



Contents lists available at [JESA](https://jesa.aks.or.id)

Journal of Engineering and Science Application

journal homepage:
<https://jesa.aks.or.id/index.php/jesa/index>



Pengenalan Aksara Batak Toba Menggunakan Modified Direction Feature dan Learning Vector Quantization

Lola Oktavia

Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, UIN Sultan Syarif Kasim Riau, Pekanbaru 28293, Indonesia

Informasi Artikel

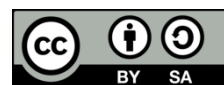
Kata Kunci:

Aksara Batak Toba
Ekstraksi Ciri
Learning Vector
Quantization (LVQ)
Modified Direction Feature
(MDF)
Thinning

ABSTRAK

Aksara Batak Toba merupakan salah satu warisan budaya Nusantara yang menghadapi tantangan kepunahan akibat minimnya pemahaman masyarakat terhadap huruf dasar (*Ina Ni Surat*). Pengembangan sistem pengenalan pola tulisan tangan aksara daerah (*Optical Character Recognition*) sangat krusial untuk pelestarian literasi digital dan edukasi budaya. Penelitian ini bertujuan untuk mengimplementasikan dan menganalisis performa metode *Modified Direction Feature* (MDF) sebagai ekstraksi ciri arah dan transisi garis, yang dikombinasikan dengan algoritma jaringan syaraf tiruan *Learning Vector Quantization* (LVQ) untuk klasifikasi karakter *Ina Ni Surat*. Dataset yang digunakan berjumlah 210 citra tulisan tangan dari 10 penulis berbeda dengan ukuran normalisasi 120×120 piksel. Tahapan pra-pengolahan citra meliputi *cropping*, konversi *grayscale*, binerisasi, dan *thinning* kerangka morfologi. Pengujian dilakukan secara sistematis dengan mengevaluasi variasi *learning rate* (0,001 hingga 0,1) dan rasio pembagian data latih terhadap data uji (90:10 hingga 50:50). Hasil eksperimen menunjukkan bahwa integrasi MDF dan LVQ mampu mengenali aksara Batak Toba dengan akurasi tertinggi sebesar 87% pada rasio data 70:30 dengan *learning rate* 0,1; 0,07; dan 0,09. Nilai rata-rata akurasi keseluruhan parameter mencapai performa optimal 83,6% pada *learning rate* 0,07.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Penulis Korespondensi:

Lola Oktavia
Program Studi Teknik Informatika
UIN Sultan Syarif Kasim Riau
Jl. HR. Soebrantas Km. 15, Simpang Baru, Tuah Madani, Kota Pekanbaru, Riau.
lola.oktavia@uin-suska.ac.id

1. PENDAHULUAN

Indonesia memiliki keanekaragaman warisan budaya yang sangat kaya. Berdasarkan komitmen pembangunan berkelanjutan (*Sustainable Development Goals* / SDGs), sektor kebudayaan memegang peranan strategis dalam menopang pendidikan bermutu dan pelestarian pengetahuan lokal [1]. Salah satu representasi

budaya Nusantara yang bernilai historis tinggi adalah aksara tradisional daerah, khususnya aksara Batak Toba yang tercatat dalam manuskrip kuno (*Pustaka*) [2]. Aksara Batak Toba memiliki struktur aksara dasar yang dikenal sebagai *Ina Ni Surat*, yang terdiri dari 19 hingga 21 karakter konsonan mandiri. Namun, di era digitalisasi saat ini, pemahaman generasi muda dalam membaca dan mengenali karakter dasar *Ina Ni Surat* mengalami penurunan signifikan karena tingkat kompleksitas bentuk guratan, sudut kemiringan, dan kemiripan antar-pola karakter [2], [3].

Tantangan visual pada tulisan tangan aksara tradisional tersebut mendorong perlunya teknologi *Optical Character Recognition* (OCR) dan pengenalan pola cerdas (*pattern recognition*) [4], [5]. Sistem OCR untuk aksara daerah memerlukan tahapan ekstraksi fitur yang tangguh dalam merepresentasikan informasi kontur serta algoritma klasifikasi yang adaptif dalam membedakan batas-batas keputusan antar-kelas karakter [6]–[8]. Beberapa penelitian terdahulu telah mengkaji pengenalan aksara Nusantara, seperti aksara Jawa menggunakan metode *k-Nearest Neighbor* dan jaringan syaraf tiruan [9], aksara Sunda dengan metode *Chain Code* dan *Support Vector Machine* [10], [11], aksara Bali [12], serta aksara Lampung [13]. Penelitian pengenalan aksara Batak Toba terdahulu juga telah mengeksplorasi penggunaan jaringan syaraf tiruan dan pencocokan pola sederhana [14], [15].

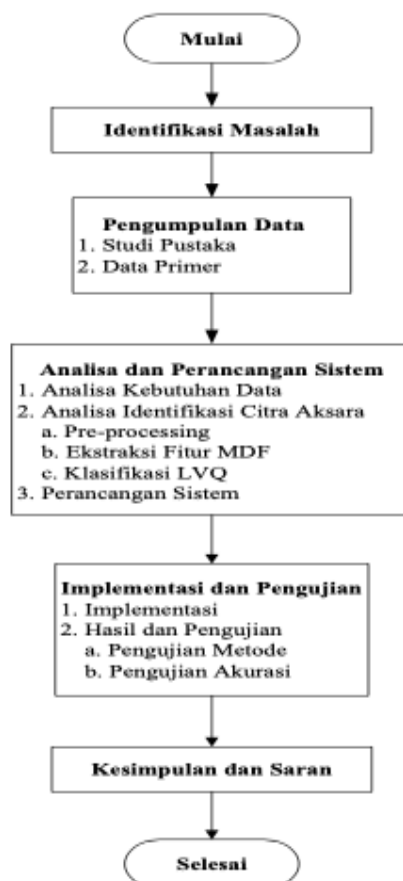
Untuk menghasilkan deskriptor bentuk yang invariant terhadap variasi ketebalan dan deformasi guratan tangan, metode ekstraksi ciri berbasis arah menjadi pendekatan yang sangat efektif [16]. *Direction Feature* (DF) dan pengembangannya, yaitu *Modified Direction Feature* (MDF), telah terbukti efektif dalam mentransformasi citra biner berkerangka tipis (*skeletonized image*) menjadi vektor fitur transisi dan arah garis dari empat arah pemindaian spasial [17], [18]. Keunggulan MDF terletak pada kemampuannya mengombinasikan nilai *Direction Feature* (DF) dan *Transition Feature* (TF) yang mencakup *Location Transitions* (LT) serta *Distance Transitions* (DT), sehingga memberikan representasi topologi struktur aksara yang sangat diskriminatif [18], [19].

Pada tahap klasifikasi, jaringan syaraf tiruan model *Learning Vector Quantization* (LVQ) memiliki keunggulan kompetitif dalam menangani pemisahan kelas nonlinear melalui penyesuaian vektor bobot acuan (*codebook vectors*) yang adaptif [20], [21]. LVQ telah sukses diterapkan pada berbagai domain pengenalan pola, seperti pengenalan huruf Hiragana [22], aksara Hijaiyah [23], pengenalan tanda tangan [24], klasifikasi dokumen historis [25], hingga klasifikasi data biomedis [26].

Meskipun MDF dan LVQ telah diterapkan pada sejumlah domain karakter, evaluasi komprehensif mengenai pengaruh parameter *learning rate* dan variasi rasio partisi data terhadap performa kombinasi MDF–LVQ pada aksara Batak Toba masih terbatas. Oleh karena itu, penelitian ini bertujuan untuk mengimplementasikan integrasi metode *Modified Direction Feature* (MDF) dan *Learning Vector Quantization* (LVQ) dalam pengenalan citra tulisan tangan *Ina Ni Surat* aksara Batak Toba, serta menganalisis pengaruh variasi parameter pelatihan terhadap tingkat akurasi pengenalan karakter.

2. METODE PENELITIAN

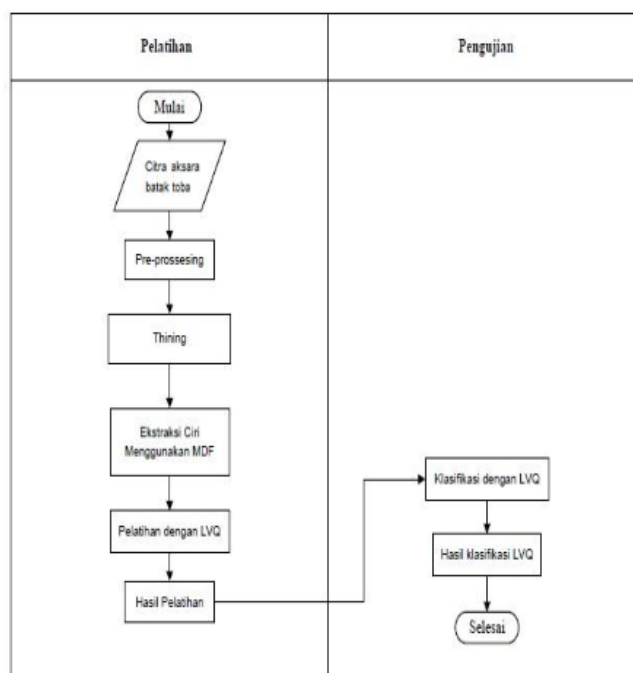
Metodologi penelitian ini berisi penjelasan langkah-langkah dan tahapan yang berurutan dalam penelitian. Berikut merupakan metodologi penelitian yang akan dilakukan dalam penelitian yang dapat dilihat pada gambar 2.1.



Gambar 2.1 Metodologi Penelitian

2.1. Analisa dan Perancangan Sistem

Proses analisa ini dilakukan untuk mendapatkan informasi dari data yang diperlukan dalam penelitian. Berikut ini merupakan alur tahapan analisa yang dapat di lihat pada gambar 2.2.



Gambar 2.2 Alur Tahapan Analisa Identifikasi Aksara Batak Toba

Berdasarkan dari alur tahapan diatas, terlihat bahwa proses pertama di lakukan adalah citra aksara batak toba dilakukan tahap *pre-processing*. Pada tahap *pre-processing* ini dilakukan mengubah ukuran gambar atau *me-resize* dengan menggunakan *Matlab* dan melakukan *cropping* menggunakan *Photoshop*. Setelah mendapatkan hasil dari *pre-processing*, maka langkah selanjutnya yaitu melakukan proses *thinning* untuk penipisan citra. Kemudian dilakukan proses ekstraksi ciri aksara batak toba dengan menggunakan metode (MDF) *Modified Direction Feature*. Selanjutnya, hasil dari ekstraksi ciri ini akan menjadi masukan untuk pelatihan dan pengujian menggunakan LVQ sehingga mendapatkan hasil pada klasifikasi.

2.2. Analisa Kebutuhan Data

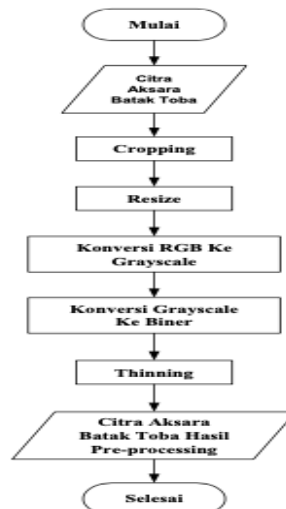
Data yang digunakan pada penelitian ini merupakan data primer, yaitu data yang didapatkan langsung oleh peneliti dalam penelitiannya. Data yang digunakan adalah data citra tulisan aksara batak toba yang berjumlah 210. Masing-masing kata dibuat dengan 10 buah citra dengan cara dituliskan dengan orang yang berbeda sebanyak 10 orang. Pengambilan data dilakukan dengan cara di photo dari tampak atas citra menggunakan kamera Hp. Dan untuk pengambilan gambar menggunakan background berwarna putih.

2.3. Analisa Proses Identifikasi Citra Aksara

Pada tahap penelitian ini, analisa proses identifikasi citra aksara merupakan proses citra yang dimulai dari tahapan *pre-processing*, kemudian lanjut ke tahapan ekstraksi ciri menggunakan metode (MDF) *Modified Direction Feature*. Hasil dari ekstraksi ciri ini akan digunakan sebagai inputan pada proses klasifikasi. Untuk proses klasifikasi menggunakan metode (LVQ) *Learning Vector Quantization*. Berikut ini adalah tahapan proses identifikasi citra aksara batak toba.

a. *Pre-processing*

Pre-processing adalah tahapan awal yang dilakukan untuk melakukan perbaikan citra. Tujuan dari *pre-processing* ini adalah untuk mendapatkan hasil citra yang lebih baik. Tahapan *pre-processing* yang dilakukan dalam tahap ini yaitu melakukan proses *cropping*, *resize*, konversi RGB ke *grayscale*, konversi *grayscale* ke biner, dan proses *thinning*. Untuk proses *resize* ini dilakukan dengan menggunakan *code matlab* sesuai dengan ukuran yang telah ditentukan yaitu, 120x120 *pixel*. Berikut ini adalah *flowchart* dari proses *preprocessing*:

Gambar 2.3 *Flowchart Pre-processing*

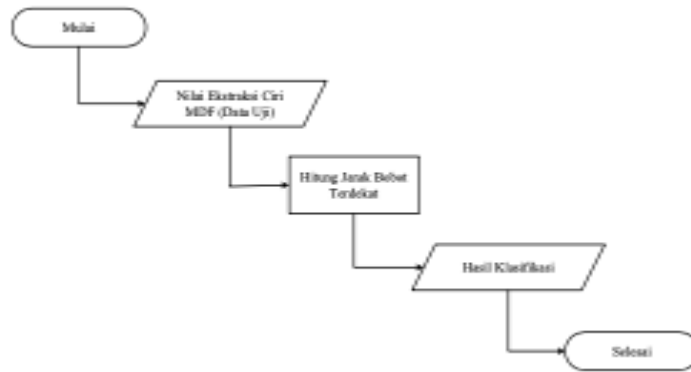
2.4. Ekstraksi Fitur MDF

Ekstraksi fitur pada MDF ini menggunakan data citra yang sudah melalui tahap *thinning* terlebih dahulu. Berikut adalah *flowchart* metode MDF yang dapat di lihat pada gambar berikut:

Gambar 2.4 *Flowchart MDF*

2.5. Klasifikasi *Learning Vector Quantization* (LVQ)

Proses klasifikasi pada LVQ (*Learning Vector Quantization*), dimana proses klasifikasi pada penelitian ini nilai-nilai ekstraksi ciri fitur dari MDF (*Modified Direction Feature*) menjadi acuan dalam penetapan klasifikasi LVQ serta proses tahapan LVQ dalam menentukan kelas berdasarkan data uji. Berikut merupakan proses klasifikasi LVQ yang dapat dilihat pada gambar *flowchart* berikut ini.



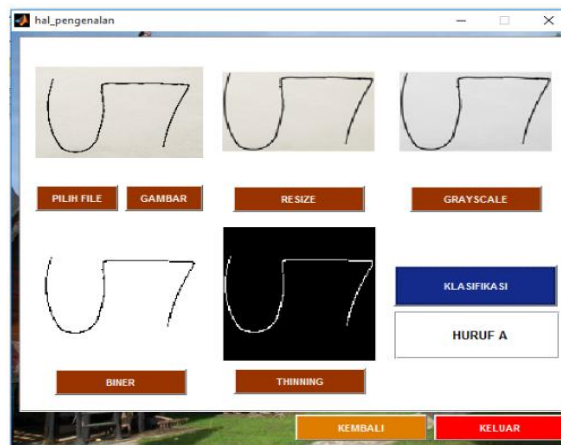
Gambar 2.5 Proses Klasifikasi

Hasil dari ekstraksi pada citra uji, dimana hasil tersebut akan dibandingkan dengan hasil ekstraksi citra latih yang telah di simpan di dalam database. Selanjutnya, untuk mengetahui kelas uji, maka perlu acuan untuk mendapatkan perbandingan kelas terhadap hasil citra latih yang telah di lakukan pelatihan sebelumnya. Tahap selanjutnya adalah melakukan proses perhitungan LVQ dari vektor data latih pada setiap kelas. Perhitungan ini dilakukan sesuai *learning rate* dan maksimal *epoch* yang digunakan pada penelitian. Untuk langkah awal yang dilakukan adalah dengan menentukan jarak minimum pada data uji.

3. HASIL DAN PEMBAHASAN

3.1. Hasil Klasifikasi Pengenalan Citra Pada Sistem

Hasil dari klasifikasi pengenalan citra aksara batak toba pada sistem menampilkan hasil akhir dari klasifikasi berupa “Huruf A” yang dapat di lihat pada gambar 3.1.



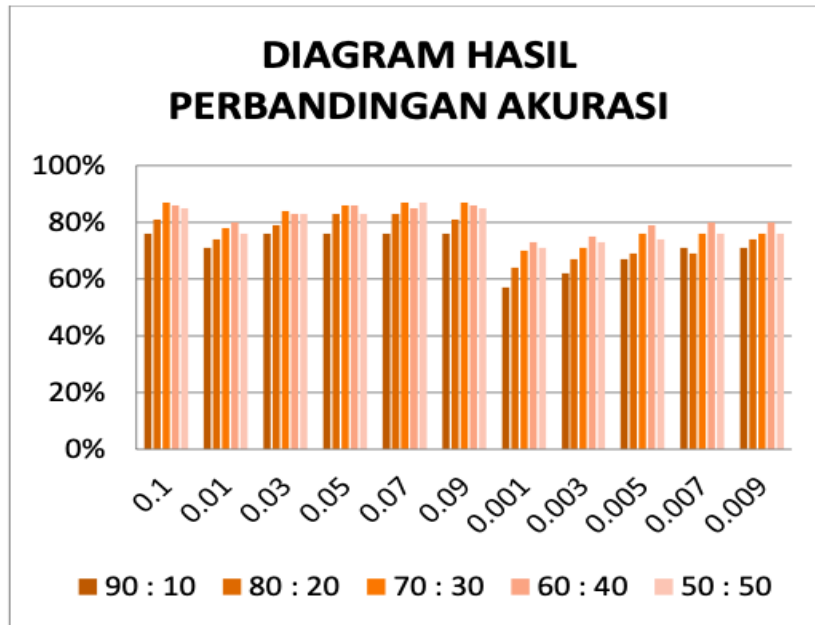
Gambar 3.1 Hasil Klasifikasi Pengenalan Citra Huruf Aksara Batak Toba

Pada penelitian ini, untuk tahapan pengujian dilakukan sesuai dengan rancangan pengujian sebagai berikut:

1. Pengujian pada penelitian ini menggunakan pengujian *whitebox* pada aplikasi, untuk mengetahui *source code* algoritma *pre-processing* MDF, LVQ, dan pengujian akurasi menggunakan *confusion matrix*.
2. Pengujian dilakukan berdasarkan hasil ekstraksi ciri MDF. Hasil ekstraksi ciri MDF ini yang akan dilakukan pelatihan dan pengujian dalam proses klasifikasi.
3. Pengujian untuk metode klasifikasi LVQ (*Learning Vector Quantization*) pengujian dilakukan dengan menggunakan nilai *learning rate* yang berbeda yaitu, 0.1, 0.01, 0.03, 0.05, 0.07, 0.09, 0.001, 0.003, 0.005, 0.007 dan 0.009.

4. Dalam pengujian citra, citra uji yang digunakan berasal dari data citra penelitian dan bukan termasuk pada citra data latih.
5. Pengujian dilakukan berdasarkan pembagian data uji dengan data latih yaitu, data uji 50%,40%,30%,20%, dan 10% dari seluruh data.

Hasil dari perbandingan akurasi yang telah dilakukan dapat di lihat pada diagram hasil berikut ini.



Gambar 4.2 Diagram Hasil Perbandingan Akurasi

Dari hasil pengujian yang telah dilakukan, maka dapat dilakukan pencarian rata-rata dari masing-masing *learning rate* atas percobaan terhadap pembagian data. Berikut adalah tabel pencarian rata-rata *learning rate* dari seluruh pengujian pembagian data yang dapat di lihat pada tabel 4.1

Tabel 4.1. Tabel Pencarian Rata-rata Learning Rate

No	Ln	90 : 10	80 : 20	70 : 30	60 : 40	50 : 50	Rata-Rata
1.	0.1	76%	81%	87%	86%	85%	83%
2.	0.01	71%	74%	78%	80%	76%	75,8%
3.	0.03	76%	79%	84%	83%	83%	81%
4.	0.05	76%	83%	86%	86%	83%	82,8%
5.	0.07	76%	83%	87%	85%	87%	83,6%
6.	0.09	76%	81%	87%	86%	85%	83%
7.	0.001	57%	64%	70%	73%	71%	67%
8.	0.003	62%	67%	71%	75%	73%	69,6%
9.	0.005	67%	69%	76%	79%	74%	73%
10.	0.007	71%	69%	76%	80%	76%	74,4%
11.	0.009	71%	74%	76%	80%	76%	75,4%

Berdasarkan hasil tabel tersebut maka dapat di lihat untuk akurasi terbaik dari 11 macam pembagian *learning rate* yang dilakukan, terdapat *Ln* pada 0.1, 0.07, dan 0.09 sebesar 83%. Untuk akurasi tertinggi pada *learning rate* 0.1, 0.07, dan 0.09 memperoleh akurasi tertinggi sebesar 87% dengan pembagian data 70:30 dan 50:50. Sedangkan *learning rate* dengan rata-rata akurasi terendah dari pengujian ini sebesar 67% dan 69%, dengan pembagian data 80:20 memperoleh akurasi sebesar 64% dan 67% yang terdapat pada *learning rate* 0.001, dan 0.003.

4. KESIMPULAN

Berdasarkan hasil dari pembahasan yang telah dilakukan dalam penelitian ini, maka dapat ditarik kesimpulan Aplikasi klasifikasi citra pengenalan huruf aksara batak toba yang dibangun dengan menerapkan konsep ekstraksi ciri *Modified Direction Feature* (MDF) dan beberapa proses lainnya seperti *resize*, konversi RGB ke *grayscale*, konversi *grayscale* ke biner, *thinning*, dan klasifikasi citra menggunakan *Learning Vector Quantization* (LVQ) mampu mengenali huruf aksara batak toba dengan matriks berukuran 120x120 yang memperoleh nilai akurasi tertinggi sebesar 87% dengan pembagian data 70:30% dan *learning rate* 0.1, 0.07, 0.09. Penggunaan nilai pada *learning rate* dapat mempengaruhi hasil dari akurasi. Untuk penggunaan nilai pada *learning rate* terbaik dari 11 jenis *learning rate* yang di uji adalah *learning rate* 0.1, 0.07, dan 0.09 dengan nilai rata-rata akurasi sebesar 83%. Akurasi terendah untuk pengenalan huruf aksara batak toba dengan matriks berukuran 120x120 terdapat pada *learning rate* 0.001 dengan nilai rata-rata akurasi sebesar 67%.

DAFTAR PUSTAKA

- [1] UNESCO, "Culture for the 2030 agenda," *United Nations Educational, Scientific and Cultural Organization*, Paris, France, Rep. CLT-2019/WS/1, 2019.
- [2] U. Kozok, *Warisan Leluhur: Sastra Lama dan Aksara Batak*, 1st ed. Jakarta: Kepustakaan Populer Gramedia, 1999.
- [3] K. Amend and M. S. Ruiz, *Handwriting Analysis: The Complete Basic Book*, 2nd ed. Franklin Lakes, NJ, USA: New Page Books, 1998.
- [4] R. Plamondon and S. N. Srihari, "Online and off-line handwriting recognition: A comprehensive survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 63–84, Jan. 2000, doi: 10.1109/34.824821.
- [5] J. Pradeep, E. Srinivasan, and S. Himavathi, "Diagonal based feature extraction for handwritten alphabets recognition system using neural network," *Int. J. Comput. Sci. Inf. Technol.*, vol. 3, no. 1, pp. 364–368, Feb. 2011.
- [6] A. Kadir and A. Susanto, *Teori dan Aplikasi Pengolahan Citra*, 1st ed. Yogyakarta: Andi Offset, 2013.
- [7] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. New York, NY, USA: Pearson, 2018.
- [8] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4th ed. Burlington, MA, USA: Academic Press, 2009.
- [9] A. C. Agustina, "Pengenalan aksara Jawa menggunakan Learning Vector Quantization (LVQ)," *J. Ilm. Teknol. Inf.*, vol. 10, no. 2, pp. 91–98, Jul. 2011.
- [10] E. F. Malik, B. Purnama, and M. D. Sulistiyo, "Analisis dan implementasi Optical Character Recognition (OCR) menggunakan Jaringan Syaraf Tiruan (JST) untuk pengenalan aksara Sunda baku," *J. Inform.*, vol. 8, no. 1, pp. 45–54, Apr. 2012.
- [11] A. A. Farhan, "Perancangan dan analisis sistem pengenalan kata aksara Sunda menggunakan metode Learning Vector Quantization berbasis pengolahan citra," *J. Elektro Telekomun. Terap.*, vol. 4, no. 1, pp. 488–496, Des. 2017.
- [12] I. M. G. Sunarya, I. G. M. Darmawiguna, and K. Agustini, "Penerapan metode Modified Direction Feature (MDF) pada pengenalan karakter aksara Bali berbasis jaringan syaraf tiruan," *J. Nas. Pendidik. Teknol. Inform.*, vol. 5, no. 2, pp. 88–95, Agu. 2016.
- [13] A. Aryantio and R. Munir, "Pengenalan aksara Lampung menggunakan ekstraksi kontur dan jaringan syaraf tiruan," dalam *Proc. Konf. Nas. Inform. (KNIF)*, Bandung, 2015, pp. 112–117.
- [14] L. K. Hutauruk, "Pengembangan perangkat lunak pengenalan pola aksara Batak Toba menggunakan pendekatan Learning Vector Quantization," Skripsi, Dept. Tek. Inform., STMIK Mikroskil, Medan, Indonesia, 2015.

- [15] F. Tampubolon, S. Sembiring, and M. Zarlis, "Implementasi jaringan syaraf tiruan backpropagation dalam mengenali karakter aksara Batak Toba," *J. Inf. Rekayasa Politek.*, vol. 2, no. 1, pp. 33–40, Jan. 2018.
- [16] S. Impedovo and P. Wang, *Machine Interpretation of Handwriting*, Singapore: World Scientific, 1993.
- [17] M. Blumenstein, B. Verma, and H. Nguyen, "A modified direction feature for cursive character recognition," dalam *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Honolulu, HI, USA, 2002, pp. 2983–2987, doi: 10.1109/IJCNN.2002.1007624.
- [18] M. Blumenstein, B. Verma, and H. Nguyen, "Neural-based solutions for handwriting recognition: An overview and comparative study," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 1, pp. 112–126, Jan. 2007.
- [19] N. Sherkat, R. J. Whitrow, and F. M. Cowell, "Feature extraction for offline recognition of handwriting," *Pattern Recognit. Lett.*, vol. 16, no. 8, pp. 843–851, Aug. 1995.
- [20] T. Kohonen, *Self-Organizing Maps*, 3rd ed. Berlin, Germany: Springer-Verlag, 2001.
- [21] J. J. Siang, *Jaringan Syaraf Tiruan dan Pemrogramannya Menggunakan MATLAB*, 2nd ed. Yogyakarta: Penerbit Andi, 2014.
- [22] B. Wirayuda, T. A. Nugroho, L. D. Krisnawati, and M. A. Wibisono, "Pengenalan pola huruf Jepang (Kana) menggunakan Direction Feature Extraction dan Learning Vector Quantization," *J. Penelit. Pengemb. Telekomun.*, vol. 13, no. 2, pp. 120–128, Des. 2008.
- [23] M. N. Muhamad and D. R. Prehanto, "Implementasi algoritma Chain Code dan Learning Vector Quantization (LVQ) pada pengenalan huruf Hijaiyah," *J. Teknol. Inform. Komput.*, vol. 3, no. 2, pp. 102–109, Okt. 2016.
- [24] F. Damayanti and W. Setiawan, "Pengenalan tanda tangan dengan metode Modified Direction Feature (MDF) dan Euclidean Distance," dalam *Proc. Conf. Smart-Green Technol. Electr. Inf. Syst. (ICSGTEIS)*, Bali, Indonesia, 2013, pp. 67–72.
- [25] T. Chuang, "Aplikasi pengenalan tulisan Mandarin kedalam Han Yu Pin Yin menggunakan metode Learning Vector Quantization (LVQ)," *J. Sist. Inform. Komput.*, vol. 1, no. 2, pp. 55–63, Des. 2011.
- [26] E. Budianita and W. Prijodiprodjo, "Penerapan Learning Vector Quantization untuk klasifikasi status gizi anak," *IJCCS (Indones. J. Comput. Cybern. Syst.)*, vol. 7, no. 2, pp. 155–166, Jul. 2013, doi: 10.22146/ijccs.3082.
- [27] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979, doi: 10.1109/TSMC.1979.4310076.
- [28] R. G. Keyes, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 6, pp. 1153–1160, Dec. 1981, doi: 10.1109/TASSP.1981.1163684.
- [29] J. Sauvola and M. Pietikäinen, "Adaptive document image binarization," *Pattern Recognit.*, vol. 33, no. 2, pp. 225–236, Feb. 2000, doi: 10.1016/S0031-3203(99)00055-2.
- [30] T. Y. Zhang and C. Y. Suen, "A fast parallel algorithm for thinning digital patterns," *Commun. ACM*, vol. 27, no. 3, pp. 236–239, Mar. 1984, doi: 10.1145/357994.358023.