

PENERAPAN METODE EKSTRASI IMAGE KE TEXT DENGAN OPTICAL CHARACTER RECOGNITION UNTUK OTOMATISASI DATA KEPENDUDUKAN

M.Makmun Effendi¹, A. Yudi Permana², Ismasari Nawangsih³

^{1,2,3}Universitas Pelita Bangsa
effendiyan@pelitabangsa.ac.id

ABSTRAK

Pemerintah sudah membuat sistem informasi kependudukan menggunakan *electronic* Kartu Tanda Penduduk yang lebih dikenal e-KTP, namun pemanfaatan e-KTP masih belum dilaksanakan secara optimal, sehingga bila dibutuhkan maka harus fotocopy kemudian data e-KTP tersebut di masukkan kembali ke *database* organisasi yang memerlukan sehingga memerlukan kerja dua kali dan penyimpanan fisik KTP memiliki kelemahan dalam asumsi yang sulit dipenuhi. Penelitian ini bertujuan untuk menerapkan metode ekstraksi image ke text dengan *optical character recognition* untuk otomatisasi data kependudukan. Untuk mengambil data e-KTP agar dapat diinput dalam database diperlukan *Tesseract Optical Character Recognition* dan bantuan *imagemagic* mampu mengekstrak citra. Metode penelitian ini dilakukan dengan deskriptif kualitatif. Simpulan adalah data yang ada dalam *image* sudah bisa berubah *text* dan bisa masuk kedalam *database* yang telah disediakan. E-KTP yang di *scan* bisa dijadikan dokumen, sehingga tanpa perlu lagi *fotocopy* sehingga mempermudah layanan administrasi.

Kata kunci : *Tesseract, Optical Character Reconigtion, Imagemagick, e-KTP*

ABSTRACT

The government has made a population information system using electronic Identity Cards, better known as e-KTPs, but the use of e-KTPs has not been implemented optimally, so that if needed then photocopies then e-KTP data will be re-entered into the database of organizations that need so requires twice the work and physical storage of ID cards has weaknesses in assumptions that are difficult to fulfill. This study aims to apply the image extraction method to text with optical character recognition to automate population data. To retrieve e-KTP data so that it can be inputted into a database, it needs Tesseract Optical Character Recognition and imagemagic assistance capable of extracting images. The method of this research is descriptive qualitative. Conclusion is the data in the image can change text and can be entered into the database that has been provided. The scanned e-KTP can be used as a document, so there is no need for photocopies to simplify administrative services.

Keyword : *Tesseract, Optical Character Reconigtion, Imagemagick, e-KTP*

PENDAHULUAN

Pemerintah sudah membuat sistem informasi kependudukan menggunakan *electronic* Kartu Tanda Penduduk yang lebih dikenal e-KTP. E-KTP ini menggunakan teknologi digital dan diharapkan dapat digunakan untuk berbagai keperluan. Menurut (Sutanta & Ashari, 2012) pada e-KTP memiliki *chip* yang digunakan sebagai *smart card* dan memuat 4 sampai dengan 8 KB, yang memuat data NIK, nama, tempat dan tanggal lahir, jenis kelamin, agama, status perkawinan, golongan darah, alamat, pekerjaan, kewarganegaraan, foto, masa berlaku, tempat dan tanggal dikeluarkan, tandatangan, serta nama dan nomor induk pegawai pejabat yang menandatangani dan ada tanda tangan pemilik e-KTP pada bagian depan. Informasi pada e-KTP telah cukup memuat data masing masing individu yang tersusun rapi sesuai dengan kolom yang telah diatur dengan

standarisasi yang sama, walaupun berbeda daerah dan juga sangat jelas sekali ada Desa, Kecamatan, Kabupaten dan propinsi.

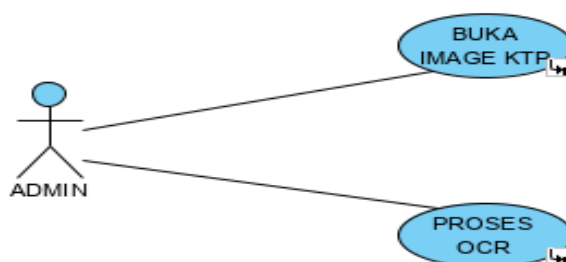
Pada saat ini pemanfaatan e-KTP masih belum dilaksanakan secara optimal, baik di pemerintahan maupun di instansi swasta ataupun perusahaan swasta sehingga bila dibutuhkan maka harus fotocopy kemudian data e-KTP tersebut di input kembali ke *database* organisasi yang memerlukan baik pemerintah maupun swasta, sehingga memerlukan kerja dua kali dan penyimpanan fisik KTP memerlukan kelemahan dalam asumsi yang sulit dipenuhi, yaitu independensi *feature* area dan penyimpanan yang tidak baik dapat merusak *fotocopy* KTP, bisa juga rusak karena lembab serta pencariannya juga agak sulit (Rahayu, Warsono, & Dwimawanti, 2014). Dengan bantuan mesin *scanner* maka pengambilan e-KTP dapat dilakukan dan disimpan pada komputer. Pengambilan data e-KTP agar dapat diinput dalam *database* diperlukan *Tesseract Optical Character Recognition* dan bantuan *imagemagic* mampu mengekstrak citra ataupun *image* dan keluarannya berupa *text* yang langsung dimasukan kolom yang sesuai dengan fungsi data tersebut dengan sistem berbasis *website* sehingga mempermudah layanan administrasi (Kadir & Susanto, 2012).

Proses *Optical Recognition Character* ini cukup efektif dalam pengklasifikasian teks yang mempunyai banyak kelasnya. (Mohamed Cheriet, Nawwaf Kharma, Cheng-Lin Liu, 2007) Penelitian ini bertujuan untuk menerapkan metode ekstraksi image ke text dengan *optical character recognition* untuk otomatisasi data kependudukan mempermudah petugas dalam melakukan *input* data, dan penyimpanan e-KTP saat dibutuhkan sudah ada di dalam komputer karena prosesnya tidak input ulang, karena saat di proses sudah direkam berbentuk gambar. Dan e-KTP lebih awet karena tidak perlu adanya pengulangan *copy/scan*.

METODE

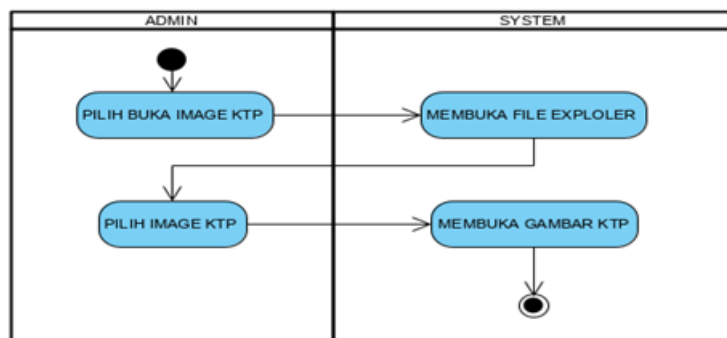
Penelitian ini menggunakan metode Deskriptif Kualitatif (Sugiyono, 2013) dengan model eksperimen berdasarkan fakta dan keadaan yang sebenarnya dari obyek yang diteliti, sementara untuk pengembangan sistem menggunakan model *Unified modeling language* (UML). Hasil desain sebagai berikut:

Use Case Diagram

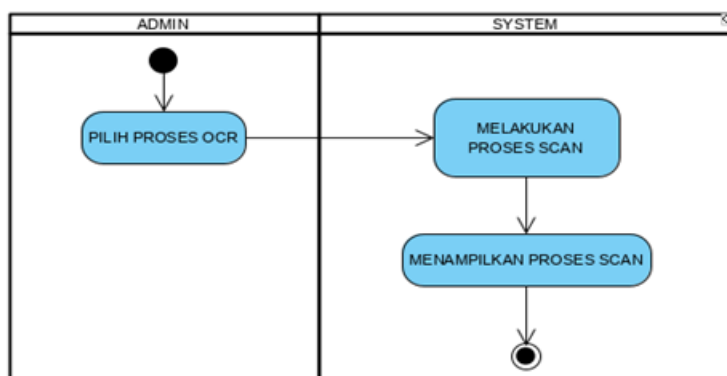


Gambar 1. Use Case Diagram

Activity Diagram

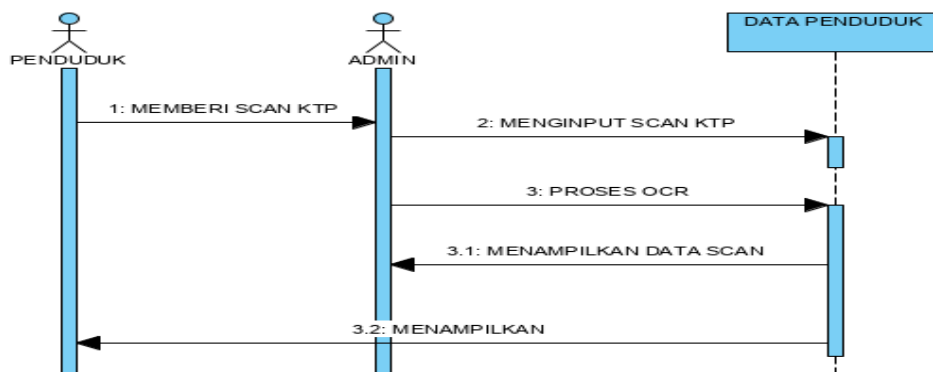


Gambar 2. Activity Diagram Buka Image KTP



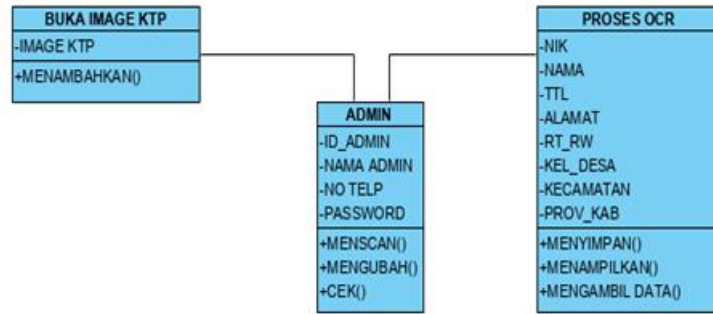
Gambar 3. Activity Diagram Proses OCR KTP

Sequence Diagram



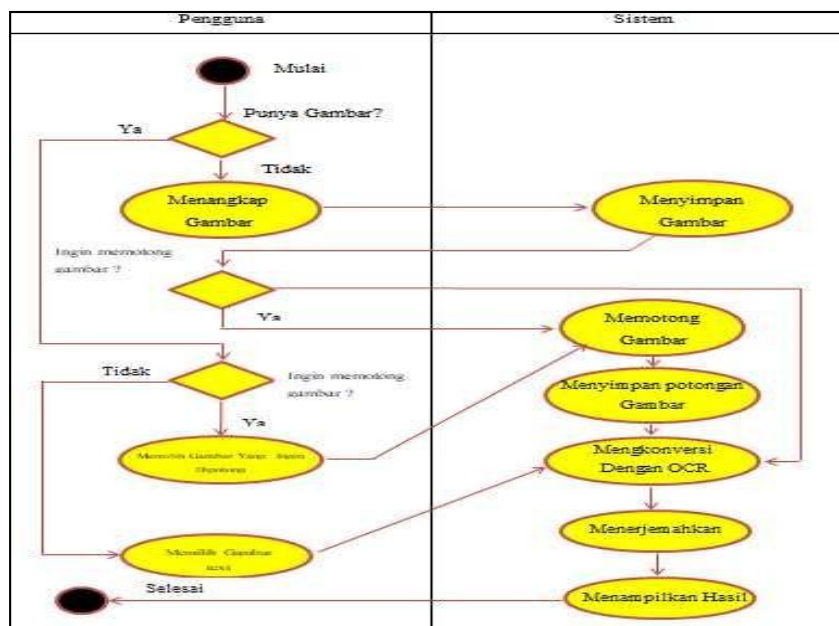
Gambar 4. Sequence Diagram

Class Diagram



Gambar 5. Class Diagram

Tahapan Dalam Image Extract



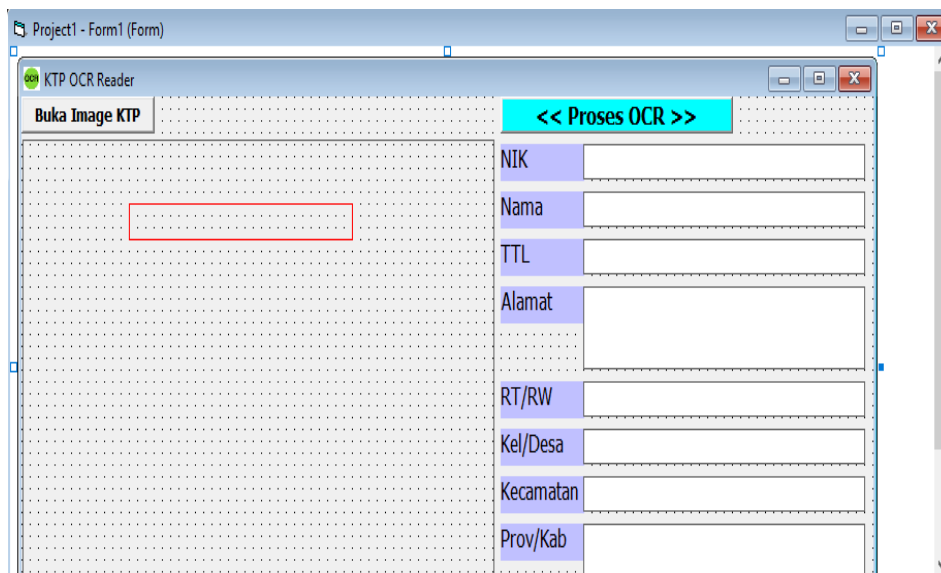
Gambar 6. Tahapan Image Extraction

HASIL

Dari hasil penelitian yang dilakukan dalam pembuatan aplikasi *Optical Character Recognition* untuk e-KTP, dilakukan uji coba untuk 10 buah KTP sampel yang hasilnya bisa di *extract* dengan baik. Hasil *Optical Character Recognition* hasil ujicoba bisa masuk ke dalam *database* sesuai dengan kolom yang telah ditentukan dan disediakan.

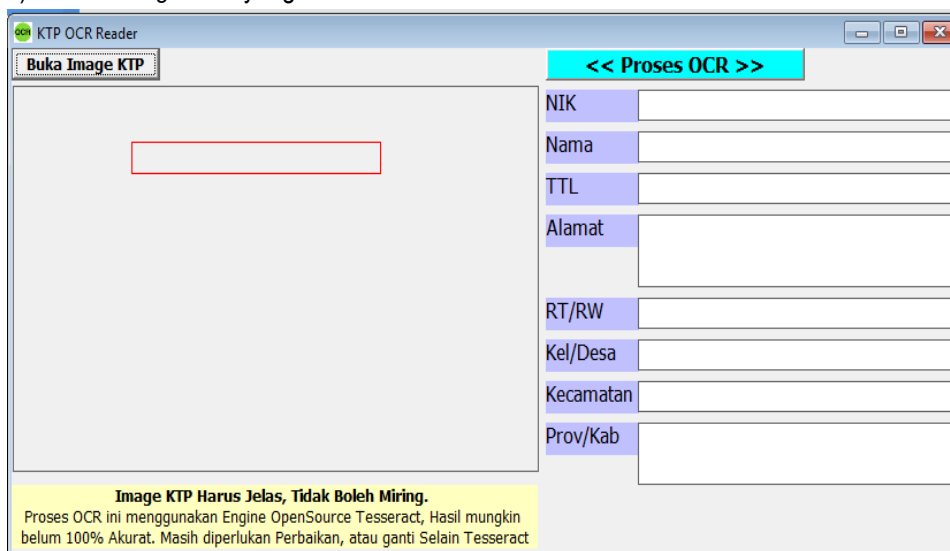
Cara Kerja dan Hasil *Image Processing* didapat dengan desain, sebagai berikut:

- 1) Desain Buka Aplikasi



Gambar 7. Tahapan Image Extraction

2) Buka Image KTP yang telah di scan



Gambar 8. Proses OCR

3) Kemudian Pilih Proses OCR



Gambar 9. Hasil OCR

SIMPULAN

Dari penelitian yang telah dilakukan dapat disimpulkan data yang ada di *image* KTP bisa di *extract* menjadi berupa *text* yang sesuai dengan *image* dalam KTP. Penggunaan aplikasi ini tanpa lagi perlu untuk menulis, hanya butuh beberapa detik maka data yang ada dalam *image* sudah bisa berubah *text* dan bisa masuk kedalam *database* yang telah disediakan. E-KTP yang di *scan* bisa dijadikan dokumen, sehingga tanpa perlu lagi *fotocopy*. Aplikasi ini masih belum sempurna dan masih perlu adanya *adjust* ukuran e-KTP dan bila kabur, atau tulisannya sedikit hilang tidak terbaca. Melakukan penelitian lanjutan untuk hasil yang lebih maksimal.

UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih yang sebesar besarnya kepada kepada pihak institusi STT pelita bangsa dan LPPM STT Pelita bangsa yang telah memberikan ruang penelitian terhadap dosen prodi teknik informatika. Kami ucapkan terima kasih atas kesediaan pengurus RT01/RW 017 BLOK E Perum KSB Desa Sukaragam Kecamatan Serang Baru Bekasi yang telah mengizinkan kami sebagai tim peneliti untuk meminta data e-KTP Indonesia, yang akan dianalisis. Pada akhirnya kami berharap banyak sebagai peneliti bisa memberikan kontribusi buat kemajuan dan pengembangan keilmuan di bidang teks processing dan klasifikasi teks.

DAFTAR RUJUKAN

- Kadir, A., & Susanto, A. (2012). *citra.pdf*.
Mohamed Cheriet, Nawwaf Kharma, Cheng-Lin Liu, C. S.-. (2007). *Character Recognition Systems: A Guide for Students and Practitioners*. Retrieved from https://books.google.ca/books?hl=en&lr=&id=txYpjiK_BmgC&oi=fnd&pg=PR7&ots=d

- FipoXvmtz&sig=90SchloWai8Rx56sbJQNkW1J4HA#v=onepage&q&f=false
- Rahayu, I. P., Warsono, H., & Dwimawanti, I. H. (2014). Analisis Kualitas Pelayanan E-KTP Di Kecamatan Gayamsari Kota Semarang Oleh : *Journal of Public Policy and Management Review*, 3(2), 271–280.
- Sugiyono. (2013). *Metode penelitian pendidikan:(pendekatan kuantitatif, kualitatif dan R & D)*. Bandung: Alfabeta.
- Sutanta, E., & Ashari, A. (2012). Jurnal Ilmu Komputer - Volume 5 - No 2 – September 2012. *Jurnal Ilmu Komputer*, 5(1), 29–35.