

OPTIMASI SISTEM PENGENALAN KARAKTER PADA *OPTICAL CHARACTER RECOGNITION* (OCR)

Nama : Siti Maimunah
NIM : 140431100044
Program Studi : Teknik Elektro
Dosen Pembimbing : Koko Joni, S.T., M. Eng.
Achmad Fiqhi Ibadillah, S.T.,
M.Sc.

ABSTRACT

Optical Character Recognition (OCR) is a branch of the Artificial Intelligence concept from the field of image processing, where character recognition is implemented in a software that recognizes many input characters in an object. This introduction will greatly assist the user in scanning non-digital text which is then processed by a system to be made into a digital text. From digital text, the user will easily process it as they wish. Therefore OCR is very appropriate to be used in this problem. OCR will convert from text images into editable text, add, even copy-paste. The technology that will be used is EmguCV library as image processing library and Tesseract as OCR library which will be implemented in character recognition with various types of character input. From these parameters the best result is that there is a white and yellow background with black writing color obtaining an average yield of approximately 80% with a distance of 10 cm to 14 cm.

Keywords: *Optical Character Recognition (OCR), Tesseract, EmguCV*

OPTIMASI SISTEM PENGENALAN KARAKTER PADA *OPTICAL CHARACTER RECOGNITION* (OCR)

Nama : Siti Maimunah
NIM : 140431100044
Program Studi : Teknik Elektro
Dosen Pembimbing : Koko Joni, S.T., M. Eng.
Achmad Fiqhi Ibadillah, S.T.,
M.Sc.

ABSTRAK

Optical Character Recognition (OCR) merupakan cabang dari konsep *Artificial Intelligence* dari bidang pengolahan citra, dimana pengenalan karakter diimplementasikan dalam sebuah perangkat lunak yang mengenali banyak inputan karakter pada suatu objek. Pengenalan ini akan sangat membantu *user* dalam memindai teks non digital yang kemudian diolah oleh sebuah sistem untuk dijadikan kedalam sebuah teks digital. Dari teks digital maka *user* akan mudah mengolahnya sesuai keinginannya. Oleh karena itu OCR sangat tepat digunakan dalam permasalahan ini. OCR sendiri akan mengkonversi dari gambar teks menjadi teks yang dapat diedit, di tambahkan bahkan di copy-paste. Teknologi yang akan digunakan yaitu *library* EmguCV sebagai *library* pengolahan citra dan *Tesseract* sebagai *library* OCR yang akan di implementasikan pada pengenalan karakter dengan berbagai jenis inputan karakter. Dari beberapa parameter tersebut hasil terbaik yaitu terdapat pada latar putih dan kuning dengan warna tulisan hitam memperoleh hasil rata-rata kurang lebih 80% dengan jarak 10 cm hingga 14 cm.

Kata kunci: *Optical Character Recogniioan* (OCR), *Tesseract*, EmguCV