Concepts

- Understand the basics of object detection and concepts of 'localization' and 'classification'.
- Learn about most commonly used algorithm including 'YOLO', 'Faster R-CNN', 'Single Shot MultiBox Detector (SSD)'.

Object Detection Architectures

- YOLO
- Faster R-CNN
- SSD
- RetinaNet

Data Annotation, Augmentation, Transfer Learning and Evaluation

- Hands-on-experience in data annotation and bounding box annotations using **LabelImg** or **CVAT** to create a custom dataset.
- Primary evaluation metrics is mAP (mean Average Precision) that determines how well the model model predicts both the class and location of the object.
- IoU (Intersection over Union), measures the overlap between the predicted bounding box and the ground truth.
- Precision and Recall, determines how well the model is classifying objects and detecting them correctly.
- Explore different types of data augments techniques during the training time.
- Adopt different pre-trained models that were trained on benchmarked large datasets and fine tune for specific tasks.

Tutorials and Documentation

- TensorFlow Object Detection API
- Detectron2 (Facebook)
- PyTorch Object Detection

Datasets

- COCO (Common Objects in Context) [Large]
- PASCAL VOC [Medium]
- Open Image Dataset [Large]
- Roboflow Custom Datasets [Custom]
- iMaterialist (Product) Challenge [Large]
- Plant CLEF [Medium]
- Custom Datasets [Utilize **LabelImg** or **CVAT** to annotate the images]

Papers

- *"Rich feature hierarchies for accurate object detection and semantic segmentation" (2014) by R. Girshick et al.*
    - Introduced R-CNN
    - CNNs, region proposals, object detection pipeline

- *"Fast R-CNN" (2015) by R. Girshick*
    - Introduce Fast-RCNN
    - RoI pooling, end-to-end training, faster detection
- *"Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks" (2015) by Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun*
    - Introduced **Region Proposal Networks (RPNs)**
    - Region Proposal Networks, end-to-end learning, real-time detection
- *"YOLO: You Only Look Once" (2016) by Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi*
    - Treats object detection as a **single regression problem**
    - Single-shot detection, real-time performance, grid-based object localization and classification
- *"YOLOv2: Better, Faster, Stronger" (2017) by Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi*
    - Refines the original YOLO architecture, introducing improvements like batch normalization, anchor boxes, and a better backbone network (Darknet-19), making it faster and more accurate
    - Anchor boxes, improved training methods, Darknet architecture
- *"SSD: Single Shot MultiBox Detector" (2016) by Wei Liu et al.*
    - Introduces a **single-shot detector** approach
    - Multi-scale detection, anchor boxes, efficient inference
- *"RetinaNet: Focal Loss for Dense Object Detection" (2017) by Tsung-Yi Lin et al*.
    - addresses the class imbalance problem (where most regions in an image do not contain objects) by introducing **focal loss**
    - Focal loss, dense detection, class imbalance
- *"Mask R-CNN" (2017) by Kaiming He, Georgia Gkioxari, Piotr Dollar, Ross B. Girshick*
    - **Mask R-CNN** extends Faster R-CNN by adding a mask branch to predict segmentation masks for each object, allowing for both **object detection** and **instance segmentation**
    - Instance segmentation, RoIAlign, mask prediction
- *"Faster R-CNN with Cascade R-CNN" (2018) by Zhaowei Cai and Nuno Vasconcelos*
    - introduces a **cascaded architecture** to improve detection performance, particularly for objects with high variability in size or aspect ratio. It's a great way to learn about multi-stage object detection
    - Cascaded networks, multi-stage detectors, better performance for hard examples
- *Detecting Objects in RGB-D Indoor Scenes" (2018) by Chih-Yao Ma, Chih-Yu Hsu, Hung-Yu Tseng, and Woei-Tseng Chen*
    - explores object detection in **RGB-D** (color and depth) images, an extension of 2D object detection that can be particularly useful in indoor environments and robotic applications
    - Depth information, 3D object detection, RGB-D sensor integration.
- *"You Only Look One-Level Feature: A Simple and Efficient Framework for Object Detection" (2020) by Wei Zhan et al.*

    - introduces a **simplified approach to detection**, focusing on using a single-level feature map for detection, significantly reducing the complexity of existing methods
    - Efficient object detection, single-level feature maps, simplification
- *"A Comprehensive Review on Object Detection Algorithms" (2020)*
    - A great review paper that covers a wide range of object detection algorithms, comparing various models, including R-CNNs, YOLO, SSD, and others. This is more of a survey paper to get a broad view of the field.
    - Survey of object detection methods, performance comparison, trends in object detection.
- *"Object Detection: A Survey" (2019) by Ali Farhadi, Muhammad Rastegari, et al.*
    - A **comprehensive overview** of various object detection approaches from a historical perspective

- ***"EfficientDet: Scalable and Efficient Object Detection"*** *by Mingxing Tan, Ruoming Pang, and Quoc V. Le in 2020.*

Model Development, Training, and Validation

- Develop a basic customize model for a custom dataset to get started with object detection tasks.
- Utilize Transfer Learning to work with pre-trained models for different types of problem.
- Explore loss functions and optimizers.
- Develop models for different types of tasks, initially starts with single object detection.

Model Compression to facilitate Deployment into large-scale production environment

- Explore both Tensorflow and PyTorch approaches.