

Informatics Institute of Technology

In Collaboration, with

University of Westminster, UK

Handwritten Data Digitizer

A dissertation by

Mr. D. P. A. De Alwis

Supervised By

Mr. Pumudu Fernando

Submitted in partial fulfilment of the requirements for the

BSc (Hons) Software Engineering degree

Department of Computing

May 2017

© The copyright for this project and all its associated products resides
with Informatics Institute of Technology

Abstract

Numerous associations still rely on paper-escalated work processes. Due to the fact that the printed and handwritten documents are all around acknowledged and perceived for any authoritative report, paper documents still have a major impact on many organizations. Most of the organizations use template based documents, that have rectangular shaped areas for user inputs to gather data from users. Many organizations already use or prefer to use digital storage since it's easy to manage and easy to access data. The major issues having handled the paper documents are the inability to monitor the lost data, storage and money and time wasted on re-keying data. It is possible to address these problems through a solution that can digitize the data in these paper documents. Currently, there is no fully automated handwritten character recognition system that can identify handwritten characters on a template-based form. The most common approach is to identify handwritten of a single person through a template matching approach. In the proposed solution, the template of the document is identified and handwritten areas are extracted through an image processing component and the identification of the handwritten characters are addressed through training the system using a convolutional neural network.

Subject Descriptors:

• **Theory of computation~Structured prediction** • **Computing methodologies~Image processing** • **Applied computing~Optical character recognition**

Keywords:

Image Classification, Handwritten Character Recognition, Image Processing, Machine Learning