

Pengenalan Karakter Tulisan Menggunakan Metode Learning Vector Quantization

Dinda Izmya Nurpadillah¹, Haviluddin², Herman Santoso Pakpahan³, Islamiyah⁴, Hario Jati Setyadi⁵

Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Mulawarman
izmydinda@gmail.com; haviluddin@unmul.ac.id; herman.pakpahan@fkti.unmul.ac.id; islamiyah@fkti.unmul.ac.id;
hario.setyadi@fkti.unmul.ac.id

INFORMASI ARTIKEL

Histori Artikel

Diterima : 2 Juni 2019
Direvisi : 17 Juli 2019
Diterbitkan : 30 Agustus 2019

Kata Kunci:

Aksara Sunda
Pengenalan Pola
LVQ
MSE

ABSTRAK

Artikel ini mengimplementasikan metode *Learning Vector Quantization* (LVQ) dalam mengenali pola aksara Sunda. Berdasarkan hasil eksperimen dengan berbagai parameter seperti *learning rate* dan jumlah *hidden layer* maka metode LVQ cukup akurat dalam mengenali pola aksara Sunda dengan nilai akurasi sebesar 6.66% dari data yang berhasil dikenali sebanyak 28 data dengan total data uji sebanyak 42 data dengan variasi *learning rate* sebesar 0.01 dan jumlah *hidden layer* sebanyak 90 *layer*. Hasil akurasi tersebut didapatkan dengan waktu pembelajaran yaitu selama 17 menit 22 detik. Adapun mean square error (MSE) yang dihasilkan sebesar 0.0408. Dari hasil akurasi, MSE dan waktu pembelajaran yang didapatkan maka dapat dikatakan metode LVQ belum optimal dalam memecahkan masalah pengenalan pola terutama aksara Sunda. Teknik optimalisasi kepada proses pembelajaran LVQ dengan algoritma-algoritma optimasi merupakan rencana penelitian selanjutnya.

2019 SAKTI – Sains, Aplikasi, Komputasi dan Teknologi Informasi.

Hak Cipta.

I. Pendahuluan

Negara Kesatuan Republik Indonesia (NKRI) merupakan Negara yang memiliki banyak sekali kebudayaan, mulai dari seni tari, seni musik, seni pahat tak terkecuali dalam hal seni tulisan atau aksara. Dalam upaya memperkenalkan kebudayaan nusantara agar masyarakat mengenal seni aksara daerah, maka pendekatan melalui teknologi dirasa cukup tepat, mengingat kemajuan teknologi sudah sangat pesat saat ini. Salah satu teknologi yang dapat digunakan adalah pengenalan pola yang merupakan salah satu cabang ilmu kecerdasan buatan. Secara umum pengenalan pola (pattern recognition) adalah suatu ilmu untuk mengklasifikasikan atau menggambarkan sesuatu berdasarkan pengukuran kuantitatif fitur (ciri) atau sifat utama dari suatu obyek (Putra, 2010).

Salah satu metode pengenalan pola seperti LVQ. Peneliti (Sambani, Uryani, & Putra, 2016) menggunakan metode LVQ untuk mengenal barcode buku di perpustakaan Universitas Galuh Ciamis. Pada penelitian ini didapatkan hasil akurasi sebesar 100% dari jumlah data yang berhasil dikenali sebanyak 30 data (total seluruh data uji) dan dengan MSE 0, hasil ini didapatkan dari learning rate sebesar 0.05. Peneliti (Afrianto & Priatama, 2013) menggunakan metode LVQ untuk membangun aplikasi mobile. Data yang digunakan pada penelitian ini berupa gambar beberapa tempat bersejarah. Dari hasil pengujian data maka diperoleh rata-rata akurasi akurasi pengenalan sebesar 78% dengan parameter pembelajaran terbaik yaitu epoch = 50, learning rate = 0.05, dan pengurangan pembelajaran = 0.1. Sistem ini juga telah mampu mengirimkan data berupa citra dari smartphone ke server dengan baik. Sementara server mampu mengolah data gambar, melakukan pengenalan objek dan memberikan jawaban. Peneliti (Ramadhan, Djamal, & Darmanto, 2016) menggunakan metode LVQ yang dikombinasikan dengan deteksi tepi metode *Canny* untuk klasifikasi identitas wajah. Hasil analisis parameter untuk pelatihan metode LVQ, dengan menggunakan MaxEpoch 1000, learning rate 0,01, pengurangan α 0,1 dan err 0.001 dari 25 data uji yang tidak dilatih memperoleh akurasi sebesar 88% dengan hasil pengenalan 3 data uji tidak dikenali dan 22 data uji yang berhasil dikenali. Sedangkan hasil klasifikasi dari 25 data uji yang telah dilatih memperoleh akurasi sebesar 92% dengan hasil pengenalan 2 data uji tidak berhasil dikenali dan 23 data uji berhasil dikenali. Peneliti (Lestari, Hidayat, & Andini, 2015) menggunakan metode LVQ untuk merancang pengenalan kata dalam aksara Sunda berbasis android. Dari hasil pengujian terhadap 10 kata dalam aksara Sunda pada citra tanpa *cropping background* diperoleh akurasi terbesar sistem bernilai 60.90% dengan jumlah data

training 9 buah. Sedangkan pada citra yang dilakukan *cropping background* diperoleh tingkat akurasi tertinggi sebesar 61.53% dengan nilai learning rate 0.01 dan nilai epoch 100. Dan untuk mendeteksi 30 kata dalam aksara Sunda didapatkan nilai akurasi tertinggi sebesar 17.78% dengan nilai learning rate dan epoch yang sama. Peneliti (Lukman, 2015) menggunakan metode LVQ untuk pengenalan pola buku. Uji coba aplikasi menggunakan 12 buku yang terdiri dari 2 judul buku, 2 untuk w1 dan w2, 4 untuk data pelatihan dan 6 untuk data uji. Inisialisasi LVQ menggunakan MaxEpoch = 10, eps = 0.001 dan learning rate = 0.05. Hasilnya, aplikasi dapat mengenali semua buku yang diujikan. Dan peneliti (Utomo & Cahyono, 2014) menggunakan metode LVQ untuk mendeteksi pra panen padi berdasarkan warna daun. Sistem ini memiliki tingkat akurasi sebesar 55% untuk pengujian data training dan 50% untuk pengujian data testing. Hasil ini didapatkan dari learning rate sebesar 0.0005. Dimana jumlah pola data yang dikenali dengan benar pada data training yang dikenali sebagai data siap panen sebanyak 30 data, dan data yang dikenali sebagai data belum siap panen sebanyak 3 data. Untuk jumlah pola data yang berhasil dikenali pada data testing yang dikenali sebagai data siap panen sebanyak 10 data, dan tidak ada data belum siap panen dikenali sebagai data belum siap panen.

Oleh karena itu, artikel ini menerapkan metode *Learning Vector Quantization* (LVQ) untuk mengenali pola aksara Sunda khususnya aksara Swara (vokal). Hasil dari penelitian ini diharapkan dapat mengenalkan kembali aksara Sunda dengan menggunakan pendekatan teknologi. Paper ini terdiri dari empat bagian. Bagian pertama adalah motivasi penelitian ini. Bagian kedua adalah menjelaskan metodologi dan teknik yang digunakan. Bagian ketiga adalah menerangkan hasil pengujian dengan metode LVQ. Dan, bagian terakhir adalah kesimpulan dan saran yang diperoleh setelah pengujian, serta rencana penelitian selanjutnya.

II. Metodologi

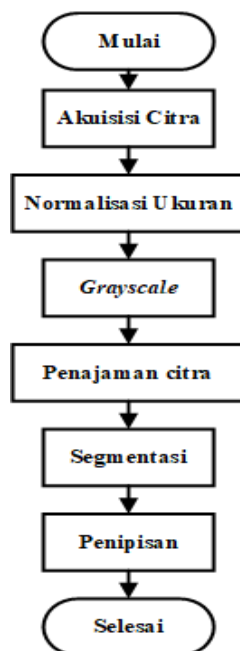
Pada bagian ini, akan dijelaskan secara singkat tentang aksara Sunda, metode pengolahan citra yang digunakan, dan metode LVQ.

A. Aksara Sunda

Aksara Ngalagena sudah digunakan dari abad ke-14 masehi oleh orang Sunda. Prasasti Kawali (Prasasti Astaga Gede) adalah salah satu peninggalan aksara Sunda. Prasasti tersebut dibuat untuk mengenang Prabu Niskala Wastukencana yang pada saat itu memerintah Kawali, Ciamis pada tahun 1271-1475 (Saleh, 2006s). Aksara Sunda dibagi menjadi 2 bagian yaitu aksara Sunda Swara (vokal mandiri) dan aksara Sunda Ngalagena (huruf konsonan) (Riansyah, Nurhasanah, & Dewi, 2017; Sundari, 2016).

B. Pengolahan Citra

Pengolahan gambar atau pengolahan citra yang sering disebut *image processing*, merupakan suatu proses yang mengubah sebuah gambar menjadi gambar lain yang memiliki kualitas lebih baik untuk tujuan tertentu. Pengolahan citra mempunyai 2 tujuan utama, yaitu (Dani, Sugiharto, & Winara, 2015; Pratiarso et al., 2011) : (1) Memperbaiki kualitas citra, dimana citra yang dihasilkan dapat menampilkan informasi secara jelas; (2) mengekstraksi informasi ciri yang menonjol pada suatu citra, hasilnya adalah informasi citra dimana manusia mendapatkan informasi ciri dan citra secara numerik. Berikut adalah tahapan proses pengolahan citra dapat dilihat pada Gambar 1.



Gambar 1. Flowchart proses pengolahan citra

C. Metode Learning Vector Quantization (LVQ)

Learning Vector Quantization (LVQ) merupakan suatu metode neural network untuk melakukan pembelajaran pada lapisan kompetitif yang terawasi. Suatu lapisan kompetitif akan secara otomatis belajar untuk mengklasifikasikan vektor – vektor input. Kelas – kelas yang didapatkan sebagai hasil dari lapisan kompetitif ini hanya tergantung pada jarak antara vektor – vektor input. Jika dua vektor input memiliki nilai yang mendekati atau hampir sama, maka dalam lapisan kompetitif akan mengenali kedua vektor input tersebut dalam kelas yang sama (Anangadipa, Hidayatno, & Zahra, 2014; Aprillia, 2018; Kusumadewi, 2003). 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. Sedangkan algoritma pengujiannya adalah menghitung nilai output (kelas vektor) yang terdekat dengan vektor input, atau dapat disamakan dengan proses pengklasifikasian (pengelompokkan). Algoritma pembelajaran jaringan LVQ dapat dijabarkan sebagai berikut (Fausett, 1994):

- Inialisasi vektor referensi.
- Inialisasi laju pelatihan (α) = 0.
- Selama kondisi berhenti belum terpenuhi, lakukan langkah 3 – 7.
- Untuk setiap pasangan data latih beserta targetnya ($x_i : t_i$) lakukan langkah 4 – 5.
- Carilah J sehingga minimum $\|x - w_j\|$
- Perbaharui bobot w_j sebagai berikut :
- Jika $T = C_j$ maka w_j (baru) = w_j (lama) + $\alpha |x - w_j|$
- Jika $T \neq C_j$ maka w_j (baru) = w_j (lama) - $\alpha |x - w_j|$
- Kurangi laju pelatihan ($\alpha = \alpha - \text{Dec } \alpha$; $\alpha = \alpha * \text{Dec } \alpha$)
- Uji syarat berhenti (cacah iterasi atau laju pelatihan).

Inialisasi bobot awal biasanya menggunakan bilangan acak atau menggunakan sebagian data latih yang mewakili setiap kelas yang ada. Adapun, *flowchart* metode LVQ dapat dilihat pada Gambar 2.

D. Performa Akurasi

Pengukuran performa akurasi yang akan digunakan pada penelitian ini yaitu menggunakan metode *Mean Square Error* (MSE). Adapun, MSE menggunakan persamaan (1).

$$MSE = \frac{1}{M} \sum_{t=1}^M (x_t - \hat{x}_t)^2 \quad (1)$$

Dimana, x_t adalah nilai data; $x_t - \hat{x}_t$ adalah nilai hasil; M adalah nilai pola..

E. Sampel Pola Karakter

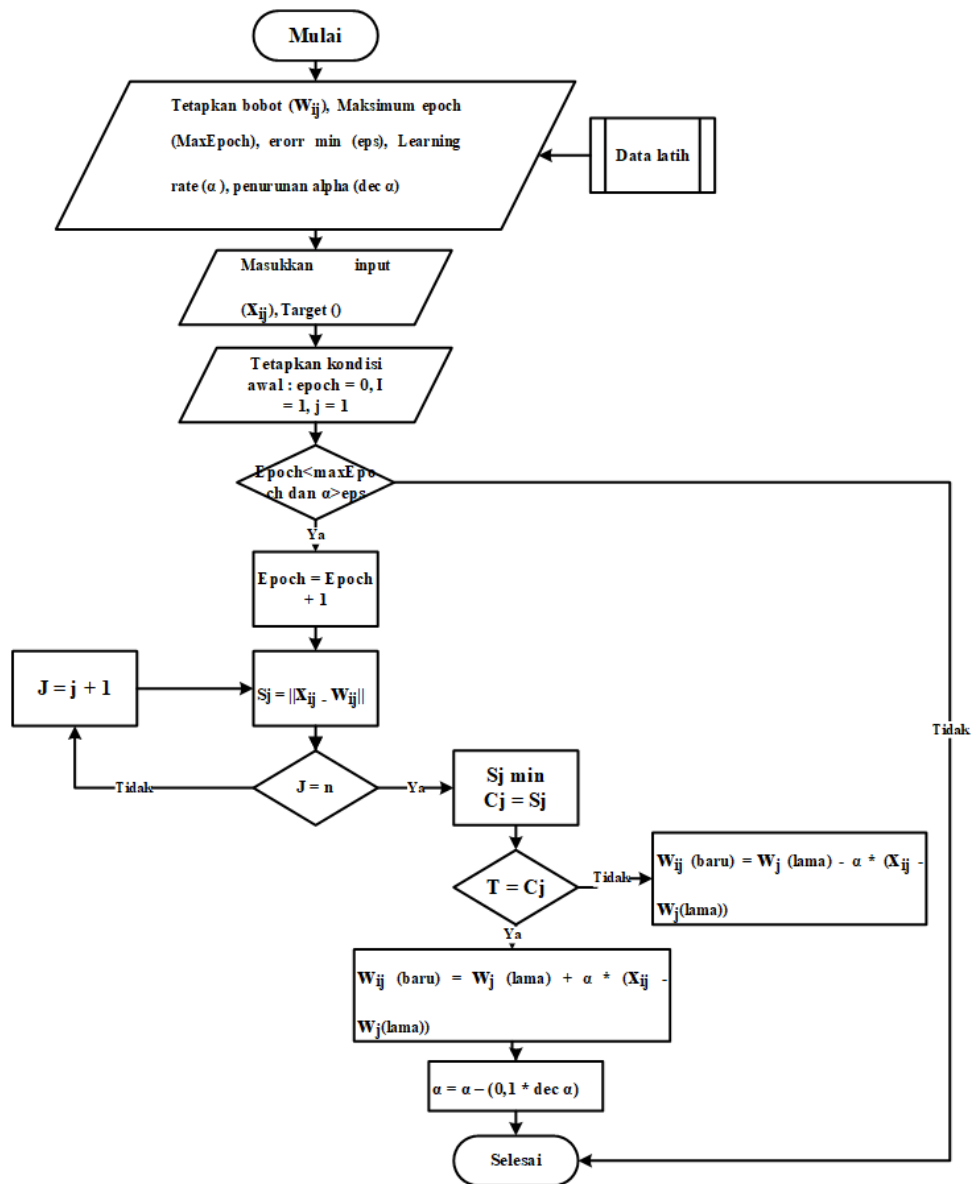
Dalam penelitian ini, data yang digunakan adalah data primer. Data diperoleh dengan cara mengambil gambar yang berupa tulisan tangan 15 orang dimana masing-masing menuliskan 7 huruf vokal aksara Sunda. Selain itu data tersebut akan digunakan untuk melakukan pelatihan program dan sebagai data uji sistem. Berikut adalah sampel data huruf vokal aksara Sunda dapat dilihat pada Gambar 3.

III. Hasil dan Pembahasan

Bagian ini adalah menerangkan hasil pengolahan citra, ekstraksi ciri yang dilanjutkan dengan pengujian menggunakan metode *Learning Vector Quantization* (LVQ).

A. Hasil Pengolahan Citra

Pengolahan citra merupakan suatu proses untuk mengubah suatu gambar menjadi gambar lain yang memiliki kualitas yang lebih baik untuk proses pengenalan pola selanjutnya . Berdasarkan tahapan proses pengolahan citra (Gambar 1), Tahap Pertama adalah akuisisi citra yaitu setelah data melewati proses *scanning* maka selanjutnya dilakukan proses *cropping* untuk memotong citra pada area yang dibutuhkan. Tahap kedua, proses normalisasi ukuran dengan dikalikan skala sebesar 0.2, citra awal berukuran 226 x 195 piksel dan ukuran setelah normalisasi menjadi 46 x 39 piksel. Tahap ketiga yaitu proses mengubah citra RGB ke citra *grayscale*. Tahap ke-4 yaitu tahap penajamn citra yang bertujuan untuk memperjelas citra. Tahap ke-5 yaitu segmentasi untuk mendapatkan citra biner dengan menggunakan metode *thresholding*. Dan yang terakhir yaitu penipisan citra untuk mengambil informasi penting dari sebuah citra. Adapun, hasil proses pengolahan citra dapat dilihat pada Gambar 4.

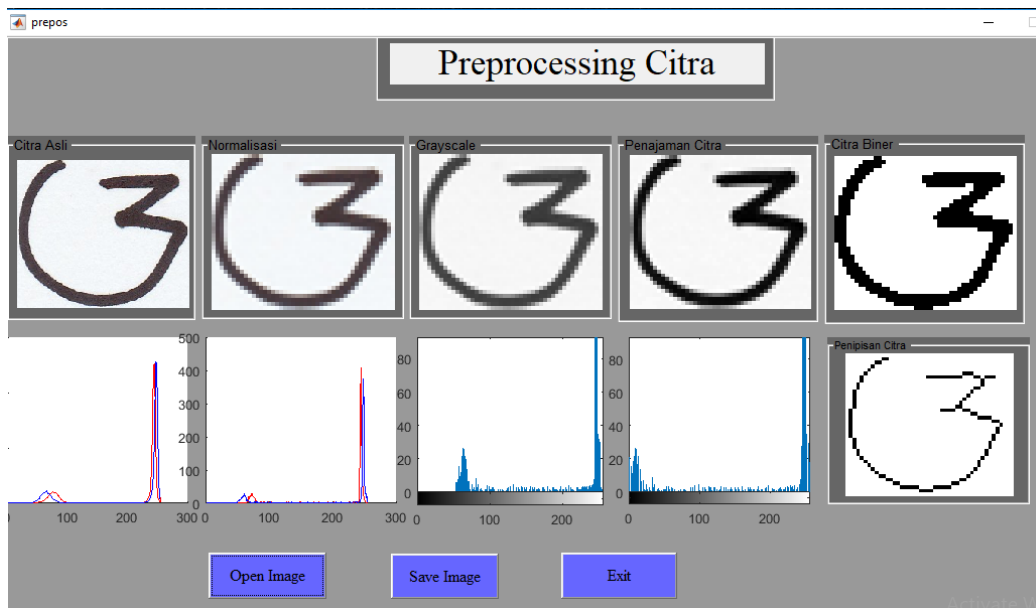


Gambar 2. Flowchart metode LVQ

ᮘ = a	ᮙ = é	ᮛ = i	ᮜ = o
ᮝ = u	ᮞ = e	ᮟ = eu	

ᮘ	ᮙ	ᮛ	ᮜ	ᮝ	ᮞ	ᮟ
a	é	i	o	u	e	eu

Gambar 3. Aksara Sunda Swara (Vokal mandiri)



Gambar 4. Hasil proses pengolahan citra

Selanjutnya yaitu masuk ke tahap ekstraksi ciri. Pada penelitian ini menggunakan algoritma intensitas karakter dan tanda arah. Intensitas karakter digunakan untuk menghitung nilai piksel hitam pada sebuah citra. Tanda arah digunakan untuk menghitung berapa banyak piksel yang memenuhi *masking* horizontal, vertical, diagonal kiri dan diagonal kanan pada semua citra. Sebelum melakukan pencirian, citra dibagi ke 9 segmen berukuran 14x14 piksel. Pembagian segmen tersebut menggunakan metode Chessboard yang membagi citra menjadi objek-objek seperti papan catur. Contohnya dapat dilihat pada gambar 5.



Gambar 5. Hasil pembagian 9 segmen

Selanjutnya yaitu mencirikan citra pada setiap segmen menggunakan tanda arah untuk menghitung *masking* vertical, horizontal, diagonal kiri, dan diagonal kanan. Lalu mencirikan intensitas karakter untuk menghitung nilai hitam pada tiap segmen. Setelah itu maka selanjutnya membuat 6 kombinasi ciri, yaitu sebagai berikut:

- Ciri 1 berisi variabel black
- Ciri 2 berisi diagonal 1
- Ciri 3 berisi variabel black dan diagonal 1
- Ciri 4 berisi variabel black dan diagonal 2
- Ciri 5 berisi variabel diagonal 1 dan diagonal 2
- Ciri 6 berisi variabel black, diagonal 1, diagonal 2, horizontal dan vertical.

Hasil ini kemudian disalin ke Excel untuk menjadikannya sebuah data set yang akan menjadi input di sistem pengenalan pola selanjutnya. Data set yang berjumlah 112 data dibagi menjadi 2 yaitu data *training* yang berjumlah 70 data dan data *testing* yang berjumlah 42 data.

B. Perhitungan Citra dengan Metode Learning Vector Quantization(LVQ)

Parameter pengujian metode LVQ dalam penelitian ini akan menggunakan 6 variasi arsitektur jaringan yang berisi 6 variasi nilai α (*learning rate*) dan 10 variasi jumlah neuron *hidden layer*, maka di setiap variasi nilai α akan diujikan ke 2 jumlah neuron *hidden layer* yang berbeda, kemudian juga menggunakan fungsi pembelajaran *trainlm*, dan fungsi aktivasi *logsig* dan *tansig*. Untuk setiap variasi akan diberikan 1000 *epoch*. Variasi yang akan dilakukan terhadap *learning rate* dan jumlah neuron *hidden layer* dapat dilihat pada Tabel 1.

Tabel 1. Variasi *learning rate* dan *hidden layer* metode LVQ

<i>Learning rate</i> (α)	Jumlah <i>hidden layer</i>	Fungsi pembelajaran
0,01	90	Learnlv1
	100	
0,02	90	
	100	
0,05	90	
	100	
0,1	90	
	100	
0,5	90	
	100	
1	90	
	100	

Pada penelitian ini akan menganalisis pengaruh perubahan *learning rate* (α) dan jumlah *hidden layer* terhadap MSE, lama waktu yang dibutuhkan, jumlah data yang berhasil dikenali dan tingkat keakurasiannya. Hasil pengujian dengan metode LVQ dapat dilihat pada Tabel 2.

Tabel 2. Hasil pengujian metode LVQ

α	Jumlah <i>hidden layer</i>	Hasil				
		Best <i>Epoch</i>	MSE	Waktu	Data Benar	Akurasi
0,01	90	241	0.0408	17 menit 22 detik	28	66,66%
	100	643	0.0449	18 menit 34 detik	24	57,14%
0,02	90	161	0.0490	18 menit 25 detik	25	59,52%
	100	26	0.0490	18 menit 42 detik	23	54,76%
0,05	90	422	0.0449	18 menit 43 detik	27	64,28%
	100	206	0.0449	18 menit 14 detik	17	40,47%
0,1	90	881	0.367	17 menit 43 detik	19	45,28%
	100	98	0.0490	16 menit 56 detik	22	52,38%
0,5	90	77	0.143	18 menit 11 detik	6	14,28%
	100	127	0.143	18 menit 8 detik	6	14,28%
1	90	2	0.176	16 menit 59 detik	6	14,28%
	100	109	0.143	20 menit	6	14,28%

IV. Kesimpulan

Dari Tabel 2 yang berisi hasil pengujian, didapatkan bahwa variasi arsitektur jaringan yang optimal dalam proses pengenalan citra aksara Sunda dengan menggunakan metode LVQ adalah dengan *learning rate* sebesar 0.01 dan jumlah *hidden layer* sebanyak 90 layer. Diperoleh *best epoch* sebesar 241, besar MSE yaitu 0.0408 dan tingkat akurasi sebesar 66.66% dengan waktu pembelajaran selama 17 menit 22 detik. Data yang berhasil dikenali sebanyak 28 data dari keseluruhan 42 data.

Daftar Pustaka

- Afrianto, I., & Priatama, D. (2013). *APLIKASI MOBILE PENGENALAN CITRA MENGGUNAKAN METODE LEARNING VECTOR*. 39–44.
- Ananggadipa, G., Hidayatno, A., & Zahra, A. A. (2014). *PENGENALAN HURUF ALFABET MENGGUNAKAN TUJUH INVARIAN MOMEN HU DAN JARINGAN SARAF TIRUAN LVQ (LEARNING VECTOR QUANTIZATION)*. *TRANSIENT*.
- Aprillia, E. M. (2018). *Aplikasi Pengolahan Citra untuk Klasifikasi Jenis Buah Mangga*. *Simki-Techsain*.

- Dani, R., Sugiharto, A., & Winara, G. A. (2015). Aplikasi Pengolahan Citra Dalam Pengenalan Pola Huruf Ngalagena Menggunakan MATLAB. *Konferensi Nasional Sistem & Informatika*.
- Fausett, L. (1994). *Fundamentals of Neural Network : architectures, algorithms, and applications*. Melbourne: Florida Institute of Technology.
- Kusumadewi, S. (2003). *Artificial Intelligence (Teknik dan Aplikasinya)* (1st ed.). Yogyakarta: Graha Ilmu.
- Lestari, D. D., Hidayat, B., & Andini, N. (2015). Perancangan Pengenal Kata Dalam Aksara Sunda Menggunakan Metode Deteksi Tepi dan LVQ Berbasis Pengolahan Citra Pada Android. 2(2), 3111–3119.
- Lukman, A. (2015). *Implementasi Pengolahan Citra dan Algoritma LVQ Untuk Pengenalan Pola Buku*. (20).
- Pratiarso, A., Hadi, M. Z. S., Sulthon, A., Endah, S. U., Elektronika, P., & Surabaya, N. (2011). *Perbandingan Metode POC, Backpropagation, Coding pada Pembacaan Plat Nomor Kendaraan Berbasis Image Processing*.
- Putra, D. (2010). *Pengolahan Citra Digital* (Westriningsih, ed.). Yogyakarta: ANDI OFFSET.
- Ramadhan, G., Djamal, E. C., & Darmanto, T. (2016). Klasifikasi Identitas Wajah Untuk Otorisasi Menggunakan Deteksi Tepi dan LVQ. *Seminar Nasional Aplikasi Teknologi Informasi (SNATi) 2016 Yogyakarta*, 37–41.
- Riansyah, R. R., Nurhasanah, Y. I., & Dewi, I. A. (2017). Sistem Pengenalan Aksara Sunda Menggunakan Metode Modified Direction Feature dan Learning Vector Quantization. *Jurnal Teknik Informatika Dan Sistem Informasi*. <https://doi.org/10.28932/jutisi.v3i1.563>
- Saleh, D. (2006). *Mencari Gerbang Pakuan dan Kajian Lainnya Mengenai Budaya Sunda* (Sundalana). Bandung: Pusat Studi Sunda.
- Sambani, E. B., Uryani, N. S., & Putra, R. A. K. (2016). *Algoritma Learning Vector Quantization Untuk Pengenalan Barcode Buku Di Perpustakaan Universitas Galuh Ciamis*. 6–7.
- Sundari, J. (2016). Melestarikan Aksara Sunda Dengan Aplikasi Multimedia. *EVOLUSI - Jurnal Sains Dan Manajemen AMIK BSI Purwokerto*.
- Utomo, P., & Cahyono, S. D. (2014). Sistem Deteksi Pra Panen Padi Berdasarkan Warna Daun Dengan Menggunakan Learning Vector Quantization (LVQ). *Prosiding Seminar Nasional Manajemen Teknologi XX*, 1–6.