LitchiInst: Instance Segmentation of the Main Fruit-Bearing Branch via Fruit-Branch Association for Robotic Litchi Harvesting

Yukun Qian[a], Wenchang Chai[b], Haitao Wang[a,*], Zhiyang Mai[a], Liangliang Zhou[a], Hejun Wu[a]

a, Department of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China

b, Department of Computing, Hong Kong Polytechnic University, Hong Kong, China

Abstract

Accurately segmenting structurally similar components in natural scenes remains a key challenge for instance segmentation models. In automated litchi harvesting, the visual similarity between main fruit-bearing branches (MFBBs) and ordinary branches complicates MFBB identification, which is critical for enabling robotic systems to harvest litchi in clusters. To address this issue, we propose LitchiInst, a query-based instance segmentation algorithm explicitly designed for accurate MFBB identification. Built upon the FastInst framework, LitchiInst introduces association-instance activation-guided (AIA-guided) queries and an association loss function. By deeply exploiting the connection between MFBBs and litchi fruits, it highlights the distinctive features of true MFBBs, enabling precise differentiation from other branches. To further enhance segmentation quality, we design an enhanced feature pyramid network (EFPN) as the pixel decoder to better preserve fine-grained details during feature fusion. Experimental results demonstrate that LitchiInst improves MFBB detection accuracy while maintaining real-time performance, achieving a 31.06% improvement in MFBB AP over the FastInst baseline. The detection capability of LitchiInst is demonstrated through real robot experiments, and its practical feasibility is further highlighted by deployment in an orchard environment.

Keywords: Litchi harvesting, Main fruit-bearing branch (MFBB), Query-based instance segmentation

1. Introduction

Accurately segmenting structurally similar components in natural scenes remains a fundamental challenge in instance segmentation. This issue is particularly pronounced in agricultural environments, where visually similar plant structures often coexist. For example, in litchi harvesting scenarios, the main fruit-bearing branch (MFBB) and ordinary branches exhibit highly similar appearances, making it difficult to distinguish between them. Such visual similarity severely hinders the accurate identification of MFBBs, a crucial step for enabling intelligent harvesting.

Unlike solitary fruits such as apples or mangoes, litchi grows in dense clusters that require harvesting as a whole. The MFBB, which supports these clusters, is crucial for guiding robotic systems in cluster-level picking. However, its high visual similarity to ordinary branches presents a major challenge for accurate identification.

This challenge reflects a broader limitation of instance segmentation models, which often struggle to distinguish structurally similar targets based solely on local appearance features. This is largely because most existing approaches optimize object masks independently, relying only on the visual characteristics of each instance. These limitations become particularly critical when the target objects, such as MFBBs, closely resemble surrounding structures in appearance. Importantly, the key differentiator of an MFBB lies in its strong semantic and physical association with multiple litchi fruits. Exploiting this relationship provides a promising direction for improving MFBB segmentation in complex agricultural environments.

In this work, we aim to accurately identify and localize litchi MFBBs within the workspace of robotic systems, as illustrated in Figure 1. We propose LitchiInst, an instance segmentation algorithm adapted from FastInst(He et al., 2023), specifically designed for

* Corresponding author. Email: wanght39@mail2.sysu.edu.cn (H. Wang)