

Sinhala Handwritten Postal Address Recognition for Postal Sorting

M.F.A. Ifhaam
Department of Industrial Management,
Faculty of Science
University of Kelaniya, Sri Lanka
ahmedifhaam1@gmail.com

S. Jayalal
Department of Industrial Management,
Faculty of Science
University of Kelaniya, Sri Lanka
shantha@kln.ac.lk

Abstract

Sri Lankan post office mail sorting process is done manually, even today. Though employees are well experienced, it takes considerable time and employees need to work overtime in places like Central Mail Exchange (CME). With major issues like unclear handwriting, having trouble to recognize some uncommon or ambiguous names, and carrying these duties twice a day create a negative impact on the efficiency of the postal delivery system. In the prevailing system, forward mails and delivery mails are the two categories of separating mails at the sorting centers. Delivery mails are the posts which can be delivered to its destination directly. Forward mails are the ones which need to be sent to an appropriate post office that can deliver the particular post to its destination. Majority of Sri Lankans use Sinhala language for their day to day activities. The primary objective of the research is to identify the automatic way of forwarding the letter to the next post office from the current post office. Proposed system is focused on the recognition of Sinhala handwriting using Optical Character Recognition (OCR) and image processing technologies. Data collected under different criteria were used for training and testing the solution. Genetic Algorithm (GA) was used to generate more optimized results faster with higher accuracy. Given addresses are written in the default format. This format can be extended to more formats as improvements in future. The algorithm shows accuracy over 92% for addresses which are recognized with 3 misrecognized characters. This algorithm can be used on practice scenario as the AI Recognition has more than 79 % of accuracy.

Keywords: *Image Processing, Machine Learning, Postal Address Sorting, Sinhala OCR, GA*