

Graphical abstract



SISTEM INFORMASI DATA SISWA BERBASIS OPTICAL CHARACTER RECOGNITION (OCR)

^{1*}Muhammad Nafsin, ¹A. Ahmad Qashlim, ¹UL Khairat.

¹Universitas Al Asyariah Mandar.

**Corresponding author*

nafsinm@gmail.com

Abstract

Student data is a unique identity that contains information about students such as name, date of birth, identity number etc. which is used for school administration purposes. OCR (Optical Character Recognition) is a computer system algorithm that is used to automatically recognize a series of text characters in an image. Using OCR (Optical Character Recognition) can help operators work in automatically inputting student data. The research method used is a questionnaire, literature study, observation and interviews at SMK Bina Harapan. In the proposed system, the admin inputs image files such as diplomas, SKHU, Birth Certificates and Family Cards to be scanned using OCR (Optical Character Recognition) where the result is data in the form of text. After testing on sample images, the success rate of diploma scans is 60%, SKHU scans are 95% successful, Family Cards are 60%, and Birth Certificates are 30%. The success rate of different scans is influenced by the level of image quality and the level of character complexity. From this level of success, this application meets the criteria for use because the text results can be processed and the application runs according to its purpose.

Keywords: Student Data, OCR (Optical Character Recognition), Information System.

Abstrak

Data siswa adalah identitas unik yang berisi informasi tentang siswa seperti nama, tanggal lahir, nomor identitas dll yang digunakan untuk keperluan administrasi sekolah. OCR (*Optical Character Recognition*) adalah *algoritma system* komputer yang digunakan secara otomatis untuk mengenali serangkaian karakter teks pada sebuah gambar. Dengan menggunakan OCR (*Optical Character Recognition*) dapat membantu pekerjaan operator dalam penginputan data siswa secara otomatis. Metode penelitian yang digunakan adalah kuesioner, studi Pustaka, observasi dan wawancara di SMK Bina Harapan. Pada sistem yang di usulkan admin melakukan penginputan file gambar seperti Ijazah, SKHU, Akta Kelahiran dan Kartu Keluarga untuk di scan dengan menggunakan OCR (Optical Character Recognition) yang mana hasilnya adalah data berupa *text*. Setelah uji coba pada sampel gambar, tingkat keberhasilan *scan* Ijazah 60%, SKHU berhasil *scan* 95%, Kartu Keluarga 60%, dan Akta Kelahiran 30%. Tingkat keberhasilan *scan* yang berbeda – beda tersebut dipengaruhi tingkat kualitas gambar dan tingkat kerumitan karakter. Dari tingkat keberhasilan tersebut, aplikasi ini memenuhi kriteria layak digunakan karena hasil *text* dapat diolah dan aplikasi berjalan sesuai tujuan.

Kata kunci: Data Siswa, OCR (*Optical Character Recognition*), Sistem Informasi.

Article history

DOI: <http://dx.doi.org/10.35329/jp.v4i1.2201>

Received : 09 Agust 2021 | Received in revised form : 27 Mei 2022 | Accepted : 31 Mei 2022

1. PENDAHULUAN

Sekolah adalah tempat berlangsungnya pendidikan dan pembelajaran siswa di dengan bimbingan guru. Sebagian besar negara memiliki sistem pendidikan yang berbeda. Di Indonesia, sistem pendidikan melalui pendataan yang akurat dan unit karena setiap siswa harus memiliki data kredensial yang valid, yang harus diintegrasikan ke dalam basisdata yang diterbitkan oleh Capil dalam bentuk kartu keluarga dan akta kelahiran. Selain itu, data siswa juga harus valid dan terintegrasi dengan sekolah sebelumnya, baik lulusan SD, SMP, dan SMA. Untuk mengurus data siswa tersebut Kepala Sekolah menunjuk seorang guru atau staf tata usaha sebagai Operator Sekolah melalui SK (surat keputusan) yang menugaskan operator sekolah untuk mengurus data siswa dan Dapodik. Lesmana, C., Arifin, A., & Santoso, D. (2017).

OCR (*Optical Character Recognition*) adalah teknologi/algorithm pemindaian. OCR (*Optical Character Recognition*) merupakan algoritma *coding* dan program yang dimanfaatkan untuk mendeteksi sebuah teks, angka, maupun sebuah pola karakter dari gambar, Alfariis, R. E. (2020).

Ia mampu membedakan antara rekaman dan gambar yang dipindai Gambar yang dipindai dapat berupa tulisan tangan atau digital dan dapat diproses. OCR (*Optical Character Recognition*) adalah sistem komputer yang dapat membaca karakter dari printer (printer atau mesin tik) atau dari tulisan tangan. Aplikasi OCR biasanya digunakan untuk mengidentifikasi gambar teks dan mengonversinya untuk menyimpan file. Utami, A. E., Nurhayati, O. D., & Martono, K. T. (2016).

Qashlim, A., & Hasruddin, H. (2015). Penelitian tentang "Implementasi Teknologi QR-Code Untuk Kartu Identitas". Penelitian ini bertujuan untuk memudahkan mahasiswa melakukan pengecekan identitas pribadi, riwayat, status kuliah dll dengan menggunakan kartu mahasiswa yang terdapat QR-Code.

Muharom, Syahri, 2019, melakukan penelitian tentang "Sistem Pengenalan Nomor Kamar Menggunakan OCR dan Teknologi Pencocokan Pola". Peneliti membuat aplikasi yang dapat mengenali setiap bagian secara digital menggunakan kamera, optical character recognition dan aplikasi pattern matching. Pengembangan sistem ini bertujuan untuk mempermudah identifikasi nomor kamar dengan kamera.

Misbah Riyandi Fauzi, dkk, 2018, melakukan penelitian tentang "Mengubah tulisan tangan menjadi text digital OCR dengan menggunakan metode segmentasi dan korelasi". Hasil dari penelitian ini adalah mengubah tulisan tangan menjadi teks tingkat keberhasilan identifikasi 81%

tergantung tingkat ketebalan dan dan kerumitan sampel tulisan yang di uji.

Setiawan, A. A., & Sudigdo, A., 2019, "Penguatan literasi siswa sekolah dasar melalui kunjungan perpustakaan". Penelitian implementasi Optical Character Recognition (OCR) dari penerjemah bahasa Indonesia ke bahasa Inggris. Ini menggunakan pengenalan karakter optik, Android dan metode terjemahan mesin.

Muhtadi Hilmi A. Tawakkal, 2018, Melakukan penelitian tentang "Pengembangan aplikasi Android untuk pengenalan citra digital menggunakan sertifikat Halal MUI menggunakan metode Tesseract Optical Character Recognition (OCR) perpustakaan." Penelitian ini dilakukan untuk membuat sistem yang menggunakan teknologi OCR dan ekstraksi teks untuk memungkinkan pemasukan data klaim secara otomatis.

Yan Puspitarani dan Yenie Syukriyah, 2020, melakukan penelitian tentang "Gunakan OCR untuk mengekstrak fitur dari teks untuk membuat database keluhan di tempat kerja. Penelitian ini dilakukan untuk membuat sistem pengisian data pengaduan secara otomatis menggunakan teknologi OCR dan kemampuan text mining.

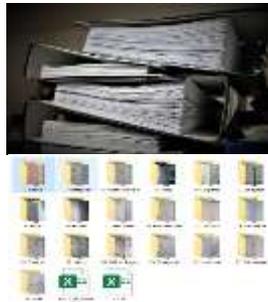
Hasil dari penelitian penulis berfokus pada aplkiasi berbasis web dengan menggunakan tekhnologi *Optical Character Recognition* (OCR) untuk mengubah (*Hardcopy*) Ijazah, SHUN, KK, dan Akta kedalam bentuk teks (*Softcopy*) serta memilah teks yang dibutuhkan dalam data siswa.

2. METODE PENELITIAN

A. Pengumpulan Data

Pengumpulan data yang dilakukan penulis untuk memperoleh informasi yang membantu proses penelitian sebagai berikut:

1. Kuesioner, peneliti membuat daftar pertanyaan untuk menggali informasi yang dibutuhkan dan memperoleh jawaban baik dari siswa, guru, kepala sekolah maupun operator sekolah itu sendiri
2. Studi Pustaka, peneliti mengumpulkan data siswa yang dibutuhkan seperti document hardcopy dan softcopy berupa Ijazah, SHUN, KK dan Akta Siswa yang akan di uji coba scan dengan aplikasi berbasis OCR.



Gambar 2.1 Data Siswa *hardcopy* dan *softcopy*

3. Observasi, peneliti mengamati, menganalisis metode pengelolaan data siswa yang sudah ada baik dari cara kerjanya serta melihat kekuatan, kelemahan, peluang dan kesempatan dalam system agar menjadi gagasan penelitian.
4. Wawancara, peneliti melakukan proses pendalaman informasi dengan cara membuat sesi wawancara terhadap narasumber yang secara teknis bersentuhan langsung dengan bidang penelitian pada tahapan ini yang menajadi narasumber ialah kepala sekolah dan operator sekolah sebagai sumber pengelolaan data siswa di SMK Bina Harapan.

Metode yang ada masih menggunakan cara manual dalam input data sehingga membutuhkan lembar kertas data siswa dan sering terjadi kesalahan dalam input data. Selain itu perkembangan teknologi *hardcopy* mulai beralih ke *softcopy* yang lebih praktis juga menjadi inovasi karena pemanfaatan pemrosesan citra untuk keperluan otomasis pendataan masih jarang digunakan. Utama, S. W., & Kusumawardhani, A. (2017).

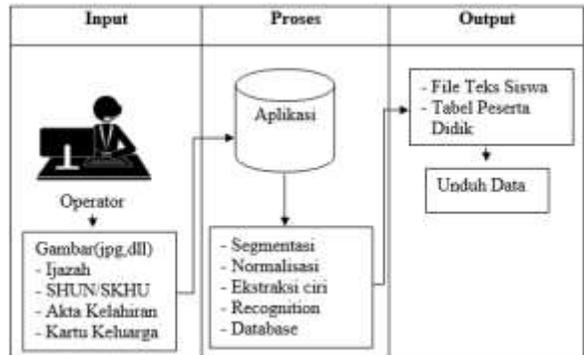
B. Tahapan Penelitian

Peneliti menggali informasi dengan melakukan proses investigasi masalah, perancangan system, pengumpulan data, revisi system, app web dan laporan evaluasi yang dilaksanakan di SMK Bina Harapan dilihat pada gambar 2.2 berikut:



Gambar 2.2 Tahapan Penelitian

C. Kerangka Sistem



Gambar 2.3 Kerangka Sistem

Kerangka sistem di atas menjelaskan alur Sistem Informasi Data Siswa berbasis OCR (Optical Character Recontion) pada SMK Bina Harapan mulai dari tahapan input data, prose OCR dan output yang berupa hasil dan laporan.

Pada tahapan input, operator tidak lagi mengetik/ input secara manual identitas siswa tetapi hanya perlu upload foto data siswa untuk di scan OCR (Optical Character Recognition). Selanjutnya, aplikasi memproses data tersebut untuk di ubah dari gambar menjadi text. Dan setelah berhasil di scan data siswa tersebut menjadi sebuah tabel berisi sistem informasi yang menyajikan data sesuai kebutuhan juga dapat di download dalam bentuk PDF dan Excel sebagai output aplikasi.

3. HASIL DAN PEMBAHASAN

Berdasarkan analisis dan perancangan sistem yang dilakukan, maka hasil dari penelitian ini adalah penerapan sistem informasi data siswa berbasis OCR (Optical Character Recognition) di SMKS Bina Harapan. Menggunakan bahasa pemrograman PHP dan MySQL sebagai databasanya.

Penelitian ini mengusulkan sistem yang dapat menscan data siswa berupa file gambar diolah menjadi gambar text yang dapat diedit kembali juga dapat sebagai database informasi siswa dan alumni yang bersekolah di SMKS Bina Harapan.

a. *Form Dashboard*

Form ini digunakan admin untuk mengelola submenu dan penyajian informasi aplikasi seperti total jumlah data yang terdapat di dalam program seperti pada gambar 3.1.



Gambar 3.1 *Form Dashboard*

b. *Form Master Data Scan OCR*

Form berikut adalah form yang digunakan operator untuk menampilkan hasil *scan* data siswa dari file gambar yang dikonversi kedalam bentuk teks, seperti pada gambar 3.2.



Gambar 3.2 *Form Master Data Scan OCR*

c. *Form Hasil Scan OCR*

Form berikut adalah form yang digunakan operator sekolah untuk menampilkan hasil *scan* data siswa dari

file gambar yang dikonversi kedalam bentuk teks, seperti pada gambar 3.3.



Gambar 3.3 *Form Scan*

Pada prose ini, Scan OCR berjalan dengan melalui proses Citra asli, *Greyscale*, *Smoothing* dan *Thresholding*. Berikut adalah gambar proses yang berlangsung :



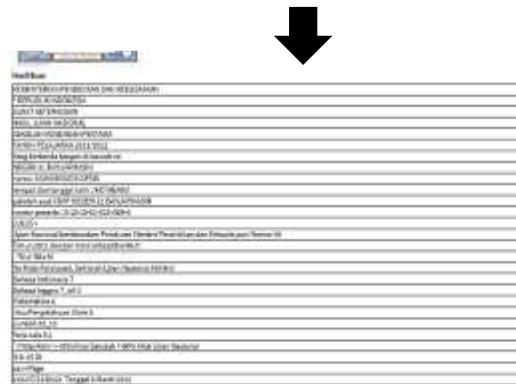
1. Citra Asli

2. Grayscale



3. Smoothing

4. Thresholding



Gambar 3.4 Hasil *Scan OCR*

d. *Form Data Siswa*

Form Data Siswa ialah data yang berhasil dikumpulkan dari hasil *scan OCR* pada

gambar 3.4. Berikut tampilan Data Siswa seperti terlihat pada Gambar 3.5.

Gambar 3.5 *Form* Data Siswa

- e. *Form* Tambah Data Siswa Baru
Form ini digunakan operator untuk menambah data siswa secara manual pada *website* seperti pada gambar 3.6.

Gambar 3.6 Tambah Data Baru

- f. *Form* Tentang Aplikasi
Form Tentang Aplikasi adalah form informasi tanggal, siapa pembuatnya, dan tujuan dibuat aplikasi lihat pada gambar 3.7.

Gambar 3.7 Tentang Aplikasi

4. SIMPULAN

Kesimpulan dari hasil penelitian ini terbangunnya Sistem Informasi Data Siswa Berbasis OCR (*Optical Character Recognition*) pada SMK Bina Harapan. Program ini akan terus dikembangkan oleh peneliti sehingga benar-benar menjadi program yang layak digunakan untuk membantu pekerjaan operator dibidang keakuratan identitas siswa dan sebagai *database* siswa yang bersekolah di SMK Bina Harapan.

Setelah di uji coba kelayakan, hasil penelitian ini berhasil men-*scan* file gambar(jpg,png dll) ke bentuk *text* yang mana hasil tersebut dapat di olah kedalam bentuk Pdf dan Excel. Adapun kekurangannya terletak pada ke akuratan hasil *scan* yang belum maksimal sehingga terdapat beberapa karakter pada gambar yang belum mampu terbaca, selain itu pada hasil *scan* kadang terdapat karakter yang seharusnya tidak ada atau kesalahan dalam proses *scan*. Setelah uji coba pada sampel gambar, tingkat keberhasilan *scan* Ijazah 60%, SKHU berhasil *scan* 95%, Kartu Keluarga 60%, dan Akta Kelahiran 30%. Tingkat keberhasilan *scan* yang berbeda – beda tersebut dipengaruhi tingkat kualitas gambar dan tingkat kerumitan karakter. Dari tingkat keberhasilan tersebut, aplikasi ini memenuhi kriteria layak digunakan karena hasil *text* dapat diolah dan aplikasi berjalan sesuai tujuan.

UCAPAN TERIMAKASIH

Ucapan terimakasih dari peneliti kepada Fakultas Ilmu Komputer, LPPM Universitas AI-Asyariah Mandar yang membantu peneliti dalam bimbingan penelitian dan administrasi penelitian. Juga ucapan terimakasih kepada SMK Bina Harapan sebagai tepat penelitian yang membantu memberikan sumber informasi yang dibutuhkan peneliti.

DAFTAR PUSTAKA

- Alfarisi, R. E. (2020). Rancang Bangun Aplikasi Terjemahan Bahasa Jepang-Indonesia Berbasis Android Menggunakan Tesseract Ocr (Doctoral dissertation, Universitas Komputer Indonesia).
- Lesmana, C., Arifin, A., & Santoso, D. (2017). Tingkat Pemahaman Operator Sekolah Pada Aplikasi Dapodik Di Kecamatan Pontianak Selatan. *Jurnal Pendidikan Informatika dan Sains*, 5(2), 184-194.
- Muhtadi Hilmi A. Tawakkal, 2018, "*Pengembangan aplikasi Android untuk pengenalan citra digital sertifikat Halal MUI menggunakan*

metode Tesseract Optical Character Recognition (OCR) perpustakaan". Program Studi Teknik Informatika, Sekolah Tinggi Teknologi Terpadu Nurul Fikri.

- Misbah Riyandi Fauzi, dkk, 2018, *"Mengubah tulisan tangan menjadi text digital OCR dengan menggunakan metode segmentasi dan korelasi"*. TRANSIENT, vo.2 No.4, Desember 2013, ISSN: 2302-9927, 1014.
- Muharom, Syahri, 2019, *"Sistem Pengenalan Nomor Ruangan menggunakan teknologi OCR dan Template Matching"*. Jurnal Ilmiah Teknologi Informasi dan Komunikasi, P-ISSN: 2502-3470, E-ISSN: 2581-0367
- Qashlim, A., & Hasruddin, H. (2015). *"Implementasi Teknologi QR-Code Untuk Kartu Identitas"*. Jurnal Ilmu Komputer Fakultas Ilmu Komputer Universitas Al Asyariah Mandar, 1(2), 1-6.
- Setiawan, A. A., & Sudigdo, A., 2019, *"Penguatan literasi siswa sekolah dasar melalui kunjungan perpustakaan"*. In Prosiding Seminar Nasional PGSD UST (Vol. 1).
- Utama, S. W., & Kusumawardhani, A. (2017). *Aplikasi Pendeteksi Plat Nomor Negara Indonesia Menggunakan OpenCV dan Tesseract OCR pada Android Studio*. no. December.
- Utami, A. E., Nurhayati, O. D., & Martono, K. T. (2016). *Aplikasi Penerjemah Bahasa Inggris-Indonesia dengan Optical Character Recognition Berbasis Android*. Jurnal Teknologi dan Sistem Komputer, 4(1), 167-177.
- Yan Puspitarani dan Yenie Syukriyah, 2020, *"Pemanfaatan OCR dan Text Feature Extraction untuk membangun Basis Data Pengaduan Tenaga Kerja"*. Vol.1 Tahun 2017 s.d Vol.5 No.3 2021