ICT AND DATA SCIENCE, IOT AND EMERGING TECHNOLOGIES(ICT)

ST107-ICT-001

Enhanced OpenCV for Text Detection Using Multi-Scale Attention Mechanism

Gedion Njoroge*, Gabriel Kamau and Anthony Maina

School of Computer Science, Dedan Kimathi University of Technology, Private Bag, Nyeri, 10134, Kenya, *Email: gedion.njoroge22@students.dkut.ac.ke,

Abstract

This research introduces an innovative strategy by merging a multi-scale attention mechanism with the OpenCV framework to enhance text detection. OpenCV, a foundational computer vision library, excels in image preprocessing and feature extraction [1]. Despite emerging deep learning frameworks, OpenCV's prowess in addressing complex text detection scenarios remains limited. To address this, a multi-scale attention mechanism is proposed, enabling the model to decode text features across diverse scales and contexts [2]. This approach improves text detection and recognition, particularly in complex scenes, demonstrated through comprehensive experiments on benchmark datasets [3]. Results highlight its superiority over conventional OpenCV methods, enhancing text-related tasks and bolstering real-time applications [4]. This integration advances text detection by combining OpenCV's processing abilities with a multi-scale attention mechanism, aligning with OCR frameworks such as Tesseract OCR for recognition [5]. The method's potential is underscored in a text-focused technological landscape.

Keywords: text detection, OpenCV, multi-scale attention, deep learning, Tesseract OCR.

ST107-ICT-002