

**PEMANFAATAN DESKTOP OCR UNTUK
MENDAPATKAN AKURASI PENGENALAN
NASKAH PADA SMARTPHONE**

TESIS



Abdul Muis Sobri

1111601124

**PROGRAM STUDI MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR**

**JAKARTA
2014**

PEMANFAATAN DESKTOP OCR UNTUK MENDAPATKAN AKURASI PENGENALAN NASKAH PADA SMARTPHONE

TESIS

Diajukan untuk memenuhi salah satu persyaratan
memperoleh gelar Magister Ilmu Komputer (MKOM)



Abdul Muis Sobri

1111601124

**PROGRAM STUDI MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR**

**JAKARTA
2014**



**PROGRAM STUDI MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR**

LEMBAR PENGESAHAN

Nama Mahasiswa : Abdul Muis Sobri
Nomor Induk Mahasiswa : 1111601124
Jenjang Studi : Magister
Konsentrasi : Komputasi Terapan
Judul : Pemanfaatan Dekstop OCR untuk Mendapatkan Akurasi Pengenalan Naskah pada Smartphone

Tgl. 20 Maret 2014

Tim Penguji:

Tanda Tangan:

Penguji I (Ketua Penguji)

Dr. Moedjiono, M.Sc

Penguji II

Mardi Hardjianto, M.Kom

Pembimbing Utama,

Dr. Ir. Nazori Az, M.T

Pembimbing Pendamping,

Mardi Hardjianto, M.Kom

Ketua Program Studi

Dr. Ir. Nazori Az, M.T



**PROGRAM STUDI: MAGISTER ILMU KOMPUTER (MKOM)
PROGRAM PASCASARJANA
UNIVERSITAS BUDI LUHUR**

LEMBAR PERNYATAAN

Nama Mahasiswa : Abdul Muis Sobri
Nomor Induk Mahasiswa : 1111601124
Program Studi : Magister Ilmu Komputer
Konsentrasi : Rekayasa Komputasi Terapan
Fakultas/Program : Pascasarjana

menyatakan bahwa tesis yang berjudul:

Pemanfaatan Desktop OCR untuk mendapatkan
Akurasi Pengenalan Naskah pada Smartphone

1. Merupakan hasil karya tulis ilmiah sendiri dan bukan merupakan karya yang pernah diajukan untuk memperoleh gelar akademik oleh pihak lain,
2. Saya ijin untuk dikelola oleh Universitas Budi Luhur sesuai dengan norma hukum dan etika yang berlaku.

Pernyataan ini saya buat dengan penuh tanggung jawab dan saya bersedia menerima konsekuensi apapun sesuai aturan yang berlaku apabila di kemudian hari pernyataan ini tidak benar.

Jakarta, 20 Maret 2014.



(Abdul Muis Sobri)

LEMBAR PENGESAHAN

Nama Mahasiswa : Abdul Muis Sobri
Nomor Induk Mahasiswa : 1111601124
Jenjang Studi : Magister
Konsentrasi : Komputasi Terapan
Judul : Pemanfaatan Dekstop OCR untuk Mendapatkan
Akurasi Pengenalan Naskah pada Smartphone

Jakarta,

Tim Penguji:

Tanda Tangan:

Moderator,
Dr. Ir. Nazori Az, M.T

Penguji I (Ketua Penguji)
Dr. Moedjiono, M.Sc

Penguji II
Mardi Hardjianto, M.Kom

Pembimbing Utama,
Dr. Ir. Nazori Az, M.T

Pembimbing Pendamping,
Mardi Hardjianto, M.Kom

Ketua Program Studi

Dr. Ir. Nazori Az, M.T

ABSTRAK

Meningkatnya kemampuan peralatan komunikasi bergerak dari sekadar peralatan untuk telepon dan berkirim pesan singkat menjadi perangkat yang bisa melakukan banyak hal, dan dikenal kemudian dengan *smartphone*, menjadikan aplikasi yang berjalan pada *smartphone* tersebut berkembang dengan pesat, salah satunya adalah aplikasi Optical Character Recognition (OCR). Saat ini di pasaran terdapat beberapa aplikasi OCR yang berjalan pada *smartphone*, baik pada platform android maupun iOS (*iPhone*, *iPad*, *iPod*) seperti OCR Kit, Mobile OCR, Scanthing dan Rocsan. Teks yang dihasilkan dari aplikasi OCR *smartphone* tersebut sudah mampu mengenali karakter/ huruf dengan baik, namun masih belum mampu mengenali dengan baik format karakter berupa format normal, *italic*, dan *bold* pada naskah. Sementara OCR pada komputer desktop menggunakan Abby Fine Reader ataupun Cvision OCR mampu mengenali dengan baik format karakter, format paragraf dan tabel. Dalam tesis ini, penulis mengajukan rancangan pemanfaatan desktop OCR agar *smartphone* dapat mengenali dengan baik variasi format karakter dengan baik. Penulis melakukan penelitian dan analisis mengenai akurasi dalam OCR terutama OCR yang digunakan untuk membantu proses cetak ulang naskah atau reproduksi buku. Dengan menggunakan rancangan yang diajukan, akurasi format naskah dapat dikenali dengan baik, serta diproses dalam waktu yang relatif cepat, sehingga rancangan yang penulis ajukan dapat digunakan membantu proses reproduksi naskah, terutama untuk naskah yang memerlukan ketelitian akurasi format.

Kata kunci: OCR, smartphone, akurasi, format karakter, server, android, iOs.

ABSTRACT

The increasing of mobile communications capabilities than just a communication tool for telephone and sending short messages become a device that can do many things--later known as smartphone--makes the applications running on the smartphone are growing rapidly, one of the applications is the Optical Character Recognition (OCR). There are some OCR applications running on smartphone, both on android and iOS platforms (iPhone, iPad, iPod) as OCR Kit, Mobile OCR, Scanthing and Rocsan. The text generated from those OCR application has been able to recognize the characters/ letters properly, but still has not been able to identify character formats such as normal, italic, and bold, underline on the manuscript. While OCR on a desktop computer using Abby Fine Reader or Cvision OCR is able to perform character recognition with good format accuracy. In this thesis, the author proposes a framework of desktop OCR utilization using android smartphone, so that the smartphone can do OCR with good character recognition. The author conducts research and analyzes about importance of OCR accuracy, primarily the OCR that used to assist book reproduction. By using the proposed design, the format characters can be recognized well, and processed quickly enough. So, the proposed design can be used in the manuscript reproduction, especially for manuscript that requires precision accuracy.

Keywords— OCR, smartphone, akurasi, character format, server, android.

KATA PENGANTAR

Assalamualaikum Wr. Wb.

Puji dan syukur senantiasa kami panjatkan kehadirat Allah SWT yang tiada pernah berhenti melimpahkan Rahmat dan Ridho kepada hamba-hamba-Nya, sehingga penulis dapat menyelesaikan penulisan tesis yang berjudul “PEMANFAATAN SERVER-BASED MOBILE OCR UNTUK MENDAPATKAN AKURASI PENGENALAN NASKAH PADA SMARTPHONE”. Shalawat serta salam senantiasa tercurahkan kepada Rasulullah SAW yang telah membawa umatnya dari zaman kegelapan menuju zaman yang penuh dengan risalah Islam, sehingga penulis dapat melakukan kewajiban menuntut ilmu sebagai salah satu kewajiban umat Rasulullah SAW.

Judul tesis ini merupakan salah satu upaya jalan keluar dari permasalahan yang sering dihadapi penulis dalam bekerja di dunia penerbitan. Dimana ketika proses produksi naskah, seringkali terkendala oleh lamanya proses pengetikan dan koreksi. Penulis berharap, dari apa yang penulis sampaikan dalam tesis ini, bisa menjadi solusi atas permasalahan yang sering penulis hadapi khususnya, dan bagi pembaca yang sering bekerja dengan naskah pada umumnya.

Penulisan tesis tidak serta merta selesai begitu saja tanpa melibatkan banyak pihak yang baik secara langsung maupun tidak langsung ikut membantu menyelesaikan. Sudah sepantasnyalah penulis mengucapkan terima kasih yang tulus kepada:

1. Istri tercinta, Ellis Ferdianingsih, yang senantiasa memberikan semangat dan dukungan, yang juga sering ditinggalkan penulis karena mengejar pelajaran dan perkuliahan.
2. Ayah, Ibu yang tiada pernah lupa akan kasih sayang, yang senantiasa mendorong untuk segera menyelesaikan studi ini.
3. Anak-anaku tersayang, Ali Wafie, Hasna Aliya, dan Muhammad Afdhal Bilal yang harus sering memendam keinginan bermain bersama karena waktu yang banyak dipakai untuk menyelesaikan studi dan penulisan tesis ini.

4. Adik-adik dan kakakku yang tidak pernah jemu mendorong dan memberi semangat kepada penulis untuk terus belajar.
5. Bapak Dr. Moedjiono, M.Sc, Bapak Dr. Ir. Nazori AZ, Bapak Mardi Hardjianto, M.Kom dan Bapak-bapak Dosen yang tidak bisa saya sebutkan satu-persatu, yang banyak membimbing dan memberi masukan dalam penyusunan tesis ini baik secara langsung maupun tidak langsung.
6. Kawan-kawan kelompok belajar “Ciputat”; Ahmad Farisi, Agnes Aryasanti, Nurmalia yang selalu saling berbagi dan membantu dalam belajar, terutama Bapak Suwato Komala yang tiada bosan meluangkan waktu hampir 5 jam setiap minggu menembus kemacetan Jakarta demi belajar kelompok bersama, sungguh sebuah motivasi dan semangat yang luar biasa.

Tidak lupa pula, kepada pihak-pihak lain yang tidak dapat penulis sebutkan satu-persatu yang sedikit banyak ikut memberikan kontribusi demi selesainya tesis ini, penulis juga menyampaikan terima kasih dan penghargaan yang sebesar-besarnya.

Akhir kata, karena manusia memang tempatnya salah dan ketidaksempurnaan, karena itulah penulis menyadari sepenuhnya bahwa tesis ini masih jauh dari sempurna. Untuk itu, penulis berharap kritik dan saran yang membangun demi kesempurnaan tulisan ini.

Wassalam.

Jakarta, Juni 2013

Abdul Muis Sobri

DAFTAR ISI

	Halaman
ABSTRAK	i
KATA PENGANTAR	iii
DAFTAR ISI	v
DAFTAR GAMBAR	viii
DAFTAR TABEL	ix
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Masalah Penelitian	3
1.2.1. Identifikasi Masalah	3
1.2.2. Batasan Masalah	3
1.2.3. Rumusan Masalah	3
1.3. Tujuan dan Manfaat Penelitian	4
1.3.1. Tujuan Penelitian	4
1.3.2. Manfaat Penelitian	4
1.4. Sistematika Penulisan	4
1.5. Daftar Istilah	5
BAB II LANDASAN PEMIKIRAN	6
2.1. Tinjauan Pustaka	6
2.1.1. OCR	6
2.1.2. Android	13
2.1.3. OOP	20
2.1.4. OOAD (Object-Oriented Analysis and Design)	21

2.1.5. Jaminan Kualitas Perangkat Lunak (SQA)	31
2.1.6. ISO 9126	35
2.2. Tinjauan Studi	37
2.3. Tinjauan Objek Penelitian	42
2.3.1. Perangkat Keras	42
2.3.2. Perangkat Lunak	43
2.4. Kerangka Konsep/ Pola Pikir Pemecahan Masalah	45
2.5. Hipotesis Penelitian	46
BAB III RANCANGAN PENELITIAN	47
3.1. Metode Penelitian	47
3.2. Metode Pengumpulan Data	47
3.3. Sumber Data	48
3.4. Data Penelitian	48
3.5. Teknis Analisa Data	48
3.6. Langkah Penelitian	48
3.7. Jadwal Penelitian	60
BAB IV PEMBAHASAN HASIL PENELITIAN	62
4.1. Temuan, Analisis, dan Interpretasi	62
4.1.1. Pengujian OCR	62
4.1.2. Pengujian Rancangan	73
4.2. Implikasi Penelitian	79
4.2.1. Aspek Penelitian	80
4.2.2. Aspek Manajerial	81
4.2.3. Aspek Penelitian Lanjutan	81
4.3. Rencana Implementasi	82

BAB V	PENUTUP	84
5.1.	Kesimpulan	84
5.2.	Saran	85
DAFTAR PUSTAKA		87
LAMPIRAN-LAMPIRAN		90
DAFTAR RIWAYAT HIDUP		113



DAFTAR GAMBAR

Gambar	Halaman
II-1. Contoh hasil perekaman gambar menggunakan kamera smartphone ..	10
II-2. Tampilan Android <i>Play Store</i>	17
II-3. Arsitektur Android	18
II-4. Diagram <i>Use Case</i>	25
II-5. Elemen-elemen <i>Use Case</i>	25
II-6. Contoh <i>Class</i>	27
II-7. <i>Class state: attributes</i>	27
II-8. <i>Class behaviour : operations</i>	27
II-9. Contoh <i>class diagram</i>	28
II-10. Contoh <i>activity diagram</i>	29
II-11. Contoh interaksi pada <i>sequence diagram</i>	31
II-12. Bentuk panah <i>sequence diagram</i>	31
II-13. Enam Karakteristik Kualitas ISO/IEC 9126	36
II-14. Spesifikasi <i>Tes Quality Model</i>	37
II-15. Bagan pola pikir pemecahan masalah	45
III-1. Langkah Penelitian	49
III-2. Arsitektur <i>Server-Based mobile OCR</i>	51
III-3. <i>Use case Diagram</i>	52
III-4. <i>Activity Diagram</i>	55
III-5. <i>Squence Diagram</i>	56
III-6. <i>User Interface</i> rancangan	58
III-7. <i>Online Text Compare</i>	53
IV-1. Spesifikasi koneksi yang digunakan	57
IV.2. Contoh perbandingan naskah asli dan hasil OCR rancangan	65
IV.3. Contoh perbandingan naskah asli dan hasil OCR Instantly Pro	67
IV.4. Contoh perbandingan naskah asli dan hasil OCR Mobile OCR	68
IV.5. Contoh perbandingan naskah asli dan hasil OCR Kit	70
IV-6. Grafik Perbandingan Hasil Pengujian	71

DAFTAR TABEL

Tabel	Halaman
II-1. Diagram konsep dasar UML	23
II-2. Sasasaran, atribut, dan metrik kualitas perangkat lunak	33
II-3. Tinjauan Studi	39
III-1. <i>Use Case Description</i>	53
III-2. Jadwal Penelitian	61
IV-1. Tabel Objek Naskah	63
IV-2. Hasil pengujian menggunakan rancangan usulan	66
IV-3. Hasil pengujian menggunakan OCR Instantly Pro	67
IV-4. Hasil pengujian menggunakan Mobile OCR	69
IV-5. Hasil pengujian menggunakan OCR Kit	70
IV-6. Perbandingan Hasil Pengujian	71
IV-7. Pengujian interface sistem	73
IV-8. Pengujian fungsi dasar sistem	74
IV-9. Rentang Skor	75
IV-10. Skor aspek <i>functionality</i>	75
IV-11. Skor aspek <i>reliability</i>	76
IV-12. Skor aspek <i>usability</i>	77
IV-13. Skor aspek <i>portability</i>	78
IV-14. Rating Akhir	79
IV-15. Rencana Implementasi	82

DAFTAR PUSTAKA

- [**A. Pender 2002**] A. Pender, T., 2002. *UML Weekend Crash Course*, New York: Wiley Publishing, Inc.
- [**Android Anon 2013**] Anon, 2013. Android KitKat 4.4 Smart, simple, and truly yours. Available at: <http://www.android.com/versions/kit-kat-4-4/> [Accessed December 30, 2013].
- [**Anon ISO n.d.**] Anon, ISO 9126: The Standard of Reference. Available at: <http://www.cse.dcu.ie/essiscope/sm2/9126ref.html> [Accessed February 8, 2014a].
- [**Anon n.d.**] Anon, ISO9126 - Software Quality Characteristics. Available at: <http://www.sqa.net/iso9126.html> [Accessed February 8, 2014b].
- [**Azis 2005**] Azis, M.F., 2005. *Object Oriented Programming Dengan PHP 5*, Jakarta: PT. Elex Media Komputindo.
- [**Bae et al. 2005**] Bae, K.S. et al., 2005. Character Recognition System for Cellular Phone with Camera. *29th Annual International Computer Software and Applications Conference (COMPSAC'05)*, 1, pp.539–544. Available at: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1510081>.
- [**Bassil & Alwani 2012**] Bassil, Y. & Alwani, M., 2012. OCR Post-Processing Error Correction Algorithm Using Google's Online Spelling Suggestion. *Journal of Emerging Trends in Computing and Information Sciences*, 3(1), pp.90–99.
- [**Bieniecki et al. 2007**] Bieniecki, W., Grabowski, S. & Rozenberg, W., 2007. Image Preprocessing for Improving OCR Accuracy. *2007 International Conference on Perspective Technologies and Methods in MEMS Design*, pp.75–80. Available at: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4283429>.
- [**Canedo-Rodríguez 2012**] Canedo-Rodríguez, A., 2012. Efficient Text Extraction Algorithm Using Color Clustering for Language Translation in Mobile Phone. *Journal of Signal and Information Processing*, 03(02), pp.228–237. Available at: <http://www.scirp.org/journal/Paper-Download.aspx?DOI=10.4236/jsip.2012.32031> [Accessed December 15, 2013].
- [**Cheriet et al. 2007**] Cheriet, M. et al., 2007. *CHARACTER RECOGNITION SYSTEMS A Guide for Students and Practioners*, New Jersey: JohnWiley & Sons, Inc.
- [**Dantes 2012**] Dantes, N., 2012. *Metode Penelitian P*. Christian, ed., Yogyakarta: CV. Andi Offset.