

SKRIPSI

**IMPLEMENTASI PROGRAM PENGOLAHAN
TEKS PADA SEBUAH GAMBAR PADA
PT.GAPURA ANGKASA**



POLITEKNIK NEGERI BALI

Oleh :

Dewa Gede Sidan Bagus Taruna Ismayanatha Putra

NIM. 2115354027

**PROGRAM STUDI SARJANA TERAPAN
TEKNOLOGI REKAYASA PERANGKAT
LUNAK JURUSAN TEKNOLOGI INFORMASI
POLITEKNIK NEGERI BALI**

2025

ABSTRAK

PT. Gapura Angkasa menghadapi tantangan dalam proses input data sertifikat karyawan yang dikirim dalam bentuk gambar (JPG/PNG), karena penginputan dilakukan secara manual dan menyita banyak waktu. Permasalahan ini mendorong perancangan sistem pengolahan teks dari gambar menggunakan metode *Optical Character Recognition (OCR)* berbasis web, untuk meningkatkan efisiensi dan akurasi pengolahan data. Penelitian ini bertujuan untuk merancang dan mengimplementasikan sistem yang mampu mengekstraksi teks dari gambar sertifikat agar dapat diproses lebih lanjut secara otomatis oleh bagian *Controller* dan *HRD*. Metodologi yang digunakan adalah model pengembangan perangkat lunak *waterfall*, dengan tahapan analisis kebutuhan, desain sistem menggunakan UML dan ERD, implementasi menggunakan *Python* dan *framework Flask*, serta pengujian menggunakan metode *white-box testing*.

Data dikumpulkan melalui observasi dan pengumpulan dataset sertifikat asli dari pegawai. Hasil implementasi menunjukkan bahwa sistem mampu mengekstrak teks dari gambar sertifikat dengan akurasi yang baik, sehingga mempercepat proses penginputan data dan meminimalkan kesalahan pengetikan manual. Sistem juga mempermudah pengguna dalam memperoleh hasil ekstraksi yang dapat langsung disalin (*copy-paste*). Dengan demikian, sistem ini dapat menjadi solusi efektif untuk mendukung proses digitalisasi dokumen di lingkungan PT. Gapura Angkasa, sekaligus menjadi contoh penerapan teknologi OCR dalam bidang administrasi perusahaan.

Kata Kunci: *Optical Character Recognition, OCR, pengolahan teks, sertifikat digital, Python Flask*

ABSTRACT

PT. Gapura Angkasa faces challenges in the process of inputting employee certificate data, which is often submitted in image formats (JPG/PNG), causing manual data entry to become time-consuming and prone to errors. To address this problem, a web-based text processing system using the Optical Character Recognition (OCR) method is proposed to improve the efficiency and accuracy of data entry. This study aims to design and implement a system capable of extracting text from certificate images to support automated input by the Controller and Human Resource Development (HRD) departments. The research employs the waterfall development model, consisting of requirement analysis, system design using UML and ERD diagrams, implementation using Python and the Flask framework, and testing through white-box testing methods.

Data collection involved direct observation and the compilation of real employee certificate images. The implementation results show that the system effectively extracts text from images with a high degree of accuracy, thereby accelerating the data input process and reducing manual entry errors. The system also allows users to directly copy and paste extracted text, further enhancing workflow efficiency. This solution offers practical support for document digitalization efforts at PT. Gapura Angkasa and demonstrates the application of OCR technology in corporate administrative processes.

Keywords: Optical Character Recognition, OCR, text processing, digital certificate, PythonFlask

DAFTAR ISI

ABSTRAK.....	iv
ABSTRACT	v
KATA PENGANTAR.....	vi
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Perumusan Masalah	2
1.3. Batasan Masalah	2
1.4. Tujuan.....	3
1.5. Manfaat Penelitian	3
1.6. Sistematika Penulisan.....	3
BAB II TINJAUAN PUSTAKA.....	4
2.1. Penelitian Sebelumnya	4
2.2. Landasan Teori.....	5
2.2.1. Website.....	5
2.2.2. Framework Flask.....	5
2.2.3. Database	6
2.2.4. Unified Modelling Language (UML)	6
2.2.5. Entity Relationship Diagram (ERD).....	6
2.2.6. Metode Pengembangan Waterfall.....	6
2.2.7. Metode Tesseract OCR (<i>Optical Character Recognition</i>).....	7
2.2.8. Metode Pengujian <i>Black-Box Testing</i>	8
BAB III METODE PENELITIAN	9
3.1. Metode Pengembangan	9
3.2. Rancangan Sistem (<i>Software</i>)/Pengukuran/Pengambilan Data	10
3.2.1. Spesifikasi Perangkat.....	10
3.2.2. Software	10
3.2.3. Teknik Pengambilan Data.....	10
3.3. Analisis Kondisi Eksisting	11
3.3.1. Analisa Sistem Berjalan.....	12
3.3.2. Analisa Sistem Baru	13
3.4. Perancangan Sistem	15
3.4.1. Flowchart Sistem.....	15
3.4.2. Use Case Diagram	16
3.4.3. <i>Entity Relationship Diagram</i> (ERD).....	17
BAB IV HASIL DAN PEMBAHASAN.....	18
4.1. Hasil Implementasi Sistem.....	18

4.1.1. Implementasi Aplikasi	18
.....	18
4.1.2. Implementasi Penyimpanan Data.....	20
4.2. Hasil Pengujian Sistem	21
4.2.1. Pengujian Sistem.....	21
4.2.2. Pengujian Penyimpanan Data	25
4.2.3. Pengujian parameter-parameter yang diamati.....	27
4.2.4 Usability Testing	33
4.3. Pembahasan Hasil Implementasi dan Pengujian.....	36
4.3.1. Analisis Implementasi Sistem	36
4.3.2. Analisis Pengujian Sistem.....	36
4.3.3. Analisis Uji Usability Testing.....	36
4.3.4. Hasil Uji Coba Dengan Dataset Manual	37
4.3.5. Analisis Perbandingan Hasil terhadap Acuan yang Dipakai di Tinjauan Pustaka.	38
BAB V KESIMPULAN DAN SARAN	39
5.1. Kesimpulan	39
5.2. Saran.....	39
Daftar Pustaka	41

DAFTAR GAMBAR

Gambar 2 1 Simbol & Penjelasan Flow Chart.....	8
Gambar 3 1 Simbol & Penjelasan Flow Chart.....	12
Gambar 3 2 Flowmap Sistem Baru	14
Gambar 3 3 Flowchart Sistem.....	15
Gambar 3 4 Use Case Diagram	16
Gambar 3 5 Entity Relationship Diagram (ERD)	17
 Gambar 4 1 tampilan UI program.	18
Gambar 4 2 tampilan proses input.	19
Gambar 4 3 tampilan hasil output	19
Gambar 4 4 tampilan Riwayat OCR dengan fitur hapus riwayat.....	20
Gambar 4 5 struktur database.....	20
Gambar 4 6 fitur “Choose Files	21
Gambar 4 7 hasil dari fitur “Choose Files”	22
Gambar 4 8 hasil dari fitur “Proses OCR”	22
Gambar 4 9 hasil dari fitur “Proses OCR”	23
Gambar 4 10 fitur “Hapus Hasil OCR ”.....	23
Gambar 4 11 hasil dari fitur “Hapus Hasil OCR ”	24
Gambar 4 12 fitur hapus Riwayat OCR.	24
Gambar 4 13 hasil pengujian fitur hapus Riwayat OCR.	25
Gambar 4 14 hasil penginputan data kedalam database.....	26
Gambar 4 15 berhasil penginputan data kedalam database.....	26
Gambar 4 16 menghapus data didalam database melalui fitur hapus.	27
Gambar 4 17 data dalam database telah terhapus.	27
Gambar 4 18 menampilkan hasil proses OCR	28

DAFTAR TABEL

Tabel 4. 1 Tabel Pengujian Black Box.....	29
---	----

DAFTAR LAMPIRAN

Lampiran 1 Form Bimbingan 1.....	43
Lampiran 2 Form Bimbingan 2.....	44
Lampiran 3 Form Menyelesaikan Bimbingan.....	45

BAB I PENDAHULUAN

1.1. Latar Belakang

PT. Gapura Angkasa merupakan anak usaha dari *Injourney Aviation Services* yang menyediakan layanan penanganan pesawat terbang selama berada di bandara. Untuk mendukung kegiatan bisnisnya, hingga akhir tahun 2018, perusahaan ini memiliki 39 kantor cabang yang tersebar di Indonesia. Menurut Ketut Dian Puspitasari dan Made Yeni Latrini, anak perusahaan adalah perusahaan yang dikendalikan oleh perusahaan induk [1]. Kesuksesan PT. Gapura Angkasa juga bergantung kepada orang yang bergerak di belakangnya. Hubungan baik dengan karyawan dapat menciptakan iklim kerja yang kondusif, dan produktifitas (kinerja) karyawan akan meningkat. Kinerja merupakan hasil kerja yang dicapai oleh seorang karyawan atau pegawai dalam melakukan tugas sesuai dengan tanggungjawab yang diberikan kepadanya [2].

Untuk menjaga kinerja karyawan tetap sesuai pada SOP atau *Standard Operating Procedure* nya, maka PT. Gapura Angkasa selalu melakukan ujian kepada setiap pegawai pegawai yang nantinya akan diberikan sebuah sertifikat sebagai tanda bukti bahwa karyawan tersebut sudah atau masih kompeten pada bidangnya. Pegawai yang mendapat sertifikat itu diminta untuk mengirimkan sertifikat mereka pada seorang yang bertugas sebagai *Controller* dan *HRD* yaitu *Human Resource Development*, lalu sertifikat itu akan diinputkan kedalam sistem oleh *Controller* dan juga *HRD*.

Namun Proses penginputan sertifikat pegawai sering kali memakan cukup banyak waktu karena pihak yang bertugas untuk melakukan input detail sertifikat sering kali harus menuliskan detail sertifikat secara manual tanpa bisa melakukan *Copy* dan *Paste* karena sertifikat yang diberikan oleh pegawai berupa *Image* seperti *JPG* atau *PNG*.

Permasalahan diatas dapat dibantu dengan adanya sistem atau program yang dapat mengolah semua teks atau mengambil setiap teks pada sertifikat yang berupa *Image* secara utuh agar bisa dilakukan proses *Copy* dan *Paste*. Dengan menerapkan sistem yang menggunakan metode *Optical Character Recognition* (OCR).Diharapkan dapat meningkatkan efisiensi, objektivitas dalam proses penginputan detail sertifikat pegawai PT. Gapura Angkasa.Dalam hal ini akan dibangun sebuah sistem pengubah sebuah gambar menjadi sebuah teks tergantung pada teks yang terkandung didalam gambar tersebut.Dengan adanya sistem *Image To Text Processing* ini seharusnya penginputan data akan lebih efisien dan menghemat waktu pengerjaan karena orang yang bertugas

menginputkan detail sertifikat tidak harus menulis secara manual lagi.

Alur pengembangan yang digunakan adalah alur pengembangan *waterfall*. *Waterfall* merupakan model yang umum digunakan pada proses pengembangan suatu sistem. *Waterfall* menjelaskan pengembangan model yang mewakili proses aturan siklus perangkat lunak/*software* dengan sistem yang dapat dipanggil secara berurutan dengan urutan proses analisis, desain, pengkodean, pengujian, dan bagian pendukung [3]. Program ini juga menggunakan metode *Optical Character Recognition* (OCR). Dengan metode pengujian *black-box testing*. *Blackbox testing* merupakan pengujian untuk menunjukkan kesalahan pada system aplikasi seperti kesalahannpada fungsi system aplikasi, serta menu aplikasi yang hilang. Jadi *Black Box testing* merupakan metode uji fungsionalitas system aplikasi. Dalam melakukan pengujian menggunakan masukan data acak dengan tujuan untuk mendapatkan hasil yang pasti. Dikatakan pasti artinya bila salah ,maka di tolak oleh sistem informasi atau data input tersebut tidak dapat disimpan dalam *data base*, sedangkan bila data input benar maka dapat di terima / masuk di *database* sistem informasi. [4].

1.2. Perumusan Masalah

Berdasarkan latar belakang yang telah dituliskan sebelumnya, maka dapat dirumuskan beberapa permasalahan diantaranya sebagai berikut :

1. Bagaimana merancang sistem pengolahan teks pada sebuah gambar menggunakan metode OCR?
2. Bagaimana cara kerja sistem pengolahan teks pada sebuah gambar untuk mendapatkan hasil yang optimal?

1.3. Batasan Masalah

Diperlukan sebuah batasan-batasan masalah yang akan digunakan sebagai pedoman untuk mencapai target dari penelitian. Batasan masalah tersebut antara lain:

1. Sistem pengolahan teks pada gambar dikembangkan dengan bahasa pemrograman *Python* dan *Framework Flask*.
2. Metode pengolahan teks pada sebuah gambar menggunakan metode *OCR*, yang akan mendeteksi teks pada sebuah gambar secara akurat
3. Sistem memerlukan pencahayaan yang baik agar scan lebih optimal

1.4. Tujuan

Tujuan diperlukan agar penelitian ini memiliki arah dan tujuan yang akan dicapai, maka tujuan dalam penelitian ini adalah sebagai berikut:

1. Merancang sistem pengolahan teks pada gambar dengan metode OCR.
2. Sistem pengolahan teks pada sebuah gambar bekerja melalui akuisisi, pra-pemrosesan, deteksi teks, ekstraksi fitur, pengenalan karakter, dan pasca-pemrosesan untuk hasil yang optimal.

1.5. Manfaat Penelitian

1. Bagi Institusi : Menambah koleksi penelitian yang relevan dengan perkembangan teknologi terkini, khususnya di bidang pengolahan teks dan visi komputer.
2. Bagi Perusahaan terkait : Memberikan solusi praktis untuk mempermudah pengolahan data teks yang berasal dari gambar, sehingga mengurangi waktu operasional.
3. Bagi Peneliti : Memberikan pengalaman langsung dalam mengembangkan dan mengimplementasikan teknologi pengolahan teks dan gambar.

1.6. Sistematika Penulisan

Struktur penulisan dalam skripsi ini disusun sebagai berikut:

a. BAB I PENDAHULUAN

Berisi latar belakang, rumusan masalah, Batasan masalah, tujuan, manfaat, dan sistematika penulisan skripsi.

b. BAB II TINJAUAN PUSTAKA

Memuat ringkasan penelitian terdahulu yang relevan serta teori-teori pendukung seperti metode OCR, dan teknologi yang digunakan.

c. BAB III METODE PENELITIAN

Menjelaskan metode penelitian, perancangan sistem, kebutuhan perangkat, serta diagram pendukung.

d. BAB IV IMPLEMENTASI SISTEM

Membahas hasil implementasi sistem, pengujian aplikasi, serta analisis hasil berdasarkan pengujian yang dilakukan.

e. BAB V PENUTUP

Berisi kesimpulan dan saran untuk sistem kedepannya

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, maka diperoleh kesimpulan sebagai berikut:

1. Perancangan sistem pengolahan teks pada sebuah gambar menggunakan metode OCR berhasil dilakukan. Sistem yang dirancang berbasis web dengan framework Flask dan menggunakan bahasa pemrograman Python. Sistem ini mampu menerima input berupa gambar sertifikat karyawan dan mengubahnya menjadi teks digital. Proses perancangan meliputi tahap analisis kebutuhan, desain sistem dengan UML dan ERD, serta pembuatan antarmuka yang memudahkan pengguna dalam mengunggah gambar dan melihat hasil ekstraksi teks.
2. Sistem pengolahan teks bekerja dengan optimal melalui tahapan pemrosesan OCR. Hasil implementasi dan pengujian menunjukkan bahwa sistem mampu mengekstrak teks dari gambar dengan akurasi tinggi. Fitur-fitur seperti “Proses OCR”, “Riwayat OCR”, dan “Hapus Riwayat” berfungsi dengan baik. Sistem ini mempercepat proses penginputan data dan meminimalkan kesalahan pengetikan manual. Walau belum mencapai tingkat akurasi hingga 100% sebagaimana penelitian terdahulu, sistem sudah memenuhi kebutuhan operasional perusahaan.

5.2. Saran

Berdasarkan hasil implementasi dan kendala yang ditemukan selama penelitian, maka beberapa saran yang dapat diberikan adalah :

1. Pengembangan fitur koreksi teks otomatis.

Sistem saat ini belum menyediakan validasi atau koreksi otomatis terhadap hasil teks OCR. Fitur secaman ini dapat membantu pengguna untuk mempercepat pengecekan dan mengurangi kesalahan ketik akibat noise dalam gambar.

2. Optimasi terhadap variasi kualitas gambar. Kinerja sistem masih bergantung pada kualitas gambar input (pencahayaan, kontras, dan resolusi). Diperlukan penelitian lanjutan yang mengintegrasikan preprocessing gambar otomatis agar hasil OCR tetap akurat meskipun input gambar tidak ideal.

3. Penambahan dukungan format file lain dan multibahasa

Sistem saat ini hanya mendukung gambar dengan format JPG/JPEG. Perlu ditambahkan dukungan untuk format PNG, PDF, atau bahkan fitur deteksi bahasa otomatis bila terdapat dokumen dengan berbagai bahasa.

Dengan adanya pengembangan lebih lanjut berdasarkan saran tersebut, sistem ini diharapkan dapat semakin optimal, fleksibel, dan mendukung digitalisasi administrasi perusahaan secara menyeluruh.

Daftar Pustaka

- [1] B. C. Das, S. D. Gero, and E. Lederer, "N-methylation of N-acyl oligopeptides," *Biochem. Biophys. Res. Commun.*, vol. 29, no. 2, pp. 211–215, 1967, doi: 10.1016/0006-291X(67)90589-X.
- [2] "Abstraksi fungsi hrd dalam meningkatkan kinerja karyawan dengan menggunakan strategi pr di pt. telkom regional iv."
- [3] A. Nurseptaji, "Implementasi Metode Waterfall Pada Perancangan Sistem Informasi Perpustakaan," *J. Dialekt. Inform.*, vol. 1, no. 2, pp. 49–57, 2021, doi: 10.24176/detika.v1i2.6101.
- [4] M. Yusuf, "Pengujian fungsional Perangkat Lunak Sistem Informasi Perpustakaan Dengan Metode Black Box Testing Bagi Pemula," *J. Siliwangi*, vol. 8, no. 1, pp. 1–6, 2022.
- [5] Kusnantoro, T. Rohana, and D. S. Kusumaningrum, "Implementasi Metode Tesseract OCR(Optical Character Recognition)untuk Deteksi Plat Nomor Kendaraan Pada Sistem Parkir," *Sci. Student J. Information, Technol. Sci.*, vol. III, pp. 59–67, 2022.
- [6] F. O. Rahmalisty, S. Aulia, and S. Hadiyoso, "Perancangan Language Translator Image To Text Menggunakan Metode Optical Character Recognition Berbasis Pengolahan Citra Language Translator Image To Text Design Using Optical Character Tecognition Method Based On Image Processing," vol. 9, no. 1, pp. 187–195, 2023,[Online].Available:<https://openlibrarypublications.telkomuniversity.ac.id/index.php/appliedscience/article/view/19395>
- [7] A. Science, "[1] , [2] , [3]," vol. 6, no. 1, pp. 629–636, 2020.
- [8] N. Giantara and I. Anandita, "Web Pada Sistem Penjadwalan Mengajar Di Smk Dwijendra Denpasar," *J. Teknol. Inf. Dan Komput.*, vol. 5, no. 1, pp. 130–138, 2019.
- [9] F. Kindarya *et al.*, "Penerapan Aplikasi Klasifikasi Hukum Tajwid Menggunakan Image Processing," *El-Mujtama J. Pengabdi. Masyarakat* , vol. 4, no. 2, pp. 660– 669, 2024, doi: 10.47467/elmujtama.v4i2.1930.
- [10] N. Eyni Alfia and B. Waseso, "Perancangan Aplikasi Retensi Data Pada Database MySQL (Studi Kasus: PT. Telkomsigma)," *Maret*, vol. 2, no. 3, pp. 2655–7541, 2020, [Online]. Available: <https://jurnal.ikhafi.or.id/index.php/jusibi/364>

- [11] D. W. T. Putra and R. Andriani, “Unified Modelling Language (UML) dalam Perancangan Sistem Informasi Permohonan Pembayaran Restitusi SPPD,” *J. TeknoIf*, vol. 7, no. 1, p. 32, 2019, doi: 10.21063/jtif.2019.v7.1.32-39.
- [12] S. M. Pulungan, R. Febrianti, T. Lestari, N. Gurning, and N. Fitriana, “Analisis Teknik Entity-Relationship Diagram Dalam Perancangan Database,” *J. Ekon. Manaj. dan Bisnis*, vol. 1, no. 2, pp. 98–102, 2023, doi: 10.47233/jemb.v1i2.533.
- [13] A. A. Wahid, ““Analisis Metode Waterfall Untuk Pengembangan Sistem Informasi,’ ,” *J. Ilmu-ilmu Inform. dan Manaj. STMIK*, vol. 1, no. November, 2020.
- [14] A. Setiawan, H. Sujaini, and A. B. Pn, “Implementasi Optical Character Recognition (OCR) pada Mesin Penerjemah Bahasa Indonesia ke Bahasa Inggris,” *J. Sist. dan Teknol. Inf.*, vol. 5, no. 2, pp. 135–141, 2017.