

Pengembangan API Ekstraksi Data dari Foto KTP = Development of API for Extracting Data from KTP Photo

Irfan Maulana Nasution, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920520914&lokasi=lokal>

Abstrak

Penggunaan Kartu Tanda Penduduk (KTP) sebagai kartu identitas aplikasi sudah umum diimplementasikan, terutama pada sektor teknologi finansial (fintech) yang sudah banyak diadopsi masyarakat. Meskipun begitu, modul ekstraksi data dari KTP yang bersifat terbuka (open source) dan siap guna belum tersedia. Penelitian ini memiliki tujuan membuat modul tersebut, untuk penggunaan penelitian dan/atau membantu bisnis startup dengan memberikan opsi gratis perihal ekstrak data KTP pelanggan mereka. Penelitian ini juga dibuat dengan harapan dapat menggunakan penelitian sebelumnya sebagai pembelajaran dan referensi, dan memperbaiki kekurangan penelitian tersebut.

Modul ini memiliki bentuk akhir docker image yang dapat digunakan langsung dengan utilisasi docker engine, dengan harapan proses kontainerisasi tersebut dapat mempermudah layanan - layanan yang menggunakan container orchestration seperti kubernetes, yang sangat umum digunakan e-commerce, untuk mengadopsi modul ini. Data dari gambar KTP akan melalui tahap preprocessing, pengenalan karakter, pengelompokan data, dan pembersihan data. Hasil eksperimen menggunakan 30 sampel KTP asli menunjukkan bahwa penggunaan masukan yang sudah ter crop lebih baik karena tingkat keberhasilan otomatis cropping hanya 43,3%. Selain itu, dari hasil bacaan pada sampel 75% berhasil terbaca dan dikelompokkan dengan tepat. Dari data yang sudah berhasil dibaca dan dikelompokkan, 17% value dari data mengandung kesalahan. Dari kesalahan yang terjadi, 52.94% kesalahan berhasil dikoreksi oleh algoritma pembersihan data. Secara keseluruhan sistem berhasil membaca dengan tingkat keberhasilan 74,6%The use of Indonesian Citizen's Identity Cards (KTP) as Identifiers in apps has been commonly implemented in apps, particularly those that work in the financial technology (fintech) sector, something that the masses have adopted to use in recent years. Despite this, an open (open source) and free module for Identity Card data extraction isn't available for immediate use. This research aims to make such a module, with the intent to help research and/or small businesses and startups by giving them a solution in the form of a free and ready to use identity card data extraction module. This paper also aims to learn from past papers, and hopefully improve upon them on some aspects.

This module will take the form of a docker image that can immediately be used as a standalone container with the use of docker engine. With this containerization approach, we hope that services using container orchestration such as kubernetes, a very commonly used platform used by e-commerces, will have an easier time adopting this module. Data from Identity Cards will go through several stages, including preprocessing, character recognition, data classification, and data cleaning. Experiments using 30 real life Identity Card samples resulted in cropped input being better since cropping automation only resulted in 43,3% success rate. The experiment also found that the reading and categorizing success rate are 75%. Out of all the categorized data, 17% of the values contained inaccuracies, 52.94% of which were successfully corrected by

the cleaning algorithm. Overall, the system successfully extracted 74,6% of the data.