

Peningkatan Akurasi OCR dalam Pemrosesan Formulir Keuangan melalui *Fine-Tuning* Transformer dan Strategi Prapemrosesan Data

Enhancing OCR Accuracy in Financial Form Processing through Transformer Fine Tuning and Data Preprocessing

Rafi Surya^{1)*} 

¹Magister Teknik Informatika, Universitas Bina Nusantara, Jakarta, Indonesia

*email korespondensi: rafi.surya@binus.ac.id

Abstrak

Permintaan layanan kartu kredit yang terus meningkat mendorong kebutuhan untuk metode pemrosesan formulir yang lebih efisien. Artikel ini mengusulkan penggunaan model Optical Character Recognition (OCR) berbasis Transformer, TrOCR, yang telah disempurnakan melalui *fine-tuning* dan prapemrosesan data untuk meningkatkan akurasi dalam pemrosesan formulir pendaftaran kartu kredit. Hasil eksperimen menunjukkan peningkatan signifikan dalam kinerja model setelah penerapan strategi prapemrosesan data dan penyetelan hiperparameter. Model dengan konfigurasi terbaik menghasilkan Character Error Rate (CER) sebesar 0,0607 pada dataset IAM dan 0,3620 pada formulir Bank XYZ.

Kata kunci: OCR berbasis Transformer; Pengenalan Formulir Tulisan Tangan; Prapemrosesan Data; Tingkat Kesalahan Karakter (CER); Digitalisasi Layanan Keuangan

Abstract

The increasing demand for credit card services drives the need for more efficient form processing methods. This paper proposes the use of a Transformer-based Optical Character Recognition (OCR) model, TrOCR, refined through fine-tuning and data preprocessing to improve the accuracy of credit card registration form processing. Experimental results show significant improvements in model performance after the application of data preprocessing strategies and hyperparameter tuning. The best-configured model produces a Character Error Rate (CER) of 0.0607 on the IAM dataset and 0.3620 on the XYZ Bank form.

Keywords: *Transformer-based OCR; Handwritten Form Recognition; Data Preprocessing; Character Error Rate (CER); Financial Services Digitization*

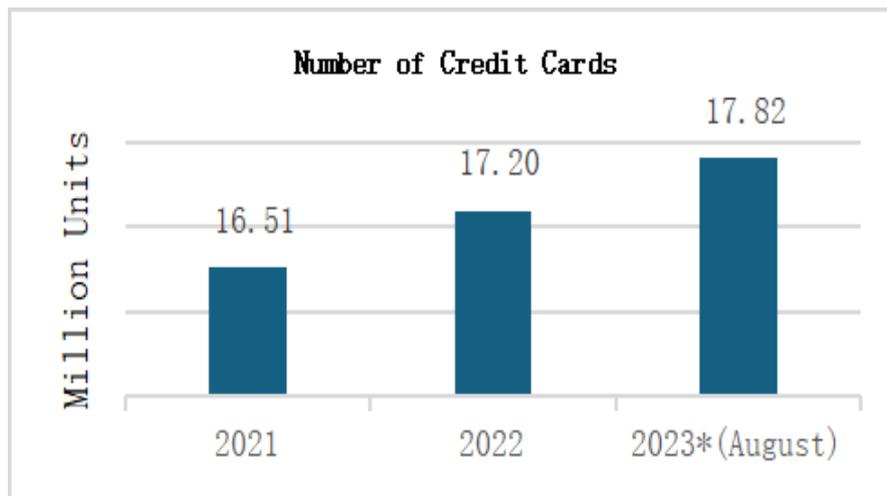
Article history: Received 15 April 2025, Accepted 28 April 2025, Available online 30 April 2025

1 PENDAHULUAN

Kartu kredit telah menjadi salah satu metode pembayaran yang paling banyak digunakan, seiring dengan meningkatnya transaksi di Indonesia. Berdasarkan data dari Bank Indonesia (2023), jumlah kartu kredit yang beredar mencapai 17,82 juta unit pada tahun 2023, yang meningkat dari 16,51 juta kartu pada tahun 2021 (Bank Indonesia, 2023). Volume transaksi kartu kredit juga mengalami peningkatan signifikan, dari Rp 244,516 juta pada tahun 2021 menjadi Rp 323,602 juta pada tahun 2022, dengan prediksi angka yang terus berkembang pada tahun 2023 (Bank Indonesia, 2023). Proses pendaftaran kartu kredit menjadi langkah

penting dalam memperoleh layanan ini, yang melibatkan pengisian formulir berisi data pribadi dan informasi keuangan.

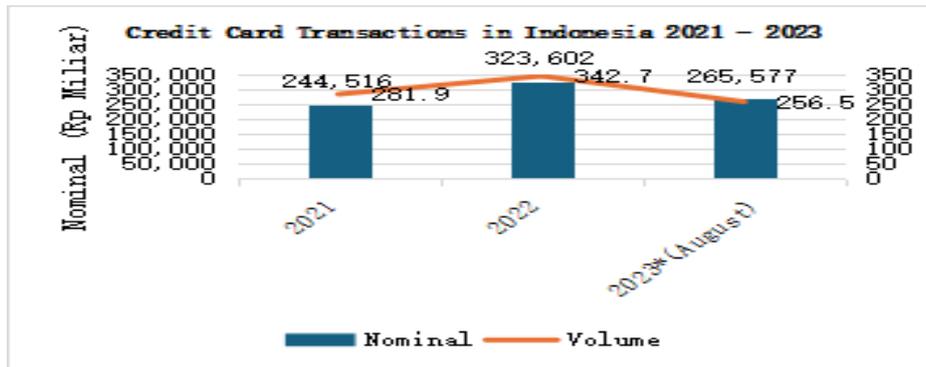
Namun, pengolahan formulir pendaftaran kartu kredit secara manual masih sangat rentan terhadap kesalahan dan membutuhkan waktu yang cukup lama. Oleh karena itu, perlu adanya solusi otomatisasi yang lebih efisien untuk mempercepat proses pendaftaran dan meminimalkan kesalahan manusia dalam entri data. Salah satu teknologi yang dapat diandalkan untuk tujuan ini adalah Optical Character Recognition (OCR), yang mampu mengotomatisasi ekstraksi data dari formulir tulisan tangan.



GAMBAR 1. JUMLAH PENGGUNA KARTU KREDIT DI INDONESIA
(Bank Indonesia, 2023)

Penelitian ini berfokus pada penggunaan model OCR berbasis Transformer, TrOCR, yang telah diperbaharui dengan teknik *fine-tuning* dan strategi prapemrosesan data untuk meningkatkan akurasi dalam memproses formulir pendaftaran kartu kredit. Teknologi OCR ini telah dipilih karena memiliki keunggulan dalam pengenalan teks tulisan tangan, dengan kemampuan untuk mengenali pola yang lebih kompleks dibandingkan dengan pendekatan sebelumnya (Vaswani et al., 2017).

Model TrOCR, yang memanfaatkan model Vision Transformer (ViT) sebagai encoder dan BERT sebagai *decoder*, telah terbukti efektif dalam mengolah teks panjang dan struktur kompleks tanpa ketergantungan pada jaringan saraf konvolusional (Li et al., 2021). Penelitian ini dilaksanakan di Bank XYZ, sebuah institusi keuangan yang memiliki infrastruktur teknologi yang memadai dan komitmen terhadap inovasi dalam pelayanan perbankan. Diharapkan, penerapan model TrOCR yang dioptimalkan ini dapat mempercepat proses entri data formulir pendaftaran kartu kredit, sekaligus meningkatkan akurasi pengenalan tulisan tangan pada data yang sangat bervariasi (Li et al., 2021).



GAMBAR 2. TRANSAKSI PENGGUNA KARTU KREDIT DI INDONESIA
 (Bank Indonesia, 2023)

2 TINJAUAN PUSTAKA

Berbagai penelitian telah dilakukan terkait penggunaan teknologi Optical Character Recognition (OCR) dalam sektor keuangan, dengan fokus pada pengenalan karakter tulisan tangan pada dokumen keuangan. Beberapa pendekatan yang telah dicoba melibatkan penggunaan model Convolutional Neural Networks (CNN) dan Recurrent Neural Networks (RNN), yang telah menunjukkan hasil yang bervariasi tergantung pada metode dan dataset yang digunakan.

Shuiqi et al. (2021) menggunakan model VGG16 untuk pengenalan nomor kartu bank, yang menghasilkan akurasi sebesar 86%. Dalam penelitian lain, Xin et al. (2019) mengimplementasikan OpenCV dalam OCR kartu bank, namun hasil yang diperoleh masih rendah, yaitu 13,4%. Untuk meningkatkan akurasi, mereka menggabungkan metode Faster R-CNN dan LeNet-5, yang menghasilkan peningkatan menjadi 90,5%. Selain itu, pendekatan lain seperti menggabungkan Deep Convolutional Neural Networks (DCNN) dengan Convolutional Recurrent Neural Networks (CRNN) menghasilkan akurasi sebesar 89,9%, sementara metode Connected Text Proposal Network (CTPN) dengan CRNN mencatatkan akurasi 95,5% (Sun & You, 2020).

TABEL 1. PERBANDINGAN AKURASI MODEL OCR PADA PENGGUNAAN BERBAGAI PENDEKATAN

Penelitian	Model yang Digunakan	Akurasi (%)
Shuiqi et al. (2021)	VGG16	86
Xin et al. (2019)	OpenCV + LeNet-5	90.5
Sun & You (2020)	DCNN + CRNN	89.9
Watson & Liu (2020)	U-Net + DenseNet-169	91.7

Penelitian terkait penggunaan CNN dalam OCR juga banyak dilakukan, seperti yang dilakukan oleh Srivastava et al. (2019) yang mengembangkan model OCR untuk pengenalan karakter tulisan tangan pada cek bank dengan akurasi mencapai 95,71%. Dalam penelitian

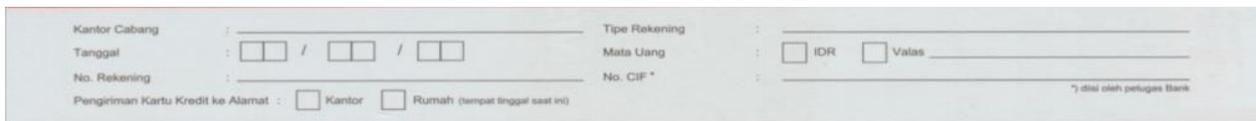
lainnya, Parthiban et al. (2020) menerapkan model Recurrent Neural Network (RNN) untuk pengenalan tulisan tangan berbahasa Inggris dengan tingkat akurasi mencapai 90%.

Namun, seiring dengan perkembangan teknologi *deep learning*, pendekatan berbasis Transformer semakin banyak diterapkan, terutama dalam pengenalan teks panjang dan struktur kompleks. TrOCR, yang merupakan model OCR berbasis Transformer, menawarkan keunggulan dalam memahami hubungan jangka panjang antar karakter dalam teks dan dapat mengatasi kesulitan yang dihadapi oleh model berbasis CNN dan RNN dalam menangani teks yang lebih kompleks (Vaswani et al., 2017). TrOCR telah menunjukkan hasil yang lebih baik dibandingkan dengan model-model OCR lainnya dalam pengenalan tulisan tangan, seperti yang telah dibuktikan dalam eksperimen yang dilakukan oleh Li et al. (2021), yang menunjukkan bahwa TrOCR dapat mengalahkan model lain dengan menggunakan dataset IAM dan mencapai CER sebesar 2,89.

Dengan keunggulan ini, penelitian ini fokus pada penerapan model TrOCR yang telah diperbarui untuk meningkatkan akurasi OCR dalam memproses formulir pendaftaran kartu kredit yang berisi teks tulisan tangan dari pengguna nyata. Penelitian ini bertujuan untuk menyesuaikan model TrOCR dengan prapemrosesan data dan tuning hiperparameter, sehingga dapat meningkatkan akurasi pengenalan karakter pada formulir yang berisi berbagai gaya tulisan tangan yang sering ditemukan pada dokumen keuangan.

3 DATASET

Dalam penelitian ini, dua dataset utama digunakan untuk pengembangan dan pengujian model OCR berbasis Transformer (TrOCR), yaitu IAM Dataset dan Formulir Bank XYZ.



GAMBAR 3. FORMULIR GAMBAR BAGIAN INFORMASI AKUN

Dataset kedua adalah Formulir Bank XYZ, sebuah kumpulan data formulir tulisan tangan dari formulir pendaftaran kartu kredit Bank XYZ. Dataset ini digunakan secara eksklusif untuk pengujian akhir model. Penggunaan dataset ini sangat penting untuk mengukur kinerja model pada data yang lebih mendekati aplikasi dunia nyata, memberikan gambaran yang lebih akurat tentang kemampuan generalisasi model. Dataset Formulir Bank XYZ terdiri dari formulir yang diisi oleh nasabah nyata, mewakili teks tulisan tangan yang akan dihadapi model dalam praktik.

GAMBAR 4. FORMULIR GAMBAR BAGIAN DATA PRIBADI

3.1 RAINING DATASET

Untuk tahap pelatihan, IAM Dataset dipilih karena sifatnya yang komprehensif, terdiri dari 13.353 gambar garis teks tulisan tangan yang dibuat oleh 657 penulis yang berbeda. Isi dataset ini berasal dari Lancaster-Oslo/Bergen Corpus Bahasa Inggris Britania, yang menyediakan korpus teks tulisan tangan yang luas. Korpus ini dibagi menjadi beberapa tingkat anotasi: kalimat, baris, dan kata, yang memungkinkan eksperimen yang mendalam pada pengenalan tulisan tangan. Data pelatihan ini menawarkan berbagai gaya dan jenis tulisan tangan, yang berkontribusi pada kemampuan model untuk mempelajari pola teks yang kompleks.

3.2 TESTING DATASET

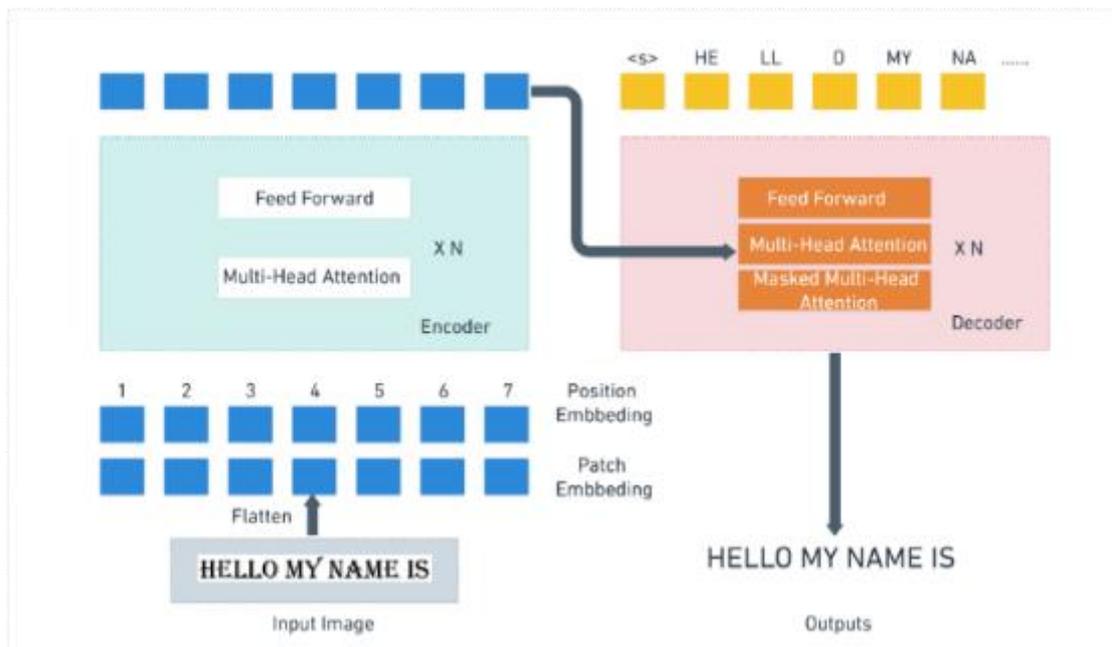
Pada fase pengujian, penelitian ini menggunakan Formulir Bank XYZ, yang merupakan formulir pendaftaran kartu kredit tulisan tangan asli. Formulir ini mencakup informasi penting nasabah seperti nama, alamat, dan detail kontak, yang harus diproses dengan akurat oleh model OCR. Berbeda dengan IAM Dataset yang berisi tulisan tangan sintetis, dataset Formulir Bank XYZ mencerminkan input pengguna nyata, yang sangat penting untuk mengevaluasi penerapan model dalam situasi dunia nyata dan kemampuan generalisasi model. Formulir yang digunakan dalam pengujian diambil secara manual, dengan fokus khusus pada informasi akun pribadi untuk memastikan kinerja model pada data tulisan tangan yang berbeda dari data pelatihan sintetis.

4 METODOLOGI

Penelitian ini mengembangkan dan menguji model OCR berbasis Transformer (TrOCR) untuk pengenalan karakter pada formulir pendaftaran kartu kredit. Proses penelitian terbagi menjadi dua tahap utama: Pelatihan Model dan Pengujian Model.

4.1 ARSITEKTUR TrOCR

Model TrOCR yang digunakan dalam penelitian ini adalah model OCR berbasis Transformer yang dirancang untuk ekstraksi teks dari gambar. Model ini pertama kali diperkenalkan oleh Li et al. (2021) dan mengintegrasikan Vision Transformer (ViT) sebagai encoder dengan BERT sebagai decoder. Arsitektur Transformer, yang memperkenalkan mekanisme *self-attention*, memungkinkan model untuk memproses informasi secara efisien tanpa ketergantungan pada pola spasial lokal seperti pada CNN. Gambar 3 menunjukkan arsitektur dasar dari TrOCR, yang memanfaatkan Vision Transformer (ViT) sebagai *encoder* dan BERT sebagai *decoder*, dengan mekanisme *self-attention* untuk pengenalan teks.



GAMBAR 5. ARSITEKTUR TRANSFORMER PADA MODEL TROCR

Mekanisme *self-attention* pada Transformer bekerja dengan cara menghitung kesamaan antara query (Q) dan key (K), kemudian menggunakan nilai tersebut untuk menghitung bobot perhatian yang diterapkan pada matriks value (V). Proses ini memungkinkan model untuk fokus pada bagian-bagian penting dari *input* dan menghasilkan *output* yang lebih akurat.

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

Di mana:

- Q (Query): Matriks query, yang merepresentasikan informasi yang ingin dicari hubungannya dalam input.
- K (Key): Matriks key, yang berfungsi sebagai referensi untuk mencari kesamaan dengan query.
- V (Value): Matriks value, yang berisi informasi yang digunakan untuk menghasilkan output setelah dihitung berdasarkan kesamaan dengan query dan key.
- d_k : Dimensi dari key (atau query), yang digunakan untuk menormalkan nilai skor kesamaan, mencegah skor menjadi terlalu besar dan memperlambat proses pembelajaran.

4.2 TAHAPAN TrOCR

Tahap *fine-tuning* pada model TrOCR dilakukan dengan menggunakan dua skenario:

1) *Fine-Tuning* dengan *Preprocessing* Data

Preprocessing dilakukan untuk meningkatkan kualitas gambar dan mengurangi *noise* yang dapat mengganggu akurasi model. Teknik yang digunakan meliputi penyesuaian kecerahan, transformasi hue, dan penerapan filter Gaussian.

2) *Fine-Tuning* tanpa *Preprocessing* Data

Pada skenario ini, model dilatih tanpa penerapan teknik *preprocessing* dan hanya mengandalkan data asli untuk pelatihan dan pengujian.

TABEL 2. PARAMETER PREPROCESSING DAN PENGATURAN HYPERPARAMETER

Parameter	Skema 1 (Dengan <i>Preprocessing</i>)	Skema 2 (Tanpa <i>Preprocessing</i>)
Format Gambar	RGB	RGB
Kecerahan	0.5	0.5
Hue	0.3	0.3
GaussianBlur	Kernel Size = (5, 9), Sigma = (0.1, 5)	Tidak Ada
Learning Rate	0.00005	0.001
Batch Size	24	32

4.3 PROSES PELATIHAN DAN PENGUJIAN

Untuk tahap pelatihan, dataset IAM digunakan untuk melatih model TrOCR yang telah dilengkapi dengan *preprocessing* data. Pengaturan hyperparameter, seperti ukuran *batch* dan *learning rate*, juga disesuaikan untuk mengoptimalkan kinerja model. Model yang telah dilatih kemudian diuji dengan menggunakan formulir pendaftaran kartu kredit Bank XYZ untuk mengevaluasi kemampuannya dalam mengenali karakter tulisan tangan pada data nyata. Proses pelatihan dan pengujian menggunakan data yang dibagi menjadi tiga subset utama:

- 1) data pelatihan
- 2) data pengujian, dan
- 3) data validasi.

Data pelatihan digunakan untuk melatih model, data pengujian untuk mengevaluasi kinerja model, dan data validasi untuk memverifikasi hasil model yang telah dilatih.

5 HASIL DAN PEMBAHASAN

Pada bab ini, kami akan memaparkan hasil evaluasi dan pembahasan mengenai performa model TrOCR yang telah disesuaikan dengan data dari IAM Dataset dan Formulir Kartu Kredit Bank XYZ. Evaluasi dilakukan dengan menggunakan Character Error Rate (CER) sebagai metrik utama untuk mengukur akurasi model dalam mengenali karakter pada gambar formulir yang diuji. Kami akan membahas hasil evaluasi model pada kedua dataset tersebut serta implikasi dari hasil yang diperoleh.

5.1 EVALUASI PADA IAM DATASET

Model TrOCR yang telah disesuaikan dengan dataset IAM diuji untuk menilai kinerjanya dalam pengenalan tulisan tangan pada teks berbahasa Inggris. Tabel 3 menunjukkan hasil evaluasi model pada dataset IAM, di mana pengujian dilakukan untuk empat model yang berbeda, dengan variasi dalam tahap *pre-processing* dan pengaturan *hyperparameter*.

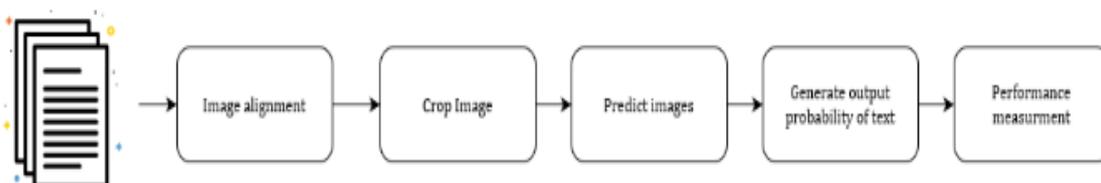
TABEL 3: KINERJA MODEL PADA PENGUJIAN DATASET IAM

<i>Experiment Model</i>	<i>Preprocessing</i>	<i>Hyperparameter Tuning</i>	CER
TrOCR_Model_1		✓	0.1150
TrOCR_Model_2		✓	0.0851
TrOCR_Model_3	✓	✓	0.0795
TrOCR_Model_4	✓	✓	0.0607

Berdasarkan hasil pengujian pada IAM dataset, model dengan konfigurasi *pre-processing* dan tuning hyperparameter menunjukkan hasil terbaik, dengan CER sebesar 0.0607. Model yang tidak menggunakan *pre-processing* dan *tuning hyperparameter* memiliki CER yang lebih tinggi, yaitu 0.1150. Ini menunjukkan bahwa pengaturan yang tepat pada tahap *pre-processing* dan pengaturan *hyperparameter* sangat penting untuk mengoptimalkan kinerja model dalam pengenalan tulisan tangan.

5.2 EVALUASI PADA FORMULIR KARTU KREDIT BANK XYZ

Pada tahap evaluasi selanjutnya, model yang telah disesuaikan diuji dengan menggunakan formulir aplikasi kartu kredit asli dari Bank XYZ. Sebanyak 50 formulir digunakan untuk mengukur akurasi model dalam mengenali informasi yang tertera pada formulir tersebut. Proses pengujian mencakup tahapan pemrosesan gambar seperti image alignment dan image cropping yang dilakukan sebelum model melakukan pengenalan karakter.

**GAMBAR 6: PROSES EVALUASI PADA FORMULIR KARTU KREDIT BANK XYZ**

Gambar 6 di atas menunjukkan langkah-langkah dalam proses evaluasi pada formulir aplikasi kartu kredit, dimulai dengan pemrosesan gambar, *alignment*, dan *cropping* untuk mendapatkan bagian-bagian formulir yang akan diuji.

TABEL 4: KINERJA MODEL PADA PENGUJIAN FORMULIR BANK XYZ

Experiment Model	CER
TrOCR_Model_1	0.5213
TrOCR_Model_2	0.4985
TrOCR_Model_3	0.4816
TrOCR_Model_4	0.3620

Berdasarkan hasil evaluasi pada formulir aplikasi kartu kredit Bank XYZ, model yang diterapkan dengan *pre-processing* dan *tuning hyperparameter* menunjukkan hasil yang lebih baik dibandingkan dengan model lainnya. Model TrOCR_Model_4, yang telah mengalami *fine-tuning* dan *preprocessing*, menghasilkan CER yang paling rendah, yaitu 0.3620. Hasil ini

menunjukkan bahwa proses *pre-processing*, seperti penyesuaian kecerahan dan hue, serta tuning parameter *beam search*, berkontribusi besar terhadap peningkatan akurasi pengenalan karakter pada formulir dengan berbagai gaya tulisan tangan.

5.3 PENGARUH *PRE-PROCESSING* DAN *TUNING HYPERPARAMETER*

Salah satu temuan penting dari eksperimen ini adalah pengaruh signifikan dari tahap *pre-processing* dan *tuning hyperparameter* terhadap kinerja model. Proses *pre-processing* membantu mengatasi gambar dengan kualitas rendah (misalnya gambar buram atau gelap) dengan menyesuaikan parameter seperti kecerahan dan hue. Selain itu, pengaturan hyperparameter seperti *beam search* memungkinkan model untuk lebih efisien dalam mencari urutan kata yang paling tepat, meningkatkan hasil pengenalan karakter secara keseluruhan.



GAMBAR 7: PROSES *IMAGE ALIGNMENT* PADA FORMULIR KARTU KREDIT

5.4 PEMBAHASAN HASIL PADA SETIAP MODEL

Dari hasil evaluasi yang diperoleh pada kedua dataset (IAM Dataset dan formulir Bank XYZ), model TrOCR yang dilatih dengan *pre-processing* dan *tuning hyperparameter* menunjukkan performa terbaik. Model ini memiliki CER yang lebih rendah, yang mengindikasikan bahwa model lebih akurat dalam mengenali karakter dan kata-kata yang ada pada formulir kartu kredit yang diuji. Hal ini menegaskan pentingnya penggunaan teknik-teknik optimisasi dalam meningkatkan akurasi OCR berbasis transformer, terutama pada formulir yang berisi tulisan tangan yang beragam.

5.5 IMPLIKASI DAN ARAH PENGEMBANGAN SELANJUTNYA

Hasil dari penelitian ini menunjukkan bahwa model TrOCR yang telah disesuaikan dapat diimplementasikan dengan baik untuk meningkatkan efisiensi dan akurasi dalam pengolahan formulir pendaftaran kartu kredit. Model ini dapat mempercepat proses input data secara otomatis, mengurangi waktu dan biaya, serta meminimalkan potensi kesalahan manusia. Namun, terdapat beberapa area yang dapat dikembangkan lebih lanjut, seperti penerapan Parameter-Efficient Fine-Tuning (PEFT) untuk mengurangi jumlah parameter yang dapat dilatih, sehingga mengurangi kebutuhan memori dan meningkatkan kecepatan pelatihan. Selain itu, pengembangan lebih lanjut pada tahap *pre-processing* dapat dilakukan dengan meningkatkan kualitas gambar dan penyesuaian tingkat kecerahan serta resolusi untuk memperoleh hasil yang lebih optimal.

6 KESIMPULAN

Penelitian ini bertujuan untuk meningkatkan efisiensi proses pendaftaran kartu kredit menggunakan model OCR berbasis transformer, yaitu TrOCR. Berdasarkan hasil evaluasi pada dua dataset yang berbeda, IAM Dataset dan Formulir Aplikasi Kartu Kredit Bank XYZ, dapat disimpulkan beberapa hal penting sebagai berikut:

- 1 Model TrOCR yang disesuaikan dengan *pre-processing* dan *tuning hyperparameter* menunjukkan hasil yang paling optimal. Pengujian pada IAM Dataset menunjukkan bahwa model dengan *pre-processing* dan *tuning hyperparameter* menghasilkan Character Error Rate (CER) terendah yaitu 0.0607. Hal ini menunjukkan bahwa proses pengolahan data yang tepat dapat meningkatkan kemampuan model dalam mengenali tulisan tangan.
- 2 Evaluasi pada formulir aplikasi kartu kredit Bank XYZ menunjukkan bahwa model dengan *pre-processing* dan *tuning hyperparameter* berhasil menurunkan CER secara signifikan. Model dengan CER 0.3620 jauh lebih baik dibandingkan dengan model lainnya, yang memiliki CER lebih tinggi (0.5213 untuk model tanpa *pre-processing*).
- 3 *Pre-processing* dan *tuning hyperparameter* memiliki dampak signifikan pada kualitas pengenalan karakter. Penyesuaian kecerahan, hue, serta penggunaan parameter beam search membantu meningkatkan ketepatan model dalam mengidentifikasi teks, khususnya dalam formulir kartu kredit yang mengandung berbagai gaya tulisan tangan.
- 4 TrOCR merupakan model yang efektif untuk pengenalan teks dalam formulir pendaftaran kartu kredit. Dengan menggunakan model ini, proses input data dapat diotomatisasi, mengurangi waktu dan biaya serta meminimalkan kesalahan manusia dalam pengolahan data.

DAFTAR PUSTAKA

- Bank Indonesia. (2023). Jumlah Pengguna Kartu Kredit di Indonesia. Diakses pada 1 Oktober 2024, dari <https://www.bi.go.id/id/statistik/ekonomi-keuangan/spip/Pages/SPIPSeptember-2023.aspx>
- Bank Indonesia. (2023). Transaksi Pengguna Kartu Kredit di Indonesia. Diakses pada 1 Oktober 2024, dari https://www.bi.go.id/id/statistik/ekonomi-keuangan/spip/Documents/TABEL_5d.pdf#search=kartu%20kredit
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Neural Information Processing Systems (NeurIPS)*, 30. <https://arxiv.org/abs/1706.03762>
- Li, M., Xu, M., Wu, Q., & Xie, X. (2021). TrOCR: Transformer-based Optical Character Recognition with Pre-trained Models. *arXiv*. <https://arxiv.org/abs/2109.10282>
- Diaz, D. H., Qin, S., Ingle, R. R., Fujii, Y., & Bissacco, A. (2021). Rethinking text line recognition models. *CoRR*, abs/2104.07787. <https://arxiv.org/abs/2104.07787>
- Kumar, B. (2021). ViT Cane: Visual assistant for the visually impaired. *CoRR*, abs/2109.13857. <https://arxiv.org/abs/2109.13857>

- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *CoRR*, abs/1810.04805. <https://arxiv.org/abs/1810.04805>
- Shuiqi, C., Yanliang, G., & Yue, C. (2021). A bank card number recognition method based on convolutional neural network. In *Proceedings of the 2021 IEEE 3rd International Conference on Communications, Information System and Computer Engineering (CISCE)* (pp. 826–830). <https://doi.org/10.1109/CISCE52179.2021.9445952>
- Xin, Y., Shi, P., & Han, S. (2019). An automatic location and recognition method for bank card number. In *Proceedings of the ACM International Conference* (pp. 728–732). <https://doi.org/10.1145/3366194.3366325>
- Sun, G., & You, F. (2020). Bank card number recognition system based on deep learning. In *Proceedings of the ACM International Conference on Bio Signals, Images, and Instrumentation* (pp. 745–749). <https://doi.org/10.1109/ICBSII51839.2021.9445136>
- Lin, Z., He, Z., Wang, P., Tan, B., Lu, J., & Bai, Y. (2020). SNRNet: A deep learning-based network for banknote serial number recognition. *Neural Process Letters*, 52(2), 1415–1426. <https://doi.org/10.1007/s11063-020-10313-9>
- Watson, W., & Liu, B. (2020). Financial table extraction in image documents. In *Proceedings of the First ACM International Conference on AI in Finance, ICAIF 2020* (pp. 1–8). <https://doi.org/10.1145/3383455.3422520>
- Rusli, F. M., Adhiguna, K. A., & Irawan, H. S. (2020). Indonesian ID card extractor using optical character recognition and natural language post-processing. *arXiv*. <http://arxiv.org/abs/2101.05214>
- Zhang, E., Putra, V. A., & Kusuma, G. P. (2022). Improving optical character recognition accuracy for Indonesia identification card using generative adversarial network. *Journal of Theoretical and Applied Information Technology*, 100(8). <https://www.jatit.org>
- Ma, X., & Yan, W. Q. (2021). Banknote serial number recognition using deep learning. *Multimedia Tools and Applications*, 80(12), 18445–18459. <https://doi.org/10.1007/s11042-020-10461-z>
- Srivastava, S., Priyadarshini, J., Gopal, S., Gupta, S., & Dayal, H. S. (2019). Optical character recognition on bank cheques using 2D convolution neural network. In *Advances in Intelligent Systems and Computing* (Vol. 697, pp. 589–596). Springer Verlag. https://doi.org/10.1007/978-981-13-1822-1_55
- Parthiban, R., Ezhilarasi, R., & Saravanan, D. (2020). Optical character recognition for English handwritten text using recurrent neural network. In *2020 International Conference on*

- System, Computation, Automation and Networking (ISSCAN) (pp. 1–6). IEEE.
<https://doi.org/10.1109/ICSCAN49426.2020.9262379>
- Narayan, A., & Muthalagu, R. (2021). Image character recognition using convolutional neural networks. In Proceedings of 2021 IEEE 7th International Conference on Bio Signals, Images and Instrumentation, ICBSII 2021 (pp. 1–5). IEEE.
<https://doi.org/10.1109/ICBSII51839.2021.9445136>
- Gao, Y., Xu, C., Shi, Z., & Zhang, H. (2019). Bank card number recognition system based on deep learning. In Proceedings of the ACM International Conference (pp. 1–5). ACM.
<https://doi.org/10.1145/3358331.3358383>
- Marti, U.-V., & Bunke, H. (2002). The IAM database: An English sentence database for offline handwriting recognition. *International Journal on Document Analysis and Recognition*, 5(1), 39–46. <https://www.tummy.com/xvscan/>
- Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A., & Jégou, H. (2020). Training data-efficient image transformers & distillation through attention. arXiv.
<https://arxiv.org/abs/2012.12877>
- Liu, Y., Ott, M., Goyal, N., Du, J., & Xiong, W. (2019). RoBERTa: A robustly optimized BERT pretraining approach. arXiv. <https://arxiv.org/abs/1907.11692>
- Chang, D., & Li, Y. (2024). DLoRA-TrOCR: Mixed text mode optical character recognition based on transformer. arXiv. <https://arxiv.org/abs/2404.12734>
- Stahlberg, F., & Byrne, B. (2019). On NMT search errors and model errors: Cat got your tongue? CoRR, abs/1908.10090. <https://arxiv.org/abs/1908.10090>
- Koehn, P., & Knowles, R. (2017). Six challenges for neural machine translation. CoRR, abs/1706.03872. <https://arxiv.org/abs/1706.03872>

Kutipan Artikel

Rafi Surya (2025), *Peningkatan Akurasi OCR dalam Pemrosesan Formulir Keuangan melalui Fine-Tuning Transformer dan Strategi Pra-pemrosesan Data*, JII, Vol: 07, No: 02, Hal: 01-12: April. DOI: <http://doi.org/10.51170/jii.v7i2>.