



中国科学院大学
University of Chinese Academy of Sciences

硕士学位论文

基于深度卷积神经网络的星系形态分类研究

作者姓名：戴加明

指导教师：佟继周 副研究员 中国科学院国家空间科学中心

学位类别：工程硕士

学科专业：计算机技术

培养单位：中国科学院国家空间科学中心

2018 年 6 月

Galaxy Morphology Classification with
Deep Convolutional Neural Networks

A thesis submitted to the
University of Chinese Academy of Sciences
in partial fulfillment of the requirement
for the degree of
Master of Engineering
in Computer Technique

By

DAI Jiaming

Supervisor: Professor TONG Jizhou

National Space Science Center, Chinese Academy of
Sciences

June, 2018

中国科学院大学 研究生学位论文原创性声明

本人郑重声明：所呈交的学位论文是本人在导师的指导下独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明或致谢。

作者签名：

日 期：

中国科学院大学 学位论文授权使用声明

本人完全了解并同意遵守中国科学院有关保存和使用学位论文的规定，即中国科学院有权保留送交学位论文的副本，允许该论文被查阅，可以按照学术研究公开原则和保护知识产权的原则公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存、汇编本学位论文。

涉密及延迟公开的学位论文在解密或延迟期后适用本声明。

作者签名：

日 期：

导师签名：

日 期：

摘要

随着观测技术的进步、观测仪器的发展，大天区星系图像巡天计划如斯隆数字巡天（Sloan Digital Sky Survey, SDSS），COSMOS 巡天（Cosmic Evolution Survey, COSMOS）等逐步实施，星系观测数据呈现爆炸式增长趋势，新兴观测手段使得天文学迈入“大数据”时代。在星系物理研究方面，海量数据使得人工进行大规模星系形态分类已经绝无可能。如何在自动、快速、高效、准确地区分星系大规模样本不同形态的同时，从海量、高维数据中挖掘隐含的信息，甚至是发现新的科学问题是当前一项意义非凡而又充满挑战的任务。

基于传统机器学习的星系形态分类方法，需要复杂的特征工程，需要领域知识精巧地设计和提取特征，特征提取的好坏直接决定了最后分类器的分类性能，主观性较强。深度学习可以自动找出复杂且有效的高阶特征，避免了人工提取特征的繁琐过程和主观性造成的分类问题。本文在比较分析了 AlexNet、VGG、Inception 和 ResNet 等经典卷积神经网络（Convolutional Neural Network, CNN）的基础上，提出了一种基于改进的深度卷积神经网络的星系形态分类（Galaxy Morphology Classification）方法。该方法以深度残差网络（Residual Network, ResNet）为基础，改进残差单元，加入 Dropout，减少网络深度，加宽网络宽度，以求在模型的泛化能力上获得最佳结果。

改进的网络模型以星系原始图片像素作为输入，自动实现对星系形态特征的自动提取、识别与分类。星系图片选择星系动物园 2 中的干净样本，以 25911 张星系图片作为训练集，进行 5 个星系类别的分类训练；以 2879 张星系图片为测试集，测试模型分类性能；经超参数选择，26 层网络结构呈现出最优性能，将该网络结构称之为 ResNet-26。在相同环境下，将 ResNet-26 与 Dieleman、AlexNet、VGG-16、Inception V3 和 ResNet-50 5 种流行的 CNN 模型进行训练和测试，ResNet-26 取得了较优的分类性能。测试集上的平均分类准确率达到 95.2083%，平均 F1 值为 0.9515，AUC 值为 0.9823。

本文将高维数据可视化技术引入星系形态分类的后续分析，从训练好的卷积神经网络模型的最后一层全连接层（除了输出层），或最后一层卷积层，或最后一层平均池化层抽取星系高维抽象表征，运用 t-SNE 降维技术，将星系高维表征降至二维空间，以散点图的方式呈现，以进一步挖掘星系形态分类网络所学习到的高维抽象表征，直观呈现海量星系图片中隐含的全局结构和局部结构信息，并

初步探究星系形态高维抽象表征内在规律与联系。

本文提出的 ResNet-26 可应用于未来的大天区星系图像巡天计划，如大型综合巡天望远镜 (Large Synoptic Survey Telescope, LSST) 的大规模星系分类系统；星系形态高维表征可视化可以作星系分类过程离群点检测，以及特殊星系类别发现等星系形态分类后续分析过程。

关键词：星系形态分类，卷积神经网络，残差网络，表征，可视化

Abstract

Along with the development of the observation technology and observation instrument, large scale surveys such as the Sloan Digital Sky Survey (SDSS), Cosmic Evolution Survey (COSMOS) have gradually carried out and resulted in the availability of very large collections of images, which show an explosive growth trend. The new observation means make astronomy enter the era of big data. In the study of Galaxy, massive data makes it impossible to manually classify large-scale galaxies images. Currently, it is a significant and challenging task to mine hidden information from massive and high-dimensional data, even discover new scientific problems, while classifying galaxies automatically, quickly, efficiently and accurately.

Galaxy morphology classification based on traditional machine learning, depends on complex feature engineering, which needs specific domain knowledge and human ingenuity to extract good data representation. Feature extraction directly determines the classification performance of the final classifier. While the most important value of neural network is the automatic extraction and abstraction of features, which eliminates the complexity of manual feature extraction and can automatically find complicated and effective higher level representations. This paper proposes a modified deep convolutional neural network (CNN) for galaxy morphology classification based on classical CNN models such as AlexNet, VGG, Inception and ResNets. We try to improve residual unit, add a dropout, reduce the depth of the network and widen the width of the network in order to achieve the best generation performance of model.

The network is fed with raw galaxy images pixels directly and automatically extracts galaxy morphological features to identification and classification. A sample of 25911 galaxy images from Galaxy Zoo 2 dataset are applied to classify galaxies into five classes and the remaining 2879 galaxy images are testing set to test the classification performance of our model. Through the selection of super-parameter, a 26-layers network is adopted and called ResNet-26. A variety of metrics, such as accuracy, precision, recall, F1 value and AUC, show that ResNet-

26 achieves the state-of-the-art classification performance among 5 popular CNNs, namely, Dieleman, AlexNet, VGG-16, Inception V3 and ResNet-50. The overall classification accuracy of our network on the testing set is 95.2083%, the average F1 value is 0.9515 and average AUC value is 0.9823.

Further, this paper digs into the networks to analyse what features are learned and how the learned features can help better understanding the data itself. We extract the activations from the last fully connected layer or the last convolutional layer or the last average pooling layer of CNNs to study the high-dimensional abstract feature representations of galaxy images. Then we apply t-Distributed Stochastic Neighbor Embedding (t-SNE), a popular dimensionality reduction technique, to visualize the high-dimensional galaxy feature representations in two-dimensional scatter plots. From the visualization, We try to understand the galaxy images data itself and obtain some highly valuable insights. For instance, we try to discover the implied global and local structure information, explore the internal laws and relations of galaxy high-dimensional and abstract representation.

ResNet-26 can be applied to large-scale galaxy classification in forthcoming surveys such as the Large Synoptic Survey Telescope (LSST). And the visualization of galaxy morphological high-dimensional representation can be applied to outlier detection, and look for abnormal galaxies in the process of subsequent analysis of galaxy morphological classification.

Keywords: Galaxy Morphology Classification, CNNs, ResNets, Representation, Visualization

目 录

第 1 章 绪论	1
1.1 研究背景及意义	1
1.2 国内外研究现状	3
1.3 研究内容与章节安排	5
1.3.1 研究内容	5
1.3.2 章节安排	5
第 2 章 星系形态分类标准	7
2.1 目视分类系统	7
2.2 模型化分类系统	8
2.3 非模型化分类系统	8
2.4 星系动物园决策树	9
2.5 本章小结	12
第 3 章 深度卷积神经网络	13
3.1 深度卷积神经网络简介	13
3.1.1 前馈神经网络	13
3.1.2 深度卷积神经网络	16
3.2 LeNet5	17
3.3 AlexNet	18
3.4 VGG	19
3.5 Google Inception	20
3.6 ResNet	21
3.7 本章小结	23
第 4 章 ResNet-26 网络结构与实验结果分析	25
4.1 网络结构设计	25
4.2 数据集选取	27

4.3	数据预处理	30
4.4	实验结果及分析	31
4.4.1	实验评价指标	31
4.4.2	超参数选择	32
4.4.3	不同网络模型分类性能对比	37
4.4.4	实验环境	39
4.5	本章小结	39
第 5 章 星系图像表征与结果分析		41
5.1	T 分布随机近邻嵌入 (t-SNE)	41
5.2	星系图像数据集的 t-SNE 可视化及分析	42
5.3	星系形态分类模型的 t-SNE 可视化及分析	43
5.4	本章小结	51
第 6 章 总结与展望		53
附录		55
参考文献		57
作者简历及攻读学位期间发表的学术论文与研究成果		63
致 谢		65

第 1 章 绪论

1.1 研究背景及意义

恒星物理、星系形成与演化,以及宇宙的起源是现代天体物理学的三个主要研究方向。研究星系结构随着时间的演化,可以知道宇宙早期的原初物质是如何通过引力作用逐渐塌缩形成原初星系,以及高红移星系是通过怎样的过程(并合、吸积?)演变为现在能够观测到的各类星系,从而了解和约束星系的形成与演化过程(汪敏等, 2007)。星系形态是表示星系结构最直观的观测特征,它是不同运动状态的恒星轨道在天空中的投影。星系形态与星系的形成与演化有着密切的联系,是探究星系物理的重要参数。如早型星系在测光上颜色偏红,星族年龄(Stellar Population)偏老年,中心多存在以速度弥散为主导的核球(Bulge),并且有其历史上经历了多次星系并合留下的壳层结构(Shell)。大质量早型星系多处于致密环境中,形态上表现为椭圆星系。而晚型星系则在测光上表现为颜色偏蓝,星族成分较为年轻,含有大量冷气体,有剧烈的恒星形成活动。这些晚型星系中恒星的运动状态以绕中心旋转为主导,形态上表现为有恒星盘以及旋臂结构的盘状星系。故而为了研究不同类型星系的物理特征,首先要做的是有效区分星系的形态(方官文等, 2015)。

宇宙中的星系形状没有完全相同的,即每一个星系都有着独特的外貌。但是由于星系都是在一个相对有限的条件范围内形成的,所以它们有一些共同的特点,人们可以对它们进行大体的分类。20 世纪 70 年代以前,限制于当时的观测仪器,只能利用照相底片观测一些距离比较近的、亮的大星系图像,然后采用目视的方法对它们的形态进行分类,称为目视分类系统。其中,最著名的是哈勃于 1926 年提出的“哈勃序列”星系形态分类系统(Hubble, 1926),基于哈勃序列,学者们对星系形态分类进行了发展和细化,后发展出模型化分类系统、非模型化分类系统和双色图分类系统等不同的分类标准。

在众多的星系分类方法中,尽管哈勃序列仅根据星系的光学波段形态特征对星系进行分类,但哈勃序列至今还具有很强的生命力。研究表明哈勃序列和许多星系物理参数都有关,如:中性氢的含量、动力学特征(恒星速度弥散和转动曲线)、星系的积分颜色、星系光度、星系质量和环境等(汪敏等, 2007)。2007 年以英国牛津大学为首的研究机构推出星系动物园项目(Galaxy Zoo, GZ),该项目分类系统基于哈勃序列的思想而设计(Zooniverse),它采用“众包”的方式

邀请大量的天文学志愿者参与星系分类，项目开始第 1 天，每小时约完成 7 万次分类投票；项目第 1 年，约有超过 15 万完成了 5000 万次分类投票。据统计，90 万张图片花费数月被标注完成分类。星系动物园项目得到的星系分类结果与专业天文学家给出的结果一致 (Lintott et al., 2008)。

随着观测技术的进步、观测仪器的发展，大天区星系图像巡天计划如斯隆数字巡天 (Sloan Digital Sky Survey, SDSS)¹，COSMOS 巡天 (Cosmic Evolution Survey, COSMOS)²，大口径全天巡视望远镜 LSST (Large Synoptic Survey Telescope, LSST)³(Abell et al., 2009) 等逐步实施，星系观测数据呈现爆炸式增长趋势，以 SDSS 为例，其完成三分之一天区观测后就获取到 12 亿张观测图像；2018 年，LSST 投入运行后计划每晚可产生 15 TB 的原始观测数据。海量、复杂的图像、时序、星表、光谱等数据犹如一个深邃的数字宇宙，为天文学家提供了广阔的挖掘空间，也使研究工作的复杂度进一步提升。很多事件和物理过程持续的时间很短，要求近乎实时完成目标证认、特征提取、天体识别、随动观测优先级确定等工作。这与传统的数据处理和分析要求很不相同，必须通过创新的手段实现对海量数据的鲁棒、灵活、实时处理和挖掘 (崔辰州 等, 2015)。

在星系物理研究方面，海量数据使得人工进行大规模星系形态分类已经绝无可能。在过去的 20 年里，星系形态自动化分类方法发展迅速。Naim et al. (1995) 提取了星系的中心凸出部分大小、旋臂数目等形态学参数，利用神经网络 (Artificial Neural Network) 进行完全自动化地星系形态分类。Owens et al. (1996) 以 1700 星系样本为训练集，设计了 13 种参数作为星系特征，作用于决策树 (Decision Tree) 进行星系分类训练，用 3517 星系图片测试分类器性能，并与人工神经网络的分类结果进行了比较，两者的结果相似。Bazell et al. (2001) 以 800 星系样本为数据集，先单独训练了朴素贝叶斯 (Naive Bayes)、神经网络和决策树模型，然后使用集成分类方法，结果表明单个模型实验，人工神经网络的错误率最低，而集成方法的分类性能最好。De La Calleja et al. (2004) 使用了主成分分析将 310 张星系图片降维，分别抽取了 8、13 和 25 个主成分，分别训练神经网络、局部加权回归 (Locally Weighted Regression) 和集成分类器，结果表明集成分类器的性能最好，分成两个类的准确率为 95%，3 个星系类 (椭圆、旋涡和不规则) 的准确率为 91%。Banerji et al. (2010) 以星系动物园数

¹<http://www.sdss.org/>

²<http://cosmos.astro.caltech.edu/>

³<https://www.lsst.org/>

数据集为样本，以颜色、形状、聚集度指数等参数训练 3 层人工神经网络，将星系分成 3 类（椭圆星系、旋涡星系和点源/人造物体），准确率超过 90%；Gauci et al. (2010) 以星系动物园数据集为训练集和测试集，使用决策树（C4.5、CART 和随机森林）和模糊逻辑（Fuzzy Logic）算法以 SDSS DR 7 中抽取的测光数据和光谱数据为输入参数将星系分成椭圆星系、旋涡星系和未知类；Ferrari et al. (2015) 测量星系的形态学参数包括聚集度指数（Concentration, C ）、非对称指数（Asymmetry, A ）、簇聚指数（Smoothness, S ）、基尼系数（Gini coefficient, G ）、矩指数（Moment, M_{20} ）、熵（Entropy）和螺旋率（Spirality）等，分别使用 3 个数据集，每个数据集的图片分别为 4478、14123 和 779235，使用线性判别分析（Linear Discriminant Analysis, LDA）将星系分成椭圆星系和旋涡星系 2 个类别，准确率超过 90%。

上述星系形态自动化分类均基于传统机器学习算法实现，一般分成特征工程和分类器两个过程，分类器的分类性能取决于特征提取的好坏 (LeCun et al., 2015)，特征设计和提取需要具有资深专业领域知识和背景的研究人员完成，设计过程耗时，且有研究表明，当分类种类的增加时，自动分类准确率迅速下降，比如 De La Calleja et al. (2004) 发现将星系分类的数目从 3 类增加到 5 类，其人工神经网络模型结果的准确率从 95.66% 降到了 56.33%。新兴观测手段使得天文学迈入“大数据”时代，如何在自动、快速、高效、准确地区分星系大规模样本不同形态的同时，从海量、高维数据中挖掘隐含的信息，甚至是发现新的科学问题是当前一项意义非凡而又充满挑战的任务。

1.2 国内外研究现状

近两年来，基于深度学习算法的星系形态分类方法是一个非常活跃的研究方向。神经网络允许直接输入原始数据，实现特征自动提取和抽象，免去了人工提取特征的繁琐，自动找出复杂且有效的高阶特征，越深层的数据特征越抽象且越有益于识别和分类 (Bengio et al., 2013; LeCun et al., 2015)，在极大程度上解决了基于传统机器学习的自动化星系形态分类方法人工提取特征的繁琐，减少了人为设计特征的主观性。

Dieleman et al. (2015) 经过 100 多次的尝试，首次提出一个 7 层深度卷积神经网络，包括 4 层卷积层（每层卷积核大小分别为 6×6 、 5×5 、 3×3 和 3×3 ）和 3 个全连接层（每层神经元数量为 2018、2048 和 37）（Dieleman 模型见图 1.1），总计 4200 万参数，以 5 万多张星系图片为训练集，进行模型训练，

最终以 0.07492 的 RMS 取得了 Galaxy Zoo-The Galaxy Challenge⁴比赛的冠军。Dieleman 模型首次将深度学习技术应用于星系形态分类，具有重要的意义。

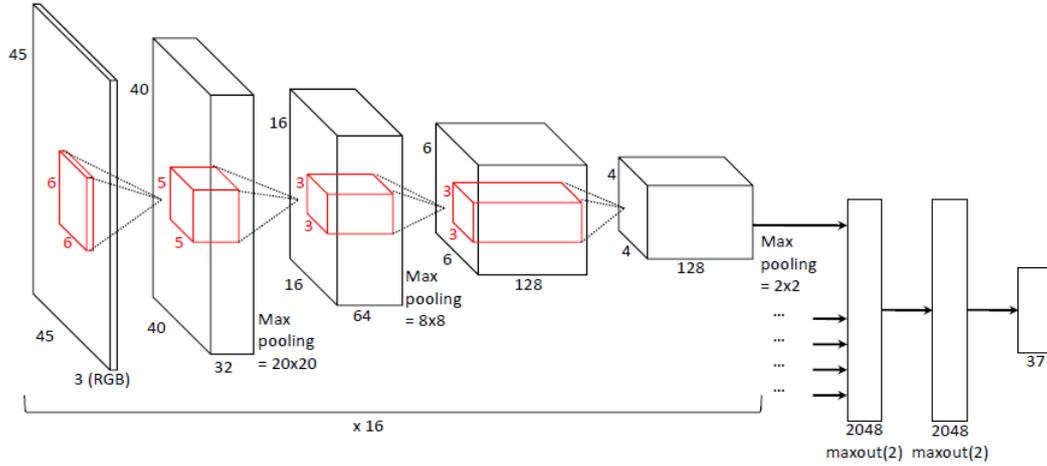


图 1.1 Dieleman 结构图

Figure 1.1 Schematic overview of Dieleman

Gravet et al. (2015) 利用 CANDELS (Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey) 的高红移星系图像，使用 Dieleman 模型将星系分成 5 个类别 (盘状类、球状、不规则类、点源和不可分的类)，测试集上的 RMSE 为 0.13。

Kim et al. (2016) 利用 SDSS DR12 (17344 张恒星图片和 47656 张星系图片) 和 CFHTLenS (Canada-France-Hawaii Telescope Lensing Survey) (8545 张恒星图片和 57843 张星系图片) 的数据集，提出一个类似 VGG (Simonyan et al., 2014) 的 11 层深度卷积神经网络对恒星/星系进行分类，测试集上的 AUC 值分别可以达到 0.9952 和 0.9948。

Aniyan et al. (2017) 利用射电星系图像 (FRI 178 张, FRII 284 张和 Bent-tailed 254 张)，分别使用深度卷积神经网络两两分类，训练了 3 个神经网络模型，然后进行集成，Bent-tailed 的精确率可以达到 0.95，FRI 精确率为 0.91，FRII 精确率为 0.75，FRI 和 FRII 的召回率都为 0.91，Bent-tailed 的召回率为 0.79。

神经网络训练后可以获得高维星系表征，这些表征除了用于星系形态分类，还有哪些用途？还可以分析出哪些问题？对星系形态分类的后续研究是否有帮助？目前国内外尚未有人展开这方面的讨论。可视化是数据理解、分析和展示

⁴<https://www.kaggle.com/c/galaxy-zoo-the-galaxy-challenge>

的重要环节，是数据挖掘不可分割的一部分，是科学数据所蕴含的定量内容和对
其直观理解的桥梁。对高维星系表征数据进行降维和可视化，尝试挖掘其中的隐
含的其他信息，是本文提出的一个创新性的想法。

1.3 研究内容与章节安排

1.3.1 研究内容

本文的主要研究内容为：

1. 调研星系形态分类标准，主要包括目视分类系统、模型化分类系统、非模
型化分类系统和星系动物园决策树，分析其优缺点，选择一种形态分类方
法作为星系形态分类模型的分类标准。
2. 调研以深度卷积神经网络模型（LeNet5、AlexNet、VGG、Inception 和
ResNet）为代表的深度学习技术的发展情况，分析上述模型的算法与技
术先进性、训练技巧以及优缺点等。
3. 基于深度残差网络，尝试改进残差单元，探索网络宽度与网络性能的关系，
结合数据样本特点，设计一种改进的基于深度残差网络的星系形态分类模
型，并进行实验验证，并与其他模型比较分析。
4. 开展高维星系表征数据可视化研究，运用高维数据可视化技术，针对神经
网络训练得到的高维星系表征数据进行降维，可视化呈现星系图片本身潜
在的全局结构和局部结构信息，寻找离群点，尝试分析星系形态高维抽象
表征的内在联系与规律。

1.3.2 章节安排

第 1 章绪论，概述星系形态分类的研究背景和意义，介绍了国内外研究现状，
提出本文主要研究内容。

第 2 章星系形态分类标准，介绍了主要的星系形态分类标准，重点介绍了星
系动物园决策树。

第 3 章深度卷积神经网络，简要介绍了深度卷积神经网络基本概念，详细介
绍了 LeNet5、AlexNet、VGG、Inception 和 ResNet 等 5 类模型及其各自的特
点。

第 4 章 ResNet-26 网络结构设计与实验结果分析，详细介绍 ResNet-26 网
络模型的网络结构设计、数据集选取、数据预处理和实验结果及分析。

第 5 章星系图像表征与结果分析，对高维数据可视化技术（t-SNE）进行了概要介绍，着重描述了星系图像数据集的 t-SNE 可视化以及结果分析情况。

第 6 章总结与展望，对本文研究工作进行总结，分析创新点及不足，展望未来工作。

第 2 章 星系形态分类标准

不同的星系分类方法依据的分类标准不尽相同，主要的星系形态分类标准有目视分类系统、模型化分类系统和非模型化分类系统等 (汪敏 等, 2007)。此外，以英国牛津大学为首的研究机构推出的基于哈勃序列思想而设计 (Zooniverse) 的星系动物园也取得了很好的应用效果。

2.1 目视分类系统

直接凭借眼睛来对星系的形态进行分类，称为目视分类系统。目视分类系统中最著名的是 1926 年哈勃提出的“哈勃序列”，又称“哈勃音叉图”，如图 2.1 所示。他凭借不到 400 张星系图像，对星系形态进行了分类 (Hubble, 1926)。随后，一系列工作对哈勃序列进行了发展和细化，使得哈勃分类日趋完善并被广泛采用 (Hubble, 1936; Sandage, 1961; Sandage et al., 1994; 汪敏 等, 2007)。

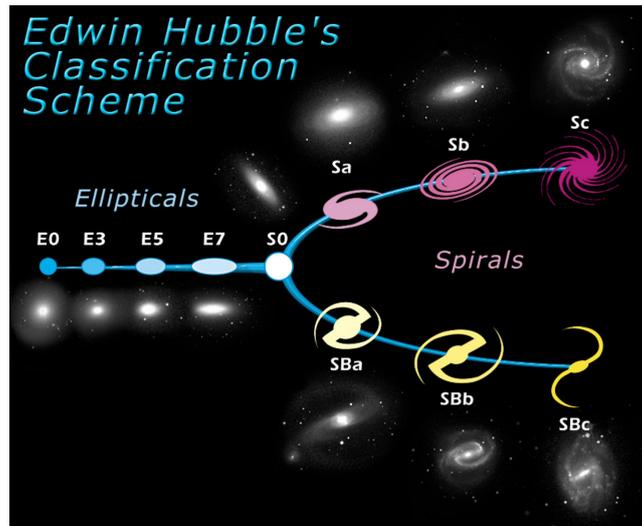


图 2.1 哈勃序列

Figure 2.1 "Hubble tuning-fork" galaxy classification scheme[ESA/Hubble]

哈勃序列可以概括为：1) 从形态上星系可以分为椭圆星系 (E)、透镜星系 (S0)、旋涡星系 (S) 和不规则星系 (Ir)；2) 椭圆星系按照外观的扁平程度从 E0 (圆) 至 E7 (扁)，“E”代表椭圆，数字 $n = 10(a - b)/a$ ， a, b 分别椭圆的半长轴和半短轴；3) 根据星系中心是否有棒 (bar) 存在分为正常旋涡星系 (S) 和棒旋星系 (SB)；4) 正常旋涡星系和棒旋星系根据其中旋臂的束缚程度和核球的相对大小等可分为 a、b 和 c 3 个子类；5) 透镜星系为从椭圆星系 E7 到旋涡星系

Sa/SBa 的过渡星系 (Van den Bergh, 1998; 汪敏 等, 2007)。

尽管哈勃序列仅仅是根据星系的光学波段形态特征来进行分类, 但随后的研究发现, 哈勃序列和许多星系的物理参数有关, 如: 中性氢的质量、动力学特征 (恒星速度弥散和转动曲线)、星系的积分颜色、星系光度和环境等 (Lahav et al., 1994), 这也是哈勃序列至今仍被广泛使用的一个重要原因。然而, 哈勃序列也存在着缺陷, 如: 哈勃序列不适合用于低面亮度星系分类, 对螺旋星系分类比较粗糙; 而且, 哈勃分类也不能对活动星系、相互作用星系进行分类 (Abraham et al., 2003)。后来的天文学家在哈勃序列的基础上进行发展和细化, 提出了德沃库勒分类系统 (De Vaucouleurs System)、叶凯士分类系统 (Yerkes Scheme) 和范登伯分类系统 (van den Bergh Scheme) 等分类方法 (汪敏 等, 2007)。

目视分类系统可以很好地用于近邻星系形态分类, 但存在一些缺点: 一, 对于同一星系, 由于人为的主观性, 不同的人给出的分类结果可能不同; 二, 大规模巡天使得观测数据极大地丰富, 当存在大样本星系数据集 (大于 1000) 的样本时, 目视分类非常耗时; 最后, 受到望远镜空间分辨率的限制, 从图像上很难区分高红移星系的结构, 此方法不适用于高红移星系的分类 (汪敏 等, 2007)。

2.2 模型化分类系统

模型化分类系统使用星系的面亮度轮廓来对星系形态进行分类, 不同形态星系的面亮度 (即星系每单位面积发射的视光强度) 分布是不同的。比如椭圆星系的面亮度轮廓可以用 $r^{1/4}$ (r 是星系面亮度分布半径) 律拟合, 旋涡星系的核球部分面亮度轮廓和椭圆星系相似, 可以用 $r^{1/4}$ 律拟合; 星系盘部分可以用指数律拟合。所以可以用不同的经验律 ($r^{1/4}$ 律或 $r^{1/4}$ 律 + 指数律) 来拟合星系的面亮度轮廓, 判断出星系是核球主导的, 还是盘主导的, 从而有效地区分早型星系和晚型星系 (汪敏 等, 2007)。

模型化分类系统可以用于大样本星系形态分类, 但它也有着自身的局限性: 一, 它依赖于假设的函数分布; 二, 它没有考虑星系中存在棒、透镜或环等成份。最后, 对于高红移星系, 限制于望远镜的观测分辨率, 很难区分核球和盘 (汪敏 等, 2007)。

2.3 非模型化分类系统

非模型化分类系统是基于表示星系形态的结构参数来对星系形态进行分类, 其参数包括聚集度指数 (Concentration, C)、非对称指数 (Asymmetry, A)、簇

聚指数 (Smoothness, S)、基尼系数 (Gini coefficient, G) 和矩指数 (Moment, M_{20}) 等。这种方法不需要假设星系光度分布满足一定的函数分布, 由计算机自动测量参数, 适合大星系样本的星系形态的研究, 具有丰富的物理内涵, 与星系的恒星形成历史 (如 C)、星系相互作用 (如 A 、 M_{20}) 和小尺度结构 (如 S) 等相关, 早型椭圆状星系有较大的 G 和 C 值、很小的 A 、 M_{20} 和 S 值 (汪敏等, 2007), 但是仍然属于半自动化分类方法, 计算机测量后的某一个或几个形态学参数, 仍需要天文学家进行综合判断, 确定星系所属类别。

2.4 星系动物园决策树

星系动物园 (Galaxy Zoo, GZ) 是一个在线众包项目, 它邀请大量的志愿者根据所给星系彩色图像的形态进行分类 (Lintott et al., 2008, 2010)。由于本文采用星系动物园 2 (Galaxy Zoo 2, GZ2) 的数据集和分类标准, 所以以 GZ2 为例, 介绍 GZ2 决策树, 见图 2.2 和表 2.1。对于一张星系图片的完整分类过程就是从决策树的根节点问题回答起, 根据答案顺着决策树往下走, 一直走到决策树的叶子节点。例如, 首先会被问 T01 “Is the galaxy simply smooth and rounded, with no sign of a disk”, 有 3 个答案: “smooth” (平滑的)、“features or disk” (有特征或者盘状) 和 “star or artifact” (恒星或者人造物体)。如果选择 “smooth”, 然后就会被问到 T07 “How rounded is it?” 即圆的程度如何, 分为 “completely round” (完全圆的)、“in between” (介于两者之间的) 和 “cigar-shaped” (雪茄状的); 如果没有选择 “smooth”, 那么 “smooth” 下面的问题不会被问到。选择 “features or disk”, 继续会被问到 T02 “Could this be a disk viewed edge-on” (是否为侧向星系), 依次回答下去, 直到决策树叶子节点。每一个志愿者回答的问题都是全部问题的一个子集。

GZ2 决策树有 11 个分类任务, 总共 37 个应答, 每一个问题只能选择一个答案应答。其中 T01 和 T06 是每一张图片分类时都会遇到的问题。每个志愿分类的图片都是随机出现的, 大约有 50 万志愿者参与到 GZ2 的分类中, 每一张图片大概被 40 到 50 人分类, 计算累计得分值。星系动物园的分类结果与专业天文学家的分类结果高度一致, 说明了其结果的可靠性和权威性 (Lintott et al., 2008)。GZ2 决策树是传统哈勃序列的一个更细粒度的分类标准, 反映着星系更精细的形态结构。

GZ2 的数据发布白皮书 Willett et al. (2013) 中关于数据的使用有 2 种方式: 一、使用校正后的阈值选取干净样本, 如表 2.2 所示。这些阈值满足于一定数目

表 2.1 GZ2 决策树 (Willett et al., 2013)

Table 2.1 The GZ2 decision tree

Task	Question	Responses	Next
01	Is the galaxy simply smooth and rounded, with no sign of a disk?	smooth	07
		features or disk	02
		star or artifact	end
02	Could this be a disk viewed edge-on?	yes	09
		no	03
03	Is there a sign of a bar feature through the centre of the galaxy?	yes	04
		no	04
04	Is there any sign of a spiral arm pattern?	yes	10
		no	05
05	How prominent is the central bulge, compared with the rest of the galaxy?	no bulge	06
		just noticeable	06
		obvious	06
		dominant	06
06	Is there anything odd?	yes	08
		no	end
07	How rounded is it?	completely round	06
		in between	06
		cigar-shaped	06
08	Is the odd feature a ring, or is the galaxy disturbed or irregular?	ring	end
		lens or arc	end
		disturbed	end
		irregular	end
		other	end
		merger	end
dust lane	end		
09	Does the galaxy have a bulge at its centre? If so, what shape?	rounded	06
		boxy	06
		no bulge	06
10	How tightly wound do the spiral arms appear?	tight	11
		medium	11
		loose	11
11	How many spiral arms are there?	1	05
		2	05
		3	05
		4	05
		more than four	05
		can't tell	05

注: GZ2 决策树中包含 11 个任务和 37 个应答, 表 2.1来自Willett et al. (2013) 中的表 2。

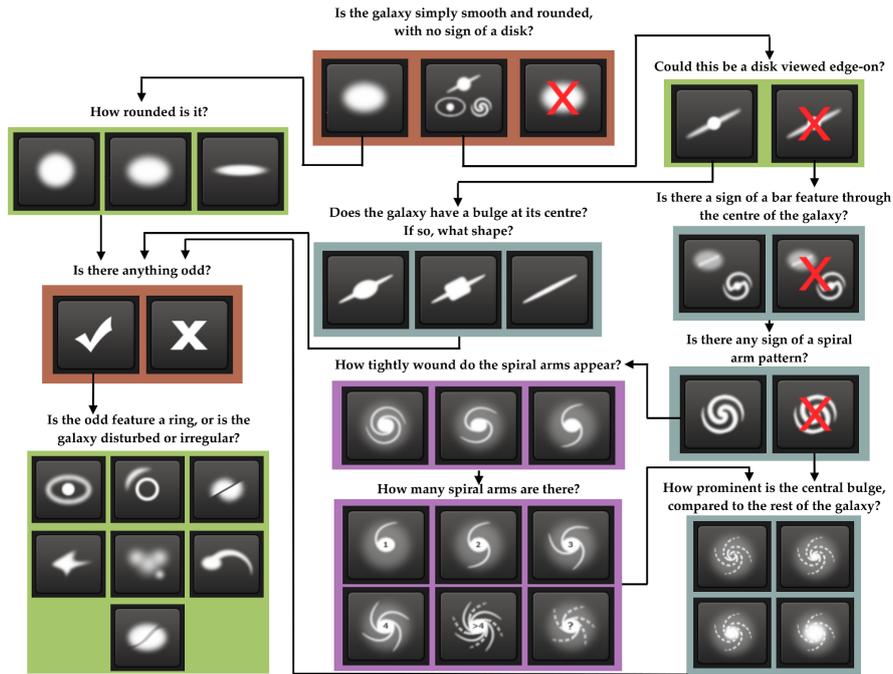


图 2.2 GZ2 决策树 (Willett et al., 2013)

Figure 2.2 The Galaxy Zoo 2 decision tree

注：图 2.2来自Willett et al. (2013) 中的图 1。其中褐色的问题每一张星系图片都会被问到。表 2.1对应图中的每一个任务和应答。

的志愿者投票数累计，如必须满足 20 人对同一张星系图片分类，然后计算累计投票分数，然后作校正。最后，如果要选取旋涡星系，则必须满足 3 个条件： $f_{features/disk} \geq 0.430$ ， $f_{edge-on,no} \geq 0.715$ ， $f_{spiral,yes} \geq 0.619$ 。阈值的选取规则通常被认为是保守的，可以保证选取到干净的样本数据。二、直接使用发布的概率值。

星系动物园项目依次发布了第一代 (GZ1)、第二代 (GZ2)、第三代 (Galaxy Zoo: Hubble) (Willett et al., 2016) 和第四代星系动物园 (Galaxy Zoo: CANDELS) (Simmons et al., 2016)。其分类决策树设计也更加细粒度化，更多的星系形态结构可以被识别出来，同时产生了带标签的大规模高质量星系图像数据集。

表 2.2 GZ2 干净样本阈值 (Willett et al., 2013)

Table 2.2 Thresholds for determining well-sampled galaxies in GZ2

Task	Previous task	Vote fraction
01	-	-
02	01	0.430
03	01,02	0.715
04	01,02	0.715
05	01,02	0.715
06	-	-
07	01	0.469
08	06	0.420
09	01,02	0.602
10	01,02,04	0.619
11	01,02,04	0.619

注：如果选取漩涡星系，则必须同时满足 3 个条件： $f_{features/disk} \geq 0.430$ ， $f_{edge-on,no} \geq 0.715$ ， $f_{spiral,yes} \geq 0.619$ ，

表 2.2来自Willett et al. (2013) 中的表 3。

2.5 本章小结

目视分类系统使用了星系的形态信息，比较明确地对星系形态进行分类。但是其存在着人为的主观性，对同一星系，不同的人可能给出不同分类结果，不适合大样本星系分类。模型化分类系统相对客观一些，它克服了目视分类系统的不足，但它需要假设星系光度分布满足一定函数分布，但实际上每个星系的光度分布都是不相同的，很难以某个或某些函数拟合。非模型化分类系统与目视分类系统和模型化分类系统相比，更客观、更有效，可以定量地测量星系形态，并能在一定程度上反映星系的形成与演化过程，但仍然属于半自动化分类方法，计算机自动测量某一个或几个形态学参数后，仍需要天文学家进行综合判断，确定星系最终所属类别。星系动物园决策树分类标准基于哈勃分类法的思想而设计，其优点是较目视分类系统分类更加细粒度化，更多的星系形态结构通过该决策树被识别出来。基于决策树的标准，产生带标签的大规模高质量星系图像数据集，可为基于深度学习方法的星系形态分类提供较好的数据集样本。

第 3 章 深度卷积神经网络

本章调研了以卷积神经网络 (Convolutional Neural Network, CNN) 为代表的深度学习技术的发展和最新研究成果, 简要介绍人工神经网络 (Artificial Neural Network, ANN) 算法原理, 卷积神经网络算法原理, 并介绍 5 种经典的 CNN 结构, 及其各自的优缺点与适用场景, 包括 LeNet5 (Cun et al., 1989)、AlexNet (Krizhevsky et al., 2012)、VGG (Simonyan et al., 2014)、Google Inception (Szegedy et al., 2015; Ioffe et al., 2015; Szegedy et al., 2016, 2017) 和 ResNet (He et al., 2016a,b; Zagoruyko et al., 2016)。

3.1 深度卷积神经网络简介

3.1.1 前馈神经网络

人工神经网络 (Artificial Neural Network, ANN) 是由具有适应性的简单单元组成的广泛并行互连的网络, 其组织可以模拟生物系统对真实世界物体所作出的交互反应 (Kohonen, 1988; 周志华, 2016)。

一个简单的 N 层前馈神经网络 (Multi-layer Feedforward Neural Network) 示意图如图 3.1 所示。网络分为输入层、隐藏层和输出层。其中, layer 1 到 layer $N-1$ 为隐藏层, 加上输入层和输出层, 总共 N 层, 称为 N 层前馈神经网络。假设 x_0 为网络的输入, x_N 为网络的输出, t 为期望输出。

设向量 x_{N-1} 为第 N 层输入, W_N 为权重矩阵, b_N 为偏置向量, 则第 N 层的输出可以表示为

$$x_N = f(W_N x_{N-1} + b_N) \quad (3.1)$$

式中, f 为激活函数 (Activation Function), 常见的激活函数有 Sigmoid 函数 ($f(x) = (1 + e^{-x})^{-1}$)、tanh、修正线性单元 ReLU ($f(x) = \max(x, 0)$) (Nair et al., 2010)、ELU (Clevert et al., 2015), PReLU (He et al., 2015) 和 RReLU (Xu et al., 2015) 等, 各种激活函数如图 3.2 所示。

神经网络的学习过程, 就是根据所拥有的数据集不断地调整更新网络权重矩阵 W_N 和偏置向量 b_N , 使得网络输出 x_N 尽可能地接近期望输出 t 。假设预测误差为 $e(x_N, t)$, 即要使目标函数 (损失函数) $e(x_N, t)$ 最小。使用梯度下降算法 (Gradient descent) 更新权重公式:

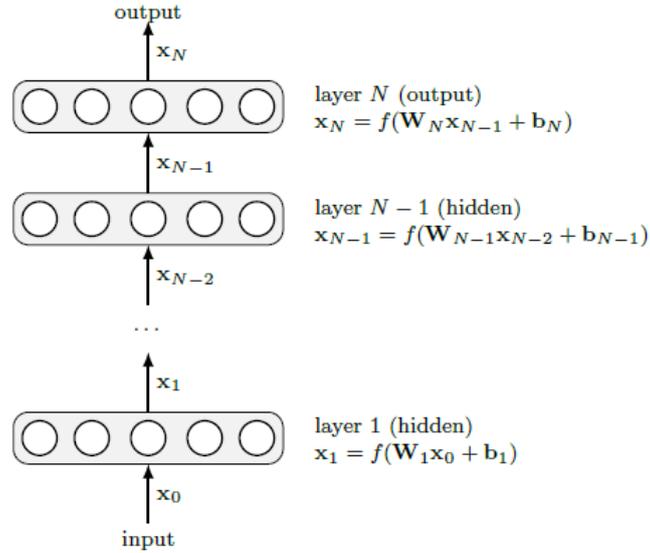


图 3.1 N 层前馈神经网络示意图

Figure 3.1 Schematic representation of a feed-forward neural network with N layers

$$W_n \leftarrow W_n - \eta \frac{\partial e(x_N, t)}{\partial W_n} \quad (3.2)$$

$$b_n \leftarrow b_n - \eta \frac{\partial e(x_N, t)}{\partial b_n} \quad (3.3)$$

式中， η 为学习率（learning rate）， $\eta \in (0, 1)$ ，学习率控制着迭代步长，若太大则容易使得训练损失函数发生震荡，太小则收敛速度慢，不容易找到最优解。对于分类问题，损失函数通常为交叉熵损失（Cross Entropy）；对于回归问题，损失函数通常为平方误差（Squared Error）。

在过去，想要训练多层神经网络是很困难的，因为通常会遭遇梯度弥散（Gradient Vanishment）、过拟合（Overfitting）、参数难以调试等问题。

梯度弥散又称梯度消失，在 ReLu 激活函数出现之前，神经网络的激活函数基本上都是用 Sigmoid 函数，当神经网络的层数较多时，Sigmoid 函数在反向传播中梯度值会逐渐减小，经过多层的传递之后会呈现指数级急剧减小，所以梯度值在传递到前面的几层时就变得非常小。这种情况下，根据训练数据的反馈来更新神经网络的权重将会非常缓慢，甚至起不到训练的作用。所以在 2006 年以前，实际应用中的神经网络通常不超过 2 个隐藏层。随着 2006 年，Hinton 等人提出逐层非监督预训练（pre-training）的方法来初始化神经网络的权重，单独训练多个受限玻尔兹曼机（Restricted Boltzmann Machine, RBM）得到相应的网络

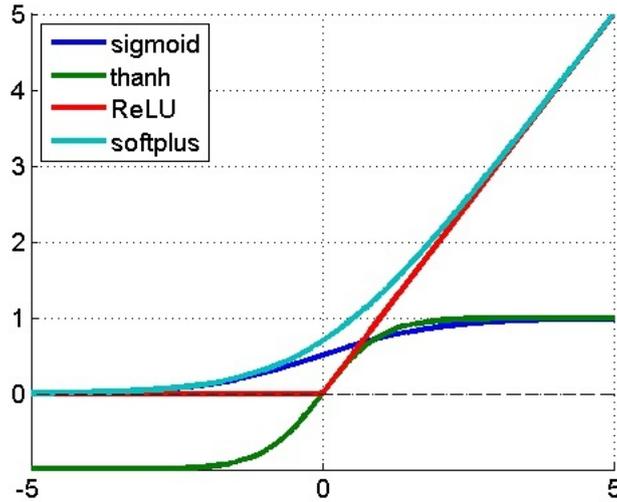


图 3.2 各种激活函数曲线图

Figure 3.2 Activation function curve

参数，然后将多个 RBM 堆叠，用训练得到的网络参数作为整体网络的初始化参数，然后通过微调方式，多层神经网络的训练就变得容易。这种预训练加微调的方式在一定程度上解决了梯度弥散问题，当然多层神经网络训练的成功也要归功于当年计算机计算能力的提升。直到 ReLU 激活函数的出现，梯度弥散问题才被完美解决 (Nair et al., 2010; Glorot et al., 2011)。ReLU 是一个非常简单的非线性函数 $f(x) = \max(x, 0)$ ，它可以比较好地传递梯度，多层反向传播之后，梯度仍然不会大幅缩小，所以非常适合用于训练深度神经网络。ReLU 函数从正面解决了梯度弥散问题，使得非监督的逐层预训练的方法变得不再必要。

过拟合问题是机器学习领域中的常见问题，指模型预测准确率在训练集上不断增长，但是在测试集上反而下降，这说明模型的泛化能力不好，模型只是记忆了当前数据的特征，不具备推广能力。在神经网络中，经常出现网络参数比训练数据还要多，就会出现过拟合现象。过拟合解决方法之一为 Dropout (Hinton et al., 2012; Srivastava et al., 2014)。Dropout 是一种正则化方法，在训练时，将神经网络的某一层的输出节点以概率 $1 - p$ 随机丢弃 (设置为 0)， p 值一般为 0.5。测试时，所有值都被保留。这种方法的本质是创造大量新的随机样本，增大样本量、减少特征数量以防止过拟合。同时防止过拟合的方法还有 L2 正则化、数据增强 (Data Augmentation) (Krizhevsky et al., 2012) 和批正则化 (Batch Normalization, BN) (Ioffe et al., 2015) 等。

神经网络参数难以调试是神经网络的另一大痛点。神经网络不是一个凸优化

问题，使用随机梯度下降法（Stochastic Gradient Descent, SGD）得到的解通常不是全局最优解，而是局部最优解，而且初始学习率设置不同，最终得到的解有可能差异巨大。有研究表明，神经网络可能有多个局部最优解可以达到很好的分类效果，但是全局最优解却反而容易引起过拟合。通常在训练时，学习率设置，一般在初始时比较大一些，可以帮助加速收敛，一段时间后，逐步减小学习率，可以使神经网络比较稳定地达到一个局部最优解。常见的梯度下降算法有 SGD、Adagrad、Adam (Kingma et al., 2014)、Adadelata (Zeiler, 2012) 等。

3.1.2 深度卷积神经网络

卷积神经网络（Convolutional Neural Network, CNN）是一种用来处理具有类似网格结构数据，如时间序列数据（时间轴上有规律采样形成的一维网格）和图像数据（二维像素网格）的神经网络 (LeCun et al., 1989; Goodfellow et al., 2016)。CNN 可以直接使用图像的原始像素作为输入，自动地提取有效特征，与传统机器学习算法相比，不需要将特征提取和分类两个过程分开，避免了复杂的特征工程，同时具有对缩放、平移、旋转等畸变不变性，有着很强的泛化性。

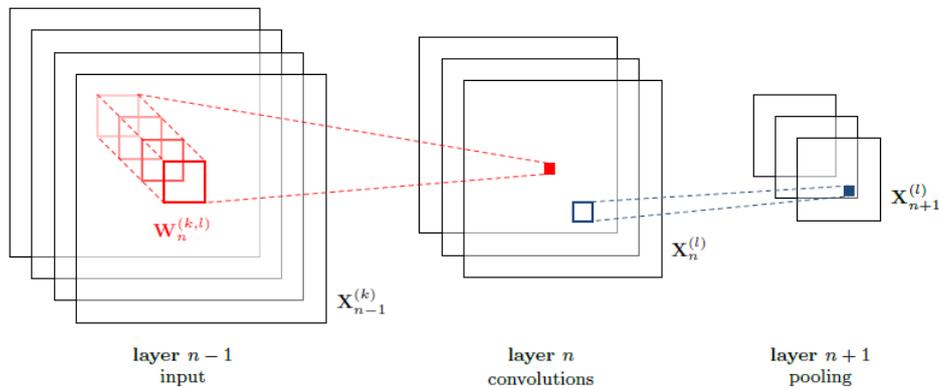


图 3.3 卷积层和池化层示意图

Figure 3.3 Schematic of a convolutional layer followed by a pooling layer

卷积神经网络一般包含卷积层、池化层和全连接层：

卷积层（Convolutional Layer）在卷积层，上一层的特征映射（Feature Map）输入后，每个卷积核都去与它做卷积，卷积核以一定步长在特征映射图上进行“滑动”，每滑动一次做一次卷积运算，卷积操作是一种特殊的线性运算，最终就得到此层的一个 Feature Map，这样每个特征映射可能与上层的几个特征映射建立关系。每一个卷积核可提取一种特征，有 k 个卷积核就能得到 k 个特征映射，就表示提取了 k 种特征。假设第 n 层的输入为 K 个矩阵 $X_{n-1}^{(k)}$ ，其中，

$k = 1 \dots K$ 。每一个矩阵表示不同的输入特征映射。输出特征映射 $X_n^{(l)}$, $l = 1 \dots L$, 可以表示为

$$X_n^{(l)} = f(\sum_{k=1}^K W_n^{(k,l)} * X_{n-1}^{(k)} + b_n^{(l)}) \quad (3.4)$$

式中, f 为激活函数, 通常使用 ReLU, $*$ 表示二维卷积操作, 矩阵 $W_n^{(k,l)}$ 表示第 n 层的卷积核, $b_n^{(l)}$ 为特征映射 l 的偏置项。图 3.3 左部分表示卷积层操作。

池化层 (Pooling Layer) 池化函数使用某一位置相邻输出的总体特征来代替网络在该位置的输出。例如, 最大池化函数 (max pooling) 给出相邻矩形区域内的最大值, 保留最显著的特征。其他常用的池化函数如平均池化函数 (average pooling) 等。图 3.3 右部分表示池化层操作。

全连接层 (Fully Connected Layer) 全连接层通常在最后一个池化层或卷积层之后, 全连接层的每个神经元将和上层的全部神经元相连接, 这也是全连接层名字得由来。

卷积神经网络有局部连接 (Local connection)、权值共享 (Weight Sharing) 和下采样 (Down-Sampling) 三大特点。局部连接和权值共享降低参数量, 使训练复杂度大大下降, 并减轻了过拟合。而且权值共享还使 CNN 模型具有对平移的容忍性, 而下采样进一步降低了网络参数, 并使模型具有对轻度形变的容忍性, 提高了模型的泛化能力。

目前卷积神经网络已经成为计算机视觉领域主流的方法, 其主要版本包括 LeNet5 (Cun et al., 1989; LeCun et al., 1998)、AlexNet (Krizhevsky et al., 2012)、VGG (Simonyan et al., 2014)、Inception (Szegedy et al., 2015; Ioffe et al., 2015; Szegedy et al., 2016, 2017)、ResNet (He et al., 2016a,b) 和 DenseNet (Huang et al., 2016) 等。

3.2 LeNet5

LeNet5 由 CNN 之父 Yann LeCun 于 1997 年提出, 是最早的卷积神经网络之一, 奠定了现代卷积神经网络的基础, 并且推动了深度学习领域的发展 (Cun et al., 1989; LeCun et al., 1998)。LeNet5 第一次提出了多层级联的卷积结构, 来对手写数字进行有效分类。LeCun 认为, 图像的特征分布在整张图像上, 可学习参数的卷积层可以用少量参数就能在图像的多个位置上提取相似特征。由于图像具有很强的空间相关性, 像素不应该被使用在第一层, 而使用图像中独立的像素作为不同的输入特征则利用不到这些相关性。LeNet5 如图 3.4 所示。

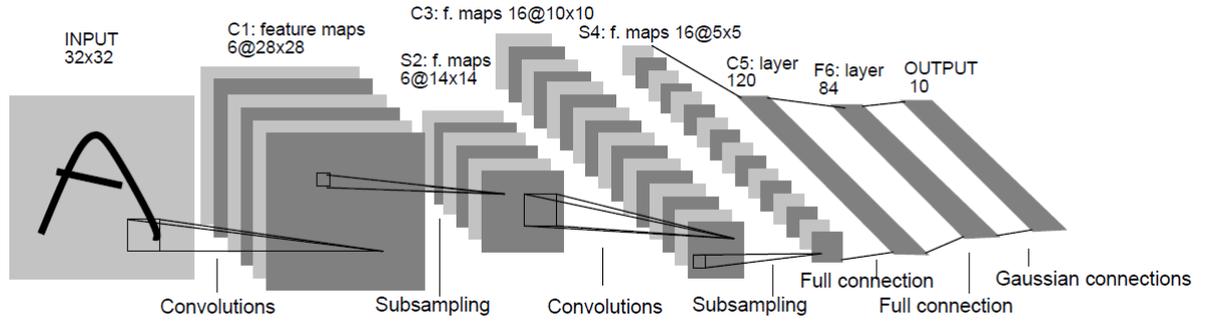


图 3.4 LeNet-5 结构示意图 (Cun et al., 1989)

Figure 3.4 Schematic representation of LeNet-5

LeNet5 的主要特点有：

- 每个卷积层包括 3 部分：卷积、池化和非线性激活函数；
- 使用卷积提取空间特征；
- 下采样的平均池化层；
- S 型（Sigmoid）或双曲正切（tanh）的非线性激活函数；
- 多层神经网络（MLP）作为最后的分类器；
- 层与层之间稀疏连接，减少计算复杂度。

3.3 AlexNet

2012 年,Hiton 的学生 Alex Krizhevsky 提出了深度卷积神经网络模型 AlexNet,使深度学习取得了突破性发展 (Krizhevsky et al., 2012), 确立深度卷积神经网络在计算机视觉领域的统治地位。AlexNet 有 6000 万个参数和 65 万个神经元, 5 个卷积层, 其中 3 个卷积层后连接了最大池化层, 最后还有 3 个全连接层。AlexNet 模型如图 3.5所示。

AlexNet 的主要特点有：

- 使用 ReLu 激活函数, 成功解决 Sigmoid 激活函数在网络较深时的梯度弥散问题;
- 使用 Dropout 随机丢弃一部分神经元, 以避过拟合;
- 使用重叠最大池化, 以避免平均池化的模糊化效果;

- 提出 LRN (Local Response Normalization) 层, 对局部神经元的活动创建竞争机制, 使得其中响应比较大的值变得相对更大, 并抑制其他反馈小的神经元, 增强了模型的泛化能力;
- 使用 CUDA 加速神经网络训练, 利用 GPU 强大的并行计算能力, 处理大量矩阵运算;
- 使用了数据增强, 包括裁剪、水平翻转等, 防止过拟合。

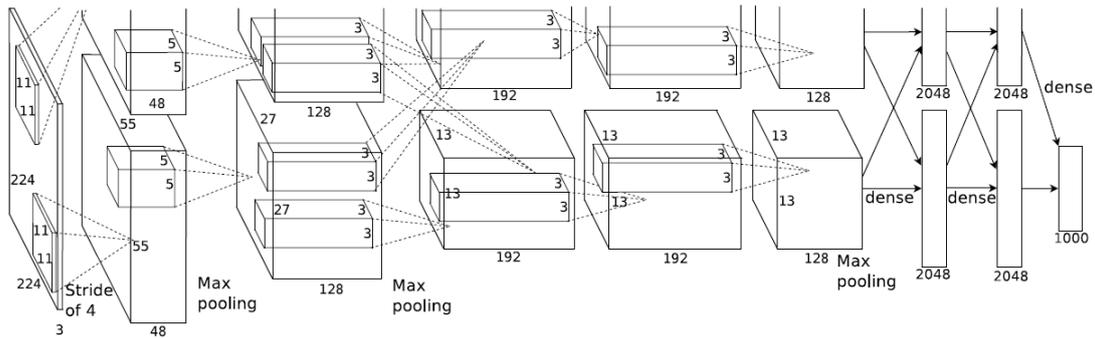


图 3.5 AlexNet 网络结构 (Krizhevsky et al., 2012)

Figure 3.5 AlexNet architecture

3.4 VGG

VGG 是由牛津大学计算机视觉组 (Visual Geometry Group) 和 Google DeepMind 2014 年提出的深度卷积神经网络结构 (Simonyan et al., 2014)。VGG 主要研究了卷积神经网络的深度与其性能之间的关系, 结构非常简单, 全部使用相同大小的 3×3 的小型卷积核和 2×2 的最大池化层, 形成了著名的 VGG-16 和 VGG-19。VGG 的迁移性很强, 拓展到其他图片数据上的泛化性也非常好。VGG 模型如图 3.6 所示。

虽然 VGG 模型参数比 AlexNet 多很多, 但是反而只需要较少的迭代次数就可以收敛, 主要是因为更小的卷积核和更深的网络带来的隐式的正则化效果, 凭借其相对不算很高的复杂度和优秀的分类性能, 成为一代经典的卷积神经网络, 直到现在依旧被广泛使用。

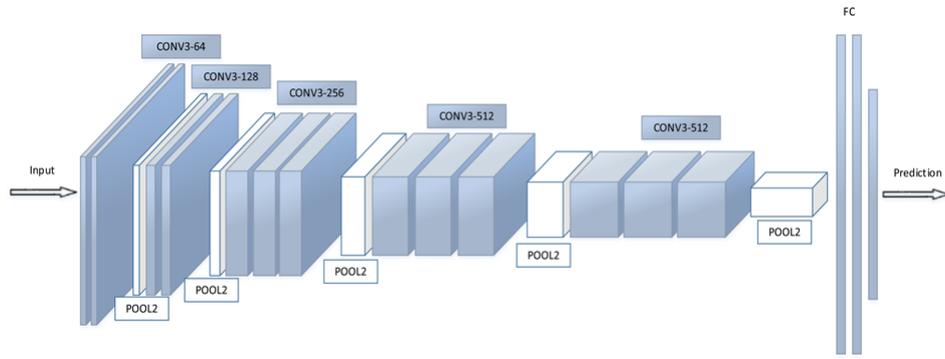


图 3.6 VGG 结构图

Figure 3.6 VGG architecture

3.5 Google Inception

Inception 由 Google 公司于 2014 年首次提出，在 ILSVRC 2014 比赛中以较大优势获得冠军 (Szegedy et al., 2015)。Inception V1 的主要特点是在控制计算量和参数量的同时，获得了非常好的分类性能。Inception V1 提出的 Inception Module 可以反复堆叠，组成高效的卷积网络结构，有 22 层深，15 亿次浮点运算，500 万参数量，仅为 AlexNet 的 1/12，却远胜于 AlexNet 的准确率。Inception V1 如图 3.7 所示。Inception 现已发展形成 4 个版本：Inception V1、V2、V3 和 V4 (Szegedy et al., 2015; Ioffe et al., 2015; Szegedy et al., 2016, 2017)。

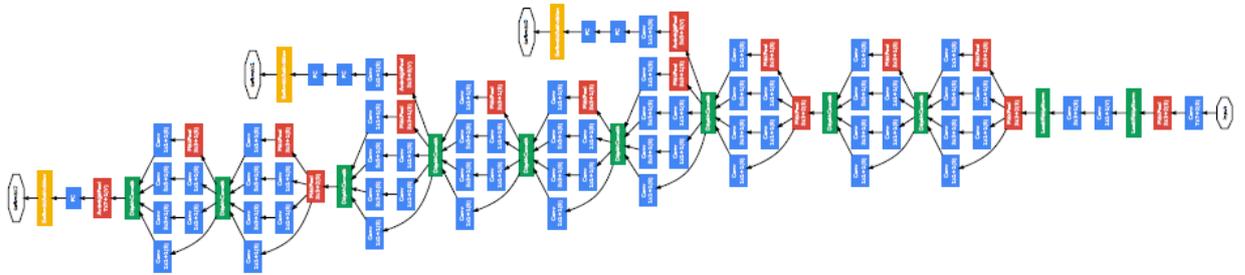


图 3.7 Inception V1 结构 (Szegedy et al., 2015)

Figure 3.7 Inception V1 architecture

Inception V1 的主要特点有：

- 模型层数深，22 层；
- 去除最后的全连接层，使用平均池化层替代，即将图片尺寸变为 1×1 ，减少了大量参数，并且减轻了过拟合；
- 精心设计了 Inception Module，提高了参数的利用率，使用 1×1 、 3×3 和

5×5 卷积核的并行组合，扩充网络的深度和宽度，提升准确率，防止过拟合。Inception Module 结构图见图 3.8；

- 使用了 Multi-Scale、Multi-Crop 等数据增强方法；
- 添加辅助分类节点（Auxiliary Classifiers），作模型融合。

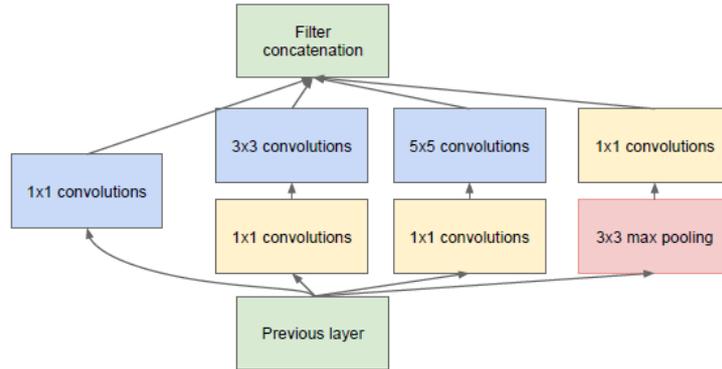


图 3.8 Inception Module 结构图 (Szegedy et al., 2015)

Figure 3.8 Inception Module architecture

Inception V2 学习了 VGG，用两个 3×3 卷积代替了 5×5 的大卷积，同时提出了著名的批正则化 BN，大大加速了训练过程，大幅提高收敛后的分类准确率，有效提升了网络性能。批正则化可以对每一个 mini-batch 的数据进行标准化（Normalization）处理，使输出标准化到 $N(0, 1)$ 的高斯分布。

Inception V3 为 42 层的深度卷积神经网络结构，继续优化网络结构，提出了 Factorization in Small Convolutions 的思想，将一个较大尺寸的二维卷积拆成多个小卷积乃至一维卷积，如将 3×3 卷积拆分成 1×3 卷积核 3×1 卷积。同时优化了 Inception Module 的结构，使用了 35×35 、 17×17 和 8×8 三种不同结构。

Inception V4 融合了 Inception 优良的网络结构，和残差网络的残差学习模块，集两个方向之长，大大提升了网络的性能。

3.6 ResNet

深度残差网络（Residual Network, ResNet）(He et al., 2016a,b) 由微软研究院的何凯明等人提出，通过使用残差单元（Residual Unit），成功训练 152 层深的神经网络，在 ILSVRC 2015 比赛中获得了冠军。残差网络通过构建极深的神

神经网络，可以极快地加速网络训练，使模型的准确率得到很大的提升。ResNet 是深度学习领域中一个里程碑式的突破，真正意义上支持了极深神经网络的训练。ResNet 以其优异的性能已经被广泛应用于图像分类、图像检测、图像定位、图像分割等任务中。

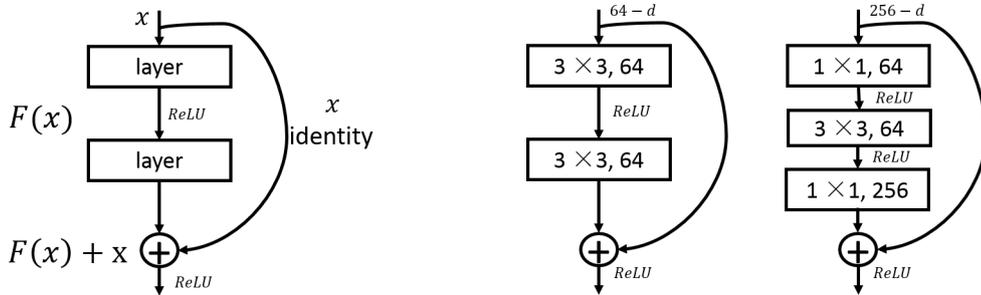


图 3.9 ResNet 的残差单元 (He et al., 2016a) 图 3.10 两层及三层的 ResNet 残差单元 (He et al., 2016a)
 Figure 3.9 Residual learning: a building block Figure 3.10 A deeper residual function F

ResNet 的核心在于残差学习模块，其基本思想为：允许原始输入信息直接传输到后面的层中，从而避免卷积层数过多引起的准确率先上升然后达到饱和，再持续增加深度导致下降的问题。假设某段神经网络的输入是 x ，期望输出是 $H(x)$ ，如果直接把输入 x 传到输出作为初始结果，那么需要学习的目标就是 $F(x) = H(x) - x$ 。如图 3.10所示，就是一个残差网络的残差学习单元 (Residual Unit)，ResNet 不再是学习一个完整的输出 $H(x)$ ，而是学习输出和输入的差别 $H(x) - x$ ，即残差。 $F(x) + x$ 可以通过短连接将输入直接连到输出，使得后面的层可以直接学习残差，这种结构被称为短连接 “short connection”。这些短连接跳过两层或者更多层，直接执行全等映射 (Identity Mapping)。残差单元可以定义为

$$x_{l+1} = f(x_l + F(x_l, W_l)). \tag{3.5}$$

其中， x_l, x_{l+1} 分别为第 l 个残差单元的输入和输出， f 是激活函数，一般为 ReLU, $ReLU = \max\{0, x\}$ 。 F 为残差函数， W 为卷积核。

图 3.11为 34 层的残差网络 ResNet-34。从中可以看出，残差网络与原来的卷积神经网络最大的区别在于，残差网有很多旁路的支线将输入直接连到后面的层，使得后面的层可以直接学习残差，这种结构也被称为 shortcut 或 skip connections。传统的卷积层或全连接层在信息传递上，或多或少会存在信息丢失、损耗等问题。残差网在某种程度上解决了这个问题，通过直接将输入信息绕道传

到输出，保护信息的完整性，整个网络则只需要学习输入、输出差别的那一部分，简化学习目标和难度。

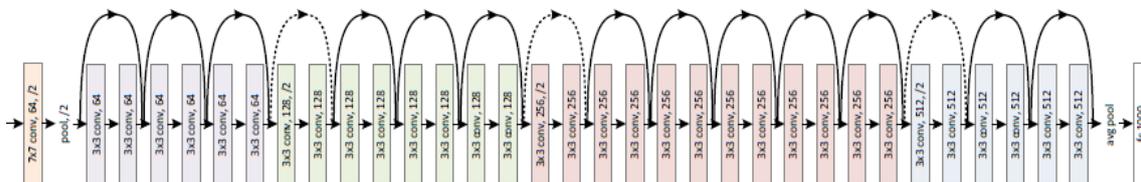


图 3.11 ResNet-34 结构示意图

Figure 3.11 Schematic representation of ResNet-34

在 ResNet 中，有两种形式的残差单元：两层残差学习单元和三层残差学习单元（又称为“bottleneck”），如图 3.10所示。两层残差学习单元中包含两个相同输出通道数的 3×3 卷积；而三层残差学习单元加入了 1×1 卷积，并且是在中间 3×3 卷积前后都使用了 1×1 卷积，有先降维再升维的作用。

随后何凯明等人提出 ResNet V2(He et al., 2016b)，发现前馈和反馈信号可以直接传输，所以 skip connection 的非线性激活函数（如 ReLu）替换为全等映射 Identity Mappings ($y = x$)。同时，每一层都使用了 Batch Normalization。新的残差单元更容易训练且泛化性更强。新的残差单元可以定义为

$$x_{l+1} = x_l + F(x_l, W_l). \quad (3.6)$$

而且 ResNet V2 中，采用预激活（“pre-activation”），即用“BN-ReLU-Conv”取代了传统的“Conv-BN-ReLU”。目前残差网络已经发展了很多版本，如 ResNet-50/101/152/200，更深的达到了 1001 层。由于残差网络瘦且高，通常应用过程中会面临特征复用减少等问题 (Zagoruyko et al., 2016)。

3.7 本章小结

通过对 LeNet5、AlexNet、VGG、Inception 和 ResNet 五个 CNN 模型的分析，我们可以看到它们都在各自的年代率先使用了先进的卷积结构，深度和复杂度依次递进，模型的准确率也越来越高。ResNet 是目前最先进的卷积神经网络，但也存在着特征复用减少、大量的残差单元堆叠并没有学到同等丰富特征等问题。为了设计适合星系形态分类的网络结构，可以尝试减少网络深度，增加网络宽度，同时使用一些正则化方法防止过拟合，以达到提升网络性能的目的。

第 4 章 ResNet-26 网络结构设计与实验结果分析

在上一章中详细介绍了深度卷积神经网络模型，本章将在深度残差网络的基础上，设计了基于改进的深度残差网络结构 ResNet-26，并进行星系形态分类实验验证，同时实现 5 种流行的卷积神经模型与之对比，分析实验结果。

4.1 网络结构设计

本文在 ResNet V2 的基础上，尝试改进残差单元，减少网络层数，增加网络宽度，同时结合星系图片本身的特点，提出了一种改进的深度残差网络模型应用于星系形态分类。改进的模型结构如表 4.1 和图 4.1 所示。

表 4.1 改进的深度残差网络结构

Table 4.1 Architecture of our model for Galaxy in this study

Group name	Output size	Depth
Conv 1	64×64	$6 \times 6, 64$
Max-pooling	32×32	2×2 , stride 2
Conv 2	16×16	$\begin{bmatrix} 1 \times 1, 64 \times k \\ 3 \times 3, 64 \times k \\ 1 \times 1, 256 \times k \\ 1 \times 1, 128 \times k \end{bmatrix} \times N$
Conv 3	8×8	$\begin{bmatrix} 3 \times 3, 128 \times k \\ 1 \times 1, 512 \times k \end{bmatrix} \times N$
Conv 4	4×4	$\begin{bmatrix} 1 \times 1, 256 \times k \\ 3 \times 3, 256 \times k \\ 1 \times 1, 1024 \times k \end{bmatrix} \times N$
Conv 5	4×4	$\begin{bmatrix} 1 \times 1, 512 \times k \\ 3 \times 3, 512 \times k \\ 1 \times 1, 2048 \times k \end{bmatrix} \times N$
Avg-pooling	1×1	$4 \times 4, 5 - d, softmax$

注：括号中为残差单元。 k 为加宽因子， N 为每组中的残差单元数量。本文采用 $k = 2$ ， $N = 2$ ，即网络为 26 层的残差网络。下采样被执行在 conv2, conv3, conv4 的最后一层，取步长为 2 的卷积。

第一个卷积层吸收了 Dieleman 模型优点，使用一个相对大的卷积核，尺寸为 6×6 ，通道数为 64，步长为 1，紧跟着一个 2×2 ，步长为 2 的最大池化层。最大池化层之后为 4 个卷积组：Conv2, Conv3, Conv4 和 Conv5，其中每个卷积组包含 N 个残差单元，比如，卷积组 Conv2 中的第 1 个残差单元为： 1×1 ，通道数为 $64 \times k$ 的卷积（ k 为加宽因子）， 3×3 ，通道数为 $64 \times k$ 的卷积， 1×1 ，

通道数为 $256 \times k$ 的卷积，如此堆叠成卷积组 Conv2 的 N 个残差单元。下采样被执行在卷积组 Conv2, Conv3, Conv4 的最后一层，取步长为 2 的卷积。最后一层是平均池化层，大小为 4×4 ，池化层的输出为 $1 \times 1 \times 4096$ ，最后作用于一个 5 个神经元的全连接层 *softmax*。

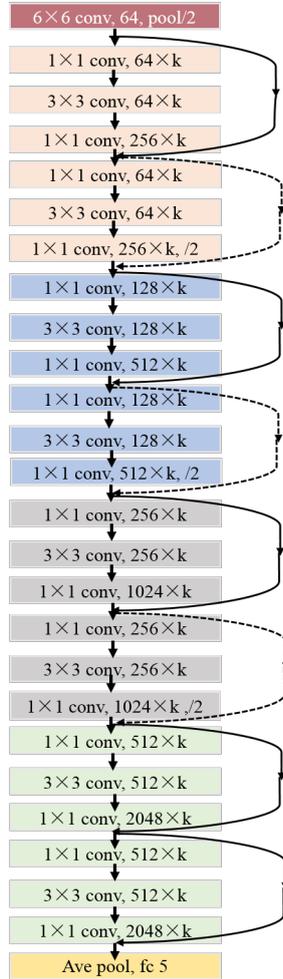


图 4.1 改进的深度残差网络结构

Figure 4.1 Our network architecture for Galaxy in this study

注： k 为加宽因子，虚线连接表示升维操作。表 4.1 显示更多细节。

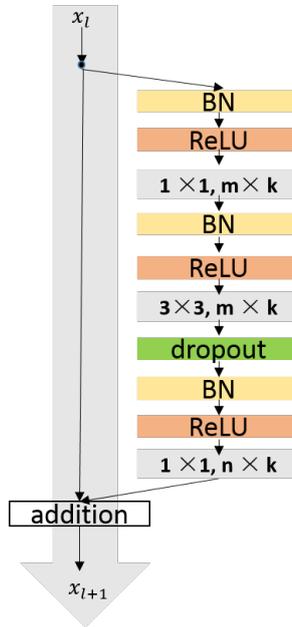


图 4.2 改进的三层残差单元

Figure 4.2 Full pre-activation residual unit in our study

注： m, n 为通道数量， k 为加宽因子。本文采用 $1 \times 1, 3 \times 3, 1 \times 1$ 卷积以及标准的预激活“BN-ReLU-Conv”方式。

改进的残差单元如图 4.2 所示，使用 3 个卷积层，分别为 $1 \times 1, 3 \times 3, 1 \times 1$ 卷积。 1×1 卷积的作用是先降维再升维，减少模型的参数并加速训练。 3×3 卷积的使用继承了 VGG 模型，可以减少模型的参数，获得更好的性能 (Simonyan et al., 2014)。前 2 个卷积层的通道数一样，第 3 个卷积层的通道数一般为前 2 个卷积层通道数的 4 倍。使用预激活“pre-activation”方式，即“BN-ReLU-Conv

”。并且在其中的 3×3 卷积之后加入 dropout(Srivastava et al., 2014), 以减轻过拟合现象。改进的残差单元可以定义为

$$x_{l+1} = x_l + W_3\sigma(W_2\sigma(W_1\sigma(x_l))) \quad (4.1)$$

其中, x_l, x_{l+1} 分别为第 l 层的输入和输出, σ 包含批正则化和激活函数 ReLU, W_1, W_2, W_3 分别表示 3 个卷积核, dropout 被放在 W_2 操作之后, 偏置 (bias) 项为了简化符号被省略。

除了最后一层全连接层使用 *softmax* 以外, 所有的卷积层都使用 ReLU 作为激活函数。训练使用的代价函数为交叉熵函数 (Cross Entropy), 采用动量梯度下降算法, Nesterov momentum 值为 0.9, batch 大小为 128。初始学习率设置为 0.1, 然后每 30k 步下降 10 倍, 即 30k 次迭代学习率变为 0.01, 60k 时降为 0.001, 72k 次迭代时停止训练。权重延迟 (weight decay) 为 0.0001, Dropout 保留率为 0.8, 网络初始化权重值采用 He et al. (2015) 的策略。在使用激活函数和卷积操作之前, 先使用批正则化。

在设计残差网络模型中, 改进了残差单元, 减少了网络的层数, 加宽了网络的宽度 (通道数目), 同时融合了 Dieleman 模型的优点, 网络的总层数为 $2N + 2$ 层。

4.2 数据集选取

本文采用的是 GZ2¹提供的数据和分类标准。其中, 实验数据集具体来自 Galaxy Zoo-The Galaxy Challenge², 这是星系动物园在 Kaggle 平台上举办的比赛。The Galaxy Challenge 的训练集全集中共包含来自于 SDSS DR7³的 61578 张带标签的星系观测 RGB 图片, 图片的尺寸大小为 $424 \times 424 \times 3$ 像素。SDSS 对星系的观测包括 5 个光学波段 (u、g、r、i 和 z), 一般取前 3 个波段 (u、g 和 r) 合成相应的 RGB 星系图像。每一张图片的标签为 1×37 的向量, 来自于 GZ2 志愿者投票分数的修正累计频率值。

在数据选取过程中, 依据 GZ2 干净样本 (well-sampled galaxies) 的定义 (Willett et al., 2013), 如表 2.2 所示。每个阈值都必须满足充足的投票数目, 即对同一张星系图片, 如必须满足 20 人以上对其分类, 然后计算累计投票分数, 原

¹<https://www.galaxyzoo.org/>

²<https://www.kaggle.com/c/galaxy-zoo-the-galaxy-challenge>

³<http://www.sdss.org/>

始的投票分数经过 Willett et al. (2013) 的校正，统计规定满足一定的阈值，就可以被认为属于某个星系类别。如某张待分类的星系是旋涡星系，在图 2.2 中，必须依次经过 T01 中的“features or disk”（有特征或者盘状），进入 T02 回答“no edge-on”（非侧向），然后进入 T03，无论回答是“yes”还是“no”，都进入 T04，回答“spiral, yes”（是旋臂），最后进入 T10，回答旋臂的松紧度，即到 T10 就可以说明这张待分类的星系为旋涡星系。这个过程对应于表 2.2 中的 Task10，所以先前的 Task 依次为 T01、T02 和 T04，以及对应的阈值分别为 0.430、0.715 和 0.619。

表 4.2 干净样本选取 (Willett et al., 2013)

Table 4.2 Clean samples selection in Galaxy Zoo 2

简称	英文名称	中文名称	任务	阈值	N_{sample}
0	Completely round smooth	圆形星系	T01	$f_{smooth} \geq 0.469$	8434
			T07	$f_{completely\ round} \geq 0.50$	
1	In between smooth	中间星系	T01	$f_{smooth} \geq 0.469$	8069
			T07	$f_{in\ between} \geq 0.50$	
2	Cigar-shaped smooth	雪茄状星系	T01	$f_{smooth} \geq 0.469$	578
			T07	$f_{cigar-shaped} \geq 0.50$	
3	Edge-on	侧向星系	T01	$f_{features/disk} \geq 0.430$	3903
			T02	$f_{edge-on,yes} \geq 0.602$	
4	Spiral	旋涡星系	T01	$f_{features/disk} \geq 0.430$	7806
			T02	$f_{edge-on,no} \geq 0.715$	
			T04	$f_{spiral,yes} \geq 0.619$	

注：T01-T11 为 GZ2 中的 11 个分类任务， f_{smooth} 为一张图片属于平滑星系的频率， $f_{completely\ round}$ 、 $f_{in\ between}$ 、 $f_{cigar-shaped}$ 刻画的是平滑星系圆的程度，分别为完全圆的、处于中间的、雪茄状的， $f_{features/disk}$ 表示有特征/盘状结构的频率， $f_{edge-on,yes}$ 、 $f_{edge-on,no}$ 分别表示侧向星系和非侧向星系的频率， $f_{spiral,yes}$ 表示旋涡星系的频率， N_{sample} 为选取到的干净样本的数目。

依据干净样本的阈值选取规则，选取 5 类星系，包括圆形星系（completely round smooth galaxy，完全圆的平滑星系）、中间星系（in between smooth galaxy，介于完全圆的平滑星系和雪茄状平滑星系之间）、雪茄状星系（cigar-shaped smooth galaxy，雪茄状平滑星系）、侧向星系（edge-on galaxy）和旋涡星系（spiral galaxy）。由于星系动物园 2 干净样本的阈值选取规则过于保守 (Willett et al., 2013)，为了选取到一定数量的样本进行模型训练与测试，将平滑星系（圆形星系，中间星系和雪茄状星系）的阈值选取标准适当放宽，从 0.8 降到了 0.5，其他所有阈值标准均依旧为干净样本中的默认取值 (Willett et al., 2013)。表 4.2 为

5 个星系类别干净样本数据的选取规则。

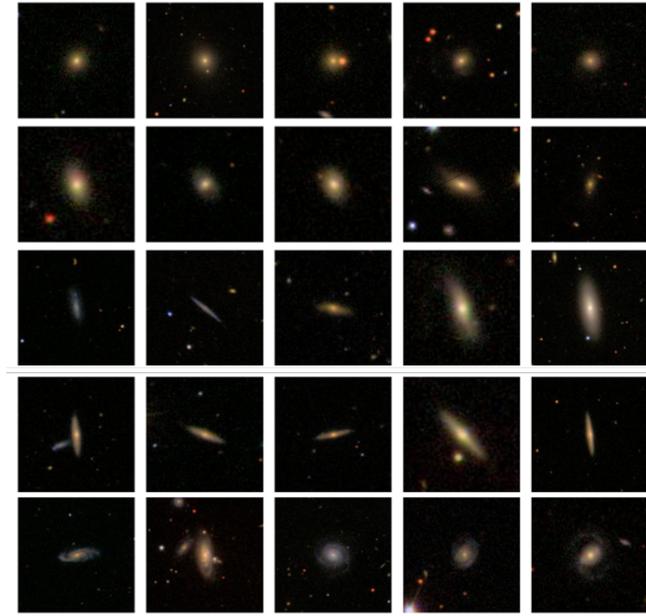


图 4.3 GZ2 中随机抽取的星系图片

Figure 4.3 Example galaxy images from the dataset

注：星系图片被从数据集中随机抽取。每一行代表一个星系类别，从上到下，GZ2 的标签分别为圆形星系（completely round smooth galaxy）、中间星系（in between smooth galaxy）、雪茄状星系（cigar-shaped smooth galaxy）、侧向星系（edge-on galaxy）和旋涡星系（spiral galaxy），以及依次被简称为类 0、1、2、3 和 4。

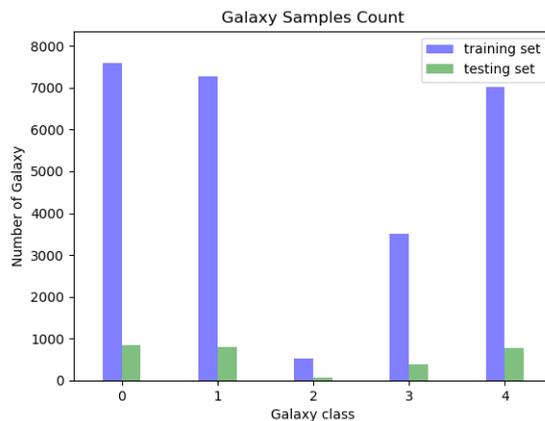


图 4.4 训练集和测试集中星系图片样本统计

Figure 4.4 Galaxy Samples Counts

经选取处理，最终获得干净样本共 28790 张星系图片，其中圆形星系、中间星系、雪茄状星系、侧向星系和旋涡星系 5 类星系分别被简称为 0, 1, 2, 3 和 4，选取后每类星系的图片数量分别为 8434, 8069, 578, 3903, 7806 张。图 4.3 为

表 4.3 训练集和测试集中星系图片样本统计

Table 4.3 Number of galaxy images in each morphological class in each set

	圆形星系 0		中间星系 1		雪茄状星系 2		侧向星系 3		旋涡星系 4		总计
	N_1	占比	N_2	占比	N_3	占比	N_4	占比	N_5	占比	
训练集	7591	29%	7262	28%	520	2%	3513	14%	7025	27%	25911
测试集	843	29%	807	28%	58	2%	390	14%	781	27%	2879
数据集	8434	29%	8069	28%	578	2%	3903	14%	7806	27%	28790

从原始数据集中随机抽取的 5 类星系图像，每一行代表一个星系类别，从上到下，依次为圆形星系、中间星系、雪茄状星系、侧向星系和旋涡星系。

将 28790 张干净样本按 9: 1 的比例划分训练集和测试集，即训练集包含 25911 张图片，测试集包含 2879 张图片。训练集用于训练模型，而测试集用于测试模型的泛化能力，即分类性能。表 4.3和图 4.4分别展示训练集和测试集中每一类星系图片数目，其中训练集和测试集中星系图片数目是同分布的。

4.3 数据预处理

选取后的训练集、测试集星系图片需进行数据预处理后方作为模型的训练和测试输入。训练集和测试集的数据预处理过程是不同的。

训练集预处理流程如图 4.5 所示。由于网络结构有上千万参数，训练集图片数量有限，仅有不到 3 万张，为了减轻过拟合现象，需使用数据增强处理方法，尽量增大训练集数量。训练集预处理的第一步是中间裁剪处理。中间裁剪到一个范围值 S ，称为训练尺寸抖动 (Training Scale Jittering) (Simonyan et al., 2014)。这样图片被随机裁剪到 S 范围中的某个值，会使不同的图片被裁剪到不同的尺寸，即使是同一张图片在不同的迭代次数时也会被裁剪到不同的尺寸。裁剪不仅可以起到降维的作用，还可以消除一些噪声如次级物体。由于数据集中几乎所有星系都位于图片的正中间，四周包含了大量的深色背景，所以选择从图片中间裁剪到一个范围值 $S = [170, 240]$ ，这样处理可在保证星系完整信息的情况下，将图片的尺寸降到了原来的四分之一。第二步，下采样处理，将训练集中的图片下采样至 $80 \times 80 \times 3$ 像素，以达到降维和节省计算资源的效果。第三步，数据增强处理，包括随机裁剪、随机旋转、水平翻转和光学畸变等。其中，随机裁剪，将图片随机裁剪至 $64 \times 64 \times 3$ ，此处理过程可将训练集图片在不改变原星系图片标签的情况下，训练集数目增大至原来的 256 倍；随机旋转，对图片随机旋转 0° , 90° , 180° , 270° ，由于星系图片具有旋转不变性，旋转之后星系图片的标签不会

改变，数据量可增大至原来的 4 倍。水平翻转亦不会改变星系图片的标签，数据量可增大至原来的 2 倍；光学畸变，包括改变亮度、对比度、饱和度、色调等。最后一步，图像白化处理，图像中相邻像素之间具有很强的相关性，作图像白化处理以降低输入数据的冗余性。最后网络输入的图片大小为 $64 \times 64 \times 3$ 。

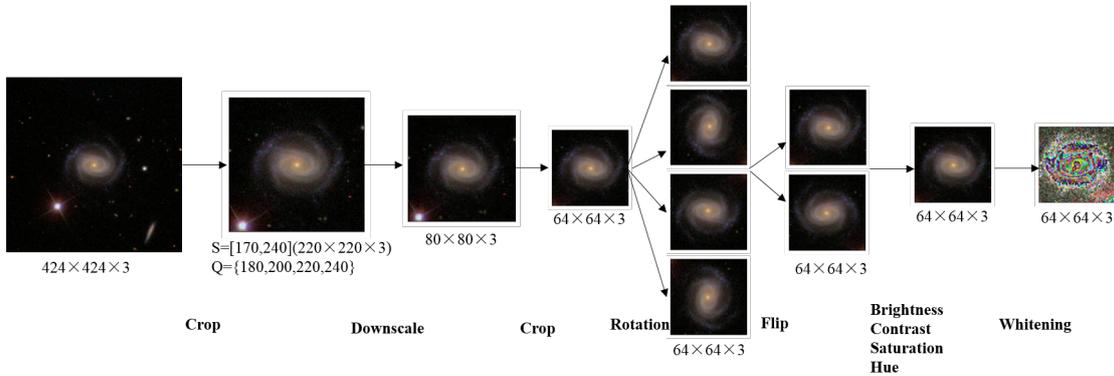


图 4.5 训练集预处理流程图

Figure 4.5 Preprocessing procedure

注：原始的图片训练时被中间裁剪到一个范围值 $S = [170, 240]$ （测试时， $Q = \{180, 200, 220, 240\}$ ）。例如，图中的旋涡星系（GalaxyID: 237308）首先被裁剪到 $220 \times 220 \times 3$ 像素，然后下采样至 $80 \times 80 \times 3$ 像素，随机裁剪到 $64 \times 64 \times 3$ ，随机旋转 $0^\circ, 90^\circ, 180^\circ, 270^\circ$ ，水平翻转，光学畸变，图片白化。最后图片的尺寸为 $64 \times 64 \times 3$ ，作为网络的输入。

测试集预处理步骤相对简单。第一步，中间裁剪处理，将图片裁剪到 4 个具体的值 $Q = \{180, 200, 220, 240\}$ ，称为测试尺度抖动（Testing Scale Jittering）。第二步，下采样处理，将测试集中的图片下采样至 $80 \times 80 \times 3$ 像素。第三步，中间裁剪处理，将图片裁剪到 $64 \times 64 \times 3$ 。最后一步，图像白化处理。最后得到网络的测试输入。

4.4 实验结果及分析

4.4.1 实验评价指标

在机器学习领域，分类模型主要用准确率（Accuracy）、精确率（Precision）、召回率（Recall）、F1 值、混淆矩阵（Confusion Matrix, CM）、ROC 曲线（receiver operating characteristic）和 AUC（Area under curve）等评价指标进行其性能评价。

准确率(Accuracy): 设 $Y = \{y_1, y_2, \dots, y_i, \dots, y_n\}$, $\hat{Y} = \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_i, \dots, \hat{y}_n\}$, y_i 、 \hat{y}_i 分别为第 i 个样本的真实标签和预测标签， n 为总的样本数量。准确率可

以定义为:

$$Accuracy(Y, \hat{Y}) = \frac{1}{n} \sum_{i=0}^{n-1} 1(\hat{y}_i = y_i). \quad (4.2)$$

精确率 (Precision), **召回率 (Recall)** 与 **F1 值** (Ceri et al., 2013): 给定真正类 (True Positive, TP): 预测为正, 实际为正; 假正类 (False Positive, FP): 预测为正, 实际为负; 真负类 (True Negative, TN): 预测为负, 实际为负; 假负类 (False Negative, FN): 预测为负, 实际为正。三者可以定义为:

$$P = \frac{TP}{TP + FP}. \quad (4.3)$$

$$R = \frac{TP}{TP + FN}. \quad (4.4)$$

$$F1 = \frac{2PR}{P + R}. \quad (4.5)$$

混淆矩阵 (Confusion Matrix, CM): $CM_{ij}(i, j = 1, 2, \dots, n_{samples})$ 可以定义为真实标签为类别 i 却被预测给类别 j 的数量。

ROC 曲线与 AUC: ROC 曲线 (receiver operating characteristic) 是接收器操作特性曲线, 描述的是真正类 (TP) 与假正类 (FP) 之间的相互关系。AUC 为 ROC 曲线下的面积, AUC 的值越接近于 1, 表示分类性能越好。

4.4.2 超参数选择

在改进的深度残差网络中, 各超参数的设置是决定模型最终分类性能的关键问题之一。本文针对第一个卷积层卷积核大小、加宽因子 k 、每组残差单元数 N 和 Dropout 大小等超参数开展一系列实验以确定其最佳取值。具体实验过程为针对每类超参数选取数个不同取值进行模型训练, 将训练得到的模型在测试集的 4 个测试尺度上分别测试 10 次, 取每个测试尺度的最好结果进行分类性能的比较分析。

(1) 第一个卷积层卷积核大小与网络性能的关系

为了探索第一个卷积层卷积核大小与网络性能的关系, 分别设计了 3 种卷积核大小: 3×3 、 6×6 和 7×7 。 3×3 的灵感来源于 VGG 中采用的小型卷积核的思路, 6×6 是 Dieleman 模型中的第一个卷积层的卷积核, 7×7 是原 ResNet 中的策略, 其他参数保持不变。实验结果如表 4.4 所示。

表 4.4 第一层卷积核尺寸与网络性能的关系

Table 4.4 The relationship between networks performance and kernel size of Conv 1

Conv 1 卷积核	准确率 (%)
3×3	92.1181
6×6	95.2083
7×7	93.7153

从表中可以看出，当第一层卷积核选择 6×6 时，模型的准确率最高。所以，第一层卷积核的尺寸定为 6×6 。

(2) 加宽因子 k 和每组残差单元数 N 与网络性能的关系

为了探索加宽因子 k 和每组残差单元数 N 与网络性能的关系，设计了 4 组实验： $k = 1, N = 1$ 、 $k = 1, N = 2$ 、 $k = 2, N = 1$ 和 $k = 2, N = 2$ ，其他参数控制不变，实验结果如表 4.5 所示。

表 4.5 加宽因子 k 和每组残差单元数 N 与网络性能的关系Table 4.5 The relationship between networks performance and k and N

k, N	层数	准确率 (%)
$k = 1, N = 1$	14	93.4028
$k = 1, N = 2$	26	94.7569
$k = 2, N = 1$	14	93.0556
$k = 2, N = 2$	26	95.2083

从表中可以得到，当 $k = 2, N = 2$ 时，网络的分类准确率最高。

(3) Dropout 大小与网络性能的关系

为了探索 Dropout 大小与网络性能的关系，Dropout 被放置在每个残差单元中的 3×3 卷积之后，用来防止过拟合。其他参数保持不变，实验结果如表 4.6 所示。

从表中可以得到，当 Dropout 为 0.8 时，网络的分类准确率最高。

通过上述一系列实验，当选取第一层卷积核的尺寸为 6×6 ，加宽因子 k 为 2，每组残差单元数为 2 以及 Dropout 为 0.8 时得到的模型具有相对最优的分类性能。由于整个网络共计 26 层（包括 25 个卷积层和 1 个全连接层），将该模型称为 ResNet-26。ResNet-26 共计 2600 万参数，在测试尺度为 220、训练迭代次

表 4.6 Dropout 大小与网络性能的关系

Table 4.6 The relationship between networks performance and dropout

Dropout	准确率 (%)
0.5	92.8819
0.7	94.2917
0.8	95.2083

数为 42k 时取得最好结果。

(4) 星系图像类型与网络性能的关系

此外，为了验证星系图像类型对模型结果的影响，本文将星系 RGB 图像转化成灰度图像，然后使用 ResNet-26 对 RGB 图像和灰度图像进行比较实验，实验结果如表4.7所示。从表中可以看出，面向灰度图像的模型在测试尺度为 240 时取得最好结果，但最高准确率仅为 93.4722%，而面向 RGB 图像的模型最高准确率达到 95.2083%，在 F1、AUC 等其他评价指标上，RGB 图像的模型结果也均优于灰度图像的模型结果。上述结果标明星系 RGB 图像的 3 个颜色通道值对模型的性能提升是有利的，因此本文模型使用星系 RGB 图像作为输入。

表 4.7 星系 RGB 图像与灰度图像单测试尺度最高准确率，F1 值和 AUC 值

Table 4.7 Test accuracy, F1 and AUC of ResNet-26 based on RGB and gray images

	测试尺度 (Q)	准确率 (%)	F1 值	AUC 值
灰度图像	240	93.4722	0.9342	0.9786
RGB 图像	220	95.2083	0.9515	0.9823

(5) ResNet-26 的分类结果

图 4.6为训练集和测试集的损失函数变化曲线图。其中，橙色代表训练集，蓝色代表测试集。训练集上每迭代 50 次记录一次，测试集每迭代 200 次记录一次，每个批 (batch) 大小为 128。从图中可以看出训练集在 35k 时开始收敛，loss 值维持在 0.05 左右，测试集在 35k 时有上升趋势，表示开始出现过拟合现象。图 4.7为训练集和测试集准确率变化曲线图，测试集的准确率最终在 94%-96% 之间震荡。震荡现象是正常的，因为记录的值只是一个 batch 大小的准确率，而不是全部测试集的平均准确率，但其总体趋势是一致的。

表 4.8给出了测试集中每一类星系的精确率，召回率和 F1 值。圆形星系、中

间星系、雪茄状星系、侧向星系和旋涡星系这 5 类星系分类的平均精确率，召回率和 F1 值分别为 0.9512, 0.9521 和 0.9515。其中，圆形星系取得最高的精确率值 0.9611，旋涡星系取得最好的召回率 0.9782 和 F1 值 0.9515。

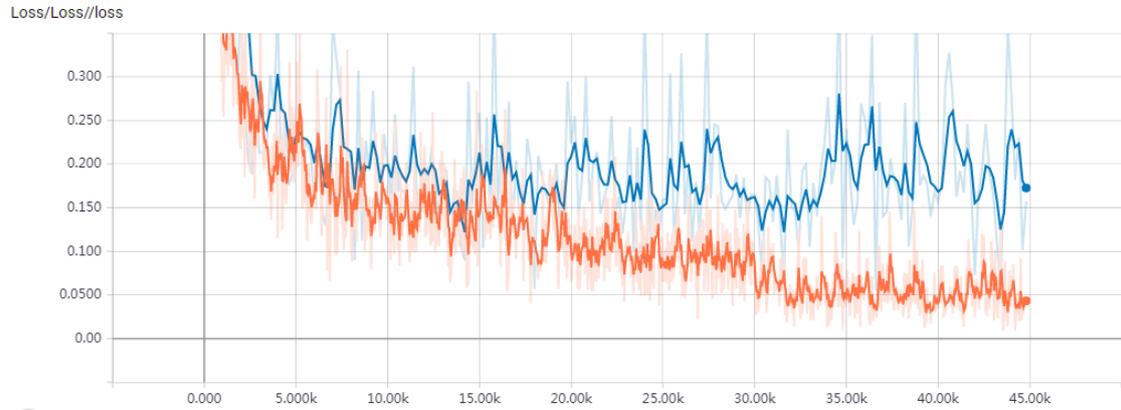


图 4.6 训练集和测试集损失函数曲线图

Figure 4.6 Training loss curve and testing loss curve for our network

注：橙色代表训练集，蓝色代表测试集。训练集上每迭代 50 次记录一次，测试集每迭代 200 次记录一次，迭代 45k 时停止训练。

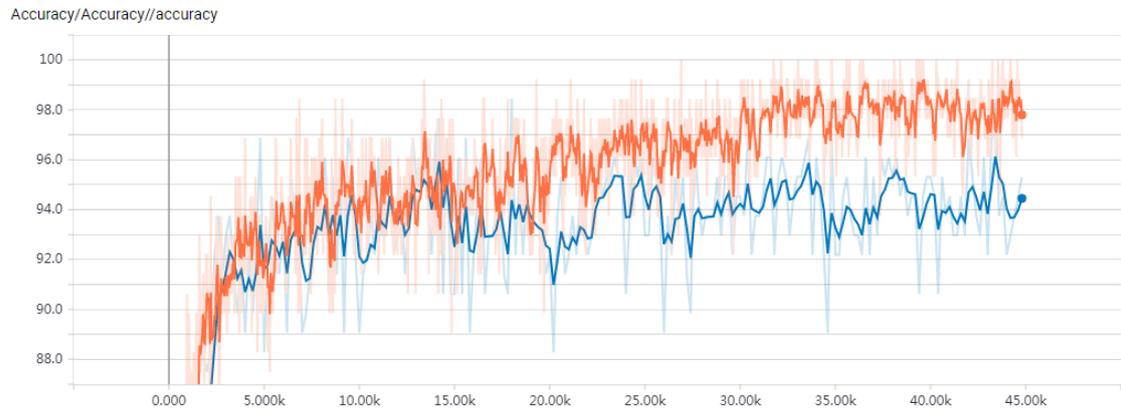


图 4.7 训练集和测试集准确率曲线图

Figure 4.7 Training accuracy curve and testing accuracy curve for our network

注：橙色代表训练集，蓝色代表测试集。训练集上每迭代 50 次记录一次，测试集每迭代 200 次记录一次，迭代 45k 时停止训练。

表 4.9 为测试集的混淆矩阵。表中，列为真实标签，行为预测标签。从混淆矩阵中可以看到，有 815 张圆形星系图片，762 张中间星系图片，34 张雪茄状星系图片，368 张侧向星系图片，763 张旋涡星系图片被正确分类，由此可以计算出 5 个类别星系的分类准确率分别为：圆形星系，96.6785%；中间星系，94.4238%；

表 4.8 测试集的精确率，召回率，F1 值

Table 4.8 Precision, Recall and F1 of our model for each class on testing set

类别	精确率	召回率	F1 值
0	0.9611	0.9634	0.9622
1	0.9561	0.9431	0.9495
2	0.7234	0.5862	0.6476
3	0.9412	0.9485	0.9448
4	0.9573	0.9782	0.9677
平均	0.9512	0.9521	0.9515

注：0、1、2、3、4 分别表示圆形星系、中间星系、雪茄状星系、侧向星系和旋涡星系。高亮标识最好结果。

雪茄状星系，58.6207%；侧向星系，94.3590% 和旋涡星系，97.6953%。；有 29 张圆形星系图片被错分给中间星系，12 张雪茄状星系被错分给侧向星系，而 18 张侧向星系被错分给雪茄状星系。分析认为，圆形星系图片被错分给中间星系属于正常现象，因为这两个星系都属于平滑星系，且二者没有严格的区分界限；雪茄状星系和侧向星系错分则较为意外，经多次实验对比初步判定为两个星系在形状上确实有些相似所致。

表 4.9 测试集的混淆矩阵

Table 4.9 Confusion matrix of our model for each class on testing set

	0	1	2	3	4
0	815	21	0	0	10
1	29	762	0	0	17
2	0	4	34	18	2
3	0	3	12	368	5
4	4	7	1	5	763

注：列为真实标签，行为预测标签。0、1、2、3、4 分别表示圆形星系、中间星系、雪茄状星系、侧向星系和旋涡星系。

图 4.8为测试集的 ROC 曲线。真正率越接近于 1，假正率越接近于 0，表示曲线预测得越好，也就是曲线越接近左上角代表结果越好。从图中看出，每一类都表现得很优秀，其中侧向星系预测得最好，其 AUC 值为 0.9939，雪茄状星系预测得相对差些，其 AUC 值为 0.9481。模型的平均 AUC 为 0.9813，说明 ResNet-26 模型的整体预测性能很好。

从上面 6 个指标可以看到，模型在圆形星系、中间星系、侧向星系和旋涡星

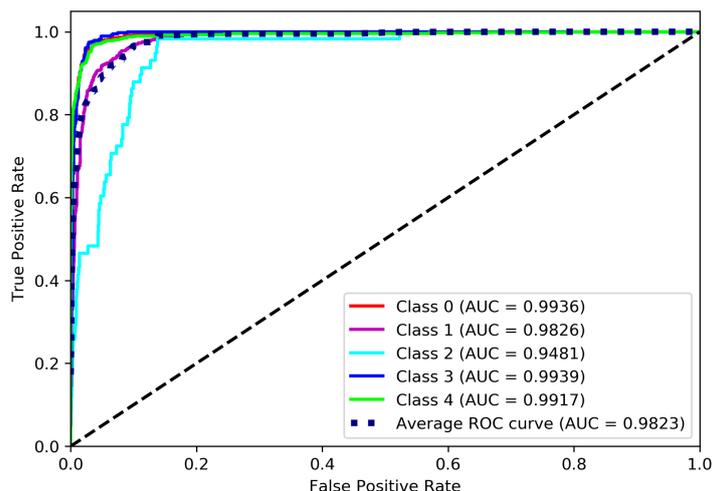


图 4.8 测试集上 5 个星系类的 ROC 曲线

Figure 4.8 ROC curve of our model for 5 classes galaxies on testing set

注：每一种颜色代表一个星系类别。

系的分类识别过程中均表现优秀，F1 值都在 0.94 以上；其中圆形星系、侧向星系和旋涡星系的 AUC 值都超过 0.99。模型针对旋涡星系的分类性能最好，其分类准确率达到了 97.6953%；针对雪茄状星系的分类性能表现一般，训练集中的雪茄状星系图片数量相对较少是造成分类性能一般的一个原因。

4.4.3 不同网络模型分类性能对比

本小节进行 Dieleman 模型、AlexNet、VGG、Inception、ResNet-50 和 ResNet-26 的分类性能对比实验。Dieleman 模型权重初始化和训练算法全部来自 [Dieleman et al. \(2015\)](#)；AlexNet 权重初始化和训练算法全部来自 [Krizhevsky et al. \(2012\)](#)；VGG 采用的是 VGG-16，权重初始化和训练算法全部来自 [Simonyan et al. \(2014\)](#)；Inception 使用的是 Inception V3 版本（42 层），具体权重参数和训练算法见 [Szegedy et al. \(2016\)](#)；ResNet-50 的权重初始化和训练算法来自 [He et al. \(2016a,b\)](#)。在相同的 GPU 服务器上，使用相同的训练集和测试集，测试不同模型的准确率、精确率、召回率、F1 值和 AUC 值。6 个模型的对比实验结果如表 4.10 和表 4.11 所示。其中，表 4.10 为不同模型在多测试尺度上的平均准确率，即在每一个测试尺度上测试 10 次，记录每个测试尺度上取得的最好结果，然后取其平均值，这样可以使模型的测试结果更加稳定。表 4.11 为不同模型在单测试尺度上的最高准确率，精确率，召回率，F1 值和 AUC 值，即在每一个测试尺度上测试 10 次，取每个测试尺度上取得的最好结果，其结果比表 4.10 中的结果更高些。

由表 4.10 可以看到, ResNet-26 取得最高的准确率 94.6875%。Dieleman 模型是专门为星系图片设计的卷积神经网络结构, 获得 93.8800% 的准确率。尽管 AlexNet、VGG、Inception 和 ResNet 是专为 ImageNet 数据集而设计的卷积神经网络, 但在星系数据集上仍然表现出很不错的泛化能力, Inception 和 ResNet 的准确率都超过了 94%。

表 4.10 不同模型的测试平均准确率

Table 4.10 Test accuracy of different methods at multiple testing scales

模型	图片尺寸		准确率 (%)
	训练 (S)	测试 (Q)	
Dieleman(Dieleman et al., 2015)	[170,240]	180,200,220,240	93.8800
AlexNet(Krizhevsky et al., 2012)	[170,240]	180,200,220,240	91.8230
VGG(Simonyan et al., 2014)	[170,240]	180,200,220,240	93.1336
Inception(Szegedy et al., 2016)	[170,240]	180,200,220,240	94.2014
ResNet-50(He et al., 2016b)	[170,240]	180,200,220,240	94.0972
ResNet-26	[170,240]	180,200,220,240	94.6875

注: 结果取自每一个测试尺度运行 10 次取最大值, 然后再求平均值。高亮标识最好结果。

表 4.11 不同模型的测试准确率, 精确率, 召回率, F1 值和 AUC 值

Table 4.11 Test accuracy, precision, Recall, F1 and AUC of different methods

模型	测试 (Q)	准确率 (%)	精确率	召回率	F1 值	AUC 值
Dieleman(Dieleman et al., 2015)	180	94.6528	0.9455	0.9465	0.9456	0.9793
AlexNet(Krizhevsky et al., 2012)	200	92.2569	0.9207	0.9226	0.9215	0.9809
VGG(Simonyan et al., 2014)	220	93.6458	0.9348	0.9365	0.9353	0.9846
Inception(Szegedy et al., 2016)	220	94.5139	0.9447	0.9451	0.9448	0.9852
ResNet-50(He et al., 2016b)	220	94.6875	0.9458	0.9469	0.9461	0.9823
ResNet-26	220	95.2083	0.9512	0.9521	0.9515	0.9823

注: 结果基于每一个测试尺度运行 10 次取最大值, 高亮标识最好结果。

从表 4.11 可以看出, ResNet-26 在测试尺度为 220 时取得了最高的准确率 95.2083%, 最大的精确率 0.9512, 最大召回率 0.9521, 最大 F1 值 0.9515 和最大 AUC 值 0.9823。Dieleman 模型在测试尺度 180 时取得其最好结果, 最高准确率为 94.6528%。Inception V3 在测试尺度 220 时取得其最好结果, 同时获得了最高的 AUC 值 0.9852。

总之, 通过与 Dieleman 模型、AlexNet、VGG、Inception 和 ResNet-50 的

对比实验，ResNet-26 取得了较优的分类结果。

4.4.4 实验环境

为实施 ResNet-26 模型及同类型模型的训练和测试，配置一台 GPU 服务器，服务器配置为 2 块 NVIDIA Tesla K80 GPU，CentOS Linux 7 操作系统，CUDA 为 8.044 版本，CUDNN 为 5.1 版本，采用 Python, Pandas, scikit-learn(Pedregosa et al., 2012), scikit-image(Van et al., 2014), TensorFlow(Abadi et al., 2016) 等语言和 API。本文所有代码可以在<https://github.com/Adaydl/>上下载。

4.5 本章小结

在深度残差网络的基础上，重点对残差单元进行改进，加入了 Dropout，减少了网络的深度，加宽了网络的宽度，结合星系图像本身的特点，设计了一个深度残差网络，通过对学习率、第一层卷积核大小、加宽因子、每组残差单元数目和 Dropout 大小等超参数的选择和大量实验，确定了最优的网络模型为 26 层，称之为 ResNet-26；在数据选取和预处理方面，采用 GZ2 决策树的标准，从 GZ2 里面选取干净样本数据，划分为训练集和测试集，进行了裁剪、下采样、旋转、水平翻转、光学畸变和白化等预处理过程，成为最终模型训练和测试的输入数据。为了验证 ResNet-26 的分类性能，选择 Dieleman、AlexNet、VGG-16、Inception V3 和 ResNet-50 模型进行比较实验，实验过程采用的训练集和测试集一样。结果表明，ResNet-26 取得了较优的分类性能，如取得了最高的分类准确率、最高的 F1 值等。

第 5 章 星系图像表征与结果分析

基于深度学习的星系形态分类已经取得了一定进展，但是这些进展更多关注的是星系形态分类本身这个任务。如何进一步挖掘神经网络所学习到的表征是什么以及这些表征如何促进于更好地理解星系数据本身，也同样是一个非常重要的研究课题。本章利用 t-SNE 降维技术，可视化从神经网络学习到的高维星系表征，从而得到一些有价值的反馈。

5.1 T 分布随机近邻嵌入 (t-SNE)

高维数据可视化是数据分析的方法之一，有助于发现很多高价值的反馈。由于真实世界数据绝大多数都是高维数据，所以为了可视化高维数据，众多的降维技术被提出 (Van der Maaten, 2007; Van Der Maaten et al., 2009)。降维技术包括线性降维技术，如主成分分析 (PCA) 和经典多维尺度变换 (Classical Scaling) 等，和非线性降维技术，如等距特征映射 (Isomap)、局部线性嵌入 (LLE)、拉普拉斯特征映射 (Laplacian Eigenmaps)、自编码器 (Auto-Encoder) 和 T 分布随机近邻嵌入 (t-SNE) 等。

T 分布随机近邻嵌入 (t-Distributed Stochastic Neighbor Embedding, t-SNE) 是一种非线性降维方法 (Maaten et al., 2008; Van Der Maaten, 2014)，以其具有保留数据局部特征以及揭示全局特征的优点，在机器学习领域被广泛使用。t-SNE 通常应用于高维数据的可视化，即将高维数据降维至二维或三维空间，以散点图的方式呈现。

假设数据集 X 包含 N 个样本 $X = \{x_1, x_2, \dots, x_N\}$ ，每个样本为 D 维向量。t-SNE 的目标是计算投影 $Y = \{y_1, y_2, \dots, y_N\}$ 。通常 $y_i \in \mathbb{R}^d$ 对应于 $x_i \in \mathbb{R}^D$ 。典型地， $d = 2$ 且 $D \gg d$ 。

首先，计算条件概率 $p_{j|i}$ 。 $p_{j|i}$ 表示数据点 x_i 和 x_j 的相似性，当 x_i 和 x_j 邻近时， $p_{j|i}$ 值越大。其定义为

$$p_{j|i} = \frac{\exp(-\|x_i - x_j\|^2 / 2\sigma_i^2)}{\sum_{k \neq i} \exp(-\|x_i - x_k\|^2 / 2\sigma_i^2)}. \quad (5.1)$$

其中 σ_i 是是以数据点 x_i 为均值的高斯分布标准差。

然后，定义高维空间中的联合分布 p_{ij} ，使其为对称的条件概率：

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2N}. \quad (5.2)$$

下一步，在低维空间中，使用更重长尾分布的 t 分布（自由度为 1）定义联合分布 q_{ij} ：

$$q_{ij} = \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k \neq l} (1 + \|y_k - y_l\|^2)^{-1}}. \quad (5.3)$$

低维空间中数据点 y_i 由最小化分布 Q 和 P 的 Kullback-Leibler 散度（KL 散度）求得，即 t-SNE 的目标函数 C 为

$$C = KL(P|Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}}. \quad (5.4)$$

P 和 Q 的 KL 散度使用如下梯度公式求解：

$$\frac{\delta C}{\delta y_i} = 4 \sum_j (p_{ij} - q_{ij})(y_i - y_j)(1 + \|y_i - y_j\|^2)^{-1}. \quad (5.5)$$

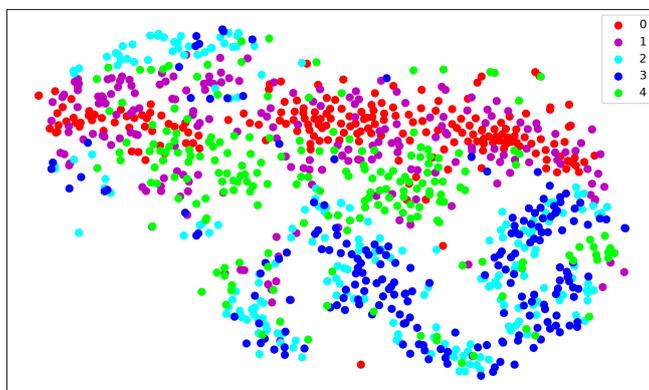
Van Der Maaten (2014) 基于 MNIST(LeCun, 1998) 和 SVHN(Netzer et al., 2011) 数据集，训练 CNN 模型，然后从最后一层全连接层（除了输出层）提取图片高维表征，用 t-SNE 降维，可视化高维数据，从散点图中获得了一些有价值的反馈和定性分析。Rauber et al. (2017) 基于 3 个传统的图片分类标准数据集（MNIST, SVHN 和 CIFAR-10）系统地训练了多层感知器（Multilayer Perceptrons, MLPs）和 CNN 模型，从最后一层全连接层抽取激活值（Activation），然后使用 t-SNE 降维至二维空间，以散点图的方式呈现学到的高维表征，得到了一些高价值的发现。这些工作启发了星系形态高维表征可视化的研究，将从神经网络中学习到的高维星系表征应用于星系形态分类的后续分析，探索星系高维抽象表征的物理意义，更好地理解星系图片数据本身。

5.2 星系图像数据集的 t-SNE 可视化及分析

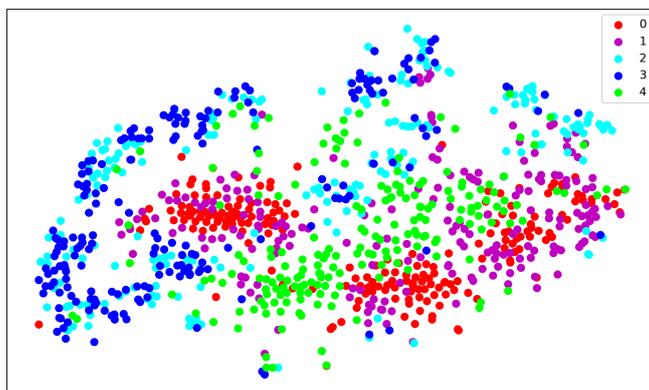
本文从训练集中随机选择 1000 张星系图片，每个类别 200 张，组成训练子集，称为 training-1000。使用同样的方法从测试集中随机选择 1000 张图片，组成测试子集，称为 testing-1000。

首先直接对原始星系图片进行 t-SNE 降维，即直接将 $424 \times 424 \times 3$ (539328) 维像素降至 2 维，结果如图 5.1 所示。图中数据点的颜色分别代表相应的类别，每个数据点代表一张星系图片。从图 5.1a 可以看出，training-1000 子集中星系类别

之间区分度很差，散乱无章，每个类别交织在一起，没有成簇现象。图 5.1b testing-1000 子集结果和图 5.1a 相似。之所以出现上述结果，是因为没有使用监督信息，即标签信息。同时，这两个图也说明此星系数据集是非常复杂和充满挑战的。



(a) training-1000 原始样本映射



(b) testing-1000 原始样本映射

图 5.1 原始样本可视化

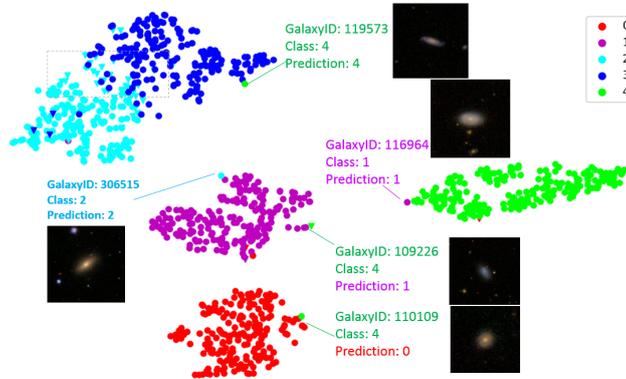
Figure 5.1 Visualizations of raw samples

注：数据点的颜色分别代表相应的类别，红色、品红、蓝绿色、蓝色和绿黄色分别代表圆形星系（0）、中间星系（1）、雪茄状星系（2）、侧向星系（3）和漩涡星系（4），每个数据点代表一张星系图片。

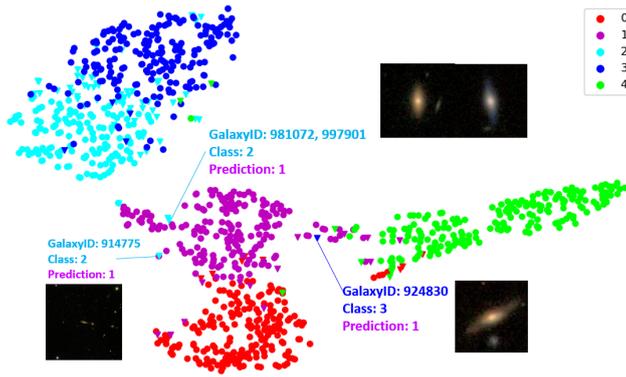
5.3 星系形态分类模型的 t-SNE 可视化及分析

对于 Dieleman 模型，从最后一个全连接层（除了输出层）抽取激活值形成 2048 维高维表征；对于 AlexNet，从最后一个全连接层（除了输出层）抽取激活值形成 4096 维高维表征；对于 VGG，从最后一个全连接层（除了输出层）抽取激活值形成 4096 维高维表征；对于 Inception，分别从最后一个卷积层和最后平均池化层抽取激活值形成 18432、2048 维高维表征；对于 ResNet-50，分别从

最后一个卷积层和最后平均池化层抽取激活值形成 8192、2048 维高维表征；对于 ResNet-26，从最后平均池化层抽取激活值形成 4096 维高维表征。然后分别使用 t-SNE 降维，可视化，6 个模型结果分别如图 5.2、图 5.3、图 5.4、图 5.5、图 5.6和图 5.7所示。



(a) 基于 training-1000 的 Dieleman 模型最后一层全连接层映射



(b) 基于 testing-1000 的 Dieleman 模型最后一层全连接层映射

图 5.2 Dieleman 模型最后一层全连接层激活值可视化

Figure 5.2 Visualizations of activations of the last fully connected layer of Dieleman

注：数据点的颜色分别代表相应的类别，红色、品红、蓝绿色、蓝色和绿黄色分别代表圆形星系（0）、中间星系（1）、雪茄状星系（2）、侧向星系（3）和旋涡星系（4），每个数据点代表一张星系图片，三角形表示错分类。

图 5.2为 Dieleman 模型最后一层全连接层激活值可视化。数据点的颜色分别代表相应的类别，每个数据点代表一张星系图片，三角形表示错分类。从图中可以看出，training-1000 和 testing-1000 经过神经网络训练之后，每一个星系类别的区分度得到了显著改善。正如LeCun et al. (2015) 所言，更深层的表征放大了对分类和识别十分关键的输入特征。图 5.2a和图 5.2b显示，Dieleman 模型最

后一层全连接层学到了原始数据的一种更好的表征，可以使类与类之间更加分离，类内的数据点更加紧密。也就是每一个星系类成簇分布，相同类别的星系距离更近，不同类别的星系尽量分散开来。这是由于相同类别的星系图片具有相似的内在结构，这使他们在散点图中趋于聚集，成簇分布，不同类别的星系内在结构不同，会分散开来。从图 5.2b 还可以发现，圆形星系和中间星系，和其他星系相比，距离更近。这是由于圆形星系和中间星系同属于一个更大的类：平滑星系。

然后，进行离群点检测。图 5.2a 中，观察蓝色簇（侧向星系）下面的绿黄色的数据点，它属于旋涡星系，也被分给了旋涡星系，但是却出现在蓝色簇中，这说明它们的结构相似。经检查，这个数据点在数据集中代号（GalaxyID）为 119573，看起来和侧向星系很相似。图中别的离群点，如代号 306515 的雪茄状星系和代号为 116964 的中间星系，也有着和上例相似的现象。另一个在红色簇（圆形星系）附近的绿黄色数据点，它属于旋涡星系，却被分给了圆形星系。经检查，这个数据点为代号 110109 的旋涡星系，其实是圆形星系，也就是说，它的标签是错误的。图 5.2b 中，一些离群点，如代号为 981072 和 997901 的雪茄状星系被预测给了中间星系，代号为 924830 的侧向星系被预测给了中间星系。检查之后，发现这些离群点的图片都和中间星系很像，即它们不是典型的原属标签星系。

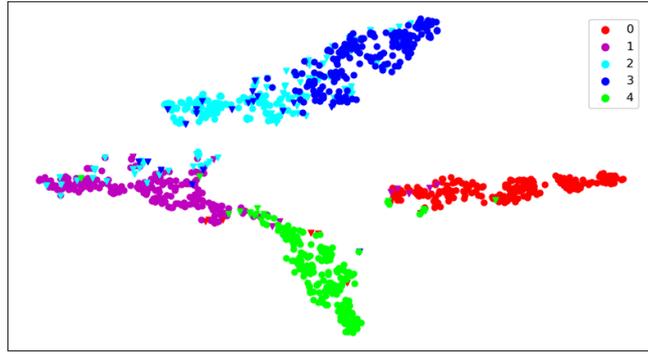
图 5.3 为 AlexNet 模型最后一层全连接层激活值可视化。每一个类都成簇分布，不同的类尽量分隔开来。而且每一个类别有着相对固定的坐标，即结构，旋涡星系坐标总是在图的下方，圆形星系在图的右方。

图 5.4 为 VGG 模型最后一层全连接层激活值可视化。从图 5.4a 和图 5.4b 中可以看出，星系类之间的分离几乎是完美的，每一个类别呈条带状分布。图 5.4a 中的 2 个用虚线框出得椭圆中分别有 9 张中间星系被错分给圆形星系，16 张雪茄状星系被错分给中间星系。这很容易理解，因为圆形星系、中间星系和雪茄状星系都属于平滑星系，它们之间并没有严格的界限去区分它们。可以解释它们是被错分的，也可以解释它们的标签是错误的，但却被模型分类正确。

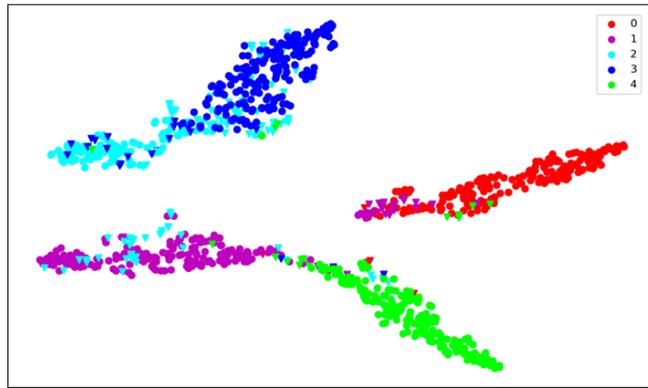
图 5.5 为 Inception V3 最后一个卷积层和最后一层平均池化层激活值可视化。可以看出，从最后一个卷积层输出的激活值得到的映射和继续平均池化（ 3×3 ）后的映射效果差不多。

图 5.6 为 ResNet-50 最后一个卷积层和最后一层平均池化层激活值可视化。testing-1000 中最后一层平均池化层的映射（右下）中圆形星系、中间星系和旋涡星系之间的分散情况比右上图要好些。

图 5.7 为 ResNet-26 最后一层平均池化层激活值可视化。图 5.7a 左上角 4 个



(a) 基于 training-1000 的 AlexNet 最后一层全连接层映射



(b) 基于 testing-1000 的 AlexNet 最后一层全连接层映射

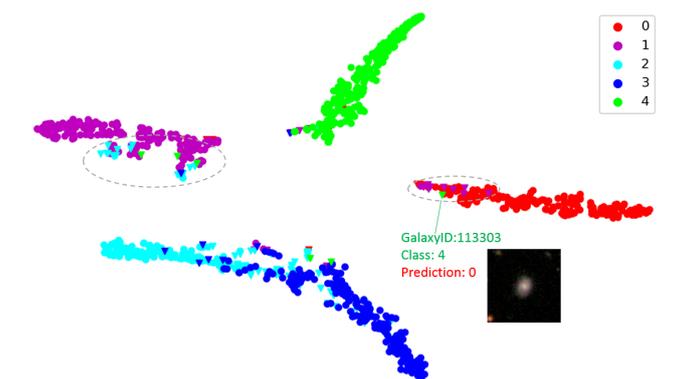
图 5.3 AlexNet 最后一层全连接层激活值可视化

Figure 5.3 Visualizations of activations of the last fully connected layer of AlexNet

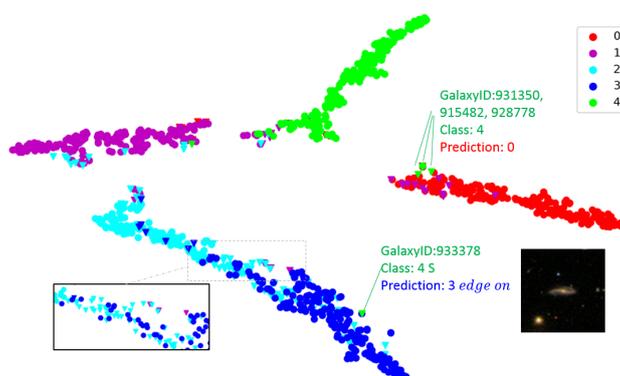
注：数据点的颜色分别代表相应的类别，红色、品红、蓝绿色、蓝色和绿黄色分别代表圆形星系（0）、中间星系（1）、雪茄状星系（2）、侧向星系（3）和旋涡星系（4），每个数据点代表一张星系图片，三角形表示错分类。

红色的离群点（GalaxyID: 102977, 107539, 122674, 106544），它们属于圆形星系，也被预测给了圆形星系，但是它们却出现在绿黄色簇所代表的旋涡星系中，检查之后，发现其中有三张真的圆形星系（GalaxyID: 102977, 107539, 106544）却跟旋涡星系有着相似的结构，另一张代号为 122674 的图片其实是旋涡星系，却被标记成圆形星系。图中还有另外 4 张旋涡星系（GalaxyID: 114957, 119150, 119573, 109366）出现在侧向星系簇和雪茄状星系簇附近，检查之后发现，它们都是瘦长的旋涡星系，看起来非常像侧向星系和雪茄状星系。

上述图中可以发现 2 张特别的星系图片，代号分别为 109266 和 914775。



(a) 基于 training-1000 的 VGG 最后一层全连接层映射



(b) 基于 testing-1000 的 VGG 最后一层全连接层映射

图 5.4 VGG 最后一层全连接层激活值可视化

Figure 5.4 Visualizations of activations of the last fully connected layer of VGG

注：数据点的颜色分别代表相应的类别，红色、品红、蓝绿色、蓝色和绿黄色分别代表圆形星系（0）、中间星系（1）、雪茄状星系（2）、侧向星系（3）和旋涡星系（4），每个数据点代表一张星系图片，三角形表示错分类。

109266 在 training-1000 子集中被标记为旋涡星系，在图 5.2a和图 5.7a中全都被错分给中间星系。检查之后发现，它是一张非常模糊的旋涡星系以至于很难识别它。914775 在 testing-1000 子集中被标记为雪茄状星系，在图 5.2b中被 Dieleman 模型错分给了中间星系，在图 5.7b中被改进的残差网错分给了圆形星系。检查之后发现，它是一张非常小的图片，星系处于图片中央非常小的位置，很难识别它属于圆形星系、中间星系还是雪茄状星系。

从以上 6 个网络模型的图中可以看出，训练子集上星系类别的分离程度比测试子集效果好。这是正常的，因为训练子集有着更高的准确率，学到了更好的表征。

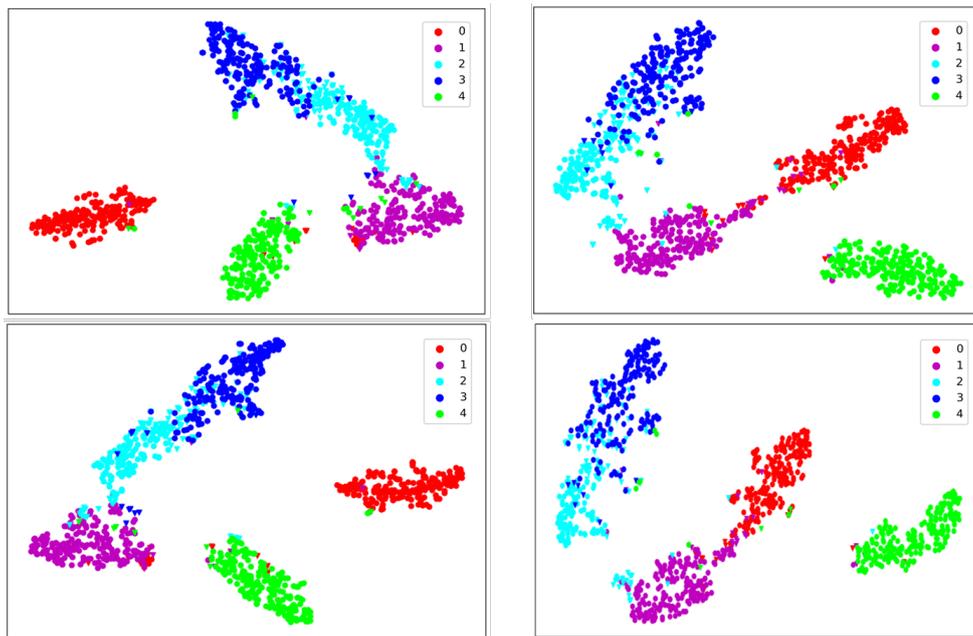


图 5.5 Inception V3 最后一个卷积层和最后一层平均池化层激活值可视化

Figure 5.5 Visualizations of activations of the last convolutional layer and last average-pooling layer based on Inception V3

注：数据点的颜色分别代表相应的类别，红色、品红、蓝绿色、蓝色和绿黄色分别代表圆形星系（0）、中间星系（1）、雪茄状星系（2）、侧向星系（3）和旋涡星系（4），每个数据点代表一张星系图片，三角形表示错分类。左上：training-1000 最后一个卷积层映射，左下：training-1000 最后一层平均池化层映射，右上：testing-1000 最后一个卷积层映射，右下：testing-1000 最后一层平均池化层映射

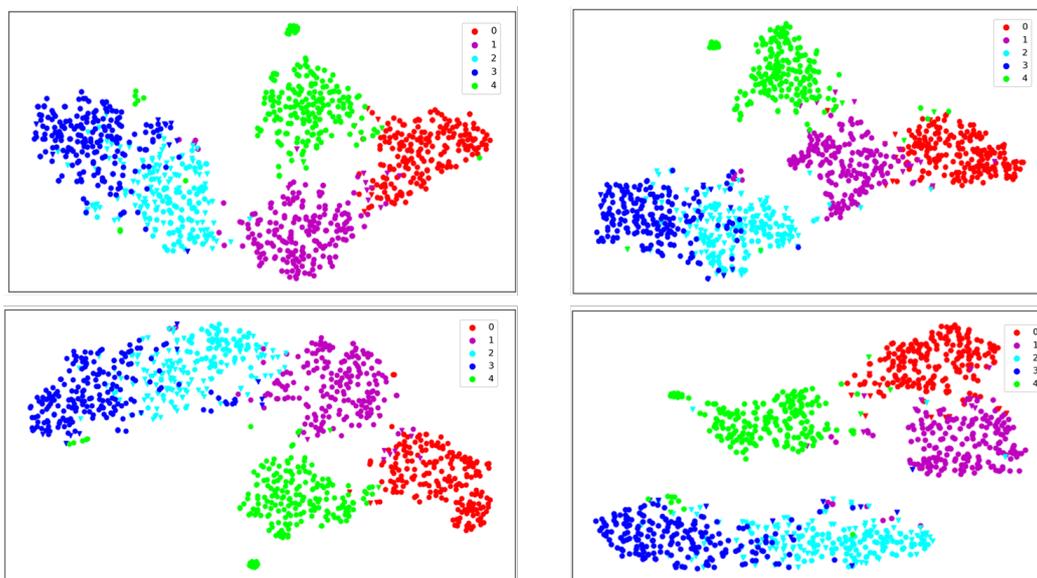
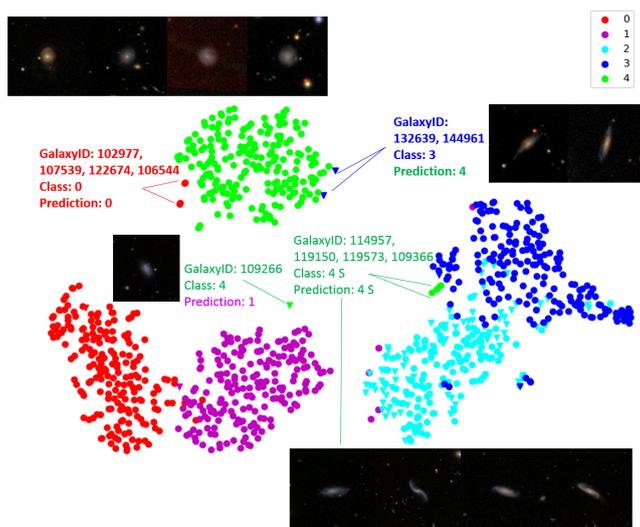


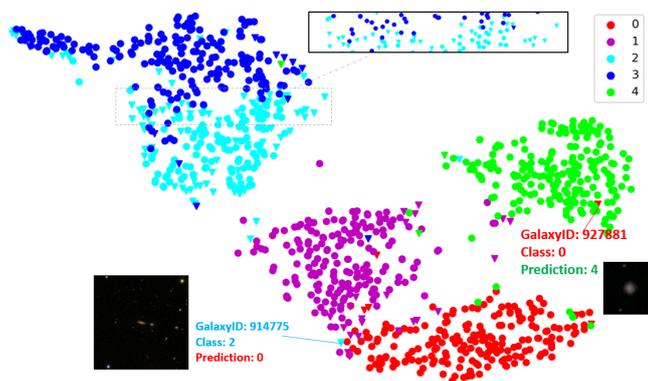
图 5.6 ResNet-50 最后一个卷积层和最后一层平均池化层激活值可视化

Figure 5.6 Visualizations of activations of the last convolutional layer and last average-pooling layer based on ResNet-50

注：数据点的颜色分别代表相应的类别，红色、品红、蓝绿色、蓝色和绿黄色分别代表圆形星系（0）、中间星系（1）、雪茄状星系（2）、侧向星系（3）和旋涡星系（4），每个数据点代表一张星系图片，三角形表示错分类。左上：training-1000 最后一个卷积层映射，左下：training-1000 最后一层平均池化层映射，右上：testing-1000 最后一个卷积层映射，右下：testing-1000 最后一层平均池化层映射



(a) 基于 training-1000 的 ResNet-26 最后一层平均池化层映射



(b) 基于 testing-1000 的 ResNet-26 最后一层平均池化层映射

图 5.7 ResNet-26 最后一层平均池化层激活值可视化

Figure 5.7 Visualizations of activations of the last average pooling layer of ResNet-26

注：数据点的颜色分别代表相应的类别，红色、品红、蓝绿色、蓝色和绿黄色分别代表圆形星系（0）、中间星系（1）、雪茄状星系（2）、侧向星系（3）和旋涡星系（4），每个数据点代表一张星系图片，三角形表示错分类。

每一个小的簇，如圆形星系簇、旋涡星系簇，通常都是成簇分布并且聚集得很紧密，几乎完美。大的星系类，如平滑星系包括圆形星系、中间星系和雪茄状星系，也趋于聚集在一起，距离更近些，这些可以从图 5.4b、图 5.7a和图 5.7b看出。

从图 5.2、图 5.3、图 5.4、图 5.5、图 5.6和图 5.7都可以发现，雪茄状星系和侧向星系交织缠绕在一起，这暗示着它们具有相似的内在结构。但是雪茄状星

系属于平滑星系的一种，它本应该和圆形星系、中间星系更相似，结果却和侧向星系相似，这一点是令人意外的。这个意外的发现或许可以有助于 GZ2 决策树的优化。

5.4 本章小结

从 Dieleman、AlexNet、VGG-16、Inception V3、ResNet-50 和 ResNet-26 等六个卷积神经网络的最后一层全连接层（除了输出层）、最后一个卷积层或最后一层平均池化层抽取高维抽象的数据表征，运用 t-SNE 降维，利用 1000 张训练子集和 1000 张测试子集进行 t-SNE 可视化，从二维散点图中发现了一些特殊的现象，如相同类别的星系成簇分布，距离更紧密；不同类别的星系则分散开来；发现原始数据集本身标记错误的图片。此外，雪茄状星系和侧向星系意外地交织缠绕在一起，是否暗示着它们的内在结构具有一定的相似性？

第 6 章 总结与展望

将深度卷积神经网络应用于星系形态分类是解决天文大数据时代下星系图像需要快速、高效的自动分类的一个有效途径。本文充分比较了 5 种经典的卷积神经网络模型：LeNet5、AlexNet、VGG、Inception 和 ResNet，在深度残差网络的基础上，改进残差单元，尝试减少网络深度，加宽网络宽度，并结合星系图片本身的特点，设计了基于改进的深度残差网络的星系形态分类框架 ResNet-26。并搭建基于 TensorFlow 的深度学习框架，实现 ResNet-26 与其他 5 种卷积神经网络，验证了 ResNet-26 模型能有效提高星系分类的准确率。在此基础上，利用高维数据可视化技术（t-SNE），从训练后的模型最后一层全连接层（除了输出层）、最后一个卷积层和最后一层平均池化层抽取高维抽象的星系表征进行降维可视化分析，发现了相同类别的星系成簇分布，不同类别的星系彼此分散、标签错误的星系图片离群点等有价值的反馈。

总结而言，本文研究中的主要创新点为：

1. 将深度学习领域最新的研究成果引入星系形态分类领域，提出并设计了基于改进的深度残差网络的星系形态分类框架 ResNet-26，并通过实验验证了模型优良的分类性能；同时实现其他 5 种流行的卷积神经网络模型，分别包括 Dielman、AlexNet、VGG-16、Inception V3 和 ResNet-50，进行对比实验，证明了 ResNet-26 网络结构较优的分类性能和泛化能力；
2. 将高维数据可视化技术（t-SNE）引入星系形态分类的后续研究中，可视化从神经网络中所学习到的高维星系表征，得到了一些有价值的发现。

本文虽然在一定程度上推动了星系形态分类问题的解决，取得了相应的进展，然而仍然存在一些不足之处有待进一步探索：

1. 深度学习模型需使用大规模数据集进行训练才能达到好的分类性能。本文选取的数据集源自星系动物园，数据集中各类星系样本数量分布不均衡，如雪茄状星系图片数量仅占数据集总图片数量的 2%。在数据集有限的情况下，ResNet-26 模型对圆形星系、中间星系、雪茄状星系、侧向星系和旋涡星系等 5 类星系进行了训练。经测试，模型针对旋涡星系的测试准确率最高，为 97.6953%，针对数据样本较少的雪茄状星系测试准确率不高，仅为

58.6207%。后续可考虑从大规模数据集，如第 4 代星系动物园项目，中获取星系类别更全面、样本数量更充分、数据质量更优的星系图片作为训练样本，进一步训练模型，测试模型针对更多星系类别的分类性能。

2. ResNet-26 模型目前采用深度残差网络设计，可自动提取星系形态特征，进行星系形态分类；未来除了可尝试运用 Inception Module 进一步改进网络结构外，还可考虑将非模型化分类系统中的结构参数，如聚集度指数、非对称指数、簇聚指数、基尼系数和矩指数等加入到模型中，构建专家系统与神经网络相结合的混合模型，即神经网络专家系统，以提升模型的分类性能。
3. 在星系图像表征可视化方面，因学科专业背景限制，目前仅利用 t-SNE 降维方法，进行了星系高维表征可视化并对可视化后的表象信息进行了初步分析，未来可与天文学家进行更深层次的探讨，在星系图片可视化信息挖掘方面作更多尝试，如发现异常星系等。

附录

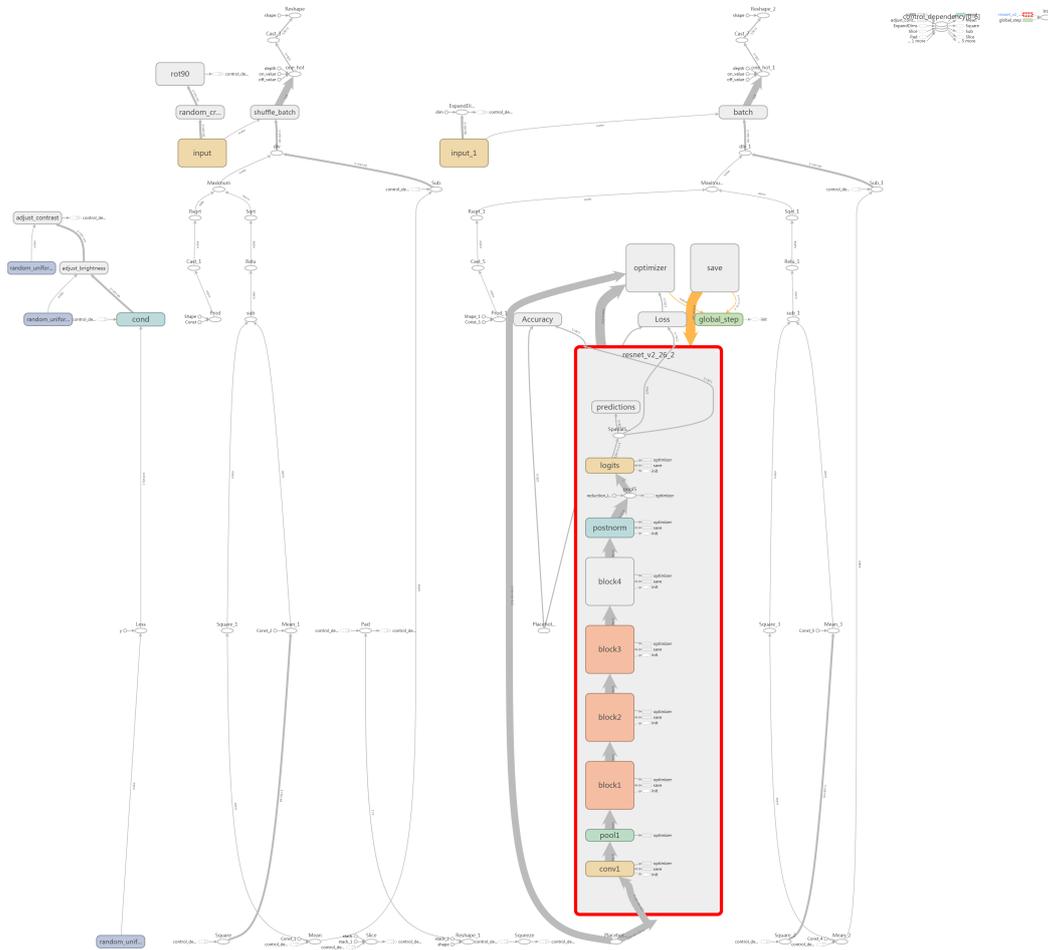
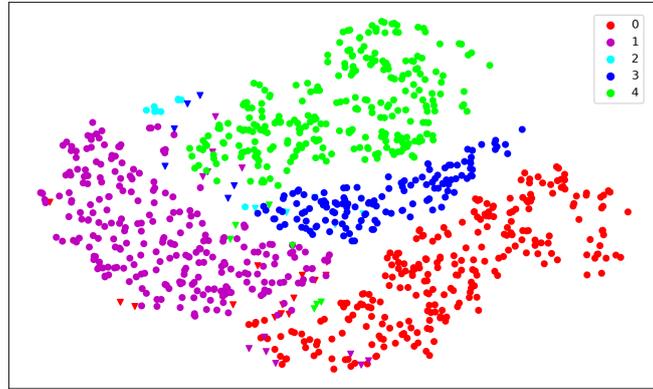
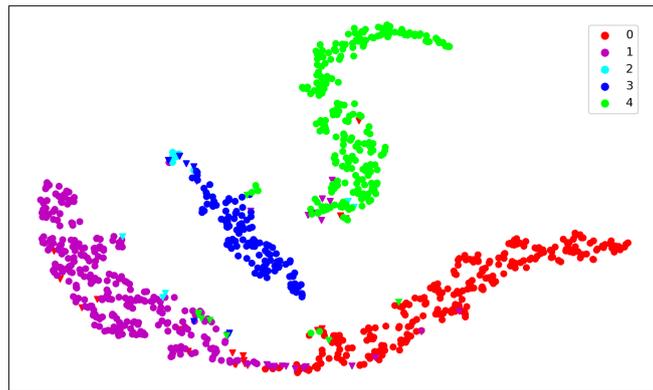


图 A.1 TensorBoard-整体网络结构

Figure A.1 TensorBoard-Our overall network architecture



(a) 全体训练集 (25911 images) VGG 最后一层全连接层映射



(b) 全体测试集 (2879 images) VGG 最后一层全连接层映射

图 A.2 VGG 最后一层全连接层激活值可视化

Figure A.2 Visualizations of activations of the last fully connected layer of VGG

注：数据点的颜色分别代表相应的类别，红色、品红、蓝绿色、蓝色和绿黄色分别代表圆形星系（0）、中间星系（1）、雪茄状星系（2）、侧向星系（3）和旋涡星系（4），每个数据点代表一张星系图片，三角形表示错分类。

参考文献

- 汪敏, 孔旭, 2007. 星系形态分类的研究进展[J]. 天文学进展, 25(3): 215–225.
- 周志华, 2016. 机器学习[M]. [出版地不详]: Qing hua da xue chu ban she.
- 崔辰州, 于策, 肖健, 等, 2015. 大数据时代的天文学研究[J]. 科学通报(5): 445–449.
- 方官文, 马仲阳, 孔旭, 2015. Cosmos 场中高红移大质量星系的形态分类研究[J]. 天文学报, 56(5): 425–435.
- ABADI M, AGARWAL A, BARHAM P, et al., 2016. Tensorflow: Large-scale machine learning on heterogeneous distributed systems[J]. arXiv preprint arXiv:1603.04467.
- ABELL P A, ALLISON J, ANDERSON S F, et al., 2009. Lsst science book, version 2.0 [J]. arXiv preprint arXiv:0912.0201.
- ABRAHAM R G, VAN DEN BERGH S, NAIR P, 2003. A new approach to galaxy morphology. i. analysis of the sloan digital sky survey early data release[J]. The Astrophysical Journal, 588(1): 218.
- ANIYAN A, THORAT K, 2017. Classifying radio galaxies with the convolutional neural network[J]. The Astrophysical Journal Supplement Series, 230(2): 20.
- BANERJI M, LAHAV O, LINTOTT C J, et al., 2010. Galaxy zoo: reproducing galaxy morphologies via machine learning[J]. Monthly Notices of the Royal Astronomical Society, 406(1): 342–353.
- BAZELL D, AHA D W, 2001. Ensembles of classifiers for morphological galaxy classification[J]. The Astrophysical Journal, 548(1): 219.
- BELKIN M, NIYOGI P, 2002. Laplacian eigenmaps and spectral techniques for embedding and clustering[C]//Advances in neural information processing systems. [S.l.: s.n.]: 585–591.
- BENGIO Y, COURVILLE A, VINCENT P, 2013. Representation learning: A review and new perspectives[J]. IEEE transactions on pattern analysis and machine intelligence, 35(8): 1798–1828.
- CERI S, BOZZON A, BRAMBILLA M, et al., 2013. An introduction to information retrieval[M/OL]. Berlin, Heidelberg: Springer Berlin Heidelberg: 3–11. https://doi.org/10.1007/978-3-642-39314-3_1.
- CLEVERT D A, UNTERTHINER T, HOCHREITER S, 2015. Fast and accurate deep network learning by exponential linear units (elus)[J]. arXiv preprint arXiv:1511.07289.

- CUN Y L, JACKEL L D, BOSER B, et al., 1989. Handwritten digit recognition: applications of neural network chips and automatic learning[J]. *IEEE Communications Magazine*, 27(11): 41–46.
- DE LA CALLEJA J, FUENTES O, 2004. Machine learning and image analysis for morphological galaxy classification[J]. *Monthly Notices of the Royal Astronomical Society*, 349(1): 87–93.
- DIELEMAN S, WILLETT K W, DAMBRE J, 2015. Rotation-invariant convolutional neural networks for galaxy morphology prediction[J]. *Monthly notices of the royal astronomical society*, 450(2): 1441–1459.
- FERRARI F, DE CARVALHO R R, TREVISAN M, 2015. Morfometryka—a new way of establishing morphological classification of galaxies[J]. *The Astrophysical Journal*, 814(1): 55.
- GAUCI A, ADAMI K Z, ABELA J, 2010. Machine learning for galaxy morphology classification[J]. arXiv preprint arXiv:1005.0390.
- GLOROT X, BORDES A, BENGIO Y, 2011. Deep sparse rectifier neural networks[C]// *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*. [S.l.: s.n.]: 315–323.
- GOODFELLOW I, BENGIO Y, COURVILLE A, 2016. *Deep learning*[M]. [S.l.]: MIT press.
- GRAVET R, CABRERA-VIVES G, PÉREZ-GONZÁLEZ P G, et al., 2015. A catalog of visual-like morphologies in the 5 candels fields using deep learning[J]. *The Astrophysical Journal Supplement Series*, 221(1): 8.
- HE K, ZHANG X, REN S, et al., 2015. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification[C]//*Proceedings of the IEEE international conference on computer vision*. [S.l.: s.n.]: 1026–1034.
- HE K, ZHANG X, REN S, et al., 2016a. Deep residual learning for image recognition [C]//*Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.]: 770–778.
- HE K, ZHANG X, REN S, et al., 2016b. Identity mappings in deep residual networks [C]//*European Conference on Computer Vision*. [S.l.]: Springer: 630–645.
- HINTON G E, ROWEIS S T, 2003. Stochastic neighbor embedding[C]//*Advances in neural information processing systems*. [S.l.: s.n.]: 857–864.
- HINTON G E, SALAKHUTDINOV R R, 2006a. Reducing the dimensionality of data with neural networks[J]. *science*, 313(5786): 504–507.

- HINTON G E, OSINDERO S, TEH Y W, 2006b. A fast learning algorithm for deep belief nets[J]. *Neural computation*, 18(7): 1527–1554.
- HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al., 2012. Improving neural networks by preventing co-adaptation of feature detectors[J]. *arXiv preprint arXiv:1207.0580*.
- HOTELLING H, 1933. Analysis of a complex of statistical variables into principal components.[J]. *Journal of educational psychology*, 24(6): 417.
- HUANG G, LIU Z, WEINBERGER K Q, et al., 2016. Densely connected convolutional networks[J]. *arXiv preprint arXiv:1608.06993*.
- HUBBLE E P, 1926. Extragalactic nebulae.[J]. *The Astrophysical Journal*, 64.
- HUBBLE E P, 1936. The realm of the nebulae: volume 25[M]. [S.l.]: Yale University Press.
- IOFFE S, SZEGEDY C, 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]//International Conference on Machine Learning. [S.l.: s.n.]: 448–456.
- KIM E J, BRUNNER R J, 2016. Star-galaxy classification using deep convolutional neural networks[J]. *Monthly Notices of the Royal Astronomical Society*: stw2672.
- KINGMA D P, BA J, 2014. Adam: A method for stochastic optimization[J]. *arXiv preprint arXiv:1412.6980*.
- KOHONEN T, 1988. An introduction to neural computing[J/OL]. *Neural Networks*, 1(1): 3 – 16. <http://www.sciencedirect.com/science/article/pii/0893608088900202>. DOI: [https://doi.org/10.1016/0893-6080\(88\)90020-2](https://doi.org/10.1016/0893-6080(88)90020-2).
- KRIZHEVSKY A, SUTSKEVER I, HINTON G E, 2012. Imagenet classification with deep convolutional neural networks[C]//Advances in neural information processing systems. [S.l.: s.n.]: 1097–1105.
- LAHAV O, NAIM A, BUTA R, et al., 1994. Galaxies, human eyes and artificial neural networks[J]. *arXiv preprint astro-ph/9412027*.
- LECUN Y, 1998. The mnist database of handwritten digits[J]. <http://yann.lecun.com/exdb/mnist/>.
- LECUN Y, BOTTOU L, BENGIO Y, et al., 1998. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 86(11): 2278–2324.
- LECUN Y, BENGIO Y, HINTON G, 2015. Deep learning[J]. *Nature*, 521(7553): 436–444.
- LECUN Y, OTHERS, 1989. Generalization and network design strategies[J]. *Connectionism in perspective*: 143–155.

- LINTOTT C, SCHAWINSKI K, BAMFORD S, et al., 2010. Galaxy zoo 1: data release of morphological classifications for nearly 900 000 galaxies[J]. *Monthly Notices of the Royal Astronomical Society*, 410(1): 166–178.
- LINTOTT C J, SCHAWINSKI K, SLOSAR A, et al., 2008. Galaxy zoo: morphologies derived from visual inspection of galaxies from the sloan digital sky survey[J]. *Monthly Notices of the Royal Astronomical Society*, 389(3): 1179–1189.
- MAATEN L V D, HINTON G, 2008. Visualizing data using t-sne[J]. *Journal of Machine Learning Research*, 9(Nov): 2579–2605.
- NAIM A, LAHAV O, SODRE L, et al., 1995. Automated morphological classification of apm galaxies by supervised artificial neural networks[J]. *Monthly Notices of the Royal Astronomical Society*, 275(3): 567–590.
- NAIR V, HINTON G E, 2010. Rectified linear units improve restricted boltzmann machines[C]//*Proceedings of the 27th international conference on machine learning (ICML-10)*. [S.l.: s.n.]: 807–814.
- NETZER Y, WANG T, COATES A, et al., 2011. Reading digits in natural images with unsupervised feature learning[C]//*NIPS workshop on deep learning and unsupervised feature learning: volume 2011*. [S.l.: s.n.]: 5.
- OWENS E, GRIFFITHS R, RATNATUNGA K, 1996. Using oblique decision trees for the morphological classification of galaxies[J]. *Monthly Notices of the Royal Astronomical Society*, 281(1): 153–157.
- PEDREGOSA F, GRAMFORT A, MICHEL V, et al., 2012. Scikit-learn: Machine learning in python[J]. *Journal of Machine Learning Research*, 12(10): 2825–2830.
- RAUBER P E, FADEL S G, FALCAO A X, et al., 2017. Visualizing the hidden activity of artificial neural networks[J]. *IEEE transactions on visualization and computer graphics*, 23(1): 101–110.
- ROWEIS S T, SAUL L K, 2000. Nonlinear dimensionality reduction by locally linear embedding[J]. *science*, 290(5500): 2323–2326.
- SANDAGE A, 1961. The hubble atlas of galaxies: volume 618[M]. [S.l.]: Carnegie Institution of Washington Washington, DC.
- SANDAGE A, BEDKE J, 1994. The carnegie atlas of galaxies. volumes i, ii.[J]. *The Carnegie Atlas of Galaxies. Volumes I, II., by Sandage, A.; Bedke, J.. Carnegie Institution of Washington Publ., No. 638., 638.*
- SIMMONS B D, LINTOTT C, WILLETT K W, et al., 2016. Galaxy zoo: quantitative visual morphological classifications for 48,000 galaxies from candels[J]. *Monthly Notices of the Royal Astronomical Society*: stw2587.

- SIMONYAN K, ZISSERMAN A, 2014. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556.
- SRIVASTAVA N, HINTON G E, KRIZHEVSKY A, et al., 2014. Dropout: a simple way to prevent neural networks from overfitting.[J]. Journal of Machine Learning Research, 15(1): 1929–1958.
- SZEGEDY C, LIU W, JIA Y, et al., 2015. Going deeper with convolutions[C]// Proceedings of the IEEE conference on computer vision and pattern recognition. [S.l.: s.n.]: 1–9.
- SZEGEDY C, VANHOUCKE V, IOFFE S, et al., 2016. Rethinking the inception architecture for computer vision[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.: s.n.]: 2818–2826.
- SZEGEDY C, IOFFE S, VANHOUCKE V, et al., 2017. Inception-v4, inception-resnet and the impact of residual connections on learning.[C]//AAAI. [S.l.: s.n.]: 4278–4284.
- TENENBAUM J B, DE SILVA V, LANGFORD J C, 2000. A global geometric framework for nonlinear dimensionality reduction[J]. science, 290(5500): 2319–2323.
- TORGERSON W S, 1952. Multidimensional scaling: I. theory and method[J]. Psychometrika, 17(4): 401–419.
- VAN S D W, SCHONBERGER J L, NUNEZIGLESIAS J, et al., 2014. scikit-image: image processing in python.[J]. Peerj, 2(2): e453.
- VAN DEN BERGH S, 1998. Galaxy morphology and classification[M]. [S.l.: Cambridge University Press.
- VAN DER MAATEN L, 2014. Accelerating t-sne using tree-based algorithms.[J]. Journal of machine learning research, 15(1): 3221–3245.
- VAN DER MAATEN L, POSTMA E, VAN DEN HERIK J, 2009. Dimensionality reduction: a comparative[J]. J Mach Learn Res, 10: 66–71.
- VAN DER MAATEN L, 2007. An introduction to dimensionality reduction using matlab [J]. Report, 1201(07-07): 62.
- WIKIPEDIA. Galaxy morphological classification[Z]. [S.l.: s.n.].
- WILLETT K W, LINTOTT C J, BAMFORD S P, et al., 2013. Galaxy zoo 2: detailed morphological classifications for 304 122 galaxies from the sloan digital sky survey[J]. Monthly Notices of the Royal Astronomical Society: stt1458.
- WILLETT K W, GALLOWAY M A, BAMFORD S P, et al., 2016. Galaxy zoo: morphological classifications for 120 000 galaxies in hst legacy imaging[J]. Monthly Notices of the Royal Astronomical Society, 464(4): 4176–4203.

- XU B, WANG N, CHEN T, et al., 2015. Empirical evaluation of rectified activations in convolutional network[J]. arXiv preprint arXiv:1505.00853.
- YOSINSKI J, CLUNE J, NGUYEN A, et al., 2015. Understanding neural networks through deep visualization[J]. arXiv preprint arXiv:1506.06579.
- ZAGORUYKO S, KOMODAKIS N, 2016. Wide residual networks[J]. arXiv preprint arXiv:1605.07146.
- ZEILER M D, 2012. Adadelta: an adaptive learning rate method[J]. arXiv preprint arXiv:1212.5701.
- ZEILER M D, FERGUS R, 2014. Visualizing and understanding convolutional networks [C]//European conference on computer vision. [S.l.]: Springer: 818–833.
- ZOONIVERSE. Edwin hubble, the man behind hst[Z]. [S.l.: s.n.].

作者简历及攻读学位期间发表的学术论文与研究成果

作者简历

2011 年 09 月–2015 年 06 月，在山东大学软件学院获得学士学位。

2015 年 09 月–2018 年 06 月，在中国科学院国家空间科学中心攻读硕士学位。

已发表 (或正式接受) 的学术论文:

[1] 戴加明, 佟继周. 基于深度残差网络的星系形态分类. 天文学进展, 2018

致 谢

同学说，论文最难写的不是正文，而是致谢部分。我虽没有回应他，但也把致谢拖到了最后一刻，可能难的不是如何写，而是不想和自己的三年，不想和身边的人告别吧。情不知何起，一往而深！请允许我向多年来给予我关心和帮助的老师、学长、同学、朋友和家人表示衷心的感谢！

首先衷心感谢我的导师佟继周老师对我的谆谆教诲和悉心指导。由于我执拗的性格和直白的语言，有时会不经意间顶撞了她，但每次她都不会生气，总是温柔地平心静气地和我沟通，非常耐心地帮我逐字逐句地修改小论文和大论文，看到她每次发我的修改意见时，已是凌晨一点多，真的让我万分愧疚。请原谅我的鲁莽和不善言语，再次向佟老师表示感谢！谢谢您对我的包容，谢谢您对我的谆谆教诲，谢谢您对我无私的帮助！

衷心感谢邹自明老师对我的悉心指导。邹老师看着不苟言笑，但其实是一个很棒的人。他是一个对学术非常认真，或者说较真的人。尽管他不是我的导师，但轮到我汇报工作时，他也会百忙之中抽空参加，他总是能够抓住问题的本质，言辞犀利，句句切中要害，指出我的问题，寻求解决方案。总之，非常敬佩邹老师的学术范！在此向邹老师致以崇高的敬意和诚挚的感谢，愿邹老师保重身体，工作顺利！

特别感谢李云龙博士。感谢李博士对我选题的支持和平日里的指导，他真的教会了我很多东西，无论科研中的还是生活中的。尽管他的工作很忙，他还是会百忙之中抽出时间听我汇报，帮我修改论文，感谢李博士无私的帮助！

同时还要感谢实验室的郑岩学长，研二进实验室时拿到的服务器完全是新的裸机，是他为我分配账号，配置环境，每次服务器出现问题时，总是很乐于帮助我，谢谢他的温文尔雅！并感谢空间科学数据中心提供的硬件支持，做深度学习是很烧钱的，没有 GPU 服务器，巧妇难为无米之炊，感谢数据中心的豪华配置！实验室中的其他同事也在我的研究生生活中给予了很多帮助，感谢胡晓彦学姐、钟佳学长、纪珍学姐等，他们在我生活和学习中遇到困难时给予了无私的帮助，让我感受到了实验室的温暖。

感谢中国科学院信息化专项“大数据驱动的空间科学领域创新示范平台”！

感谢运控室的李虎学长对我科研和生活上的帮助，每次遇到了解决不了的代码问题，虎哥总会耐心地帮我寻找 bug，生活中的烦恼也会经常向他倾诉，祝早

日博士毕业，工作顺利。

感谢我的室友邓翔、段永强、韩传奇、马吉宁、刘盼、熊安、朱晓攀和姚开一，感谢贾立丽、王元超、余文艳同学，遇见 432，是我的幸运，遇见你们，真好！愿我们的友谊长存！

感谢研究生部的李燕秋老师、许贺楠老师、云婷老师和何京勇老师对我的帮助和支持。

最重要的还要感谢我的父母，谢谢你们一直陪伴我，关心我，支持我，我爱你们！愿父母身体健康，万事顺心。

感谢各位专家百忙之中抽出时间参加我的论文评审和毕业答辩。

祝每一个人身体健康，工作顺利，一切平安！