Deep Learning Based Distance Estimation of Galaxies

Anjitha John William

Supervisors: **Prof. Maciej Bilicki, Prof. Wojciech Hellwing**

In collaboration with

Dr. Priyanka Jalan,

Center For Theoretical Physics, PAN, Warsaw

Hareesh Thuruthipilly

National Center for Nuclear Research, Warsaw

ML in PL
Warsaw (26-29 October)

Distance to Galaxies

- Hubble-Lemaitre Law

$$v = H_0 d$$

v - recessional velocity
H_0 - Hubble's constant
d - proper distance

$$z = \frac{v}{c}$$

c - speed of light

- Distance measured by using **Redshift**

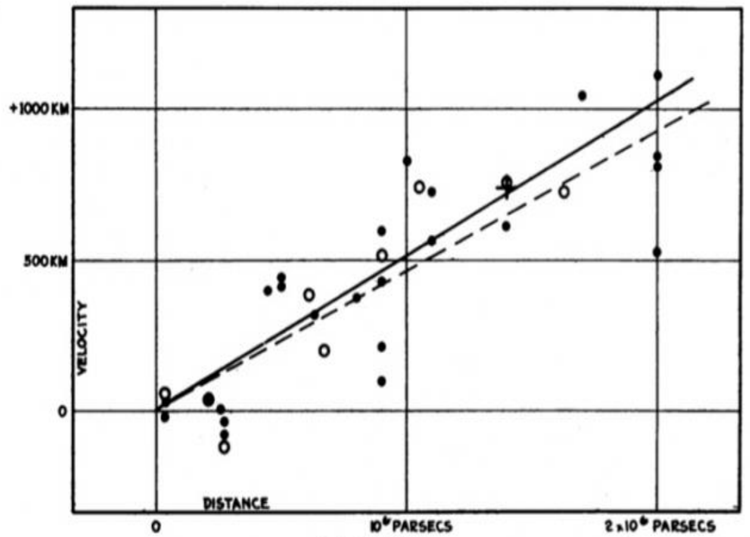


FIGURE 1
Velocity-Distance Relation among Extra-Galactic Nebulae.

Credit: Edwin Hubble, Proceedings of the National Academy of Sciences, vol. 15 no. 3, pp.168-173

Cosmological Expansion and Redshift

- Proper physical distance between a pair of well-separated events is increasing with time.
- The light from galaxies shifted to longer wavelengths - **Redshift (z)**

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

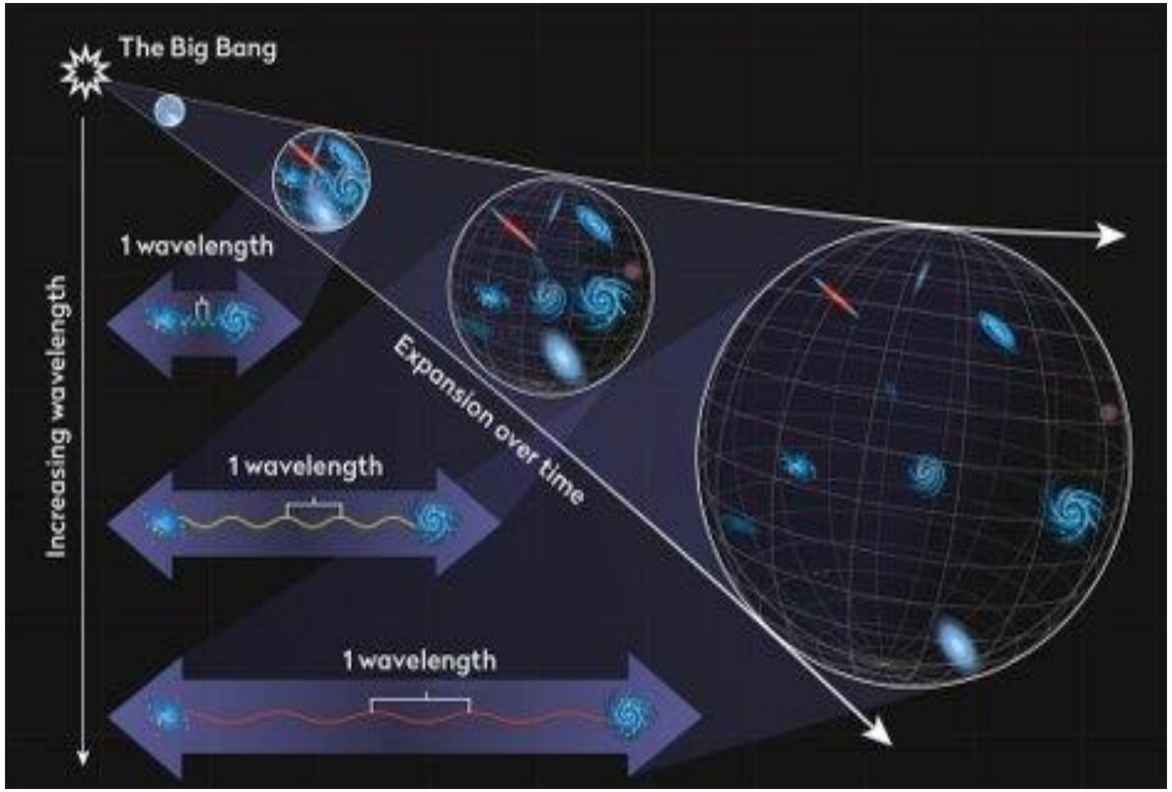


Image credit: https://www.skyatnightmagazine.com/

How to measure Redshift?

Spectroscopic way

- Measuring the shift in spectral lines
- Spec-z

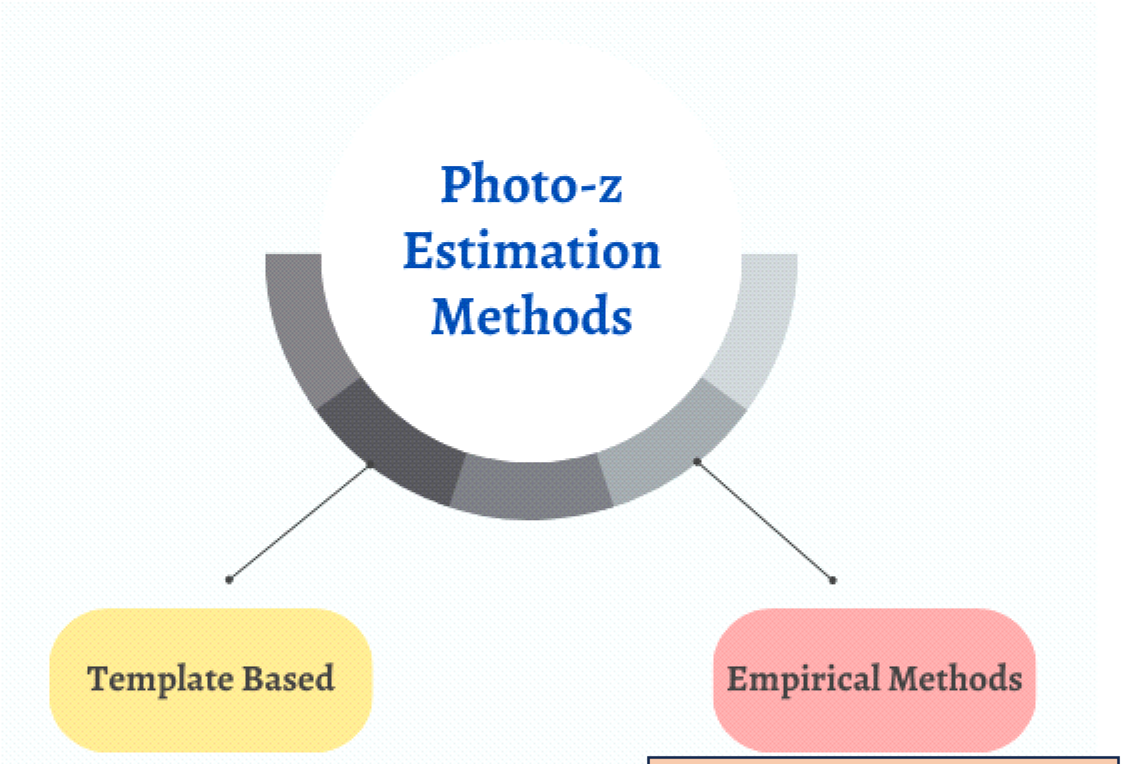
Photometric way

- Based on observed **photometric quantities**
- Mapping of photometric space into redshift space.
- Photo-z

→

- The measurable quantities or characteristics of an object's light.
- Derived from the object's intensity or flux measurements.

Photo-z estimation methods



Comparing observations to spectral energy distribution templates

Empirical relations of photometry to redshift

- Deep learning based estimation is an empirical method.
- Convolutional Neural Network (CNN) is used for photo-z estimation of galaxies in Kilo-Degree Survey (KiDS).
- KiDS is a wide-angle image campaign using broad band optical filters (u,g, r and i).

Convolution

- Small matrix of weights – Kernel/Filter
- Convolved with input data to extract features such as edges, corners etc. of input data.

a	b	c
d	e	f
g	h	i

Input Image

★

w	x
y	z

Kernel

=

aw+bx+dy+ez	bw+cx+ey+fz
dw+ex+gy+hz	ew+fx+hy+iz

Feature Map

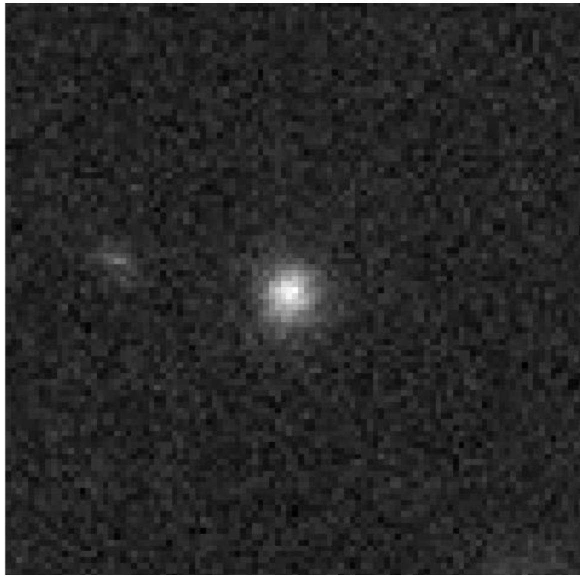
- Activation function is applied in feature map to introduce non-linearity into the network.

Pooling and Dense layers

- Pooling layers are used to reduce the dimension of feature map.
- Feature map is a tensor; multi-dimensional array
- It is flattened into 1D array – Input for dense layer
- Performs high-level feature extraction based on low-level features learned by the convolutional layers and pooling layers.

Input Data

- Input images are galaxy cutouts from KiDS DR4 bright sample.
- Network is trained by **galaxy 4- band images** and their corresponding spectroscopic redshifts (spec-zs).
- **Spec-zs** are obtained from Galaxy And Mass Assembly (GAMA) survey.
- Images are supplement with **9-band magnitudes** (u, g, r, i, Z,Y,J,H and, Ks)
- Cutout size = (36,36,4)
    - Height = 36 pixels
    - Width = 36 pixels
    - Number of bands = 4; (u, g, r and, i)



36x36 cutout of Galaxy image in i-band and its rmag=19.37

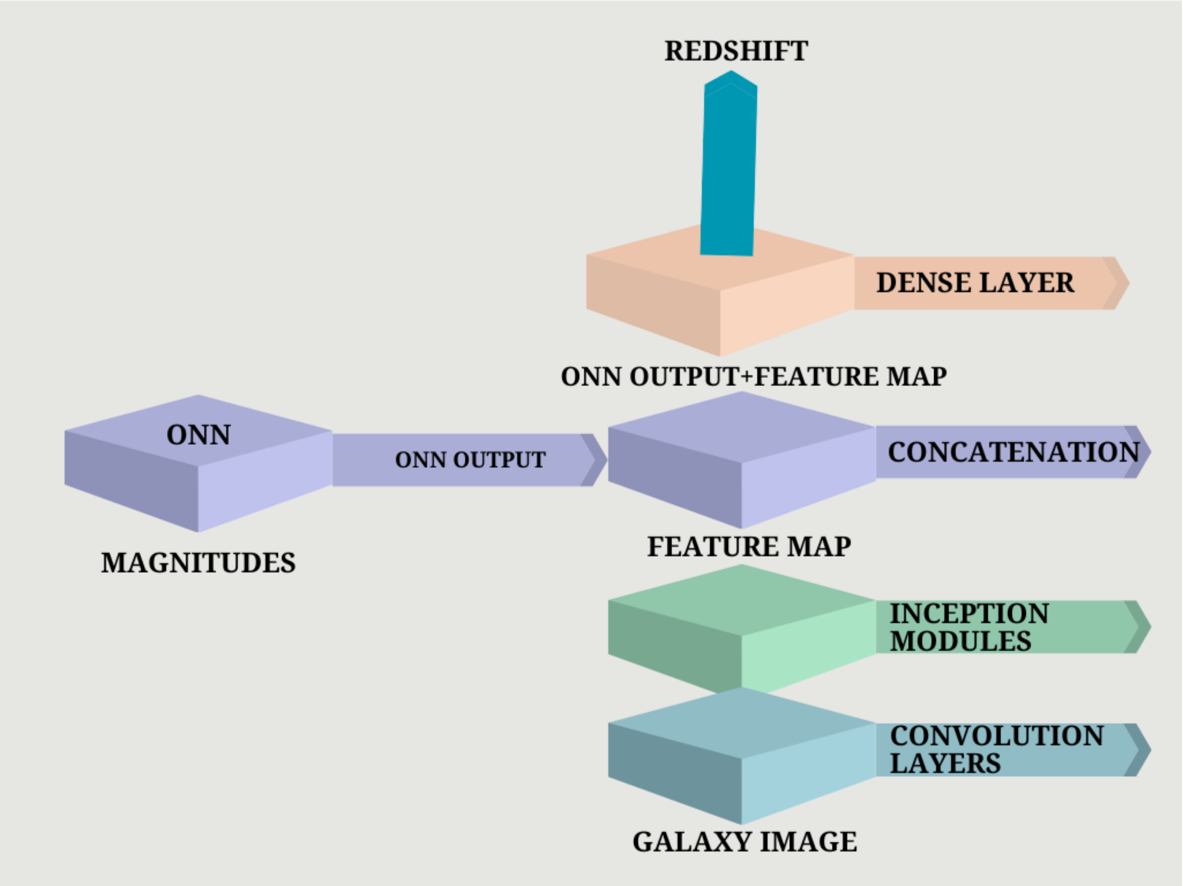
Inception

- Inception is a CNN architecture for image recognition.
- Parallel convolution operation

Inception Module

- Input layer
- 1x1 convolution layer
- 3x3 convolution layer
- 5x5 convolution layer
- Max pooling layer
- Concatenation layer

Inception based photo-z



- Developed a deep learning model based on Inception.
- Treated as a regression problem.

- This model uses two inputs:
  - Galaxy images
  - Magnitudes of galaxies

- ONN – Ordinary Neural Network
- Concatenation – Combines two outputs

Training

- Network predicts redshift.
- This predicted redshift is compared with the true redshift by loss function.
- Huber loss function is used.
- It is the combination of Mean Squared Error (MSE) and Mean Absolute Error (MAE).

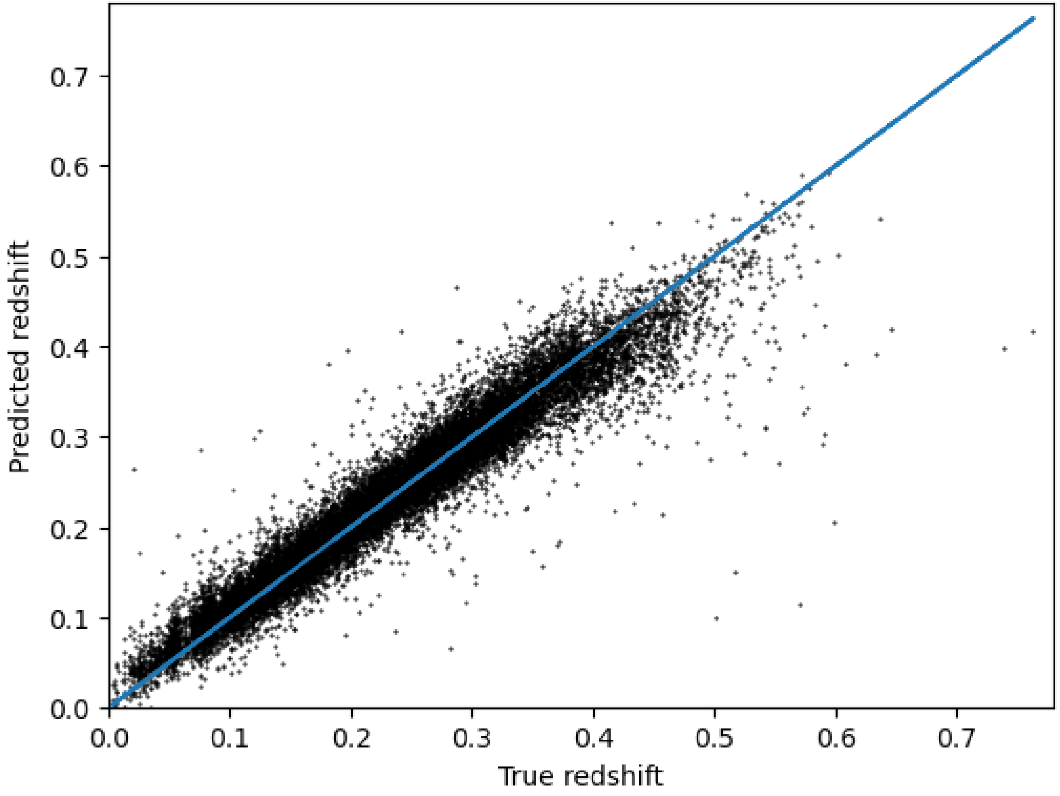
$$L_h = \begin{cases} \frac{1}{2}(e)^2, & |e| \leq \alpha \\ \alpha(|e| - \frac{1}{2}\alpha), & \text{otherwise} \end{cases}$$

  - E= true redshift – predicted redshift
  - α is a hyperparameter that determines the transition between MSE and MAE
- During training, the network tries to minimize this loss function by adjusting the weights in kernel.
- Training : Validation: Testing = 70:15:15

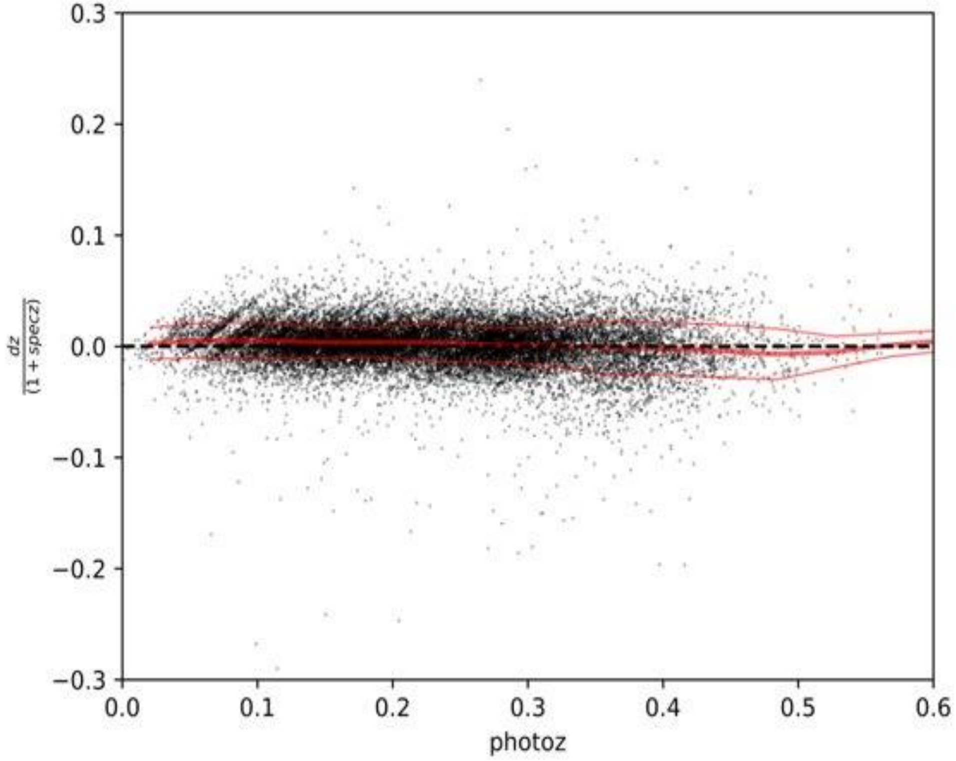
Statistics

- Bias
  - **dz = Photoz – Specz**
- Normalized bias
  - **normdz = dz / (1+specz)**
- Standard Deviation of normdz, **SD(normdz)**
- **Median (normdz)**
- Scale Median Absolute Deviation of normdz, SMAD(normdz)
  - Where, **SMAD(x) = 1.4826*median(|x-median(x)|)**

Result



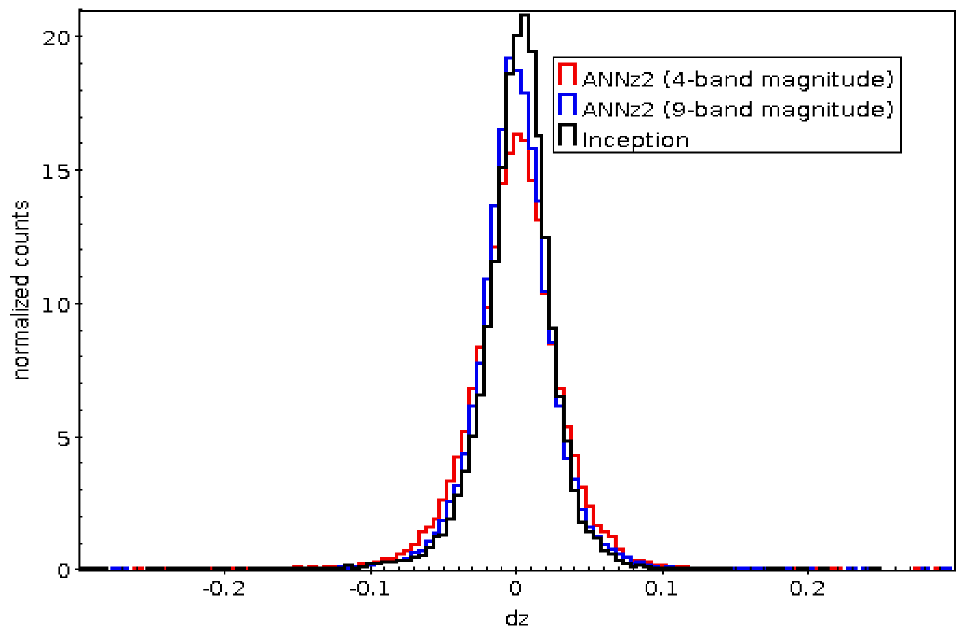
Predicted redshift
True redshift
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7



dz / (1 + specz)
photoz
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
0.4
0.5
0.6

Comparison with ANNz2 Result

- ANNz2 is a photo-z estimation method based on ordinary neural network. (*Bilicki et al. 2018, A&A 616, A69).*



Method	SMAD(normdz)
ANNz2 (4-band magnitude)	0.021
ANNz2 (9-band magnitude)	0.018
Inception	0.016

Future Work

- Apply to KiDS-DR5 bright sample. KiDS-DR5 is the final data release.


Thank You