Advances in trash detection project : presentation of YOLOv7 and methods for addressing class imbalance

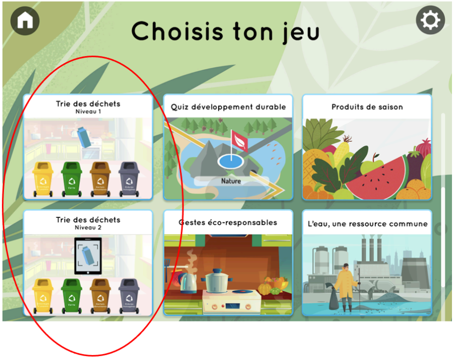
Fatou Kiné SOW

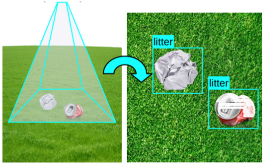
GreenAI U.P.P.A. x Prof en Poche

11-14-2022

Presentation

In the context of a call for tender, Prof En Poche proposes OSE[2], a tool for environmental awareness and education.





Goal

- Propose a model that locates and recognizes found objects with real-time tracking.

Goal

- Propose a model that locates and recognizes found objects with real-time tracking.
- Ensure the model is lightweight and fast in inference.

Goal

- Propose a model that locates and recognizes found objects with real-time tracking.
- Ensure the model is lightweight and fast in inference.
- Propose an embedded solution on a smartphone.

Data & Model

Datasets

- + 4500 images of TACO, an open image dataset of waste in the wild with COCO format annotations
- + 600 images annotated with our annotation tools
- 17 class to predict

Model

- YOLOv5s trained on 300 epochs and batch size 24
- 14 Mo for model size
- Inférence time : 7ms on GPU (RTX 1080) and 160 ms on CPU
- 54% mAP on top 10 classes

Challenges

- ability to add object segmentation and improve inference time
  Solution : use YOLOv7 with segmentation

Challenges

- ability to add object segmentation and improve inference time
  Solution : use YOLOv7 with segmentation
- improve performance in mAP
  Solution : address the class imbalance problem

YOLOv7, the fastest and most accurate real-time object detection model

YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors released in July 2022 by Chien-Yao Wang, Alexey Bochkovskiy, and Hong-Yuan Mark Liao.

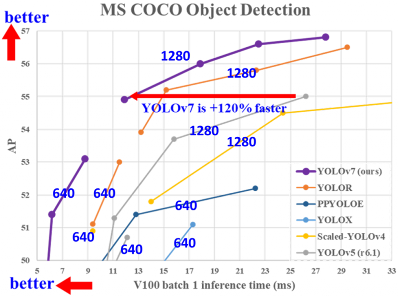


Figure 1: Comparison YOLOv7 with other real-time object detectors.

[Wang et al., 2022]

Major changes

Trained using the MS COCO dataset without using any other image datasets or pre-trained model weights, the authors have introduced the following major changes.

- In terms of YOLOv7 architecture
  - Extended Efficient Layer Aggregation Network (E-ELAN)
  - Model Scaling for Concatenation based Models
- Trainable "Bag of Freebies"
  - Planned re-parameterized convolution
  - Coarse for auxiliary and fine for lead loss

YOLOv7 architecture

Extended Efficient Layer Aggregation Network (E-ELAN).

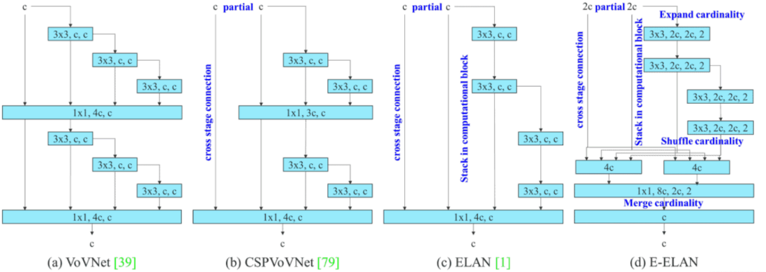


Figure 2: E-ELAN and previous work on maximal layer efficiency

[Wang et al., 2022]

YOLOv7 architecture

Model Scaling for Concatenation based Models.

Model scaling is performed to generate models that meet the needs of different application requirements.

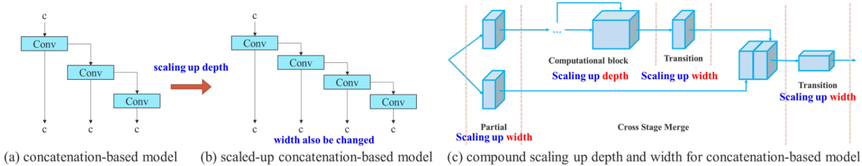


Figure 3: Model scaling for concatenation-based models.

[Wang et al., 2022]

Planned Re-parameterized Convolution

Re-parameterization is a technique used after training to improve the model. It increases the training time but improves the inference results.
There are two types of re-parametrization used to finalize models Model level and Module level ensemble [Kukil and Rath, 2022].
Model level re-parametrization can be done in the following two ways.

- Different training data to train multiple model with the same parameters.
- The average of the weights of models at different epochs.

In module level re-parametrization, the model training process is split into multiple modules.

Planned Re-parameterized Convolution

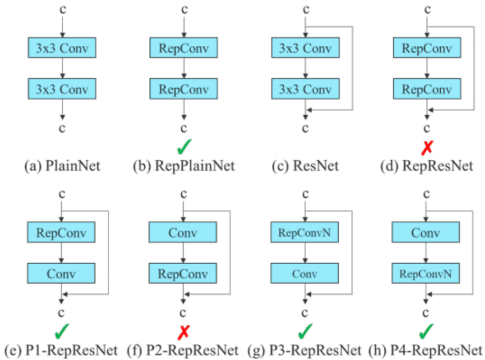
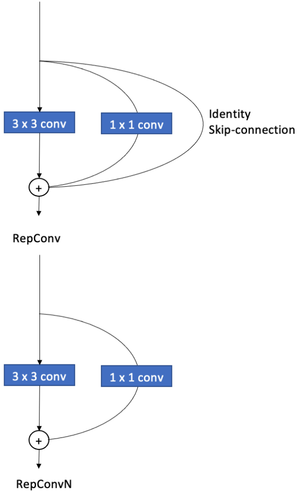


Figure 4: Planned re-parameterized model.

[Wang et al., 2022]



Coarse for auxiliary and fine for lead loss

.

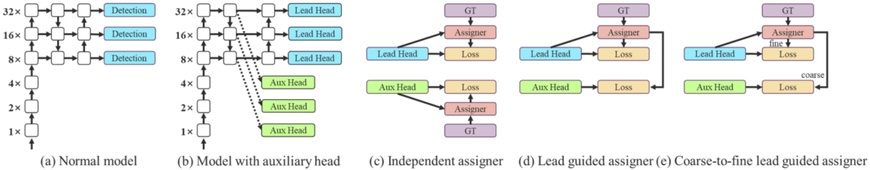


Figure 5: Coarse for auxiliary and fine for lead head label assigner.

[Wang et al., 2022]

Application to trash detection

We performed a comparison of the two models on the 10 best represented classes in our dataset.

Yolov5

	precision	recall	f1-score	support
Other plastic	0.70	0.73	0.72	201
Plastic bag & wrapper	0.74	0.86	0.80	176
Plastic bottle	0.82	0.90	0.86	115
Unlabeled litter	0.67	0.66	0.66	112
Other paper	0.70	0.82	0.75	77
Drink can	0.79	0.91	0.84	53
Other metal	0.66	0.70	0.68	44
Glass bottle	0.87	0.97	0.92	34
Food can / tupperware	0.73	0.86	0.79	28
Plastic cup	0.88	0.85	0.86	26
micro avg	0.74	0.81	0.77	866
macro avg	0.75	0.83	0.79	866
weighted avg	0.74	0.81	0.77	866

0.5345152455486654

Yolov7

	precision	recall	f1-score	support
Other plastic	0.86	0.83	0.85	203
Plastic bag & wrapper	0.80	0.92	0.86	174
Plastic bottle	0.91	0.98	0.95	115
Unlabeled litter	0.82	0.79	0.80	112
Other paper	0.90	0.92	0.91	78
Drink can	0.81	0.89	0.85	54
Other metal	0.79	0.82	0.80	45
Glass bottle	0.89	0.97	0.93	34
Food can / tupperware	0.86	0.89	0.88	28
Plastic cup	1.00	0.85	0.92	26
micro avg	0.85	0.88	0.87	869
macro avg	0.87	0.89	0.87	869
weighted avg	0.85	0.88	0.87	869

0.7488119524735997

mAP YOLOv5 : 53%, mAP YOLOv7 : 74%

Paper

A systematic study of the class imbalance problem in convolutional neural networks
[Mateusz Buda, Atsuto Maki and Maciej A. Mazurowski]

In this study, the authors investigate the impact of class imbalance on classification performance of convolutional neural networks (CNNs) and compare frequently used methods to address the issue. Class imbalance is a common problem that has been comprehensively studied in classical machine learning, yet very limited systematic research is available in the context of deep learning.

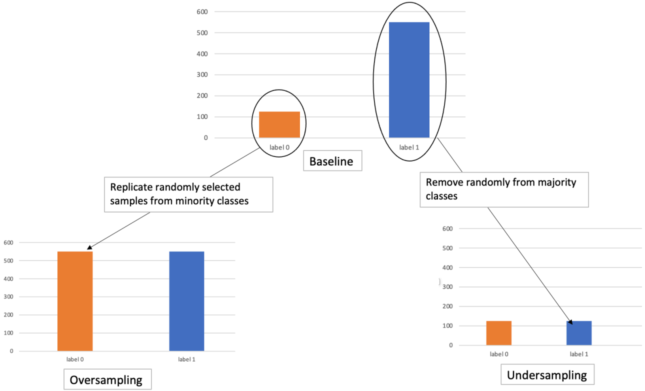
Paper

The authors mention two categories of methods for addressing the class imbalance problem.

- Data-level methods that modify the distribution of classes in the training set.
- Classifier-level methods that keep the training set unchanged and adjust the learning and inference algorithms.

Data Level Méthods

The best and simplest method is to add new images for minority classes.
However, there are other methods like oversampling and undersampling



Classifier Level Méthods

- Thresholding
  Adjustment of the decision threshold of a classifier.

Classifier Level Méthods

- Thresholding
  Adjustment of the decision threshold of a classifier.
- Cost sensitive learning
  This method assigns a different cost to the misclassification of examples from different classes and can be implemented in different ways.

Classifier Level Méthods

- Thresholding
  Adjustment of the decision threshold of a classifier.
- Cost sensitive learning
  This method assigns a different cost to the misclassification of examples from different classes and can be implemented in different ways.
- Hybrid of methods

Methods of addressing imbalance compared in this paper

- Random minority oversampling
- Random majority undersampling
- Thresholding with prior class probabilities
- Oversampling with thresholding
- Undersampling with thresholding

Methods of addressing imbalance compared on MNIST and CIFAR-10

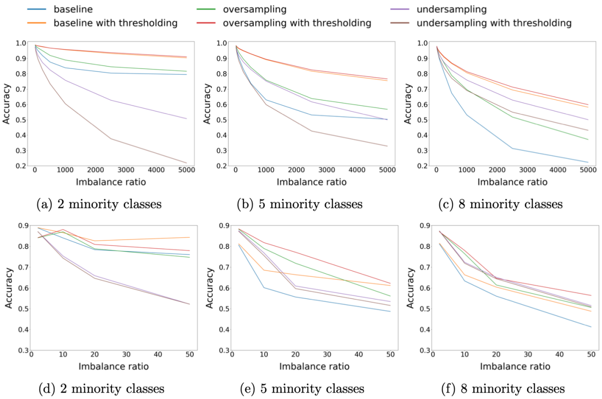


Figure 6: Comparison of methods with respect to accuracy on MNIST (a - c) and CIFAR-10 (d - f).

[Buda et al., 2017]

Methods of addressing imbalance compared on ImageNet

Method	$\mu = 0.1, \rho = 10$	$\mu = 0.8, \rho = 50$	$\mu = 0.9, \rho = 100$		
Baseline	99.41	96.31	90.74	90.46	90.05
Oversampling	99.35	95.06	88.38	88.39	88.17
Undersampling	96.85	94.98	88.35	84.08	83.74

Figure 7: Comparison of results on ImageNet with respect to multi-class ROC AUC.

[Buda et al., 2017]

$\mu = 0.1,\ \rho = 10, \mu = 0.8,\ \rho = 50\ and\ \mu = 0.9, \rho = 100$ correspond to 100 minority classes with imbalance ratio of 10, 800 minority classes with imbalance of 50, and 900 minority classes with imbalance ratio of 100, respectively.

Next steps

- Adding new images to the dataset for the minority classes

Next steps

- Adding new images to the dataset for the minority classes
- Applying the class imbalance management methods

Next steps

- Adding new images to the dataset for the minority classes
- Applying the class imbalance management methods
- Implementation of an embedded prototype

Thanks!

References I

[Buda et al., 2017] Buda, M., Maki, A., and Mazurowski, M. A. (2017).
A systematic study of the class imbalance problem in convolutional neural networks.
CoRR, abs/1710.05381.

[Kukil and Rath, 2022] Kukil and Rath, S. (2022).
Modélisation et évaluation du poids carbone de produits de consommation et biens déquipements.
Source here.

[Wang et al., 2022] Wang, C.-Y., Bochkovskiy, A., and Liao, H.-Y. M. (2022).
Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors.