

基于反贝叶斯学习的 WDMS 光谱自动识别研究

姜 斌, 赵梓良, 黄 灏, 钟云鹏, 赵永健, 曲美霞*

山东大学(威海)机电与信息工程学院, 山东 威海 264209

摘 要 天体光谱是天体物理学重要的研究对象, 通过光谱可以获取天体的许多物理、化学参数如有效温度、金属丰度、表面重力加速度和视向速度等。白矮主序双星是一类致密的双星系统, 对研究致密双星的演化特别是公共包层的演化有着重要的意义。国内外的大型巡天望远镜如美国斯隆望远镜以及中国的郭守敬望远镜, 每天都产生大量光谱数据。如此海量的光谱数据无法完全用人工进行分析。因此, 使用机器学习方法从海量的天体光谱中自动搜索白矮主序双星光谱, 有着非常现实的意义。目前的光谱自动识别方法主要通过已有的标签样本进行分析, 通过训练得到分类器, 再对未知目标进行识别。这类方法对样本的数量有明确的要求。白矮主序双星的实测光谱数量有限。若要通过有限的样本集准确学习白矮主序双星的光谱特征, 不仅需要扩大样本数量, 还需要提高特征提取和分类算法的精度。在前期工作中, 通过机器学习等方法在海量巡天数据中识别了一批白矮主序双星的光谱, 为该实验提供了数据源。使用对抗神经网络生成新的白矮主序双星光谱, 扩大训练数据量至原数据集约两倍的数量, 增强了分类模型的泛化能力。通过反贝叶斯学习修正损失函数, 将损失函数的大小与样本的方差相关联, 抑制了异常数据对模型造成的影响, 提升了模型的鲁棒性, 解决了由于训练样本集偏差带来的梯度消失以及训练陷入局部最优解等问题。该实验基于 Tensorflow 深度学习库。使用 Tensorflow 搭建的生成对抗网络具有较好的鲁棒性, 并且封装了内部实现细节, 使得算法得以更好地实现。除此之外, 由 Tensorflow 搭建的卷积神经网络在该实验中用于分类准确度测试。实验结果表明, 二维卷积神经网络能够利用卷积核有效地提取白矮主序双星的卷积特征并进行分类。基于反贝叶斯学习策略的卷积神经网络分类器在白矮主序双星原始数据及对抗神经网络生成光谱的识别任务中达到了约 98.3% 的准确率。该方法也可用于在巡天望远镜的海量光谱中搜索其他特殊和稀少天体如激变变星、超新星等。

关键词 白矮主序双星; 生成对抗网络; 反贝叶斯学习策略; 卷积神经网络

中图分类号: TP29 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2019)06-1829-05

引 言

白矮-主序双星(white dwarf-main sequence, WDMS)是一类密近双星, 主星是一颗白矮星(white dwarf, WD), 伴星通常是一颗小质量的主序星^[1-2]。在光学波段, WDMS 光谱特征非常明显, 可以分为两部分: 蓝端呈现白矮星特征, 出现了巴尔末线系的吸收线; 红端呈现小质量主序星特征, 出现了 TiO 等分子带。WDMS 对研究致密双星的演化特别是公共包层的演化有重要的意义。WDMS 的光谱数量相对较少, 应用机器学习方法在海量光谱中搜索 WDMS 并保证结果的完备性具有非常现实的意义。

图 1 为 SDSS(the sloan digital sky survey)^[3]中的一条 WDMS 光谱, 其波长范围为 3 800~9 000 Å(在图中分别对应于横坐标 0~3 522)。

在针对 WDMS 自动识别的研究中, 任娟娟等对郭守敬望远镜^[4]的巡天数据进行了系统的搜寻^[5]。郭格霖等使用 RBF(radial Basis function, RBF)的方法发现了一批新的候选体^[6]。Thomas 和 Oommen^[7]基于反贝叶斯学习策略(anti-Bayesian learning paradigm, ALP), 提出使用类别条件分布的分位数而不是平均值的信息来进行分类。该方法可以获得各种单维对称分布的最优分类以及不对称分布的近似最优精度。

生成对抗网络(generative adversarial nets, GAN)于

收稿日期: 2018-04-18, 修订日期: 2018-10-16

基金项目: 国家自然科学基金项目(11473019), 山东省自然科学基金项目(ZR2018MA032, ZR2017MA046)资助

作者简介: 姜 斌, 1977 年生, 山东大学(威海)机电与信息工程学院副教授 e-mail: jiangbin@sdu.edu.cn

* 通讯联系人 e-mail: whkunyushan@163.com

2014 年被 Ian J Goodfellow^[8]等提出, GAN 是一种无监督学习方法, 利用“对抗”的思想来训练生成模型, 并通过训练好的模型生成全新的数据样本。目前, GAN 已被广泛运用在图像处理领域^[9-10], 但其也可以被用于任何类型数据的生成。

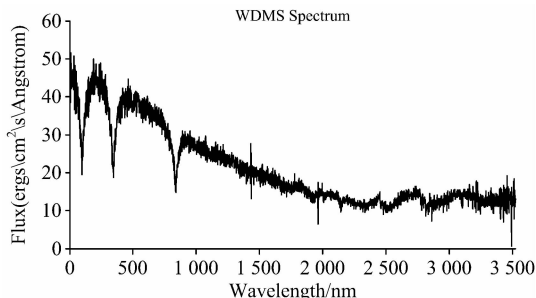


图 1 WDMs 光谱
Fig. 1 WDMs Spectrum

为提高对 WDMs 光谱特征提取的准确率, 使用反贝叶斯学习策略(ALP)修改了二维卷积神经网络的损失函数。与贝叶斯范式相反, ALP 将测试样本与分布的中心点相比较。ALP 技术基于远离平均值样本的信息, 通过使用基于指数移动平均的估计器或滑动窗口估计器来动态地跟踪每个类别的均值。除此以外, 实验过程中将原始数据与 GAN 的生成数据共同进行训练, 并在二维卷积神经网络模型上取得了较好的结果。

1 反贝叶斯学习策略

在海量天体光谱分类过程中, 训练集中光谱的信噪比和类型分布对分类模型至关重要。WDMs 光谱相对于其他类型的恒星光谱数量较少, 且具有信噪比不高、分布不均匀等特点, 使得分类模型在训练中容易产生过拟合现象。针对这一问题, 基于 ALP 来进行光谱卷积模型的训练。

ALP 假设不同光谱的类别是正态分布并使用次序统计量的一阶矩 $y \pm \sigma_D / \sqrt{2\pi}$ 作为均值。

使用均方误差(mean-square error, MSE)作为原始损失函数。在上述 ALP 假设下, 通过最大化数据似然, 导出带有正则化项的代价函数

$$J(\theta) = \frac{1}{N} \sum_{i=1}^N (y_i - \theta_i x_i)^2 + \gamma \sum_{i=1}^N |\theta_i| \quad (1)$$

式(1)中, $\gamma = 2\sigma_D / \sqrt{2\pi}$, 决定了 ALP 的强度, 同时也决定分布的偏倚程度。此时负对数似然性被重新定义为

$$\begin{aligned} -\log p(D|q) &= -\log \left[\frac{1}{\sqrt{2ps_D}} \exp \left\{ -\frac{(t-y+s_D/\sqrt{2p})^2}{2s_D^2} \right\} \right] \\ &= \frac{(t-y)^2}{2s_D^2} + \frac{t-y}{s_D \sqrt{2p}} + \frac{1}{4p} + C_1 \\ &= \frac{1}{2s_D^2} (t-y)^2 + \frac{1}{s_D \sqrt{2p}} (t-y) + C \end{aligned}$$

图 2 为 ALP 的概率模型示意图, 其中 $p^* = P\left(\frac{y-\sigma_D}{\sqrt{2\pi}} \leq T \leq y\right)$ 。

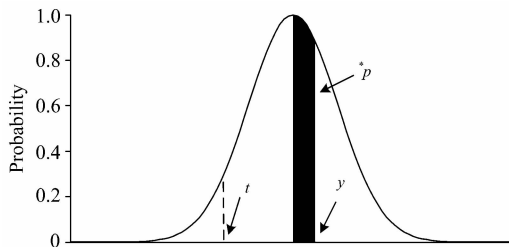


图 2 ALP 示意图
Fig. 2 ALP Schematic diagram

2 生成对抗网络

2.1 光谱获取及预处理

通过 Casjob^[11]获取 SDSS 光谱库中不同信噪比(signal noise ratio, SNR)的 WDMs 光谱共 1 746 条, 波长范围为 3 800~9 000 Å, 信噪比范围为 1~50。经过数据清洗除去异常数据以及信噪比过低的数据后, 保留光谱 1 600 条。

在实测光谱由神经网络的输入层进行正向传播运算前, 对原始光谱 S 通过最值归一化方法[式(1)]进行标准化预处理, 将其映射到[0, 1]区间内

$$S = \frac{S - S_{\min}}{S_{\max} - S_{\min}} \quad (2)$$

2.2 网络模型搭建

GAN 通常由两部分组成: 生成器 $G(z)$ 以及判别器 $D(z)$, 图 3 展示了 GAN 的基本原理。

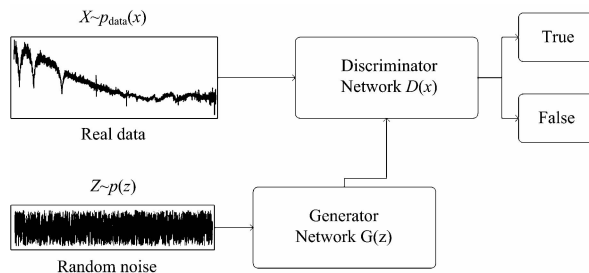


图 3 GAN 基本原理图

Fig. 3 GAN Basic schematic

表 1 生成对抗网络结构

Table 1 Network Structure for GAN

net	layer	neurons	af
$G(z)$	input	100	ReLU
$G(z)$	hidden	128	ReLU
$G(z)$	output	3 522	Sigmoid
$D(x)$	input	3 522	ReLU
$D(x)$	hidden	128	ReLU
$D(x)$	output	1	Sigmoid

$G(z)$ 从概率分布 $p(z)$ 中接收输入 z , 并试图生成样本数据; 而生成器以真实数据或生成数据作为输入, 并试图预测当前输入数据。在网络的平衡点, 判别器网络认为生成器网络输出的结果是真实数据的概率为 0.5

$$\min_G \max_D V(D, G) = E_{x \sim p_{\text{data}}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (3)$$

建立如表 1 所示的 GAN，各层的参数如下所示(位于网络 net，网络层次 layer，神经元个数 neurons，激活函数 af)。

3 实验过程及结论

3.1 生成光谱数据

由于光谱已经进行了归一化处理，将其直接送入 GAN 进行训练，并观察生成器和判别器的损失函数的变化。图 4 表明训练迭代 10 000 次之后，损失函数的值已经趋于平稳。

训练结束后，在生成的光谱集中挑选出生成质量较好的光谱，共选出 1 400 条，WDMS 样本集由原来的 1 600 条光谱扩大到了 3 000 条光谱。部分实测光谱和生成光谱如图 5

所示，左列的两个光谱为真实数据，右列为生成光谱。

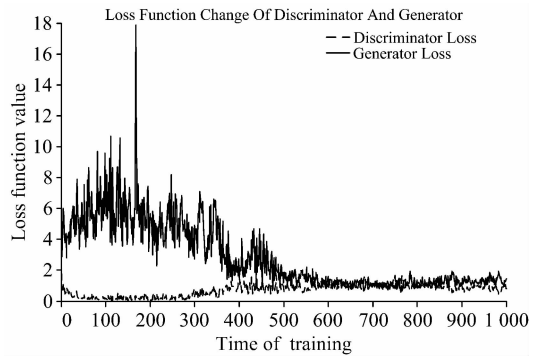


图 4 判别器和生成器的损失函数变化图

Fig. 4 Loss function change of discriminator and generator

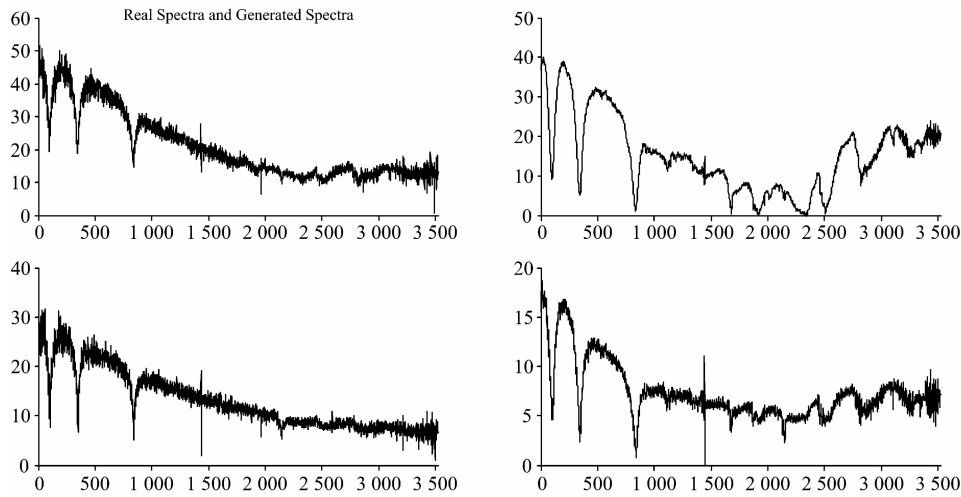


图 5 真实光谱和生成光谱图

Fig. 5 Real spectra and generated spectra

3.2 分类方法

选取 7 000 条负样例样本，与 3 000 条原始及生成的 WDMS 数据共同形成训练数据集。对光谱使用插值算法从 3 522 维扩增至 5 000 维后，将其堆叠为 50 × 100 大小的数据格式。如图 6 所示，折叠后绘制出的灰度图通过人眼很难区分是否为 WDMS 数据，但是通过 Pavel^[12] 等设计的二维卷积神经网络可以有效地提取光谱的特征。

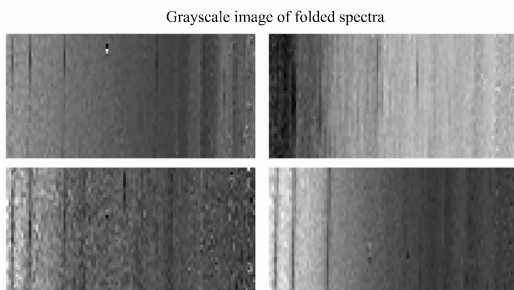


图 6 光谱经折叠后的灰度图

Fig. 6 Grayscale image of folded spectra

3.3 实验结论

使用结构化后的卷积神经网络进行 10 000 次迭代后，模型在测试数据集上达到了 98.3% ± 0.5% 的准确率。分类准确率随着训练次数的变化见图 7。

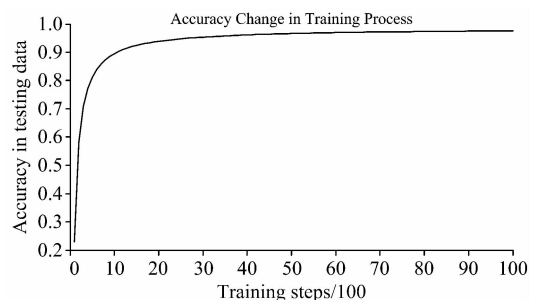


图 7 分类准确率随训练次数的变化

Fig. 7 Accuracy varies with training steps

在本实验中，由于 WDMS 光谱本身数量较少，使用 GAN 生成的部分数据依然存在一定的噪声。这使得生成的 WDMS 数据与原有数据存在一定的偏差，但由于谱线特征

提取较为明显,其并未明显影响分类准确度,反而能够使得模型在应对存在一定噪声的数据的分类问题上仍然有较好的表现。

4 结 论

重点研究如何通过反贝叶斯学习策略修改原有损失函数的正则化项,克服了由于训练集中的光谱抽样有偏引起泛化误差过大的问题,增强了模型的泛化能力。为使分类模型在针对光谱的训练过程中具有更好的抽样鲁棒性,使用生成对抗网络扩大了 WDMS 数据的样本范围。最终将光谱进行折

叠并使用二维卷积神经网络进行分类实验,取得了较好的实验结果。可以考虑和其他机器学习方法相结合,进一步提高分类精度。

实验证明, CNN 在测试数据集上能够达到 $98.3\% \pm 0.5\%$ 的分类准确率。二维卷积神经网络对于光谱分类是有效的,反贝叶斯学习策略对损失函数的修改有助于提高二维卷积神经网络的泛化能力,同时对 GAN 的使用也能够扩大训练样本,解决光谱数量不足的问题。

WDMS 的实测光谱数量相对较少,实验数据来自 SDSS 光谱,但本方法也适用于郭守敬望远镜这样的大型巡天项目。

References

- [1] Zasowski G, Johnson Jennifer A, Frinchaboy P M, et al. *The Astronomical Journal*, 2013, 146(4): 81.
- [2] Jiang Bin, Luo Ali, Zhao Yongheng, et al. *Monthly Notices of the Royal Astronomical Society*, 2013, 430(2): 986.
- [3] Wei Peng, Luo Ali, Wang Fengfei, et al. *The Astronomical Journal*, 2014, 147: 101.
- [4] Cui X Q, Zhao Y H, Chu Y Q, et al. *Research in Astron. Astrophys*, 2012, 12(9): 1197.
- [5] Ren J, Luo A, Li Y, et al. *AJ*, 2013, 146(4): 82.
- [6] WANG Wen-yu, GUO Ge-lin, JIANG Bin(王文玉, 郭格霖, 姜 斌, 等). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2016, 36(10): 3360.
- [7] Thomas A, John Oommen B. *Pattern Recognition*, 2013, 46: 376.
- [8] Goodfellow Ian J, Pouget-Abadie Jean, Mirza Mehdi, et al. *Generative Adversarial Networks*. arXiv: 1406. 2661.
- [9] Wu Xian, Xu Kun, Hall Peter, et al. *Tsinghua Science and Technology*, 2017, 22(6): 660.
- [10] TANG Xian-lun, DU Yi-ming, LIU Yu-wei, et al(唐贤伦, 杜一铭, 刘雨微, 等). *Acta Automatica Sinica(自动化学报)*, 2018, 44(5): 855.
- [11] Alam S, Albareti F D, Allende Prieto C, et al. *Astrophysical Journal Supplement*, 2015, 219: 12.
- [12] Pavel Hála. *Computer Science*, 2014, 1412: 8341.

Automatic Identification of WDMS Spectra Based on Anti-Bayesian Learning Paradigm

JIANG Bin, ZHAO Zi-liang, HUANG Hao, ZHONG Yun-peng, ZHAO Yong-jian, QU Mei-xia*

School of Mechanical, Electrical & Information Engineering, Shandong University, Weihai, Weihai 264209, China

Abstract Astronomical spectrum is an important research object in astrophysics. Many physical and chemical properties such as effective temperature, metal abundance, surface gravity and radial velocity can be inferred according to the spectra. The white dwarf main sequence binary star (WDMS) is a kind of binary star system, which is of great significance to the study of the evolution of binary stars, especially the evolution of post-common envelope. Domestic and foreign survey telescopes such as SDSS and LAMOST generate massive spectral data every day and such a large amount of spectral data cannot be analyzed manually. Therefore, it is very practical to use the machine learning method to automatically search for the WDMS spectra from the massive survey spectra. Current automatic spectral identification methods mainly depends on the existing labeled samples. Nevertheless, the number of WDMS spectra is limited. To accurately study the spectral features of WDMS spectra through a limited sample set, it is necessary to increase the number of samples and improve the accuracy of the feature extraction algorithm simultaneously. In the previous work, a batch of WDMS spectra was identified through machine learning methods in the sky survey data, providing data source for the experiment. In this paper, the generative adversarial network (GAN) is used to generate new WDMS spectra and expand the training data volume to about twice the original data set, which enhances the generalization ability of the classification model. By modifying the loss function by Anti-Bayesian learning method, the value of the loss function is correlated with the variance of the sample, which suppresses the influence of abnormally large data on the model. It improves the robustness of the model and solves the problems like vanishing gradient and getting stuck in a local optimal solution caused by the deviation of

the training sample. The experiments in this paper are based on the Tensorflow deep learning library. The GAN built by Tensorflow is robust and encapsulates the internal implementation details, making the algorithm itself better represented. In addition, the Convolutional Neural Network (CNN) built by Tensor flow was used in this experiment for classification accuracy testing. The experimental results show that the two-dimensional convolutional neural network can use the convolution kernel to effectively extract the convolution characteristics of WDMS spectra and classify them. The convolutional neural network classifier based on the anti-Bayesian learning strategy achieves an accuracy of about 98.3% in the identification task of original WDMS spectra and GAN generated data. The method can also be used to search for other specific targets such as cataclysmic variable stars or supernova in the massive spectra of the telescope.

Keywords WDMS; GAN; Anti-Bayesian; CNN

(Received Apr. 18, 2018; accepted Oct. 16, 2018)

* Corresponding author

《光谱学与光谱分析》对来稿英文摘要的要求

来稿英文摘要不符合下列要求者，本刊要求作者重写，这可能要推迟论文发表的时间。

1. 请用符合语法的英文，要求言简意明、确切地论述文章的主要内容，**突出创新之处**。
2. 应拥有与论文同等量的主要信息，包括四个要素，即研究目的、方法、结果、结论。其中后两个要素最重要。有时一个句子即可包含前两个要素，例如“用某种改进的 ICP-AES 测量了鱼池水样的痕量铅”。但有些情况下，英文摘要可包括研究工作的主要对象和范围，以及具有情报价值的其他重要信息。在结果部分最好有定量数据，如检测限、相对标准偏差等；结论部分最好指出方法或结果的优点和意义。
3. 句型力求简单，尽量采用被动式，建议经专业英语翻译机构润色，与中文摘要相对应。用 A4 复印纸单面打印。
4. 摘要不应有引言中出现的内容，换言之，摘要中必须写进的内容应尽量避免在引言中出现。摘要也不要对论文内容作解释和评论，不得简单重复题名中已有的信息；不用非公知公用的符号和术语；不用引文，除非该论文证实或否定了他人已发表的论文。缩略语、略称、代号，除相邻专业的读者也能清楚地理解外，在首次出现时必须加以说明，例如用括号写出全称。