

深度学习中的星系图像特征提取

沈世银

中国科学院上海天文台

2025 CSST暑期学校 北京大学 07/05

提纲

- 天文图像
- 深度学习简介
- 深度学习中星系特征的提取
 - 已有研究案例
- 天文研究的需求

星系天文学的需求：既要又要

- 更远，更早期（时间跨度）
 - 目前詹姆斯韦伯望远镜看到最远的星系：
- 更多（空间尺度）
 - 天区更大：大样本巡天
 - 统计的power
- 看得更清楚（空间分辨率）
 - 更多的细节
- 看的波段更多（光谱维度）

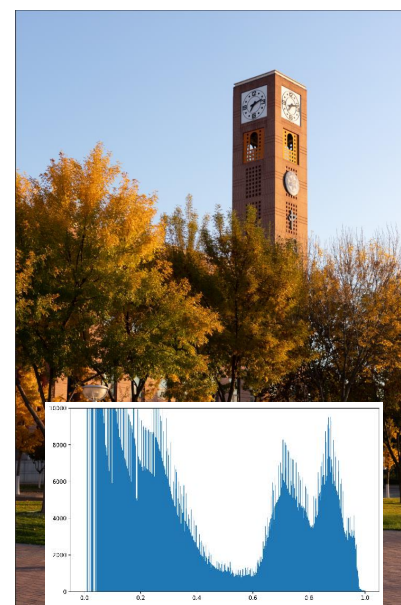
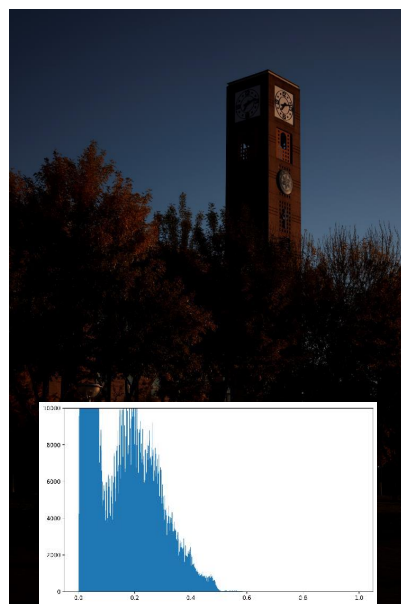
星系天文学的挑战

- 低信噪比数据中信号的探测
 - 不浪费任何一个光子
 - 引入比较强的prior
- 海量数据的处理和存储 (CSST~100PB)
 - 代替人工 (如传统的形态分类)
 - GPU加速 (传统的模型估计)
 - 压缩即智能?
- 三维数据 (图像+光谱) 中高维信息的提取
 - 数据驱动模型 (知识) 构建

(可见) 天文图像

- 单一能段通道上对光子数的直接记录
 - CCD: 光子 $\rightarrow$ 电子 $\rightarrow$ ADU, 动态范围: ~ 16 位
 - 可以用3个能段的FITS文件, 进行非线性变换之后转化为日常的PNG图像

AI的很多CV算法是否只对日常图像才适用? 在天文上是否还能使用?



天文图像的特点

- 共同点：图像中不同像素之间存在相关（信息冗余）
 - 可以认为不同像素点上的噪声是相互独立的（不可预测性）
- 低信噪比
- 仪器效应（PSF）
 - 不同仪器效应不同

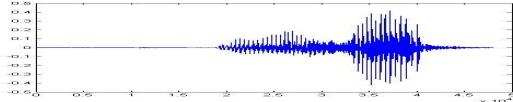
深度学习：模仿人类原始的学习过程

人类的原始学习	深度学习
技能的获得过程	特征（模式）的提取过程
反馈机制	反馈机制
一个不断纠错的过程	损失下降的过程
学会接到一个棒球	成功的进行人脸识别
没有先验的知识	不需要先验模型
重力场抛物线的知识	眼睛，鼻子的概念



学习的本质 $\approx$ 寻找映射关系

- 语音辨识

$$f(\text{  }) = \text{"How are you"}$$

- 图像辨识

$$f(\text{  }) = \text{"Cat"}$$

学习是把一个不熟悉的概念装换（映射）为熟悉概念的过程

如何建立映射关系？

1. 用很多的参数构建一个很大自由度的函数
2. 给一组参数，用这个函数算一下结果，看看结果好不好（猜）
3. 如果结果不好，调整这组参数，回到2

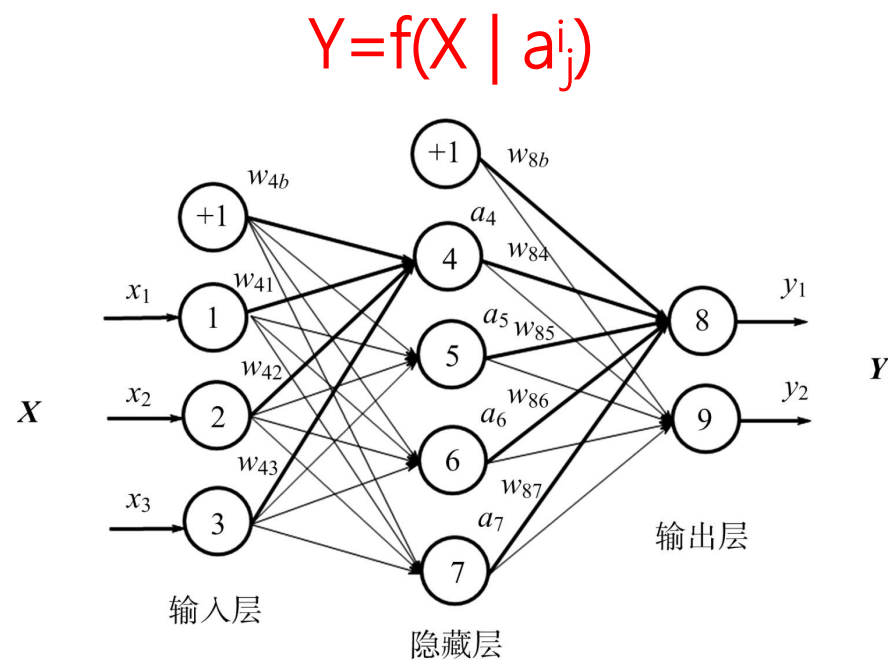
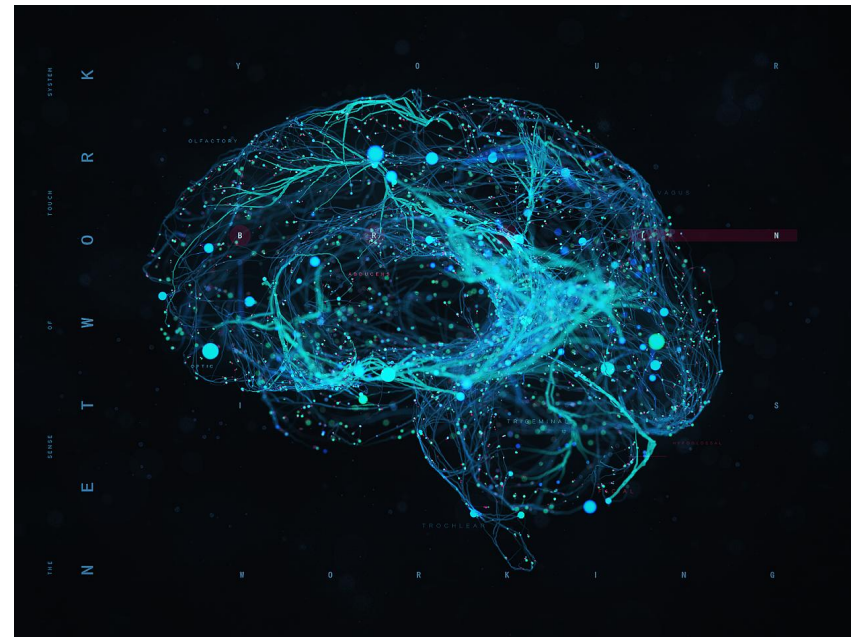


如何构建这个映射函数？

1. 神经网络（模仿人类大脑神经元）

- 网络：简单功能的堆叠
 - 非线性过程（激活函数）
- → 超强表达力
- 数学上可导

网络的架构：基本与天文学家无关



$y=f(x)$ 当中, $y(\text{label})$ 和 $x(\text{image})$ 已知, 如何构造映射函数? Or 如何贴近真实情况?

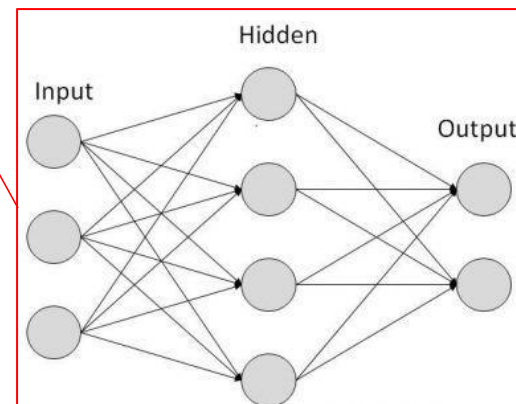
神经网络(Neural Networks)

- 向前传播
- 构建损失函数
- 损失函数修正模型 (误差反向传播)

$f($



$) = \text{"Cat"}$



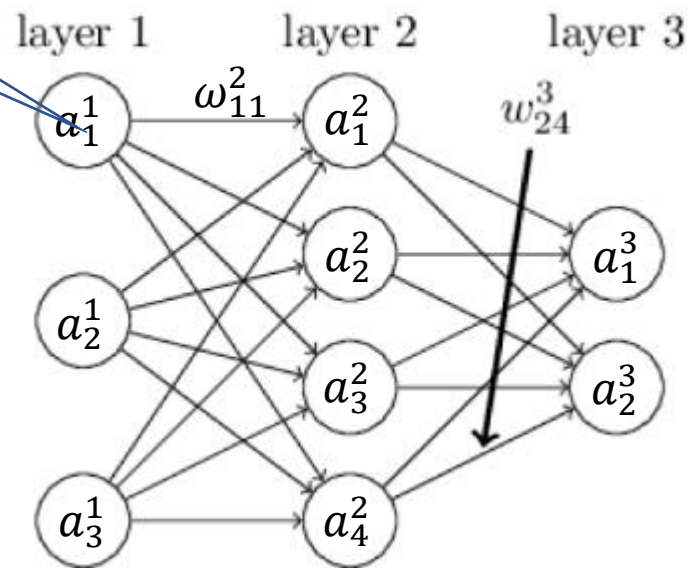
神经网络的构成

图像某个像素的值

记 w_{jk}^l 为第 $l-1$ 层第 k 个神经元到第 l 层第 j 个神经元的权重

b_j^l 为第 l 层第 j 个神经元的偏置

a_j^l 为第 l 层第 j 个神经元的激活值



全连接层, Fully Connected

第 l 层的值取决于 $l-1$ 层神经元的构建:

$$a_j^l = \sigma \left(\sum_k w_{jk}^l a_k^{l-1} + b_j^l \right)$$

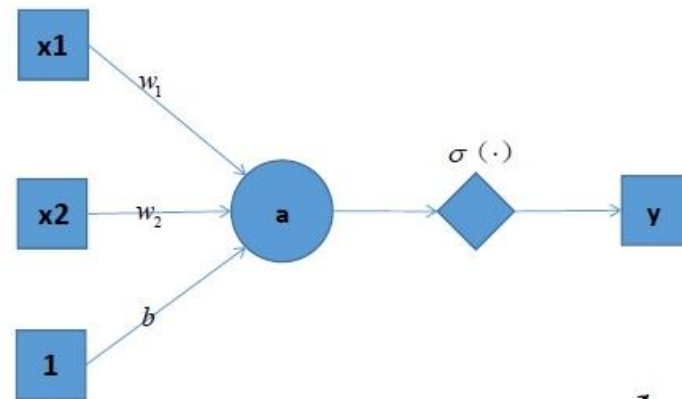
写成计算机能够理解的矩阵运算: $a^l = \sigma(w^l a^{l-1} + b^l)$

• 最简单的神经网络

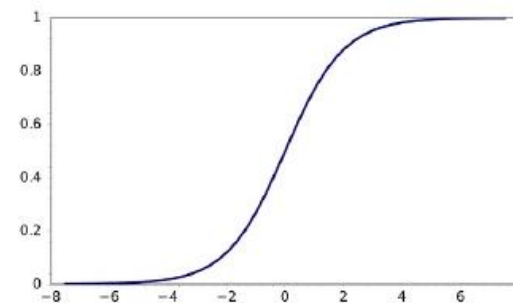
- 近似于一个线性变换
- 每一层的输出可以是任意大小
- 同一层中, 神经元是等价的

激活函数

- 引入非线性过程，增加网络的表达能力
- 模拟大脑神经元的激活过程
 - 降噪
- 有多种形式
- 没有自由参数



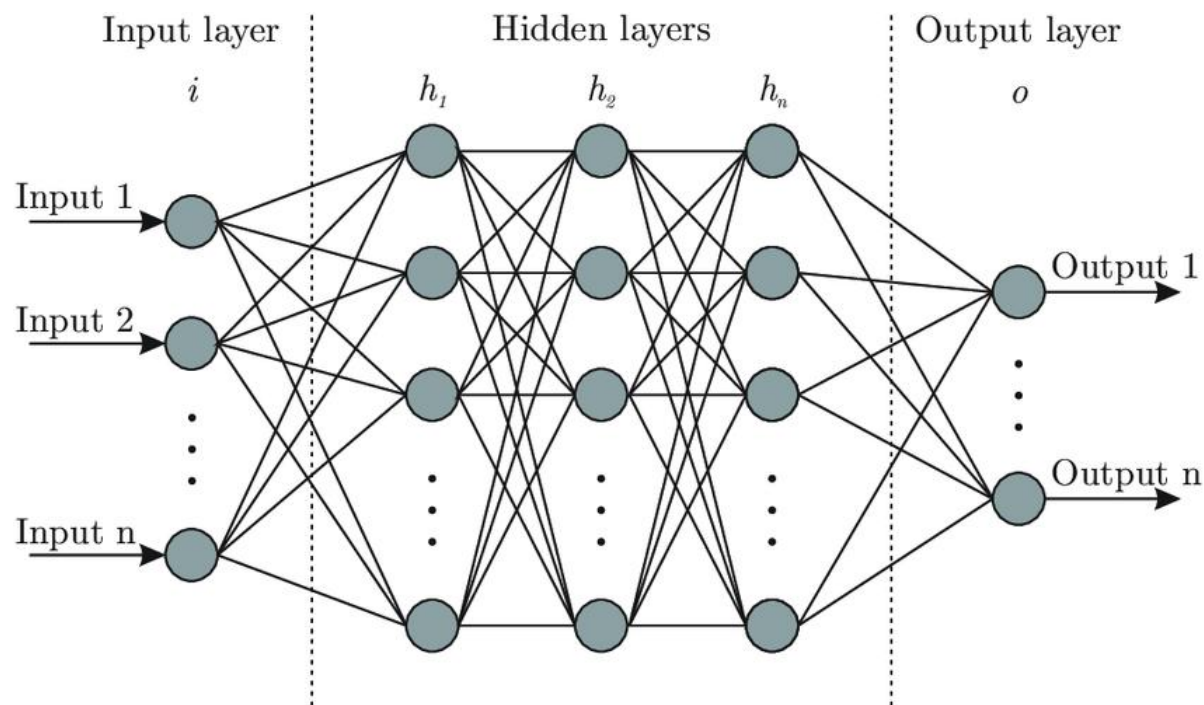
$$a = w_1 x_1 + w_2 x_2 + b$$
$$y = \sigma(a)$$



$\sigma(\cdot)$ is a non-linear activation function, sigmoid was the most popular one,

$$\sigma(y) = \frac{1}{1 + e^{-y}}$$

- 神经网络由神经元之间的关系所定义，其具体形状是自由的
 - 全连接网络由 w 和 b 的组合（权重）所构成
- 输入（图像）和输出（标签）改变网络的形状，但是不是网络的一部分



2. 猜得对不对：损失函数

```
a = loss_function(Y_true, Y_pred)  
    = loss_function(Y_true, f(X_true))
```

- 无监督 VS 有监督学习：没有本质区别
- 损失函数中包含了问题的本质：你想学到什么？

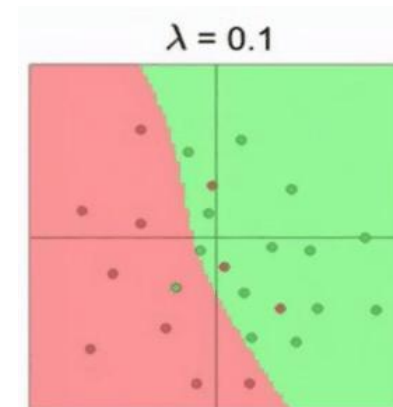
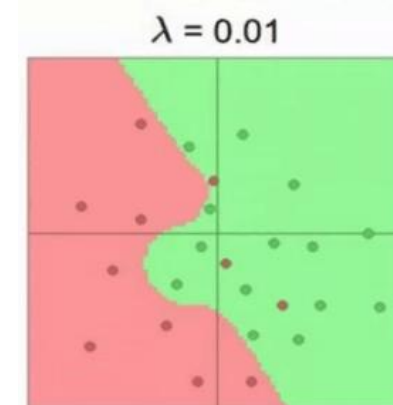
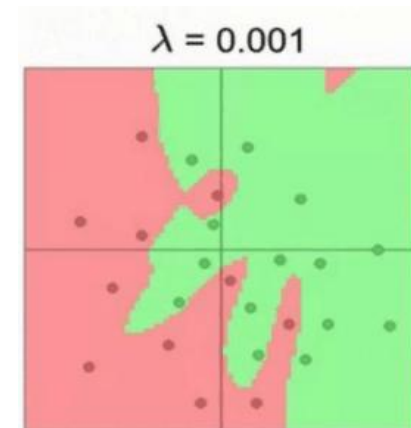
Question：损失函数中的变量是什么？



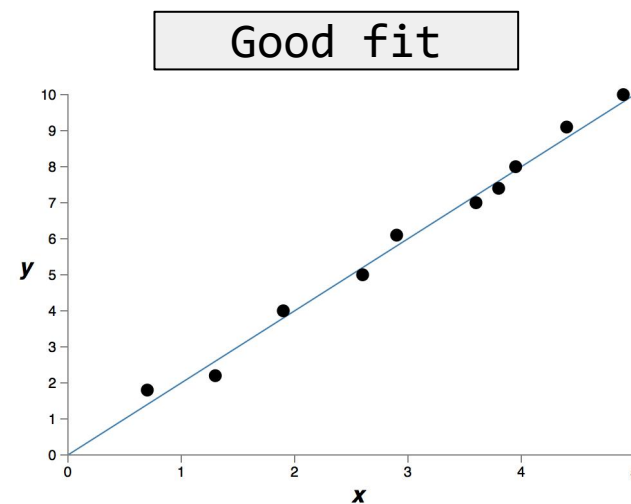
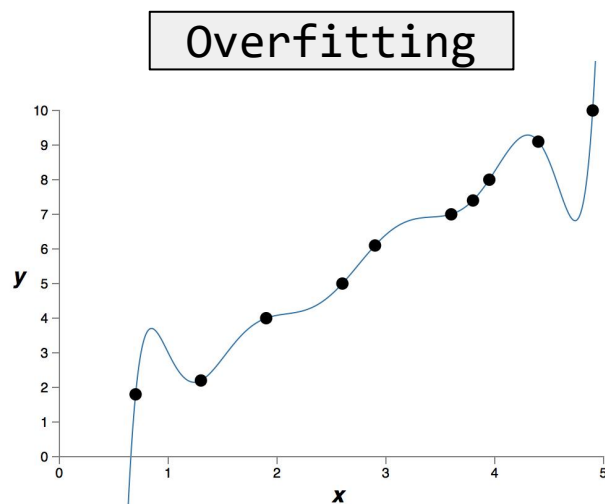
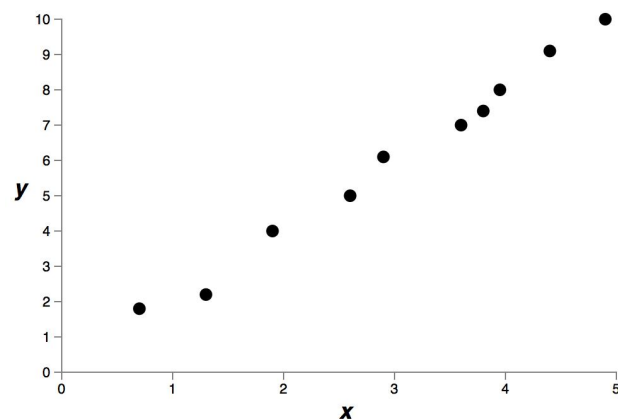
损失函数： 学习过程中的关键要素

- 损失函数由学习者（甲方）确定
- 统计学知识
 - 类似于最大似然，最小Chi²
 - 如何定义两张图片（两种统计分布）的相似度？
 - KL散度
- 损失函数中的额外信息
 - 防止过拟合：惩罚项
 - 网络的自由度可能是无限大
 - 模型的Prior：比如特征量服从某种假设的分布

$$Loss = \frac{1}{n} \sum_x \|y_{true} - y_{pred}\|^2 + \frac{\lambda}{2n} \sum_{w \in Weight} \|w\|_2^2$$



如何限制过拟合 (*overfitting*)



$$f(x) = w_1x^9 + w_2x^8 + \dots + w_8x^2 + w_9x^1 + w_{10}x^0$$

在上图中有十个点，这意味着我们可以找到一个9次多项式完美拟合！
但是我们也可以使用线性模型 $y=2x$ 来得到一个较好的拟合。
如何限制拟合条件呢？

3. 调整参数使得损失下降!

- 如何调整?

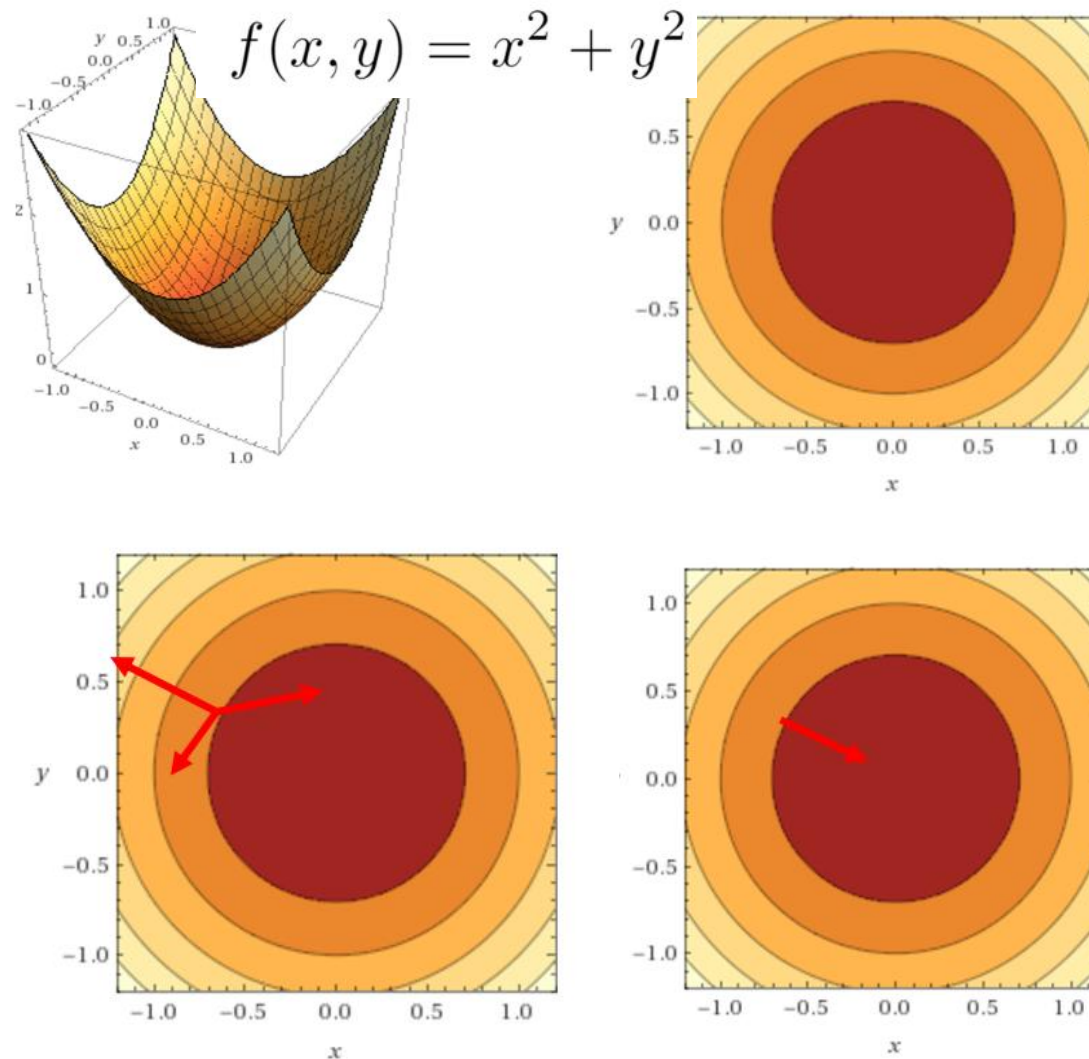
- 本质一种反馈机制

X 随机寻找 (MCMC)

- 参数太多!!!

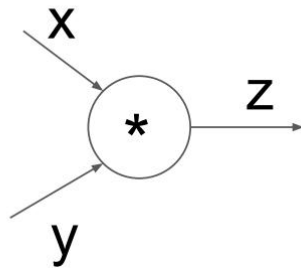
$$x_{t+1} = x_t - \alpha \nabla f(x_t)$$

```
while True:
    dx = compute_gradient(x)
    x -= learning_rate * dx
```



深度学习框架：自动求导与模块化

- Pytorch, Tensorflow
- 网络单元层 (layer) : 模块化
 - 支持导数的反向传递



(x,y,z are scalars)

From Fei-Fei Li, CS231

```
class Multiply(torch.autograd.Function):  
    @staticmethod  
    def forward(ctx, x, y):  
        ctx.save_for_backward(x, y)  
        z = x * y  
        return z  
    @staticmethod  
    def backward(ctx, grad_z):  
        x, y = ctx.saved_tensors  
        grad_x = y * grad_z # dz/dx * dL/dz  
        grad_y = x * grad_z # dz/dy * dL/dz  
        return grad_x, grad_y
```

Need to stash
some values for
use in backward

Upstream
gradient

Multiply upstream
and local gradients

深度学习：GPU运算

- 网络的正向计算过程比较简单
- 大量的矩阵运算：适合并行
- 网络训练需要大的训练样本：
并行计算



Numpy

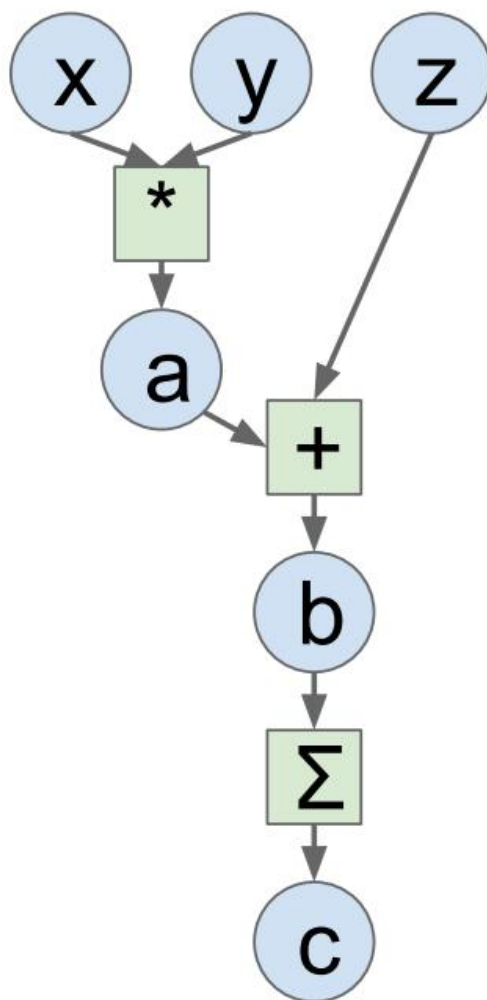
```
import numpy as np
np.random.seed(0)

N, D = 3, 4

x = np.random.randn(N, D)
y = np.random.randn(N, D)
z = np.random.randn(N, D)

a = x * y
b = a + z
c = np.sum(b)

grad_c = 1.0
grad_b = grad_c * np.ones((N, D))
grad_a = grad_b.copy()
grad_z = grad_b.copy()
grad_x = grad_a * y
grad_y = grad_a * x
```



PyTorch

```
import torch
device = 'cuda:0'
N, D = 3, 4
x = torch.randn(N, D, requires_grad=True,
                 device=device)
y = torch.randn(N, D, device=device)
z = torch.randn(N, D, device=device)

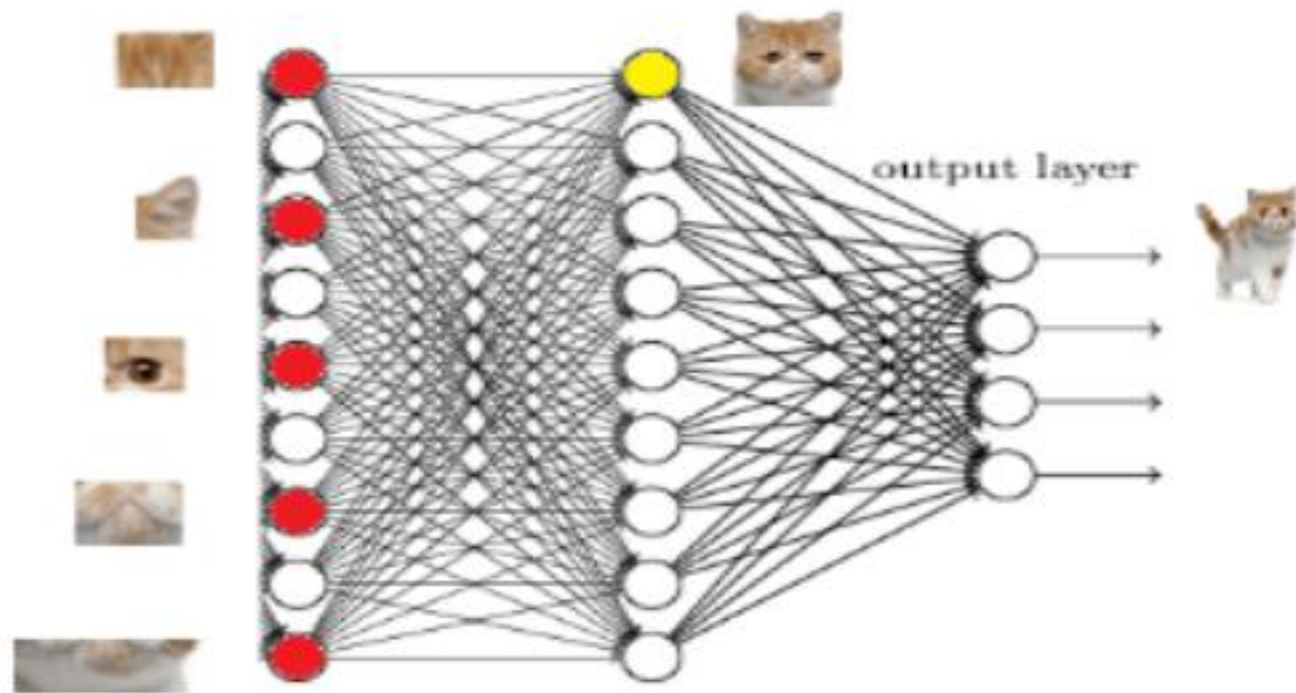
a = x * y
b = a + z
c = torch.sum(b)

c.backward()
print(x.grad)
```

Trivial to run on GPU - just construct arrays on a different device!

4. 复杂神经网络模型

- 图像不是简单像素的总和
 - 轮廓特征（卷积核提取）
 - 组合特征（多层网络）
- 卷积神经网络（CNN）
- Transformer (self-attention)
- 主要由AI学者开发



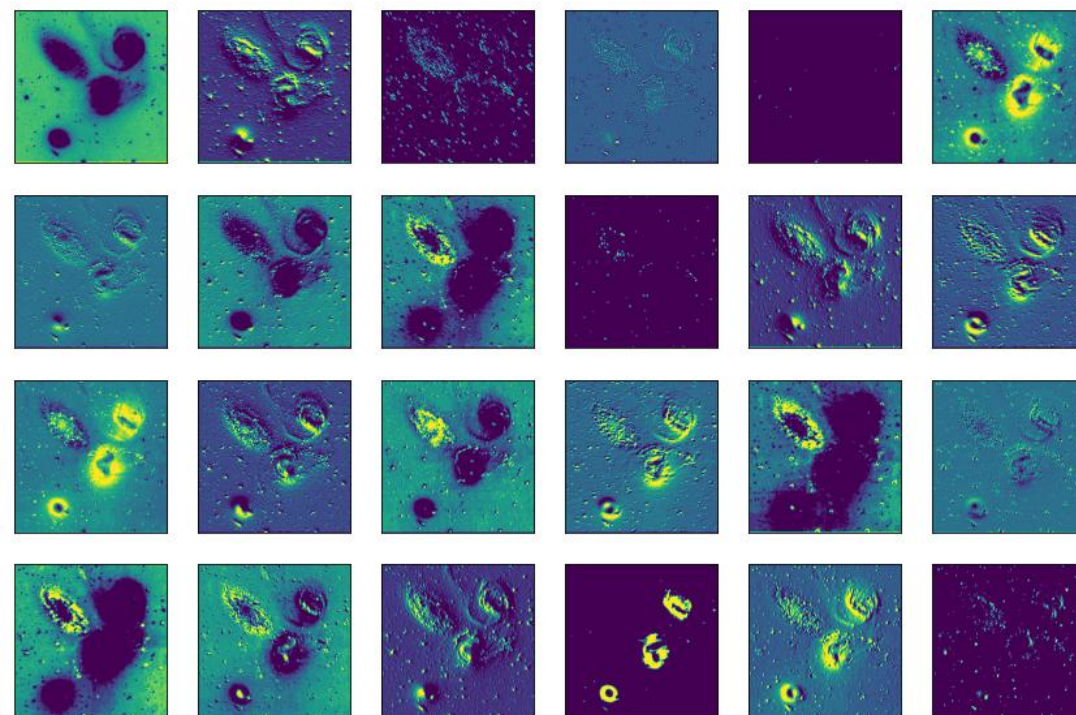
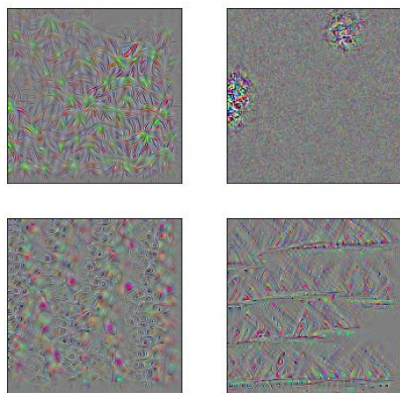
星系假色图经过卷积神经网络提取特征

分类或检测: 用探测特定纹路的卷积得到对应的特征响应程度，表征在图像上即为对应的**特征图**。

卷积神经网络是目前主流的图像分类和识别方法，可以有效地找寻和探测图像中的细节特征。

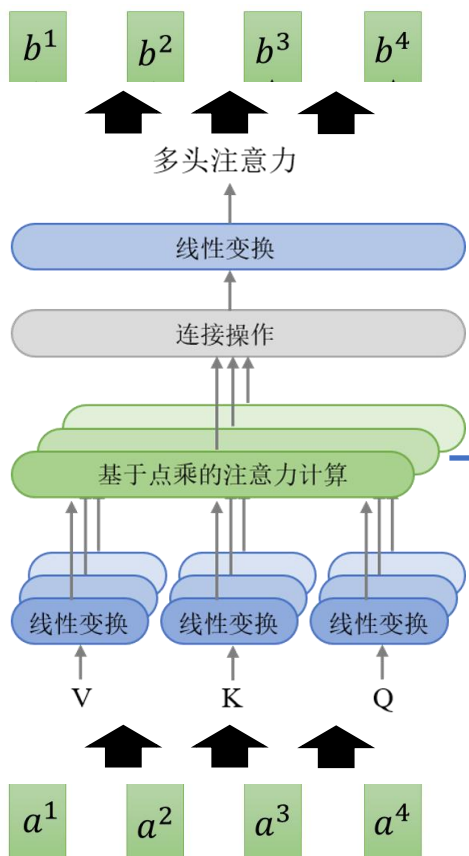


卷积: 经过**特定纹路的卷积核**得到的**特征图**

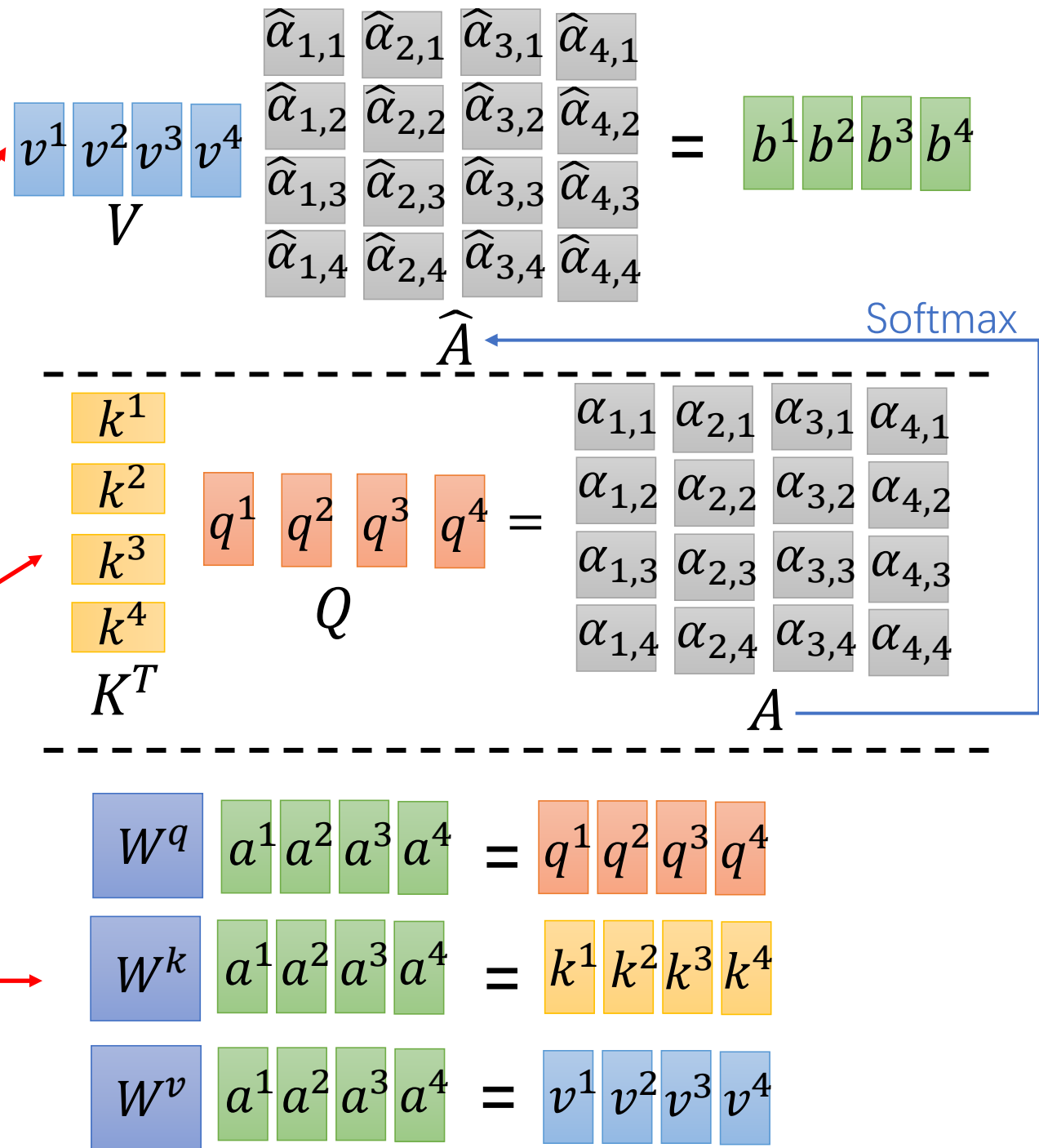
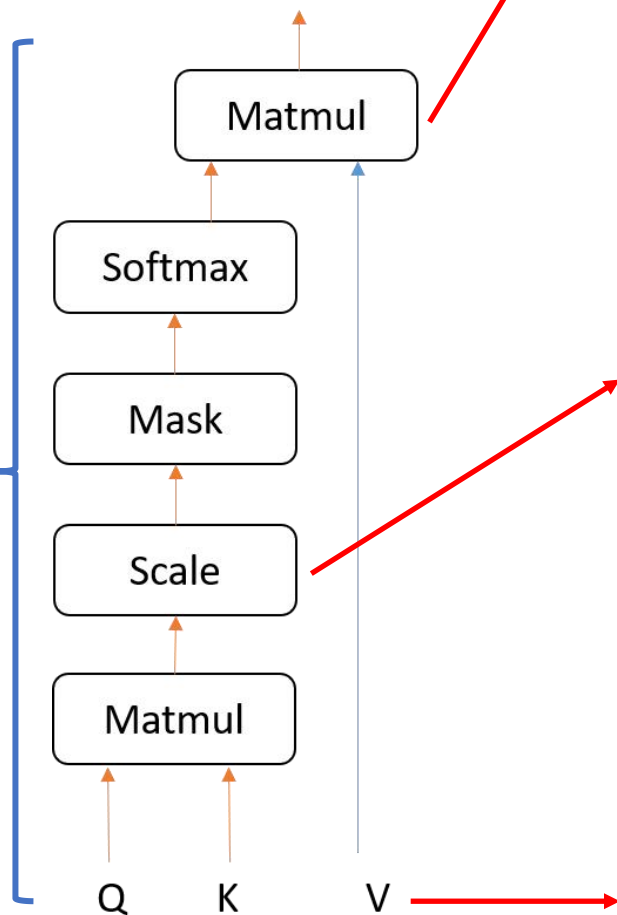


Transformer

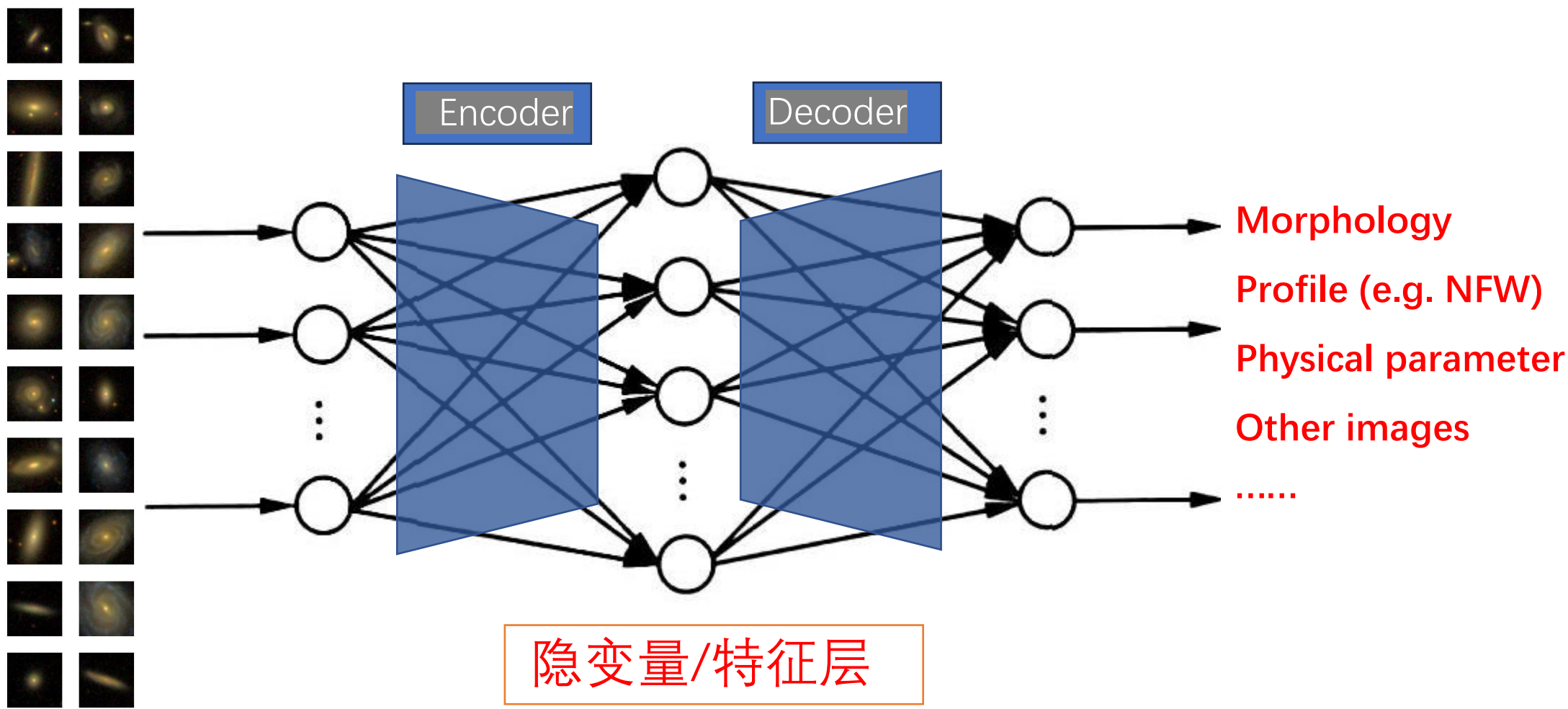
算法核心：自相关



Self-attention



计算机视觉模型的基础架构



隐变量/特征层

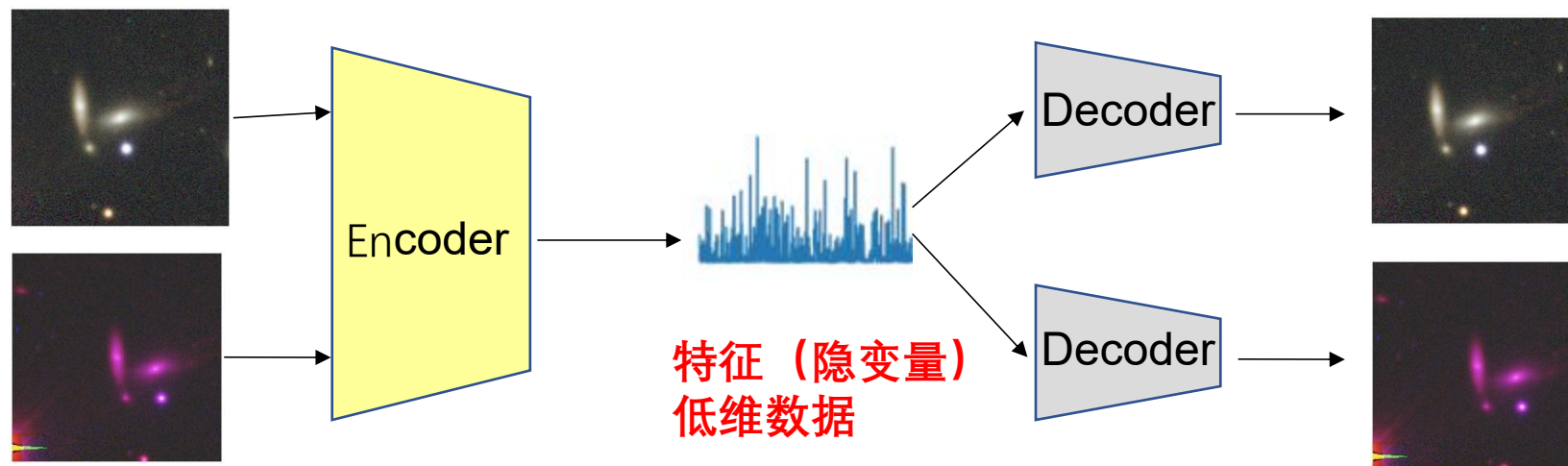
- 对输入的一种映射（提取）
- 可以基于特征量执行不同的下游任务
 - 形态分类，参数测量，相似图像查找生成等
- 压缩即智能

从天文图像到AI空间中的特征：VAE方法

(Xu, Shen*+ 2023)



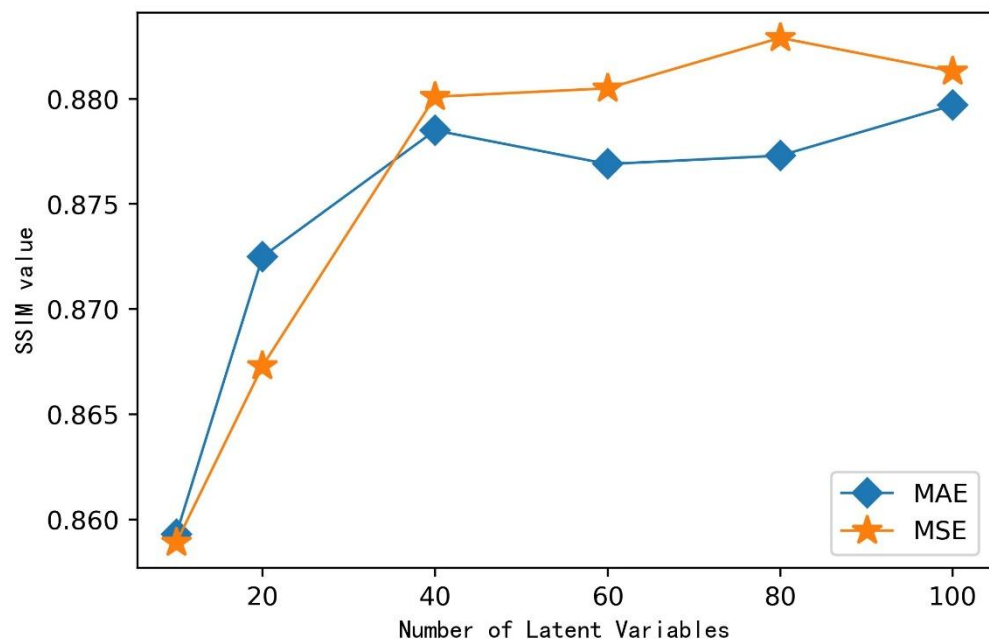
VAE方法的核心假设：隐变量服从一个多元的简单分布（比如高斯）



- 特征量的提取是一种数据降维过程
 - 降到多少维合适（超参）？

- 特征量通常不具有直接的可解释性
 - 可以做什么？

一张256*256星系的图像可以压缩到多少维度?



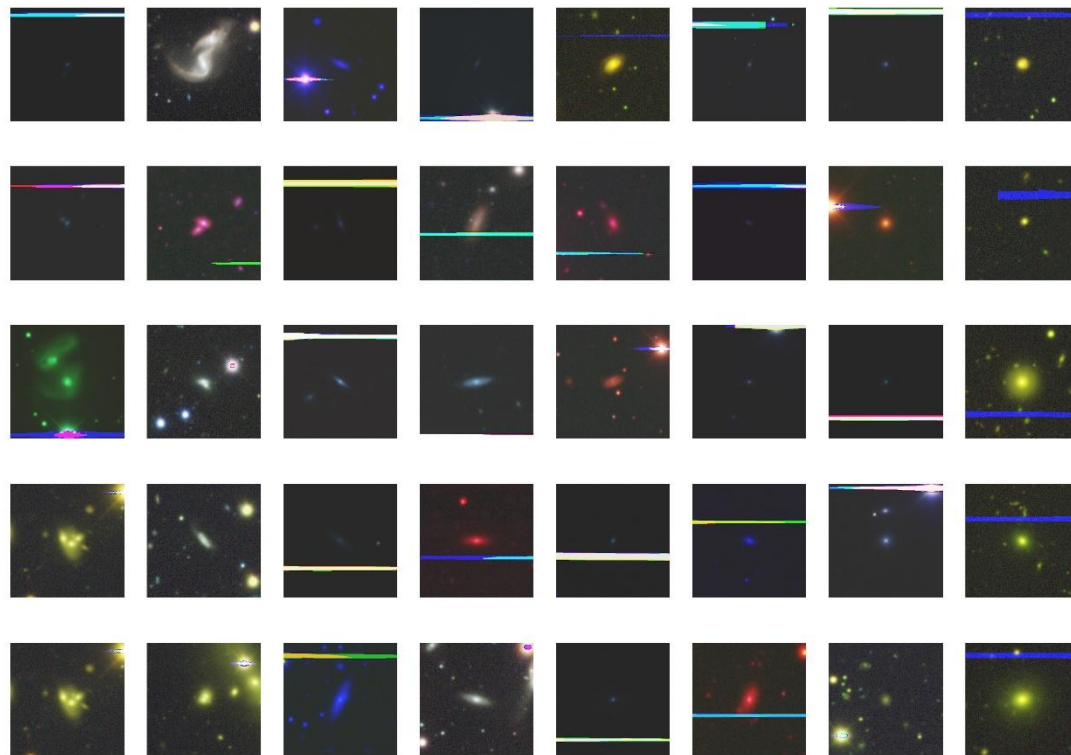
N ~ 40

利用这40个参数就可以对星系形态进行分类!

Feature	Classes	Test acc	Precision	Recall	F1
How round	3	90.68 %	90.68 %	90.68 %	90.68 %
Edge-on	2	84.25 %	98.78 %	84.43 %	91.04 %
Bulge Shape	3	84.08 %	84.08 %	84.08 %	84.08 %
Bar	3	76.21 %	76.21 %	76.21 %	76.21 %
Have Arm	2	91.91 %	73.38 %	96.68 %	83.44 %
Arm Tightness	3	75.84 %	75.84 %	75.84 %	75.84 %
Arm Count	4	86.32 %	86.32 %	86.32 %	86.32 %
Bulge Size	4	82.60 %	82.60 %	82.60 %	82.60 %
Merger	2	96.36 %	99.90 %	96.45 %	98.15 %

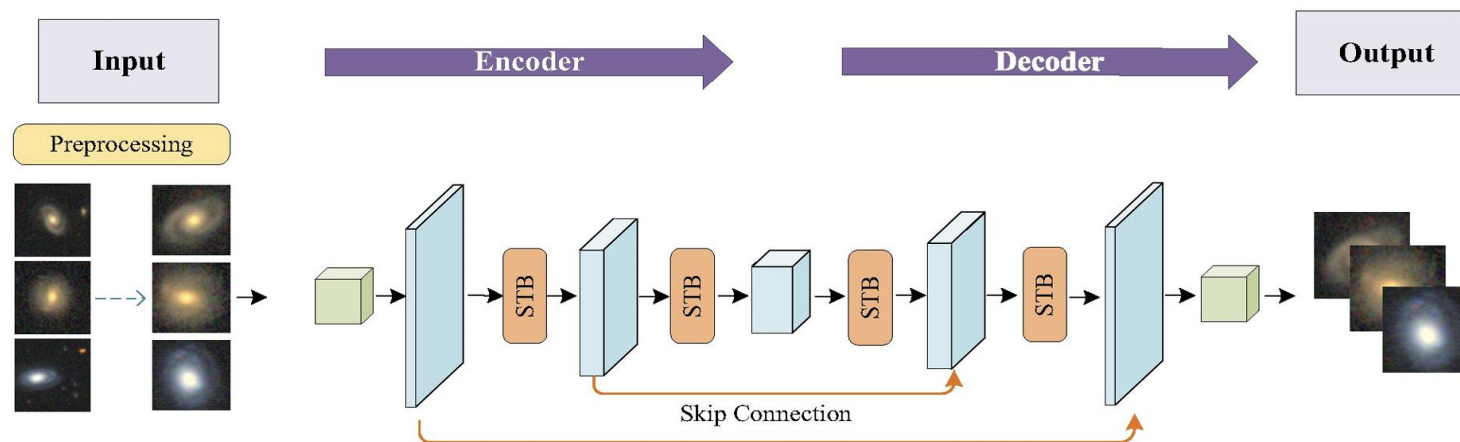
特征量的重要应用：异常目标快速检测

- VAE方法的先验假设
 - 特征参数（隐变量）服从高维的高斯分布
- 在特征空间偏离多个sigma的图片

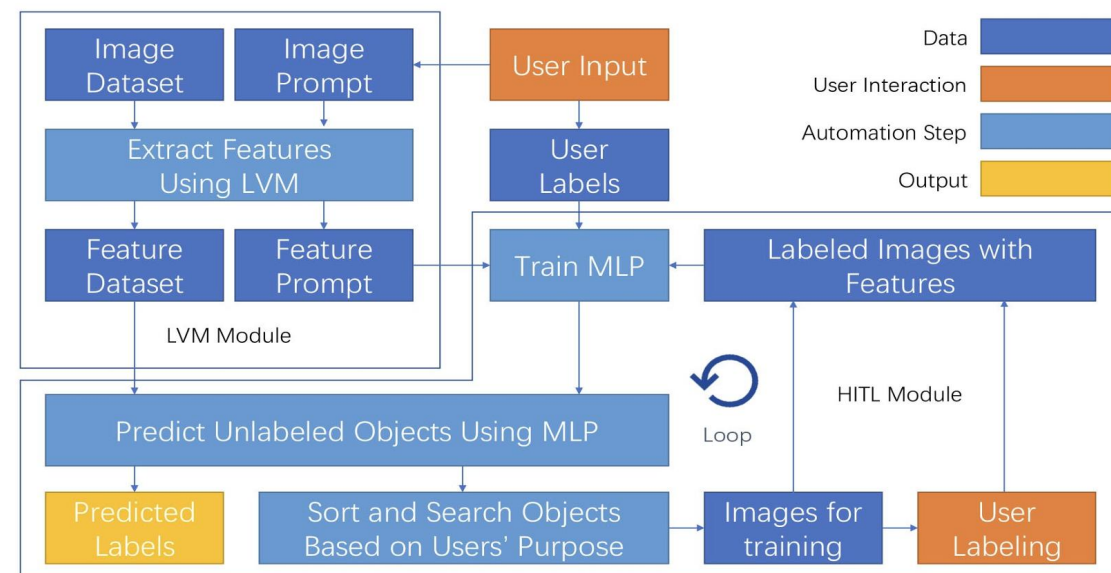
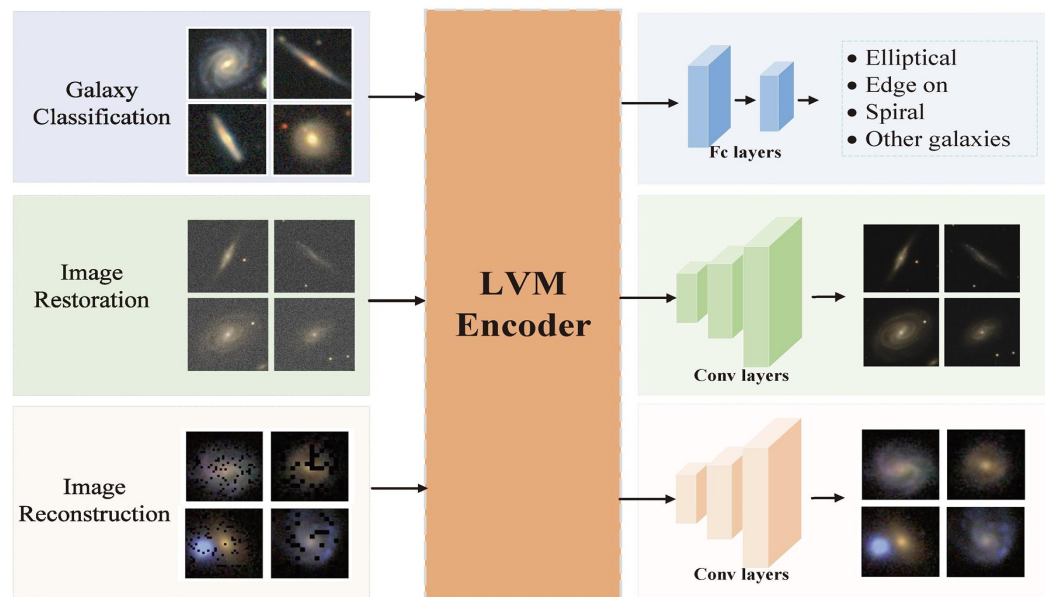


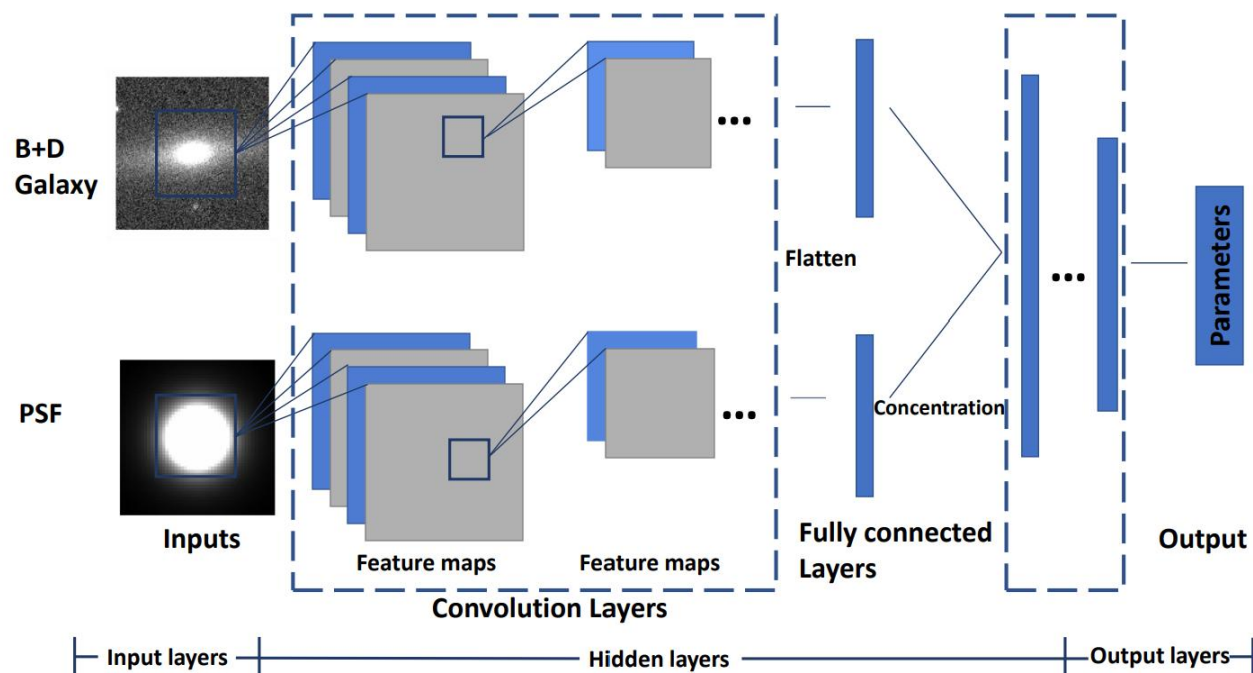
隐变量可以接多种不同的下游任务

- large vision model (LVM) trained on DESI galaxy images (Fu et al. arXiv 2405.10890)



human in loop





隐变量不能带来图像之外的额外信息

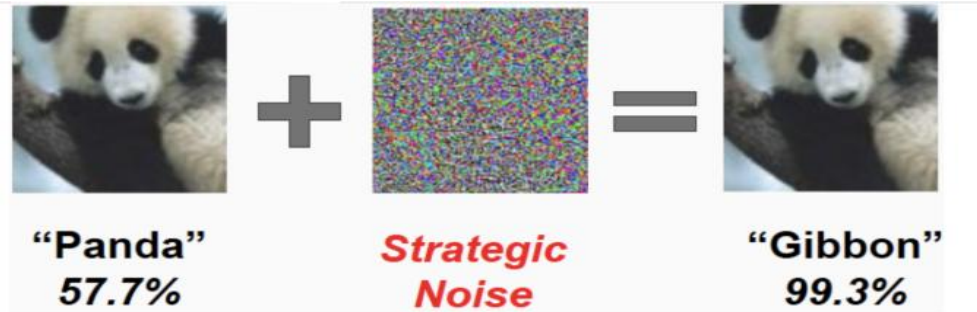
输入中附加PSF信息

输出可以做定量的形态参数测量

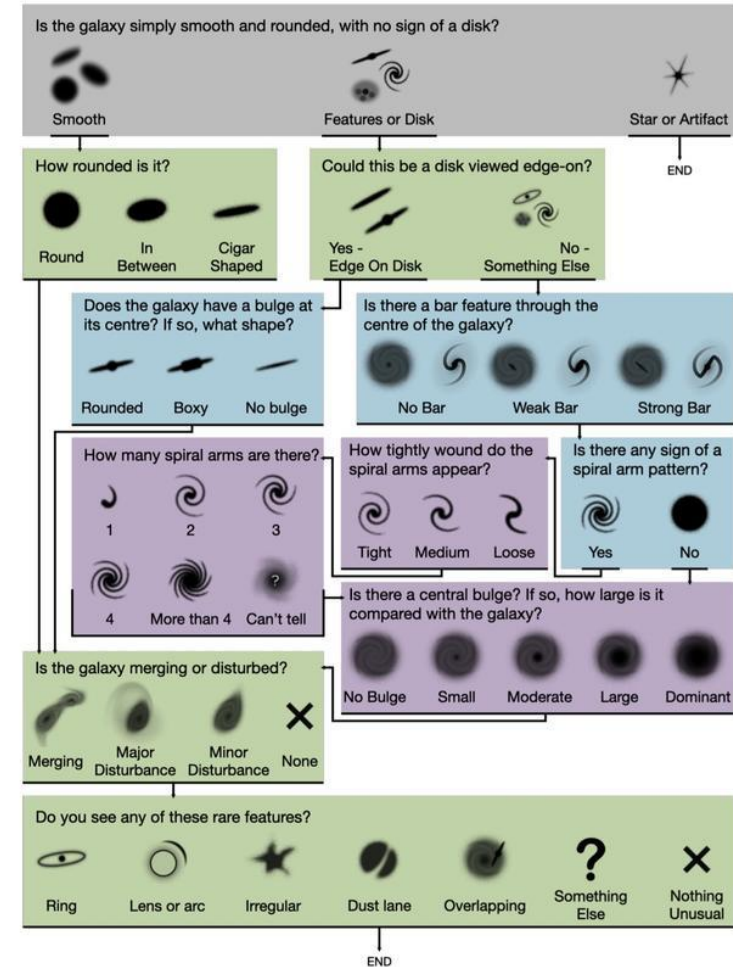
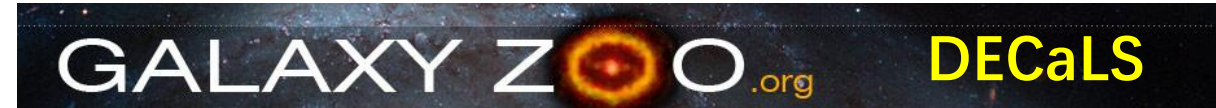
Galaxy Light profile neural Networks (GaLNet). II. Bulge-Disc decomposition in optical space-based observations

CHEN QIU,¹ NICOLA R. NAPOLITANO ,^{1,2} RUI LI,^{3,4} YUEDONG FANG,⁵ CRESCENZO TORTORA,⁶ SHIYIN SHEN,^{7,8}
 LUIS C. HO ,^{9,10} WEIPENG LIN,^{1,2} LEYAO WEI,¹ RAN LI,⁴ ZUHUI FAN,¹¹ YANG WANG,^{2,12} GUOLIANG LI,¹³ HU ZHAN,^{14,9}
 AND DEZI LIU¹¹

域适应问题: From Survey A to Survey B

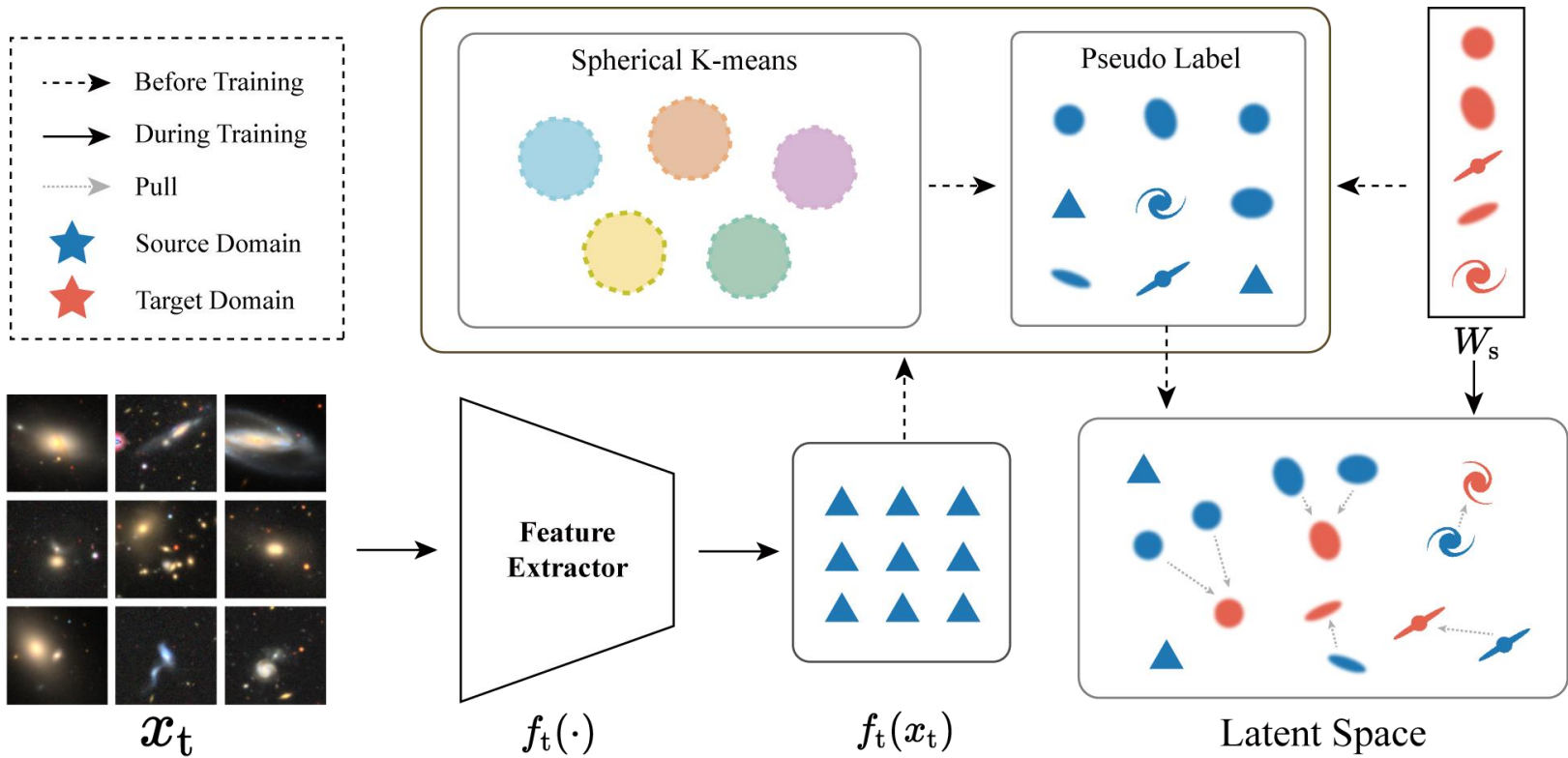


Survey	Instrument	Bands	Area	Pixel Scale
DECaLS	Blanco 4m/DECam	g, r, z	$\sim 9000^\circ$	$0''.262$
BASS	Bok 2.3m/90Prime	g, r	$\sim 5000^\circ$	$0''.454$
MzLS	Mayall 4m/MOSAIC-3	z	$\sim 5000^\circ$	$0''.262$



How to transfer AI classification from DECaLS to BASS/MzLS?

无监督的域迁移：对齐特征空间中的分布



Ye, Shen* et al. 2025

Feature	Accuracy
bar	82.1%
DECaLS数据的源域模型	



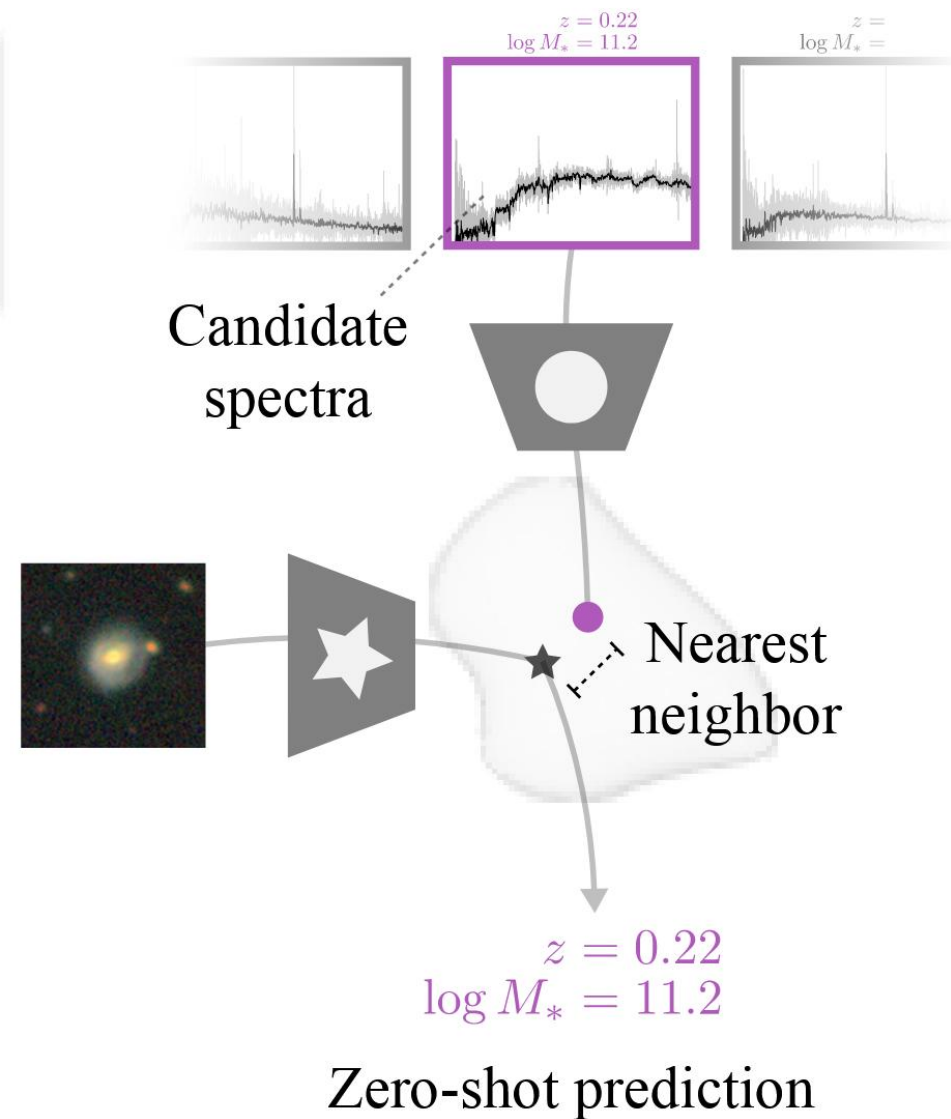
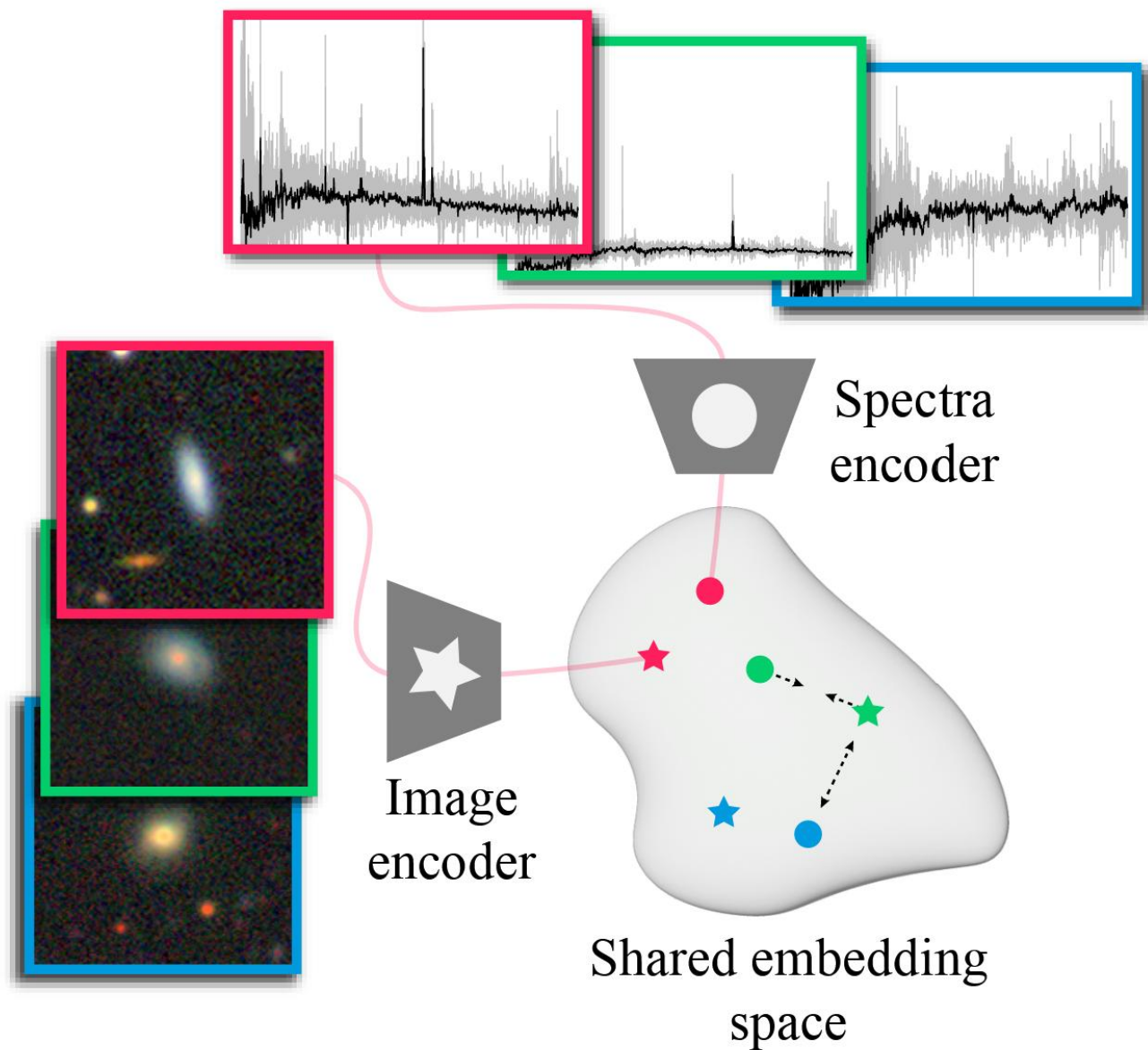
Feature	Accuracy
bar	75.9%
源域模型直接应用到BASS/MzLS	



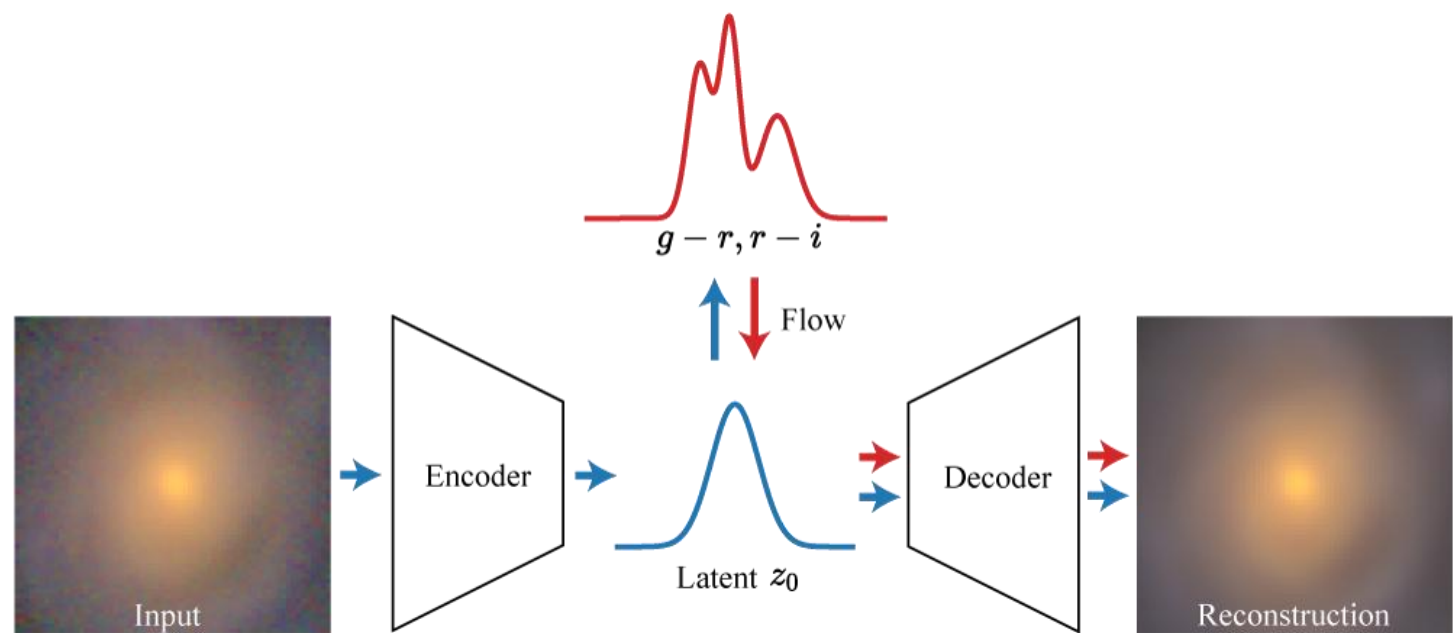
Feature	Accuracy
bar	83.4%
经过非监督域适应过程之后	

多模态数据

AstroCLIP



基于隐变量的图像生成

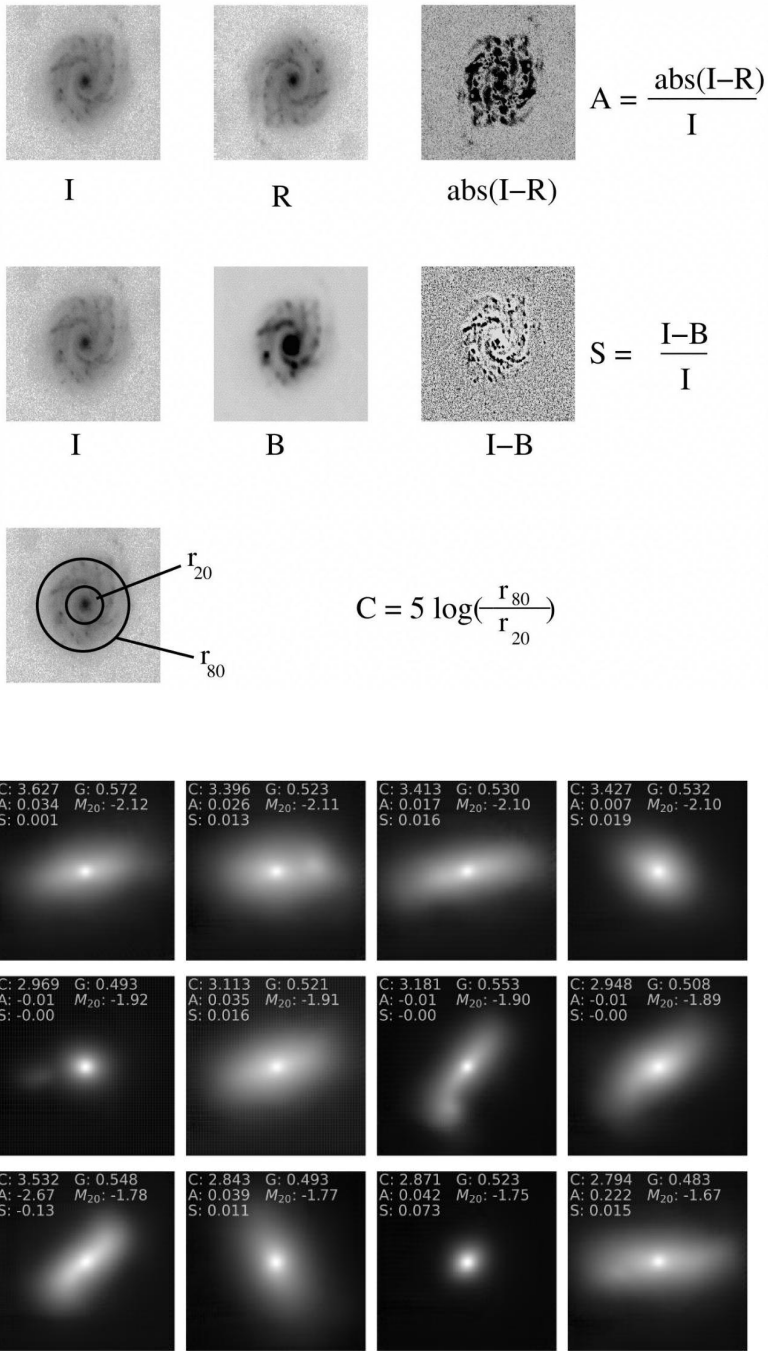


- **方法: Flow VAE**

- 原理: 在 (图像参数+图像) 的分布与隐变量分布之间进行映射

- **结果 (Xu et al. 2025, in prep)**

- 接基于非形态参数生成图像 (传统方法无法完成)



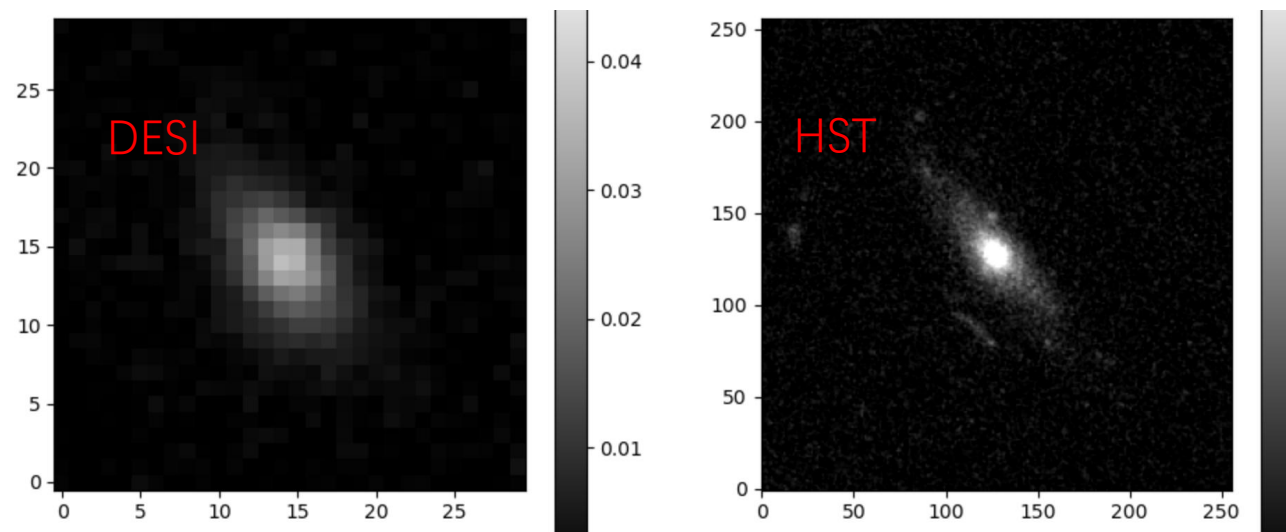
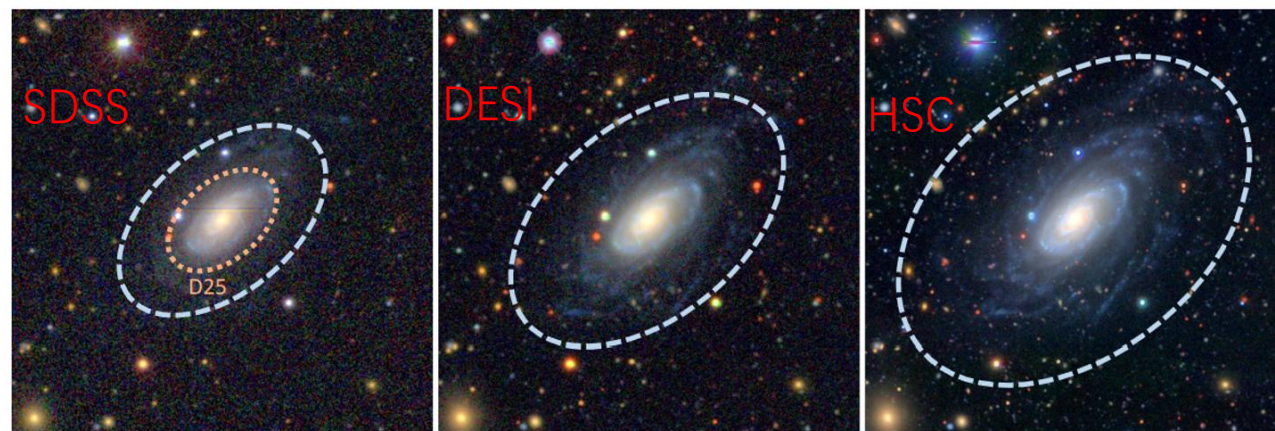
展望：天文图像基础大模型（星系，光学波段）

- ✓ 不同的望远镜观测的同一个星空（基本没有时变），积累了海量的（冗余）信息
- ✓ 宇宙学原理：宇宙中没有特殊的方向（天区）

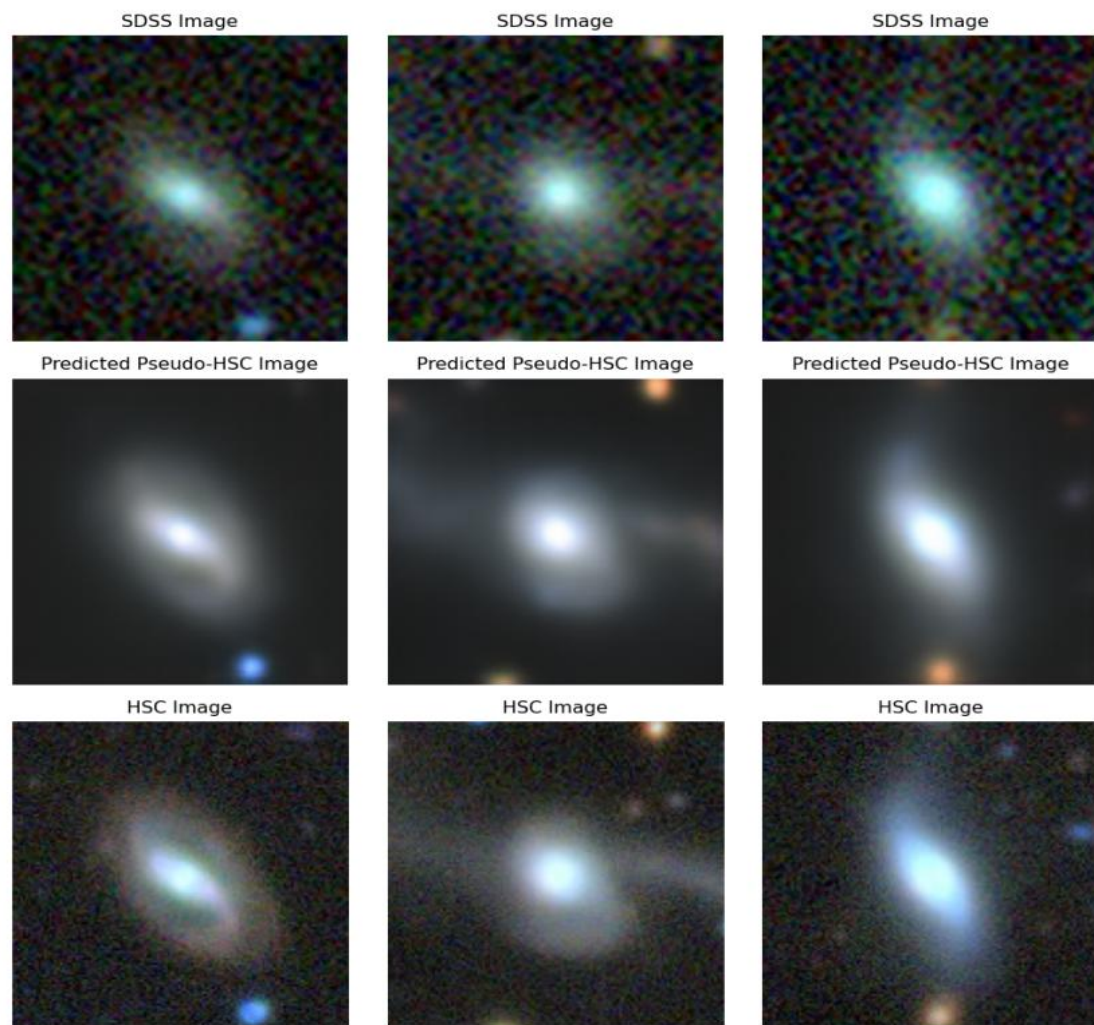
所有的星系巡天数据 → 现代的AI技术能够从数据中获取什么？

需求1：不同观测设备图像之间的映射

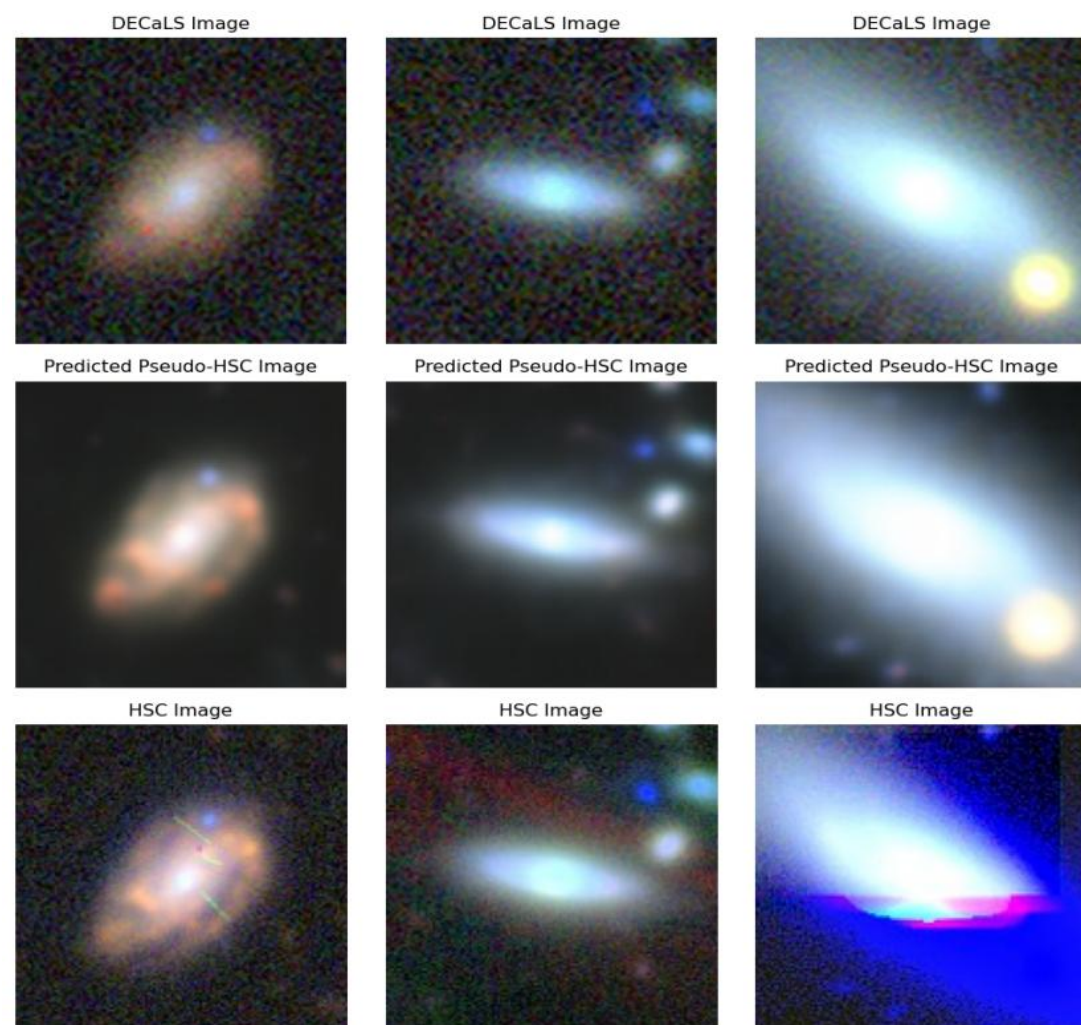
- 浅曝光数据 → 深曝光数据
 - 信息淹没在背景中（去噪）
- 低分辨率数据 → 高分辨率数据
 - 分辨率：点扩展函数（和采样率匹配）
 - 去噪，反卷积（仪器效应）
- 关键：不确定性的准确评估
- → 所有星系的高分辨率深场图像
 - 星系形态、结构、（引力透镜）研究领域的革命性工具



SDSS→HSC



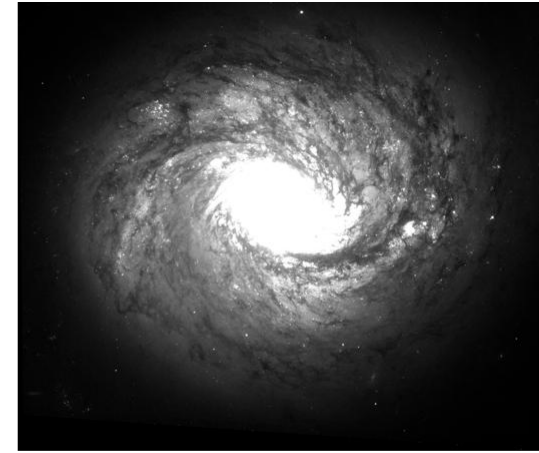
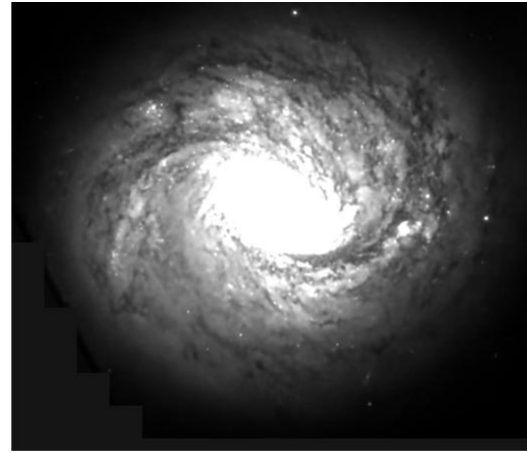
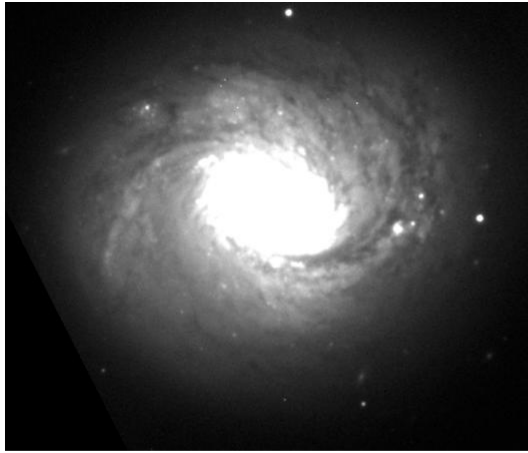
DECaLS→HSC



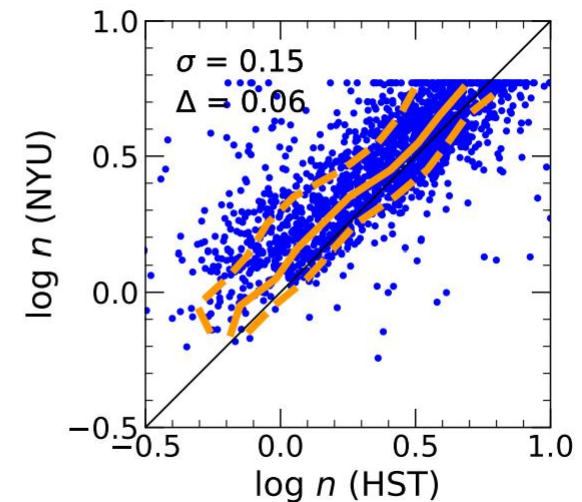
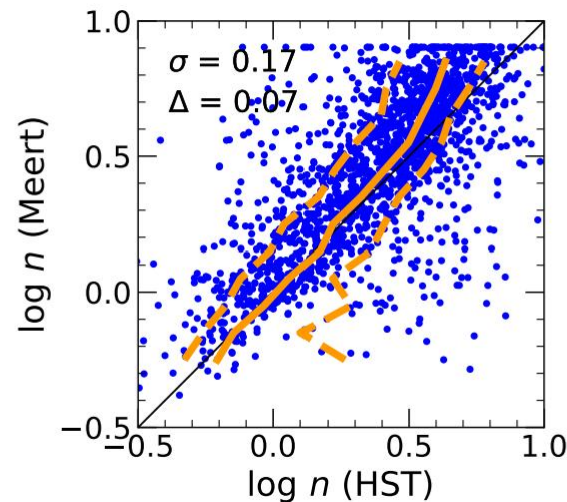
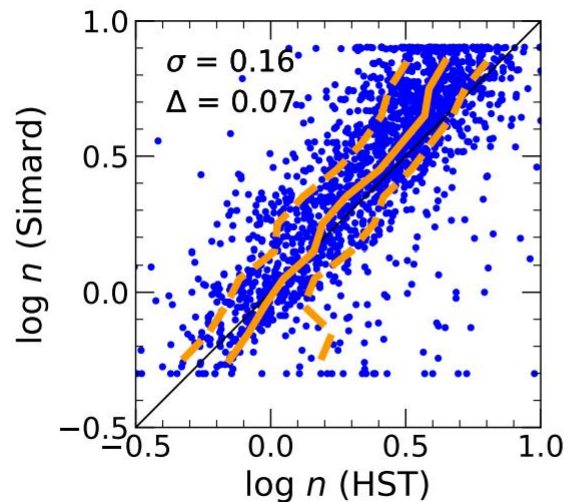
Cascade Pix2WGAN

Luo et al. 2024, arXiv 2410.20025

SDSS→HST (collab with Shanghai AI lab)



Which one is generated by AI?

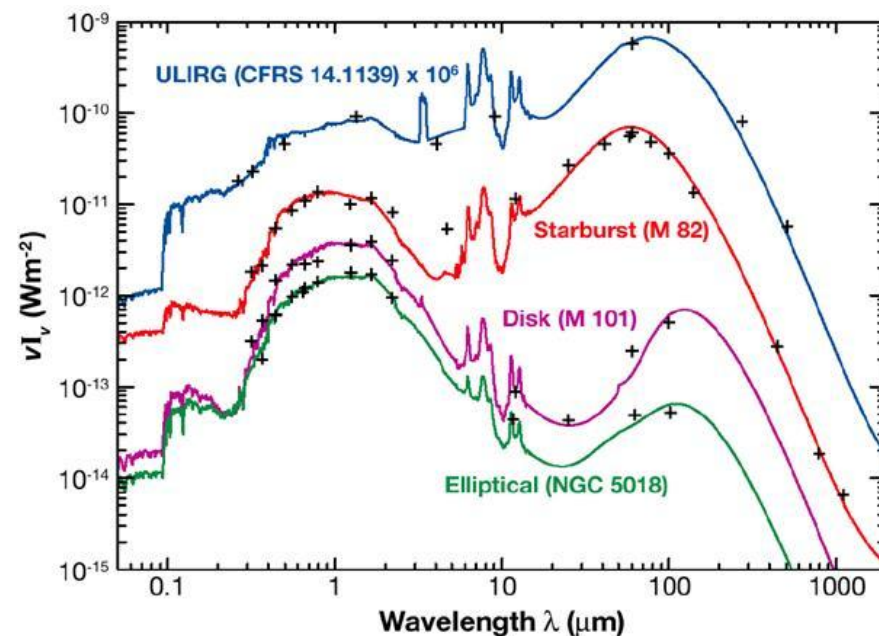
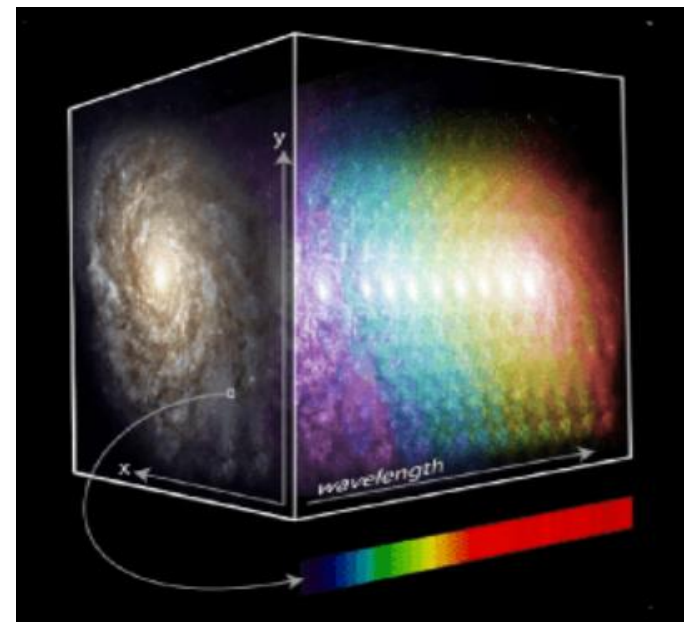


AI VS 观测

- AI的超分，超深生成是基于已有数据模式给出的概率采样
- 深场、高分辨率观测是对真实数据的一次观测采样

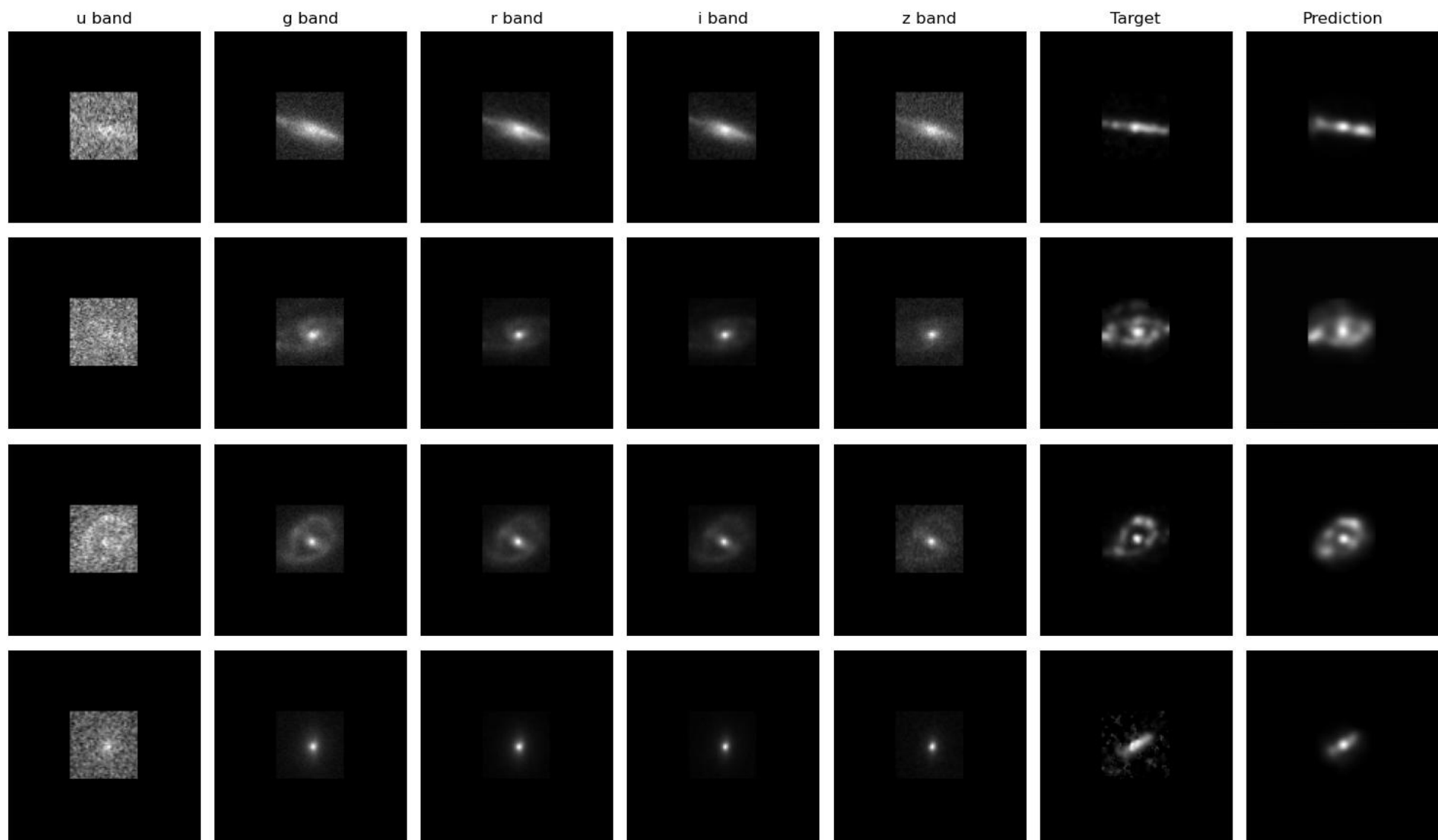
需求2：不同波段图像之间的映射

- 星系图像在不同波段的差异(除了观测效应之外)
 - 包含了极为丰富的物理信息 (如星族, 气体等)
 - 星系内禀的光谱 (SED) 变化并不复杂 ($n(\text{PCA}) < 10$)
- 积分视场光谱数据 (贵)
 - 已有数据样本 ~ 15000
- 经典的SED拟合
 - 需要克服不同仪器的观测效应的差异
- 少量测光数据 $\rightarrow$ 所有星系的3维数据立方
 - 星系物理研究的基础性工具



Predict Ha image from SDSS u,g,r,i,z images

preliminary



陈思远

Trained on MaNGA data

需求3：基于**AI**的天文图像处理

- 天文图像处理的核心任务：

- 去背景（噪声），退卷积(PSF, 仪器效应)

噪声来源：光子涨落，探测器读出噪声，天光背景

噪声：空间上各点独立，近似为高斯分布，最大熵，无信息量，不可预测

- 天文图像的噪声在统计上是已知的！ PSF也是已知的
 - 退卷积是困难的ill-posed problem（计算耗时）

传统的图像退卷积方法

- 频域空间滤波 (抑制噪声)

- Wiener滤波

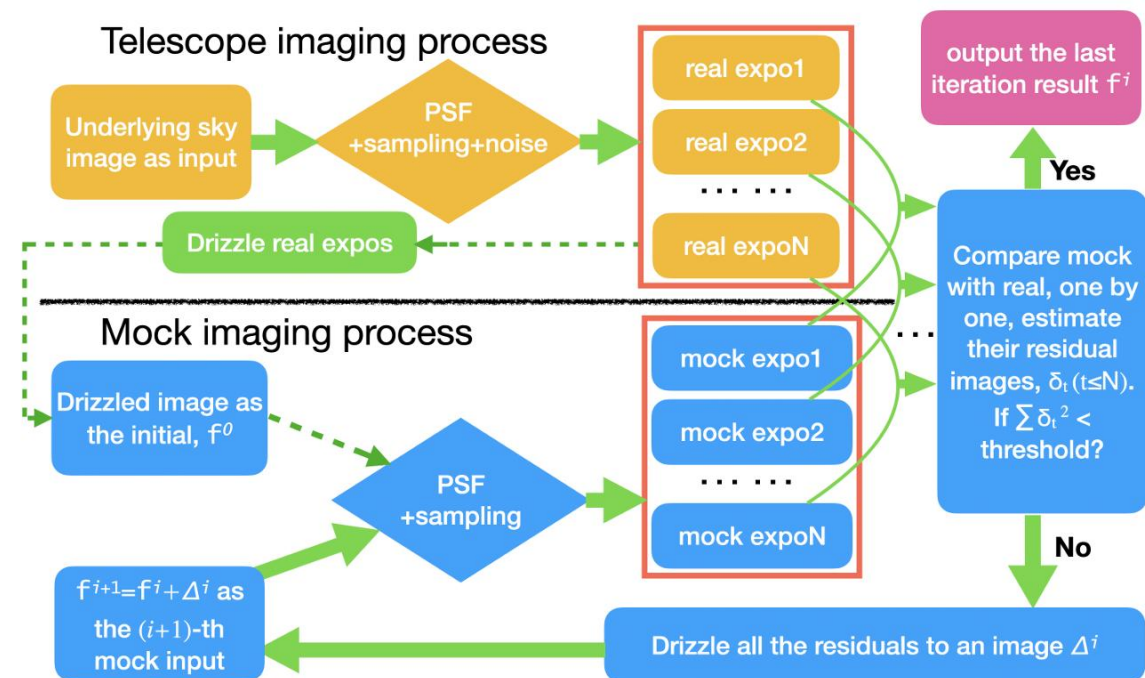
$$\tilde{I}(\mathbf{k}) = \Phi(\mathbf{k}) \frac{\tilde{D}(\mathbf{k})}{\tilde{H}(\mathbf{k})}. \quad (13)$$

It can be shown (e.g., Press et al. 2002) that the optimal filter, which minimizes the difference (in the least squares sense) between the filtered noisy data and the true signal, is the Wiener filter, expressed in Fourier space as

$$\Phi(\mathbf{k}) = \frac{\langle |\tilde{D}_0(\mathbf{k})|^2 \rangle}{\langle |\tilde{D}_0(\mathbf{k})|^2 \rangle + \langle |\tilde{N}(\mathbf{k})|^2 \rangle} = \frac{\langle |\tilde{D}_0(\mathbf{k})|^2 \rangle}{\langle |\tilde{D}(\mathbf{k})|^2 \rangle}. \quad (14)$$

- 图像域的迭代过程

- Up-sampling and PSF Deconvolution Coaddition (Wang et al. 2022,2025)



AI中的数据处理

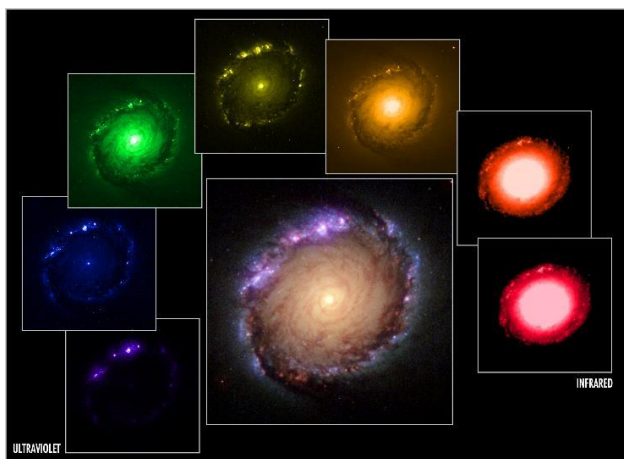
- 万能函数：能否直接 “猜” 出真实图像信息？
- 图像中存在大量冗余信息
- 训练集可以提供大量的先验信息 --> 解决ill-posed问题

Self2Self: 一幅图直接去噪

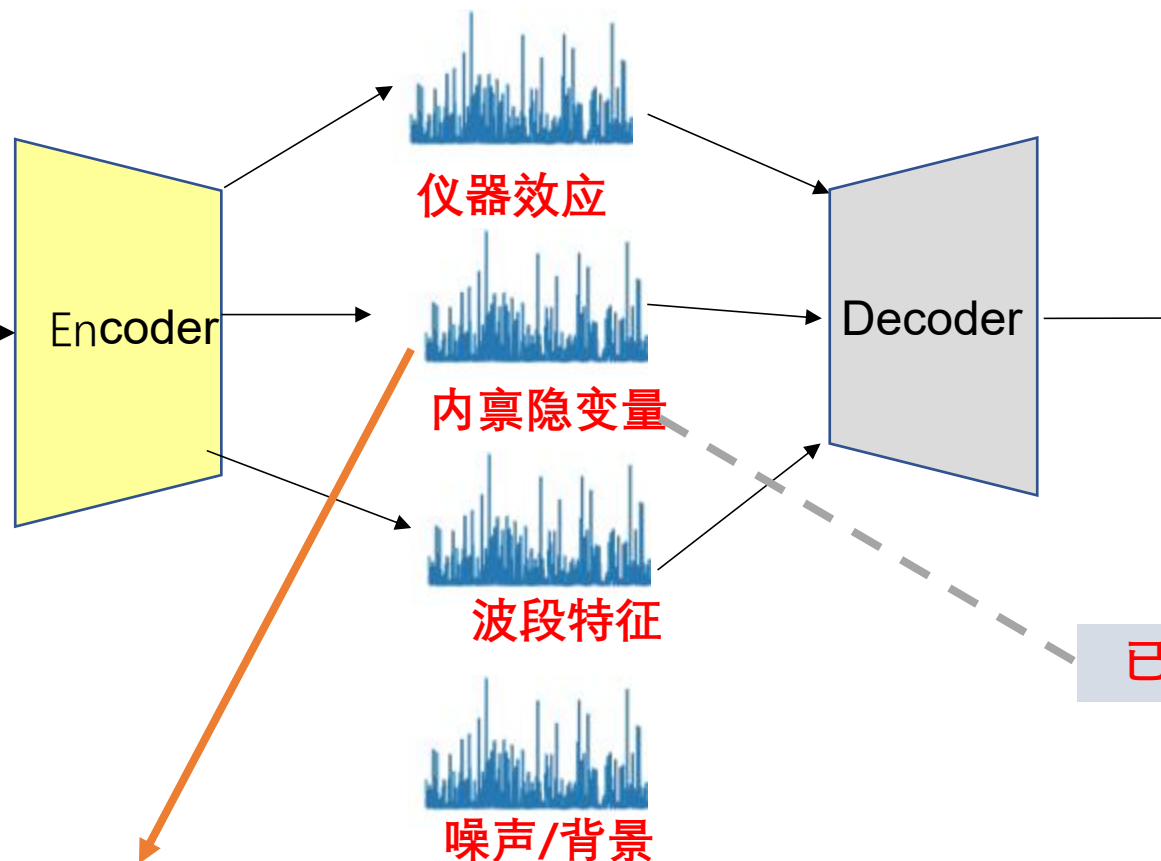
- 原理

- 图像进行随机Mask，由图像其余部分预言Mask区域
 - 图像信息是相关，冗余的
- 噪声是不可预测（独立）的，预言的Mask区域的值是期望值（去噪）

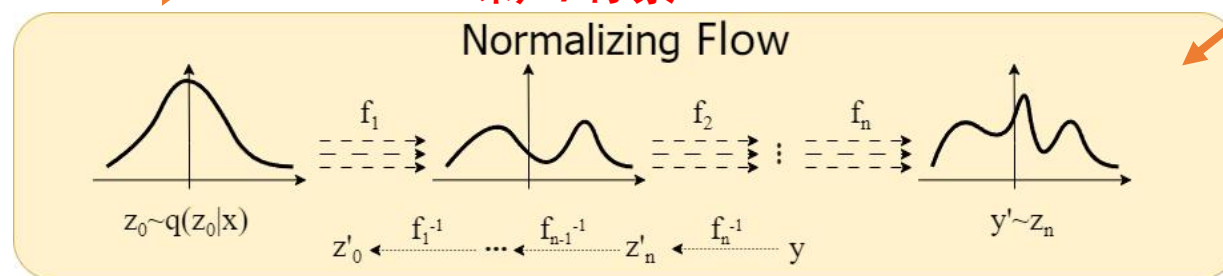
星系图像基础大模型



已有的星系图像（观测）



已有的物理参数测量（知识）



总结



- 基于深度学习中的图像特征提取技术和已有的海量星系巡天数据，有望（需要）建立一个基础的星系图像大模型
 - 超分，超深，全波段星系图像及精度评估
 - 实现图像/仪器/波段/噪声特征的分离、提取及语义映射

