangkan dan dimanfaatkan untuk membantu aktivitas manusia, seperti pengenalan wajah, pengenalan suara, sampai pada pengenalan tulisan. Selain itu pengenalan pola juga digunakan untuk membangun sistem lain dimana pengenalan pola itu digunakan. Pengenalan tulisan merupakan salah satu dari pengenalan pola yang umum dikenal orang. Seiring pesatnya perkembangan pengenalan tulisan banyak bermunculan teori dan algoritma baru yang dapat digunakan. Namun dari sekian banyak teori dan algoritma tidak semua teori mampu memberikan hasil yang maksimal tergantung dari jenis dan keunikan dari tulisan yang digunakan. Keunikan dari setiap tulisan dan sifat dari setiap teori menghasilkan sebuah permasalahan baru yang menarik untuk diangkat.

Indonesia kaya akan budaya yang beragam. Setiap daerah memiliki budaya yang berbeda dan tentunya memiliki kelebihan tersendiri dalam hal tulisan. Kekayaan ini menjadi penelitian tersendiri dimana dilakukan pengenalan terhadap pola dari tulisan daerah ini. Masing masing daerah mempunyai bentuk huruf yang beragam dan memiliki keunikan tersendiri. Pengenalan tulisan yang dibuat pada tugas akhir ini adalah pengenalan tulisan daerah bali. Dimana dalam tulisan Bali memiliki keunikan dari bentuknya yang hampir sama dan beberapa tulisan hanya dibedakan oleh satu guratan garis hal ini yang kadang menjadi permasalahan tersendiri dalam mempelajari tulisan Bali.

Pada Tugas Akhir ini pengenalan huruf Bali dikerjakan dengan menerapkan Algoritma berbasis *Vertex Chain Code (VCC)* sebagai teknik ekstraksi ciri dan Algoritma *Sequence Alignment* sebagai teknik klasifikasinya. Dasar dari teknik *Vertex Chain Code* adalah menelusuri garis batas dari objek pada citra, kemudian mengambil nilai jumlah sudut sel pada garis batas tersebut yang akan lebih mudah dilihat jika citra divisualisasikan kedalam *rectangular cell*[11]. Nilai-nilai ini akan membentuk suatu rantai kode. Rantai kode yang dihasilkan akan berbeda untuk setiap karakter sehingga diharapkan dapat memberikan hasil yang unik untuk tiap huruf yang dibaca dan hasil pengenalan huruf pun menjadi maksimal.

Pada TA ini teknik klasifikasi yang digunakan adalah *Sequence Alignment*. Karena berdasarkan pada [7] metode klasifikasi dengan *Sequence alignment* menghasilkan akurasi yang baik untuk mengklasifikasikan ciri citra yang berupa *chain code* atau sekuen-sekuen tertentu. *Sequence alignment* merupakan metode yang menyusun *sequence-sequence* huruf atau nilai tertentu (dalam bioinformatika digunakan untuk menyusun DNA, RNA, atau protein) untuk mencari kesamaan struktur antara 2 *sequence* [7]. Pada TA ini, algoritma *sequence alignment* diterapkan untuk mencari kesamaan antara 2 *vertex chain code*. Artinya, bila terdapat sebuah *vertex chain code* yang berasal dari sebuah huruf uji, *vertex chain code* ini akan diperiksa kesamaannya dengan semua *vertex chain code* pada basis pengetahuan. Metode *alignment* ini mencari sebanyak mungkin kode yang sama sehingga meskipun titik awal penelusuran *chain code* berubah, *vertex chain code* dengan sendirinya menyesuaikan diri untuk mencari pasangannya sehingga posisinya sejajar.

## 1.2 Perumusan masalah

Pada Tugas Akhir ini dilakukan Analisis dan Implementasi Pengenalan Huruf Bali menggunakan Algoritma berbasis *Vertex Chain Code* sebagai kerangka penelitiannya, sehingga dapat dirumuskan beberapa masalah antara lain :

1. Bagaimana membangun sistem yang mengenali huruf Bali yang ditulis tangan dengan output berupa pelafalan huruf yang benar berupa teks.
2. Bagaimana melakukan ekstraksi ciri menggunakan Algoritma berbasis *Vertex Chain Code*.
3. Faktor apa yang dapat mempengaruhi tingkat akurasi pengenalan huruf bali menggunakan algoritma berbasis *Vertex Chain Code* dan *Sequence Alignment*.

Adapun batasan masalah dalam Tugas Akhir ini adalah sebagai berikut :

1. Sistem yang dibangun adalah sistem yang sifatnya *offline*. Dimana input berupa image dari salah satu huruf Bali yang ditulis tangan.
2. Huruf bali yang digunakan adalah huruf Wianjana yang berjumlah 18 huruf.
3. File citra yang dipakai untuk training dan testing adalah dengan format \*.jpg dengan menggunakan background putih dengan ukuran citra bebas.
4. Citra input adalah citra huruf bali hasil scan dari huruf bali yang ditulis tangan dan tanpa noise.
5. Data citra terdiri dari 2 kelompok yaitu data training dan data uji. Data training diperoleh dari tulisan tangan 5 orang yang berbeda dan data uji diperoleh dari tulisan tangan 3 orang yang berbeda. Sehingga total data citra untuk data training berjumlah  $5 \times 18 = 90$  huruf dan untuk data uji berjumlah  $3 \times 18 = 54$  huruf. Dan total data citra berjumlah  $90 + 54 = 144$  citra huruf.

## 1.3 Tujuan

Tujuan penyusunan Tugas Akhir ini adalah :

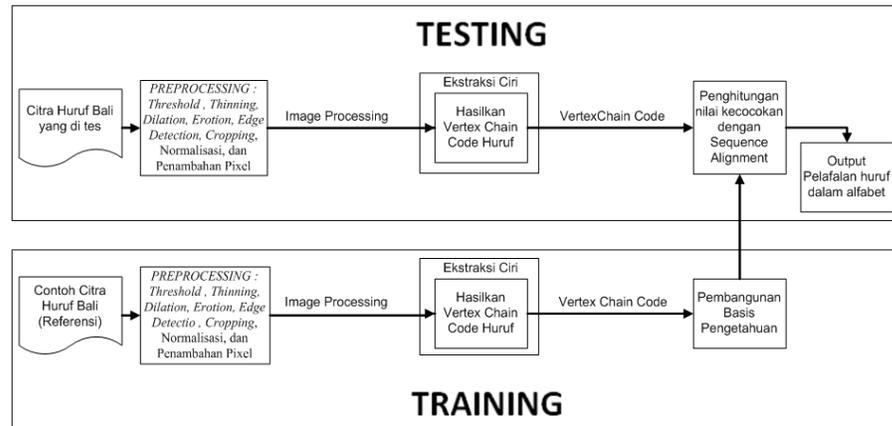
1. Menerapkan Algoritma berbasis *Vertex Chain Code* dan *Sequence Alignment* dalam sistem pengenalan huruf Bali.
2. Menganalisis tingkat akurasi dari sistem pengenalan huruf Bali.
3. Menganalisis faktor-faktor yang mempengaruhi tingkat akurasi dari sistem pengenalan huruf Bali.

## 1.4 Metodologi penyelesaian masalah

Metodologi penyelesaian masalah yang digunakan adalah :

- a. Study literature  
Tahapan ini bertujuan untuk mengumpulkan bahan-bahan, data, dan mendapatkan deskripsi yang jelas dan dasar teori yang kuat tentang *Vertex Chain Code*, *Sequence Alignment*, dan referensi Matlab sebagai simulasi software yang digunakan untuk membangun aplikasi pengenalan huruf Bali.
- b. Analisis dan desain  
Tahap ini meliputi analisis kebutuhan untuk merancang sistem pengenalan huruf Bali dengan *Vertex Chain Code*. Desain dari sistem menggunakan pendekatan struktural.

Gambaran proses pengenalan secara umum dapat dilihat pada bagan berikut:



Gambar 1-1 : Proses Pengenalan Huruf menggunakan Algoritma berbasis Vertex Chain Code

Gambar 1-1 didasarkan pada dasar teori dan konsep pengenalan objek[1]. Penggunaan algoritma dan metode dalam setiap proses juga disesuaikan dengan Tugas Akhir yang dikembangkan. Berbeda dengan[11] yang melakukan proses *thinning* dan transformasi ke rectangular cell, pada tahap preprocessing dilakukan proses *Threshold*, *Thinning*, *Dilation*, *Erosion*, *Edge Detection*, *Cropping*, Normalisasi, dan Penambahan Pixel terhadap image/citra input.

- c. Implementasi  
Tahap ini meliputi pembangunan sistem yang telah dirancang pada tahap sebelumnya. Pada tahap ini diimplementasikan perancangan yang telah dilakukan menjadi sebuah sistem dengan menggunakan software pemrograman Matlab.
- d. Testing dan Analisis Hasil  
Pada tahap ini dilakukan pengujian terhadap sistem yang telah dibangun dan sekaligus melakukan analisis terhadap hasil dari sistem. Output dari sistem ini dianalisis akurasi. Testing dilakukan dengan melihat akurasi dari output yang dihasilkan dengan cara menerapkan beberapa skenario yang dijelaskan lebih detail pada bab 4.
- e. Penyusunan laporan  
Pada tahap ini dilakukan penyusunan laporan hasil penelitian yang telah dilakukan dan membuat kesimpulan dari hasil penelitian tersebut.  
Laporan yang disusun terdiri dari:
  1. Pendahuluan  
Bab ini menjelaskan mengenai latar belakang penelitian, permasalahan yang diteliti, tujuan penelitian, dan metodologi penyelesaian masalah.
  2. Landasan Teori  
Landasan teori memberikan dasar-dasar pengetahuan yang dapat membantu penelitian.
  3. Rancangan dan Gambaran Sistem

Bab ini memberikan gambaran dan penjelasan mengenai tahap-tahap pengerjaan sistem pengenalan huruf menggunakan algoritma berbasis *vertex chain code* dan *sequence alignment*.

4. Pengujian dan Analisis

Bab ini berisi hasil-hasil penelitian dan pengujian sistem yang telah dibangun. Kemudian dijelaskan juga hal-hal yang dianalisis serta bagaimana hasil analisis dari pengujian yang dilakukan.

5. Penutup

Bab terakhir ini memberikan kesimpulan yang didapatkan berdasarkan penelitian yang telah dikerjakan. Selain itu diberikan saran-saran yang dapat membantu pelaksanaan penelitian berikutnya yang dapat dilakukan.

## 2. Landasan Teori

### 2.1 Huruf (Aksara) Bali

Tulisan Bali pada awalnya berpangkal dari aksara/huruf Dewa Nagari/Palawa dengan bahasanya sansekerta. Seiring dengan berkembangnya jaman maka tulisan bali mengalami beberapa perubahan seperti bentuknya dan ucapannya. Karena banyaknya perubahan pada tulisan dasar dari tulisan bali maka diadakanlah sebuah Pesamuhan Agung yang membahas perubahan-perubahan yang terjadi pada tulisan bali dan dibuatlah buku ejaan huruf bali dengan huruf latin[3].

Pesamuhan Agung dilakukan sebanyak 2 kali yaitu Pesamuhan Agung besar pada tahun 1957 yang bertujuan untuk mengeja Bahasa Bali dengan huruf Latin berdasarkan lontar Ramayana oleh Dr. Kern dan Baratayuda oleh Dr. Gunning dan Pesamuhan Agung Kecil pada tahun 1963 untuk meninjau kembali hasil keputusan Pesamuhan Agung Bahasa Bali tahun 1957, yaitu Ejaan Bahasa Bali dengan huruf Latin dan huruf Bali, karena di dalamnya masih terdapat kesalahan-kesalahan kecil dan beberapa kesukaran pada waktu mengajar, terutama di sekolah Dasar[2].

Dari keputusan Pesamuhan Agung Kecil tahun 1963 tersebut maka ditentukan huruf-huruf yang dibuat untuk menuliskan huruf Bali dengan huruf Latin dimana terdiri dari

- a. Aksara Suara (Vokal) : a,i,u,e,o
- b. Aksara Wianjana (Konsonan) : h, n, c, r, k, d, t, s, w, l, m, g, b, ng, p, j, y, ny

Pada Tugas Akhir mengambil studi kasus untuk mengenali Huruf Wianjana. Huruf Wianjana yang terdiri dari 18 huruf yang dapat ditunjukkan pada Gambar 2-1 :



Gambar 2-1 Huruf wianjana

### 2.2 Citra Digital

Citra (*image*) adalah gambar pada bidang 2 dimensi. Ditinjau dari sudut pandang matematis, citra merupakan fungsi yang kontinu dari intensitas cahaya pada bidang 2 dimensi. Sumber cahaya menerangi objek, objek memantulkan kembali sebagian dari berkas cahaya tersebut. Pantulan cahaya ini ditangkap oleh alat-alat optik, misalnya mata pada manusia, kamera, scanner, dsb, sehingga bayangan objek yang disebut citra tersebut terekam[8].

Citra digital adalah representasi dari citra dua dimensi menjadi sekumpulan nilai-nilai digital yang direpresentasikan sebagai fungsi intensitas cahaya dimana

x dan y menyatakan koordinat *spatial*[7]. Sebuah citra digital direpresentasikan dalam bentuk *array* pixel. Pixel merupakan elemen terkecil dari sebuah gambar atau citra. Ukuran pixel menentukan tingkat kehalusan sebuah citra. Semakin kecil ukuran sebuah pixel maka citra akan semakin halus.

Data digital pada umumnya diperoleh dengan cara melakukan proses kuantisasi dan sampling terhadap sinyal analog[10]. Sampling merupakan digitalisasi koordinat spatial, sedangkan Kuantisasi merupakan digitalisasi amplitudo[8]. Data citra dalam bentuk digital pada umumnya menyimpan informasi mengenai data intensitas dengan rentang nilai 0 -255 atau dalam bentuk bilangan riil 0 – 1 [10].

## 2.3 Sistem Pengenalan

### Pengenalan Pola

Pengenalan pola merupakan salah satu aplikasi pengolahan citra yang saat ini berkembang pesat. Pengenalan pola bertujuan untuk mengelompokkan data (pola-pola) berdasarkan pengetahuan yang sudah ada sebelumnya, atau informasi yang telah didapatkan dari pola-pola tersebut[7]. Beberapa teknologi pengenalan yang telah mendunia antara lain sistem OCR, pengenalan struktur jaringan dalam bidang biomedis dan ilmu kedokteran, pengenalan wajah dan iris mata untuk keperluan identifikasi dan keamanan, serta pengenalan barcode di aspek pemasaran.

Sistem pengenalan pola yang lengkap terdiri dari beberapa tahap yaitu: akuisisi pola, menerapkan mekanisme ekstraksi ciri yang akan mengukur informasi pola, dan melakukan skema klasifikasi yaitu berupa pengelompokan pola ke sebuah kelas tertentu berdasarkan hasil ekstraksi ciri. Dari ketiga tahap tersebut, proses pemilihan dan ekstraksi ciri yang benar merupakan permasalahan paling besar dalam pengenalan pola[7]. Sehingga dapat disimpulkan bahwa bila tahap ini mampu menghasilkan ciri yang benar, kemungkinan besar hasil pengenalan akan maksimal.

### Pengenalan Huruf

OCR adalah terjemahan citra tulisan baik dari tulisan tangan, hasil ketik, maupun hasil percetakan menjadi tulisan yang dapat diubah (*editable*) oleh mesin[7]. OCR merupakan bentuk *instance* dari *off-line character recognition* atau pengenalan karakter *off-line* dimana sistem akan mengenali karakter dengan bentuk statis. OCR sangat penting dalam perkembangan teknologi karena OCR merupakan teknologi dasar untuk mengembangkan aplikasi-aplikasi yang lebih besar.

## 2.4 Preprocessing

Sebelum melakukan langkah pengolahan citra, gambar-gambar yang telah ditangkap sebaiknya melewati tahap persiapan citra atau *preprocessing*. Tahap ini menjadi penting karena dapat menghasilkan gambar yang lebih seragam baik dari ukuran, warna, dan ketebalan sehingga bisa memperbaiki kinerja proses komputasi dan menyederhanakan langkah pengolahan citra berikutnya. Tahap ini meliputi 5 proses yaitu *threshold*, normalisasi, *cropping*, *edge detection* dan penambahan pixel.