Pengenalan Tulisan Tangan Aksara Bali Menggunakan Algoritma CRNN

I Kadek Bisma Yoga[a1], Cokorda Pramartha[a2]

Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia
[1]bismayoga432@gmaill.com
[2]cokorda@unud.ac.id

Abstract

This research proposes the development of a handwritten Balinese script recognition system using the Convolutional Recurrent Neural Network (CRNN) algorithm as a vital step in preserving this intangible cultural heritage from the threat of obsolescence in the digital era. Given the inherent complexity and variability of handwritten Balinese script, which distinguishes it from printed text, a deep learning approach is essential. CRNN was chosen for its ability to integrate the strengths of CNN in extracting spatial visual features with the power of RNN (specifically BiLSTM) in modeling sequential dependencies. Primary handwritten data was meticulously validated by a Balinese language expert, then processed through grayscale conversion, pixel normalization, and resizing for standardization. The model was constructed with convolutional layers, recurrent BiLSTM layers, and a Connectionist Temporal Classification (CTC) transcription layer, which is effective in translating sequential features into character labels. Performance evaluation of the model using the Character Error Rate (CER) on separate test data showed an average accuracy of 89.9%. These results significantly affirm the great potential of CRNN in supporting the digitalization efforts of Balinese script, as well as facilitating its broader integration into modern environments.

Keywords: *Balinese Script, Handwritten Recognition, CRNN, Deep Learning, Character Error Rate*

1. Pendahuluan

Aksara Bali, sebagai salah satu warisan budaya takbenda Indonesia, memiliki posisi sentral dalam menjaga identitas dan nilai-nilai luhur masyarakat Bali. Aksara ini bukan hanya sekadar sistem penulisan, melainkan juga cerminan kekayaan historis dan filosofis yang termanifestasi dalam naskah-naskah kuno (lontar), ritual keagamaan, serta beragam bentuk seni. Namun, di tengah pesatnya arus globalisasi dan dominasi aksara Latin dalam ranah digital, keberadaan Aksara Bali menghadapi tantangan kepunahan. Oleh karena itu, upaya pelestarian dan adaptasi teknologi menjadi krusial agar aksara ini tidak tergerus zaman dan tetap relevan dalam konteks *modern* [1]. Salah satu upaya adaptasi yang signifikan adalah melalui digitalisasi dan pengembangan teknologi Pengenalan Karakter Optis (OCR). Meskipun teknologi OCR untuk aksara Latin telah mencapai tingkat akurasi yang sangat tinggi yaitu 95% [2], implementasinya untuk aksara daerah yang kompleks seperti Aksara Bali masih memerlukan eksplorasi mendalam, khususnya pada variasi tulisan tangan. Pengenalan tulisan tangan secara inheren lebih menantang dibandingkan tulisan cetak karena adanya variabilitas gaya penulisan individu yang sangat besar, termasuk pada naskah yang sudah terdegradasi[1], komponen-komponen Aksara Bali memiliki struktur yang unik dan berbeda dari sistem penulisan modern [1]. Proses pengenalannya secara otomatis menghadapi berbagai kendala, seperti kemiripan bentuk antar beberapa karakter, variasi ketebalan garis, dan inkonsistensi goresan dari satu penulis ke penulis lainnya. Metode konvensional dalam pengenalan pola seringkali tidak mampu mengekstraksi fitur yang cukup kuat untuk mengatasi variasi ini contohnya yaitu menggunakan algoritma SVM dengan akurasi 82,1% [3]. Oleh karena itu, diperlukan pendekatan komputasi yang lebih canggih yang mampu mempelajari fitur-fitur relevan secara otomatis dari data citra. Dalam konteks ini,

Algoritma Deep Learning, khususnya *Convolutional Recurrent Neural Network* (CRNN), telah menunjukkan kapabilitas luar biasa dalam berbagai tugas pengenalan sekuensial visual, termasuk pengenalan teks tulisan tangan [4]. Arsitektur CRNN secara efektif menggabungkan keunggulan *Convolutional Neural Network* (CNN) dalam mengekstraksi fitur spasial dari gambar dengan kekuatan Recurrent Neural Network (RNN) dalam memodelkan dependensi sekuensial dari fitur-fitur tersebut. Kombinasi ini menjadikannya sangat ideal untuk tugas pengenalan tulisan tangan Aksara Bali, di mana pengenalan karakter tidak hanya bergantung pada bentuk visualnya semata, tetapi juga pada konteks sekuensial dari fitur-fitur yang membentuk karakter tersebut. Penelitian ini diharapkan dapat berkontribusi pada upaya digitalisasi dan pelestarian Aksara Bali, membuka jalan bagi integrasinya yang lebih luas di era digital.

2. Metode Penelitian

Penelitian ini dilaksanakan melalui empat tahap utama:

a. Pengumpulan data: Proses akuisisi data mentah yang diperlukan.
b. Pemrosesan data: Tahap persiapan dan transformasi data untuk analisis
c. Pembuatan model: Pengembangan arsitektur dan implementasi model komputasi.
d. Evaluasi model: Penilaian komprehensif terhadap kinerja dan keefektifan model yang dibuat.

Gambar 1. Diagram Alur Penelitian

Alur detail dari tahapan penelitian ini, seperti yang diilustrasikan pada Gambar 1, akan dijelaskan secara rinci di bagian selanjutnya.

2.1 Pengumpulan Data

Sumber data untuk penelitian ini diperoleh melalui akuisisi data primer yang berasal dari hasil tugas menulis angka Aksara Bali oleh tiga orang siswa, di mana seluruh citra tulisan tangan yang terkumpul kemudian melewati sebuah tahap validasi yang krusial oleh seorang guru Bahasa Bali untuk memastikan keabsahan dan kesesuaian setiap karakter dengan kaidah penulisan yang benar. Proses verifikasi ahli ini menjadi filter utama, di mana hanya data yang dinyatakan valid dan jelas secara visual yang akan digunakan pada tahap selanjutnya, sedangkan semua data yang mengandung kesalahan penulisan atau memiliki tingkat ambiguitas yang tinggi akan disisihkan secara cermat demi menjaga integritas dan kualitas dataset. Selanjutnya, lembar tugas yang telah lolos validasi tersebut didigitalisasi dengan cara dipindai (*scanned*) ke dalam format digital beresolusi tinggi untuk memastikan setiap detail goresan dapat ditangkap secara optimal. Dari gambar hasil pindaian yang utuh, dilakukan proses ekstraksi untuk memisahkan dan mengisolasi setiap citra angka Aksara Bali menjadi file gambar individual, yang pada tahap akhir masing-masing file gambar tersebut diberi label secara sistematis sesuai dengan nilai angkanya. Setelah melalui proses ini, dataset yang terkumpul memiliki ukuran data latih (Train Data Size) sebanyak 519 dan data validasi (Valid Data Size) sebanyak 130, dengan total keseluruhan data sebesar 649 detail pembagian penulisan dataset bisa dilihat pada tabel 1.

Tabel 1. Deskripsi Data

Penulis	Variasi Aksara	Format	Jumlah
Penulis 1	Baseline	JPG	320
Penulis 2	Baseline	JPG	98
Penulis 3	Baseline	JPG	231

2.2 Pemrosesan Data

Setiap citra dari dataset melewati serangkaian transformasi yang telah didefinisikan sebelum siap digunakan untuk pelatihan. Berdasarkan parameter dan fungsi yang telah ditetapkan, tahapan ini meliputi:

Gambar 2. Dataset Sebelum Diproses

a. Dari data yang belum di proses seperti yang telihat pada gambar 2 dilakukan pembacaan dan Konversi Grayscale untuk setiap file citra dibaca dari direktori penyimpanannya dan di-decode menjadi format tensor. Pada tahap ini, citra juga langsung dikonversi menjadi format *grayscale* (kanal tunggal atau N_CHANNELS = 1). Langkah ini bertujuan untuk mereduksi kompleksitas data dengan menghilangkan informasi warna dan hanya mempertahankan informasi luminans (intensitas piksel) yang esensial untuk pengenalan bentuk dan struktur karakter.
b. Normalisasi nilai piksel yang semula berada dalam rentang integer [0, 255], dinormalisasi menjadi rentang *floating-point* [0.0, 1.0]. Normalisasi ini merupakan praktik standar dalam pelatihan model *deep learning* karena dapat membantu menstabilkan proses optimisasi dan mempercepat laju konvergensi model selama pelatihan.
c. Pengubahan ukuran (Resizing) untuk memastikan seluruh data input memiliki dimensi yang seragam sebelum dimasukkan ke lapisan konvolusional, setiap citra diubah ukurannya (*resize*) menjadi dimensi yang telah ditentukan, yaitu tinggi 160 piksel dan lebar 320 piksel (IMAGE_HEIGHT = 160, IMAGE_WIDTH = 320).

Gambar 3. Dataset Setelah Diproses

Hasil akhir dari serangkaian proses ini adalah sebuah tensor berformat (320, 160, 1) dengan tipe data *float32* sehingga tampilan dataset terlihat seperti gambar 3. yang sepenuhnya siap untuk dijadikan input bagi model CRNN yang akan dibangun.

2.3 Pembuatan Model

Proses pembuatan model untuk pengenalan aksara Bali ini melibatkan perancangan dan implementasi arsitektur deep learning yang dikenal sebagai Convolutional Recurrent Neural Network (CRNN). CRNN adalah arsitektur hibrida yang menggabungkan kemampuan dari beberapa komponen kunci. Pertama, Convolutional Neural Network (CNN) digunakan untuk mengekstrak fitur visual penting dari gambar masukan, mengubahnya menjadi representasi data yang lebih abstrak yang disebut peta fitur[7]. Selanjutnya, Recurrent Neural Network (RNN), khususnya Long Short-Term Memory (LSTM) atau Bidirectional LSTM (BiLSTM), memproses peta fitur ini sebagai serangkaian data berurutan untuk menangkap hubungan kontekstual dan dependensi antar fitur dalam sekuens tersebut. Tahap terakhir dalam arsitektur CRNN adalah

lapisan transkripsi yang memanfaatkan Connectionist Temporal Classification (CTC). CTC bertanggung jawab menerjemahkan *output* dari RNN menjadi label sekuens akhir yang koheren, secara cerdas mengatasi masalah penyesuaian panjang dan prediksi berulang [6]. Berkat arsitektur terintegrasi ini, CRNN sangat efektif dalam berbagai aplikasi praktis seperti pengenalan teks pada gambar (scene text recognition), Optical Character Recognition (OCR), pengenalan tulisan tangan, dan identifikasi plat nomor kendaraan.

2.4 Evaluasi Model

Character Error Rate (CER) adalah metrik evaluasi standar yang digunakan untuk mengukur kinerja sistem pengenalan sekuens, khususnya dalam konteks pengenalan karakter seperti OCR dan pengenalan tulisan tangan [5]. CER menghitung persentase kesalahan pada tingkat karakter antara urutan karakter yang diprediksi oleh model dan urutan karakter referensi (ground truth) yang sebenarnya. Secara matematis, CER dihitung menggunakan rumus *Levenshtein distance* (edit distance) yang memperhitungkan jumlah minimum operasi (penyisipan, penghapusan, atau substitusi karakter) yang diperlukan untuk mengubah urutan prediksi menjadi urutan referensi. Rumus CER adalah sebagai berikut:

$$CER = \frac{S + D + I}{N} \quad (1)$$

di mana:

a. S adalah jumlah substitusi (karakter yang salah diprediksi).
b. D adalah jumlah penghapusan (karakter dalam referensi yang tidak ada di prediksi).
c. I adalah jumlah penyisipan (karakter di prediksi yang tidak ada di referensi).
d. N adalah jumlah total karakter dalam urutan referensi.

Semakin rendah nilai CER, semakin baik kinerja model pengenalan. Metrik ini memberikan gambaran yang jelas mengenai akurasi model dalam memprediksi setiap karakter individual dalam sebuah sekuens, menjadikannya alat yang efektif untuk membandingkan kinerja berbagai algoritma pengenalan tulisan tangan.

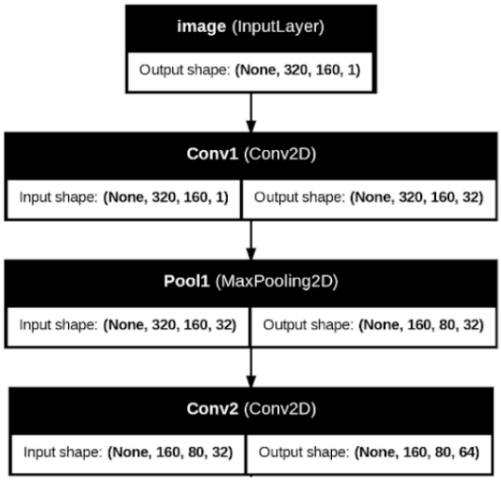
3. Hasil dan Diskusi

Pembuatan model pengenalan tulisan tangan Aksara Bali menggunakan algoritma Convolutional Recurrent Neural Network (CRNN) melibatkan beberapa tahapan kunci dalam perancangan dan implementasi arsitektur jaringan. Proses ini dimulai dengan persiapan data yang telah dijelaskan sebelumnya, yaitu dataset citra karakter Aksara Bali yang telah dilabeli dan dibagi menjadi data latih serta data validasi.

Arsitektur CRNN akan dibangun dengan mengintegrasikan tiga komponen utama:

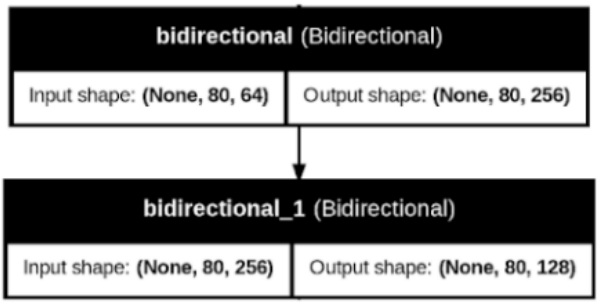
a. Lapisan Konvolusional (Convolutional Layers): Bagian ini berfungsi sebagai ekstraktor fitur visual dari citra masukan. Serangkaian lapisan konvolusional digunakan untuk secara progresif mengekstraksi fitur-fitur hierarkis dari citra, mulai dari tepi, tekstur, hingga pola yang lebih kompleks. Output dari lapisan konvolusional ini adalah peta fitur (feature map) yang merepresentasikan informasi spasial dari citra susunan visualisasi dari lapisan konvolusionial akan terlihat seperti pada gambar 4.

p-ISSN: 2986-3929
e-ISSN: 3032-1948



Gambar 4. Lapisan Konvolusional

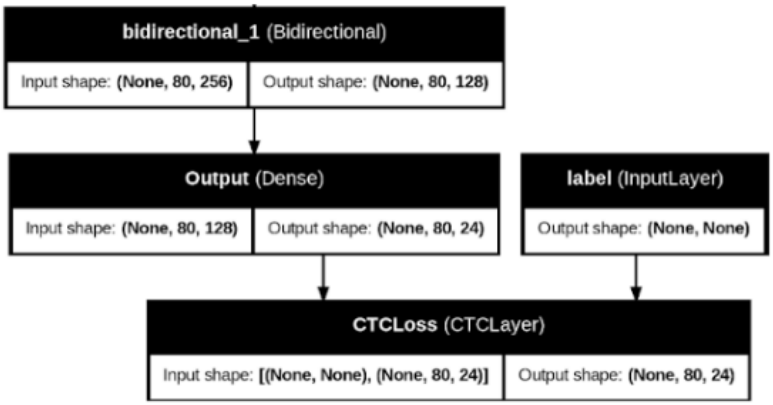
b. Lapisan Rekuren (Recurrent Layers): Peta fitur yang dihasilkan oleh CNN kemudian diubah menjadi urutan fitur yang akan diumpankan ke lapisan rekuren. Umumnya, lapisan Bi-directional Long Short-Term Memory (BiLSTM) digunakan di sini. BiLSTM sangat efektif dalam menangkap dependensi kontekstual jangka panjang dalam urutan, baik dari arah maju maupun mundur. Ini krusial untuk pengenalan tulisan tangan di mana konteks antar goresan atau bagian karakter sangat penting. Output dari lapisan BiLSTM adalah representasi sekuensial dari fitur-fitur yang telah diproses yang divisualisasikan seperti pada gambar 5.



Gambar 5. Lapisan Rekuren

c. Lapisan Transkripsi (Transcription Layer): Tahap terakhir adalah lapisan Connectionist Temporal Classification (CTC). CTC bertanggung jawab untuk menerjemahkan output sekuensial dari lapisan rekuren menjadi urutan label karakter yang sebenarnya. Seperti pada informasi Gambar 6 CTC memungkinkan model untuk belajar penjajaran (alignment) antara fitur yang diekstraksi dan label target tanpa memerlukan segmentasi eksplisit setiap karakter dalam citra, yang sangat menguntungkan untuk data tulisan tangan yang bervariasi.

p-ISSN: 2986-3929
e-ISSN: 3032-1948



Gambar 6. Lapisan Connectionist Temporal Classification (CTC)

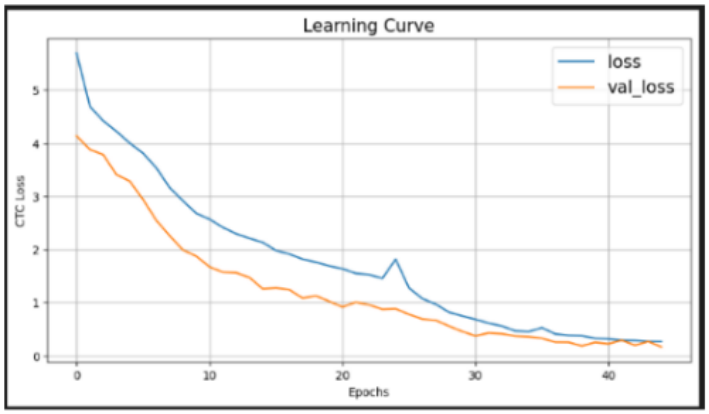
Secara umum code program pembuatan model sebagai berikut:

```python
# Image input layer
input_image = layers.Input(shape=(IMAGE_WIDTH, IMAGE_HEIGHT, N_CHANNELS), name="image", dtype=tf.float32)

# Label input layer
input_label = layers.Input(shape=(None, ), dtype=tf.float32, name="label")

# Convolutional network part for capturing the image features.
conv_1 = layers.Conv2D(
    filters=32,
    kernel_size=3,
    activation='relu',
    padding='same',
    kernel_initializer='he_normal',
    name="Conv1"
)(input_image)

pool_1 = layers.MaxPool2D(pool_size=(2,2), name="Pool1")(conv_1)

conv_2 = layers.Conv2D(
    filters=64,
    kernel_size=3,
    activation='relu',
    padding='same',
    kernel_initializer='he_normal',
    name="Conv2"
)(pool_1)

image_embedding = layers.MaxPool2D(pool_size=(2,2), name="Pool2")(conv_2)

# Reshaped inputs carefully for making it compatible for the recurrent neural network.
reshaped = layers.Reshape(target_shape=(IMAGE_WIDTH//4, (IMAGE_HEIGHT//4)*64), name="Reshape")(image_embedding)
encoding = layers.Dense(
    64,
    activation='relu',
    kernel_initializer='he_normal',
    name="DenseMapping"
)(reshaped)
encoding = layers.Dropout(0.2, name="DenseDrop")(encoding)

# Pass these representations to a recurrent neural network
rnn_1 = layers.Bidirectional(layers.LSTM(128, return_sequences=True, dropout=0.25))(encoding)
text_embedding = layers.Bidirectional(layers.LSTM(64, return_sequences=True, dropout=0.25))(rnn_1)

# Output layer for model predictions
output_preds = layers.Dense(units=len(char_to_num.get_vocabulary())+1, activation='softmax', name="Output")(text_embedding)

# CTC loss layer
Output = CTCLayer(name="CTCLoss")(input_label, output_preds)
```

Gambar 7. Code Pembuatan Model

Model akan dilatih menggunakan data latih dengan mengoptimalkan fungsi kerugian CTC Loss. Dengan susunan code yang tedapat pada cuplikan gambar 7 Proses pelatihan melibatkan iterasi berulang (epoch) 70 di mana model memproses data, menghitung kerugian, dan memperbarui bobotnya untuk meminimalkan CTC Loss. Kinerja model akan dipantau secara berkala pada data validasi untuk mencegah *overfitting* dan menyesuaikan *hyperparameter* jika diperlukan.

p-ISSN: 2986-3929
e-ISSN: 3032-1948



Gambar 8. Grafik Kinerja model

Kinerja model selama pelatihan tidak terjadi *overfitting* seperti terlihat pada gambar 8. Setelah proses pelatihan, kinerja model CRNN dievaluasi secara menyeluruh menggunakan data *testing* yang terpisah dan belum pernah dilihat oleh model selama proses pelatihan maupun validasi. Tujuan utama evaluasi ini adalah untuk mengukur seberapa baik model dapat menggeneralisasi dan memprediksi karakter Aksara Bali dari tulisan tangan yang benar-benar baru. Metrik utama yang digunakan untuk evaluasi ini adalah Character Error Rate (CER), yang memberikan indikasi langsung tentang tingkat kesalahan pada tingkat karakter.

Gambar 9. Data Pengujian

p-ISSN: 2986-3929
e-ISSN: 3032-1948

Tabel 2. Hasil Pengujian

Nama Data	Hasil Prediksi	Label Sebenarnya	Decode Hasil Prediksi	Nilai CER	Akurasi
datates1.jpg	mgl	mgl	magala	0	100%
datates2.jpg	crk	crk	caraka	0	100%
datates3.jpg	klkl	klkl	kalakala	0	100%
datates4.jpg	dho	dhr	daho	0,333	66,6%
datates5.jpg	dtt	dtt	datata	0	100%
datates6.jpg	mkk	mrk	makaka	0,333	66,6%
datates7.jpg	zpl	zhl	nyapala	0,333	66,6%
datates8.jpg	k	k	ka	0	100%
datates9.jpg	phn	phn	pahana	0	100%
datates10.jpg	skl	skl	sakala	0	100%
Rata Rata Akurasi					89,9%

Berdasarkan hasil pengujian pada Tabel 2, terhadap data pengujian yang bisa dilihat pada gambar 9 model CRNN menunjukkan kinerja yang baik dalam memprediksi karakter Aksara Bali, dengan sebagian besar data uji berhasil diprediksi 100% akurat (CER 0), membuktikan kemampuan model mengenali karakter-karakter tersebut dengan sangat baik. Meskipun demikian, beberapa kasus, seperti datates4.jpg, datates6.jpg, dan datates7.jpg, menunjukkan CER 0,333 dengan akurasi 66,6% karena adanya satu kesalahan karakter dalam prediksinya misalnya, prediksi dho untuk label sebenarnya dhr yang mungkin disebabkan oleh variasi tulisan tangan yang kompleks atau ambiguitas visual. Meskipun terdapat tantangan ini, model mencapai rata-rata akurasi keseluruhan 89,9%, menegaskan kemampuannya menggeneralisasi pada data baru dan menunjukkan potensi besar untuk aplikasi pengenalan karakter Aksara Bali tulisan tangan, perbaikan di masa depan dapat difokuskan pada diversifikasi data pelatihan untuk karakter-karakter yang rentan kesalahan.

4. Kesimpulan

Penelitian ini berhasil mengimplementasikan dan mengevaluasi sistem pengenalan tulisan tangan Aksara Bali menggunakan algoritma Convolutional Recurrent Neural Network (CRNN). Pendekatan *deep learning* ini dipilih secara khusus untuk mengatasi kompleksitas dan variabilitas inheren dalam karakter tulisan tangan Aksara Bali. Melalui proses pengumpulan data yang divalidasi ahli dan tahapan pemrosesan yang sistematis (konversi *grayscale*, normalisasi, dan *resizing*), model CRNN yang mengintegrasikan lapisan konvolusional, rekuren BiLSTM, dan transkripsi CTC mampu mengekstraksi serta menerjemahkan fitur visual menjadi label karakter. Hasil evaluasi menunjukkan kinerja yang kuat dengan akurasi rata-rata 89,9% pada data pengujian yang belum pernah dilihat sebelumnya, menegaskan potensi besar CRNN dalam pengenalan Aksara Bali tulisan tangan. Capaian ini berkontribusi signifikan pada upaya digitalisasi dan pelestarian Aksara Bali, membuka jalan bagi integrasi yang lebih luas di era digital. Meskipun demikian, masih ada ruang untuk peningkatan akurasi lebih lanjut, terutama pada karakter-karakter yang menunjukkan nilai CER lebih tinggi, yang dapat dieksplorasi melalui penambahan diversitas data pelatihan.

Daftar Pustaka

[1] W. Sundaram, J. Pendidikan Seni dan Budaya, and N. Wayan Budiasti, “Perjalanan Digitalilasi Aksara Bali: Dari Lontar ke Layar dalam Konteks Pelestarian dan Modernisasi

Budaya (Studi Tahun 2010-2015)".
[2] Y. Darmi, M. F. Sepriansyah, and Y. Darnita, "Penerapan Metode Optical Character Recognition (OCR) untuk Mengidentifikasi Teks Pada Identitas Dokumen Surat Izin Mengemudi (SIM)," 2025.
[3] J. Elektronik and I. Komputer Udayana, "Transliterasi Aksara Bali menjadi Huruf Latin pada Citra Tulisan Tangan," vol. 11, no. 4, pp. 2654–5101, 2023.
[4] K. Banu, D. Andreas, W. Anggoro, and A. Setiawan, "OCR: Masa Depan Pengenalan Karakter Optik dan Dampaknya pada Kehidupan Modern," *Jurnal Teknologi Informasi*, vol. 9, 2023, [Online]. Available: https://ejournal.urindo.ac.id/index.php/TI/index
[5] T. D. K, J. James, D. P. Gopinath, and M. K. Ashraf, "Advocating Character Error Rate for Multilingual ASR Evaluation." [Online]. Available: https://github.com/
[6] Y. Teytaut, B. Bouvier, and A. Roebel, "A study on constraining Connectionist Temporal Classification for temporal audio alignment," pp. 5015–5019, 2022, doi: 10.21437/Interspeech.2022-10940ï.
[7] A. S. Gaafar, J. M. Dahr, and A. K. Hamoud, "Comparative Analysis of Performance of Deep Learning Classification Approach based on LSTM-RNN for Textual and Image Datasets," *Informatica (Slovenia)*, vol. 46, no. 5, pp. 21–28, 2022, doi: 10.31449/inf.v46i5.3872.

p-ISSN: 2986-3929
e-ISSN: 3032-1948

Halaman ini sengaja dibiarkan kosong