







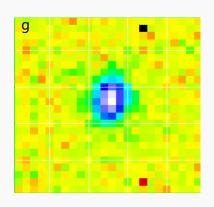
类星体



类星体(quasar)是极度明亮的活动星系核(AGN, active galactic nucleus),是宇宙中最明亮、活动最剧烈的天体之一。其中心是超大质量黑洞,这些黑洞虽然自身不发光,但由于其强大的引力,周围物质在快速落向黑洞的过程中以类似"摩擦生热"的方式释放出巨大的能量,使得类星体成为宇宙中最耀眼的天体。绝大多数类星体都有非常大的红移值。



类星体巡天图像



在所有天文数据中,图像是较为直接简单的数据获取方式。巡天图像的分析,可为光谱望远镜提供更加高效的选源指导和红移估算,有助于正在建设的中国空间站光学巡天望远镜(CSST)低像素巡天图像的研究,对宇宙历史和演化的研究都有着积极意义。然而类星体的巡天图像成像很小,约为5*5个像素,这对天体识别和参数估计等工作来说是一个挑战。





相关研究

作者	年份	模型	研究对象	使用数据
Hoyle	2016	DNN(Deep Neural Networks)	galaxy	image
Pasquet et al.	2019	CNN(Convolutional Neural Networks)	galaxy	image
Desprez G. et al.	2020	Hybrid Systems	galaxy	photometry & image
Razim et al.	2021	MLPQNA (Multi-Layer Perceptron with Quasi- Newton Algorithm)	galaxy	photometry & spectral
Y`eche et al.	2010	ANN(Artificial Neural Networks)	quasar	photometry
D'Isanto & Polsterer	2018	DCMDN(Deep Convolutional Mixture Density Network)	quasar	image
Curran et al.	2020	KNN(K-Nearest Neighbor)	quasar	photometry
Yao et al.	2023	Q-PreNet(Quasar Photometric Estimation Network)	quasar	photometry & image
Hong et al.	2023	MML(Muti-modal Machine Learning)	quasar	photometry & spectral

厚供外稅市長為行

相关研究

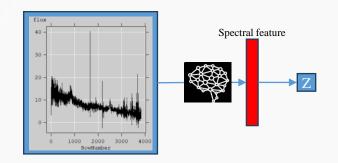
总结:

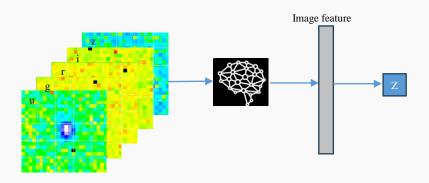
相比于星系红移,现有对类星体使用巡天图像进行研究的工作较少,大部分工作集中于测光、光谱等数据的研究,但这些数据的获取不管是复杂程度亦或是成本相比于巡天图像的获取都要高,而且在天体目标的识别阶段,若是缺乏对应天体的相关测量数据,上述方法都不再适用。为了解决上述问题,本课题组想要实现一个在应用阶段只使用图像数据就能够预测出让人较为满意的红移值的模型。





模拟光谱特征的生成思想

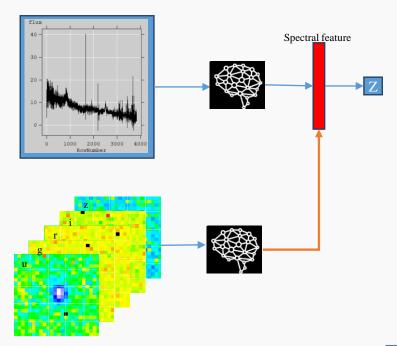




前人的工作已经证明:光谱数据可以很好地预测出 红移,而类星体巡天图像在预测红移时又受限于像素过 少、特征不足的问题。那么光谱的特征信息是否对利用 巡天图像的红移预测有所帮助?



模拟光谱特征的生成思想





数据预处理

数据来源: SDSS DR16Q

数据类型: ugriz五波段图像

光谱flux数据

红移值



斯隆数字巡天



数据预处理

图像处理: 原始尺寸→(24, 24, 5)→归一化



光谱flux处理: 原始长度→(3600, 1)→归一化

类星体星体选择原则:

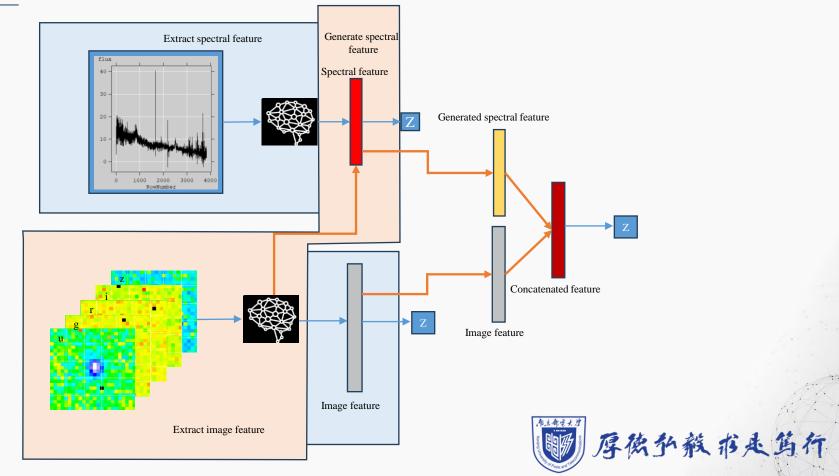
- 1.红移值小于4
- 2.切割后形状满足(24, 24, 5)
- 3.存在光谱数据

最终数据集大小: 729,674

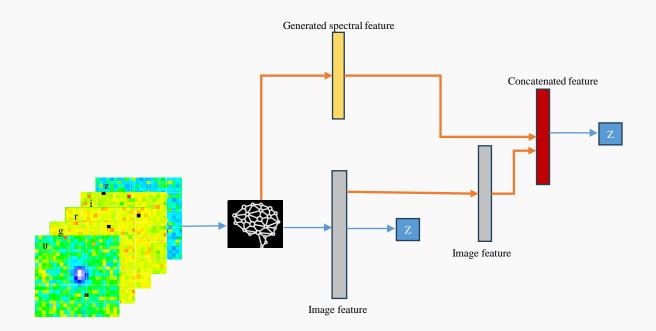


厚供外裁市是笃行

模拟光谱特征的生成模型



模拟光谱特征生成模型的应用







研究进展

评估指标 :
$$|\Delta z| = \left| \frac{z_{phot} - z_{spec}}{1 + z_{spec}} \right|$$

Experiment conducted on a dataset consisting of 100,000 quasars.

	$ \Delta z < 0.1$	$ \Delta z < 0.15$	$ \Delta z < 0.3$
D'Isanto & Polsterer (2018, image)	/	0.620	0.847
Yao et al.(2018, image&photometry)	/	0.856	0.958
Spectra(ours)	0.972	0.9841	0.988
Image(CNN, ours)	0.774	0.831	0.888
Image with generated spectral feature(CNN, ours)	0.793	0.841	0.891



研究进展

评估指标 :
$$|\Delta z| = \left| \frac{z_{\text{phot}} - z_{\text{spec}}}{1 + z_{\text{spec}}} \right|$$

Experiment conducted on a larger dataset

	$ \Delta z < 0.1$	$ \Delta z < 0.15$	$ \Delta z < 0.3$
D'Isanto & Polsterer (2018,image)	/	0.620	0.847
Yao et al.(2018,image&photometry)	/	0.856	0.958
Image_A(300,000, ours)	0.751	0.808	0.863
Image_B(429,674, ours)	0.687	0.781	0.867
Image_All(729,674, ours)	0.714	0.795	0.883



总结

- 1. 在10万条数据集上已经初步证明,图像与图像生成的模拟光谱特征的多模态结果优于仅使用图像的单模态结果。
- 2. 在使用不同数据样本时,|**Δz**| < **0**.**1**的占比有较为明显波动,分析为所用数据集样本分布不同导致,启发后面可以尝试使用无监督模型分类后,再分类别进行针对性训练。

