

**APLIKASI EKSTRAKSI TEXT DENGAN OCR (*OPTICAL  
CHARACTER RECOGNITION*) PADA KARTU TANDA  
MAHASISWA UNIVERSITAS NURUL JADID**

**SKRIPSI**



OLEH :

**ZAINUL HASAN**

**16010222**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK UNIVERSITAS NURUL JADID  
PAITON PROBOLINGGO**

**2020**

**APLIKASI EKSTRAKSI TEXT DENGAN OCR (*OPTICAL  
CHARACTER RECOGNITION*) PADA KARTU TANDA  
MAHASISWA UNIVERSITAS NURUL JADID**

**SKRIPSI**

Diajukan untuk Memenuhi Persyaratan Menyelesaikan Studi Strata  
Satu (S-1) dan Memperoleh Gelar Sarjana Komputer (S. Kom.)



OLEH :

**ZAINUL HASAN**

**16010222**

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNIK UNIVERSITAS NURUL JADID  
PAITON PROBOLINGGO  
2020**



## UNIVERSITAS NURUL JADID

### PENGESAHAN STATUS SKRIPSI

JUDUL : APLIKASI EKSTRAKSI TEXT DENGAN OCR (OPTICAL CHARACTER RECOGNITION) PADA KARTU TANDA MAHASISWA UNIVERSITAS NURUL JADID

SAYA : ZAINUL HASAN

Mengijinkan Skripsi Strata Satu Komputer ini disimpan di Perpustakaan Universitas Nurul Jadid dengan syarat-syarat kegunaan sebagai berikut:

1. Skripsi adalah hak milik Universitas Nurul Jadid.
2. Perpustakaan Universitas Nurul Jadid dibenarkan membuat salinan untuk tujuan referensi saja.
3. Perpustakaan juga dibenarkan membuat salinan Skripsi ini sebagai bahan pertukaran antar institusi pendidikan tinggi.
4. Berikan tanda  sesuai dengan kategori Skripsi.

Sangat Rahasia (Mengandung isi tentang keselamatan atau kepentingan Negara Republik Indonesia)

Rahasia (Mengandung isi tentang kerahasiaan dari suatu organisasi/badan tempat penelitian Skripsi ini dikerjakan)

Biasa

Zainul Hasan

Disahkan oleh,

Fathur Rizal, M.Kom.

Alamat Tetap Mahasiswa:

Desa Sidodadi, Kecamatan Paiton

Kab. Probolinggo Jawa Timur

Tanggal: 15 Agustus 2020

Tanggal: 15 Agustus 2020

**HALAMAN PERSETUJUAN**

**SKRIPSI**


**APLIKASI EKSTRAKSI TEXT DENGAN OCR  
(OPTICAL CHARACTER RECOGNITION) PADA KARTU TANDA  
MAHASISWA UNIVERSITAS NURUL JADID**


Oleh:

**ZAINUL HASAN**

Dipertahankan di depan penguji  
Pada Tanggal : 15 Agustus 2020  
Dan dinyatakan memenuhi syarat

**Komisi Pembimbing,**

  
**FATHUR RIZAL, M.Kom**  
Pembimbing I

  
**MOH FURQAN, M.Kom**  
Pembimbing II

**Tim Penguji,**

  
**AHMAD KHAIRI, M.Kom**  
Penguji I

  
**GULPI QORIK OKTAGALU P., S.Pd., M.Kom**  
Penguji II

Paiton, 21 Oktober 2020

Fakultas Teknik Universitas Nurul Jadid  
Ketua Program Studi **TEKNIK INFORMATIKA**

  
**GULPI QORIK OKTAGALU P., S.Pd., M. Kom.**

**HALAMAN PENGESAHAN**

**APLIKASI EKSTRAKSI TEXT DENGAN OCR  
(OPTICAL CHARACTER RECOGNITION) PADA KARTU TANDA  
MAHASISWA UNIVERSITAS NURUL JADID**

**SKRIPSI**

Telah dipertahankan dihadapan Dewan Penguji Skripsi Program Studi Teknik Informatika  
Fakultas Teknik Universitas Nurul Jadid Probolinggo  
dan diterima untuk memenuhi salah satu syarat memperoleh gelar

Sarjana Komputer (S. Kom) pada :

Hari : Sabtu 15 Agustus 2020

**Disusun Oleh :**

**ZAINUL HASAN**  
**16010222**

Dewan Penguji SKRIPSI

Nama Penguji I : **AHMAD KHAIRI, M.Kom**

NIDN : **0718018902**

()

Nama Penguji II : **GULPI QORIK OKTAGALU P., S.Pd., M.Kom**

NIDN : **0730109002**

()

Ketua Sidang : **ABDUL KARIM, S.Kom, M.T**

NIDN : **0015057502**

()

## PERNYATAAN ORISINALITAS SKRIPSI

Saya yang bertanda tangan dibawah ini :

Nama : **ZAINUL HASAN**  
NIM : **16010222**  
Judul Skripsi : **“APLIKASI EKSTRAKSI TEXT DENGAN OCR  
(OPTICAL CHARACTER RECOGNITION) PADA  
KARTU TANDA MAHASISWA UNIVERSITAS  
NURUL JADID”**

Menyatakan dengan sebenarnya bahwa penulisan Skripsi ini berdasarkan hasil penelitian, pemikiran dan pemaparan asli saya sendiri, baik untuk naskah laporan maupun kegiatan *programming* yang tercantum sebagai bagian dari skripsi ini, jika terdapat karya orang lain, saya akan mencantumkan sumber literasi yang jelas.

Demikian pernyataan ini saya buat dengan sesungguhnya dan apabila dikemudian hari terdapat penyimpangan dan ketidak benaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya tulis ini dan sanksi lain sesuai dengan peraturan yang berlaku di Universitas Nurul Jadid.

Demikian surat pernyataan ini saya buat dalam keadaan sadar tanpa paksaan dari pihak manapun.

Paiton, 15 Agustus 2020  
Yang membuat pernyataan



**ZAINUL HASAN**  
**16010222**

**MOTTO**

**JANGAN JADIKAN PENDIDIKAN SEBAGAI ALAT UNTUK  
MENDAPATKAN HARTA, DEMI MEMPEROLEH UANG UNTUK  
MEMPERKAYA DIRIMU. BELAJARLAH SUPAYA TIDAK MENJADI  
ORANG BODOH DAN DIBODOHI OLEH ORANG**

**\*\*\* Zainul hasan 2020 \*\*\***



## ABSTRAK

### APLIKASI EKSTRAKSI TEXT DENGAN OCR (*OPTICAL CHARACTER RECOGNITION*) PADA KARTU TANDA MAHASISWA UNIVERSITAS NURUL JADID

Zainul hasan, Fathur Rizal, M.Kom Moh. Furqan, M.Kom  
[zainulhasan502@gmail.com](mailto:zainulhasan502@gmail.com); [fathurrizal@unuja.ac.id](mailto:fathurrizal@unuja.ac.id); [moh.furqan@unuja.ac.id](mailto:moh.furqan@unuja.ac.id).

Kartu Tanda Mahasiswa atau KTM merupakan kartu tanda pengenal pada universitas, sama halnya seperti Kartu Tanda Penduduk atau KTP berfungsi sebagai tanda pengenal bahwa mahasiswa tersebut merupakan mahasiswa dari suatu universitas. KTM berisikan biodata inti pada mahasiswa, umumnya nama mahasiswa, nim, dan program studi. Pada Penelitian ini bertujuan membangun aplikasi yang dapat memindahkan text yang ada pada Kartu Tanda Mahasiswa nama, nim dan prodi dengan konversi citra kartu nama ke teks dengan melakukan ekstraksi teks. Algoritma yang digunakan untuk ekstraksi teks adalah OCR (*Optical Character Recognition*). Pengujian sistem yang dilakukan dengan menggunakan dataset KTM Universitas Nurul Jadid sebanyak 200 buah dengan menggunakan kecerahan terang dan gelap. aplikasi dirancang dengan Bahasa pemrograman Java dengan IDE Android Studio dan library Tesseract. Hasil akhir dari sistem yang dilakukan didapatkan nilai akurasi rata-rata dengan menggunakan cahaya terang 200 kali percobaan adalah 94,55% untuk data nama, nim, dan prodi.

**Kata Kunci :** Ekstraksi Teks, Kartu Tanda Mahasiswa (KTM), Optical Character Recognition (OCR), Android studio.

## KATA PENGANTAR

Segenap puji syukur penulis sampaikan kepada Allah karena dengan Rahmat dan Hidayahnya, perencanaan, pelaksanaan, dan penyelesaian skripsi, sebagai salah satu syarat menyelesaikan Program Sarjana di Fakultas Teknik Universitas Nurul Jadid Paiton Probolinggo dapat terselesaikan dengan lancar. Seiring dengan itu, penulis sangat berterima kasih kepada kedua orang tua karena orang yang paling berjasa dalam hidupku, cucuran keringat dan air mata mereka sebagai pengorbanan yang tak terhingga.

Kesuksesan ini dapat penulis raih karena dukungan banyak pihak. Oleh karena itu, penulis menyadari dan menyampaikan terima kasih yang sedalam-dalamnya kepada semua pihak yang telah ikut andil dalam menyelesaikan skripsi/tesis ini, terutama kepada :

1. Bapak KH. Abd. Hamid Wahid, M.Ag. selaku Rektor Universitas Nurul Jadid Paiton Probolinggo.
2. Bapak Moh. Furqan, M.Kom selaku Dekan Fakultas Teknik Universitas Nurul Jadid Paiton Probolinggo.
3. Bapak Gulpi Qorik Oktagalu Pratamasunu, S.Pd., M,Kom. Selaku Ketua Prodi Teknik Informatika Fakultas Teknik Universitas Nurul Jadid Paiton Probolinggo.
4. Bapak Fathur Rizal, M.Kom, dan Moh. Furqan, M.Kom selaku pembimbing dalam penulisan skripsi.
5. Semua pihak yang turut membantu mensukseskan selesainya penulisan skripsi.

Akhirnya, semoga segala amal yang telah bapak berikan kepada penulis mendapat balasan yang sebaik mungkin dari Allah SWT. Amin.

Paiton, 21 Oktober 2020

Penyusun

## DAFTAR ISI

|  |             |
|--|-------------|
| <b>HALAMAN SAMPUL</b>                                  |             |
| <b>HALAMAN JUDUL</b>                                   |             |
| <b>PERSETUJUAN PEMBIMBING SKRIPSI.....</b>             | <b>iii</b>  |
| <b>PENGESAHAN TIM PENGUJI SKRIPSI.....</b>             | <b>iv</b>   |
| <b>MOTTO.....</b>                                      | <b>vii</b>  |
| <b>ABSTRAK.....</b>                                    | <b>vii</b>  |
| <b>KATA PENGANTAR.....</b>                             | <b>viii</b> |
| <b>DAFTAR ISI.....</b>                                 | <b>9</b>    |
| <b>DAFTAR TABEL.....</b>                               | <b>10</b>   |
| <b>DAFTAR LAMPIRAN.....</b>                            | <b>12</b>   |
| 1.1 Latar Belakang.....                                | 1           |
| 1.2 Rumusan Masalah.....                               | 2           |
| 1.3 Tujuan penelitian.....                             | 2           |
| 1.4 Manfaat Penelitian.....                            | 3           |
| 1.5 Batasan Masalah.....                               | 3           |
| 2.1 Penelitian Relavan.....                            | 4           |
| 2.2 Landasan Teori.....                                | 7           |
| 2.2.1 Kartu Tanda Mahasiswa KTM.....                   | 7           |
| 2.2.2 Android.....                                     | 7           |
| 2.2.3 Android Studio.....                              | 8           |
| 2.2.4 <i>Optical Character Recognition (OCR)</i> ..... | 8           |
| 2.2.5 Konversi Citra ke Teks.....                      | 9           |
| 2.2.6 Tesseract OCR.....                               | 10          |
| 2.2.7 Akuisisi Citra.....                              | 10          |
| <b>BAB III METODE PENELITIAN.....</b>                  | <b>11</b>   |
| 3.1 Kerangka Penelitian.....                           | 11          |
| 3.2 Model pengembangan.....                            | 12          |
| <b>3.2.1 Studi Literatur.....</b>                      | <b>12</b>   |
| <b>3.2.2 Pengumpulan dataset.....</b>                  | <b>12</b>   |
| <b>3.2.3 Pre-Processing.....</b>                       | <b>12</b>   |

|                                  |  |           |
|----------------------------------|--|-----------|
| 3.2.4                            | Deteksi teks.....                                | 12        |
| 3.2.5                            | Implementasi OCR.....                            | 13        |
| 3.2.6                            | Uji coba.....                                    | 13        |
| 3.2.7                            | Penarikan Kesimpulan.....                        | 14        |
| <b>BAB IV.....</b>               |  | <b>15</b> |
| <b>HASIL DAN PEMBAHASAN.....</b> |  | <b>15</b> |
| 4.1                              | Hasil pengumpulan data.....                      | 15        |
| 4.2                              | Hasil analisa Desain.....                        | 21        |
| 4.3                              | Deteksi Teks.....                                | 22        |
| 4.4                              | <i>Optical Character Recognition (OCR)</i> ..... | 23        |
| 4.5                              | Ekstraksi teks.....                              | 23        |
| 4.6                              | Implementasi.....                                | 24        |
| 4.7                              | Uji coba.....                                    | 26        |
| <b>BAB V.....</b>                |  | <b>37</b> |
| <b>KESIMPULAN DAN SARAN.....</b> |  | <b>37</b> |
| 5.1                              | Kesimpulan.....                                  | 37        |
| 5.2                              | Saran .....                                      | 57        |
| <b>DAFTAR PUSTAKA.....</b>       |  | <b>38</b> |

## DAFTAR TABEL

|   |    |
|---|----|
| Tabel 4. 1 Hasil uji coba Ekstraksi tex.....                  | 25 |
| Tabel 4. 2 Nilai <i>Approximate Value</i> Ekstraksi text..... | 27 |



## DAFTAR GAMBAR

|   |    |
|---|----|
| Gambar 2. 1 Kartu Tanda Mahasiswa UNUJA.....                    | 7  |
| Gambar 2. 2 Skema proses OCR.....                               | 9  |
| Gambar 3. 1 Kerangka Penelitian                                 | 11 |
| Gambar 4. 1 Dataset Citra KTM.....                              | Z  |
| Gambar 4. 2 Arsitektur sistem.....                              | 19 |
| Gambar 4. 3 Diagram Alur ekstraksi teks.....                    | 20 |
| Gambar 4. 4 Ilustrasi Pemisahan.....                            | 20 |
| Gambar 4. 5 Deteksi teks.....                                   | 21 |
| Gambar 4. 6 Hasil ekstraksi teks.....                           | 22 |
| Gambar 4. 7 Tampilan Menu Utama.....                            | 23 |
| Gambar 4. 8 Tampilan <i>Optical Character Recognition</i> ..... | 24 |
| Gambar 4. 9 Hasil ekstrak text KTM.....                         | 24 |

## DAFTAR LAMPIRAN

Lampiran.1. Dataset50

Lampiran.2. Hasil Plagiasi

Lampiran.3. Biodata diri

