ISSN 2362–0412

ABS-073

High-Volume Automated MCQ Marking: An Image Processing Approach for Enhanced Efficiency and Accuracy

S.H.K.K. Gunawickrama[1], U.G.R.S. Bandara[1], D.M.S.C. Dissanayaka[1], D.M.M.I.T. Dissanayake[1], H.L.D.U.G. Perera[1]*

[1] University of Ruhuna

**Corresponding author:* keerthi@eie.ruh.ac.lk

This study proposes the development of an AI-powered optical mark recognition (OMR) system that aims to automate and enhance the grading of multiple-choice question (MCQ) answer sheets. Traditional OMR systems rely on shaded bubbles or pencil marks, which can cause recognition errors, especially when students use different marking styles, such as crosses instead of shading. The manual marking process can take a considerable amount of time for large numbers of students. To overcome these limitations, the proposed system combines advanced image processing techniques with a Convolutional Neural Network (CNN) to accurately classify bubbles as crossed, shaded, or empty. A mathematically calculated and defined A5 bubble-sheet with a 50-question layout was designed to support two paper structures, enabling shuffled question papers for enhanced exam security. These structures make the system more practical, user-friendly, and suitable for large-scale academic and institutional assessments.

The system was implemented using Python with PySide6 for the graphical user interface. In addition, OpenCV, TensorFlow, NumPy, and Pandas were used for processing and model training. A custom dataset was developed using real student-marked papers representing different marking behaviors and image conditions. The final CNN model achieved a validation accuracy of 98.87%. The model demonstrated strong generalization and consistent performance across different lighting and image quality conditions. Furthermore, the Gemini Pro API was incorporated to automatically extract student registration numbers from scanned answer sheets, ensuring accurate mapping between responses and student identities.

The desktop application provides features for automated scoring, manual verification by examiners, and exporting results in Excel format. It also supports evaluating shuffled MCQ papers, enabling examiners to mark answer sheets that belong to different paper versions with different question orders. Additionally, the system includes registration number detection. The system substantially reduces grading time while maintaining human-verified reliability through dual examiner review features. The solution provides a practical, scalable, and cost-efficient alternative to commercial OMR systems for large-scale assessments in universities.

Keywords: MCQ marking, Image processing, Enhancing efficiency, Accuracy