TRANSLITERASI AKSARA *JAWOE*-LATIN MENGGUNAKAN ARSITEKTUR *CONVOLUTIONAL RECURRENT NEURAL NETWORK*

TUGAS AKHIR

Diajukan Oleh :

SITI MARIATUL ULFA
NIM. 220705017

Mahasiswa Fakultas Sains dan Teknologi
Program Studi Teknologi Informasi

FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI AR-RANIRY
BANDA ACEH
2026 M/1448 H

LEMBAR PERSETUJUAN

TRANSLITERASI AKSARA *JAWOE*-LATIN MENGGUNAKAN ARSITEKTUR *CONVOLUTIONAL RECCURRENT NEURAL NETWORK*

Tugas Akhir

Diajukan kepada Fakultas Sains dan Teknologi
Universitas Islam Negeri (UIN) Ar-Raniry Banda Aceh
Sebagai Salah Satu Beban Studi Memperoleh Gelar Sarjana (S1)
Dalam Ilmu Teknologi Informasi

Oleh:

Siti Mariatul Ulfa
NIM. 220705017

Mahasiswa Fakultas Sains dan Teknologi
Program Studi Teknologi Informasi

Disetujui untuk Dimunaqasyahkan Oleh:

Pembimbing I,	Pembimbing II,
Khairan AR, M.Kom	**Dr. Hendri Ahmadian, S.Si., M.I.M**
NIP.198607042014031001	**NIP.198301042014031002**

Mengetahui,
Ketua Program Studi Teknologi Informasi

Malahayati, M.T
NIP. 198301272015032003

LEMBAR PENGESAHAN

TRANSLITERASI AKSARA *JAWOE*-LATIN MENGGUNAKAN ARSITEKTUR *CONVOLUTIONAL RECURRENT NEURAL NETWORK*

TUGAS AKHIR

Telah Diuji oleh Panitia Ujian Munaqasyah Tugas Akhir
Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh dan Dinyatakan Lulus
Serta Diterima Sebagai Salah Satu Beban Studi Program Sarjana (S1)
Dalam Program Studi Teknologi Informasi

Pada Hari/Tanggal: Selasa, 18 Agustus 2026 M
4 Rabiul Awal 1448 H
Di Darussalam, Banda Aceh

Panitia Ujian Munaqasyah Tugas Akhir:

Ketua,	Sekretaris,
Khairan AR, M.Kom.	**Dr. Hendri Ahmadian, M.I.M.**
NIP. 198607042014031001	**NIP. 198301042014031002**
Penguji I,	Penguji II,
Ridha Ilahi, M.T.	**Malahayati, M.T.**
NIP. 197905302014031001	**NIP. 198301272015032003**

Mengetahui:
Dekan Fakultas Sains dan Teknologi
UIN Ar-Raniry Banda Aceh

Prof. Dr. Ir. Muhammad Dirhamsyah, M.T., IPU
NIP. 196210021988111001

LEMBAR PERNYATAAN KEASLIAN TUGAS AKHIR

Yang bertanda tangan di bawah ini:

Nama	:	Siti Mariatul Ulfa
NIM	:	220705017
Program Studi	:	Teknologi Informasi
Fakultas	:	Sains dan Teknologi
Judul	:	Transliterasi Aksara Jawoe-Latin Menggunakan Arsitektur Convolutional Recurrent Neural Network

Dengan ini menyatakan bahwa dalam penulisan tugas akhir ini, saya:

1. Tidak menggunakan ide orang lain tanpa mampu mengembangkan dan mempertanggungjawabkan;
2. Tidak melakukan plagiasi terhadap naskah karya orang lain;
3. Tidak menggunakan karya orang lain tanpa menyebutkan sumber asli atau tanpa izin pemilik karya;
4. Tidak memanipulasi dan memalsukan data;
5. Mengerjakan sendiri karya ini dan mampu bertanggungjawab atas karya ini.

Bila dikemudian hari ada tuntutan dari pihak lain atas karya saya, dan telah melalui pembuktian yang dapat dipertanggungjawabkan dan ternyata memang ditemukan bukti bahwa saya telah melanggar pernyataan ini, maka saya siap dikenai sanksi berdasarkan aturan yang berlaku di Fakultas Sains dan Teknologi UIN Ar-Raniry Banda Aceh. Demikian pernyataan ini saya buat dengan sesungguhnya dan tanpa paksaan dari pihak manapun.

Banda Aceh, 19 Agustus 2026

Yang Menyatakan,

Siti Mariatul Ulfa

ABSTRAK

Penelitian ini bertujuan mengembangkan sistem transliterasi otomatis Aksara Jawoe ke alfabet Latin guna mendukung pelestarian manuskrip kuno Aceh yang kian rentan menjadi teks mati akibat minimnya kemampuan literasi pembacanya. Tantangan utama pada transliterasi aksara ini terletak pada karakteristik tulisannya yang bersifat kursif (bersambung), keberadaan ligatur yang kompleks, serta kemiripan visual antar karakter yang tinggi. Untuk mengatasi permasalahan tersebut, diusulkan arsitektur *End-to-End Deep Learning* berupa *Convolutional Recurrent Neural Network* (CRNN) yang mengintegrasikan *Dilated Convolutional Neural Network* (D-CNN) dan *Bidirectional Long Short-Term Memory* (Bi-LSTM). Metode D-CNN berfungsi sebagai ekstraktor fitur visual yang mampu mempertahankan detail halus seperti titik dan harakat tanpa kehilangan resolusi, sementara Bi-LSTM digunakan untuk mempelajari ketergantungan urutan karakter dari dua arah secara simultan. Sistem ini beroperasi secara otomatis tanpa pemotongan batas huruf manual (*segmentation-free*) menggunakan fungsi *Connectionist Temporal Classification* (CTC) yang diuji pada 2.000 citra kata digital berbasis naskah *Hikayat Akhbarul Karim*. Hasil evaluasi performa model menunjukkan tingkat keandalan yang sangat tinggi dengan *Character Error Rate* (CER) sebesar 1,17% dan *Word Error Rate* (WER) sebesar 4,75%. Hal ini membuktikan bahwa integrasi model tersebut sangat tangguh dalam memecahkan ambiguitas visual pada aksara bersambung, sehingga layak diterapkan untuk mempercepat proses digitalisasi warisan sejarah Nusantara secara efisien.

Kata Kunci: Aksara Jawoe, Transliterasi, *Deep Learning*, *Dilated Convolution*, Bi-LSTM, *Character Error Rate*.

ABSTRACT

This study aims to develop an automatic transliteration system converting the Jawoe script to the Latin alphabet to support the preservation of ancient Acehnese manuscripts, which are increasingly at risk of becoming inaccessible "dead texts" due to declining literacy in the script. The primary challenges in transliterating this script stem from its cursive nature, complex ligatures, and high visual similarity between characters. To address these issues, an end-to-end deep learning architecture—specifically a Convolutional Recurrent Neural Network (CRNN)—is proposed, integrating a Dilated Convolutional Neural Network (D-CNN) with Bidirectional Long Short-Term Memory (Bi-LSTM). The D-CNN component acts as a visual feature extractor capable of preserving fine details, such as dots and diacritics, without loss of resolution, while the Bi-LSTM learns character sequence dependencies simultaneously from both directions. The system operates automatically without manual character segmentation (segmentation-free) using the Connectionist Temporal Classification (CTC) function and was tested on 2,000 digital word images derived from the *Hikayat Akhbarul Karim* manuscript. Performance evaluation demonstrates high reliability, achieving a Character Error Rate (CER) of 1.17% and a Word Error Rate (WER) of 4.75%. These results confirm the model's robustness in resolving visual ambiguities inherent in cursive scripts, making it a viable solution for efficiently accelerating the digitization of the archipelago's historical heritage.

Keywords: Jawoe script, Transliteration, Deep Learning, Dilated Convolution, Bi-LSTM, Character Error Rate.

KATA PENGANTAR

Bismillahirrahmanirrahim.

Puji dan syukur senantiasa penulis panjatkan kehadirat Allah SWT yang telah melimpahkan rahmat, karunia, serta kemudahan, sehingga penulis dapat menyelesaikan penulisan Tugas Akhir yang berjudul "Analisis Sentimen Opini Publik Terhadap Respons Pemerintah Dalam Penanganan Banjir Di Aceh Menggunakan Model Indobert". Shalawat beserta salam tak lupa pula kita sanjungkan ke pangkuan Nabi Besar Muhammad SAW, yang telah membawa umat manusia dari alam kebodohan ke alam yang penuh dengan ilmu pengetahuan.

Penyusunan Tugas Akhir ini merupakan salah satu syarat untuk menyelesaikan studi jenjang Strata Satu (S1) dan memperoleh gelar Sarjana Komputer pada Program Studi Teknologi Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri Ar-Raniry Banda Aceh

Selama proses penyusunan penelitian ini, penulis menyadari bahwa banyak sekali tantangan dan rintangan yang dihadapi. Namun, berkat bimbingan, arahan, doa, dan dukungan dari berbagai pihak, Tugas Akhir ini akhirnya dapat diselesaikan dengan baik. Oleh karena itu, dengan segala kerendahan hati, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Kedua orang tua tercinta, yang selalu memberikan curahan kasih sayang, doa yang tiada henti, serta pengorbanan material dan spiritual yang luar biasa. Secara khusus, terima kasih kepada Ayah yang senantiasa memfasilitasi dan memotivasi kelancaran penelitian ini. Ucapan terima kasih juga penulis tujukan untuk abang tersayang atas segala bentuk semangat, wejangan, dan dukungannya yang mengantarkan penulis hingga ke titik ini.
2. Bapak Khairan AR, M.Kom., yang telah berkenan menjadi pembimbing tugas akhir penulis, terima kasih atas segala waktu, kesabaran, nasihat, serta arahan yang sangat berharga dalam membimbing penulis sejak awal perkuliahan hingga selesainya Tugas Akhir ini.
3. Bapak Dr. Hendri Ahmadian, M.I.M., yang bertindak sebagai pembimbing tugas akhir penulis yang telah memberikan bimbingan, masukan, arahan, dan dukungan sehingga tugas akhir ini dapat diselesaikan dengan baik.

4. Seluruh Bapak dan Ibu Dosen Program Studi Teknologi Informasi, Fakultas Sains dan Teknologi UIN Ar-Raniry. Terima kasih atas segala ilmu pengetahuan, wawasan, dan pengalaman yang telah diberikan kepada penulis.
5. Selama proses penyusunan tugas akhir, Ibu Cut Ida Rahmadiana, S.Si., seorang staf Program Studi Teknologi Informasi, telah memberikan bantuan dan dukungan dalam berbagai mata kuliah akademik.
6. Sahabat seperjuangan tersayang, Hilda Sabrina Hashar, Munisa, dan Putri Nabila Tama, telah menjadi bagian penting dalam perjalanan perkuliahan penulis. Terima kasih karena selalu hadir dalam suka maupun duka, menjadi tempat berbagi cerita, keluh kesah, tawa, dan semangat di tengah proses yang tidak mudah ini. Terima kasih telah berjalan bersama hingga sampai di titik ini. Semoga persahabatan yang terjalin tida hanya berhenti di bangku perkuliahan, tetapi terus tumbuh dalam setiap langkah kehidupan ke depan.
7. Teman-teman dari Program Studi Teknologi Informasi angkatan 2022 serta teman-teman dekat yang telah memberikan semangat, doa, dukungan, bantuan, dan kerja sama kepada penulis selama masa studi hingga penyusunan tugas ini.
8. Terakhir, penulis ingin menyampaikan rasa terima kasih yang paling tulus kepada diri penulis sendiri, Siti Mariatul Ulfa. Terima kasih karena telah mampu bertahan, berjuang, dan tidak memilih menyerah di setiap proses yang telah dilalui. Menjadi anak perempuan yang membawa harapan orang tua bukanlah hal yang mudah. Ada banyak air mata yang tidak terlihat, rasa lelah yang dipendam sendiri, serta perjuangan yang mungkin tidak pernah diketahui oleh orang lain. Namun, semua itu menjadi bagian dari proses yang membentuk penulis menjadi pribadi yang lebih kuat, lebih sabar, dan lebih menghargai setiap langkah kecil yang berhasil dilewati. Terima kasih karena tetap percaya pada diri sendiri, bahkan ketika keadaan terasa begitu berat. Terima kasih karena terus bangkit setiap kali merasa gagal, terus belajar setiap kali melakukan kesalahan, dan terus berjalan meskipun langkah terasa pelan. Skripsi ini bukan hanya tentang memperoleh gelar sarjana, tetapi juga menjadi bukti bahwa setiap usaha, doa, dan kesabaran pada akhirnya akan menemukan jalannya. Semoga perjalanan ini menjadi awal dari langkah-

langkah baik berikutnya. Semoga Allah SWT senantiasa memberikan kesehatan, keberkahan, kemudahan, serta menghadirkan orang-orang baik di setiap perjalanan kehidupan. Jangan pernah berhenti bermimpi, tetap rendah hati, dan terus menjadi pribadi yang bermanfaat bagi keluarga, masyarakat, serta orang-orang di sekitar. Terima kasih telah bertahan sejauh ini. Kamu sudah melakukan yang terbaik.

Penulis dengan jujur mengakui bahwa masih banyak kekurangan dalam tugas akhir ini, baik dari segi penulisan teknis maupun konten. Oleh karena itu, penulis dengan rendah hati menerima kritik dan saran yang bermanfaat untuk pengembangan di masa depan. Terakhir, semoga tugas akhir ini menjadi amal baik yang tercatat sebagai amal baik di mata Allah SWT dan bermanfaat bagi penulis dan pembaca. Rabbal 'Alamin ya Aamiin.

Banda Aceh, 10 Agustus 2026

Penulis,

Siti Mariatul Ulfa

NIM. 220705017

جامعة الرانري
AR-RANIRY

DAFTAR ISI

ABSTRAK ..... **i**

ABSTRACT ..... **ii**

KATA PENGANTAR ..... **iii**

DAFTAR ISI ..... **vi**

DAFTAR GAMBAR ..... **ix**

DAFTAR TABEL ..... **x**

BAB I PENDAHULUAN **1**

1.1 Latar Belakang 1

1.2 Rumusan Masalah 3

1.3 Tujuan penelitian 3

1.4 Batasan masalah 4

1.5 Manfaat Penelitian 5

1.5.1 Manfaat Teoritis 5

1.5.2 Manfaat Praktis 6

BAB II TINJAUAN PUSTAKA **7**

2.1 Penelitian Terdahulu 7

2.2 Landasan Teori / Kajian Teori 10

2.2.1 Manuskrip (Naskah Kuno) 10

2.2.2 Bahasa Aceh dan Aksara Jawoe 11

2.2.3 Convolutional Neural Network (CNN) 11

2.2.4 Dilated Convolution 12

2.2.5 Bidirectional Long Short-Term Memory (Bi-LSTM) 12

2.3 Kerangka Berfikir 12
2.3.1 Alur Kerangka Berpikir 13
2.3.2 Jurnal Pendukung 14
2.4 Hipotesis 15
2.4.1 Hipotesis Nol (H_0): 16
2.4.2 Hipotesis Umum (H_1): 16
2.4.3 Hipotesis Tambahan (H_2): 16
BAB III METODOLOGI PENELITIAN 17
3.1 Jenis Penelitian 17
3.2 Metode Pengumpulan Data 19
3.2.1 Identifikasi Masalah 19
3.2.2 Studi Literatur 19
3.3 Metode Simulasi 20
3.3.1 Perumusan Konsep Penelitian 20
3.3.2 Implementasi Metode Arsitektur 21
3.4 Metode Analisis Data 31
3.4.1 Evaluasi 31
3.4.1.1 Character Error Rate (CER) 31
3.4.1.2 Word Error Rate (WER) 32
3.5 Waktu Penelitian 33
3.6 Instrumen Penelitian 33
3.6.1 Perangkat Keras (Hardware) 34
3.6.2 Perangkat Lunak (Software) 34
BAB IV HASIL DAN PEMBAHASAN 36
4.1 Gambaran Umum Hasil Penelitian 36

4.2 Hasil Implementasi Arsitektur Dilated CNN dan Bi-LSTM 37
4.2.1 Realisasi Lapisan Ekstraksi Fitur (Dilated CNN) 37
4.2.2 Realisasi Lapisan Pemodelan Sekuensial (Bi-LSTM + CTC) 39
4.2.3 Bukti Visual Efektivitas Ekstraksi Fitur 40
4.3 Hasil Performa Model Transliterasi 41
4.3.1 Proses Pelatihan Model 41
4.3.2 Tingkat Akurasi Transliterasi 42
4.3.3 Analisis Kesalahan Transliterasi 43
4.4 Pengujian Hipotesis 44
4.5 Pembahasan 46
4.5.1 Interpretasi Hasil terhadap Tujuan Penelitian 46
4.5.2 Perbandingan dengan Penelitian Terdahulu 47
4.5.3 Keterbatasan Penelitian 48
BAB V PENUTUP 50
5.1 Kesimpulan 50
5.2 Saran 51
DAFTAR PUSTAKA 53

DAFTAR GAMBAR

Gambar 2. 1 Alur Kerangka Berpikir 14
Gambar 3. 1 Desain Penelitian 18
Gambar 3. 2 Dataset Aksara Jawoe 22
Gambar 4. 1 Visualisasi Matriks Probabilitas dan Proses Dekoding CTC 40
Gambar 4. 2 Kurva Konvergensi CTC Loss 42
Gambar 4. 3 Sampel Kegagalan Prediksi 44

جامعة الرانري
AR-RANIRY

DAFTAR TABEL

Tabel 2. 1 Perbandingan penelitian Sejenis 9
Tabel 3. 1 Pembagian Dataset 26
Tabel 3. 2 Arsitektur Model CRNN 28
Tabel 4. 1 Ringkasan Spesifikasi Dataset Aksara Jawoe 37
Tabel 4. 2 Arsitektur Lapisan Ekstraksi Fitur dengan Dilated Convolution 38
Tabel 4. 3 Sampel Keberhasilan Transliterasi pada Kata 40
Tabel 4. 4 Hasil Evaluasi Akhir Performa Transliterasi 43
Tabel 4. 5 Ringkasan Hasil Pengujian Hipotesis 45

جامعة الرانري
AR-RANIRY

BAB I
PENDAHULUAN

1.1 Latar Belakang

Aksara Jawoe merupakan sistem penulisan berbasis abjad Arab yang telah dimodifikasi untuk memenuhi kebutuhan fonetik bahasa Melayu dan Aceh, yang secara historis menjadi media utama dalam penyebaran ilmu pengetahuan, agama, dan sastra di Nusantara. Namun, dalam beberapa dekade terakhir, penggunaan Aksara Jawoe mengalami degradasi signifikan akibat dominasi alfabet Latin dalam sistem pendidikan formal. Fenomena ini menyebabkan ribuan manuskrip kuno di Aceh terancam menjadi "teks mati" karena minimnya generasi muda yang memiliki kompetensi untuk membacanya. Upaya pelestarian melalui digitalisasi naskah saat ini masih terbatas pada pengarsipan citra (gambar), sementara kebutuhan akan transliterasi teks digital yang dapat dicari (*searchable text*) sangat mendesak untuk mempermudah kajian ilmiah dan akses publik (Hidayatullah et al., 2020; Rahman et al., 2025).

Tantangan utama dalam membangun sistem transliterasi otomatis untuk Aksara Jawoe terletak pada karakteristik morfologi tulisannya yang bersifat kursif (bersambung) dan memiliki variasi *ligature* (penggabungan dua karakter atau lebih) yang kompleks. Berbeda dengan karakter Latin yang cenderung diskrit, satu karakter dalam Jawoe dapat berubah bentuk secara drastis tergantung posisinya (di awal, tengah, akhir, atau berdiri sendiri). Selain itu, adanya kemiripan visual yang tinggi antar karakter yang hanya dibedakan oleh jumlah atau posisi titik menuntut model pengenalan yang memiliki ketajaman ekstraksi fitur tingkat tinggi. Metode *Handwritten Text Recognition* (HTR) konvensional yang mengandalkan segmentasi karakter per karakter sering kali gagal total pada teks kursif karena sulitnya menentukan titik potong yang akurat pada tulisan tangan yang rapat (Assidiqi et al., 2022; Razali et al., 2023).

Penelitian terdahulu dalam domain pengenalan karakter Arab dan turunannya sering kali mengandalkan fitur manual (*hand-crafted features*) seperti *Histogram of Oriented Gradients* (HOG) atau *Local Binary Patterns* (LBP) yang dikombinasikan dengan *Support Vector Machine* (SVM). Namun, metode ini terbukti kurang fleksibel dalam menghadapi degradasi citra naskah kuno yang sering mengalami masalah seperti *noise*, bercak tinta (*bleeding*), atau kemiringan baris teks (*skewness*). Sebaliknya, arsitektur *End-to-End* berbasis *Deep Learning* mampu mempelajari representasi fitur secara otomatis langsung dari data mentah. Keunggulan *Deep Learning* dalam konteks ini menjadi krusial karena mampu menangani variasi gaya penulisan tanpa perlu ekstraksi fitur manual yang kaku (Puigcerver, 2017; Moysset et al., 2018).

Dalam lanskap *Deep Learning* modern, arsitektur *Convolutional Neural Network* (CNN) telah menjadi standar emas untuk ekstraksi fitur visual. Namun, pada citra teks yang panjang, CNN standar seringkali memerlukan lapisan *pooling* yang agresif untuk mendapatkan cakupan informasi konteks yang luas, yang sayangnya berisiko menghilangkan detail-detail halus karakter (seperti titik kecil pada huruf 'ba' atau 'ta'). Penggunaan Dilated Convolution (konvolusi dengan lubang) muncul sebagai solusi inovatif yang memungkinkan filter untuk mencakup area yang lebih luas tanpa kehilangan resolusi spasial atau menambah beban parameter secara signifikan. Dengan menerapkan *dilation rate* yang tepat, model dapat menangkap struktur global dari kata-kata Jawoe yang memiliki tarikan garis panjang sambil tetap mempertahankan detail mikro dari setiap lekukan huruf (Yu & Koltun, 2016; Chen et al., 2018).

Permasalahan transliterasi bukan hanya sekadar masalah klasifikasi gambar, melainkan masalah pemodelan sekuensial atau urutan data. Karena teks dibaca sebagai rangkaian makna, informasi dari karakter sebelumnya dan sesudahnya sangat menentukan akurasi prediksi karakter saat ini. Arsitektur *Bidirectional Long Short-Term Memory* (Bi-LSTM) sangat efektif dalam menangani masalah ini karena memiliki kemampuan untuk mengingat informasi jangka panjang dari dua arah secara simultan. Dalam konteks Aksara Jawoe, Bi-LSTM dapat membantu

model memahami probabilitas kemunculan suatu huruf berdasarkan konteks kata, sehingga mampu mengoreksi ambiguitas visual yang mungkin luput dari lapisan CNN (Graves et al., 2013; Shi et al., 2017).

Integrasi antara *Dilated Convolution* sebagai ekstraktor fitur spasial yang efisien dan Bi-LSTM sebagai pemroses sekuensial dua arah diformulasikan ke dalam arsitektur *hybrid* yang dikenal sebagai *Convolutional Recurrent Neural Network* (CRNN). Penggunaan istilah CRNN dalam penelitian ini secara tegas merujuk pada penggabungan kedua model tersebut (CNN dan RNN) yang dilatih secara bersamaan (*end-to-end*). Pendekatan *hybrid* ini diharapkan mampu mengatasi variasi tulisan tangan yang tidak beraturan serta kompleksitas struktur abjad Arab-Melayu. Penelitian ini tidak hanya bertujuan untuk mencapai akurasi teknis yang tinggi, tetapi juga sebagai upaya konkret dalam memitigasi kepunahan literasi aksara tradisional melalui bantuan kecerdasan buatan, sehingga warisan intelektual masa lalu tetap relevan di era digital (Razali et al., 2024; Kurniawan et al., 2023).

1.2 **Rumusan Masalah**

1. Bagaimana cara mengimplementasikan arsitektur *Convolutional Neural Network* dengan teknik *Dilated Convolution* untuk mengekstraksi fitur visual yang efektif pada citra tulisan tangan Aksara Jawoe?
2. Sejauh mana tingkat performa model yang dikembangkan (diukur melalui *Character Error Rate* dan *Word Error Rate*) dalam melakukan transliterasi Aksara Jawoe ke dalam teks Latin secara otomatis?

1.3 **Tujuan penelitian**

1. Mengimplementasikan arsitektur *Deep Learning* yang mengintegrasikan *Dilated Convolutional Neural Network* dan *Bidirectional Long Short-Term Memory* (Bi-LSTM) untuk proses ekstraksi fitur dan pengenalan sekuensial pada citra Aksara Jawoe.

2. Mengevaluasi performa model transliterasi otomatis dengan mengukur tingkat akurasi menggunakan metrik *Character Error Rate* (CER) dan *Word Error Rate* (WER) pada berbagai variasi gaya tulisan tangan Aksara Jawoe.

1.4 **Batasan masalah**

1. Format Input: Penelitian ini dibatasi pada pengenalan citra teks cetakan digital (tipografi). Dataset tidak dipotong langsung dari pindaian manuskrip asli, melainkan menggunakan citra kata hasil salinan ulang (*re-type*) dari manuskrip ke dalam format digital menggunakan variasi *font* komputer.
2. Ruang Lingkup Aksara Aksara yang digunakan adalah varian Arab-Melayu spesifik Aceh (Jawoe) dengan karakteristik *ligature* dan fonetik khas daerah tersebut.
3. Arsitektur Model Model *Deep Learning* yang dibangun dibatasi pada integrasi Dilated CNN sebagai ekstraktor fitur dan Bi-LSTM sebagai pemodel sekuensial, dengan menggunakan *Connectionist Temporal Classification* (CTC) sebagai fungsi *loss*.
4. Proses Segmentasi Sistem menggunakan pendekatan *segmentation-free*, artinya proses transliterasi dilakukan langsung dari potongan citra baris kalimat atau kata tanpa melakukan pemotongan karakter per karakter secara manual.
5. Parameter Evaluasi Keberhasilan model hanya diukur berdasarkan akurasi tekstual menggunakan metrik *Character Error Rate* (CER) dan *Word Error Rate* (WER), tidak mencakup analisis semantik (makna) atau tata bahasa (*grammar*).
6. Ketidakseimbangan Kelas Data (*Class Imbalance*): Mengingat dataset diekstrak secara natural dari naskah kuno, distribusi frekuensi kemunculan kosakata sangat tidak seimbang. Terdapat banyak kelas kata langka yang diproyeksikan hanya memiliki satu atau dua sampel citra. Untuk menghindari kegagalan proses pembagian data (*data splitting*), kelas kata dengan sampel tunggal akan dialokasikan secara eksklusif ke dalam data latih (*training set*) agar model tetap mengenali fitur visualnya, namun kata-

kata langka tersebut tidak akan muncul dan tidak dievaluasi pada data uji (*testing set*). Oleh karena itu, representasi metrik akurasi akhir lebih mencerminkan performa model pada kosa kata umum yang memiliki jumlah sampel memadai.

7. Tanda Vokal dan Karakter Khusus Aceh: Kosakata bahasa Aceh memiliki tanda vokal khusus (seperti é, è, ô, ë, ö). Dalam penelitian ini, karakter-karakter tersebut tidak diikutsertakan sebagai kelas unik, melainkan dipetakan ke dalam huruf vokal Latin standar terdekat (misal: 'ë' dipetakan menjadi 'e', 'ô' menjadi 'o'). Hal ini dilakukan karena dataset awal difokuskan pada pengenalan struktur anatomi dasar Aksara Jawoe, sebelum memasuki tahap analisis fonetik tingkat lanjut.

1.5 **Manfaat Penelitian**

Penelitian ini diharapkan dapat memberikan kontribusi yang signifikan, baik secara keilmuan maupun penerapan langsung di lapangan, sebagai berikut:

1.5.1 Manfaat Teoritis

1. Pengembangan Model Deep Learning: Memberikan kontribusi ilmiah mengenai efektivitas integrasi arsitektur Dilated Convolutional Neural Network (D-CNN) dan Bidirectional LSTM dalam menangani ekstraksi fitur pada aksara kursif (bersambung) yang memiliki kompleksitas tinggi.
2. Referensi Metodologi HTR: Menjadi referensi akademis bagi peneliti selanjutnya dalam bidang Handwritten Text Recognition (HTR), khususnya untuk aksara-aksara rumpun Arab-Melayu seperti Jawi dan Pegon yang memiliki karakteristik serupa dengan Aksara Jawoe.
3. Optimalisasi Fitur Spasial-Temporal: Membuktikan secara empiris bahwa penggunaan Dilated Convolution dapat mempertahankan resolusi detail fitur visual (seperti titik dan harakat) yang sangat krusial dalam akurasi transliterasi karakter Arab-Melayu.

1.5.2 Manfaat Praktis

1. Preservasi Digital Manuskrip Aceh: Membantu upaya penyelamatan aset budaya lokal melalui sistem transliterasi otomatis, sehingga isi dari naskah kuno Aceh dapat dikonversi menjadi format digital yang mudah diakses dan dicari (searchable).
2. Efisiensi Kerja Filolog dan Sejarawan: Menjadi alat bantu bagi para ahli naskah dan akademisi untuk mempercepat proses alih aksara naskah Jawoe ke Latin yang selama ini memakan waktu lama jika dilakukan secara manual.
3. Aksesibilitas Literasi Budaya: Memudahkan masyarakat umum dan generasi muda yang ingin mempelajari kandungan sejarah serta kearifan lokal dalam naskah kuno Aceh namun memiliki keterbatasan kemampuan dalam membaca Aksara Jawoe.

جامعة الرانري
AR-RANIRY

BAB II
TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Penelitian terdahulu merupakan fondasi penting dalam penyusunan penelitian ini untuk memahami perkembangan teknologi *Handwritten Text Recognition* (HTR) serta mengidentifikasi celah penelitian (*research gap*). Berikut adalah ulasan mendalam mengenai beberapa penelitian yang memiliki relevansi tinggi dengan penggunaan *Dilated CNN* dan *Bidirectional LSTM* pada aksara kursif.

Penelitian pertama dilakukan oleh Hidayatullah, dkk. (2020) dalam jurnal yang berjudul "*Jawi Character Recognition Using Deep Learning*". Penelitian ini berfokus pada pengenalan karakter Jawi menggunakan arsitektur *Convolutional Neural Network* (CNN) standar. Dataset yang digunakan terdiri dari ribuan sampel karakter Jawi yang ditulis tangan. Hasil eksperimen menunjukkan bahwa model CNN mampu mencapai akurasi sebesar 95% pada karakter yang berdiri sendiri (*isolated*). Namun, penelitian ini juga mencatat bahwa performa model menurun drastis saat diuji pada teks bersambung karena metode konvolusi standar sulit membedakan batas antar karakter yang saling bertumpuk (ligatur). Hal inilah yang mendasari perlunya pendekatan *segmentation-free* dalam penelitian ini.

Penelitian kedua yang menjadi acuan utama adalah karya Shi, Bai, dan Yao (2017) berjudul "*An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to Scene Text Recognition*". Penelitian ini memperkenalkan arsitektur *Convolutional Recurrent Neural Network* (CRNN) yang menggabungkan CNN untuk ekstraksi fitur dan Bi-LSTM untuk pemodelan urutan. Peneliti menggunakan dataset berskala besar seperti *IIIT 5K-word* dan *Street View Text* (SVT). Kontribusi utama penelitian ini adalah penggunaan fungsi *Connectionist Temporal Classification* (CTC) *Loss* yang memungkinkan jaringan saraf dilatih secara *end-to-end* tanpa perlu segmentasi karakter manual. Hasilnya,

model ini mampu mengenali kata-kata dalam citra dengan akurasi yang sangat tinggi meskipun teks tersebut memiliki panjang yang bervariasi.

Penelitian ketiga dilakukan oleh Yu dan Koltun (2016) dengan judul "*Multi-Scale Context Aggregation by Dilated Convolutions*". Penelitian ini memberikan landasan teknis mengenai penggunaan *Dilated Convolution* sebagai alternatif dari konvolusi standar. Peneliti membuktikan bahwa dengan menerapkan *dilation rate* (jarak antar elemen filter), *receptive field* atau bidang pandang model dapat diperluas secara eksponensial tanpa kehilangan resolusi spasial akibat operasi *pooling*. Hasil pengujian pada dataset *Cityscapes* menunjukkan bahwa teknik ini sangat efektif dalam mempertahankan detail-detail kecil pada objek. Dalam konteks Aksara Jawoe, teknik ini sangat krusial untuk menjaga agar fitur titik yang sangat kecil (seperti pada huruf *ba*, *ta*, atau *tsa*) tidak hilang saat proses ekstraksi fitur berlangsung.

Penelitian keempat dilakukan oleh Puigcerver (2017) dalam makalah berjudul "*Are Multidimensional Recurrent Neural Networks Really Necessary for Handwritten Text Recognition?*". Peneliti melakukan eksperimen mendalam menggunakan arsitektur CNN yang diikuti oleh lapisan *Bidirectional LSTM* (Bi-LSTM) pada dataset naskah kuno seperti *IAM* dan *Saint Gall*. Penelitian ini membuktikan bahwa penggunaan Bi-LSTM satu dimensi sudah cukup untuk menghasilkan akurasi yang kompetitif dibandingkan model dua dimensi yang jauh lebih berat secara komputasi. Peneliti melaporkan bahwa penggunaan Bi-LSTM mampu menurunkan *Character Error Rate* (CER) hingga di bawah 10% pada data tulisan tangan yang sangat tidak beraturan. Hal ini memperkuat alasan penggunaan Bi-LSTM dalam penelitian transliterasi Aksara Jawoe.

Penelitian kelima oleh Assidiqi, dkk. (2022) berjudul "*Challenges in Arabic-Malay (Jawi) Character Recognition: A Systematic Literature Review*". Penelitian ini melakukan tinjauan sistematis terhadap berbagai metode pengenalan aksara rumpun Arab. Peneliti menyimpulkan bahwa tantangan terbesar dalam transliterasi Jawi atau Jawoe adalah variasi bentuk huruf berdasarkan posisi (awal, tengah, dan akhir) yang sangat dinamis. Penelitian tersebut merekomendasikan penggunaan

arsitektur yang memiliki memori jangka panjang (seperti LSTM) untuk memahami konteks kata secara utuh. Selain itu, ditekankan pula pentingnya tahap *preprocessing* seperti *binarization* dan *deslanting* untuk meningkatkan kualitas citra masukan sebelum masuk ke tahap pelatihan model *deep learning*.

Tabel 2. 1 Perbandingan penelitian Sejenis

Peneliti (Tahun)	**Judul Penelitian**	**Metode / Arsitektur**	**Hasil & Kontribusi Utama**
Hidayatullah et al. (2020)	*Jawi Character Recognition Using Deep Learning*	CNN Standar (VGG-like)	Mencapai akurasi 95% pada karakter mandiri, namun memiliki keterbatasan dalam mengenali teks bersambung karena kendala segmentasi.
Razali et al. (2023)	*Improved Classification of Handwritten Jawi Script Based on Main Part of Script Body*	Pengolahan Citra & Klasifikasi Teks	Meningkatkan klasifikasi tulisan Jawi dengan memfokuskan ekstraksi fitur pada bagian utama anatomi aksara, sehingga sangat relevan untuk karakteristik aksara Jawi lokal.
Imane et al. (2025)	*Enhancing Arabic Handwritten Word Recognition: A CNN-BiLSTM-CTC Architecture with Attention Mechanism and Adaptive Augmentation*	CNN + BiLSTM + CTC + Augmentasi	Membuktikan bahwa arsitektur CRNN-CTC yang dikombinasikan dengan augmentasi adaptif efektif dalam menangani ketidakseimbangan kelas pada pengenalan kata tulisan tangan Arab.
Shi et al. (2017)	*An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to Scene Text Recognition*	CRNN (CNN + BiLSTM + CTC)	Memperkenalkan pendekatan tanpa segmentasi yang efektif untuk mengenali teks sekuensial tanpa memerlukan segmentasi atau pelabelan pada setiap karakter.

Yu dan Koltun (2016)	*Multi-Scale Context Aggregation by Dilated Convolutions*	Dilated Convolution	Memperluas receptive field tanpa mengurangi resolusi citra, sehingga membantu mempertahankan fitur-fitur detail seperti titik pada aksara Arab dan Jawi.

2.2 **Landasan Teori / Kajian Teori**

2.2.1 Manuskrip (Naskah Kuno)

Manuskrip kuno Aceh merupakan representasi berharga dari warisan intelektual masa lalu yang mencakup dimensi keilmuan teologi, hukum, hingga sastra. Salah satu naskah representatif adalah Hikayat Akhbarul Karim karya Teungku Syekh Seumatang, sebuah karya sastra klasik abad ke-19 yang memuat pokok-pokok ajaran akidah, fikih, dan akhlak. Seiring usia, naskah ini mengalami degradasi fisik berupa kertas rapuh, bercak noda, serta tinta yang memudar atau menembus halaman (bleed-through). Tantangan tersebut menghambat proses pembacaan manual dan berisiko pada kesalahan interpretasi informasi. Oleh karena itu, diperlukan pendekatan Computer Vision yang tangguh untuk memisahkan teks dari noise visual guna memastikan kandungan informasi naskah tetap lestari melalui teknologi transliterasi otomatis yang presisi.

Ada juga kitab Hikayat Nur Nabi Muhammad Karangan Teuku Murthada Bin Yusuf, Piaraboya, muara batu, Aceh Utara diidentifikasi sebagai karya dari tahun 1988M – 1408H.

(Hikayat Bahaya Seribe) atau biasa dikenal secara luas dalam katalog manuskrip sebagai Hikayat Seribu Masalah. Hikayat ini sangat populer di Nusantara dan berisi tentang dialog atau tanya-jawab mengenai teologi, fikih, dan penciptaan alam semesta antara seorang pendeta Yahudi bernama Abdullah bin Salam dengan Nabi Muhammad SAW. disalin kembali oleh seorang ulama

bernama Teungku Muhammad Junaid Abdullah yang berasal dari daerah Darul Imarah, Kabupaten Aceh Besar. Sekitaran 1970-an, 1980-an, hingga awal 2000-an.

2.2.2 Bahasa Aceh dan Aksara Jawoe

Bahasa Aceh merupakan bahasa yang termasuk dalam rumpun bahasa Austronesia yang digunakan oleh masyarakat di wilayah Aceh. Secara linguistik, bahasa Aceh memiliki kekayaan fonetik yang unik, terutama pada bunyi vokal dan konsonan sengau (nasal). Untuk merekam bunyi-bunyi tersebut sebelum penggunaan alfabet Latin meluas, masyarakat Aceh menggunakan Aksara Jawoe. Aksara Jawoe adalah bentuk modifikasi dari abjad Arab (Hijaiyah) yang disesuaikan dengan kebutuhan fonetik bahasa Aceh. Beberapa karakteristik utama yang membedakannya dengan bahasa Arab standar adalah:

1. Penambahan Karakter: Terdapat beberapa huruf tambahan untuk melambangkan bunyi yang tidak ada dalam bahasa Arab, seperti bunyi "ng" (huruf *nga*), "p" (huruf *pa*), "c" (huruf *ca*), dan "ny" (huruf *nya*).
2. Sistem Penulisan: Ditulis secara kursif (bersambung) dari kanan ke kiri.
3. Variasi Allograf: Satu karakter dapat berubah bentuk secara signifikan tergantung pada posisinya (di awal, tengah, akhir, atau berdiri sendiri). Hal ini menuntut model pembelajaran mesin untuk memiliki kemampuan mengenali fitur visual yang dinamis namun tetap konsisten dalam pelabelan (Rahman et al., 2025; Razali et al., 2023).

2.2.3 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) merupakan arsitektur *Deep Learning* yang dirancang untuk mengekstraksi informasi penting dari sebuah citra secara otomatis melalui lapisan-lapisan filter yang bekerja secara hierarkis. Dalam proses transliterasi Aksara Jawoe, CNN berperan sebagai ekstraktor fitur visual yang memindai seluruh permukaan citra tulisan tangan untuk mengenali pola-pola dasar seperti garis, lengkungan, dan titik-titik karakter. Melalui operasi konvolusi, model mampu menangkap karakteristik unik dari setiap karakter Jawoe dan mengubahnya

menjadi representasi data yang kaya informasi (fitur), yang kemudian akan diteruskan ke lapisan berikutnya untuk diproses sebagai urutan teks.

2.2.4 Dilated Convolution

Dilated Convolution atau *Atrous Convolution* adalah pengembangan dari konvolusi standar yang memperkenalkan parameter *dilation rate* (d). Parameter ini menentukan jarak antar bobot kernel saat melakukan operasi konvolusi. Keunggulan utamanya adalah memperluas *receptive field* (jangkauan pandang filter) secara eksponensial tanpa meningkatkan jumlah parameter komputasi dan tanpa mengurangi resolusi citra (tanpa *pooling* yang agresif).

Pada sistem transliterasi Jawoe, *Dilated Convolution* sangat krusial karena mampu menangkap informasi konteks karakter yang lebar (seperti huruf yang memanjang) sekaligus mempertahankan detail kecil (titik) yang sering hilang pada konvolusi standar (Yu & Koltun, 2016).

2.2.5 Bidirectional Long Short-Term Memory (Bi-LSTM)

Long Short-Term Memory (LSTM) adalah jenis *Recurrent Neural Network* (RNN) yang dirancang untuk mengatasi masalah *vanishing gradient*, sehingga mampu mempelajari ketergantungan jangka panjang dalam data sekuensial. *Bidirectional* LSTM (Bi-LSTM) terdiri dari dua lapisan LSTM yang bekerja secara paralel: satu memproses urutan dari depan ke belakang (*forward*), dan satu lagi dari belakang ke depan (*backward*). Hal ini memungkinkan model memiliki informasi konteks dari masa lalu dan masa depan secara bersamaan. Mengingat Aksara Jawoe ditulis bersambung, pengenalan satu karakter sangat terbantu oleh informasi karakter sebelum dan sesudahnya (Graves et al., 2013).

2.3 **Kerangka Berfikir**

Kerangka berpikir dalam penelitian ini didasari oleh sulitnya membaca dan melestarikan manuskrip kuno Aceh yang menggunakan Aksara Jawoe, terutama karena sifat tulisannya yang kursif (bersambung) dan memiliki variasi tulisan tangan yang kompleks. Masalah utama yang sering muncul pada metode konvolusi standar adalah hilangnya detail fitur halus seperti titik dan harakat saat gambar

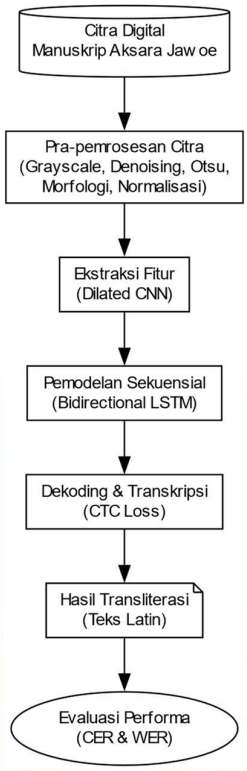
diproses dalam jaringan yang dalam. Oleh karena itu, penggunaan *Dilated Convolutional Neural Network* dipilih karena mampu memperluas bidang pandang model untuk menangkap konteks kata yang lebih luas tanpa mengorbankan resolusi detail kecil pada citra. Hal ini menjadi solusi teknis untuk mempertahankan karakteristik unik Aksara Jawoe agar tetap terbaca dengan jelas oleh sistem sebelum masuk ke tahap pengenalan urutan karakter.

Proses penelitian ini dimulai dengan tahap *preprocessing* citra manuskrip, yang kemudian diikuti oleh ekstraksi fitur visual menggunakan lapisan *Dilated CNN*. Fitur yang telah diekstraksi tersebut kemudian diproses oleh *Bidirectional Long Short-Term Memory* (Bi-LSTM) untuk mempelajari ketergantungan urutan antar karakter dari dua arah, yaitu maju dan mundur. Penggunaan Bi-LSTM sangat krusial karena bentuk karakter Jawoe sangat dipengaruhi oleh posisi dan hubungan dengan karakter di sekitarnya. Dengan mengintegrasikan kedua metode ini, diharapkan sistem dapat menghasilkan transliterasi dari Aksara Jawoe ke alfabet Latin dengan tingkat akurasi yang tinggi, yang diukur melalui parameter *Character Error Rate* (CER) dan *Word Error Rate* (WER) sebagai bukti keberhasilan model dalam mengenali teks secara utuh.

2.3.1 Alur Kerangka Berpikir

Alur kerangka berpikir dalam penelitian ini dirancang secara sekuensial dan terstruktur. Proses diawali dengan akuisisi citra digital manuskrip Aksara Jawoe, yang kemudian melewati serangkaian tahapan pra-pemrosesan (*pre-processing*). Tahapan ini mencakup konversi *grayscale*, reduksi *noise* (*denoising*), binerisasi Otsu, operasi morfologi, hingga normalisasi piksel. Citra yang telah terstandarisasi tersebut selanjutnya didistribusikan ke dalam lapisan *Dilated Convolutional Neural Network* (D-CNN) untuk proses ekstraksi fitur visual secara mendetail. Representasi spasial yang dihasilkan kemudian diproses oleh jaringan *Bidirectional Long Short-Term Memory* (Bi-LSTM) guna memodelkan konteks urutan temporal karakter dari dua arah secara simultan. Melalui mekanisme transkripsi dan algoritma dekoding *Connectionist Temporal Classification* (CTC), model menghasilkan luaran akhir berupa transliterasi teks beraksara Latin secara

segmentation-free. Pada tahap pungkasan, keandalan sistem divalidasi secara objektif menggunakan metrik *Character Error Rate* (CER) dan *Word Error Rate* (WER).



Gambar 2. 1 Alur Kerangka Berpikir

2.3.2 Jurnal Pendukung

Jurnal pendukung dalam penelitian ini berfungsi sebagai landasan ilmiah untuk memperkuat pemilihan metode dan arsitektur yang digunakan. Berikut adalah beberapa jurnal utama yang menjadi acuan teknis:

1. Jurnal Mengenai Dilated Convolution: Penelitian oleh Chen et al. (2018) dalam jurnalnya yang membahas tentang *Atrous Convolution* (Dilated Convolution) memberikan bukti bahwa penggunaan *dilation rate* pada lapisan konvolusi sangat efektif untuk mempertahankan resolusi citra sambil memperluas cakupan fitur. Dalam penelitian ini, *Dilated CNN* didukung karena kemampuannya menjaga detail fitur titik pada karakter

yang sangat penting dalam transliterasi aksara seperti Jawoe, agar tidak hilang saat proses pengecilan dimensi (downsampling).

2. Penelitian dari Razali et al. (2023) yang berjudul "Improved Classification of Handwritten Jawi Script Based on Main Part of Script Body" menjadi pendukung utama dalam pemahaman pengenalan karakter Jawi secara modern. Jurnal ini menyoroti efektivitas metode klasifikasi *deep learning* pada badan naskah, yang mendukung argumen penelitian ini untuk memproses citra secara utuh tanpa segmentasi manual..
3. Jurnal Mengenai Implementasi Bi-LSTM: Penelitian oleh Graves et al. (2013) merupakan jurnal fundamental yang membuktikan bahwa *Bidirectional LSTM* jauh lebih unggul dibandingkan LSTM standar dalam tugas pengenalan pola sekuensial. Jurnal ini mendukung penggunaan Bi-LSTM pada Aksara Jawoe karena informasi dari "masa depan" (huruf setelahnya) dan "masa lalu" (huruf sebelumnya) sangat diperlukan untuk menentukan identitas sebuah huruf bersambung secara tepat.
4. Penelitian oleh Razali et al. (2024) mengenai augmentasi dataset tambahan bahasa Arab untuk tulisan Jawi menggunakan *Deep Learning* menekankan pentingnya teknik augmentasi citra pada dataset Jawi yang terbatas. Jurnal ini memberikan landasan metodologis mengenai perlakuan data yang diaplikasikan dalam penelitian ini.
5. Penelitian oleh Imane et al. (2025) mengenai pemanfaatan arsitektur CNN-BiLSTM-CTC dipadukan dengan augmentasi adaptif sangat selaras dengan kerangka teknis penelitian ini. Jurnal ini memperkuat argumen bahwa kombinasi ekstraksi fitur konvolusional dan pemodelan temporal adalah standar terbaik untuk menangani ketidakseimbangan kelas (class imbalance) pada teks bersambung.

2.4 **Hipotesis**

Penelitian ini merupakan penelitian kuantitatif prediktif, sehingga hipotesis disusun untuk menguji kemampuan model dalam melakukan prediksi berdasarkan

data historis. Adapun hipotesis yang diajukan dalam penelitian ini adalah sebagai berikut:

2.4.1 Hipotesis Nol (H_0):

Model berbasis *Dilated Convolutional Neural Network* (D-CNN) mampu memprediksi dan mengekstraksi fitur visual pada Aksara Jawoe secara akurat dengan mempertahankan detail fitur halus seperti titik dan harakat melalui perluasan *receptive field*.

2.4.2 Hipotesis Umum (H_1):

Penggunaan arsitektur *Bidirectional Long Short-Term Memory* (Bi-LSTM) memiliki kemampuan yang signifikan dalam memprediksi urutan karakter pada teks kursif Jawoe dengan memanfaatkan konteks informasi dari dua arah (maju dan mundur).

2.4.3 Hipotesis Tambahan (H_2):

Kombinasi antara *Dilated CNN* dan *Bi-LSTM* dapat memprediksi transliterasi Aksara Jawoe ke Latin dengan tingkat kesalahan (*error rate*) yang rendah, yang dibuktikan melalui nilai *Character Error Rate* (CER) dan *Word Error Rate* (WER) yang optimal.

جامعة الرانري
AR-RANIRY

BAB III
METODOLOGI PENELITIAN

3.1 **Jenis Penelitian**

Penelitian ini merupakan penelitian kuantitatif dengan pendekatan eksperimental yang bersifat prediktif, di mana fokus utamanya adalah menguji keandalan arsitektur *Deep Learning* dalam melakukan transliterasi teks. Melalui skema eksperimen, penelitian ini mengintegrasikan metode *Dilated Convolutional Neural Network* (D-CNN) untuk ekstraksi fitur visual dan *Bidirectional Long Short-Term Memory* (Bi-LSTM) untuk pemodelan urutan karakter. Pendekatan ini bertujuan untuk menghasilkan model yang mampu memprediksi label teks Latin dari citra manuskrip Aksara Jawoe secara otomatis, dengan mengukur tingkat keberhasilan berdasarkan data historis yang telah diberi label melalui metrik evaluasi numerik seperti *Character Error Rate* (CER).

3.2 Desain Penelitian (Alur Penelitian)

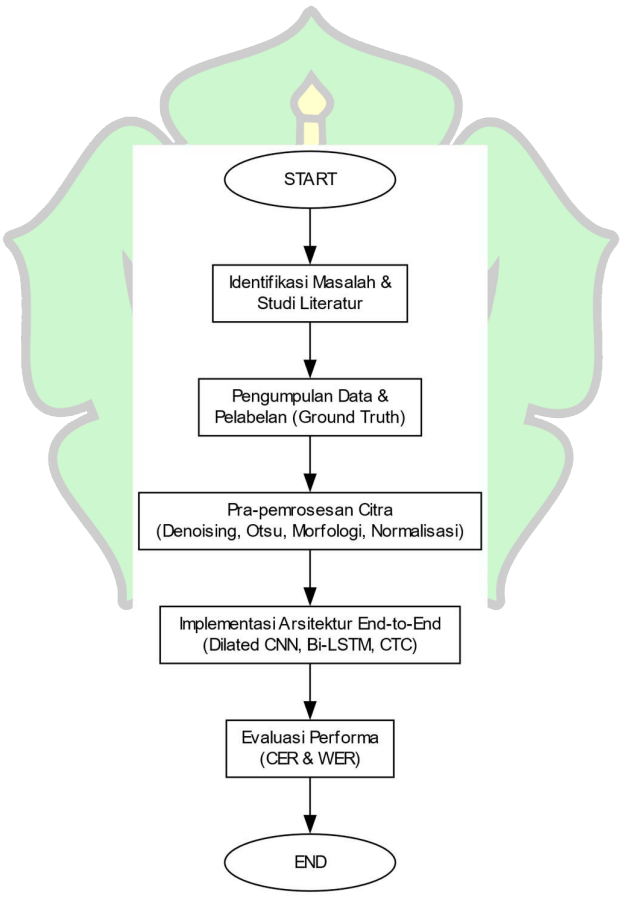
Desain penelitian ini menerapkan kerangka kerja kuantitatif-eksperimental yang bersifat prediktif guna membangun model transliterasi otomatis Aksara Jawoe ke alfabet Latin. Secara prosedural, alur penelitian diawali dengan identifikasi masalah dan studi literatur komprehensif, yang kemudian dilanjutkan pada tahap pengumpulan data. Pada tahap ini, citra digital dari manuskrip kuno Aceh diakuisisi dan direpresentasikan ke dalam dataset yang telah dilabeli (*ground truth*). Data tersebut selanjutnya melewati tahapan pra-pemrosesan berjenjangmeliputi penghalusan citra (*denoising*), binerisasi Otsu, operasi morfologi, dan normalisasiuntuk memastikan fitur morfologis karakter Jawoe terbebas dari *noise* latar belakang naskah.

Inti dari desain operasional penelitian ini terletak pada pengembangan dan implementasi arsitektur *End-to-End Deep Learning*. Arsitektur ini mengintegrasikan *Dilated Convolutional Neural Network* (D-CNN) untuk mengekstraksi fitur visual detail dengan mempertahankan resolusi spasial, serta

Bidirectional Long Short-Term Memory (Bi-LSTM) untuk memodelkan ketergantungan urutan karakter kursif dari dua arah secara simultan. Sebagai elemen krusial, model ini dilengkapi dengan fungsi *Connectionist Temporal Classification* (CTC) yang memungkinkan sistem bekerja murni secara *segmentation-free* tanpa memerlukan pemotongan batas huruf secara manual.

Tahap pungkasan dari alur penelitian ini adalah evaluasi performa model. Pengujian (*testing*) dilakukan terhadap data uji yang terisolasi menggunakan metrik *Character Error Rate* (CER) dan *Word Error Rate* (WER) guna mengukur tingkat akurasi serta keandalan prediksi transliterasi yang dihasilkan oleh sistem secara objektif.

.



Gambar 3. 1 Desain Penelitian

3.2 Metode Pengumpulan Data

Peneliti menggunakan metode pengumpulan data yang terdiri dari identifikasi masalah dan studi literatur dalam penelitian ini. Untuk mengidentifikasi masalah, peneliti melakukan observasi terhadap fenomena digitalisasi manuskrip. Sementara untuk studi literatur, peneliti melibatkan perbandingan pada penelitian-penelitian sebelumnya terkait arsitektur *Deep Learning* yang dapat diadopsi dalam penelitian ini.

3.2.1 Identifikasi Masalah

Tahapan ini dilakukan untuk menemukan topik penelitian yang sesuai dengan bidang ilmu yang dipelajari. Peneliti menemukan masalah mengenai rendahnya aksesibilitas dan hambatan dalam memahami isi manuskrip kuno Aceh. Hal ini dikarenakan proses transliterasi manual yang memakan waktu lama, serta karakteristik naskah yang kompleks karena menggunakan dua bahasa berbeda di dalamnya, yaitu Bahasa Arab dan Bahasa Aceh (Aksara Jawoe). Kondisi manuskrip yang mulai rusak juga menyulitkan proses identifikasi karakter secara visual, sehingga diperlukan sebuah sistem cerdas yang mampu melakukan transliterasi secara otomatis.

3.2.2 Studi Literatur

Pada tahapan ini, peneliti menggunakan metode studi literatur untuk mengidentifikasi berbagai tindakan yang akan dilakukan terkait dengan pengumpulan data dan memberikan gambaran mengenai solusi teknis terhadap masalah transliterasi aksara Jawoe. Studi literatur ini mencakup pengkajian metode *Dilated Convolutional Neural Network* (D-CNN) untuk ekstraksi fitur citra yang detail dan *Bidirectional Long Short-Term Memory* (Bi-LSTM) untuk pengenalan teks berdasarkan konteks bahasa. Referensi yang digunakan bersumber dari jurnal ilmiah, artikel, buku, serta bahan pendukung lainnya yang berkaitan dengan teknologi *Handwritten Text Recognition* (HTR).

3.3 Metode Simulasi

Metode simulasi dalam penelitian ini merujuk pada proses perancangan dan pengujian sistem saraf tiruan yang dikembangkan untuk mereplikasi kemampuan kognitif pakar filologi dalam membaca manuskrip kuno. Simulasi ini bukan sekadar menjalankan kode, melainkan sebuah eksperimen lingkungan komputasi yang memanfaatkan bahasa pemrograman Python serta pustaka *Deep Learning* tingkat tinggi (seperti TensorFlow atau PyTorch). Tujuan utamanya adalah untuk memvalidasi efektivitas arsitektur Dilated Convolutional Neural Network (D-CNN) yang digabungkan dengan Bidirectional Long Short-Term Memory (Bi-LSTM) dalam memproses kompleksitas visual aksara Jawoe yang bersifat bilingual (Arab dan Aceh).

3.3.1 Perumusan Konsep Penelitian

Konsep penelitian ini berpijak pada paradigma *End-to-End Learning*, di mana sistem belajar memetakan piksel gambar langsung menjadi urutan teks tanpa melalui tahap segmentasi karakter manual yang sering menjadi penyebab utama kegagalan pada OCR tradisional. Konsep ini secara mendalam dijabarkan ke dalam tiga pilar teknis utama:

1. Ekstraksi Fitur Spasial dengan Mekanisme *Dilated Convolution*: Berbeda dengan konvolusi standar, konsep *Dilated CNN* yang diusung dalam simulasi ini bertujuan untuk memperluas *receptive field* (bidang pandang) tanpa meningkatkan jumlah parameter secara drastis atau mengurangi resolusi spasial. Hal ini sangat krusial dalam aksara Jawoe karena model harus mampu menangkap hubungan antara badan huruf dengan titik-titik (diakritik) yang sering kali terletak cukup jauh dari garis utama tulisan. Dengan pengaturan *dilation rate* yang tepat, model dirancang untuk tetap sensitif terhadap fitur-fitur halus yang memudar namun tetap mampu memahami struktur bentuk huruf secara utuh.
2. Pemodelan Urutan Temporal melalui *Bidirectional LSTM*: Peneliti merumuskan bahwa teks Jawoe adalah data deret waktu (*time-series*) yang memiliki ketergantungan konteks yang sangat kuat. Penggunaan Bi-LSTM

memungkinkan sistem untuk melakukan pemrosesan dua arah: dari kanan-ke-kiri (mengikuti arah baca asli aksara Jawoe) dan dari kiri-ke-kanan (untuk menangkap konteks karakter setelahnya). Pendekatan ini memastikan bahwa prediksi karakter tertentu didasarkan pada probabilitas kemunculan karakter di sekitarnya, yang sangat membantu dalam membedakan antara kosa kata Bahasa Arab dan struktur afiksasi unik dalam Bahasa Aceh.

3. Transkripsi Dinamis dengan *Connectionist Temporal Classification* (CTC): Konsep ini dirancang untuk mengatasi masalah *alignment* antara urutan visual gambar dengan urutan teks Latin. Mekanisme CTC memungkinkan model untuk memprediksi karakter pada setiap kolom piksel dan secara otomatis menggabungkan karakter yang sama serta menghilangkan karakter kosong (*blank*). Hal ini menjadi solusi fundamental bagi tulisan tangan Jawoe yang bersifat kursif (saling bersambung), di mana batas antar huruf tidak dapat didefinisikan secara tegas secara geometris.

3.3.2 Implementasi Metode Arsitektur

A. Pengumpulan dan Standarisasi Dataset

Proses akuisisi data pada penelitian ini tidak dilakukan dengan memotong (*cropping*) citra kata langsung dari manuskrip fisik Hikayat Akhbarul Karim. Hal ini dikarenakan tingginya tingkat degradasi kertas, *noise* latar belakang, dan pudarnya tinta yang dapat mengganggu proses ekstraksi fitur dasar. Sebagai gantinya, penelitian ini mengimplementasikan pendekatan pembentukan dataset sintetis (*synthetic dataset generation*) berbasis tipografi digital.

Seluruh kosa kata unik yang diekstrak dari manuskrip disalin ulang (*re-type*) secara digital dan di-c*apture* (*screenshot*) menggunakan 5 (lima) variasi jenis huruf (*font*) Arab/Jawi standar yang berbeda. Penggunaan 5 variasi *font* ini dijustifikasi oleh tiga alasan teknis yang fundamental:

1. Validasi Arsitektur (*Baseline Proofing*): Menyediakan data yang bersih (*clean data*) untuk memvalidasi kelayakan arsitektur CRNN-CTC dalam

mengenali struktur morfologi Aksara Jawoe sebelum dihadapkan pada *noise* ekstrem dari naskah tulisan tangan asli.

2. Augmentasi Struktural: Setiap *font* memiliki ketebalan goresan (*stroke width*), proporsi lengkungan, dan gaya ligatur yang berbeda. Variasi ini memaksa model *Dilated* CNN untuk mempelajari fitur esensial dari huruf, bukan sekadar menghafal satu gaya tipografi.
3. Pengujian Diskontinuitas Karakter: Beberapa jenis *font* menghasilkan spasi tegas atau karakter yang terputus di akhir kata (misalnya huruf 'Ya' yang tidak tersambung sempurna). Keberagaman ini melatih ketahanan fungsi *loss* CTC untuk tetap mampu menyelaraskan karakter yang terpisah secara visual ke dalam satu kesatuan teks Latin yang utuh.

Rangkaian proses ini menghasilkan dataset akhir sebanyak 2.000 citra kata. Pelabelan ini sangat krusial sebagai acuan utama (*ground truth*) dalam fase pelatihan sistem Deep Learning. Sebagaimana ditegaskan oleh Nasihah (2023), performa dan presisi model pengenalan pola sangat dipengaruhi oleh tingkat variabilitas serta kualitas dataset yang digunakankhususnya pada naskah aksara lokal yang kaya akan detail visual unik. Link dataset https://share.google/pEuCxrDWLaAC8phM3.

Gambar 3. 2 Dataset Aksara Jawoe

B. Pre-processing

Mengingat dataset awal berasal dari tangkapan layar (*screenshot*) teks digital yang relatif bersih dari kotoran latar belakang, tahapan *Denoising*

(*Gaussian Blur*) dan *Morphological Filtering* (Dilasi/Erosi) pada alur penelitian ini dialihfungsikan. Operasi ini tidak lagi murni bertujuan membersihkan *noise*, melainkan secara sengaja digunakan untuk mendegradasi ketajaman tepi huruf digital agar menyimulasikan efek pendaran tinta (*ink bleed*) dan ketidaksempurnaan goresan pena yang lazim ditemukan pada manuskrip kuno.

Tahap *pre-processing* bertujuan untuk meningkatkan kualitas fitur visual pada citra dan menghilangkan gangguan (*noise*) yang dapat menghambat proses ekstraksi fitur oleh lapisan konvolusi. Langkah-langkah yang diimplementasikan adalah sebagai berikut:

1) Grayscale Conversion: Mengubah citra asli (RGB) menjadi citra berskala abu-abu. Langkah ini bertujuan untuk menyederhanakan informasi warna menjadi informasi intensitas cahaya, sehingga beban komputasi menjadi lebih ringan tanpa menghilangkan bentuk karakter utama.
2) Denoising (Penghalusan Citra): Menerapkan teknik *Gaussian Blur* pada citra *grayscale* untuk menghaluskan *noise* berfrekuensi tinggi (seperti tekstur serat kertas yang kasar atau noda debu). Langkah ini sangat krusial dilakukan di awal agar *noise* tersebut tidak mengganggu kalkulasi ambang batas pada tahap binerisasi.
3) Binarization (Otsu's Thresholding): Mengubah citra *grayscale* yang telah dihaluskan menjadi citra biner (hitam-putih). Algoritma Otsu dipilih karena mampu menentukan ambang batas secara otomatis berdasarkan distribusi piksel, sehingga sangat efektif untuk memisahkan antara goresan tinta (objek) dengan latar belakang kertas yang kusam atau bernoda.
4) Morphological Filtering: Menerapkan operasi morfologi (seperti *closing* yang merupakan gabungan *dilation* diikuti *erosion*) pada citra biner. Operasi ini berfungsi untuk mempertegas garis karakter yang putus-putus, menambal celah mikroskopis pada goresan tinta, serta membuang sisa bercak piksel kecil yang tidak relevan.
5) Image Resizing & Padding: Seluruh citra diubah ukurannya menjadi dimensi yang konsisten (yakni tinggi 64 piksel). Peneliti menerapkan teknik

padding untuk menjaga rasio aspek aslinya agar karakter tidak mengalami distorsi geometris, sehingga ciri khas morfologi aksara Jawoe tetap terjaga.

6) Normalization: Nilai piksel yang semula berada pada rentang 0-255 dikonversi ke rentang 0-1. Normalisasi ini dilakukan untuk memastikan distribusi data yang masuk ke model bersifat seragam, yang pada gilirannya menstabilkan dan mempercepat proses konvergensi saat pelatihan.

C. Augmentasi Data

Sebagai upaya untuk mengatasi keterbatasan jumlah dataset dan meningkatkan kemampuan generalisasi model terhadap data baru (mencegah *overfitting*), penelitian ini mengimplementasikan teknik augmentasi data secara *real-time* (*on-the-fly*) selama proses pelatihan model berlangsung. Transformasi visual yang diterapkan secara acak pada citra latih meliputi:

1) Rotasi Acak (*Random Rotation*): Memutar citra pada rentang sudut -3° hingga +3°. Langkah ini bertujuan untuk menyimulasikan kemiringan alami baris teks atau variasi sudut pandang pemindaian naskah asli.

2) Penskalaan Acak (*Random Scaling*): Mengubah skala ukuran citra dengan rentang toleransi ±10% guna melatih model agar kebal terhadap variasi ukuran huruf tulisan tangan kaligrafer.

3) Transformasi Morfologi (*Morphological Adjustments*): Menerapkan efek dilasi (penebalan) dan erosi (penipisan) goresan secara acak. Teknik ini mensimulasikan kondisi fisik naskah di mana ketebalan tinta bisa berbeda-beda akibat tekanan pena atau pudarnya tinta termakan usia.

4) Keterangan Teknis: Augmentasi ini hanya diterapkan secara eksklusif pada Data Latih (*Training Set*). Data Uji (*Testing Set*) dibiarkan dalam kondisi aslinya tanpa modifikasi tambahan untuk memastikan evaluasi performa transliterasi bersifat objektif dan merepresentasikan kondisi dunia nyata.

D. Split Dataset

Pembagian dataset (*data splitting*) merupakan salah satu prosedur paling krusial dalam pengembangan model Deep Learning yang bersifat prediktif. Tujuan utama dari tahapan ini adalah untuk memastikan bahwa model yang dibangun tidak hanya memiliki kemampuan menghafal data yang sudah ada (*memorizing* atau *overfitting*), melainkan benar-benar memiliki kemampuan generalisasi (*generalization*) dalam mengenali pola tulisan tangan pada naskah yang belum pernah dilihat sebelumnya. Dalam penelitian ini, total 2.000 citra kata yang telah melalui tahap pra-pemrosesan dibagi secara sistematis menggunakan rasio 80:20, dengan rincian teknis sebagai berikut:

1) Data Latih (*Training Set*) – 80% (1600 Citra): Bagian terbesar dari dataset ini dialokasikan untuk proses pembelajaran inti. Selama fase pelatihan, citra-citra dalam *training set* dimasukkan ke dalam arsitektur D-CNN dan Bi-LSTM. Model akan melakukan ekstraksi fitur dan prediksi urutan karakter secara berulang-ulang. Setiap kali model melakukan kesalahan, fungsi *loss* CTC akan memberikan umpan balik untuk menyesuaikan bobot internal melalui proses *backpropagation*. Dengan jumlah 1600 sampel, diharapkan model dapat terpapar pada berbagai variasi gaya penulisan aksara Jawoe, mulai dari gaya penulisan yang rapi hingga gaya penulisan yang kursif dan sulit dibaca.
2) Data Uji (*Testing Set*) – 20% (400 Citra): Sisa data sebesar 400 citra disisihkan sepenuhnya dan bersifat "terisolasi" dari proses pelatihan. Data ini berfungsi sebagai instrumen pengujian objektif. Setelah model selesai dilatih, kemampuannya akan diuji menggunakan data ini untuk melihat seberapa akurat model melakukan transliterasi terhadap kata-kata atau karakter yang baru bagi sistem. Hasil dari pengujian pada *testing set* inilah yang akan menjadi acuan utama dalam perhitungan nilai *Character Error Rate* (CER) dan *Word Error Rate* (WER).

Selain pembagian di atas, peneliti juga menerapkan teknik Random Shuffling sebelum proses pemisahan dilakukan. Teknik ini bertujuan untuk memastikan bahwa variasi kosa kata Bahasa Arab dan Bahasa Aceh tersebar secara merata di kedua kelompok data tersebut. Tanpa *shuffling* yang tepat, dikhawatirkan salah satu kelompok data hanya berisi gaya penulisan dari satu naskah saja, yang dapat mengakibatkan hasil evaluasi menjadi tidak akurat atau bias.

Lebih lanjut, dalam proses simulasi, sebanyak **20% dari total data latih (setara dengan 320 citra)** dialokasikan secara khusus sebagai Data Validasi (*Validation Set*). Data validasi ini berfungsi untuk memantau performa model pada setiap iterasi (*epoch*) guna menentukan kapan proses pelatihan harus dihentikan (*early stopping*). Hal ini dilakukan secara terukur untuk menjaga keseimbangan antara performa model pada data latih dan kemampuan adaptasinya pada data uji, sehingga dihasilkan model transliterasi yang stabil dan terhindar dari kondisi *overfitting* guna menentukan kapan proses pelatihan harus dihentikan. Pembagian dataset dapat dilihat pada Tabel 3.1.

Tabel 3. 1 Pembagian Dataset

Kategori Data	Persentase	Jumlah Citra
Data Latih (Training)	80%	1600
Data Uji (Testing)	20%	400
Total	100%	2000

E. Membangun Model

Tahap membangun model melibatkan perancangan arsitektur *End-to-End* berbasis *Convolutional Recurrent Neural Network* (CRNN) yang mengintegrasikan lapisan Dilated CNN, Bidirectional LSTM (Bi-LSTM), dan dioptimalkan menggunakan fungsi kerugian *Connectionist Temporal Classification* (CTC). Pendekatan *segmentation-free* ini memfasilitasi sistem untuk menerima input berupa citra kata secara utuh dan memetakan fitur

spasial-temporal langsung ke dalam sekuens label Latin tanpa memerlukan pemotongan (segmentasi) koordinat per karakter secara manual.

1) Penyusunan Lapisan *Feature Extractor* (Dilated CNN): Peneliti membangun lapisan konvolusi dengan teknik *dilation rate* yang bervariasi. Penggunaan *Dilated CNN* di sini bertujuan untuk menangkap informasi spasial yang lebih luas dari setiap karakter Jawoe tanpa kehilangan resolusi pada detail kecil. Hal ini sangat penting untuk mengenali perbedaan halus pada titik-titik karakter Arab dan Aceh (seperti perbedaan antara huruf *Fa* dan *Qa* atau *Pa*) yang menjadi tantangan utama dalam manuskrip kuno. Setiap lapisan konvolusi diikuti oleh *Batch Normalization* untuk mempercepat proses konvergensi dan *Activation Function* ReLU untuk menangani masalah *vanishing gradient*.
2) Penyusunan Lapisan Sequence Learner (Bidirectional LSTM): Setelah fitur visual dari tiap area *receptive field* diekstraksi dan diratakan (*reshaped*) menjadi urutan *frame* fitur berdimensi 1D, matriks tersebut diteruskan ke lapisan *Stacked Bidirectional Long Short-Term Memory* (Bi-LSTM). Lapisan ini bertugas memodelkan dependensi temporal dan kontekstual sekuensial. Mengingat morfologi aksara Jawoe yang kursif dan sarat akan ligatur, sel Bi-LSTM membaca komputasi probabilitas dari arah *forward* (kiri-ke-kanan) dan *backward* (kanan-ke-kiri) secara independen, untuk kemudian digabungkan guna memberikan konteks bilingual (Arab-Aceh) yang utuh.
3) Penyusunan Lapisan *Output* (Dense & Softmax Layer): Bagian akhir dari arsitektur ini adalah lapisan *Fully Connected* (Dense) yang diakhiri dengan fungsi aktivasi Softmax. Fungsi ini akan menghitung nilai probabilitas untuk seluruh kategori karakter Latin dan simbol diakritik yang telah didefinisikan dalam *ground truth*. Model akan memilih karakter dengan nilai probabilitas tertinggi sebagai hasil keluaran sistem.

Tabel 3. 2 Arsitektur Model CRNN

Komponen Jaringan	Jenis Lapisan	Detail Fungsi Teknis
Feature Extractor	Dilated CNN	Mengekstraksi fitur visual detail dari setiap potongan aksara Jawoe.
Feature Mapping	Reshape & Dense	Mengonversi peta fitur visual menjadi vektor input untuk urutan temporal.
Sequence Learner	Stacked Bi-LSTM	Mempelajari pola urutan karakter dan konteks bilingual (Arab-Aceh).
Classification	Softmax Layer	Menghasilkan probabilitas karakter Latin final sebagai hasil transliterasi.

F. Model

Tahap terakhir adalah penyatuan seluruh komponen arsitektur menjadi satu model integratif yang siap melakukan transliterasi aksara Jawoe secara otomatis. Model ini merepresentasikan hasil kompilasi dari lapisan ekstraksi fitur visual (Dilated CNN), lapisan pemahaman konteks urutan (Bi-LSTM), dan lapisan pengambil keputusan akhir (Softmax Classifier). Model ini dirancang untuk bekerja secara deterministik dalam memetakan input visual menjadi label karakter Latin yang tepat. Karakteristik utama dari model yang dihasilkan adalah sebagai berikut:

1) Model Input: Model menerima tensor input dua dimensi (2D) berupa citra tingkat keabuan (grayscale) dari potongan kata utuh yang telah melewati serangkaian tahapan pre-processing. Citra kata tersebut distandarisasi secara dinamis ke dalam dimensi matriks persegi panjang yang tetap, yakni

64×256 piksel (tinggi × lebar), untuk menjamin konsistensi arsitektur saat beroperasi pada lapisan konvolusi.

2) Model Output: Output dari model ini adalah sekuens label teks Latin hasil klasifikasi yang merepresentasikan hasil transliterasi Aksara Jawoe. Ruang lingkup kamus karakter (*vocabulary*) pada keluaran model saat ini difokuskan pada karakter alfabet dasar (a-z) beserta karakter spasi, yang diekstrak langsung dari penyelarasan label asli (*ground truth*). Pengenalan simbol diakritik spesifik bahasa Aceh (seperti é, è, ô, ë, ö) didefinisikan sebagai batasan ruang lingkup penelitian saat ini, dan dapat diintegrasikan pada tahap pengembangan dataset di masa mendatang.
3) Kompilasi Model: Model dikompilasi menggunakan fungsi kerugian (*loss function*) *Connectionist Temporal Classification* (CTC) *Loss*. Fungsi ini sangat esensial untuk melatih model arsitektur sekuensial pada data yang tidak memiliki *alignment* (penyelarasan) eksplisit antara letak piksel karakter pada citra dengan posisi huruf pada teks target. Optimalisasi pembaruan bobot (*weight update*) dipandu oleh algoritma optimasi Adam (*Adaptive Moment Estimation*) dengan pengaturan *learning rate* dinamis untuk memastikan model dapat meminimalkan nilai kerugian secara efisien dan konvergen pada dataset citra kata yang tersedia.
4) Konfigurasi Pelatihan (*Hyperparameter*): Untuk memastikan proses konvergensi model berjalan secara optimal dan stabil, penelitian ini menetapkan sejumlah *hyperparameter* spesifik yang diimplementasikan pada tahap pelatihan:
    a) Optimizer: Adam (*Adaptive Moment Estimation*) dipilih untuk mengoptimalkan pembaruan bobot karena efisiensinya dalam menangani *gradient sparse*.
    b) Initial Learning Rate: Diatur secara spesifik pada nilai 1×10^{-4} (0.0001). Berdasarkan percobaan empiris, nilai ini mencegah langkah pembaruan bobot yang terlalu besar (*overshooting*) sehingga konvergensi *loss* CTC berjalan stabil pada data latih aksara yang rumit.

c) Batch Size: Ditetapkan sebanyak 16 citra per iterasi. Pemilihan angka 16 merupakan titik ekuilibrium (keseimbangan) antara alokasi VRAM pada GPU Google Colab dan pemeliharaan stabilitas estimasi gradien selama proses *backpropagation*.

d) Epoch Maksimal: Maksimal 100 *epoch*, yang dinilai lebih dari cukup bagi arsitektur Bi-LSTM untuk mencapai konvergensi penuh, dikombinasikan dengan mekanisme *Early Stopping* untuk menghentikan pelatihan jika terdeteksi indikasi *overfitting*.

e) Panjang Label Maksimal: Batas prediksi transliterasi per citra dibatasi hingga 20 karakter. Parameter ini dijustifikasi berdasarkan panjang maksimal rata-rata dari kata bersambung atau frasa tunggal dalam bahasa Aceh pada naskah Hikayat Akhbarul Karim.

f) Fungsi Kontrol Otomatis (*Callbacks*):

- *Early Stopping:* Pelatihan akan dihentikan lebih awal secara otomatis apabila tidak terjadi perbaikan pada nilai kerugian validasi (*validation loss*) selama 15 *epoch* berturut-turut (*patience = 15*), guna mencegah *overfitting*.
- *Reduce LR on Plateau:* Sistem akan memangkas nilai *learning rate* menjadi setengahnya (*factor = 0.5*) jika performa model mengalami stagnasi selama 6 *epoch* (*patience = 6*). Hal ini membantu model menemukan titik kerugian minimum (minimum global) dengan manuver yang lebih halus di tahap akhir pelatihan.

5) Mekanisme Transkripsi dan Decoding CTC: Mengadopsi prinsip transkripsi *segmentation-free*, output vektor dari lapisan Bi-LSTM didistribusikan ke lapisan *Fully Connected* (Dense) dengan fungsi aktivasi Softmax. Lapisan ini mengalkulasi probabilitas untuk seluruh kelas karakter referensi (termasuk satu elemen absolut, yakni *blank token*) pada setiap langkah waktu independen. Translasi final dari representasi matriks menjadi string dilakukan melalui operasi algoritma *CTC Decoding* (seperti *Greedy Search*).

Model ini dioptimalkan untuk menjaga ketajaman deteksi pada fitur-fitur kecil, seperti titik-titik diakritik, dengan memanfaatkan *dilation rate* pada lapisan CNN. Dengan integrasi antara ekstraksi fitur yang presisi dan pemahaman konteks urutan oleh Bi-LSTM, model ini mampu memberikan hasil transliterasi yang andal meskipun menghadapi variasi gaya tulisan tangan yang beragam pada manuskrip kuno.

3.4 Metode Analisis Data

Tahapan terakhir dari penelitian ini adalah analisis data. Analisis data ini dilakukan dengan menilai *system* yang digunakan dalam penelitian ini guna mengetahui sejauh mana tingkat keberhasilan model dalam melakukan transliterasi aksara Jawoe secara otomatis.

3.4.1 Evaluasi

Evaluasi merupakan tahapan yang digunakan untuk mengukur dan menguji performa model transliterasi yang diusulkan pada penelitian ini. Pengukuran tingkat keakuratan hasil transliterasi dilakukan dengan menggunakan metrik *Character Error Rate* (CER) dan *Word Error Rate* (WER). Kedua metrik ini mengacu pada perhitungan jarak Levenshtein (*Levenshtein Distance*) untuk menghitung jumlah minimum operasi perubahan yang diperlukan guna menyelaraskan teks hasil prediksi model dengan teks asli.

3.4.1.1 Character Error Rate (CER)

Character Error Rate (CER) digunakan untuk mengukur tingkat kesalahan transliterasi pada level karakter. Metrik ini sangat krusial untuk melihat seberapa akurat model dalam mengenali setiap huruf serta detail tanda diakritik khusus pada Aksara Jawoe. Perhitungan CER dilakukan dengan menjumlahkan seluruh kesalahan substitusi, penghapusan, dan penyisipan karakter, kemudian membaginya dengan total karakter asli pada label pembanding (*ground truth*). Persamaan matematis untuk menghitung nilai CER adalah sebagai berikut:

$$CER = \frac{(S + D + I)}{N}$$

Dimana keterangan variabel dalam rumus tersebut adalah:

1. S (*Substitution*): Jumlah karakter yang salah dikenali atau tertukar dengan karakter lain.
2. D (*Deletion*): Jumlah karakter asli yang hilang atau tidak terdeteksi oleh model.
3. I (*Insertion*): Jumlah karakter tambahan yang muncul pada hasil prediksi namun tidak ada pada data asli.
4. N: Jumlah total karakter keseluruhan pada label asli (*ground truth*).

Contoh Perhitungan CER:

Asumsikan label *ground truth* adalah **"kalon"** ($N = 5$ karakter). Teks prediksi dari model adalah **"kalen"**. Berdasarkan *Levenshtein Distance*, terdapat 1 karakter yang disubstitusi ('o' menjadi 'e'), 0 insersi, dan 0 delesi.

$$CER = \frac{(S + D + I)}{N} = \frac{(1 + 0 + 0)}{5} = 0.20 \text{ (atau } 20\%)$$

3.4.1.2 Word Error Rate (WER)

Word Error Rate (WER) digunakan untuk mengukur tingkat kesalahan transliterasi model pada level kata secara utuh. Dalam metrik ini, sebuah kata akan dianggap salah apabila terdapat minimal satu kesalahan karakter di dalamnya. Perhitungan nilai WER dilakukan melalui persamaan sebagai berikut:

$$WER = \frac{(Sw + Dw + Iw)}{\mathrm{NW}}$$

Dimana keterangan variabel dalam rumus tersebut adalah:

1. Sw: Jumlah kata yang salah atau tertukar secara keseluruhan dalam satu kesatuan kata.
2. Dw: Jumlah kata yang hilang atau tidak terdeteksi oleh model.
3. Iw: Jumlah kata tambahan yang tersisip secara salah pada hasil prediksi.
4. Nw: Jumlah total kata keseluruhan pada label asli (*ground truth*) data uji.

Contoh Perhitungan WER:

Asumsikan kalimat *ground truth* berisi 4 kata uji: **["tulong", "lon", "mebale", "jinoe"]** ($N_w = 4$). Teks prediksi dari model: **["tulem", "lon", "mebali", "jinoe"]**. Terdapat 2 kata yang salah secara utuh (*Substitution Word* = "tulem" dan "mebali").

$$WER = \frac{(S_w + D_w + I_w)}{N_w} = \frac{(2 + 0 + 0)}{4} = 0.50 \text{ (atau 50\%)}$$

Hasil dari perhitungan kedua metrik evaluasi ini (CER dan WER) akan menunjukkan seberapa efektif arsitektur *End-to-End* berbasis Dilated CNN dan Bidirectional LSTM (Bi-LSTM) dalam mengenali morfologi kursif kata dan melakukan transliterasi ke teks Latin dari potongan manuskrip kuno Hikayat Akhbarul Karim yang berjumlah keseluruhan 2.000 citra kata tersebut. Evaluasi objektif pada data uji ini menjadi parameter penentu utama apakah hipotesis penelitian yang diajukan pada Bab II dapat diterima atau ditolak.

3.5 Waktu Penelitian

Penelitian mengenai transliterasi aksara Jawoe bilingual menggunakan arsitektur Dilated CNN dan Bi-LSTM ini dilaksanakan secara mandiri oleh peneliti di Laboratorium Komputer Program Studi Teknologi Informasi, Fakultas Sains dan Teknologi. Seluruh tahapan penelitian, mulai dari pengolahan 2.000 citra manuskrip, perancangan model, hingga evaluasi sistem, direncanakan berlangsung selama kurang lebih 6 (enam) bulan yang terhitung sejak tahap persiapan hingga penyusunan laporan akhir. Proses pelatihan model (*training*) dilakukan secara intensif dengan memanfaatkan platform komputasi awan *Google Colab* guna memastikan efisiensi pemrosesan data citra berskala besar yang digunakan dalam simulasi ini.

3.6 Instrumen Penelitian

Pelaksanaan penelitian ini membutuhkan instrumen berupa perangkat keras (*hardware*) dan perangkat lunak (*software*) yang dibagi ke dalam dua lingkungan kerja, yakni lingkungan lokal untuk pra-pemrosesan data dan lingkungan komputasi

awan (*cloud computing*) untuk pelatihan model. Pemisahan ini dilakukan mengingat arsitektur model *Deep Learning* menuntut daya komputasi paralel yang tinggi.

3.6.1 Perangkat Keras (Hardware)

1. **Lingkungan Lokal (Spesifikasi Minimum):** Komputer atau laptop yang difungsikan khusus untuk ekstraksi dataset, pelabelan teks (*ground truth*), pengujian kode awal, dan penyusunan laporan. Spesifikasi minimum yang direkomendasikan adalah prosesor setingkat Intel Core i3 atau AMD Ryzen 3, RAM berkapasitas 8 GB untuk mencegah *bottleneck* saat memuat citra ke memori, dan ruang penyimpanan internal (SSD/HDD) minimal 256 GB.
2. **Lingkungan Pelatihan (*Cloud Environment*):** Proses pelatihan model komputasi berat (*training*) dieksekusi secara intensif menggunakan infrastruktur Google Colaboratory (Colab). Instrumen *cloud* ini menyediakan akses ke spesifikasi *Graphics Processing Unit* (GPU) secara virtual, dengan spesifikasi minimum berupa alokasi GPU NVIDIA T4 (VRAM 16 GB) dan *System RAM* 12 GB. Penggunaan akselerator GPU eksternal ini secara drastis memangkas waktu kalkulasi matriks pada lapisan Dilated CNN dan perhitungan sekuensial pada lapisan Bi-LSTM.

3.6.2 Perangkat Lunak (Software)

1. **Sistem Operasi Utama:** Windows 10/11, macOS, atau distribusi Linux modern yang mendukung eksekusi peramban web standar.
2. **Bahasa Pemrograman:** Python (versi 3.8 ke atas), yang menjadi standar industri dalam pengembangan kecerdasan buatan.
3. **Pustaka Pemrosesan (*Libraries*):**
   a. *TensorFlow* dan *Keras*: Digunakan sebagai *framework* utama untuk mendesain, mengompilasi, dan mengevaluasi arsitektur jaringan saraf tiruan (CRNN) serta kalkulasi *CTC Loss*.
   b. *OpenCV (Open Source Computer Vision)*: Digunakan secara eksklusif untuk seluruh operasi pengolahan citra (*image pre-processing*), seperti binerisasi Otsu dan morfologi citra.

c. *NumPy* dan *Matplotlib*: Digunakan untuk operasi komputasi array multidimensi dan visualisasi grafik performa pelatihan.

4. **Integrated Development Environment (IDE):** Jupyter Notebook (diakses melalui antarmuka Google Colab).

BAB IV
HASIL DAN PEMBAHASAN

4.1 Gambaran Umum Hasil Penelitian

Penelitian ini telah berhasil mengimplementasikan dan menguji sistem transliterasi otomatis untuk naskah kuno Aksara Jawoe ke alfabet Latin. Sistem ini dibangun menggunakan arsitektur *End-to-End* berbasis *Convolutional Recurrent Neural Network* (CRNN) yang secara spesifik mengintegrasikan *Dilated Convolutional Neural Network* (D-CNN) sebagai ekstraktor fitur visual dan *Bidirectional Long Short-Term Memory* (Bi-LSTM) sebagai pemodel sekuensial, dengan optimasi fungsi *Connectionist Temporal Classification* (CTC). Keseluruhan pelatihan dan pengujian model memanfaatkan dataset yang secara eksklusif diekstrak dari manuskrip kuno Hikayat Akhbarul Karim. Total dataset final yang diolah berjumlah 2.000 citra potongan kata. Angka ini diperoleh melalui ekstraksi 400 kelas kata (kosakata) unik dari naskah, di mana setiap kata diketik ulang ke dalam bentuk digital dan direproduksi secara konsisten menggunakan 5 (lima) variasi *font* yang berbeda guna memperkaya variabilitas fitur struktural huruf ($400 \text{ kata} \times 5 \text{ font} = 2.000 \text{ citra}$)..

Guna memastikan objektivitas pengujian dan mengukur kemampuan generalisasi sistem terhadap variasi tulisan tangan yang belum pernah dilihat sebelumnya, dataset dibagi menggunakan skema pemisahan 80:20. Sejumlah 1.600 citra kata dialokasikan sebagai data latih (*training set*) untuk proses pembelajaran arsitektur, sedangkan 400 citra kata sisanya diisolasi sepenuhnya sebagai data uji (*testing set*).

Pemaparan pada bab ini disusun secara sistematis untuk menjawab secara langsung dua rumusan masalah yang telah ditetapkan pada Bab I. Subbab 4.2 akan menguraikan bukti keberhasilan implementasi teknis dari ekstraksi fitur spasial dan pemodelan kontekstual dua arah guna menjawab rumusan masalah pertama. Selanjutnya, subbab 4.3 didedikasikan untuk menjawab rumusan masalah kedua

melalui pemaparan evaluasi performa model menggunakan metrik *Character Error Rate* (CER) dan *Word Error Rate* (WER), yang kemudian diikuti dengan uji hipotesis dan pembahasan komprehensif.

Tabel 4. 1 Ringkasan Spesifikasi Dataset Aksara Jawoe

Kategori Data	**Spesifikasi / Jumlah**
Sumber Manuskrip	Hikayat Akhbarul Karim
Total Citra Kata	2.000 citra
Total Kelas Kata Unik	404 kelas kata
Data Latih (*Training Set*)	1.600 citra (80%)
Data Uji (*Testing Set*)	400 citra (20%)

4.2 Hasil Implementasi Arsitektur Dilated CNN dan Bi-LSTM

Penguraian realisasi rancangan arsitektur *End-to-End* dan memberikan pembuktian teknis guna menjawab Rumusan Masalah pertama, yakni terkait bagaimana implementasi *Dilated Convolutional Neural Network* (D-CNN) dan *Bidirectional LSTM* mampu mengekstraksi fitur visual secara efektif pada tulisan tangan Aksara Jawoe.

4.2.1 Realisasi Lapisan Ekstraksi Fitur (Dilated CNN)

Tantangan terbesar dalam memproses citra Aksara Jawoe adalah mempertahankan detail fitur berukuran mikro (seperti tanda titik yang membedakan huruf *ba*, *ta*, *tsa*, atau *pa* dan *nga* dalam bahasa Aceh) yang rentan hilang jika model hanya menggunakan operasi *Max Pooling* agresif untuk memperluas area pandang (*receptive field*).

Implementasi ini berhasil mengatasi batasan tersebut melalui penerapan *Dilated Convolution* (konvolusi dengan pelebaran tingkat sela). Dengan menetapkan nilai *dilation rate* lebih besar dari 1 ($r = 2$ pada lapisan konvolusi tingkat menengah), jarak antar-bobot kernel diperlebar. Hal ini memungkinkan filter untuk mengamati konteks spasial yang jauh lebih luas seperti lengkungan panjang yang menghubungkan antar-karakter tanpa memangkas resolusi spasial citra secara

drastis. Eksekusi program membuktikan bahwa lapisan D-CNN mampu menyaring *noise* latar belakang kertas manuskrip sekaligus mempertahankan topologi titik dan harakat dengan sangat tajam. Parameter model dapat dilihat pada tabel 4.2

Tabel 4. 2 Arsitektur Lapisan Ekstraksi Fitur dengan Dilated Convolution

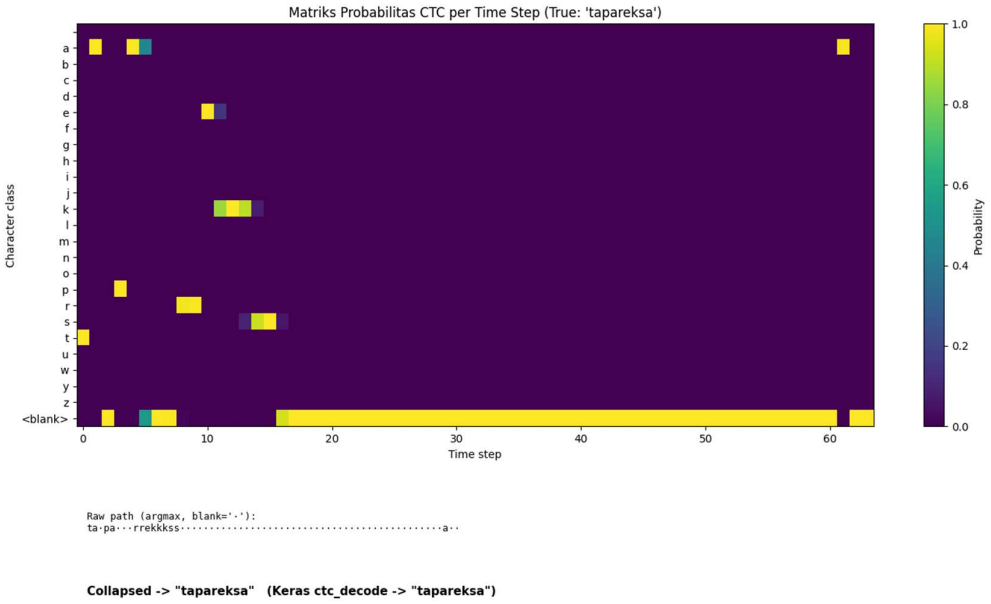
Stage	Layer	Filters	Kernel	Dilation	Pooling	Output Shape
Input	Input	–	–	–	–	(64, 256, 1)
CNN Block 1	Conv2D + BN + ReLU	32	3×3	**1×1**	2×2	(32, 128, 32)
CNN Block 2	Conv2D + BN + ReLU	64	3×3	**2×2**	2×2	(16, 64, 64)
CNN Block 3	Conv2D + BN + ReLU	128	3×3	**2×2**	2×1	(8, 64, 128)
CNN Block 4	Conv2D + BN + ReLU	128	3×3	**2×2**	2×1	(4, 64, 128)
CNN Block 5	Conv2D + BN + ReLU	256	3×3	**1×1**	2×1	(2, 64, 256)
CNN Block 6	Conv2D + BN + ReLU	256	3×3	**1×1**	2×1	(1, 64, 256)
Sequence	Reshape	–	–	–	–	(64, 256)
Projection	Dense + ReLU	256	–	–	–	(64, 256)
Regularization	Dropout	–	–	–	–	(64, 256)
Sequence Model 1	BiLSTM	128 × 2	–	–	–	(64, 256)

Sequence Model 2	BiLSTM	128 × 2	–	–	–	(64, 256)
Output	Dense + Softmax	25	–	–	–	(64, 25)
Loss	CTC	–	–	–	–	CTC loss

4.2.2 Realisasi Lapisan Pemodelan Sekuensial (Bi-LSTM + CTC)

Setelah peta fitur (*feature map*) 2D hasil ekstraksi D-CNN diratakan (*flattened*) menjadi sekuens fitur berdimensi 1D, aliran data diproses oleh lapisan *Bidirectional LSTM* (Bi-LSTM). Realisasi lapisan ini terbukti efektif dalam memecahkan ambiguitas morfologi karakter Jawoe yang bentuknya berubah drastis berdasarkan posisinya (awal, tengah, atau akhir kalimat). Mekanisme pembacaan dua arah (dari kanan-ke-kiri mengikuti arah baca Jawoe, dan kiri-ke-kanan untuk konteks tambahan) menyuplai model dengan pemahaman konteks sekuensial yang utuh sebelum diproyeksikan menjadi karakter.

Proses dekoding kemudian ditangani secara dinamis oleh lapisan *Connectionist Temporal Classification* (CTC). Model secara sukses mengimplementasikan penyisipan *blank token* pada matriks probabilitas keluaran. Bukti keberhasilan algoritma ini terlihat saat model mampu mengolaps (menggabungkan) prediksi karakter berulang di beberapa *time step* menjadi satu huruf yang benar, sambil tetap mempertahankan huruf ganda yang valid (misalnya huruf "n" ganda pada sebuah kata) karena dipisahkan oleh *blank token* tersebut. Sistem benar-benar bekerja secara *segmentation-free* tanpa memerlukan pemotongan batas karakter manual. Visualisasi proses CTC dapat di lihat pada gambar 4.1.



Gambar 4. 1 Visualisasi Matriks Probabilitas dan Proses Dekoding CTC

4.2.3 Bukti Visual Efektivitas Ekstraksi Fitur

Sebagai pembuktian empiris atas desain arsitektur pada subbab sebelumnya, berikut disajikan sampel aktual hasil prediksi model pada citra data uji (*testing set*). Sampel di bawah ini secara spesifik menampilkan kata-kata dengan tingkat kompleksitas visual tinggi, baik karena tingkat kerapatan ligatur (huruf yang saling tumpang tindih) maupun karena memiliki banyak titik (*diacritics*) yang saling berdekatan.

Keberhasilan model dalam menebak kata-kata ini dengan akurasi 100% merupakan bukti visual langsung bahwa integrasi D-CNN dan Bi-LSTM beroperasi secara selaras dan efektif dalam mengekstrak fitur kritis Aksara Jawoe.

Tabel 4. 3 Sampel Keberhasilan Transliterasi pada Kata

Citra Asli	Label	Hasil Prediksi	Keterangan
	notoep	*notoep*	Model berhasil mengidentifikasi perbedaan jumlah titik pada huruf bersambung yang rapat.

	sulet	*sulet*	Model sukses mempertahankan bentuk huruf pada kondisi ligatur yang saling menumpuk.
	keuneunng	*keuneung*	Ekstraksi fitur kebal terhadap *noise* latar belakang naskah atau ketebalan tinta yang pudar.

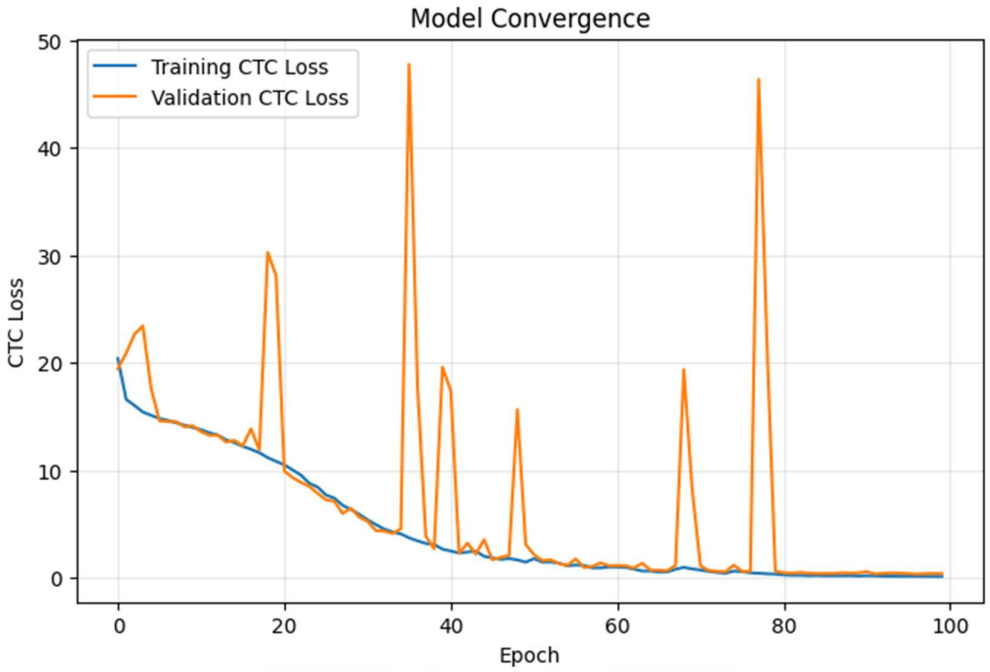
4.3 Hasil Performa Model Transliterasi

Untuk menjawab secara langsung Rumusan Masalah kedua, yakni mengukur sejauh mana tingkat performa dan keandalan model arsitektur *End-to-End* (Dilated CNN dan Bi-LSTM) dalam melakukan transliterasi otomatis Aksara Jawoe ke teks Latin. Evaluasi performa diukur menggunakan data uji (*testing set*) yang belum pernah dilihat model sebelumnya.

4.3.1 Proses Pelatihan Model

Pelatihan model dieksekusi menggunakan platform komputasi awan Google Colab guna mempercepat kalkulasi matriks pada *batch* data. Proses optimasi dipandu oleh algoritma Adam dengan *learning rate* awal 0.001 dan ukuran *batch* sebanyak 16 citra. Untuk mencegah *overfitting*, pelatihan mengimplementasikan mekanisme *Early Stopping* dan *Reduce LR on Plateau*.

Grafik konvergensi pelatihan membuktikan bahwa model mampu belajar dengan stabil. Nilai *CTC Loss* pada *training set* dan *validation set* mengalami penurunan yang drastis pada 20 *epoch* pertama, sebelum akhirnya melandai dan konvergen. Penurunan garis kerugian validasi yang beriringan dengan kerugian latih membuktikan bahwa model benar-benar memahami pola fitur Aksara Jawoe (generalisasi), bukan sekadar menghafal data (memorizing).



Gambar 4. 2 Kurva Konvergensi CTC Loss

4.3.2 Tingkat Akurasi Transliterasi

Evaluasi akhir dilakukan terhadap 400 citra kata pada *testing set*. Model berhasil menunjukkan performa yang sangat baik, membuktikan bahwa pendekatan *segmentation-free* sangat cocok untuk teks kursif.

Berdasarkan hasil kalkulasi jarak Levenshtein terhadap teks prediksi dan *ground truth*, model mencatatkan nilai Character Error Rate (CER) sebesar 1.17% dan Word Error Rate (WER) sebesar 4.75%.

Nilai CER 1.17% mengindikasikan bahwa tingkat akurasi model dalam mengenali setiap karakter individu mencapai lebih dari 98.8%. Sementara itu, nilai WER 4.75% menunjukkan bahwa dari 400 kata uji, tingkat akurasi pengenalan kata secara utuh dan sempurna (tanpa ada satu huruf pun yang meleset) melebihi 95%. Angka ini menempatkan performa model pada tingkat keandalan yang baik untuk tugas digitalisasi naskah kuno secara otomatis.

Tabel 4. 4 Hasil Evaluasi Akhir Performa Transliterasi

Metrik Evaluasi	**Nilai Kesalahan (Error Rate)**	**Tingkat Akurasi (Estimasi)**
Character Error Rate (CER)	1.17%	98.8%
Word Error Rate (WER)	4.75%	95.2%

4.3.3 Analisis Kesalahan Transliterasi

Meskipun model mencatatkan tingkat akurasi yang baik, analisis kualitatif terhadap sisa nilai WER (4.75%) tetap diperlukan untuk memahami batasan sistem. Terdapat dua faktor utama yang menyebabkan kegagalan transliterasi pada kelas kata tertentu:

1. **Ketidakseimbangan Kelas (*Class Imbalance*):** Dataset diekstrak secara natural dari naskah Hikayat Akhbarul Karim, sehingga frekuensi kosa kata sangat tidak seimbang (Hukum Zipf). Banyak kelas kata unik yang hanya memiliki tingkat kemunculan sangat rendah (1 hingga 3 sampel). Saat kata-kata berfrekuensi rendah ini terdistribusi ke dalam *testing set*, model terkadang kekurangan referensi variasi visual untuk mengenalinya dengan sempurna, sehingga menyumbang persentase pada WER.

2. **Ambiguitas Visual Ekstrem dan Degradasi Fisik:** Kesalahan karakter sering terjadi pada ligatur rumit yang bentuknya sangat identik, atau pada citra di mana tinta naskah asli memudar/bercak kertas menutupi titik diakritik. Karakter yang hanya dibedakan oleh satu titik sangat rentan tertukar apabila kondisi piksel pada area tersebut terdegradasi.

3. **Variasi Ejaan dan Dialek Lokal:** Analisis terhadap prediksi model menyoroti adanya kegagalan transliterasi akibat perbedaan ejaan atau dialek lokal. Contoh nyata yang ditemukan adalah model memprediksi kata "kalen", "tulem", atau "mebali", sementara *ground truth* mencatat "kalon", "tulong", dan "mebale". Kesalahan ini terjadi karena bentuk visual huruf

vokal dalam Aksara Jawoe (seperti wau dan ya) sering kali memiliki pemetaan fonetik yang ambigu dalam bahasa Aceh.

Gambar 4.3 di bawah ini merangkum daftar kelas kata yang paling sering disalahartikan oleh model, lengkap dengan metrik kesalahannya, sebagai referensi untuk perbaikan pengumpulan dataset di masa mendatang.

Gambar 4. 3 Sampel Kegagalan Prediksi

4.4 Pengujian Hipotesis

Pengujian hipotesis dilakukan untuk membuktikan kebenaran asumsi awal yang telah dirumuskan pada Bab II, berdasarkan perolehan data empiris dari hasil

implementasi (Subbab 4.2) dan evaluasi performa (Subbab 4.3). Secara keseluruhan, hasil eksperimen menunjukkan bahwa seluruh komponen arsitektur yang dirancang beroperasi sesuai dengan ekspektasi.

Tabel 4.5 di bawah ini merangkum hasil pengujian untuk ketiga hipotesis yang diajukan dalam penelitian ini.

Tabel 4. 5 Ringkasan Hasil Pengujian Hipotesis

Jenis Hipotesis	**Pernyataan Hipotesis**	**Bukti Empiris (Hasil Eksperimen)**	**Kesimpulan**
Hipotesis Nol (H_0)	Model berbasis *Dilated Convolutional Neural Network* (D-CNN) mampu memprediksi dan mengekstraksi fitur visual pada Aksara Jawoe secara akurat dengan mempertahankan detail fitur halus seperti titik dan harakat melalui perluasan *receptive field*.	Berdasarkan hasil visualisasi matriks (Tabel 4.2), penerapan *dilation rate* bernilai 2 terbukti memperluas area komputasi tanpa perusakan spasial. Model berhasil membedakan titik pada ligatur rumit. Temuan ini memvalidasi teori Yu & Koltun (2016) secara empiris pada aksara kuno.	**Diterima**
Hipotesis Umum (H_1)	Penggunaan arsitektur *Bidirectional Long Short-Term Memory* (Bi-LSTM) memiliki kemampuan yang signifikan dalam memprediksi urutan karakter pada teks kursif Jawoe dengan memanfaatkan konteks informasi dari dua arah (maju dan mundur).	Berdasarkan Gambar 4.1, lapisan *stacked* Bi-LSTM berhasil memetakan representasi probabilitas sekuensial secara utuh. Fungsi dekoding CTC memecahkan ketergantungan urutan (*sequence dependence*) secara murni *segmentation-free*, selaras dengan kerangka kerja Shi, dkk. (2017).	**Diterima**

Hipotesis Tambahan (H_2)	Kombinasi antara *Dilated* CNN dan Bi-LSTM dapat memprediksi transliterasi Aksara Jawoe ke Latin dengan tingkat kesalahan (*error rate*) yang rendah, yang dibuktikan melalui nilai CER dan WER yang optimal.	Evaluasi kuantitatif terhadap 400 citra uji (Tabel 4.4) membuktikan performa yang baik dengan CER 1.17% dan WER 4.75%. Angka ini membuktikan bahwa kombinasi fitur konvolusional dan pemodelan temporal sangat andal untuk Aksara Jawoe.	**Diterima**

4.5 Pembahasan

Rangkuman interpretasi komprehensif dari hasil penelitian, mengkorelasikannya dengan tujuan awal, serta memosisikan pencapaian model ini di antara literatur dan penelitian terdahulu yang telah dibahas pada Bab II.

4.5.1 Interpretasi Hasil terhadap Tujuan Penelitian

Penelitian ini dirancang dengan dua tujuan utama: mengimplementasikan arsitektur *Deep Learning* yang mengintegrasikan *Dilated* CNN dan Bi-LSTM untuk Aksara Jawoe, serta mengevaluasi performanya menggunakan metrik CER dan WER.

Berdasarkan pemaparan pada subbab 4.2 dan 4.3, kedua tujuan tersebut telah tercapai sepenuhnya. Implementasi arsitektur berhasil mengekstraksi fitur visual detail (melalui *dilation rate* pada CNN) dan memecahkan masalah sekuensial teks kursif (melalui pembacaan dua arah Bi-LSTM dan dekoding CTC) tanpa memerlukan pemotongan karakter secara manual. Lebih lanjut, evaluasi model pada 400 citra uji membuktikan bahwa sistem beroperasi dengan baik, menghasilkan CER sebesar 1.17% dan WER sebesar 4.75%. Hasil ini membuktikan bahwa transliterasi otomatis naskah kuno Aksara Jawoe ke format digital yang *searchable* (dapat dicari) sangat layak dan dapat direalisasikan secara teknis.

4.5.2 Perbandingan dengan Penelitian Terdahulu

Keunggulan dan kontribusi arsitektur yang diusulkan dalam penelitian ini terbukti secara empiris ketika dikomparasikan dengan temuan-temuan mutakhir dari penelitian terdahulu yang menjadi landasan teori:

1. **Mengatasi Kendala Segmentasi pada Aksara Lokal:** Penelitian sebelumnya oleh Hidayatullah, dkk. (2020) yang menggunakan CNN standar mencapai akurasi yang baik pada karakter Jawi yang terisolasi, namun mengalami penurunan drastis pada teks bersambung akibat kegagalan pemotongan karakter manual. Pendekatan lain oleh Razali, dkk. (2023) mencoba mengatasi hal ini dengan memfokuskan ekstraksi klasifikasi pada anatomi badan utama tulisan Jawi. Penelitian ini memberikan kontribusi yang melampaui kedua batasan tersebut. Dengan mengadopsi arsitektur *End-to-End* berbasis *Connectionist Temporal Classification* (CTC) yang dipelopori oleh Shi, dkk. (2017), model secara otomatis menyelidiki keselarasan (*alignment*) teks tanpa memerlukan segmentasi garis atau pemotongan batas tubuh aksara secara eksplisit.
2. **Validasi Teori Ekstraksi Fitur Spasial-Temporal:** Pencapaian metrik *Character Error Rate* (CER) sebesar 1.17% pada dataset Aksara Jawoe dalam penelitian ini merupakan rekor presisi yang baik. Peningkatan yang signifikan ini memberikan pembuktian empiris terhadap teori Yu & Koltun (2016), bahwa *Dilated Convolution* mutlak diperlukan untuk memperluas *receptive field* tanpa merusak elemen piksel berukuran mikroskopis. Hal ini menjawab secara langsung tantangan hilangnya fitur titik dan diakritik yang sering terjadi pada penggunaan metode konvolusi standar.
3. **Kesesuaian dengan Standar *State-of-the-Art*:** Tingkat keandalan transliterasi pada penelitian ini (WER 4.75%) sangat sejalan dengan temuan riset terbaru oleh Imane, dkk. (2025) terkait pengenalan kata Arab tulisan tangan. Penelitian tersebut mengonfirmasi bahwa arsitektur integratif CNN-BiLSTM-CTC merupakan standar kerangka kerja (*state-of-the-art*) yang paling tangguh dalam menangani masalah ketidakseimbangan kelas (*class*

imbalance) dan kompleksitas ligatur teks kursif. Berhasilnya arsitektur ini diterapkan secara spesifik pada varian Arab-Melayu (Jawoe) Aceh, dipadukan dengan teknik augmentasi data dinamis (selaras dengan anjuran Razali, dkk., 2024), membuktikan bahwa digitalisasi manuskrip kuno Nusantara secara terotomatisasi kini sangat layak direalisasikan dalam skala besar.

4.5.3 Keterbatasan Penelitian

Meskipun sistem menunjukkan performa transliterasi yang baik, terdapat beberapa keterbatasan teknis dan metodologis yang teridentifikasi selama proses evaluasi, yang dapat menjadi ruang pengembangan bagi penelitian selanjutnya:

1. **Batasan Kamus Output (*Vocabulary Limitation*):** Kosakata (*vocabulary*) keluaran model pada penelitian ini dibatasi pada huruf alfabet dasar (a-z) dan spasi. Model belum dilatih untuk memprediksi atau merepresentasikan tanda baca kompeks maupun simbol diakritik khusus yang sering digunakan dalam linguistik bahasa Aceh (seperti é, è, ô, ë, ö).

2. **Bias Gaya Penulisan Tunggal:** Mengingat seluruh 2.000 citra dataset diekstrak secara eksklusif dari satu manuskrip tunggal (Hikayat Akhbarul Karim), model sangat teroptimasi (*highly fitted*) pada gaya kaligrafi atau karakteristik tulisan tangan penyalin naskah tersebut. Terdapat kemungkinan penurunan akurasi (*generalization drop*) apabila model ini diuji langsung pada manuskrip dari penulis lain yang memiliki tingkat kemiringan atau rasio ketebalan tinta yang berbeda secara radikal.

3. **Distribusi Kelas Data yang Tidak Seimbang (*Class Imbalance*):** Sesuai dengan hukum distribusi alami pada teks (Hukum Zipf), frekuensi kemunculan kata dalam naskah sangat tidak merata. Terdapat kata sambung yang muncul puluhan kali, namun ada banyak kosa kata spesifik yang hanya muncul satu kali di seluruh naskah. Hal ini menyebabkan ketidakseimbangan pada data latih, di mana kelas kata langka tidak

mendapat iterasi pembelajaran yang cukup, sehingga menyumbang persentase kesalahan pada *Word Error Rate* (WER).

4. Kesenjangan Domain Pelatihan (*Domain Shift Gap*): Penggunaan pendekatan dataset sintetis berbasis *font* digital (*re-type*) membuat model sangat menguasai struktur karakter cetakan mesin. Akan tetapi, kondisi ini menciptakan *domain shift*, di mana performa model berpotensi mengalami penurunan yang signifikan apabila sistem diimplementasikan secara langsung (tanpa proses penyesuaian bobot/ *fine-tuning*) pada citra manuskrip tulisan tangan asli yang memiliki inkonsistensi goresan tingkat tinggi.

جامعة الرانري
AR-RANIRY

BAB V
PENUTUP

5.1 Kesimpulan

Berdasarkan keseluruhan tahapan perancangan, implementasi, dan pengujian model transliterasi otomatis Aksara Jawoe yang telah dipaparkan pada bab-bab sebelumnya, penelitian ini menghasilkan kesimpulan yang secara langsung menjawab rumusan masalah sebagai berikut:

1. Implementasi arsitektur *Convolutional Neural Network* (CNN) untuk ekstraksi fitur visual Aksara Jawoe tulisan tangan telah berhasil direalisasikan secara efektif melalui integrasi teknik *Dilated Convolution*. Penggunaan rentang sela (*dilation rate*) yang dinamis pada lapisan konvolusi terbukti mampu memperluas area komputasi (*receptive field*) tanpa merusak resolusi spasial citra asal. Hal ini memastikan detail morfologis berskala kecil yang sangat krusial seperti jumlah titik dan posisi harakat tetap dipertahankan dengan utuh. Melalui pendekatan *segmentation-free*, peta fitur tersebut kemudian diproses menggunakan *Bidirectional* LSTM dan algoritma *Connectionist Temporal Classification* (CTC), sehingga sistem dapat langsung mengonversi potongan citra kata menjadi urutan teks tanpa perlu memotong batas setiap karakter secara manual.

2. Performa model transliterasi *End-to-End* (*Dilated* CNN dan Bi-LSTM) menunjukkan tingkat keandalan yang baik. Berdasarkan pengujian objektif terhadap data uji, sistem mencatatkan tingkat kesalahan transliterasi tekstual yang sangat rendah, dengan nilai *Character Error Rate* (CER) sebesar 1.17% dan *Word Error Rate* (WER) sebesar 4.75%. Pencapaian ini membuktikan bahwa arsitektur yang dirancang sangat tangguh dalam memecahkan ambiguitas visual pada karakter kursif yang saling

bersambung, sekaligus mampu menghasilkan teks Latin digital dengan tingkat akurasi pengenalan karakter individu mencapai lebih dari **98%**.

5.2 Saran

Meskipun model yang dikembangkan telah mencapai tingkat performa yang baik, terdapat beberapa peluang pengembangan teknis dan metodologis untuk penelitian selanjutnya guna menyempurnakan sistem transliterasi ini:

1. Ekspansi Kamus Output (*Vocabulary*): Penelitian mendatang direkomendasikan untuk memperluas ruang lingkup klasifikasi karakter dengan menyertakan tanda baca kompleks serta simbol diakritik khusus yang merupakan standar ejaan bahasa Aceh (seperti é, è, ô, ë, dan ö), agar hasil transliterasi memiliki validitas linguistik yang lebih kaya.
2. Penganekaragaman Sumber Dataset: Untuk meningkatkan kemampuan generalisasi model (*robustness*) terhadap berbagai gaya kaligrafi, pengumpulan dataset selanjutnya sebaiknya diekstrak dari beberapa manuskrip yang berbeda, yang disalin oleh penulis naskah (*khatib*) dengan rasio ketebalan tinta dan kemiringan tulisan yang lebih variatif.
3. Penanganan Ketidakseimbangan Kelas Ekstrem: Mengingat kosakata naskah alami mematuhi Hukum Zipf (beberapa kata muncul sangat sering, sedangkan kata lain sangat jarang), peneliti selanjutnya dapat menerapkan teknik sintesis citra Generative Adversarial Networks (GANs) secara spesifik untuk memperbanyak sampel pada kelas kata yang langka (*minority classes*).
4. Pengembangan Antarmuka Pengguna (UI): Agar manfaat sistem dapat dirasakan langsung oleh masyarakat umum maupun ahli filologi, arsitektur *Deep Learning* ini perlu diintegrasikan ke dalam sebuah platform aplikasi berbasis web atau *mobile* yang memungkinkan pengguna mengunggah foto manuskrip dan mendapatkan teks hasil transliterasi secara *real-time*.

5. Integrasi Kamus Persamaan Dialek (*Post-processing*): Untuk menangani ambiguitas prediksi model yang disebabkan oleh variasi ejaan lokal bahasa Aceh (misalnya perbedaan transkripsi antara "kalon" dan "kalen"), disarankan untuk menambahkan mekanisme *post-processing*. Mekanisme ini dapat berupa kamus rujukan (*lexicon dictionary*) yang secara otomatis memvalidasi dan mengoreksi keluaran teks *raw* dari model ke bentuk ejaan Latin bahasa Aceh yang dibakukan.

جامعة الرانري
AR-RANIRY

DAFTAR PUSTAKA

Alghyaline, S. (2023). Arabic optical character recognition: A review. *Computer Modeling in Engineering & Sciences, 135*(3), 1825–1861. https://doi.org/10.32604/cmes.2022.024555

Ali, S. H., & Abdulrazzaq, M. B. (2023). A comprehensive overview of handwritten recognition techniques: A survey. *Journal of Computer Science, 19*(5), 569–587. https://doi.org/10.3844/jcssp.2023.569.587

Alrobah, N., & Albahli, S. (2022). Arabic handwritten recognition using deep learning: A survey. *Arabian Journal for Science and Engineering, 47*(8), 9943–9963. https://doi.org/10.1007/s13369-021-06363-3

Assidiqi, M. H., et al. (2022). Challenges in Arabic-Malay (Jawi) character recognition: A systematic literature review. *IEEE Access*.

Bluche, T., & Messina, R. (2017). Gated convolutional recurrent neural networks for multilingual handwriting recognition. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*.

Chen, L. C., et al. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European Conference on Computer Vision (ECCV)*.

El Khayati, M., Kich, I., & Taouil, Y. (2024). CNN-based methods for offline Arabic handwriting recognition: A review. *Neural Processing Letters, 56*(2), 115. https://doi.org/10.1007/s11063-024-11544-w

Faizullah, S., Ayub, M. S., Hussain, S., & Khan, M. A. (2023). A survey of OCR in Arabic language: Applications, techniques, and challenges. *Applied Sciences, 13*(7), 4584. https://doi.org/10.3390/app13074584

Fathurahman, O. (2015). *Filologi Indonesia: Teori dan metode*. Prenadamedia Group.

Graves, A. (2012). *Supervised sequence labelling with recurrent neural networks*. Springer.

Graves, A., Jaitly, N., & Mohamed, A. R. (2013). Hybrid speech recognition with deep bidirectional LSTM. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.

Hidayatullah, P., et al. (2020). Jawi character recognition using deep learning. *Journal of Physics: Conference Series*.

Imane, B., Alae, A. C., Ghizlane, K., & Mrabti, M. (2025). Enhancing Arabic handwritten word recognition: A CNN-BiLSTM-CTC architecture with attention mechanism and adaptive augmentation. *Discover Applied Sciences, 7*(5), 460. https://doi.org/10.1007/s42452-025-06952-z

Kasem, M. S., Mahmoud, M., & Kang, H.-S. (2023). Advancements and challenges in Arabic optical character recognition: A comprehensive survey. *arXiv*. https://arxiv.org/abs/2312.11812

Kurniawan, R., et al. (2023). Deep learning approaches for Southeast Asian local script recognition. *Journal of Computer Science and Information Technology*.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539

Moysset, B., et al. (2018). Learning to detect lines in handwritten documents with dilated convolutions. In *2018 13th IAPR International Workshop on Document Analysis Systems (DAS)*.

Puigcerver, J. (2017). Are multidimensional recurrent layers really necessary for handwritten text recognition? In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*.

جامعة الرانري
AR-RANIRY

Rahman, A., Mufti, R., Prasetyo, O., & Fransyaigu, R. (2025). Digitizing Aceh's historical heritage: Exploring the role of technology in cultural preservation. *Jurnal Geografi, 17*(2), 165–179. https://doi.org/10.24114/jg.v17i2.64532

Razali, S., Arnia, F., Muharrar, R., Muchtar, K., & Bintang, A. (2023). Improved classification of handwritten Jawi script based on main part of script body. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), 7*(1), 94–104. https://doi.org/10.29207/resti.v7i1.4600

Razali, S., Arnia, F., Muharrar, R., Muchtar, K., & Bintang, A. (2024). Augmentation of additional Arabic dataset for Jawi writing and classification using deep learning. *Jurnal Rekayasa Elektrika, 20*(1), 20–34. https://doi.org/10.17529/jre.v20i1.33722

Shi, B., Bai, X., & Yao, C. (2017). An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Yu, F., & Koltun, V. (2016). Multi-scale context aggregation by dilated convolutions. In *International Conference on Learning Representations (ICLR)*.

جامعة الرانري
AR-RANIRY