Lab- 6

CSET340- Advanced Computer Vision and Video analytics

Task-1:- Image segmentation and Object identification/ detection using Hough transform

- T1.1 Edge-based segmentation
  - Detects boundaries between objects by marking discontinuities in color or gray level (Canny popular edge detection technique)

- T1.2 Region-based segmentation
  - Regions based (group pixels) on criteria like color, texture, or intensity.

- T1.3 Hough transform to detect specific shapes like lines and circles within an image.
  - Hough Transform is a feature extraction technique can identify geometric shapes in images by converting them into a mathematical representation in parameter space.
  - Hough Transform is primarily used to detect straight lines.
  - Hough Transform can be used to detect circular objects within an image.

Note:- Take any road image, or other images with circle and apply hough transform to image.

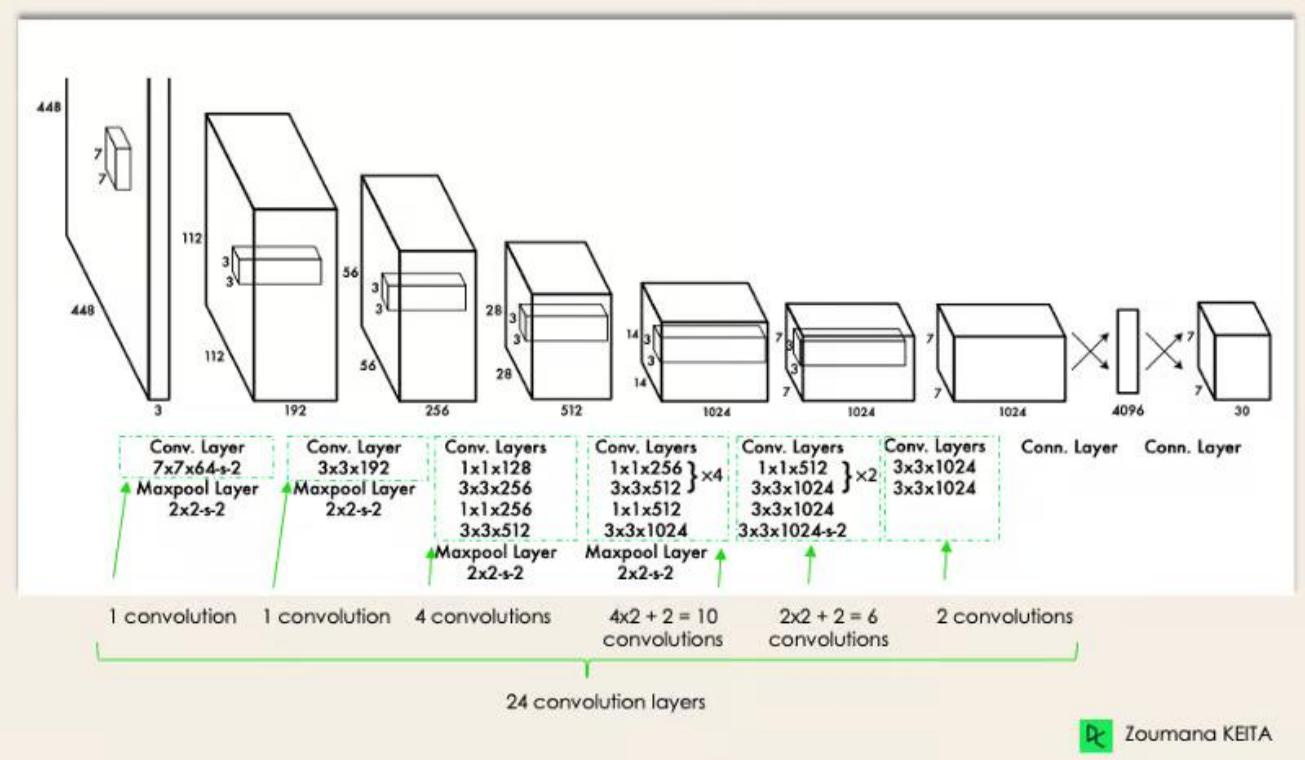
#Detect points that form a line

lines = cv2.HoughLinesP(edges, 1, np.pi/180, 68, minLineLength=15, maxLineGap=250)

Task-2:- Object Detection using following techniques.

- T2.1 YOLO (You Only Look Once)

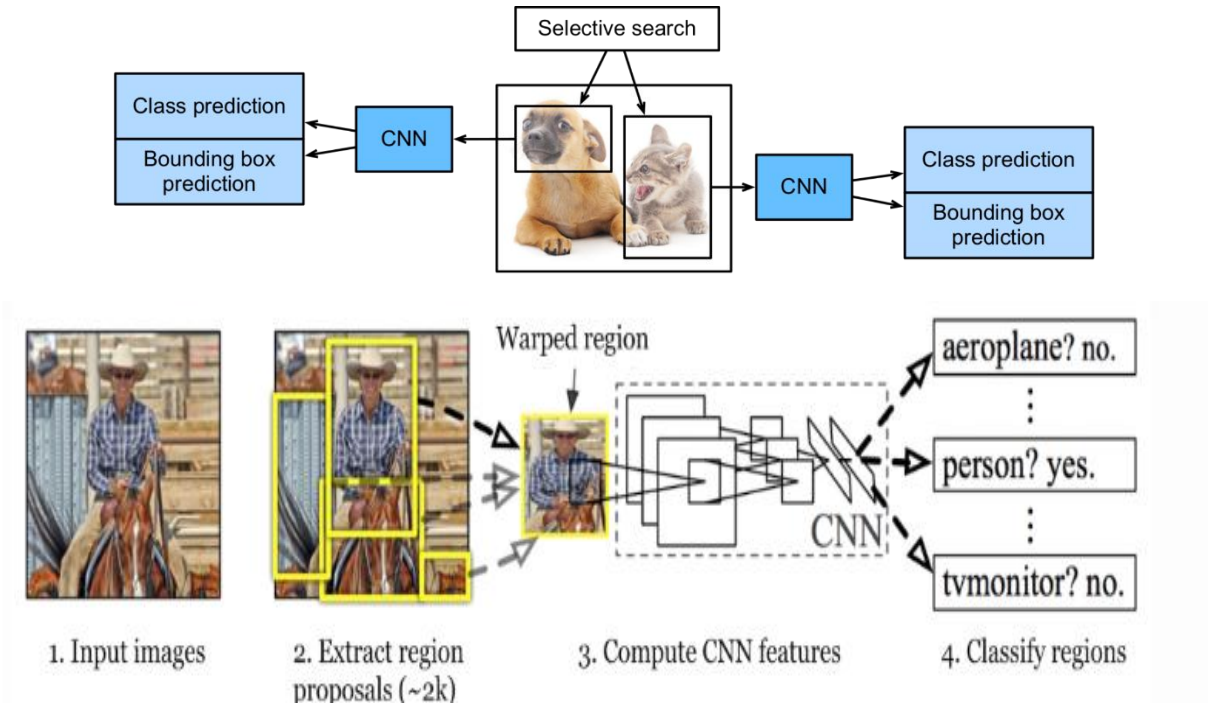
- YOLO is a variant of a Convolutional Neural Network (CNN) as it utilizes a single CNN architecture to perform real-time object detection by predicting bounding boxes and class probabilities.



Key Steps:-

1. **Step 1: Setup the Environment** : pip install opencv-python
   pip install ultralytics
   import cv2
   from ultralytics import YOLO
2. **Step 2: Load YOLO : yolo = YOLO('yolo.pt') .**
3. Optional :- %%load any variant of YOLO for fast convergence.

- T2.2 Region based Convolution Neural Network (RCNN)
  - RCNN utilizes a convolutional neural network (CNN) architecture to extract features for object detection in computer vision tasks.
  - Given an input image, R-CNN begins by applying selective search to extract regions of interest (ROI), where each ROI is a rectangle that may represent the boundary of an object in image.
  - Depending on the scenario, there may be as many as two thousand ROIs.
  - After that, each ROI is fed through a neural network to produce output features.



Steps-

1. Image Input
2. Divide image with regions
3. Consider each region as separate image
4. Pass all these regions (images) to CNN and classify them into classes
5. We can combine all these regions (images) into one and its corresponding classes

Optional:- Try some variant of R-CNN like **Fast RCNN** or **MASK RCNN** or **Faster-RCNN.**

Datasets-

1. Fashion MNIST dataset (The dataset of 70,000 28x28 labeled fashion images)
   Fashion-MNIST is a dataset of Zalando's article images—consisting of a training set of 60,000 examples and a test set of 10,000 examples. Each example is a 28x28 grayscale image, associated with a label from 10 classes.
2. CIFAR-100 dataset
   This dataset is just like the CIFAR-10, except it has 100 classes containing 600 images each. There are 500 training images and 100 testing images per class. The 100 classes in the CIFAR-100 are grouped into 20 superclasses. Each image comes with a "fine" label (the class to which it belongs) and a "coarse" label (the superclass to which it belongs).

Note:-

1. Submit your assignment on LMS in due time.
2. Marks will be deducted for late submission.