

Pengenalan Sekuens Kata Aksara Jawa Tanpa Segmentasi Menggunakan Arsitektur CRNN dan CTC

^{1*}Moch. Saefudin Yuhri, ²Ahmad Bagus Setiawan, ³Intan Nur Faridha

^{1,2,3} Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: ¹_saefudinyuhri123@gmail.com, ²ahmadbagus@unpkediri.ac.id, ³intannf@unpkediri.ac.id

Penulis Korespondens : Moch. Saefudin Yuhri

Abstrak— Terdapat penurunan kemampuan masyarakat dalam membaca Aksara Jawa, yang sebagian besar disebabkan oleh kompleksitas aturan penulisan *Sandhangan* dan *Pasangan* yang posisinya dapat bertumpuk secara vertikal. Metode *Optical Character Recognition* (OCR) konvensional yang berbasis segmentasi karakter sering mengalami kegagalan karena pemotongan karakter justru merusak informasi visual dari aksara yang bertumpuk. Penelitian ini mengusulkan pendekatan *segmentation-free* menggunakan arsitektur *Convolutional Recurrent Neural Network* (CRNN) yang dipadukan dengan *Connectionist Temporal Classification* (CTC) Loss. Model dilatih menggunakan lebih dari 15.000 citra *dataset* sintesis kata Aksara Jawa yang telah melalui proses augmentasi. Hasil pengujian dievaluasi menggunakan metrik *Character Error Rate* (CER) dan *Word Error Rate* (WER) berdasarkan jarak Levenshtein. Penelitian ini membuktikan bahwa pendekatan pengenalan sekuens kata secara utuh mampu menerjemahkan Aksara Jawa dengan tingkat akurasi karakter mencapai 99,71%, sehingga dapat menjadi solusi digitalisasi aksara tradisional yang efektif.

Kata Kunci— Aksara Jawa, CRNN, CTC, Deep Learning, Transliterasi

Abstract— *There is a decline in the public's ability to read Javanese script, largely due to the complexity of the writing rules for Sandhangan and Pasangan, which can be stacked vertically. Conventional Optical Character Recognition (OCR) methods based on character segmentation often fail because cutting characters destroys the visual information of the stacked scripts. This study proposes a segmentation-free approach using a Convolutional Recurrent Neural Network (CRNN) architecture combined with a Connectionist Temporal Classification (CTC) Loss. The model is trained using over 15,000 synthetic dataset images of Javanese script words that have undergone augmentation. The test results are evaluated using the Character Error Rate (CER) and Word Error Rate (WER) metrics based on the Levenshtein distance. This research proves that the entire word sequence recognition approach can translate Javanese script with a character accuracy rate of up to 99.71%, making it an effective solution for the digitalization of traditional scripts.*

Keywords— CRNN, CTC, Deep Learning, Javanese Script, Transliteration

This is an open access article under the CC BY-SA License.



I. PENDAHULUAN

Aksara Jawa merupakan salah satu warisan budaya Nusantara yang sarat akan nilai historis dan filosofis. Namun, seiring perkembangan zaman, terjadi penurunan kemampuan literasi masyarakat dalam membaca maupun menulis Aksara Jawa. Kendala utama dalam digitalisasi dan pembelajaran Aksara Jawa terletak pada kompleksitas struktur penulisannya, di mana penerapan *Sandhangan* (tanda baca vokal/konsonan) dan *Pasangan* (penekan vokal) sering kali ditulis secara bertumpuk (vertikal) dengan karakter utamanya (*Nglegena*) [1].

Upaya digitalisasi Aksara Jawa telah banyak dilakukan melalui pendekatan *Optical Character Recognition* (OCR) tradisional [2]. Namun, OCR konvensional umumnya mengandalkan teknik segmentasi, yaitu memotong citra kata menjadi karakter tunggal sebelum dikenali oleh sistem. Pada kasus Aksara Jawa, proses pemotongan ini sering kali merusak struktur visual aksara yang bertumpuk, sehingga menghasilkan tingkat kesalahan transliterasi yang tinggi [3].

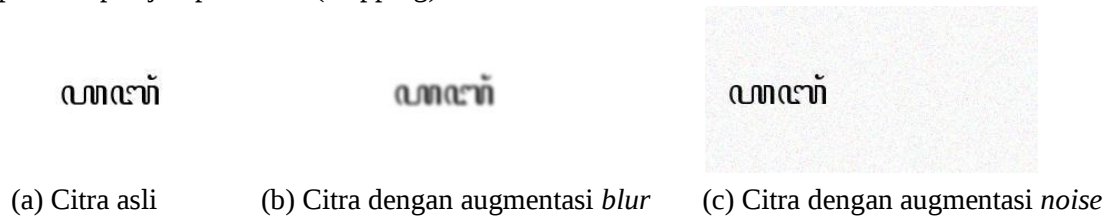
Untuk mengatasi batasan tersebut, penelitian ini mengusulkan penerapan metode *segmentation-free* (tanpa segmentasi karakter) menggunakan arsitektur *Convolutional Recurrent Neural Network* (CRNN) yang dipadukan dengan *Connectionist Temporal Classification* (CTC) [4]. Pendekatan ini memungkinkan sistem untuk memproses citra kata secara utuh sebagai sebuah urutan (sekuens) temporal. Ekstraksi fitur visual dilakukan oleh lapisan *Convolutional Neural Network* (CNN), analisis konteks sekuensial dilakukan oleh *Bidirectional LSTM*, dan *decoding* ke dalam bentuk teks difasilitasi oleh CTC. Penelitian ini bertujuan untuk membuktikan efektivitas CRNN dan CTC dalam meminimalisir kesalahan pembacaan akibat tumpukan aksara tanpa memerlukan algoritma pemisahan huruf yang rumit.

II. METODE

Penelitian ini menggunakan pendekatan eksperimental komputasi dalam bidang kecerdasan buatan (*Artificial Intelligence*) untuk merancang dan membangun model transliterasi Aksara Jawa. Subjek utama yang diteliti adalah kumpulan citra digital berformat JPG yang memuat sekuens teks Aksara Jawa, meliputi variasi huruf dasar (*Nglegena*), tanda baca (*Sandhangan*), dan penekan vokal (*Pasangan*). Perangkat lunak yang digunakan dalam eksperimen ini meliputi bahasa pemrograman Python beserta kerangka kerja *Deep Learning TensorFlow/Keras* dan pustaka pengolahan citra *OpenCV*, yang dioperasikan pada perangkat keras berspesifikasi CPU AMD Ryzen 5 dan RAM 8 GB. Desain eksperimen dirancang secara terstruktur meliputi serangkaian tahap: pembangkitan data sintetis (*data generation*), pra-pemrosesan citra (*preprocessing*), pelatihan model (*training*), dan evaluasi (*testing*). Data diambil dan dihasilkan secara komputasi untuk membentuk korpus berukuran 15.000 sampel yang kemudian didistribusikan secara acak menjadi data latih (80%), data validasi (10%), dan data uji (10%). Pemodelan analisis menggunakan arsitektur *Convolutional Recurrent Neural Network* (CRNN) yang dipadukan dengan *Connectionist Temporal Classification* (CTC) Loss [4]. Adapun variabel dependen yang diukur untuk mengevaluasi kinerja model adalah tingkat kesalahan pembacaan yang dikalkulasi menggunakan persamaan matematis *Character Error Rate* (CER) dan *Word Error Rate* (WER) [8].

A. Pengumpulan dan Pemrosesan Data

Karena keterbatasan *dataset* publik untuk kata Aksara Jawa beranotasi lengkap, penelitian ini menggunakan pembangkitan data sintetis (*synthetic data generation*). *Dataset* terdiri dari 15.000 citra kata Aksara Jawa yang diekstrak dari font cetak. Untuk meningkatkan ketahanan (*robustness*) model, dilakukan teknik augmentasi citra berupa variasi ketajaman (*blur*), penambahan *noise*, serta penyesuaian rotasi [5]. Seluruh citra input diubah menjadi skala abu-abu (*grayscale*) dan distandardisasi ke dimensi ukuran tetap, yaitu 128x32 piksel, sebelum diumpankan ke jaringan saraf. Setiap citra kata Aksara Jawa dipasangkan dengan label teks berhuruf Latin yang merepresentasikan hasil transliterasinya secara sekuensial, sehingga model dapat mempelajari pemetaan (mapping) fitur visual ke teks secara *end-to-end*.



Gambar 1. Contoh variasi citra *dataset* sintetis Aksara Jawa berukuran 128x32 piksel.

B. Arsitektur Model CRNN

Model CRNN yang diusulkan terdiri dari dua komponen utama. Komponen pertama adalah ekstraktor fitur berbasis CNN. Citra input berukuran (128x32x1) diproses melalui serangkaian lapisan Conv2D dan MaxPooling2D guna menghasilkan representasi peta fitur (*feature maps*). Komponen kedua adalah pemodelan sekuens, di mana peta fitur diubah orientasinya (melalui fungsi *Reshape* dan *Permute*) menjadi urutan waktu (*time-steps*) dan diumpankan ke dalam dua lapisan *Bidirectional LSTM* (Bi-LSTM) dengan *dropout* sebesar 0.25 untuk mencegah *overfitting* [6].

C. Connectionist Temporal Classification (CTC)

Keluaran dari Bi-LSTM berupa matriks probabilitas (*logits*) yang kemudian diproses menggunakan fungsi *CTC Loss* pada tahap pelatihan. CTC berperan krusial dalam menyelaraskan prediksi urutan karakter dengan label teks asli tanpa memerlukan informasi kesejajaran (*alignment*) pada tingkat piksel secara eksplisit. Metode ini secara otomatis menangani pengulangan karakter dan spasi kosong (*blank index*) [7].

D. Evaluasi Kinerja

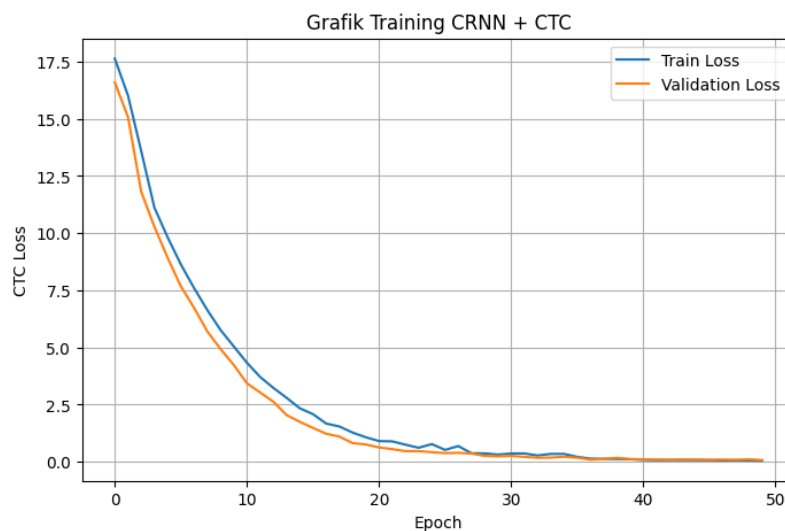
Pengujian keandalan model dievaluasi menggunakan metode jarak Levenshtein (*Levenshtein distance*) yang menghitung jumlah minimum operasi penyisipan (*insertion*), penghapusan (*deletion*), atau penggantian (*substitution*) karakter agar teks prediksi identik dengan label asli [8]. Evaluasi diterapkan secara kuantitatif melalui kalkulasi metrik *Character Error Rate* (CER) dan *Word Error Rate* (WER).

III. HASIL DAN PEMBAHASAN

Data yang disajikan pada bagian ini diproduksi melalui proses inferensi model *Convolutional Recurrent Neural Network* (CRNN) terhadap 1.500 citra *dataset* uji sintesis yang belum pernah dipelajari sebelumnya oleh sistem. Pengujian diekstraksi dari matriks probabilitas (*logits*) menggunakan algoritma *Greedy Decoding* untuk melihat akurasi transliterasi Aksara Jawa ke huruf Latin secara *segmentation-free* (tanpa pemotongan karakter).

A. Hasil Pelatihan Model

Pelatihan model dilakukan menggunakan *optimizer* Adam dengan *learning rate* 0.001. Hasil dari proses percobaan ini disajikan secara kuantitatif dalam bentuk grafik penurunan nilai *Connectionist Temporal Classification* (CTC) *loss* yang dapat dilihat pada Gambar 2.



Gambar 2. Grafik Penurunan CTC Loss pada Data Latih dan Validasi

Berdasarkan visualisasi pada Gambar 2 *dataset*, model tidak menunjukkan adanya gejala *overfitting*. Algoritma mencapai titik konvergensi yang optimal pada *epoch* ke-50, mencatatkan nilai akhir *training loss* sebesar 0.0454 dan *validation loss* sebesar 0.0717.

B. Hasil Evaluasi Pengujian dan Pembahasan Teoritis

Hasil ringkasan pengujian pada data uji untuk mengukur tingkat kesalahan prediksi ditampilkan melalui parameter jarak Levenshtein. Hasil inferensi transliterasi memperoleh nilai rata-rata *Character Error Rate* (CER) sebesar 0.0029 dan *Word Error Rate* (WER) sebesar 0.0167. Melalui kalkulasi tersebut, sistem mencapai tingkat akurasi pengenalan karakter sebesar 99,71%.

Tabel 1. Contoh Hasil Prediksi Transliterasi Model pada Data Uji

No.	Citra Input	Nama File	Label Asli	Prediksi	CER	WER
1.		abot_Nawatura_beta_rotk iri.jpg	abot	abot	0.0	0.0
2.		cangkem_nyk_Ngayogya n_Jejeg_rotkanan.jpg	cangkem	cangkem	0.0	0.0
3.		pang_NotoSansJavanese- VariableFont_wght_rotka nan.jpg	pang	pang	0.0	0.0
4.		mireng_Nawatura_beta_a ndroid2_asli.jpg	mireng	mireng	0.0	0.0
5.		numpak_nyk_Ngayogyan _Jegeg_blur.jpg	numpak	numpak	0.0	0.0
6.		sugih_NotoSansJavanese- VariableFont_wght_asli.j pg	sugih	sugih	0.0	0.0
7.		gumuyu_TuladhaJegeg_as li.jpg	gumuyu	gumuyu	0.0	0.0
8.		nesu_nyk_Ngayogyan_Jej eg_asli.jpg	nesu	nesu	0.0	0.0
9.		dalam_nyk_Ngayogyan_Je jeg_noise.jpg	dalam	dalam	0.0	0.0
10.		arit_Nawatura_beta_andr oid2_blur.jpg	arit	arit	0.0	0.0

Tingkat akurasi yang nyaris sempurna ini membuktikan hipotesis awal bahwa penerapan arsitektur *Bidirectional LSTM* mampu menangkap konteks spasial dari *Sandhangan* dan *Pasangan* Aksara Jawa secara utuh dari dua arah. Lebih lanjut, temuan ini memberikan argumen yang mematahkan kelemahan metode *Optical Character Recognition (OCR)* konvensional berbasis segmentasi yang dilaporkan oleh peneliti sebelumnya [9].

Pemotongan karakter eksplisit pada OCR tradisional terbukti merusak informasi visual dari aksara yang bertumpuk secara vertikal, sementara fungsi CTC pada penelitian ini berhasil menyelaraskan (*alignment*) fitur visual ke sekuens teks tanpa merusak topologi aksara asli [4]. Implikasi dari keberhasilan pengenalan sekuens kata secara utuh ini memungkinkan algoritma diterapkan langsung pada perangkat lunak berskala produksi (*production-scale*), seperti sistem transliterasi dokumen atau aplikasi penerjemah waktu nyata (*real-time*). Kendati demikian, batasan pada penelitian saat ini adalah model hanya dilatih menggunakan citra yang diekstrak dari *font* cetak digital (*computer printed*). Tantangan yang perlu dieksplorasi pada penelitian selanjutnya adalah penerapan arsitektur CRNN dan CTC untuk mengenali teks pada naskah kuno, manuskrip, atau Aksara Jawa berupa tulisan tangan (*handwritten*), yang memiliki variasi dan distorsi goresan yang jauh lebih tidak teratur dan kompleks [10].

IV. KESIMPULAN

Pendekatan *segmentation-free* menggunakan arsitektur *Convolutional Recurrent Neural Network* (CRNN) dan *Connectionist Temporal Classification* (CTC) terbukti berhasil memecahkan permasalahan transliterasi Aksara Jawa, khususnya pada kegagalan pengenalan struktur aksara bertumpuk seperti *Sandhangan* dan *Pasangan*. Penelitian ini memberikan peningkatan yang signifikan pada bidang sains komputer dan pengenalan pola (*pattern recognition*) dengan mengeliminasi proses pemotongan karakter secara eksplisit yang selama ini menjadi kelemahan utama pada sistem *Optical Character Recognition* (OCR) konvensional. Keberhasilan algoritma dalam menyelaraskan fitur visual menjadi sekuens teks secara utuh dibuktikan dengan tingkat kesalahan prediksi (CER dan WER) yang sangat minimal. Sebagai saran untuk pengembangan ilmu pengetahuan lebih lanjut, arsitektur ini perlu dievaluasi dan dioptimasi menggunakan korpus data berupa manuskrip kuno atau tulisan tangan (*handwritten*) asli, sehingga keandalan sistem dapat diimplementasikan sepenuhnya pada upaya pelestarian literatur budaya di skenario dunia nyata.

DAFTAR PUSTAKA

- [1] E. K. Aribowo, "DIGITALISASI AKSARA JAWA DAN PEMANFAATANNYA SEBAGAI MEDIA PEMBELAJARAN BAGI MUSYAWARAH GURU MATA PELAJARAN BAHASA JAWA SMP KABUPATEN KLATEN," *Warta LPM*, vol. 21, no. 2, pp. 59–70, Jun. 2018, doi: 10.23917/warta.v21i2.5620.
- [2] Erviana, "Pemanfaatan Teknologi Optical Character Recognition (Untuk Mengenali Alfabet Aksara Jawa Berbasis Android Dengan Algoritma Convolutional Neural Network)," UNIVERSITAS BUDDHI DHARMA TANGERANG, 2023. [Online]. Available: <https://repositori.buddhidharma.ac.id/1988/2/COVER-BAB3.pdf>
- [3] Y. Sugianela and N. Suciati, "CHARACTER IMAGE SEGMENTATION OF JAVANESE SCRIPT USING CONNECTED COMPONENT METHOD," *Jurnal Ilmu Komputer dan Informasi*, vol. 12, no. 2, pp. 67–74, Jul. 2019, doi: 10.21609/jiki.v12i2.677.

- [4] A. Akoushideh, A. Ranjkesh Rashtehroudi, and A. Shahbahrami, "Persian/Arabic Scene Text Recognition With Convolutional Recurrent Neural Network," *IET Smart Cities*, vol. 7, no. 1, Jan. 2025, doi: 10.1049/smc2.70001.
- [5] M. Nagaraju, P. Chawla, and N. Kumar, "Performance improvement of Deep Learning Models using image augmentation techniques," *Multimed. Tools Appl.*, vol. 81, no. 7, pp. 9177–9200, Mar. 2022, doi: 10.1007/s11042-021-11869-x.
- [6] R. Siringoringo, J. Jamaluddin, R. Perangin-angin, E. J. G. Harianja, G. Lumbantoruan, and E. N. Purba, "MODEL BIDIRECTIONAL LSTM UNTUK PEMROSESAN SEKUENSIAL DATA TEKS SPAM," *METHOMIKA Jurnal Manajemen Informatika dan Komputerisasi Akuntansi*, vol. 7, no. 2, pp. 265–271, Oct. 2023, doi: 10.46880/jmika.Vol7No2.pp265-271.
- [7] Z. Zhang *et al.*, "Self-Distillation Regularized Connectionist Temporal Classification Loss for Text Recognition: A Simple Yet Effective Approach," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 7, pp. 7441–7449, Mar. 2024, doi: 10.1609/aaai.v38i7.28575.
- [8] R. P. Hati, "EVALUASI KINERJA SPEECH-TO-TEXT DENGAN REDUKSI KEBISINGAN SPECTRAL GATING DAN WIENER FILTERING PADA AUDIO BAHASA LAMPUNG," UNIVERSITAS LAMPUNG, 2025. Accessed: Jun. 04, 2026. [Online]. Available: <http://digilib.unila.ac.id/id/eprint/88775>
- [9] Muhammad Nashiruddin, F. D. Noor Praditya, Agiel Faiz Mufazzal, and A. Ardiansyah, "Implementation of Tesseract-based OCR for UMKLA student card data extraction," *JKTI Jurnal Keilmuan Teknologi Informasi*, vol. 1, no. 1, pp. 11–14, Jun. 2025, doi: 10.61902/jkti.v1i1.1688.
- [10] L. Abdiansah, S. Sumarno, A. Eviyanti, and N. L. Azizah, "Penerapan Algoritma Convolutional Neural Networks untuk Pengenalan Tulisan Tangan Aksara Jawa," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 2, pp. 496–504, Mar. 2025, doi: 10.57152/malcom.v5i2.1814.