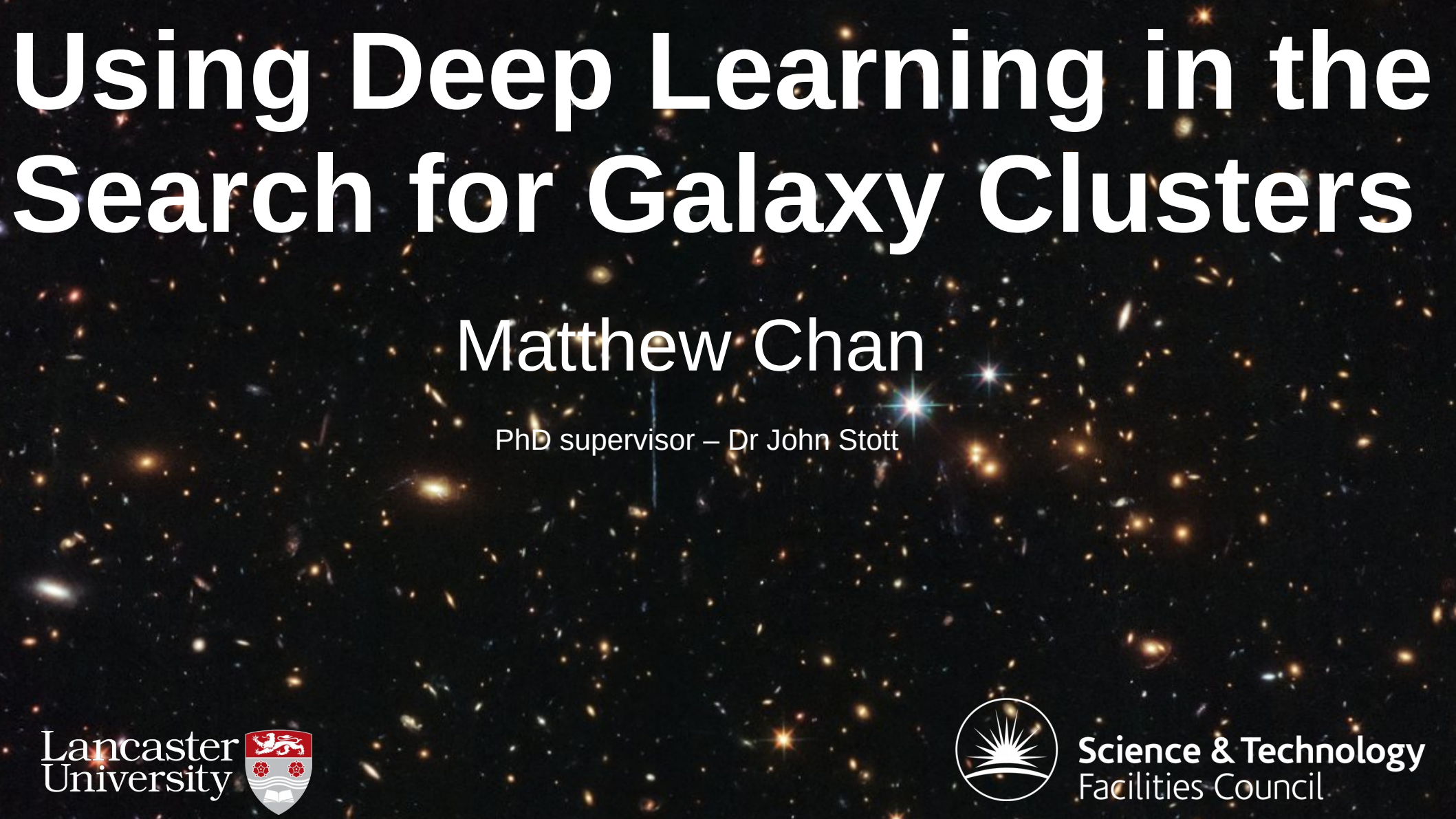
# Using Deep Learning in the Search for Galaxy Clusters

Matthew Chan

PhD supervisor – Dr John Stott

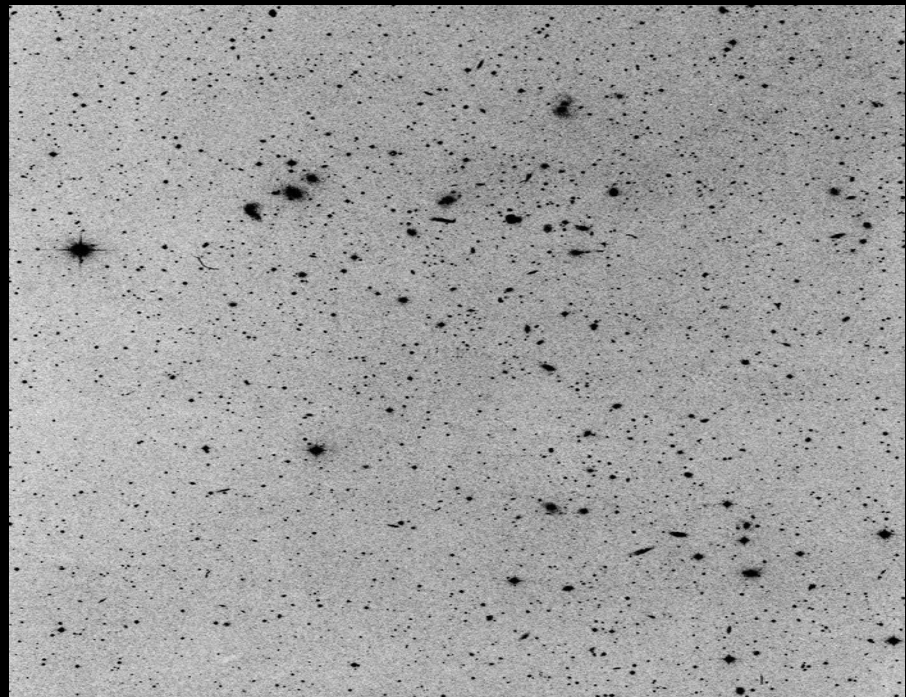


Lancaster University

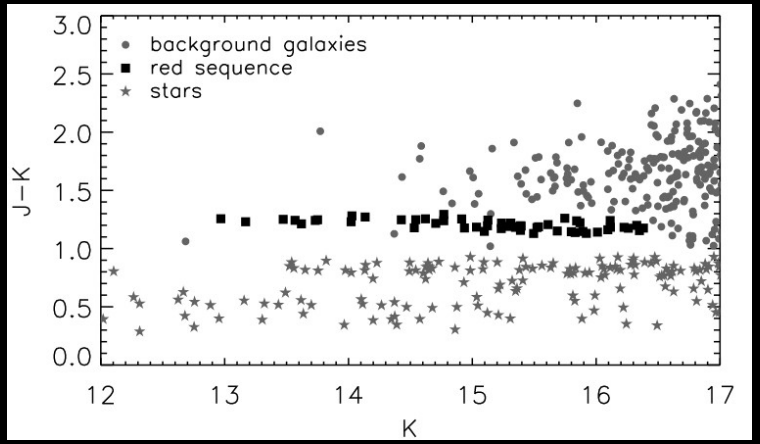
Science & Technology Facilities Council

# Introduction

- Observations of galaxy clusters are important to study the evolution of galaxies in extreme environments and determine cosmological parameters.
- George Abell created the Abell catalogue containing 4,073 galaxy clusters (Abell et al 1989).

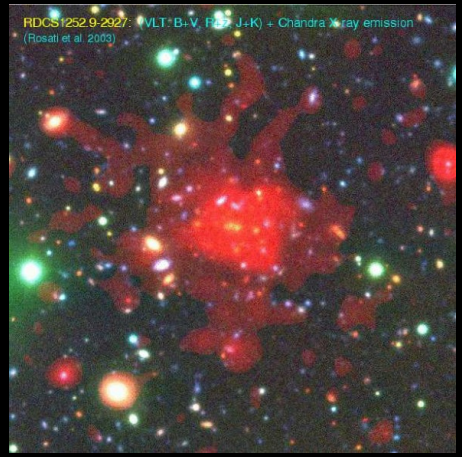



# Red Sequence Fitting



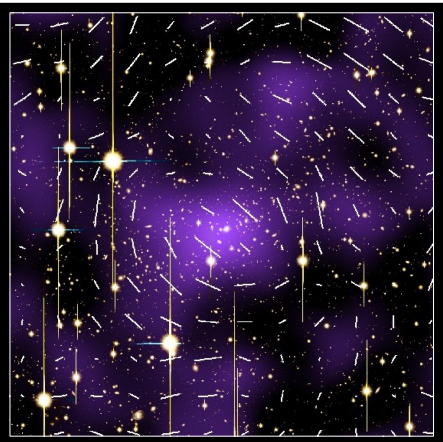

J. Stott et al 2009.

# X-ray Emissions



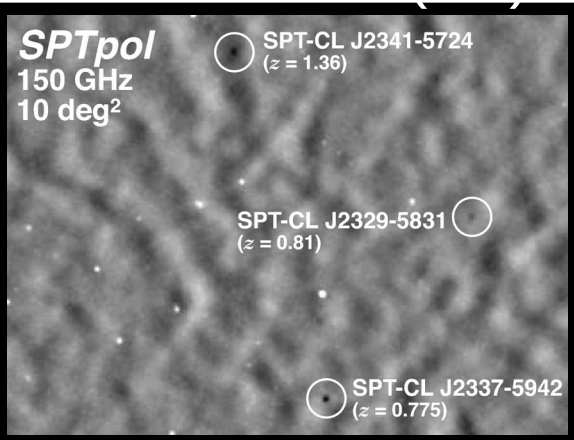

P. Tozzi 2007.

# Weak Gravitational Lensing



M. Oguri et al. 2010

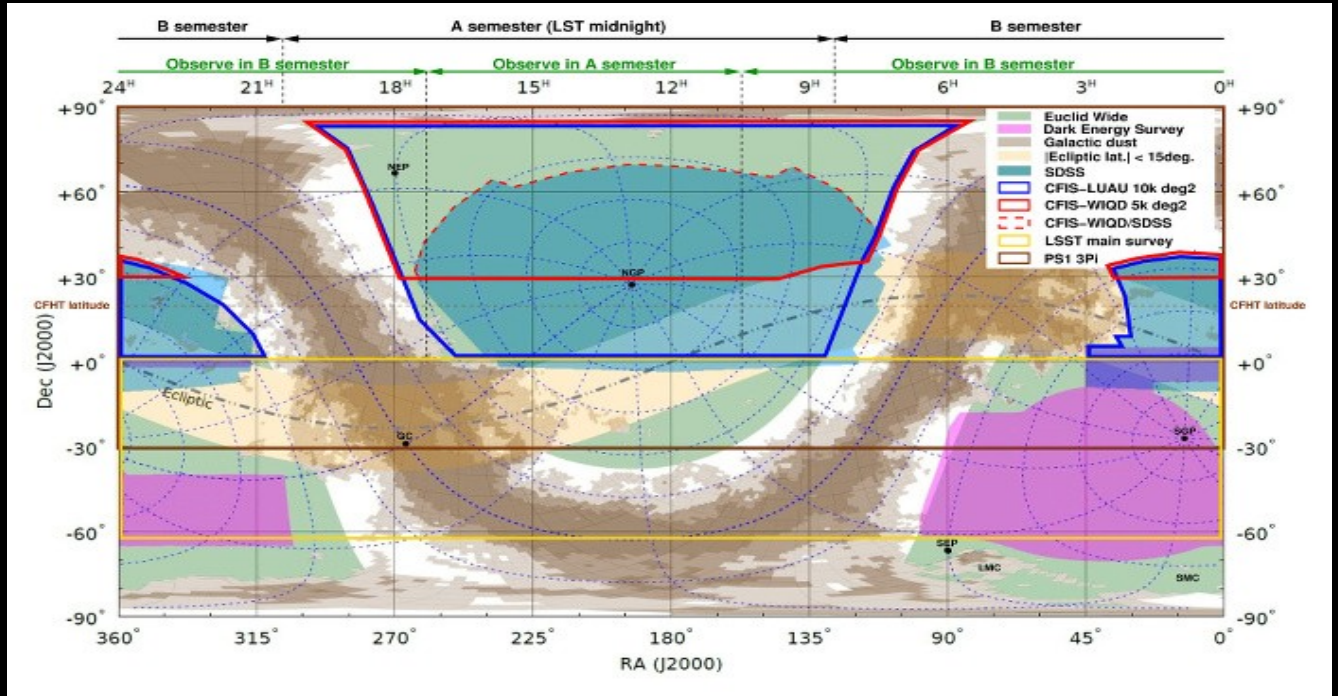
# Sunyaev-Zeldovich (SZ) Effect




B.A. Benson et al. 2014.

# Deep Learning Approach with LSST

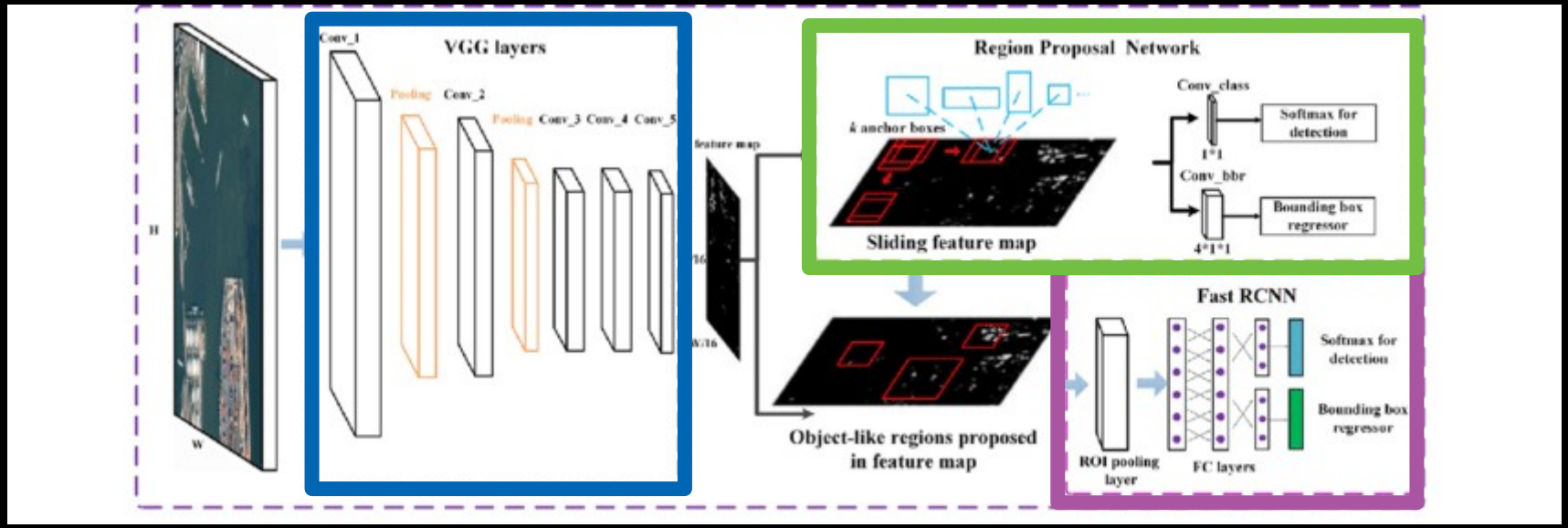
- LSST will scan the entire southern sky with an estimated 15TB of data generated per night.
- We can apply a deep learning approach to search for galaxy clusters as we will have lots of imaging data to work with.




Credit: T. Dwelly

# Deep Learning Model

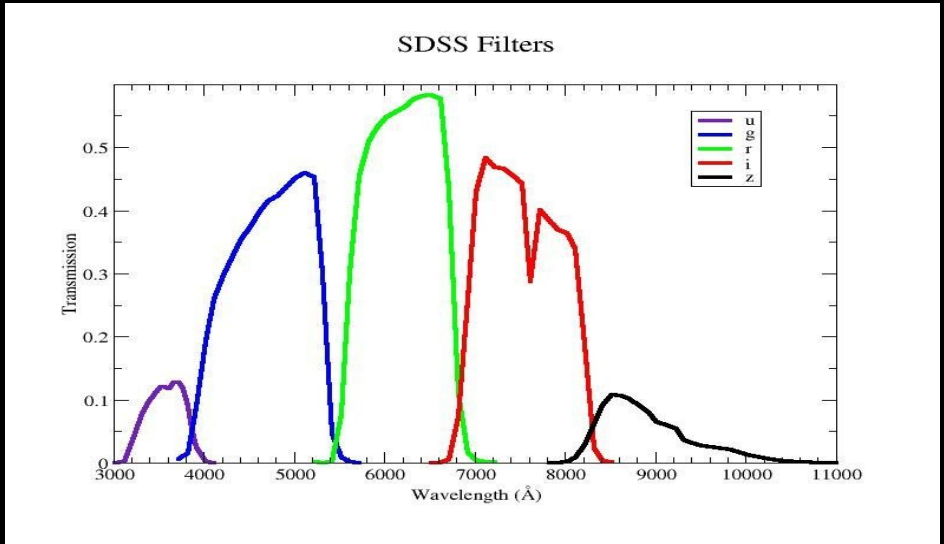
- We use a popular object detection algorithm known as Faster-RCNN (S. Ren et al 2015).
- Faster-RCNN is composed of three individual networks that work collectively to localize and identify an object.




Z. Deng et al. 2018

# Training and Test Sets

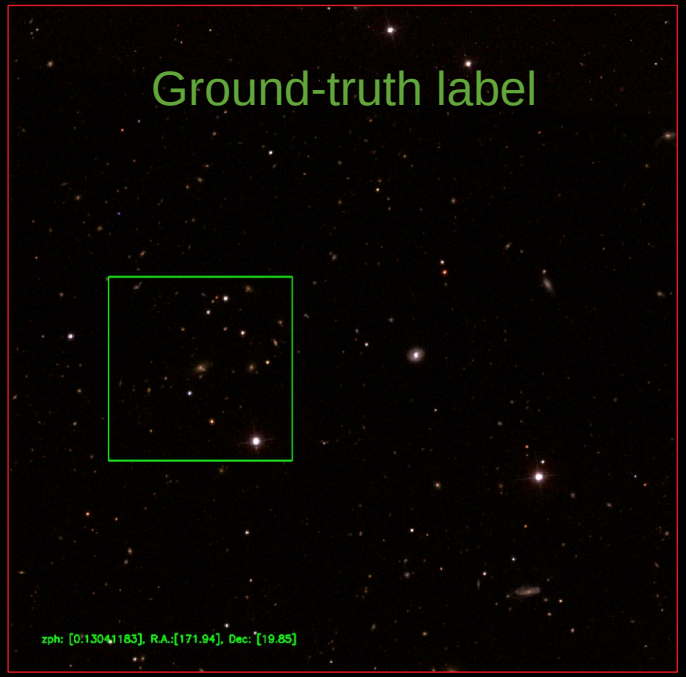
- We extract RA and Dec coordinates of the Abell galaxy clusters identified in Wen et al. 2012 catalogue.
- We select galaxy clusters between $0.1 < z < 0.2$ and contains $20 \geq$ members inside an $R_{200}$ radius from the cluster centre.
- We use the i, r, g filters on the SDSS camera as the RGB channels to form coloured images.




Credit: A. Faisst

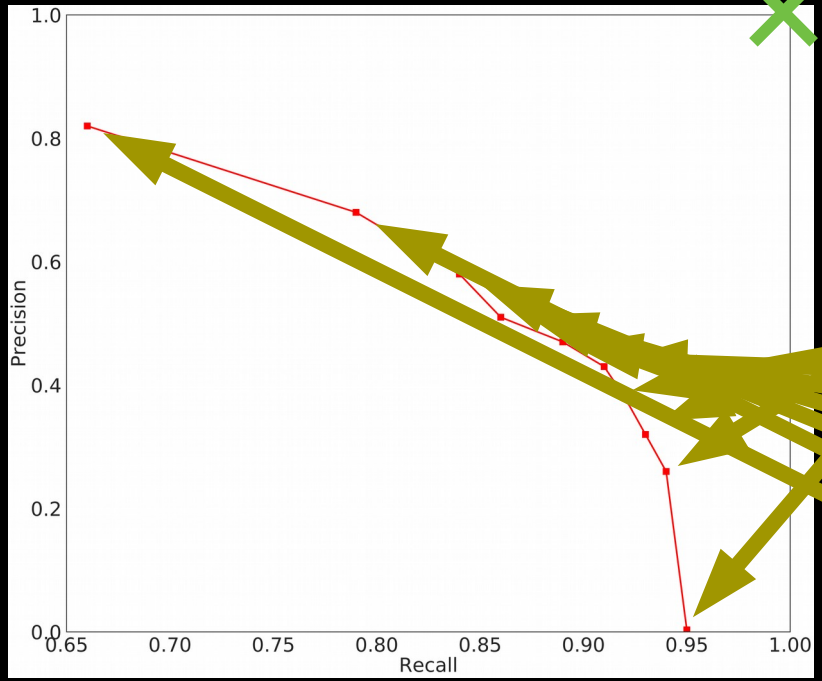
# Image Pre-Processing

- We apply a random offset sampled from a uniform distribution.
- Set fixed image sizes of 2000x2000 pixels (approximately 1443×1443 kpc at z=0.1).
- We use a non-linear transformation to 'stretch' the contrast of the image.



# Model Analysis on Test Set

- We use common evaluation metrics such as precision, recall and F1-score.
- We want a confidence score threshold that maximizes F1-score as a higher F1-score signifies a better balance between precision and recall.




| Confidence threshold (%) | Precision | Recall | F1-score |
|---|---|---|---|
| 0 | 0.0032 | 0.95 | 0.0063 |
| 10 | 0.26 | 0.94 | 0.41 |
| 20 | 0.32 | 0.93 | 0.48 |
| 30 | 0.38 | 0.92 | 0.54 |
| 40 | 0.43 | 0.91 | 0.58 |
| 50 | 0.47 | 0.89 | 0.61 |
| 60 | 0.51 | 0.86 | 0.64 |
| 70 | 0.58 | 0.84 | 0.69 |
| 80 | 0.68 | 0.79 | 0.73 |
| 90 | 0.82 | 0.66 | 0.73 |
| 100 | 0.00 | 0.00 | 0.00 |

# Summary

- We present a novel approach to detect galaxy clusters from coloured images.
- We could potentially discover many more galaxy clusters that have never been seen before at higher redshift or with lower mass.
- We also want to build upon our existing model by including methods to examine further intrinsic properties of galaxy clusters.