

Underwater Coral Imaging

Presented By:

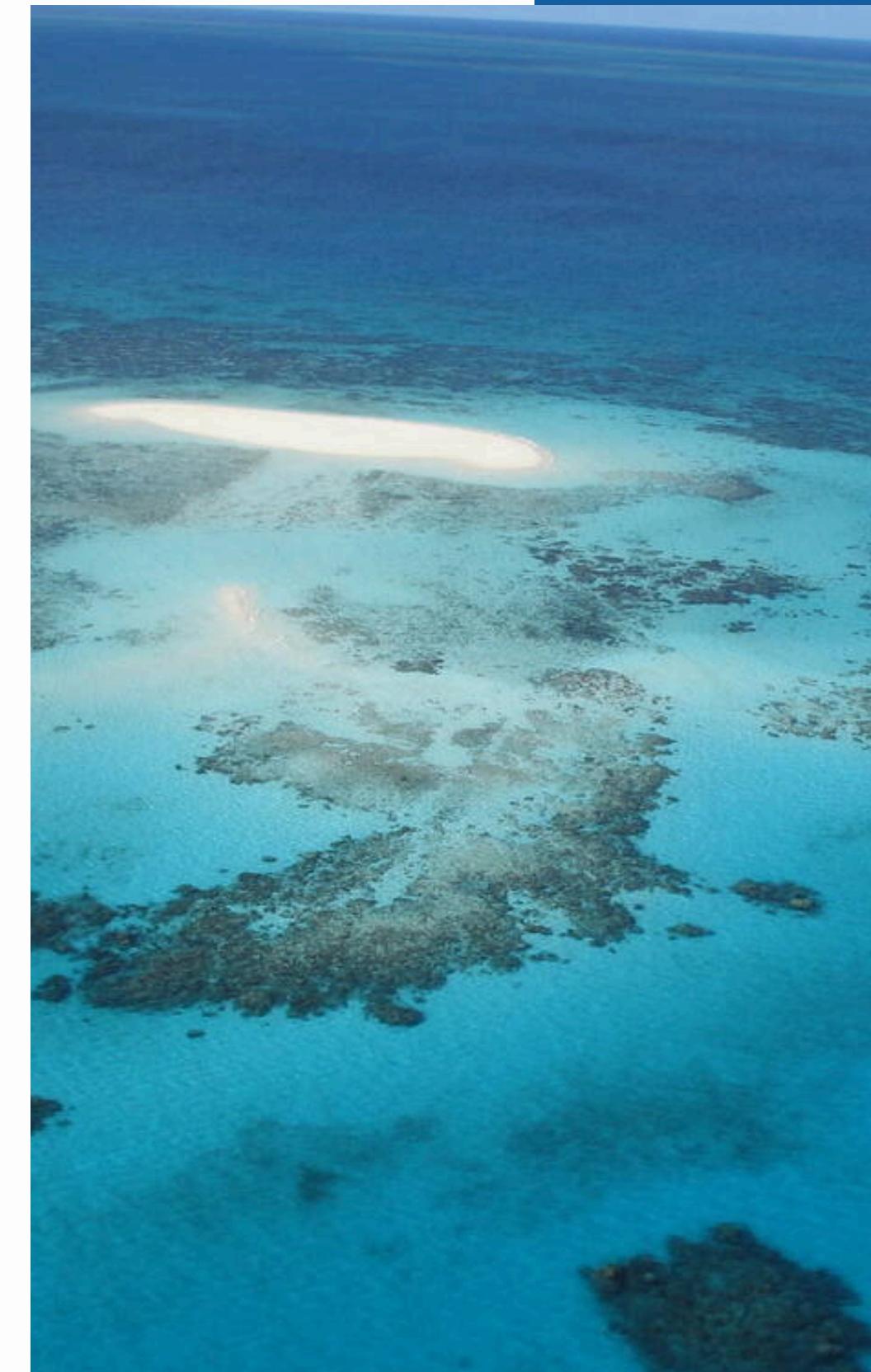
Ayushman Baghel
Prakhar Gupta

Supervised By:

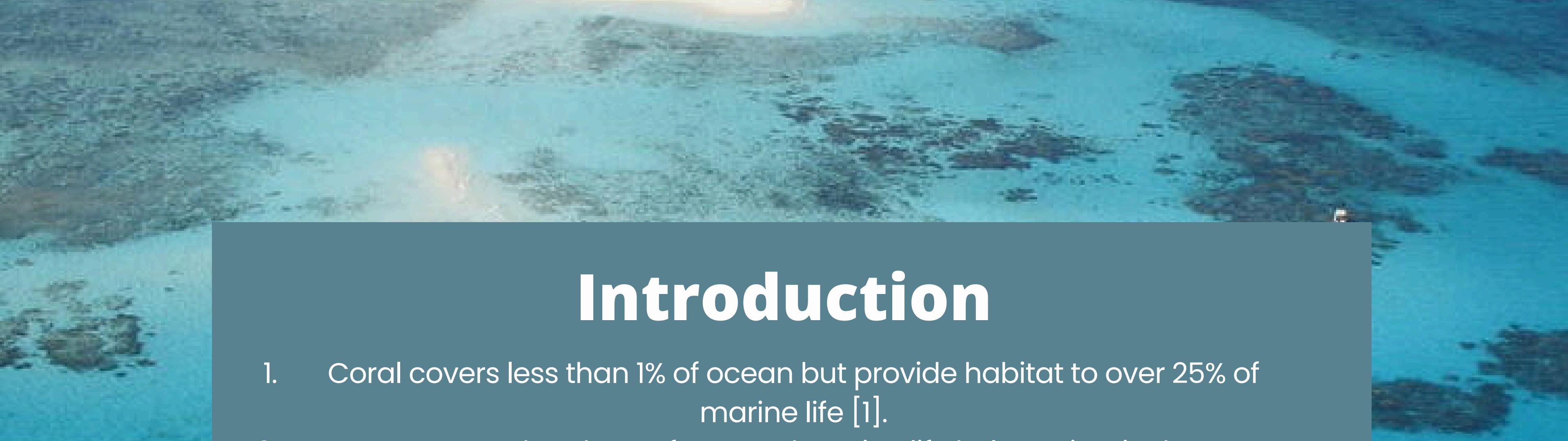
Dr. Shitala Prasad

Overview

▶ Introduction	01
▶ Literature Review	02
▶ Challenges	03
▶ Datasets	04
▶ Models	04
▶ Results	05
▶ Discussion	06
▶ Conclusion	07
▶ References	08



Great Barrier Reef



Introduction

1. Coral covers less than 1% of ocean but provide habitat to over 25% of marine life [1].
2. Oceans are net absorbers of CO₂ and marine life help maintain the ecosystem.
3. Due to global warming and acidification of oceans these ecosystem are increasingly stressed which result in imbalance of their ecosystem.
4. Monitoring and analysis of these ecosystems are required for identifying potential threats and to enable timely intervention.
5. These monitoring and analysis done by domain experienced diver are expensive, time consuming and resource intensive.

Literature Review

Toward Highly Accurate Coral Texture Images Classification [2] review

1. MultiClass Dataset: EILAT and RSMAS.
2. Using various CNNs Architectures mainly:
 - a. ResNet 50
 - b. ResNet 152
 - c. Inception v3
 - d. DenseNet-121
 - e. DenseNet-161
3. EILAT: 8 Classes
4. RSMAS: 14 Classes

Our Aim:

- 1. Larger Dataset,**
- 2. More Classes,**
- 3. Bigger Models,**
- 4. New Architectures**
- 5. Extend to Detection**

Towards Highly Accurate Coral Texture Images Classification Using Deep Convolutional Neural Networks and Data Augmentation

Anabel Gómez-Ríos^{a,*}, Siham Tabik^a, Julián Luengo^a, ASM Shihavuddin^b, Bartosz Krawczyk^c, Francisco Herrera^a

^aDept. of Computer Science and Artificial Intelligence, University of Granada, C/ Periodista Daniel Saucedo Aranda, s/n, 18071, Granada, Spain

^bDept. of Applied Mathematics and Computer Science, Technical University of Denmark (DTU), Kgs. Lyngby, Denmark

^cDept. of Computer Science, Virginia Commonwealth University, USA

Abstract

The automatic identification of coral species based on underwater texture images is a challenging task due to the difficulty for machine learning algorithms, due to the inherent characteristics of the data. These characteristics are deeply embedded in the nature of this data: 1) dataset size; 2) the complex global structure of the coral; 2) the variability of the textures; and 3) defining the boundaries between different coral species.

Challenges

Problem 01

Under water images are riddled with visual distortion, blur, occlusion, lightening variation which significantly effect learning.

Problem 02

Coral Species although can be visual identified, many coral species need electron microscope for 100% accurate identification.

Problem 03

Even among coral, many species look visual similar, and need to be identified on close range for accurate prediction.

These challenges requires high quality image taken in close range.

Images might be needed to be preprocessed heavily for restoration of hue and lightening.

Domain Specialist for accurate annotation of dataset.

Datasets

We contributed 3 datasets each for specific task with aim to help formulate standard coral dataset. We have manually labelled (some), augmented, cleaned and combined multiple datasets.



Binary Classification

Dataset for classifying whether an image is coral or not.

Coral: 32,187

Non-Coral: 27,887

Total: 60,074



Multiclass Classification

Dataset for classifying coral species in different categories.

27 coral species classes.

1 Marine class

1 Other class



Coral Detection

Dataset for detection underwater coral.

Binary Detector

Binary Classification Dataset

625x625

3 Channels

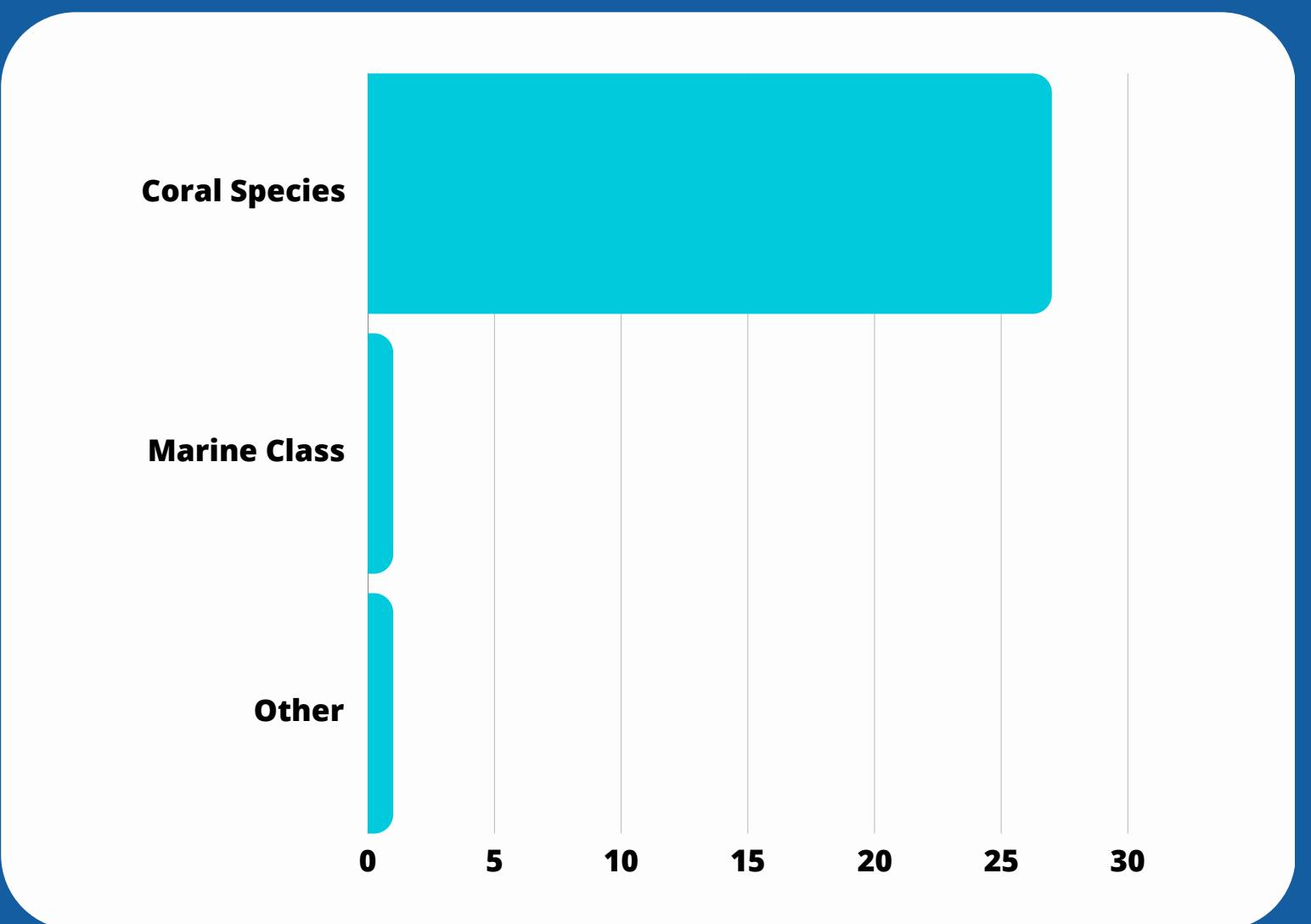
Coral	32,187
Non Coral	27,887
Total	60,074



Augmentation by Rotation, Horizontal Flip and Brightness adjustment

Multiclass Classification Dataset

Total 29 Classes



Coral Detection Dataset

Training set: 49,533

Test set: 3,081

Valid set: 4,372

Total: 56,986

Models

For Models we used 2 types of Architectures specifically CNN and ViT.

Each Architecture type has its advantage and disadvantage, each its learning paradigm.

01

ResNet-50

02

EfficientNet

03

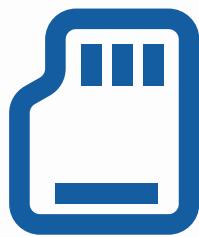
ViT-G

04

CoCa ViT

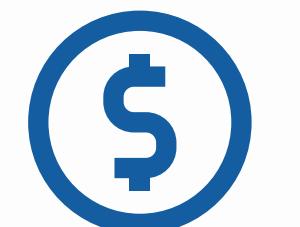
ResNet-50

Introduces Skip/Residual Connections for training deeper networks



Layers

- Consists of 50 convolution/pool layers.
- 4 Layered Structure with each layer containing multiple bottleneck blocks.
- Each block contains convolution.



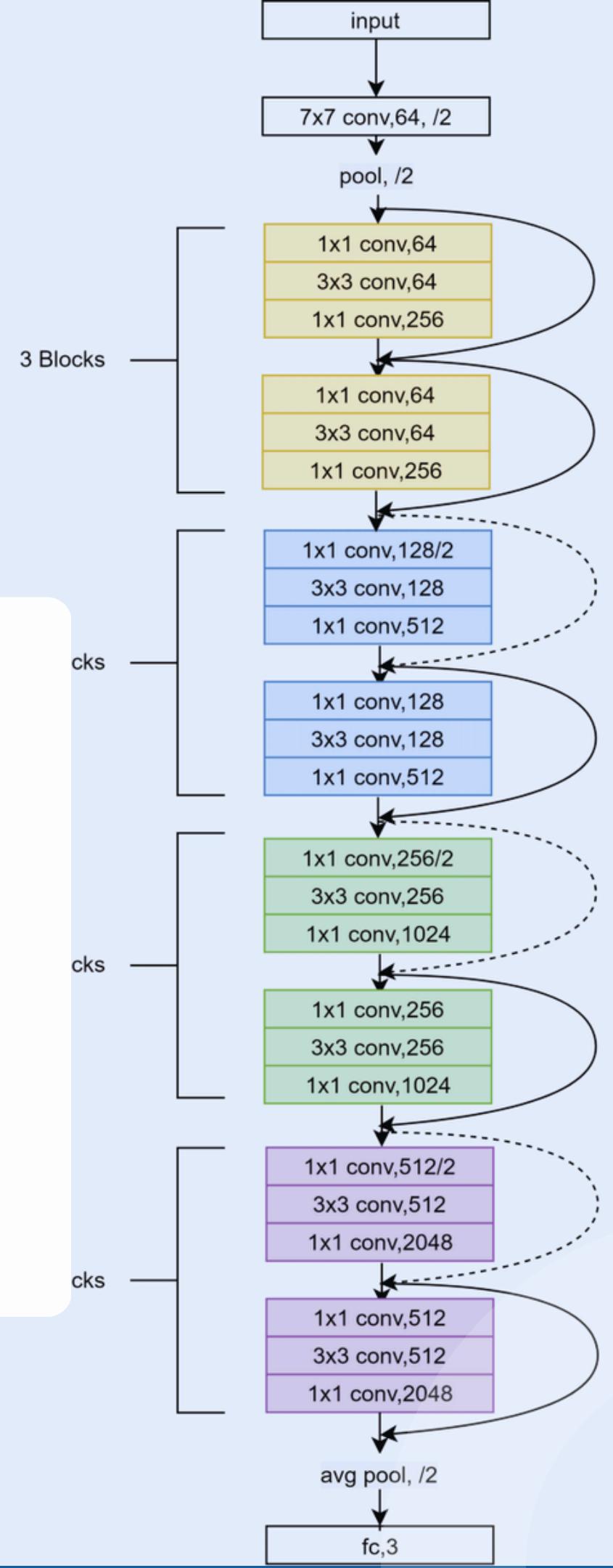
Parameters

25.6 million trainable parameters



Architecture

CNN based Architecture



*Image source [3]

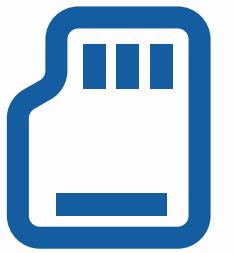
EfficientNet

Introduces Compound Scaling Method for efficient scaling

depth: $d = \alpha^\varphi$, width: $w = \beta^\varphi$, resolution: $r = \gamma^\varphi$

$$\text{st } \alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$$

$$\alpha \geq 1, \beta \geq 1, \gamma \geq 1$$



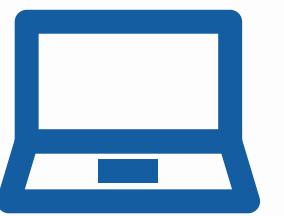
Depth Multiplier

B4	1.8
B5	2.2
B6	2.6
B7	3.1



Width Multiplier

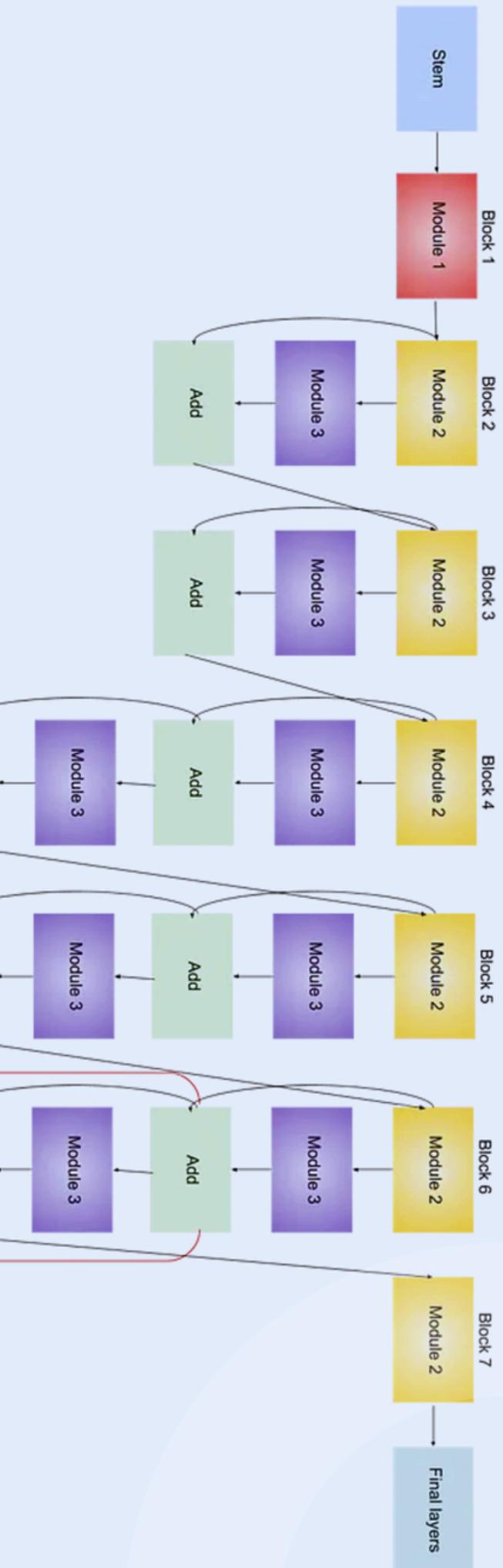
1.4
1.6
1.8
2.0



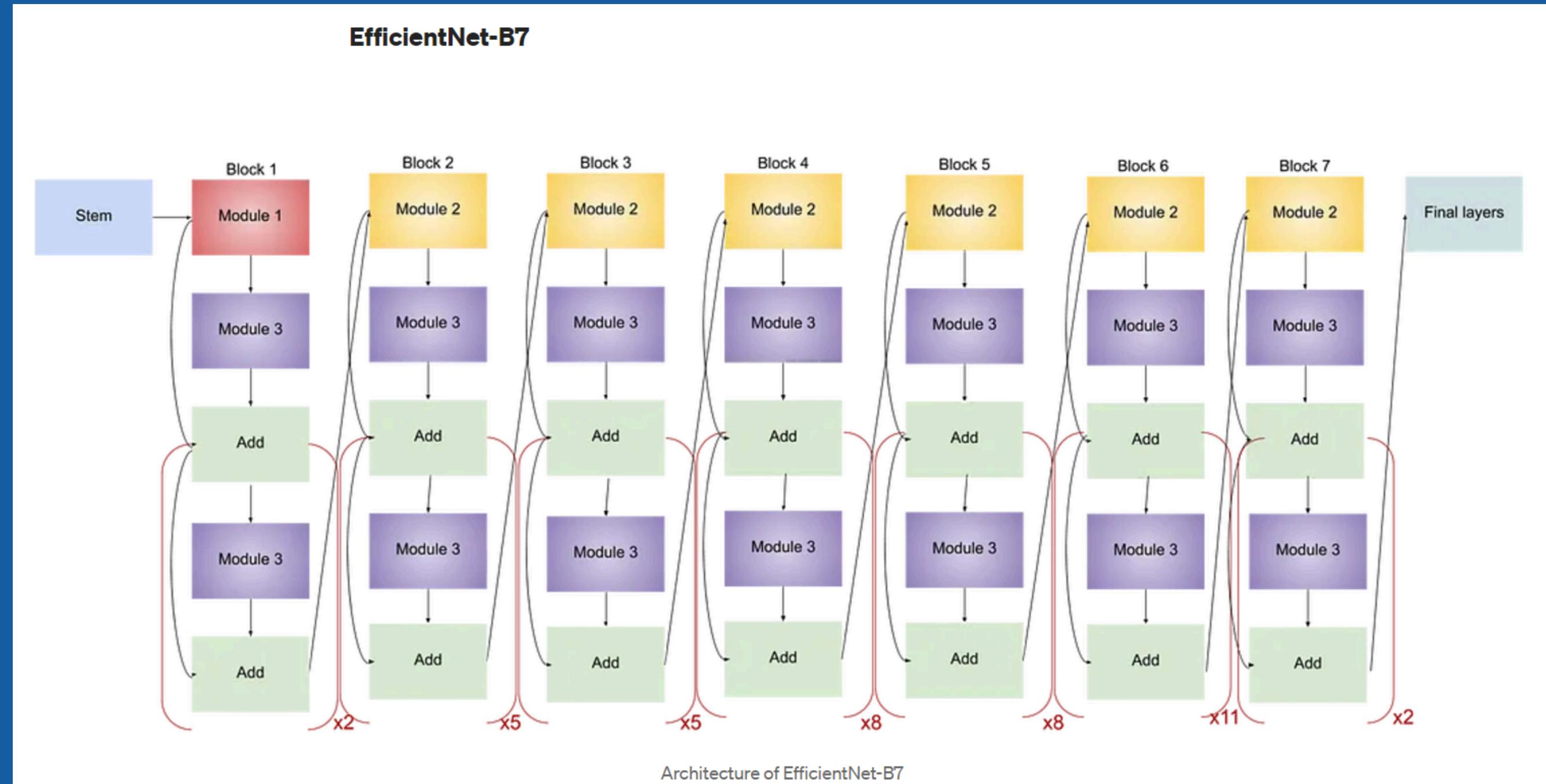
Resolution

380x380
456x456
528x528
600x600

*Image source [4]



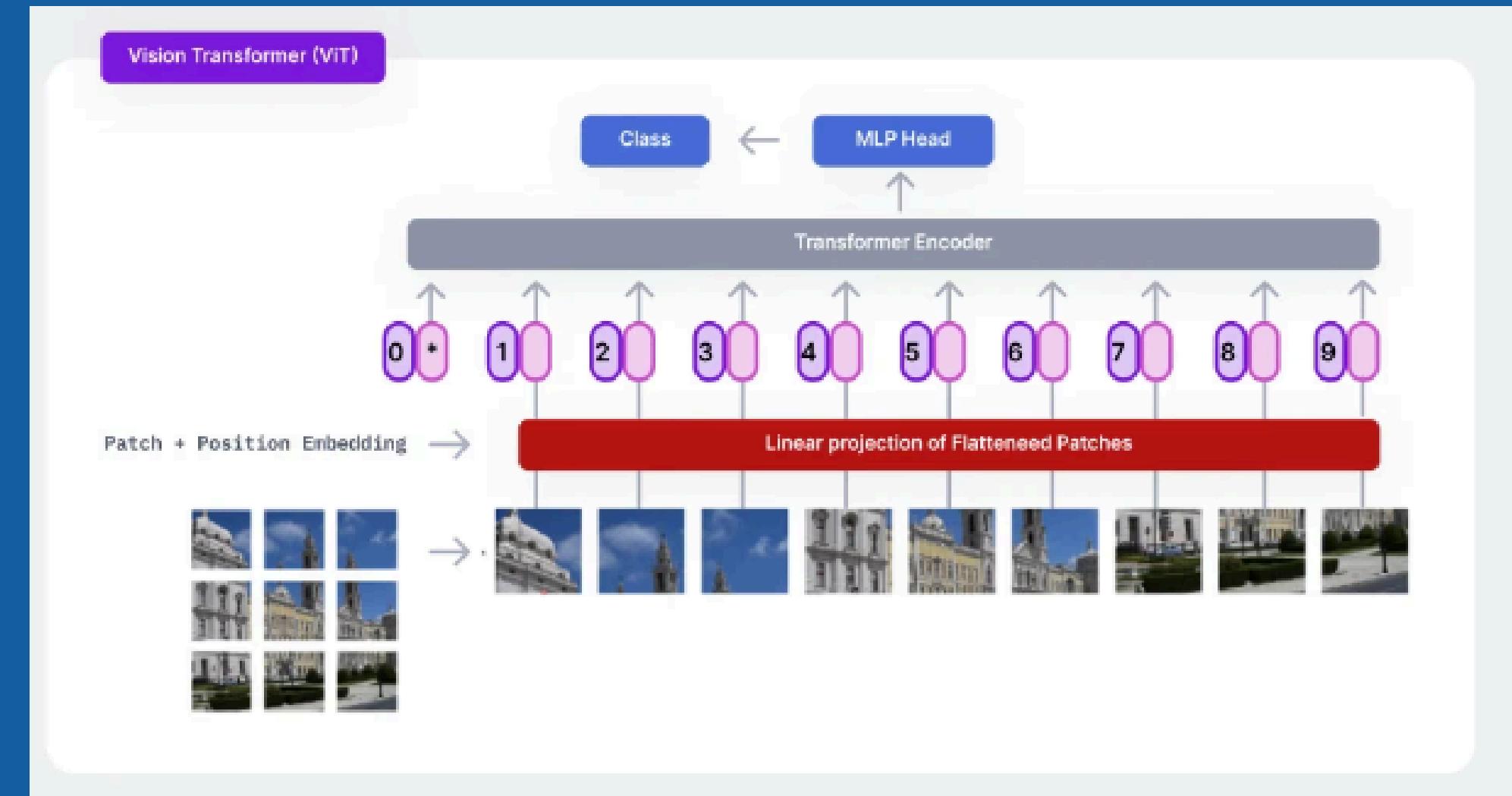
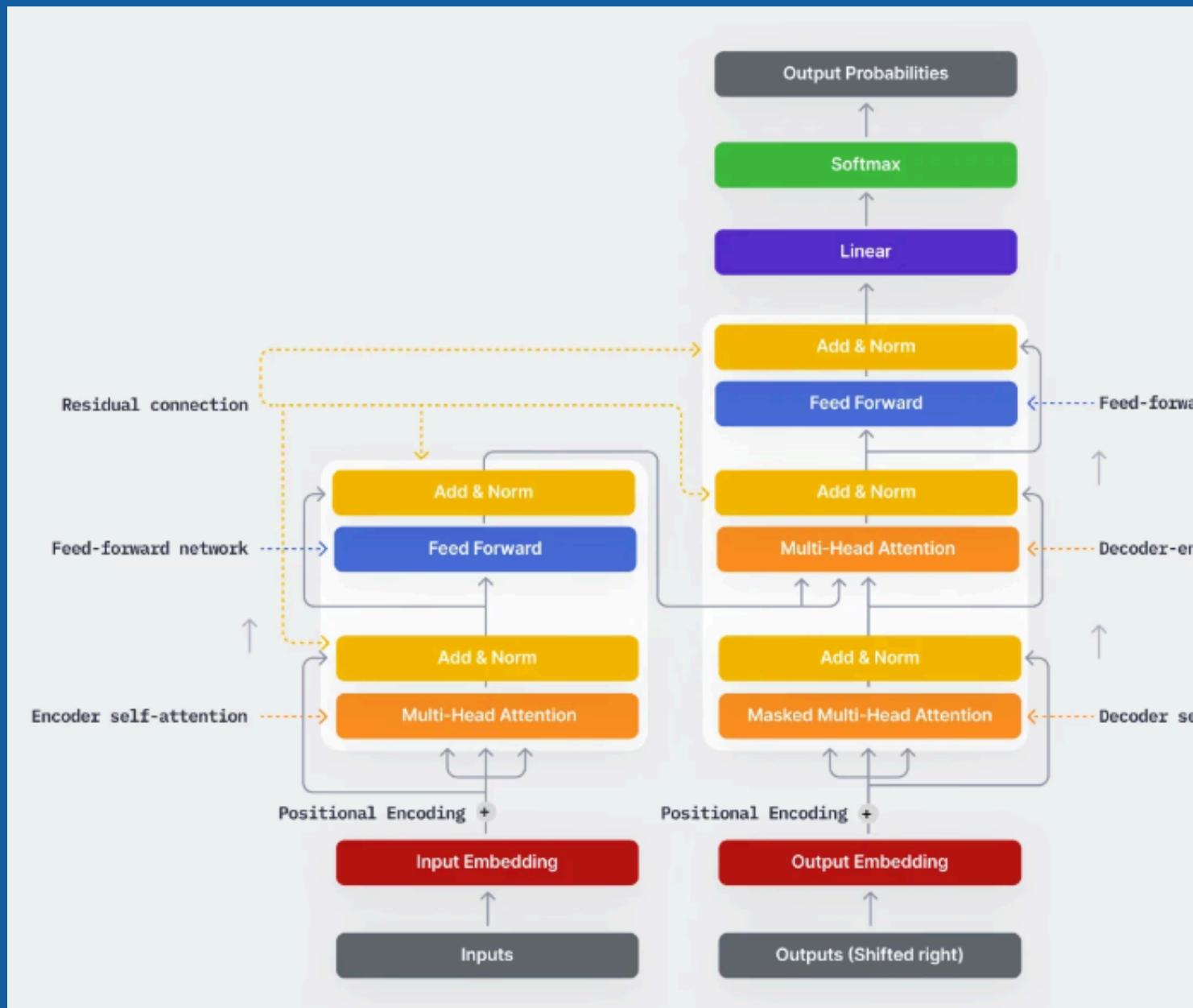
EfficientNet



*Image source [4]

Transformer

Vision Transformer



*Image source [5] [

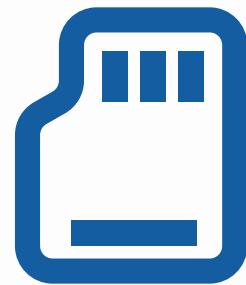
ViT-G-Base

Part of Google Research for Scaling ViT to very large model.

Sample Efficient (pretrained on large dataset, regularization prevent overfitting)

Power Law pattern for Compute and Performance (error $\propto 1/\text{compute}$)

Extra class token to produce final representation is removed to save memory



Layers

- 12 Transformer Encoder Layers
- 12 Attention Heads
- 768 dimension for patch Embedding
- Patch Size: 16x16



Parameters

86 million



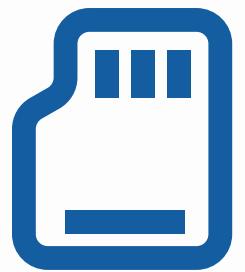
Architecture

ViT Based

CoCa ViT

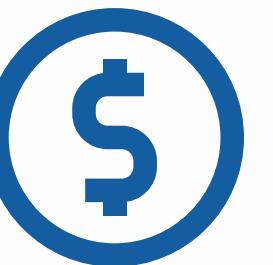
Multi-Modal Architecture designed for both text and image data.

Constrastive Pretraining[6]
Foundation Model



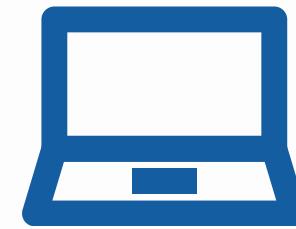
Layers

- 12 Transformer Encoder Layers
- 12 Attention Heads
- 768 dimension for patch Embedding
- Patch Size: 16x16



Parameters

220 million



Architecture

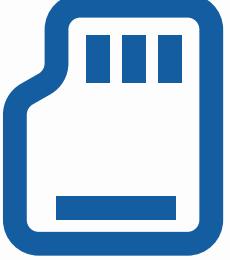
ViT Based

Yolo v8

Employs FPN (Feature Pyramid Network) and PAN (Path Aggregation Network)

Predicts bounding boxes, object classes and confidence score using anchor-free design.

Non Maximum Suppression applied to filter overlapping predictions

	 Layers	 Parameters	 Use case
n	30	6.2	Limited Resource, High fps
l	50	46.5	Balanced Accuracy, Good fps
x	70	68.2	High Accuracy, Low fps

RESULTS

Coral Binary Classification

Coral Multiclass Classification

Coral Detection

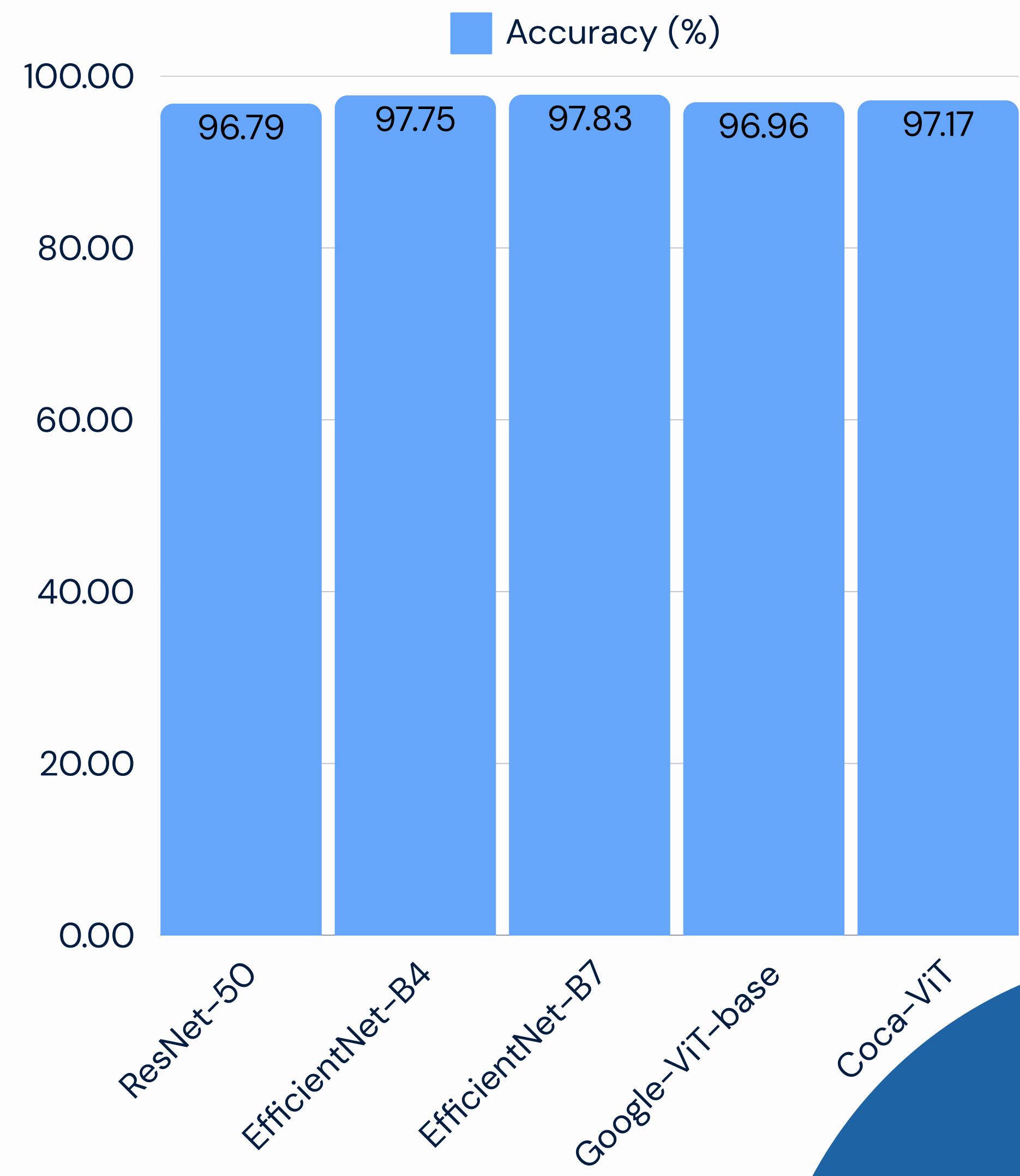
Binary Classification

In Binary Classification the task was to detect whether an image is of coral or not

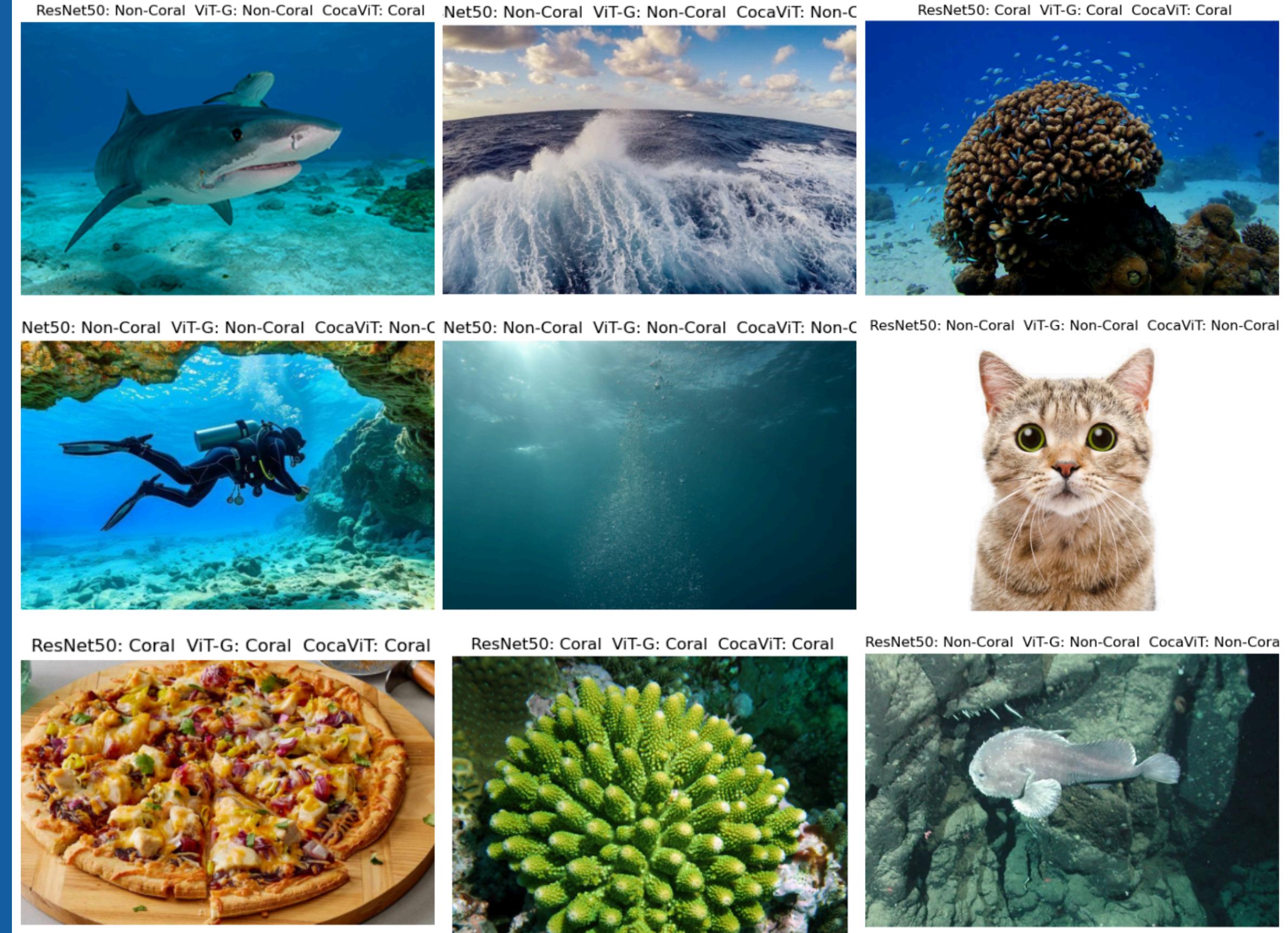
Best Model -> EfficientNet-B7

Accuracy -> 97.83%

The Google ViT-Base and CoCa ViT-Large, which leverage vision transformers, performed competitively, achieving accuracies of 96.96% and 97.17%, respectively, even with fewer epochs trained. ViT-G is significantly smaller and more computationally efficient than CoCa ViT, with only a 0.21% drop in accuracy while requiring considerably less training time.



RESCUE AI



Multiclass Classification

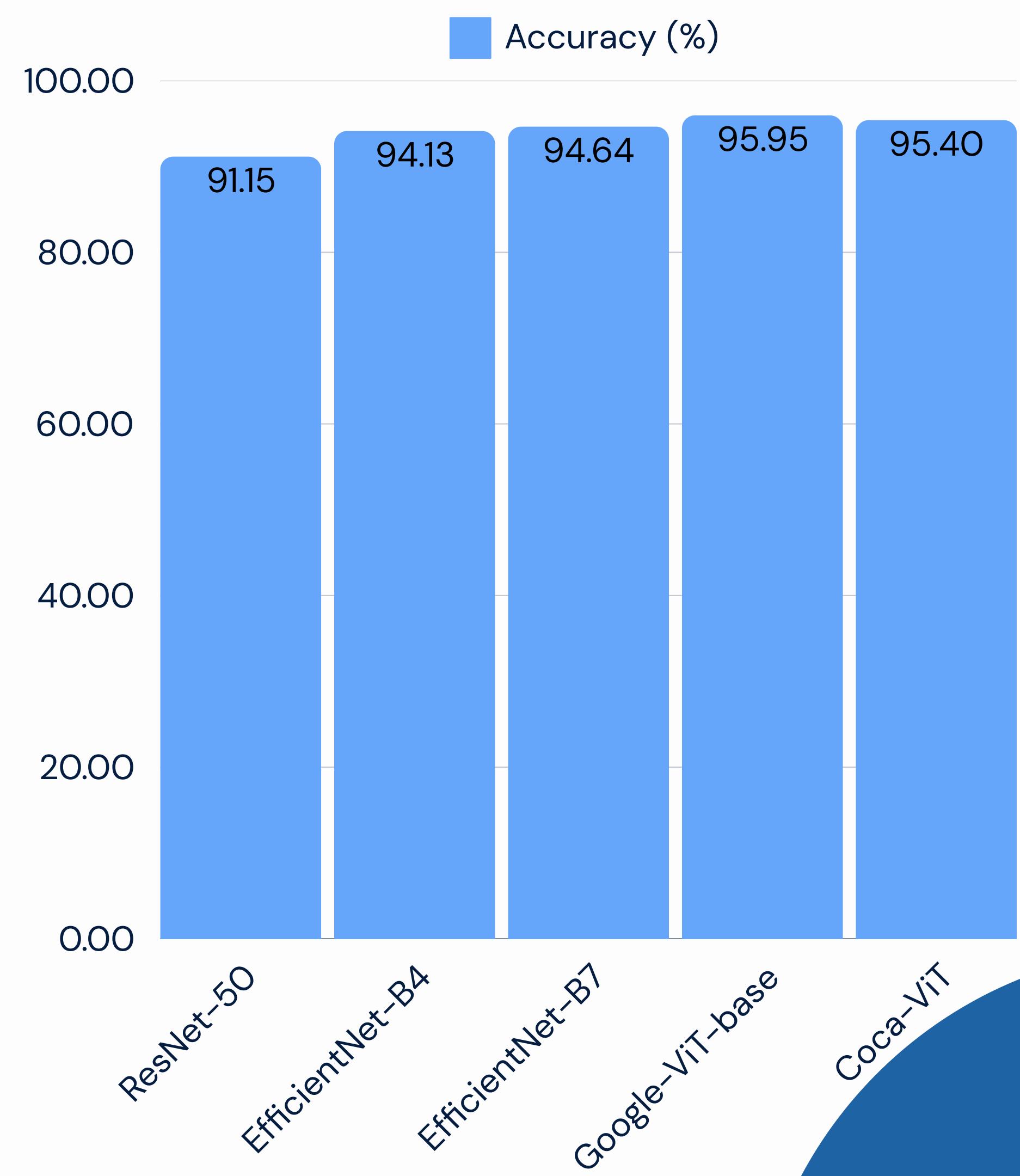
In Multiclass Classification task was to classify image into 29 classes out of which 27 were coral species and 2 classes were Marine for underwater non coral images and other class was for non marine non coral images.

Best Model -> Google-ViT-base

Accuracy -> 95.95%

The CoCa ViT-Large, which combines the strengths of contrastive learning and vision transformers, achieved a notable accuracy of 95.4%.

The EfficientNet-B4 and B7 models demonstrated consistent performance, with accuracies ranging from 94.31% to 94.64%, indicating their suitability for multiclass tasks as well.



RESULTS

ResNet50:Dendrogyra cylindrus ViT-G:Acropora Palmata CocaViT:Dendrogyra cylindrus



ResNet50:Montipora ViT-G:Montipora CocaViT:Montipora



ResNet50:Marine ViT-G:Marine CocaViT:Marine



ResNet50:Others ViT-G:Others CocaViT:Others



ResNet50:Colpophyllia Natans ViT-G:Colpophyllia Natans CocaViT:Colpophyllia Natans



ResNet50:Pocillopora meandrina ViT-G:Pocillopora meandrina CocaViT:Montipora



ResNet50:Siderastrea Siderea ViT-G:Leptoseris CocaViT:Leptoseris



ResNet50:Porites ViT-G:Pocillopora meandrina CocaViT:Pocillopora meandrina



Object Detecton

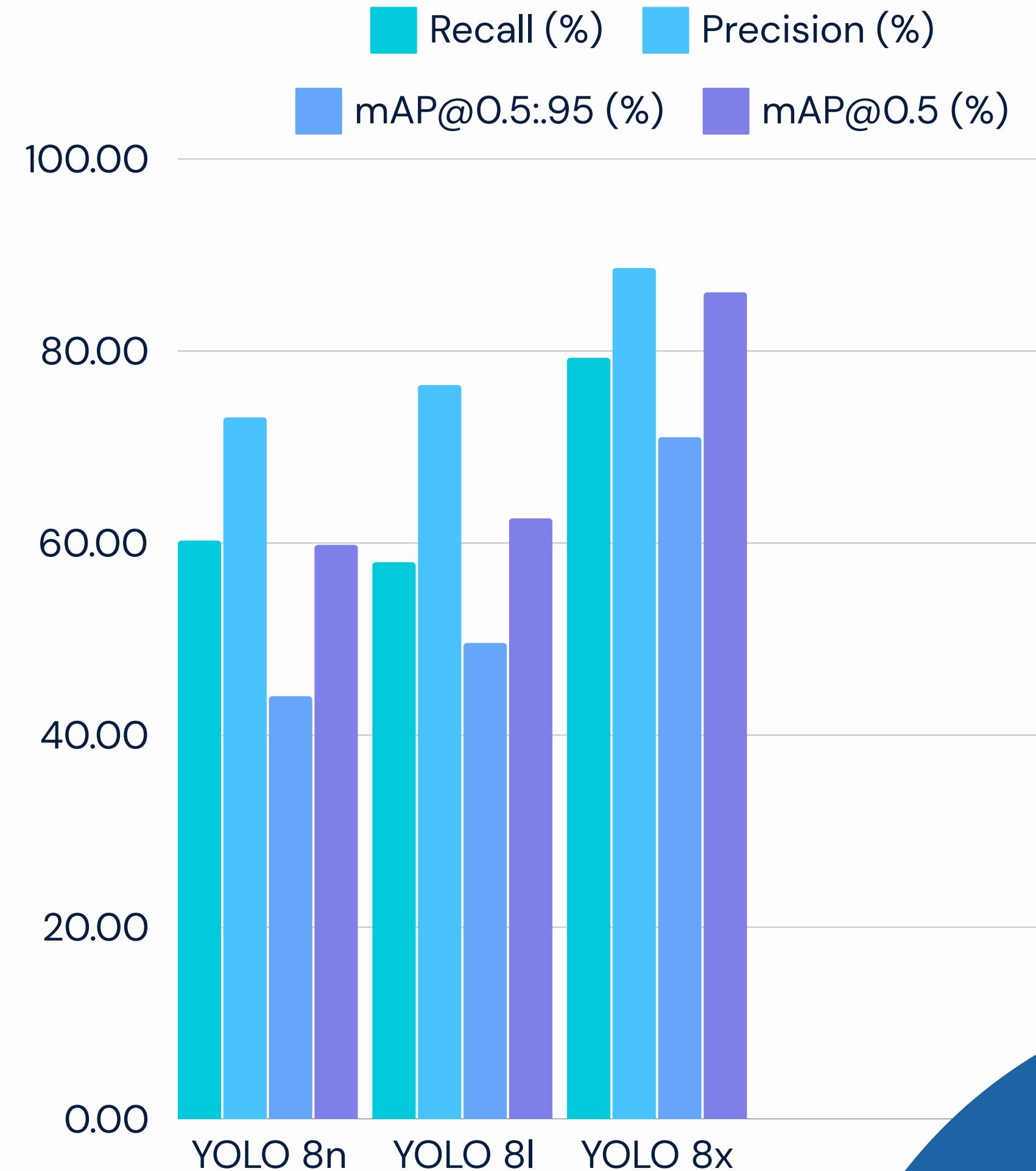
The aim of the task was to detect coral from images and therefore can be identified further if need be

Best Model -> YOLO 8x

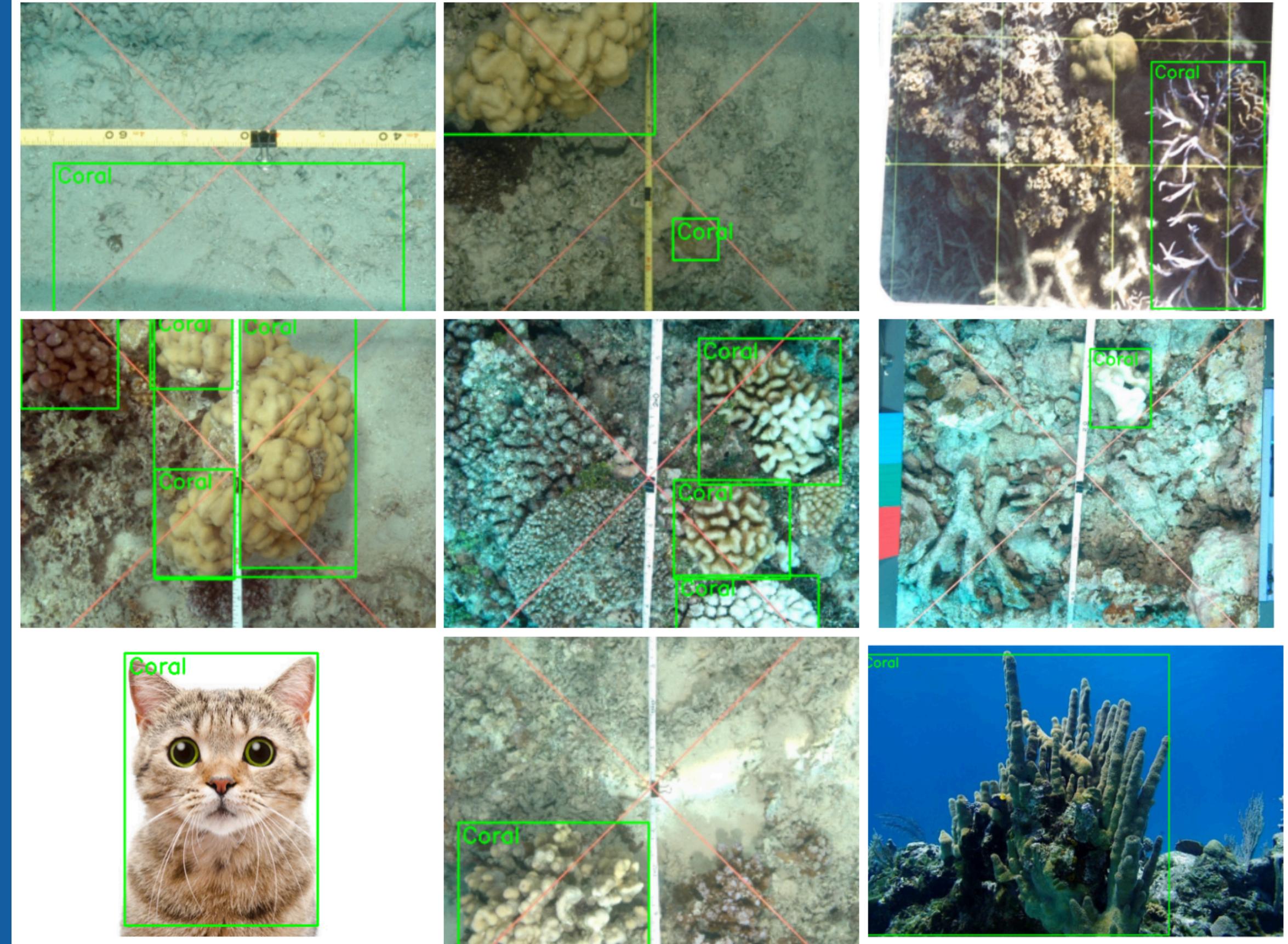
F1 score -> 83.86%

YOLOv8l showed balanced performance, achieving better results than YOLOv8n in all metrics, particularly in mAP@0.5:0.95 with a value of 49.59\%.

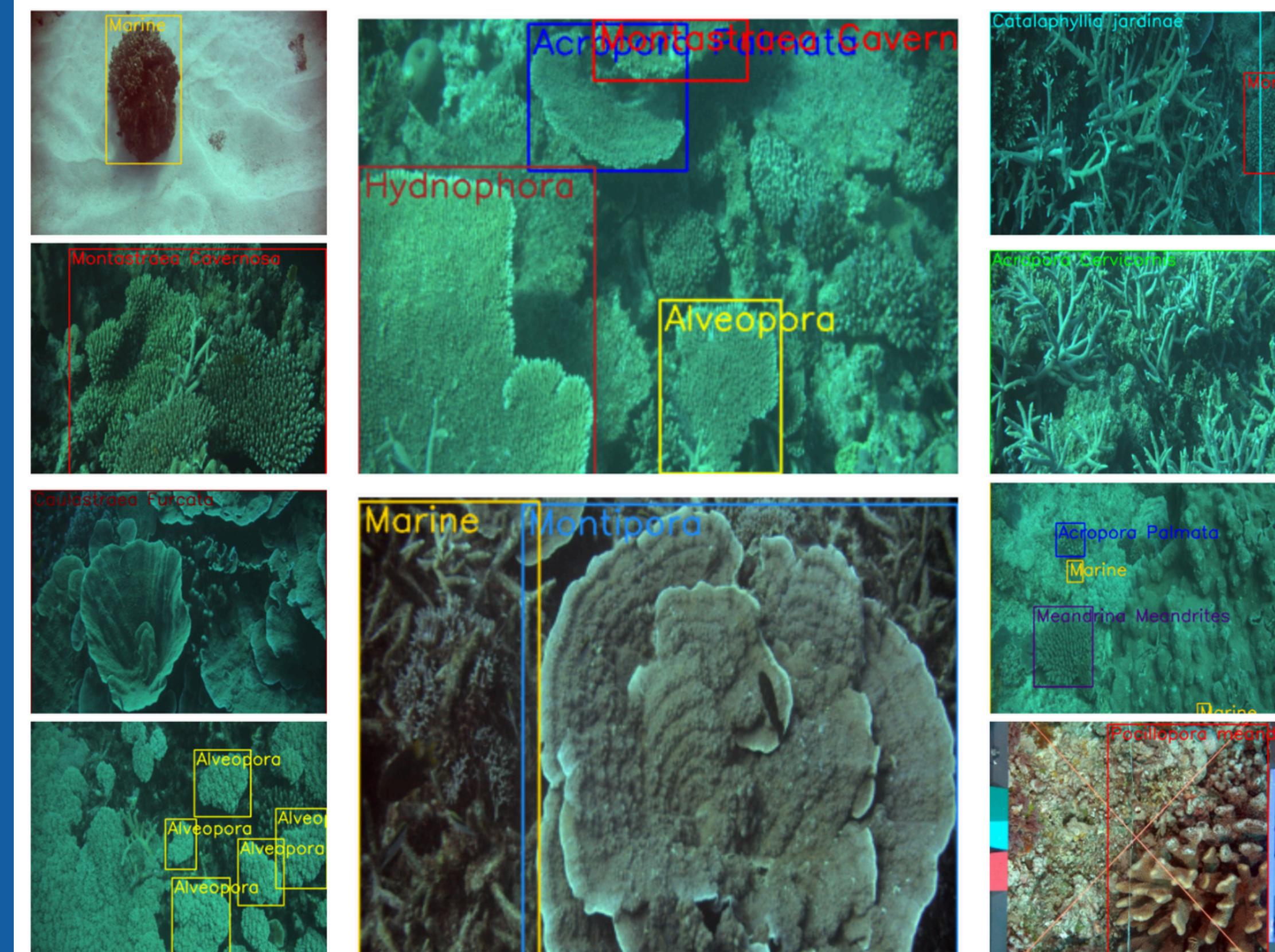
YOLOv8n, the smallest and most lightweight model, demonstrated the fastest computation but with the lowest accuracy metrics, achieving a mAP@0.5:0.95 of 44.05\%.



RESULT VISUAL



RESCUE FINAL



CONCLUSION

- **ViT-G performs exceptionally across binary and multiclass classification in terms of both compute efficiency and accuracy.**
- **Followed closely by EfficientNet-B7 which also set highly benchmarking result.**
- **CoCaViT although accurate fail to provide substantial improvement given its size may be due to limited data.**
- **Yolo8x performs best among yolo8 variants due to sheer size and ability to learn complex patterns.**



Future Work

1. Focus on improving dataset quality for more accurate predictions.
2. Seeking professional help in labeling of dataset for misclassified data.
3. Extending this study to Image segmentation which would be more accurate in real life given coral shape and size.
4. Deployment of such models to autonomous submersible to real time detection and classification of coral for monitoring and analysis.

Human-in-the-Loop Segmentation of Multi-species Coral Imagery

Scarlett Raine^{1,2} *Graduate Student Member, IEEE*, Ross Marchant³,
 Brano Kusy² *Member, IEEE*, Frederic Maire¹, Niko Sünderhauf¹ *Member, IEEE*
 and Tobias Fischer¹ *Senior Member, IEEE*

Marine surveillance and monitoring of coral reefs is a challenging task due to the complex and dynamic nature of coral reef imagery, high inter-class similarity between coral species, and the need for accurate segmentation results.

label propagation methods have been proposed to address this challenge. However, these methods often fail to correctly segment coral reefs due to the presence of artifacts and noise in the images. To overcome this limitation, we propose a novel approach that combines point-based labeling with a dense segmentation network. This approach allows for accurate segmentation of coral reefs even in challenging environments, such as those with low lighting conditions or poor visibility.

10 point labels. When human-in-the-loop labeling is not available, using the denoised DINov2 features with a KNN still improves on the prior state-of-the-art by 2.7% for pixel accuracy and 5.8% for mIoU (5 grid points). On the semantic segmentation task, we outperform the prior state-of-the-art by 8.8% for pixel accuracy and by 13.5% for mIoU when only 5 point labels are used for point label propagation. Additionally, we perform a comprehensive study into the impacts of the point label placement style

CoralVOS: Dataset and Benchmark for Coral Video Segmentation

Ziqiang Zheng^{1*}, Yaofeng Xie², Haixin Liang³, Zhibin Yu², Sai-Kit Yeung^{1,3}



Next Stop Image Segmentation

on manual efforts is significantly time-consuming, the existing coral analysis algorithms compromise and opt for performing down-sampling and only conducting sparse point-based coral analysis within selected frames. However, such down-sampling will inevitable introduce the estimation bias or even lead to wrong results. To address this issue, we propose to perform dense coral video segmentation, with no down-sampling involved. Through video object segmentation, we could generate more reliable and

Coral reefs represent one of the planet's most diverse and productive ecosystems, providing habitat and shelter for a vast range of marine species. Performing underwater coral reef monitoring [1], [2], [3], [4], [5], [6], [7] can identify and track changes in coral reef health, understand the impacts of human activities on coral reefs, and help maintain the

THANK YOU!

Contact:
ayushman.baghel.21033@iitgoa.ac.in
prakhar.gupta.21031@iitgoa.ac.in

References

1. <https://www.gy4es.org/post/the-coral-reefs-and-the-coral-crisis#:~:text=Biodiversity%20Hotspots-,Coral%20reefs%20are%20often%20referred%20to%20as%20the%20rainforests%20of,25%25%20of%20all%20marine%20species.>
2. <https://doi.org/10.48550/arXiv.1804.00516>
3. <https://wandb.ai/mostafaibrahim17/ml-articles/reports/The-Basics-of-ResNet50---VmIIdzo2NDkwNDE2>
4. <https://towardsdatascience.com/complete-architectural-details-of-all-efficientnet-models-5fd5b736142>
5. <https://www.v7labs.com/blog/vision-transformer-guide>
6. <https://www.v7labs.com/blog/contrastive-learning-guide>
7. <https://doi.org/10.48550/arXiv.2404.09406>