Underwater Coral Imaging

Presented By:

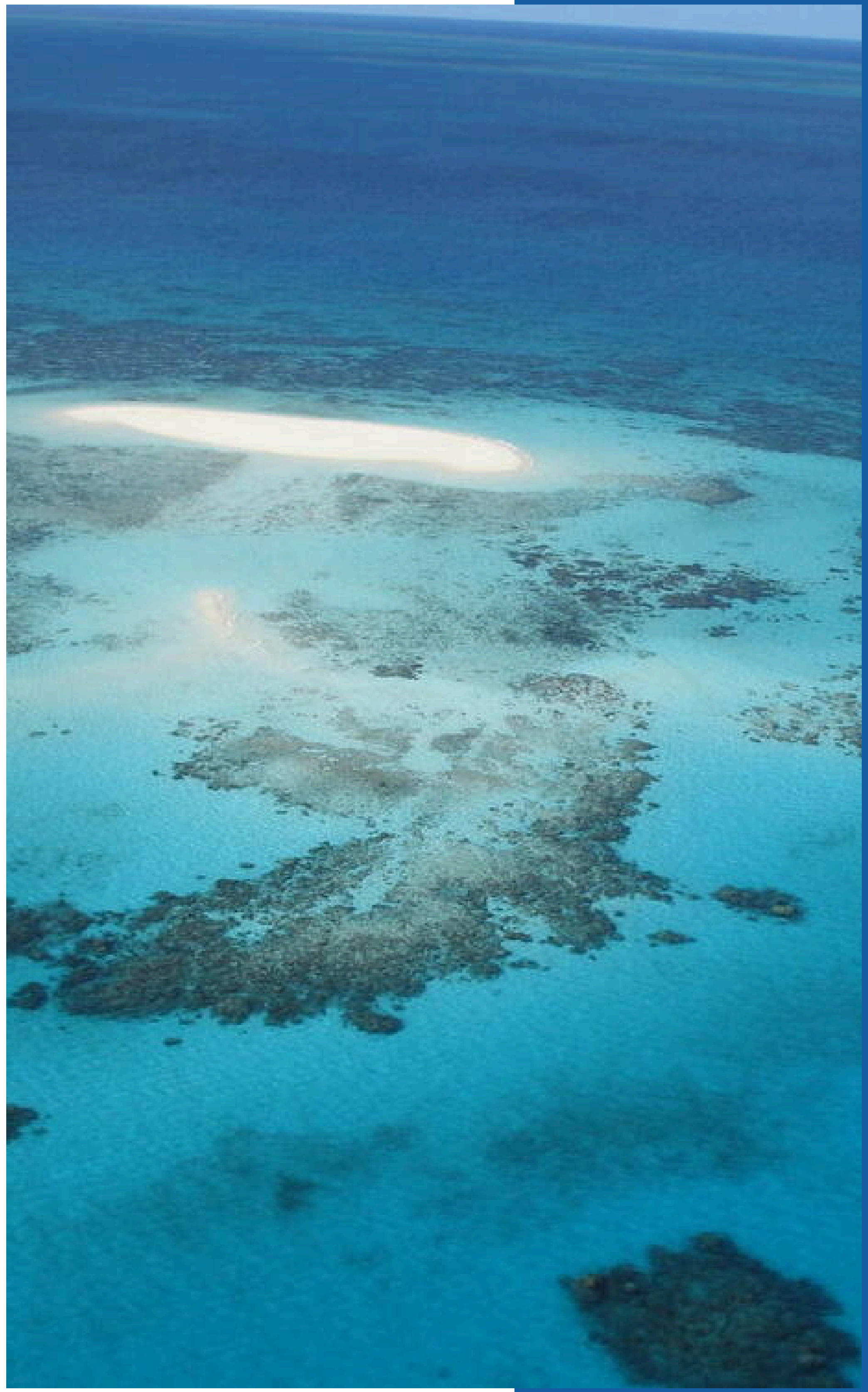
Ayushman Baghel

Prakhar Gupta

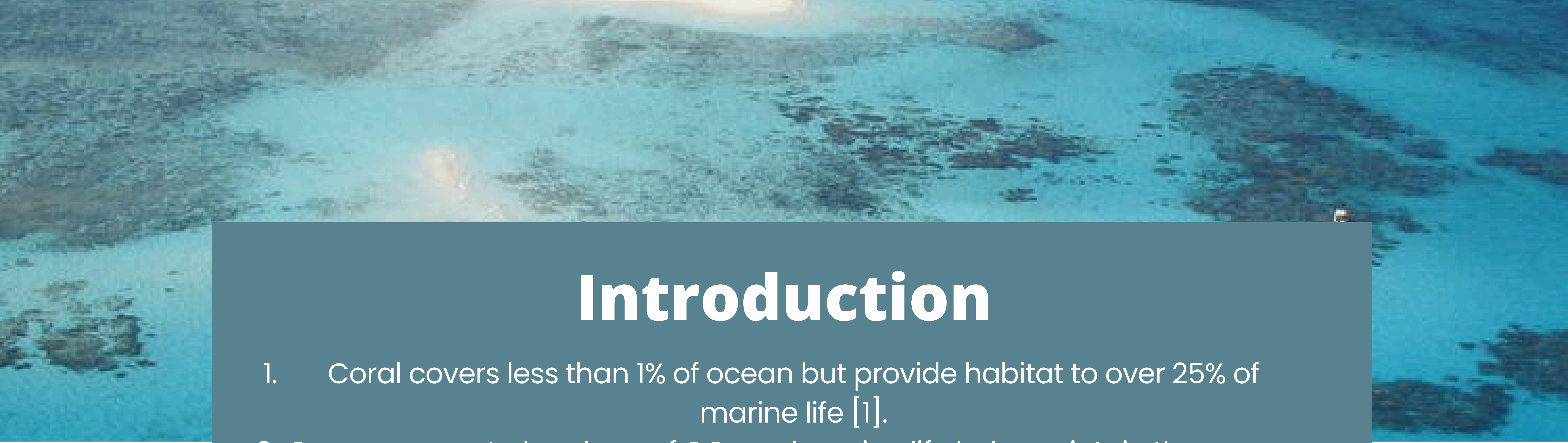
Supervised By:

Dr. Shitala Prasad

Overview



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 CO_2 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.

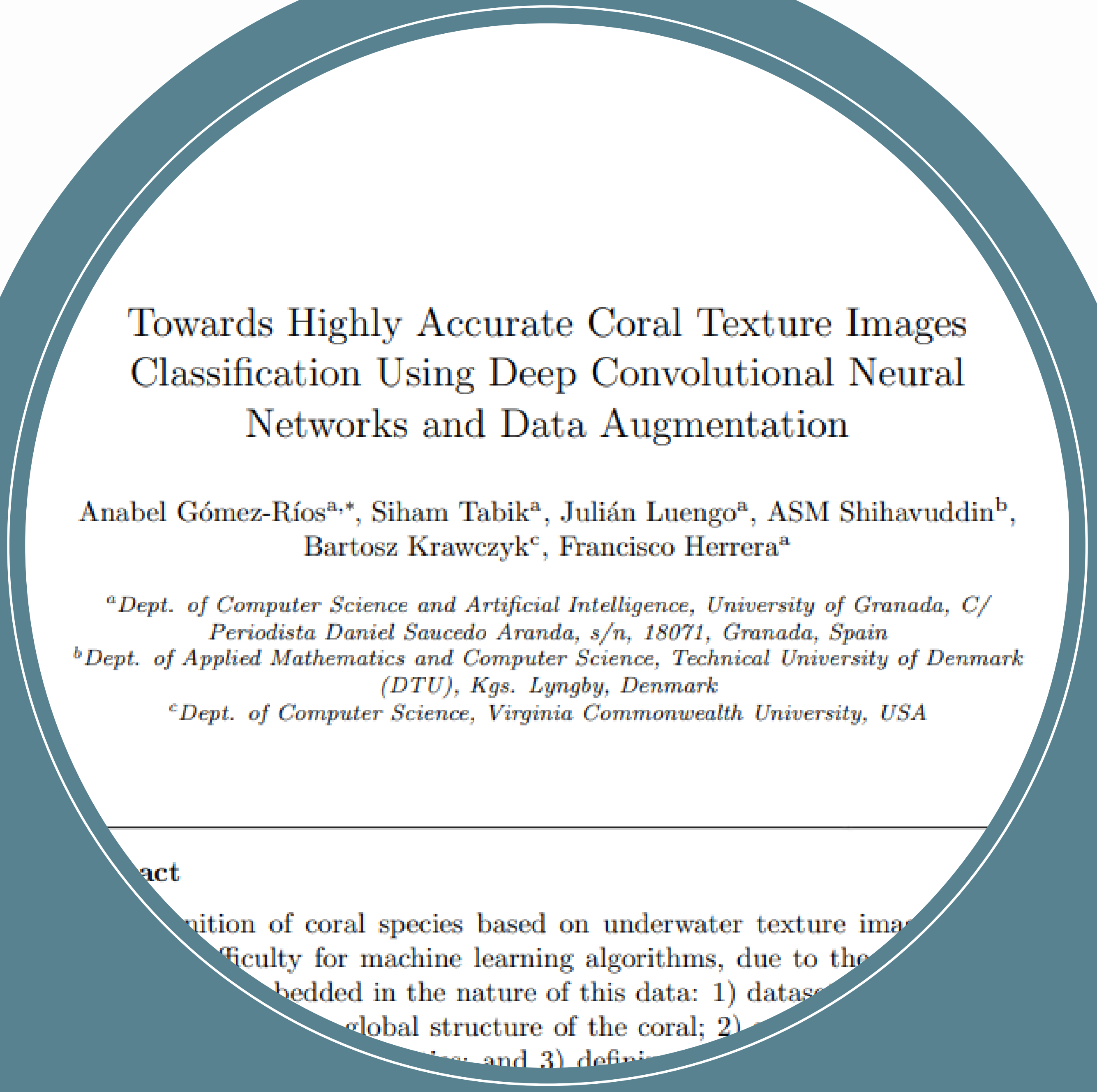
Literature Review

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]

[a]Dept. of Computer Science and Artificial Intelligence, University of Granada, C/ Periodista Daniel Saucedo Aranda, s/n, 18071, Granada, Spain
[b]Dept. of Applied Mathematics and Computer Science, Technical University of Denmark (DTU), Kgs. Lyngby, Denmark
[c]Dept. of Computer Science, Virginia Commonwealth University, USA

act

nition of coral species based on underwater texture ima
ficulty for machine learning algorithms, due to the
bedded in the nature of this data: 1) datas
global structure of the coral; 2)

Challenges

Problem 01

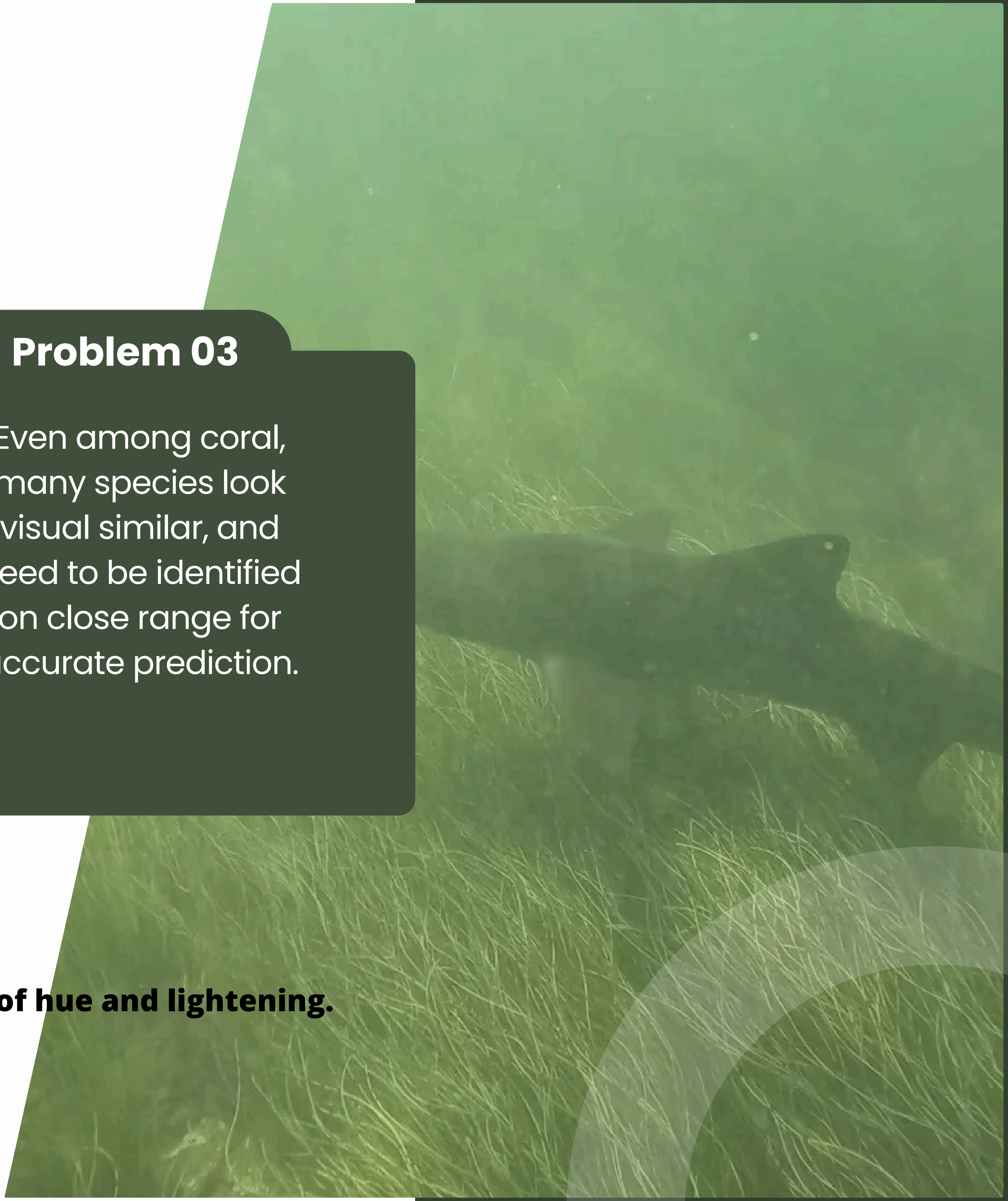
Under water images are riddled with visual distortion, blur, occlusion, lighening variation which significantly effect learning.

Problem 02

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

Problem 03

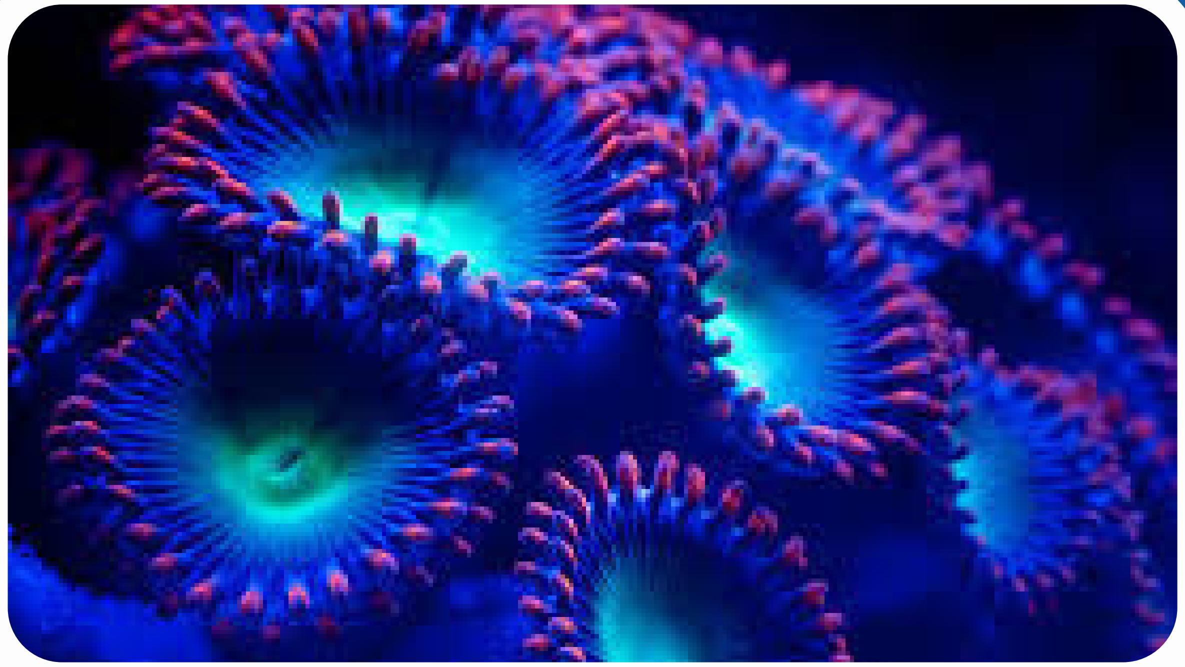
Even among coral, many species look visual similar, and need to be identified on close range for accurate prediction.



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

*Above images are from our own dataset that we made.

Binary Classification Dataset

625x625

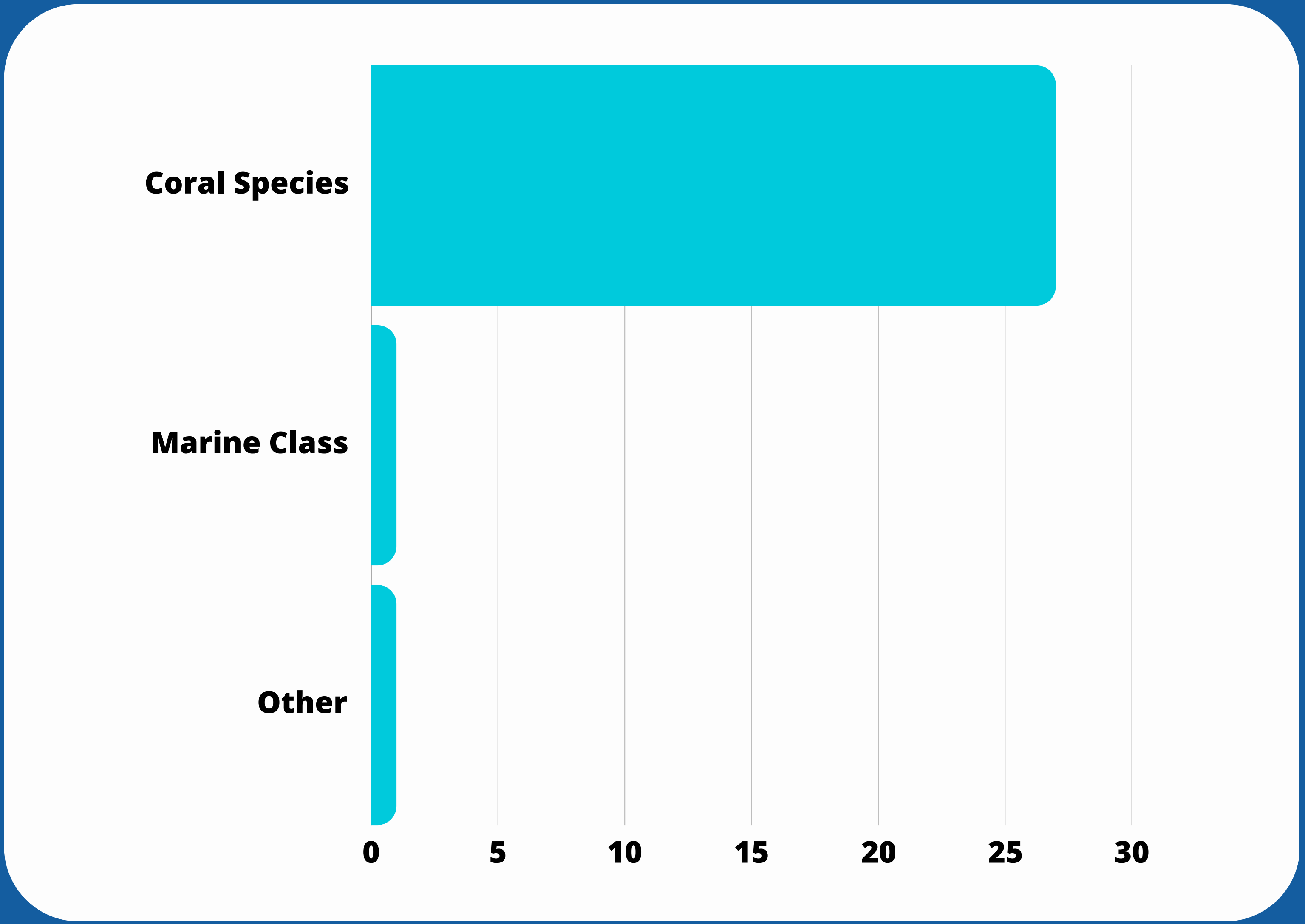
3 Channels

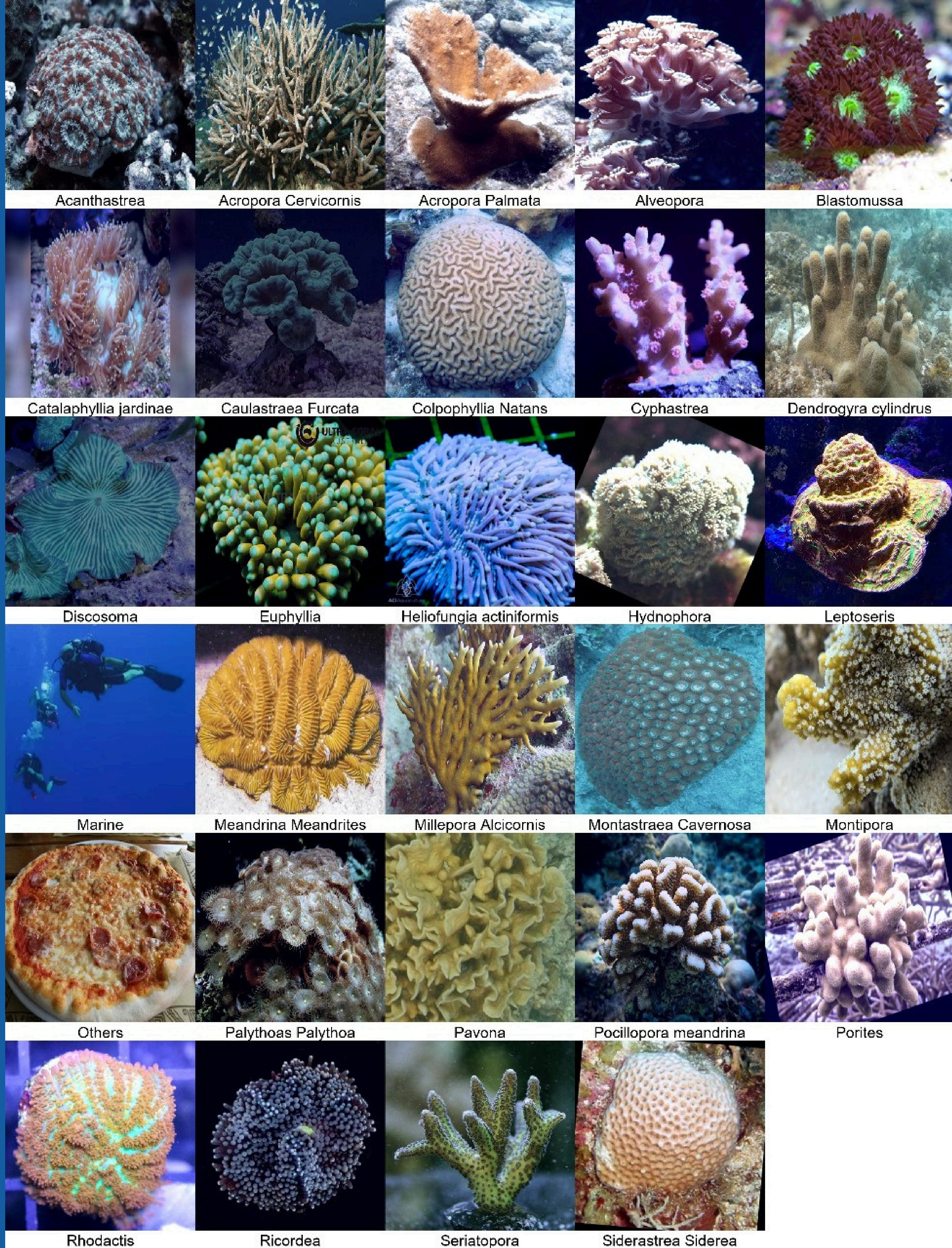
Coral	32,187
Non Coral	27,887
Total	60,074

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 3 convolution.

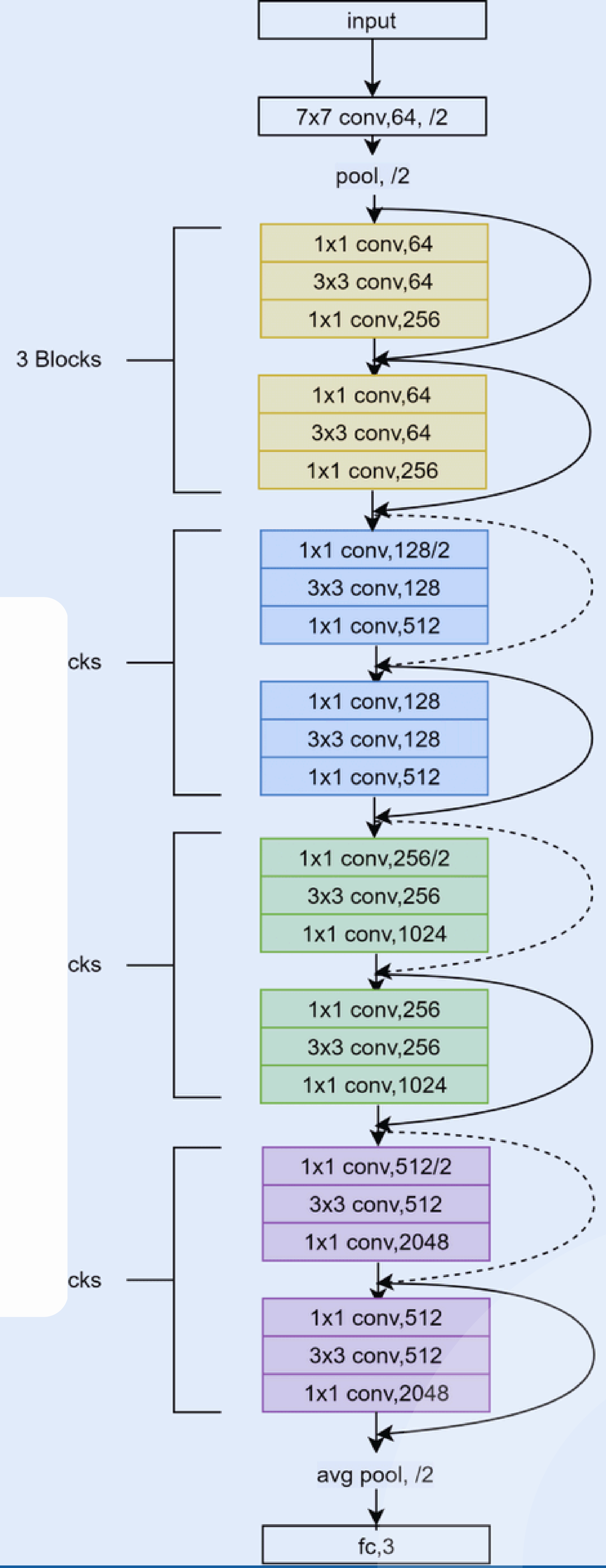
Parameters

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: $r = \gamma^\varphi$

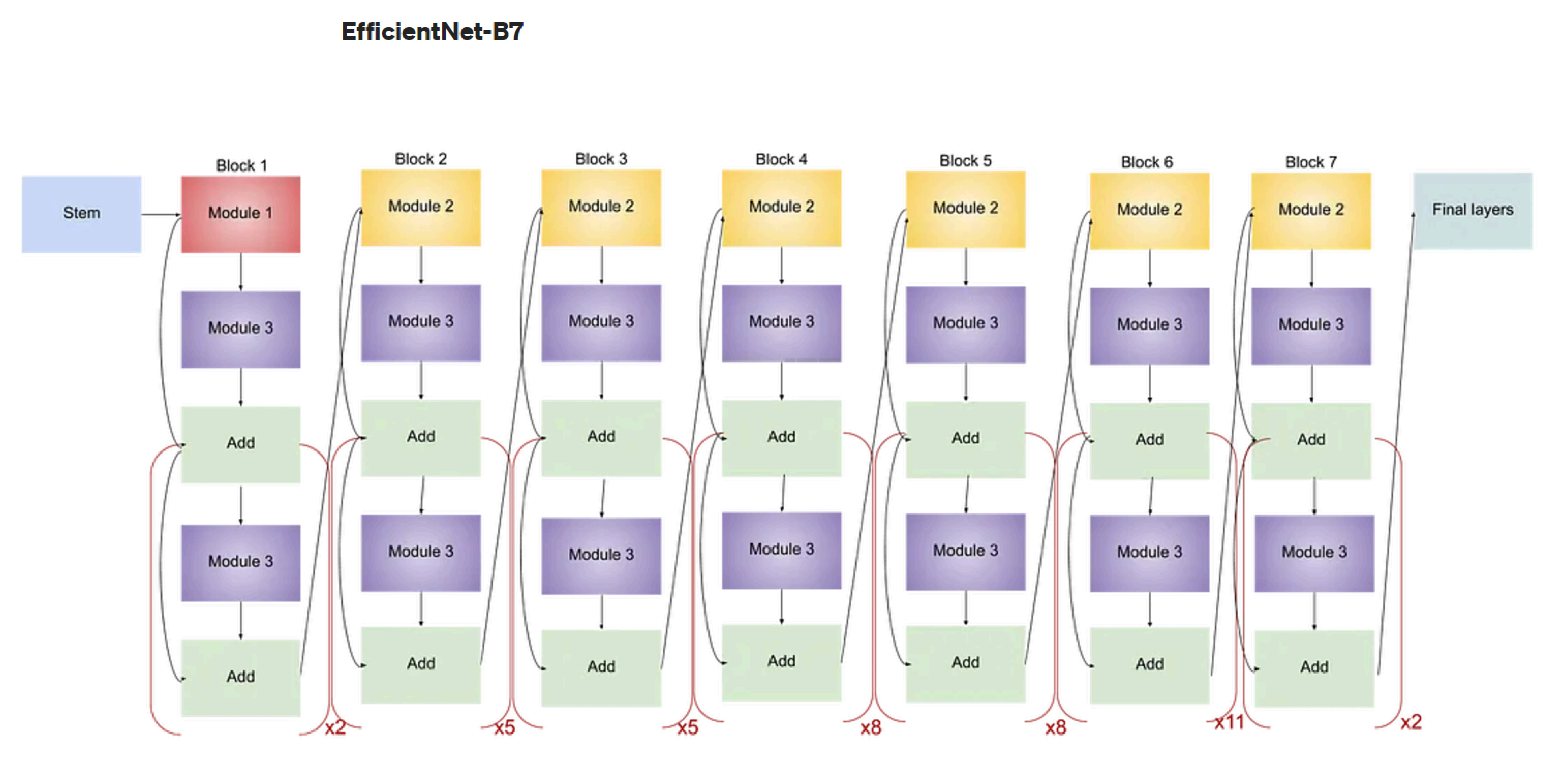
st $\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$

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

	Depth Multiplier	Width Multiplier	Resolution
B4	1.8	1.4	380x380
B5	2.2	1.6	456x456
B6	2.6	1.8	528x528
B7	3.1	2.0	600x600

*Image source [4]

EfficientNet

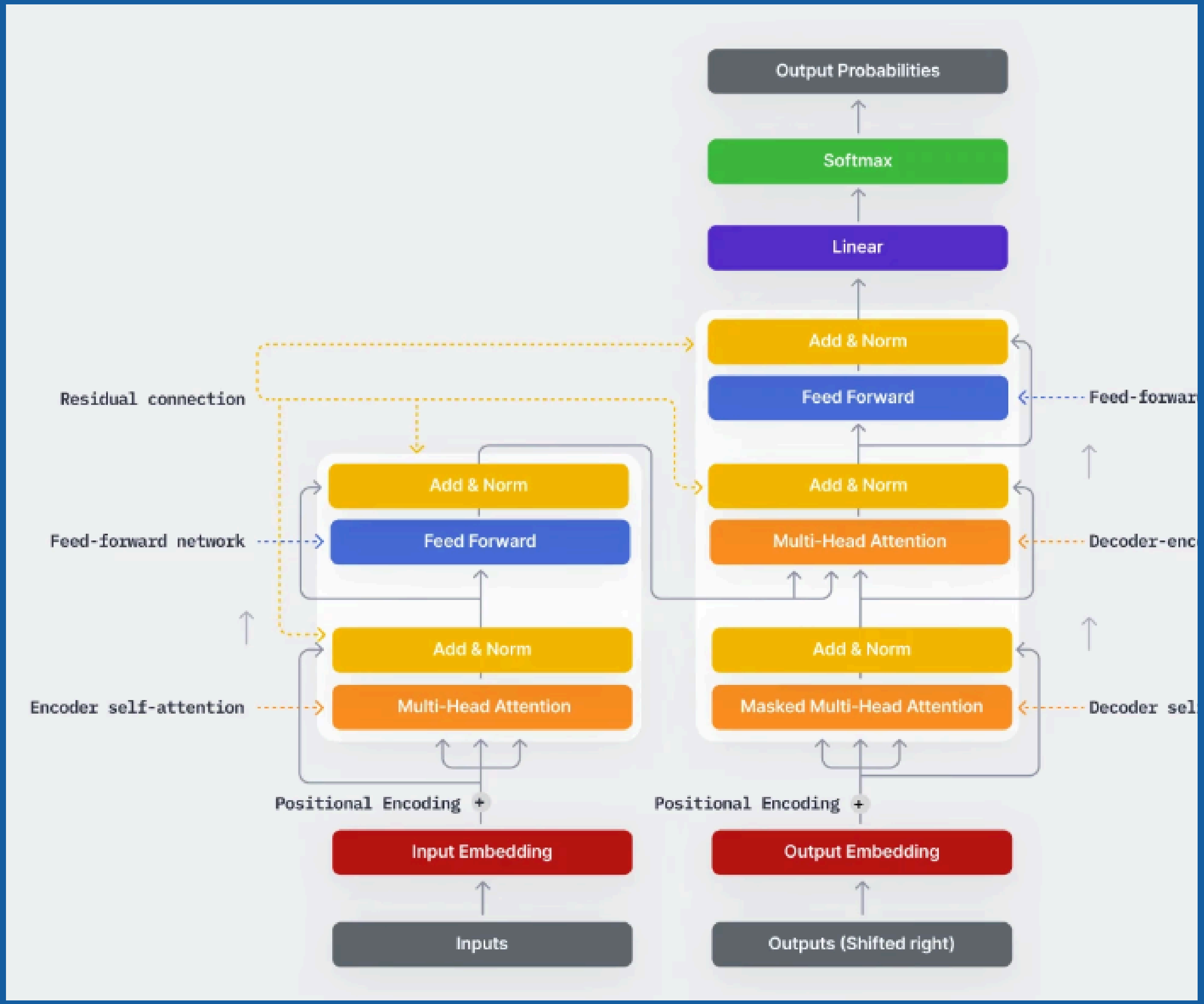


Architecture of EfficientNet-B7

*Image source [4]

Transformer

Vision Transformer



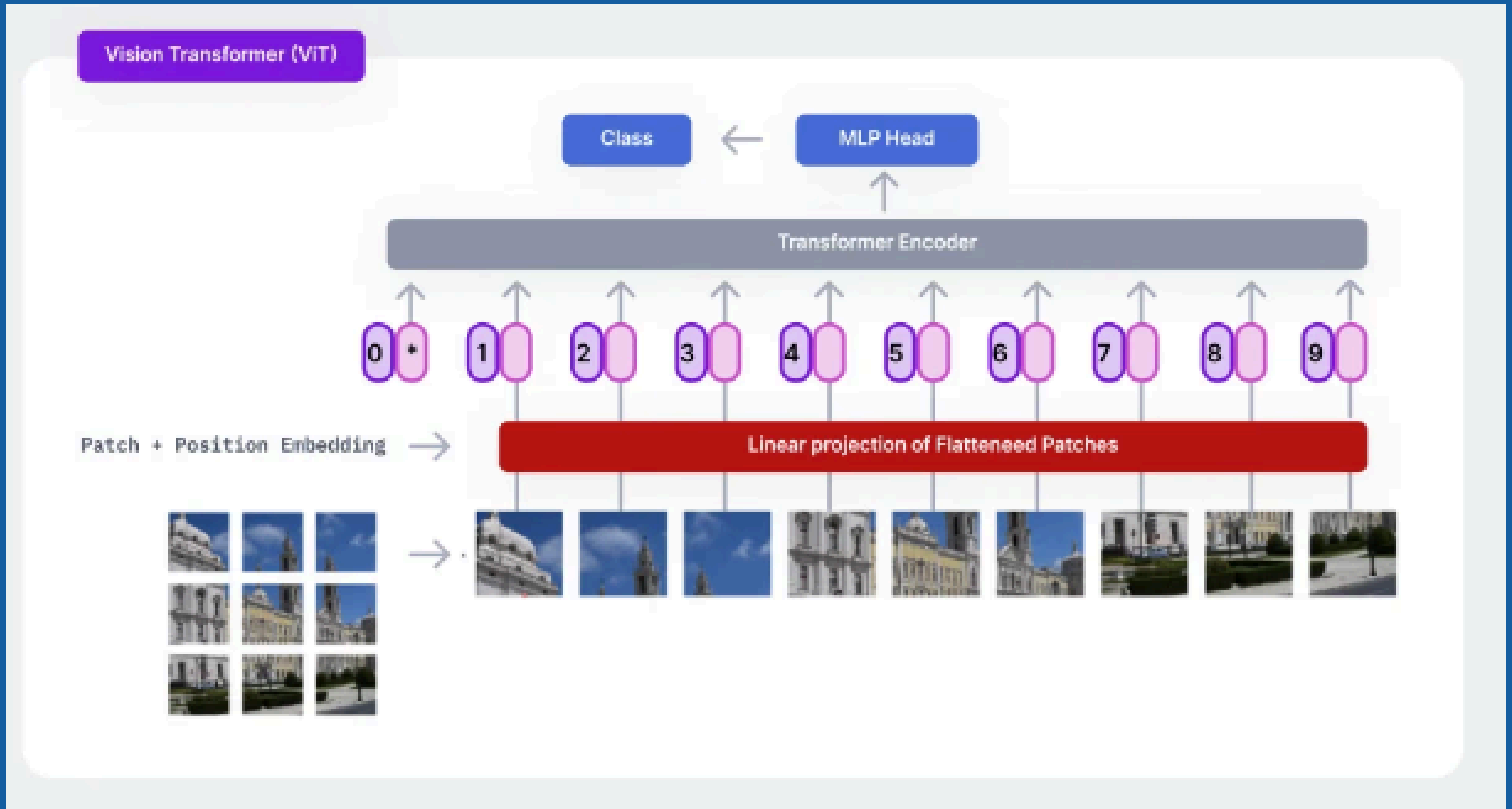


*Image source [5] [

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/compute)
Extra class token to produce final representation is removed to save memory

Layers

- 12 Transformmer 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 Pretaining[6]
Foundation Model

Layers

- 12 Transformmer 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

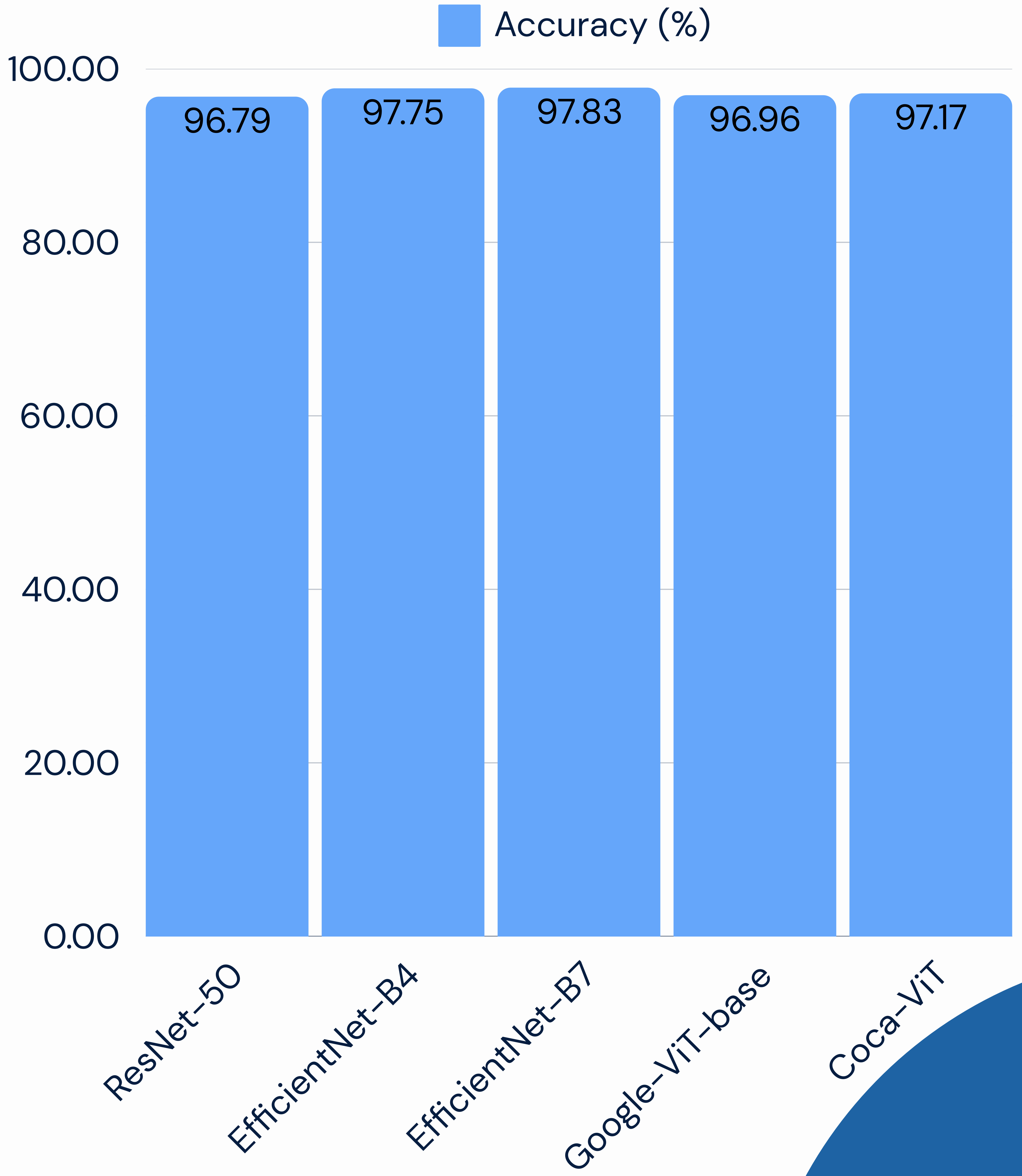
Binary Classification

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

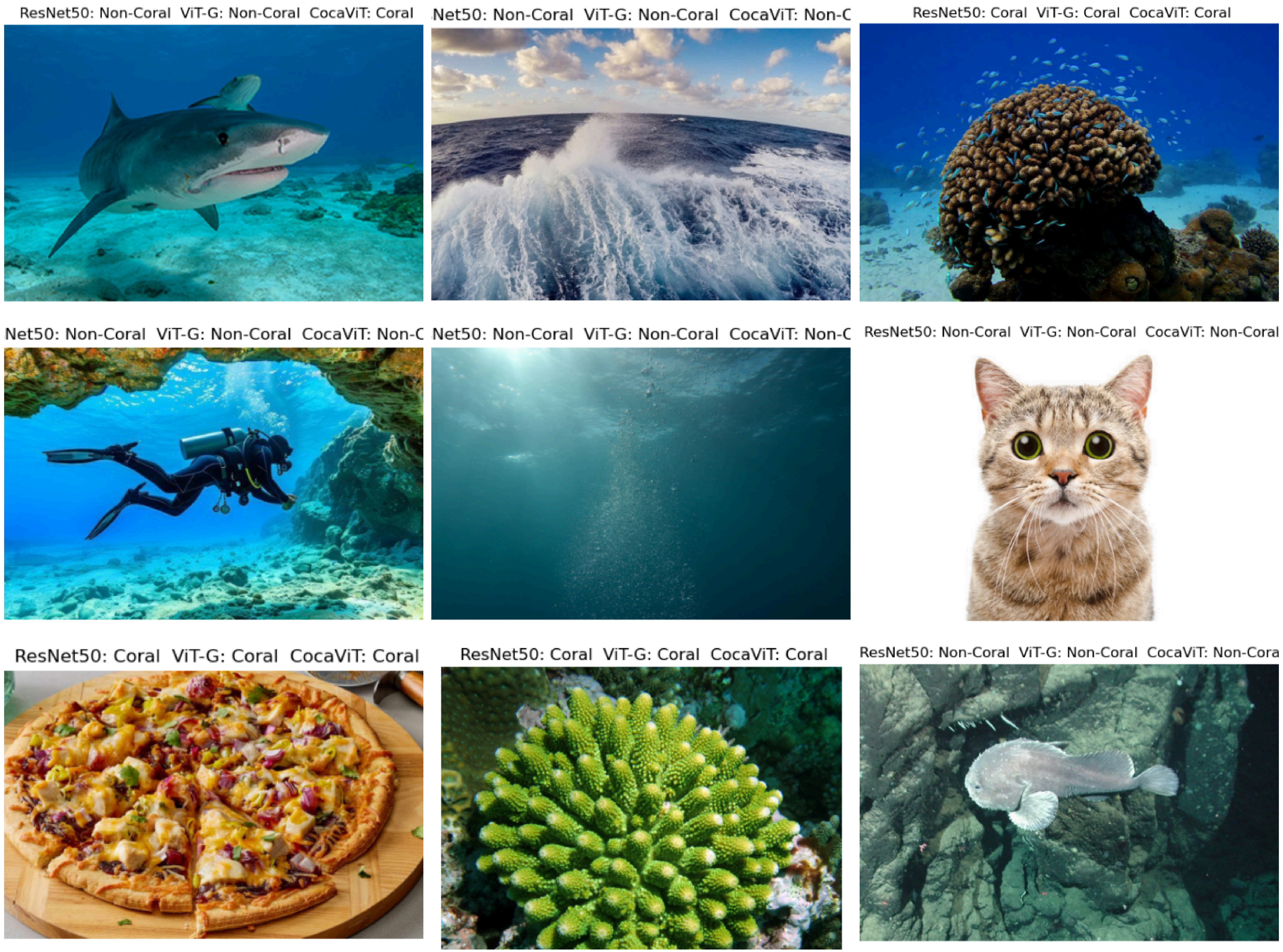
Best Model -> EfficientNet-B7

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.



VISUAL RESULT



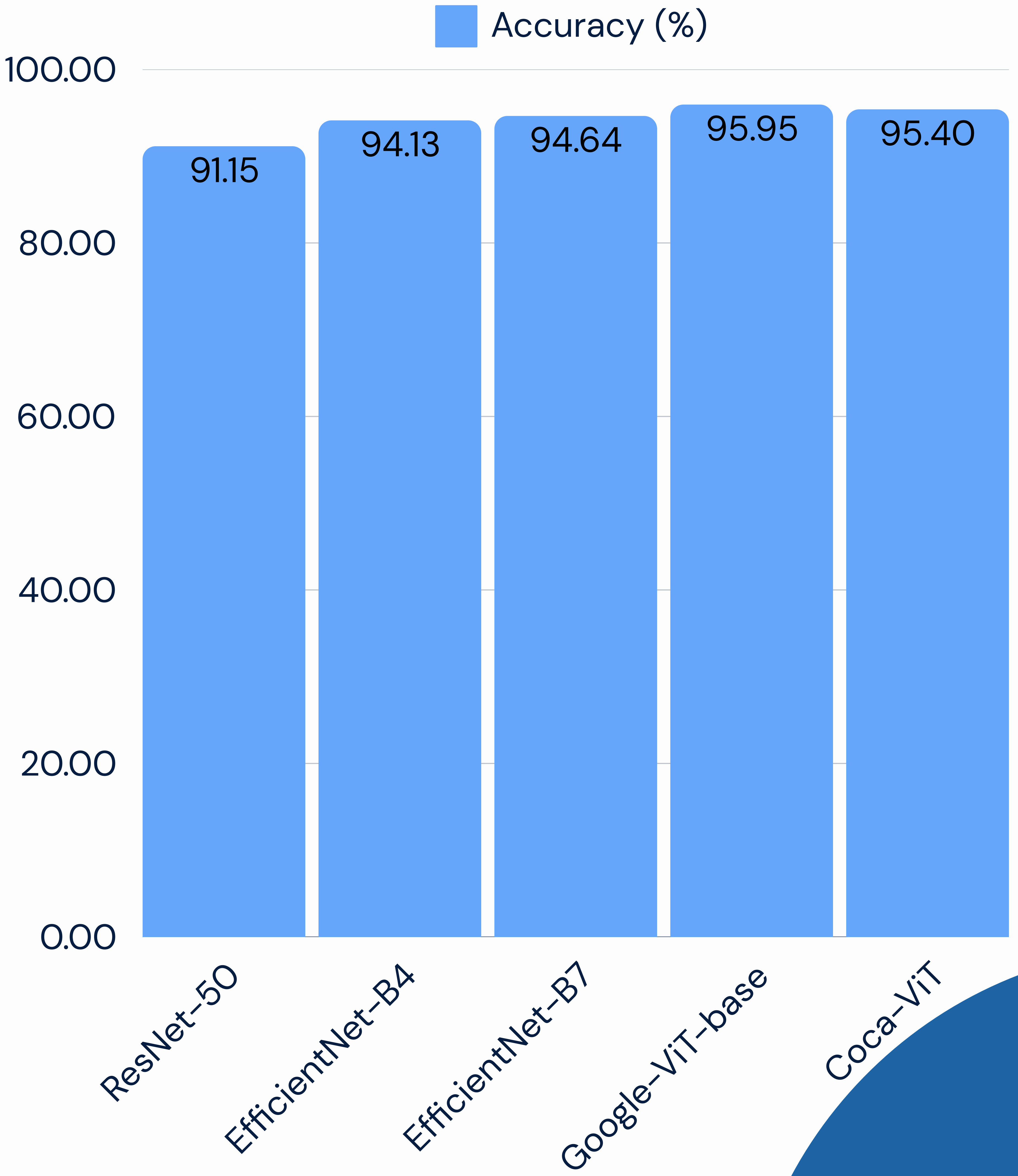
Multiclass Classification

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.

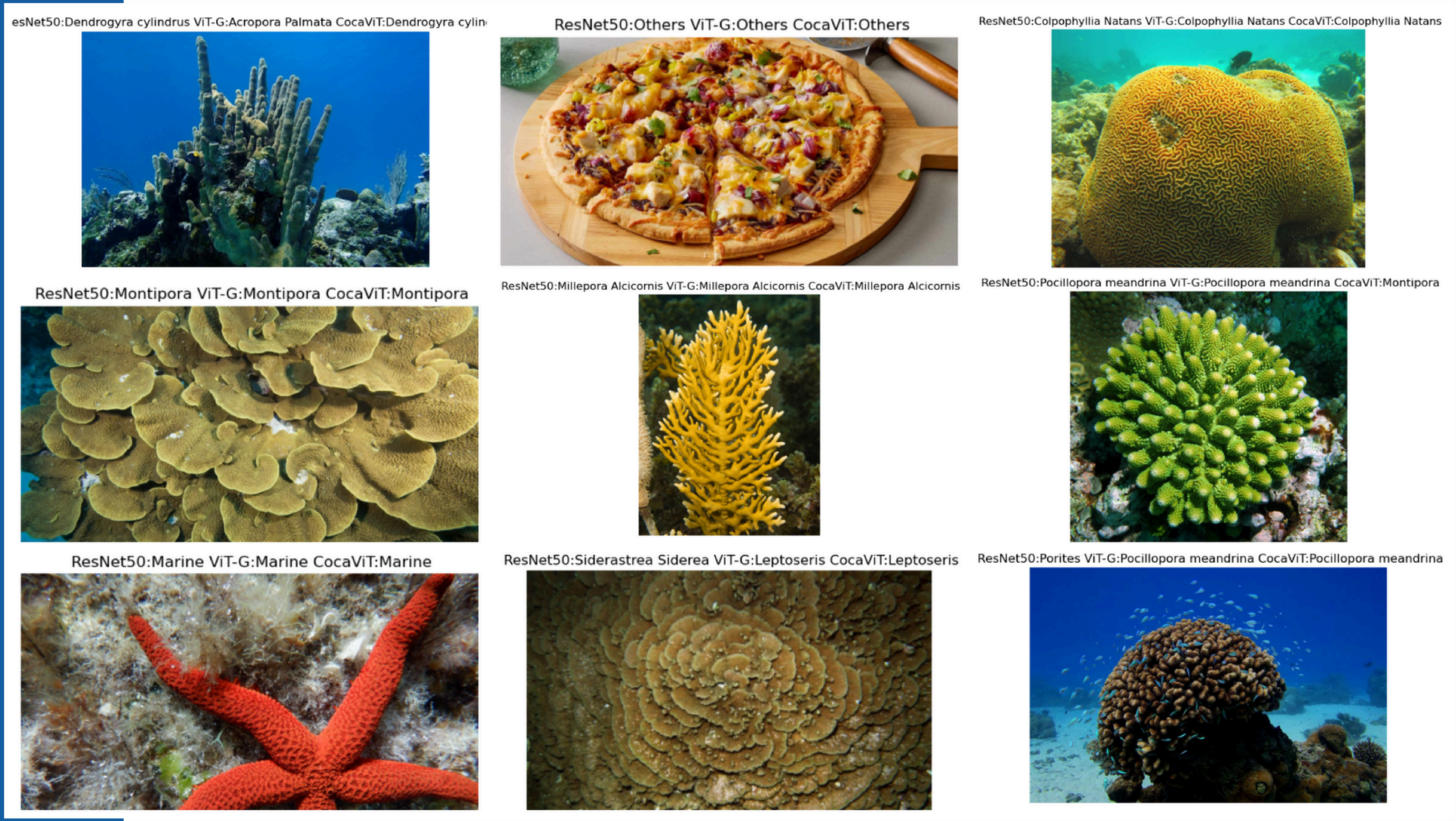
Best Model -> Google-ViT-base

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.



VISUAL RESULT



Object Detecton

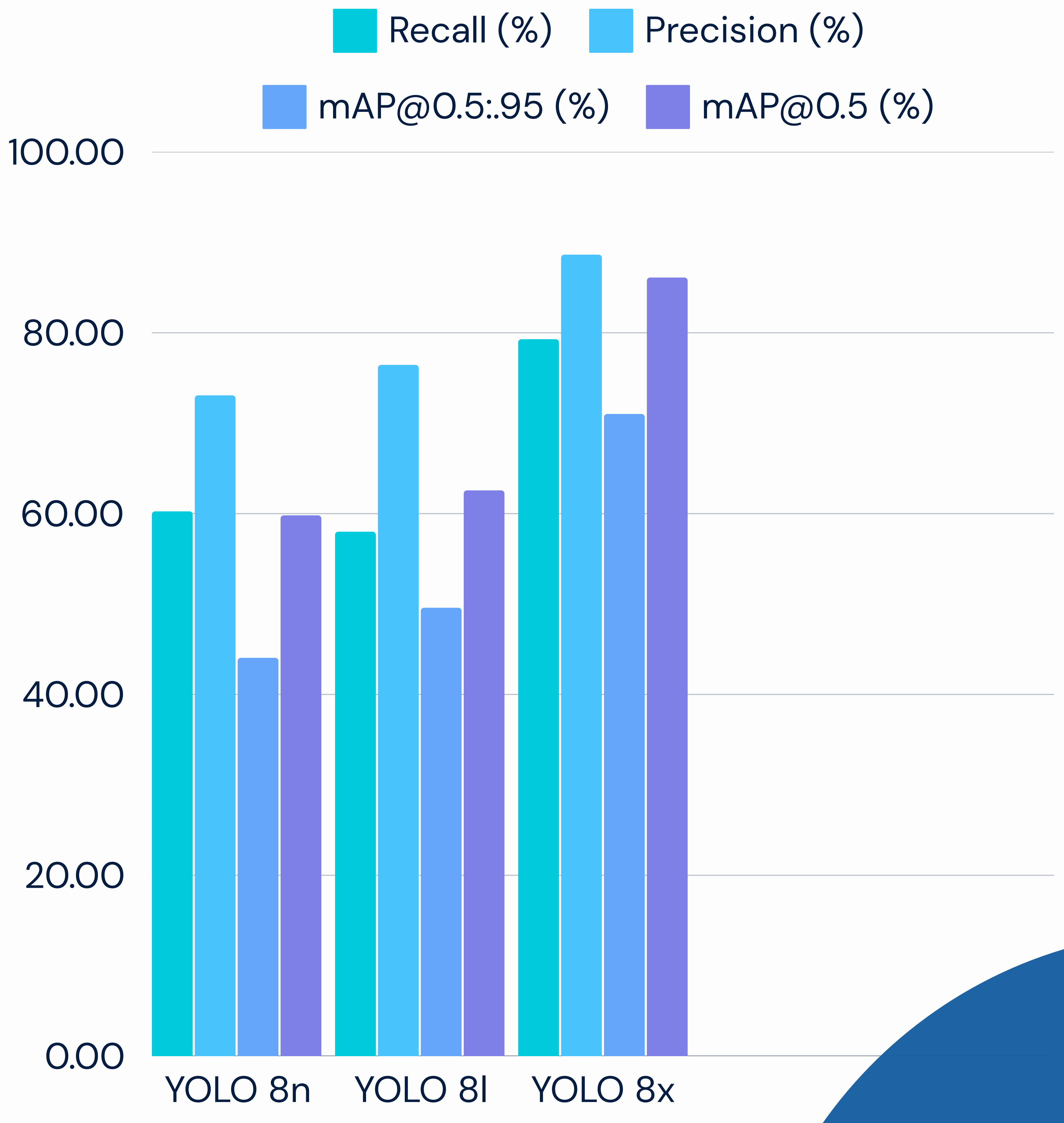
The aim of the task was to detect coral from images and therefore can be identified further if need be

Best Model -> YOLO 8x

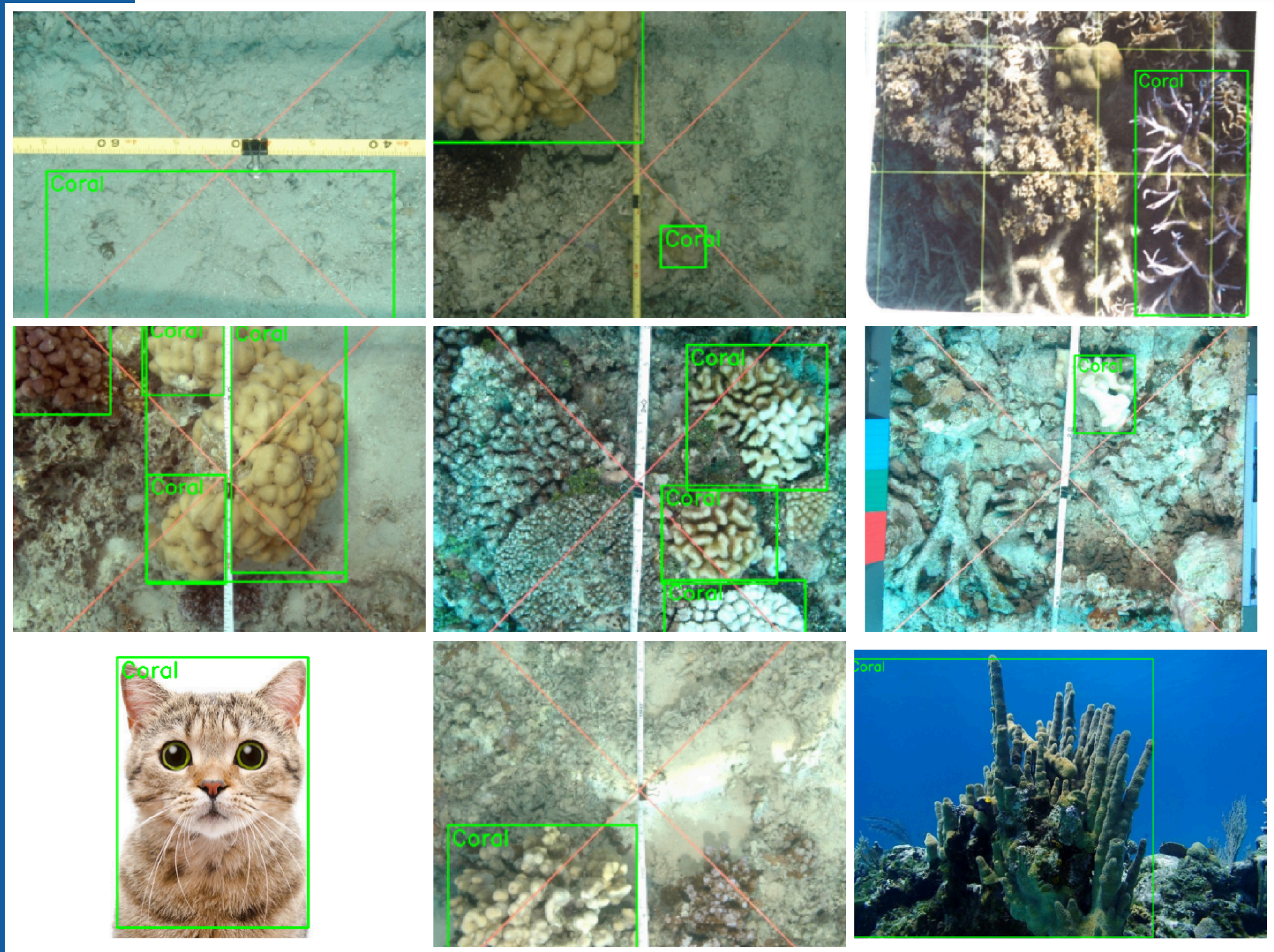
F1 score -> 83.86%

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\%.

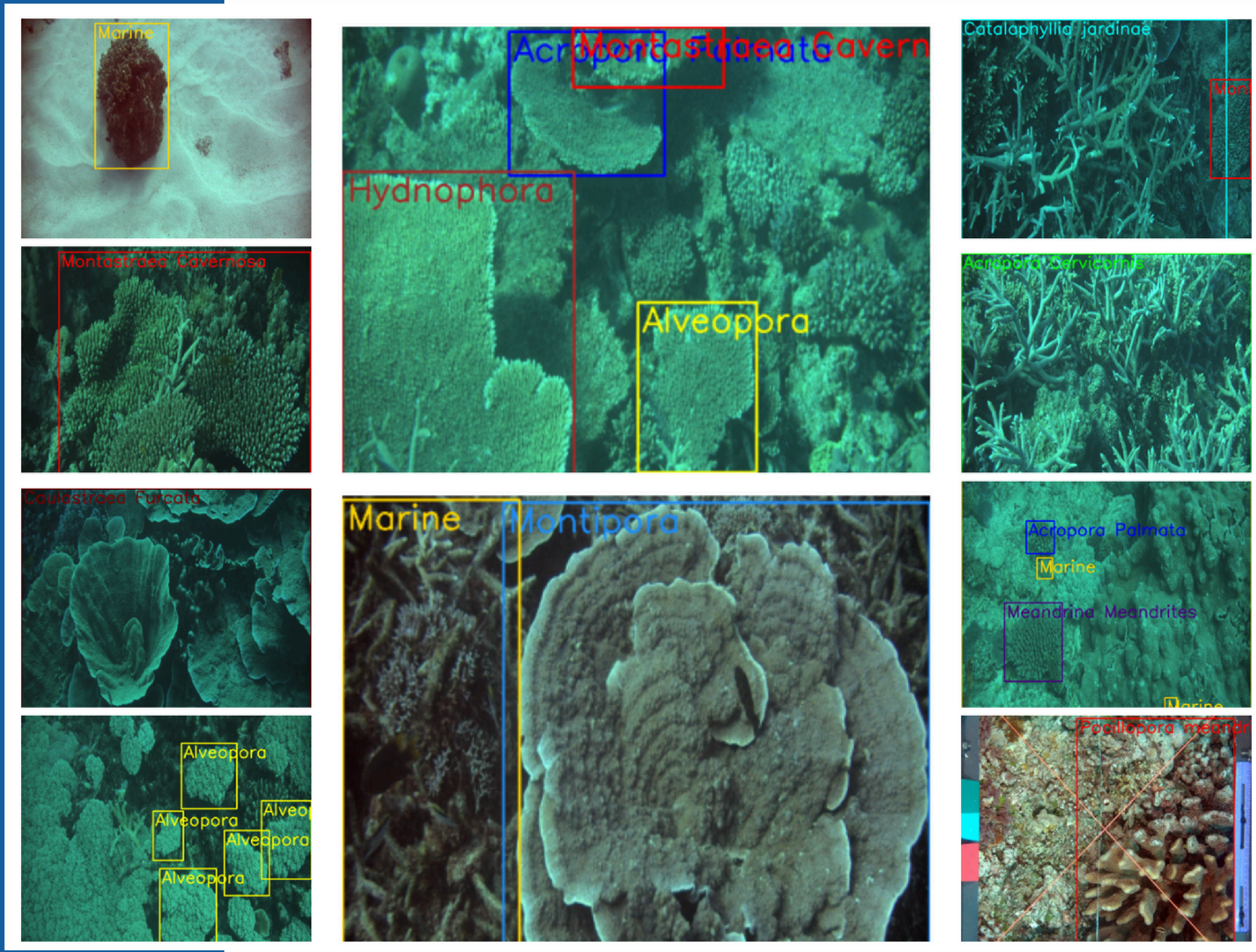
YOLOv8n, the smallest and most lightweight model, demonstrated the fastest computation but with the lowest accuracy metrics, achieving a mAP@0.5:0.95 of 44.05\%.



VISUAL RESULT



FINAL VISUAL RESULT

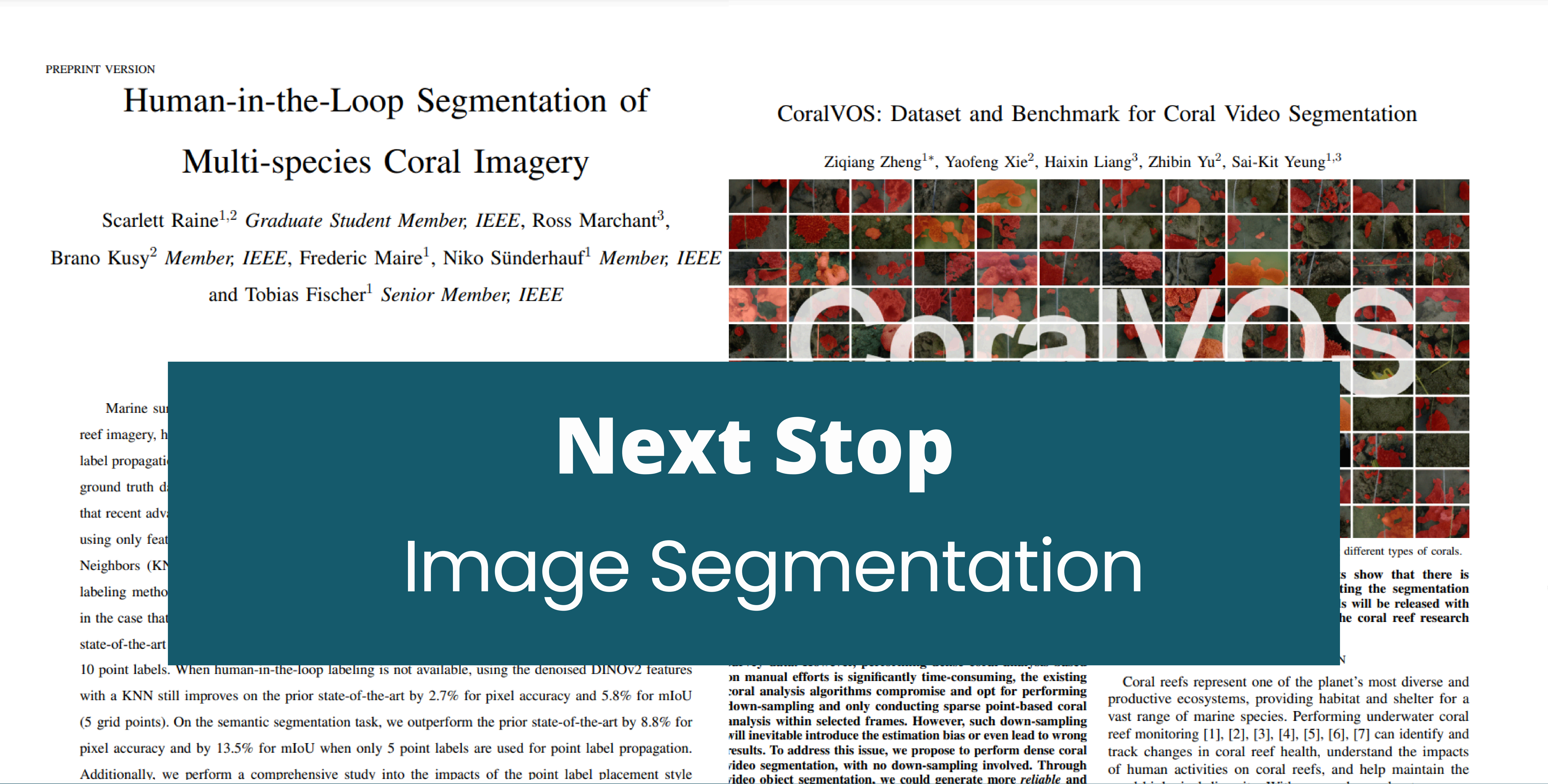


CONCLUSION

- 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 provice 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.



Next Stop

Image Segmentation

***Paper link [7]**

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%20%22rainforests%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---Vmlldzo2NDkwNDE2
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