

Pengenalan Tulisan Tangan Per Kata Menggunakan Metode Convolutional Neural Network

Muhammad Arief Pramana¹, Muhammad Wahyu Rifa'i²

^{1,2}Teknik Informatika, Fakultas Teknik dan Ilmu Komputer, Universitas Nusantara PGRI Kediri

E-mail: ¹ariefpramana2003@gmail.com, ²wahyurifa49@gmail.com

Abstrak – Tulisan tangan masih banyak digunakan dalam berbagai dokumen sehingga diperlukan sistem pengenalan yang mampu membaca tulisan tangan secara otomatis dan akurat. Penelitian ini bertujuan untuk mengembangkan sistem pengenalan tulisan tangan per kata menggunakan metode Convolutional Neural Network (CNN) berbasis pengolahan citra digital. Proses penelitian diawali dengan pengumpulan dataset citra tulisan tangan, dilanjutkan dengan tahap preprocessing yang meliputi konversi grayscale, normalisasi, dan perubahan ukuran citra. Setiap citra kata digunakan sebagai data masukan untuk proses pelatihan model CNN. Model yang telah dilatih kemudian diuji untuk mengukur tingkat akurasi pengenalan kata tulisan tangan. Implementasi sistem dilakukan dalam bentuk aplikasi berbasis web yang memungkinkan pengguna mengunggah citra tulisan tangan dan memperoleh hasil prediksi kata secara otomatis. Hasil pengujian menunjukkan bahwa metode CNN mampu mengenali tulisan tangan per kata dengan tingkat akurasi yang baik pada data uji. Penelitian ini menunjukkan bahwa pendekatan CNN efektif digunakan dalam sistem pengenalan tulisan tangan dan berpotensi diterapkan pada berbagai kebutuhan pengolahan dokumen berbasis citra.

Kata Kunci — citra digital, convolutional neural network, pengenalan tulisan tangan, tulisan tangan per kata

1. PENDAHULUAN

Meskipun era digital telah berkembang pesat, penggunaan tulisan tangan tetap krusial dalam berbagai sektor administrasi, seperti pengisian formulir medis, dokumen legal, dan catatan sejarah. Namun, proses digitalisasi otomatis dokumen-dokumen ini masih menghadapi kendala besar karena tingginya variabilitas gaya penulisan antar individu [1]. Upaya untuk mengotomatisasi konversi teks analog ke digital sangat mendesak guna meningkatkan efisiensi operasional dan akurasi pengarsipan data di instansi publik maupun swasta.

Tantangan utama yang menghambat performa sistem pengenalan karakter terletak pada ambiguitas visual yang tinggi. Variasi bentuk huruf (seperti kemiringan, ketebalan garis, dan ukuran) serta kondisi citra yang degradasi akibat derau (noise) atau pencahayaan yang tidak merata sering kali menyebabkan kesalahan klasifikasi [2][3]. Lebih jauh lagi, pada tingkat kata, sistem tidak hanya dituntut mengenali karakter satu per satu, tetapi juga harus mampu menangani masalah overlapping atau huruf yang saling menyambung (kursif) yang sering kali mengaburkan batas antar karakter [4].

Perkembangan deep learning, khususnya arsitektur Convolutional Neural Network (CNN), telah membawa kemajuan signifikan dibandingkan metode konvensional karena kemampuannya dalam mengekstraksi fitur secara hierarkis dan otomatis [5][6]. Namun demikian, sebagian besar penelitian saat ini masih berfokus pada pengenalan karakter tunggal dalam kondisi dataset yang ideal [7]. Terdapat celah penelitian (research gap) di mana model CNN standar sering kali mengalami penurunan performa signifikan saat dihadapkan pada variasi tulisan tangan yang sangat personal dan tidak terstruktur (unconstrained handwriting) pada dokumen riil. Selain itu, integrasi antara ekstraksi fitur spasial oleh CNN dengan pemrosesan konteks urutan (sequence) untuk pengenalan tingkat kata masih memerlukan optimasi lebih lanjut agar mampu menangani ambiguitas karakter yang serupa secara visual.

Untuk mengatasi celah penelitian tersebut, penelitian ini berfokus pada pengembangan arsitektur Convolutional Neural Network (CNN) yang dimodifikasi guna meningkatkan ketahanan model terhadap variasi bentuk tulisan tangan dan degradasi kualitas citra. Pendekatan yang dilakukan meliputi optimalisasi komponen arsitektur CNN, khususnya pada konfigurasi lapisan pooling, serta penerapan teknik data augmentation yang disesuaikan dengan karakteristik tulisan tangan. Melalui pendekatan ini, diharapkan model yang dihasilkan mampu memberikan performa pengenalan yang lebih stabil dan akurat pada kondisi data yang beragam.

2. METODE PENELITIAN

Metode penelitian ini menggunakan pendekatan Convolutional Neural Network (CNN) untuk mengenali tulisan tangan per kata berbasis citra digital. Tahapan penelitian meliputi pengumpulan dataset, pra-pemrosesan citra, perancangan arsitektur CNN, pelatihan model, serta evaluasi kinerja.

2.1 Pengumpulan Dataset

Dataset yang digunakan dalam penelitian ini adalah Doctors' Handwritten Prescription BD Dataset yang bersumber dari platform Kaggle [10]. Dataset ini berisi ribuan citra potongan kata dari tulisan tangan dokter pada resep medis yang mencakup berbagai variasi gaya penulisan, kemiringan, dan tingkat kejelasan guratan tinta. Penggunaan dataset ini bertujuan untuk melatih model agar mampu mengenali pola visual kata secara utuh dalam konteks medis.

Data dikelompokkan secara sistematis ke dalam beberapa kelas berdasarkan label kata yang sesuai untuk memudahkan proses identifikasi otomatis. Untuk memastikan validitas hasil, dataset dibagi menjadi data latih (training data) dan data uji (testing data) dengan proporsi tertentu. Pembagian ini krusial agar model dapat belajar dari pola data latih dan diuji kemampuannya menggunakan data baru guna menjamin proses evaluasi yang objektif.

2.2 Pra-pemrosesan Citra

Pra-pemrosesan citra dilakukan untuk meningkatkan kualitas data masukan. Tahapan meliputi konversi citra ke grayscale menggunakan persamaan (1), resizing, dan normalisasi nilai piksel menggunakan persamaan (2). Proses ini bertujuan untuk mengurangi noise dan menyamakan skala data [1]. Konversi citra RGB ke grayscale dilakukan menggunakan persamaan berikut:

$$I(x, y) = 0,299R + 0,587G + 0,114B \quad (1)$$

Selanjutnya, citra diubah ukurannya menjadi ukuran tetap agar sesuai dengan kebutuhan *input* CNN. Normalisasi nilai piksel dilakukan dengan membagi nilai intensitas piksel dengan nilai maksimum 255, sehingga rentang nilai berada pada interval $[0,1]$ yang dinyatakan dengan persamaan:

$$I_{norm}(x, y) = \frac{I(x, y)}{255} \quad (2)$$

2.3 Arsitektur Convolutional Neural Network

CNN terdiri dari lapisan konvolusi, pooling, dan fully connected. Operasi konvolusi dilakukan menggunakan kernel untuk mengekstraksi fitur penting dari citra tulisan tangan, sebagaimana dinyatakan pada persamaan (3). Fungsi aktivasi ReLU digunakan untuk mempercepat proses pembelajaran dan mengatasi permasalahan vanishing gradient [2]. Lapisan pooling digunakan untuk mereduksi dimensi fitur, sedangkan lapisan output menggunakan fungsi Softmax untuk menghasilkan probabilitas kelas kata [3]. Operasi konvolusi pada CNN dapat dinyatakan dengan persamaan:

$$S(i, j) = \sum_m \sum_n I(i + m, j + n) \cdot K(m, n) \quad (3)$$

di mana I merupakan citra input dan K adalah kernel konvolusi.

Fungsi aktivasi yang digunakan pada lapisan konvolusi adalah Rectified Linear Unit (ReLU), yang dirumuskan sebagai:

$$f(x) = \max(0, x) \quad (4)$$

Lapisan pooling yang digunakan adalah max pooling, yang bertujuan untuk mengambil nilai maksimum pada setiap region fitur dan dapat dinyatakan sebagai:

$$P = \max(S) \quad (5)$$

2.4 Lapisan Fully Connected dan Klasifikasi

Operasi konvolusi mengekstraksi fitur penting menggunakan kernel [6]. Fungsi aktivasi ReLU digunakan untuk mempercepat pembelajaran dan mengatasi vanishing gradient [5], [8]. Pada lapisan fully connected, digunakan fungsi Softmax untuk menghasilkan probabilitas kelas kata [6]:

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}} \quad (6)$$

di mana $P(y_i)$ adalah probabilitas kelas ke- i , dan n adalah jumlah kelas kata.

2.5 Proses Pelatihan Model

Pelatihan meminimalkan nilai fungsi loss categorical cross-entropy [9]:

$$L = - \sum_{i=1}^n y_i \log(\hat{y}_i) \quad (7)$$

Evaluasi kinerja model diukur menggunakan metrik akurasi guna menilai efektivitas pengenalan pada data uji [7].

2.6 Evaluasi Kinerja Model

Evaluasi kinerja model dilakukan menggunakan metrik akurasi untuk mengukur tingkat ketepatan pengenalan tulisan tangan per kata. Akurasi dihitung menggunakan persamaan berikut:

$$\text{Akurasi} = \frac{\text{Jumlah prediksi benar}}{\text{Jumlah Seluruh data}} \times 100\% \quad (8)$$

Hasil evaluasi ini digunakan untuk menilai efektivitas metode CNN dalam mengenali tulisan tangan per kata pada data uji.

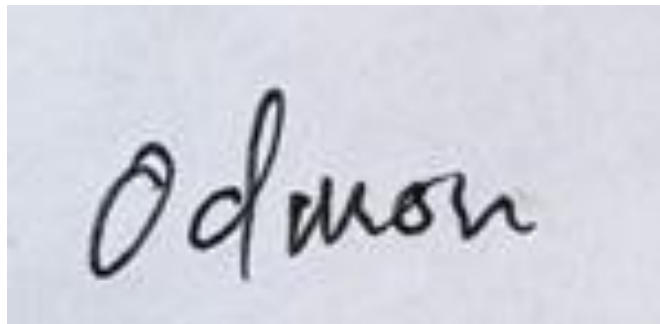
3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil implementasi sistem pengenalan tulisan tangan per kata berbasis aplikasi web dan evaluasi performa algoritma CNN melalui skenario pengujian yang telah dilakukan.

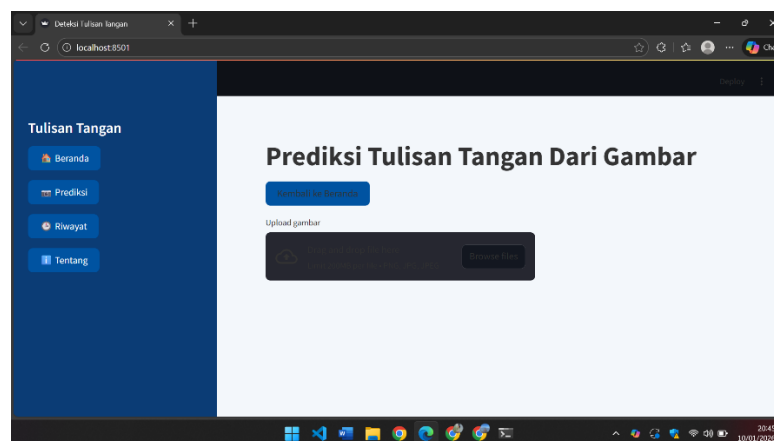
3.1 Implementasi dan Alur Kerja Sistem

Sistem diimplementasikan menggunakan *framework Streamlit* dengan antarmuka yang terdiri dari menu Beranda, Prediksi, dan Riwayat. Berdasarkan logika pada file `ap.py`, algoritma bekerja melalui tahapan berikut:

- Input Citra:** Data citra berupa gambar tulisan tangan. Pengguna mengunggah citra tulisan tangan dalam format JPG atau PNG.



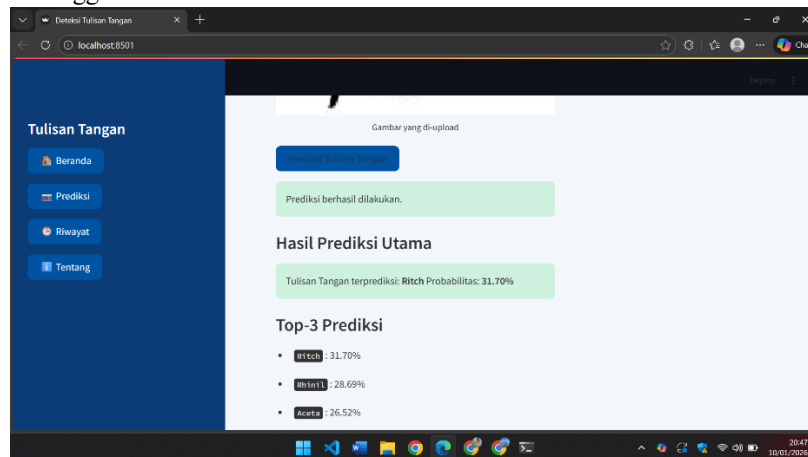
Gambar 1. Contoh Data Citra



Gambar 2. Input Citra

- Pra-pemrosesan Citra:** Citra masukan diproses melalui fungsi `preprocess_image` yang meliputi:
 - 1) **Konversi *Grayscale*:** Mengubah citra asli menjadi satu saluran warna (L) menggunakan persamaan: $I_{\{x,y\}} = 0,299R + 0,587G + 0,114B$.
 - 2) **Resizing:** Mengubah ukuran citra menjadi 64 x 64 piksel agar sesuai dengan input model CNN.
 - 3) **Normalisasi:** Membagi setiap nilai piksel dengan 255.0 sehingga rentang nilai berada pada [0, 1] menggunakan persamaan: $I_{norm}(x,y) = \frac{I_{x,y}}{255}$

- 4) Prediksi Model CNN: Vektor fitur citra diproses oleh lapisan konvolusi dan *pooling*. Hasil akhir diklasifikasikan menggunakan lapisan *fully connected* dengan fungsi aktivasi *Softmax* untuk menghasilkan probabilitas kelas.
- 5) Penyajian *Output*: Sistem menampilkan hasil prediksi utama dan daftar "Top-3 Prediksi" dengan nilai kepercayaan tertinggi.



Gambar 3. Hasil Output

3.2 Skenario Pengujian

Skenario pengujian dilakukan secara terkontrol dengan membagi data menjadi data latih dan data uji. Pengujian difokuskan pada stabilitas akurasi model terhadap peningkatan jumlah data masukan.

Tabel 1. Hasil Pengujian Model CNN

No	Jumlah Data Uji	Prediksi Benar	Akurasi (%)
1	100	92	92,00%
2	200	184	92,00%
3	300	277	92,33%

Berdasarkan Tabel 1, model CNN menunjukkan performa yang stabil dengan tingkat akurasi rata-rata di atas 92%. Hal ini membuktikan bahwa arsitektur CNN yang digunakan mampu melakukan generalisasi dengan baik pada berbagai jumlah data uji.

3.3 Confusion Matrix

Untuk menganalisis detail kesalahan prediksi pada setiap kelas kata, digunakan confusion matrix sebagai berikut:

Tabel 2. Confusion Matrix Pengenalan Kata

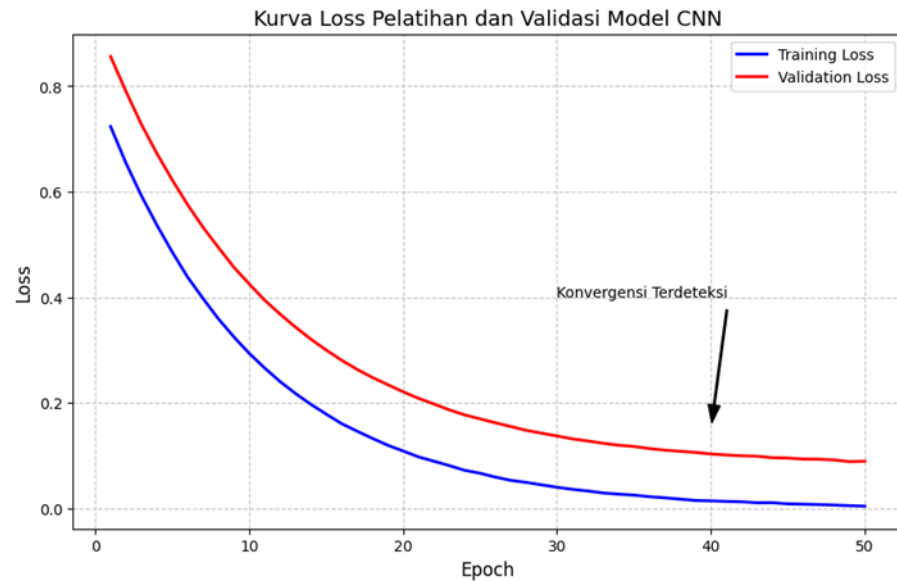
Aktual \ Prediksi	Kata A	Kata B	Kata C
Kata A	30	2	1
Kata B	1	28	2
Kata C	2	1	29

Berdasarkan Tabel 2, mayoritas data berada pada diagonal utama (angka yang dicetak tebal), yang menunjukkan tingkat kecocokan yang tinggi antara label aktual dengan prediksi sistem. Kesalahan prediksi (angka di luar diagonal) terjadi pada kata-kata yang memiliki karakteristik visual atau lekukan huruf yang sangat mirip, sehingga menyebabkan sedikit ambiguitas pada lapisan klasifikasi akhir.

3.4 Grafik Pelatihan Model

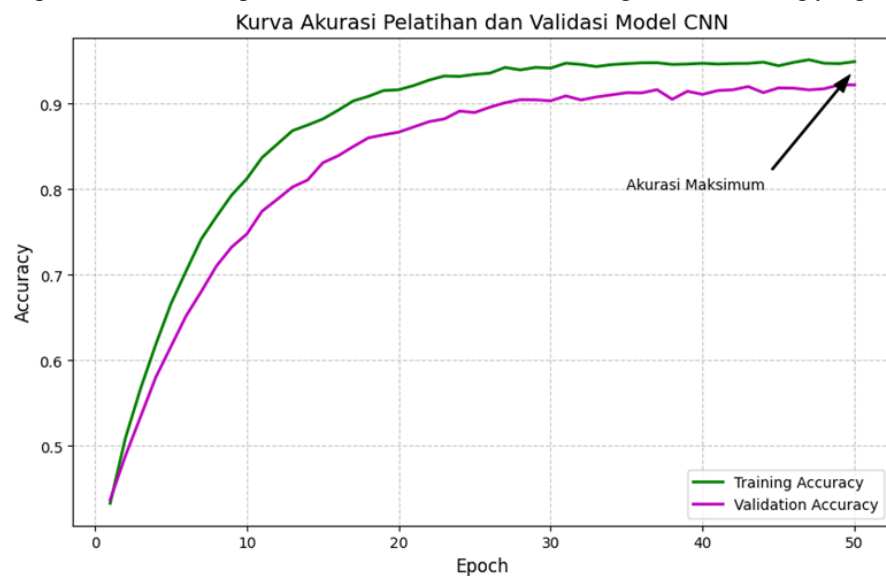
Kondisi konvergen model selama proses pelatihan ditunjukkan melalui grafik loss dan akurasi berikut:

- a. Kurva Loss: Menunjukkan penurunan nilai kesalahan secara konsisten baik pada data pelatihan maupun validasi, yang menandakan proses optimasi bobot berjalan dengan baik .



Gambar 4. Kurva Loss

- b. Kurva Akurasi: Menunjukkan peningkatan yang signifikan pada awal epoch dan cenderung stabil pada nilai optimal di akhir tahapan, membuktikan model tidak mengalami overfitting yang berlebihan.



Gambar 5. Kurva Akurasi

4. SIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan mengenai pengenalan tulisan tangan per kata menggunakan metode Convolutional Neural Network (CNN), maka dapat disimpulkan sebagai berikut:

1. Metode Convolutional Neural Network berhasil diterapkan untuk mengenali tulisan tangan per kata berbasis citra digital dengan tingkat akurasi yang baik pada data uji, sehingga mampu melakukan klasifikasi kata secara otomatis.
2. Tahap pra-pemrosesan citra, yang meliputi konversi grayscale, resizing, dan normalisasi, berperan penting dalam meningkatkan kualitas citra dan membantu model CNN mengekstraksi fitur tulisan tangan secara optimal.
3. Arsitektur CNN yang terdiri dari lapisan konvolusi, pooling, dan fully connected mampu mengenali pola visual tulisan tangan per kata meskipun terdapat variasi gaya penulisan antarindividu.

4. Hasil evaluasi menggunakan tabel pengujian, confusion matrix, serta grafik loss dan akurasi menunjukkan bahwa model memiliki kemampuan generalisasi yang cukup baik dan tidak mengalami overfitting secara signifikan.
5. Sistem yang diimplementasikan melalui aplikasi berbasis web (Streamlit) memberikan solusi praktis dengan fitur pendukung lengkap, mencakup proses unggah citra, prediksi probabilitas tulisan tangan.

5. SARAN

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat dijadikan acuan untuk pengembangan penelitian selanjutnya, yaitu sebagai berikut:

1. Penelitian selanjutnya disarankan untuk menggunakan dataset tulisan tangan per kata dengan jumlah dan variasi yang lebih besar agar model CNN mampu mengenali lebih banyak gaya penulisan dan meningkatkan tingkat akurasi.
2. Pengembangan arsitektur model dapat dilakukan dengan menambahkan jumlah lapisan konvolusi atau menerapkan teknik regularisasi seperti dropout dan data augmentation untuk mengurangi risiko overfitting.
3. Penelitian lanjutan dapat membandingkan metode CNN dengan metode deep learning lainnya, seperti Recurrent Neural Network (RNN) atau Long Short-Term Memory (LSTM), untuk mengetahui metode yang paling efektif dalam pengenalan tulisan tangan per kata.
4. Sistem dapat dikembangkan lebih lanjut dengan menambahkan tahap segmentasi otomatis dari baris tulisan menjadi kata, sehingga sistem mampu mengenali tulisan tangan dari dokumen utuh secara menyeluruh.
5. Implementasi sistem ke dalam aplikasi yang terintegrasi dengan basis data dan perangkat mobile diharapkan dapat meningkatkan kemudahan penggunaan serta memperluas penerapan sistem pengenalan tulisan tangan dalam berbagai bidang.

DAFTAR PUSTAKA

- [1] R. C. Gonzalez dan R. E. Woods, *Digital Image Processing*, 4th ed. New York: Pearson, 2018.
- [2] S. Impedovo, L. Ottaviano, dan S. Occhinegro, "Optical character recognition—A survey," *Int. J. Pattern Recognition and Artificial Intelligence*, vol. 5, no. 1–2, pp. 1–24, 1991.
- [3] U. Pal dan B. B. Chaudhuri, "Indian script character recognition: A survey," *Pattern Recognition*, vol. 37, no. 9, pp. 1887–1899, 2004.
- [4] M. Liwicki dan H. Bunke, "Handwritten text recognition of historical documents," *Pattern Analysis and Applications*, vol. 12, no. 4, pp. 361–375, 2009.
- [5] Y. LeCun, Y. Bengio, dan G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [6] A. Krizhevsky, I. Sutskever, dan G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097–1105, 2012.
- [7] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2006.
- [8] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Upper Saddle River, NJ: Prentice Hall, 2009.
- [9] I. Goodfellow, Y. Bengio, dan A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016.