

ABSTRACT

IMPLEMENTATION OF BOYER-MOORE ALGORITIHM FOR APPLICATION A DICTIONARY LANGUANGE SUNDA- INDONESIA BY USING TESSERACT TO FEATURE OPTICAL CHARACTER RECOGNITION (OCR)

By:

Risma Nurfalah

1127050132

Mobile applications have been used by people of both adults and children, because it helps to jobs and education. one of the media in education of language is dictionary that can be used to find some words that are not known. The purpose of making this Sundanese-Indonesian dictionary application is to create a translator application based on mobile platform from picture or image using algorithm to search words. An algorithm is a sequence of steps which are explained in a clear and unambiguous to solve some of the problems within a specified time. It means, every step should be done and have some effect. One of algorithms that can be used to search processing is Boyer-Moore algorithm. Booyer-Moore algorithm uses String matching started from right position to the last left position. Boyer-Moore algorithm apply the good sufix and bad character principles. With these two principles, String information obtained in search processing will be coming more and more, so the output will come out better. Tesseract for Optical Character Recognition (OCR) feature is used because it can take pictures into text that taken by camera to help users to translating language.

Keywords: *Android device, Boyer-Moore algorithm, Dictionary, Tesseract Optical CharacterRecognition (OCR).* **ABSTRAK**

**IMPLEMENTASI ALGORITMA *BOYER-MOORE*
PADA APLIKASI KAMUS BAHASA SUNDA-INDONESIA**

DENGAN MENGGUNAKAN *TESSERACT* UNTUK FITUR *OPTICAL CHARACTER RECOGNITION (OCR)*

Oleh:

Risma Nurfalih

1127050132

Aplikasi berbasis mobile telah marak digunakan di lingkungan masyarakat baik kalangan anak maupun dewasa karena telah banyak membantu dalam masalah pekerjaan maupun dalam pembelajaran. Salah satu media dalam pembelajaran bahasa adalah kamus yang dapat digunakan untuk mencari arti kata-kata yang tidak diketahui. Tujuan dari pembuatan aplikasi kamus bahasa sunda-indonesia ini adalah untuk membuat sebuah aplikasi penerjemah berbasis *mobile* dari citra atau gambar dengan melibatkan algoritma dalam proses pencarian kata. Algoritma merupakan urutan langkah-langkah yang dinyatakan dengan jelas dan tidak rancu untuk memecahkan suatu masalah dalam rentang waktu tertentu. Artinya, setiap langkah harus dapat dikerjakan dan mempunyai efek tertentu. Salah satu algoritma yang digunakan dalam proses pencarian adalah algoritma *Boyer-Moore*. Algoritma *Boyer-Moore* menggunakan metode pencocokan *string* dimulai dari posisi kanan hingga akhirnya sampai pada posisi paling kiri. Algoritma *Boyer-Moore* menerapkan prinsip *good suffix* serta prinsip *bad character*. Dengan kedua prinsip ini, informasi *string* yang diperoleh dalam pencarian akan semakin banyak sehingga *output* yang dihasilkan akan menjadi lebih baik. Fitur *Optical Character Recognition (OCR)* dengan bantuan *tesseract library* digunakan agar *image* terkonversi dengan baik kedalam bentuk teks yang ditangkap oleh kamera *handphone* untuk memudahkan pengguna dalam menterjemahkan bahasa.

Kata Kunci: Android, Algoritma *Boyer-Moore*, kamus, *Tesseract Optical Character Recognition (OCR)*.