

基于卷积神经网络的快速射电暴候选体分类*

刘艳玲^{1,2,3,4} 陈卯蒸^{1,2,3,4†} 李健^{1,3,4} 闫浩^{1,3,4} 袁建平^{1,2,3}

(1 中国科学院新疆天文台 乌鲁木齐 830011)

(2 中国科学院大学 北京 100049)

(3 中国科学院射电天文重点实验室 南京 210023)

(4 新疆微波技术重点实验室 乌鲁木齐 830011)

摘要 针对目前从海量的快速射电暴(Fast Radio Burst, FRB)候选体中人工筛选FRB事件难以为继的现状,提出了一种基于卷积神经网络(Convolutional Neural Networks, CNN)的FRB候选体分类方法. 首先,通过真实的观测数据和仿真FRB组成训练和测试样本集. 其次,建立了二输入的深度卷积神经网络模型,并对其进行训练、测试和优化,获取FRB候选体分类器. 最后,利用来自脉冲星的单脉冲数据对该分类器的有效性和性能进行了验证. 实验结果表明,该方法可以快速从大量候选体中准确识别出单脉冲事件,极大地提高了FRB候选体的处理速率和效率.

关键词 射电连续谱; 暂现源; 方法: 数据分析; 方法: 分类

中图分类号: P127; **文献标识码:** A

1 引言

快速射电暴(Fast Radio Burst, FRB)是一种明亮的、持续时间仅为数毫秒的射电天文瞬变现象^[1-3]. 由于其爆发时间极短且鲜少重复,捕捉快速射电暴事件非常困难. 到目前,FRB的许多属性仍然未知. 因此,需要基于大量的FRB观测数据进行研究,以回答与其起源和发射机理有关的问题. FRB是目前射电天文领域的热点前沿之一,近几年搜寻FRB的观测数据量急剧增多,特别是多波束接收机、相控阵接收机和望远镜阵列所产生的数据量呈指数级增长,如500 m口径球面射电望远镜(Five-hundred-meter Aperture Spherical Telescope, FAST)、澳大利亚平方公里阵探路者(Australian Square Kilometre Array Pathfinder, ASKAP)等,这给FRB的搜寻工作带来了前所未有

的挑战.

传统技术上,通常采用自动化、高性能的基于消色散理论的软件管道进行FRB事件搜寻. 如HEIMDALL(取自漫威漫画中的同名角色)^[4]、FDMT(Fast Dispersion Measure Transform)^[5]、Presto(Pulsar Exploration and Search Toolkit)^[6]和BEAR(Burst Emission Automatic Roger)^[7]等. 由于射频干扰(Radio Frequency Interference, RFI)的影响,这些算法面临着噪声和RFI导致的假阳性挑战. 最初,从大量候选体中筛选FRB事件的审查工作是由人工进行的. 但是,随着FRB观测数据量的急剧增加,假阳性候选体数量也相应增长. 因此,目前急需开发一种自动的FRB候选体高精度识别分类方法,解决人工处理时面对大量候选体难以为继的现状.

2021-10-08收到原稿, 2021-11-05收到修改稿

*国家自然科学基金项目(11903071), 中国科学院天文台站设备更新及重大仪器设备运行专项资助

†chen@xao.ac.cn

基于机器学习的目标检测与分类识别在射电天文领域已有诸多应用场景,如脉冲星候选体识别分类、天体光谱分类、引力透镜识别等.这也为快速、高效地筛选FRB事件提供了思路.针对FRB瞬态电波脉冲事件, Wagstaff等^[8]、Farah等^[9]、Foster等^[10]和Michilli等^[11]基于传统的机器学习方法实现了FRB候选体的识别分类.近些年,由于计算机技术的发展,特别是GPU (Graphics Processing Unit)技术的发展,使得深度学习在所有数据科学领域的信号分类、模式识别等方面取得了卓越的成绩.深度学习相较于传统的机器学习方法,它不需要有经验的专家花大量时间提取特征,避免了人工设计、选取特征造成的不完备性,且深度学习对于大数据集的学习能力已经远远超过了传统的机器学习. Connor等^[12]、Zhang等^[13]和Agarwal等^[14]基于深度学习技术实现了FRB事件的高精度、快速搜寻与识别,为FRB电压数据转储和多波段跟踪观测提供了条件. Connor等^[12]基于简单的两维卷积神经网络(Convolutional Neural Network, CNN)、1维CNN和前馈神经网络(Feedforward Neural Network, FNN)建立了四输入的网络模型,对FRB候选体的消色散动态谱图和色散度量(Dispersion Measure, DM)-时间图,时序轮廓图以及多波束探测信噪比信息分别提取特征,然后合并特征层,最后再进行识别. Agarwal等^[14]基于VGG (Visual Geometry Group) 16、VGG19、Densenet (Dense Convolutional Network) 121/169/201等经典的CNN复杂网络模型,组合建立了11个二输入(消色散动态谱和DM-时间强度数据)的二分类网络模型架构,分别对其进行了训练与测试. Zhang等^[13]则将深度学习分类器直接应用于色散动态谱数据中进行FRB事件搜寻.

本文综合Connor等^[12]和Agarwal等^[14]的研究结果,结合新疆天文台(Xinjiang Astronomical Observatory, XAO)南山26 m射电望远镜XFB (XAO Filter Bank)观测终端产生的FRB搜寻观测数据,基于深度学习模型开发了FRB候选体的自动分类处理算法.该方法能够节省人工筛选时间成本,提高FRB候选体的处理速度和效率.在后续章节中,将详细介绍样本库的建立、网络模型的架构设计、训练和测试及分类器性能的验证与分析等实验过

程.

2 数据集和数据预处理

2.1 数据集

目前,已检测到的FRB信号数量不足以为深度学习模型提供足够的训练样本.因此,也还没有现成的公共训练集可用.来自银河系脉冲星的单脉冲信号与FRB信号相似,但是直接采用脉冲星的单脉冲作为主要的训练集,可能会导致训练模型对脉冲星特性的过拟合^[12].本文参考Connor等^[12]的仿真方法,用高斯函数与散射轮廓卷积来生成FRB信号在每个频率通道上的脉冲轮廓,然后,注入到仅含有噪声和干扰的经过随机DM值消色散后的天文观测数据中,生成大量的消色散FRB仿真样本,DM在100–2000之间随机取值.天文观测数据是从filterbank或者FITS (Flexible Image Transport System)格式数据包中随机抽取.最终,选取其中信噪比大于8的FRB仿真样本组成正样本集. FRB仿真样本生成过程如图1所示.负样本集是由大量的真实观测数据通过HEIMDALL软件进行单脉冲搜寻处理后,在生成的候选体中提取其中信噪比大于8的干扰数据组成.

本文选择消色散动态谱数据和DM-时间强度数据来训练和测试网络模型.动态谱图信息最为丰富,可以展现FRB事件的频率结构信息. DM-时间强度图类似于蝴蝶结的形状,辨识度也比较高.通常它是通过对动态谱数据进行不同DM值的消色散处理之后,沿频率轴求和(或平均)生成的,DM在0–2 dm之间连续等间隔取值(dm为候选体的色散量).即DM轴对应的是动态谱数据经指定DM值消色散处理后,沿频率轴求和(或平均)的时间数据流.本文没有选择1维的时序轮廓数据作为网络输入的原因是它在形态上与较强的瞬态干扰区分度不高.另外, Connor等^[12]提到的多波束信噪比信息对于我们的望远镜观测系统也是不适用的.

图2是两例FRB仿真样本(a)和两例负样本(b)的示意图.每幅图从上到下依次是时序轮廓图,消色散动态谱图,DM-时间强度图.图2中的图(c)是为了更好地展示仿真的FRB形态,将仿真FRB叠加到高斯噪声上生成的样本.

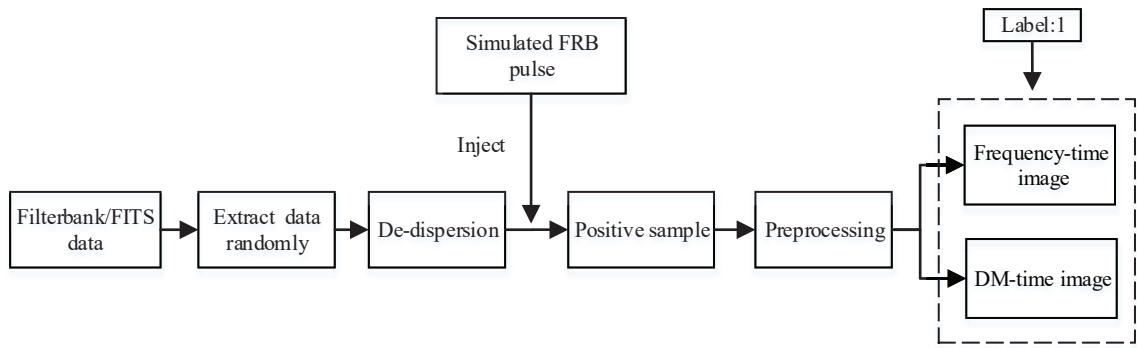


图1 FRB仿真样本生成过程

Fig. 1 Schematic diagram of simulated FRBs generation process

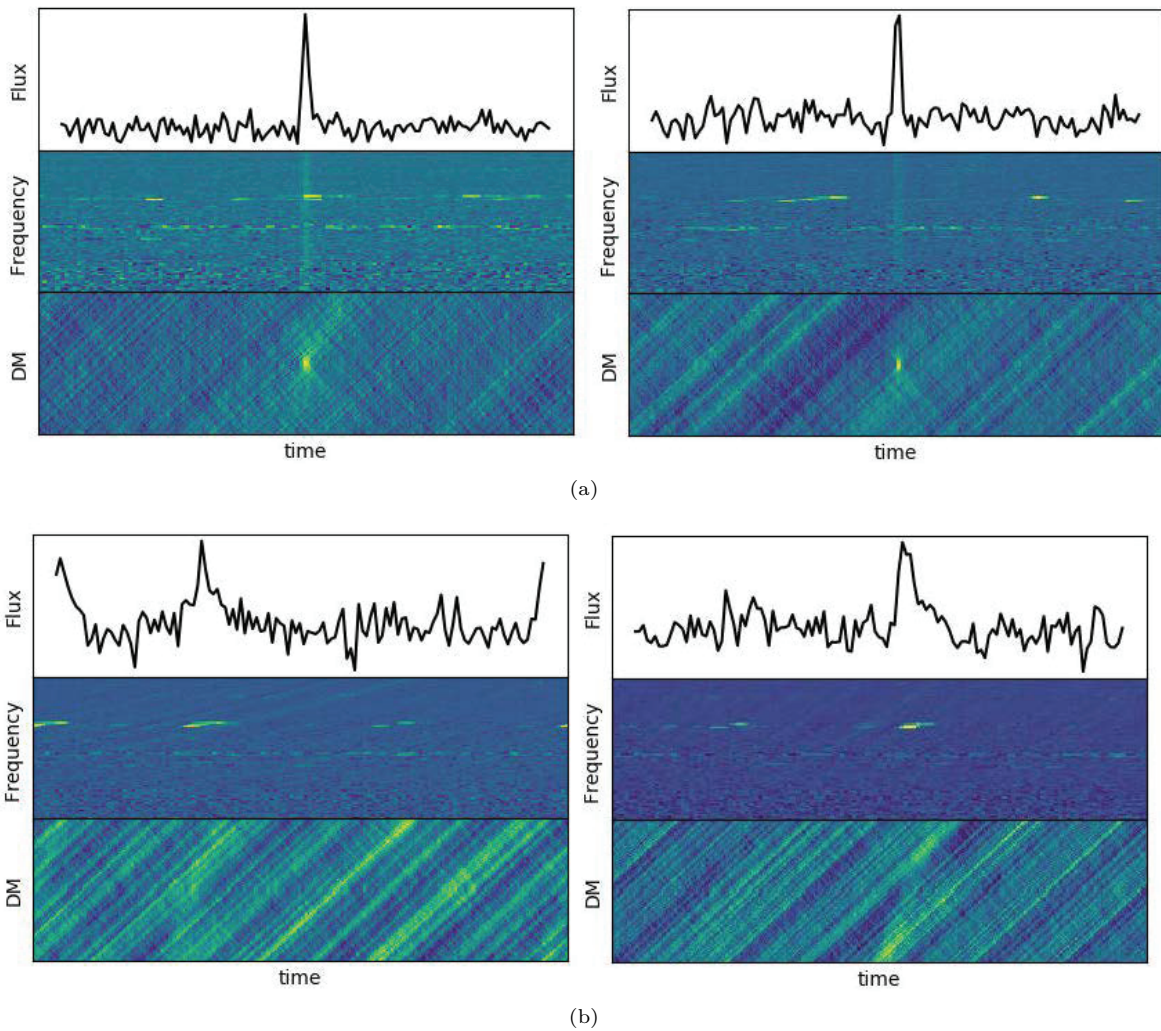
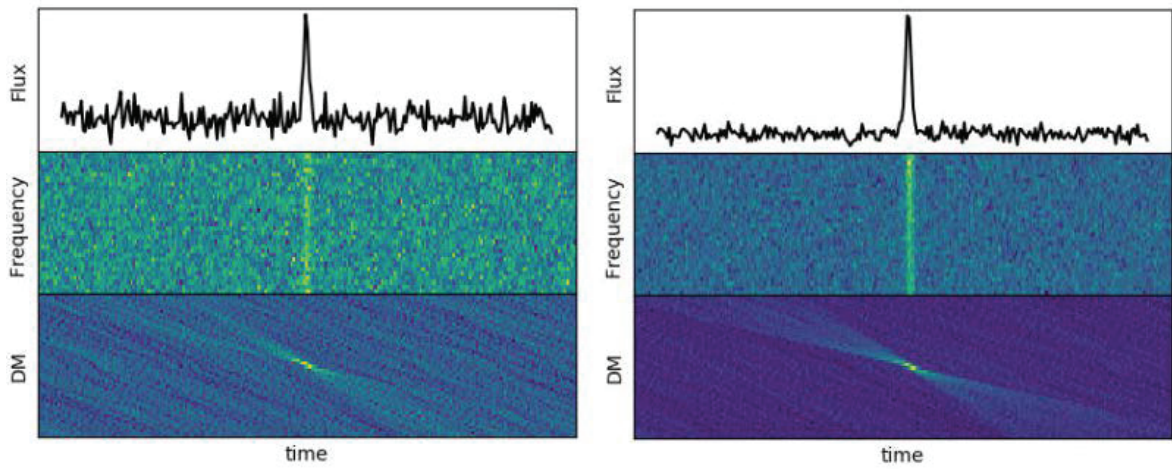


图2 仿真的FRB样本和来自真实观测数据的负样本. (a)仿真的FRB样本, (b)干扰样本(负样本), (c)仿真FRB叠加到高斯噪声上生成的样本.

Fig. 2 Examples of simulated FRBs and negative samples from real observation data. (a) The examples of simulated FRBs on observations, (b) RFI samples (negative samples), (c) the examples of simulated FRBs on Gaussian noise.



(c)

图 2 续

Fig. 2 Continued

2.2 数据预处理

本文所用数据为新疆天文台南山基地26 m射电望远镜观测数据,接收机和终端分别为L波段接收机和XFB数字终端. XFB数字终端的频率带宽是512 MHz,频率分辨率可设置为0.5 MHz和1 MHz,时间分辨率为64 μ s. 由于L波段接收机工作带宽为320 MHz (1400–1720 MHz),因此,在预处理过程中,会对超出接收机工作带宽的数据进行删除. 对于频率分辨率和时间分辨率,我们的卷积神经网络模型是不敏感的,可以通过在预处理过程中进行降

采样处理,达到调整数据尺寸的目的.

本文验证用的候选体样本是采用HEIMDALL软件对观测数据进行搜寻处理后生成的,其处理流程见图3. FRB观测数据经HEIMDALL搜寻处理之后,会生成1个FRB单脉冲候选体的文档. 该文档对每个候选体都提供了信噪比、样本位置、DM等信息. 通过该文档信息,从原始观测数据中提取候选体的动态谱数据,并对动态谱数据进行消色散处理以及DM变换,分别生成消色散动态谱数组和DM-时间强度数组. 最后,将其输入到训练好的卷积神经网络分类器中进行识别分类.

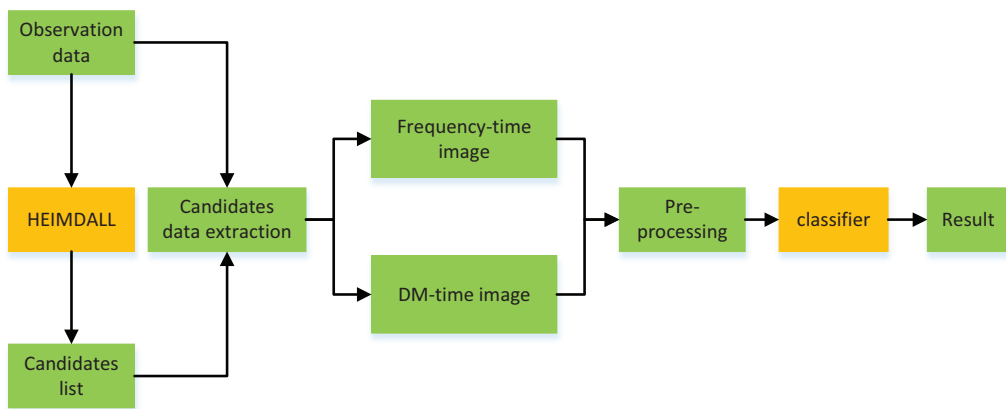


图 3 FRB候选体处理流程

Fig. 3 The processing flow chart of FRB candidates

无论是前期网络模型训练、测试所需的样本, 还是后期验证所需的候选体样本, 在其输入网络之前都需要进行预处理. 首先对样本数据进行统一的 z -score 标准化处理, 以消除奇异样本数据可能会对模型训练导致的不良影响. z -score, 也叫标准分数 (standard score), 是样本与平均数的差再除以标准差的过程. z -score 标准化处理计算如下:

$$z = \frac{x - \mu}{\sigma}, \quad (1)$$

其中, x 为样本数据, μ 为样本数据的均值, σ 为样本数据的标准差, 即

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}, \quad (2)$$

其中, N 为样本总的的数据量, i 为数值的序号. 经 z -score 标准化处理后的样本符合标准正态分布, 即均值为 0, 标准差为 1. 然后再对每个样本进行降采样和边缘修剪, 调整消色散动态谱数组大小为 160×125 和 DM-时间强度数组大小为 50×125 . 这样的尺寸大小既保证了信号质量, 同时又可以降低网络模型运算的复杂度.

3 网络模型

3.1 网络模型介绍

卷积神经网络是 1 类包含卷积计算且具有深度结构的前馈神经网络, 是深度学习的代表算法之一. 由于其在图像识别方面的出色表现, 目前已被广泛应用到众多领域中. 目前常见的卷积神经网络结构有 AlexNet (以第 1 作者 Alex 命名)^[15]、ZFNet (以两位作者 Zeiler 和 Fergus 的名字首字母命名)^[16]、VGGNet^[17]、Inception 系列^[18-21]、残差神经网络 (ResNet, 由算法 Residual 命名)^[22] 等, 这些模型都取得了不错的成绩. LeNet (以作者 LeCun 命名)^[23] (或称 LeNet-5) 是最早的卷积神经网络之一, 利用卷积、参数共享、池化等操作提取特征, 避免了大量的计算成本, 最后再使用全连接神经网络进行分类识别. 该网络虽小, 但是它包含了深度学习的基本模块: 卷积层、池化层以及全连接层, 是卷积神经网络最基础的模型. 对于深度神经网络

而言, 增加网络层数可以提升网络提取复杂特征的能力, 但是随着网络深度的增加, 网络会出现梯度消失或者爆炸问题. ResNet 是在 2015 年被提出的, 可以称作是里程碑式的创新. 其引入了残差模块结构 (residual block) 和残差模块的快捷连接 (shortcut connections) 解决了其他网络模型的训练精度随着网络深度增加而退化的问题. 目前基于该模型已经提出了许多功能与结构的改进, 发展出了许多变体. 2016 年, He 等^[24] 修改了 ResNet 的残差模块结构. 将批量标准化 (Batch Normalization, BN) 和修正线性单元 (Rectified Linear Units, ReLU) 移到了 Conv (convolution 的缩写) 层的前面, 即用 BN-ReLU-Conv 替代了传统的 Conv-BN-ReLU 结构, 起到了稳定网络收敛和防止过拟合的作用, 改进后的 ResNet 被称为 ResNet-V2. 此外, 更多的网络结构还有如宽残差神经网络^[25] (Wide Residual Network, WRN) 通过增加宽度提升网络性能以及 DenseNet^[26] 通过更密集的连接实现特征重用从而提高网络性能等.

3.2 模型架构

本文需要建立的是二输入的二分类网络架构. 两路分支网络分别对消色散动态谱数据和 DM-时间强度数据提取特征, 然后合并特征层. 在后续的实验中也证明, 针对本文的应用, 二输入模型比单独使用 1 个输入的效果要好. 类比于人类大脑, 当增加 1 个实物的多种形态展示时, 会提高大脑对该实物正确判断的概率.

相较于常见的人脸、动植物识别等应用场景, FRB 的图像信息相对简单. 虽然增加卷积神经网络的层数可以提高模型的性能, 但是也增加了模型的复杂度和训练难度, 另外, 层数的增加对模型性能的贡献也可能会达到饱和. 因此, 相较于 Agarwal 等^[14] 应用的复杂网络模型, 本文基于相对简单的 18 层的 ResNet-V2 网络结构和 7 层的 LeNet 网络结构建立模型, 分别用于完成对消色散动态谱数据和 DM-时间强度数据的特征提取. 在本文实验过程中, 起初尝试训练 ResNet18 网络模型识别 DM-时间强度数据, 其效果并不好, 网络呈现严重过拟合特点, 因此, 改用更简单的 LeNet 网络. LeNet 由 1 个输入

层、两个卷积层、两个池化层和3个全连接层组成, 结构虽简单却达到了我们想要的效果. ResNet18是由单独1个卷积层、8个残差块和1个全连接层构成. ResNet18包含两种残差块结构, 分别为恒等块和卷积块, 主要区别在于快捷连接是否含有卷积模块. 当快捷连接的两组数据通道数不一致时, 需要通过卷积块进行通道数调整.

基于ResNet18和LeNet的网络模型分别完成消色散动态谱数据和DM-时间强度数据的特征提取之后, 连接特征层, 再增加1个分类层, 最后网络输出分类结果. 本文采用基于TensorFlow为后端引擎的Keras快速建立网络模型. TensorFlow是端到端的开源机器学习平台, 由Tensor和Flow两个英文单词构成. Keras是用Python编写的开源人工神经网络库, 命名来自古希腊语, 意为将梦境化为显示的“牛角之门”. 模型架构如图4所示, 全连接层Dense后面括号内为选择的激活函数.

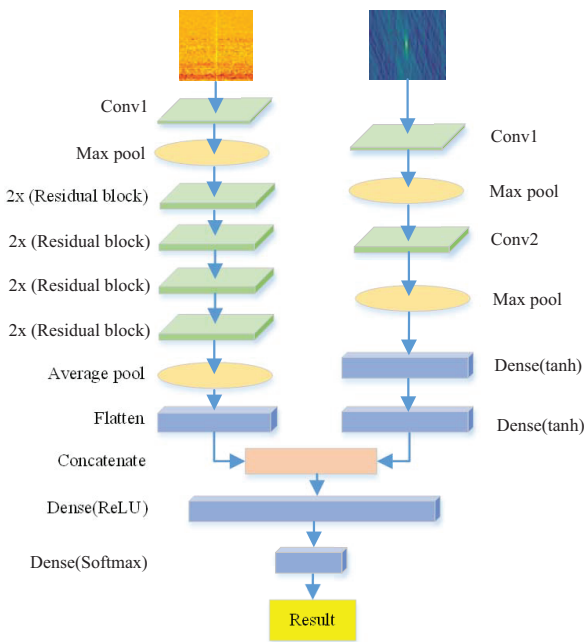


图 4 设计的二输入深度神经网络模型结构示意图

Fig. 4 The schematic diagram of our two-input deep convolutional neural network model

4 训练结果与分析

4.1 评估指标

本文研究的是针对FRB-RFI的二分类问题. 基于混淆矩阵方法, 采用准确率(Accuracy)、召回率(Recall)、精确率(Precision)来评估模型性能. 但是召回率和精确度是1对矛盾的度量, 在两个指标不能兼得的情况下, 还可以选用F1值(F-Score)指标来确定平衡点. 针对网络模型的测试结果, TP (True Positive)代表正样本被正确分类的数量, T-N (True Negative)代表负样本被正确分类的数量, FN (False Negative)代表被错分的负样本数量, FP (False Positive)代表被错分的正样本数量. Accuracy反映了模型正确分类样本的能力, 其计算公式如下:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}. \quad (3)$$

本文的研究目标倾向于尽可能少地错过FRB事件, Recall反映了模型识别正样本的能力, 其计算公式:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (4)$$

另外, 在分类结果中应尽量少地出现假阳性候选体, 从而降低实时电压数据转储的压力和后续人工审查的工作量, 对应的评价指标是Precision, 其计算公式:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (5)$$

F1是对召回率和精确率的调和平均值, 当召回率和精确率平衡的时候, F1值取得较高值, 取值范围为[0,1], 其计算公式:

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (6)$$

4.2 模型训练

为了提高网络训练效率和性能, 在反复的训练实验中, 采取了多种措施改进网络. Dropout层可以提升神经网络对未知数据处理的鲁棒性, 尽管抛弃了部分神经元, 网络也依然可以朝目标进行训练.

本文在LeNet的两个池化层和第2个全连接层后增加了Dropout层,用于预防网络过拟合. BN的使用可以避免因隐含层数据分布不均,导致梯度消失或不起作用的情况发生,进一步提高模型训练过程的稳定性,还可以加速模型收敛,减少训练时间. 除了ResNet18网络模型本身包含的BN处理外,在Flatten层后又增加了BN的使用. 训练过程采用简单交叉验证方法,将训练集与验证集的样本比例设置为8:2. 经过多次训练与调试,选定随机梯度下降

(Stochastic Gradient Descent, SGD)和自适应学习率(Adaptive Moment Estimation, Adam)分别作为ResNet18和LeNet的优化器来优化模型,表1显示了本文最终设置的网络训练参数. 另外,为了测试网络模型性能,在实验过程中,应用ResNet18和LeNet的网络模型分别对消色散动态谱数据和DM-时间强度数据进行了训练与测试. 其网络模型训练参数与测试准确率也一并在表1中展示. 通过对比,我们发现二输入网络模型的准确率是最高的.

表 1 网络模型的训练参数配置
Table 1 Training parameter configuration of network model

Parameter type	ResNet18	LeNet	Merge
Optimizer	SGD	Adam	Adam
Learning rate	0.01	/	/
Loss function	Binary cross entropy	Binary cross entropy	Binary cross entropy
Batch size	32	48	48
Number of epochs	15	20	10
BN momentum	0.9	/	/
Accuracy	99.0%	99.3%	99.8%

本文的训练和测试任务是基于1个GPU卡完成的,其型号是NVIDIA GTX 1080. 基于ResNet18网络模型的消色散动态谱数据训练耗时约25 min,基于LeNet网络模型的DM-时间强度数据训练耗时约几十秒. 本文采用的二输入合并网络模型在30 min之内即可以完成训练过程. 训练好的分类器完成单个样本检测的时间约为4 ms,运行速度非常快,这为后续实现电压数据转储提供了速度上的优势.

本文基于2019年6月2日的观测数据,建立了2.08万个样本,其中正样本和负样本的比例是1:1. 训练集与测试集的比例是8:2. 通常情况下,网络模型在经过6次以上的迭代训练之后,即能达到较高的性能,且训练误差和验证误差都能够收敛,两者之间相差很小. 如图5显示了某次模型训练过程中准确率变化情况(acc)和损失函数变化情况(loss).

在置信阈值设置为0.9时测试结果的混淆矩阵

如表2所示. 计算准确率为99.7%,召回率为99.8%,精确率为99.6%. 召回率和精确率均达到了较高值,F1值约为0.997.

表 2 具有仿真FRB样本的测试集的混淆矩阵
Table 2 Confusion matrix for test set with simulated FRBs

True class \ Predicted results	Positive	Negative	Total
	FRB	2016	8
RFI	4	2132	2136

4.3 模型验证与结果分析

由于选用的FRB训练样本由仿真生成,本文选取了新疆天文台南山26 m射电望远镜于2019年1月24日观测的脉冲星PSR J0332+5434的单脉冲共计300个,对网络模型进行测试验证. 图6为提取的

两个PSR J0332+5434的单脉冲样本示意图, 每幅图从上到下依次是色散动态谱图、消色散动态谱图和DM-时间强度图. 利用§2.2的方法, 从历史观测数据中提取了230个负样本, 与300个来自PSR J0332+5434的单脉冲组成验证集, 输入§4.2中所述训练好的网络模型中进行识别分类, 置信阈值设置

为0.9时的分类结果混淆矩阵如表3所示. 该模型针对脉冲星单脉冲的识别准确率为94.9%、召回率为99.6%、精确率为92.0%. 模型能够对几乎所有的脉冲星单脉冲进行正确识别和分类, 模型的召回率完全达到了我们的预期.

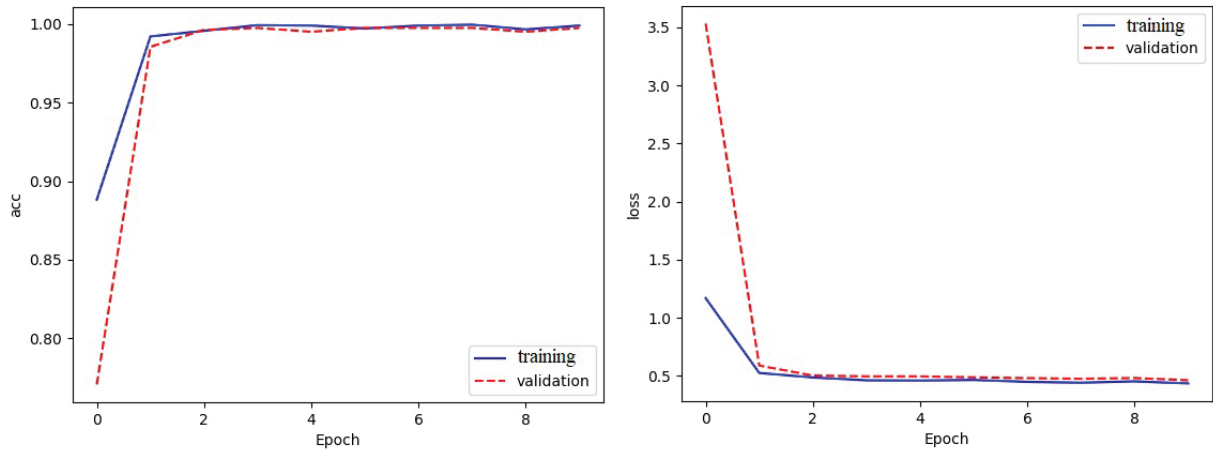


图 5 训练过程中的准确率变化曲线(左)和损失函数变化曲线(右)

Fig. 5 Accuracy (left) and loss (right) change curves with epoch of training process

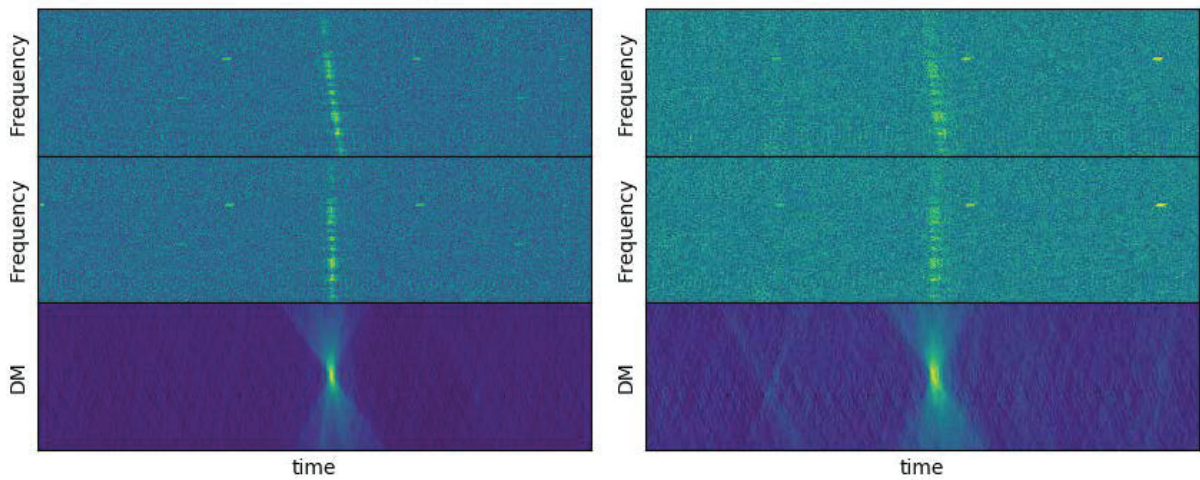


图 6 来自脉冲星PSR J0332+5434的单脉冲

Fig. 6 Single pulses from PSR J0332+5434

分类结果显示, 分类错误的发生主要是将部分RFI误判为FRB. 为了分析模型错判的原因, 对错分样本单独进行了分析. 发现在此次验证实验中选

用的RFI样本较之前训练时所用的样本增加了许多变化情况. 本次验证实验的RFI样本是从2020年5月4日的观测数据中提取的, 而训练样本选用的是

2019年6月2日的观测数据. 当我们将RFI样本重新从2019年6月的观测数据中提取, 再次进行验证实验时, 分类网络对RFI的召回率达到了98%以上.

表 3 具有脉冲星单脉冲的验证集的混淆矩阵
Table 3 Confusion matrix for validation set with pulses from pulsar

True class \ Predicted	Predicted results			Total
	Positive	Negative		
Single pulse	299	26		325
RFI	1	204		205

5 总结与展望

本文基于卷积神经网络实现了对FRB候选体的识别与分类. 网络模型的性能也达到了实验期望效果. 目前的实验中, 本文对于样本中的RFI并没有采取任何处理, 完全依赖卷积神经网络的训练和学习, 提高网络模型区分RFI和FRB的能力. 在实验中, 我们也发现训练好的卷积神经网络分类器的性能对于台址内外RFI环境或者观测系统改变引起的观测数据中的RFI变化非常敏感. 然而, 干扰种类繁多、来源复杂、特征呈多样性, 单纯使用简单粗暴的删除方法并不能取得明显的效果. 因此, 考虑针对训练样本中占少数的RFI类型或者新出现的RFI类型, 对训练样本做进一步的类别平衡处理, 可以改善网络模型识别RFI的能力. 后续, 将围绕观测数据中的RFI问题开展研究工作, 以期进一步提高分类器的泛化能力和鲁棒性.

致谢 感谢审稿人对文章提出的宝贵建议, 使得文章内容更加严谨和充实. 感谢新疆天文台南山基地26 m射电望远镜为本文的研究与实验提供了观测数据.

参考文献

- [1] Lorimer D R, Bailes M, Mclaughlin M A, et al. *Science*, 2007, 318: 777
- [2] Thornton D, Stappers B, Bailes M, et al. *Science*, 2013, 341: 53
- [3] Petroff E, Hessels J W T, Lorimer D R. *A&ARv*, 2019, 27: 4
- [4] Barsdell B R, Bailes M, Barnes D G, et al. *MNRAS*, 2012, 422: 379
- [5] Zackay B, Ofek E O. *ApJ*, 2017, 835: 11
- [6] Ransom S M. *New Search Techniques for Binary Pulsars*. Cambridge: Harvard University, 2001: 74
- [7] Men Y P, Luo R, Chen M Z, et al. *MNRAS*, 2019, 488: 3957
- [8] Wagstaff K L, Tang B Y, Thompson D R, et al. *PASP*, 2016, 128: 084503
- [9] Farah W, Flynn C, Bailes M, et al. *MNRAS*, 2018, 478: 1209
- [10] Foster G, Karastergiou A, Golpayegani G, et al. *MNRAS*, 2018, 474: 3847
- [11] Michilli D, Hessels J W T, Lyon R J, et al. *MNRAS*, 2018, 480: 3457
- [12] Connor L, Van Leeuwen J. *AJ*, 2018, 156: 256
- [13] Zhang Y G, Gajjar V, foster G, et al. *ApJ*, 2018, 866: 149
- [14] Agarwal D, Aggarwal K, Burke-Spolaor S, et al. *MNRAS*, 2020, 497: 1661
- [15] Krizhevsky A, Sutskever I, Hinton G E. *Proceedings of the 25th International Conference on Neural Information Processing Systems*. Lake Tahoe: Curran Associates Inc., 2012, 1: 1097
- [16] Zeiler M D, Fergus R. *Proceedings of 13th European Conference on Computer Vision*. Zurich: Springer, 2014: 818
- [17] Simonyan K, Zisserman A. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. *Proceedings of the 3rd International Conference on Learning Representations*, San Diego, May 7-9, 2015
- [18] Szegedy C, Liu W, Jia Y Q, et al. *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston: IEEE, 2015: 1
- [19] Ioffe S, Szegedy C. *Proceedings of the 32nd International Conference on International Conference on Machine Learning*. Lille: JMLR.org, 2015, 37: 448
- [20] Szegedy C, Vanhoucke V, Ioffe S, et al. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 2818
- [21] Szegedy C, Ioffe S, Vanhoucke V, et al. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*. San Francisco: AAAI Press, 2017: 4278
- [22] He K M, Zhang X Y, Ren S Q, et al. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 770
- [23] LeCun Y, Bottou L, Bengio Y, et al. *Proceedings of the IEEE*, 1998, 86: 2278
- [24] He K M, Zhang X Y, Ren S Q, et al. *Proceedings of the 14th European Conference on Computer Vision*. Amsterdam: Springer, 2016: 630
- [25] Zagoruyko S, Komodakis N. *Proceedings of the British Machine Vision Conference*. York: BMVA Press, 2016: 87
- [26] Huang G, Liu Z, Van Der Maaten L, et al. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 2261

Fast Radio Burst Candidate Classification with Convolutional Neural Networks

LIU Yan-ling^{1,2,3,4} CHEN Mao-zheng^{1,2,3,4} LI Jian^{1,3,4} YAN Hao^{1,3,4} YUAN Jian-ping^{1,2,3}

(1 Xinjiang Astronomical Observatory, Chinese Academy of Sciences, Urumqi 830011)

(2 University of Chinese Academy of Sciences, Beijing 100049)

(3 Key Laboratory of Radio Astronomy, Chinese Academy of Sciences, Nanjing 210023)

(4 Xinjiang Key Laboratory of Microwave Technology, Urumqi 830011)

ABSTRACT Manually identifying fast radio burst (FRB) events from the massive candidates by a human is a laborious and time-consuming task. It's an unsustainable working mode for the constantly growing volume of observation data. In this paper, we present a method of FRB candidates classification based on convolutional neural networks (CNN). First, we build training and test sets with real observation data and simulated FRBs. Second, a two-input deep convolutional neural network model is constructed, trained and optimized, and the FRB candidate classifier is obtained. Then, the effectiveness and performance of the classifier are tested and verified by using single pulses from pulsar. Experiment results show that this method can quickly and accurately identify single pulse events from candidates, which greatly improves the processing speed and efficiency of FRB candidates.

Key words radio continuum: transients, methods: data analysis, methods: classification