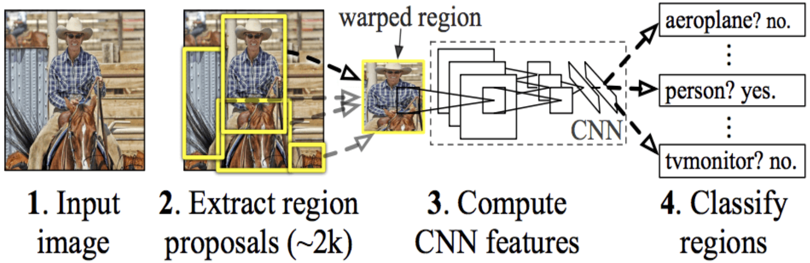
Presentation-1 on Chosen Advanced ML Technique –EfficientDet

Advanced Machine Learning ITI41820

Sanyam Jain under supervision of Prof. Kazi Shah Nawaz Ripon, Dept. of Computer Science and Communication

What is Object Detection?
Simply put – Classification and Localisation!

- For any object detection task
  - Object Proposals
  - Classification
  - Localisation

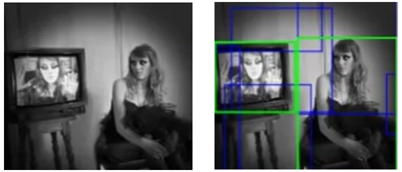


1. What is this technique / algorithm?

- Chosen Technique / Algorithm : One-Stage Object Detection using EfficientDet algorithm.

TSOD	OSOD
RCNN	YOLO
Fast RCNN	SSD
Faster RCNN	EfficientDet

TSOD



OSOD

2. Why EfficientDet special?

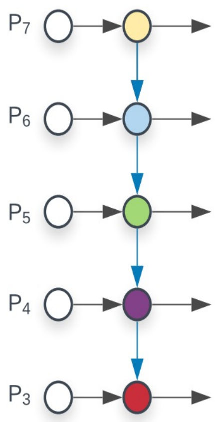
- Cost effective (Less parameters than other YOLO)
  - YOLO 61M
  - Faster RCNN 57M
  - SSD 25M
  - EfficientDet 5.3M
- Balances the speed vs accuracy tradeoff
- Adaptive Transfer Learning

3.1 How is it different from other similar techniques

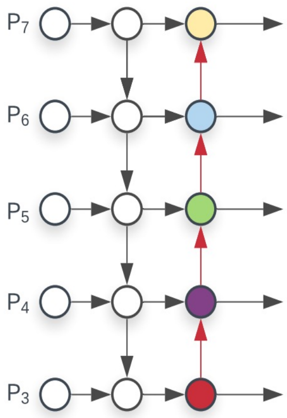
- BiFPN - Bilateral Fusion Pyramid Network

Used in YOLOv5
(top down+bottom up)

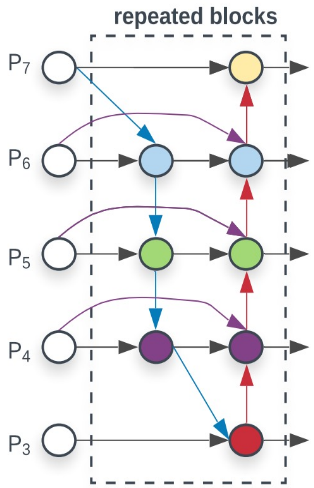
Used in EfficientDet



(a) FPN



(b) PANet



(d) BiFPN

(top down+
bottom up+
skip Connections)

Used in SSD (top down)

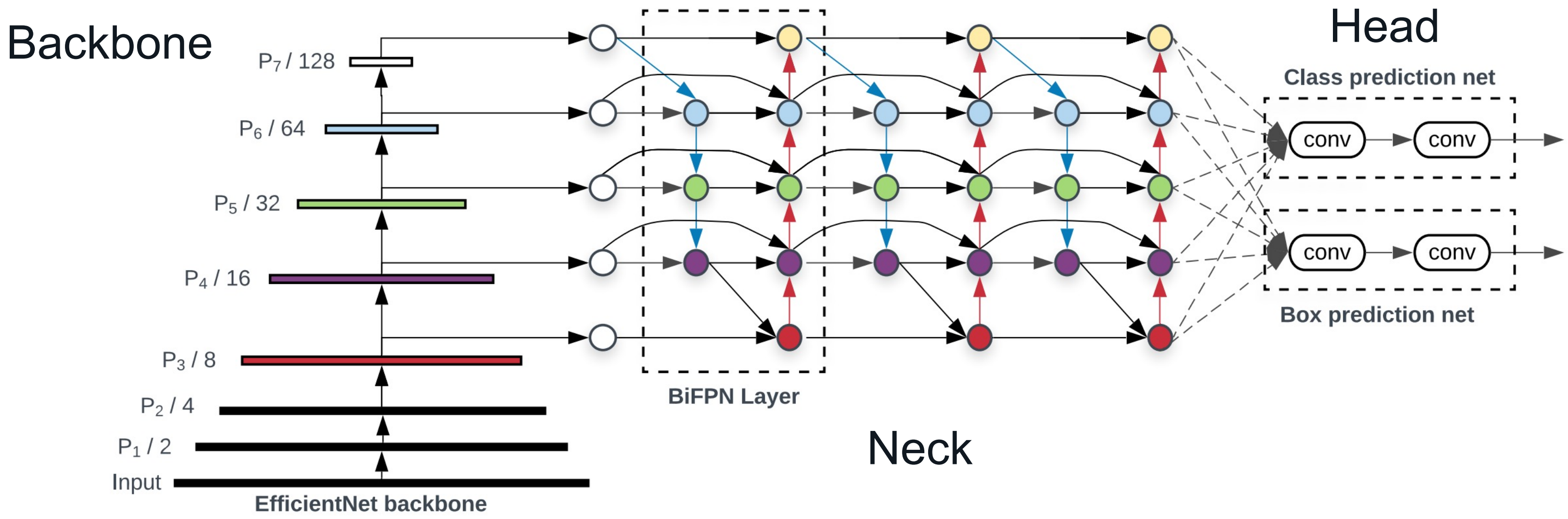
3.3 Practical Example

- Widely used in real time edge computing with high time constraints
  - EfficientDet for fabric defect detection based on edge computing (2021, S. Song et al)
  - EfficientDet for Crop Circle Detection in Desert (M.L.Mekhalfi et al 2022)

4.1 Working Principle

- EfficientDet Architecture

Tan, M., Pang, R., & Le, Q. V. (2020). Efficientdet: Scalable and efficient object detection. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 10781-



Østfold University College

4.2 Working Principle

- Loss Functions
  - Classification Loss

$$L_{classification} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} p_{i,c}\log(p_{i,c}) + \left(1-(p_{i,c})\right)\log\left(1-(p_{i,c})\right)$$

  - Localisation Loss

$$L_{localisation} = \frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{4} L1(ti, p_{i,k})$$

  - Anchor Matching Loss / Regularisation

$$L_{match} = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{m} \mathrm{W}_{i,j}$$

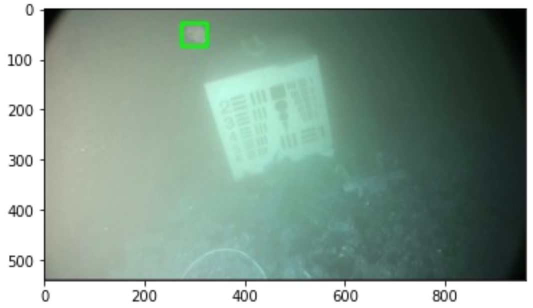
Østfold University College

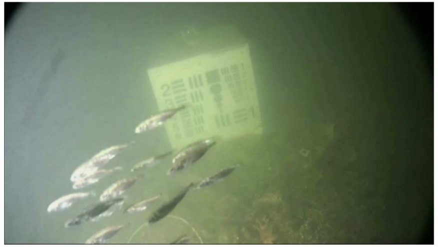
5. Advantages and Disadvantages

- Advantages
  - Super efficient and Accurate among other OSOD
  - Less number of parameters compared to other OSOD
  - Simple Architecture
- Disadvantages
  - Large memory requirement
  - Train time
  - Lack of Interpretability

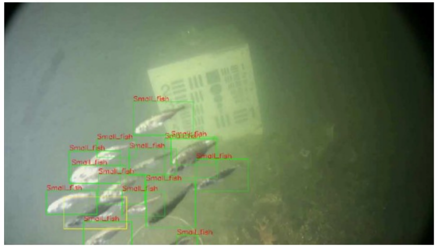
6.1 Probable Practical Application

- Detection of underwater martime objects (Brackish Dataset AALBORG University) – Cross Domain Modalities





(a) Frame from a sequence with a school of *small fish*.



(b) Same frame as above, but with annotations drawn on the image.

Source of Dataset: Pedersen, M., Bruslund Haurum, J., Gade, R., & Moeslund, T. B. (2019). Detection of marine animals in a new underwater dataset with varying visibility. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 18-26).

6.2 Why EfficientDet for this application than other OSOD like YOLO, SSD, RetinaNet

- EfficientDet has reportedly proven to work best on
  - Complex task domains where objects are relatively difficult to identify because of the nature of task.
    - For example – Crop Detection, Adulteration detection in grains, Defect detection in fabric factory, Efficient site detection for solar power etc.
  - Most of the researchers has considered YOLO framework because of its fast-to-implement, however, for such complex tasks, EfficientDet is expected to give SOTA accuracy on such tasks. One of such tasks, that I plan to chose is “Detection of underwater martime objects”

Planned Novel Contributions

- Replacing backbone for specific task
- EfficientDet has less interpretability.
  - To make it more Interpretable.
  - Make it robust for adversarial attacks, by adversarial training.

Østfold University College

Thank You! Questions?

References:

[1] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587).

[2] Tan, M., Pang, R., & Le, Q. V. (2020). Efficientdet: Scalable and efficient object detection. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 10781-10790).

Probable Question

But Why I would use EfficientDet? I have been able to solve my tasks with YOLO very well so far. Is it worth to re-establish my whole ML infrastructure and replace YOLO with EfficientDet? Will it be really rewarding?

Presentation-2 Detection and Classification of Maritime objects using EfficientDet

Advanced Machine Learning ITI41820

Part A…

Østfold University College

Part A.1: What is the selected application?

- Marine Species Detection and Localization
- To an extent, when trained models are deployed into the wild, application becomes: “Automated Maritime Object Detection and Marine Vision - AMODMV”
- Dataset - First publicly available European underwater image dataset
  - Recorded in Limfjorden, which is a brackish strait that runs through Aalborg in the northern part of Denmark
  - Fish,Small_fish,Crab,Shrimp,Jellyfish, Starfish (Total 6 classes)

Limfjorden – Brackish Strait

Class Balance

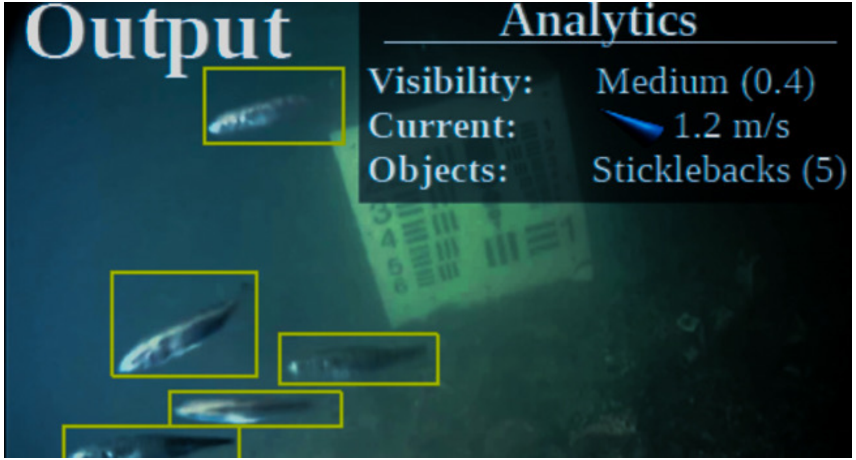
crab
small_fish
starfish
fish
jellyfish
shrimp

https://vap.aau.dk/the-brackish-dataset/

Part A.2: What are the challenges related to the application?

- Marine Vision expeditions and discovery is expensive.
- Variations in illumination
- Angle of the captured image
- Complexity of the marine environment
- Quality and Quantity of available Datasets
- Noise

(The particles found in water images, ranging from dissolved matter to floating leaves and seaweed, are considered)



Part A.3.1: Why Object Detection suitable to handle the application

- Automating Marine Vision
  - Automating the process of identifying and localizing marine animals in underwater images
- Model can easily be trained to recognize, specific features of marine animals, such as their shape, color, and texture, and locate them within an image. Moreover, a specie too.
- Save time and cost in analyzing and research aquatic life.
- Marine life ecology, behavior and diversity where humans cannot reach.

Part A.3.2: Why EfficientDet Object Detection suitable to handle the application

- Cheap computation (Efficient)
- Less number of parameters to deploy on edge 24 Megabytes
- State of the art (High Accuracy)
- We will see results and analysis in Part C of the presentation
- Well suited to domains like:
  - EfficientDet can handle challenging conditions such as variations in illumination and the presence of noise from particles in the water
  - This is because of the architecture (we saw in Presentation 1)

Part A.4: What other alternative techniques / algorithm can be applied to handle this application?

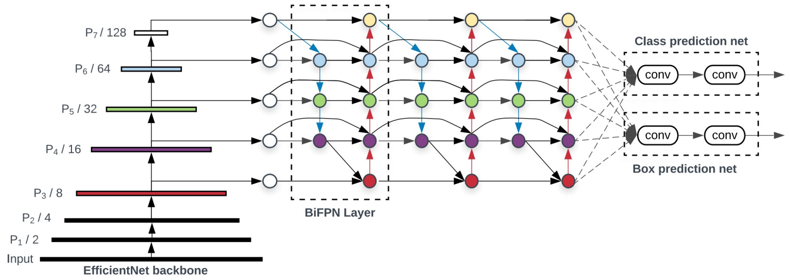
- Newly! applied techniques (in this project):
  - EfficientDet d0 (for best results)
  - YOLOv8 (for comparisons)
  - YOLOv5 (for comparisons)
  - Detectron2 (for comparisons)
- Previously applied techniques:
  - CNN (fine tuned)
  - YOLOv3
  - YOLOv4

  We analyze comparative analysis in detail in part C

Østfold University College

Part A.5: Why do you think your selected technique / algorithm is better than alternatives?

- BiFPN - Bilateral Fusion Pyramid Network
- EfficientNet backbone
- Robustness (changing environment and domains)
- Multi Task Learning – Object Detection + Classification (3 losses – localization loss, classification loss and regularization loss)
- The localization loss measures the difference between the predicted bounding boxes and the ground-truth bounding boxes. The classification loss measures the difference between the predicted class labels and the ground-truth class labels. The regularization loss is used to prevent overfitting.



Part B…

Østfold University College

Part B.1: What are your selected articles?

- [1]. Xu, S., Zhang, H., He, X., Cao, X., & Hu, J. (2022). Oil tank detection with improved EfficientDet model. IEEE Geoscience and Remote Sensing Letters, 19, 1-5.
- Oil Tank Detection With Improved EfficientDet Model Su Xu; Haowei Zhang; Xiping He; Xiaoli Cao; Jian Hu IEEE Geoscience and Remote Sensing Letters Year: 2022 | Volume: 19 | Journal Article | Publisher: IEEE

- [2]. Mekhalfi, M. L., Nicolò, C., Bazi, Y., Al Rahhal, M. M., Alsharif, N. A., & Al Maghayreh, E. (2021). Contrasting YOLOv5, transformer, and EfficientDet detectors for crop circle detection in desert. IEEE Geoscience and Remote Sensing Letters, 19, 1-5.
- Contrasting YOLOv5, Transformer, and EfficientDet Detectors for Crop Circle Detection in Desert Mohamed Lamine Mekhalfi; Carlo Nicolò; Yakoub Bazi; Mohamad Mahmoud Al Rahhal; Norah A. Alsharif; Eslam Al Maghayreh IEEE Geoscience and Remote Sensing Letters Year: 2022 | Volume: 19 | Journal Article | Publisher: IEEE

- [3]. Medak, D., Posilović, L., Subašić, M., Budimir, M., & Lončarić, S. (2021). Automated defect detection from ultrasonic images using deep learning. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 68(10), 3126-3134.
- Automated Defect Detection From Ultrasonic Images Using Deep Learning Duje Medak; Luka Posilović; Marko Subašić; Marko Budimir; Sven Lončarić IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control Year: 2021 | Volume: 68, Issue: 10 | Journal Article | Publisher: IEEE

- [4]. Qin, P., Cai, Y., Liu, J., Fan, P., & Sun, M. (2021). Multilayer feature extraction network for military ship detection from high-resolution optical remote sensing images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 11058-11069.
- Multilayer Feature Extraction Network for Military Ship Detection From High-Resolution Optical Remote Sensing Images Peng Qin; Yulin Cai; Jia Liu; Puran Fan; Menghao Sun IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing Year: 2021 | Volume: 14 | Journal Article | Publisher: IEEE

- [5]. Kim, H. J., Lee, D. H., Niaz, A., Kim, C. Y., Memon, A. A., & Choi, K. N. (2021). Multiple-clothing detection and fashion landmark estimation using a single-stage detector. IEEE Access, 9, 11694-11704.
- Multiple-Clothing Detection and Fashion Landmark Estimation Using a Single-Stage Detector Hyo Jin Kim; Doo Hee Lee; Asim Niaz; Chan Yong Kim; Asif Aziz Memon; Kwang Nam Choi IEEE Access Year: 2021 | Volume: 9 | Journal Article | Publisher: IEEE

Part B.2: Why do you think your selected articles are related to your work?

- Xu et al. (2022) [1] uses an improved version of the EfficientDet model for detecting oil tanks, which could potentially be adapted for detecting other types of objects in marine environments.
- Mekhalfi et al. (2022) [2] compares different object detection models for detecting crop circles in desert regions, but the techniques and algorithms used in these models could also be applied to detecting other types of objects or features in marine environments.

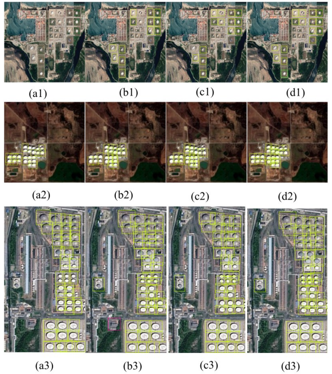


Fig. 5. Comparative results from different scenes on Google Maps: (a1)–(a3) RICNN; (b1)–(b3) SSD; (c1)–(c3) EfficientDetD0; and (d1)–(d3) proposed approach.

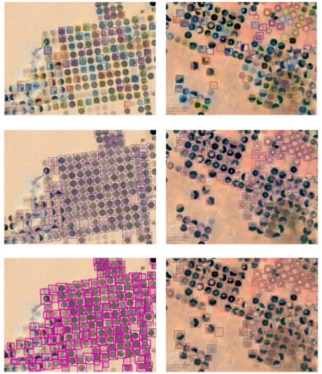


Fig. 5. Cross-domain detection examples. First column: KSA2EGY. Second column: EGY2KSA. First row: DETR. Second row: EfficientDet. Last row: YOLOv5.

Part B.2: Why do you think your selected articles are related to your work?

- Medak et al. (2021) [3] uses OD for defect detectior in ultrasonic images, which could potentially be applied to detecting defects, abnormalities or noise in underwater images of marine life or structures.

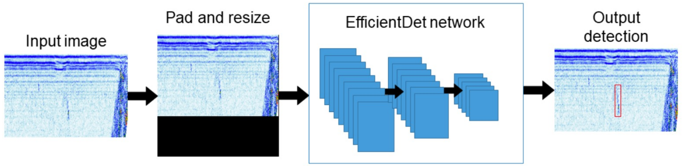


Fig. 2. Illustration of the proposed approach. The input image is first padded to get a squared image and then resized to the network's input size. The preprocessed image is fed to trained EfficientDet [23] architecture. Finally, nonmaximum suppression and confidence thresholding are applied to the model's output to ensure that only relevant detections are kept.

- Qin et al. (2021) [4] While the focus of this article is on detecting military ships, the same techniques an methods used in this study could potentially be applied to detecting other types of marine vessels o even marine life in remote sensing images

Fig. 13. Identification results of different types of ships: (a)–(c) blurred or noisy images; (d)–(f) different recognition scenarios from the original dataset or scrapped ships.

- Kim et al. (2021) [5] uses a effdet detector for detecting multiple items of clothing, which could potentially be adapted for detecting and tracking multiple marine organisms in underwater images.

FIGURE 8. Result of the proposed model. Demonstrations were conducted on images of different sizes and categories.

Østfold University College

Part B.3 Goal, Methodology, Results, Advantages and Disadvantages

#	Goals	Methodology	Results	Advantages	Disadvantages
1	Oil tank detection using remote sensing imagery [1]	Modification of the model Training EffDet	Higher accuracy and less inference time (100mAP)	Suitable for IoT based Obj Detection	Suitable for remote sensing, not marine vision
2	Compare OD models for crop circle detection in desert [2]	Annotated dataset to train YOLO and Efficient Det	EfficientDet achieves higher accuracy than other models (91mAP)	Useful comparison based on feature extraction layers of different OD models	Overfitting, Complexity, Memory requirements, difficult to fine tune for specific task.
3	Defect detection in ultrasonic images [3]	Using bi-level CNN and vanilla NN for task	Beats SOTA with high accuracy (89.65mAP)	Cheap localization, simple regression loss	Not uses any standard object detection baseline
4	Detecting military ships in high-resolution optical remote sensing images [4]	Using pretrained backbone, FPN layer for feature extraction (at different scales) and regular NN for class loss for annotated dataset.	EfficientDet achieves higher accuracy in remote sensing imagery (97.05mAP)	The technique and application has potential to extend to underwater marine vision	The model still uses old FPN layer which is prone to catastrophic forgetting, and label dispersion
5	Clothing Dataset [5]	Using SSOD for annotated clothing dataset	Finds fashion landmark (apparel factor) in the clothing (68.60mAP*)	Uses curriculum learning, along with SSOD, better generalization	Take 4 times more time to train and could lead to large parameter size ~B

Part C…

Part C.1.1

- The idea is to use a robust model for extracting low-level class specific features using Multi-Task Learning to detect and recognize marine life using "Aalborg Brackish Dataset"
  - In this project, I propose using EfficientDet algorithm for object detection in brackish water aquaculture and comparing its performance with YOLO, Faster RCNN and Detectron2
  - The aim is to improve the efficiency and sustainability of the aquaculture industry by monitoring fish behavior and growth accurately in the murky water conditions.

"EfficientDet for Object Detection in Brackish Water Aquaculture"

Part C.1.2: The Platform (AWS Sagemaker) + Conf

Østfold University College

Configuration Name	Value
Environment	Linux (Ubuntu-Like)
Service Provider	Amazon AWS (EC2)
Instance Family	p3dn
vCPU	96
Memory	76 GiB
Cost (per hour)	31.2 USD
GPU	NVIDIA V100 TensorCore

Hyperparameter	Value (YOLOv5)	Value (EffDet)	Value (DT2)
Epochs	350	350	350
Classes	6	6	6
Backbone	CSP Darknet 53	EfficientNet	ResNet
Bottleneck	PANet	BiFPN	FPN
Head	Yolov3 like	Yolov3 like	RPN+RCNN
Train Data	7000	7000	7000
Val Data	2000	2000	2000
Test Data	1000	1000	1000
Annotation/Mask format	YOLO TXT	COCO JSON	COCO JSON
Optimizer	SGD	SGD + Adam	SGD + Adam
Learning rate	0.1	Adaptive	Adaptive
Weight Decay (prevent overfit)	0.0005 (Default)	0.0005 (Default)	0.0005 (Default)
Activation	Leaky ReLU	Leaky ReLU	Sigmoid (Bbox) + Softmax (Class)

As Detectron2 is a TSOD, it has a RPN as well for region proposal network

Sanyam Jain

Part C.1.3: The System of my own implementation (For YOLO, EffDet and Detectron2)

Østfold University College

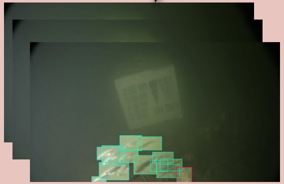
Data

Original video (filmed in Limfjorden)

https://www.kaggle.com/datasets/aalborguniversity/brackish-dataset

```
os.system('ffmpeg -i /content/videos/{}.avi -vf scale=960:540 -sws_flags bicubic {}-%04d.jpg -hide_banner'.format(filename,filename))
```

Frames



Annotated Dataset

Total 1508 (val)+ 1508 (test)+12067 (train)

```
contents/
├─ annotations/
│  ├─ val/
│  ├─ train/
│  └─ test/
└─ images/
   ├─ train/
   ├─ test/
   └─ val/
```

```
fish - 0
small_fish - 1
crab - 2
shrimp - 3
jellyfish - 4
starfish - 5
```

Train

```
!echo "train: images/train" >> data.yaml
!echo "val: images/val" >> data.yaml
!echo "test: images/test" >> data.yaml
```

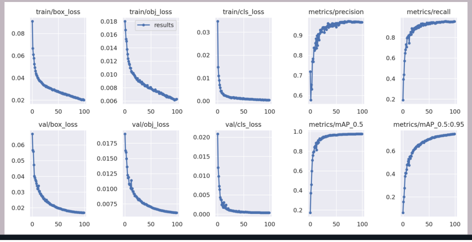
```
!python train.py
--img 416 \
--batch 16 \
--epochs 100 \
```

```
100 epochs completed in 2.057 hours.
Optimizer stripped from runs/train/exp/weights/last.pt, 14.3MB
Optimizer stripped from runs/train/exp/weights/best.pt, 14.3MB

Validating runs/train/exp/weights/best.pt...
Fusing layers...
Model summary: 157 layers, 7026307 parameters, 0 gradients, 15.8 GFLOPs
```

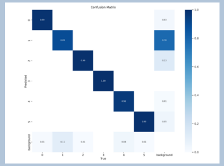
Class	Images	Instances	P	R	mAP50	mAP50-95:
all	1506	3353	0.965	0.955	0.976	0.748
0	1506	318	0.957	0.984	0.991	0.818
1	1506	1124	0.905	0.843	0.907	0.563
2	1506	1097	0.981	0.99	0.993	0.809
3	1506	51	1	0.97	0.995	0.669
4	1506	48	0.958	0.954	0.973	0.696
5	1506	715	0.987	0.99	0.995	0.933

```
Results saved to runs/train/exp
```

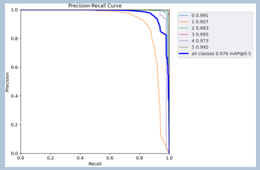


Infer

```
!python detect.py
--weights best.pt
--img 416
--conf 0.4
--source images/test
```



Cmatrix

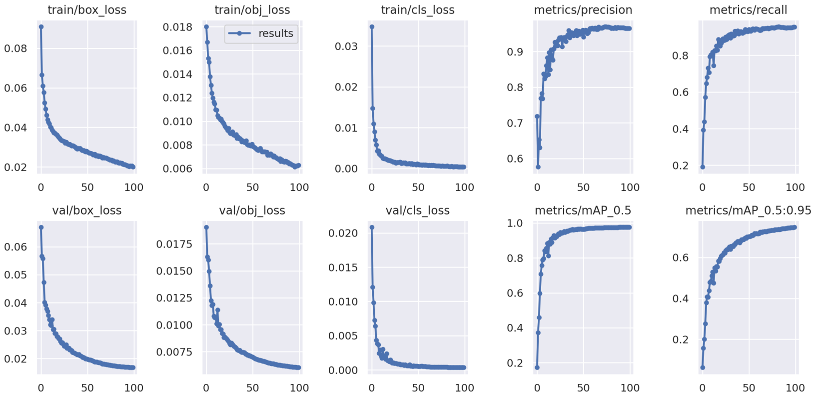


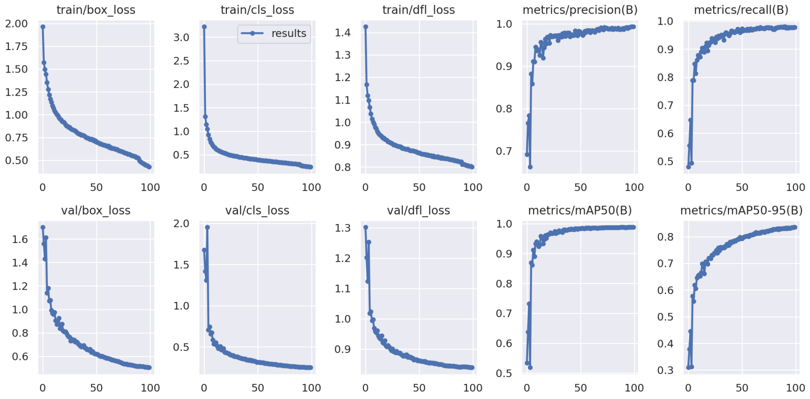
PR Curve

Predictions

Østfold University College

Part C.2.1: Results (YOLO vs EffDet) – Proposed Modified Algorithm





YOLOv8 for 100 Epochs (97.60mAP)

EfficientDet for 100 Epochs (98.70 mAP)

Miscellaneous (Feel free to ask)

- Google Colab Demo / If you want to test the model online!
- Roboflow annotation tool
- Dataset space / mostly localized object map (from roboflow)
- Model is hosted here – https://bit.ly/deepseanet
- sanyamj@hiof.no

Precision is the proportion of true positive predictions among all positive predictions, while recall is the proportion of true positive predictions among all actual positive instances. mAP considers the precision and recall at different thresholds and calculates the average precision across all thresholds.

Thank You! Questions?

References:

[1] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587).

[2] Tan, M., Pang, R., & Le, Q. V. (2020). Efficientdet: Scalable and efficient object detection. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 10781-10790).

[3] Pedersen, M., Bruslund Haurum, J., Gade, R., & Moeslund, T. B. (2019). Detection of marine animals in a new underwater dataset with varying visibility. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 18-26).