

**OPTICAL CHARACTER RECOGNITION (OCR) BERBASIS API
UNTUK MENGEKSTRAK TEXT PADA GAMBAR DENGAN
TELEGRAM BOT DAN HEROKU**

SKRIPSI



disusun oleh

Muhammad Yusuf Abdurrohman

17.11.1039

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

**OPTICAL CHARACTER RECOGNITION (OCR) BERBASIS API
UNTUK MENGEKSTRAK TEXT PADA GAMBAR DENGAN
TELEGRAM BOT DAN HEROKU**

SKRIPSI

untuk memenuhi persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



disusun oleh

Muhammad Yusuf Abdurrohman

17.11.1039

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

PERSETUJUAN

SKRIPSI

OPTICAL CHARACTER RECOGNITION (OCR) BERBASIS API UNTUK MENGEKSTRAK TEXT PADA GAMBAR DENGAN TELEGRAM BOT DAN HEROKU

yang dipersiapkan dan disusun oleh

Muhammad Yusuf Abdurrohman

17.11.1039

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 14 Maret 2021

Dosen Pembimbing,

Haryoko, S.Kom, M.Cs
NIK. 190302286

PENGESAHAN

SKRIPSI

**OPTICAL CHARACTER RECOGNITION (OCR) BERBASIS API
UNTUK MENGEKSTRAK TEXT PADA GAMBAR DENGAN
TELEGRAM BOT DAN HEROKU**

yang dipersiapkan dan disusun oleh

Muhammad Yusuf Abdurrohman

17.11.1039

telah dipertahankan di depan Dewan Penguji
pada tanggal 21 April 2021

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Haryoko, S.Kom, M.Cs
NIK. 190302286

Bayu Setiaji, M.Kom
NIK. 190302216

Agus Fatkhurohman, M.Kom
NIK. 190302249

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 21 April 2021

DEKAN FAKULTAS ILMU KOMPUTER

Krisnawati, S.Si, MT.
NIK. 190302038

PERNYATAAN

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 29 April 2021



Muhammad Yusuf Abdurrohman

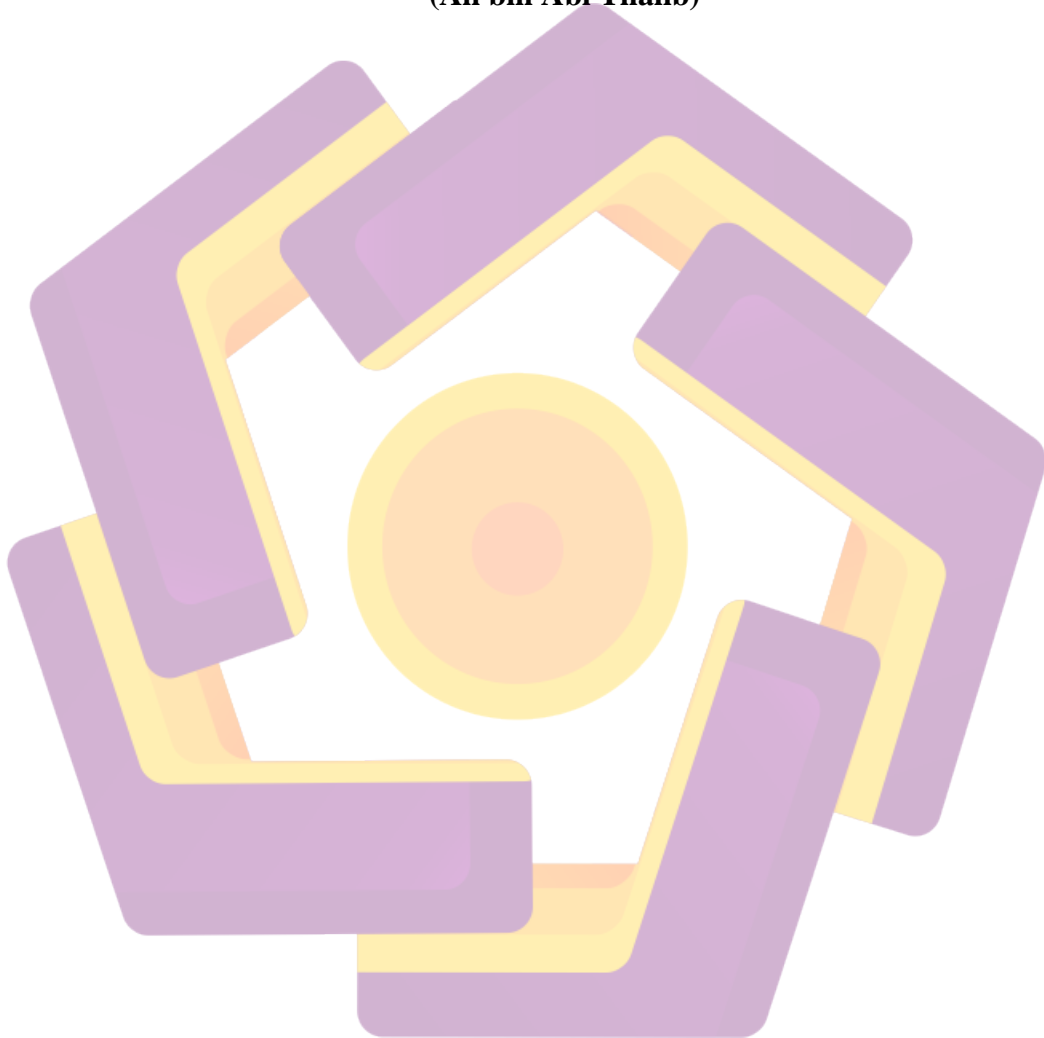
NIM. 17.11.1039

MOTTO

”Kebahagiaan dan kesuksesan semua akan tercapai, tergantung pada dirimu sendiri.”

”Kesalahan terburuk adalah ketertarikan kita dengan kesalahan orang lain.”

(Ali bin Abi Thalib)



PERSEMBAHAN

Pada halaman persembahan ini penulis ingin menyampaikan rasa syukur dan terima kasih kepada :

1. Allah SWT & Rasulullah SAW, Puji Syukur penulis panjatkan kepada Allah SWT, yang telah memberikan kesehatan, rahmat dan hidayah, sehingga penulis masih diberikan kesempatan untuk menyelesaikan skripsi ini, sebagai salah satu syarat untuk mendapatkan gelar kesarjanaan. Walaupun jauh dari kata sempurna, namun penulis bangga telah mencapai pada titik ini, yang akhirnya skripsi ini bisa selesai diwaktu yang tepat. Shalawat serta salam selalu terlimpahkan pada junjungan dan suri tauladan kita Nabi Muhammad SAW, keluarganya dan para sabahatnya.
2. Orangtua saya tercinta (Bapak M. Nur Salim) yang senantiasa mendoakan dan memberi dukungan sehingga skripsi ini bisa terselesaikan seperti seharusnya.
3. Saudara saya Fajar Rizkiantoro yang sudah membantu saya dalam menyelesaikan skripsi ini.
4. Bapak Haryoko, S.Kom, M.Cs saya ucapkan terimakasih banyak kepada bapak, karena sudah membimbing saya dalam menyusun dan menyelesaikan dengan penuh kesabaran yang luar biasa, sehingga skripsi ini dapat selesai dan tersusun dengan baik.
5. Teman-teman yang tidak bisa saya sebutkan satu persatu, yang selama ini sudah berjuang bersama-sama selama perkuliahan.
6. Sahabat saya Fendi Rahman Saputro, Ricky Syah Nanda, Adi Yoga Prakasa dan Fajar Wibowo yang berjuang dan saling betukar pikiran dalam menyelesaikan skripsi ini.

KATA PENGANTAR

Puji dan syukur penulis haturkan kepada Allah Subhanahu Wa Ta'ala atas limpahan serta rahmat-Nya lah penulis diberikan kesempatan dan kemudahan dalam menyelesaikan skripsi ini. Sholawat serta salam senantiasa tercurah kepada Rasul junjungan kita Nabi Muhammad Shalallahu 'Alaihi Wasalam.

Tujuan pembuatan tugas akhir ini adalah sebagai syarat kelulusan dari perguruan tinggi program Studi Strata-1 Sistem Informasi Fakultas Ilmu Komputer di Universitas AMIKOM Yogyakarta untuk meraih gelar S.Kom. Dengan judul skripsi **“OPTICAL CHARACTER RECOGNITION (OCR) BERBASIS API UNTUK MENGEKSTRAK TEXT PADA GAMBAR DENGAN TELEGRAM BOT DAN HEROKU”**. Skripsi ini berhasil terselesaikan karena bantuan dan kerjasama seluruh pihak. Oleh sebab itu penulis ingin menyampaikan rasa hormat dan terima kasih kepada :

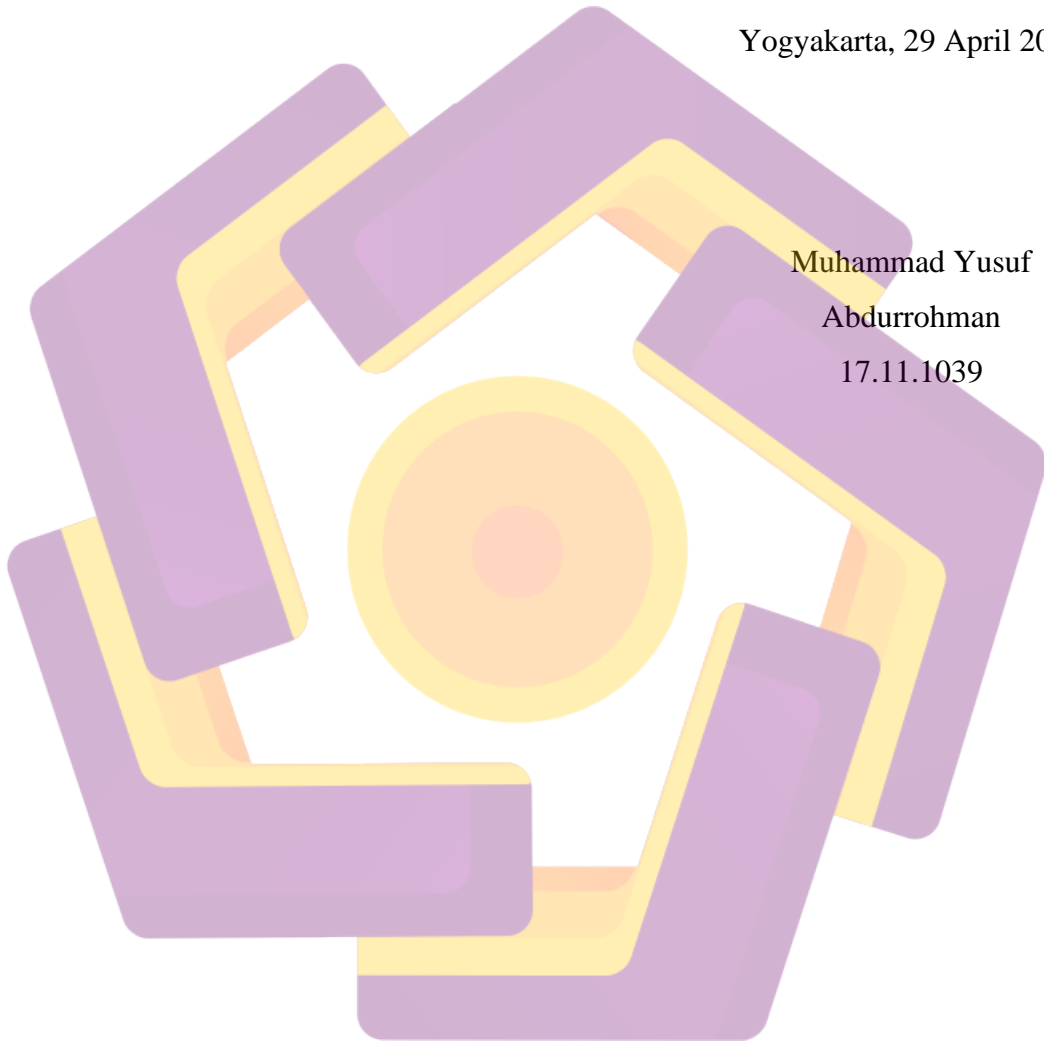
1. Bapak M. Suyanto, Prof., Dr., M.M. selaku Rektor Universitas AMIKOM Yogyakarta.
2. Ibu Krisnawati, S.Si, MT selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Ibu Windha Mega Pradnya D, M.Kom selaku Ketua Jurusan Informatika Universitas Amikom Yogyakarta.
4. Bapak Haryoko, S.Kom, M.Cs selaku dosen pembimbing, karena bimbingan, arahan dan masukan beliau sehingga skripsi ini dapat terselesaikan dengan hasil yang terbaik.
5. Bapak dan Ibu Dosen yang telah banyak memberikan ilmunya selama masa perkuliahan serta segenap Staf Universitas Amikom Yogyakarta.
6. Bapak, Ibu dan seluruh keluarga besar penulis yang senantiasa memberikan kasih sayang, doa dan dukungan kepada penulis.
7. Seluruh pihak yang tidak dapat penulis sebutkan satu persatu yang telah terlibat dan ikut membantu selama proses pengerjaan hingga skripsi ini selesai.

Penulis juga memohon maaf apabila penyusunan skripsi ini masih memiliki banyak kekurangan dan masih jauh dari kata sempurna. Oleh sebab itu, penulis mengharapkan kepada seluruh pihak untuk memberikan kritik, dan saran yang membangun untuk kesempurnaan skripsi ini.

Semoga skripsi ini dapat menambah pengetahuan dan memberikan manfaat kepada pembaca serta dapat digunakan sebagaimana mestinya.

Yogyakarta, 29 April 2021

Muhammad Yusuf
Abdurrohman
17.11.1039



DAFTAR ISI

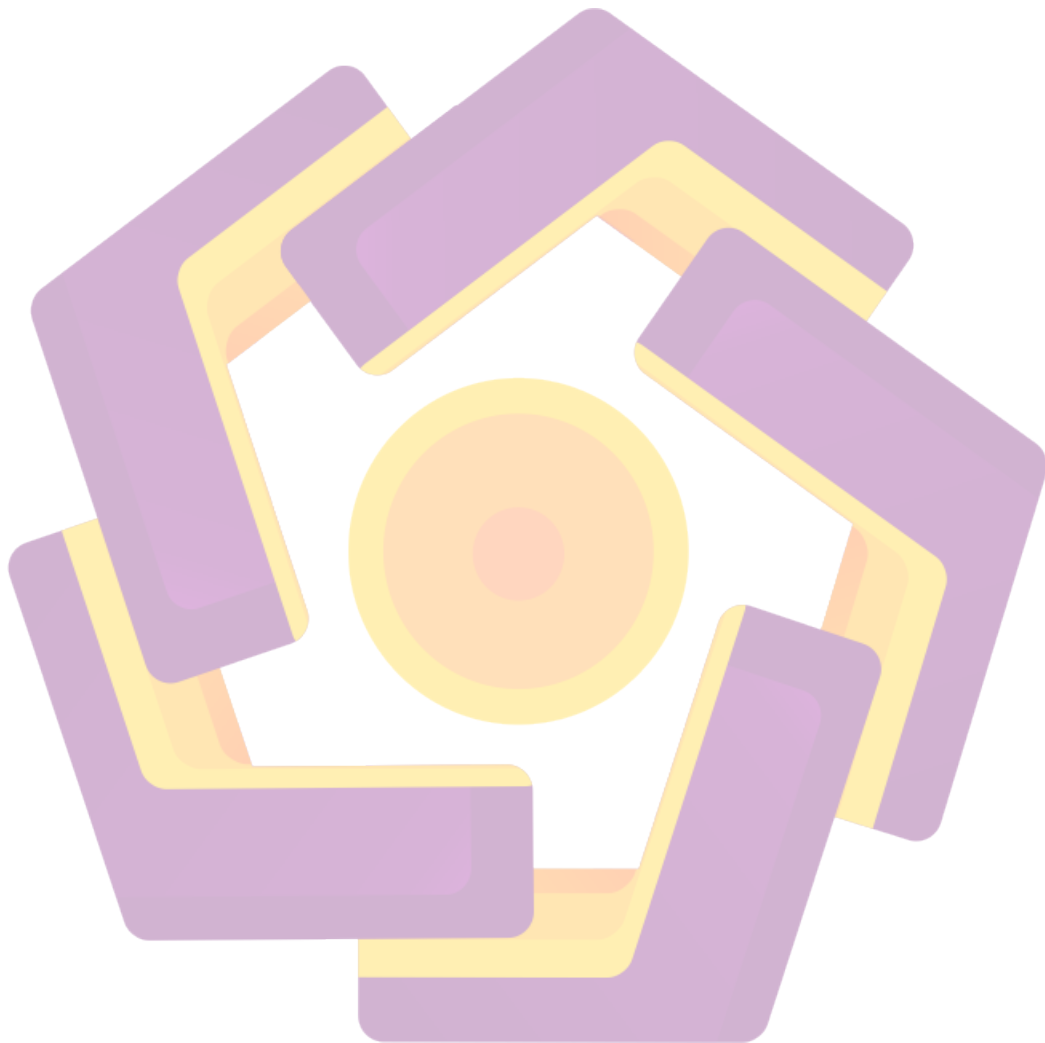
JUDUL.....	I
PERSETUJUAN.....	II
PENGESAHAN.....	III
PERNYATAAN.....	IV
MOTTO.....	V
PERSEMBAHAN.....	VI
KATA PENGANTAR.....	VII
DAFTAR ISI.....	IX
DAFTAR TABEL.....	XII
DAFTAR GAMBAR.....	XIII
DAFTAR LAMPIRAN.....	XV
INTISARI.....	XVI
ABSTRACT.....	XVII
BAB I PENDAHULUAN.....	1
1.1 LATAR BELAKANG.....	1
1.2 RUMUSAN MASALAH.....	2
1.3 BATASAN MASALAH.....	2
1.4 MAKSUD DAN TUJUAN PENELITIAN.....	3
1.5 MANFAAT PENELITIAN.....	3
1.6 METODE PENELITIAN.....	4
1.7 SISTEMATIKA PENULISAN.....	5
BAB II LANDASAN TEORI.....	7
2.1 TINJAUAN PUSTAKA.....	7
2.2 DASAR TEORI.....	22
2.2.1 Optical Character Recognition (OCR).....	22
2.2.2 Proses Umum Optical Character Recognition (OCR).....	22
2.2.3 BOT (Build Operate Transfer).....	28

2.2.4	Telegram Bot.....	28
2.2.5	Application Programming Interface (API).....	29
2.2.6	Cloudmersive OCR API.....	30
2.2.7	PaaS Heroku.....	31
2.2.8	DFD (Data Flow Diagram).....	31
2.2.9	Bahasa Pemrograman Python.....	33
2.2.10	Microsoft Visual Studio Code.....	34
BAB III METODE PENELITIAN.....		35
3.1	GAMBARAN UMUM PENELITIAN.....	35
3.2	ANALISIS KEBUTUHAN SISTEM.....	35
3.2.1	Kebutuhan Fungsional.....	36
3.2.2	Kebutuhan Non-fungsional.....	36
3.3	PERANCANGAN SISTEM.....	37
3.3.1	Context Diagram (DFD Level 0).....	37
3.3.2	Data Flow Diagram Level 1 (DFD Level 1).....	38
3.3.3	Flowchart Sistem.....	41
3.4	SKENARIO PENGUJIAN DAN EVALUASI SISTEM.....	43
3.4.1	Skenario Pengujian.....	43
3.4.2	Evaluasi Sistem.....	44
BAB IV HASIL DAN PEMBAHASAN.....		45
4.1	IMPLEMENTASI SISTEM.....	45
4.1.1	Requirements Analysis.....	46
4.1.2	Design.....	48
4.1.3	Implementation.....	52
4.1.4	Deployment.....	61
4.1.5	Testing.....	64
4.1.6	Maintenance.....	66
4.2	PENGUJIAN DAN EVALUASI SISTEM.....	66
4.2.1	Pengujian Sistem.....	66
4.2.2	Evaluasi Sistem.....	72
4.3	PENGUJIAN LANJUTAN.....	76
BAB V PENUTUP.....		85
5.1	KESIMPULAN.....	85

5.2 SARAN85

DAFTAR PUSTAKA.....87

DAFTAR LAMPIRAN.....90



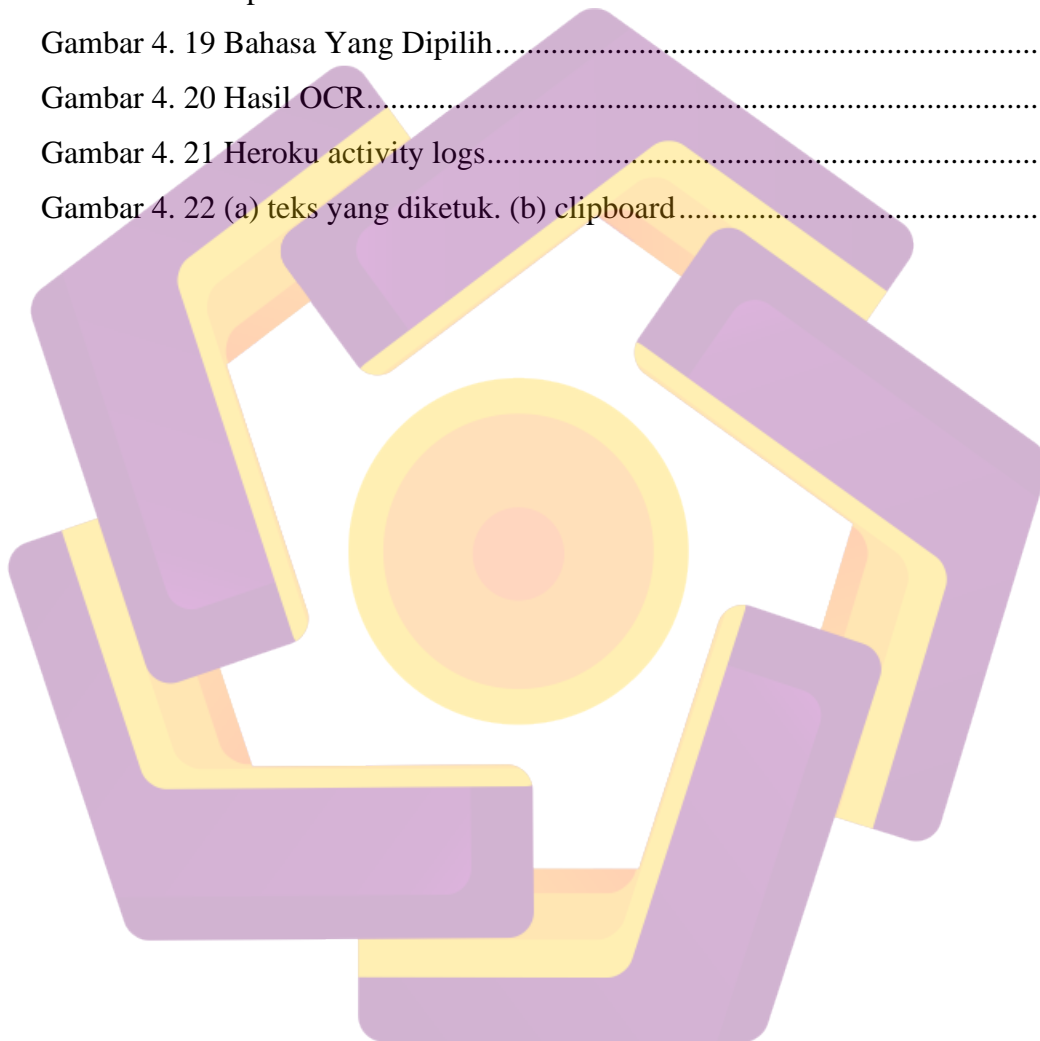
DAFTAR TABEL

Tabel 2. 1 Penelitian yang menjadi referensi	12
Tabel 3. 1 Skenario Pengujian.....	43
Tabel 4. 1 Black-box testing.....	66
Tabel 4. 2 Kriteria pengambilan nilai.....	67
Tabel 4. 3 UAT aspek kemudahan penggunaan.....	68
Tabel 4. 4 UAT aspek pencapaian tujuan.....	69
Tabel 4. 5 UAT aspek penilaian aplikasi.....	69
Tabel 4. 6 Hasil pengujian UAT aspek kemudahan penggunaan.....	70
Tabel 4. 7 Hasil pengujian UAT aspek pencapaian tujuan.....	71
Tabel 4. 8 Hasil pengujian UAT aspek penilaian aplikasi	71
Tabel 4. 9 Hasil Pengujian Telegram OCR Bot	72
Tabel 4. 10 Menghitung nilai Confidence	75
Tabel 4. 11 Uji hasil OCR pada dokumen baru (print preview).....	78
Tabel 4. 12 Uji kamera smartphone	80
Tabel 4. 13 Uji pemilihan opsi bahasa yang salah	83

DAFTAR GAMBAR

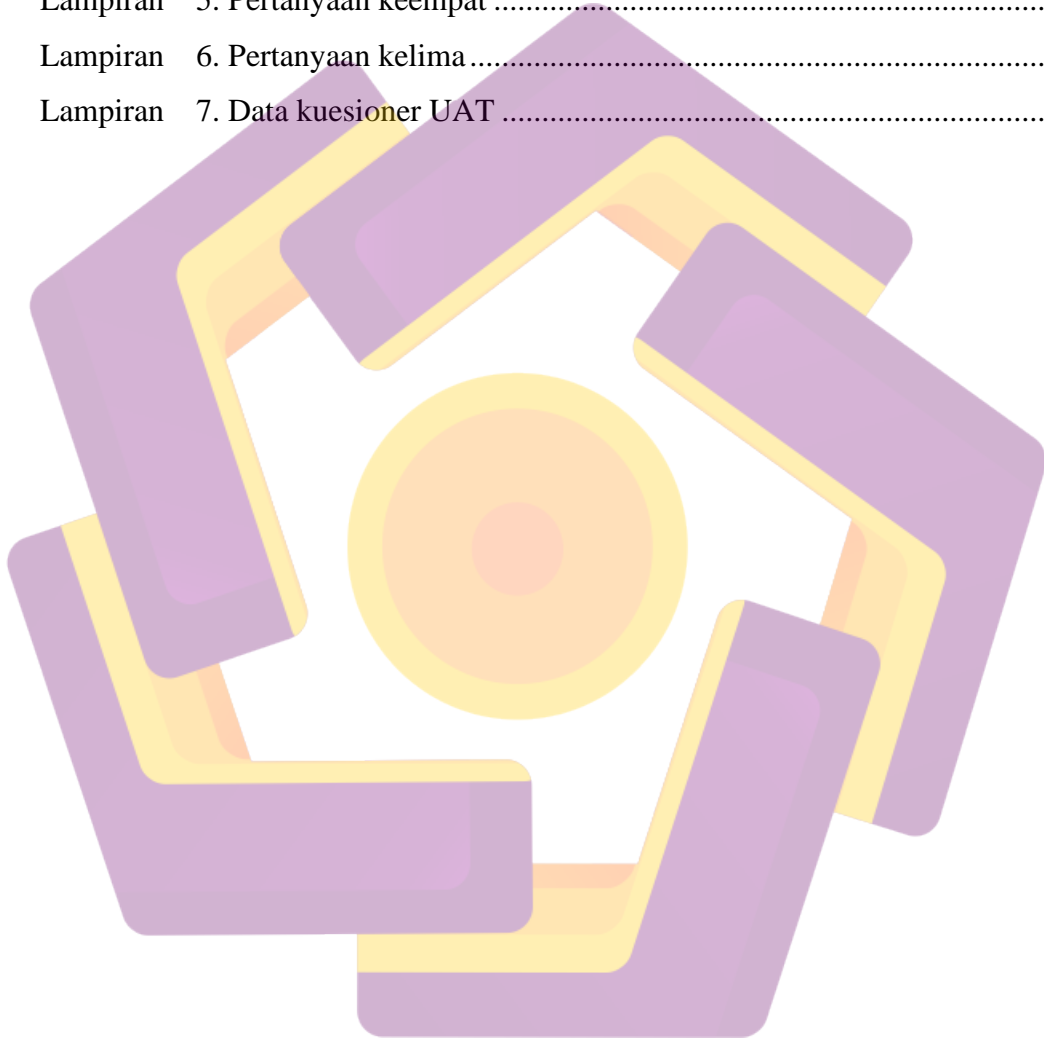
Gambar 2. 1 Flowchart proses OCR.....	23
Gambar 2. 2 Skew	24
Gambar 2. 3 Despeckled (noise removal)	24
Gambar 2. 4 Binarized.....	25
Gambar 2. 5 Feature detection.....	25
Gambar 2. 6 Pattern recognition identifikasi berdasarkan baris.....	26
Gambar 2. 7 Pattern recognition identifikasi berdasarkan kolom	26
Gambar 2. 8 Pattern recognition matriks biner.....	27
Gambar 2. 9 Pattern recognition matriks biner.....	27
Gambar 2. 10 Application Programming Interface (API).....	29
Gambar 2. 11 Alur OCR pada Cloudmersive OCR.....	31
Gambar 2. 12 External Entity	32
Gambar 2. 13 Proses	32
Gambar 2. 14 Aliran data (data flow).....	32
Gambar 2. 15 Penyimpanan data (data store).....	33
Gambar 3. 1 Alur aplikasi Telegram OCR Bot berbasis API.....	35
Gambar 3. 2 Context Diagram (DFD Level 0).....	38
Gambar 3. 3 Data Flow Diagram Level 1 (DFD Level 1).....	39
Gambar 3. 4 Flowchart sistem Telegram OCR Bot.....	41
Gambar 4. 1 SDLC iterative model	45
Gambar 4. 2 Create new Heroku app	46
Gambar 4. 3 Dashboard API key Cloudmersive OCR	47
Gambar 4. 4 BotFather & command /start	47
Gambar 4. 5 Halaman utama PC/Web dan Mobile	49
Gambar 4. 6 Halaman Intro PC/Web dan Mobile	50
Gambar 4. 7 Halaman Opsi Bahasa PC/Web dan Mobile.....	51
Gambar 4. 8 Halaman Hasil OCR PC/Web dan Mobile	52
Gambar 4. 9 Instalasi add-on package telegram.....	53
Gambar 4. 10 Instalasi add-on package python-telegram-bot.....	53
Gambar 4. 11 Instalasi add-on package cloudmersive_ocr_api_client	54

Gambar 4. 12 Heroku Login.....	61
Gambar 4. 13 Heroku Procfile.....	62
Gambar 4. 14 Heroku Requirements.txt.....	63
Gambar 4. 15 Heroku Requirements.txt.....	63
Gambar 4. 16 Push source code program ke Heroku	64
Gambar 4. 17 Menjalankan Command /start.....	64
Gambar 4. 18 Opsi Pilih Bahasa.....	65
Gambar 4. 19 Bahasa Yang Dipilih.....	65
Gambar 4. 20 Hasil OCR.....	65
Gambar 4. 21 Heroku activity logs.....	66
Gambar 4. 22 (a) teks yang diketuk. (b) clipboard.....	77



DAFTAR LAMPIRAN

Lampiran	1. Data nama responden	90
Lampiran	2. Pertanyaan pertama	90
Lampiran	3. Pertanyaan kedua	91
Lampiran	4. Pertanyaan ketiga	91
Lampiran	5. Pertanyaan keempat	92
Lampiran	6. Pertanyaan kelima	92
Lampiran	7. Data kuesioner UAT	93



INTISARI

Optical Character Recognition (OCR) merupakan sebuah teknologi yang banyak digunakan dalam berbagai bidang, salah satunya yaitu pada bidang kesehatan dan perbankan. OCR dalam penerapannya bertujuan untuk mengekstrak atau mengkonversi setiap karakter/teks pada dokumen citra menjadi teks digital.

Telegram *Messenger* merupakan aplikasi *chatting* yang memungkinkan pengguna dapat bertukar pesan, dokumen, foto, *video* dan sebagainya. telegram *messenger* memiliki banyak sekali fitur yang ditawarkan, salah satunya adalah telegram bot. telegram bot merupakan sebuah akun telegram yang dioperasikan oleh perangkat lunak yang dapat ditambahkan *source code* bahkan A.I didalamnya. telegram bot banyak digunakan untuk membuat aplikasi *chatbot*, manajemen grup, penerapan algoritma, game dan sebagainya.

Oleh karena itu, dalam penelitian ini berfokus pada pembuatan *platform* dengan memanfaatkan fitur telegram bot sebagai media dalam pemanfaatan teknologi OCR (*Optical Character Recognition*) berbasis API yang diterapkan dan menggunakan heroku untuk hosting aplikasi OCR. Adapun hasil dari penelitian berupa aplikasi OCR berbasis API yang diterapkan pada telegram bot yang dapat diakses pada semua *platform*, sehingga penggunaan aplikasi dapat digunakan secara luas.

Kata Kunci: OCR, Cloudmersive OCR API, Telegram Bot, Heroku

ABSTRACT

Optical Character Recognition (OCR) is a technology that is widely used in various fields, one of which is healthcare and banking. OCR in its application aims to extract or convert every character/text in an image document into digital text.

Telegram Messenger is a chat application that allows users to exchange messages, documents, photos, videos etc. Telegram Messenger has a lot of features to offer, one of which is the Telegram Bot. Telegram Bot is a Telegram account operated by software that can add source code and even A.I (Artificial Intelligence) in it. Telegram Bot is widely used to create chatbot applications, group management, algorithm implementation, games etc.

Therefore, in this case the research focuses on creating a platform by utilizing the telegram bot feature as a place for OCR (Optical Character Recognition) technology based on API and Heroku as a hosting for OCR application. The research results is a OCR application based on API that is applied to Telegram Bot which can be accessed on all platforms, so that application can be widely used.

Keyword: OCR, Cloudmersive OCR API, Telegram Bot, Heroku