

# Object Detection for E-Commerce: A Survey

Suraj Khanna<sup>1</sup> Aditya Jaiswal<sup>2</sup> Ravi Kothari<sup>3</sup> Piyush Patil<sup>4</sup> Kunal Ahire<sup>5</sup>

<sup>1,2,3,4</sup>Student <sup>5</sup>Assistant Professor

<sup>1,2,3,4,5</sup>Department of Information Technology

<sup>1,2,3,4,5</sup>MET's Bhujbal Knowledge City, Institute of Engineering, Adgaon, Nashik, India

**Abstract**— Object Detection for E-Commerce will merely focus on objects that are available to buy from the E-Commerce websites. Object Detection for E-Commerce will be an interface-based application which will accept a user uploaded video and will scan the whole video thoroughly for the objects that are available to buy from the market. As soon as the object gets detected, the user can get the relevant links from the E-Commerce websites to buy that object. This interface, in simple words, will do the work of finding that object or the item over the E-Commerce market. The project will be based on Machine Learning and will be created in Python programming language. For Image Detection and classification, which is the core part of the project, different Image recognition and classification algorithms will be used. Thus, for helping the users in shopping and providing them with a touch-to-shop experience this interface application would be very beneficial.

**Keywords:** Object Detection for E-Commerce, Python Web Scraping, Image Recognition, Image Classification, YOLO Algorithm, OpenCV, Python Qt Designer, TensorFlow, Deep Learning, Convolutional Neural Network (CNN)

## I. INTRODUCTION

There are many occurrences in our day-to-day life where we are watching a video and we liked an item from that video very much and so we think of buying it. The traditional way is to go to any random search engine – search for the object using simple keywords defining it – and then buy that random object. Object Detection for E-Commerce provides one click solution for this task. The project will be specific to videos i.e. Object Detection for E-Commerce will work by detecting the objects in the video.

## II. AIM AND OBJECTIVE

### A. Aim of the application:

- 1) Application can help play a major role for E-Commerce Market by enhancing the user shopping experience.
- 2) With the help of Object Detection, users can increase productivity in various E-Commerce fields like Automated Product Tagging, Scan-to-Shop Experience, etc.

### B. Objective of the application:

- 1) Increasing the User-Shopping Experience.
- 2) Touch and go shopping experience for the users.

## III. LITERATURE SURVEY

### A. Google Lens

It is developed by Google. The application has an integrated object detection feature that can identify various objects in the input of the live camera. To be more specific, Google

Lens can be said as a Super-Powered version of Google Goggles, and it's quite similar to Samsung's Bixby Vision. Google Lens is an android application that will enable someone to do things such as point your phone at something specific, example - a specific flower, and then ask what is the object which you are pointing at to the Google Assistant. You'll not only be told the answer, but you'll get suggestions based on the object, like nearby florists, in the case of a flower. <sup>[1]</sup>

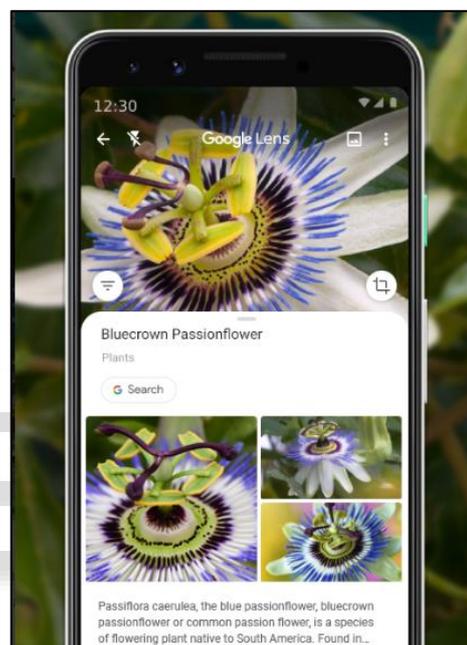


Fig. 1: Working of Google Lens

### B. KnoIT

The implementation of the program demonstrates the exactness of the object. The app also displays the percentage of accuracy the object was detected with. KnoIT is an application which works in Offline Mode and gets a shot from your camera and then uses TensorFlow (TensorFlow is a machine learning framework that is created by Google and it is used to design and train deep learning models. TensorFlow library is basically used for doing the numerical computations, which are done with data flow graphs <sup>[2]</sup>) for predicting and classifying the object which it belongs to. It is trained using Deep Learning CNN (A Convolutional Neural Network, also known as CNN, is designed to process pixel data from images and is used especially for image recognition and processing.<sup>[3]</sup>) over hundred categories of objects. The most important thing about the app is that it works OFFLINE and merely consumes any Space. People tend to forget the name of the common things that surround them in this busy world. This program is best suited for people like these. All they need to do is remove their mobile from their pocket and launch it, point the camera of the Mobile over the thing, then they're set and on the go!

By opening the app, you can simply point your smartphone to an object and it will give you results in no time. This is very fast and works the best on every mobile phone.<sup>[4]</sup>



Fig. 2: Working of KnoIT Application

### C. CamFind

This is an Android application based on shopping where a user clicks a picture and the app detects a similar object that you can buy as well. CamFind is a mobile app created in 2013 by Image Searcher, Inc. for visual search and image recognition. CamFind is operated by the developer's API - CloudSight and allows users to recognize any object simply by taking a photo with their smartphone, offering a variety of information including related images, local shopping results, price comparisons and site results. It has been downloaded over 2 million times, and over 50 million images have been identified.<sup>[5]</sup>

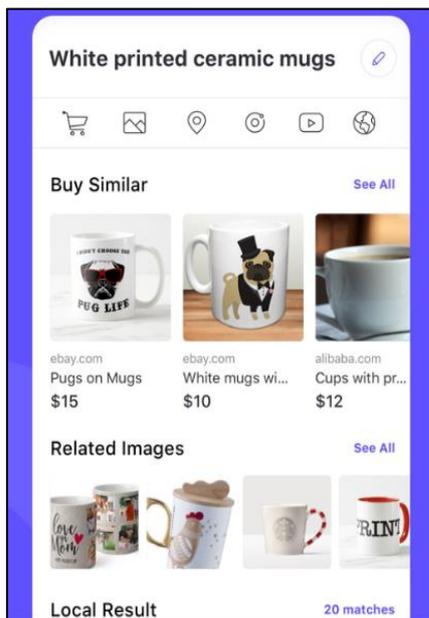


Fig. 3: Working of CamFind

### D. Screenshop

Screenshop is a Fashion based app. Screenshop provides user with powerful machine learning created fashion lenses to intuitively tailor the result with the imagination. The application will work in such a way that the user will upload an image of the look for the fashion product. This AI based app will now look for the similar looking products available in the apps market and show the relevant data to the user. Users can also change the color, category or can even choose the products based on the celebrities who wore it.<sup>[6]</sup>

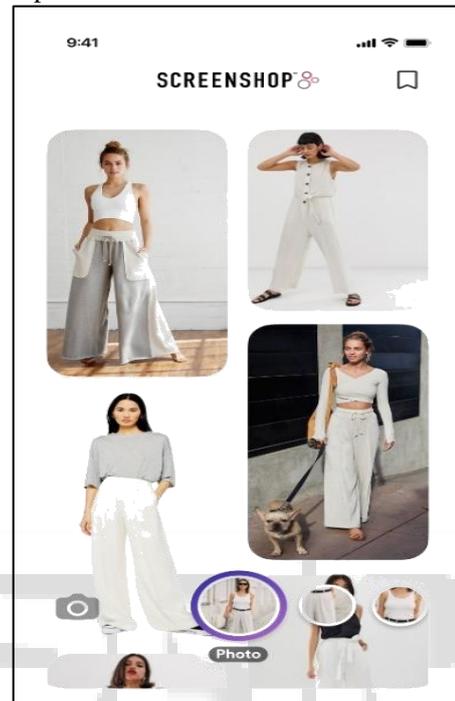


Fig. 4: Screenshop application interface<sup>[7]</sup>.

### E. Flow Powered

Flow Powered by Amazon is an Augmented Reality App that will let you discover information about items in the world around you. With this app, the user can identify millions of products varying from various categories including books, DVD's, and packaged household items. The app work in such a way that as soon as you open your app, aim your camera towards the item that the user wants to identify. Amazon's flow uses A9.com's continuous scan technology to automatically recognize items. If there are more than one items on the screen, the information is automatically delivered to the screen. The app also provides an history feature which will give you access to all your scanned items, sorted by date, item name, scan type or product category<sup>[8]</sup>

All these applications and softwares work with either direct camera input or image-based input. None of them works by detecting object from videos. Object Detection for E-Commerce is an application based on providing touch-and-shop experience for user from videos.

## IV. EXPERIMENTAL METHODOLOGY

- 1) After starting the software application, upload the video from which the object is to be detected.

- 2) As soon as the video is uploaded, a normal time-consuming processing will happen and the video will start playing.
- 3) As soon as the video continue to move forward, object classification and object detection will occur using Support Vector Machine (SVM) and YOLO Algorithm respectively.
- 4) The object will get detected in the video itself and simultaneously, the links to buy that object will be shown from different E-Commerce websites using Python Web Scraping (Web scraping is a term used to describe the use of a program or algorithm to extract and process large amounts of data from the web<sup>[9]</sup>)

## V. ARCHITECTURE

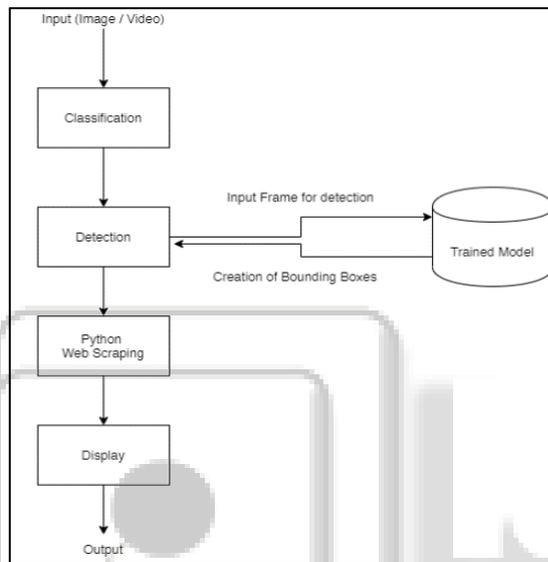


Fig. 5: Architecture of the project.

- 1) As soon as the video (or more specifically, the frame of the video) is given as an input, Object classification algorithm will work by classifying multicategory objects present in the same. Typically, SVM can be used as an Object classification algorithm. (A support vector machine (SVM) is a type of deep learning algorithm that performs supervised learning for classification or regression of data group. Support vector machines are used by like classification for sorting two groups of data. The algorithms draw lines (hyperplanes) to separate the groups according to patterns<sup>[10]</sup>).
- 2) Object Detection i.e. creation of Bounding Boxes will occur using YOLO Algorithm and the object detection using a pre-trained dataset. (YOLO Algorithm, stands for You Only Look Once algorithm, takes the entire image in a single instance and predicts the bounding box coordinates and class probabilities for these boxes. The biggest advantage of using YOLO is its superb speed – it's incredibly fast and can process 45 frames per second. YOLO also understands generalized object representation<sup>[11]</sup>)
- 3) Python Web Scraping will scrap relevant data from E-Commerce websites. The basic steps/working of Python Web Scraping is as follows:
  - a) Find the website that you want to scrape.

- b) Inspecting the page (typically using robots.txt)
  - c) Find the data you want to extract.
  - d) Write the code.
  - e) Run the code and extract the data.
  - f) Store the data in the required format.
- 4) The Bounding Boxes created and the links scraped will be displayed in the application.

## VI. GUI IMPLEMENTATION

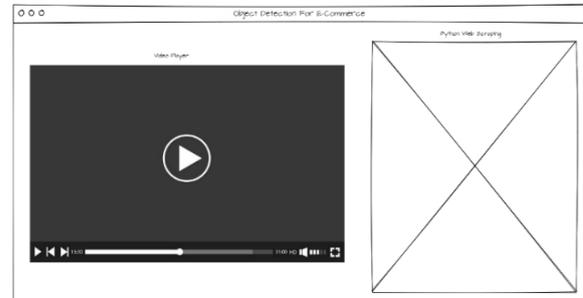


Fig. 6: Typical Implementation of the interface.

- 1) Python QtDesigner<sup>[12]</sup> will be used to create the GUI for the application.
- 2) The application developed will be a standalone application. No installation of the application software is going to be required.

## REFERENCES

- [1] <https://www.pocket-lint.com/apps/news/google/141075-what-is-google-lens-and-how-does-it-work-and-which-devices-have-it>
- [2] <https://www.datacamp.com/community/tutorials/tensorflow-low-tutorial>
- [3] <https://searchenterpriseai.techtarget.com/definition/convolutional-neural-network>
- [4] <https://play.google.com/store/apps/details?id=com.knoit.flitecamerademo>
- [5] <https://en.wikipedia.org/wiki/CamFind>
- [6] <https://apps.apple.com/us/app/screenshop/id1254964391>
- [7] <https://screenshopit.com/>
- [8] [https://play.google.com/store/apps/details?id=com.a9.flow&hl=en\\_IN](https://play.google.com/store/apps/details?id=com.a9.flow&hl=en_IN)
- [9] <https://www.datacamp.com/community/tutorials/web-scraping-using-python>
- [10] <https://whatis.techtarget.com/definition/support-vector-machine-SVM>
- [11] <https://www.analyticsvidhya.com/blog/2018/12/practical-guide-object-detection-yolo-framework-python/>
- [12] <https://build-system.fman.io/qt-designer-download>