

The nut job

1st Priteshkumar Gohil *Department of Computer Science
Hochschule Bonn-Rhein-Sieg
Bonn, Germany
priteshbgohil@gmail.com*

Abstract

Object detection is the task in computer vision to make computer intelligent to recognize and locate the objects. Some of the domains in the object detection are less explored and one such is the food domain. Food packaging industry e.g. peanut packaging service might require to detect bad quality nuts. This paper presents an approach to detect nuts in the scene with cluttered objects, varying illumination condition, and static images. We use Convolution Neural Network (CNN), a deep learning approach to learn features, classify and localize nuts in the scene. Nuts to detect are peanut, walnut and hazelnut. This paper also presents an algorithm to detect stable frames in the video input, which are based on colour histograms and their difference. The frame detection algorithm can correctly detect the stable frame with an accuracy of 98.94%. Our deep learning approach can detect nuts in different challenging situations.

Index Terms

Object Detection, Deep Learning, Event Detection, Convolution Nerual Network

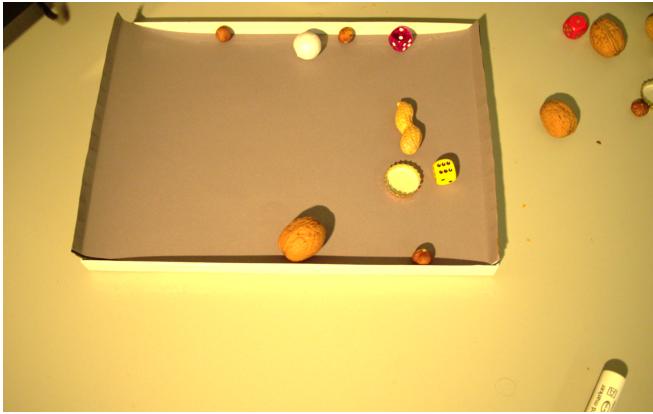
I. INTRODUCTION

The visual information that humans and animals perceive has played an important role in interacting with their environment. Giving such capabilities to machines and computer to extract the information of interest from the image helped many sectors like healthcare, driving, robotics, food and beverage etc. Therefore, extracting information of interest from images is the central focus of this project.

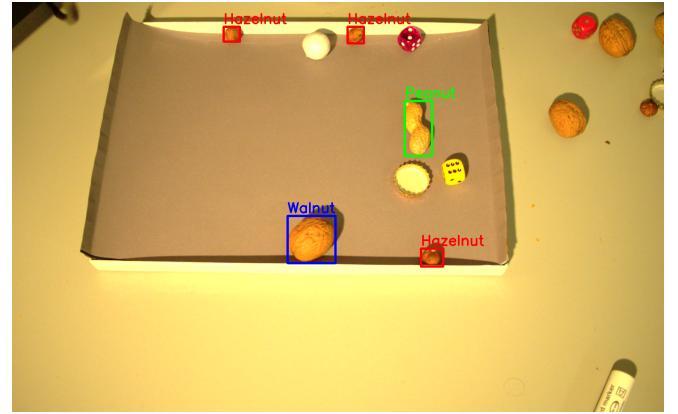
Computer vision addressed the problem of automatically extracting useful information from images such as segmentation, classification and object detection, pattern recognition, and many more. One of the active fields of research is object detection aiming to find the answer of two questions. 1. Which objects are in the image, also known as classification and 2. Where objects are located in the image, also known as localization.

Objects in the image are classified based on the features extracted from the bounding box. Figure 1a and 1b shows the original image and ground truth bounding box. The motivation for selecting features and not pixels for the classification are many. First, features extracted from the object of interest can encode the domain knowledge with finite data. Second, object detection process is faster while using feature based system [1]. Some of the widely used feature extraction techniques include, Histogram of Oriented Gradients (HOG) [2], Speeded-Up Robust Features (SURF) [3], haar wavelet [1] and convolution using CNN [4].

Detecting objects of interest in the real-world such as peanut, walnut and hazelnut pose different challenging situations. Fundamental characteristic of the nuts is its shape, colour, texture and size. For example, nuts in our experiment shares similar hue but slightly different saturation and value. The texture features of the peanut and walnut can easily be confused by the classifier. Also, shape of peanut might look same but the width and height might vary. Despite challenges from the object properties, camera and surrounding illumination plays an important role. The camera and the colour information from it impose several challenges. Figure 2 illustrates the different challenging situations such as low illumination, high illumination, shadow, occlusion and nuts outside the experimental tray in figure 1a.

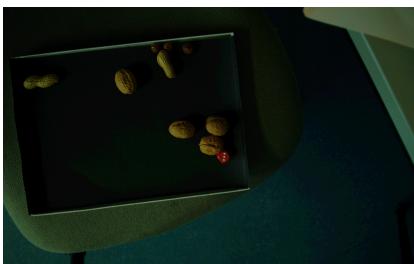


(a) Image to detect nuts inside the rectangle tray.



(b) Image annotated with bounding box.

Fig. 1: Input image left side and the ground truth annotation on the right side to show the object of interest.



(a) Poor illumination



(b) High illumination



(c) Object shadow



(d) Cluttered objects

Fig. 2: Input image left side and the ground truth annotation on the right side to show the object of interest.

Intro: Gentle start with basic terminology
Problem statement: Still unsolved because of number of challenges. Give images of occlusion, shadow, illumination, . Challenge is takes more time.
Motivation : Object detection is used in our daily life. Give many examples.
Proposed approach: Two stage: frame detection and object detection. Explain Frame detection Explain Object detection

Frame detection: # Logic

1. find color histogram
 2. calculate difference of color histogram between previous and current frame. ($cv2.HISTCMP_CORREL \& cv2.HISTCMP_KL_DIV$)
 3. we have two list of difference for entire video using two differnt methods
 4. perform element wise subtraction of correlation with $kLDivergence - i$
 new list
 5. find value ; 0. this is the frame with hands most of the time . add offset of 10 or 15 and you got it.
 6. With too dark video you will never get. So suggestion is amplify $kLDiv$ by 10 and check if (value > 0 and index < 15) perfect situation - 267, 303, 316 low illumination - 282, 371, 12 High illumination - 336 shadow - 315

II. RESULTS

- A. Speed of NN model*
- B. stable frame detection*
- C. Image processing*
rescaled, normalized, augmentation
- D. Training*
- E. Classification results*

III. CONCLUSION

Cocclusion here

ACKNOWLEDGMENT

WRITE YOUR ACK here

REFERENCES

- [1] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 1, pages I–I, Dec 2001.
- [2] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR05) - Volume 1 - Volume 01*, CVPR 05, page 886893, USA, 2005. IEEE Computer Society.
- [3] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In Aleš Leonardis, Horst Bischof, and Axel Pinz, editors, *Computer Vision – ECCV 2006*, pages 404–417, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.
- [4] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.