

TEXT DETECTION IN SCENE IMAGES AND POSTERS

V.A.C. Premachandra^{1*} and M.A.C. Akmal Jahan¹

*¹Department of Computer Science, Faculty of Applied Sciences,
South Eastern University of Sri Lanka, Sammanthurai, Sri Lanka.*

**chathuz196@gmail.com*

Scene images, name boards, road names and posters enriched with Sinhala are crucial issue for locals whose mother tongue is another language or foreigners who visit Sri Lanka often. Since tourists cannot understand the local language, there are only guides to help them find places and understand the things around them. To overcome the issue, this research aims to introduce Sinhala text detection and character recognition in scene images and posters. This helps the users to clarify the process of understanding Sinhala text content in scene images and posters. It facilitates the understanding process while translating the content into their native languages. However, the complexity of character detection and recognition in scene images is extremely challenging compared to classical character detection and recognition on plain background images. This has been overcome by introducing projection profiles. In this work, vertical and horizontal projection profiles were exploited to extract the text lines and words, while the OCR technique was used to recognize the characters. The extracted sentences with a set of words were then translated into the user's native language. A real-time dataset was acquired from recent protests, road labels, cutouts, name boards and sign boards for the experiment. The experimental results of each phase show that there is potential to use the system for Sinhala language character recognition in scenes and posters that exist in public places.

Keywords: *character recognition, character detection, scene image*