

BAB I

PENDAHULUAN

1.1. Latar Belakang

Jepang merupakan negara yang memiliki keanekaragaman budaya yang sangat unik dan menarik. Menurut Nugraha, kebudayaan Jepang berhasil disebarkan ke berbagai negara lain terutama negara Indonesia dan mendapatkan respon yang positif terutama dikalangan remaja, sehingga membuat minat kalangan remaja untuk mengenal budayanya. Apabila ingin mengenal budaya negara Jepang, sebaiknya terlebih dahulu memahami bahasa Jepang (Nugraha dan Hidayat 2016). Bahasa terdiri dari bentuk lisan dan tulisan (Susetyo, Warnandi, dan Ridwan 2009). Menurut Raneriah, memahami bahasa dalam bentuk tulisan, terlebih dahulu harus memahami hurufnya, begitu juga dengan bahasa Jepang. Huruf dalam bahasa Jepang terkelompok dalam tiga jenis huruf yakni huruf *kanji*, huruf *hiragana* dan huruf *katakana* (Renariah 2002).

Salah satu pakar dibidang huruf Jepang yang bernama Yoshiaki Takabe dalam buku *Kanji no Yoohoo* yang dikutip Renariah mengemukakan bahwa sistem penulisan di dunia dapat dikelompokkan menjadi dua, yakni *hyo i moji* dan *hyo on moji*. Penulisan Jepang dalam *hyo i moji* merupakan huruf *kanji* karena huruf *kanji* merupakan huruf yang menyatakan arti dan *hyo on moji* merupakan huruf *hiragana* dan *katakana* karena hanya melambangkan bunyi saja tanpa menyatakan arti (Renariah 2002). Huruf *hiragana* dan *katakana* merupakan dua jenis huruf dasar dalam bahasa Jepang. Huruf *hiragana* dan huruf *katakana* sering disebut sebagai huruf *kana*. Huruf *hiragana* biasanya digunakan untuk menulis kata-kata Jepang yang murni atau asli. Huruf *katakana* biasanya digunakan untuk kata-kata serapan dari bahasa asing (Lensun 2004).

Salah satu yang menjadi kesulitan dalam mempelajari bahasa Jepang terdapat pada sistem penulisannya. Huruf Jepang yang kompleks dan sangat berbeda dengan huruf latin, sehingga cukup sulit untuk mempelajari dan menghafal hurufnya. Huruf Jepang juga berbeda pada pelafalannya, karena pelafalan huruf Jepang biasanya

terdiri dari dua suku kata (Fitria, Usman, dan Mursita 2010). Huruf Jepang dibentuk dengan guratan garis-garis berarah yang berurutan dan sudut-sudut dengan keunikan tersendiri dari setiap huruf (Kurnia, Romadhony, dan Huda 2012). Karakter huruf Jepang merupakan karakter yang cukup kompleks dan memiliki karakteristik yang unik, terlebih bila ditulis dengan tulisan tangan karena variasi bentuk tulisan yang dihasilkan oleh setiap orang akan memiliki pola yang berbeda-beda (Masril 2013).

Penelitian terkait yang mengangkat kasus huruf Jepang khususnya huruf *hiragana* Jepang sudah pernah dilakukan sebelumnya. Salah satunya penelitian yang dilakukan oleh Trianto dengan menerapkan metode *Template Matching Correlation* dengan mendapatkan tingkat akurasi sebesar 75% (Trianto et al. 2014). Penelitian lainnya juga dilakukan oleh Riyadi, Wirayuda dan Wibowo dengan menggunakan metode *Independent Component Analysis* (ICA) dan *Learning Vector Quantization* (LVQ) dengan mendapatkan tingkat akurasi 43,9% (Riyadi, Wirayuda, dan Wibowo 2011).

Melihat tingkat akurasi pada penelitian terkait sebelumnya dalam pengenalan huruf *hiragana* Jepang yang cukup rendah, maka pada penelitian ini akan diterapkan metode yang berbeda dari metode sebelumnya dan diharapkan metode tersebut mampu lebih baik dalam mengenali pola huruf *hiragana* Jepang. Sehingga mempermudah dalam mengenali dan mempelajari huruf *hiragana* Jepang tersebut. Metode yang diangkat dalam penelitian ini yaitu dengan melakukan penerapan dari pengolahan citra digital dan jaringan syaraf tiruan (JST).

Pengolahan citra merupakan suatu pemrosesan citra dengan menggunakan komputer untuk mejadikan suatu citra yang memiliki kualitas yang lebih baik. Hasil dari pengolahan citra yaitu suatu keputusan dimana salah satunya yaitu pengenalan pola (*pattern recognition*) (Herlambang, Nurhayati, dan Martono 2016). Salah satu metode pada pengolahan citra yang digunakan yaitu *Principal Component Analysis* (PCA). PCA adalah suatu teknik linier reduksi yang dimana menggunakan teori sederhana dari suatu statistik. Teori tersebut antara lain ialah varian, standar deviasi, *zero mean*, kovarian, dan Persamaan karakteristik (Anam, Widodo, dan Sul-ton 2016). PCA ini sendiri dapat mempermudah proses perhitungan, mengkompresi

data, dan menyederhanakan suatu data yang diamati yaitu dengan cara mereduksi atau menyusutkan dimensinya tanpa banyak kehilangan informasi (Mulyadi dan Suryanto 2016).

Penelitian yang menerapkan metode PCA sudah banyak dilakukan. Salah satunya pernah dilakukan oleh Diyah Puspitaningrum dan kawan-kawan tentang dampak reduksi sampel menggunakan PCA mendapatkan tingkat akurasi 86,75% (Puspitaningrum, Kemala Sari, dan Susilo 2014). Penelitian selanjutnya dilakukan oleh Rizal Isnanto, Ajub Zahra, dan Eko Widiyanto tentang pengenalan telapak tangan menggunakan PCA mendapatkan tingkat akurasi 90% (Isnanto, Zahra, dan Widiyanto 2015). Selanjutnya penelitian yang dilakukan oleh Yasser Fouad Hassan dan Nora Habeb tentang klasifikasi pada wajah menggunakan PCA mendapatkan tingkat akurasi 93% (Hassan dan Habeb 2012).

JST merupakan suatu sistem pemrosesan informasi yang mempunyai karakteristik menyerupai jaringan syaraf manusia. JST tercipta sebagai suatu generalisasi model matematis dari pemahaman manusia sendiri (*human cognition*) (Dessy dan Irawan, 2012). Salah satu metode pada JST yang digunakan yaitu *Learning Vector Quantization* (LVQ). LVQ merupakan JST dalam mengklasifikasikan pola yang dimana setiap keluarannya itu memiliki sebuah kelas. Vektor Perwakilan untuk sebuah keluaran didapatkan dari *training*. Setelah dilakukan *training*, LVQ mengklasifikasikan sebuah vektor masukan kedalam sebuah kelas yang mempunyai Vektor Perwakilan vektor terdekat dengan Vektor Perwakilan vektor yang telah di *training*. (Solichin dan Rahman, 2015). LVQ sendiri sudah memiliki beberapa variasi metode diantaranya LVQ atau LVQ1, LVQ2, LVQ2.1 dan LVQ3 (Jatmiko 2013). Pada LVQ2 proses pembelajaran memperhitungkan vektor *runner-up* dimana proses pembelajaran ditentukan berdasarkan jika vektor masukan memiliki jarak yang hampir sama antara pemenang dan *runner-up*, maka kedua vektor perwakilan seharusnya diperbarui secara simultan jika masukan berada pada bagian kelas yang salah. Pada LVQ2.1 algoritma ini merupakan penyempurnaan dari LVQ2 dimana algoritma ini mengabaikan aturan dari kondisi *update* vektor LVQ2. Pada algoritma LVQ2.1, kategori dari sampel data tidak harus sama dengan vektor *runner-up*. Persyaratannya adalah minimal salah satu dari vektor berasal dari

kategori yang sama dengan kategori masukan. Sedangkan aturan *update* yang lain masih tetap sama dengan sebelumnya (Jatmiko 2013). Pada LVQ3 terjadi koreksi terhadap LVQ2.1. Dimana pada algoritma LVQ2.1 memiliki kelemahan yaitu dimana pada vektor kemungkinan mengalami divergensi selama proses pembelajaran dilakukan. Adapun koreksi yang dilakukan terhadap LVQ2.1 memastikan vektor agar selalu mendekati distribusi dari kelas. Aturan *update* vektor perwakilan sama dengan LVQ2.1, hanya saja terdapat aturan tambahan dimana jika kedua vektor w_1 dan w_2 berasal dari kelas yang sama, maka *update* vektor perwakilnya (Jatmiko 2013). Adapun pada penelitian ini variasi LVQ yang digunakan adalah LVQ3. Ide pengembangan algoritma LVQ yaitu jika masukan memiliki hitungan jarak yang sama dengan vektor pemenang dan *runner-up*, maka masing-masing vektor tersebut harus melakukan pembelajaran (Budianita dan Prijodiprodjo 2012). Metode ini termasuk ke dalam metode pembelajaran terarah (*supervised*) (Ekayama, C. Djamal, dan Komarudin 2015).

Penelitian yang menerapkan metode LVQ3 pernah dilakukan oleh Elvia Budianita dan Widodo Prijodiprodjo tentang klasifikasi status gizi anak menggunakan LVQ1 dan LVQ3. Pada penelitian tersebut membuktikan bahwa LVQ3 memiliki kinerja lebih baik dari LVQ1 dimana LVQ1 mendapatkan tingkat akurasi 88% dan sedangkan LVQ3 mendapatkan tingkat akurasi 95,2% (Budianita dan Prijodiprodjo 2012). Penelitian lainnya yang menerapkan LVQ3 lainnya dilakukan oleh Mohammed Azara, Tamer Fatayer, dan Alaa El-Halees tentang klasifikasi teks arab dengan mendapatkan tingkat akurasi 84,09% (Azara, Fatayer, dan El-Halees 2012).

Berdasarkan permasalahan dan melihat dari penelitian sebelumnya yang memanfaatkan metode PCA dan LVQ3, maka dalam penelitian ini akan diterapkan metode PCA dan LVQ3 dalam pengenalan pola huruf *hiragana* Jepang. Kedua metode tersebut diharapkan mampu mendapatkan tingkat akurasi dan ketepatan dalam mengenali pola huruf *hiragana* Jepang yang lebih baik dari penelitian terkait sebelumnya.

1.2. Rumusan Masalah

Berdasarkan latar belakang masalah yang telah dijelaskan diatas, maka diperoleh rumusan masalah yaitu bagaimana menerapkan *Principal Component Analysis* (PCA) dan *Learning Vector Quantization 3* (LVQ3) dalam pengenalan huruf *hiragana* Jepang dan menghitung akurasi dalam ketepatan pengenalan pola huruf *hiragana* Jepang.

1.3. Batasan Masalah

Adapun batasan masalah dari penelitian yang dibahas antara lain:

1. Huruf yang digunakan pada penelitian ini yaitu huruf *hiragana seion*.
Jumlah data yang digunakan yaitu 460 huruf. Dimana 1 huruf terdiri dari 10 data.
2. *Output* kelas yang digunakan sebanyak 46 kelas.
3. Jumlah *pixel* dari citra yang digunakan berdasarkan penelitian sebelumnya adalah 40 *pixel* x 40 *pixel*.
4. Data yang digunakan berformat .PNG.

1.4. Tujuan Penelitian

Adapun penelitian ini bertujuan untuk menerapkan metode PCA dan LVQ 3 untuk pengenalan pola huruf *hiragana* Jepang dan menghitung akurasi dalam ketepatan pengenalan pola huruf *hiragana* Jepang.

1.5. Sistematika Penulisan

Secara garis besar laporan tugas akhir ini dibagi atas enam bab yang terdiri dari beberapa subbab antara lain:

BAB I PENDAHULUAN

Pada bab ini berisi latar belakang, batasan masalah, rumusan masalah, tujuan, dan sistematika penulisan.

BAB II LANDASAN TEORI

Pada bab ini peneliti menjelaskan teori singkat tentang hal-hal yang berhubungan dengan judul, model pengembangan aplikasi serta

tentang teori-teori yang mendukung pembuatan aplikasi. Teori yang diangkat yaitu mengenai metode *Principal Component Analysis* (PCA) dan *Learning Vector Quantization 3* (LVQ3).

BAB III METODOLOGI PENELITIAN

Pada bab ini berisi tentang beberapa rangkaian tahapan dalam pembuatan aplikasi, mulai dari melakukan pengumpulan data, analisa dan perancangan, tahap implementasi dan pengujian yang digunakan, hingga kesimpulan dan saran.

BAB IV ANALISA DAN PERANCANGAN SISTEM

Pada bab ini berisi tentang analisa dari aplikasi yang akan dibangun dan metode *Principal Component Analysis* (PCA) dan *Learning Vector Quantization 3* (LVQ3) yang digunakan dalam tugas akhir ini.

BAB V IMPLEMENTASI DAN PENGUJIAN SISTEM

Pada bab ini berisi tentang hasil implementasi dari perancangan yang telah dibuat sebelumnya, yaitu meliputi implementasi basis data, implementasi metode yang digunakan dan implementasi form-form antarmuka aplikasi.

BAB VI PENUTUP

Pada bab ini berisikan tentang kesimpulan dan saran dari hasil penelitian yang telah diperoleh.