

Fine-Grained Grocery Product Recognition by One-Shot Learning

Weidong Geng, Feilin Han, Jiangke Lin, Liuyi Zhu, Jieming Bai,
Suzhen Wang, lin He, Qiang Xiao, Zhangjiong Lai

State Key Lab of CAD&CG, College of Computer Science, Zhejiang University

{gengwd,hanfeilin,jiangke_lin,jiemingbai,suzhenwang,helin9,x-qiang,lzjiong}@zju.edu.cn;zhu liuyi@hotmail.com

ABSTRACT

Fine-grained grocery product recognition via camera is a challenging task to identify the visually similar products with subtle differences by using single-shot training examples. To address this issue, we present a novel hybrid classification approach that combines feature-based matching and one-shot deep learning with a coarse-to-fine strategy. The candidate regions of product instances are first detected and coarsely labeled by recurring features in product images without any training. Then, attention maps are generated to guide the classifier to focus on fine discriminative details by magnifying the influences of the features in the candidate regions of interest (ROI) and suppressing the interferences of the features outside, improving the accuracy of fine-grained grocery products recognition effectively. Our framework also performs a good adaptability which allows existing classifier to be refined without retraining for new coming product classes. As an additional contribution, we collect a new grocery product database with 102 classes from 2 stores. Extensive experiments demonstrate that our approach outperforms the state-of-the-art methods.

CCS CONCEPTS

• **Computing methodologies** → **Object recognition**; *Visual content-based indexing and retrieval*; • **Applied computing** → E-commerce infrastructure;

KEYWORDS

Product Categorization; Fine-Grained Object Recognition

ACM Reference Format:

Weidong Geng, Feilin Han, Jiangke Lin, Liuyi Zhu, Jieming Bai, Suzhen Wang, lin He, Qiang Xiao, Zhangjiong Lai. 2018. Fine-Grained Grocery Product Recognition by One-Shot Learning. In *2018 ACM Multimedia Conference (MM '18)*, October 22–26, 2018, Seoul, Republic of Korea. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3240508.3240522>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MM '18, October 22–26, 2018, Seoul, Republic of Korea

© 2018 Association for Computing Machinery.

ACM ISBN 978-1-4503-5665-7/18/10...\$15.00

<https://doi.org/10.1145/3240508.3240522>

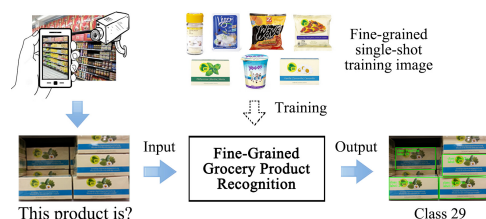


Figure 1: Illustration of fine-grained grocery product recognition, which can detect and recognize grocery products captured via mobile devices by using one-shot learning.

1 INTRODUCTION

The rapid growth of mobile devices equipped with high quality cameras enables an associated increase of visual recognition applications in daily life. The detection and recognition of grocery products by smart mobile devices has recently gained increasing interest in intellectualized store scenario, shown as Figure 1. The applications include automatic shopping, self-assisted shopping, product tracking, and real-time management of inventory [6, 16, 39]. Amazon Go is a typical application of automatic shopping, which is an unmanned supermarket and is regarded as a retail prototype of futuristic grocery stores [3]. Under intellectualized store scenario, there usually is only one product photo for each product instance in the SKU product database. Therefore, grocery product recognition issue can be regarded as one-shot learning due to the limited training data.

Fine-grained object recognition refers to recognizing the subordinate category of an observed object. Most of the existing works aim to build a generalized framework for fine-grained object recognition tasks, such as identifying animals [17, 24]) and plant species [4, 12]. As a special case of fine-grained object recognition, fine-grained grocery product recognition [18, 23] is to recognize subordinate categories of retail products, which is a challenging task due to intra-class variance and inter-class similarity [23, 43].

Grocery products are hard to be recognized since many products are visually similar in terms of shape, color, texture, and metric size [6], e.g., same products with different flavors may only have minor differences of packaging details. Another peculiarity of grocery product recognition in retail scenarios is that training data is not acquired in real environments but in ideal conditions [16]. The images of product instances in real environments are quite different from products, due to the changes in scale, pose, viewpoint, lighting conditions and occlusion [18]. Thus, the classification framework for grocery products should be able to differentiate subtle details in appearance of products.