IAP colloquium: Debating the potential of machine learning in astronomical surveys - 18/10/2021

Machine learning-infused cluster cosmology

Stéphane Ilić
PSL / LERMA / APC (France)

In collaboration with
S. Mei (APC, France),
L. Daudet, A. Chatelain (LightOn), F. Krzakala (ENS/EPFL)

LERMA l'Observatoire de Paris | PSL

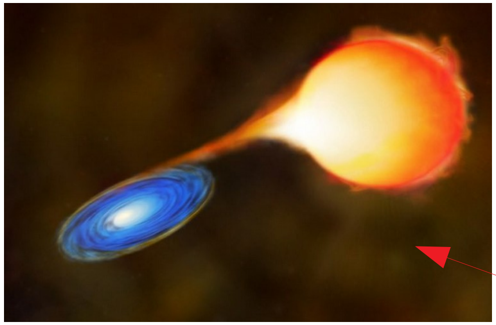
Constraining our cosmological model

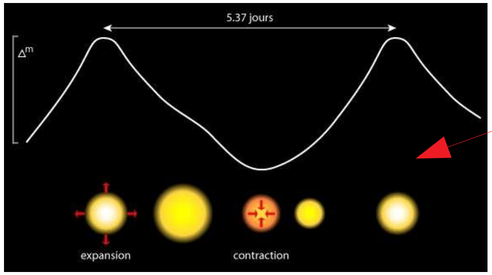
Probing two different “sectors”:

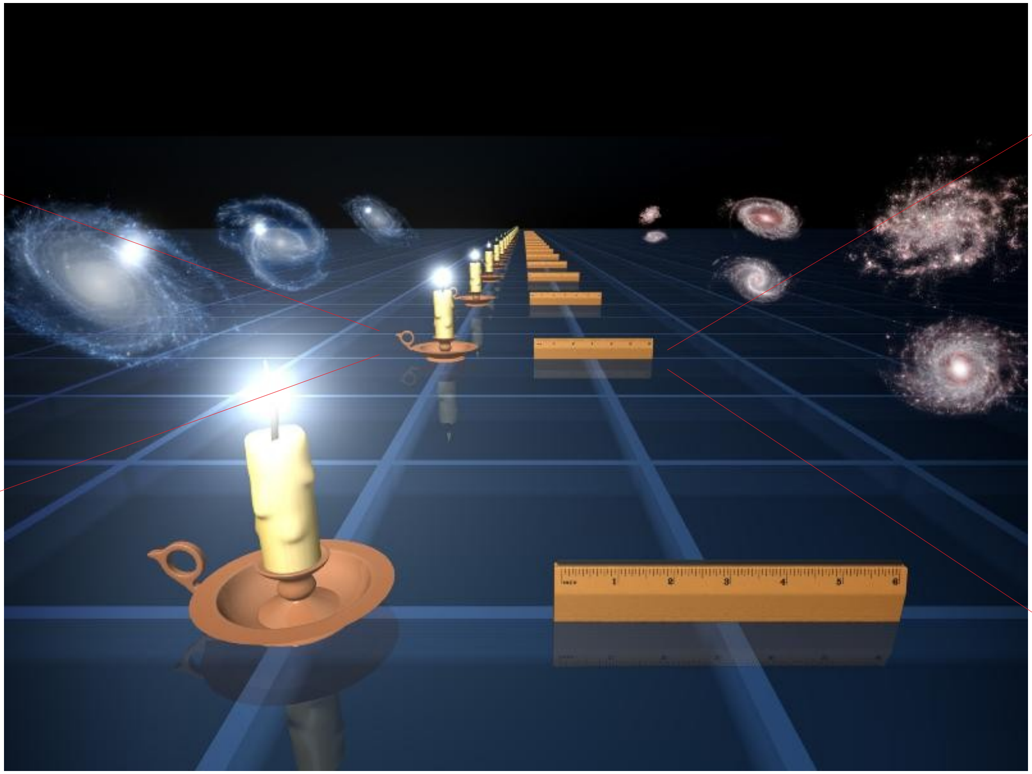
Constraining our cosmological model

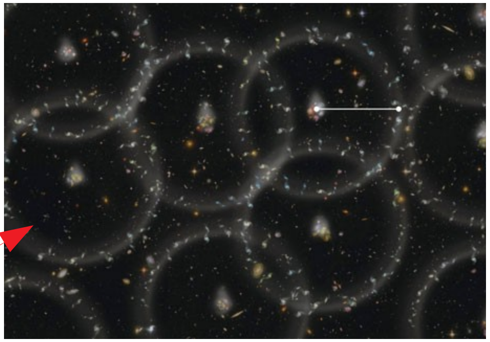
Probing two different “sectors”:

- Background evolution: all standard rulers/candles





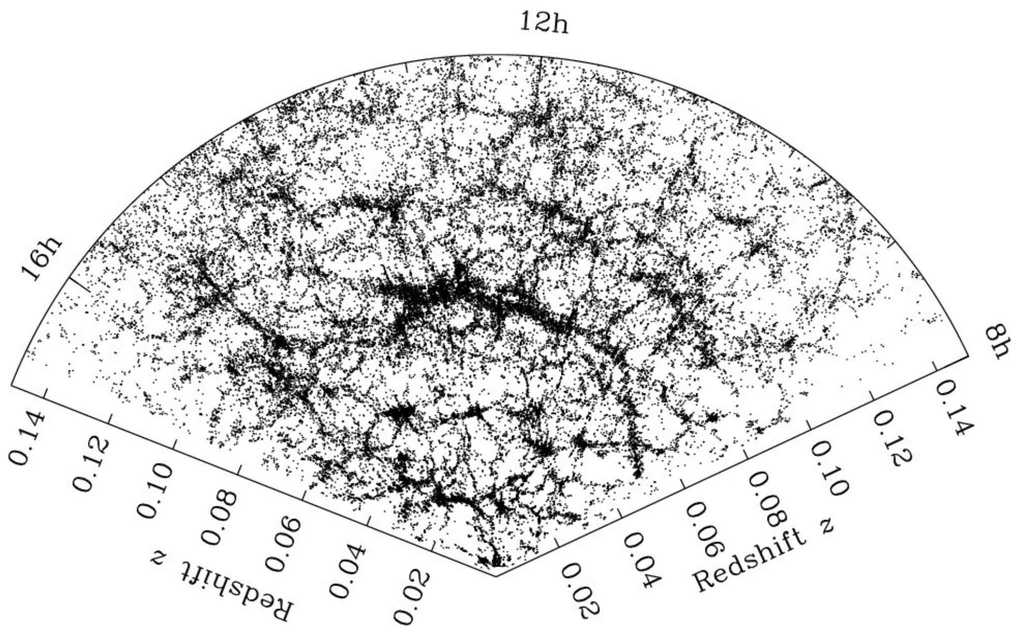




Constraining our cosmological model

Probing two different "sectors":

- Background evolution: all standard rulers/candles
- Perturbations: probes of structure growth



Large-scale structures

- Statistical properties depend on cosmology
- Main observable: 2pt-correlation function $\xi(r)$, $P(k,z)$, $C_{ell}(z_1,z_2)$
- What about 1pt-CF ?

→ Hard to predict because galaxies = non-linear

Clusters as cosmological probes

Clusters of galaxies:

- Largest structures in the Universe → closer to <u>linear regime</u>
- Exponentially <u>sensitive to growth rate</u> of structure → great probes of DE
- Affected by volume effects → <u>sensitivity to background</u>

Use as cosmological probe:

- Main principle : compare predicted and observed N(M,z)
- (Fairly) robust framework for predicting abundances ↔ “mass function” (Press & Schechter 1974 and “successors”)

Clusters as cosmological probes

Clusters of galaxies:

- Largest structures in the Universe → closer to linear regime
- Exponentially sensitive to growth rate of structure → great probes of DE
- Affected by volume effects → sensitivity to background

Use as cosmological probe:

- Main principle : compare predicted and observed N(M,z)
- (Fairly) robust framework for predicting abundances ↔ “mass function” (Press & Schechter 1974 and “successors”)

Challenges:

Clusters as cosmological probes

Clusters of galaxies:

- Largest structures in the Universe → closer to linear regime
- Exponentially sensitive to growth rate of structure → great probes of DE
- Affected by volume effects → sensitivity to background

Use as cosmological probe:

- Main principle : compare predicted and observed N(M,z)
- (Fairly) robust framework for predicting abundances ↔ “mass function” (Press & Schechter 1974 and “successors”)

Challenges:

- Detecting/identifying clusters in data (what even *is* a cluster ??)

Clusters as cosmological probes

Clusters of gal

- Largest struct … ime
- Exponentially … eat probes of DE
- Affected by v

Use as cosmol

- Main principle …)
- (Fairly) robust … mass function"
  (Press & Sche

Challenges:

- Detecting/iden … uster ??)

Clusters as cosmological probes

Clusters of galaxies:

- Largest structures in the Universe → closer to linear regime
- Exponentially sensitive to growth rate of structure → great probes of DE
- Affected by volume effects → sensitivity to background

Use as cosmological probe:

- Main principle : compare predicted and observed N(M,z)
- (Fairly) robust framework for predicting abundances ↔ “mass function” (Press & Schechter 1974 and “successors”)

Challenges:

- Detecting/identifying clusters in data (what even *is* a cluster ??)
- **Total mass is not an observable** : proxies (temperature, richness, ...) & “scaling laws” required → no consensus on those laws

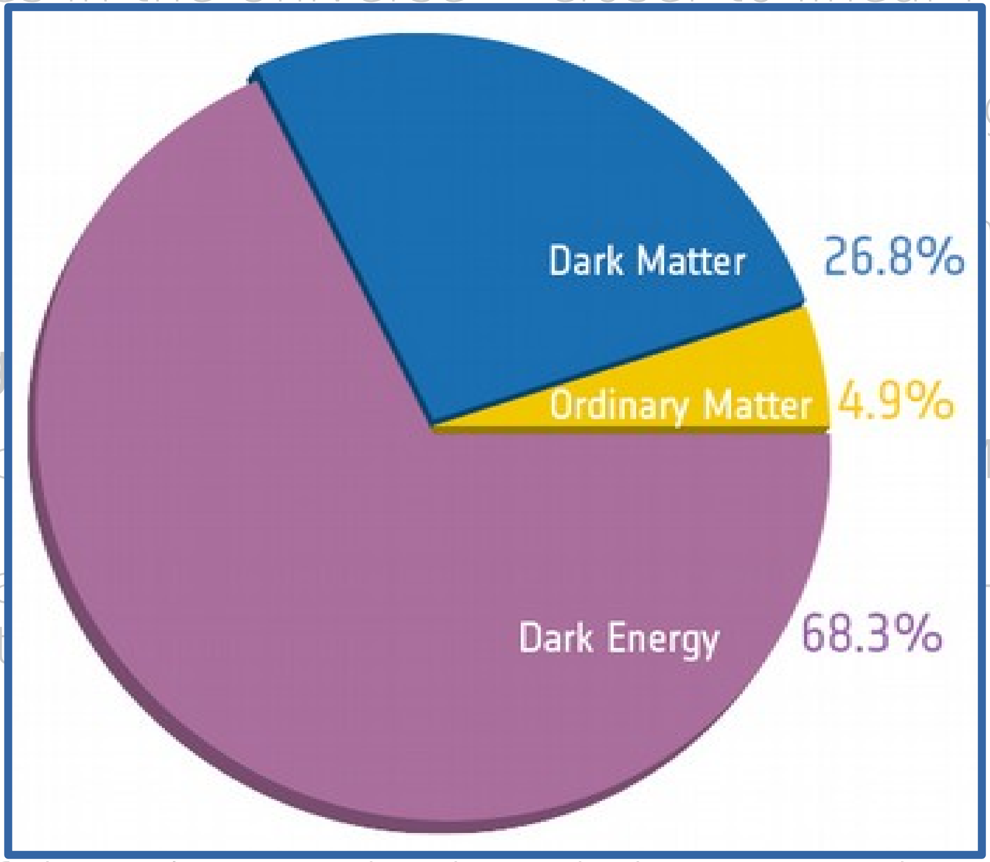
Clusters as cosmological probes

Clusters of galaxies:

- Largest structures in the Universe → closer to linear regime
- Exponentially se … great probes of DE
- Affected by volu … nd

Use as cosmolog

- Main principle : c … l,z)
- (Fairly) robust fra … → "mass function" (Press & Schecht



Challenges:

- Detecting/identifying clusters in data (what even *is* a cluster ??)
- **Total mass is not an observable** : proxies (temperature, richness, ...) & "scaling laws" required → no consensus on those laws

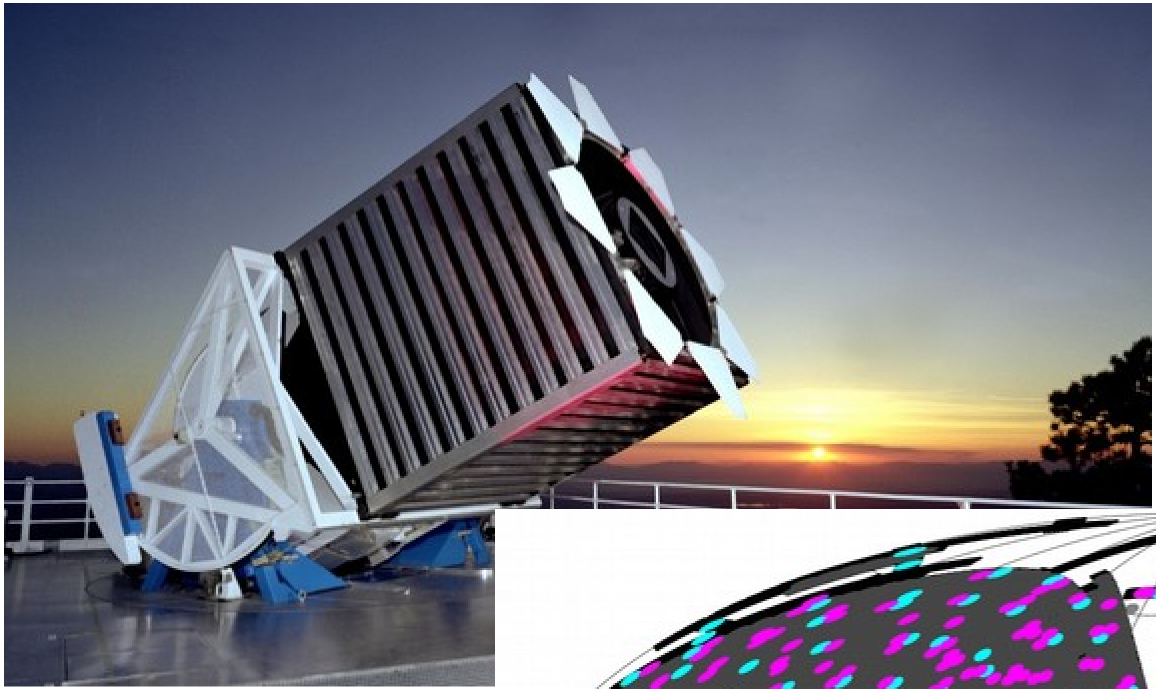
Objective

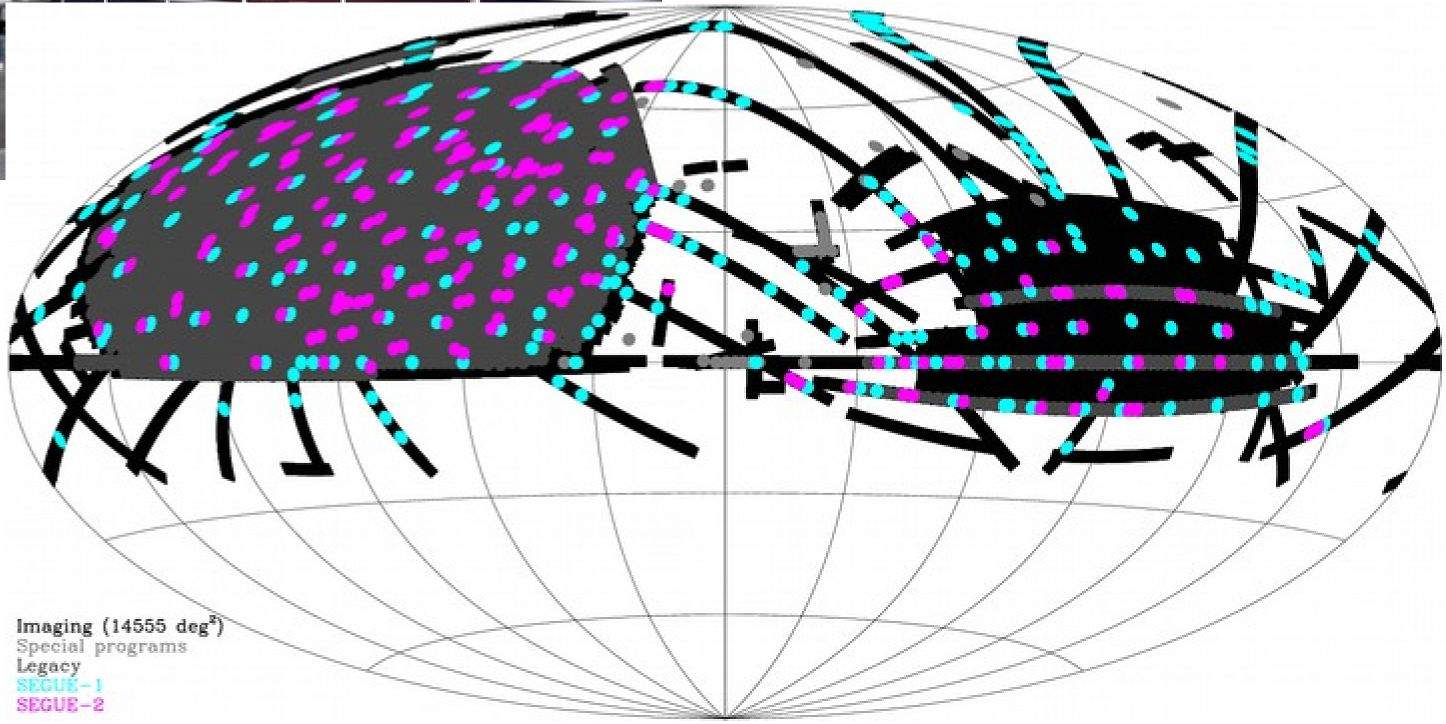
How can machine-learning algorithms
help us here ?

- Neural networks for **detecting clusters** in telescope images and/or galaxy catalogues

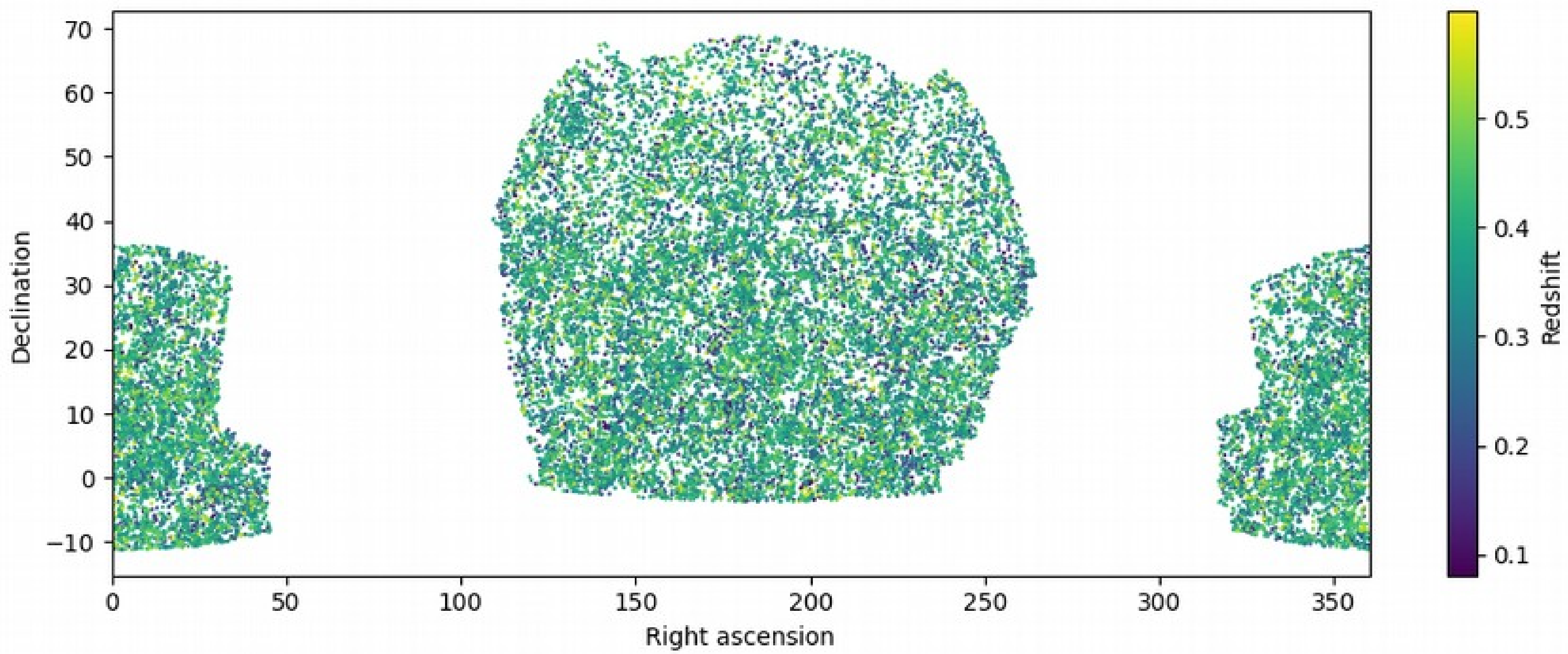
- Neural networks for **characterizing clusters** (mass, redshift, etc)

Sloan Digital Sky Survey (SDSS)





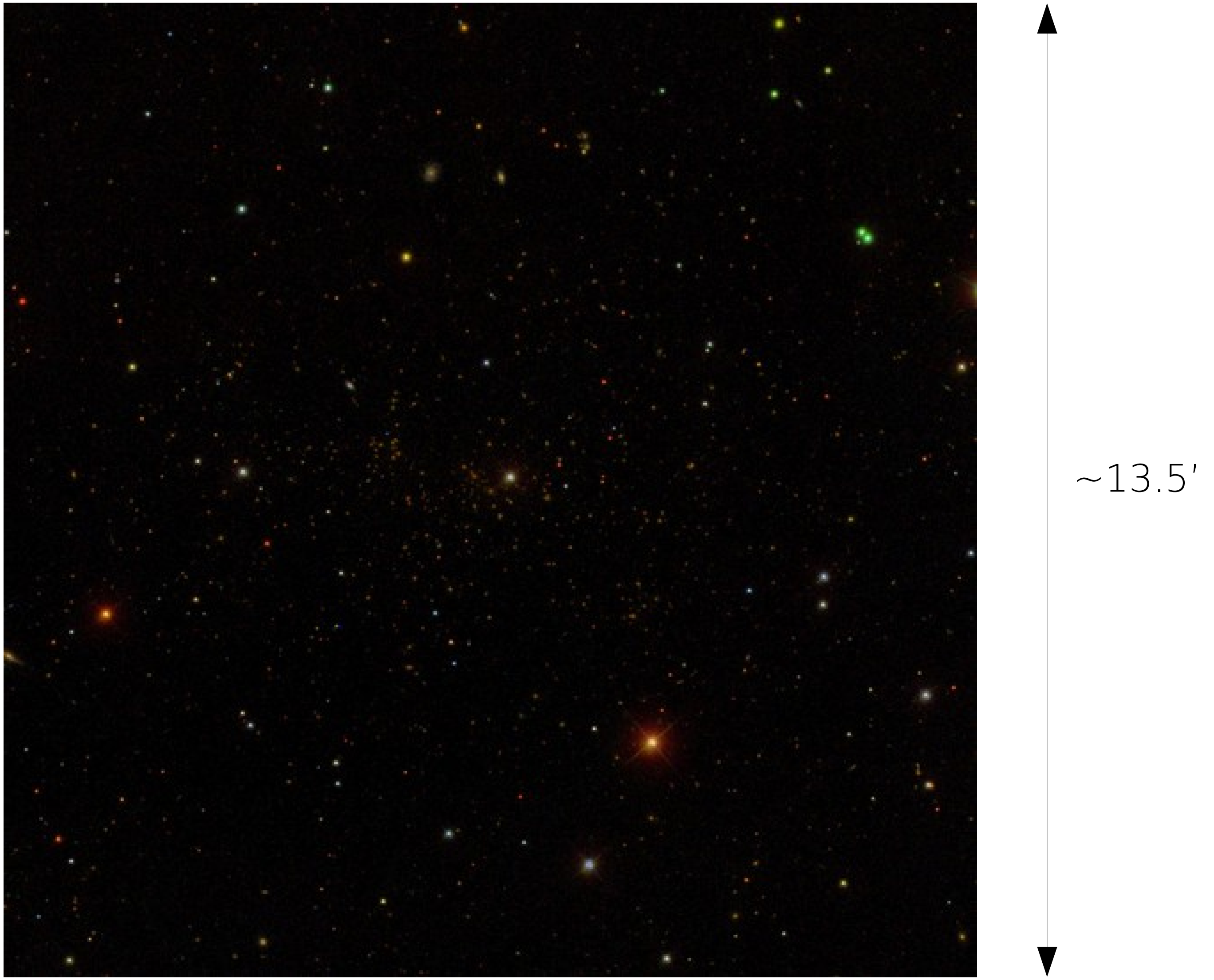
The redMaPPer cluster catalogue



- Catalogue of 26,111 clusters with characteristics (position, redshift & “richness” estimates)

- Associated catalogue of 1,703,685 “member galaxies” with characteristics (position & photometry)

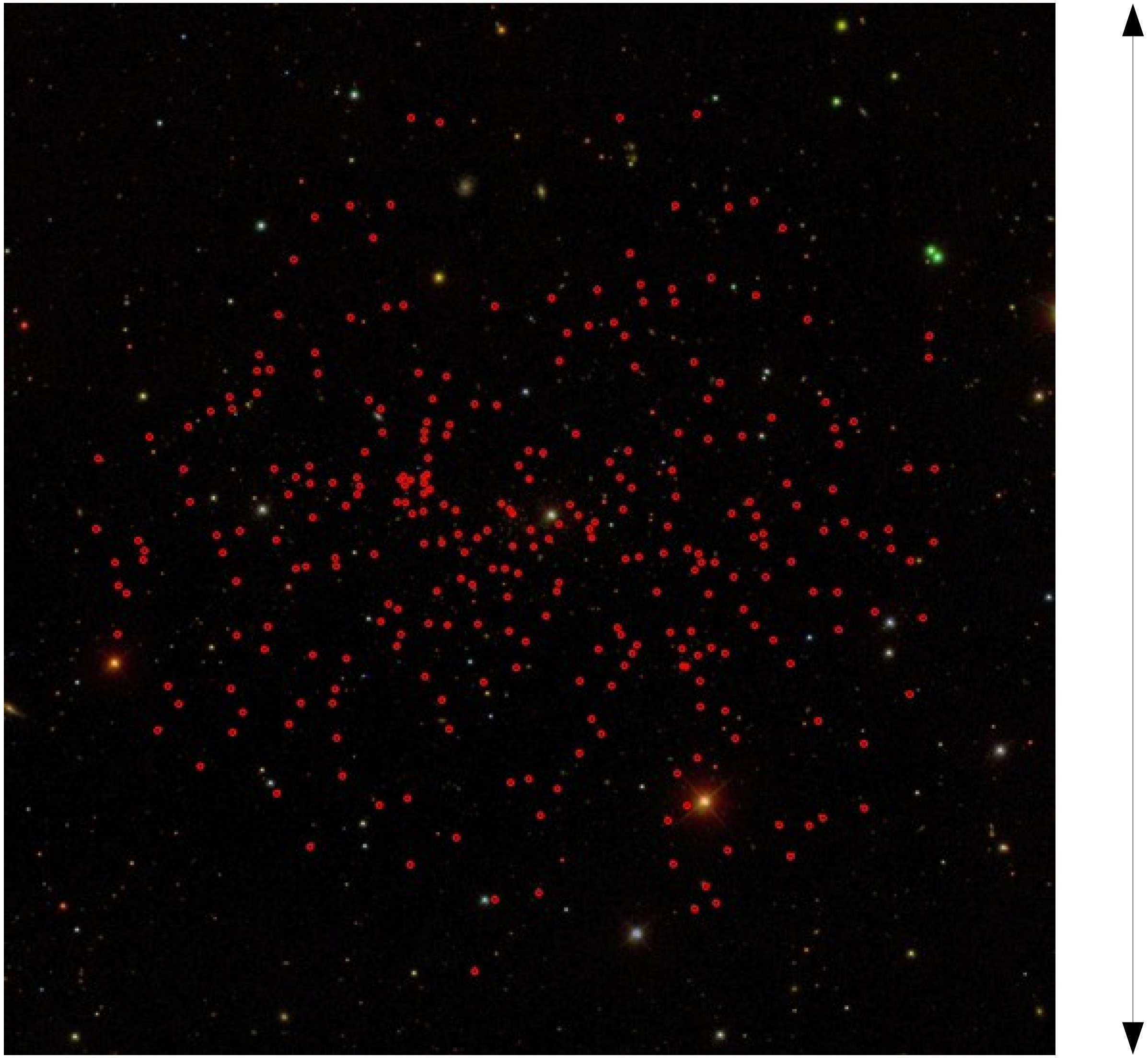
Cluster images



2048x2048 RGB images, ~0.396′′/pixel

(RGB channels roughly mapped to i-r-g frequency bands)

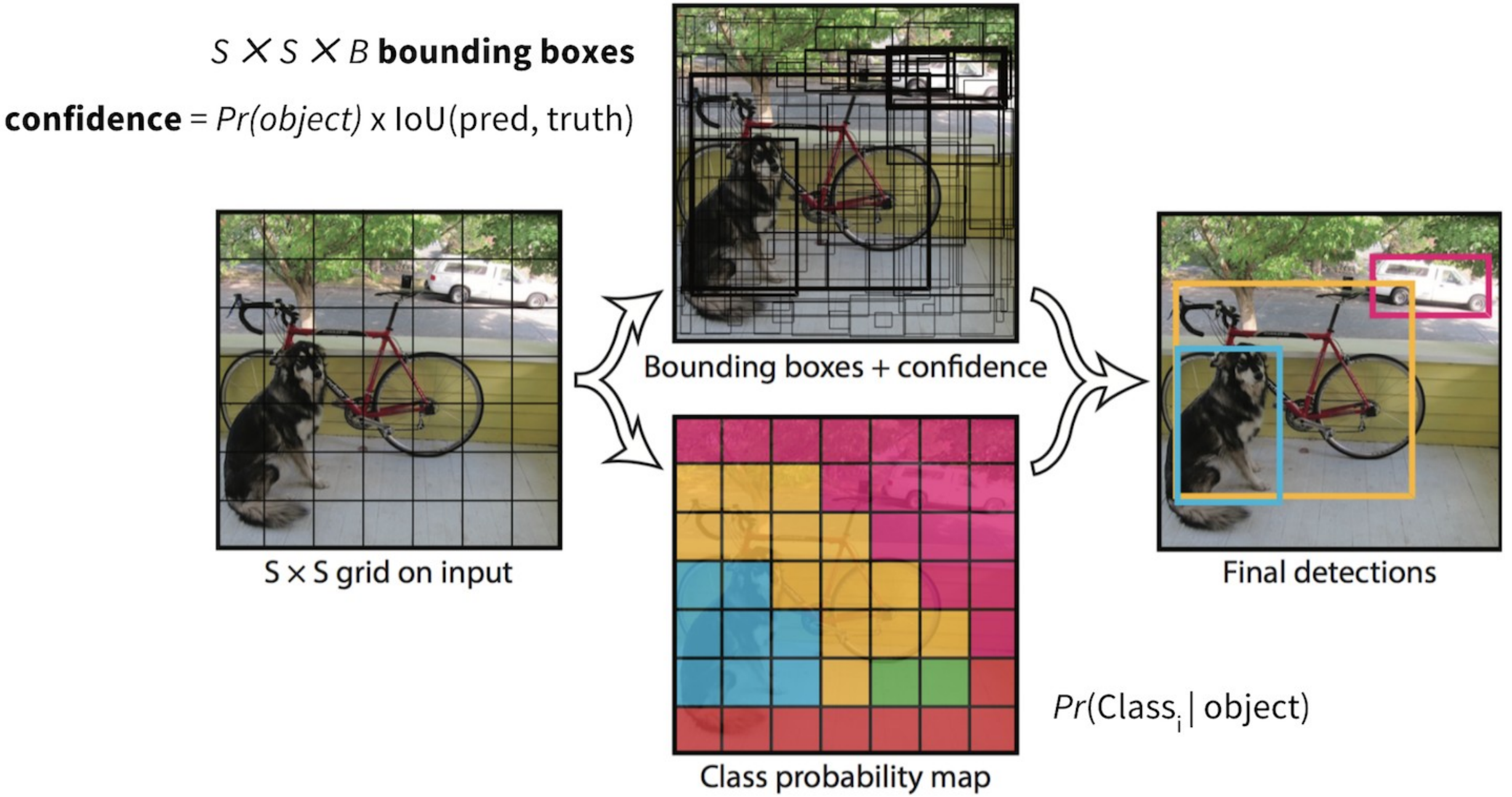
Cluster images



2048x2048 RGB images, ~0.396′′/pixel

(RGB channels roughly mapped to i-r-g frequency bands)

NN architecture: YOLOv3 (Redmon & Farhadi '18)



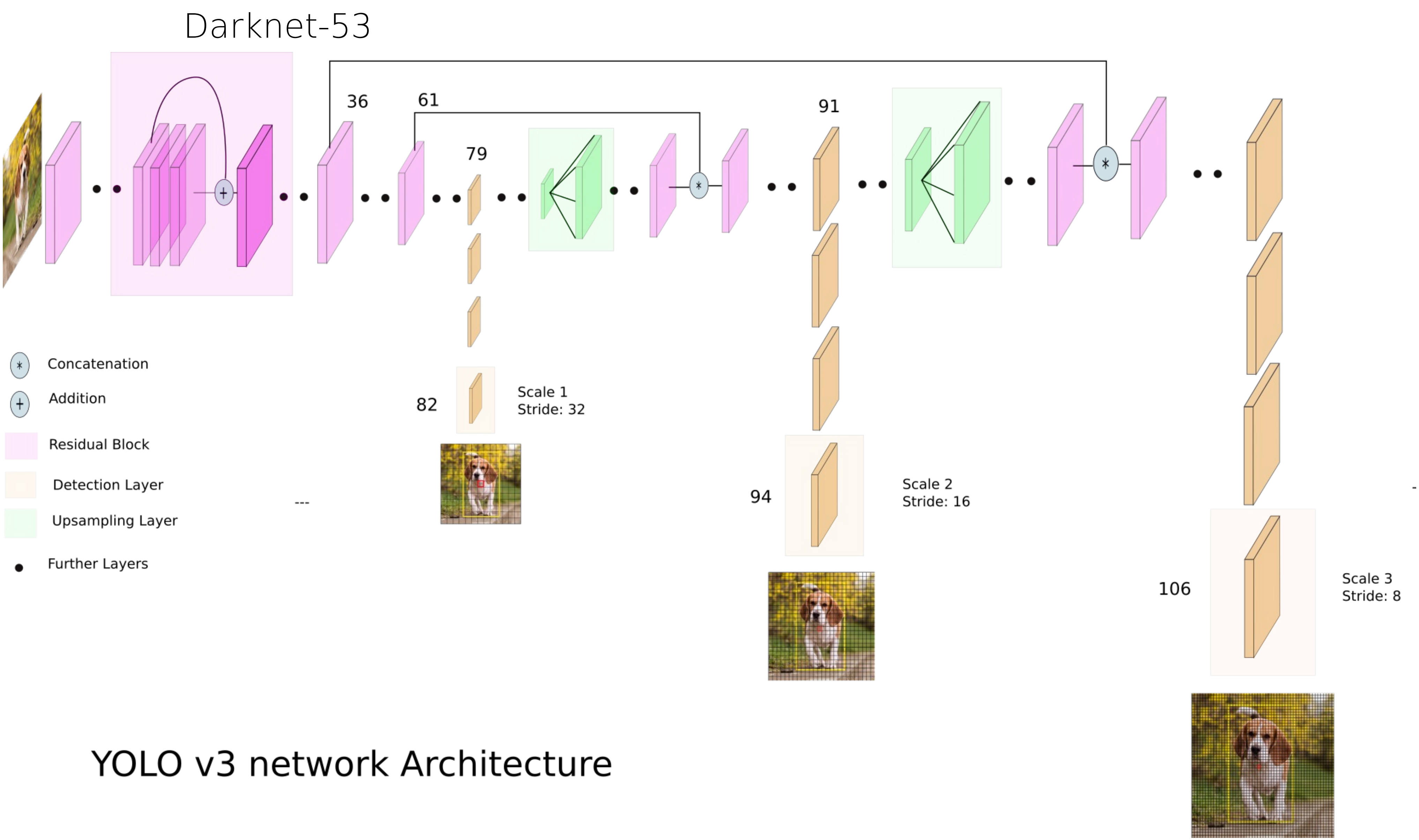
- Split image into S×S cells
- If object centre falls into cell → cell is "responsible" for detecting object
- Each cell predicts :

  (a) location of B bounding boxes (bbox)

  (b) confidence score for each bbox

  (c) probability of object class (conditioned on existence of object in bbox)

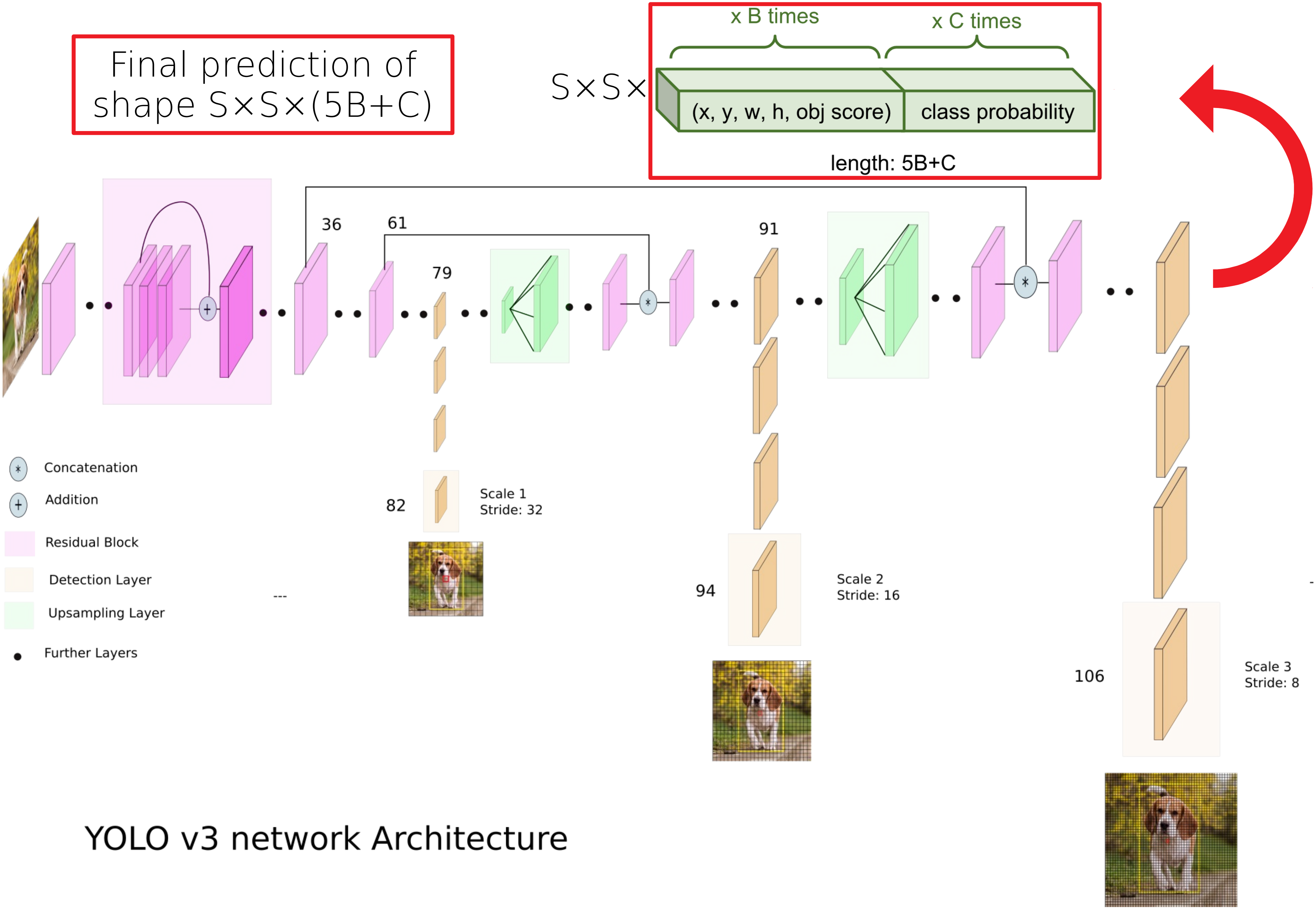
NN architecture: YOLOv3 (Redmon & Farhadi '18)



YOLO v3 network Architecture

NN architecture: YOLOv3 (Redmon & Farhadi '18)

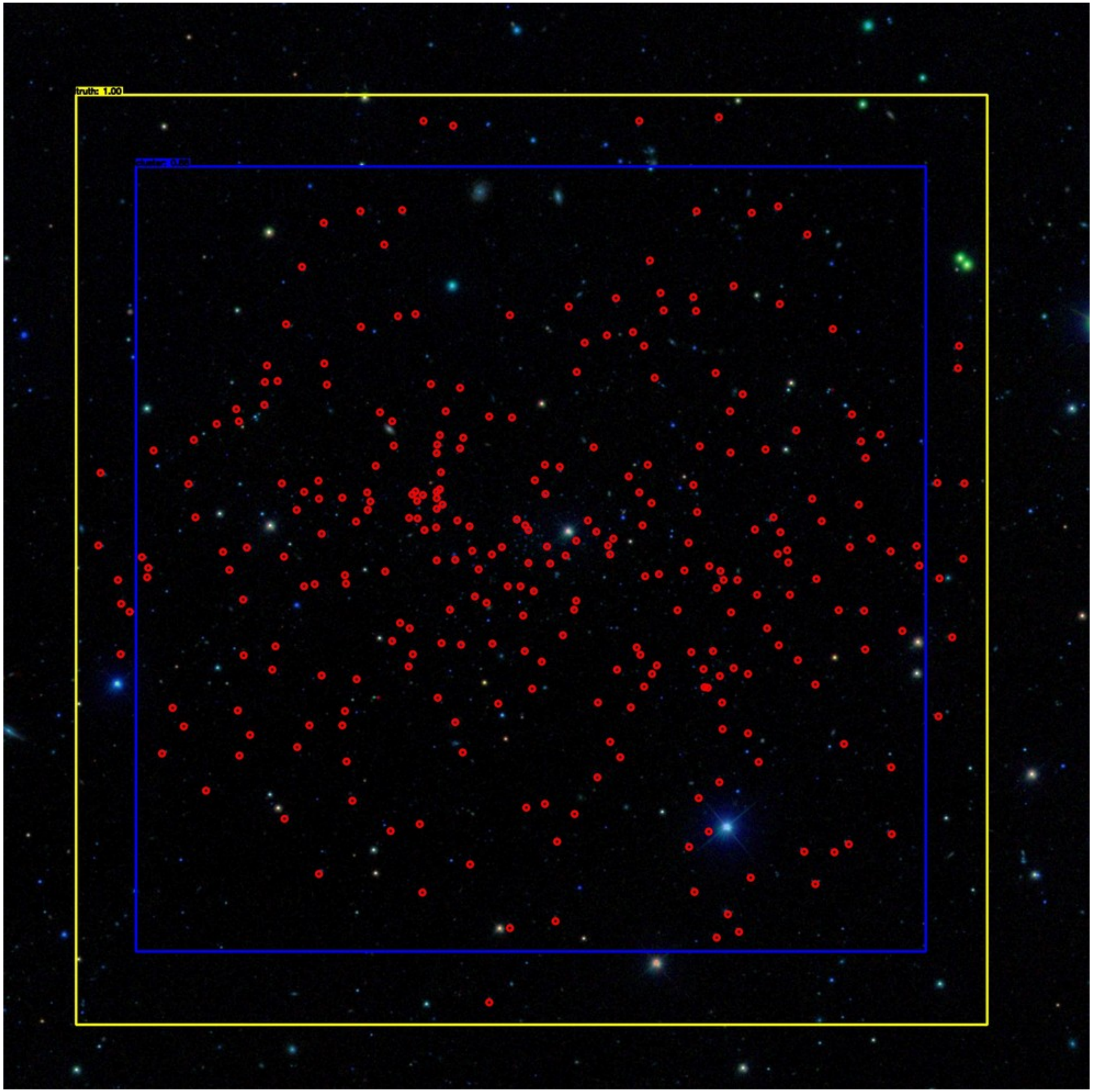
Final prediction of shape S×S×(5B+C)



YOLO v3 network Architecture

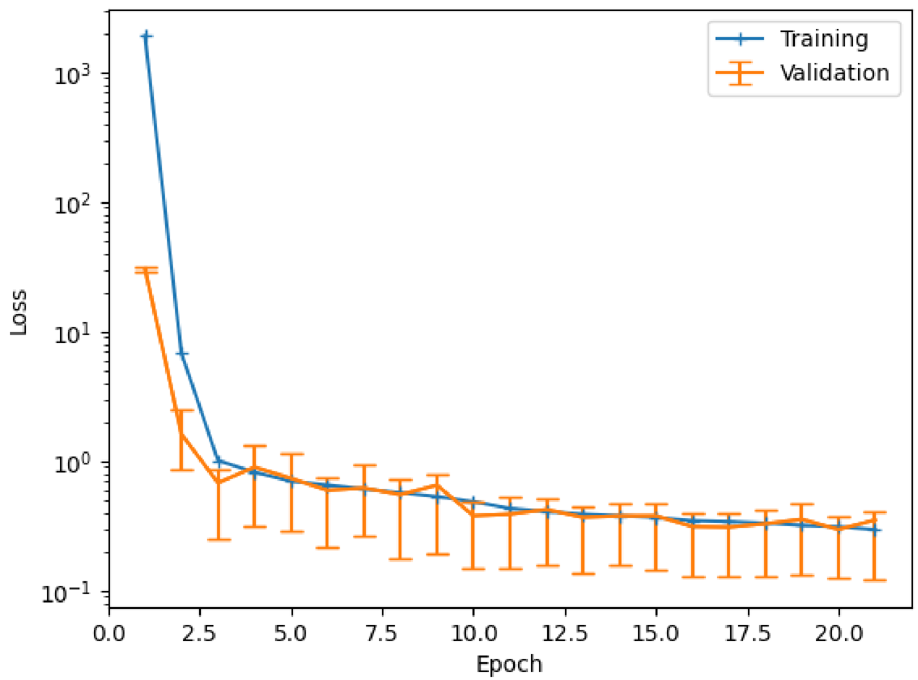
Example of YOLO application

Ilić et al. 2021, to be submitted

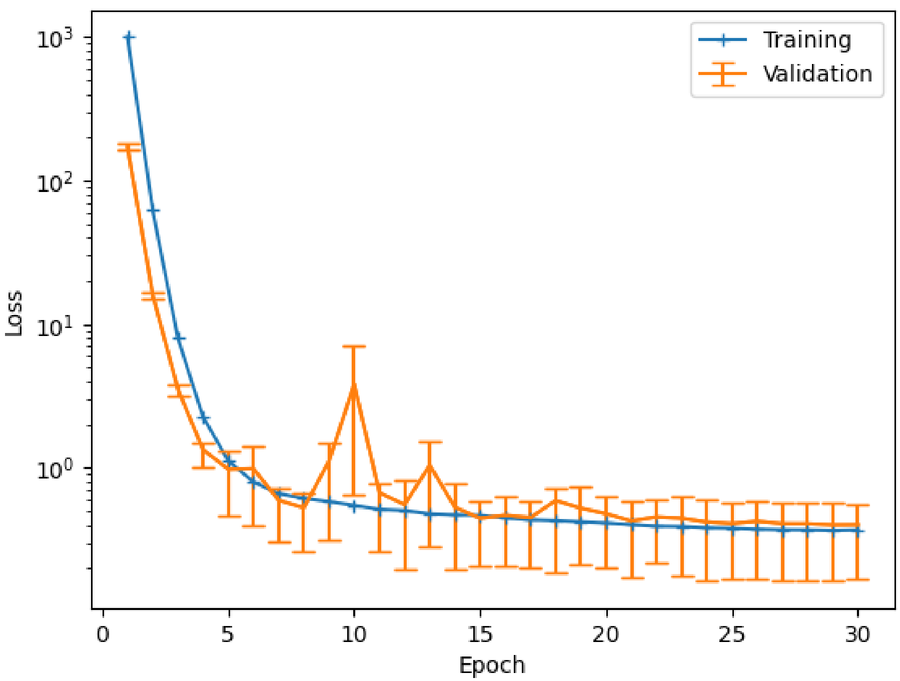


Losses

Ilić et al. 2021, to be submitted



1024x1024

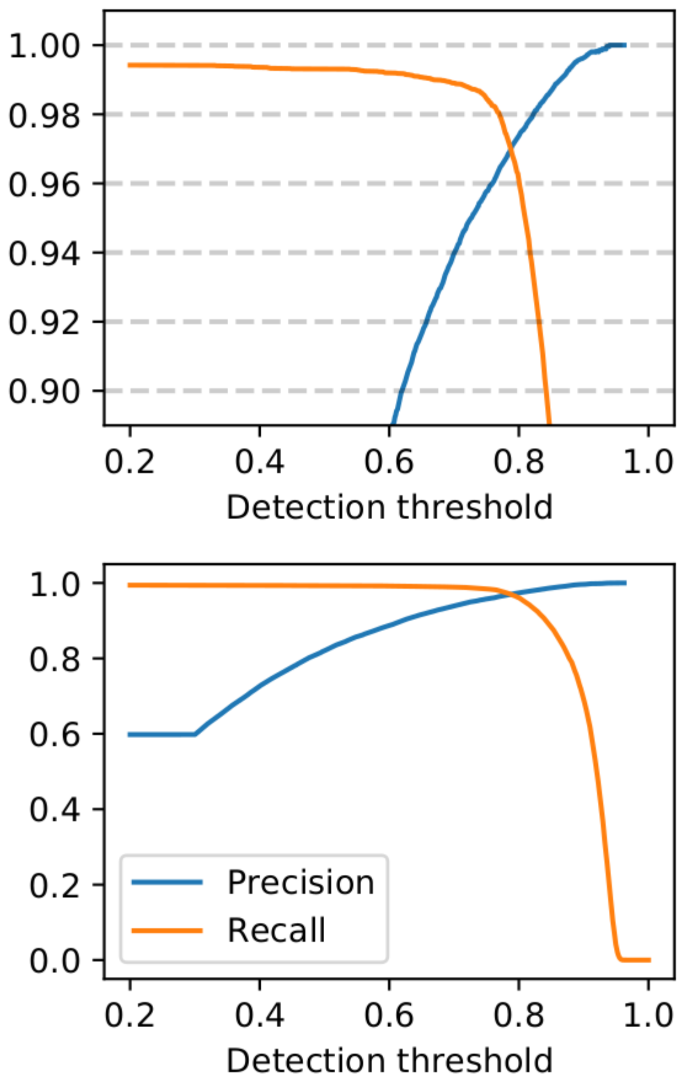


512x512

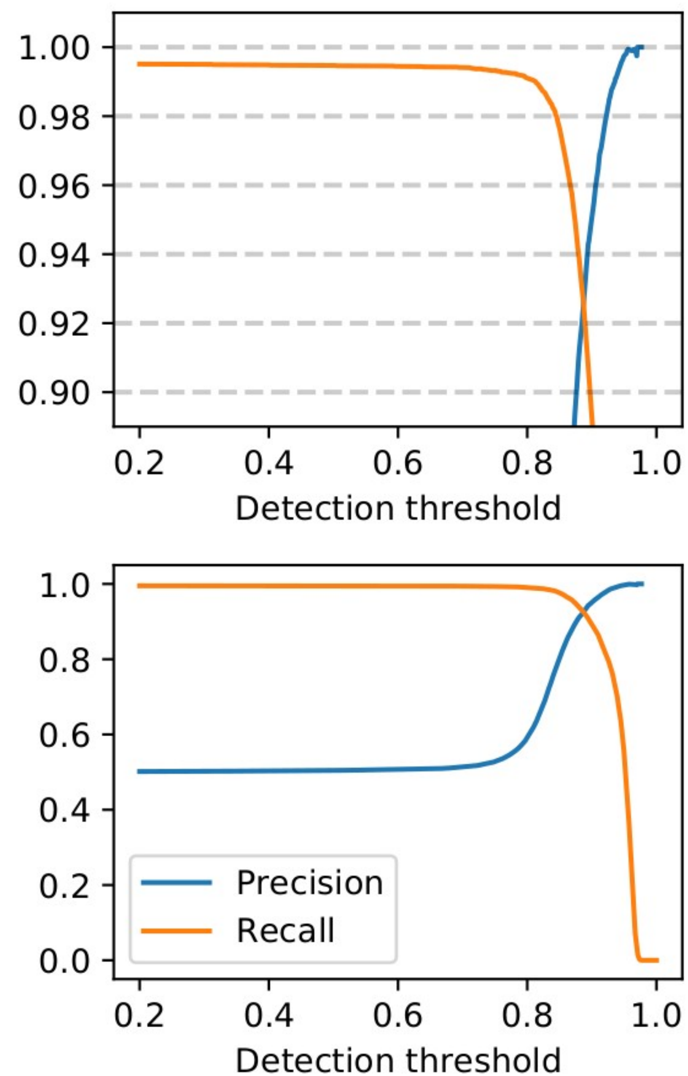
(tuning of “input size” = dimensions of 1^{st} YOLO layer)

Recall/precision

Ilić et al. 2021, to be submitted



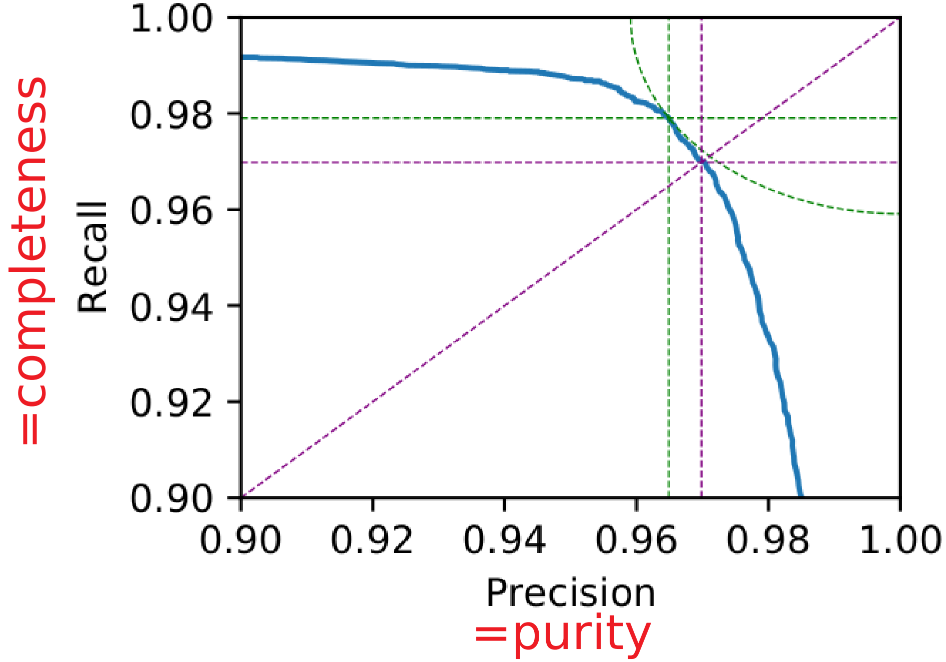
1024x1024



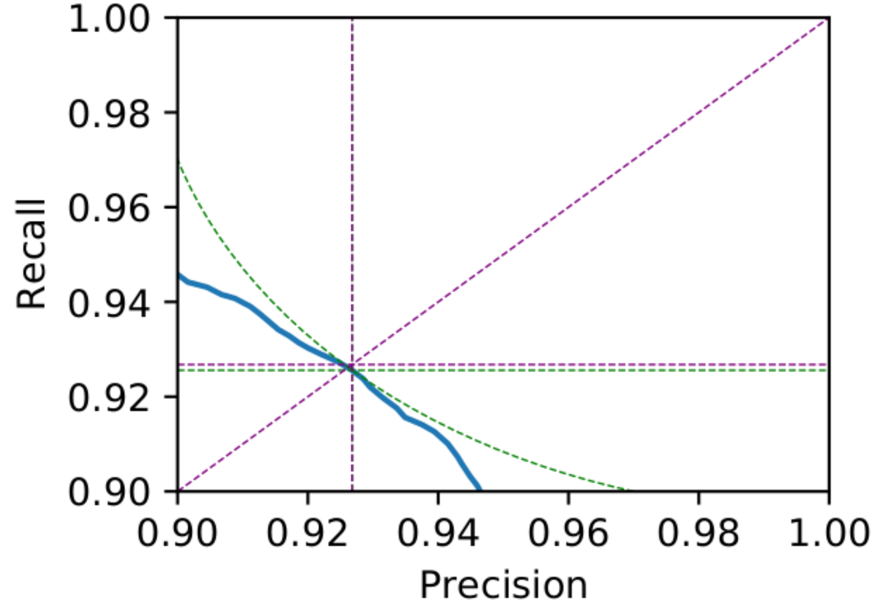
512x512

Recall/precision

Ilić et al. 2021, to be submitted



1024x1024



512x512

Conclusions and perspectives

- Fast and accurate performance from YOLOv3
- New avenue of research for cluster detection + domain-specific customisation
- Performance ultimately conditioned on training sample ($\Leftrightarrow$ redMaPPer algo)
- Eventually: training on other samples and/or simulations
- Secondary permutation-invariant NN for characterisation from galaxy list

Backup slides

YOLOv3 training: technicalities

- Pure Tensorflow implementation
- Run on NVIDIA Tesla P100 16 GB
- Usual hyper-parameters to be tuned: batch size, learning rate,...
- Additional tuning: “input size” (first layer) of YOLO network
  → we did 512x512, 1024x1024, 2048x2048

YOLOv3 training: technicalities

- Training/Validation split: 50/50
- Bbox defined as minimal box encompassing all member galaxies
- For each image/cluster, bbox of “main” cluster has to be fully in image
- For secondary clusters, bbox has to have center in image to be considered
- Total training/validation : ~12000/12000 (+ aug)
- For testing: equivalent amount of empty images

Recall/precision

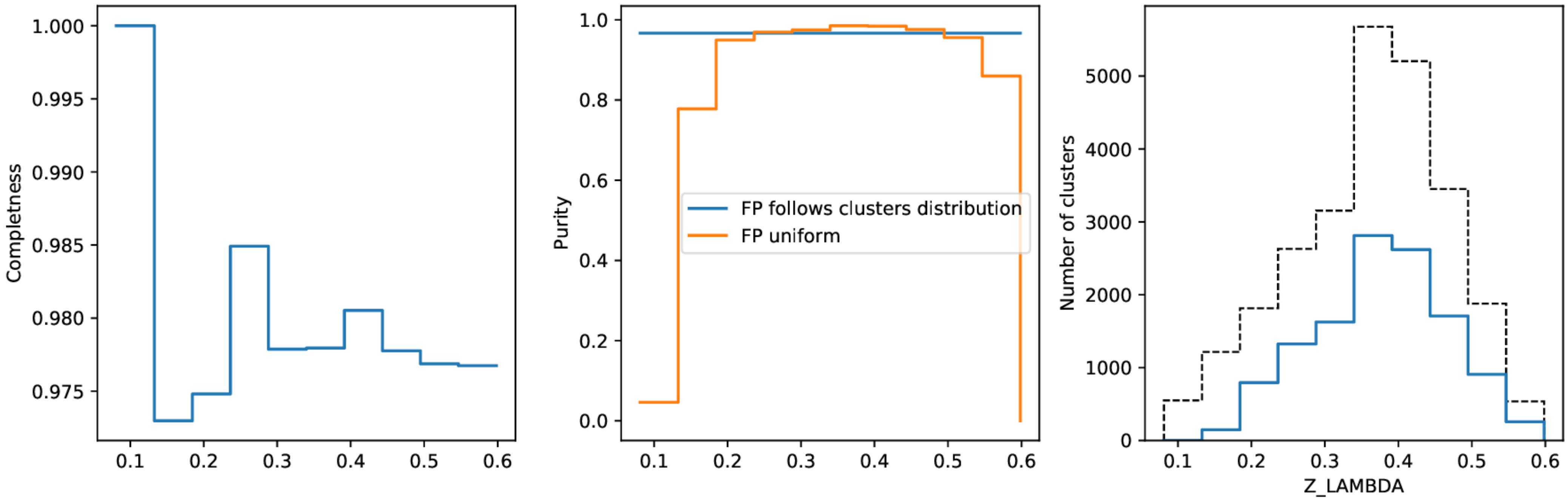
<u>"Pure counting" performance:</u>

	YOLO detects bbox	YOLO does not detect bbox
Cluster is in image	TP	FN
No cluster is in image	FP	TN

- Precision (purity) = TP / (TP + FP)
- Recall (completeness) = TP / (TP + FN)

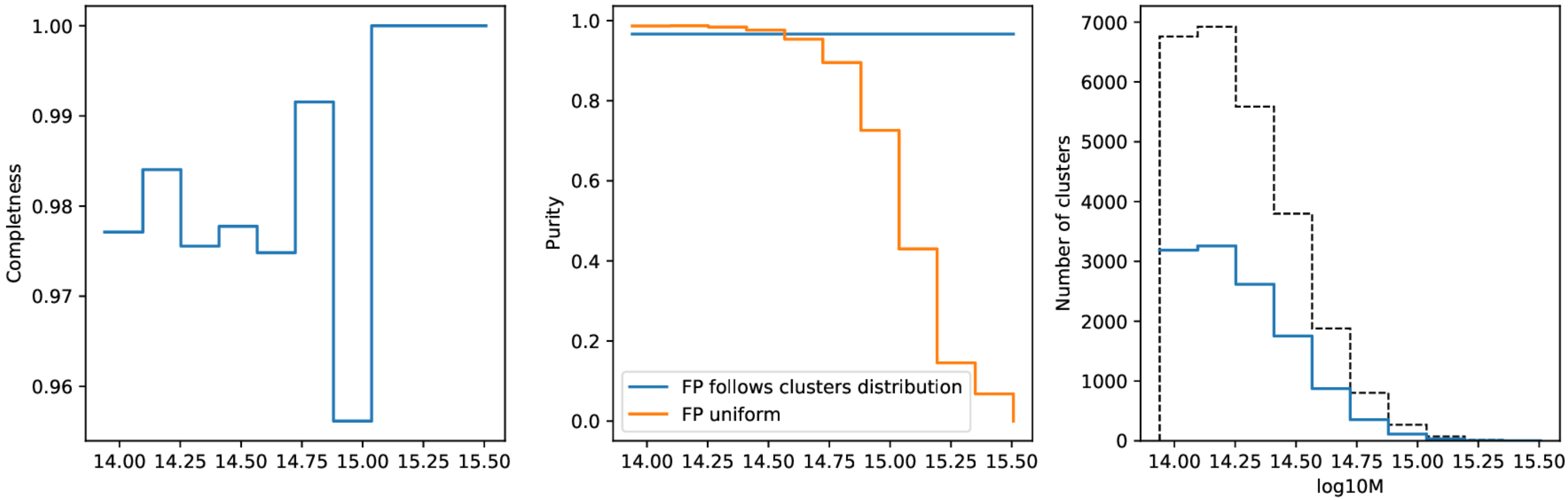
Performance as function of redshift

Ilić et al. 2021, to be submitted

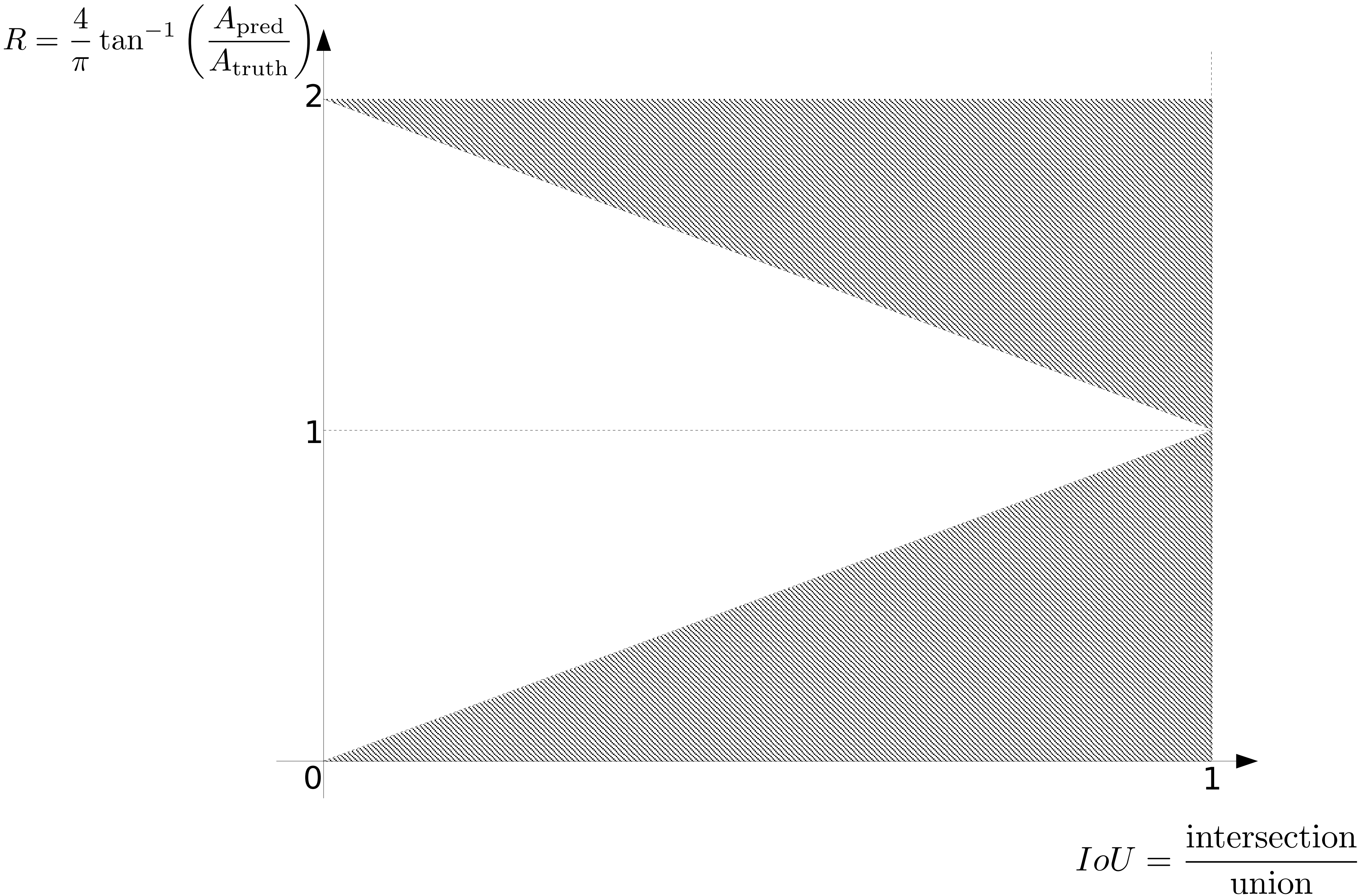


Performance as function of ($\log_{10}$) mass

Ilić et al. 2021, to be submitted



Metric for performance



Metric for performance