Weighing Black Holes Using Deep Learning

Sneh Pandya, Devanshi Pratap
Under Instruction of: Joshua Yao-Yu Lin, Xin Liu

Introduction

Goal

To develop a convolutional neural network that can predict mass of supermassive black holes using time series spectra and redshift data.

Supermassive black holes (SMBH) are ubiquitously found at the centers of most galaxies. Measuring SMBH mass is important for understanding their origin and evolution. We train Deep Learning (DL) algorithms that learn from the Sloan Digital Sky Survey (SDSS) Stripe 82 and DR7 data for a sample of $\sim$10,000 quasars to map out the nonlinear encoding between black hole mass and multi-color optical light curves (Sun, 2020). We apply data augmentation on spectral data to increase our SMBH light curve and mass sample to a total $\sim$100,000 quasars from the Dark Energy Survey Supernova fields.

Motivation

Traditional methods of weighing SMBH require spectral data which are expensive to gather, as well as tedious. Our results have direct implications for efficient applications with future observations from the Vera Rubin Observatory (LSST).

- Our neural network loss function is as follows

$$L = \frac{1}{n}\sum_{i=0}^{n}(y_i - \hat{y}_i)^2 \exp(-s_i) + s_i \,,$$

Data and Methodology

Data Matching

We adopt multi-color photometric light curves from SDSS Stripe 82 as our spectral data. Our baseline sample consists of $\sim$10,000 quasars in the Stripe 82 survey. We assume the virial black hole mass estimates from the SDSS DR7 catalog as the ground truth, and match the two according to their equatorial coordinates. Spectra measure both how fast a gas cloud is moving and how far it is from the black hole. The black hole mass is obtained using laws of Keplerian motion similar to that of weighing stellar systems.

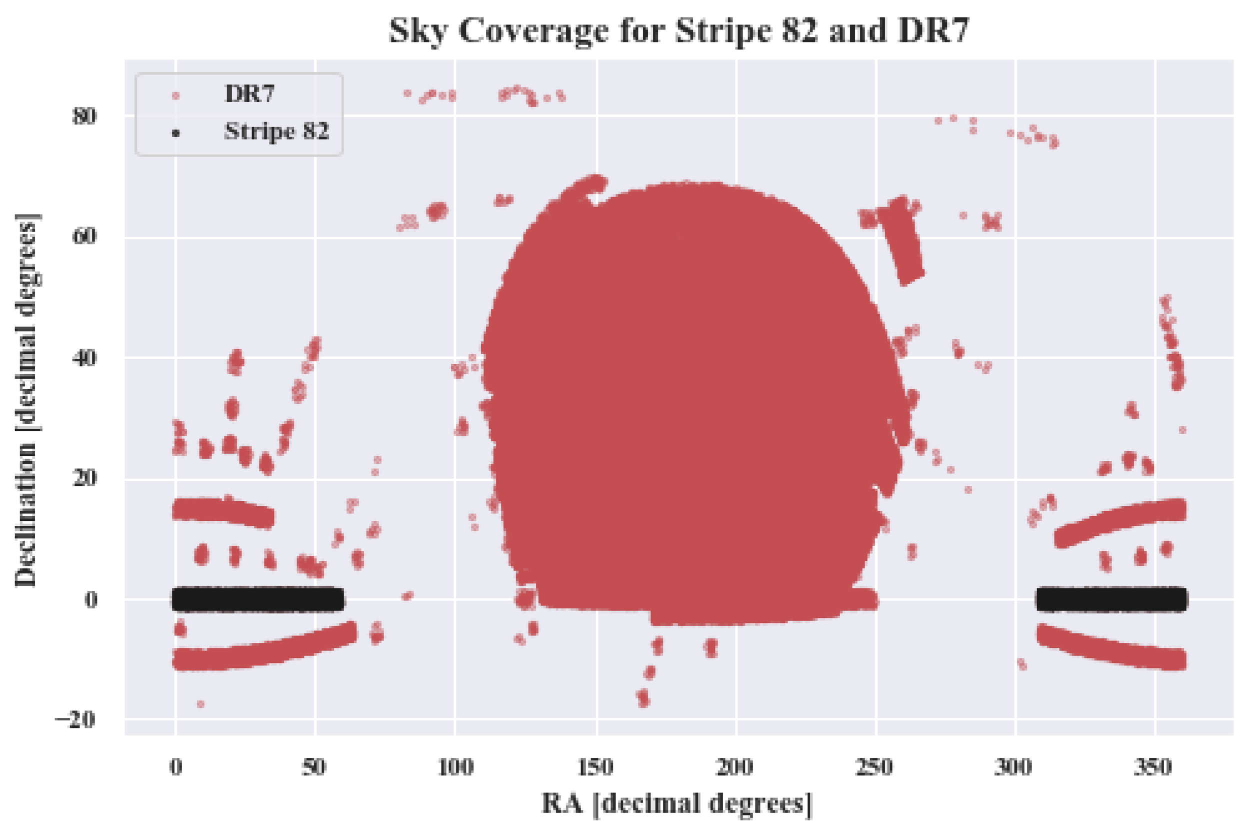

Figure 1: *Matched DR7 and Stripe 82 Data Sets*

Data Augmentation

To eliminate small sample size as a factor in neural network performance, we use data augmentation to simulate 10x new light curves using random seeds. Simulated light curves share redshift and ID information with initial objects, and we simulate new magnitudes for 5 bands (u,g,r,i,z) using error information within 1σ of original magnitude.

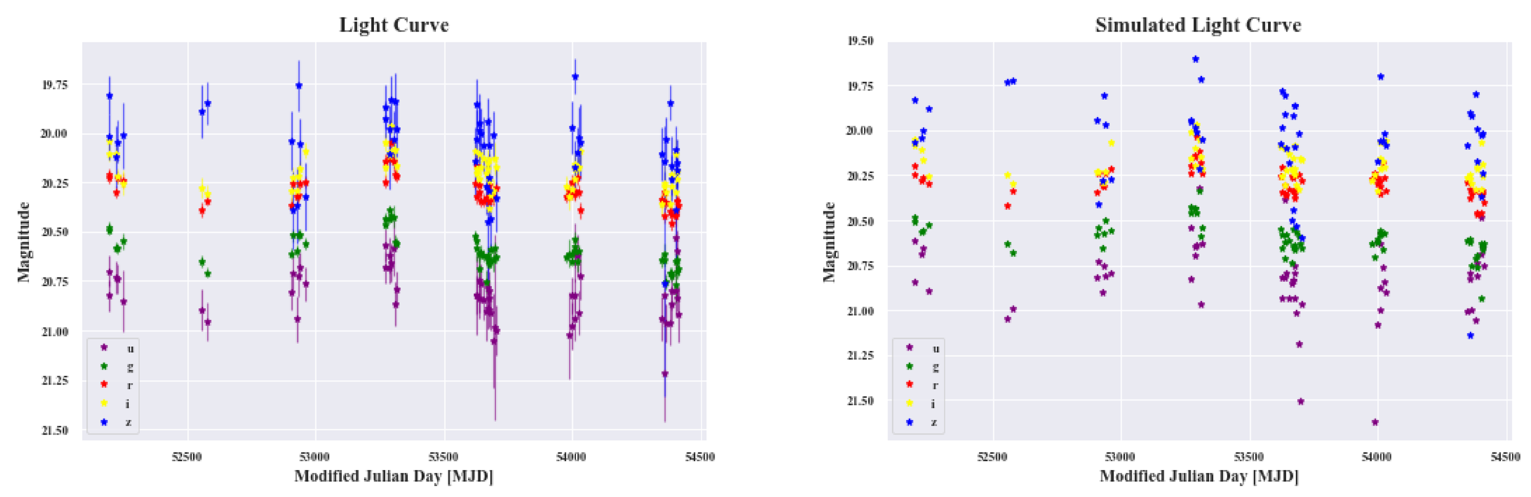

Figure 2: *Original and Simulated Light Curve*

Network Implementation

We first reshape our light curves into 224 x 224 numpy images to feed our network. We use deep convolutional neural networks (CNNs) to predict black hole mass directly from quasar light curves, employing *Pytorch*, a deep learning Python library. We use the standard 18/34/50 layers of deep residual network architecture as our baseline, and further modify the last layer by adding a fully connected layer so that it outputs the number of parameters we desire. The skip connection helps the neural network to be trained without the *vanishing gradient problem*. It's been shown that neural networks with skip connections have a better ability to approach the minimum of highly non-convex loss functions with a smoother loss surface. We use a pre-trained ResNet and add an output neuron at the last layer for outputting the value we wish to obtain.

Find our code *AGNet* at: https://github.com/devanshipratap/DeepLearningAGN/ (QR code above)

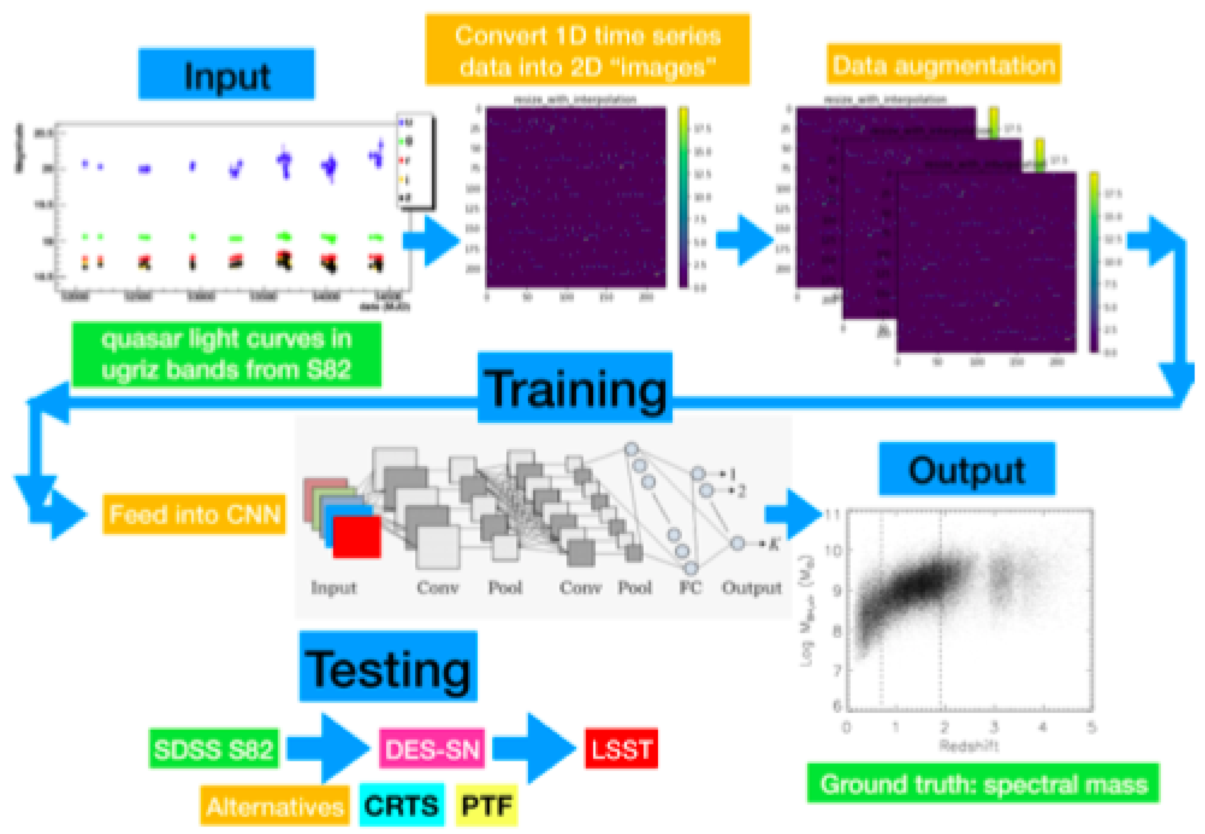

Figure 3: *Network Implementation Flowchart*

Statistical Analysis and Testing

We split our data set of $\sim$100,000 quasars into a 80/20 training and testing set. We test our data for an initial neural network framework proven to be effective in predicting redshift, and in our altered network architecture for obtaining mass values. We also analyze the error distribution of our quasar catalog to validate any concerns about measurement in error. A degree of error is apparent, however most errors are small ($<$0.25) and are confined to our lower mass quasars (Figure 4).

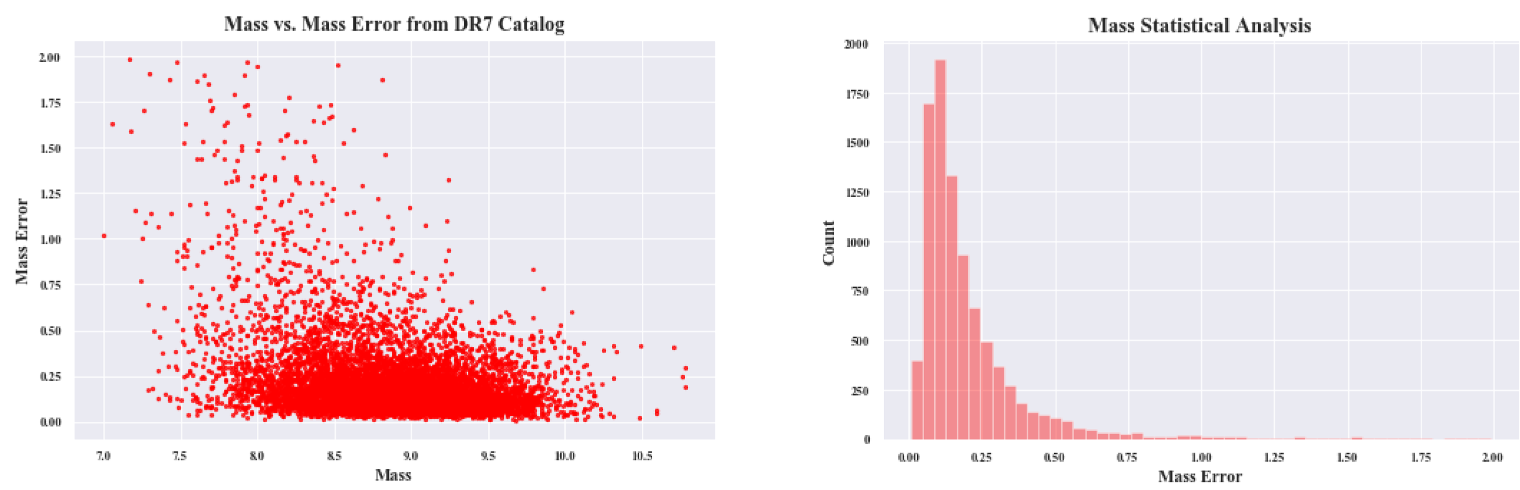

Figure 4: *Error Information Regarding Quasar Masses*

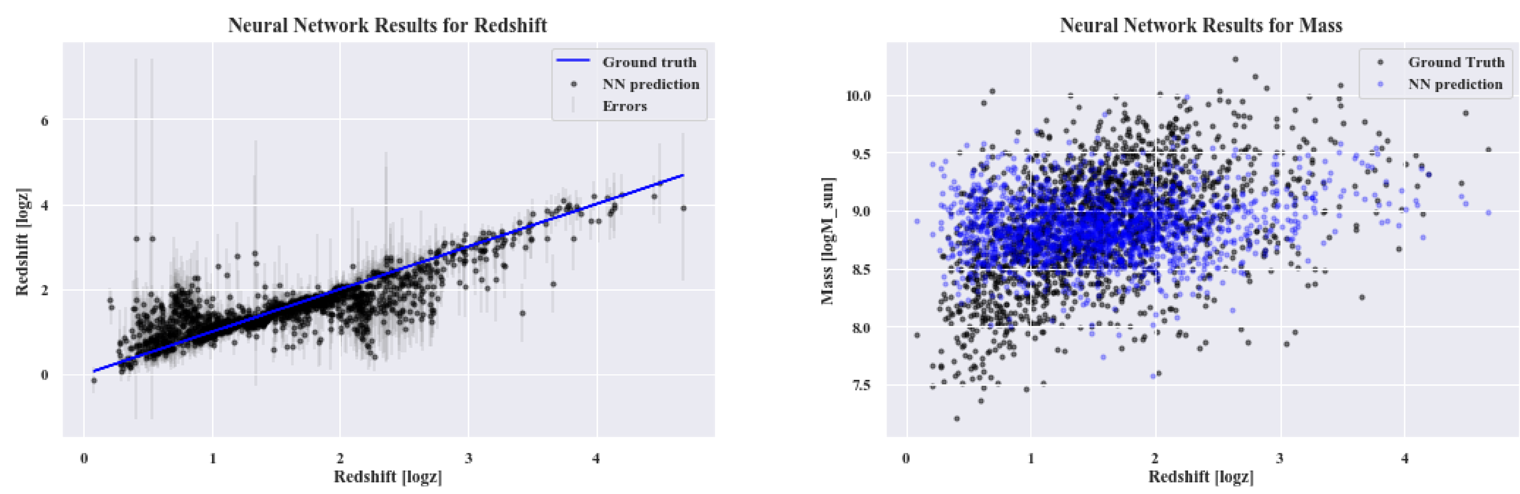

Figure 5: *Neural Network Redshift and Mass Results*

Discussion and Next Steps

It is evident from our results that our network performs well in predicting redshift, which was initially shown in Pasquet-Itam & Pasquet, 2018. In mass estimations (Figure 5) the network shows signs of learning but estimations tend to consolidate around the mean. In the future, modified network architectures would be advisable. Sampling datasets outside of SDSS Stripe 82 and DR7, such as the Dark Energy Survey (DES) and Catalina Real-Time Transient Survey (CRTS) and eventually LSST may also improve network results.

Support for this project was provided by the National Center for Supercomputing Applications SPIN Program and the Department of Physics, Department of Astronomy at the University of Illinois at Urbana-Champaign. We also acknowledge the NCSA HAL cluster, without which most of our computation could not be possible.

References

Pasquet-Itam, J., and J. Pasquet. âDeep Learning Approach for Classifying, Detecting Andpredicting Photometric Redshifts of Quasars in the Sloan DigitalSky Survey Stripe 82.â Astronomy amp; Astrophysics, vol. 611, 2018,

Sun, Mouyuan et al. âCorona-Heated Accretion-Disk Reprocessing: A Physical Model to Decipher the Melody of AGN UV/Optical Twinkling.â The Astrophysical Journal 891.2 (2020): 178. Crossref. Web.