**PROJECT REPORT ON**

**OBJECT DETECTION MODEL DEVELOPMENT**

Submitted by:

Vinay Varma Hemadri (101709710)

Ram Srikar Prava (1017101360)

**Under Guidance of:**

Professor Dr. Mehdi Zargham

Chair of Department of Computer Science

Department of Computer Science

University of Dayton

Dayton, Ohio 45469

**Introduction:**

In computer vision, the problem of estimating the class and location of objects contained within an image is known as object detection problem. Unlike classification, every instances of objects are detected in the task of object detection. So here we are designing and training a deep learning model to classify the images of dogs and cats on their features and the output is identifying whether the animal is dog or cat. After developing we have deployed the same model compatible with handheld device which is android.

**Project Description:**

The aim of this project is to build an object detection with a simple pre-trained Convolutional neural network model used for image classification. Here, we are using the YOLO approach for object detection, as the name suggests, looks at the image only once(You Only Look Once), i.e., there is only a single network evaluation unlike the previous systems like the R-CNN approach which required thousands of evaluations for a single image. It was developed by Joseph Redmon with a real-time object recognition system that can recognize multiple objects in a single frame. This CNN model takes the number of images with different features as input and gets trained accordingly with maximum accuracy, this pre-trained model is extremely light weight and can be deployed to the mobile devices, which is our ultimate aim.

**Purpose of the Project:**

The objective of this project is to detect the cat and dog respectively in each given images and display the name of the respective animal along the bounding box.

**Source of Data:**

We have downloaded the dataset from Kaggle that includes 2000 images with equal number of labels for cats and dogs. The images are from all possible angles and directions.

<https://www.kaggle.com/tarunbisht11/yolo-animal-detection-small?select=yolo-animal-detection-small>

**Sample Images:**

A picture containing cat, domestic cat, mammal

Description automatically generatedA cat sitting on a ledge

Description automatically generated with medium confidenceA cat with green eyes

Description automatically generated with medium confidence

A white dog sitting on grass

Description automatically generated with medium confidenceA puppy lying on a bed

Description automatically generated with low confidenceA dog lying on a bed

Description automatically generated with medium confidence

**Proposed method:**

We are using the YOLO approach for object detection, as the name suggests, looks at the image only once(YOLO - You Only Look Once), i.e., there is only a single network evaluation unlike the previous systems like the R-CNN approach which required thousands of evaluations for a single image. It was developed by Joseph Redmon with a real-time object recognition system that can recognize multiple objects in a single frame. YOLO v5 is almost as accurate as YOLO v4 but a bit less efficient than YOLO v4. It is the fastest YOLO version and is very easy to implement.

**Steps of Implementation:**

* Downloaded cats and dogs images from a public dataset in Kaggle.
* Labelled all of these images by using open source website <https://www.makesense.ai/>. Selected the object (face of dog or cat) in each image and exported the annotations data in the txt and xml format.
* Cloned the yolov5 git repository to our notebook and additionally we had to import a lot of related dependencies to get this up and running.
* Installed roboflow and imported roboflow to upload my dataset to the roboflow.
* Uploaded our local dataset to the roboflow. We can label and annotate the images in roboflow as well. But in our case, we have already labelled and annotated the images in the makesense.ai website. Hence uploading the images and their respective annotated files to this Roboflow.
* We will be given an API key and value pair code once we export our dataset from the roboflow. By running this in our notebook we will be able to import the dataset (train-70%, valid-20% and test-10%)
* By feeding this dataset to our yolo model we got an accuracy of (mAP) 88% with 100 epochs in total.

**Challenges Faced:**

* We were unable to get the data.yaml file which holds crucial to dataset configuration details such as train and valid paths; class names. We overcome this issue by replacing latest Roboflow API code with the stable code.
* There were too many inconsistencies in our data. Using Roboflow, we made the labels consistent.

**Conclusion:**

Therefore we were able to develop and implement the state of art YOLOv5 model in our Object Detection project on dog and cat images and achieved the below results with very good accuracy.

A cat with a green string in its mouth

Description automatically generated with low confidence A person holding a dog

Description automatically generated with medium confidence

**References:**

* <https://www.youtube.com/watch?v=x0ThXHbtqCQ>
* <https://www.youtube.com/watch?v=MdF6x6ZmLAY&t=1258s>
* <https://www.youtube.com/watch?v=VDqsK3FDIsQ&t=590s>
* <https://colab.research.google.com/drive/1gDZ2xcTOgR39tGGs-EZ6i3RTs16wmzZQ#scrollTo=Ug_PhK1oqwQA>
* <https://www.kaggle.com/tarunbisht11/yolo-animal-detection-small?select=yolo-animal-detection-small>