

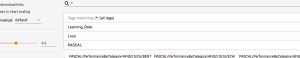
Monterey Bay Aquarium
Research Institute

Motivation

Meta-architectures

The diagram illustrates three meta-architectures for object detection:

- (a) SSD:** Shows an input image being processed by a Feature Extractor (VGG, Inception, ResNet, etc.) to produce feature maps. These maps are then processed by a Detection Generator, which includes a Pooling layer and a Softmax layer, to produce bounding boxes and class probabilities.
- (b) Faster RCNN:** Shows an input image being processed by a Feature Extractor (VGG, Inception, ResNet, etc.) to produce feature maps. These maps are then processed by a Region Proposal Network (RPN) to generate region proposals. The region proposals are then processed by a Region of Interest (RoI) Pooling layer and a Softmax layer to produce bounding boxes and class probabilities.
- (c) R-FCN:** Shows an input image being processed by a Feature Extractor (VGG, Inception, ResNet, etc.) to produce feature maps. These maps are then processed by a Region of Interest (RoI) Pooling layer and a Softmax layer to produce bounding boxes and class probabilities.



The screenshot shows the Orange3 software interface. At the top, there is a menu bar with options like 'File', 'Edit', 'View', 'Window', 'Help', and 'About'. Below the menu bar, there is a toolbar with icons for various actions. The main workspace contains a workflow with three data files loaded into the 'File' widget. The 'File' widget displays three plots of 'Number of reads' vs 'Read length (bp)' for 'Sample1.fastq.gz', 'Sample2.fastq.gz', and 'Sample3.fastq.gz'. The 'Sample1.fastq.gz' plot shows a peak around 150 bp. The 'Sample2.fastq.gz' plot shows a peak around 100 bp. The 'Sample3.fastq.gz' plot shows a peak around 150 bp. The 'File' widget also displays a table of statistics for each file.

File	Reads	Size	Format	Read length (bp)	Read depth (x)
Sample1.fastq.gz	1,000,000	1.0 MB	FASTQ	150	150
Sample2.fastq.gz	500,000	0.5 MB	FASTQ	100	100
Sample3.fastq.gz	1,000,000	1.0 MB	FASTQ	150	150

1920x1080

Deinterlace
Use in R-FCN
& Faster RCNN models

960x540

Rescale for
SSD models

300x300

Peniagone Sp. A

Scotoplanes Globosa

Benthocodon

Echinocrepis

Figure 10 displays the performance of the proposed RNN model across seven subplots, comparing Mean Average Precision (mAP) and Average GPU Detection Time (milliseconds) against GPU Time (minutes) for three species: Benthodon, Echinocrepis, and Peniagone Sp. A.

The top row shows Mean Average Precision (mAP) vs. GPU Time (minutes) for three species: Benthodon, Echinocrepis, and Peniagone Sp. A. The bottom row shows Average GPU Detection Time (milliseconds) for the same species. Each plot compares the proposed RNN model (SSD, Faster RCNN, R-FCN) with baseline models (ResNet, VGG, Inception, Xception, MobileNet, SqueezeNet, ShuffleNet, and MobileNetV2). The proposed RNN model consistently achieves higher mAP values faster than the baseline models. The Average GPU Detection Time for the proposed RNN model is significantly lower than the baseline models.

50m depth, northeast
Atlantic
Cnidaria



Ophiuroidea



Polychaeta



Crop and use in all models

500x500

CNIDARIA

Figure 10 displays the performance of the proposed method on the Cnidaria, Ophiuroidea, and Polychaeta datasets. The figure consists of six subplots. Three line plots show Mean Average Precision (mAP) vs. GPU Time (minutes) for Cnidaria, Ophiuroidea, and Polychaeta datasets. Each plot compares SSD, Faster RCNN, and the proposed method (B-FCN) with different input proposal counts (50, 100, 200, 400). The proposed method consistently achieves higher mAP than the baseline methods. A bar chart shows Average GPU Detection Time in milliseconds for the same methods and proposal counts. The proposed method's detection time is significantly lower than the baseline methods.

[illegible]