**IMPLEMNTASI OPTICAL CHARACTER RECOGNITION UNTUK MENGINPUT JUDUL DAN ABSTRAK SKRIPSI (STUDI KASUS : INSTITUT BISNIS DAN INFORMATIKA KWIK KIAN GIE)**

**Ivan Immanuel1), Joko Susilo2)**

Program Studi Teknik Informatika

InstitutBisnis dan Informatika Kwik Kian Gie

Jl. YosSudarso Kav.87 Sunter Jakarta Utara 14350

1 59150092@student.kwikkiangie.ac.id, 2joko.susilo@kwikkiangie.ac.id

**ABSTRAK**

Ivan Imanuel / 59150092 / 2019 / Implementasi *Optical* *Character* *Recognition* (OCR) Untuk Menginput Judul dan Abstrak Skripsi Berbasis Web (Studi Kasus : Institut Bisnis dan Informatika Kwik Kian Gie)/ Pembimbing: Joko Susilo,S.Kom., M.M., M.Kom Perkembangan teknologi di Indonesia saat ini semakin maju dalam kehidupan masyarakat dan tidak dapat dihindarkan. Penggunaan *Artificial* *Intelligence* dalam membantu manusia dalam menangani masalah semakin berkembang. Manusia dapat memanfaatkan media komputer/smartphone di era teknologi sekarang ini. Salah satu pemanfaatannya adalah *Optical* *Character* *Recognition*. Penelitian ini dilatarbelakangi oleh permasalahan dimana sistem yang berjalan memerlukan pengembangan dari segi teknologi, karena pada sistem yang terdahulu masi melakukan input manual. Fitur ini juga dibuat untuk menangani masalah yang akan dihadapi kedepannya. *Prototipe* ini dibuat untuk memudahkan mahasiswa yang telah wisuda agar tidak perlu jauh – jauh datang kekampus hanya untuk memberikan *soft* *copy* judul dan abstrak skripsi dan membantu pihak Bagian Administrasi Akademik dan Kemahasiswaan Institut Bisnis dan Informatika Kwik Kian Gie dalam menginput data – data judul dan abstrak skripsi yang mahasiswa berikan.

# **Kata Kunci**: *Optical* *Character* *Recognition*, *OCR*, *Prototipe*, *OCR*.*Space* *API*, *Cloud* *API*.

**ABSTRACT**

Ivan Imanuel / 59150092/ 2019 / Optical Character Recognition (OCR) Application To Input Thesis Titles and Abstracts (Case Study: Kwik Kian Gie Institute of Business and Informatics) / Advisor: Joko Susilo, S.Kom., M.M., M.Kom Technology developments in Indonesia are currently advancing in life and cannot be avoided. The use of Artificial Intelligence in helping humans to deal with problems is growing. Human also use computer / smartphone media in this technology era. One of the uses is Optical Character Recognition. This research is motivated by the problems in which the system running requires development in terms of technology, because of the system we did previously, was still manually input. This feature is also made to handle problems that will be faced in the future. This prototype is made to make it easier for students who have graduated so they don't need to come so far to the campus just to give a soft copy of the title and thesis abstract and help the Academic and Financial Administration Agency Kwik Kian Gie's Business and Informatics Institute in inputting title data and thesis abstracts that students give.

**Keywords**: *Optical* *Character* *Recognition*, *OCR*, *Prototipe*, OCR.*Space* *API*, *Cloud* *API*.

**PENDAHULUAN**

Pada abad ke-21 ini terutama didunia kerja hingga pendidikan, menginput data kedalam data base atau memori internal komputer merupakan suatu hal yang wajar. Hal tersebut dilakukan untuk menghemat waktu pekerjaan. Namun dengan seiring berkembangnya dunia teknologi menginput data mengunakan cara manual sudah mulai meresahkan masyarakat terutama pada data yang sukar didapat. Tentunya, hal ini dapat memperlama waktu suatu proses pengerjaan.

Institut Bisnis dan Informatika Kwik Kian Gie, setiap tahunnya akan meluluskan mahasisa dan mahasiswi yang telah menyelesaikan pendidikan jenjang Sarjana (S1) dan Magister (S2), mahasiswa – mahasiswa yang telah lulus Sarjana harus memberikan judul skripsi dan abstrak skripsi kepada pihak kampus yaitu Bagian Administrasi Akademik dan ( BAAK ) setelah itu pihak BAAK akan mengirimkan Google Form ke mahasiswa . Kejadian ini akan terus berlangsung setiap tahun saat kelulusan mahasiswa, dan pihak kampus juga harus bekerja ekstra dalam menginput data judul dan abstrak skripsi yang diberikan mahasiswa kedalam data base.

Seiring dengan berkembangnya teknologi di masa ini, banyak sekali cabang ilmu teknologi yang bermuncul dan berfungsi untuk membantu manusia dalam bidang yang berbeda – beda. Beberapa cabang ilmu yang berkembang di antaranya adalah, Software Engineering, Web Development, IT Security, Artificial Intelligence (AI), dan lain- lain. Perkembang teknologi yang terbilang pesat adalah Artificial Intelligence. Hasil produk dari teknologi AI ini pun juga semakin bertambah, seperti Robot, Expert Systems, Natural Language Processing, Computer Vision , Optical Character Recognition ( OCR ) dan lain- lain.

*Optical* *Character* *Recognition* ( *OCR* ) merupakan salah satu aplikasi dalam mengelolah informasi. OCR digunakan untuk mengenali karkter optis pada suatu gambar atau foto yang kemudian akan dianalisa dan menghasilkan suatu karater – karakter yang dapat diolah kedalam suatu sistem. Dengan adanya *OCR* banyak sekali aplikasi – aplikasi yang dapat diciptakan untuk membantu pekerjaan manusia dalam menginput data, seperti menginput data KTP, membaca karakter pada gambar, dan lain – lain. Dengan fungsi yang seperti itu *OCR* dapat digunakan untuk membantu user BAAK dalam menginput data judul dan abstrak skripsi mahasiswa yang sebenarnya sukar didapat, karena mahasiswa harus datang dari rumah hingga ke-kampus. Penerapan teknologi *OCR* dalam hal ini dimaksudkan untuk penghematan waktu, tenaga, dan biaya yang dikeluarkan dari kedua pihak, dari kampus maupun mahasiswa. Dengan adanya beberapa kebutuhan untuk penerapan *OCR* pada pihak BAAK yang telah dijabarkan di atas, serta adanya fitur *OCR* yang memiliki Implementasi *Optical* *Character* *Recognition* Untuk Menginput Judul dan Abstrak Skripsi Pada Pihak Bagian Administrasi Akademik dan Kemahasiswaan.

***PROTOTYPING***

Menurut Car dan Verner dalam buku Pedro Isaias dan Tomayess Issa, 2015 : 33).Model *prototyping* adalah kerangka kerja berulang yang menjadi pusat dari banyak pendekatan yang lebih cepat untuk pengembangan perangkat lunak, sejak awal 1980-an. Pada 1997, Carr dan Verner mengamati bahwa model *prototyping* lebih dinamis dan lebih responsif terhadap kebutuhan klien, serta lebih efisien. seperti juga kurang berisiko dan lebih efisien. Karena alasan itu, mereka berusaha merangkum model *prototyping* dalam satu kerangka kerja yang konsisten.

Model prototyping digunakan untuk menciptakan sistem yang masih versi percontohan (sample), yang disebut prototype. Tujuannya adalah untuk membangun berbagai versi dan secara konsisten memperbaiki versi-versi tersebut hingga produk akhir tercapai, Carr dan Verner (1997). Pembuatan versi – versi sistem tersebur lebih ditekankan kepada respon atau umpan balik pengguna, hal ini sangat mendasar untuk mengembangkan *prototype* berikutnya, hingga produk akhir.

Berikut ini merupakan tahapan proto.

**OPTICAL CHARACTER RECOGNITION**

### Menurut Suryo Hartanto, (2015 : 16). *Optical* *Character* *Recognition* (*OCR*) adalah sebuah aplikasi komputer yang digunakan untuk mengidentifikasi citra huruf maupun angka untuk dikonversi ke dalam bentuk file tulisan. Sistem pengenal huruf ini dapat meningkatkan fleksibilitas atau kemampuan dan kecerdasan sistem komputer.

Secara umum proses *OCR* memiliki beberapa tahap, berikut adalah tahap – tahapnya :

#### File Input

#### Menerima inputan dari user yang merupakan foto yang akan diproses dalam format ( PDF, screenshoot, JPG)

#### Prepoccessing Preprocessing merupakan suatu proses untuk menghilangkan bagian-bagian yang tidak diperlukan pada gambar input untuk proses selanjutnya.

#### Segmentasi Segmentasi adalah proses memisahkan area pengamatan (region) pada tiap karakter yang dideteksi.

#### Normalisasi Normalisasi adalah proses merubah dimensi region tiap karakter dan ketebalan karakter.

#### Ekstrasi Ekstraksi ciri adalah proses untuk mengambil ciri-ciri tertentu dari karakter yang diamati.

#### Recognition Recognition merupakan proses untuk mengenali karakter yang diamati dengan cara membandingkan ciri-ciri karakter yang diperoleh dengan ciri-ciri karakter yang ada pada basis data.

Tahapan – tahapan dalam Optical Character Recognition: 

Gambar 2‑1: Optical Character Recognition Architecture

Sumber : Onyejegbu L. N. Department of Computer Science(N., 2016)

1. *Input Image* : Sistem OCR menerima input berupa gambar
2. *Adaptive* *Thresholding* : meng-skalasi ulang warna dari gambar sebagai sebuah input.
3. *Connected* *component* *Analysis* : Menghubungkan masing – masing komponen yang telah di analisis dari *Adaptive* *Thresholding*.
4. *Find* *Lines* and *Words* : Menentukan masing – masing baris dan menentukan dimana huruf – huruf tersebut harus diletakan.
5. *Recognize* *Word* : mengidentifikasi huruf dan kata yang telah diterjemahkan dari *binary* menjadi huruf yang dapat dikenali manusia.

**TEKNIK MATRIKS MACTHING**

matching adalah salah satu metode yang efektif untuk melakukan pengenalan pada input memiliki kesamaan bentuk. Cara kerja dari template matching adalah mengukur tingkat kesamaan antara input dengan contoh karakter yang tersimpan, kemudian mengambil nilai kesamaan tertinggi sebagai karakter yang benar. Cara ini akan mengambil karakter sesuai dengan bentuk karakter dan menggunakan prinsip penyamaan bentuk. Cara ini dilakukan dengan lebih cepat. Karakter input digambarkan sebagai fungsi f(x,y) dan karakter pencocok digambarkan sebagai fungsi F(x,y). Sedangkan hasil penyamaan sebagai output adalah fungsi T(x,y). Karakter pencocok diambilkan dari contoh karakter pada e-KTP.

Berikut merupakan rumus Macthing Matriks :

Pergeseran template keatas T(x, y)=f(x, y)&&F(x, y+Δ y)

Pergeseran template kebawah : T(x, y)=f(x, y +Δy )&&F(x, y)

Pergeseran template kekiri : T(x, y)=f(x, y)&&F(x+ Δ x, y)

Pergeseran template kekanan : T(x ,y)=f(x+ Δx ,y)&&F(x ,y)

Pergeseran template keatas kiri : T(x, y)=f(x, y)&&F(x+Δ x , y+ Δy)

Pergeseran kebawah kiri : T(x, y)=f(x, y +Δy)&&F(x +Δx, y)

Pergeseran keatas kanan : T(x, y)=f(x+ Δx , y)&&F(x ,y+ Δy)

Pergeseran bawah kanan :T(x, y )=f(x+ Δx , y+Δ y)&&F(x ,y)



Gambar 3‑4 : Matriks Macthing Output

**METODE PENELITIAN ( KUALITATI )**

#### Metode Kualitatif - Penelitian Lapangan (Field Research) pada tahap ini penulis memilih metode Kualitatif, yakni Penelitian Lapangan. Hal ini dikarenakan, Penulis ingin mengamati objek penelitian secara langsung serta pengumpulan data dari objek penelitian. Dalam hal ini, yaitu cover skripsi, halaman abstrak skripsi, dan kinerja BAAK dalam menginput data mahasiswa.

**HASIL PENELITIAN DAN PEMBAHASAN**

Berdasarkan hasil penelitian yang dilakukan oleh peneliti, maka peneliti dapat mengambil kesimpulan sebagai berikut:

### Dengan adanya implementasi *Optical Character Recognition* mahasiswa tidak perlu datang ke-kampus untuk mentranslasi data judul dan abstrak skripsi ke BAAK.

### Dengan menggunakan inputan berupa JPG dan PDF (*1 MegaBytes*) *Prototipe* ini dapat mengurangi waktu yang dibutuhkan dalam menginput judul dan abstrak skripsi.

**REFERENSI**

Ariefin, M. R. (2014). Sistem Real-Time Untuk Manajemen. *Coding Sistem Komputer Universitas Tanjungpura* , 21-22.

Brata, D. W. (2015). Perancangan Sistem KHS Mobile di STMIK ASIA Malang Menggunakan Android Programming dan JSON. *Ilmiah Teknologi dan Informasia ASIA* , 1-10.

Hartanto, S. (2015). Optical Character Recognition Menggunakan Algoritma Template Matching Correlation. *FSM Universitas Diponegoro* , 1-18.

Isaias, P. (2015). *High Level Models and Methodologies for Information Systems.* New York, London: Springer New York Heidelberg Dordrecht.

Jackson, W. (2017). *Android Apps for Absolute Beginners: Covering Android 7.* Apress.

Leonardo. (2018). Cara bikin Skripsi. *Tata Cara* .

Madria, S., & Hara, T. (2016). *Big Data Analytics and Knowledge Discovery: 18th International Conference, DaWaK 2016, Porto, Portugal, September 6-8, 2016, Proceedings.* Rolla, MO: Springer International Publishing.

Mandal, D., Kar, R., Das, S., & Panigrahi , B. K. (2015). *Intelligent Computing and Applications: Proceedings of the International Conference on ICA, 22-24 December 2014.* India: Springer .

Marakas, J. A. (New York). *Management Information Systems, Edisi ke-10.* 2011: The McGraw-Hill Companies, Inc.

Martin, K. (2017). *Everyday Cryptography: Fundamental Principles and Applications (2nd Edition).* United Kingdom: Oxford University Press.

N., O. L. (2016). Optical Character Recognition as a Cloud Service in Azure Architecture . *Department of Computer Science* , 1-7.

Rao, V. (2016). Optical Character RecognitionTechnique Algorithms. *Theoretical and Applied Information Technology* , 1-8.

ResearchGate. (2014). Optical Character Recognition for Indonesian Electronic Id-Card Image. *IC-TECH* , 15-16.

Seidl, M., Scholz, M., Huemer, C., & Gerti, K. (2015). *UML @ Classroom: An Introduction to Object-Oriented Modeling.* Heidelberg: Springer International Publishing.

Sena, S. A. (2013). Perancangan dan pembuatan Application Programming Interface Server untuk Arduino. *Brawijaya* , 1-2.

Siyoto, S. (2015). *Dasar Metodologi Penelitian.* Yogyakarta: Literas Media.

Vainikka, J. (2018). Full-stack web development using Django REST framework and React. *Metropolia* , 1-46.

Yener, M., & Dundar, O. (2016). *Expert Android Studio.* Indianapolis: Wrox.

Zackova, J. R. (2015). *Beyond Artificial Intelligence : The Disappearing Human-Machine Divide Vol 9.* New York, London: Springer Cham Heidelberg New York Dordrecht London.