德州学院
DEZHOU UNIVERSITY

基于多模态深度学习的星系物理参数预测

汇报人： 王丽丽

单位：德州学院

电话：15865348734

邮箱：jsjxwll@126.com

2025-11-28

海南，海口

目录 /CONTENTS

01 研究背景

02 研究内容与方法

03 结果分析

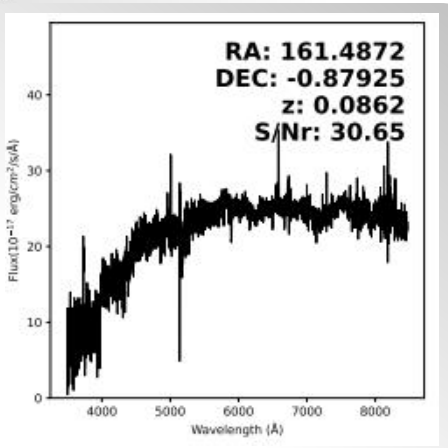
04 总结



01

研究背景

研究背景





星系

星系年龄：记录恒星形成与演化历程，是揭示星系构建历史的时间刻度。

金属丰度：衡量星系内部重元素含量，反映化学富集和恒星形成效率。

➡ **估计年龄与金属丰度，对理解星系形成与演化至关重要。**

光谱

传统方法通过连续谱、吸收线解析物理性质，方法成熟，但存在以下问题：

- 光谱拟合计算量比较大 ->基于卷积神经网络预测星族 https://doi.org/10.1093/mnras/stad3756
- **光谱获取成本高，数量远少于测光图像**

测光图像

图像包含形态、颜色、亮度等信息，观测量大，但是：

与年龄金属丰度相关的信息量远少于光谱->参数估计精度不高

仅依靠图像或有限的图像–光谱信息，准确预测星系年龄与金属丰度



02

研究内容与方法

- 研究内容
- 模型整体框架
- 数据
- 研究方法

研究内容

提出一种多模态学习框架，在训练时融合图像和光谱特征对星系年龄和金属丰度进行学习，在推理阶段只需要输入图像即可得到预测值。旨在克服只使用图像预测参数不准确的问题，从而得到更海量的星系年龄和金属丰度，有助于大样本统计研究。

Li Ping; Wang Li-Li; Yang Guang-Jun; et al. Astronomy and Astrophysics, 2025, 698, A222
https://doi.org/10.1051/0004-6361/202553751

创新点

01

构建图像→光谱特征转换模型

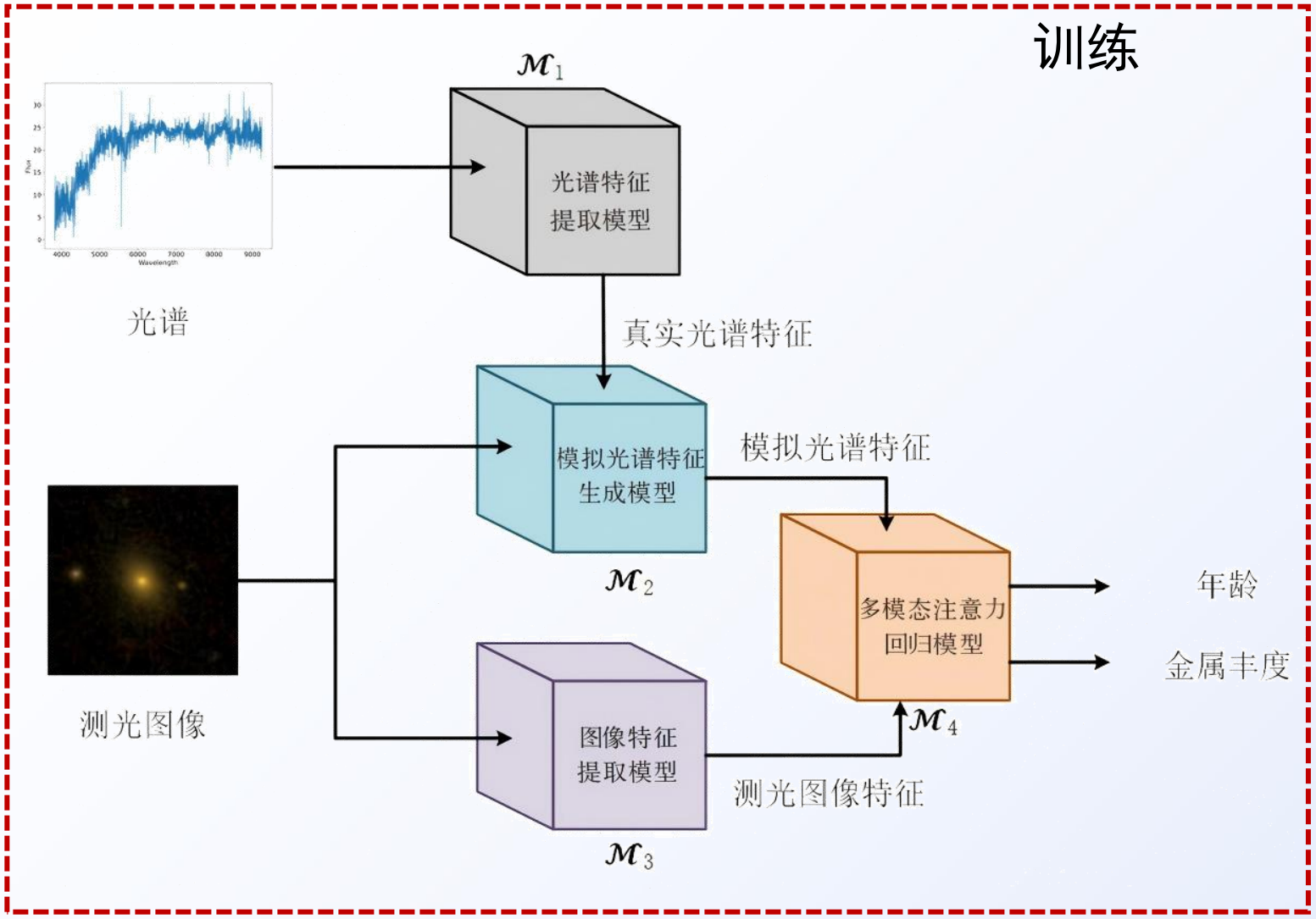
在天文观测中，许多星系有测光图像而没有相应的观测光谱。该转换模型通过从图像中生成与年龄和金属丰度相关的模拟光谱特征来弥补多模态学习中缺少光谱模态的问题。

基于多模态特征融合的星系参数回归预测

引入注意力机制对图像和光谱两种模态的特征进行融合，使其更加关注关键信息。

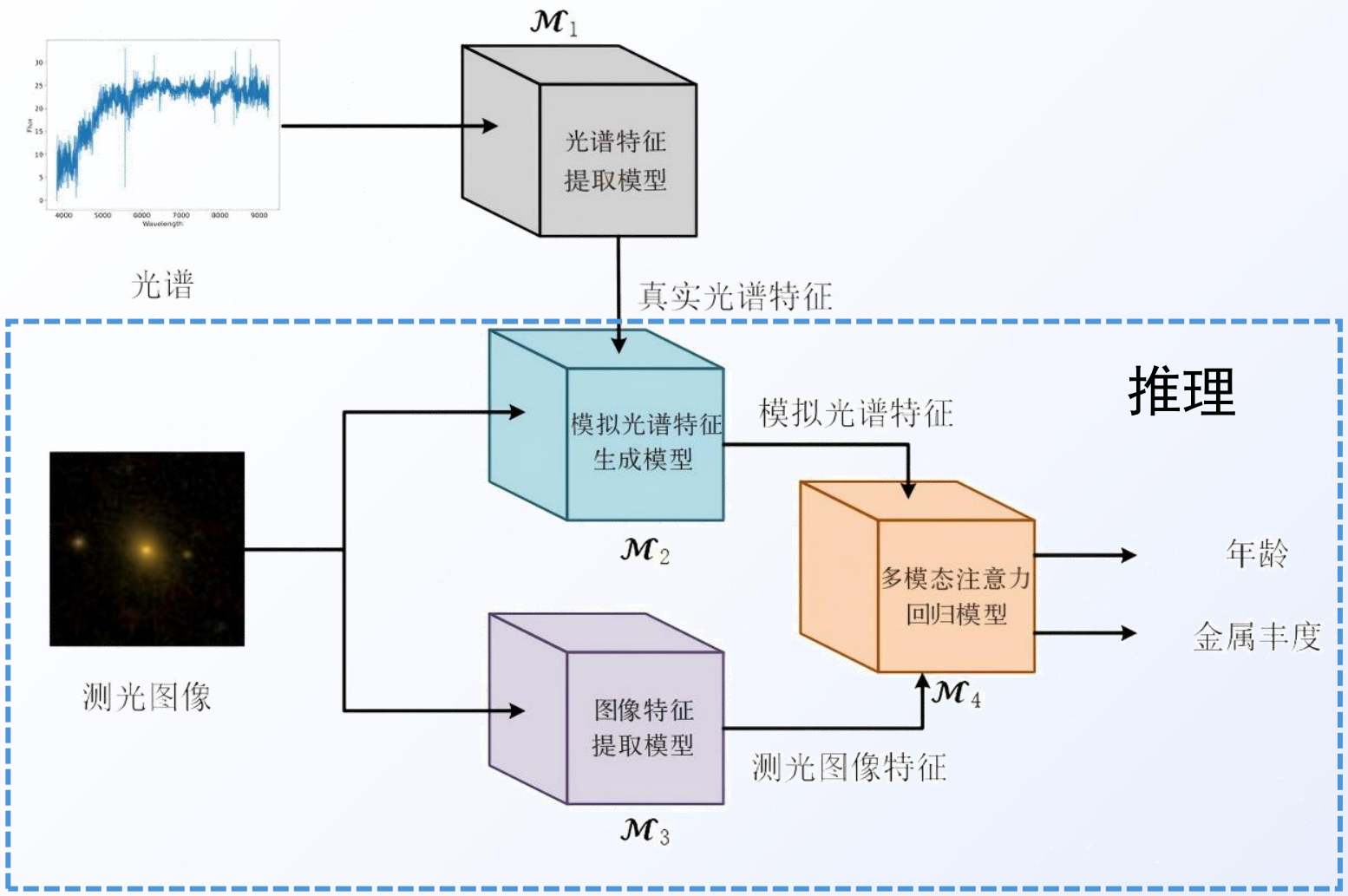
02

模型整体框架



- $\mathcal{M}_1$光谱特征提取器：光谱 → 光谱特征
- $\mathcal{M}_2$模拟光谱生成器：图像 → 模拟光谱特征
- $\mathcal{M}_3$图像特征提取器：图像 →图像特征
- $\mathcal{M}_4$多模态注意力融合：→ 年龄 & 金属丰度

模型整体框架



- $\mathcal{M}_1$光谱特征提取器：光谱 → 光谱特征
- $\mathcal{M}_2$模拟光谱生成器：图像 → 模拟光谱特征
- $\mathcal{M}_3$图像特征提取器：图像 →图像特征
- $\mathcal{M}_4$多模态注意力融合：→ 年龄 & 金属丰度

数据

多模态数据集获取

- **光谱选择——SDSS DR7**

（1）对红移范围进行限定：**0.002≤z≤0.3**，且$zWarning = 0$

（2）r波段的信噪比**S/Nr≥5**

通过上述筛选标准，最终获得了约80万条光谱。考虑到计算资源的限制，随机选取了5万条光谱用于进一步分析。

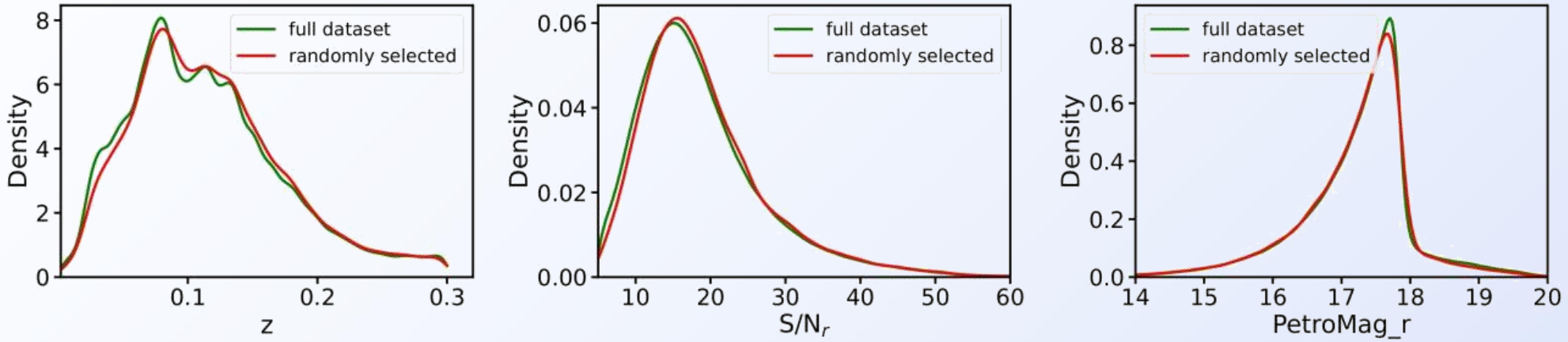


Fig. 1. Comparison of the redshift, S/N, and magnitude distributions between the randomly selected sample and the full dataset

数据

多模态数据集获取

- **图像获取**

SDSS图像由ImgCutout web服务提供jpeg格式，

图像分辨率为128×128像素。

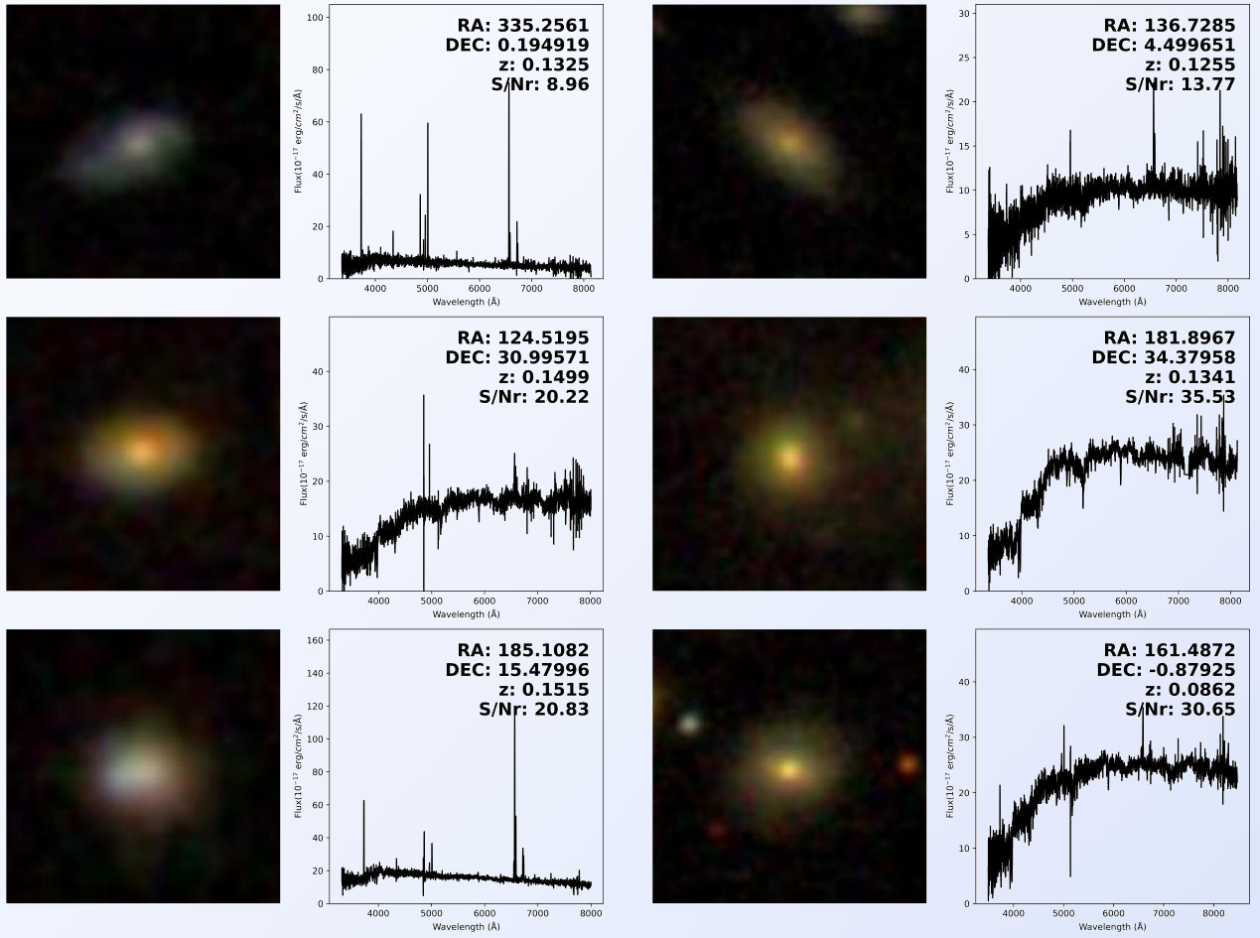


Fig. 2. Examples of photometric images and their corresponding spectra

数据

多模态数据集的预处理——图像

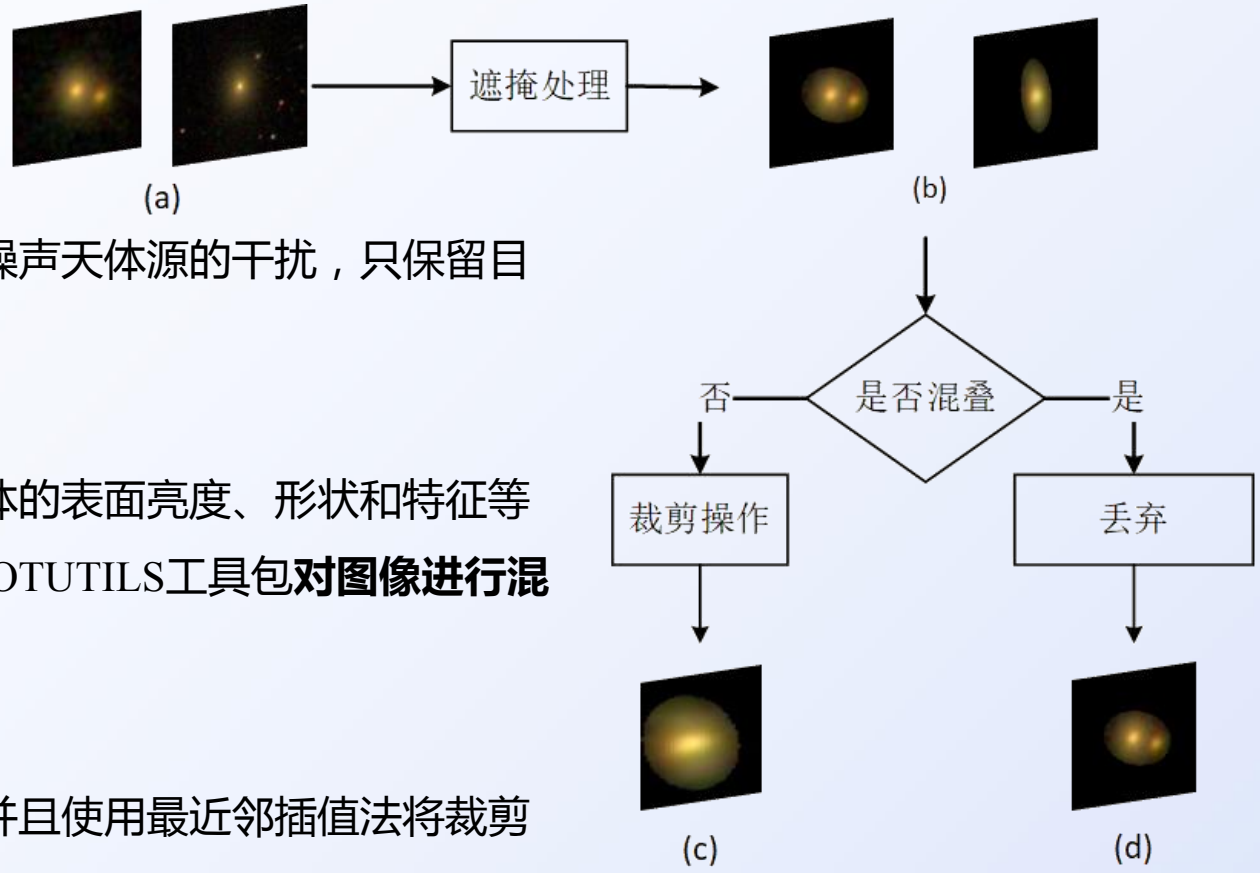


Fig. 3. Workflow of photometric image pre-processing

➢ 遮掩处理

利用自适应阈值和轮廓拟合技术对图像进行遮掩，去除其他噪声天体源的干扰，只保留目标星系。

➢ 混叠判断

混叠星系是指目标天体和其他天体相互重叠，这会把目标天体的表面亮度、形状和特征等信息掩盖掉，从而影响星系参数估计的准确性，所以使用PHOTUTILS工具包**对图像进行混叠判断与识别**，并**将混叠星系的图像丢弃。**

➢ 裁剪操作

通过检测非零像素区域的边界，仅保留含目标星系的区域。并且使用最近邻插值法将裁剪后的图像**重新调整到128×128像素的大小。**

数据

多模态数据集的标签生成—基于 pPXF 的全光谱拟合

➢ 模板库

使用 Vazdekis et al. (2010) 的 SSP 模板（36 个）

- 9 个年龄（0.06，0.12，0.25，0.5，1.0，2.0，4.0，8.0，15）
- 4 个金属丰度（−1.71，−0.71，0，0.22）

得到：星系光度加权平均的年龄、星系金属丰度

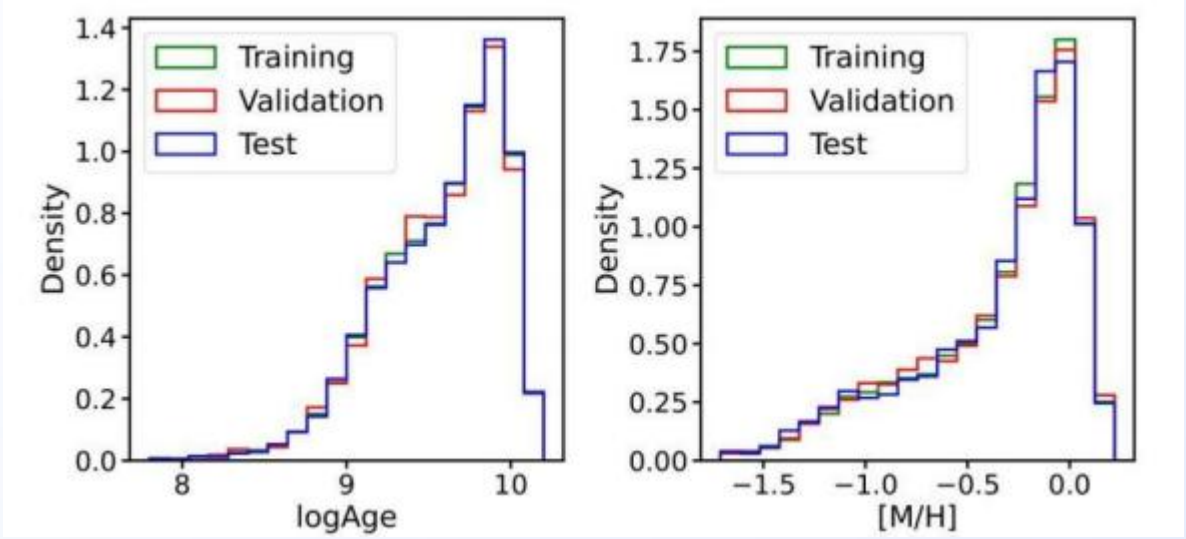


Fig. 4. Distributions of age and metallicity in the training, validation, and test sets(8:1:1)

研究方法

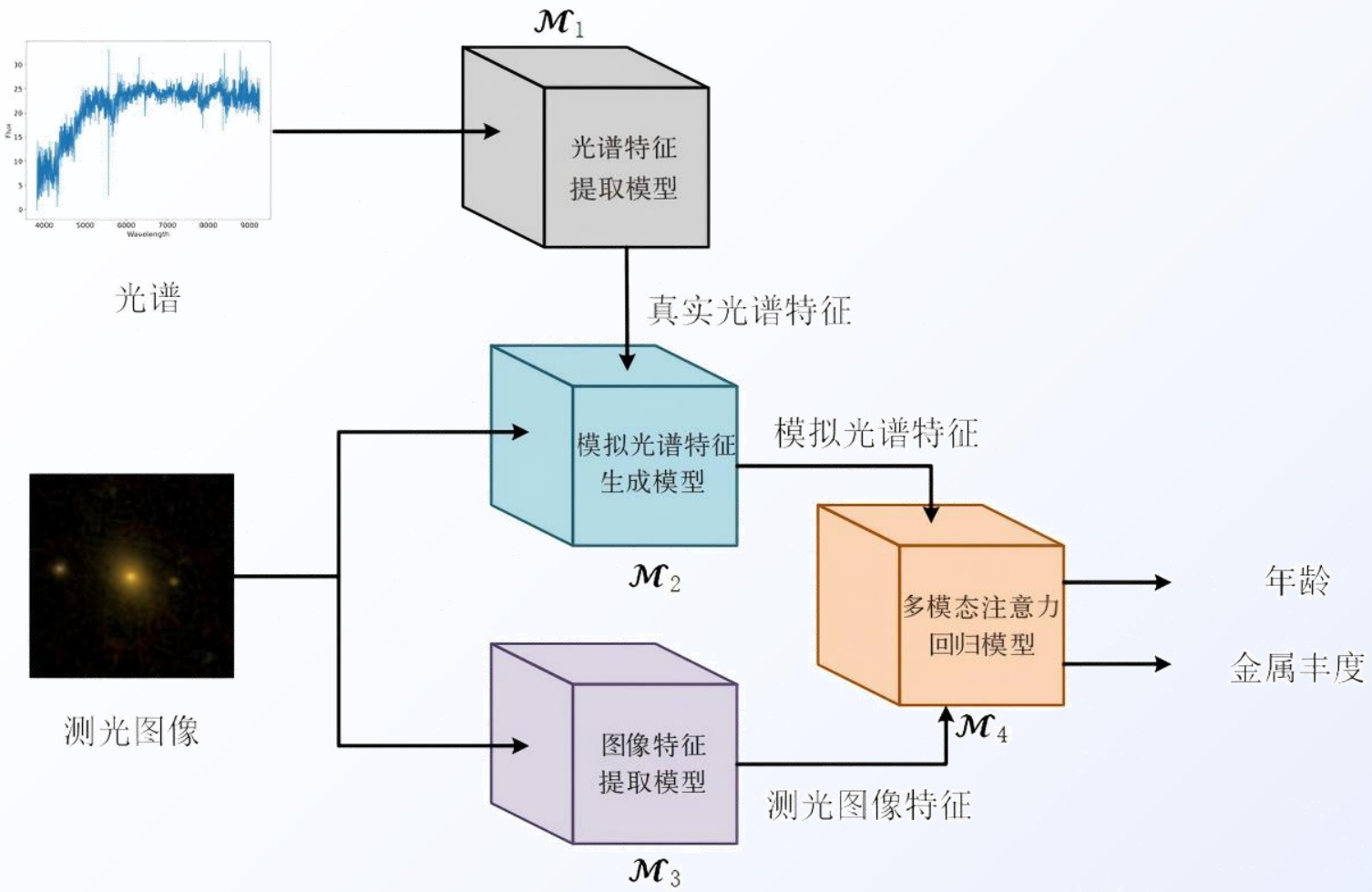


Fig. 5. Framework of the MMLforGalAM

- $\mathcal{M}_1$光谱特征提取器：光谱 → 光谱特征
- $\mathcal{M}_2$模拟光谱生成器：图像 → 模拟光谱特征
- $\mathcal{M}_3$图像特征提取器：图像 →图像特征
- $\mathcal{M}_4$多模态注意力融合：→ 年龄 & 金属丰度

研究方法

M1 —— 光谱特征提取模型

目的：M1 的任务是从真实光谱中提取与年龄金属丰度相关性强的光谱特征。这些特征将用于**监督 M2**（模拟光谱生成）。

结构：M1 是一个典型的 1D CNN + 全连接网络结构：

1. 输入：预处理后的 4171 维光谱
2. (卷积层×2 + 最大池化) ×4
3. Flatten
4. 全连接层（4171 → 1024）

输出 1024 维光谱特征向量（物理信息压缩表示），将用于训练 M2 的光谱特征。

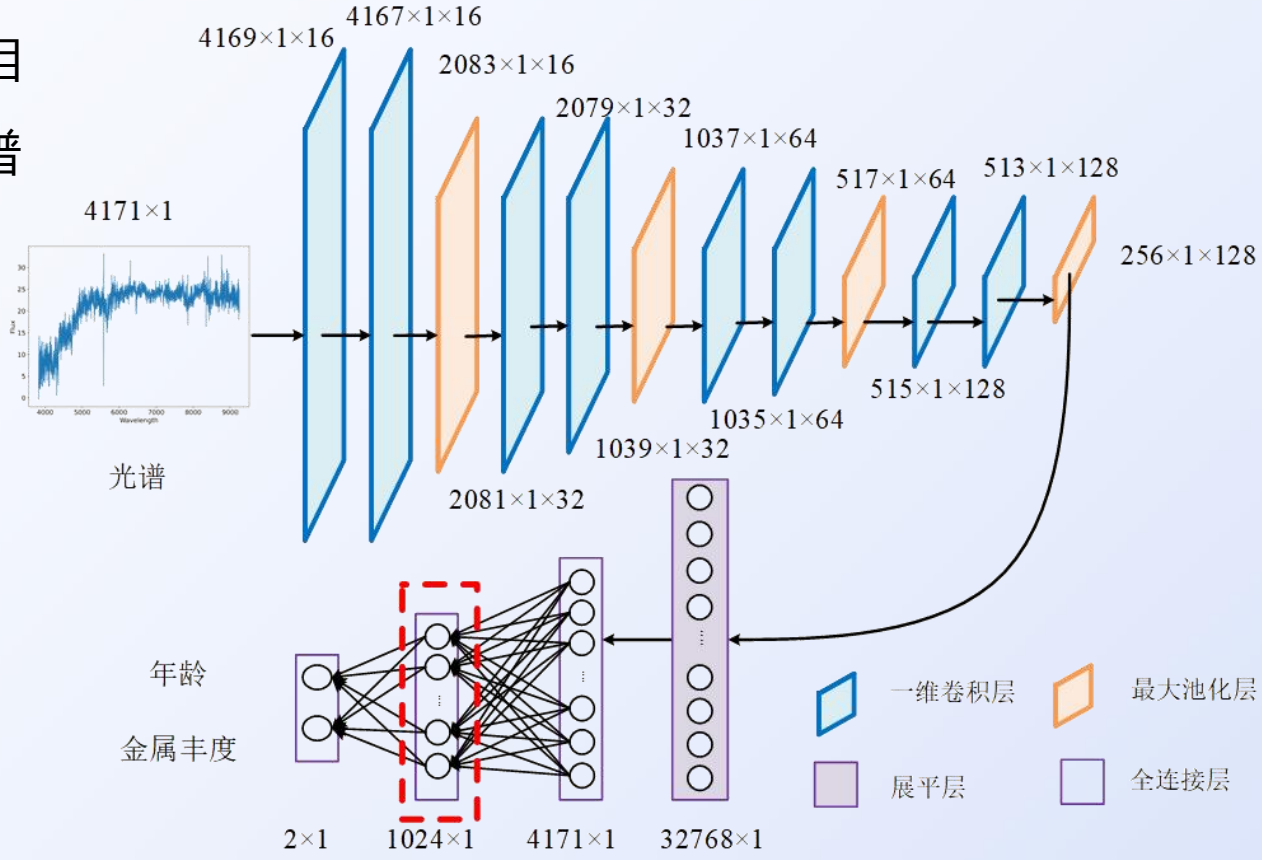


Fig. 6. Architecture of the spectral feature extraction model (M1)

研究方法

M2 —— 模拟光谱特征生成模型

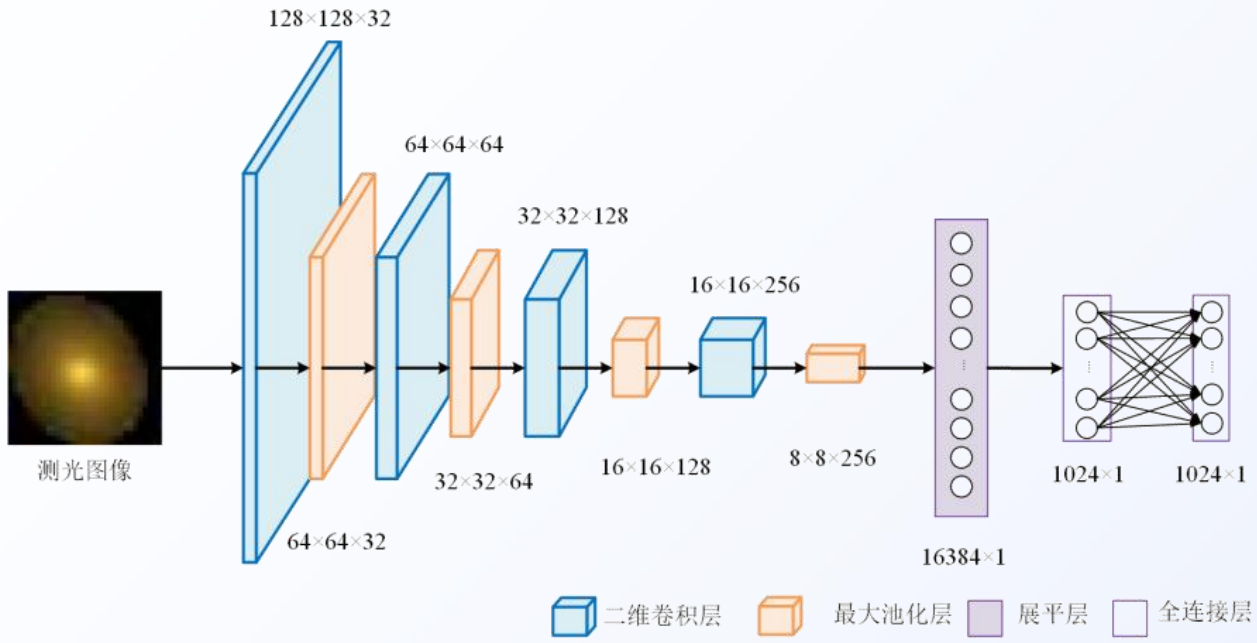


Fig. 7. Architecture of the simulated spectral feature generation mode(M2)

目的：M2 的任务是从图像中生成一个**模拟光谱特征**，可以补齐光谱模态，为多模态融合提供重要信息。

结构：M2 是图像 → 光谱特征的映射网络：

1. 输入：128×128 测光图像，M1 的光谱特征作为标签
2. (2D 卷积 + 池化)×4
3. Flatten 展平
4. 全连接层
5. 输出：1024 维模拟光谱特征，与 M1 提取的真实光谱特征在同一向量空间。

研究方法

M3 —— 图像特征提取模型

目的：图像特征提取，用作多模态融合的输入。

结构：

1. 输入：128×128 图像
2. 卷积块 × 4（每块含 3 个 3×3 卷积）
3. 最大池化降低空间维度
4. Flatten
5. 全连接层（4171 → 1024）

得到 1024 维图像特征向量

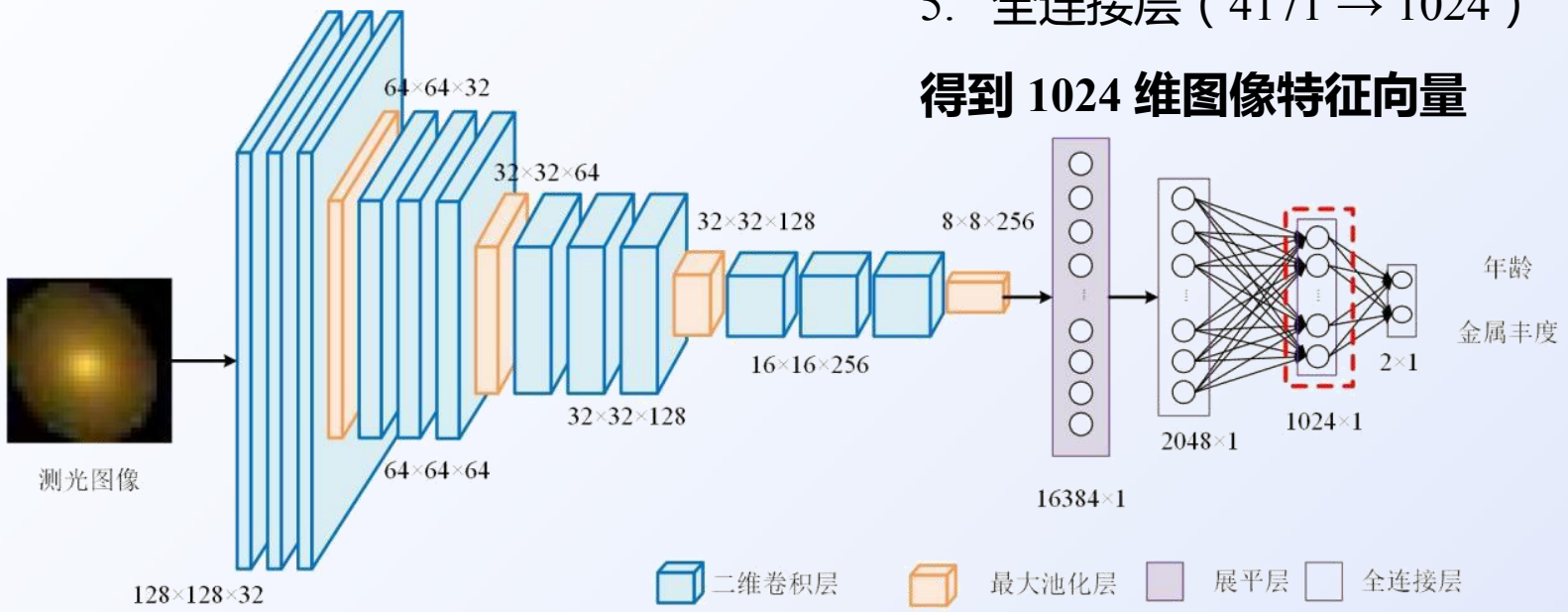


Fig. 8. Architecture of the image feature extraction model (M3)

研究方法

M4 —— 多模态注意力回归模型

模型采用基于**多头注意力机制**的加权融合策略。

该机制通过并行运行多个注意力头，使模型能够在输入特征的不同子空间中捕获注意力分布。

结构：

1. Query = 模拟光谱特征
2. Key / Value = 图像特征
3. 融合特征 → 全连接网络

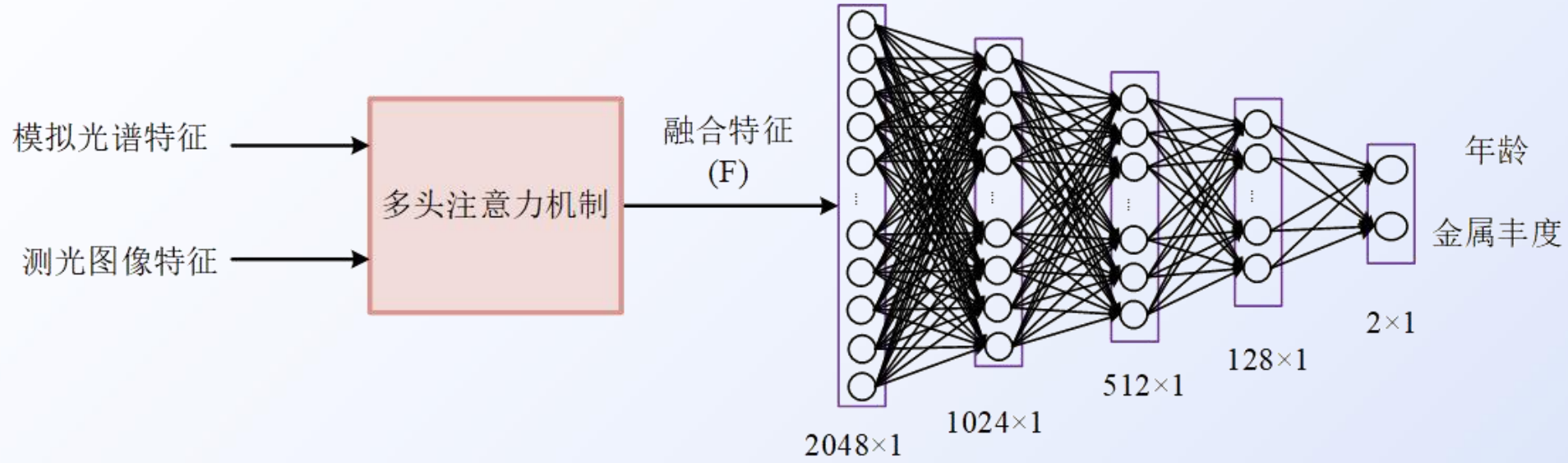
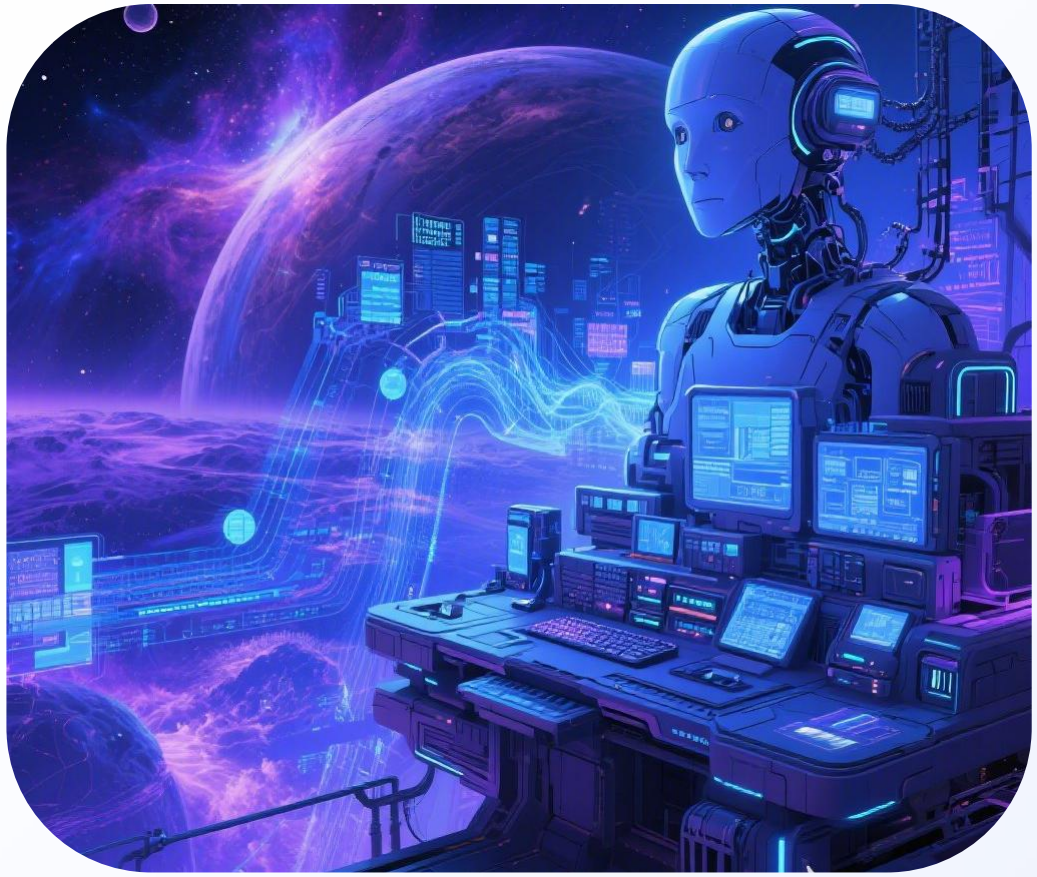


Fig. 9. Architecture of the multi-modal attention regression model(M4)



03

研究结果

- 模型的预测效果评估
- 模拟光谱特征生成的效果评估
- 多模态与单模态的比较

结果分析

● 多模态模型的预测值与真实值的对比

Table 1. Comparison of age and metallicity prediction among different models

模型	MSE		MAE		Bias(SD)	
	年龄	金属丰度	年龄	金属丰度	年龄	金属丰度
仅图像模型	0.04	0.03	0.168	0.13	0.042(0.21)	0.012(0.164)
仅光谱模型	0.019	0.016	0.113	0.102	0.036(0.14)	-0.006(0.127)
MMLforGalAM	0.02	0.02	0.1215	0.112	0.003(0.15)	-0.009(0.14)
MMLforGalAMwithReal	0.017	0.016	0.106	0.104	-0.009(0.13)	-0.017(0.127)

MMLforGalAM模型对年龄和金属丰度的预测值与真实值基本一致，标准差分别为0.15dex和0.14dex。

结果分析

模拟光谱特征生成模型的准确性验证

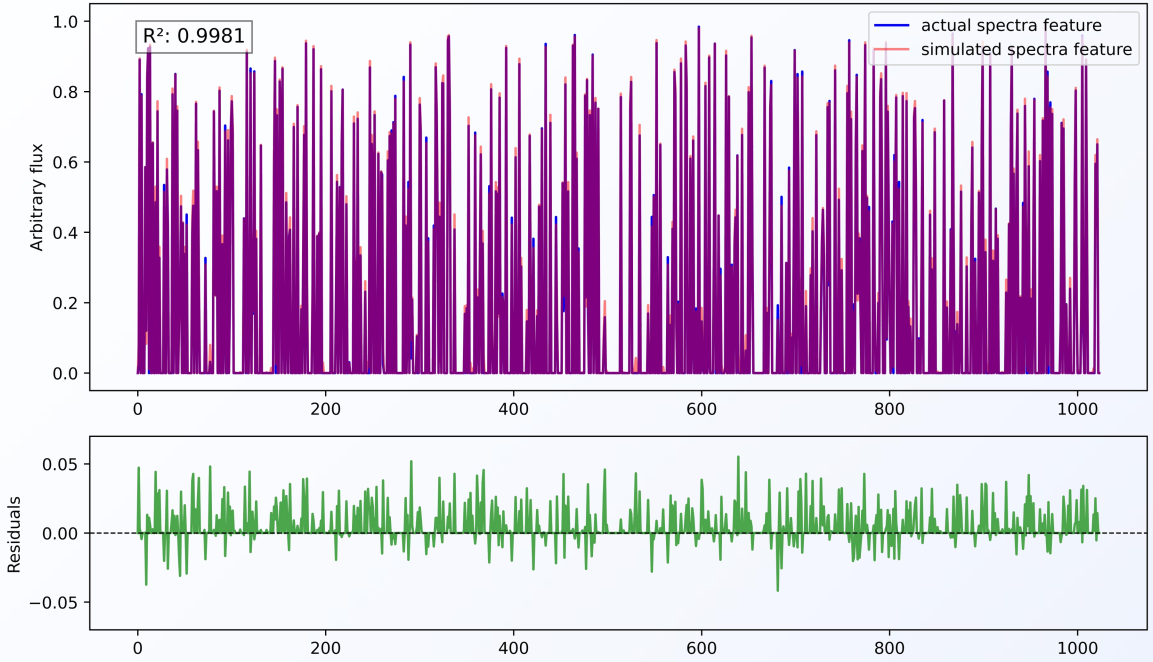


Fig. 10. Example of spectral features generated by model M2 and true spectral features extracted from observed galaxy spectra

示例图：二者之间的残差在[-0.05，0.05]

定量评估：

$$R^2 = 1 - SS_res / SS_tot$$

使用test set中的图像，得到R^2=0.97

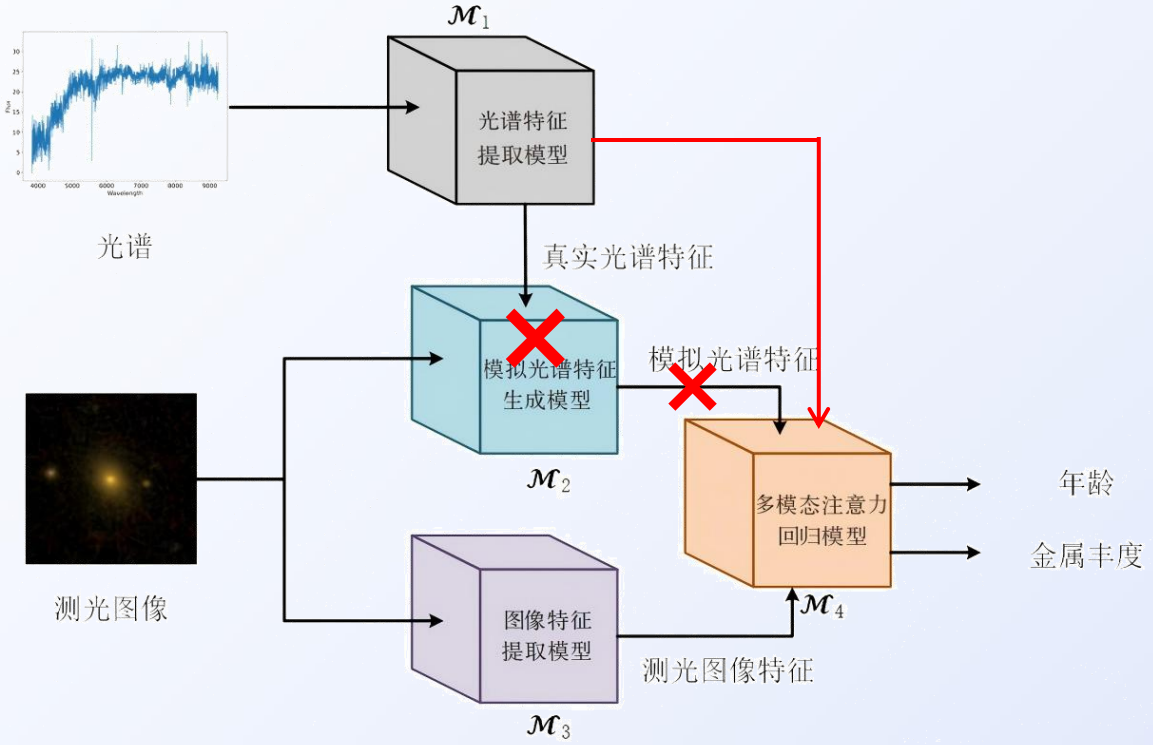
结果分析

模拟光谱特征生成模型的准确性验证

多模态：真实光谱特征+图像特征。

学习框架：直接输入提取的真实光谱特征与图像特征进行多模态融合（称为**MMLforGalAMwithReal**）。

结论：使用模拟光谱特征的多模态与使用真实光谱特征的多模态预测效果略有下降。



模型	MSE		MAE		Bias(SD)	
	年龄	金属丰度	年龄	金属丰度	年龄	金属丰度
仅图像模型	0.04	0.03	0.168	0.13	0.042(0.21)	0.012(0.164)
仅光谱模型	0.019	0.016	0.113	0.102	0.036(0.14)	-0.006(0.127)
MMLforGalAM	0.02	0.02	0.1215	0.112	0.003(0.15)	-0.009(0.14)
MMLforGalAMwithReal	0.017	0.016	0.106	0.104	-0.009(0.13)	-0.017(0.127)

结果分析

● 多模态模型与单模态模型的对比

单图像模态：训练一个仅图像模态的模型，该模型的网络结构及超参设置与M3一致。

测试结果表明：与仅图像模型相比，多模态的年龄预测误差降低了27%，金属丰度预测误差降低了15%，这表明图像特征融合了光谱特征，提高了从图像中预测年龄和金属丰度的精度。

单光谱模态：训练一个仅光谱模态的模型，该模型的网络结构及超参设置与M1一致。

测试结果表明：仅光谱模型的预测效果略优于多模态模型。

模型	MSE		MAE		Bias(SD)	
	年龄	金属丰度	年龄	金属丰度	年龄	金属丰度
仅图像模态模型	0.04	0.03	0.168	0.13	0.042(0.21)	0.012(0.164)
仅光谱模态模型	0.019	0.016	0.113	0.102	0.036(0.14)	-0.006(0.127)
MMLforGalAM	0.02	0.02	0.1215	0.112	0.003(0.15)	-0.009(0.14)
MMLforGalAMwithReal	0.017	0.016	0.106	0.104	-0.009(0.13)	-0.017(0.127)

结果分析

模型使用建议

提供两个训练好的模型

M1：有光谱 → 使用单光谱模型
（精度高且时间复杂度低）

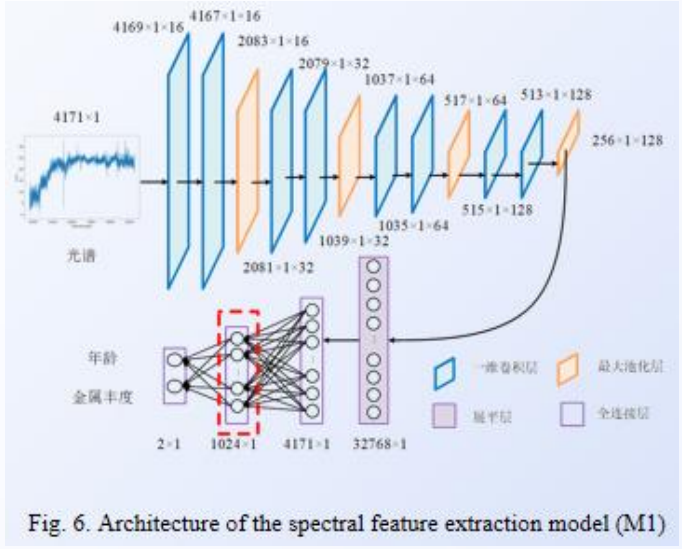
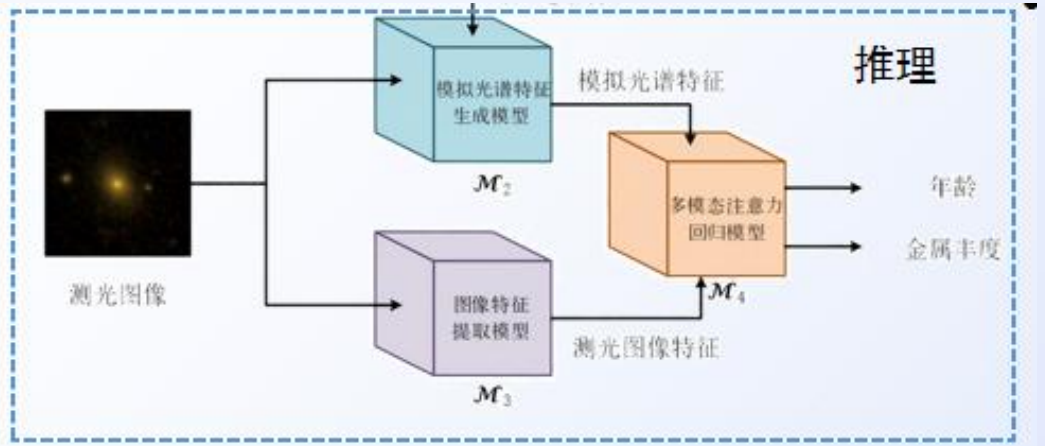


Fig. 6. Architecture of the spectral feature extraction model (M1)

MMLforGalAM：只有图像





04

总结

总结

研究目标

提出一种多模态深度学习框架，融合图像特征与光谱特征预测年龄和金属丰度，解决使用图像预测参数精度不高的问题。

核心方法：MMLforGalAM

- ✓ M1：从光谱中提取年龄与金属丰度相关特征。
- ✓ M2：基于图像生成模拟光谱特征，弥补没有光谱的问题。
- ✓ M3：从测光图像中提取物理参数相关特征。
- ✓ M4：通过注意力机制融合图像与模拟光谱特征，进行最终回归预测。

下一步工作

- ✓ 框架拓展：更多物理参数估计；其他巡天数据
- ✓ 物理参数相关的图像特征重要性分析

德州学院
DEZHOU UNIVERSITY

谢谢！

王丽丽
jsjxwll@126.com