Deep Learning for Blended Source Identification in Galaxy Survey Data

Samuel Farrens









Outline



- Cosmology
- Weak Gravitational Lensing
- Deblending

Deep Learning

- Simulations
- Networks
- Results
- Future Work













Years after the Big Bang



Background

ter the Big Bang



- 10
- edshift + 1















Cosmology



Evolution of the Universe

Content of the Universe

- Dark matter does not interact electromagnetically, it only interacts gravitationally.
- How can we work out how much dark matter there is in the Universe?





Dark Matter



Cosmic Web

Galaxy Cluster

- Galaxies and gas are tracers of dark matter haloes, but with significant biases.
- The shapes of background galaxies provide a more reliable tracer.





Gravitational Lensing





Gravitational Lensing

Strong Gravitational Lens

- Background galaxies are distorted by foreground dark matter.
- Strongly lensed systems can be seen by eye, but weak lensings is a 1% effect.





Galaxy Shapes































































Euclid Mission

- 6-year space mission designed to measure galaxy shapes to highest possible accuracy.
- The Euclid consortium consists of >1500 members from 16 countries!





http://www.euclid-ec.org/















http://www.cfht.hawaii.edu/Science/CFIS/





CFIS Pipeline







Alexandre Bruckert

Identification of blended sources in CFIS images.





CANADA-FRANCE IMAGING SURVEY





Unlabelled Data







Unlabelled Data







Simulations

- Galaxy images simulated using GalSim (Rowe et al. 2015)
- Mimic the properties of CFIS *r*-band.







Simulations

- Galaxy images simulated using GalSim (Rowe et al. 2015)
- Mimic the properties of CFIS *r*-band.





Simulations



Mandelbaum et al. (2015)

- Large variety of simulated galaxy images.
- Artificially blended half of the sample and added noise.
- Created training, validation and testing sets.
- Transfer learning.



SExtractor





- Community standard for source extraction.
- Performs star-galaxy separation.
- Includes deblending system that identifies blends using fixed thresholds.



Siamese Network



- Finds similarities between two images.
- Shared weights means fewer parameters to train.
- Usually most robust to imbalance.
- Does not require huge number of labelled training images.



Siamese Network



- Finds similarities between two images.
- Shared weights means fewer parameters to train.
- Usually most robust to imbalance.
- Does not require huge number of labelled training images.



Siamese Network



- Finds similarities between two images.
- Shared weights means fewer parameters to train.
- Usually most robust to imbalance.
- Does not require huge number of labelled training images.



Siamese Network



- Finds similarities between two images.
- Shared weights means fewer parameters to train.
- Usually most robust to imbalance.
- Does not require huge number of labelled training images.





Siamese Network



- Finds similarities between two images.
- Shared weights means fewer parameters to train.
- Usually most robust to imbalance.
- Does not require huge number of labelled training images.



One-Class Classification



- Trained to find one particular class.
- Robust to transfer learning.
- Does not require huge number of labelled training images.







Simonyan & Zisserman (2018)

- Pre-trained network (e.g. using IMAGENET)
- Simple architecture.















Simonyan & Zisserman (2018)

- Pre-trained network (e.g. using IMAGENET)
- Simple architecture.









Simonyan & Zisserman (2018)

- Pre-trained network (e.g. using IMAGENET)
- Simple architecture.





Noise-less Simulations







Noisy Simulations







CFIS Simulations



Sextractor	45.31 %
VGG16	74.85 %
Siamese Networks	55.23 %
One-Class	61.07 %



Results

Future Work

- Build deblending module into CFIS pipeline.
- Investigate how blends identified with VGG-16 impact science.
- Add multi-class classification.
- Investigate benefits of segmentation of identified blends.

