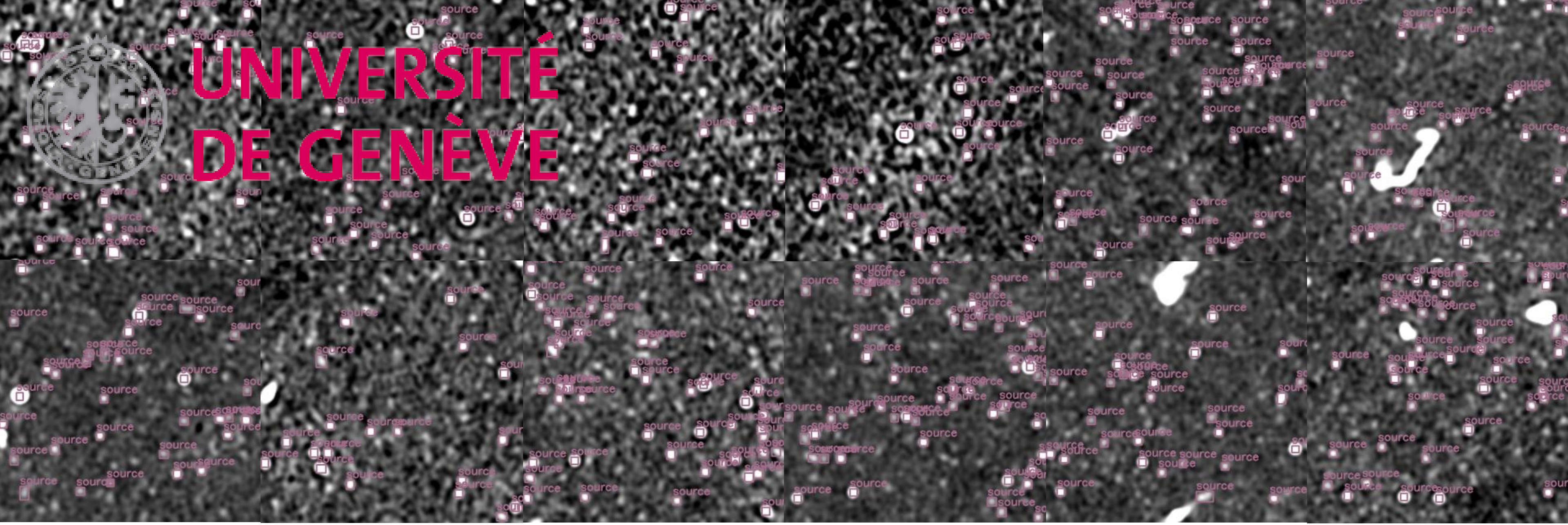


USE OF FOUNDATION MODELS FOR SOURCE FINDING IN RADIO IMAGES

E. Lastufka

M. Audard, O. Bait, M. Dessauges-Zavadsky, M. Drozdova, T. Holotyak, V. Kinakh, D. Piras, O. Taran, D. Schaerer, S. Voloshynovskiy

SKACH

How to find sources in images?

Analytic methods

Needs a lot of time

ML-based methods

Needs a lot of data

How to find sources in images?

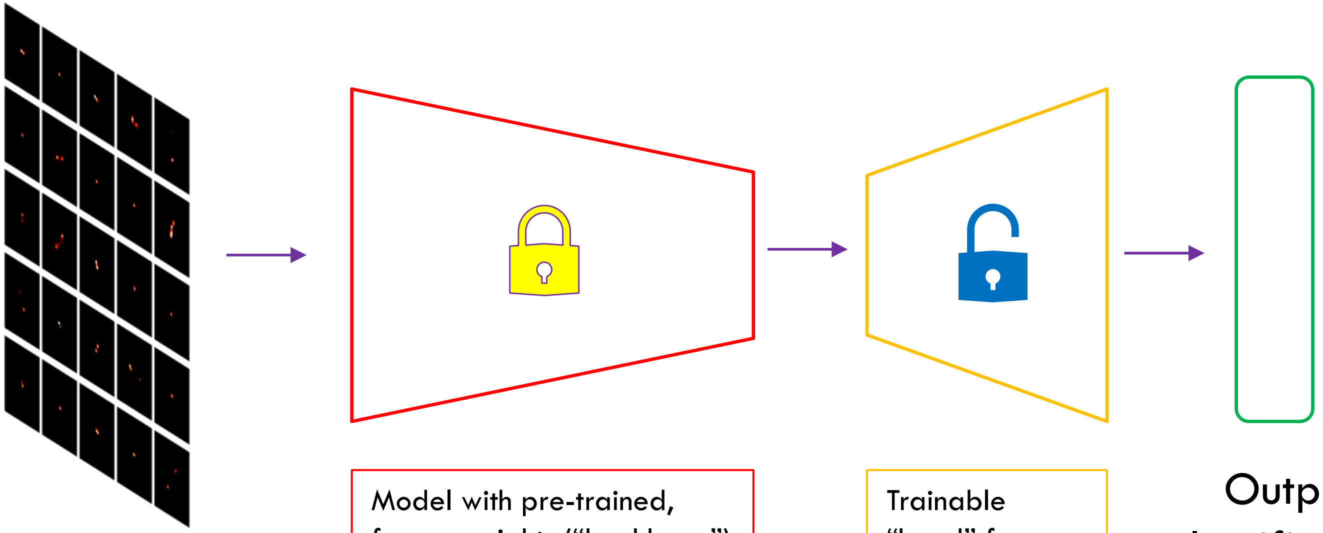
Analytic methods

Needs a lot of time

ML-based methods

Foundation models

FOUNDATION MODELS & TRANSFER LEARNING



Input data

Model with pre-trained, frozen weights ("backbone")

Trainable "head" for downstream task

Output: classification, detection, etc.

Using foundation models, can we bypass the need for large training datasets?

SOURCE FINDING: ANALYTIC METHODS

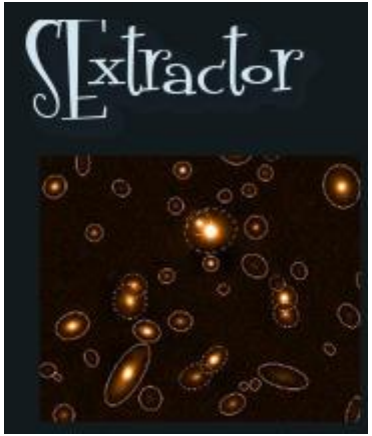
Background estimation and subtraction

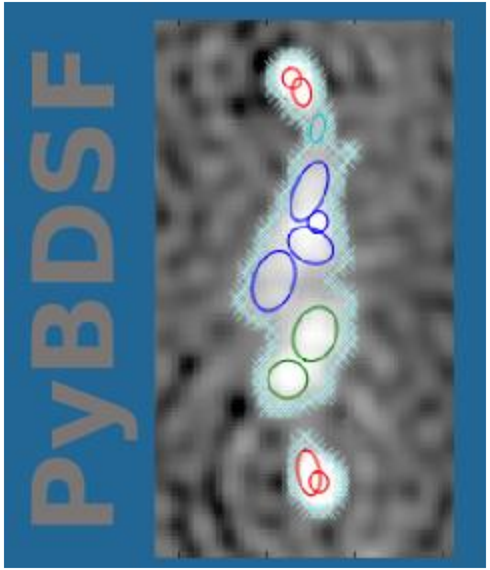
Source identification

Source characterization

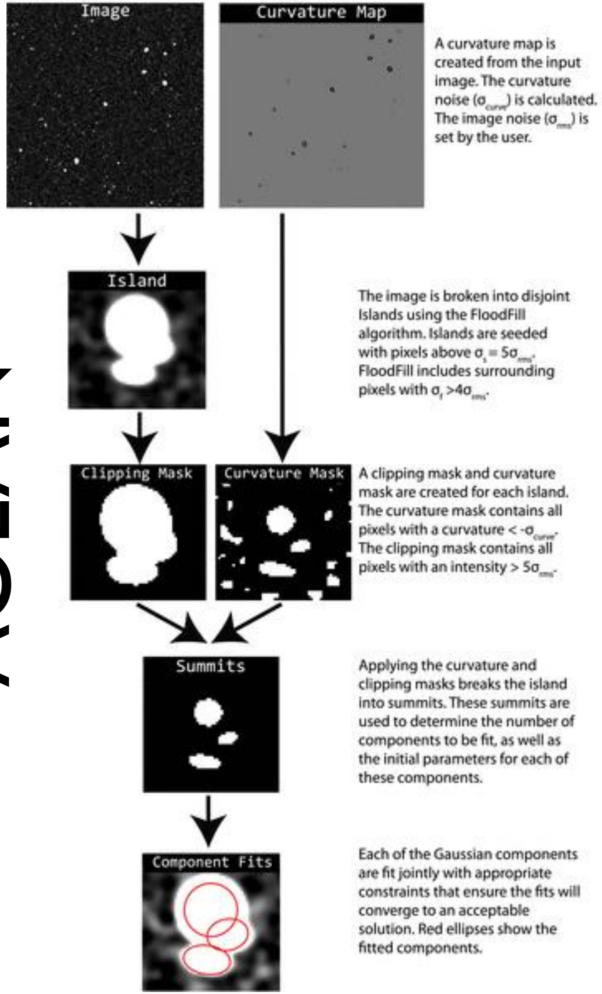
Cataloging

SOURCE FINDING: ANALYTIC METHODS





AGEAN

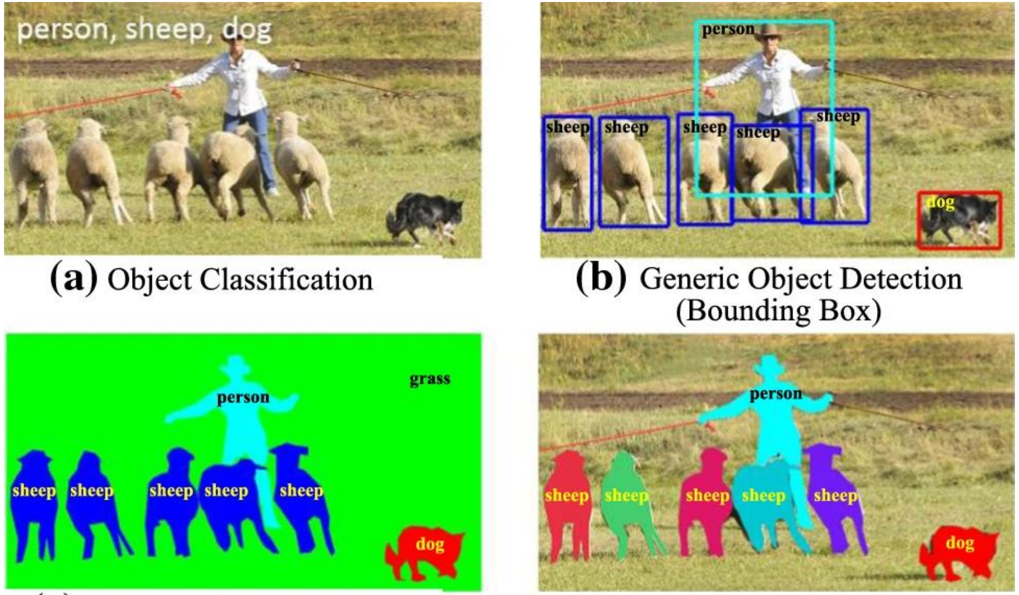


SOURCE FINDING: ML-BASED METHODS

Source identification

Source characterization

SOURCE FINDING: ML-BASED METHODS



(a) Object Classification

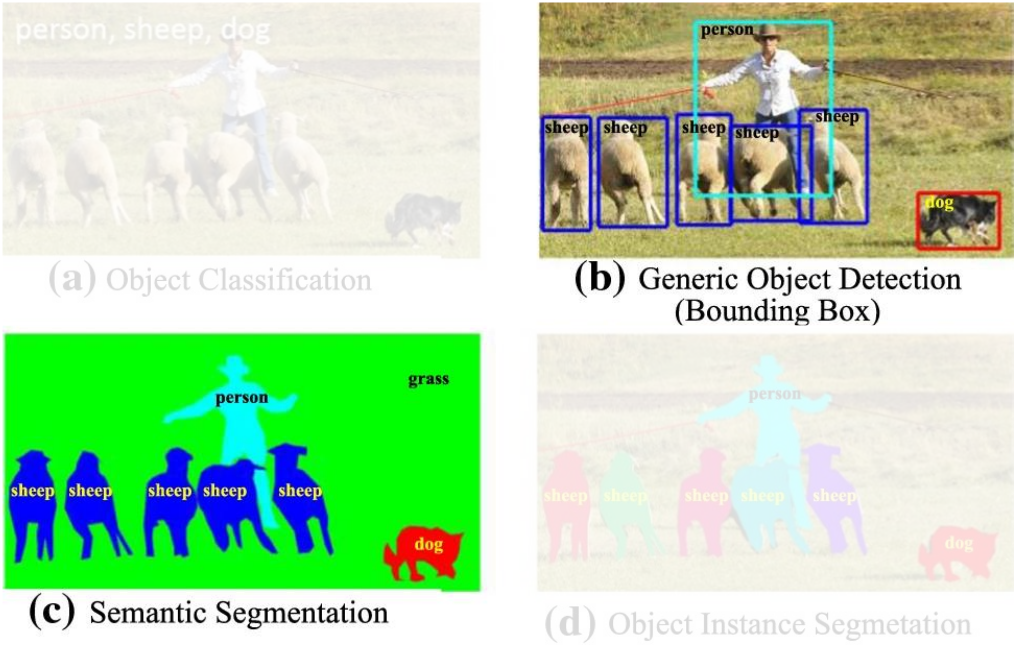
(b) Generic Object Detection (Bounding Box)

(c) Semantic Segmentation

(d) Object Instance Segmetation

Liu et al 2019

SOURCE FINDING: ML-BASED METHODS



Liu et al 2019

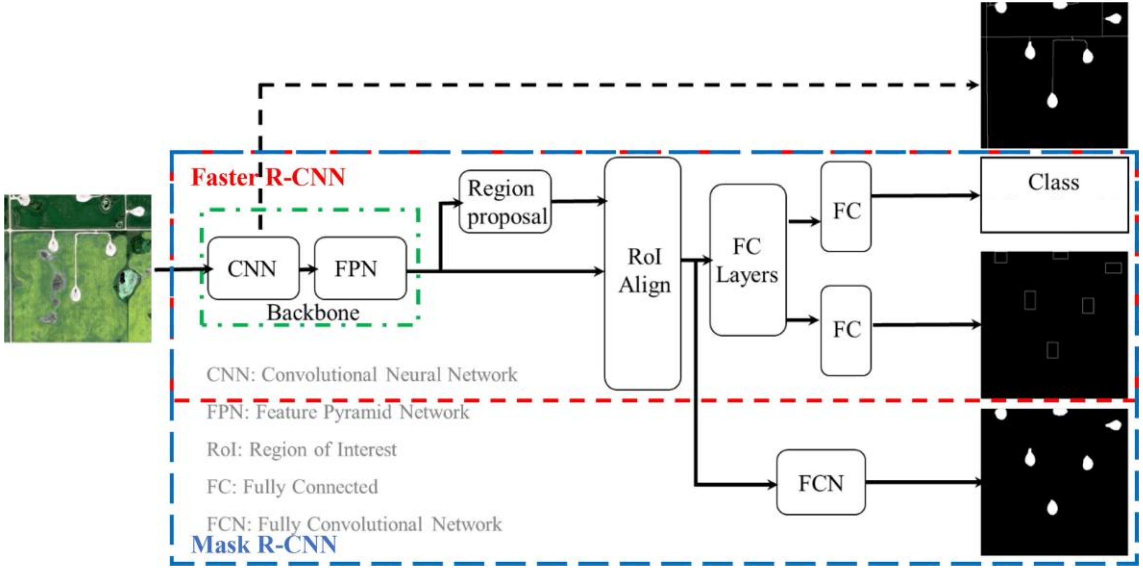
SOURCE FINDING: ML-BASED METHODS

Faster/Mask-RCNN

ViT-Det

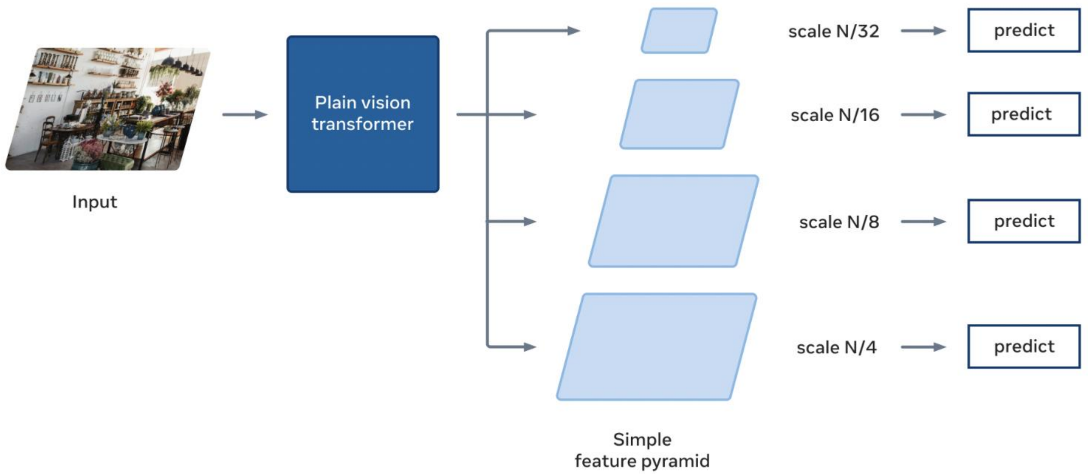
YOLO

FASTER/MASK-RCNN



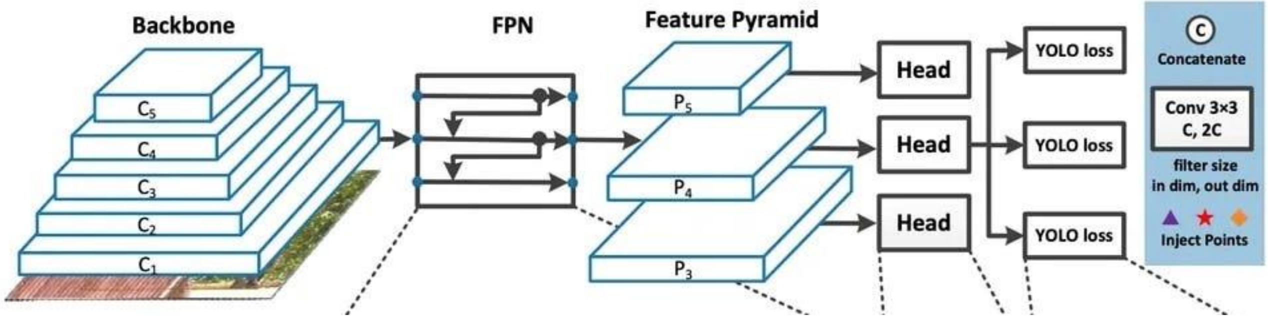
He et al 2022

VIT-DET



Input
Plain vision transformer
scale N/32
scale N/16
scale N/8
scale N/4
predict
predict
predict
predict
Simple
feature pyramid

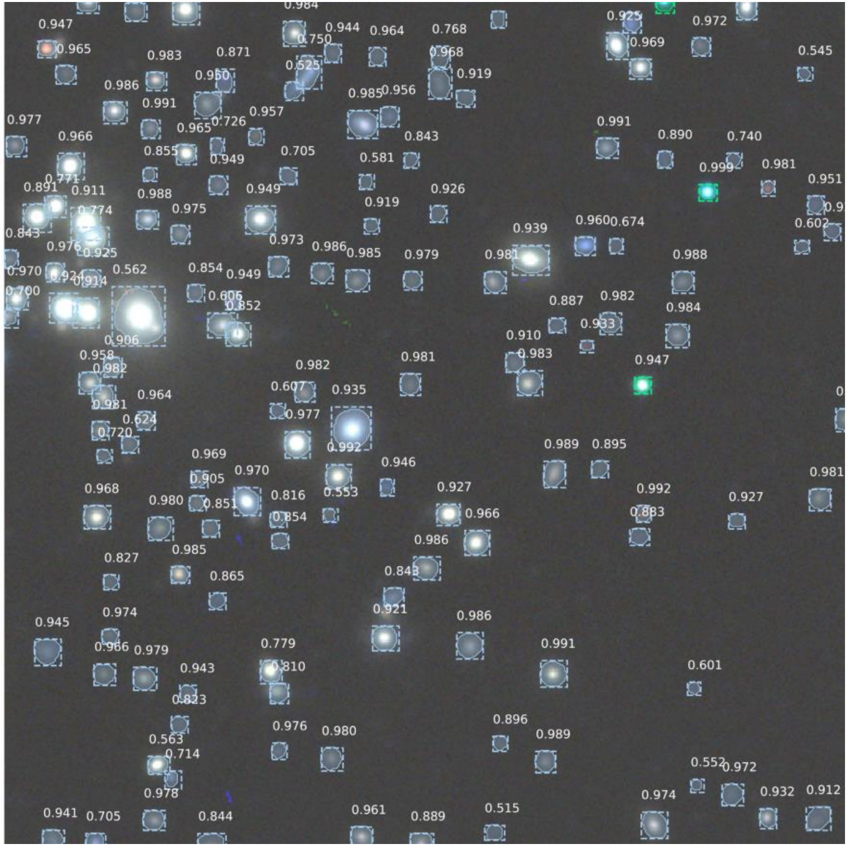
YOLO



Backbone
C_5
C_4
C_3
C_2
C_1
FPN
Feature Pyramid
P_5
P_4
P_3
Head
Head
Head
YOLO loss
YOLO loss
YOLO loss
Concatenate
Conv 3×3
C, 2C
filter size
in dim, out dim
Inject Points

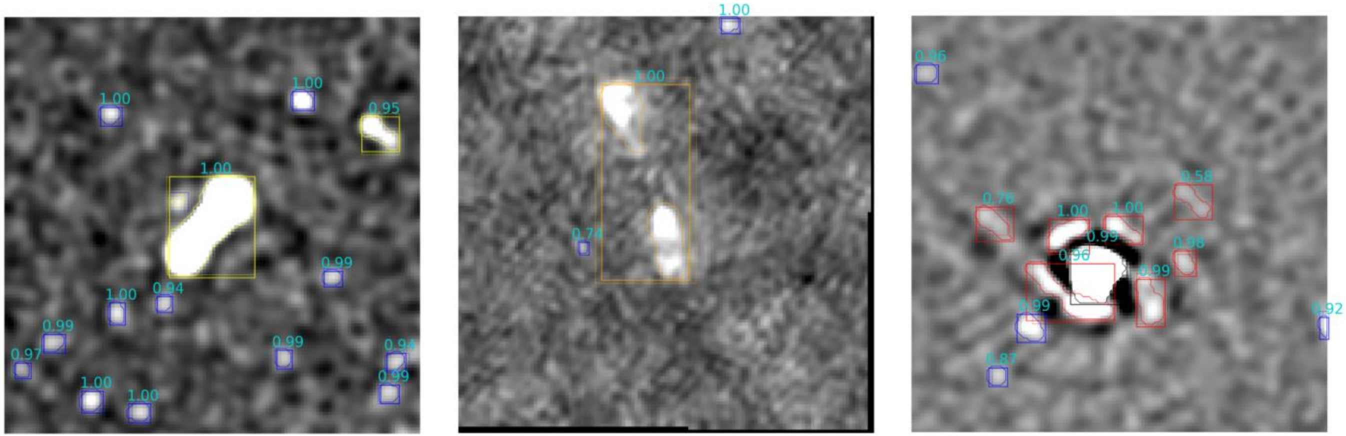
BURKE ET AL 2019

- Optical data (simulated DECam images)
- Mask-RCNN for detection and segmentation
- Classification into stars and galaxies



RIGGI ET AL 2022

- Small, mixed dataset of radio continuum images (ASKAP, VLA, ATCA)
- Mask-RCNN for detection and segmentation



UNIVERSITÉ
DE GENÈVE

SORTINO ET AL 2023

Comparison of different object detection and segmentation methods from computer vision

Table 2: Detection metrics. YOLOv4 shows the best reliability, but this high value is given by a high IoU threshold. BS stands for batch size

Model	BS	Reliability			Completeness			F1-Score			mAP
		Compact	Extended	Total	Compact	Extended	Total	Compact	Extended	Total	Total
Mask R-CNN	32	48.7%	88.8%	52.0%	82.3%	77.0%	79.5%	61.2%	82.5%	62.9%	70.2%
Detectron2	64	59.8%	62.9%	59.1%	**83.7%**	**90.9%**	**83.9%**	69.7%	74.3%	69.4%	**83.9%**
DETR	2	75.0%	84.6%	76.4%	76.6%	84.9%	76.8%	**75.8%**	84.8%	**76.6%**	79.0%
Yolo v4	64	**97.4%**	**95.9%**	**97.2%**	48.3%	85.5%	50.2%	64.5%	**90.4%**	66.2%	53.8%
Yolo v7	32	87.5%	87.7%	87.4%	60.0%	86.6%	61.0%	69.1%	87.2%	71.7%	61.6%
YOLOS	2	55.9%	78.1%	58.0%	75.0%	84.8%	75.5%	64.1%	81.3%	65.6%	76.3%
EffDet-D1	64	96.1%	0.0%	64.9%	42.2%	0.0%	33.7%	58.6%	0.0%	44.4%	53.5%
EffDet-D2	32	96.7%	0.0%	69.8%	48.5%	0.0%	39.1%	64.6%	0.0%	50.1%	53.8%

SORTINO ET AL 2023

Transformer-based methods benefit from pre-training

Table 2: Detection metrics. YOLOv4 shows the best reliability, but this high value is given by a high IoU threshold. BS stands for batch size

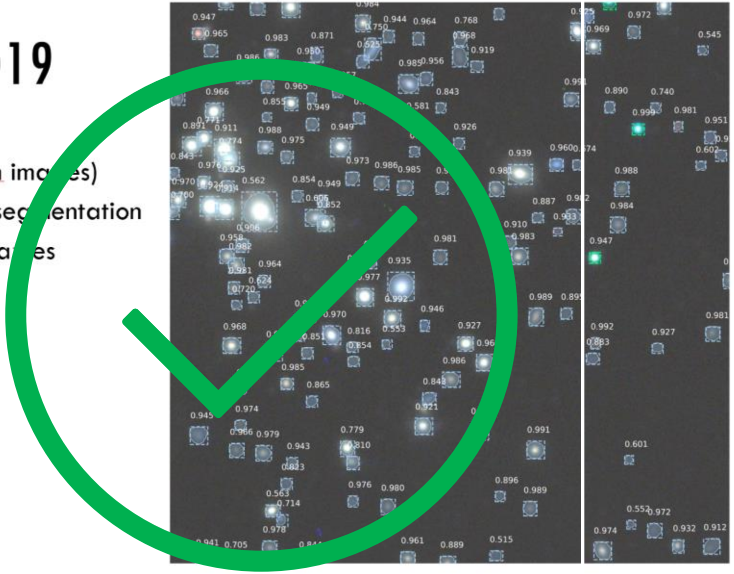
Model	BS	Reliability			Completeness			F1-Score			mAP
		Compact	Extended	Total	Compact	Extended	Total	Compact	Extended	Total	Total
Mask R-CNN	32	48.7%	88.8%	52.0%	82.3%	77.0%	79.5%	61.2%	82.5%	62.9%	70.2%
Detectron2	64	59.8%	62.9%	59.1%	**83.7%**	**90.9%**	**83.9%**	69.7%	74.3%	69.4%	**83.9%**
DETR	2	75.0%	84.6%	76.4%	76.6%	84.9%	76.8%	**75.8%**	84.8%	**76.6%**	79.0%
Yolo v4	64	**97.4%**	**95.9%**	**97.2%**	48.3%	85.5%	50.2%	64.5%	**90.4%**	66.2%	53.8%
Yolo v7	32	87.5%	87.7%	87.4%	60.0%	86.6%	61.0%	69.1%	87.2%	71.7%	61.6%
YOLOS	2	55.9%	78.1%	58.0%	75.0%	84.8%	75.5%	64.1%	81.3%	65.6%	76.3%
EffDet-D1	64	96.1%	0.0%	64.9%	42.2%	0.0%	33.7%	58.6%	0.0%	44.4%	53.5%
EffDet-D2	32	96.7%	0.0%	69.8%	48.5%	0.0%	39.1%	64.6%	0.0%	50.1%	53.8%

UNIVERSITÉ DE GENÈVE

TRANSFER LEARNING: DOES IT WORK?

BURKE ET AL 2019

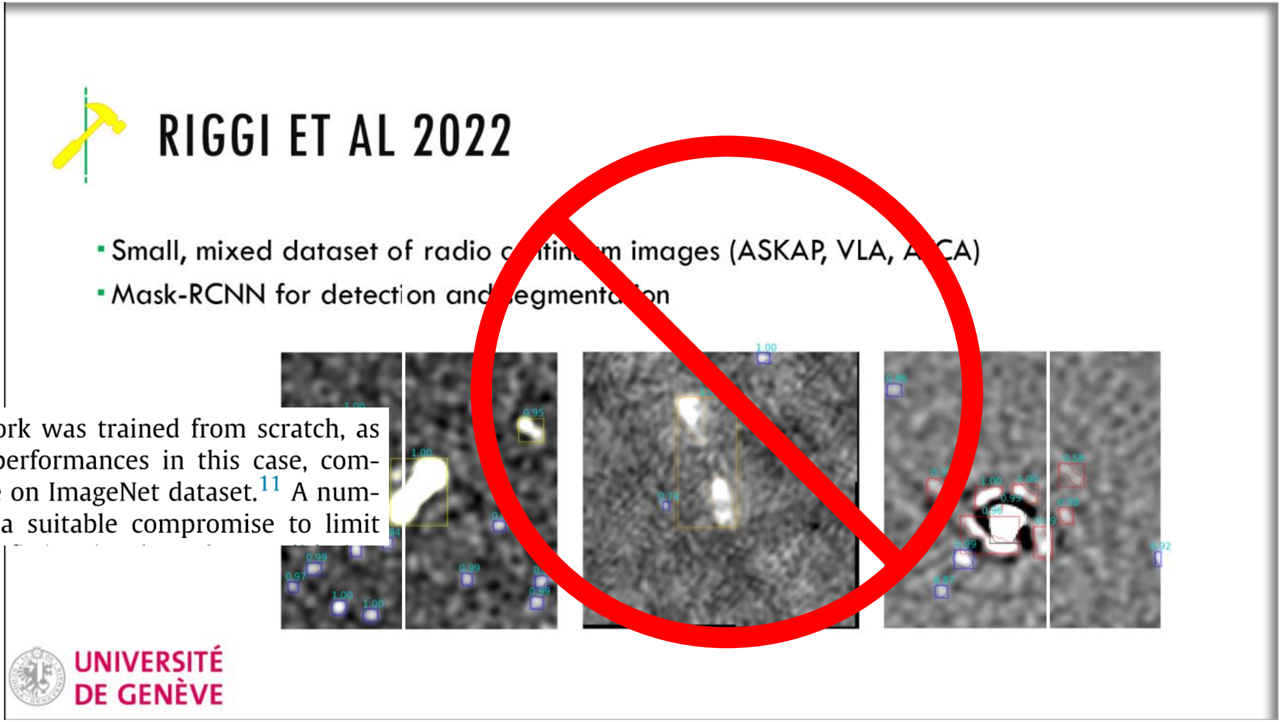
- Optical data (simulated DECam images)
- Mask-RCNN for detection and segmentation
- Classification into stars and galaxies



overfitting of the network. We use Mask R-CNN weights provided by Abdulla (2017) trained on the Microsoft Common Objects in Context (MS COCO) data set (Lin et al. 2014) as the starting point for our training procedure. MS COCO is a data set of ~328 000

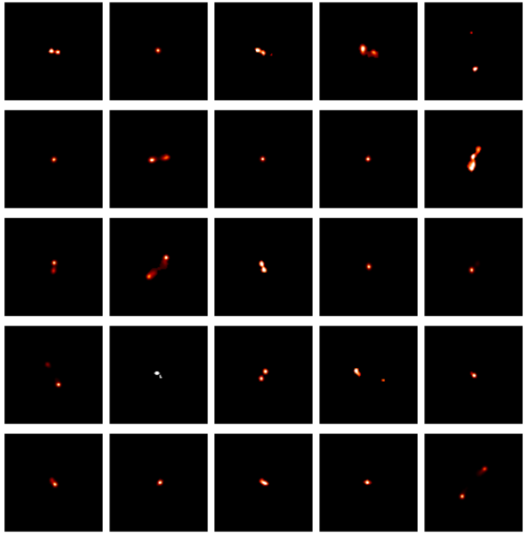
UNIVERSITÉ DE GENÈVE

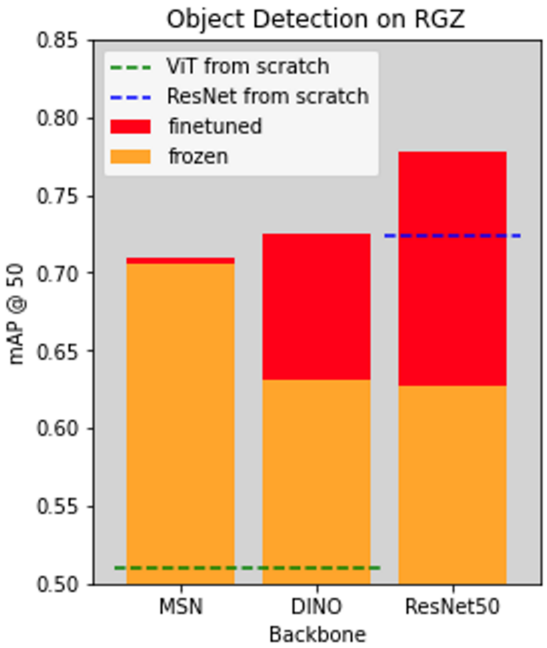
TRANSFER LEARNING: DOES IT WORK?



The *ResNet-101* backbone network was trained from scratch, as we obtained slightly superior performances in this case, compared to a pre-trained backbone on ImageNet dataset.[11] A number of 150 epochs was found a suitable compromise to limit

TRANSFER LEARNING: RADIO GALAXY ZOO

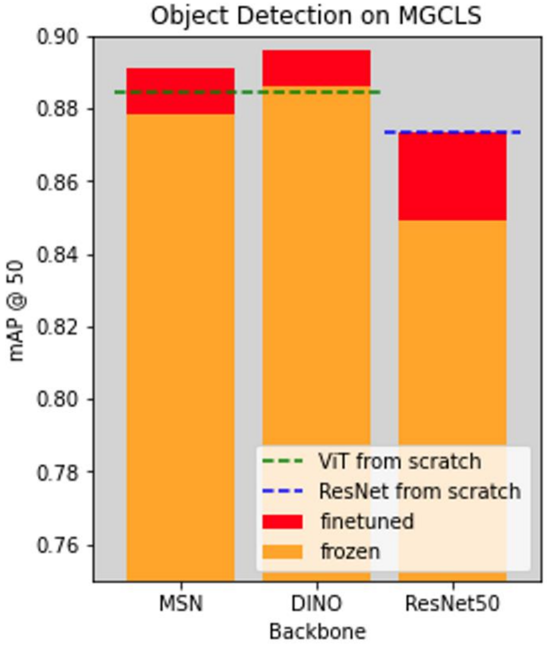




Vision transformers

- Transfer learning with frozen Vision Transformer backbones gives better performance than trained-from-scratch Vision Transformers!
- Transfer learning less successful with ResNet50 backbone
- Fine-tuning improves performance

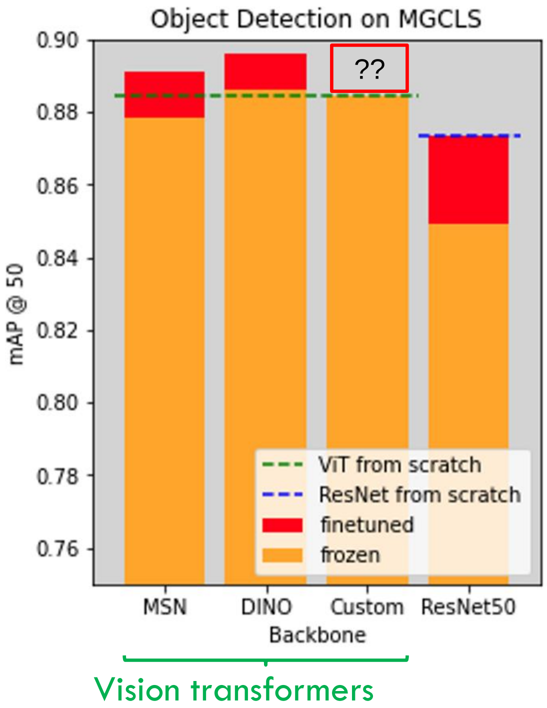
TRANSFER LEARNING: MEERKAT SURVEYS



Vision transformers

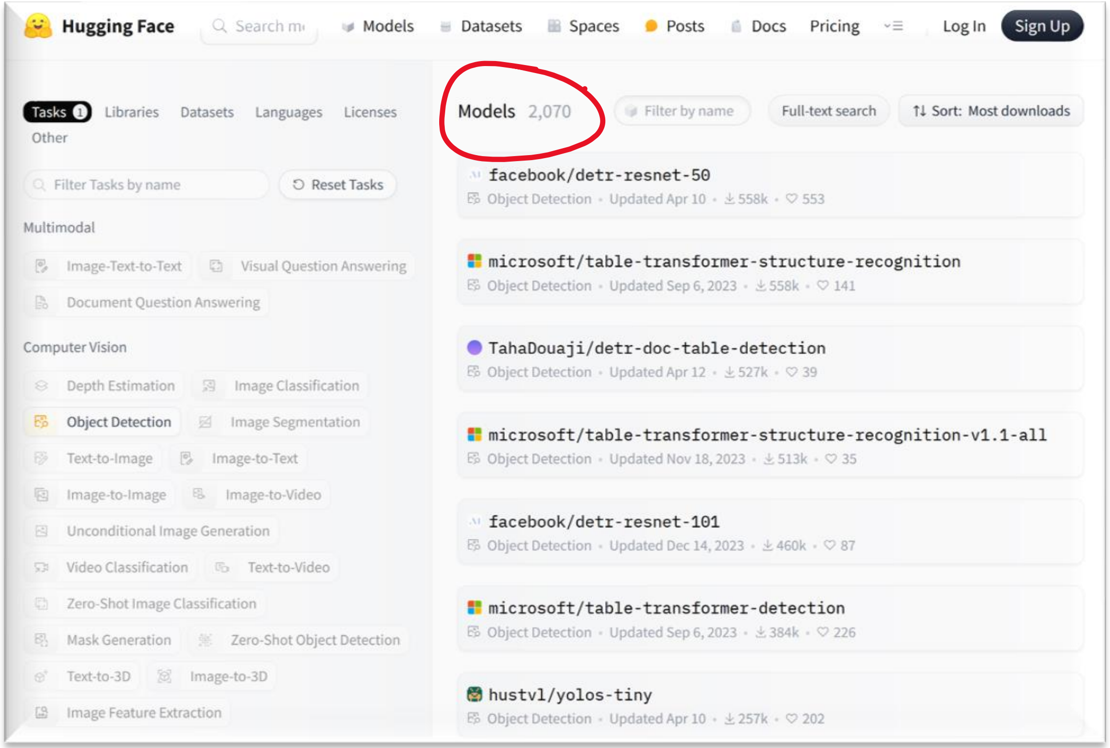
- Fine-tuning required to match or surpass trained-from-scratch performance

TRANSFER LEARNING: MEERKAT SURVEYS



- Will using a foundation model pre-trained with radio astronomy images bring significant performance benefits?

TRANSFER LEARNING: HOW TO?



TRANSFER LEARNING: HOW TO?

Data Characteristics

Model architecture

Model training method

Data Amount

Model training dataset

Downstream task

TRANSFER LEARNING: A PRACTICAL GUIDE FOR ASTROPHYSICS DATA