# Project on

IDENTIFYING OF STARFISH UNDER SEA USING DEEP LEARNING TECHNOLOGY

# CONTENT

- Introduction
- Literature Survey
- Motivation

- Project Architecture
- Components

    -> Dataset

    ->Model building

    ->Model deployment

- Conclusion
- References

# INTRODUCTION

Australia's stunningly beautiful Great Barrier Reef is the world’s largest coral reef and home to 1,500 species of fish, 400 species of corals, 130 species of sharks, rays, and a massive variety of other sea life.

Unfortunately, the reef is under threat, in part because of the overpopulation of one particular starfish – the coral-eating crown-of-thorns starfish (or COTS for short). Scientists, tourism operators and reef managers established a large-scale intervention program to control COTS outbreaks to ecologically sustainable levels.

To know where the COTS are, a traditional reef survey method, called "Manta Tow", is performed by a snorkel diver. While towed by a boat, they visually assess the reef, stopping to record variables observed every 200m. While generally effective, this method faces clear limitations, including operational scalability, data resolution, reliability, and traceability.

Using Deep Learning technology we can correctly predict the position of COTS and take necessary steps to save GBR.

# LITERATURE SURVEY

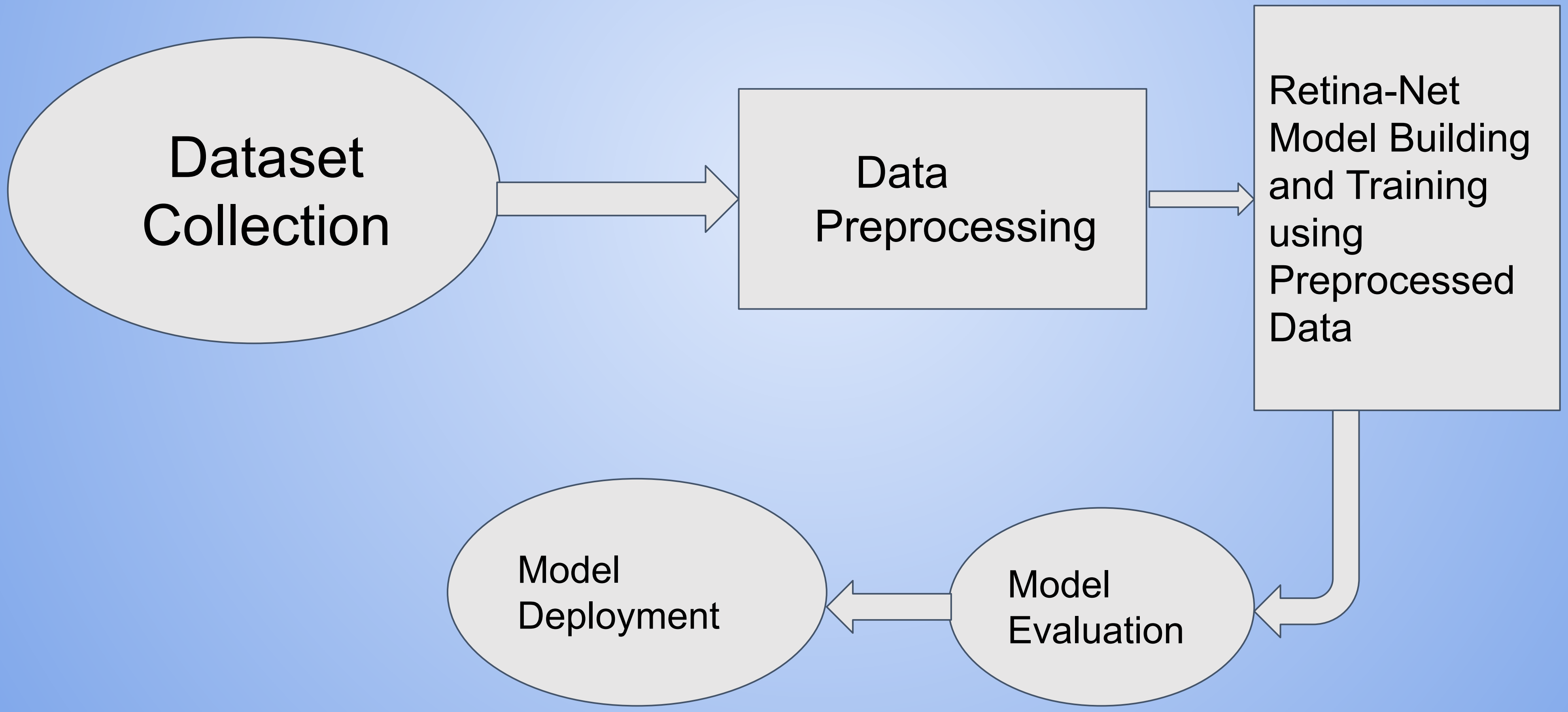
Object detection is a common term for computer vision techniques classifying and locating objects in an image. Modern object detection is largely based on use of convolutional neural networks Some of the most relevant system types today are Faster R-CNN, R-FCN, Multibox Single Shot Detector (SSD) and YOLO (You Only Look Once), Original R-CNN method worked by running a neural net classifier on samples cropped from images using externally computed box proposals. This approach was computationally expensive due to many crops.

Object detection models can be broadly classified into "single-stage" and "two-stage" detectors. Two-stage detectors are often more accurate but at the cost of being slower. Modern research outcome suggest that RetinaNet, a popular single-stage detector, which is accurate and runs fast. RetinaNet uses a feature pyramid network to efficiently detect objects at multiple scales and introduces a new loss, the Focal loss function, to alleviate the problem of the extreme foreground-background class imbalance. It is also used for classifying dense object clearly.

# MOTIVATION

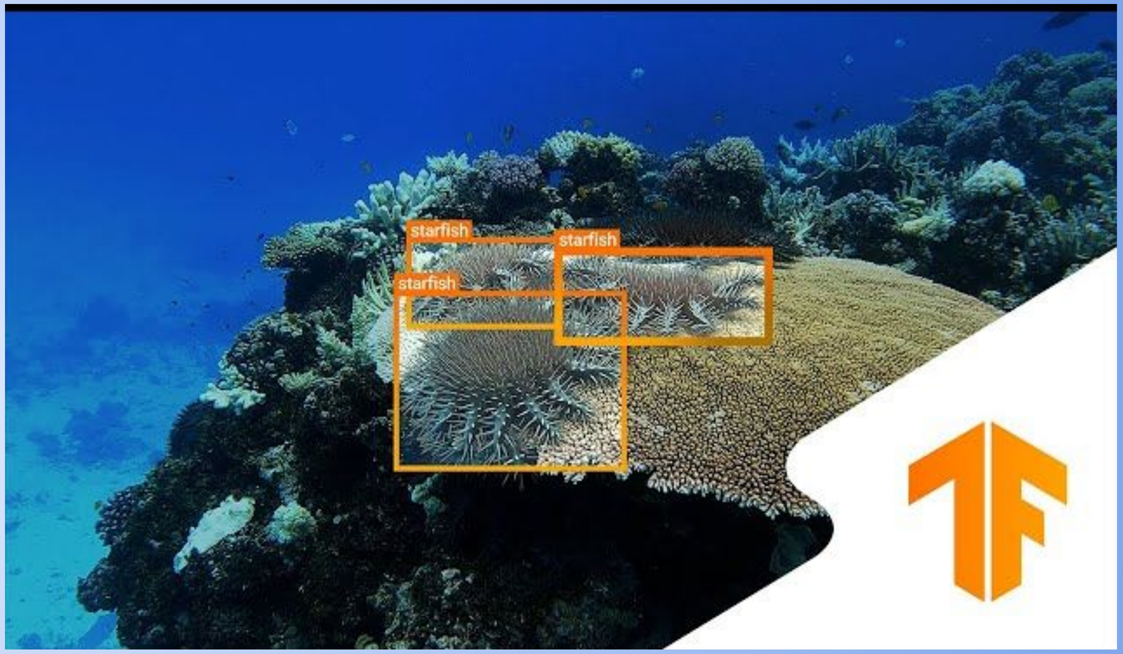
- Environmental Issue.
- Protection of Environment.
- Huge impact
- Using AI to solve the Problem.

# PROJECT ARCHITECHTURE



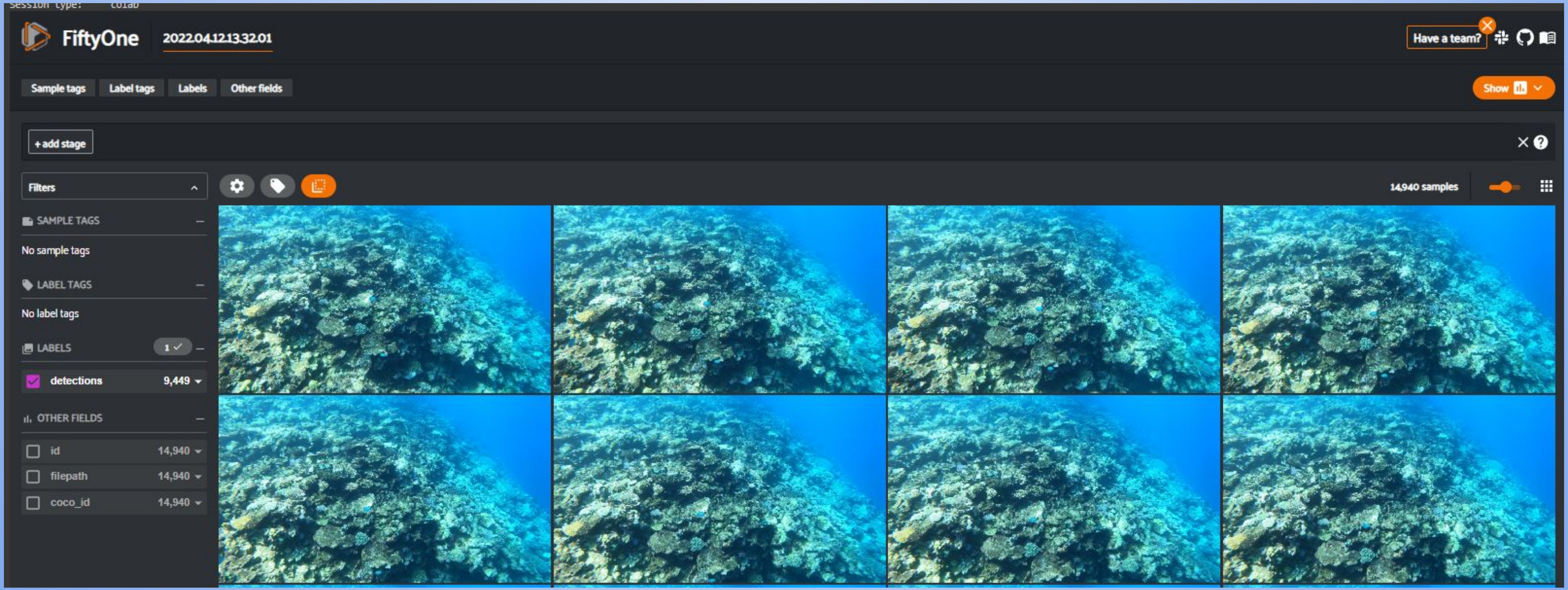
# DATASET COLLECTION

- Dataset is collected from kaggle.
- Kaggle is a open source Machine learning platform.
- Data is about 15.23 GB in size.
- Dataset contains 23501 images in total.
- Dataset is fully annotated.



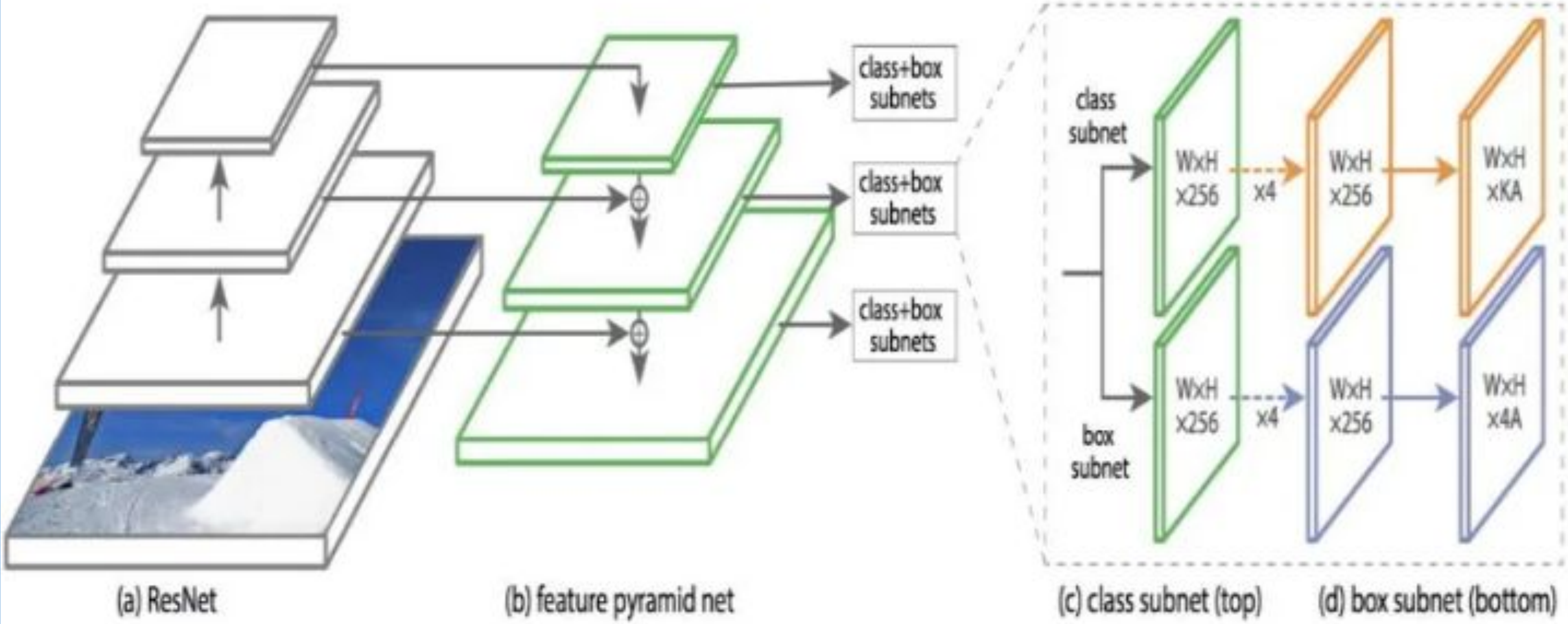
# DATA PREPROCESSING

- Converting data from csv format to COCO format.
- Removing duplicates files by checking file hashes.
- Converting data to a specific format
- Checking null value and replacing it.
- 



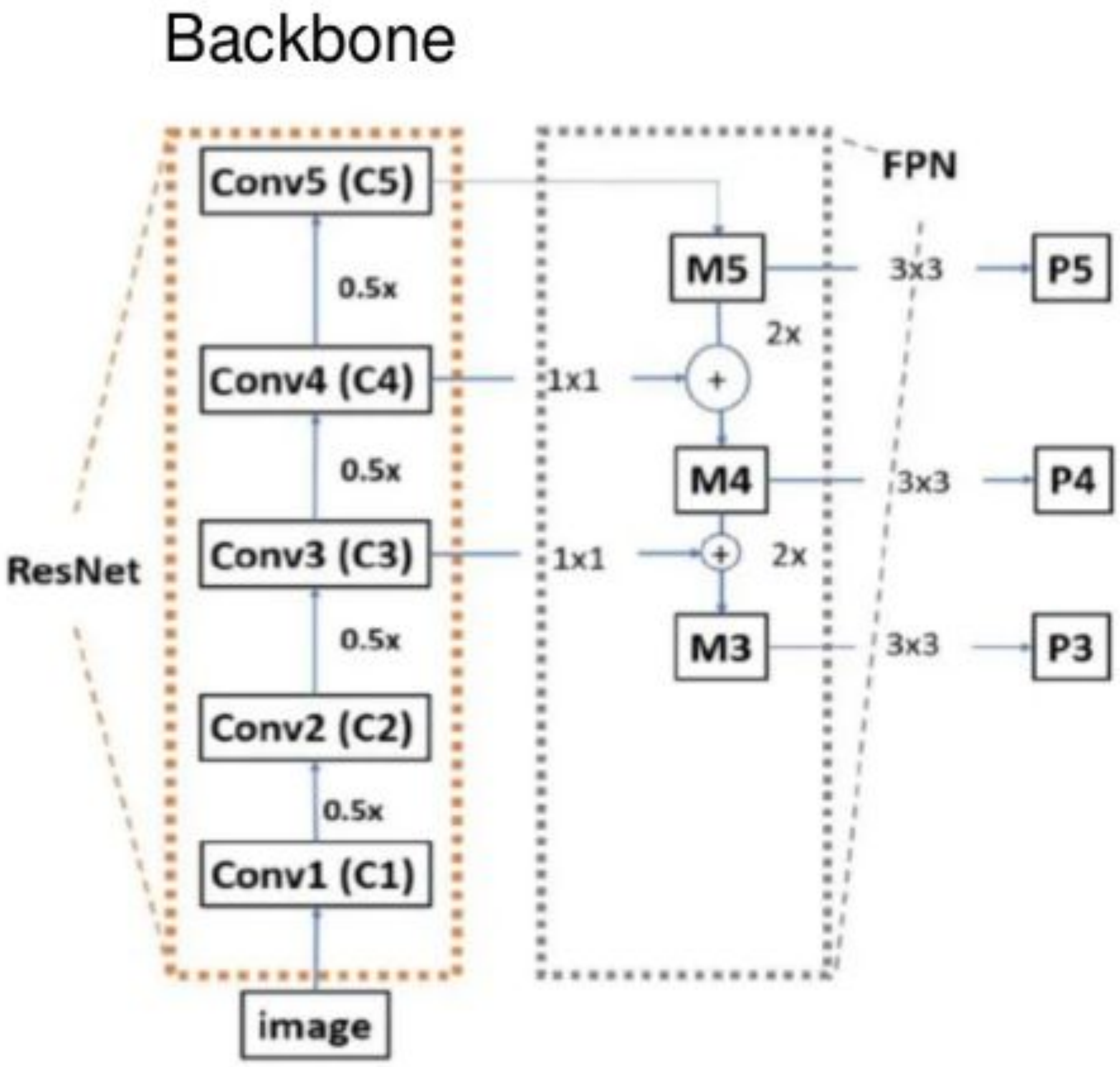
# MODEL BUILDING

# Retinanet - Architecture




RetinaNet

# Retinanet

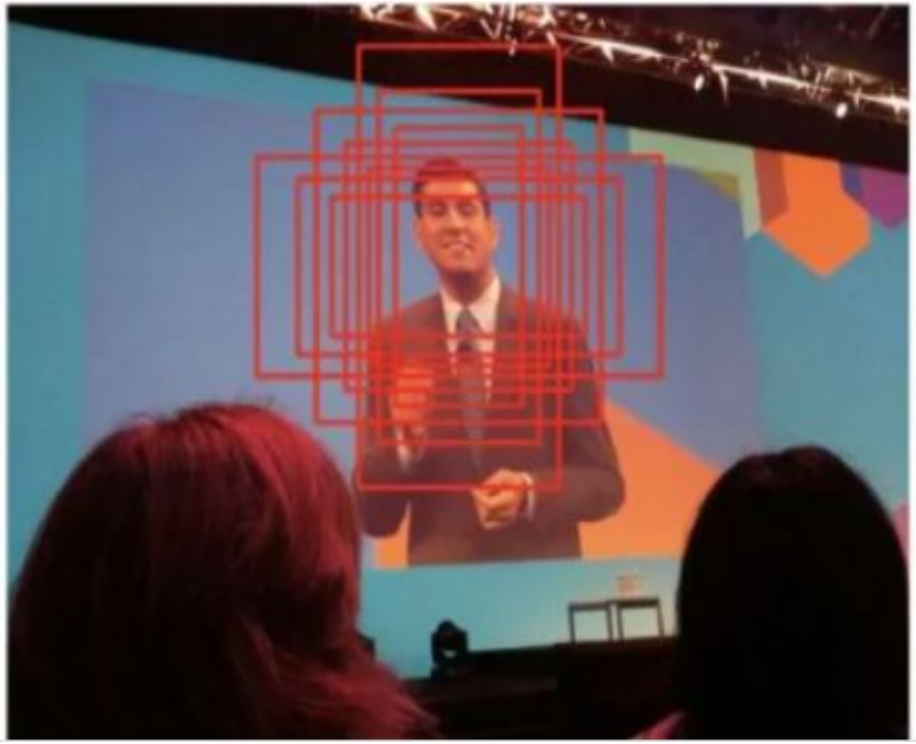



Can be:
Densenet
VGG
MobileNet

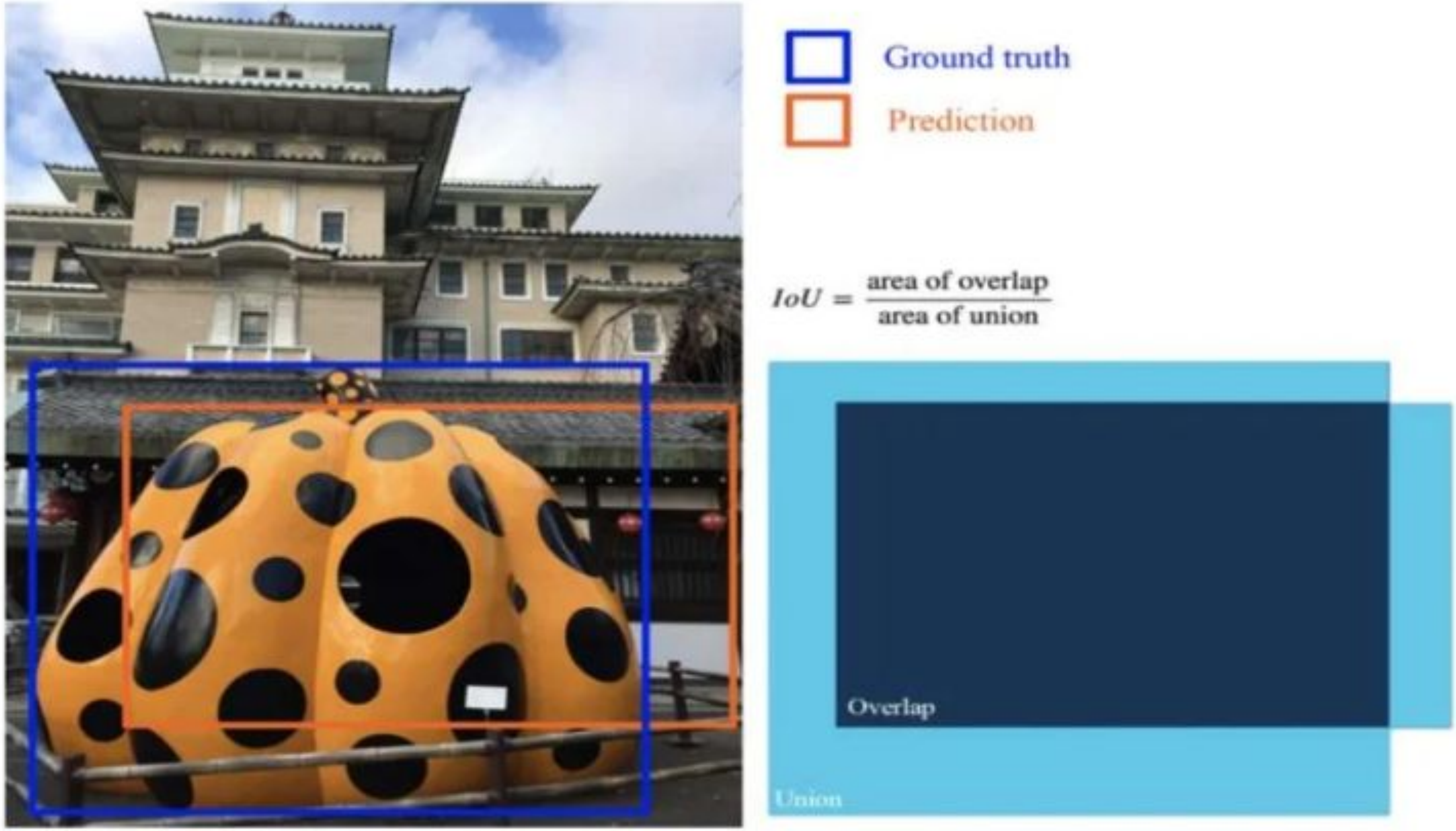
Activation maps at different pyramid levels

# ANCHORS BOX

- Anchor boxes are a set of predefined bounding boxes of a certain height and width.
- These boxes are defined to capture the scale and aspect ratio of specific object classes you want to detect and are typically chosen based on object sizes in your training datasets.
- Aspect ratios :0.5,1,2
- Scales:1,1.25,1.58
- Strides:8,16,32,64,128
- Sizes:32,64,128,256,512
- Total(A):ratios*scales=3*3=9 anchors/pixel location
- (K) object classes



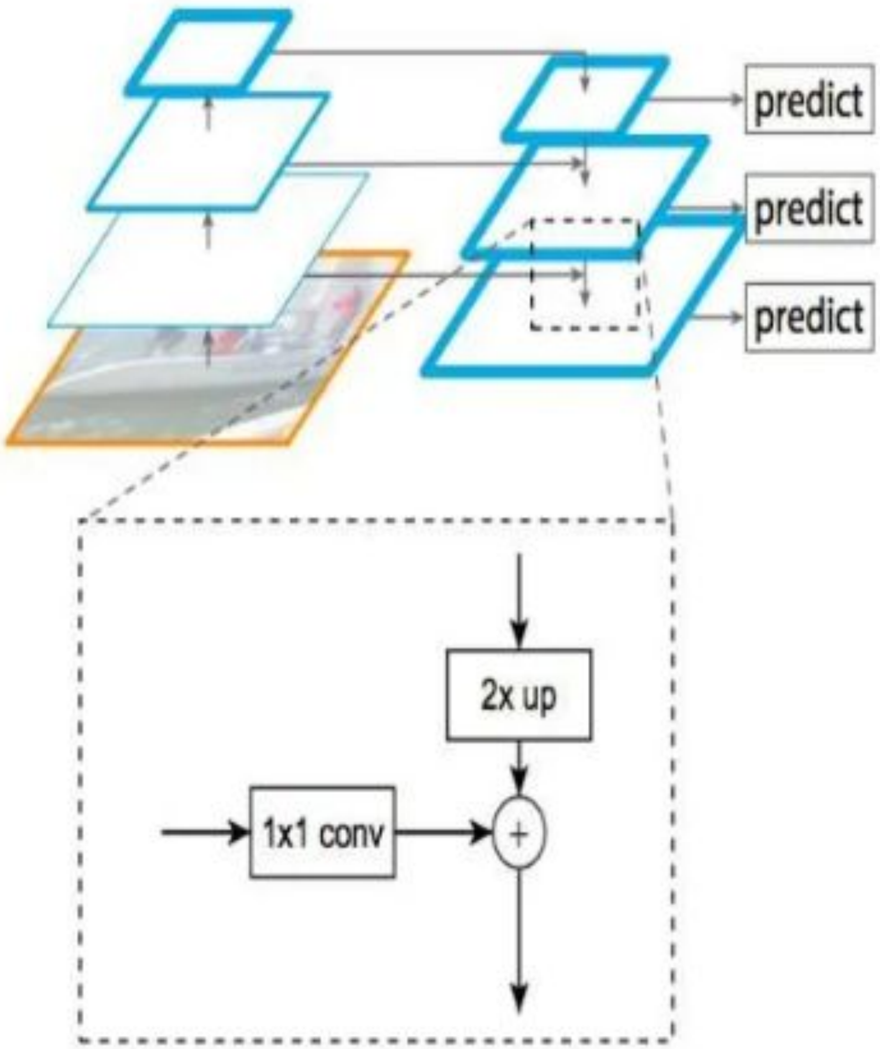
# IOU: Intersection over union (refresher)




IoU definition

# Feature pyramid networks (FPN)

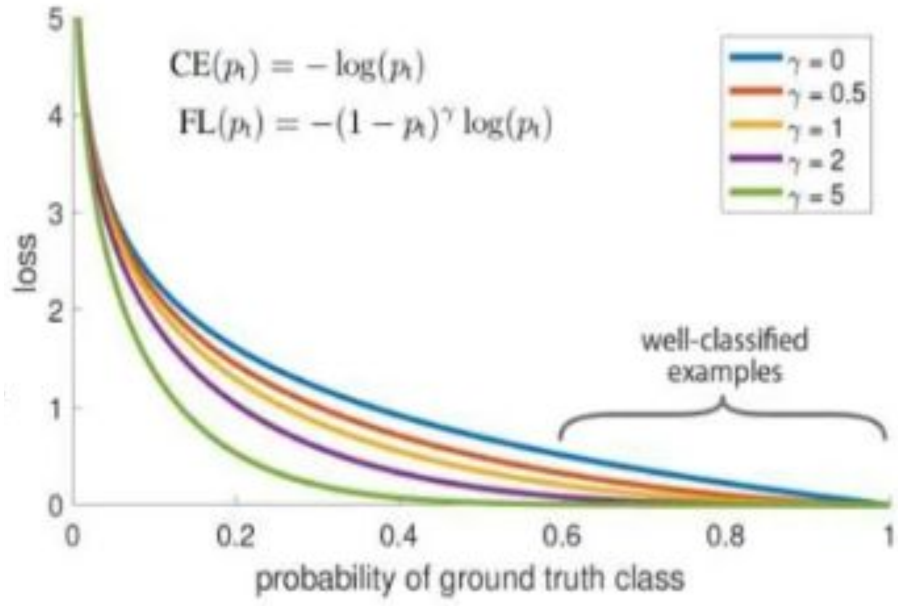
- Improve predictive power of lower-level feature maps by adding contextual information from higher-level feature maps




Top-Down+Lateral connections

# Focal loss

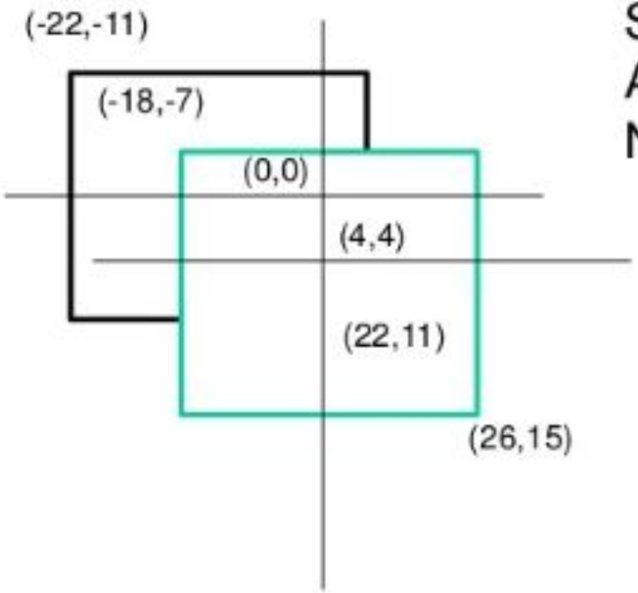
$$\mathrm{FL}(p_t) = -\alpha_t(1-p_t)^\gamma \log(p_t).$$




- Every sample weighted according to its error!
- Modulating factor added
- Focusing parameter gamma smoothly adjusts the rate at which easy examples are downweighted.
- Misclassified,pt is small, modulating factor is near 1, loss is unaffected
- As pt->1 the factor goes to 0 and the loss for well classified examples is down-weighted
- with gamma=2 examples classified with pt=0.9 would have 100* lower loss compared to CE and with pt= 0.968 it would have 1000*lower loss. This in turn increases the importance of correcting misclassified examples.

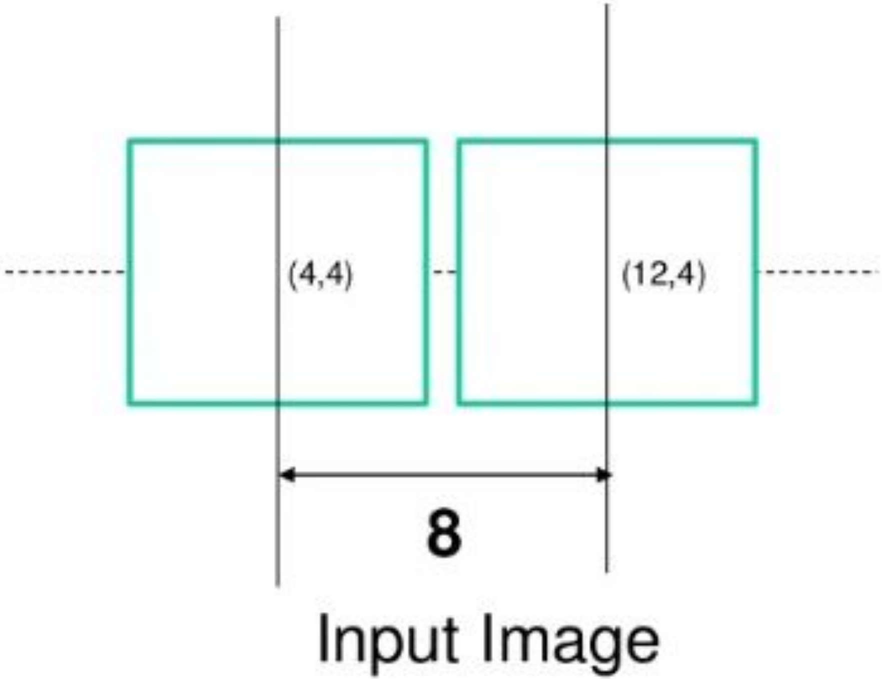
# Shift anchors

Shift anchors according to input image from activation map




Shift anchor centered at (0,0) on P3 (stride 8)
Activation map by [ 4. 4. 4. 4.]
Next shift [ 12. 4. 12. 4.], [ 20. 4. 20. 4.] , ….

**Anchors applied wrt to input image!**



# Model Deployment

- Model has been deployed in cloud platform Google Cloud, for online and batch predictions.
- API has been generated to produce real-time predictions using valid key.
- Constantly monitoring model performances and updating model within certain interval for better accuracy.

# Conclusion

- Using Deep-learning technology we can easily detect COTs and protect GBR.
- Fine-tuning model and more quality data can increase its accuracy.
- Using of advanced Object detection Model (More computationally expensive) can increase model accuracy and performances.

THANK YOU