

# 1 Pendahuluan

## 1.1 Latar Belakang

Huruf Arab merupakan salah satu huruf yang dipakai pada penulisan dokumen-dokumen rujukan utama di seluruh dunia [1] selain huruf latin. Contoh dokumen rujukan utama yang menggunakan Huruf Arab adalah *Al-Quran* yang juga menjadi kitab suci umat Islam. Dalam penulisannya *Al-Quran* modern menggunakan Huruf Arab yang *berharakat* (diakritik atau tanda baca yang menunjukkan cara pengucapan Huruf Arab) untuk meminimalisir kesalahan interpretasi.

Pengenalan pola huruf merupakan bagian dari bidang pembelajaran mesin yang mencoba memanfaatkan karakteristik unik pada huruf untuk dapat dinyatakan dalam serangkaian objek pengamatan, kemudian dikenali oleh suatu sistem. Inti dari sistem pengenalan pola huruf ini terdiri dari tahapan ekstraksi ciri dan klasifikasi.

Pendekatan pengenalan pola Huruf Arab yang sudah pernah dilakukan lebih menitikberatkan pada Huruf Arab yang tidak *berharakat*, dan dari hasil pengerjaan sebelumnya [16], dihasilkan akurasi sebesar 74,2857% untuk data *testing*, dan 76,1904% untuk data *training*. Untuk Huruf Arab yang *berharakat*, tantangannya bertambah dengan adanya beberapa variasi bunyi pada sebuah huruf yang sama. Oleh karena itu menjadi menarik untuk mencoba mengenali pola Huruf Arab *berharakat*, dimulai dari Huruf *Hijaiyah* (Huruf Arab yang tidak bersambung).

Adapun pengenalan pola Huruf *Hijaiyah berharakat* pada Tugas Akhir ini akan dilakukan dengan metode *Modified Direction Feature Extraction* (M-DFE) untuk tahap ekstraksi cirinya. M-DFE merupakan pengembangan dari DFE. Pada DFE, ciri yang diekstrak dari sebuah citra *digital* adalah nilai arah. Sedangkan pada M-DFE, selain diekstrak nilai arah (*direction feature*), juga diekstrak nilai transisi (*transition feature*) dari *background* ke *foreground* pada suatu citra [5]. Huruf *Hijaiyah* memiliki karakter goresan (*stroke*) yang tetap untuk semua huruf dan *harakat*, oleh karena itu M-DFE yang mengkombinasikan ciri arah dan transisi dari struktur global suatu huruf menjadi pertimbangan untuk diimplementasikan pada pengenalan pola Huruf *Hijaiyah*.

Untuk tahap klasifikasi akan digunakan *Hidden Markov Models* (HMM) yang telah berhasil diimplementasikan pada beberapa kasus pengenalan Huruf Arab

tidak berharakat, baik yang menggunakan akuisisi citra secara *online* [4], *offline* [2][3], maupun memakai *multilevel classifier* dalam proses klasifikasinya [17].

Hidden Markov Model merupakan sebuah model statistik, dimana model sistem diasumsikan sebagai Markov proses dengan parameter yang tersembunyi. Tujuannya adalah untuk menentukan parameter-parameter tersembunyi (*hidden*) tersebut berdasarkan parameter-parameter yang dapat diamati. Adapun pemodelan Huruf *Hijaiyah* dilakukan melalui proses estimasi berulang parameter HMM. Proses pengenalannya dilakukan dengan mencari huruf yang memberikan peluang terbesar terhadap kemunculan urutan pengamatan tersebut pada model HMM.

## 1.2 Perumusan Masalah

Berdasarkan pemaparan latar belakang di atas, dapat dirumuskan beberapa permasalahan terkait pengenalan pola Huruf *Hijaiyah* berharakat, diantaranya:

1. Tahapan *pre-processing* yang tepat sehingga pola huruf akan mudah di segmentasi antara huruf dengan *harakatnya*.
2. Melakukan segmentasi dan menentukan objek mana yang merupakan huruf utama, mana yang merupakan *harakat*.
3. Implementasi M-DFE dan HMM untuk mengenali pola Huruf *Hijaiyah* berharakat.
4. Membuat kelas-kelas untuk dapat mengklasifikasikan Huruf *Hijaiyah* yang berharakat.
5. Mengukur akurasi dari pengenalan pola Huruf *Hijaiyah* yang mengimplementasikan M-DFE dan HMM.

## 1.3 Tujuan

Tujuan yang ingin dicapai dari pengerjaan Tugas Akhir ini adalah:

1. Menganalisis dan mengimplementasikan M-DFE dan HMM untuk pengenalan pola Huruf *Hijaiyah* berharakat.
2. Menganalisis performansi sistem berdasarkan tingkat akurasi dari M-DFE dan HMM pada pengenalan pola Huruf *Hijaiyah* berharakat dari nilai *accuracy rate* dan waktu pemrosesan yang dihasilkan.

## 1.4 Ruang Lingkup dan Batasan Masalah

Adapun batasan masalah dalam pengenalan pola Huruf *Hijaiyah* berharakat ini adalah:

1. Pengenalan pola huruf dilakukan pada Huruf *Hijaiyah*, yakni Huruf Arab yang tidak bersambung (*isolated*) seperti yang ditunjukkan pada gambar 1.1.
2. *Lam-alif* dan *hamzah* tidak termasuk dalam cakupan pengujian dengan alasan bahwa *lam-alif* merupakan gabungan/*ligature* dari dua buah huruf (*lam* dan *alif*), sedangkan *hamzah* merupakan huruf semi konsonan, yang jika diberi tambahan *harakat* akan memiliki bunyi sama dengan huruf *alif*.
3. Huruf *Hijaiyah* yang dianalisis mengikuti gaya *Khat Naskhi*.
4. Input yang ditangani adalah huruf-huruf yang tegak dan sejajar, tidak saling bertumpuk (*overlapping*) ataupun berbentuk kaligrafi (tulisan indah).
5. Pola-pola huruf yang digunakan sebagai input untuk proses *training* dan *testing* berupa Huruf *Hijaiyah* dari tulisan tangan (*handwritten*).
6. Data yang digunakan sebanyak 7 set tulisan tangan Huruf *Hijaiyah* yang diperoleh dari 7 orang responden yang berbeda-beda. Satu set terdiri dari 2 buah *sheet* yang masing-masing memuat 28 Huruf *Hijaiyah* untuk setiap *harakat*.
7. *Harakat* yang ditangani terdiri dari *fathah* ( <sup>َ</sup> ), *dhomah* ( <sup>ُ</sup> ), dan *kasrah* ( <sup>ِ</sup> ).

ح ح ح - ج ج ج - ث ث ث - ت ت ت - ب ب ب - ا ا ا  
س س س - ز ز ز - ر ر ر - د د د - ذ ذ ذ - خ خ خ  
ط ط ط - ض ض ض - ص ص ص - ش ش ش  
ك ك ك - ق ق ق - ف ف ف - غ غ غ - ع ع ع - ظ ظ ظ  
ي ي ي - ه ه ه - و و و - ن ن ن - م م م - ل ل ل

Gambar 1-1: Huruf *Hijaiyah* Berharakat

## 1.5 Metodologi Penyelesaian Masalah

1. Identifikasi Masalah

Melakukan identifikasi permasalahan yang akan dibahas, yakni mengenai pengenalan pola Huruf *Hijaiyah* berharakat. Kesulitan-kesulitan apa yang

akan muncul, bagaimana pengerjaan terdahulu terhadap permasalahan yang satu tema, batasan masalahnya, serta tujuan apa yang ingin dicapai.

## 2. Studi Pustaka dan Literatur

Mencari referensi mengenai pengenalan pola huruf Arab yang sudah pernah dilakukan sebelumnya (*State of The Art*) baik dari paper maupun jurnal ilmiah. Selain itu dilakukan pula studi mengenai metode M-DFE dan HMM, dari prosesnya hingga pertimbangan mengapa metode tersebut akan tepat diterapkan untuk dapat mengenali pola Huruf *Hijaiyah* berharakat.

## 3. Analisis Kebutuhan dan Perancangan Sistem

Melakukan analisis terhadap kebutuhan sistem dalam mengenali pola Huruf *Hijaiyah* berharakat, kemudian merancang sistem mulai dari tahapan *pre-processing*, ekstraksi ciri, klasifikasi, hingga skenario *testing*nya.

## 4. Konsultasi dan Diskusi

Melakukan konsultasi mengenai sistem penulisan Huruf Arab kepada ahli, dalam hal ini ustad yang memiliki spesialisasi di bidang *Khatul Huruf*.

## 5. Pengumpulan Data

Pengumpulan data tulisan Huruf *Hijaiyah* dari beberapa penulis yang dipilih berdasarkan kualifikasi tertentu dalam menulis Huruf *Hijaiyah*. Dataset ini nantinya akan dibagi ke dalam data *training* dan data *testing*.

## 6. Implementasi Sistem.

Sistem yang sudah dirancang sebelumnya mulai diimplementasikan pada tahap ini. Mulai dari membangun tahapan *pre-processing*, kemudian mengimplementasikan M-DFE untuk mengekstraksi ciri, hingga akhirnya mengimplementasikan HMM untuk melakukan klasifikasi.

## 7. Pengujian Sistem dan Analisis.

Melakukan uji coba terhadap sistem sesuai dengan skenario *testing*, kemudian melakukan pengolahan data dari proses uji coba tersebut untuk selanjutnya dianalisis.

## 8. Penyusunan Laporan Tugas Akhir.

Mendokumentasikan semua tahapan pengerjaan tugas akhir, mulai dari dasar teori yang digunakan, implementasi sistem, hingga hasil pengujiannya.