

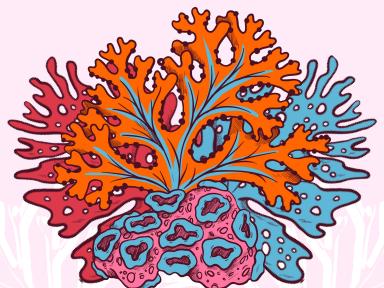
Group 7

Shashank Bidwai

Sohan Birajdar

Steffy Roselina Eben Judson

Amirthavarshini Dhanavel



Coral Detection Using Detection Transformer (DETR)



Table of contents

Introduction

Objective and implementation strategy

Approach

Explanation of DETR

01

02

03

04

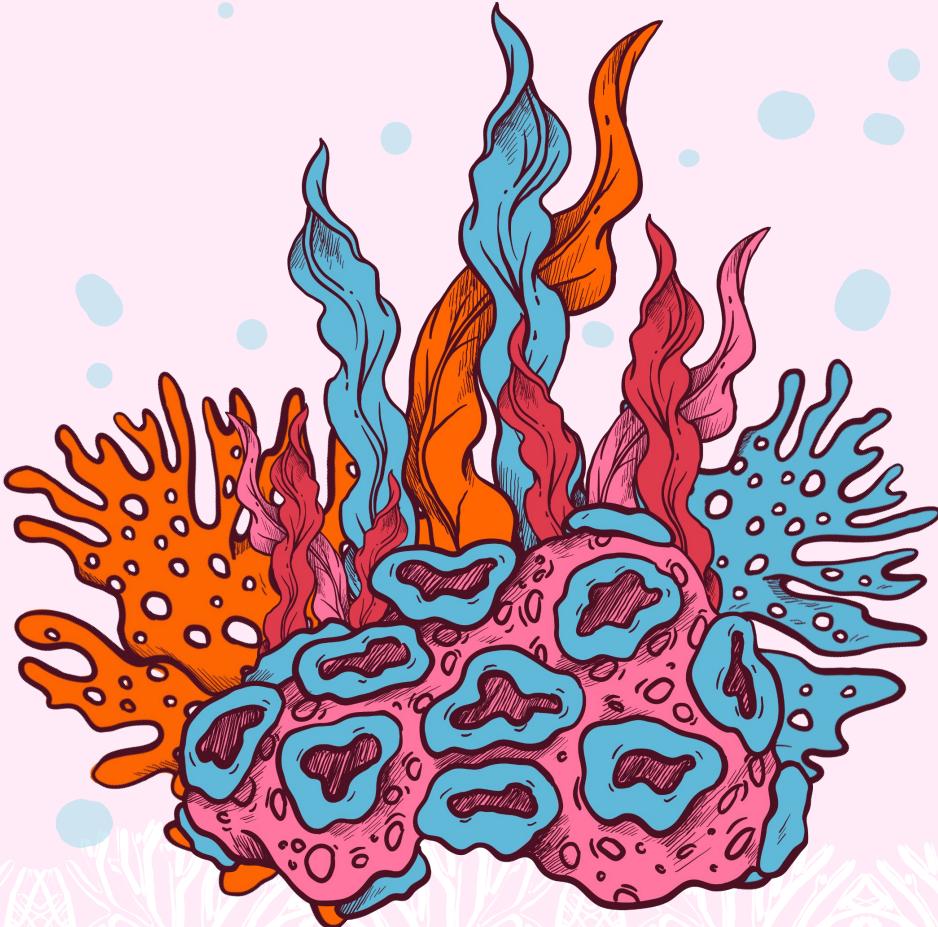
Modeling & Output

Explanation of Model Architecture

Conclusion

Output of the model and final conclusion.





01

Introduction

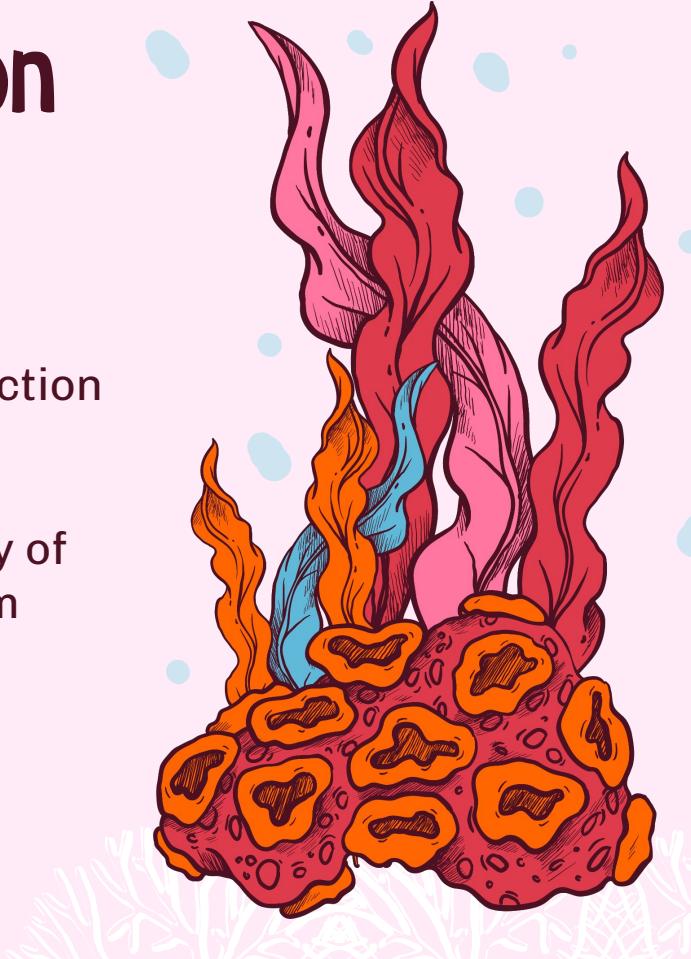
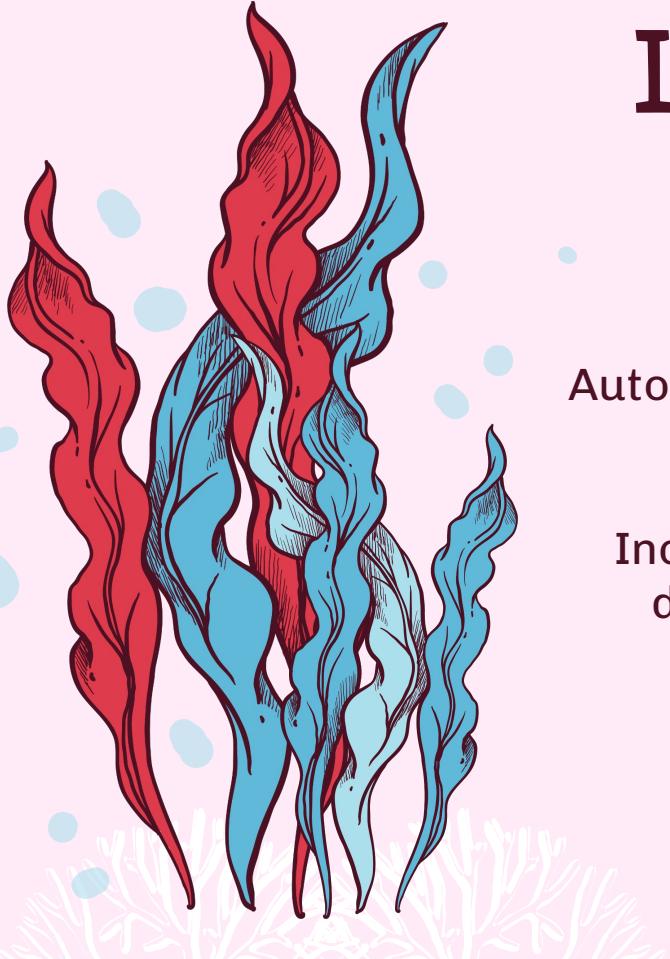
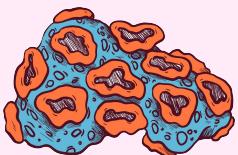
Objective and implementation
strategy

Introduction

Motivation

Automate the coral detection processes

Increase the efficiency of detection mechanism



DETR - Detection Transformer

The transformer architectures help to increase the speed up.

It is also helpful in terms of solving object detection problems.

The DETR performs faster because of its capability of Parallel processing capability.

Advantages and Drawbacks of using DETR model

Advantages:

1. The DETR model is quite simple and easy to implement and understand.
2. DETR showcases the better performance on larger datasets.
3. Uses transformer encoder-decoder structure.
4. It does the bipartite matching between predicted and ground truth object.

Drawbacks:

1. The DETR model takes a lot of time to train, it needs 500 epochs to generate the most accurate results.
2. The DETR model perform poor on smaller datasets.



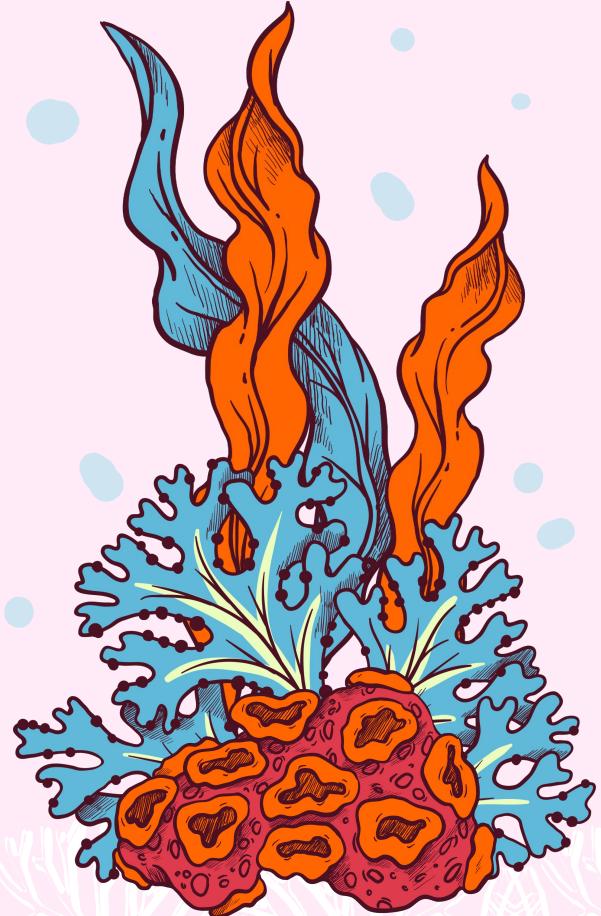
Why DETR is better than CNN

1. The DETR model is based on the encoder-decoder transformers.
2. DETR treats an object detection problem as a direct set prediction problem.
3. The DETR leverage global attention to compensate for CNN's local receptive fields.
4. It can achieve similar results with state of the art models, but using fewer parameters and floating point operations per second, which indicates efficiency.
5. Due to the global attention feature mechanism, DETR achieves better performance on large object.



02

Approach

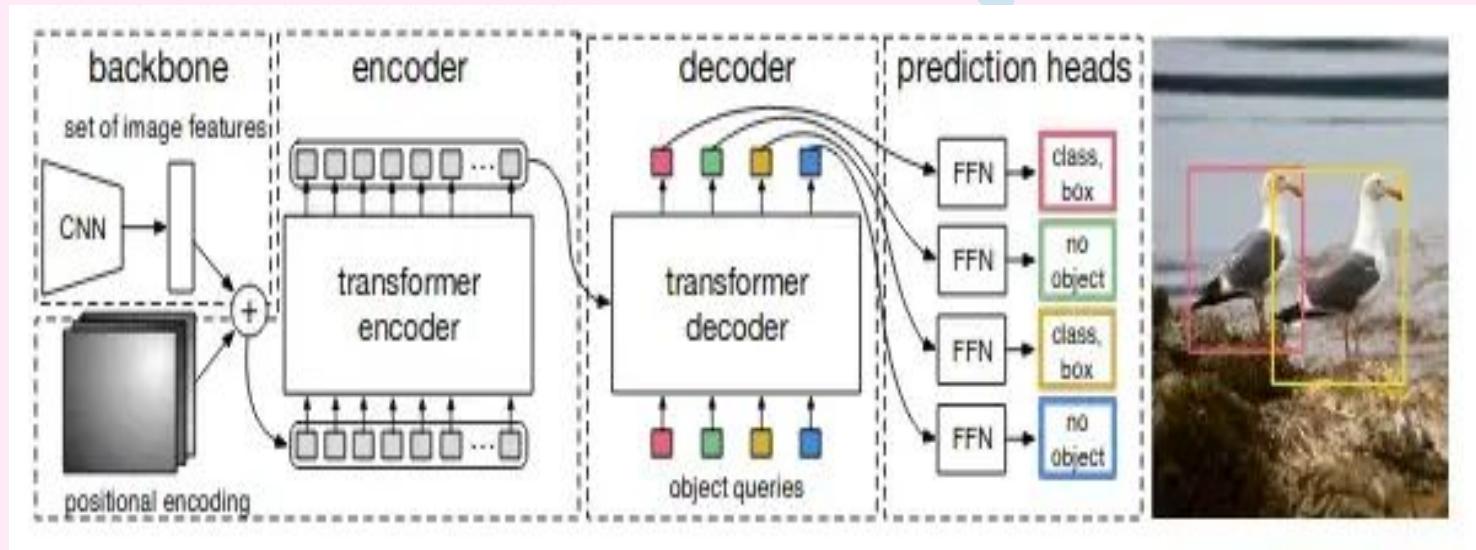


DETR (Detection Transformer) - Layers

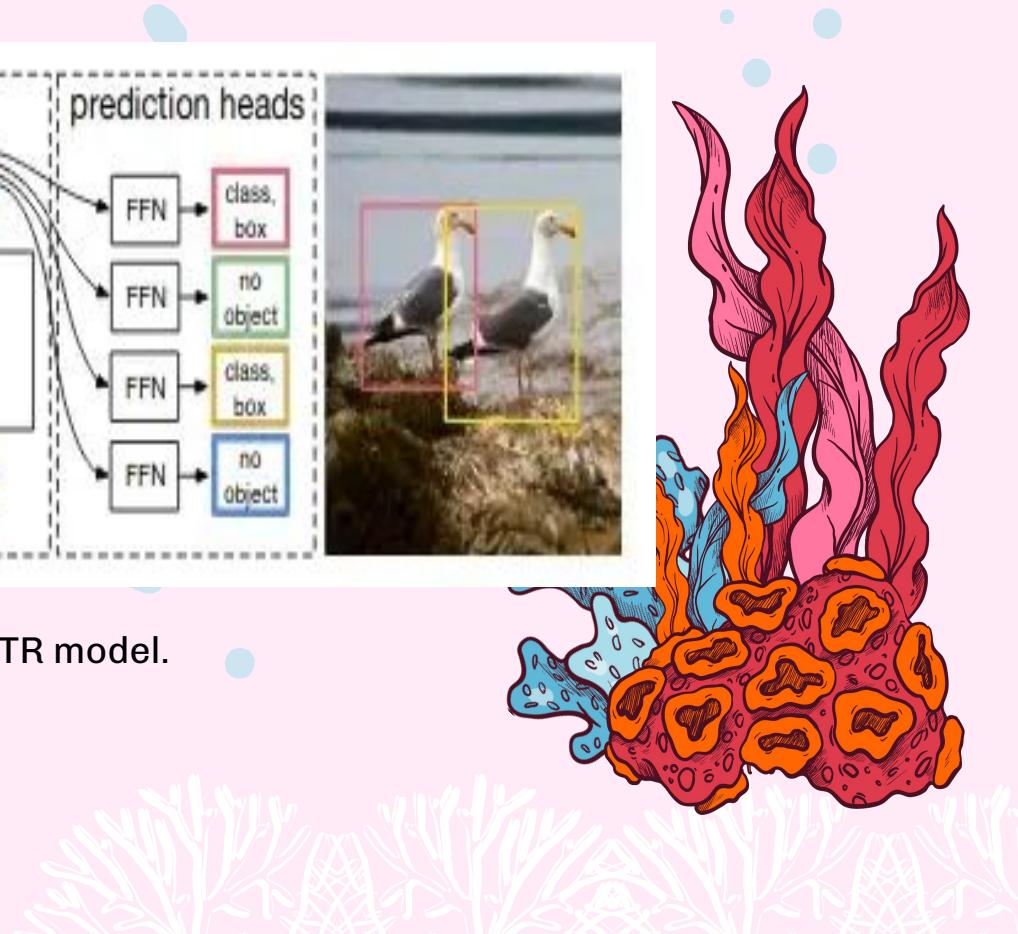
DETR treats an object detection problem as a direct set prediction problem with the help of an encoder-decoder architecture based on transformers.

Transformers are the new breed of deep learning models that have performed outstandingly in the NLP domain.

1. **The Backbone:** Here features extracted from CNN and a positional encoding are passed.
2. **The Transformer Encoder:** The incoming tensors are flattened. It transforms the sequence into an equally long sequence of features.
3. **The Transformer Decoder:** It takes input as object queries, So it's a decoder as a side input for conditioning information.
4. **Prediction Feed-Forward Network (FFN):** The output for this is going through a classifier which further outputs the class labels and bounding boxes as the final output.
5. **Evaluator:** The evaluation is done on **COCO** dataset.



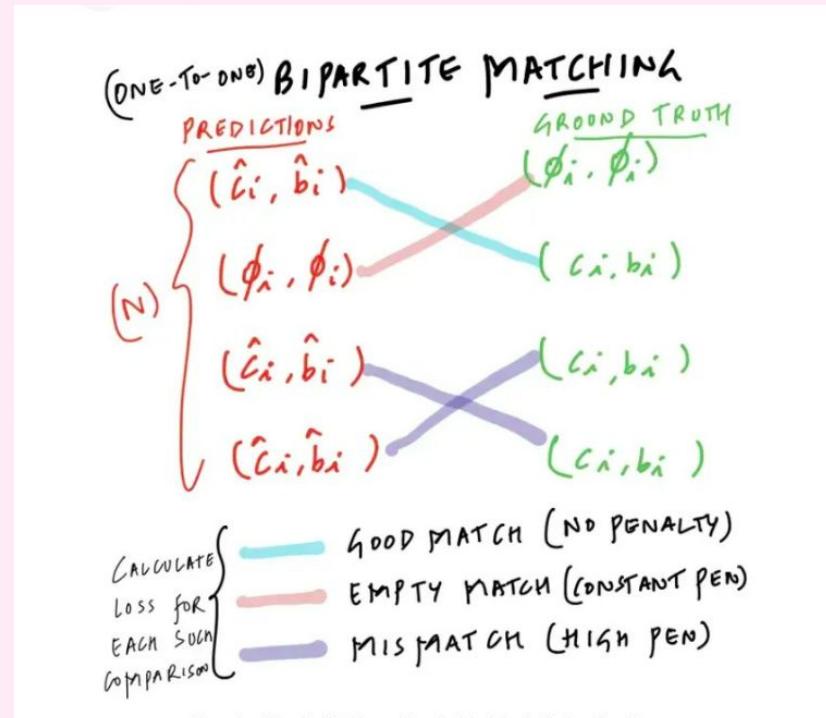
Architecture of DETR model.



Bipartite matching loss

Bipartite matching is done between predicted and ground truth objects.

This ensures one-to-one mapping between predicted and ground truth bounding boxes.

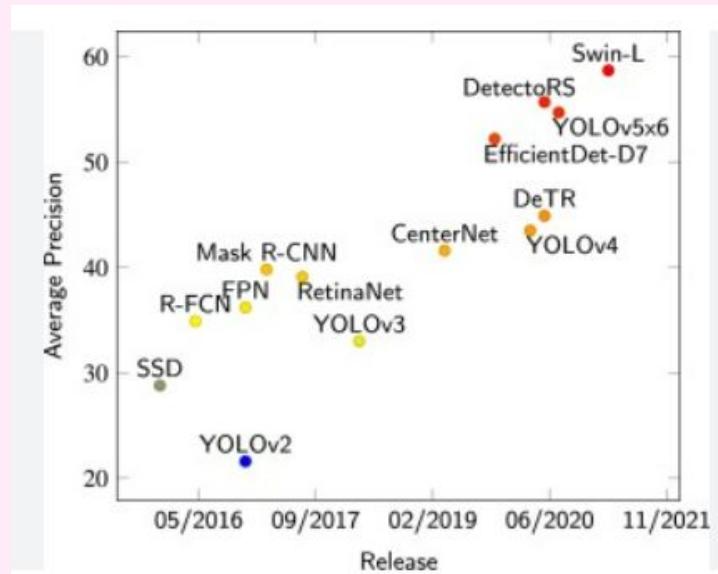


DETR vs YOLO

- Single-stage detecting methods such as YOLO make predictions with respect to anchors or a grid of possible object centers.

In DETR model, we are able to remove this hand-crafted process and streamline the detection process by directly predicting the set of detections.

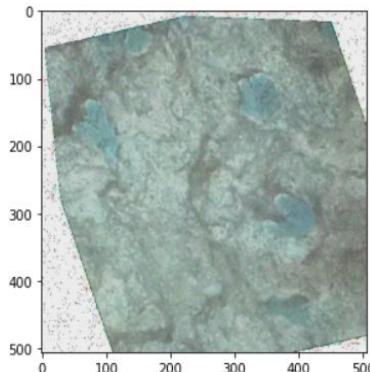
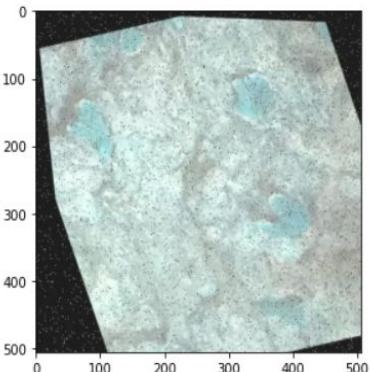
- It makes predictions with absolute box prediction with respect to the input image rather than an anchor.



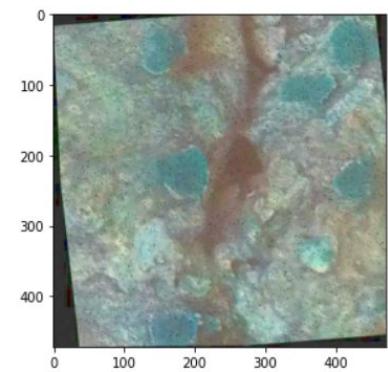
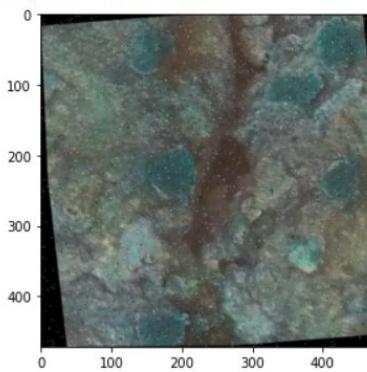
AUGMENTATION

Increasing and decreasing the brightness of images.

mean is 202.12150415721845
brightness decreased 131

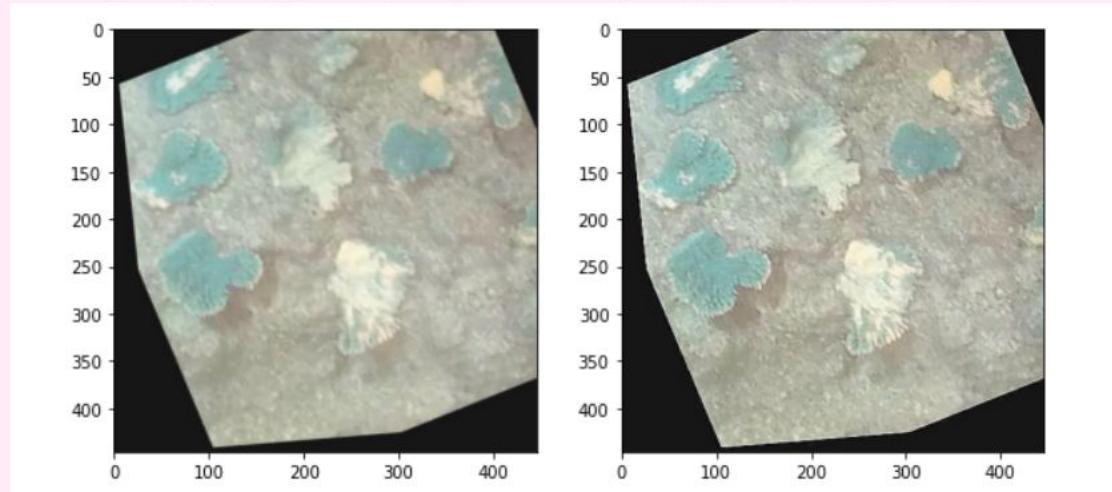


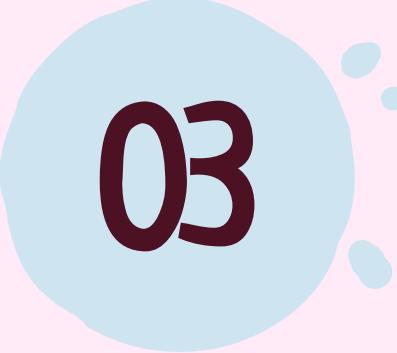
mean is 100.32069328736954
brightness increased 1



AUGMENTATION

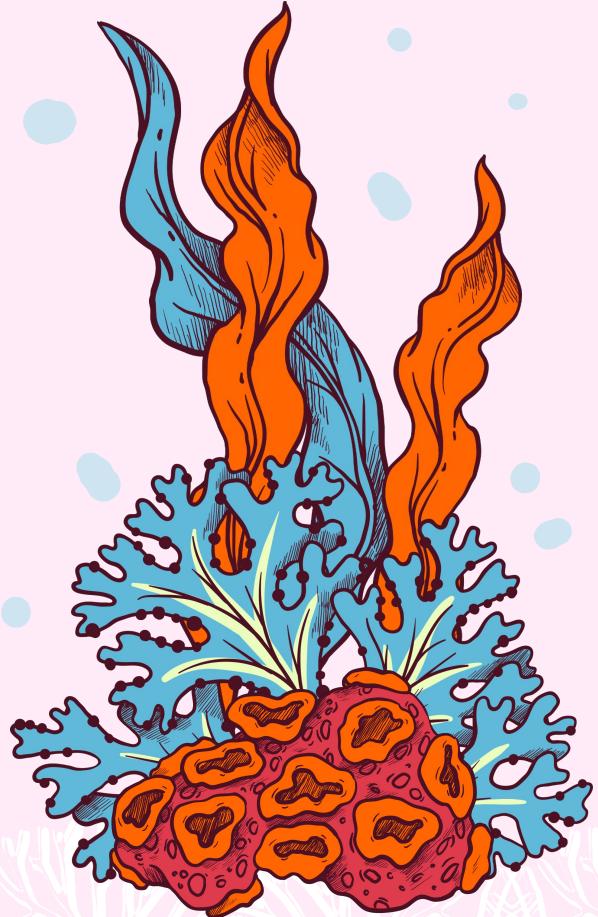
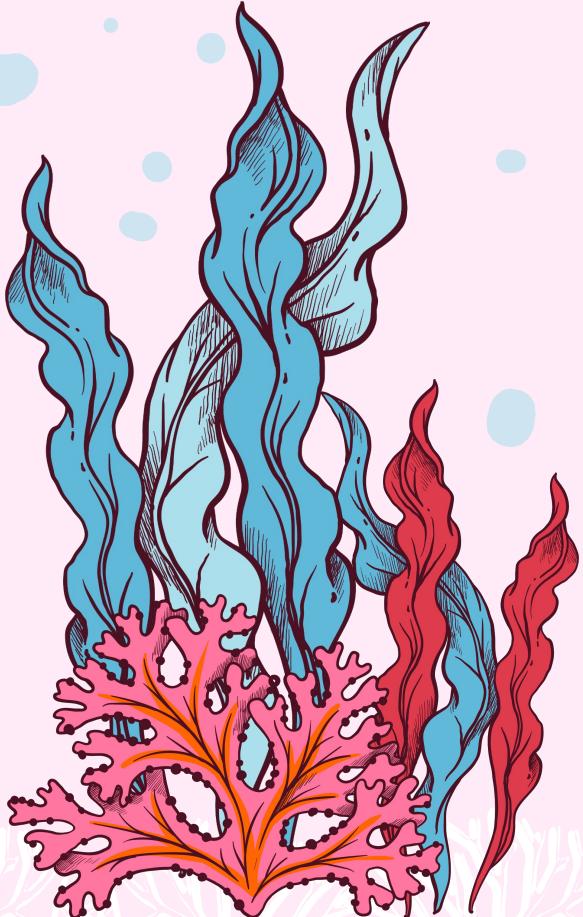
All the images are then sharpened. Resulting to about 5000 images for training.

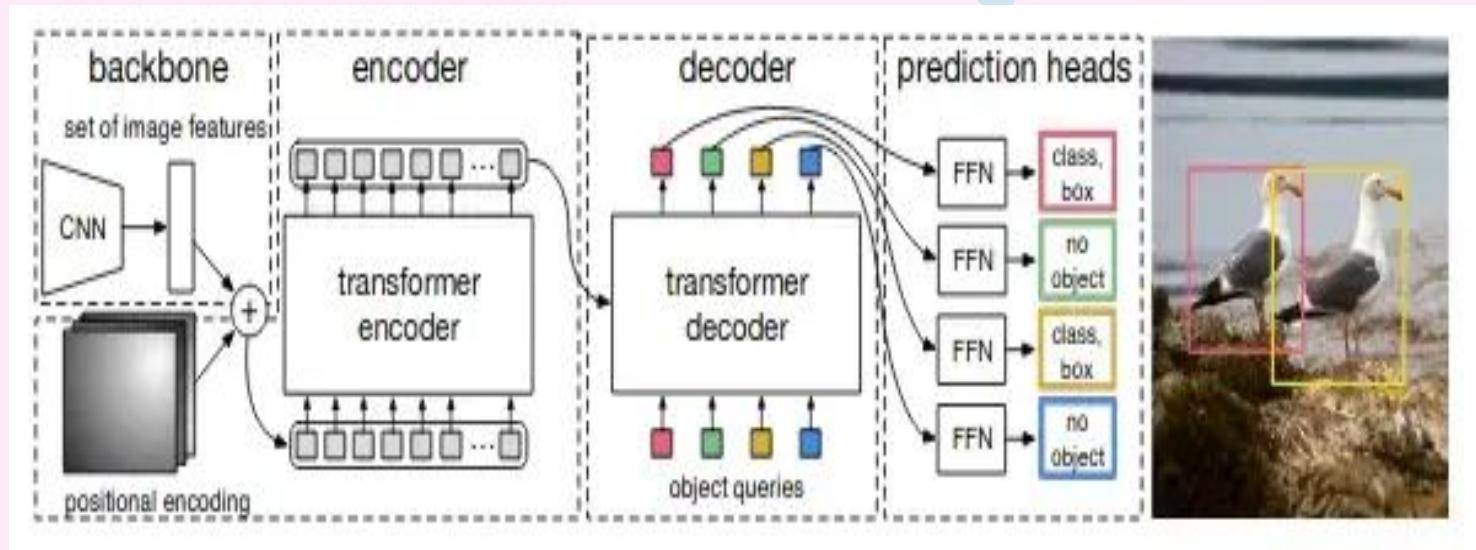




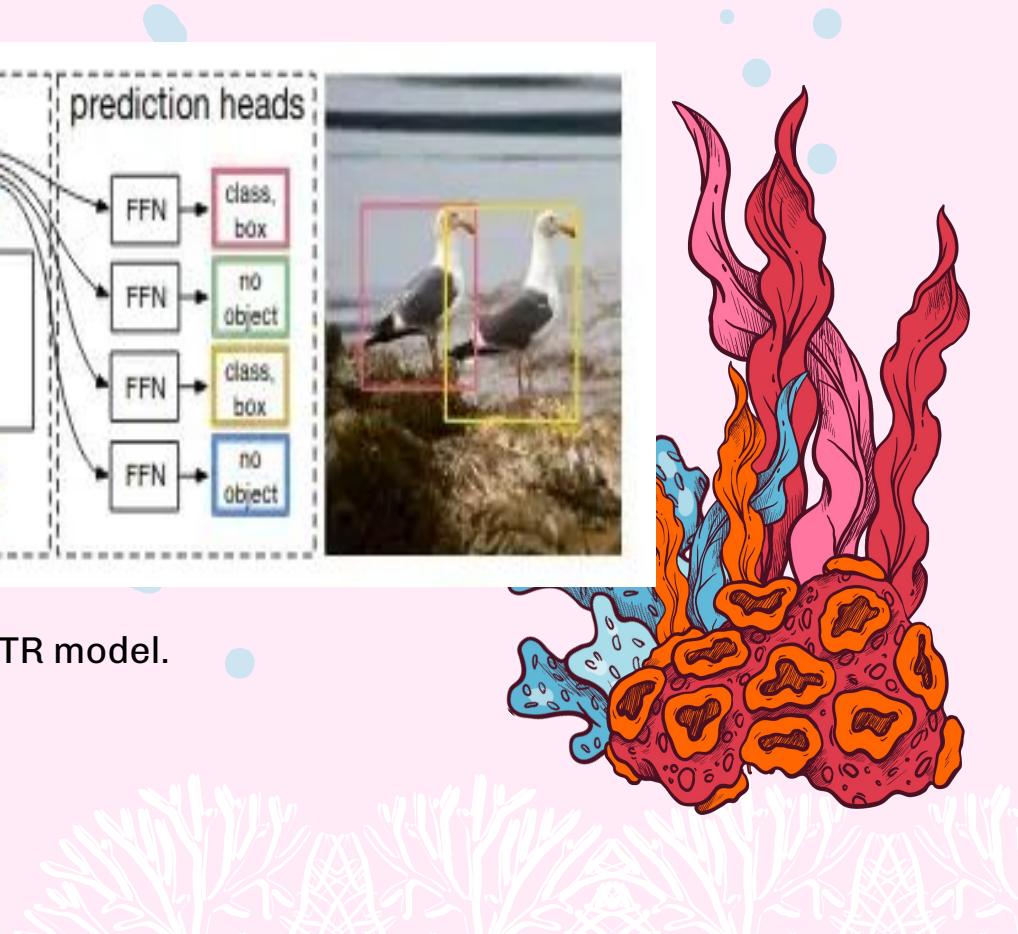
03

Modeling





Architecture of DETR model.



The background features stylized illustrations of marine life. On the left, there's a cluster of blue, flame-like seaweed and a pink, branching coral reef. On the right, there's a cluster of orange, flame-like seaweed and a red, textured coral reef. Small, light blue circular bubbles are scattered across the background.

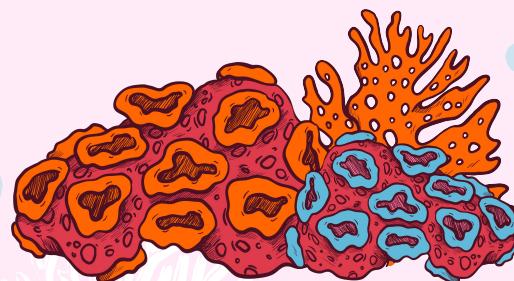
04 Results And Conclusion

IOU - Intersection over Union

Measure of similarity of
2 bounding boxes

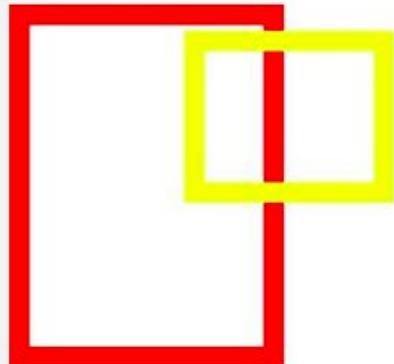
Area of intersection

—
Area of union



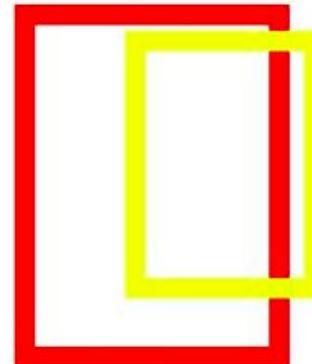
IoU - Range

IoU 0.2



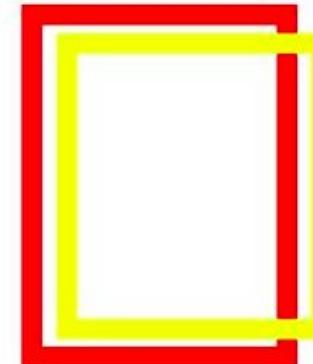
A

IoU 0.5

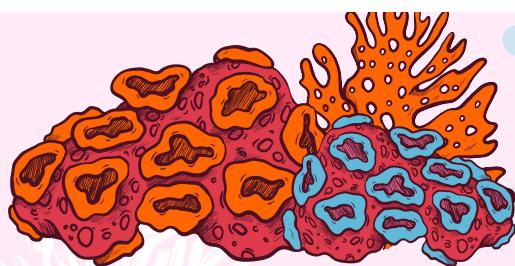
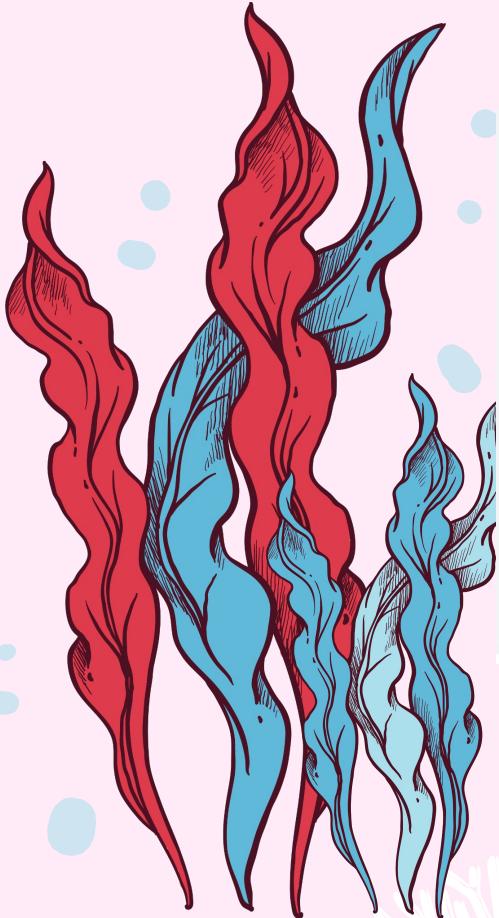


B

IoU 0.9



C



Overlapping bounding boxes



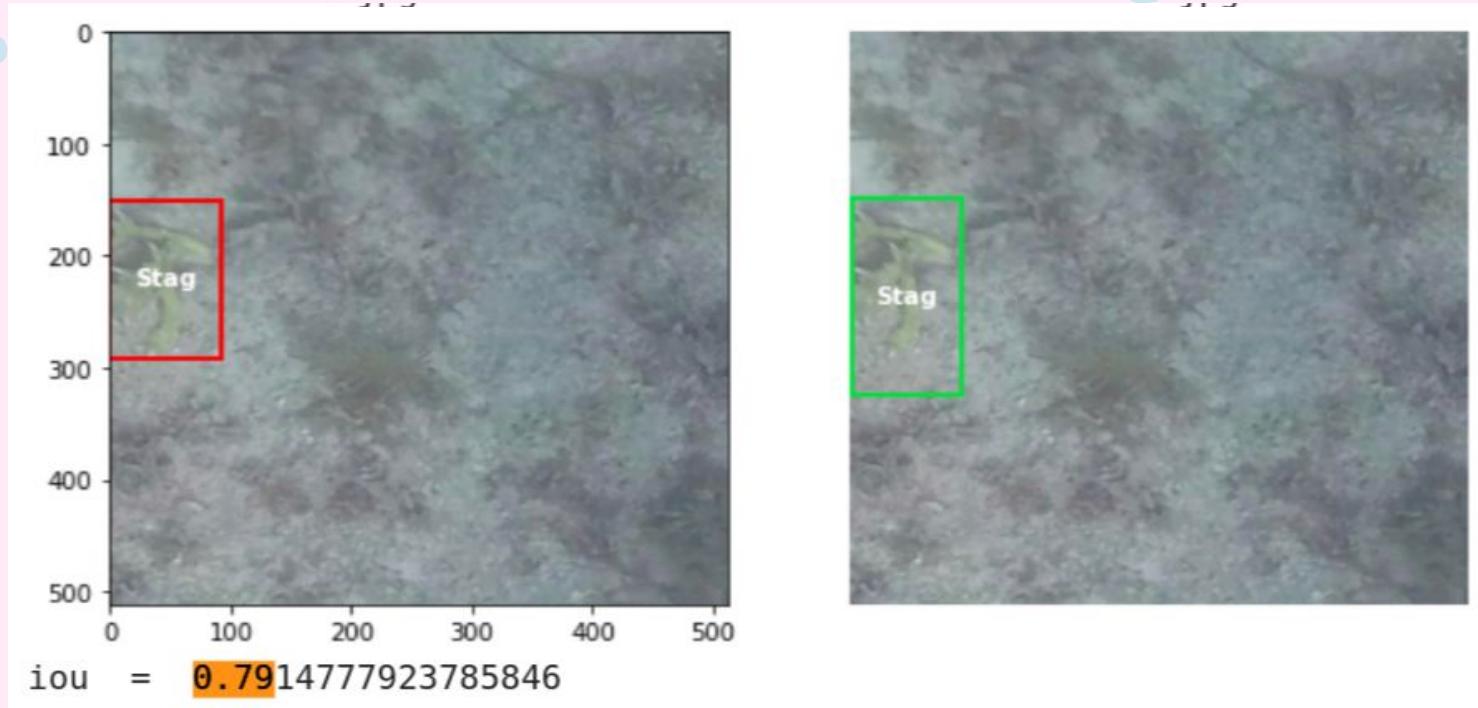
NON - Overlapping bounding boxes



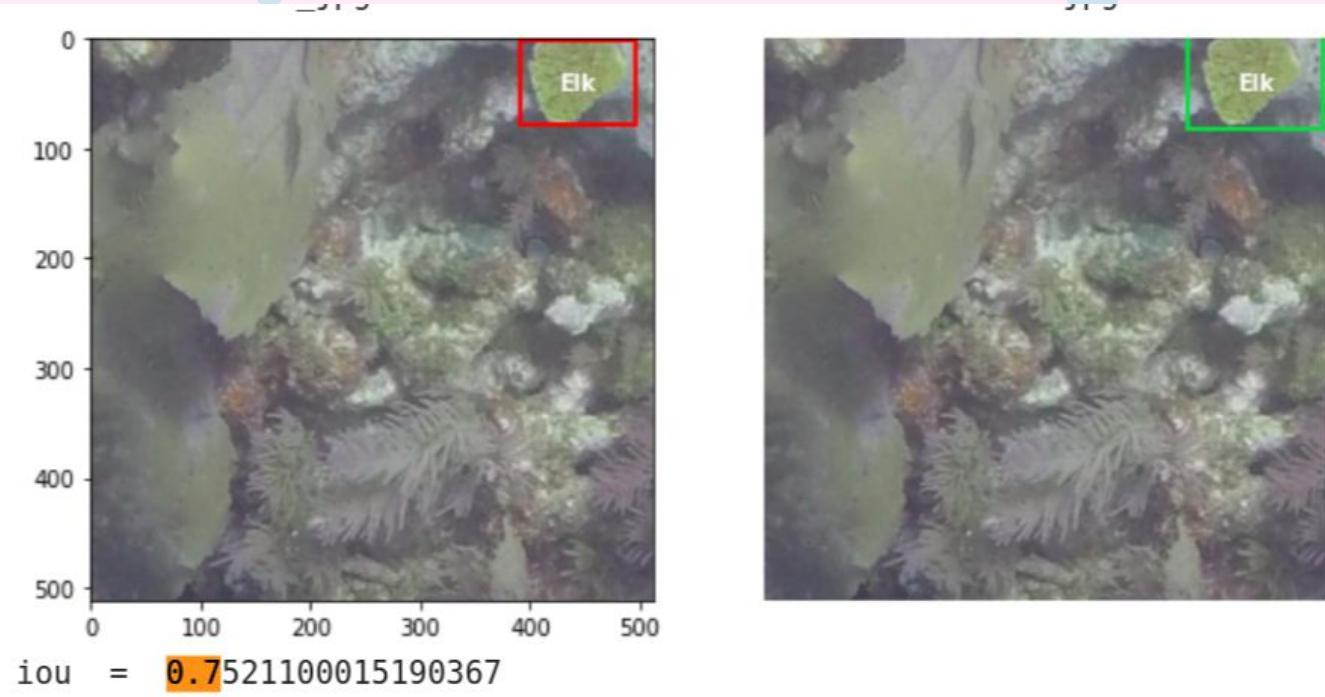
NON - Overlapping bounding boxes



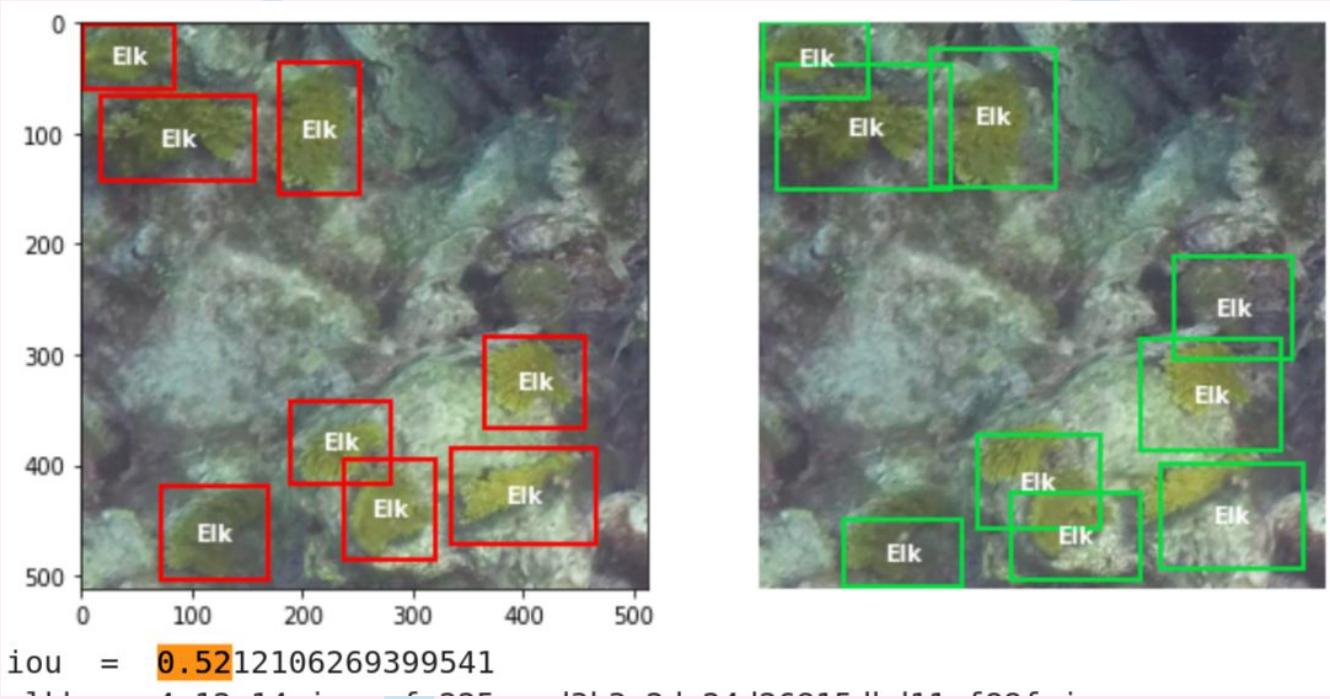
Excellent Match - IOU > 0.70



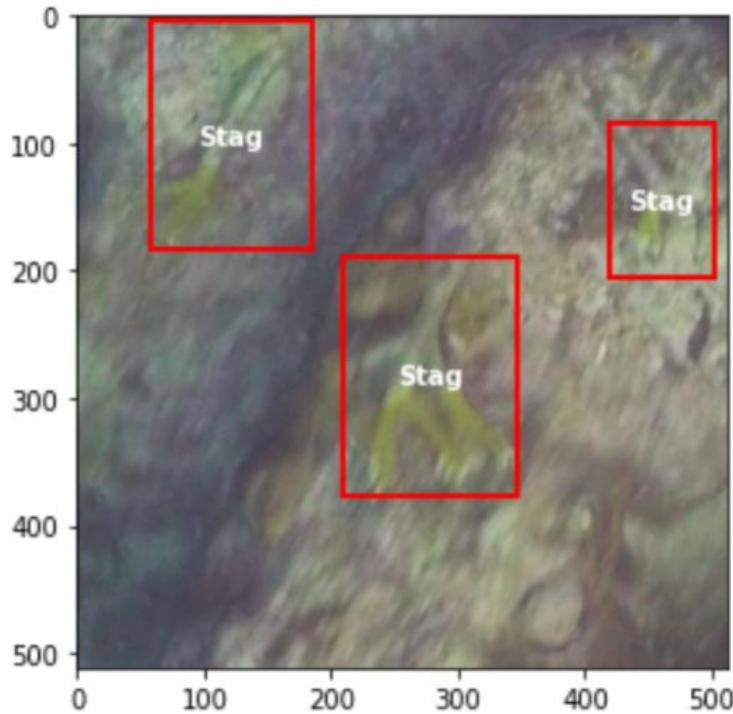
Excellent Match - IOU > 0.70



Good Match - IOU > 0.45

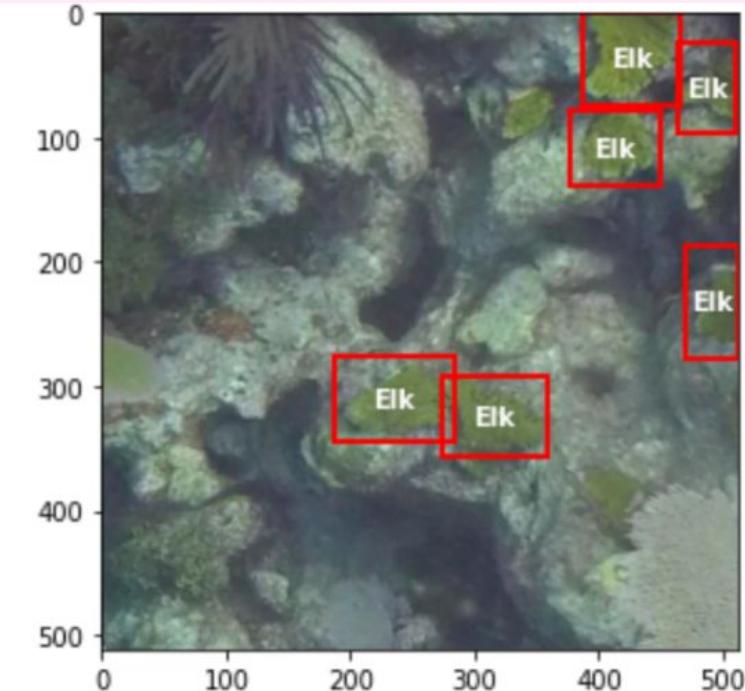


Good Match - IOU > 0.45



iou = 0.44241953349771274

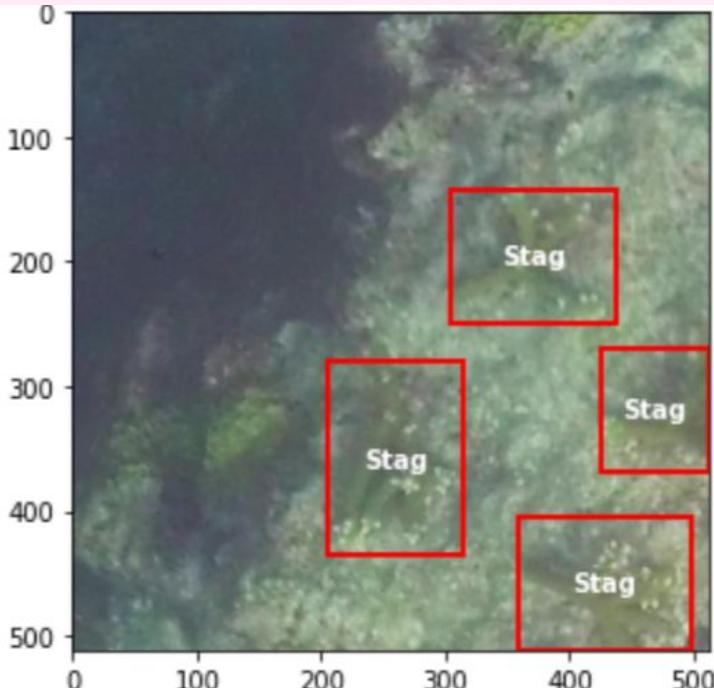
Poor Match - IOU < 0.45



iou = 0.22664389175364041

elkhorn-4-17-12.jpg.rf.d3f535f601fd7dcf9f300e63e0aa2b21.jpg

Poor Match - IOU < 0.45



iou = 0.17136727605768196

staghorn-1-31-16 jpg.rf.b26ff310f189b603ccbde48867563ff0.jpg

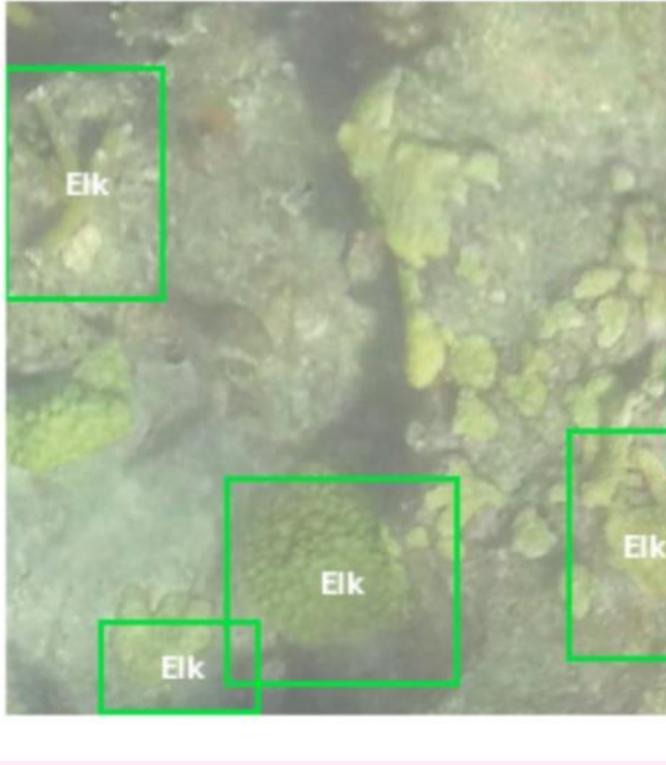
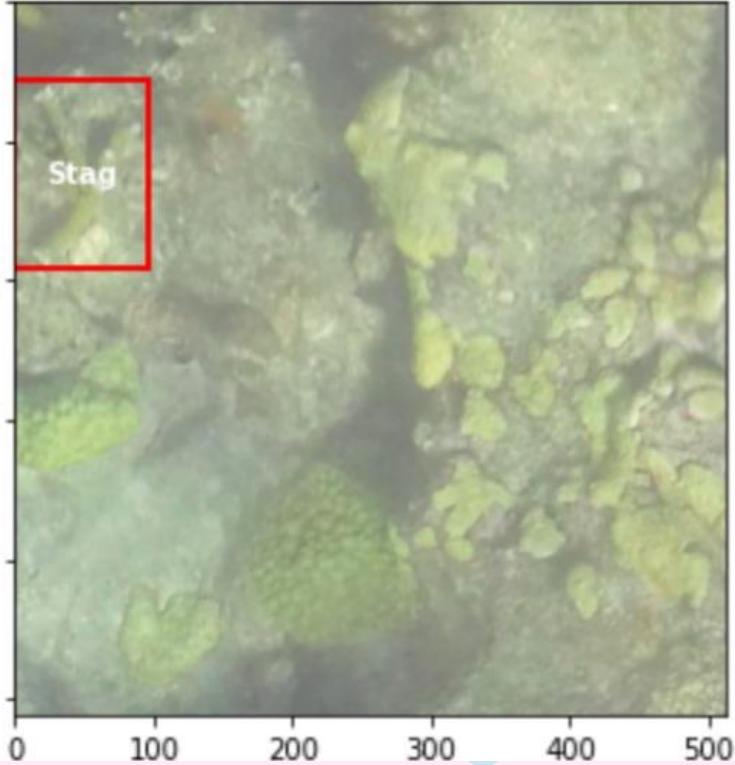


Average IOU

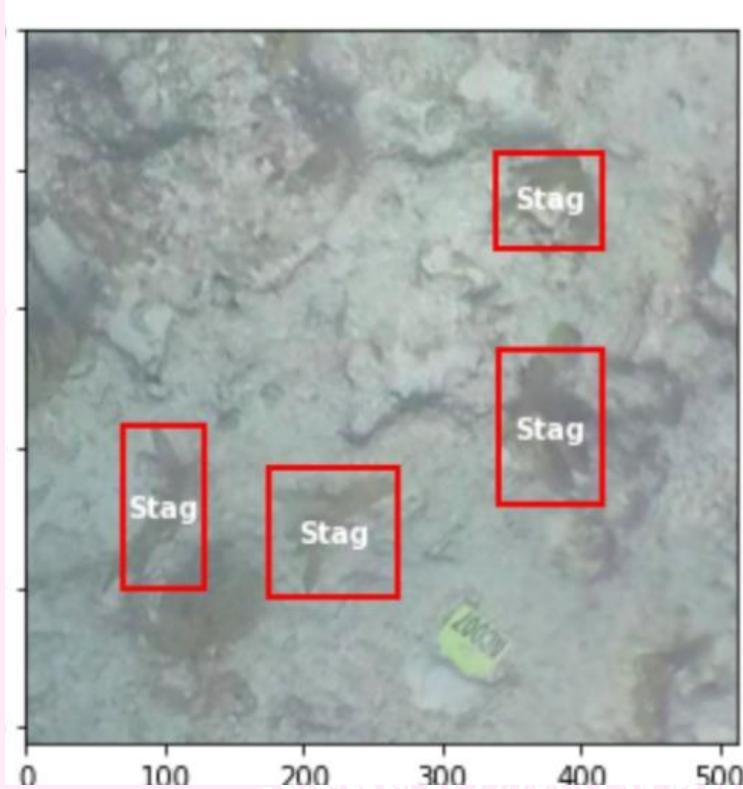
0.45



Coral Type Prediction



Coral Type Prediction

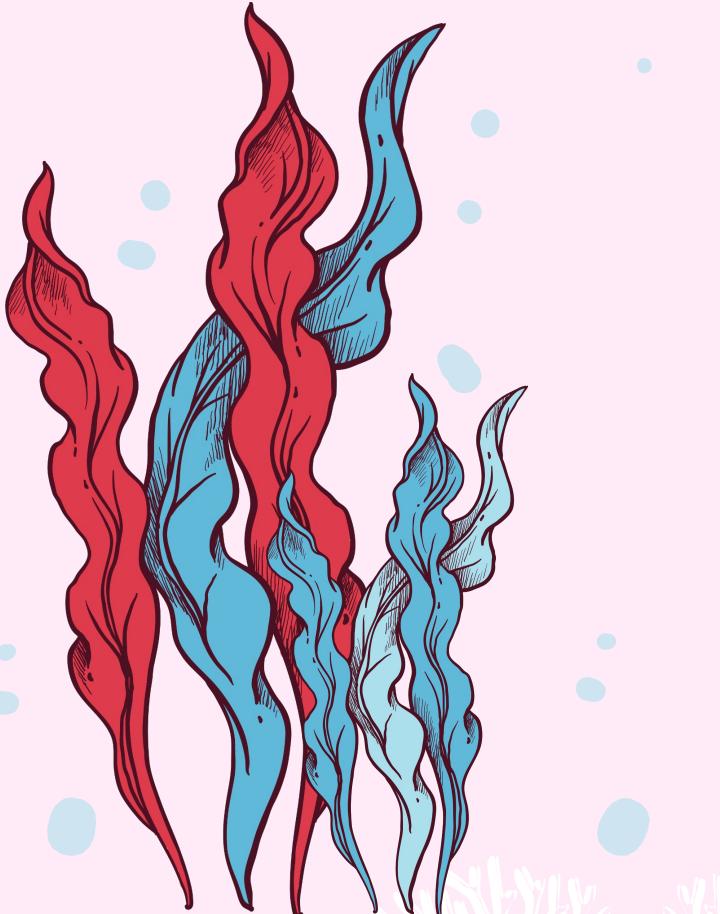


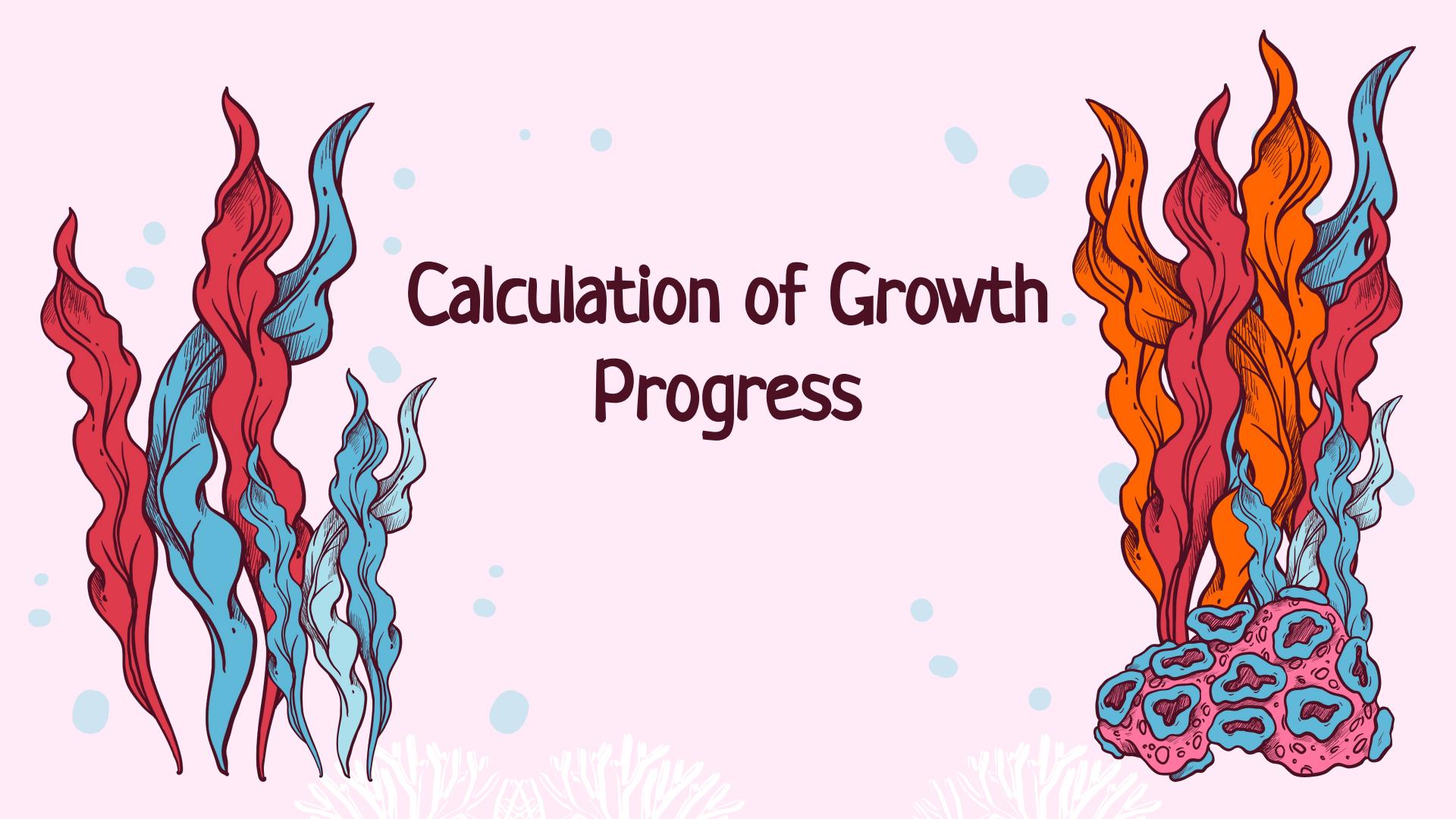
Prediction accuracy

213

215

0.99%





Calculation of Growth Progress

STAGE - 1



elkhorn-5-33-17.jpg.rf.
b988b45e5e12ecb76e395
db7fde6bcc3.jpg



staghorn-2-6-40.jpg.rf.
898ebfdfffeb3105f90b8d68
a29b01378.jpg

STAGE - 2



elkhorn-2-49-10.jpg.rf.
730725103897c7aaade039
b686897a7fd.jpg



staghorn-2-19-19.jpg.rf.
4e099ab3cde0e8b1f3042
284d08ea73b.jpg

Growth Percentage

STAGE 1 : Area of Elkhorn 58721.73

STAGE 2 : Area of Elkhorn 526335.37

STAGE 1 : Area of Staghorn 181427.21

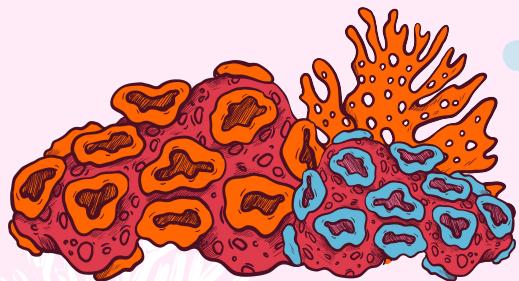
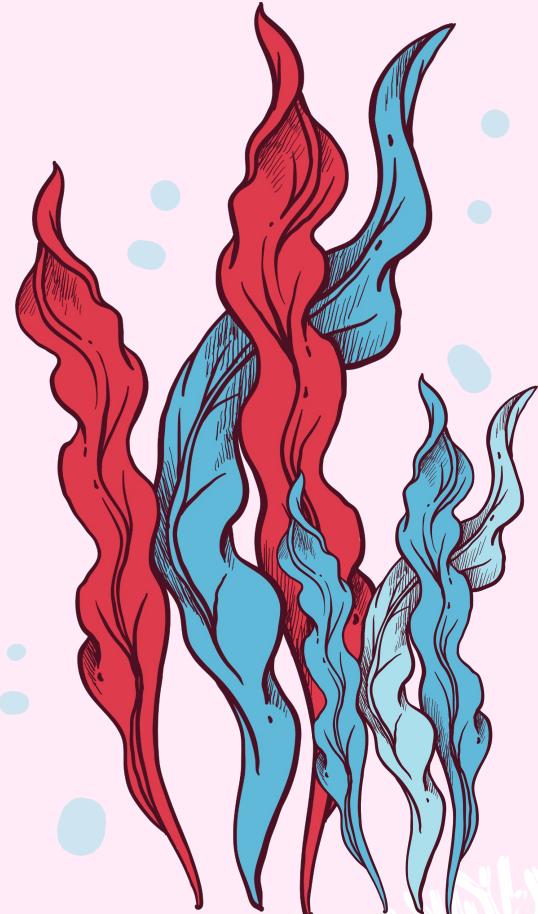
STAGE 2 : Area of Staghorn 690954.29

Growth percentage of Elkhorn 7.96 %

Growth percentage of Staghorn 2.81 %

Future Work

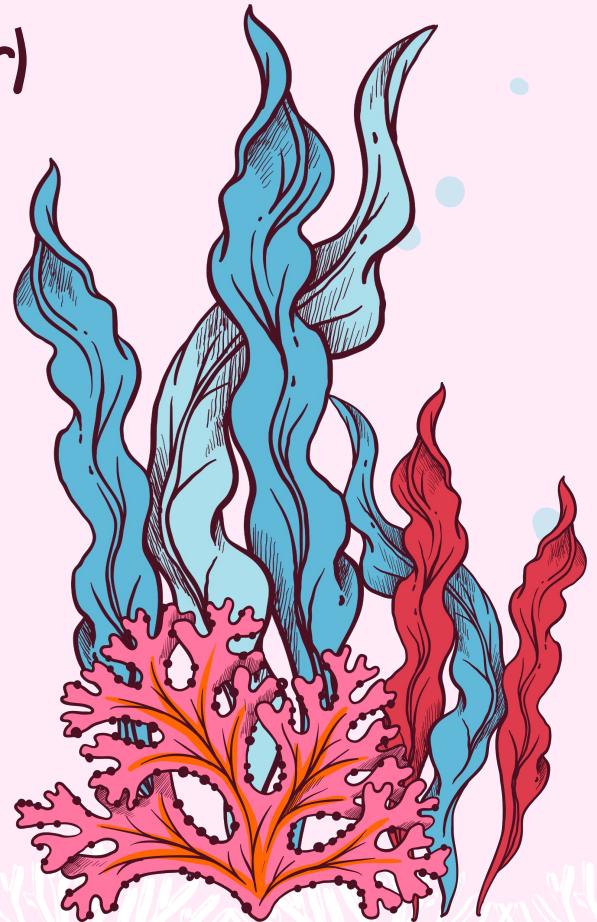
Deformable DETR

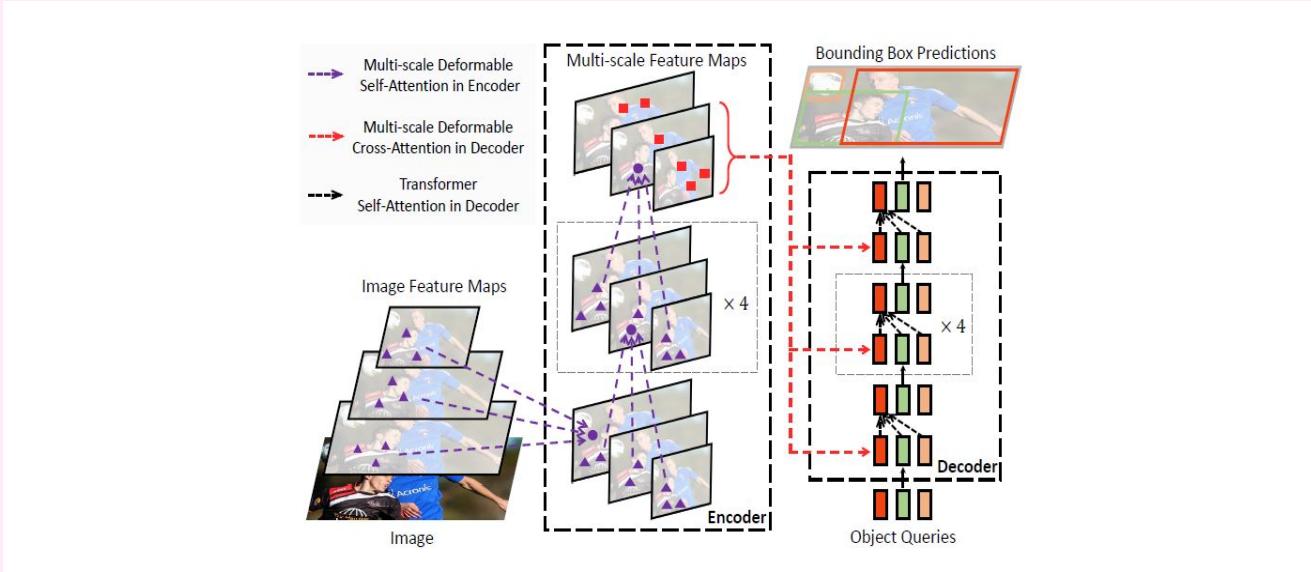


Deformable-DETR (Detection Transformer)

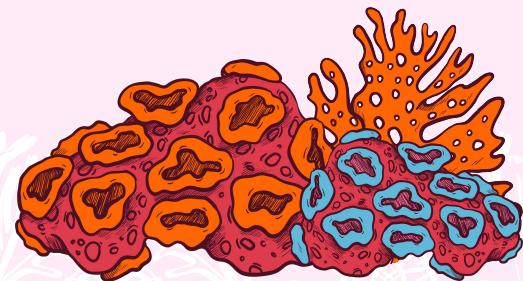
- This Deformable DETR is used to overcome the slow convergence and high complexity issues of DETR. Deformable-DETR combines the best of sparse spatial sampling of deformable convolution, and also has the capability of Transformers relational modeling.

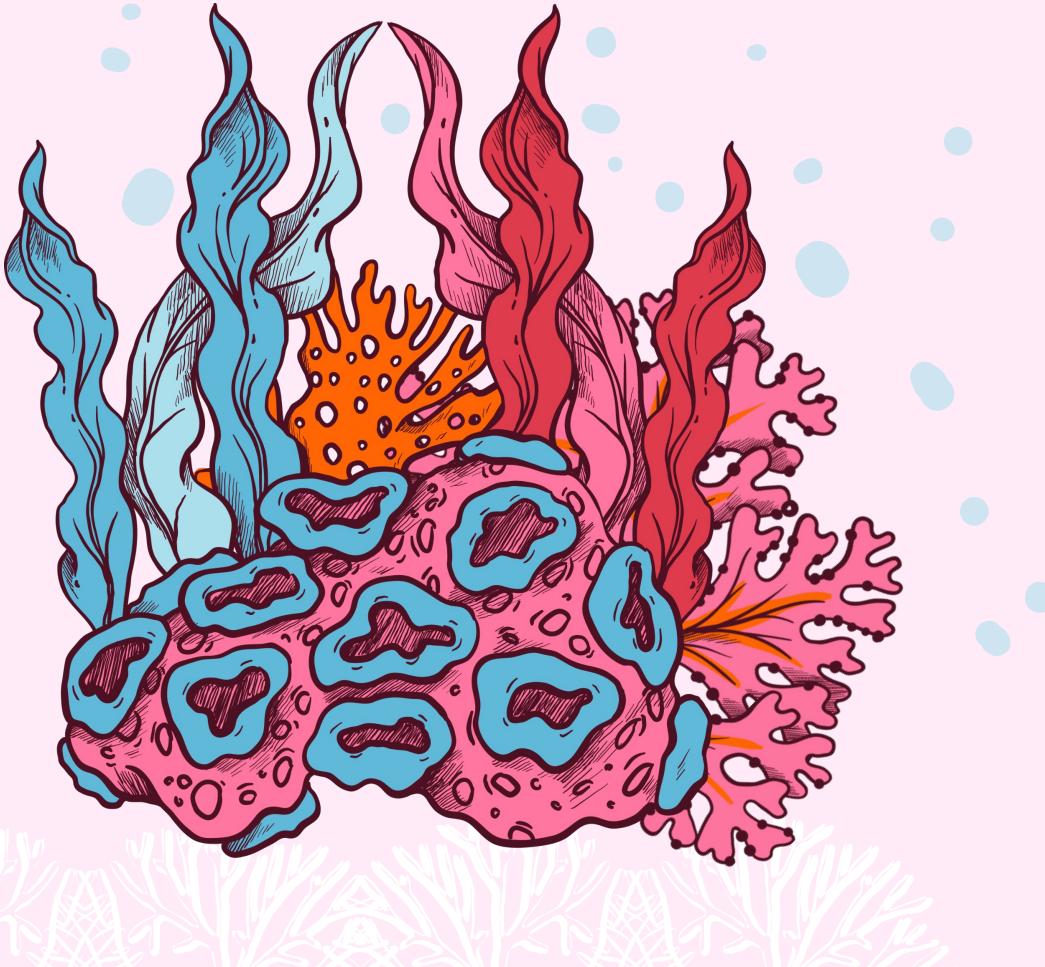
1. Deformable attention module.
2. Multi-scale deformable attention module.
3. Deformable transformer encoder.





**Deformable DETR - Deformable Transformers
for end-to-end object detection.**





Questions



Thank You!

A detailed illustration of a coral reef in shades of red, orange, and blue. The coral has various shapes, from long, sweeping tentacles to shorter, more rounded clusters. It is set against a light pink background with scattered light blue circles of different sizes.A detailed illustration of a coral reef in shades of orange, red, and blue. The coral has various shapes, from long, sweeping tentacles to shorter, more rounded clusters. It is set against a light pink background with scattered light blue circles of different sizes.