

IMPLEMENTASI OPTICAL CHARACTER RECOGNITION DALAM PENGOLAHAN DATA STATISTIK SURVEI PENDUDUK DI BADAN PUSAT STATISTIK KOTA TEBING TINGGI

Sandrina Malakiano Ritonga^{*1}, Dwi Andini², Intan Janurianty³, Rini Halila Nst⁴

^{1,2,3,4}Universitas Islam Negeri Sumatera Utara, Indonesia

Email: ¹sandrinamalakianoritonga@gmail.com, ²andinidwi097@gmail.com, ³intanpermata199005@gmail.com

Abstrak

Implementasi teknologi Optical Character Recognition (OCR) di lingkungan Badan Pusat Statistik (BPS) Kota Tebing Tinggi bertujuan untuk meningkatkan efisiensi dan akurasi dalam pengolahan data survei penduduk. Teknologi ini memungkinkan konversi dokumen fisik menjadi data digital secara otomatis, sehingga mampu mengurangi beban kerja manual dan meminimalkan kesalahan input data. Penelitian ini menggunakan pendekatan studi kasus deskriptif kualitatif melalui observasi langsung dan wawancara dengan staf BPS. Hasil penelitian menunjukkan bahwa penggunaan aplikasi OCR seperti CamScanner mampu mempercepat proses input data dan mendukung transformasi digital di institusi pemerintah. Keberhasilan implementasi dipengaruhi oleh kesiapan infrastruktur, pemilihan perangkat yang sesuai, pelatihan operator, serta integrasi sistem OCR ke dalam sistem kerja yang ada.

Kata kunci: Optical Character Recognition, BPS, CamScanner

IMPLEMENTATION OF OPTICAL CHARACTER RECOGNITION IN THE PROCESSING OF POPULATION SURVEY STATISTICAL DATA AT THE CENTRAL STATISTICS AGENCY OF TEBING TINGGI CITY

Abstract

The implementation of Optical Character Recognition (OCR) technology in the Central Statistics Agency (BPS) of Tebing Tinggi City aims to improve efficiency and accuracy in processing population survey data. This technology allows the conversion of physical documents into digital data automatically, thereby reducing manual workload and minimizing data input errors. This study uses a qualitative descriptive case study approach through direct observation and interviews with BPS staff. The results of the study show that the use of OCR applications such as CamScanner can accelerate the data input process and support digital transformation in government institutions. The success of the implementation is influenced by infrastructure readiness, selection of appropriate devices, operator training, and integration of the OCR system into existing work systems.

Keywords: Optical Character Recognition, BPS, CamScanner

1. PENDAHULUAN

Badan Pusat Statistik (BPS) sebagai lembaga pemerintah non-kementerian memiliki peran strategis dalam menyediakan data statistik yang akurat dan terpercaya. Di era digital seperti saat ini, tuntutan akan kecepatan dan ketepatan dalam pengolahan data semakin tinggi. Namun, proses entri data hasil survei di BPS masih banyak yang dilakukan secara manual, terutama pada level daerah seperti Kota Tebing Tinggi. Hal ini menyebabkan waktu pengolahan yang lama serta tingginya risiko kesalahan input. Kota Tebing Tinggi sebagai salah satu wilayah administratif di Provinsi Sumatera Utara memiliki kebutuhan statistik yang cukup kompleks, terutama dalam menyusun data sektoral yang dibutuhkan oleh berbagai instansi pemerintah. Volume dokumen survei dan sensus yang harus diproses setiap tahun cukup besar, yang apabila tidak didukung oleh

teknologi yang tepat dapat menghambat kualitas pelayanan data. Oleh karena itu, implementasi OCR dapat menjadi kunci penting untuk menjawab tantangan tersebut. Optical Character Recognition (OCR) hadir sebagai salah satu solusi teknologi yang mampu mengenali teks dari dokumen fisik lalu mengubahnya menjadi data digital. Teknologi ini telah digunakan luas di berbagai sektor untuk mendukung proses digitalisasi dokumen, termasuk dalam bidang administrasi publik dan pengolahan arsip. OCR mampu menghemat waktu entri data serta meminimalisir kesalahan dengan tingkat akurasi yang cukup tinggi jika dibandingkan dengan proses manual. Proses manual dalam entri data memerlukan ketelitian tinggi, tetapi tetap tidak luput dari kesalahan manusia seperti kesalahan ketik, duplikasi data, atau hilangnya lembar jawaban survei. Selain itu, metode ini juga tidak efisien karena

membutuhkan banyak sumber daya manusia dan waktu yang cukup panjang untuk menyelesaikan pengolahan data dalam jumlah besar. Akibatnya, keterlambatan dalam penyajian data statistik sering terjadi, yang kemudian berdampak pada lambatnya pengambilan kebijakan berbasis data. Penelitian ini dilakukan untuk mengkaji secara lebih dalam dampak dari penerapan teknologi OCR terhadap efisiensi dan akurasi pengolahan data statistik di BPS Kota Tebing Tinggi. Dengan pendekatan studi kasus, penelitian ini bertujuan memberikan gambaran nyata mengenai efektivitas OCR dalam proses digitalisasi data survei serta kontribusinya terhadap transformasi digital di instansi pemerintah. Temuan dari studi ini diharapkan dapat memberikan masukan bagi pengembangan teknologi pengolahan data di sektor publik secara lebih luas.

2. METODE PENELITIAN

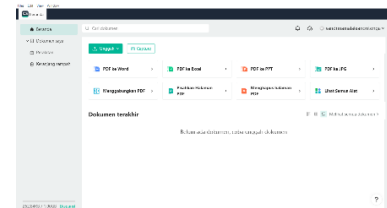
Penelitian ini menggunakan pendekatan studi kasus deskriptif kualitatif yang dilaksanakan di BPS Kota Tebing Tinggi. Teknik pengumpulan data dilakukan melalui observasi langsung terhadap alur proses pengolahan data sebelum dan sesudah penerapan OCR, serta wawancara kepada staf pengolahan data dan tim teknologi informasi.

3. HASIL DAN PEMBAHASAN

Data statistik yang akurat memiliki peran yang sangat penting sebagai elemen kunci dalam proses pengambilan keputusan yang berbasis bukti dan dapat dipertanggungjawabkan. Keputusan-keputusan ini mencakup berbagai aspek kehidupan dan dilakukan oleh beragam pihak, termasuk pemerintah yang merancang kebijakan publik, sektor swasta yang menentukan strategi bisnis, hingga masyarakat luas yang memanfaatkan data untuk berbagai kebutuhan. Dalam konteks ini, Badan Pusat Statistik (BPS) sebagai penyedia data statistik resmi memiliki tanggung jawab besar untuk menghasilkan data yang tidak hanya relevan, tetapi juga konsisten dan akurat. Data yang berkualitas tersebut diperlukan guna mendukung upaya pembangunan nasional secara menyeluruh, termasuk perencanaan kebijakan publik yang efisien, transparan, dan berorientasi pada hasil yang maksimal demi kesejahteraan masyarakat. Optical Character Recognition (OCR) adalah sebuah teknologi canggih yang dirancang untuk membaca dan mengenali teks dari dokumen fisik, seperti formulir survei, dokumen cetak, atau bahkan tulisan tangan. Teknologi ini memungkinkan teks yang terdapat pada dokumen fisik tersebut diubah menjadi format digital yang dapat dengan mudah diolah, dianalisis, dan disimpan untuk keperluan lebih lanjut. Implementasi OCR menjadi solusi yang sangat penting, terutama dalam mengatasi berbagai tantangan yang sering muncul dalam pengolahan data secara manual, seperti proses yang memakan waktu lama, biaya operasional yang tinggi, serta risiko

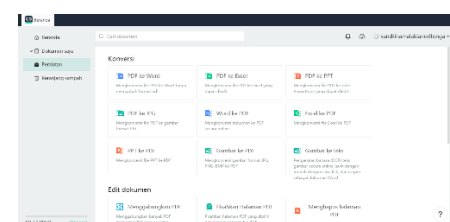
kesalahan yang lebih besar akibat faktor manusia. Dengan menggunakan teknologi OCR, instansi dapat meningkatkan efisiensi kerja, memastikan akurasi data yang lebih baik, serta mempercepat proses pengumpulan dan analisis data untuk mendukung pengambilan keputusan yang lebih cepat dan tepat. Dalam mendukung proses digitalisasi dokumen survei penduduk, CamScanner menjadi salah satu aplikasi yang direkomendasikan. Aplikasi mobile berbasis OCR ini mampu memindai dokumen fisik melalui kamera ponsel dan secara otomatis mengekstrak teks ke dalam format digital. Solusi ini praktis dan efisien, terutama bagi BPS yang menangani data dalam jumlah besar dan waktu terbatas. CamScanner sangat praktis digunakan di lapangan tanpa memerlukan perangkat tambahan seperti scanner desktop. Fitur-fiturnya mencakup koreksi perspektif, peningkatan kualitas gambar, dan ekstraksi teks yang cukup akurat, bahkan dari dokumen yang kurang jelas. Teks yang dihasilkan dapat langsung disalin atau disimpan dalam berbagai format, sehingga mendukung percepatan input data dan mengurangi kesalahan manual. Bagi BPS Kota Tebing Tinggi, aplikasi ini merupakan solusi efisien untuk mempercepat proses pengolahan data sekaligus mendorong transformasi digital. Berikut ini adalah contoh hasil penggunaan CamScanner dalam proses digitalisasi data survei:

1. Berikut merupakan tampilan awal (dashboard) aplikasi CamScanner dengan berbagai fitur utama yang tersedia.



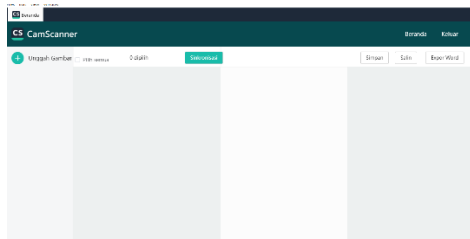
Gambar 1. Tampilan Awal Aplikasi CamScanner

2. Kemudian pilih peralatan yang tersedia dan melanjutkan dengan fitur konversi gambar ke teks (OCR)



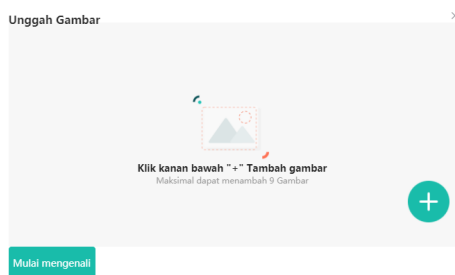
Gambar 2. Tampilan Peralatan Konversi

3. Berikut ini adalah tampilan dari fitur konversi gambar ke teks (OCR)



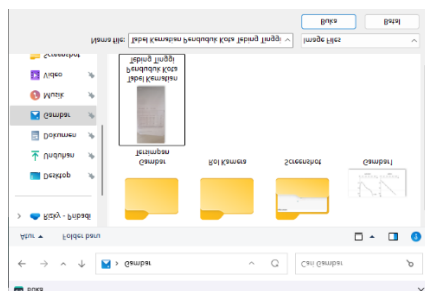
Gambar 3. Tampilan dari fitur konversi gambar ke teks (OCR)

4. Selanjutnya, pilih opsi 'Unggah Gambar' dan klik ikon tambah



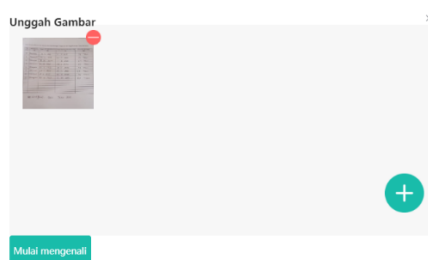
Gambar 4. Tampilan Unggah Gambar

5. Selanjutnya, pilih gambar yang akan dikonversi menjadi teks.



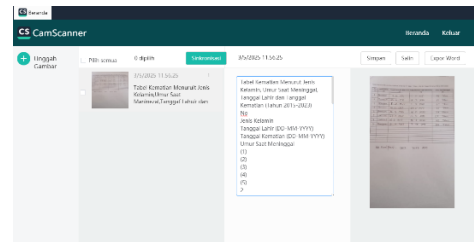
Gambar 5. Tampilan setelah memilih gambar

6. Setelah memilih gambar, akan muncul tampilan seperti berikut. Selanjutnya, klik tombol 'Mulai Mengenali' untuk memulai proses konversi teks.



Gambar 6. Tampilan proses konversi teks

7. Berikut ini adalah tampilan hasil dari konversi gambar menjadi teks. Teks yang dihasilkan dapat disalin dan langsung dimasukkan ke dalam aplikasi yang digunakan untuk pendataan penduduk di BPS Kota Tebing Tinggi



Gambar 7. Tampilan hasil dari konversi gambar menjadi teks

Implementasi teknologi *Optical Character Recognition* (OCR) memerlukan langkah-langkah yang sistematis dan terencana agar dapat memberikan hasil yang optimal, khususnya bagi institusi seperti Badan Pusat Statistik (BPS), antara lain:

a. Evaluasi Kebutuhan Teknologi

Langkah pertama yang harus dilakukan adalah melakukan analisis terhadap kebutuhan teknologi yang spesifik, yang mencakup jenis dokumen yang paling sering digunakan, seperti formulir survei atau kuesioner. Badan Pusat Statistik perlu memastikan bahwa perangkat OCR yang dipilih dapat bekerja secara efektif untuk mendigitalisasi dokumen-dokumen tersebut. Selain itu, perlu dilakukan penentuan perangkat keras yang tepat, seperti pemindai (*scanner*) berkecepatan tinggi yang dapat memproses dokumen dalam jumlah besar dengan cepat dan efisien. Tidak kalah penting, perangkat lunak OCR yang digunakan harus memiliki kemampuan untuk mengenali berbagai format teks, tabel, dan elemen lainnya, sehingga dapat menghasilkan data digital yang akurat dan berkualitas.

b. Pengadaan Perangkat

Setelah kebutuhan teknologi ditentukan, langkah berikutnya adalah melakukan proses

pengadaan perangkat keras dan perangkat lunak yang sesuai. Pengadaan perangkat ini harus mempertimbangkan spesifikasi dan kebutuhan dokumen survei yang digunakan oleh BPS. Oleh karena itu, pemilihan perangkat keras dan perangkat lunak perlu dilakukan dengan cermat, dengan mempertimbangkan kualitas, performa, serta kompatibilitas perangkat tersebut dengan sistem yang sudah ada di BPS. Dengan memilih perangkat yang tepat, proses pengolahan data akan menjadi lebih efisien dan dapat dilakukan dalam waktu yang lebih singkat.

c. Pelatihan Operator

Pelatihan operator merupakan langkah yang tidak kalah penting dalam implementasi teknologi OCR. Staf yang akan menggunakan perangkat OCR perlu mendapatkan pelatihan yang memadai terkait penggunaan perangkat tersebut. Pelatihan ini meliputi pengenalan terhadap cara memindai dan mengonversi dokumen fisik menjadi data digital, serta pemahaman mengenai cara mengatasi masalah teknis yang mungkin timbul selama proses penggunaan, seperti kesalahan dalam pemindaian atau pengenalan teks. Selain itu, pelatihan juga harus mencakup pengelolaan data digital yang dihasilkan oleh OCR, agar data tersebut dapat disimpan, diorganisir, dan digunakan untuk analisis lebih lanjut dengan efektif dan efisien.

d. Integrasi Sistem

Langkah terakhir dalam implementasi OCR adalah mengintegrasikan teknologi ini ke dalam sistem pengolahan data yang sudah ada di BPS. Hal ini bertujuan agar hasil dari proses OCR dapat langsung digunakan dalam sistem untuk validasi dan analisis data tanpa memerlukan langkah tambahan yang memakan waktu. Proses integrasi ini perlu dilakukan dengan cermat, memastikan bahwa data yang telah diproses dapat diakses dan digunakan dengan mudah dalam setiap tahap pengolahan data.

Keberhasilan implementasi teknologi seperti *Optical Character Recognition* dalam pengolahan data penduduk sangat bergantung pada kesiapan, kompetensi, dan profesionalisme SDM yang bertugas di BPS Kota Tebing Tinggi. Oleh karena itu, pengembangan kualitas SDM menjadi langkah prioritas dalam memastikan proses pengolahan data berjalan efisien, akurat, dan sesuai dengan kebutuhan masyarakat.

Pelatihan teknologi untuk operator data penduduk perlu mencakup penggunaan OCR dan sistem validasi data otomatis. Operator harus dilatih menggunakan OCR untuk membaca data dari dokumen fisik secara akurat dan mengatasi masalah seperti dokumen rusak atau buram. Selain itu, pemahaman tentang sistem validasi data otomatis penting untuk memastikan data yang diinput sesuai dan konsisten sebelum diproses. Dengan pelatihan ini, operator dapat bekerja lebih efisien dan mengurangi kesalahan dalam pengolahan data.

Perangkat lunak OCR adalah elemen penting dalam digitalisasi dokumen penduduk. Aplikasi ini harus mampu mengenali teks dari berbagai format dokumen, termasuk formulir, tabel, dan laporan. Selain itu, perangkat lunak harus mendukung bahasa yang digunakan dalam dokumen survei penduduk, seperti Bahasa Indonesia, dan memiliki kemampuan untuk membaca teks dalam format tulisan tangan maupun cetakan. Fitur tambahan, seperti konversi tabel dan validasi otomatis, juga menjadi nilai tambah untuk meningkatkan efisiensi kerja.

Sistem penyimpanan berfungsi sebagai tempat menyimpan data digital yang telah diolah. Untuk pengolahan data penduduk di BPS Kota Tebingtinggi, diperlukan penyimpanan yang aman, terorganisasi, dan memiliki kapasitas besar. Server lokal dapat digunakan untuk menjaga kendali penuh terhadap data, sementara penyimpanan berbasis cloud menawarkan fleksibilitas aksesibilitas dan perlindungan data melalui fitur backup otomatis.

Sistem penyimpanan harus dilengkapi dengan langkah-langkah keamanan, seperti enkripsi, proteksi kata sandi, dan mekanisme pemulihan data jika terjadi kerusakan atau kehilangan data. Ini penting untuk menjaga kerahasiaan dan integritas data penduduk, yang sifatnya sensitif dan berpengaruh besar terhadap kebijakan publik.

Langkah pertama adalah melakukan audit terhadap infrastruktur teknologi yang sudah ada di BPS Kota Tebingtinggi. Audit ini bertujuan untuk mengevaluasi kekuatan dan kelemahan dari perangkat keras, perangkat lunak, dan sistem penyimpanan yang sedang digunakan. Hasil audit akan membantu menentukan kebutuhan spesifik yang harus dipenuhi untuk mendukung proses pengolahan data secara lebih efisien.

Setelah kebutuhan diidentifikasi, langkah berikutnya adalah mengadakan perangkat keras dan perangkat lunak yang sesuai. Pemilihan scanner, komputer, dan aplikasi OCR harus didasarkan pada kapasitas dokumen yang diolah, kompleksitas data, dan kompatibilitas dengan sistem yang ada. Selain itu, perangkat yang dipilih harus memiliki kemampuan untuk diperbarui agar tetap relevan dengan perkembangan teknologi di masa mendatang.

4. KESIMPULAN

Implementasi teknologi *Optical Character Recognition* (OCR) melalui penggunaan aplikasi *CamScanner* telah terbukti menjadi solusi yang efektif dalam mendukung proses digitalisasi dokumen survei di BPS Kota Tebing Tinggi. Dengan kemampuan untuk mengubah dokumen fisik menjadi data digital secara otomatis, teknologi ini memberikan dampak signifikan dalam meningkatkan efisiensi kerja, akurasi data, serta kecepatan pengolahan dan analisis informasi. Proses yang sebelumnya memerlukan waktu dan tenaga besar kini dapat dilakukan dengan lebih praktis dan minim

kesalahan, sehingga mendukung pencapaian target kerja yang lebih optimal.

Keberhasilan dari penerapan OCR ini tentu tidak lepas dari berbagai faktor pendukung, antara lain kesiapan infrastruktur teknologi, pemilihan perangkat keras dan lunak yang tepat, pelatihan sumber daya manusia (SDM) yang terlibat dalam proses, serta integrasi sistem OCR dengan sistem pengolahan data yang sudah ada di lingkungan BPS. Tanpa dukungan yang kuat dari aspek-aspek tersebut, penerapan teknologi tidak akan berjalan maksimal.

5. DAFTAR PUSTAKA

- ALIF, A., 2013. Komputasi cerdas untuk pemula. Dewi, L. N., & Setiawan, A. (2021). *Digitalisasi Arsip menggunakan OCR dalam Mendukung Smart Government*. Jurnal Teknologi dan Sistem Informasi, 2(2), 45–51.
- Handayani, R., & Syahril, H. (2022). *Efektivitas Penggunaan OCR dalam Pengolahan Data Survei Sosial Ekonomi di BPS Provinsi*. Jurnal Ilmu Komputer dan Statistik, 6(1), 25–32.
- Lestari, I. N. T., & Mulyana, D. I. “*Implementation of OCR (Optical Character Recognition) Using Tesseract in Detecting Character in Quotes Text Images*”. Journal of Applied Engineering and Technological Science (JAETS), Vol. 4, No. 1, Hal 58–63, 2022.
- Li, M., Lv, T., Chen, J., Cui, L., Lu, Y., Florencio, D., Zhang, C., Li, Z., & Wei, F. “*TrOCR: Transformer-Based Optical Character Recognition with Pre-Trained Models*”. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 37, No. 11, Hal 13094–13102, 2023.

- Nurohman and Nurhayati, “Academic Information System User Satisfaction Model Using Technology Acceptance Model (TAM) With End-User Computing Satisfaction (EUCS) Modification,” *International Journal of Multi Science*, vol. 1, pp. 34–45, 2021.
- Putra, R. H., & Santoso, D. (2020). *Penerapan Optical Character Recognition (OCR) untuk Digitalisasi Dokumen Administratif di Instansi Pemerintah*. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 7(3), 211–217.
- Sholihah, R., & Indriyanti, A. D. ”Analisis Kepuasan Pengguna Aplikasi Camscanner Menggunakan Metode *Technology Acceptance Model (TAM)* dan *End-User Computing Satisfaction (EUCS)*”. *Journal of Emerging Information*, Vol. 3, No. 3 e-ISSN, 2774 – 3993, hal – 102, 2022.
- Susanty, M., & Nugroho, H. (2020). *Implementasi Optical Character Recognition (OCR) Menggunakan Tesseract untuk Ekstraksi Teks pada Gambar*. *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, Vol. 11, No. 2, Hal 123–130.
- Reswan, Y., Raffles, R., Wijaya, A., & Apridiansyah, Y. “ Penerapan Algoritma OC untuk Ekstraksi Informasi dari Citra Kartu Tanda Mahasiswa (KTM)”. *JATI (Jurnal Mahasiswa Teknik Informatika)*, Vol. 8, No. 5, Hal 11004–11009. 2024.
- Wahyuni, D., & Rahman, A. (2023). *Analisis Kinerja OCR Mobile dalam Digitalisasi Dokumen Lapangan: Studi Kasus Penggunaan CamScanner di Instansi Pemerintah Daerah*. *Jurnal Riset Teknologi Informasi dan Komputer*, 8(2), 88–94