

MENDETEKSI KEBENARAN PENGGUNAAN KATRA BENDA DAN KATA SIFAT PADA HURUF ARAB BERSAMBUNG DENGAN MENGGUNAKAN METODE TEMPLATE MATCHING

Kemas Aditya¹, Ahmad Suryan²

¹Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Ilmu nahwu dan sharaf merupakan salah satu ilmu yang mengatur proses pembentukan struktur kata yang baik dan benar pada suatu kalimat dalam bahasa Arab. Adapun jenis kata atau kelas kata secara umum terdiri atas tiga yaitu Al-Harfu, Al-Ismu (kata benda/sifat) dan Al-Fi'lu (kata kerja). Jenis kata yang akan menjadi objek dalam penelitian ini adalah kata benda dan kata sifat. Untuk mengetahui apakah struktur kata benda dan kata sifat dituliskan secara benar dapat diketahui melalui ciri - cirinya.

Pada penelitian ini, akan dilakukan pemeriksaan kebenaran penulisan kata benda dan kata sifat dengan melihat ciri - ciri dari kata tersebut. Metode yang digunakan adalah template matching. Template Matching adalah salah satu teknik dalam pengolahan citra digital yang berfungsi untuk mencocokkan tiap-tiap bagian dari suatu citra dengan citra yang menjadi template (acuan).

Metode template matching sering digunakan untuk mengidentifikasi citra karakter huruf, angka, sidik jari (fingerprint) dan aplikasi-aplikasi pencocokan citra lainnya.

Pada penelitian ini metode template matching digunakan untuk memeriksa kebenaran struktur pembentukan kata benda dan kata sifat pada huruf arab bersambung dengan melihat pola yang menjadi ciri dari pembentuk kata tersebut. Selanjutnya, dilakukan pengujian terhadap citra uji berupa kata benda dan kata sifat yang dituliskan secara benar maupun salah. Berdasarkan hasil pengujian, metode template matching memberikan hasil akurasi rata-rata sebesar 90 %.

Kata Kunci : kata benda, kata sifat, nahwu dan sharaf, template matching

Abstract

Nahwu and Sharaf is one of science that governs the formation of the structure of word correctly in a Arabic sentences. The type or classes of words are generally divided into three, namely Al-Harfu, Al-Ismu (noun/adjective) and Al-Fi'lu (verb). Types of the word that will be the object of this research is noun and adjectives. To determine whether the structure of nouns and adjectives written correctly can be determined by its characteristics.

In this research, nouns and adjectives will be verified by looking at the characteristics of the word. The method used is template matching. Template matching method is one of the techniques in digital image processing used to match each part of an image with image of the template (reference). Template matching methods often used to identify the image of the character letters, numbers, fingerprint and the other image matching application.

In this research, template matching method is used to verify the structure of forming nouns and adjectives in Arabic. Furthermore, testing the nouns and adjectives which are written correctly or incorrectly. Based on the test results, the template matching method yield an average accuracy of 90%.

Keywords : nouns, adjectives, nahwu and sharaf, template matching

1. Pendahuluan

1.1 Latar Belakang Masalah

Bahasa Arab adalah salah satu bahasa Internasional yang sekarang banyak digunakan oleh penduduk di dunia terutama di negara-negara bagian Timur Tengah. Bahasa Arab memiliki kedudukan yang istimewa diantara bahasa – bahasa yang lain di dunia karena sebagai bahasa Al-Qur'an dan Hadist serta kitab-kitab lainnya.

Tatkala Islam berkembang bahasa Arab mulai bercampur dengan bahasa-bahasa asing sehingga berdampak timbulnya banyak kekeliruan dan kesalahan dalam penggunaan bahasa Arab tersebut. Keadaan seperti ini mendorong para sahabat Rasulullah untuk membuat kaidah berbahasa sebagai rujukan dalam menggunakan bahasa Arab yang baik dan benar salah satunya adalah ilmu *Nahwu* dan *Sharaf*.

Dengan ilmu *nahwu sharaf*, dapat diketahui proses pembentukan suatu kata atau kalimat sesuai dengan kaidah penulisan yang baik dan benar. Dalam pembentukan kata, ada beberapa hal yang perlu diperhatikan seperti perubahan *harokat* akhir suatu kata dan tetapnya *harokat* akhir suatu kata dalam satu keadaan. Kata dengan arti yang sama bisa memiliki bentuk penulisan yang berbeda terutama dalam pemberian *harokat* akhir dari kata tersebut.

Bagi orang – orang yang belum memahami ilmu *nahwu sharaf* akan mengalami kesulitan untuk menentukan *harokat* akhir suatu kata beserta perubahan bentuk yang ada di dalamnya. Untuk menentukan bentuk dan *harokat* akhir suatu kata dapat diketahui dengan melihat ciri – ciri dari kata tersebut. Setiap kata memiliki ciri masing – masing, baik itu kata benda, kata sifat, kata kerja dan kata keterangan. Berdasarkan hal ini tercetuslah ide untuk memanfaatkan teknologi kecerdasan buatan (*Artificial Intelligence*) untuk mengamati pola pembentukan kata khususnya kata benda dan kata sifat.

Metode *Template Matching* dipilih untuk menyelesaikan kasus ini. *Template Matching* cocok digunakan untuk melakukan pengenalan objek-objek yang

memiliki ketetapan pola dan ciri. Metode *Template Matching* yang digunakan terdiri dari dua algoritma penyelesaian yaitu algoritma *Cross Correlation* dan *Correlation Coefficients*.

Template dibuat dengan mencari sejumlah variasi bentuk kata benda dan kata sifat. Dari *template-template* tersebut diambil ciri khas untuk setiap pembentukan kata. Selanjutnya *template* akan dicocokkan dengan data uji. Data uji berupa ragam bentuk penulisan kata benda dan kata sifat yang dituliskan secara benar maupun tidak menurut kaidah penulisannya. Hasil akhir dari penelitian ini adalah menentukan apakah data yang diujikan benar atau tidak berdasarkan kaidah – kaidah penulisan yang terdapat di dalam ilmu *nahwu sharaf* dan juga menganalisis hasil akurasi dan performansi dari sistem.

1.2 Perumusan Masalah

1. Bagaimana proses pengambilan ciri dari setiap kata yang diuji.
2. Bagaimana menentukan *template* yang digunakan untuk mencocokkan setiap ciri dari kata yang diuji.
3. Bagaimana proses pendeteksian kebenaran penggunaan kata benda dan kata sifat dengan menggunakan metode *template matching*.
4. Bagaimana mengukur tingkat akurasi keberhasilan dari data yang diuji.
5. Bagaimana menghitung *persentase* kecocokan citra data uji dengan citra *template*.

1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut :

1. Data citra uji menggunakan *file* gambar (*.jpg*) yang diambil dari tulisan Arab yang diketik menggunakan program *Arabic Pad 1.4*.
2. Tulisan arab menggunakan warna hitam dengan latar tulisan menggunakan warna putih
3. Data uji yang digunakan adalah sebagai berikut :
 - a. Jenis huruf *Traditional Arabic*, ukuran *font* 100, sebanyak 100 kata
 - b. Jenis huruf *Traditional Arabic*, ukuran *font* 200, sebanyak 100 kata

- c. Jenis huruf *Traditonal Arabic*, ukuran *font* 80, sebanyak 100 kata
 - d. Jenis huruf *Simplified Arabic*, ukuran *font* 200, sebanyak 100 kata
 - e. Jenis huruf *Arabic Typesetting*, ukuran *font* 150, sebanyak 100 kata
4. Proses pendeteksian menggunakan satu kata.
 5. Kata yang digunakan adalah kata benda dan kata sifat.
 6. Metode yang digunakan adalah *Template Matching* dengan dua jenis algoritma penyelesaian yang berbeda yaitu, *Correlation Coefficients* dan *Cross Correlation*.
 7. Data citra *template* merupakan hasil pemotongan secara manual dengan mengambil bentuk ciri dari data citra uji.
 8. Implementasi program menggunakan *Matlab*.
 9. Kaidah *nahwu* yang digunakan terbatas pada kaidah – kaidah dasar.
 10. Pendeteksian kata benda dan kata sifat menggunakan karakteristik : kata diawali *alif lam*, kata diakhiri *tanwin*, kata diakhiri *kasrah*, kata diakhiri *ta marbutah*.

1.4 Tujuan

1. Mengoreksi kebenaran penulisan kata benda dan kata sifat pada huruf Arab bersambung dengan mengimplementasikan metode *template matching*
2. Menganalisis perbandingan algoritma *cross correlation* dan *correlation coefficients* dalam mengoreksi kebenaran penulisan kata benda dan kata sifat pada huruf Arab bersambung.
3. Menganalisa parameter pada *template matching* yang dapat mempengaruhi akurasi dan performansi.

1.5 Metodologi Penyelesaian Masalah

1. Studi Literatur

Tahap awal dalam penelitian ini adalah mempelajari tentang ilmu *nahwu* dan *sharaf* dari buku dan sumber lainnya khususnya yang membahas tentang proses pembentukan kata, yaitu kata benda dan kata sifat.

Mempelajari dasar - dasar pengolahan citra dan konsep penggunaan metode *template matching* dalam berbagai kasus melalui buku, jurnal, karya ilmiah dan internet.

2. Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data berupa kata benda dan kata sifat dari tulisan bahasa arab yang diketik dengan menggunakan *software Arabic Pad 1.4*. Tipe huruf yang digunakan adalah *Traditional Arabic*, *Simplified Arabic*, dan *Arabic Typesetting* dengan beberapa ukuran font yang berbeda . Jumlah kata benda dan kata sifat yang digunakan sebanyak 500 kata. Setiap kata disimpan dalam bentuk file gambar (.jpg). Template yang digunakan pada template matching merupakan hasil pemotongan ciri dari kata benda maupun kata sifat yang telah disimpan dalam bentuk file gambar (.jpg).

3. Perancangan Sistem

Pada tahap ini merancang sistem untuk mendeteksi kebenaran penggunaan kata benda dan kata sifat pada huruf arab bersambung yang diimplementasikan pada tiga buah algoritma template matching yang berbeda.

4. Implementasi

Tahap ini mengimplementasikan rancangan yang telah dibuat dengan menggunakan aplikasi *matlab*.

5. Pengujian

Pada tahap ini dilakukan pengujian sistem yang telah dibuat. Pengujian dilakukan menggunakan data pengujian. Kemudian menghitung nilai akurasi ketepatan sistem dalam mendeteksi kebenaran penulisan kata benda dan kata sifat.

6. Analisa hasil

Pada tahap ini menganalisa data dari hasil yang telah diuji, yaitu menganalisa parameter yang mempengaruhi akurasi dan performansi sistem. Analisa juga dilakukan pada tiga jenis algoritma *template matching* yang digunakan pada penelitian. Dari tiga jenis algoritma *template*

matching tersebut akan ditentukan algoritma manakah yang paling baik digunakan untuk menyelesaikan kasus ini.

7. Kesimpulan

Pada tahap akhir ini menyimpulkan hasil dari analisis dan dimasukkan kedalam laporan hasil penelitian.

1.6 Sistematika Penulisan

Tugas akhir ini disusun dengan sistematika penulisan sebagai berikut :

1 Pendahuluan

Bab ini menjelaskan tentang tugas akhir ini secara umum, meliputi latar belakang masalah, perumusan masalah, batasan masalah, tujuan, metodologi penyelesaian masalah, dan sistematika penulisan.

2 Landasan Teori

Bab ini menjelaskan tentang dasar teori yang berhubungan dengan pembuatan tugas akhir ini, yaitu meliputi *nahwu sharaf*, citra digital, *image processing*, ekstraksi ciri, klasifikasi dan *template matching*.

3 Analisis dan Perancangan Perangkat Lunak

Bab ini menjelaskan tentang analisis sistem, yaitu kebutuhan dan masalah yang ada di dalam sistem ini. Kemudian merancang sistem tersebut secara terstruktur.

4 Implementasi dan Analisis Hasil Pengujian

Bab ini menjelaskan tentang implementasi sistem yang telah dirancang. Kemudian dilakukan pengujian terhadap perangkat lunak dengan cara membandingkan hasil yang didapat dari perangkat lunak tersebut dengan data yang sebenarnya. Kemudian menganalisis parameter pada metode *template matching* yang mempengaruhi akurasi dan performansi sistem.

5 Kesimpulan dan Saran

Berisi kesimpulan dan saran yang diajukan penulis dari tugas akhir ini.



5. Kesimpulan dan Saran

5.1 Kesimpulan

Kesimpulan yang dapat diambil dari tugas akhir ini adalah :

1. Metode *Template Matching* dapat diimplementasikan untuk mengidentifikasi kebenaran penulisan kata benda dan kata sifat pada huruf Arab bersambung.
2. Akurasi terbaik yang didapat untuk citra dengan tipe huruf *Traditional Arabic* dan ukuran *font* 100 adalah 95% dengan menggunakan algoritma *correlation coefficients* dan nilai *threshold* 0.6.
3. Akurasi terbaik yang didapat untuk citra dengan tipe huruf *Traditional Arabic* dan ukuran *font* 200 adalah 90% dengan menggunakan algoritma *correlation coefficients* dan nilai *threshold* 0.6 dan 0.7.
4. Akurasi terbaik yang didapat untuk citra dengan tipe huruf *Traditional Arabic* dan ukuran *font* 80 adalah 94% dengan menggunakan algoritma *cross correlation* dan nilai *threshold* 0.8.
5. Akurasi terbaik yang didapat untuk citra dengan tipe huruf *Simplified Arabic* dan ukuran *font* 200 adalah 96% dengan menggunakan algoritma *correlation coefficients* dan nilai *threshold* 0.6.
6. Akurasi terbaik yang didapat untuk citra dengan tipe huruf *Arabic Typesetting* dan ukuran *font* 150 adalah 100% dengan menggunakan algoritma *correlation coefficients* dan nilai *threshold* 0.5.
7. Untuk pengujian terhadap seluruh jenis tipe huruf, akurasi terbaik yang didapat adalah 87.2% dengan menggunakan algoritma *correlation coefficients* dan nilai *threshold* 0.6.
8. Penetapan nilai dilasi mempengaruhi hasil dari pemetaan gambar (*bounding box*).
9. Hasil dari proses pemetaan gambar (*bounding box*) mempengaruhi nilai akurasi sistem.
10. Penetapan nilai *threshold* pada metode *Template Matching* mempengaruhi hasil dari akurasi sistem.

11. Algoritma *cross correlation* membutuhkan waktu komputasi yang lebih cepat dibandingkan dengan algoritma *correlation coefficients*.

5.2 Saran

Saran yang penulis berikan untuk tugas akhir ini adalah :

1. Untuk mendapatkan hasil akurasi yang maksimal, proses ekstraksi ciri memiliki pengaruh yang cukup besar. Oleh karena itu, gunakan metode ekstraksi ciri lainnya yang dapat memberikan hasil lebih baik.
2. Metode *Template Matching* memiliki turunan cukup banyak. Gunakan algoritma *template matching* lainnya untuk mendapatkan hasil akurasi yang lebih baik.
3. Tugas Akhir ini dapat diteruskan dengan kasus identifikasi kebenaran penulisan huruf *jar* yang merupakan kelanjutan dari kasus ini yaitu dengan menggabungkan dua kata atau lebih untuk kemudian di cek kebenaran penulisannya sesuai dengan aturan *nahwu* dan *sharaf*.

Daftar Pustaka

- [1] Cole, Luke, dkk. 2004. *Visual Object Recognition using Template Matching*. Australia : Australian National University
- [2] Cox, G.S. 1995. *Template Matching and Measures of Match in Image Processing*. Departement of Electrical Engineering University of Cape Town
- [3] Image Processing, (http://en.wikipedia.org/wiki/Image_processing), diakses pada tanggal 20 Agustus 2012
- [4] Jacobs, David. 2005. *Correlation and Convolution*. Class Note for CMSC 426
- [5] Kusmayadi, Hendra. 2008. *Analisis Perbandingan Antara Proses Image Correlation dan Feature Extraction pada Metode Template Matching dalam Pengenalan Notasi Balok*. Bandung : Institut Teknologi Telkom
- [6] Leksono, Bowo. 2007. *Aplikasi Metode Template Matching untuk Klasifikasi Sidik Jari*. Semarang : Universitas Diponegoro
- [7] Maliki, Irfan. Bahri, Raden Sofian. 2012. *Perbandingan Algoritma dan Feature Extraction pada Optical Character Recognition*. Bandung : Universitas Komputer Indonesia
- [8] Murni, Aniati. Chahyati, Dina. 2002. *Pengolahan Citra Digital : Morfologi Citra*. Jakarta : Fakultas Ilmu Komputer Universitas Indonesia
- [9] Nugroho, Kris Adhy. 2007. *Identifikasi Cacat pada Keping PCB Menggunakan Pencocokan Model (Template Matching)*. Semarang : Universitas Diponegoro
- [10] Nurdin, Asep Supian. 2010. *Ilmu Nahwu dan Sharaf*. Bandung
- [11] Template Matching, (http://en.wikipedia.org/wiki/Template_matching), diakses pada tanggal 15 Juli 2012)