ABSTRACT

The agricultural sector is undergoing a profound transition as it seeks to meet the food demands of a growing global population while confronting labor shortages and sustainability challenges. Robotic systems offer a compelling approach to automating labor-intensive operations, such as harvesting, crop monitoring, and disease management. However, despite notable advancements, current agricultural robots remain constrained by limited mobility on uneven terrain, reliance on expensive sensing infrastructure, limited single-task functionality, and perception systems that fail to perform consistently across different crops and field conditions. This research addresses these limitations by developing a vision-guided tracked mobile manipulator that integrates deep learning-based perception with an optimized mechanical platform to enable autonomous navigation, precision harvesting, targeted disease treatment, and crop health assessment within a unified robotic framework. The mechanical system comprises a tracked base for enhanced terrain stability and a 5-Degree-of-Freedom (DOF) serial manipulator with a force-feedback gripper, designed through kinematic modeling based on the Denavit–Hartenberg (D-H) convention and systematic workspace analysis. Perception capabilities are developed using Convolutional Neural Network (CNN)–based Visual Geometry Group 16 (VGG16), Vision Transformers (ViTs), and You Only Look Once (YOLO)-based architectures for fruit detection, maturity classification, soil moisture estimation, and leaf disease identification, while a novel hybrid VGG16-ViT architecture addresses the generalization challenge by combining local texture analysis with global contextual information for plant leaf disease classification. Autonomous navigation is achieved through a vision-based control strategy that integrates Red-Green-Blue-Depth (RGB-D) sensing, pulse-based localization, and reactive path planning, enabling navigation in unstructured environments without reliance on an expensive Global Positioning System (GPS). The precision harvesting pipeline combines a deep learning–based Modified-YOLOv8 (M-YOLOv8) fruit localization algorithm with inverse kinematics–driven manipulation for fruit detection, trajectory planning, force-controlled grasping, and gentle deposition. In addition, a targeted spraying system employs two-stage perception for leaf detection and disease classification, with three-dimensional (3D) localization, to deliver precise pesticide application to infected foliage. Experimental validation through laboratory and field trials demonstrates the effectiveness of the proposed system across all functional domains. This research contributes a coherent integration of mechanical design, deep learning-based perception, and autonomous control, establishing a foundation for intelligent, multifunctional agricultural robots capable of advancing sustainable precision agriculture.

Keywords: Agricultural robotics; Deep learning; Autonomous navigation; Fruit harvesting; Targeted spraying; Mobile manipulator; Precision agriculture