GROCERY ITEM DETECTION USING ANDROID STUDIO

[1]Ms.J. Rohini AP/CSE,[2]M.Thanush,[3]K.Uthiramoorthi,[4]M.Vignesh

Department of Computer Science And Engineering
Erode Sengunthar Engineering CollegeErode,Tamilnadu,India.

Abstract

In grocery stores, product supply management is critical to employee productivity. Finding the right time to update an item in terms of design/replenishment requires real-time data on item availability. As a result, goods are always available on the shelf when customers need them. This study focuses on product display management in grocery stores to determine the specific products and their quantities on the shelves. Using deep learning (DL) to determine and identify each item, the warehouse manager compares each identified item to the previously created configured item plan. We used You Only Look Once version 5 (YOLOV5) for product detection and used both shape and size features and color features for product detection to reduce false product detection. Experimental results were performed using the dataset. Analysis shows that the proposed approach improved accuracy, precision and recall. Product recognition can shorten defect dates by including color features. It is useful to distinguish identical logos with different colors. You can achieve a functional level 75, reputation level 81 accuracy percentage. We've created a new solution that can reduce checkout and billing times by 50%. What if all the products a customer purchased together were scanned using this algorithm in less than a minute? This algorithm has excellent accuracy, precision and recognition and is considered the best algorithm used in item recognition.

I. INTRODUCTION

The main goal of developing a grocery store mobile application is to help the consumer by reducing their effort. The Perception Problem as a Classification Problem The purpose of product perception is to facilitate the management of retail products and improve the consumer's shopping experience.The most used technology nowadays for automatic product identification in industry and research alike is barcode recognition. By reading the barcode printed on each product package, product management becomes easier. Almost every product on the market usually has a corresponding barcode. However, due to the uncertainty of where the barcode will be printed, manually locating the barcode and allowing the machine to identify it at checkout is often time consuming. According to a survey, 45% of his customers complained that using barcode scanners was inconvenient.

Many grocery stores employ human labor to manipulate items on shelves, shelves, and counters. Employees manually check product availability, calculate balances, and compare locations to specifications. The whole process is costly and error prone. An important part of presentation is that products that are not displayed correctly can lead to lost sales. To increase sales of their products, many manufacturers offer attractive displays in their stores. Every merchandiser wants to display their products in a central place/place where all customers pay attention.

Customers are increasingly warned. Keeping up with the latest trends and approaches is critical to delivering a satisfying experience to your customers and ensuring they meet their rising expectations. On store shelves, customers make a serious purchase decision. Shelf identification using resumes is essential for digitizing store orders and allowing AI to capture critical consumer data. It uses deep neural networks to recognize objects in images of shelves and classify them by category, brand, and item. Automation with vision reduces the effort, errors, and risks associated with human-assisted configuration. Merchandisers must adapt these techniques to deliver a rich customer experience

.

II. WORKING PRINCIPLE

We all experience scenes like taking time to locate anticipated products and waiting in line at the register in a store on a regular basis. The development of automatic product recognition is crucial for the advancement of society and the economy since it saves time and is more dependable than manual labour.

The discipline of computer vision has a difficult issue when trying to recognise products from photographs. Due to the tremendous application potential, including automatic checkout, stock tracking, planogram compliance, and support for the visually impaired, it is receiving more and more attention. Deep learning has made enormous strides recently, with successes in object detection and image classification. The most significant and difficult issues in computer vision are those involving object detection and recognition. The development of object identification and recognition has substantially accelerated in recent years thanks to the impressive improvements in deep learning techniques

(1) Several researchers are interested in it because of the variety of uses it provides. This project focuses on identifying label text to identify several retail products piled on or beneath a grocery shop shelf

(2) . One task that will have a significant impact on society as a whole is automating the product checkout procedure in conventional retail establishments. In order to achieve this goal, trustworthy deep learning models that provide automatic product counting to facilitate quick consumer checkout can be useful.

(3) In this study, we offer a novel region-based deep learning method for automating product counting using a deep SORT algorithm and a tailored YOLOv5 item detection pipeline.Our findings on difficult real-world test films show that our approach generalises predictions with a high enough level of precision and quick enough execution times to support its usage in a real-world commercial setting.

(4) In track 4 of the AI City Challenge 2022, our suggested methodology took fourth place with an F1 score of 0.4400 in experimental validation data

An important task for computer vision-based assistive technologies is to enable visually impaired people to perceive objects in confined environments, such as identifying food items at the grocery store. In this article, we present a new dataset containing natural images of food (fruits, vegetables, and packaged products). These were all captured to resemble his scenario of shopping at the grocery store

A programme that can identify and locate food items in an image captured during a real-time scene. It can be applied to connected or standalone application frameworks. Using data sets gathered from several online sources, we trained a Single Shot Detector (SSD) setup. The InceptionV2 convolutional neural network design combined with single-shot detectors was found to be the most effective method for combining object identification models with different convolutional network architecture

In grocery stores, product supply management is critical to employee productivity. Finding the right time to update an item in terms of design/replenishment requires real-time data on item availability. As a result, goods are always available on the shelf when customers need them

The logo and wording on the merchandise are used in this content to identify the brand of the item. As an illustration, if milk packets are placed on the shelves, the product identification process will identify the boxes by placing a bounding box over each of the items. In this research, we concentrate on locating and identifying supermarket items on shelves near the user in a grocery shop

One staff may oversee many self-checkout counters, which lowers labour expenses. The customer must scan the barcode of the item, put it in the packing area, and occasionally even weigh it. shorter waits than when using checkout lane

III. MODULES

Training data is the data used to train an algorithm or machine learning model to predict the outcome of designing the model. When using supervised learning or hybrids involving this approach, the data is enriched with data labeling or annotation. The goal is to create a trained (fitted) model that generalizes well to new, unknown data. The fitted model is evaluated using "new" examples from the available datasets (validation and test datasets) to assess the model's accuracy in classifying new data.

A machine learning technology called active learning finds data that staff should label. This function is known as automatic data labelling in Ground Truth. When opposed to using data sets entirely for humans, automated data labelling can help reduce labelling costs and time.

Image classification is a supervised learning problem.Define a set of target classes (objects to be identified in images) and train a model to recognize them from labeled sample photos. Early computer vision models relied on raw pixel data as input to the model.

A computer vision technique called object recognition is used to locate occurrences of objects in pictures or movies. To generate useful results, object detection algorithms frequently use machine learning or deep learning. To be more precise, object detection creates a bounding box around these recognised things so that you can locate them in a certain scene (or how they move).

Android studio is the official Integrated Development Environment (IDE) for android application development. Android Studio provides more features that enhance our productivity while building Android apps.IDE for Android app development. It started its early access preview from version 0.1 in May 2013. The first stable built version was released in December 2014, starts from version 1.0

XML stands for Extensible Markup Language. XML is a markup language much like HTML used to describe data. It is derived from Standard Generalized Markup Language(SGML). Basically, the XML tags are not predefined in XML. We need to implement and define the tags in XML. XML tags define the data and used to store and organize data. It's easily scalable and simple to develop. In Android, the XML is used to implement UI-related data, and it's a lightweight markup language that doesn't make layout heavy. XML only contains tags, while implementing they need to be just invoked.

The Firebase Realtime Database is a cloud-hosted database. Data is stored as JSON and synchronized in realtime to every connected client. When you build cross-platform apps with our Apple platforms, Android, and JavaScript SDKs, all of your clients share one Realtime Database instance and automatically receive updates with the newest data.Firebase Hosting is production-grade web content hosting for developers. With a single command, you can quickly deploy web apps and serve both static and dynamic content to a global CDN (content delivery network).

Firebase Hosting works out-of-the-box with Firebase services, including Cloud Functions, Authentication, Realtime Database, Cloud Firestore, and Cloud Messaging. You can build powerful microservices and web apps using these complementary Firebase services.Firebase is a Backend-as-a-Service — BaaS — that started grew up into a next-generation app-development platform on Google Cloud Platform. Firebase frees developers to focus crafting fantastic user experiences. You don't need to manage servers.

You don't need to write APIs. Firebase is your server, your API and your datastore, all written so generically that you can modify it to suit most needs. Yeah, you'll occasionally need to use other bits of the Google Cloud for your advanced applications. Firebase can't be everything to everybody. But it gets pretty close. Instead of typical HTTP requests, the Firebase Realtime Database uses data synchronization-every time data changes, any connected device receives that update within milliseconds. Provide collaborative and immersive experiences without thinking about networking code.

In order to call a Google Cloud API from your app, you need to create an intermediate REST API that handles authorization and protects secret values such as API keys. You then need to write code in your mobile app to authenticate to and communicate with this intermediate service.One way to create this REST API is by using Firebase Authentication and Functions, which gives you a managed, serverless gateway to Google Cloud APIs that handles authentication and can be called from your mobile app with pre-built SDKs.This guide demonstrates how to use this technique to call the Cloud Vision API from your app. This method will allow all authenticated users to access Cloud Vision billed services through your Cloud project, so consider whether this mechanism is sufficient for your use case before proceeding.

YOLOv5 is an open-source implementation of the latest version of YOLO (for a quick test of loading YOLOv5 from PyTorch hub for inference, see here). This Object Detection with YOLOv5 Android sample app uses the PyTorch scripted YOLOv5 model to detect objects of the 80 classes trained with the model.YOLOv5 a family of compound-scaled object detection models trained on the COCO dataset, and includes simple functionality for Test Time Augmentation (TTA), model ensembling, hyperparameter evolution, and export to ONNX, CoreML and TFLite.From the YOLOv5 family, fine tuning YOLOv5m at 640 resolution will yield good results. It is capable of running at more than 80 FPS, even on an older GPU like the TESLA P100. All the while still giving an mAP of 45.4. YOLOv6m is also a pretty good model with 49.5 mAP and almost 50 FPS on the TESLAHigh-level architecture for single-stage object detectors

There is no difference between the five models in terms of operations used except for the number of layers and parameters as shown in the table below.All the YOLOv5 models are composed of the same 3 components: CSP-Darknet53 as a backbone, SPP and PANet in the model neck and the head used in YOLOv4.YOLOv5 uses CSP-Darknet53 as its backbone. CSP-Darknet53 is just the convolutional network Darknet53 used as the backbone for YOLOv3 to which the authors applied the Cross Stage Partial (CSP) network strategy.YOLO is a deep network, it uses residual and dense blocks in order to enable the flow of information to the deepest layers and to overcome the vanishing gradient problem. However one of the perks of using dense and residual blocks is the problem of redundant gradients. CSPNet helps tackling this problem by truncating the gradient flow.*CSP* network preserves the advantage of DenseNet's feature reuse characteristics and helps reducing the excessive amount of redundant gradient information by truncating the gradient flow.YOLOv5 employs CSPNet strategy to partition the feature map of the base layer into two parts and then merges them through a cross-stage hierarchy as shown in the figure bellowsource: YOLOv5 github repo Applying this strategy comes with big advantages to YOLOv5, since it helps reducing the number of parameters and helps reducing an important amount of computation (less FLOPS) which lead to increasing the inference speed that is crucial parameter in real-time object detection models.

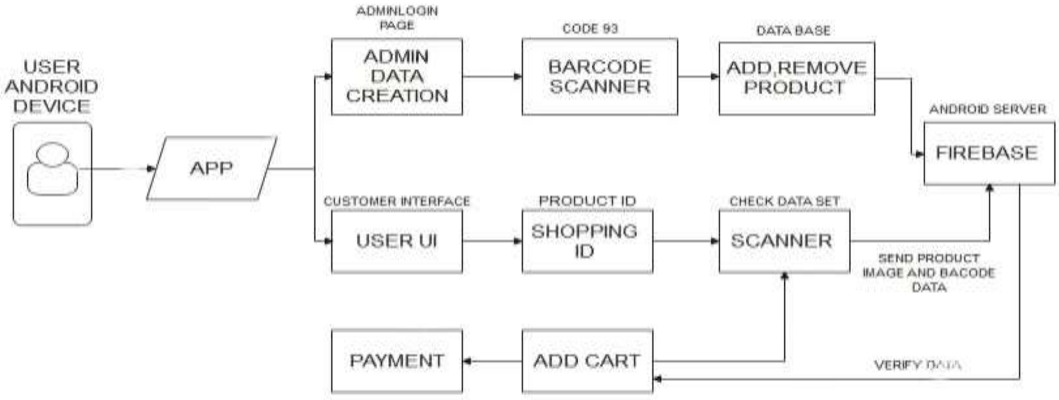
IV. METHODOLOGY

Deep learning is a subset of machine learning that has become widely popular for object detection and automatic target recognition (ATR). Convolutional neural networks (CNNs) have proven to be successful for target classification tasks without the need for feature engineering . They form the backbones of target detection and classification algorithms like Faster R-CNN and RetinaNet .

Globally-scalable Automated Target Recognition (GATR) framework serves as a deep learning tool for object detection and classification . GATR has been trained to detect many different objects including aircraft, vehicles, airstrips, and oil fracking wells. It searches large geographic regions with high accuracy and monitors sites for changes over time. GATR has also been extended to include online learning to improve detection with analyst feedback.

Deep learning–based methods now dominate detection results. To speed detection, proposal-based detectors such as R-CNN and Fast R-CNN were developed, followed by Faster R-CNN which introduced a region proposal network (RPN), then accelerated even more by RFCN . Mask-RCNN later added segmentation output and better detection pooling . We build on these methods, claiming no advantage in standard object detection tasks. Unlike us, however, these two-stage methods were not designed for crowded scenes where small objects appear in dense formations. Recently, some offered proposal-free detectors, including YOLO , SSD , and YOLO9000 .

To handle scale variance, feature pyramid network (FPN) added up-scaling layers. RetinaNet [utilized the same FPN model, introducing a Focal Loss to dynamically weigh hard and easy samples for better handling of class imbalances that naturally occur in detection datasets..



V. CONCLUSION

This paper describes a method for detecting retail products on grocery shelves. The proposed module contains two different modules and can be run independently. Color features are also included to allow product identification and help reduce error dates. It is useful to distinguish similar logos with different colors. You can get the feature level accuracy percentage as 75 and the score level as 81. Improved post-processing efficiency. Using two vocabularies improved accuracy. Experimental results show that the proposed approach provides better performance. This article proposed a quick and effective solution to the problem of detecting grocery items on store

Class-agnostic object detection to identify each item displayed in the image, detection by K-NN similarity search based on global image descriptors, and final refinements to further improve performance. All three steps employ the latest deep learning techniques to recognize objects and learn image descriptors via state-of-the-art CNNs (detectors).

REFERENCES

[1] M.Grace, Y.Zhou, Q.Zhang, S.Zou and X.Jiang, “RiskRanker: Scalable andaccuratezero-day android malware detection,”inProc.10thInt.Conf. Mobile Syst.,Appl.,Services, 2012, pp. 281–294.

[2] L. Marchesotti, C. Cifarelli and G. Csurka, “A framework for visual saliency detection with applications to image thumb nailing,” in IEEE 12th Int. Conf. on Computer Vision, Kyoto, Japan, pp. 2232–2239, 2009.

[3] L. Zhang, Y. Wei, H. Wang, Y. Shao and J. Shen, “Real-time detection of river surface floating object based on improved RefineDet,” IEEE Access, vol. 9, pp. 81147–81160, 2021

[4] M. Merler, C. Galleguillos and S. Belongie, “Recognizing groceries s using in vitro training data,” in IEEE Conf. on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, pp. 1–8, 2007..

[5] N. Hashim, N. Ibrahim, N. Saad, F. Sakaguchi, and Z. Zakaria, "bar-code recognition system", International Journal of Emerging Trends & Technology in Computer Science, Vol.2, Issue 4, PP.1-6, 201.

[6] P. Subpratatsavee, N. Janthong, P. Kuha, and C. Chintho, "HC2D bar-code reader using embedded camera in Android phone", IEEE 11th International Joint Conference on Computer Science and Software Engineering, Thailand, 2014.

[7] S. Ren, K. He, R. Girshick and J. Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137–1149, 2017

[8] T. Stahl, S. L. Pintea and J. C. V. Gemert, “Divide and count: Generic object counting by image divisions,”

IEEE Transactions on Image Processing, vol. 28, no. 2, pp. 1035–1044, 2019.

[9] AV. Nogueira, H. Oliveira, J. A. Silva, T. Vieira and K. Oliveira, “RetailNet: A deep learning approach for people counting and hot spots detection in retail stores,” in 32nd SIBGRAPI Conf. on Graphics, Patterns and Images (SIBGRAPI), Rio de Janeiro, Brazil, pp. 155–162, 2019.

[10] F. Reischach, S. Karpischek, F. Michahelles, and R. Adelmann, "Evaluation of 1D bar-code scanning on mobile phones", IEEE Conference on Internet of Things, Japan, 2010..

.