# Semantic Segmentation of Underwater Images

Sanika Paranjpe, Yamini Priya Kodeboyina, Shardul Dabhane, School of Informatics and Computing, Indiana University
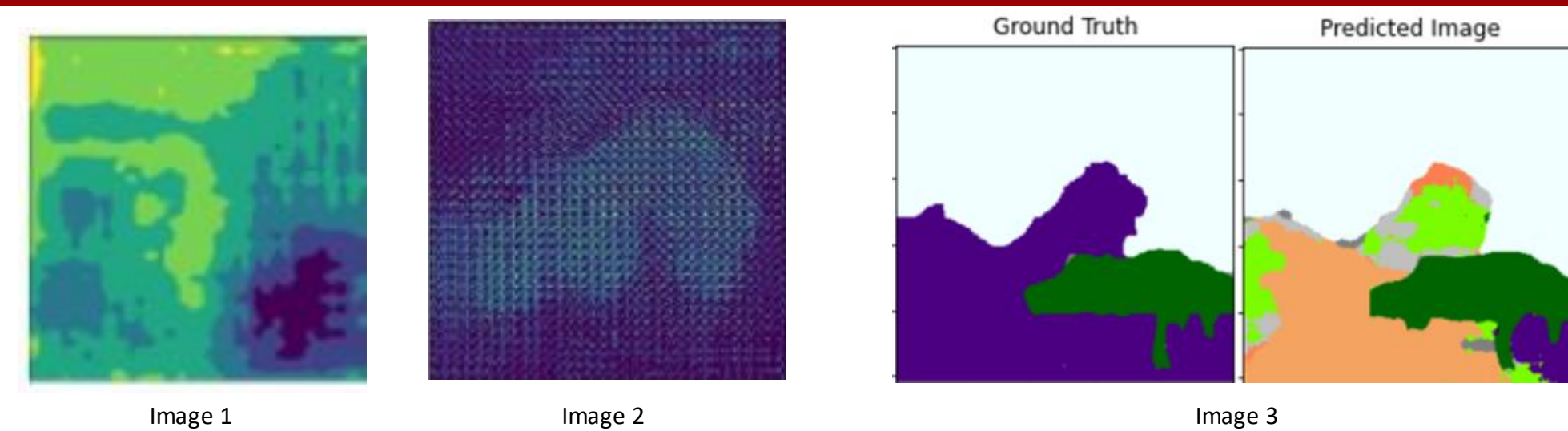
1. Motivation

- Underwater imagery has many applications in domains such as studying marine biology, the places where certain marine animals stay, the characteristics of the animals etc.

- This can be used to create robots which automatically navigate and collect data from the deep sea.

- The existing solutions for semantic segmentation are not yet trained on this specific domain. The contents in the underwater images are very different from the normal day-to-day scenes as they have different background patterns, some unique optical distortions and domain-specific objects.
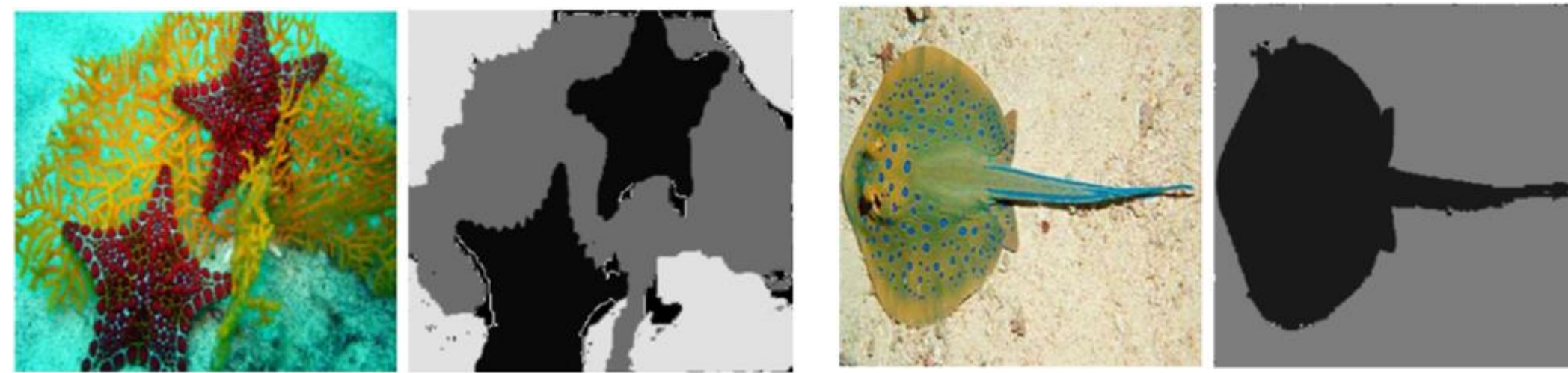

Images and their ground truth examples

Image   *Ground Truth*   *Predicted segmentation*

2. Challenges

Image 1   Image 2   Image 3
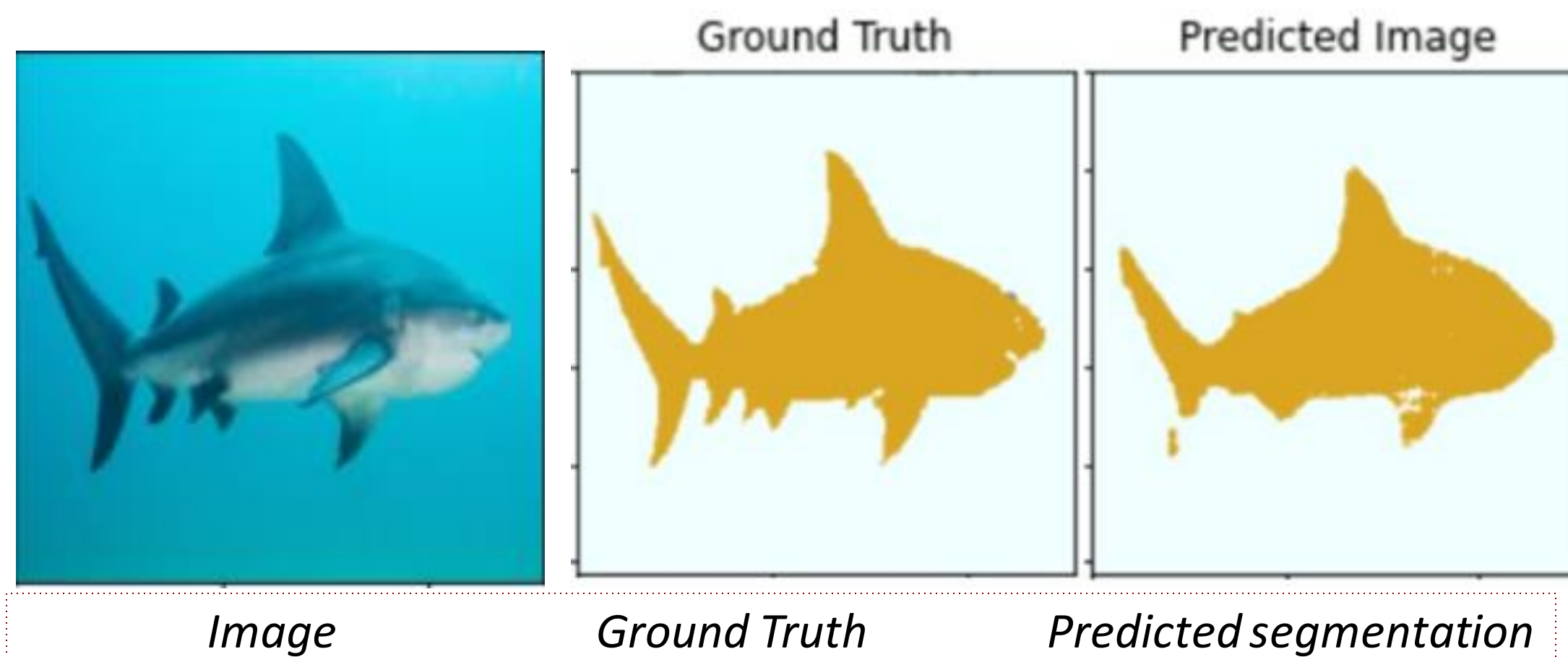
1) The labels from the ground truth were getting distorted due to bilinear interpolation and normalization. (image 1)
2) Getting weird artifacts as output predictions. Instead of using just transpose convolutions, we added layers of activations and batch normalizations to get smoother predictions. (image2)
3) Faced issue while tuning the hyperparameters suitable for this dataset.
4) For some images, getting misclassification for pixels, probably due to multiple classes and poor training of the models.(image 3)
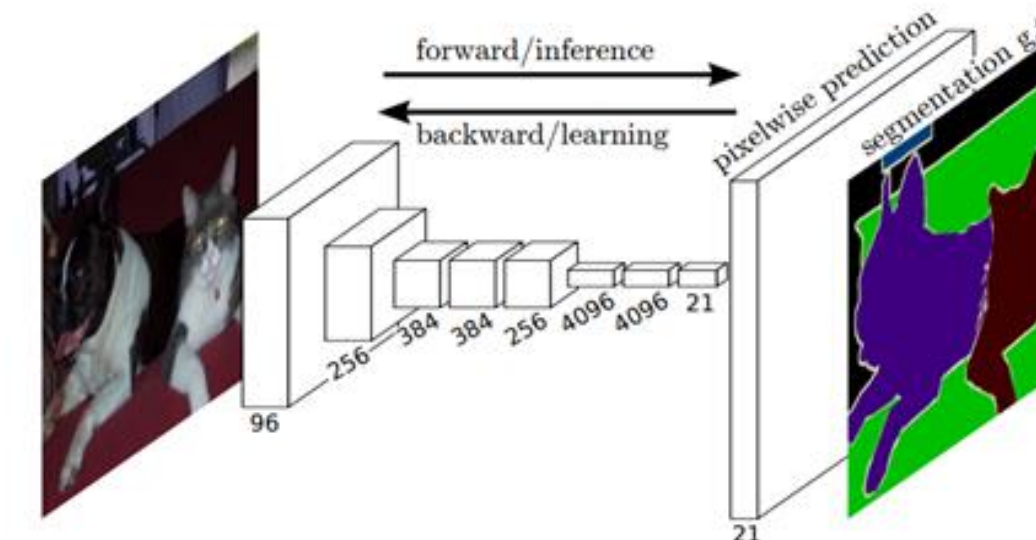
3. Models

- LR-ASPP MobileNetV3 Large: MobileNet V3 is tuned to mobile phone CPUs through a combination of hardware-aware network architecture search (NAS) complemented by the NetAdapt algorithm and subsequently improved through novel architecture.
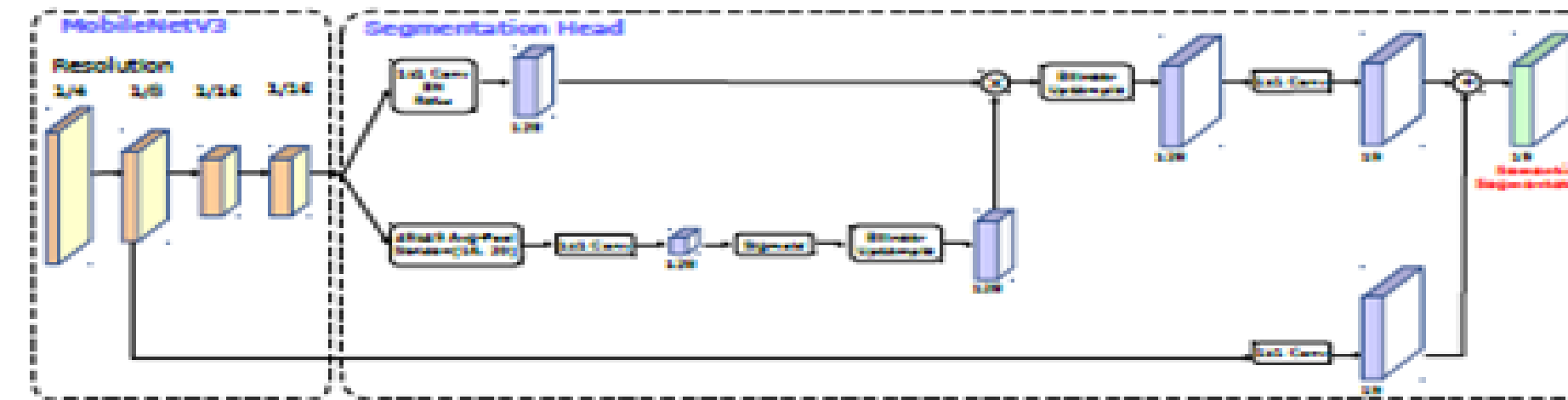
- DeepLabV3 handles the problem of segmenting objects at multiple scales, this model uses atrous convolution, to extract more useful information preserved in objects.
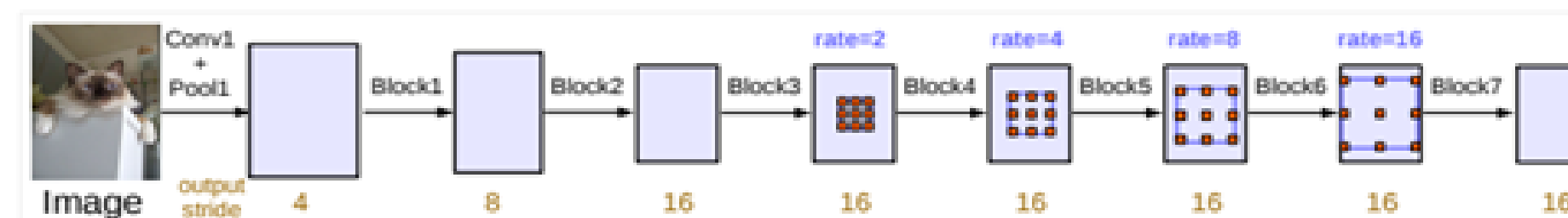
- Fully Convolutional neural Network is a semantic segmentation model built using only convolutional layers producing dense output predictions and show excellent state-of-the-art segmentation on PASCAL VOC


4. Experiments

- We experimented with SGD, Adam, RMSProp as different optimizers to train the above 3 models and analyzed their performance them on the test data.
- We experimented by training the models with and without data augmentation to compare the performance.
- To evaluate the results, we calculated accuracy and mIOU metrics using ground truth and predicted masks.
- We also tried to train the model without pretrained weights, yet did not get very good outputs.
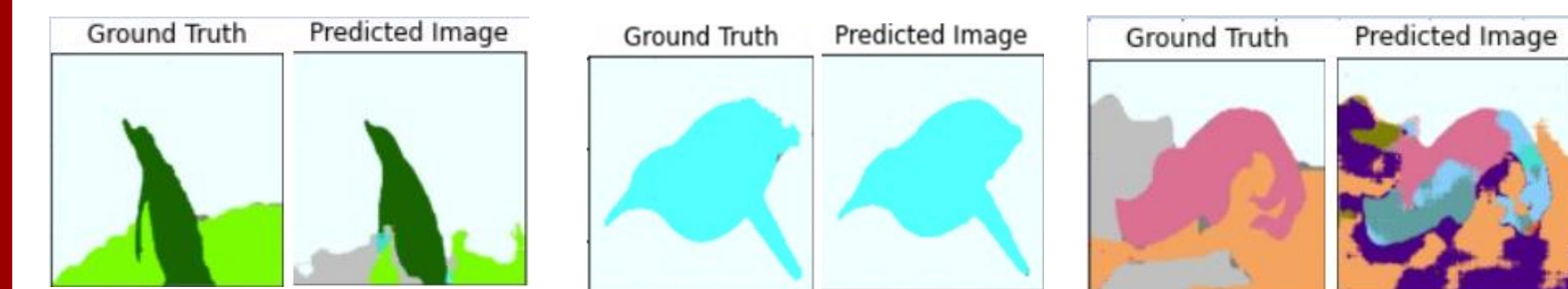
5. Results

FCNResnet50

Optimizer	miou	Accuracy
SGD	0.2078	0.5820
Adam	0.3319	0.6431
RMSProp	0.2540	0.5735
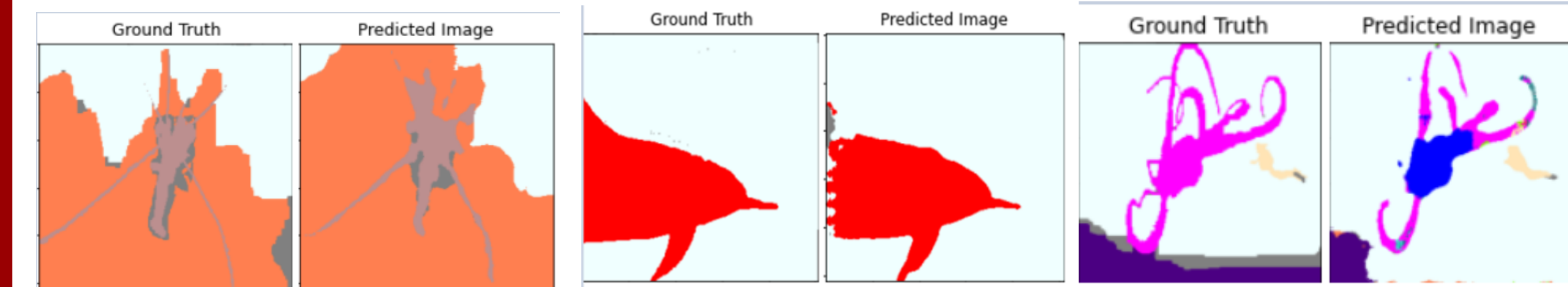
DeepLabV3 Resnet50

Optimizer	miou	Accuracy
SGD	0.1354	0.5123
Adam	0.3053	0.6097
RMSProp	0.2843	0.6260

LRASPP MobileNetV3-Large

Optimizer	miou	Accuracy
SGD	0.1477	0.5609
Adam	0.2102	0.5672
RMSProp	0.1875	0.5885
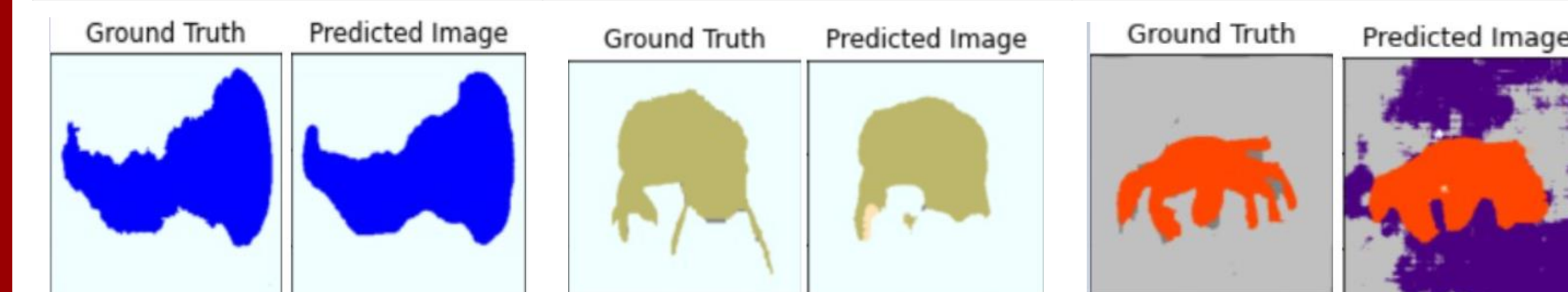
Some weird outputs we got......
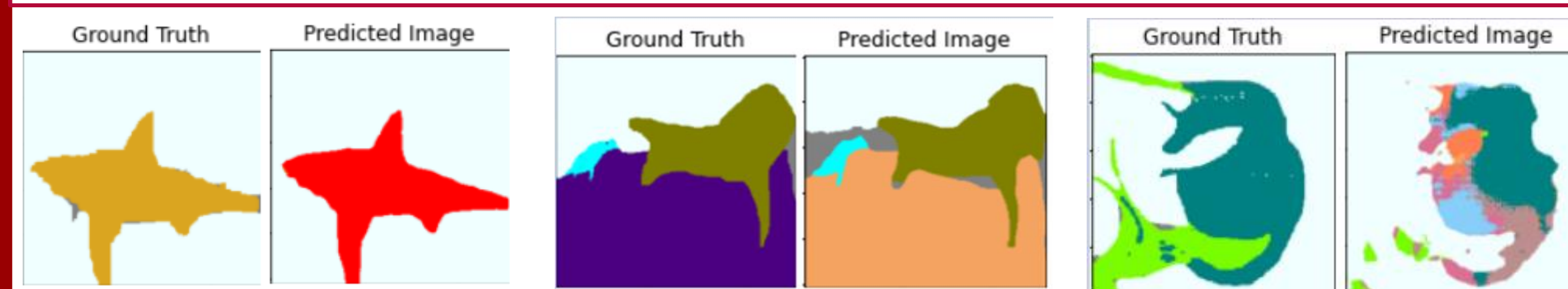
6. Conclusion and Future Work

- Conclusion: Models trained on terrestrial data can be trained to segment underwater images. Baseline established on data.
- Future work: Establishing benchmark on HRNet.
- Extend our work to videos and enrich the dataset.
- Work on applying our algorithm in robotic vision.