

PENGARUH LINEAR BINARY PATTERN (LBP) DALAM PENGENALAN CITRA AKSARA JAWA BERBASIS OPTICAL CHARACTER RECOGNITION (OCR)

Christy Atika Sari¹, Wellia Shinta Sari², Putri Mega Arum Wijayanti³

^{1,3}Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro

²Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro

e-mail: ¹christy.atika.sari@dsn.dinus.ac.id, ²wellia.shinta@dsn.dinus.ac.id, ³111201811502@mhs.dinus.ac.id

ABSTRAK

Salah satu peninggalan budaya Indonesia dari tanah Jawa yaitu Aksara Jawa. Aksara Jawa telah digunakan oleh masyarakat sejak jaman dulu untuk menulis sastra dan menulis sehari-hari. Karena memiliki bentuk yang rumit aksara Jawa menjadi jarang dikenali. Oleh karena itu pada penelitian ini akan dijadikan salah satu upaya untuk belajar sekaligus melestarikan budaya khususnya Aksara Jawa yakni melakukan suatu transliterasi Aksara Jawa. Metode yang digunakan pada penelitian ini adalah dengan menggunakan Optical Character Recognition (OCR) berbasis Template matching dan ekstraksi fitur Linear Binary Pattern (LBP). OCR akan dapat membantu dalam proses pengkonversian gambar yang berisi tulisan Aksara Jawa yang nantinya akan dilakukan proses transliterasi. Cara kerja Template Matching adalah dengan mencocokkan tiap bagian pada citra dengan citra template yang telah ditentukan. LBP akan membantu dalam pengenalan huruf aksara yang memiliki objek terpisah. Berdasarkan dari pengujian diperoleh rata-rata akurasi sebesar 89,4%. Nilai akurasi tertinggi sebesar 100% dan nilai akurasi terendah sebesar 50%. Tingkat keberhasilan dalam proses transliterasi bergantung pada kejelasan objek huruf Aksara Jawa pada citra uji.

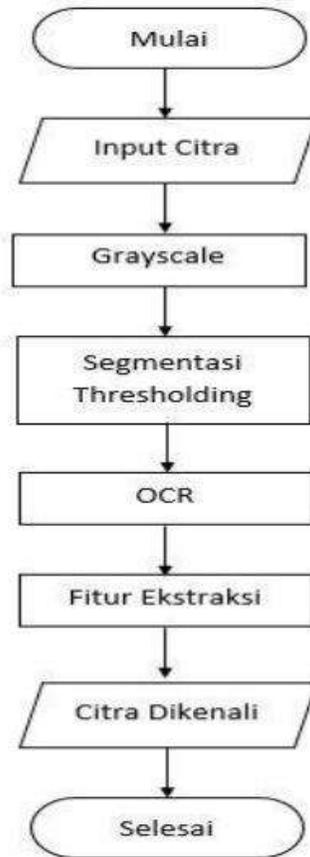
Kata Kunci: Linear Binary Pattern, Template Matching, Optical Character Recognition, Aksara Jawa.

1. PENDAHULUAN

Indonesia kaya akan bahasa dan memiliki keragaman budaya. Salah satu peninggalan budaya Indonesia dari tanah Jawa yaitu Aksara Jawa. Aksara Jawa atau lebih dikenali dengan sebutan Hanacaraka, sering digunakan untuk menulis sastra dan menulis sehari-hari dalam Bahasa Jawa, sejak pada abad pertengahan sekitar abad ke-15 sampai abad pertengahan ke-20, namun seiring perkembangannya zaman digunakan juga untuk menulis dalam berbagai macam bahasa daerah. Seiring berkembangnya zaman, Aksara Jawa seolah terlupakan dan jarang dikenali oleh masyarakat khususnya generasi muda sekarang. Namun Aksara Jawa hingga sekarang masih banyak dipelajari di banyak sekolah-sekolah di berbagai wilayah di sekitar Pulau Jawa sebagai kurikulum pembelajaran muatan lokal yang pada penggunaannya terbatas dalam rutinitas sehari – hari [1]–[5]. Walaupun begitu tidak jarang juga masih ditemukan tulisan Aksara Jawa pada tempat-tempat tertentu, seperti tempat wisata, sekolah, papan instansi pemerintah, dan papan nama jalan. Sebagai generasi muda, melestarikan budaya tulisan Aksara Jawa khususnya di Pulau Jawa sangatlah penting agar tetap dikenal sampai ke generasi berikutnya. Menurut [6], [7] dalam penelitiannya yang membahas permasalahan membagi tulisan aksara yang terdapat pada naskah menjadi bagian-bagian karakter aksara yang dapat digunakan dalam pengenalan Aksara Jawa. Rismiyati, Djalal [8] yang membahas permasalahan meningkatkan akurasi dalam pengenalan karakter Jawa menggunakan fitur HOG dan fitur Zone Base. Irham, Fairuz [7] dalam Jurnal Pattern Recognition of Javanese Letter Using Template Matching Correlation Method, membahas implementasi algoritma template matching untuk melakukan pengenalan pola serta mengenal dan transliterasi huruf Jawa dengan menggunakan metode Template Matching Correlation berbasis Optical Character Recognition dan membangun aplikasi. Aksara Jawa memiliki bentuk yang sulit serta memiliki unsur-unsur yang rumit yang terdapat pada Aksara Jawa.

Dari permasalahan tersebut perlu adanya proses pelestarian Aksara Jawa agar dapat kembali diminati terutama oleh generasi muda. Maka pada penelitian ini akan dijadikan salah satu upaya untuk belajar sekaligus melestarikan budaya Jawa khususnya Aksara Jawa yakni melakukan suatu transliterasi Aksara Jawa menjadi huruf latin/ huruf biasa. Selain dapat melakukan transliterasi, juga dapat membantu dalam memahami dan mempelajari huruf Aksara Jawa. Pengolahan Citra dikembangkan untuk tujuan memperbaiki atau meningkatkan kualitas citra yang lebih baik dari sebelumnya [9], [10]. Dalam pengolahan citra terdapat proses Grayscale yang mengubah citra berwarna menjadi citra abu-abu. Metode Thresholding adalah suatu teknik untuk melakukan pemetakan citra yang pada prosedurnya berdasarkan oleh kontras pada tingkat keabuan citra. Daerah gelap pada citra akan dijadikan menjadi lebih gelap, sedangkan daerah terang pada citra akan dijadikan menjadi lebih cerah. Hasil dari metode segmentasi threshold akan berbentuk citra biner serta memiliki nilai pixel sebesar 0 atau 1 [1], [11], [12]. Nilai intensitas piksel untuk warna hitam sempurna sebesar 0 dan untuk warna putih sempurna dengan nilai intensitas 1. Metode ini akan digunakan untuk memisahkan objek citra Aksara Jawa pada backgroundnya. Optical Character Recognition (OCR) [7], [13] merupakan

3. Template Matching. Template Matching [7], [17]–[19] merupakan salah satu teknik dalam pengolahan citra digital yang mengidentifikasi bagian-bagian pada gambar yang sesuai dengan template yang telah ditentukan. Gambar input dibandingkan dengan gambar template pada basis data, kemudian dicari persamaannya dengan menggunakan aturan tertentu. Pencocokan gambar yang menghasilkan tingkat kemiripan atau kemiripan yang tinggi menentukan suatu gambar dikenali sebagai salah satu gambar template.
4. Local Binary Pattern (LBP). Tujuan dari Local Binary Pattern (LBP) dirancang untuk klasifikasi tekstur. Keberhasilan dalam menggunakan ekstraksi LBP dalam kemampuan membedakan dan komputasi sederhana telah banyak dibuktikan melalui penelitian computer vision dalam mendeteksi. Hal ini dapat diterapkan dalam pendeteksian objek dan bukan hanya berfokus pada klasifikasi tekstur. Selama langkah dalam deteksi objek akurat, fitur LBP dapat digunakan [4], [20], [21].



Gambar 2. Tahapan Penelitian

Pada Gambar 3 menunjukkan flowchart jalannya program transliterasi Aksara Jawa. Proses dimulai dengan memasukkan citra uji Aksara Jawa, kemudian citra yang telah diinputkan dilakukan preprocessing citra grayscale, segmentasi thresholding, transliterasi menggunakan algoritma OCR berbasis template matching, ekstraksi fitur menggunakan LBP hingga diperoleh hasil transliterasi.

2.4 Pengujian Ekperimen

Ketepatan Optical Character Recognition (OCR) dalam proses transliterasi atau pengenalan Aksara Jawa akan digunakan untuk melakukan pengujian akurasi pada system. Terdapat dua cara dalam menghitung keefektifan OCR: pertama, hitung akurasi pada level karakter, dan kedua, hitung akurasi pada level kata [22] dimana *KB* adalah jumlah karakter benar dan *JK* adalah jumlah karakter salah.

$$\text{Akurasi} = \frac{KB}{JK} \times 100\% \quad (1)$$

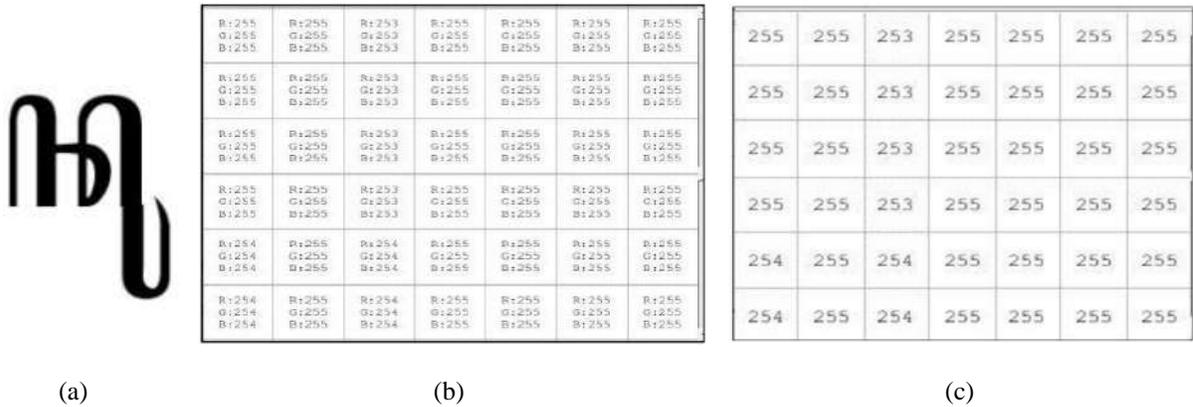
3. HASIL DAN PEMBAHASAN

3.1 Preprocessing

Preprocessing merupakan tahap pertama yang dilakukan dalam penelitian transliterasi Aksara Jawa. Pada tahap preprocessing dilakukan pengolahan citra untuk menghilangkan noise atau bagian citra yang tidak dibutuhkan yang selanjutnya akan dilakukan proses pengenalan citra. Tahapan preprocessing diantaranya:

1. Citra RGB ke Grayscale. Pada penelitian ini, citra uji berupa gambar dengan warna hitam dan putih, namun citra tersebut masih termasuk citra RGB karena belum melalui proses preprocessing. Citra yang telah diinputkan akan diubah menjadi grayscale. Citra grayscale atau keabuan merupakan citra dengan nilai piksel antara 0 sampai 255. Untuk mengubah citra RGB menjadi grayscale digunakan persamaan (2).

$$\text{Grayscale} = 0,299 \times R + 0,587 \times G + 0,114 \times B \quad (2)$$



Gambar 3. Konversi Citra : (a) Citra RGB, (b) Intensitas nilai piksel RGB, (c) Intensitas nilai piksel Grayscale

Analisis perhitungannya adalah sebagai berikut. Setelah citra RGB dilakukan proses grayscale menghasilkan citra grayscale yang hanya memiliki satu nilai keabuan 0 sampai 255.

Pixel (0,1)
Grayscale= $0.299 \times 255 + 0.587 \times 255 + 0.114 \times 255$
Grayscale= $76,245 + 149,685 + 29,07$
Grayscale= 255

Pixel (0,2)
Grayscale= $0.299 \times 255 + 0.587 \times 255 + 0.114 \times 255$
Grayscale= $76,245 + 149,685 + 29,07$
Grayscale= 255

Pixel (0,3)
Grayscale= $0.299 \times 253 + 0.587 \times 253 + 0.114 \times 253$
Grayscale= $75,647 + 148,511 + 28,842$
Grayscale= 253

Pixel (0,4)
Grayscale= $0.299 \times 255 + 0.587 \times 255 + 0.114 \times 255$
Grayscale= $76,245 + 149,685 + 29,07$
Grayscale= 255

2. Segmentasi thresholding digunakan untuk memisahkan citra objek dari background berdasarkan tingkat kecerahan citra. Daerah citra yang gelap akan dibuat semakin gelap, sebaliknya daerah citra yang terang akan dibuat semakin terang. Nilai piksel dari segmentasi thresholding adalah sebesar 0 (daerah gelap) atau 1 (daerah terang).



Gambar 4. (a) Citra grayscale, (b) Citra thresholding

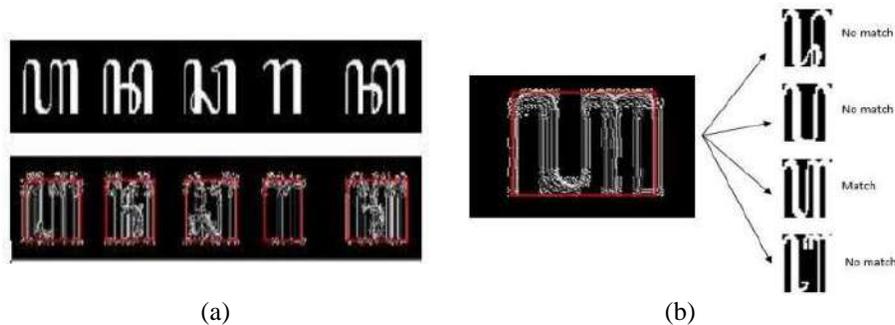
3. Citra Biner.

Fungsi dari penggunaan citra biner digunakan untuk mempermudah dalam proses transliterasi Aksara Jawa karena mempermudah dalam proses deteksi citra.



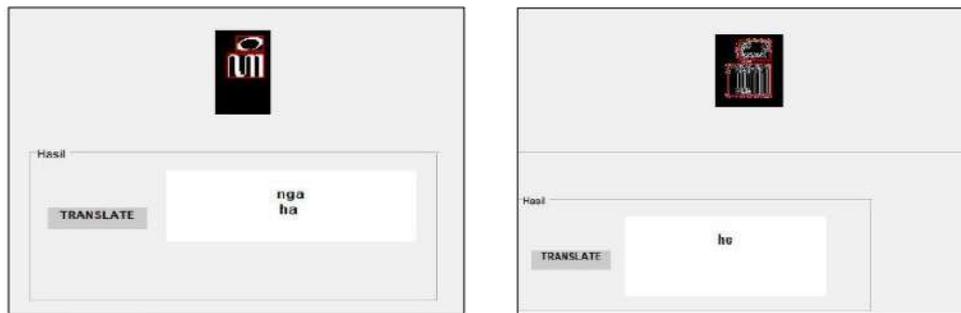
Gambar 5. Citra Biner

4. Proses Pengenalan OCR Berbasis Template Matching. Proses Pengenalan OCR Berbasis Template Matching Pengenalan karakter dengan menggunakan metode OCR berbasis Template Matching dilakukan dengan cara membandingkan citra uji dengan citra latih, selanjutnya citra yang memiliki kemiripan akan dianggap sebagai karakternya.



Gambar 6. Pengenalan OCR berbasis Template Matching: (a) Citra biner, (b) Pencocokan citra

5. Ekstraksi Fitur Local Binary Pattern (LBP). Local Binary Pattern (LBP) merupakan teknik ekstraksi yang sederhana namun sangat efisien yang melabeli piksel dari suatu gambar dengan ambang batas daerah piksel dengan menganggapnya sebagai biner [9]. Karena kemampuan LBP dalam membedakan objek citra dan komputasi yang sederhana, ekstraksi ciri LBP digunakan dalam penelitian ini untuk membantu dalam proses transliterasi Aksara Jawa. Penggunaan ekstraksi ciri LBP sangat membantu terutama untuk citra yang memiliki objek terpisah dalam penelitian seperti huruf aksara 'i', 'e', dan 'o'.



(a) Ekstraksi hasil tanpa LBP
(b) Ekstraksi hasil dengan LBP
Gambar 7. Ekstraksi Fitur Local Binary Pattern (LBP)

Sebagai contoh, dari hasil eksperimen ditemukan bahwa citra Aksara Jawa huruf “he” yang tidak menggunakan LBP akan teridentifikasi sebagai huruf “nga” dan “ha”, hal ini karena OCR mendeteksi dua karakter objek yang terpisah sehingga terdapat dua huruf yang terdeteksi dan terbaca. Sedangkan citra Aksara Jawa huruf “he” yang menggunakan ekstraksi ciri LBP akan menghasilkan huruf “he”, karena radius dalam LBP akan menghubungkan tiap piksel sehingga objek citra akan terhubung dan menjadi satu.

3.2 Pengujian Akurasi

Setelah semua citra uji dilakukan pengujian pada program, selanjutnya akan dilakukan perhitungan nilai akurasi pengujian. Dimana nilai akurasi diperoleh dari jumlah huruf benar yang terdeteksi dibagi dengan jumlah seluruh huruf dan dikalikan dengan 100%. Dengan taksiran semakin besar nilai akurasi yang didapatkan maka nilai kebenaran akan semakin tinggi. Hasil ketepatan pengujian program transliterasi Aksara Jawa pada 20 data citra uji adalah sebagai berikut.

Tabel 1. Hasil pengujian beberapa suku kata menggunakan fitur LBP

No	Citra Uji	Hasil Translate	Kata yang Benar	Akurasi
1		bahasajawa	bahasajawa	100%
2		bajubaru	bajubaru	100%
3		batako	batako	100%
4		cerita	cerita	100%
5		gulabatu	gulabatu	100%
6		kacamata	kacamata	100%
7		matematih ^a	matematika	80%
8		tusibasi	nasibasi	75%
9		rusalima	rusalima	100%
10		sahabiga	salatiga	50%
11		jepara	jepara	100%
12		mamuju	mamuju	100%
13		petani	petani	100%
14		pusaka	pusaka	100%
15		tukubuku	tukubuku	100%
16		nyangiji	nyawiji	66,7%
17		sijilorojilu	sijilorotelu	83,3%
18		merasi	merapi	66,7%
19		wisaga	wisata	66,7%
20		wonosobo	wonosobo	100%

Pada Tabel 1, terdapat huruf dengan font berwarna merah, huruf dengan warna merah tersebut merupakan hasil transliterasi yang kurang tepat. Penyebab dari kurang tepatnya hasil transliterasi citra adalah karena objek citra yang kurang jelas saat proses ekstraksi. Untuk mendapatkan hasil transliterasi yang lebih baik adalah dengan menggunakan citra dengan objek yang terlihat jelas. Sehingga dari Table 1 dapat disimpulkan bahwa transliterasi Aksara Jawa ke Latin dengan menggunakan Optical Character Recognition (OCR) berbasis Template Matching dengan menggunakan ekstraksi fitur LBP diperoleh rata-rata nilai akurasi sebesar 89,4%. Nilai akurasi sebesar 100% merupakan yang tertinggi dan 50% adalah yang terendah.

5. KESIMPULAN

Berdasarkan dari pengujian program transliterasi Aksara Jawa menggunakan Optical Character Recognition (OCR) berbasis Template Matching dengan menggunakan ekstraksi fitur Local Binary Pattern (LBP) dapat disimpulkan sebagai berikut:

1. Program transliterasi Aksara Jawa menggunakan Optical Character Recognition (OCR) berbasis Template Matching dengan menggunakan ekstraksi fitur Local Binary Pattern (LBP) telah dibuat dan berhasil untuk menerjemahkan citra uji Aksara Jawa.
2. Transliterasi dengan menggunakan Local Binary Pattern (LBP) sangat berpengaruh, karena tanpa menggunakan ekstraksi LBP program akan mendeteksi objek aksara jawa yang memiliki sandangan sebagai karakter yang memiliki lebih dari satu objek sehingga hasil deteksinya akan lebih dari satu.
3. Dari 20 data citra uji Aksara Jawa terdapat 7 citra uji menghasilkan transliterasi yang kurang tepat.
4. Hasil pengujian transliterasi Aksara Jawa dengan nilai akurasi terendah sebesar 50% dan nilai akurasi tertinggi sebesar 100%. Sehingga diperoleh rata-rata akurasi sebesar 89,4%.
5. Proses kegagalan dalam transliterasi terjadi karena objek pada citra yang kurang jelas pada saat proses ekstraksi.

6. SARAN

Setelah dilakukan penelitian dan diperoleh kesimpulan, pada penelitian ini masih banyak kekurangan sehingga terdapat beberapa saran untuk penelitian selanjutnya:

1. Proses transliterasi untuk lebih dikembangkan dengan menambahkan variasi sandangan aksara jawa taling, layer, wignyan, cecak, pangkon, pengkal, cakra, dan keret agar dapat melakukan transliterasi lebih baik lagi.
2. Objek penelitian aksara jawa lebih dikembangkan sehingga data uji tidak hanya didapatkan dari citra Ms.Word tetapi bisa dari foto dan naskah jawa.
3. Menggunakan metode yang lain untuk transliterasi.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada **Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) Universitas Dian Nuswantoro** yang telah memberi **“dukungan financial”** terhadap penelitian ini melalui Hibah Penelitian Terapan Perguruan Tinggi (PTPT) Tahun Anggaran 2022.

DAFTAR PUSTAKA

- [1] I. Prihandi, I. Ranggadara, S. Dwiasnati, Y. S. Sari, and Suhendra, “Implementation of Backpropagation Method for Identified Javanese Scripts,” *J. Phys. Conf. Ser.*, vol. 1477, no. 3, 2020.
- [2] M. A. Rasyidi, T. Bariyah, Y. I. Riskajaya, and A. D. Septyani, “Classification of handwritten javanese script using random forest algorithm,” *Bull. Electr. Eng. Informatics*, vol. 10, no. 3, pp. 1308–1315, 2021.
- [3] A. Setiawan, A. S. Prabowo, and E. Y. Puspaningrum, “Handwriting Character Recognition Javanese Letters Based on Artificial Neural Network,” *Int. J. Comput. Netw. Secur. Inf. Syst.*, no. September, pp. 39–42, 2019.
- [4] A. Susanto, D. Sinaga, C. A. Sari, E. H. Rachmawanto, and D. R. I. M. Setiadi, “A High Performace of Local Binary Pattern on Classify Javanese Character Classification,” *Sci. J. Informatics*, vol. 5, no. 1, p. 8, 2018.
- [5] C. A. Sari, M. W. Kuncoro, D. R. I. M. Setiadi, and E. H. Rachmawanto, “Roundness and eccentricity feature extraction for Javanese handwritten character recognition based on K-nearest neighbor,” *2018 Int. Semin. Res. Inf. Technol. Intell. Syst. ISRITI 2018*, pp. 5–10, 2018.
- [6] G. H. Wibowo, R. Sigit, and A. Barakbah, “Javanese Character Feature Extraction Based on Shape Energy,” *Emit. Int. J. Eng. Technol.*, vol. 5, no. 1, pp. 154–169, Jul. 2017.
- [7] I. F. Katili, F. D. Esabella, and A. Luthfiarta, “Pattern Recognition Of Javanese Letter Using Template Matching Correlation Method,” *J. Appl. Intell. Syst.*, vol. 3, no. 2, pp. 49–56, Dec. 2018.
- [8] Rismiyati, Khadijah, and A. Nurhadiyah, “Deep learning for handwritten Javanese character recognition,” in *2017 1st International Conference on Informatics and Computational Sciences (ICICoS)*, 2017, pp. 59–64.
- [9] Y. Sugianela and N. Suciati, “Javanese Document Image Recognition Using Multiclass Support Vector Machine,” *CommIT (Communication Inf. Technol. J.*, vol. 13, no. 1, p. 25, May 2019.
- [10] A. N. Handayani, H. W. Herwanto, K. L. Chandrika, and K. Arai, “Recognition of Handwritten Javanese Script using Backpropagation with Zoning Feature Extraction,” *Knowl. Eng. Data Sci.*, vol. 4, no. 2, p. 117, Dec. 2021.
- [11] Y. Harjoseputro, “Convolutional Neural Network (Cnn) Untuk Pengklasifikasian Aksara Jawa,” *Buana Inform.*, p. 23, 2018.

- [12] L. D. Krisnawati and A. W. Mahastama, "Building classifier models for on-off javanese character recognition," in *ACM International Conference Proceeding Series*, 2019.
- [13] G. A. Robby, A. Tandra, I. Susanto, J. Harefa, and A. Chowanda, "Implementation of Optical Character Recognition using Tesseract with the Javanese Script Target in Android Application," *Procedia Comput. Sci.*, vol. 157, pp. 499–505, 2019.
- [14] F. T. Anggraeny, E. P. Mandyartha, and D. S. Y. Kartika, "Texture Feature Local Binary Pattern for Handwritten Character Recognition," in *2020 6th Information Technology International Seminar (ITIS)*, 2020, pp. 125–129.
- [15] M. Nasir, N. Suciati, and A. Y. Wijaya, "Kombinasi Fitur Tekstur Local Binary Pattern yang Invariant Terhadap Rotasi dengan Fitur Warna Berbasis Ruang Warna HSV untuk Temu Kembali Citra Kain Tradisional," *Inspir. J. Teknol. Inf. dan Komun.*, vol. 7, no. 1, 2017.
- [16] N. D. A. Partiningsih, R. R. Fratama, C. A. Sari, D. R. I. M. Setiadi, and E. H. Rachmawanto, "Handwriting Ownership Recognition using Contrast Enhancement and LBP Feature Extraction based on KNN," in *2018 5th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, 2018, pp. 342–346.
- [17] I. Kusumadewi, C. A. Sari, D. R. I. Moses Setiadi, and E. H. Rachmawanto, "License Number Plate Recognition using Template Matching and Bounding Box Method," in *Journal of Physics: Conference Series*, 2019, vol. 1201, no. 1.
- [18] Y. Septiana, A. Mulyani, D. Kurniadi, and H. Hasanudin, "Handwritten recognition of Hiragana and Katakana characters based on template matching algorithm," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1098, no. 3, p. 032093, Mar. 2021.
- [19] X. Jun-bo, "Template matching algorithm based on gradient search," in *2014 International Conference on Mechatronics and Control (ICMC)*, 2014, no. Icmc, pp. 1472–1475.
- [20] I. Gogul and V. S. Kumar, "Flower species recognition system using convolution neural networks and transfer learning," *2017 4th Int. Conf. Signal Process. Commun. Networking, ICSCN 2017*, pp. 1–6, 2017.
- [21] M. H. Shakoor and R. Boostani, "A novel advanced local binary pattern for image-based coral reef classification," *Multimed. Tools Appl.*, vol. 77, no. 2, pp. 2561–2591, Jan. 2018.
- [22] K. Ayuningsih, Y. A. Sari, and P. P. Adikara, "Klasifikasi Citra Makanan Menggunakan HSV Color Moment dan Local Binary Pattern dengan Naïve Bayes Classifier," in *Proceeding of the 2019 International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, 2019, pp. 3166–3173.