

APLIKASI PENCARI DATA IJAZAH SISWA BERBASIS IMAGE DENGAN METODE OCR STUDI KASUS PADA SMP PASUNDAN 1 BANDUNG

Oscar Hadikaryana¹, Yudiana²

Program Studi Teknik Informatika, Fakultas Ilmu komputer dan Sistem Informasi, Universitas Kebangsaan
Republik Indonesia, Bandung

Email: oscarhadikaryana@ukri.ac.id dan iduybmv@gmail.com

ABSTRAK

Kata kunci:

OCR, Optical Character Recognition, Template Matching Correlation, preprocessing

OCR (Optical Character Recognition) merupakan solusi efektif untuk mengubah dokumen cetak menjadi dokumen digital. Permasalahan yang muncul dalam proses pengenalan dokumen komputer adalah bagaimana teknik pengenalan untuk mengidentifikasi berbagai jenis karakter dengan berbagai ukuran dan bentuk. Metode pengenalan yang digunakan dalam tugas akhir ini adalah metode Template Matching Correlation. Sebelum dilakukan proses pengenalan, citra masukan dengan format *bmp atau jpg* diproses terlebih dahulu pada proses preprocessing yang meliputi binerisasi, segmentasi, dan normalisasi citra. Rata-rata tingkat keberhasilan pengenalan yang dihasilkan oleh sistem ini adalah 92,40%. Hasil akhir menunjukkan bahwa penggunaan metode Template Matching Correlation sudah cukup untuk membangun sistem OCR dengan akurasi yang baik secara efektif.

ABSTRACT

Keywords:

OCR, Optical Character Recognition, Template Matching Correlation, preprocessing

OCR (Optical Character Recognition) is an effective solution for converting printed documents to digital documents. The problem that arises in the process of recognizing computer documents is how the recognition technique is to identify various types of characters with various sizes and shapes. The introduction method used in this final project is the Template Matching Correlation method. Before the recognition process, the input image with the format * bmp or jpg * is processed first in the preprocessing process, which includes binaryzation, segmentation, and image normalization. The average recognition success rate generated by this system is 92.40%. The final results show that the use of the Template Matching Correlation method is sufficient to build an OCR system with good accuracy effectively.

PENDAHULUAN

Seiring perkembangan ilmu pengetahuan dan teknologi serta teknik informatika saat ini yang menampilkan kemajuan di berbagai bidang, terlihat dari adanya komputerisasi di berbagai bidang diantaranya di bidang Pendidikan (Calista, 2019). Tidak hanya peralihan proses kerja manual ke komputer, dunia pendidikan pun telah memanfaatkan fasilitas internet. Dari mulai perguruan tinggi, sekolah menengah atas, sekolah menengah pertama, sampai ke sekolah dasar. Di SMP Pasundan 1 Bandung, tempat Penulis melakukan penelitian, komputerisasi belum sepenuhnya terlaksana, masih banyak data yang pengelolannya dilakukan secara manual diantaranya

pengelolaan data kearsipan siswa. Pengelolaan data arsip siswa masih di simpan dalam bentuk foto copy arsip seperti ijazah yang diterbitkan oleh Dinas Pendidikan. Arsip ijazah tersebut memuat keterangan bukti akhir kelulusan seorang siswa atau alumni. Foto copy arsip ijazah tersebut masih rentan terhadap kerusakan dan dinilai tidak aman.

Berdasarkan paparan tersebut diatas Penulis bermaksud merancang sebuah program aplikasi komputer untuk pemrosesan arsip data siswa berbasis image dengan metode ocr di SMP Pasundan 1 Bandung. Penulis berharap dengan adanya software (perangkat lunak) tersebut pengelolaan arsip ijazah siswa di SMP Pasundan 1 Bandung dapat lebih efektif dan efisien (Al Ma'shum & Soemitra, 2022).

(Optical Character Recognition) OCR adalah sebuah aplikasi komputer yang digunakan untuk mengidentifikasi citra huruf maupun angka untuk dikonversi ke dalam bentuk file tulisan. Sistem pengenalan huruf ini dapat meningkatkan fleksibilitas atau kemampuan dan kecerdasan sistem komputer. Sistem pengenalan huruf yang cerdas sangat membantu usaha besar-besaran yang saat ini dilakukan banyak pihak yakni usaha digitalisasi informasi dan pengetahuan, misalnya dalam pembuatan koleksi pustaka digital, koleksi sastra kuno digital, dan lain-lain. Secara umum proses OCR dapat dilihat pada gambar 1, dengan penjelasan sebagai berikut:

a. File Input

File input berupa file citra digital dengan format *.bmp atau *.jpg.

b. Preprocessing

Preprocessing merupakan suatu proses untuk menghilangkan bagian-bagian yang tidak diperlukan pada gambar input untuk proses selanjutnya.

c. Segmentasi

Segmentasi adalah proses memisahkan area pengamatan (region) pada tiap karakter yang dideteksi.

d. Normalisasi

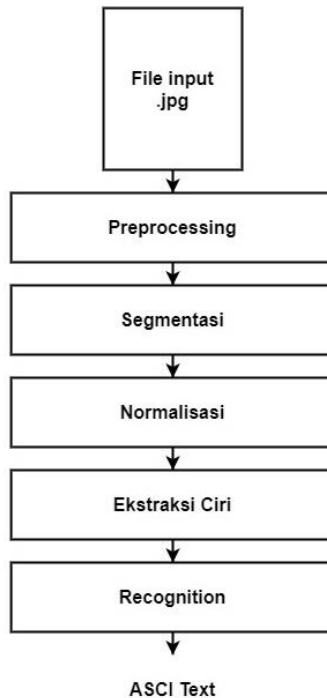
Normalisasi adalah proses merubah dimensi region tiap karakter dan ketebalan karakter.

e. Ekstraksi ciri

Ekstraksi ciri adalah proses untuk mengambil ciri-ciri tertentu dari karakter yang diamati.

f. Recognition

Recognition merupakan proses untuk mengenali karakter yang diamati dengan cara membandingkan ciri-ciri karakter yang diperoleh dengan ciri-ciri karakter yang ada pada basis data.



Gambar 1 Proses OCR secara umum.

METODE

Pengumpulan data, analisis dan perancangan merupakan suatu aktivitas yang harus dilakukan dengan sistematis dan teratur. Oleh karena itu diperlukan cara-cara khusus atau metode tertentu untuk memudahkan pelaksanaannya di lapangan sehingga diperoleh hasil yang optimal.

Metode penelitian yang Penulis gunakan dalam penyusunan laporan ini adalah metode kualitatif deskriptif (Gunawan, 2022). Penelitian kualitatif adalah penelitian dimana data yang diolah berbentuk data kualitatif seperti kata, kalimat atau gambar-gambar. Sedangkan penelitian deskriptif adalah penelitian untuk mengetahui nilai variabel mandiri baik satu atau lebih variabel tanpa membuat perbandingan atau hubungan dengan variabel lainnya.

Metode penelitian ini menggunakan data primer dan data sekunder yang diperoleh dengan melakukan penelitian lapangan. Untuk mendukung metode di atas Penulis melakukan beberapa langkah yaitu pengumpulan informasi melalui teknik kepustakaan (library technique), teknik observasi (field technique) dan teknik wawancara (interview) sebagai bahan analisis untuk memperoleh gambaran yang jelas terhadap objek penelitian. Kemudian melakukan reduksi dari data yang dikumpulkan dan disajikan dalam bentuk tabel, ataupun uraian penjelasan sehingga bisa diperoleh suatu kesimpulan dari objek yang diteliti.

HASIL DAN PEMBAHASAN

Data hasil pengenalan aplikasi OCR pada pengujian pertama (menggunakan huruf ukuran 20pt) dapat dilihat pada tabel 1.

Tabel 1. Persentase hasil pengenalan pada pengujian

Citra masukan	Jumlah karakter yang dikenali	Total karakter pada citra	Persentase keberhasilan
Citra_Uji_Arial.jpg	249	263	94,68%
Citra_Uji_Times.jpg	223	263	84,79%
Citra_Uji_Calibri.jpg	257	263	97,72%
Rata-rata persentase keberhasilan			92,40%

Dari tabel. 1 dapat diketahui bahwa rata-rata persentase keberhasilan pengenalan pada pengujian sebesar 92,40%. Tingkat keberhasilan pengenalan yang dihasilkan cukup tinggi meskipun jenis huruf yang digunakan sebagai masukan berbeda dengan template.

KESIMPULAN

Kesimpulan yang dapat diambil dalam pembuatan tugas akhir ini adalah, pertama, dihasilkan sebuah aplikasi Optical Character Recognition menggunakan algoritma Template Matching Correlation. Kedua, algoritma Template Matching Correlation cukup efektif untuk pengenalan karakter huruf cetak. Rata-rata tingkat keberhasilan pengenalan yang dihasilkan sebesar 92,40%.

Saran

Demikian jurnal ilmiah ini disusun dengan segala kekurangan dan butuh penyempurnaan lagi dikemudian hari, sehingga sistem yang dikembangkan mampu menunjukan kondisi yang sesungguhnya diberbagai tempat atau wilayah. Untuk itu dalam kesempatan ini disampaikan ucapan terima kasih kepada pihak-pihak yang telah membantu kelancaran penyusunan jurnal ilmiah ini, diantaranya: Fakultas Ilmu Komputer dan Sistem Infomasi, Program Studi Teknik Informatika UKRI dan Keluarga besar SMP Pasundan 1 Bandung yang telah memberikan dukungan moril dalam penyelesaian jurnal ilmiah ini.

DAFTAR PUSTAKA

- Al Ma'shum, N. H. H., & Soemitra, A. (2022). Analisis Peranan Badan Perencanaan Pembangunan Daerah Sumatera Utara (Bappeda Su) Terhadap Pembangunan Daerah Di Wilayah Sumatera Utara. *JIEM: Jurnal Ilmu Komputer, Ekonomi Dan Manajemen*, 2(2), 3396–3400.
- Brunelli Roberto, 2009, "Template Matching Techniques In Computer Vision", John Willey & Sons. Inc.
- Calista, A. (2019). Kemampuan dan Kedisiplinan Kerja terhadap Kepuasan Kerja Serta Dampaknya pada Kinerja Pegawai Dinas Perhubungan, Komunikasi dan Informatika Kabupaten Musi Banyuasin. *Jurnal Ilmiah Ekonomi Global Masa Kini*, 9(2), 145–150.
- Cheriet M., Kharma N., Liu C., Suen C.Y.,2006, "Character Recognition System A Guide

- for Student and Practioners”, John Willey & Sons. Inc.
- Eikvil Line, 1993, “Optical Character Recognition”, Nork Regnesentral.
- Gonzales Rafael C. dan Woods Richard E.,2001, “Digital Image Processing”, Prentice Hall.
- Gunawan, I. (2022). *Metode Penelitian Kualitatif: teori dan praktik*. Bumi Aksara.
- Hendry J., 2011, “Deteksi Kkarakter Pada Citra Digital (OCR) dengan Teknik Korelasi (Template Matching)”, \http://www.scribd.com/doc/68866773/Project-Deteksi-Karakter-Pada-Citra-DigitalOcr-Dengan-Teknik-Korelasi-TemplateMatching.
- Keys, Robert G., 1981, "Cubic Convolution Interpolation for Digital Image Processing", IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-29.
- MathWorks, _____, “Convert RGB image or colormap to grayscale”, diakses dari <http://www.mathworks.com/help/images/ref/rgb2gray.html>
- Moler, 2004, “MATLAB About”, diakses dari <http://www.mathworks.com>.
- Munir, Rinaldi, 2004, “Pengolahan Citra Digital dengan Pendekatan Algoritmik”, Informatika, Bandung.
- Nurwanto T. B., 2007, “Pengenalan Huruf Tulisan Tangan Menggunakan Logika Fuzzy Dengan Pendekatan Neural Networks Back Propagation”, STT Telkom, Bandung.
- Parno, _____, “DFD”, diakses dari http://parno.staff.gunadarma.ac.idDownloadsfiles4395SI_03_DFD.pdf.
- Purnomo, Seno, 2009, “Interpolasi”, diakses dari <http://senosuke.wordpress.com/2009/12/22/interpolasi/>
- Putra, Darma, 2010, “Pengolahan Citra Digital”, Penerbit Andi, Yogyakarta.
- Sdarsono, _____, “Flowchart”, diakses dari sdarsono.staff.gunadarma.ac.id/Downloads/files/16512/Flowchart.pdf.
- Sofani, Rachmah, 2009, “Sistem OCR”, http://digilib.itelkom.ac.id/index.php?option=com_content&view=article&id=413:sistem-ocr&catid=15:pemrosesansinyal&Itemid=14.
- Sugiharto, Aris, 2006, “Pemrograman GUI dengan MATLAB”, Penerbit Andi, Yogyakarta.
- Wardhana A. W., Prayudi Y., 2008, “Penggunaan Metode Template Matching Untuk Identifikasi Kecacatan Pada PCB”, Universitas Islam Indonesia, Yogyakarta.
- Wirayuda T. A. B., Vaulin S., dan Novi D.R. 2009,“Pengenalan Huruf Komputer Menggunakan Algoritma Berbasis Chain Code dan Algoritma Sequence Alignment”, IT Telkom, Bandung.