A Retrieval Augmented Visual Inspection Framework for Counterfeit Medicine Detection Using Image Matching and Explainable AI

Name: Yona Zakaria
Laboratory's name: Symbiotic Systems & Ubiquitous Computing Systems Laboratory
Supervisor's name: Professor Keiichi Yasumoto

Abstract

Counterfeit and substandard medicines remain a serious public health risk, especially in low and middle-income countries where laboratory testing is limited. Visual inspection is widely recommended as the first step for identifying suspicious medicines. However, its effectiveness depends heavily on the user's knowledge of authentic packaging, and many subtle design differences are difficult to notice. These limitations weaken the reliability of visual inspection in real-world medical purchasing situations.

This dissertation addresses these challenges by developing a packaging visual inspection framework that integrates image matching, explainable artificial intelligence, and retrieval augmentation to support more dependable and understandable identification of suspicious medicines. Three research tasks contribute to the framework. Task 1 focuses on fast and accurate retrieval of trusted reference images. Our proposed framework introduces an accelerated outlier-rejection method that significantly improves the speed of keypoint-based image matching across varying image capture conditions. In parallel, our framework incorporates an image retrieval augmentation approach designed to gather reliable up-to-date reference images and packaging information from the web, official announcements and regulatory alerts. Task 2 addresses the automated detection, visualization, and interpretation of visual discrepancies between the user's photo and the retrieved reference image. This involves aligning the images, identifying subtle but meaningful differences, generating clear discrepancy maps that guide user attention, and using a multimodal large language model to explain these differences in simple and understandable terms.

User evaluations on research task 2 showed a **23.3%** improvement in overall counterfeit detection accuracy, with natural-language explanations contributing an additional **18.1%** increase in decision accuracy and reducing decision time by **30.5%** (from 89.3 s to 62.1 s). These results show that the framework improves the reliability and accuracy of visual inspection and supports faster, more consistent judgments in situations where conventional methods struggle.

The broader significance of this work lies in making expert-level visual reasoning more accessible to ordinary medicine buyers and contributing to safer medicine use and better post-market surveillance.