**PROJECT REPORT**

1. **INTRODUCTION**
   1. **TITLE:** Smart Recommendation System
   2. **OBJECTIVES OF RESEARCH:**

* Image Modelling.
* Object detection using COCO dataset using optimal models.
* Displaying the number of searches made.
* Next image category prediction based on history.
* Predicting top ten web-results for the next likely category of image.
* Analyze user emotions through facial expressions and decide whether or not to display ads.
* Display ads which are related to the searches made by the user.
  1. **PROBLEM STATEMENT:**

With technology awareness taking an enormous leap each day, Visualizations are preferred to textual information. So, it is crucial that we update our existing technology to meet the most ignored needs of the users.

* Despite having highly efficient search engines such as Google, which has taken the first step towards visualization by including search by image, we see that it fails to predict the next image search that the user might make.
* With increasing browsers and social media users, the marketing field has got a brighter side, but in most cases, advertisements end up annoying users than causing the desired effect.
* Recommendation system in today’s world is an essential feature as users are often overwhelmed by choice and recommendation system help users quickly find the product they love. It leads to more sales as well as make users happier. These systems make companies dollars as recommending a product to a user might yield a sale or make the user think about whether to purchase a product or not. It’s like a salesman recommending product but a virtual one without any human curation but

a better curation based on your history.

* 1. **SOLUTION:**

The solution to the problem statement discussed here lies in most basic aspect of image classification and object detection.

* Based on the categories the previous image searches are classified into, the next category of image search can be predicted.
* This can also be implemented in collaborative filtering:
* In case of browsers, based on their browsing history, advertisement falling into the most searched category will be displayed.
* In case of social media users, based on the posts they’ve liked, advertisements falling into the most liked category will be displayed.

These recommendations are made based on the basic aspect of image classification and object detection.

* Based on the categories the previous image searches are classified into, the next category of image search can be predicted.
* This can also be implemented in collaborative filtering:
* In case of browsers, based on their browsing history, advertisement falling into the most searched category will be displayed.
* In case of social media users, based on the posts they’ve liked, advertisements falling into the most liked category will be displayed.
* “Face is the index of mind.” Human face gives away many clues to recognize the underlying emotion. So, it is necessary to consider the emotional status of the user before making a recommendation.
* If the user is undergoing a positive or a neutral emotion, then, recommendations are made.
* If the user is undergoing a negative emotion, then, recommendations are not made.

1. **LITERATURE REVIEW:**

One of the primary goals of computer vision is the understanding of visual scenes. Scene understanding involves numerous tasks including recognizing what objects are present, localizing the objects in 2D and 3D, determining the objects’ and scene’s attributes, characterizing relationships between objects and providing a semantic description of the scene. The current object classification and detection datasets help us explore the first challenges related to scene understanding. For instance, the ImageNet dataset, which contains an unprecedented number of images, has recently enabled breakthroughs in both object classification and detection research. The community has also created datasets containing object attributes, scene attributes, key points, and 3D scene information. This leads us to the obvious question: what datasets will best continue our advance towards our ultimate goal of scene understanding? We introduce a new large-scale dataset that addresses three core research problems in scene understanding: detecting non-iconic views (or non-canonical perspectives) of objects, contextual reasoning between objects and the precise 2D localization of objects. For many categories of objects, there exists an iconic view. For example, when performing a web-based image search for the object category “bike,” the top-ranked retrieved examples appear in profile, unobstructed near the center of a neatly composed photo. We posit that current recognition systems perform fairly well on iconic views, but struggle to recognize objects otherwise We introduce a large, richly-annotated dataset comprised of images depicting complex everyday scenes of common objects in their natural context. Background, partially occluded, amid clutter – reflecting the composition of actual everyday scenes. We verify this experimentally; when evaluated on everyday scenes, models trained on our data perform better than those trained with prior datasets. A challenge is finding natural images that contain multiple objects. The identity of many objects can only be resolved using context, due to small size or ambiguous appearance in the image. To push research in contextual reasoning, images depicting scenes rather than objects in isolation are necessary. Finally, we argue that detailed spatial understanding of object layout will be a core component of scene analysis. An object’s spatial location can be defined coarsely using a bounding box or with a precise pixel-level segmentation.

1. **DATA COLLECTION:**

Secondary dataset - dataset which was available on Kaggle website online.

COCO stands for Common Objects in Context. This dataset is an excellent object detection dataset with 80 classes, 80,000 training images and 40,000 validation images. As hinted by the name, images in this dataset are taken from everyday scenes thus attaching “context” to the objects captured in the scenes. This dataset was an initiative to collect natural images, the images that reflect everyday scene and provides contextual information. In everyday scene, multiple objects can be found in the same image and each should be labelled as a different object and segmented properly. This provides the labelling and segmentation of the objects in the images. A machine learning practitioner can take advantage of the labelled and segmented images to create a better performing object detection model.

1. **IMPLEMENTATION AND DATA MODELLING:**
   1. **LIBRARIES USED:**

* **ImageAI:**

**ImageAI** is a python library built to empower developers, researchers and students to build applications and systems with self-contained Deep Learning and Computer Vision capabilities using simple and few lines of code.

* Matplotlib:

Matplotlib is a plotting library for the python programming language and its numerical mathematics extension Numpy. It provides an object-oriented API for embedding plots into applications using general-purpose GUI-toolkits like TKinter, wxPython, QT, or GTK+. There is also a procedural "pylab" interface based on a state machine, designed to closely resemble that of MATLAB.

* GoogleSearch:

GoogleSearch library in python is used to obtain the web results of a particular

query.

* 1. **MODEL SELECTION FOR IMAGE PREDICTION:**

The Image Prediction class provides functions to use state-of-the-art image recognition models like

* SqueezeNet
* ResNet
* InceptionV3
* DenseNet.

All the above shown models are used for image prediction, while, the model **DenseNet** gives the highest accuracy.

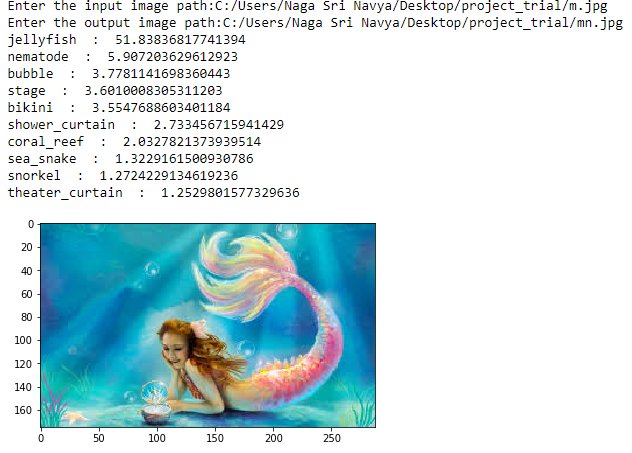
* 1. **MODEL SELECTION FOR OBJECT DETECTION:**

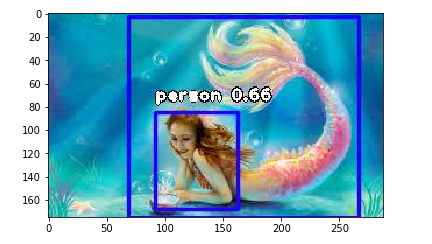
ImageAI allows us to perform all of these with state-of-the-art deep learning algorithms like

* RetinaNet
* YOLOv3
* TinyYOLOv3

**YOLOV3** algorithm of all the above-mentioned algorithms provides the highest accuracy score.

1. **OUTPUTS OBTAINED:**

****

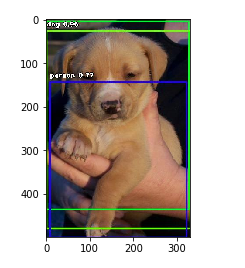
****

**FIG**

FIG I refers to the first image search made by the user. The inputs given are:

* The path where the image to be searched is present.
* The path where the image after object detection is to be saved.

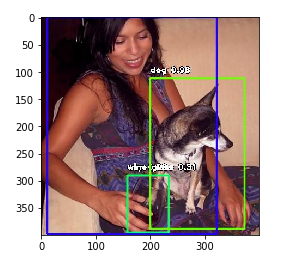
****

****

**FIG II**

From FIG II, we can observe that a provision for the user to decide whether to continue the search or not is provided. As the user chose to continue the search, second object detection took place.

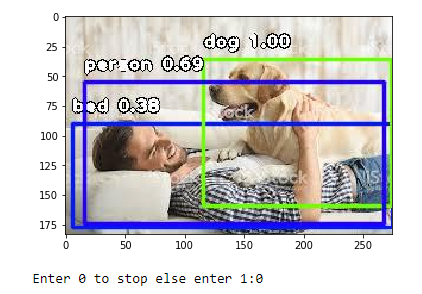


****

**FIG III**

By observing FIG III, we note that the third search also ran successfully.

****

****

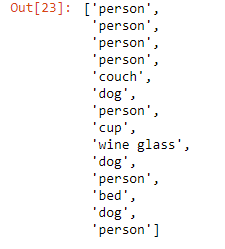
**FIG IV**

In FIG IV, we can note that the user has now decided to stop the searching process by entering 0.



**FIG V**

FIG V refers to the number of searches made by the user.



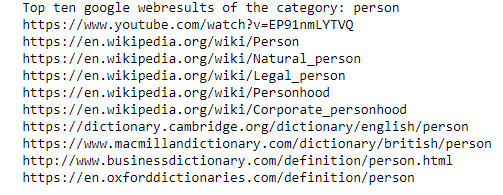
**FIG VI**

FIG VI shows the history of categories that have been searched for.



**FIG VII**

FIG VII shows the prediction of the next likely category of the image search.



**FIG VIII**

FIG VIII gives us the top ten web search results of the next likely category of the image search.

1. **CONCLUSION:**

* Image Modelling is done using python programming language.
* Object detection using COCO dataset using optimal models is done.
* The number of searches made were displayed.
* Next image category prediction based on history is made.
* The top ten web-results for the next likely category of image were predicted.
* User emotion analysis using facial expressions was done and decision regarding the displaying of ads was made.
* Ads related to searches made only were displayed.

1. **REFERENCES:**

* <https://gitlab.iit.it/pgay/tango_demo_server/blob/2c0883d0ee1c4d82b397bfd5349595e11200f584/darknet/data/coco.names>
* <https://www.kaggle.com/account/login?returnUrl=%2Fabhigupta4981%2Fcoco2017%2Fversion%2F2>

