

ANALISIS DAN IMPLEMENTASI PENGENALAN CITRA TULISAN TANGAN ALFABET MENGGUNAKAN METODE MODIFIED DIRECTION FEATURE DAN KLASIFIKASI JARINGAN SYARAF TIRUAN-LEARNING VECTOR QUANTIZATION

Insan Akbar Eryaputra¹, Yuliant Sibaroni², Fhira Nhita³

¹Ilmu Komputasi, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Alfabet adalah sebuah sistem tulisan yang berdasarkan lambang fonem vokal dan konsonan. Alfabet berasal dari bahasa Yunani yaitu, alfa dan beta. Huruf alfabet terdiri dari huruf A hingga Z. Alfabet digunakan hampir di seluruh negara sebagai sistem penulisan. Tulisan tangan alfabet seseorang dapat berbeda dengan orang lainnya. Oleh karena itu, akan terdapat banyak pola yang berbeda untuk menentukan suatu huruf alfabet. MDF adalah salah satu metode ekstraksi ciri yang dapat mengenali ciri berdasarkan bentuk pola yang terdapat dalam gambar. LVQ merupakan salah satu jenis pembelajaran pada JST yang melakukan pembelajaran pada lapisan kompetitif yang terawasi. Dalam tugas akhir ini, pola tulisan tangan manusia akan di ambil ciri nya menggunakan MDF kemudian hasil dari ekstraksi ciri dikelompokkan menggunakan LVQ. Parameter-parameter yang diuji untuk mendapatkan hasil terbaik dalam tugas akhir ini yaitu ukuran normalisasi, jumlah transisi, dan jumlah hidden neuron. Proses pencarian parameter terbaik dengan cara mencari masing-masing nilai yang terbaik kemudian digabungkan untuk mendapat akurasi yang terbaik. Dalam uji coba menggunakan gabungan huruf besar dan huruf kecil, hasil yang didapat kurang baik hanya sebesar 68,27 %. Dari hasil yang didapat, hasil tidak begitu baik dikarenakan beberapa huruf alfabet memiliki kemiripan bentuk dengan huruf lainnya seperti contoh huruf "U" dengan "u". Oleh karena itu, data dibagi menjadi dua bagian, yaitu huruf kapital dan huruf kecil. Hasil akurasi terbaik untuk pembagian huruf kapital sebesar 74,04 % dan untuk huruf kecil sebesar 75 %. Dengan proses pembagian huruf kapital dan huruf kecil menghasilkan akurasi yang lebih baik daripada penggabungan huruf kapital dan huruf kecil meskipun masih terdapat beberapa huruf yang tidak dapat terbaca dengan baik.

Kata Kunci : Modified Direction Feature, Learning Vector Quantization, Citra Digital, Pengenalan Pola, Tulisan Tangan

Abstract

Alphabet is the written system based on vocal and consonant phonem. Alphabet came from Yunani, that are alfa and beta. Alphabet consist from the letter A until Z. Almost every country around the world use alphabet for their written system. The written alphabet can be difference between one and another people. So, there will be so many ways in considering the alphabet itself. MDF is on of the extraction method to know the character based on the pattern of picture. LVQ is one of the learning type in JST in the supervised competitive learning. In this research, the written alphabet will be categorize using MDF then devided into groups using LVQ. Normalization, number of transition, and number of hidden neuron are parameters that are used in this research. Best parameter can be got through getting the best number and combine it to get the best number of accuration. The number of accuration in this research is 68.27%. That result did not give the best accuration because there are similarity between some letters, such as "U" and "u". In that case, letters divided between capital and non capital letters. The number of accuration for the capital letter is 74.04% and non-capital letter is 75%. Dividing between capital and non-capital letter is a good way in getting better result though there were some letters that can not be read.

Keywords : Modified Direction Feature, Learning Vector Quantization, Digital Image, Pattern Recognition, Written Alphabet

BAB I

PENDAHULUAN

1.1 Latar Belakang

Alfabet adalah sebuah sistem tulisan yang berdasarkan lambang fonem vokal dan konsonan. Kata alfabet diambil dari bahasa Yunani yaitu, alfa dan beta. Huruf alfabet terdiri dari huruf A hingga Z. Alfabet digunakan hampir di seluruh negara sebagai sistem penulisan.

Tulisan tangan alfabet seseorang dapat berbeda dengan orang lainnya. Oleh karena itu, akan terdapat banyak pola yang berbeda untuk menentukan suatu huruf alfabet. Apabila di sebuah perusahaan atau instansi perlu memasukan data dari tulisan tangan, maka akan membutuhkan waktu yang cukup lama untuk mengenali tulisan dan memasukan data ke dalam sistem dengan cara pengetikan *manual*.

Sebenarnya pengidentifikasian paling sederhana dapat dilakukan dengan penglihatan biasa. Namun apabila dilakukan dengan komputer, hal tersebut tidak dapat dilakukan secara otomatis. Dengan menggunakan pemodelan visual untuk membedakan antar huruf tersebut sehingga untuk mengidentifikasi huruf tersebut menggunakan pengenalan pola dari huruf tersebut. Dalam hal ini, komputer ingin meniru cara kerja otak dalam pengenalan huruf, khususnya huruf alphabet. Namun proses dalam mengenali huruf tersebut relatif cukup lama.

Pengenalan pola huruf alfabet dilakukan dalam tugas akhir ini memiliki dua tahap, yaitu ekstraksi ciri dan klasifikasi. Terdapat beberapa metode dalam menentukan ekstraksi ciri seperti *Direction Feature (DF)*, *Gabor filter*, *MDF* dan lain-lain. Kemudian terdapat beberapa metode dalam mengklasifikasi, seperti *Support Vector Machine (SVM)*, *Jaringan Syaraf Tiruan (JST)*, dan lain-lain. Dalam tugas akhir ini metode yang digunakan dalam mengekstraksi diri menggunakan metode *Modified Direction Feature (MDF)* [1]. Metode ini merupakan hasil dari memodifikasi keunggulan dari teknik *traditional direction feature extraction* dan *transition feature extraction*, yaitu mengkombinasikan ciri arah dan informasi struktur global yang ada pada karakter [1]. Metode ini

melakukan pendekatan dengan cara melakukan deteksi arah dari segmen-segmen garis pembentuk karakter dan pixel-pixel pembentuk karakter tersebut, namun kemudian digantikan dengan nilai arah yang sesuai. Akhirnya, ciri didapatkan dari karakter yang akan menyimpan lokasi dari pixel dan nilai arahnya. MDF memiliki vektor ciri pola yang cukup tinggi dan dapat memberikan keunikan pada pola. Ciri yang sudah didapat, akan dijadikan inputan untuk proses klasifikasi. Oleh karena itu dalam tugas akhir ini menggunakan metode MDF sebagai ekstraksi ciri huruf alfabet yang memiliki tingkatan kerumitan cukup tinggi.

Learning Vector Quantization (LVQ) adalah metode klasifikasi pola yang setiap *output* akan mewakili kategori atau kelas tertentu. LVQ melakukan pembelajaran dalam suatu lapisan kompetitif yang secara otomatis akan belajar untuk mengklasifikasikan vektor-vektor input. Jika dua vektor input mendekati sama, maka lapisan kompetitif akan meletakkan kedua vektor input tersebut ke dalam kelas yang sama. Algoritma LVQ bertujuan akhir mencari nilai bobot yang sesuai untuk mengelompokkan vektor-vektor input ke dalam kelas tujuan yang telah di inialisasi pada saat pembentukan Jaringan LVQ. Metode LVQ membutuhkan sedikit contoh latihan, lebih cepat dari metode klasifikasi lainnya, dan memproses data yang memiliki *missing value* [14]. Hal tersebut yang menjadikan pertimbangan dalam pemilihan LVQ sebagai pengklasifikasian huruf alfabet yang cukup baik

MDF dengan LVQ dapat digabungkan dan dapat menghasilkan akurasi yang cukup baik apabila sudah dapat menentukan konfigurasi parameter MDF yang terbaik [5] [10].

1.2 Perumusan Masalah

Perumusan masalah dari Tugas Akhir ini berdasarkan latar belakang masalah yang telah dijelaskan di atas adalah :

1. Bagaimana mengimplementasikan *Modified Direction Feature* (MDF) dan *JST Learning Vector Quantization* (LVQ) dalam mengenali huruf alphabet.

2. Bagaimana menentukan metode untuk mencari parameter fitur ekstraksi ciri dan LVQ yang terbaik.
3. Bagaimana analisis metode MDF dan LVQ pada pengenalan pola huruf alfabet berdasarkan hasil akurasi yang didapat dari parameter-parameter dari metode tersebut.

Adapun batasan masalah pada tugas akhir ini adalah sebagai berikut :

1. Input dan output merupakan file image dengan tipe *.jpg dengan *background* putih
2. Pola inputan dengan posisi tertentu, tidak menangani posisi bertumpuk
3. Citra masukan berupa citra karakter tulisan alfabet 'A' – 'Z' dan 'a'-'z'.
4. Huruf tulisan tangan bersifat *offline*.
5. Huruf *sample* yang digunakan didapat dari tulisan tangan mahasiswa dan pelajar.

1.3 Tujuan Penelitian

Tujuan dari tugas akhir ini adalah :

1. Mengimplementasikan kombinasi MDF dan LVQ dalam pengenalan pola huruf alfabet.
2. Mengembangkan metode untuk mencari parameter MDF dan LVQ yang terbaik dan dapat mempengaruhi performansi akurasi sistem dalam pengenalan huruf alfabet yang ada di dalam citra.
3. Menentukan kombinasi terbaik MDF dan LVQ dengan melihat performansi akurasi sistem berdasarkan nilai CER (*Character Error Rate*) mengenai ketepatan pengenalan pola.

1.4 Metodologi Penelitian

Metodologi yang digunakan dalam memecahkan masalah di atas adalah dengan menggunakan langkah-langkah berikut :

1. Studi Literatur

Pencarian referensi dan sumber-sumber yang berhubungan dengan *Modified Direction Feature* (MDF) dan *Learning Vector Quantization* (LVQ) dalam menyelesaikan tugas akhir ini.

2. Pengumpulan Data

Mengumpulkan data-data pola tulisan alfabet yang akan digunakan sebagai input dan proses *training*. Pengumpulan data dilakukan dengan memberikan form yang diisi oleh beberapa orang kemudian form tersebut dipindai menjadi file image (jpg) menggunakan *scanner*.

3. Analisis dan perancangan sistem

Pada tahap ini meliputi analisis kebutuhan untuk merancang sistem pengenalan huruf alfabet dengan *Modified Direction Feature* (MDF), metode dalam pencarian parameter terbaik, dan *Learning Vector Quantization* (LVQ).

4. Implementasi dan Pembangunan Sistem

Melakukan implementasi pada sistem sesuai dengan analisis kebutuhan dan perancangan yang telah dilakukan pada tahap sebelumnya.

5. Analisis Hasil Pengujian Sistem

Melakukan analisa terhadap hasil dari pengujian dan evaluasi sistem yang telah dibuat dengan membandingkan berdasarkan tingkat akurasi dan parameter-parameter yang terdapat pada *Modified Direction Feature* (MDF) dan *Learning Vector Quantization* (LVQ) yang disesuaikan dengan skenario pengujian.

6. Kesimpulan

Kesimpulan berisi tentang hasil akurasi terbaik dan parameter terbaik dari metode MDF dan LVQ yang didapat dari seluruh percobaan yang dilakukan pada bab sebelumnya. Dan menyimpulkan pengaruhnya parameter-parameter tersebut terhadap hasil akurasi.

7. Pembuatan Laporan

Membuat laporan yang berisi tentang pendahuluan, dasar teori, serta metode dan perancangan sistem dari metode *Modified Direction Feature* (MDF) dan *Learning Vector Quantization* (LVQ) dalam kasus ini.

1.5 Sistematika Penulisan

Sistematika dalam penulisan tugas akhir ini adalah sebagai berikut :

1. BAB I Pendahuluan

Bab ini berisi uraian tentang latar belakang masalah yang mendasari pentingnya diadakan penelitian, identifikasi, pembatasan dan perumusan Masalah Penelitian, Maksud dan Tujuan Penelitian, Kegunaan Penelitian yang diharapkan, Metodologi Penelitian, Sistematika penulisan, serta Jadwal kegiatan.

2. BAB II Dasar Teori

Bab ini berisi dasar teori yang mendeskripsikan pengertian dan prinsip dasar metode-metode yang digunakan.

3. BAB III Analisis dan Perancangan Sistem

Bab ini berisi uraian tentang deskripsi dan analisis sistem, perancangan sistem, serta implementasi dari metode-metode yang digunakan

4. BAB IV Pengujian dan Implementasi

Bab ini berisi tentang proses pengujian dan implementasi dari sistem yang akan dibangun. Terdapat skenario yang akan dilakukan untuk menghasilkan hasil yang terbaik.

5. BAB V Penutup

Bab ini berisi tentang kesimpulan dan saran. Kesimpulan didapat berdasarkan hasil yang telah diperoleh pada bab IV. Kemudian saran merupakan opini atau solusi dari penulis agar dapat memperbaiki hasil dari pengujian yang telah dilakukan.

6. Daftar Pustaka

Bab ini berisi tentang daftar sumber bacaan yang digunakan sebagai bahan acuan dalam penulisan dan mengerjakan penelitian ini.

1.6 Jadwal Kegiatan

Rencana kegiatan yang akan dilakukan dalam penyelesaian tugas akhir ini dijadwalkan dengan representasi tabel 1.1 berikut :

Tabel 1.1 Jadwal kegiatan tugas akhir

No.	Kegiatan	Bulan ke-1	Bulan ke-2	Bulan ke-3	Bulan ke-4	Bulan ke-5	Bulan ke-6
1	Studi literature	■	■	■	■		
2	Pengumpulan data	■					
3	Perancangan sistem	■	■	■			
4	Implementasi dan pembangunan sistem		■	■	■	■	
5	Analisis hasil dan pengujian sistem				■	■	
6	Penarikan Kesimpulan				■	■	
7	Pembuatan laporan		■	■	■	■	■

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil dari pengujian parameter metode MDF dan LVQ pada bab sebelumnya, dapat disimpulkan bahwa :

1. Parameter terbaik untuk metode MDF yaitu ukuran normalisasi sebesar 180 x 100 dengan jumlah transisi 3. Parameter MDF tersebut mempengaruhi terhadap jumlah vektor ciri yang dihasilkan.
2. Parameter terbaik untuk metode LVQ yaitu 230 untuk jumlah *hidden neuron*. Parameter LVQ tersebut mempengaruhi hasil terhadap klasifikasi.
3. Akurasi untuk gabungan huruf kapital dan huruf kecil kurang baik, hasil terbaik yang didapat hanya sebesar 68,27 % dengan nilai rata-rata sebesar 60,1 %. Hal ini disebabkan terdapat kemiripan bentuk antara huruf besar dan huruf kecil.
4. Akurasi dengan melakukan pembagian data menjadi huruf kapital dan huruf kecil memiliki peningkatan hasil akurasi dibandingkan dengan gabungan huruf. Hasil akurasi terbaik untuk huruf kapital sebesar 74,04 % dan untuk huruf kecil sebesar 75 %

5.2 Saran

Menurut hasil percobaan pada bab sebelumnya, masih terdapat beberapa huruf yang tidak dapat terbaca dengan baik. Maka diperlukan metode ekstraksi ciri yang digunakan selain MDF agar hasil yang diperoleh lebih bervariasi dengan tingkat akurasi yang lebih tinggi.

DAFTAR PUSTAKA

- [1] Blumenstein, M., Verma, B. 2007. *A Modified Direction Feature for Cursive Character Recognition*. International Joint Conference on Neural Networks.
- [2] Burrascan, Pietro. 1991. *Learning Vector Quantization for the Probabilistic Neural Network*. IEEE Transactions on Neural Networks.
- [3] Gader, Paul D., Magdi Mohamed. 1996. *Handwritten Word Recognition Using Segmentation-Free Hidden Markov Modelling and Segmentation-Based Dynamic Programming Techniques*. IEEE Transactions on Pattern Analysis and Machine Intelligence.
- [4] Gader, Paul D., Magdi Mohamed, Jung-Hsien Chiang. 1997. *Handwritten Word Recognition with Character and Inter-Character Neural Networks*. IEEE Transactions on Systems, MAN, and Cybernetics.
- [5] Hermanto, I Gede Rudy., Tjokorda Agung Budi W, Retno Novi D. 2009. *Pengenalan Huruf Bali Menggunakan Metode Modified Direction Feature (MDF) dan Learning Vector Quantization (LVQ)*. Konferensi Nasional Sistem dan Informatika.
- [6] Kurniawan Fachrul, Nurhayati Hani. *Simulasi pengenalan tulisan tangan menggunakan LVQ (Learning Vektor quantization)*. Jurusan Teknik Informatika, Sains dan Teknologi Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang
- [7] O. Matan, R. K. Kiang, C. E. Stenard, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel, Y. Le Cun. 1990. *Handwritten Character Recognition Using Neural Network Architectures*. 4th USPS Advanced Technology Conference, Washington D.C.
- [8] Putra, Darma. 2010. *Pengolahan Citra Digital*. Yogyakarta: Andi
- [9] Ranadhi, D., Indarto, W., Hidayat, T. 2006. *Implementasi Learning Vector Quantization (Lvq) Untuk Pengenal Pola Sidik Jari Pada Sistem Informasi Narapidana Lp Wirogunan*. Univeristas Islam Indonesia.

- [10] Rice, Stephen., Nagy, G., Nartker, T., 1999. *OPTICAL CHARACTER RECOGNITION : An Illustrated Guide to the Frontier*. Springer: Kluwer Academic.
- [11] Ummami, Arul., Adiwijaya., Retno Novi D. 2010. “Analisis dan Implementasi Pengenalan Huruf Arab Menggunakan Modified Direction Feature Extraction dan Learning Vector Quantization”. Program Sarjana. Institut Teknologi Telkom. Bandung
- [12] Sato, A., Yamada, K., 2002. *Generalized Learning Vector Quantization*. MIT Press, Cambridge.
- [13] V. J. Rayward-Smith., S. H. Al-Harbi., 2006. *Adapting k-means for supervised clustering*. Springer Science + Business Media, LLC
- [14] *WEKA Classification Algorithm*. Terdapat pada halaman web: <http://weka.classalgos.sourceforge.net/> (13 Januari 2013 pukul: 14.00).