CAASTRO
ARC CENTRE OF EXCELLENCE FOR ALL-SKY ASTROPHYSICS

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

DASH
Deep Learning for the Spectral Classification of Supernovae

Daniel Muthukrishna

The University of Queensland /
Australian National University

Australian Government
Australian Research Council

DARK ENERGY SURVEY

OzDES

Big Data Era in Astronomy

- Incredible time for astronomy!
- Unprecedented amounts of data
  - Dark Energy Survey, SDSS, LSST, SKA, and more to come…
- We need to be able to process the Petabytes of data quickly
- Astroinformatics will lead the new era of astronomy
  - Deep Learning
  - Parallelisation
- Open-source
  - DASH: https://github.com/daniel-muthukrishna/DASH

Outline

- OzDES: Dark Energy Survey
- Supernova spectra – why is classification difficult?
- Previous classification methods
- DASH Design Decisions
- Data description
- Why Deep learning?
  - Tensorflow
- Processing Spectra
- Interfaces
- Performance of DASH

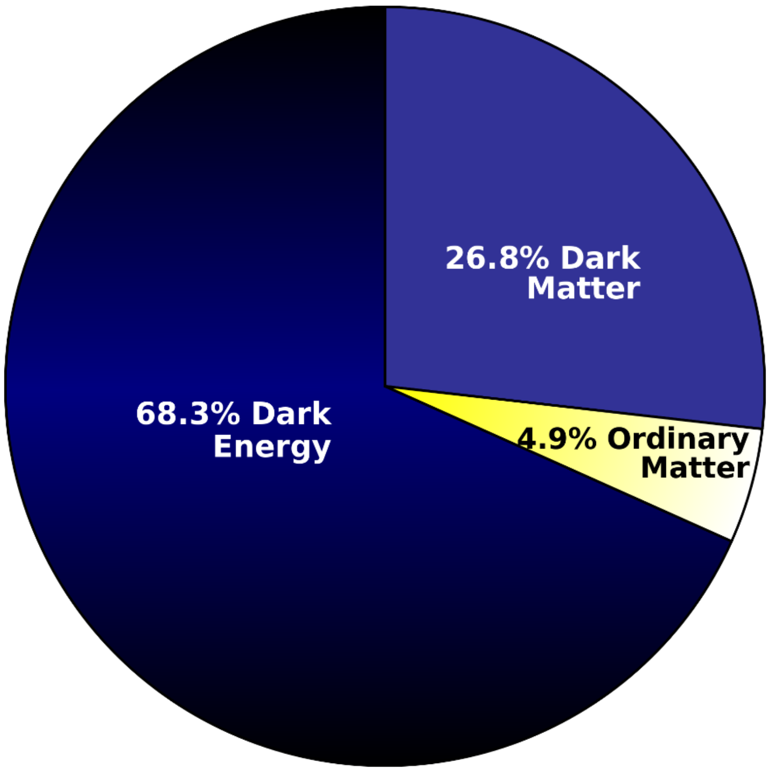
Acknowledgements

Thanks to David Parkinson, Brad Tucker, Samuel Hinton, Tamara Davis, Chris Lidman, Stephane Blondin

Obelix Supercomputer, University of Queensland

Dark Energy

- 2011 Nobel Prize
  - Perlmutter, Schmidt (ANU), Riess
- The universe's expansion is accelerating…
- This leaves a few options
  - Einstein's theory of gravity is wrong on cosmological scales?
  - Most of our universe's energy comes from Dark Energy?
  - A fifth fundamental force?
- ΛCDM fits very well – but mysterious



Type Ia Supernovae

- Type Ia Supernovae have provided the most compelling evidence of cosmic acceleration
- Standardizable Candles
- Several surveys are aiming to increase the dataset to understand Dark Energy/modified theories of gravity

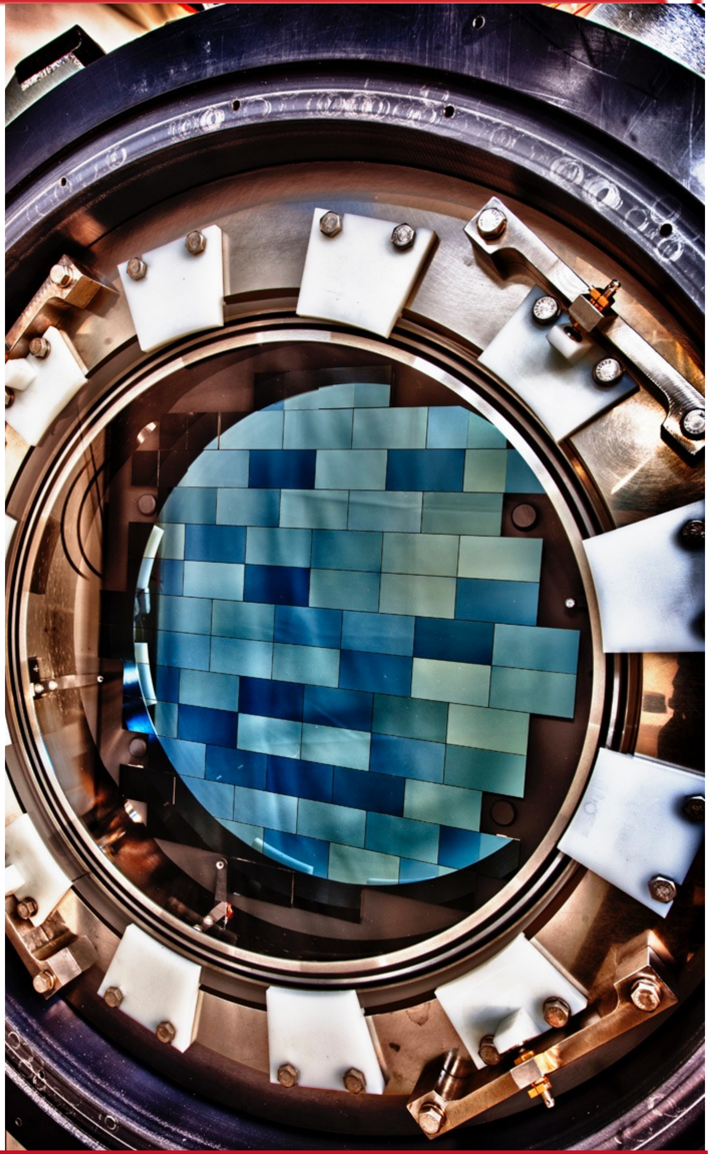
Dark Energy Survey (DES)

- Weak Lensing
- Galaxy clusters
- Baryon Acoustic Oscillations
- Supernovae

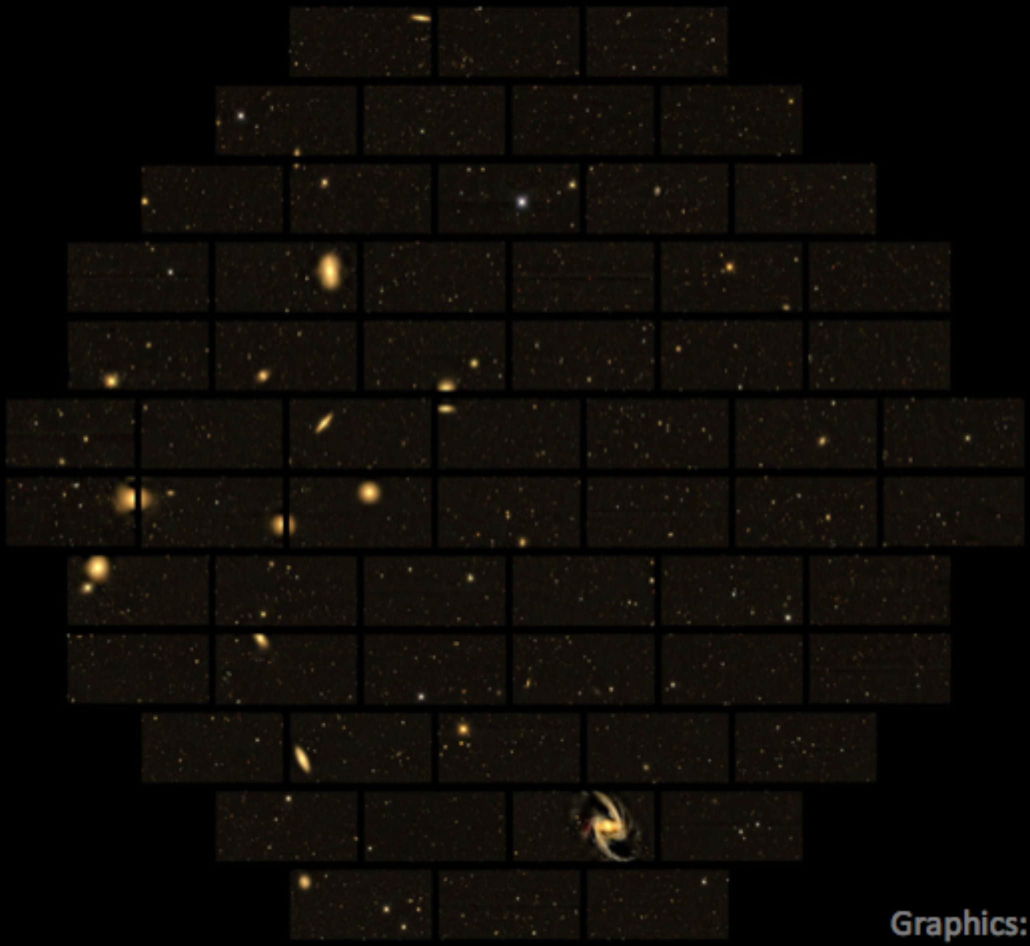
300 million galaxies!!

3000 supernovae!!

525 Nights
5000 square degrees
5 filters
Time-lapse over 30 square degrees
~500 researchers



THE DARK ENERGY SURVEY



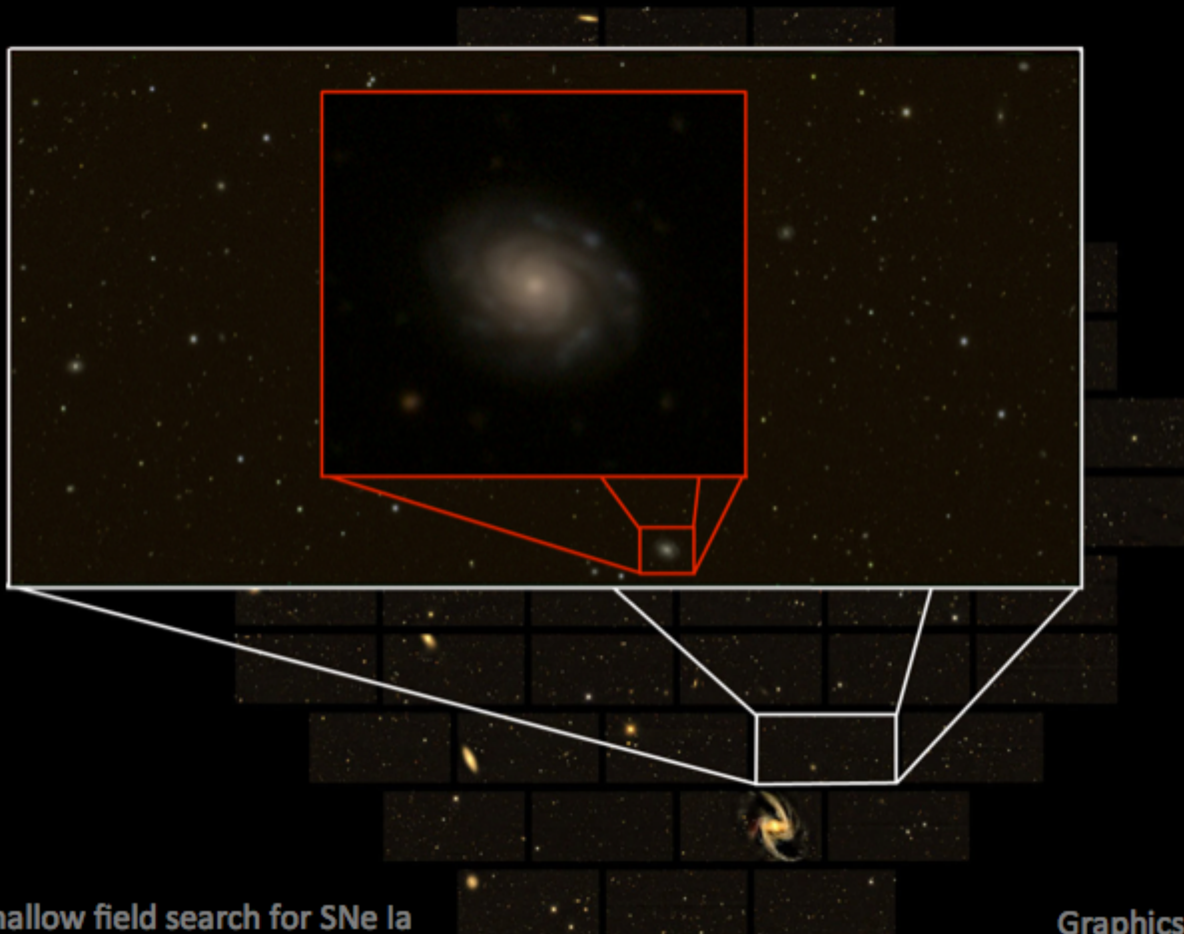
Graphics: C. D'Andrea

THE DARK ENERGY SURVEY

Shallow field search for SNe Ia

Graphics: C. D’Andrea

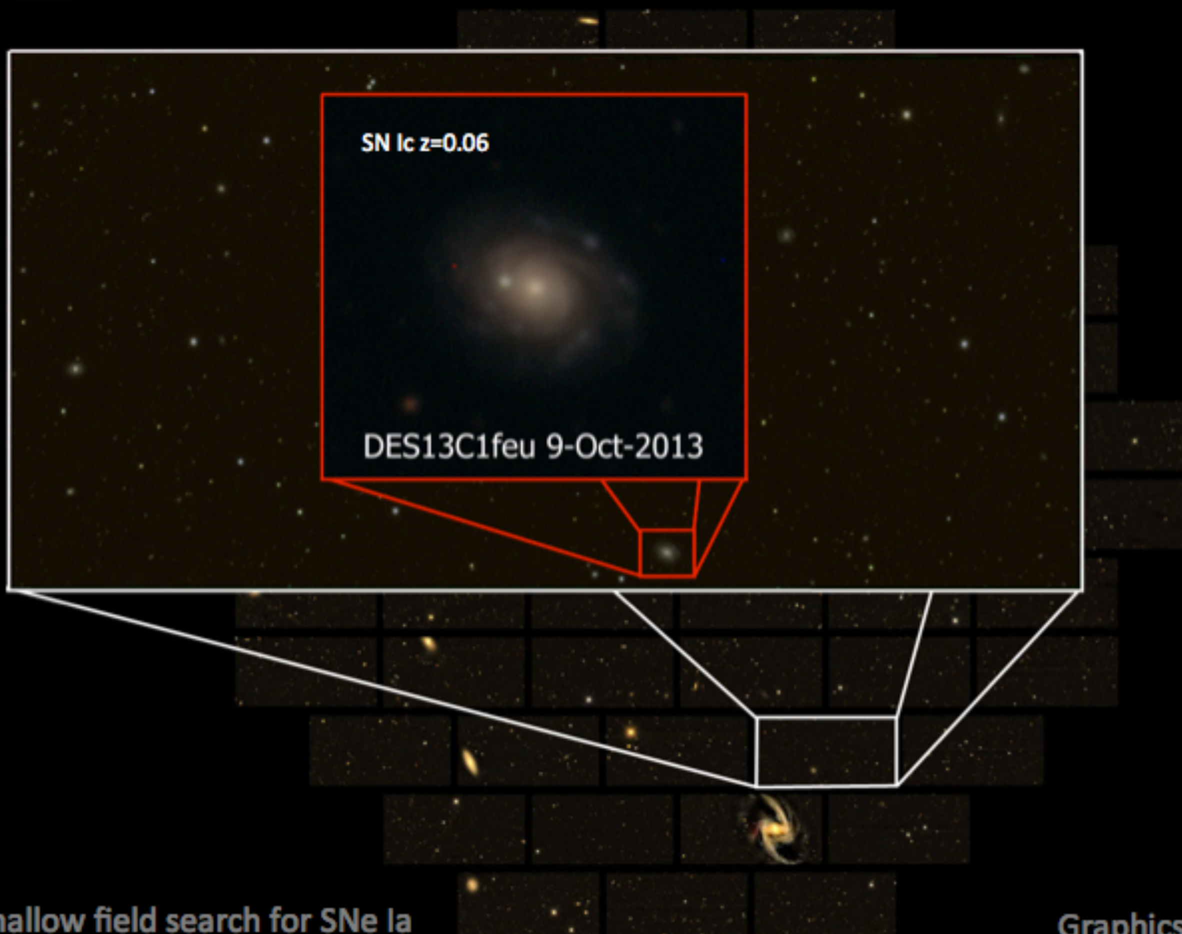
THE DARK ENERGY SURVEY



Shallow field search for SNe Ia

Graphics: C. D’Andrea

THE DARK ENERGY SURVEY

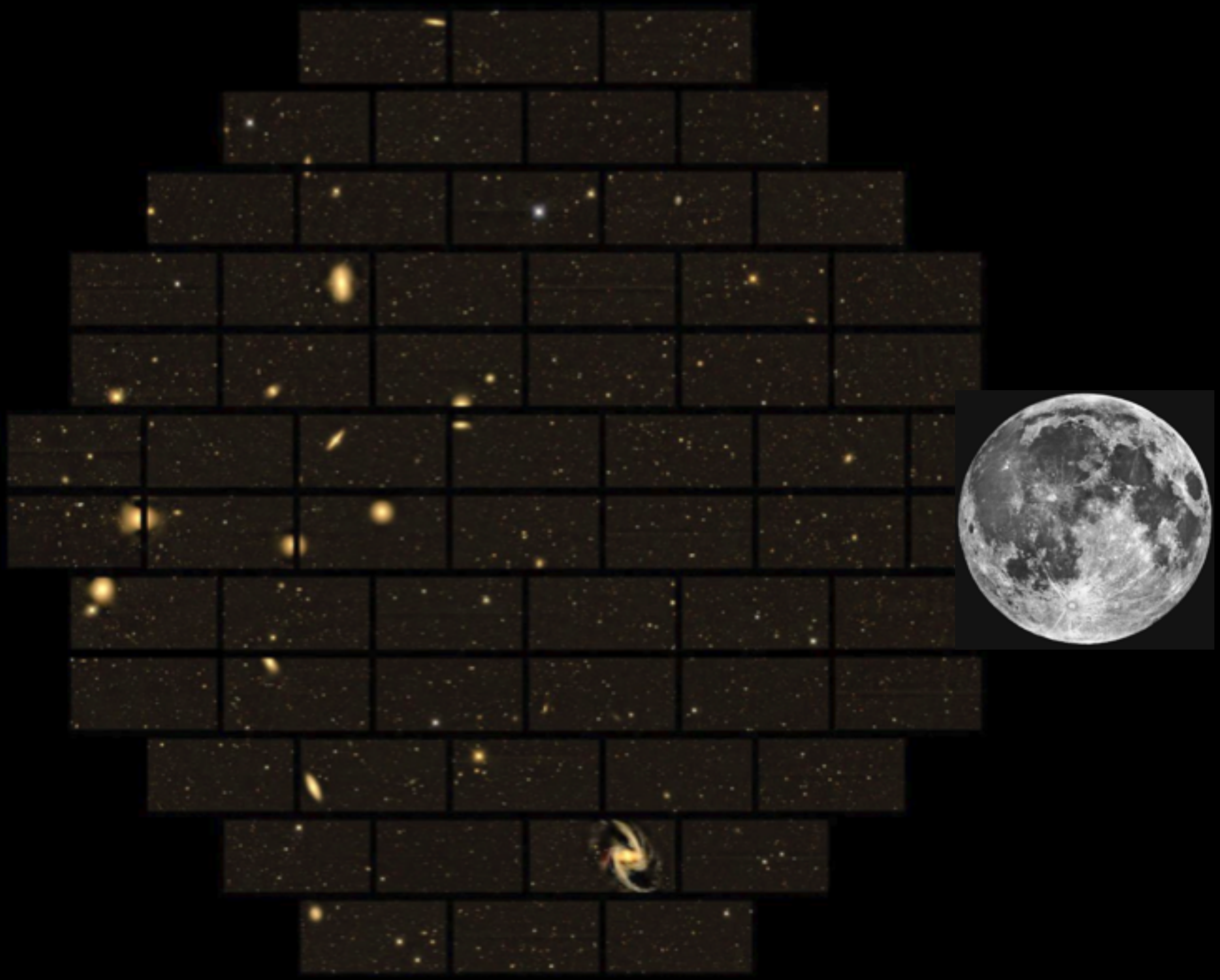


Shallow field search for SNe Ia

Graphics: C. D'Andrea

OzDES
Australian Dark Energy Survey

- 5-year 100-night spectroscopic survey on the Anglo Australian Telescope (AAT)
- Follow up on DES transients
- On track to observe 5000 SN host galaxy redshift with spectra
- Monitoring 771 AGN out to z~4 for reverberation mapping
  - Expect to measure the mass of supermassive blackholes at the centre of approx. 30-40% of these
    - The largest and most distant sample ever measured!

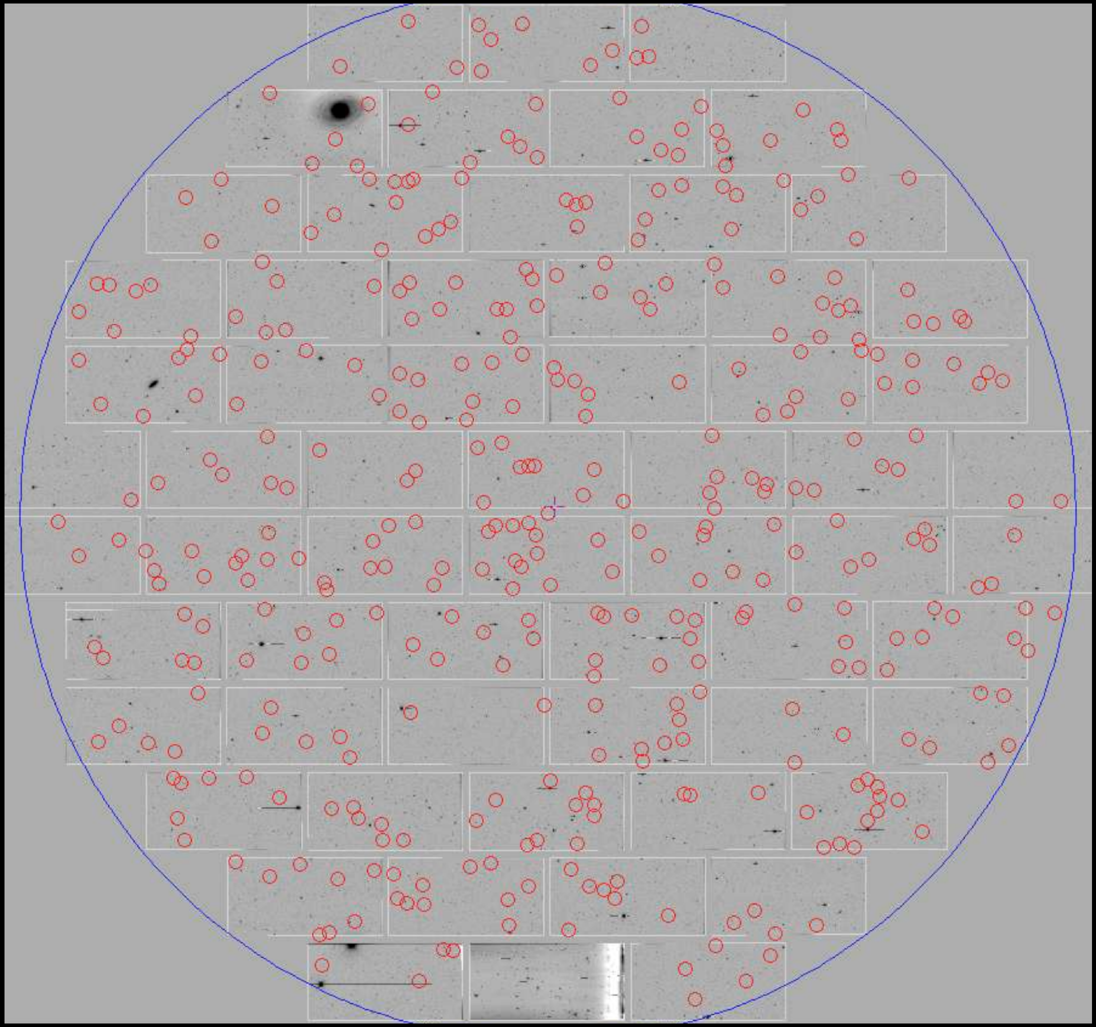


The Dark Energy Camera
(Blanco telescope, Chile)

The Dark Energy Camera
(Blanco telescope, Chile)

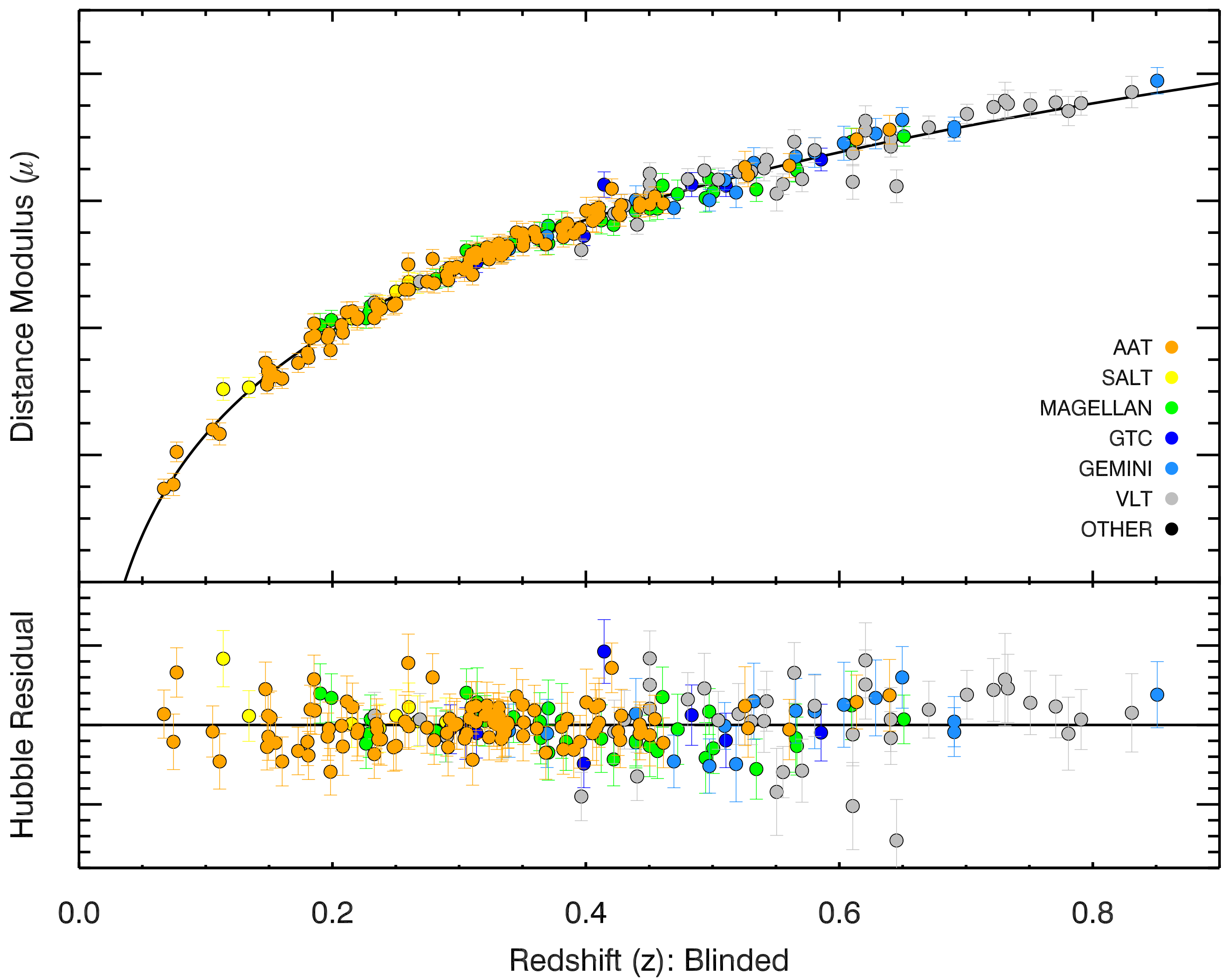
2 degree field spectrograph (Anglo-Australian Telescope)

superimposed on



The Dark Energy Camera (Blanco telescope, Chile)

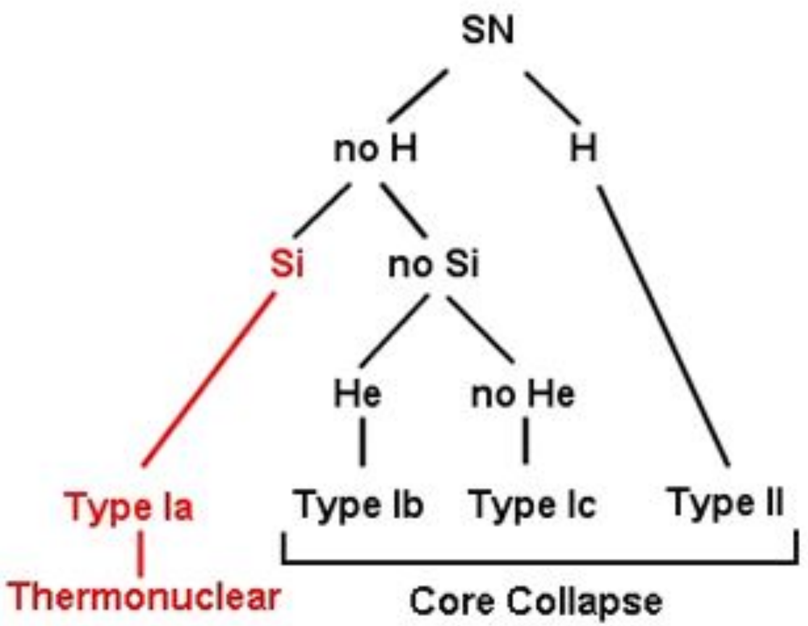
OzDES
Australian Dark Energy Survey



Why is this a difficult problem?

- Each Supernova needs to be classified in terms of its:
  - Type
  - Age
  - Redshift
- Light is intermixed with host-galaxy
- Variations/Distortions (no perfect template)
  - Interstellar Dust
  - Skylines
  - Noise
  - Dichroic jumps (equipment miscalibrations)

Supernova Types



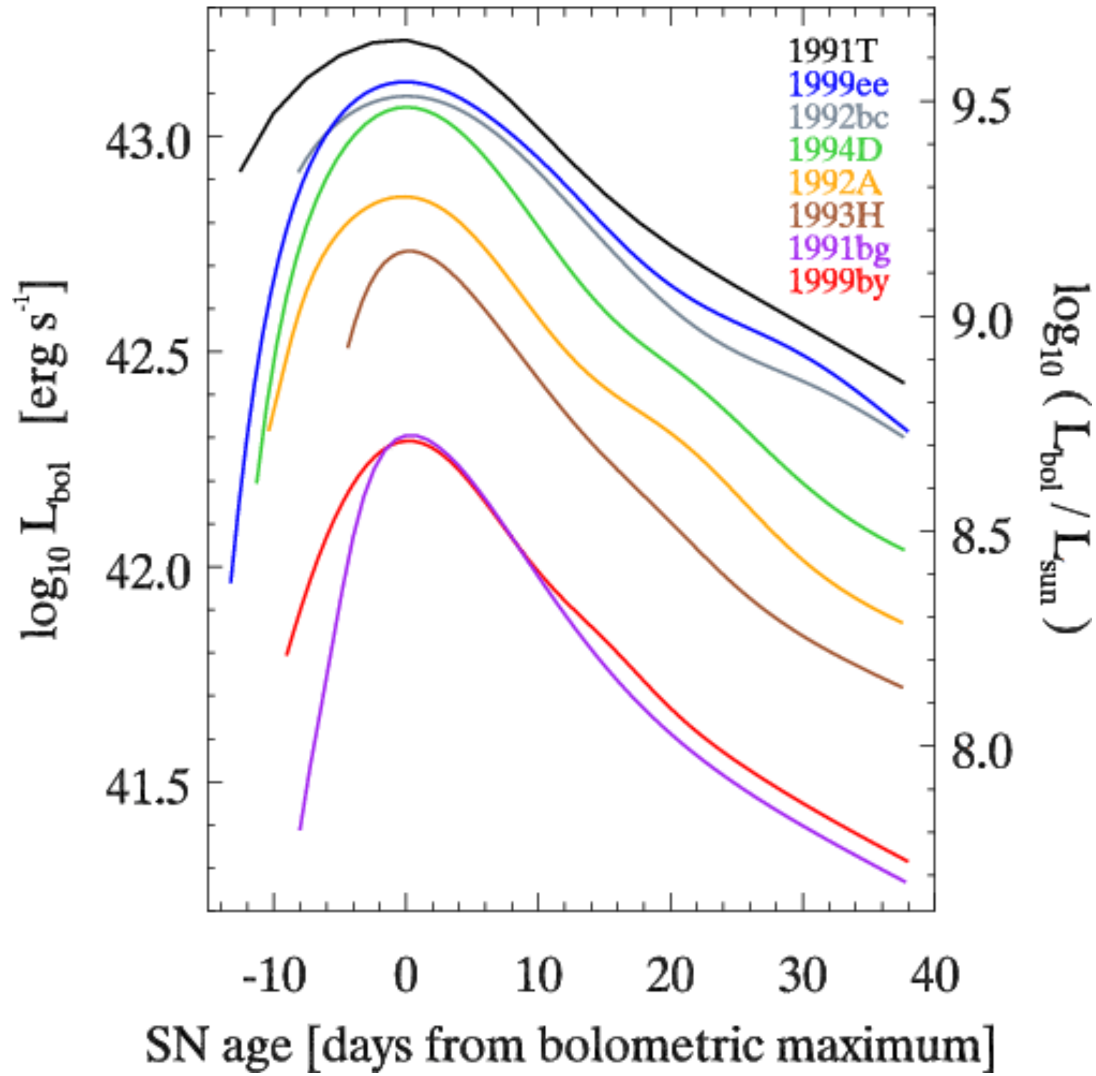
SNIa: Ia-norm, Ia-91T, Ia-91bg, Ia-02cx, Ia-csm, Ia-pec

SNIb: Ib-norm, Ibn, IIb, Ib-pec

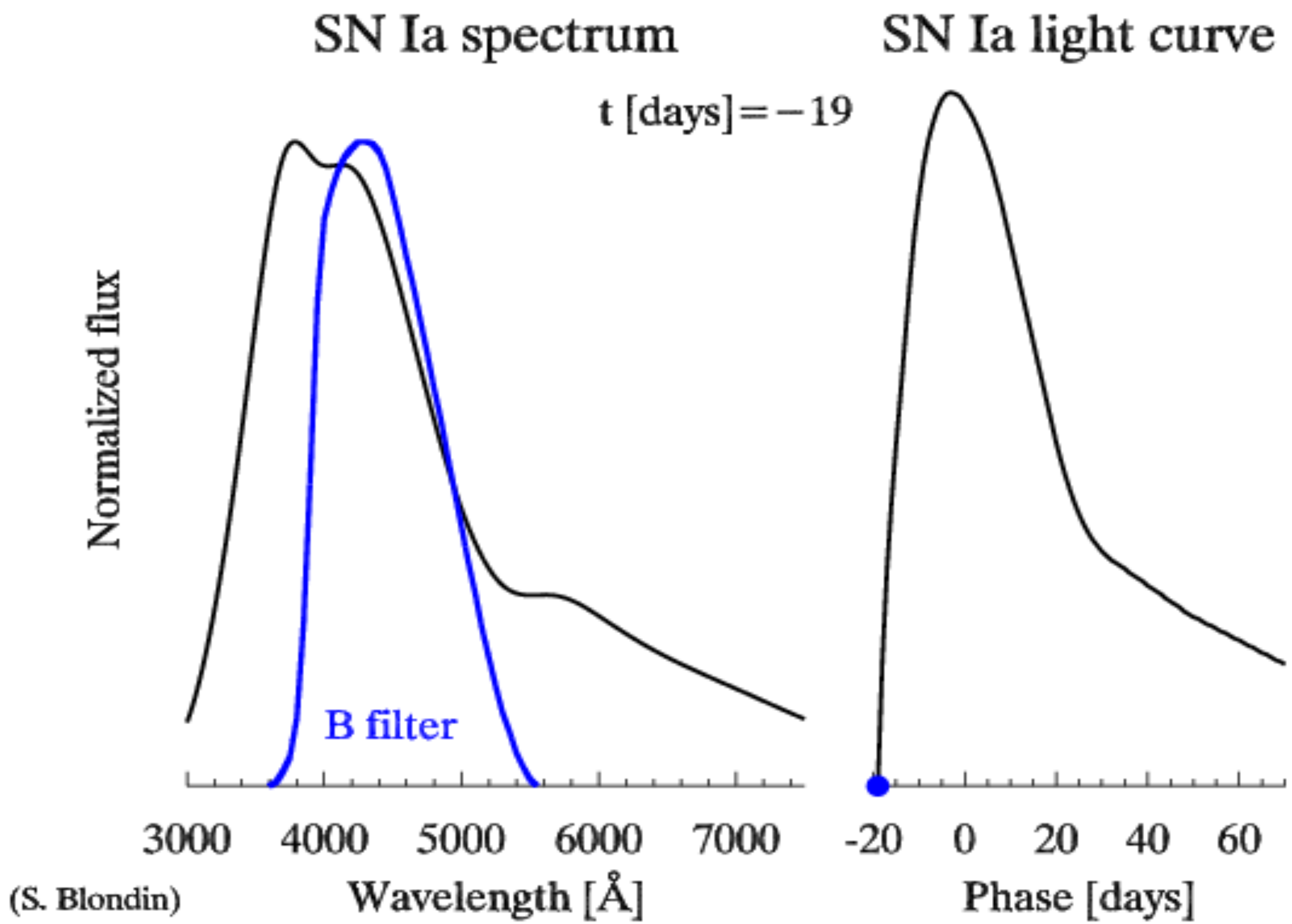
SNIc: Ic-norm, Ic-broad, Ic-pec

SNII: IIP, IIL, IIn, II-pec

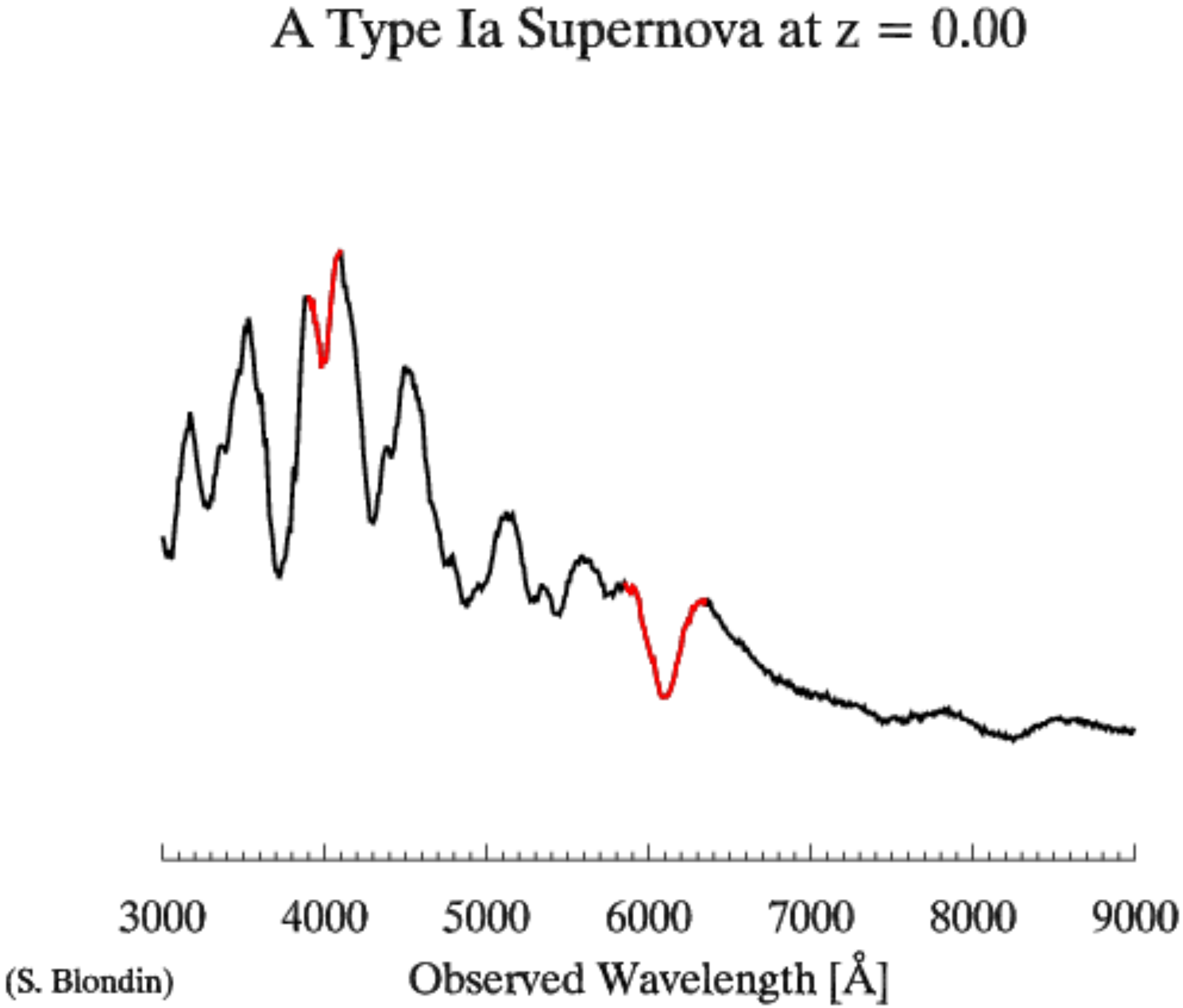
Bolometric Light-curves of SNIa



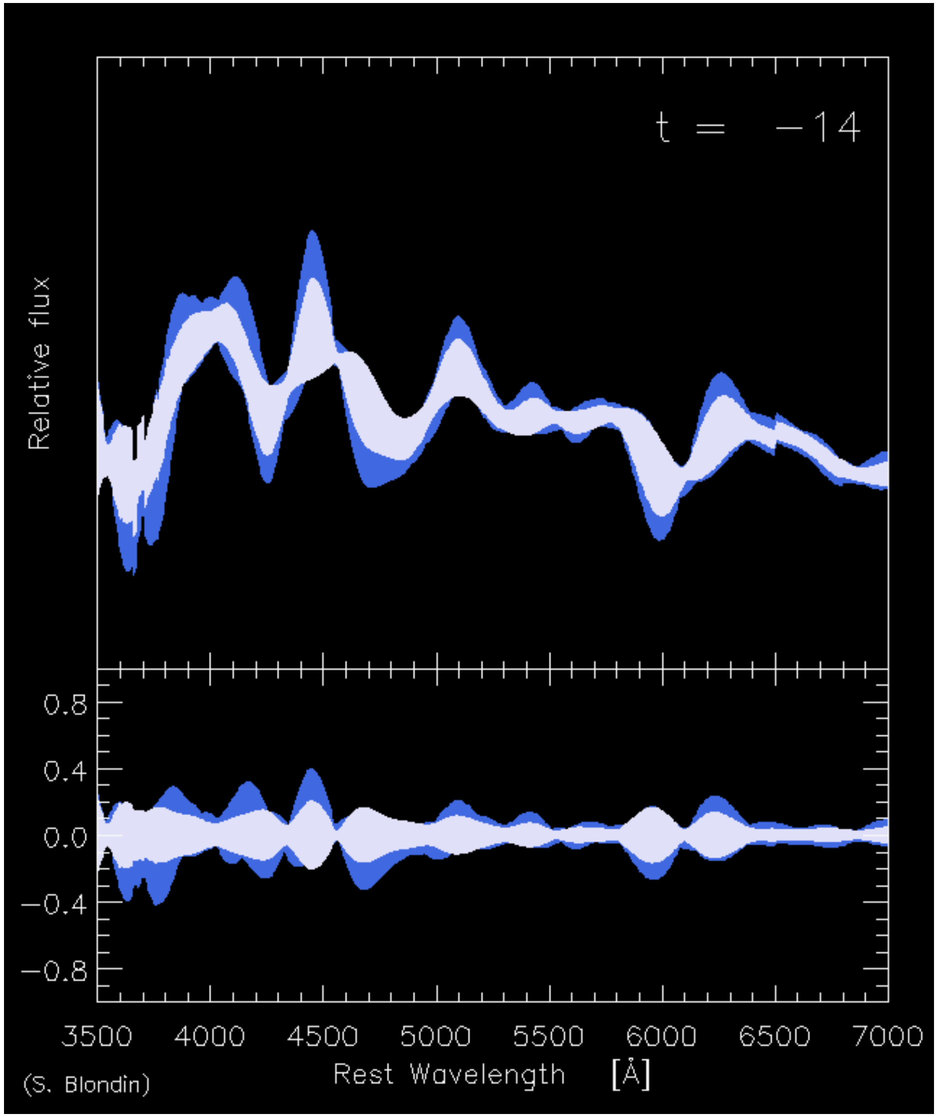
Why is this a difficult problem?
Age



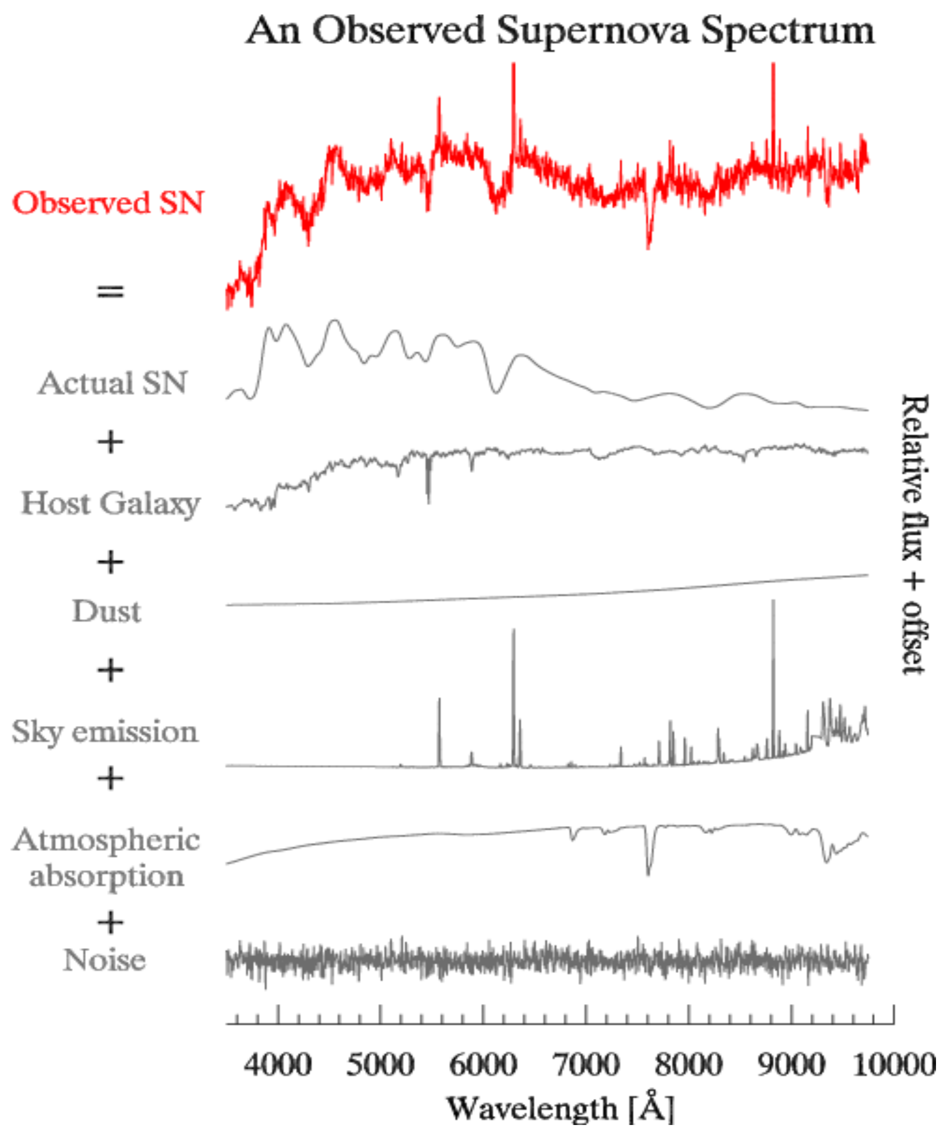
Why is this a difficult problem?
Redshift and Noise



Why is this a difficult problem?
Variations in data



What do we observe?



CAASTRO
ARC CENTRE OF EXCELLENCE FOR ALL-SKY ASTROPHYSICS

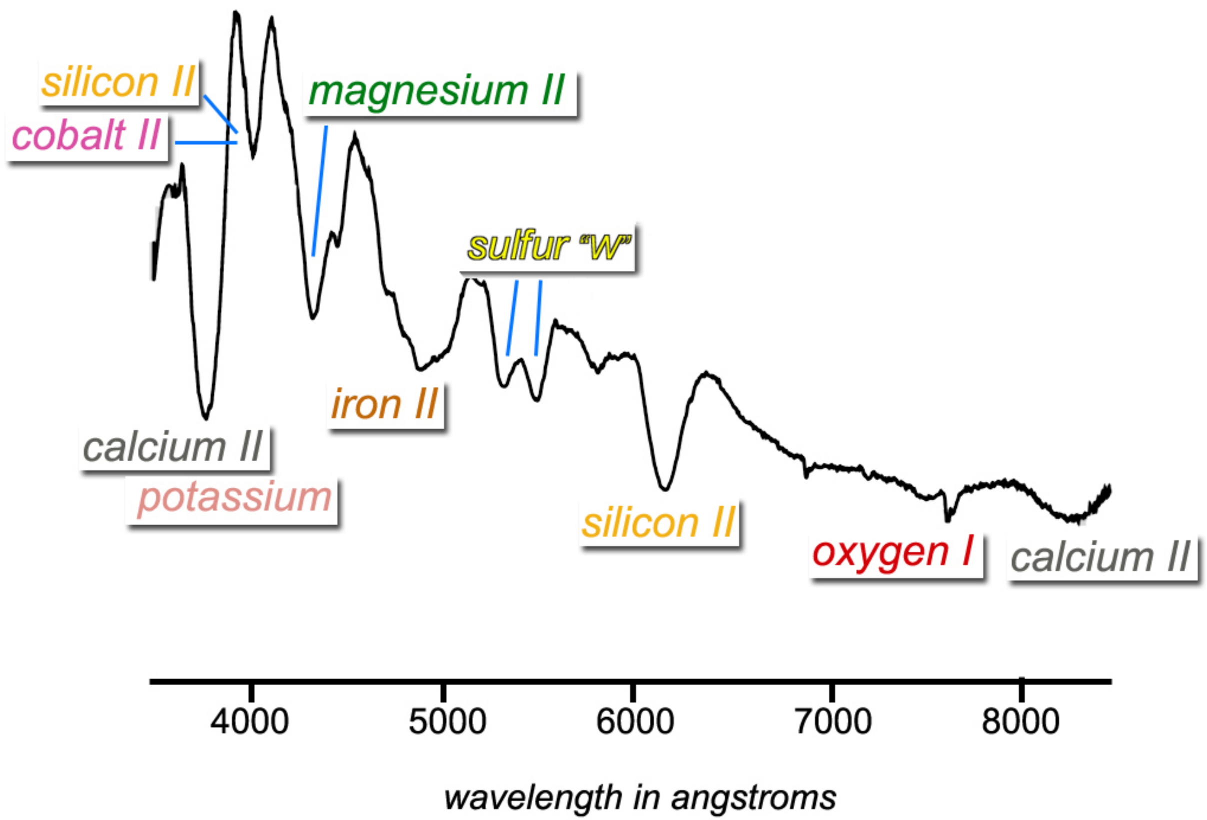
Previous classification methods

- Currently classification is slow, labour-intensive, and can take tens of minutes for a single supernova spectrum
- SNID – Stephane Blondin (Fortran)
  - Uses a cross-correlation of input with templates
  - Fast
  - Inaccurate with signals that are intermixed with host galaxy light
- Superfit – Andy Howell (IDL)
  - Uses a minimisation of chi-squared
  - Very slow, labour-intensive
  - Can deal with intermixed host galaxy light
- Good redshifting tools (from host galaxy)
  - MARZ (Samuel Hinton, University of Queensland), AUTOZ
  - Use a cross-correlation approach

Problems with current methods

- All rely on iterative template matching processes
  - Computation time increases linearly with the number of templates
  - Can only compare to one template at a time (rather than the aggregate set of each SN type)
- Chi-squared minimisations are slow
- Not autonomous: requires a lot of human-input

What type of SN is this?



Design Decisions
How DASH improves

- Speed
  - Autonomously classify several spectra at once
  - Significantly faster (example: 70 classified spectra in 18 seconds)
- Accuracy
  - DASH classifies based on *features* instead of templates
    - Uses aggregate set of templates rather than a single template
  - Softmax regression probabilities
- Precision
  - More specific classification including age and specific type
- Installation and ease of use
  - Graphical interface and python library
  - Very simple installation and use

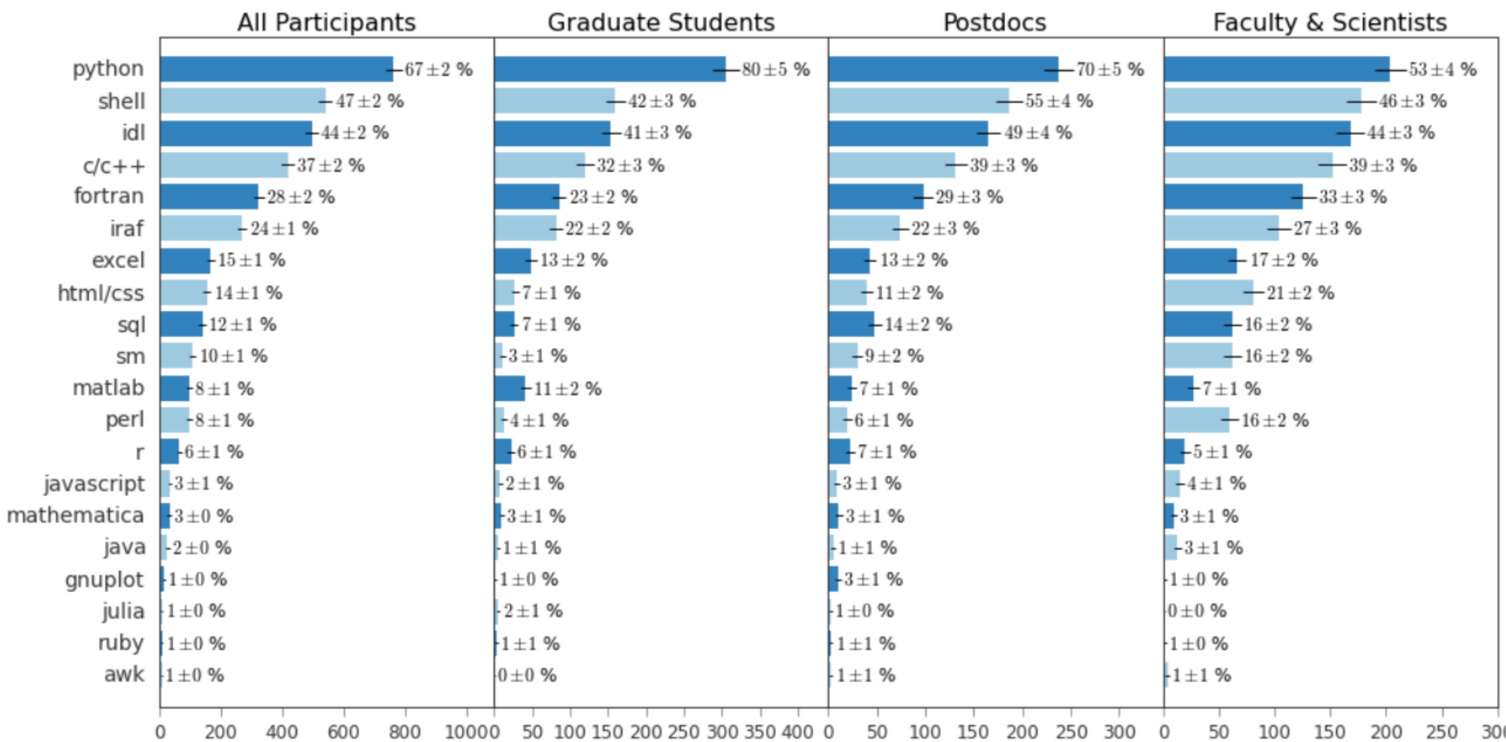
OzDES Design Requirements

- Speed and Accuracy
- Installation and Updating
  - Language decision
- Operating System
- Online vs Offline

CAASTRO
ARC CENTRE OF EXCELLENCE FOR ALL-SKY ASTROPHYSICS

Design Decisions

› Language decision: The rise of Python



Ivelina Momcheva and Erik Tollerud. “SOFTWARE USE IN ASTRONOMY: AN INFORMAL SURVEY”. In: (2015).

Design Decisions

› Online vs Offline interface

	Online Interface	Offline Interface
Installation	None	Required
Internet access	Required	Not needed
Integration	Not easy	Possible
Functionality	Fixed	More options

› Two interfaces

- Graphical interface and python library

Data Description

- Datasets:
  - SNID Templates 2.0 (Blondin, Liu, Modjaz 2016)
  - BSNIP – Berkeley Supernovae Ia Program (Silverman et al. 2012)
- 3936 spectra across 515 different supernovae
- Separated into 17 different subtypes

SNIa: Ia-norm, Ia-91T, Ia-91bg, Ia-02cx, Ia-csm, Ia-pec

SNIb: Ib-norm, Ibn, IIb, Ib-pec

SNIc: Ic-norm, Ic-broad, Ic-pec

SNII: IIP, IIL, IIn, II-pec

What about weirdos??

- Super-luminous Supernovae (SLSN)?
- New unknown transients?

Problems with Datasets

› Bias due to gaps in the template types
  - Few older non-Ia Supernovae
› Need for more templates
› WISeREP
  - Huge collection (13000 SN spectra) but no ages…

CAASTRO
ARC CENTRE OF EXCELLENCE FOR ALL-SKY ASTROPHYSICS

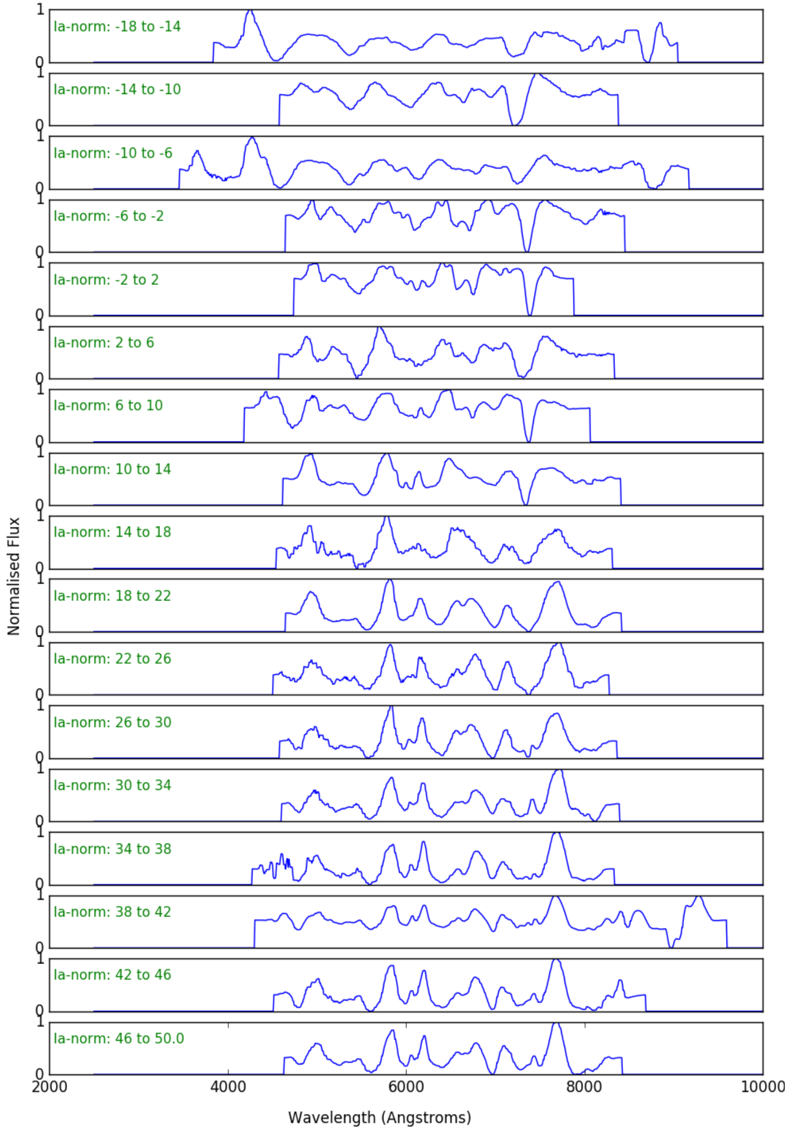
Template Sample Bias

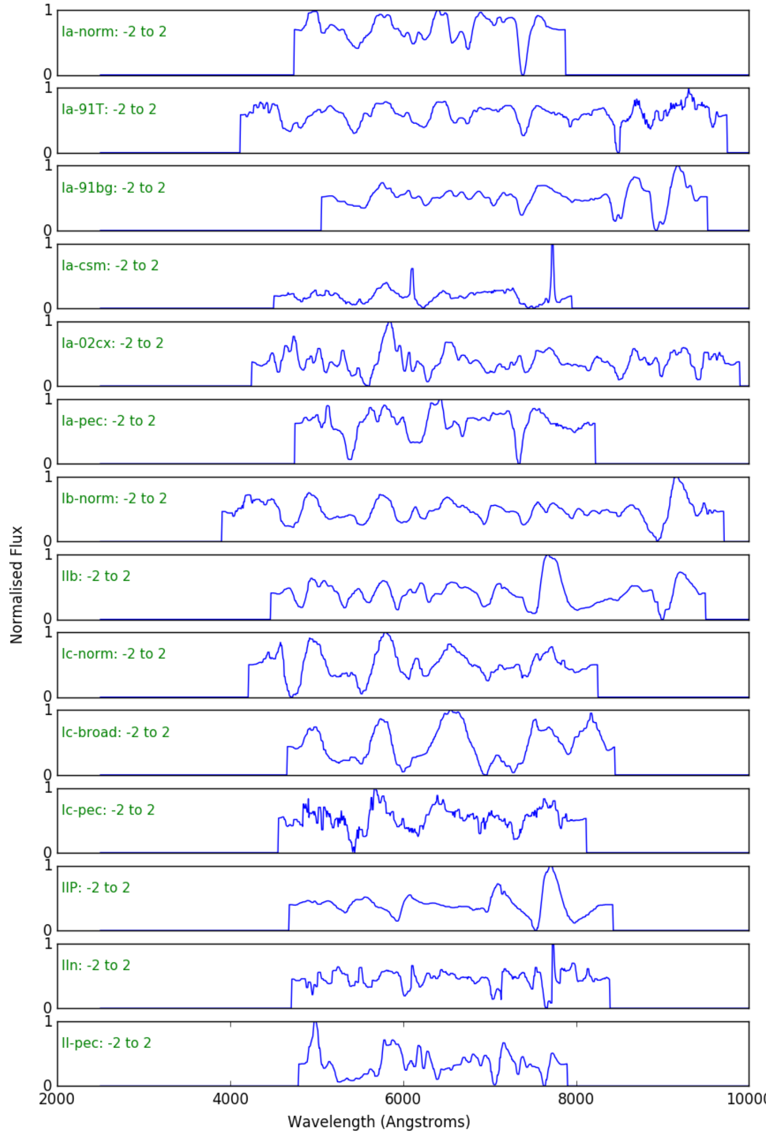
- 306 classification bins
  - 17 subtypes
  - 18 age bins
- Three common methods to deal with bias
  - Undersampling
  - Oversampling
  - Adding weights

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
Ia-norm	0	2	54	179	231	274	286	229	191	158	121	96	91	79	52	60	54	29
Ia-91T	0	0	18	51	54	32	19	25	20	27	26	21	19	12	13	9	11	15
Ia-91bg	0	0	0	10	24	25	39	23	24	17	21	10	10	13	2	2	6	5
Ia-csm	0	0	2	0	0	2	2	4	2	2	0	0	2	2	0	6	4	2
Ia-02cx	0	0	0	8	11	2	1	0	2	3	3	5	2	0	1	1	2	0
Ia-pec	0	0	5	15	18	20	13	17	10	9	8	12	8	8	7	7	4	1
Ib-norm	1	4	8	13	18	17	8	10	15	11	7	4	4	4	5	5	2	6
Ibn	0	0	0	0	0	0	3	6	3	2	3	1	1	3	1	1	2	1
IIb	4	12	15	7	13	6	13	12	15	11	8	5	3	4	7	6	3	2
Ib-pec	0	0	2	1	2	0	0	0	1	4	1	0	0	0	0	0	1	0
Ic-norm	0	1	1	11	18	18	15	9	7	12	9	12	5	10	3	4	3	8
Ic-broad	0	1	7	6	21	16	18	17	13	10	13	10	13	6	5	3	3	11
Ic-pec	0	0	0	0	3	9	7	0	1	0	4	3	0	2	0	0	0	2
IIP	0	0	0	1	12	30	23	22	12	10	4	11	10	5	13	5	5	3
IIL	0	0	0	0	0	0	0	3	4	0	0	0	0	0	0	1	2	0
IIn	0	0	4	0	0	8	2	4	6	2	0	0	0	6	0	4	6	2
II-pec	1	3	2	3	3	1	2	2	0	0	0	0	0	0	0	0	0	0

A: -20 to -18 days	G: -2 to 2 days	L: 18 to 22 days	Q: 38 to 42 days
B: -18 to -14 days	H: 2 to 6 days	M: 22 to 26 days	R: 42 to 46 days
D: -14 to -10 days	I: 6 to 10 days	N: 26 to 30 days	S: 46 to 50 days
E: -10 to -6 days	J: 10 to 14 days	O: 30 to 34 days	
F: -6 to -2 days	K: 14 to 18 days	P: 34 to 38 days	

Ages and types





Why Deep Learning?

- Deep Learning has had success in a range of new Big Data problems:
  - Image, speech, language recognition. Beating grandmasters at Chess and Go
- Accuracy improves with number of template (does not affect computation time)
- Training process is separate to testing
- Only need to train once. Then only need the trained model instead of the entire template set.
- Train based on the aggregate set of all templates in a particular SN bin
- Disadvantages
  - Deep learning is often position invariant, which makes redshifting difficult.
  - Softmax probabilities are relative, not absolute measures

CAASTRO ARC CENTRE OF EXCELLENCE FOR ALL-SKY ASTROPHYSICS

Why Deep Learning?

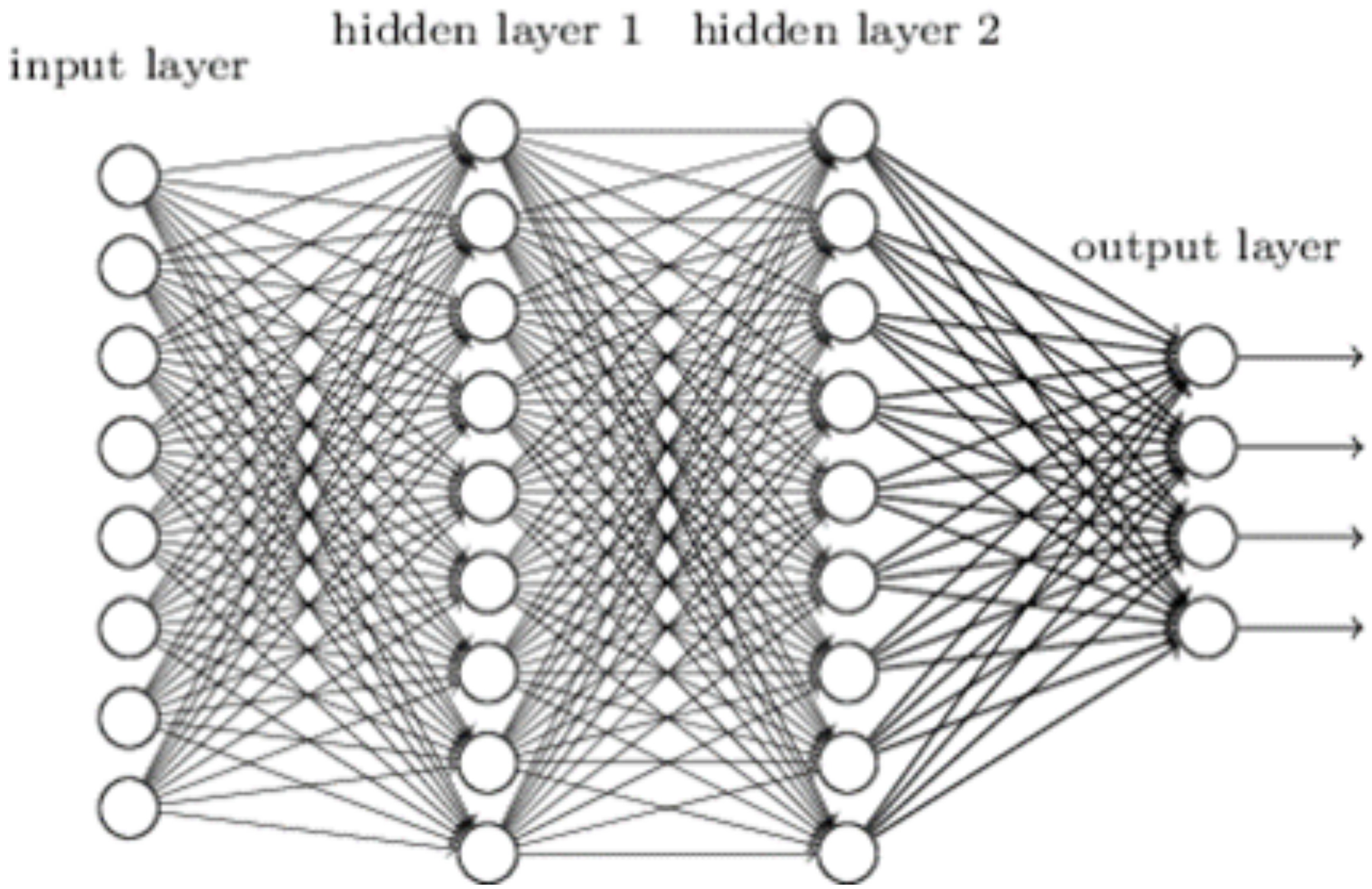
	Deep Learning	Cross-correlation matching	Chi-squared matching
Classification technique	Matches based on the combined 'features' of all templates	Iteratively compares to templates	Iteratively compares to templates
Speed	Very Fast (no change in speed with templates)	Fast (but increases linearly with number of templates)	Slow (increases linearly with number of templates)
Noise	Can train with noise	Cannot classify low S/N	OK with low S/N
Redshifting	Redshifting is unreliable	Very good at redshifting	OK redshifting
Goodness of Fit	Relative	Absolute	Absolute

Why Tensorflow?

- Highly efficient C++ backend to do its computation.
- Tensorflow has been used in several major companies
- Only recently released to the public
- Makes neural network design simple!
  - High level library that avoids low-level details
  - Flexible architecture
  - Very fast performance
- GoogleBrain is investing a lot into making this the best CNN library

CAASTRO
ARC CENTRE OF EXCELLENCE FOR ALL-SKY ASTROPHYSICS

Deep Neural Network

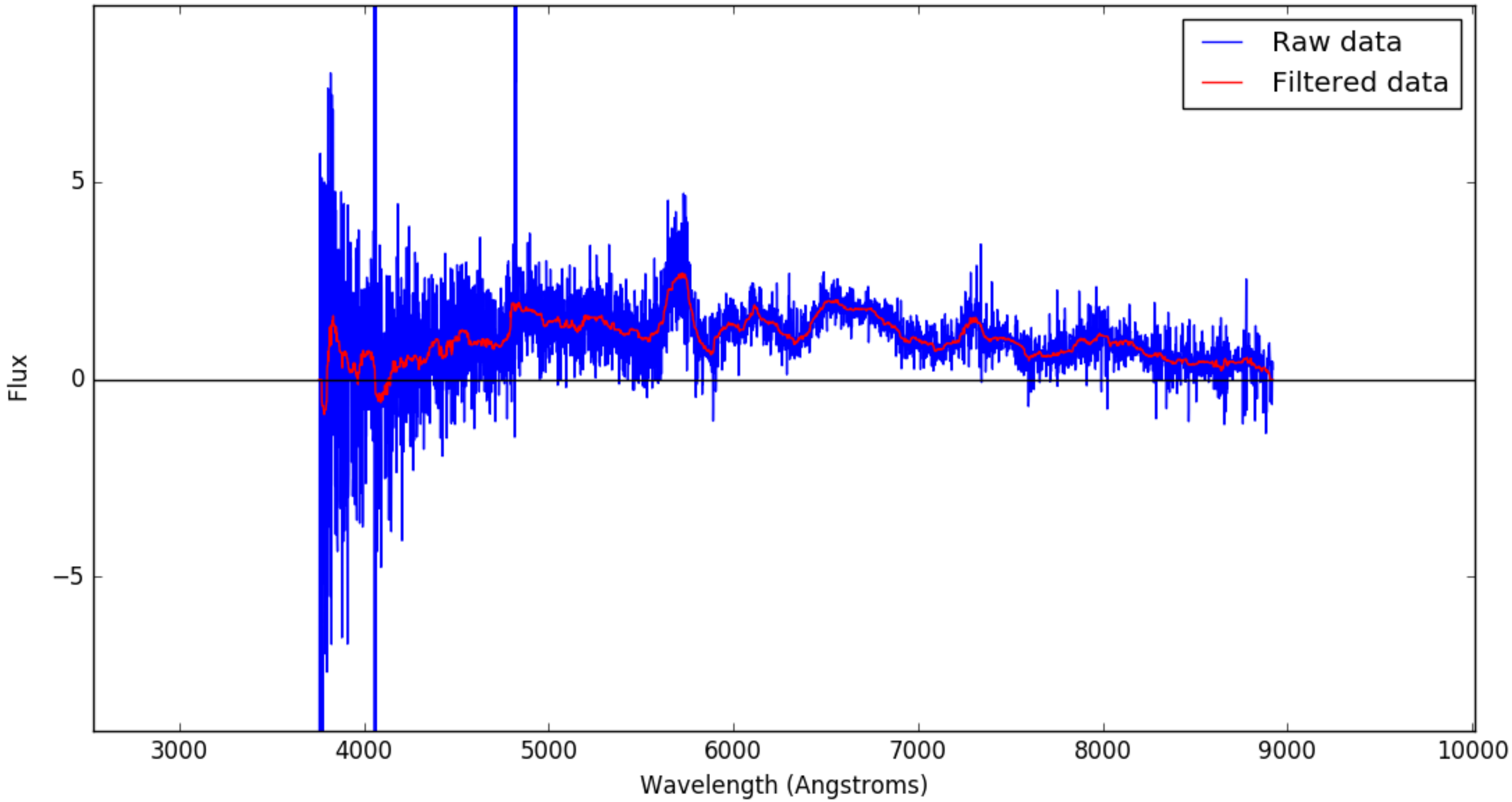


Important that the problem is well-defined, and to process the spectra in a uniform way

Pre-processing spectra

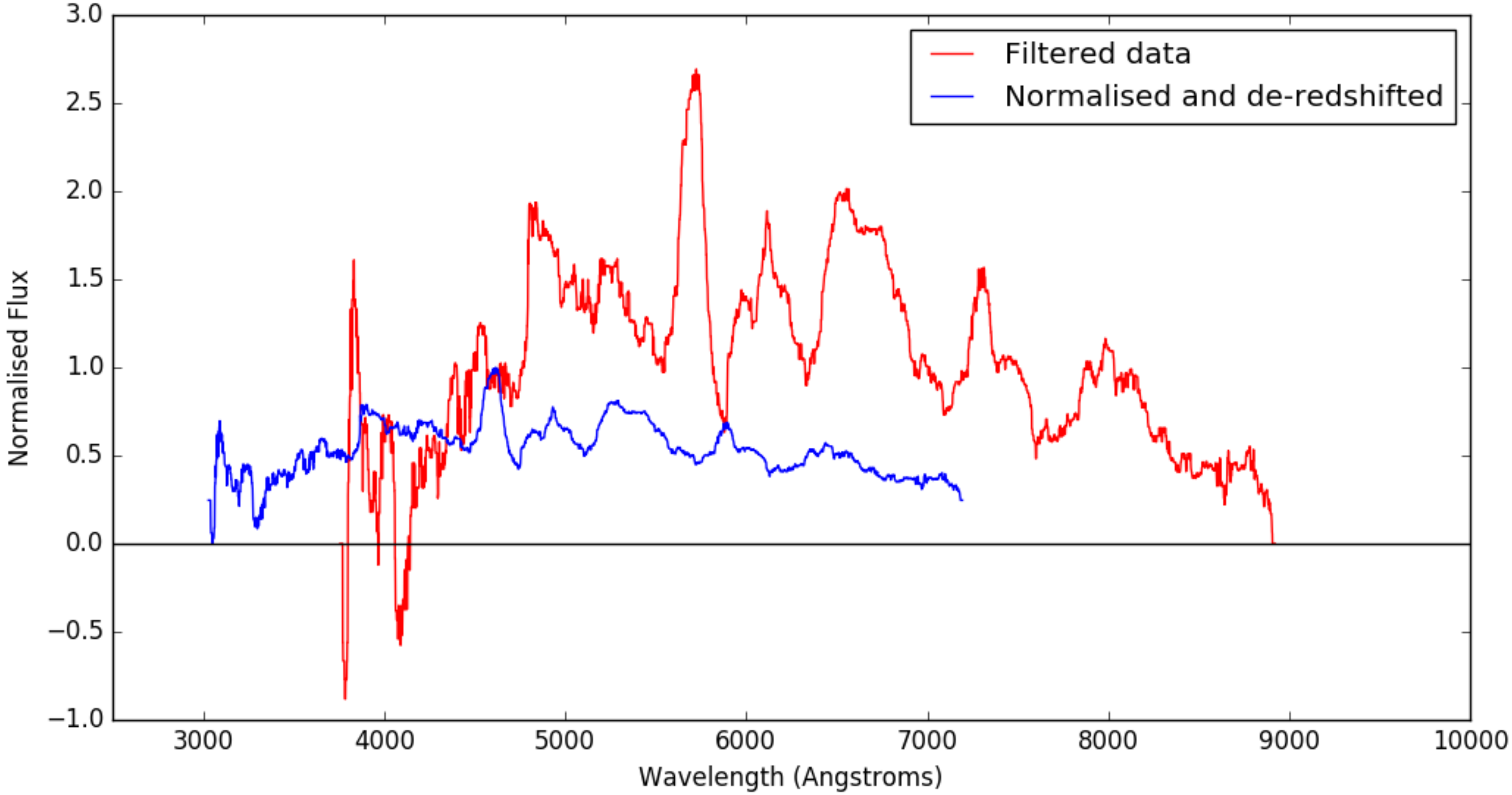
- Import to prepare signal in standard uniform way
- I treated the problem like a one-dimensional image classification problem
- Steps to processing:
  1. Low pass median filtering
  2. Normalising
  3. De-redshfting
  4. Log-wavelength binning
  5. Continuum modelling with spline interpolation
  6. Continuum subtraction

Low-pass filtering



Normalising and De-redshifting

$$\lambda_{emitted} = \frac{\lambda_{observed}}{z+1}$$



Log-wavelength binning

- Binned to 1024 points between 2500 and 10000 Angstroms
  - Enough to not lose information, but not too much for memory problems (Tonry & Davis 1979)

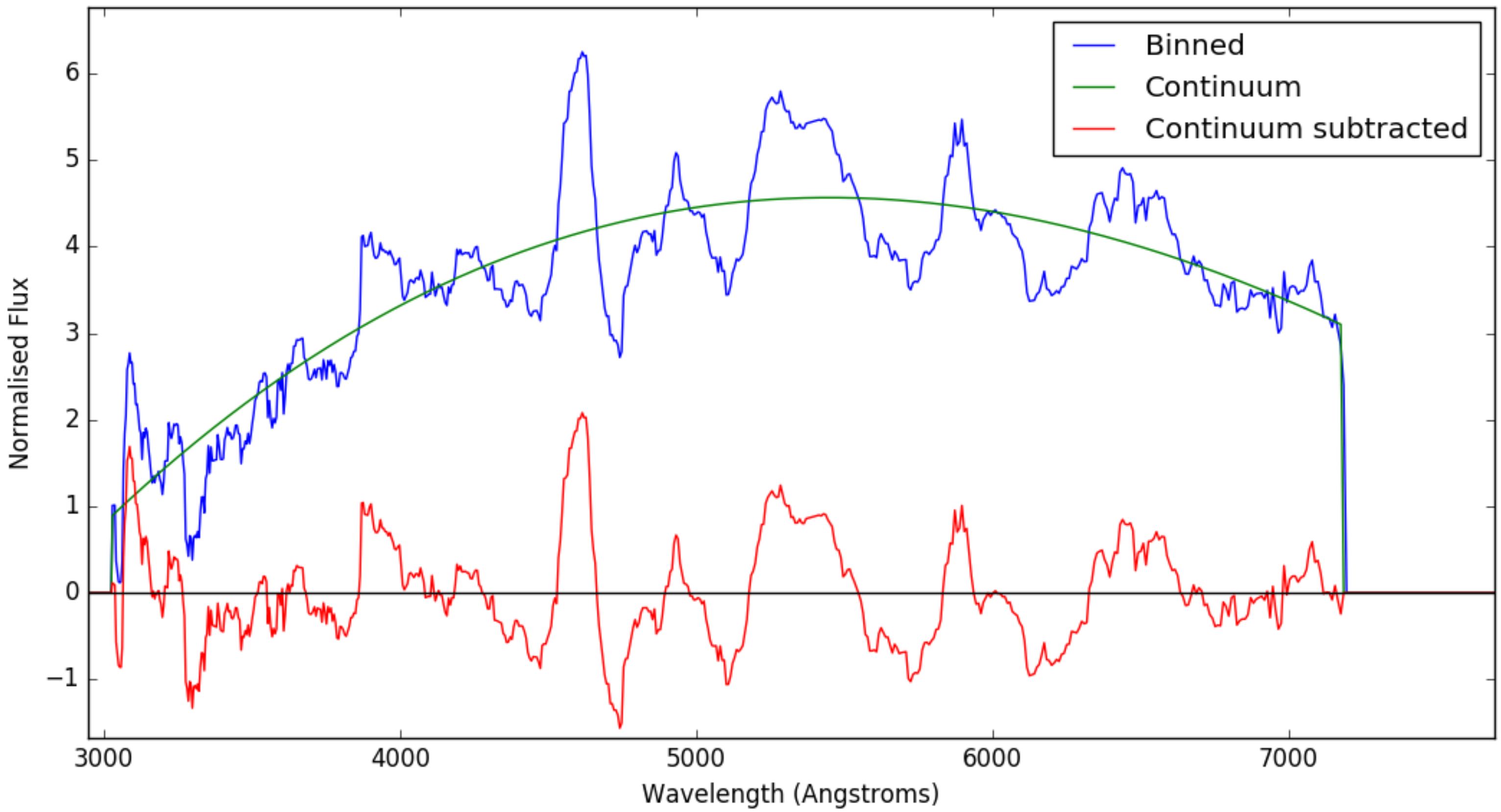
$$l_{\log,n} = l_0 \ln e^{n \times dl_{\log}}$$

$$dl_{\log} = \ln(l_1/l_0)/N$$

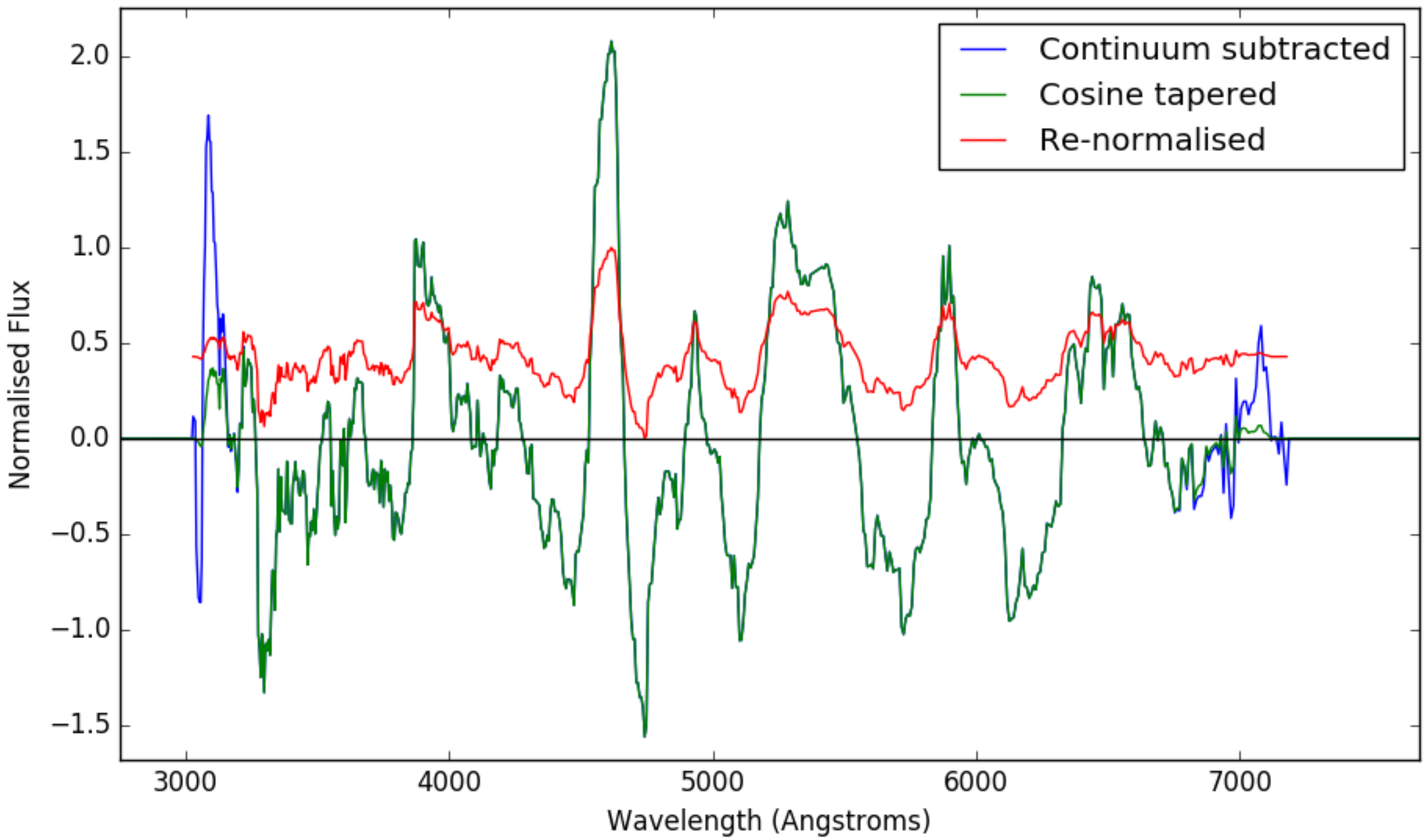
- Consistent with template format from SNID
- Can make redshifting easier
  - Multiplying by $(1+z)$ is equivalent to adding $\ln(1+z)$

Continuum Removal

- Continuum modelled with 13-point cubic spline interpolation
- Continuum then subtracted – removes spectral colour information (including flux miscalibrations)



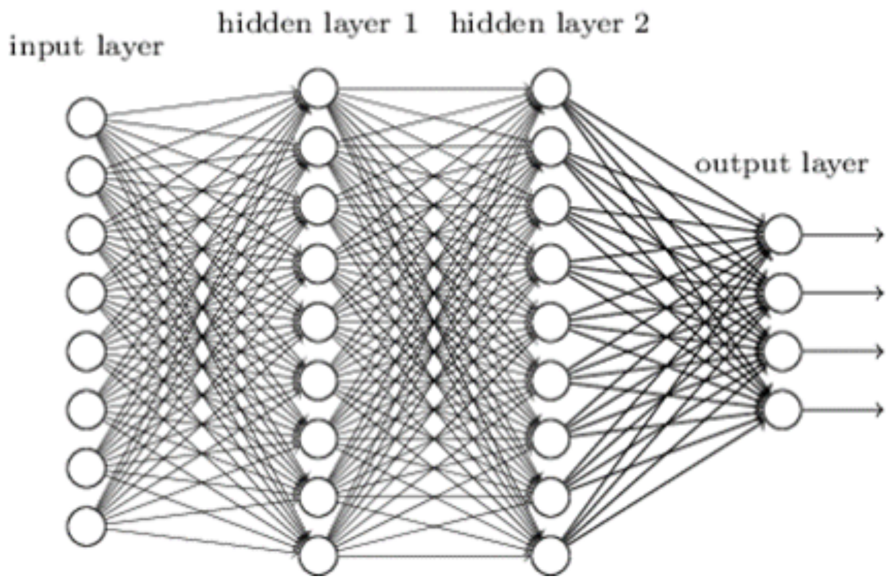
Apodising edges



Convolutional Neural Network

- Two layer neural network with tensorflow
  - 2 layers was over 10-30% better than a single layer
  - 3 layers gave no significant improvement

- 3936 spectra across 515 different supernovae
  - Training set – 80%
  - Validation set – 20%



Tensorflow training

› Label and Image Data
  - Label: one hot vector representing one of the 306 classification bins
  - Image data: 1024-point binned vector representing the spectrum flux

› Weights and Biases
  - Weights include a small amount of noise
  - Slightly positive initial bias to avoid “dead neurons”

Tensorflow Training

› Softmax Regression

$$\text{evidence}_\text{i} = \sum_{i=1}^{1024} W_{i,j} x_j + b_i$$

$$y = \text{Softmax}(\text{evidence})$$ where $$\text{Softmax}(x)_i = \frac{e^{x_i}}{\sum_j e^{x_j}}$$

› Cross-entropy:

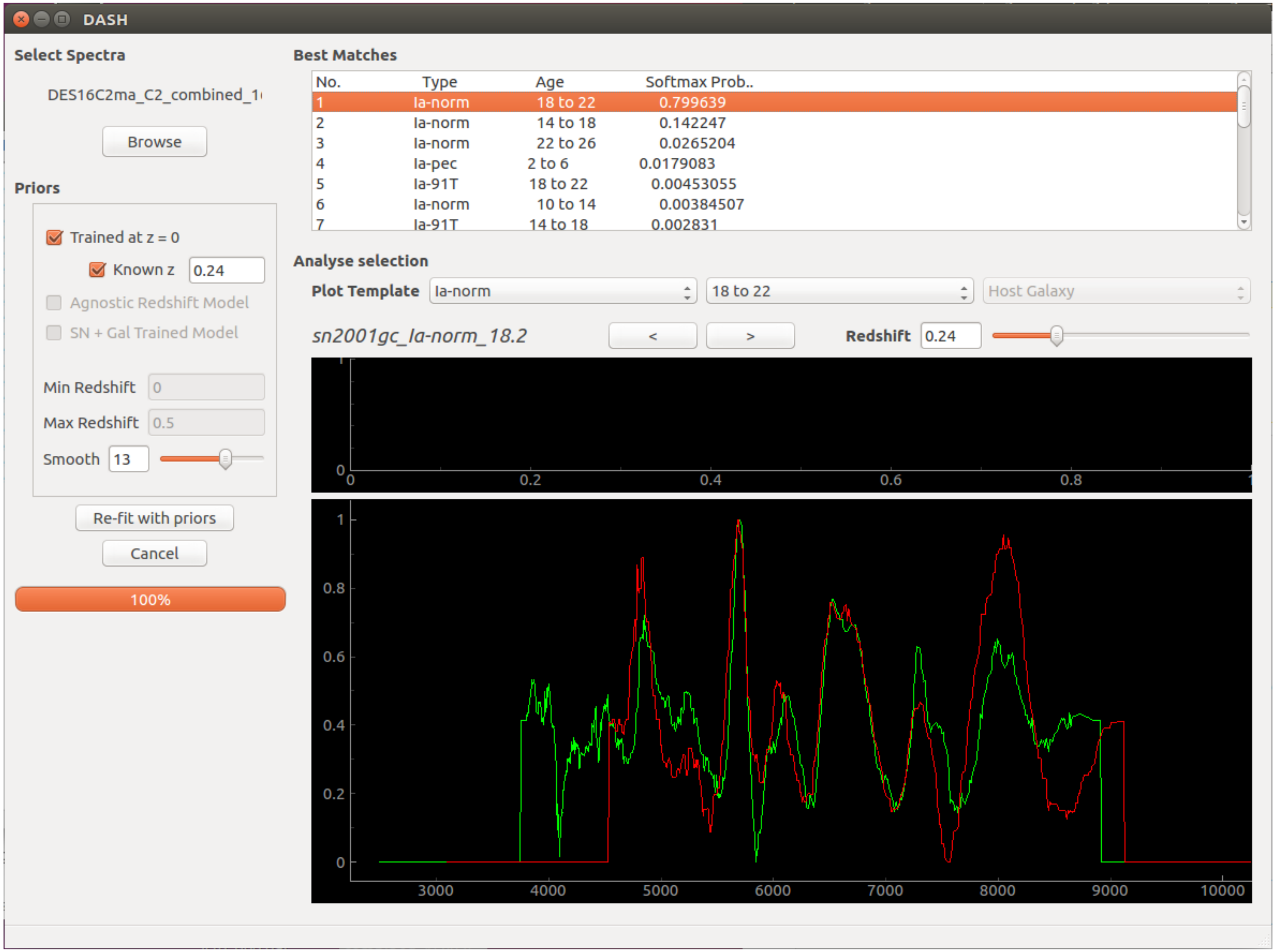
$$H'_y(y) = -\sum_{i=1}^{306} y'_i \log(y_i)$$

- Minimised with ADAM optimiser

› Building the layers

- Convolution and Pooling
- Rectified linear units (ReLU)
- Dropout layer

Graphical Interface



Python Interface

```
pip install astrodash
```

- Python 2/3
- Operating Systems: Unix/Mac/Windows

```
import dash
   classify = dash.Classify([filenames], [knownRedshifts])
   print classify.list_best_matches()
   classify.plot_with_gui(indexToPlot=0)
```

- Available open source at:

  https://github.com/daniel-muthukrishna/DASH

Validation Set Performance

- **Type:** Correct broad type (i.e. Ia, Ib, Ic, II) identified by the matching algorithm.
- **Subtype:** Correct subtype (i.e. Ia-norm, Ib-pec, Ib-norm, etc.) identified.
- **Type and Age:** Correct broad type and the correct age bin identified by the matching algorithm.
- **Subtype and Age:** Correct subtype and the correct age bin identified.

Criteria	Correctly Classified
Type	99.87%
Subtype	95.46%
Type and Age	84.24%
Subtype and Age	88.15%

CAASTRO
ARC CENTRE OF EXCELLENCE FOR ALL-SKY ASTROPHYSICS

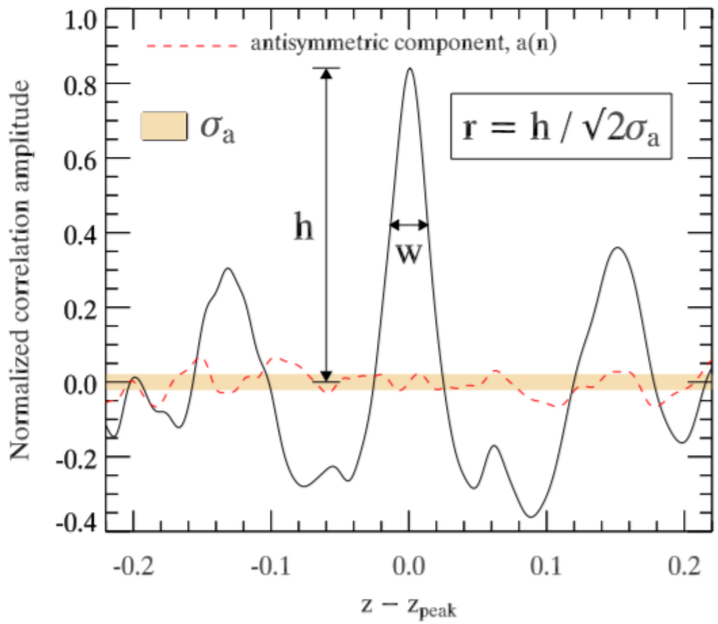
Results with OzDES Data

- OzDES data from the last couple of runs at the end of 2016.
- Matches Superfit in 100% of confirmed cases
- Classified all 23 spectra in <10 seconds!
- Able to classify more spectra
  - Precise likelihood measurements (from softmax regression)
  - More precise measurement (with age and specific type)

Name	Redshift	ATEL Classification	DASH Classification	DASH Probability	Match?
DES16E1de	0.292	Ia? (+2)	Ia-pec (+2 to +10)	91%	✓
DES16E2dd	0.0746	Ia (+3)	Ia-norm (+2 to +6)	89%	✓
DES16X3km	0.0542	II (+)	IIP (+6 to +10)	99.7%	✓
DES16X3er	0.167	Ia (+2)	Ia-91T (-2 to +6)	86%	✓
DES16X3hj	0.308	Ia (0)	Ia-norm (-2 to +2)	90%	✓
DES16X3es	0.554	Ia? (0)	IIP (+22 to +26)	92%	x
DES16X3jj	0.238	II? (+)	Ic-pec (-2 to 2)	37%	x
DES16C3fv	0.322	Ia (-6)	Ia-norm (-10 to +2)	99.8%	✓
DES16C3bq	0.241	Ia (+0)	Ia-norm (-2 to +6)	99.6%	✓
DES16E1md	0.178	Ia (0)	Ia-norm (-6 to +2)	99%	✓
DES16E1ah	0.149	II (+)	Ia-91T (+14 to +22)	75%	x
DES16C3ea	0.217	Ia (+)	Ia-norm (+10 to +26)	88%	✓
DES16X1ey	0.076	II (+)	IIb (+2 to +6)	38%	✓
DES16C3bq	0.237	Ia (+)	Ia-norm (-2 to +6)	97%	✓
DES16E2aoh	0.403	Ia (+)	Ia-norm (-2 to +6)	88%	✓
DES16X3aqd	0.033	IIP (+)	IIb (-6 to +2)	99%	✓
DES16X3biz	0.24	Ia (-)	Ia-norm (-14 to +2)	98%	✓
DES16C2aiy	0.182	Ia (+)	Ia-norm (-2 to +6)	99.99%	✓
DES16C2ma	0.24	Ia (+)	Ia-norm (+14 to +22)	99.2%	✓
DES16X1ge	0.25	Ia (+)	Ia-norm (+14 to +22)	99.7%	✓
DES16X2auj	0.144	Ia (0)	Ia-norm (-6 to +6)	84%	✓
DES16E2bkg	0.478	Ia (0)	Ia-norm (-2 to +6)	99%	✓
DES16E2bht	0.392	Ia (+3)	Ia-norm (-6 to +2)	58%	✓

False Positive Rejection

- Relative Softmax probabilities (not absolute)
- Cross-correlation with best fit templates
- Ensure that rlap > 5
  - rlap is defined by Stephane Blondin in SNID
  - r is related to the cross-correlation maximum height
  - lap is the overlap between the input and template spectra



Blondin & Tonry (2007)

Evaluation

- Speed
  - > 100 times faster
  - Autonomously classify multiple spectra at once with *Tensorflow session*
- Precision
  - More specific classification including age and specific type (rather than just the subtype)
- Accuracy
  - Softmax regression probabilities (at least as accurate as Superfit)
  - Classifies on aggregate set of templates rather than a single template
- Installation and ease of use
  - Graphical and pythonic interface
- Currently being tested by OzDES for implementation in the coming months

Future Improvements

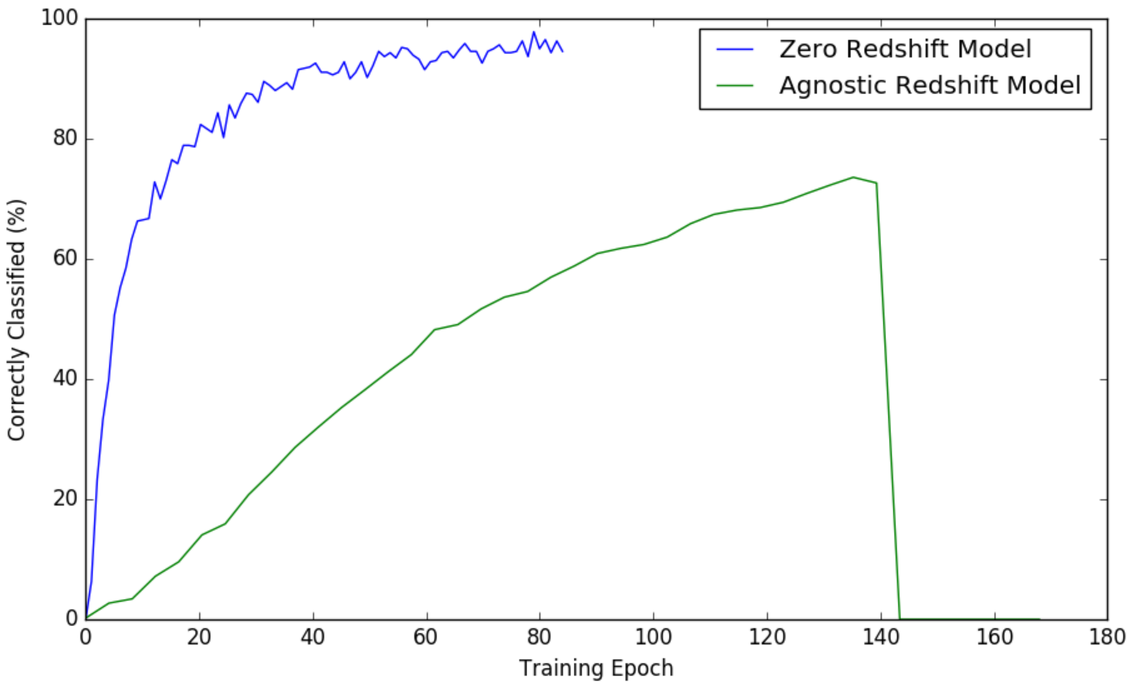
- Identification of Host Galaxy
  - Extra dimension of classification bins
  - $\text{Combined} = \alpha(\text{SN}) + \beta(\text{Host})$

- Redshifting (currently from host galaxy using MARZ)
  - Agnostic Redshift Model

- MARZ already does both redshifting and host galaxy classification (by Samuel Hinton, University of Queensland)
  - http://samreay.github.io/Marz/

Alternative Models

- Zero-redshift model
  - Trained on templates at redshift zero (requires a known redshift)
- Agnostic Redshift model
  - Trained on templates at a range of redshifts



Conclusions

- Supernovae are the most powerful probe for probing the nature of dark energy
- DASH makes use of a 2 layer neural network with Tensorflow
- Over 100 times faster and more precise than previous methods because it classifies based on aggregate features instead of individual templates
- Two interfaces:
  - Graphical interface, Python library
- Easy installation
  - pip install astrodash
- Open Source
  - https://github.com/daniel-muthukrishna/DASH

Acknowledgements

- Thanks to
  - David Parkinson
  - Brad Tucker
  - Samuel Hinton
  - Tamara Davis
  - Chris Lidman
  - Stephane Blondin
- Obelix Supercomputer, University of Queensland