

Abstrak

Huruf Arab merupakan salah satu sistem penulisan yang banyak dipakai pada penulisan dokumen-dokumen rujukan utama di seluruh dunia selain Huruf Latin. Di bidang pengenalan pola, Huruf Arab sudah cukup banyak mendapatkan perhatian, namun pendekatan-pendekatan yang sudah pernah dilakukan hingga saat ini baru mencakup pengenalan pola Huruf Arab yang tidak berharakat. *Harakat* adalah diakritik atau tanda baca yang menunjukkan cara pengucapan Huruf Arab untuk meminimalisir kesalahan interpretasi. Contoh dokumen rujukan utama yang menggunakan Huruf Arab berharakat adalah kitab suci umat Islam, *Al-Quran*.

Pada Tugas Akhir ini, dilakukan pengenalan pola Huruf *Hijaiyah* (Huruf Arab yang tidak bersambung) berharakat dengan menggunakan metode ekstraksi ciri *Modified Direction Feature Extraction* (M-DFE) dan metode klasifikasi *Hidden Markov Models* (HMM). Secara morfologis, Huruf *Hijaiyah* memiliki goresan (*stroke*) dengan arah tetap. Karena M-DFE menyimpan ciri arah dan transisi dari struktur global suatu karakter, maka dengan memakai metode M-DFE untuk ekstraksi ciri akan dihasilkan vektor ciri yang bagus untuk setiap Huruf *Hijaiyah*. Sedangkan metode klasifikasi HMM sendiri sudah berhasil diimplementasikan pada beberapa kasus pengenalan Huruf Arab tidak berharakat, baik yang menggunakan akuisisi citra secara *online*, *offline*, maupun memakai *multilevel classifier* dalam proses klasifikasinya.

Menangani *harakat* pada Huruf *Hijaiyah* menjadi suatu tantangan tersendiri. Tambahan *harakat* pada Huruf *Hijaiyah* memperbesar jumlah kelas dan secara langsung mempengaruhi tingkat akurasi sistem serta waktu prosesnya. Untuk mengatasi hal ini, dilakukan segmentasi terhadap *harakat* dan huruf utama pada citra input yang kemudian diproses secara terpisah. Dengan dilakukan segmentasi dipadukan dengan *rule-based clustering*, jumlah kelas dapat diminimalisasi sehingga waktu pemrosesan dapat berlangsung jauh lebih cepat dengan akurasi yang juga cukup baik. Setelah dilakukan beberapa skenario pengujian, diperoleh tingkat akurasi terbaik 69,4% pada *maximum transition* berjumlah 4 dan *state* HMM berjumlah 70.

Kata Kunci: *Huruf Hijaiyah, Harakat, Pengenalan Pola, Modified Direction Feature Extraction, Hidden Markov Models.*