M2 Thesis Defense

Anastasios Theodoropoulos

“Machine Learning for Photometric redshift estimation of LSST galaxies”

Supervisor: Prof. Simona Mei

Structure

Introduction:

- Vera Rubin Observatory - LSST
- Photometric redshift
- Machine Learning-Neural Networks
- Convolutional Neural Networks

Results

- Performance Metrics
- Sequential CNN
- GoogleNet
- Comparison

Conclusion and future prospects

Vera Rubin Observatory-LSST

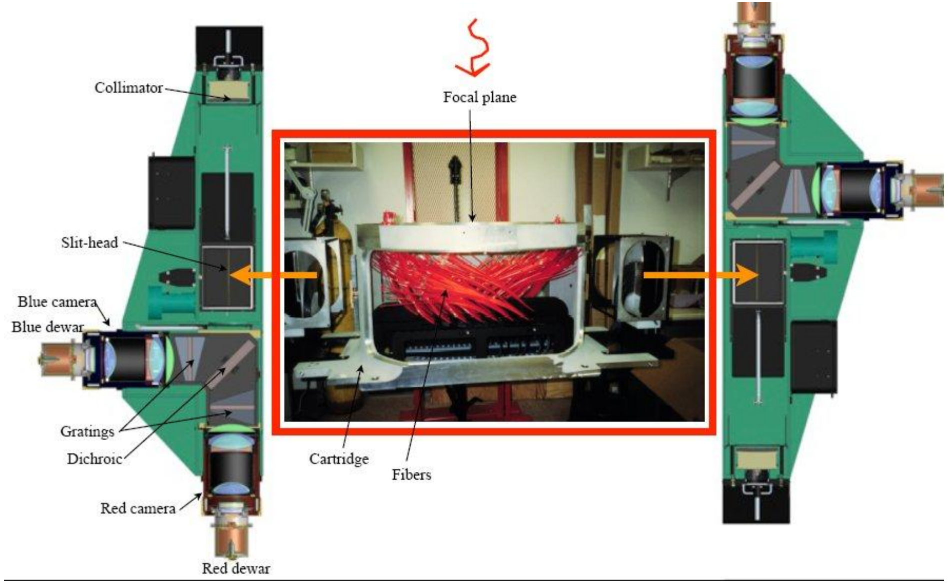
- Legacy Survey of Space and Time (LSST) is a wide field survey that is going to observe galaxies in redshifts up to approx. 1.3-1.5, in 6 broad filters (ugrizy).

- It will use the Vera Rubin Observatory and its main goal is to probe Dark Energy and Dark Matter.

- It is anticipated to image around 20 Billion galaxies during its 10 years projected run.

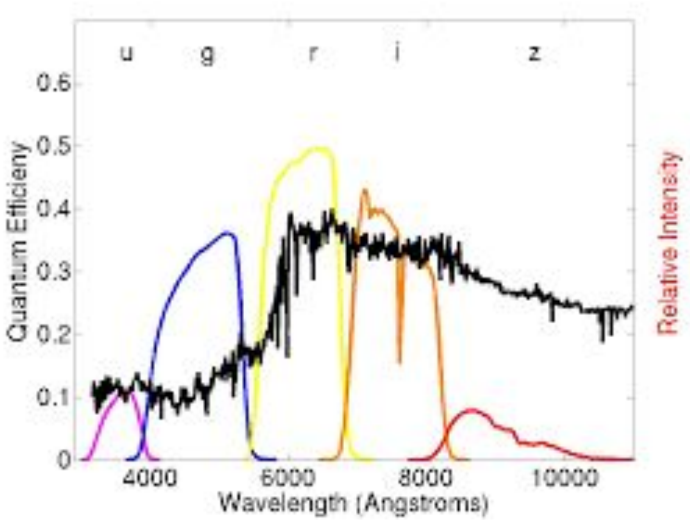
Vera Rubin Observatory-LSST

- A typical spectrograph can produce spectra of 1000 galaxies each time.
- The spectroscopic measurement of all the galaxies in the LSST survey is not possible.
- Thus, photometric redshift estimation is the only option.
- What is a photometric redshift?



Photometric redshift

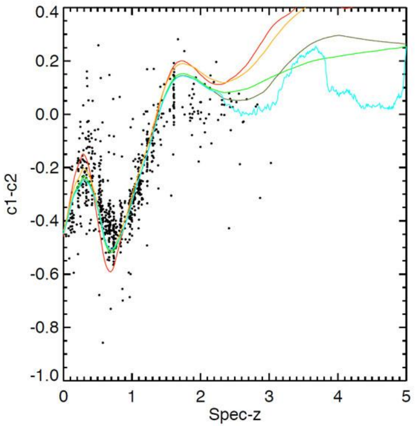
- When the redshift of an object is estimated through flux measurements in broader filters, like the ones used in imaging surveys, it is called a “Photometric redshift”.
- Essentially, it is just a mapping between the colour and the redshift.



Photometric redshift

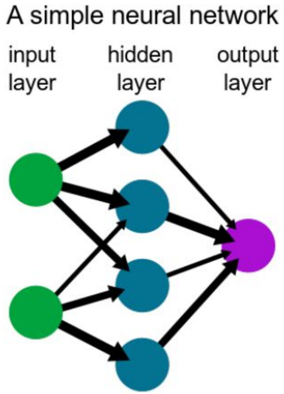
- Template fitting

The redshift-colour mapping is based on previous physical knowledge, e.g. galaxy star formation histories, extended observed spectra.



- Machine Learning

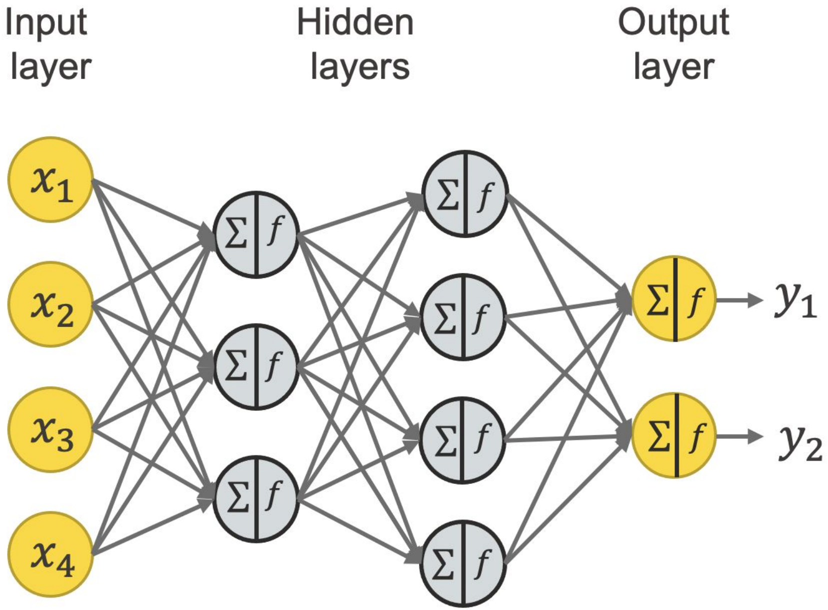
The redshift-colour mapping is obtained each time using a representative training sample with both photometry and redshift measurements.



Machine Learning–Neural Networks

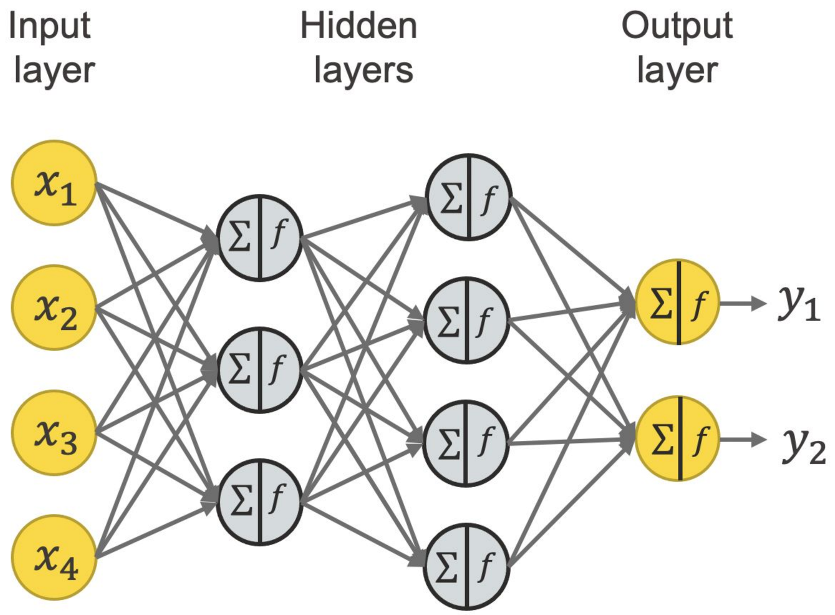
What is a Neural Network?

- An Artificial Neural Network is composed of artificial neurons which mimic the behavior of biological neurons.
- The structure consists of different layers.
- Each layer can be seen as a transformation of those which come before it.



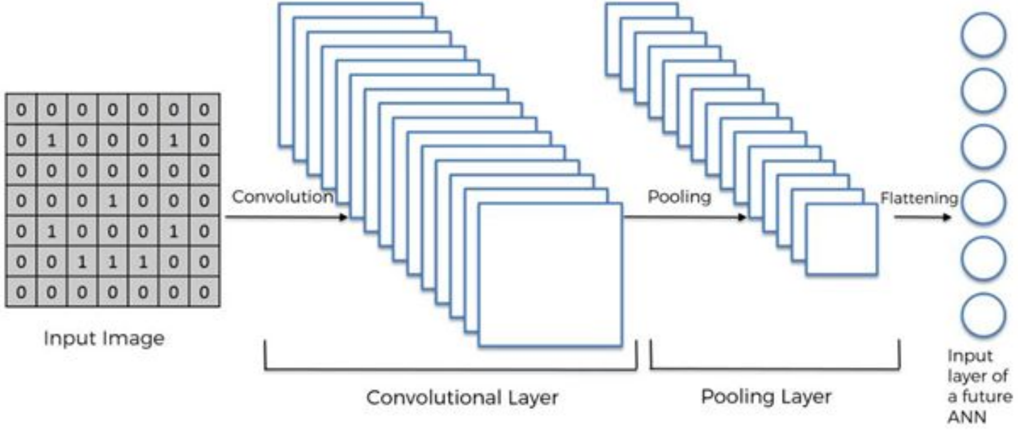
Machine Learning–Neural Networks

- The output of the NN is specified by the user and represents the target value of interest.
- The algorithm, using the training sample, tunes the transformations so that the output value minimises the residuals between the predicted and the true value.



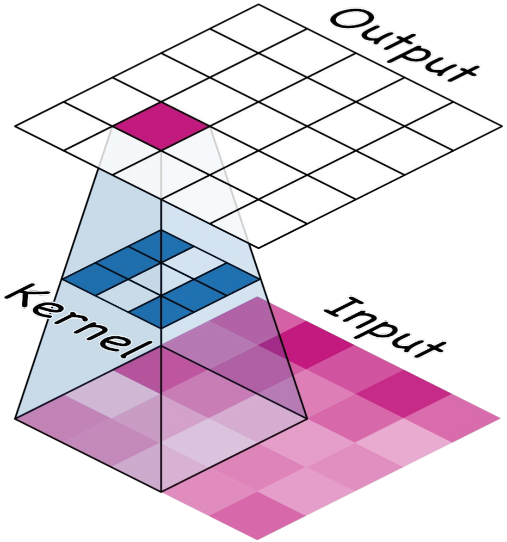
Convolutional Neural Networks (CNNs)

- CNNs are a special type of multilayered NN cosmposed of a number of convolutional and pooling layers followed by fully connected layers.



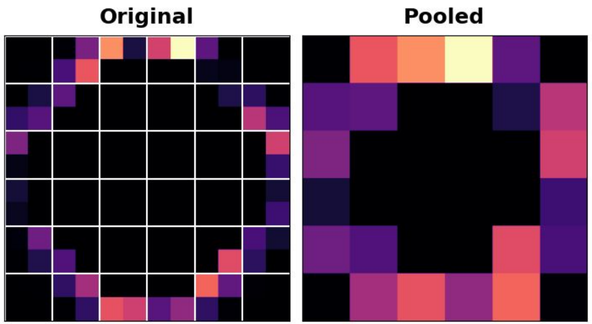
Convolutional Neural Networks (CNNs)

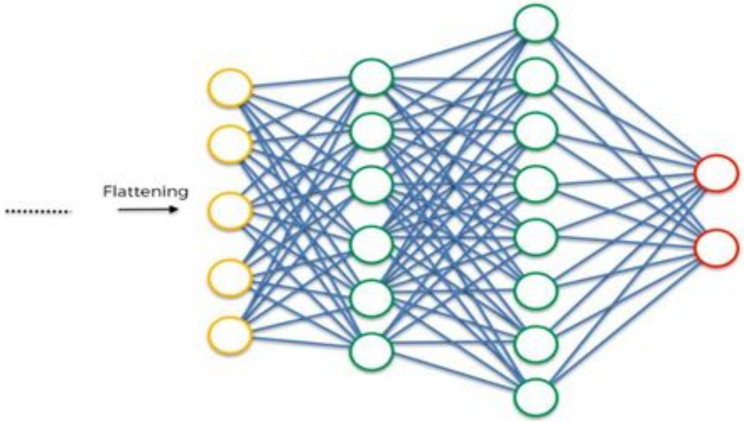
- Convolutional layers use kernels and they operate on a datacube and compute several feature maps which emphasises regions that correlate with a specific pattern represented by the kernels.
- During the training process, the kernels' weights get updated to become relevant to the classification.



Convolutional Neural Networks (CNNs)

- Pooling layers are used to reduce the size of input of the feature maps through downsampling along the spatial dimensions.
- The Fully Connected layer uses the features extracted from the previous layers for classifying the input image into various classes based on the training dataset.

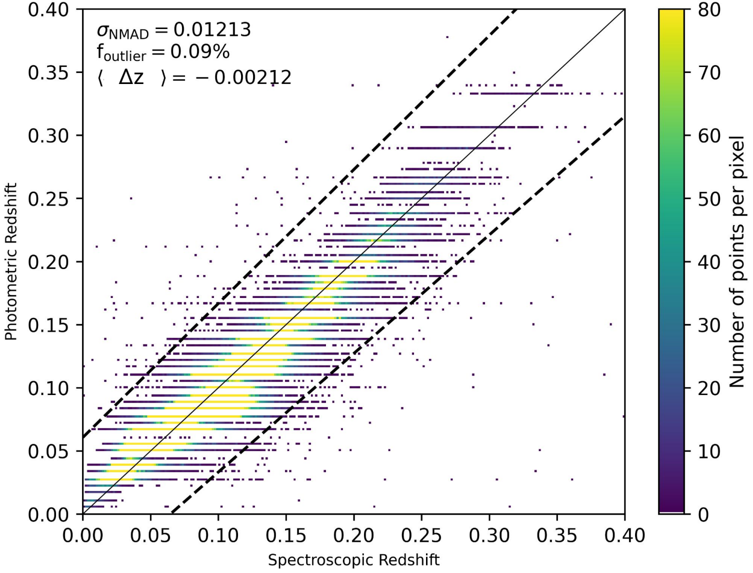




Performance Metrics

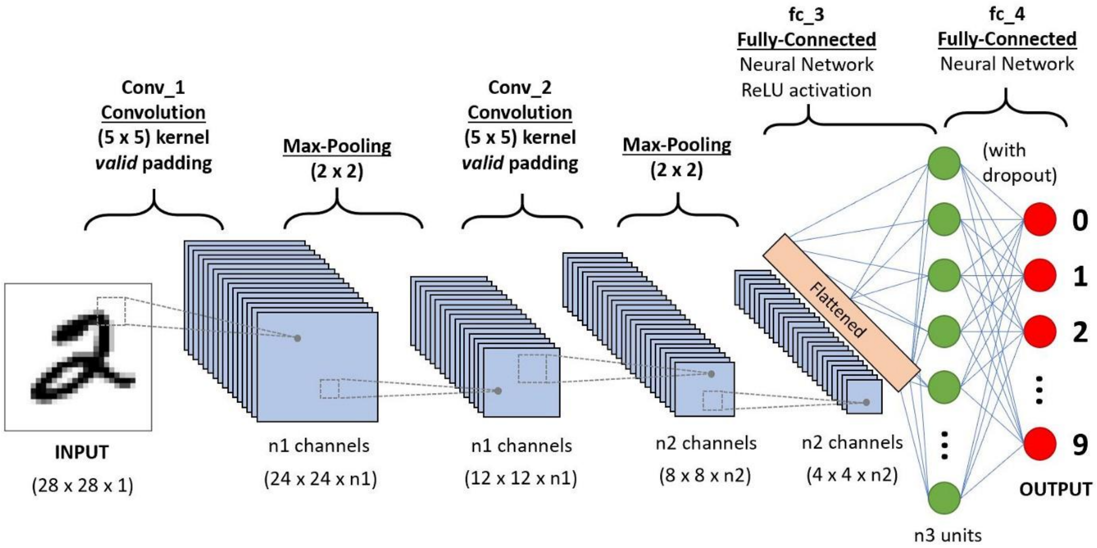
- Residuals: $$\triangle z = \frac{z_{phot} - z_{spec}}{1 + z_{spec}}$$
- Prediction bias: $$< \triangle z >$$
- MAD deviation: $$\sigma_{MAD} = 1.4826 \times MAD$$
- MAD (Median Absolute Deviation): $$Median(|\triangle z - Median(\triangle z)|)$$
- Outlier fraction:

$$\eta = \text{outliers with } \triangle z > 0.5 \; or \; (5 \times \sigma_{MAD})$$

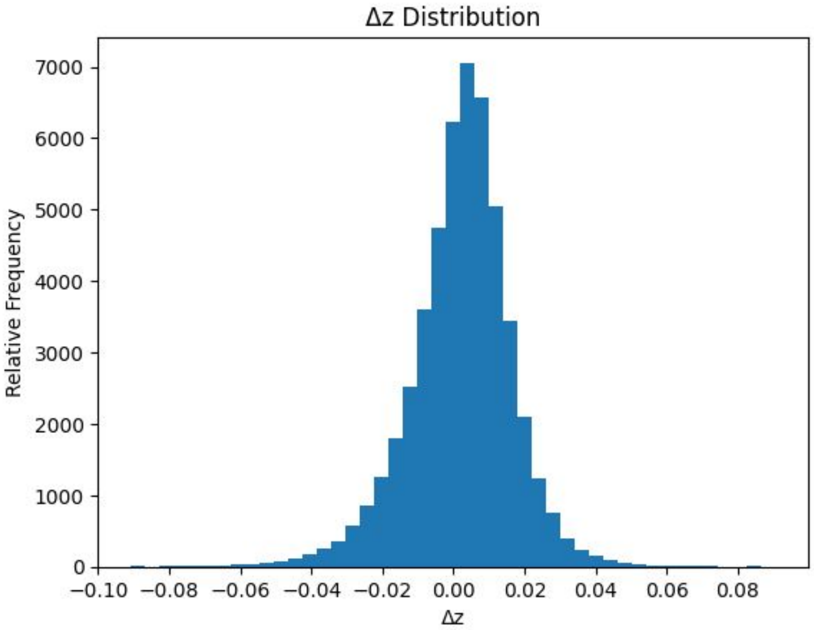


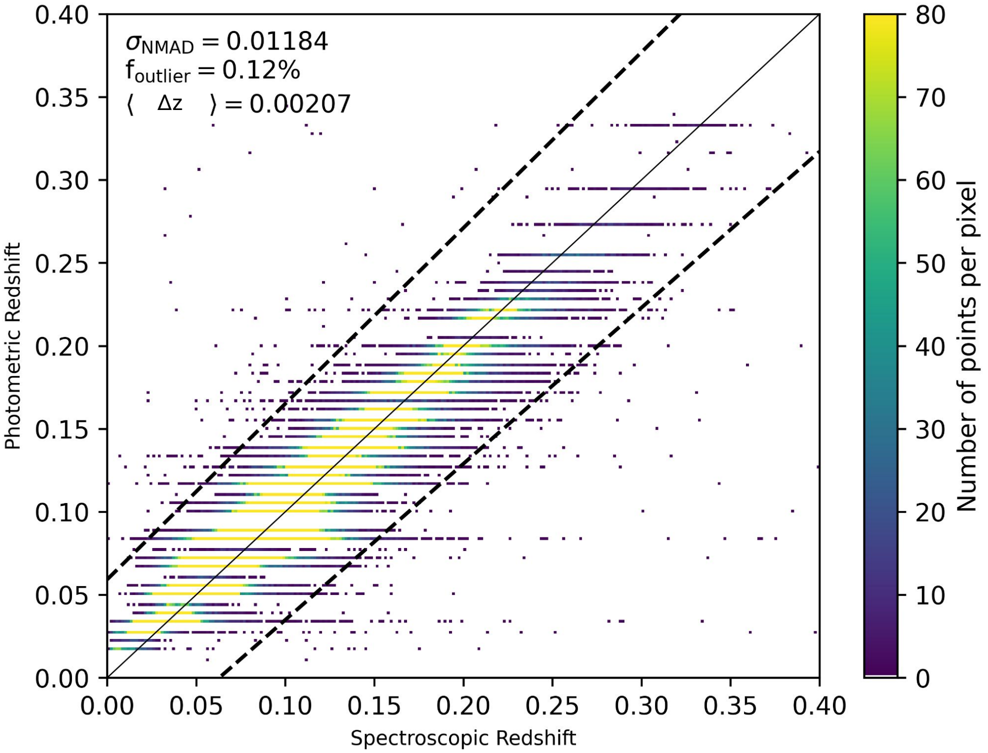
Sequential CNN

- The simplest possible CNN.



Sequential CNN – Results



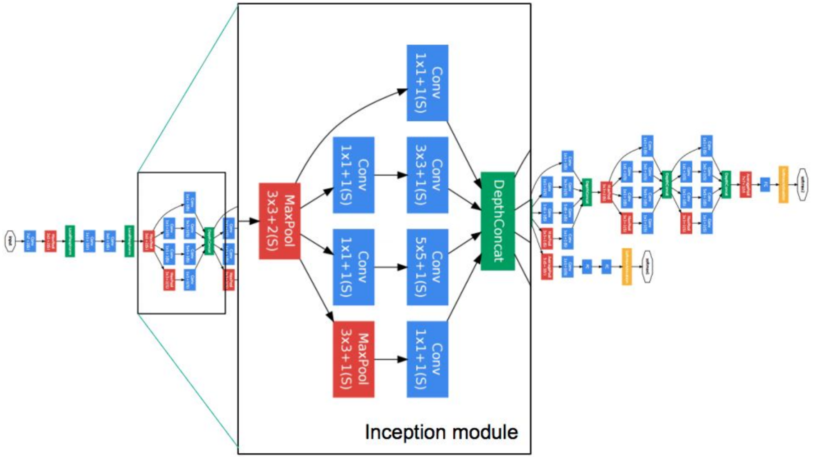


GoogleNet

- Introduced in 2015 by Szegedy et al.
- It is based on the construction of Inception modules.

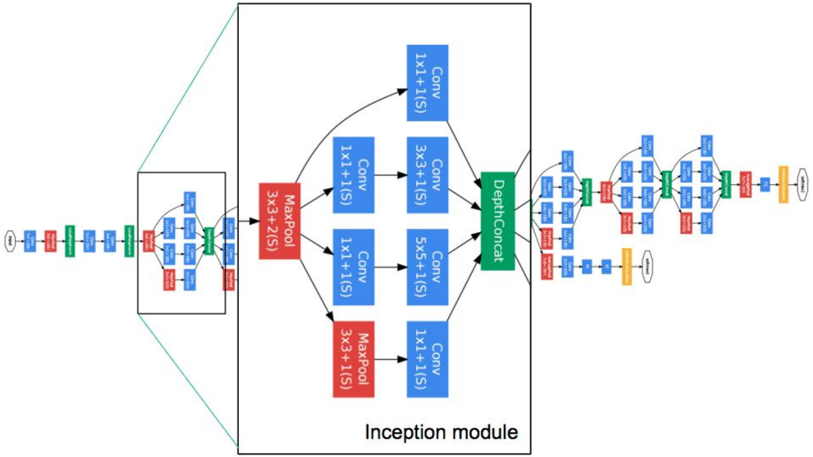
Each inception module is organised in two stages:

- At first, the feature maps are convolved by three “1x1” convolutional layers.
- Thus, they are being combined and reduced in number.

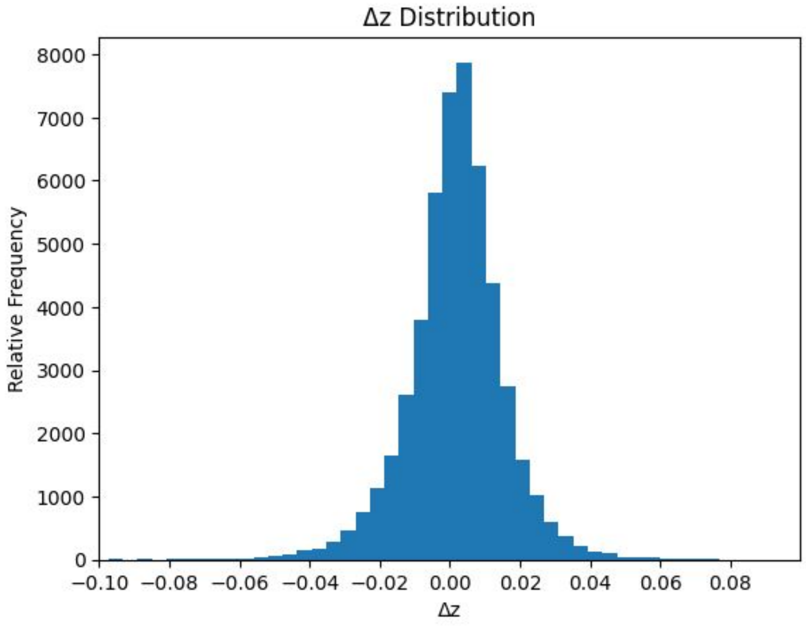


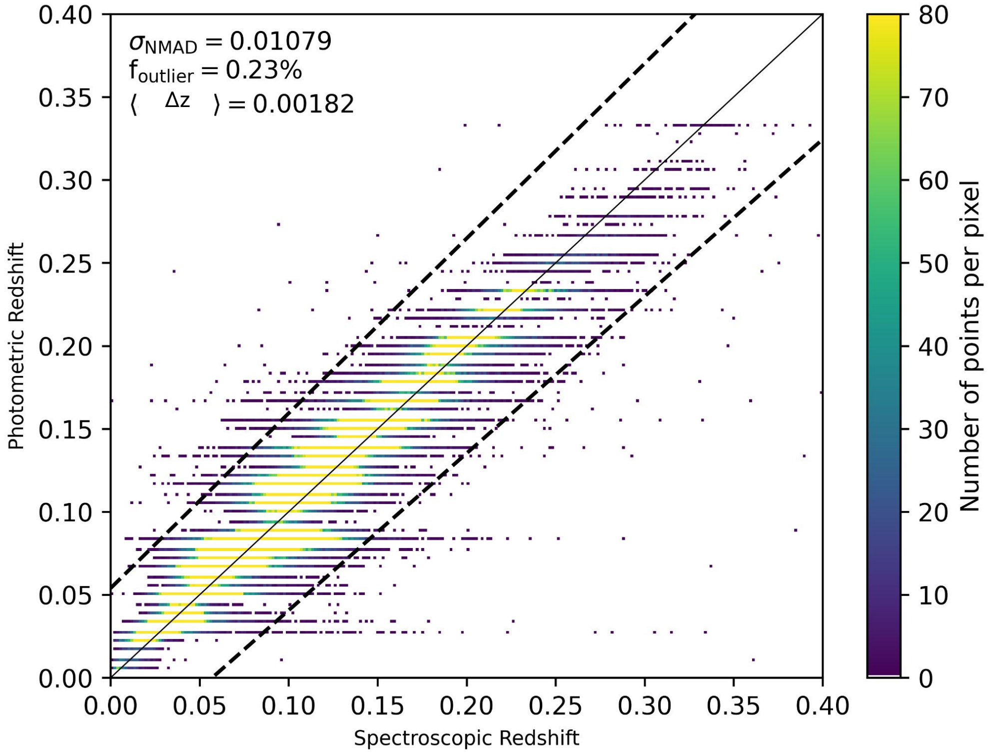
GoogleNet

- Then, they are processed in parallel in a pooling layer and a pair of larger convolution layers, such as: “3×3” and “5×5”.
- These layers help the network identify larger patterns.
- Lastly, the resulting feature maps are concatenated along the depth dimension before going to the fully connected layer.

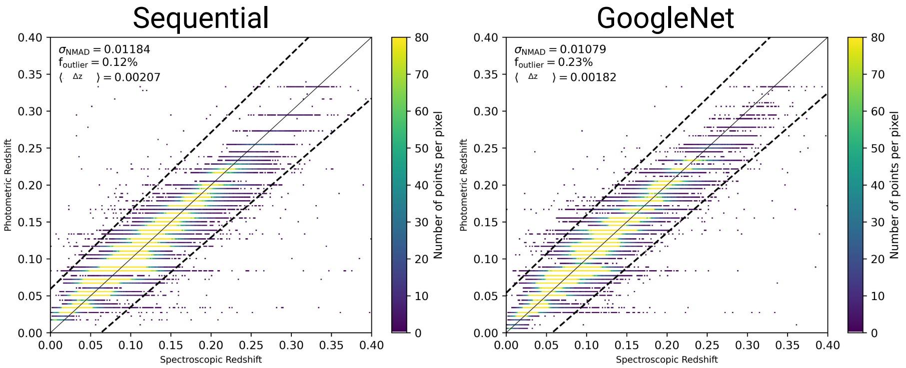


GoogleNet – Results

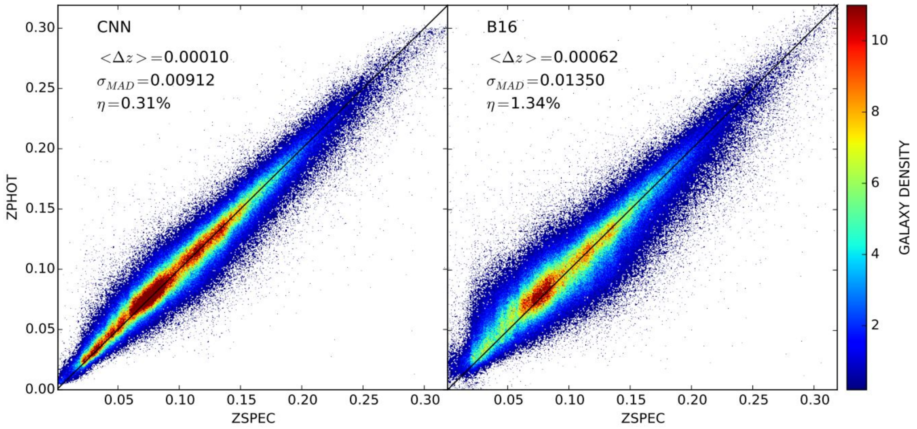




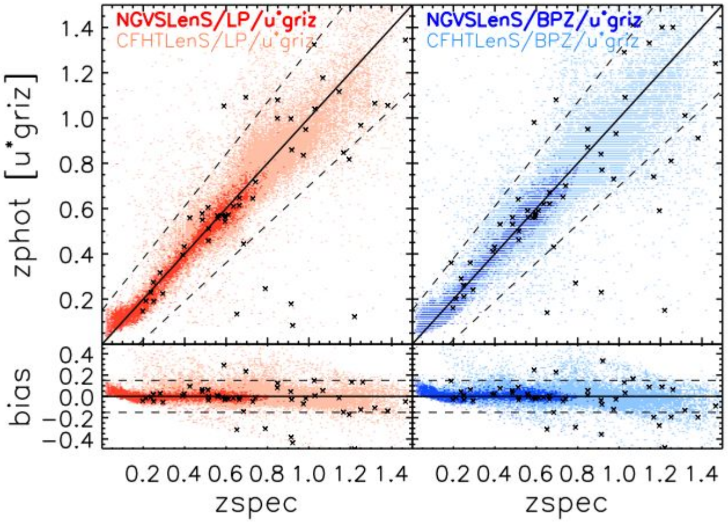
Comparison



Sequential
GoogleNet
$\sigma_{NMAD} = 0.01184$
$f_{outlier} = 0.12\%$
$\langle \Delta z \rangle = 0.00207$
$\sigma_{NMAD} = 0.01079$
$f_{outlier} = 0.23\%$
$\langle \Delta z \rangle = 0.00182$
Photometric Redshift
Spectroscopic Redshift
Number of points per pixel



CNN
$<\Delta z> = 0.00010$
$\sigma_{MAD} = 0.00912$
$\eta = 0.31\%$
B16
$<\Delta z> = 0.00062$
$\sigma_{MAD} = 0.01350$
$\eta = 1.34\%$
ZPHOT
ZSPEC
GALAXY DENSITY



NGVSLenS/LP/u*griz
CFHTLenS/LP/u*griz
NGVSLenS/BPZ/u*griz
CFHTLenS/BPZ/u*griz
zphot [u*griz]
bias
zspec

Conclusion

- In this work, I present the use of two types of CNNs in photometric redshift estimation of galaxies.
- Essentially, the information present in the images is being exploited without the need of pre-extracted spectral or image features.
- The value of σ_{MAD} I obtained is comparable with the best models published so far.
- Even when using the simplest possible CNN, the MAD deviation does not worsen to a great extent, staying comparable to the best performing models, while having much less trainable parameters.

Conclusion–Future prospects

- The bias appears to be one order of magnitude higher.
- A possible explanation is due to a trade-off between higher dispersion and lower bias as we continue the training of the models.
- The first possible extension of this work is the use of data coming from simulations of the LSST data, and also the first real data.
- Also, another possible extension is the use of current state-of-the-art networks such as RNNs and EfficientNets for Photometric Redshift estimation and the comparison with current models.

Thank you for your attention!

The code, along with the plots, produced during this internship are publicly available in my github [page]

References

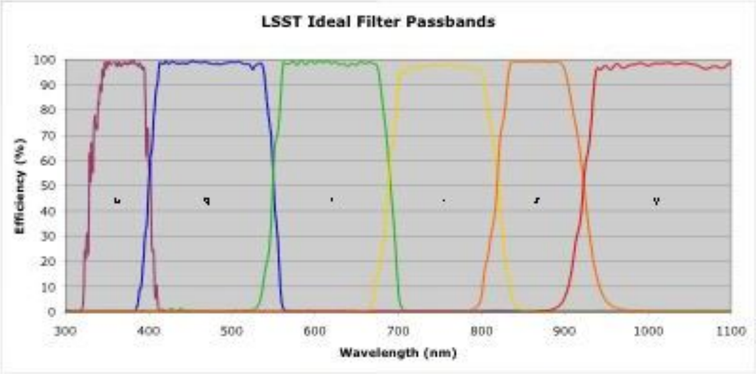
- Pasquet, J., Bertin, E., Treyer, M., Arnouts, S., & Fouchez, D. (2019). Photometric redshifts from SDSS images using a convolutional neural network. Astronomy & Astrophysics, 621, A26.
- Dey, B., Andrews, B. H., Newman, J. A., Mao, Y. Y., Rau, M. M., & Zhou, R. (2021). Photometric Redshifts from SDSS Images with an Interpretable Deep Capsule Network. arXiv preprint arXiv:2112.03939.
- Salvato, M., Ilbert, O., & Hoyle, B. (2019). The many flavours of photometric redshifts. Nature Astronomy, 3(3), 212-222.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
- Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning (pp. 6105-6114). PMLR.
- Raichoor, A., Mei, S., Erben, T., Hildebrandt, H., Ilbert, O., Licitra, R., ... & Zhang, H. (2014). The Next Generation Virgo Cluster Survey. XV. The photometric redshift estimation for background sources. The Astrophysical Journal, 797(2), 102.

Extra Slides–LSST Filters

Specs

Half-Maximum Transmission Wavelength			
	Blue Side	Red Side	Comments
U	350	400	Blue side cut-off depends on AR coating
G	400	552	Balmer break at 400 nm
R	552	691	Matches SDSS
I	691	818	Red side short of sky emission at 826 nm
Z	818	922	Red side stop before H_2O bands
Y	948	1060	Red cut-off before detector cut-off

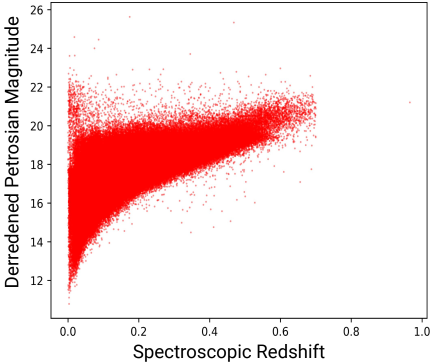
- 75 cm dia.
- Curved surface
- Filter is concentric about the chief ray so that all portions of the filter see the same angle of incidence range, 14.2° to 23.6°

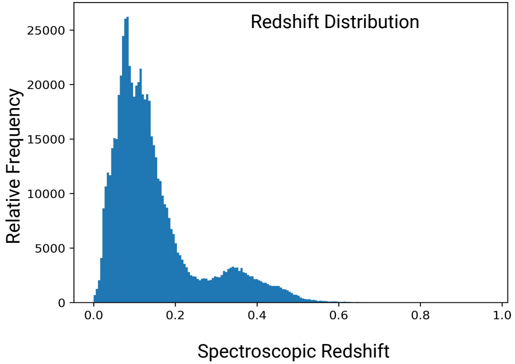


Uniform deposition required at 1% level over entire filter

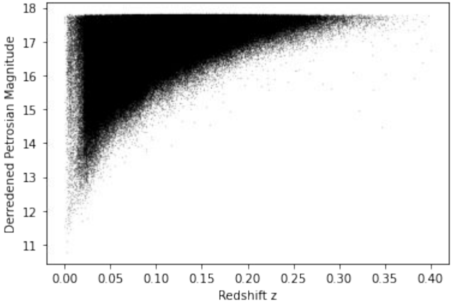
Extra Slides–Raw Data

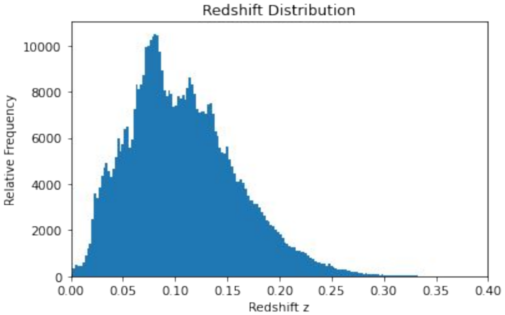
Before Cut



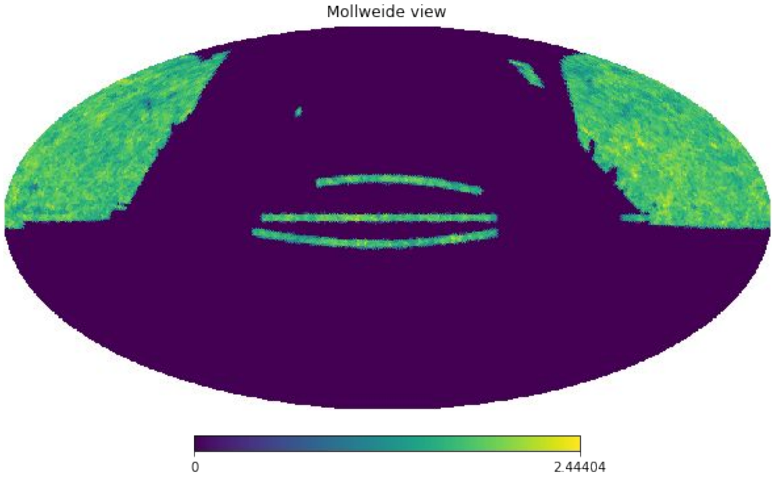


After Cut





Coordinates of the data



Extra Slides–Models Comparison