

基于卷积门控循环网络和注意力机制的短期负荷预测

赵瑞晓¹ 顾航¹
ZHAO Ruixiao GU Hang

摘要

为提高电力系统中短期负荷的预测精度，文章提出了一种基于注意力机制的卷积门控循环混合神经网络（ACGRU）的短期负荷预测模型。首先根据负荷时间序列中出现的混沌特征，利用互信息法（mutual information）和Cao法重构负荷数据的相空间；然后使用CNN对负荷相空间数据进行空间特征提取；再将提取后的空间特征序列与多变量时间序列融合，并利用GRU对其进行时间特征提取；最后利用注意力机制捕获时间序列的关键特征，输出预测结果。该模型对电工数学建模竞赛负荷预测数据集进行预测实验，与混合模型CNN-GRU、单一的网络模型CNN和LSTM的预测性能进行对比，在平均绝对误差率（MAPE）的性能评价中，所提模型算法的误差比CNN-GRU混合神经网络模型减少了0.98%，比CNN、GRU分别减少了3.22%、4.91%，负荷预测性能更为优越。

关键词

卷积神经网络；门控循环神经网络；注意力机制；混沌负荷预测

doi: 10.3969/j.issn.1672-9528.2026.02.042

0 引言

电力负荷数据具有时序性和非线性的特征，短期负荷预测领域的主要研究方法是统计学方法^[1-2]、机器学习^[3]、深度学习^[4]和混沌理论^[5]等。其中，将混沌理论应用到负荷预测中，主要是依据从混沌系统中提取出的非线性负荷时间序列，来对电力系统的未来情况展开预测^[6]，即使用相空间重构技术从负荷时间序列恢复混沌系统中的规律从而获取未来的负荷数据^[7]。文献[8]利用相空间重构将光伏功率序列投射至高维相空间后，使用增强型大脑情绪神经网络建立高维空间中的数据映射关系，并采用改进粒子群优化，实现光伏发电功率预测。文献[9]借助相空间重构（PSR）算法在重构时间序列相空间上的长处，以及双平方核（BSK）回归模型在同时兼顾原始光谱特征与空间信息方面的优势，构建了包含多维回归的PSR-BSK模型，用于单变量混沌时间序列的短期预测。

基于负荷序列的时间序列特性，许多学者将RNN引入负荷预测中，文献[10-12]提出基于LSTM算法的负荷预测，文献[11-12]结合CNN提取特征，同时在空间和时间领域使用深度学习来预测短期负荷预测。文献[13]在对混沌时间序

列进行相空间重构和聚类分析后，引入改进的粒子群算法的长短期记忆神经网络模型对负荷时间序列进行预测。近年来，有研究者将门控循环单元（GRU）运用到负荷预测领域，并且在其各自的实验数据集^[14-15]当中，都收获了较为理想的预测成效。在文献[16]中，卷积神经网络和门控循环单元（GRU）的组合模型被用来预测风速和住宅负荷，实验结果表明，该组合模型的预测准确度高于对比模型。但是，在输入时间序列比较长时，容易丢失序列信息，影响负荷模型的准确率。Attention机制是一种对不同资源赋予不同的权重，突出重点信息的资源分配机制，可以更好地处理长序列中的信息丢失问题。短期负荷序列受许多外生因素的影响，表现出复杂的特征，尤其是多维非线性，文献[4-6]为单变量时间序列的预测方法，将单一的时间序列进行相空间重构嵌入到高维空间序列后，对高维空间序列进行预测，忽略外生因素的影响。而短期负荷预测既受历史负荷时间序列数据影响，又与节假日、湿度、温度等因素密切相关。

短期负荷预测需要充分考虑负荷与其影响因素的关联关系，也需要考虑电力系统的内在随机性的因素。基于上述考虑，本文提出一种基于注意力机制（Attention）的卷积神经网络和门控循环单元混合算法模型（ACGRU）。首先在数据处理阶段，在对历史数据进行Pearson相关分析，选取最重要的特征变量的同时，还利用混沌理论中的相空间重构方法对历史负荷数据进行重构相空间。其次使用CNN捕获数据的空间特征与GRU获取数据的时间特征，使用注意力机

1. 江苏财经职业技术学院智能工程技术学院 江苏淮安 223000
[基金项目] “碳中和”背景下基于混合神经网络的城市用电负荷预测研究（HABL2023045）；江苏省软件和服务外包工程技术研究开发中心（2021607287）；基于OpenAI API智能信息检索工具的开发与应用（HABL2023039）

制进行特征分配,输出预测结果。最后将提出的模型应用于2016年电工数学建模竞赛负荷预测数据集,实验结果表明,本文方法可以很好的预测负荷数据,与其他预测方法相比,ACGRU模型预测精度更好。

1 数据预处理重构相空间

重构相空间是混沌系统的理论基础,最早是由 Packard 等^[17-18]提出在实践中一般采用延迟坐标法来实现相空间的重构。依据 Takens 定理,当嵌入维数处于合适范围时,单变量的负荷值能够实现对原负荷系统维数的恢复,同时确定嵌入维数 m 和延迟时间 τ 是重构相空间的关键。

设短期负荷时间序列为 $\{x_1, x_2, \dots, x_N\}$, $N = n - (m-1)\tau$, 其中嵌入维数 m 和延迟时间 τ 分别使用互信息法 (mutual information)^[19] 和 Cao 法^[20] 确定,重构后的相空间矩阵表示为:

$$T = \begin{pmatrix} x_1 & x_{1+\tau} & \dots & x_{1+(m-1)\tau} \\ x_2 & x_{2+\tau} & \dots & x_{2+(m-1)\tau} \\ \vdots & \vdots & \vdots & \vdots \\ x_n & x_{n+\tau} & \dots & x_{n+(m-1)\tau} \end{pmatrix} \quad (1)$$

数据集采用电工数学建模竞赛负荷预测数据集,短期负荷时间序列 x_i 是一种非线性混沌时间序列,本文对其进行相空间重构后作为模型输入。将进行相空间重构后的序列以及多变量时间序列需要进行归一化处理,标准化以及去除量纲,将其都统一到 (0,1) 的范围。本文使用的归一化函数为:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (2)$$

2 算法模型

本文采用基于注意力机制的 CNN-GRU 混合算法模型 (ACGRU),如图 1 所示。模型结构主要是由卷积神经网络、门控循环单元和注意力机制组成,首先利用 CNN 模型提取短期负荷序列重构相空间后的空间特征,并通过 GRU 获取空间特征下的时空特征以及多元特征序列的时序特征,最后通过注意力机制获得混合神经网络的关键时空特征。添加 Dropout 层,为防止训练过程中出现过拟合,提高模型的鲁棒性以及泛化能力。

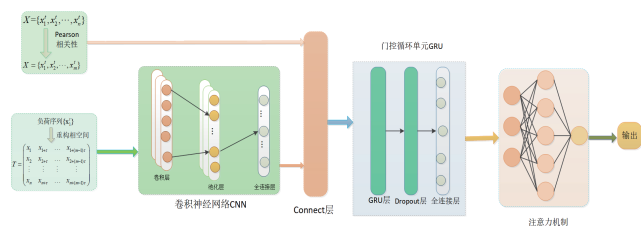


图 1 模型结构图

2.1 卷积神经网络 CNN

卷积神经网络^[21],是一种常用的深度学习网络,主要应用于时间序列、计算机视觉以及自然语言处理等领域。短期负荷序列属于时间序列数据的范畴,针对这一数据类型,主要采用的是一维卷积神经网络结构,相关结构如图 2 所示。卷积神经网络 (CNN) 的核心构成部分包含输入层、卷积层、池化层、全连接层和输出层。卷积层主要是提取时间序列数据的特征,即获取相空间数据中的空间特征,同时卷积神经网络是权值共享,可以减少网络中的训练参数的数量。

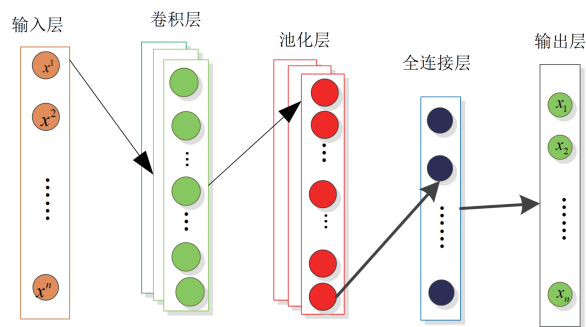


图 2 CNN 模型结构图

2.2 门控循环单元 GRU

循环神经网络 (RNN)^[22]被广泛应用于语音识别和自然语言处理,并非一次处理所有数据,而是以一种循环的方式处理数据。同时递归连接的优点是可以学习和记录序列中的历史信息,因此可以很好地预测时间序列^[23]上的数据特征。门控循环单元 (GRU) 是一种特殊的循环神经网络 (RNN),同 LSTM 一样,为了解决 RNN 在处理时序数据时,难以捕捉长期依赖,以及训练过程中梯度爆炸和梯度消失而提出^[24]。GRU 的结构比 LSTM 更加简单,结构图如图 3 所示,模型效果也比较好。GRU 的结构只有两个门结构,更新门和重置门,更新门是将 LSTM 的遗忘门和输入门合并在一起,控制前一刻信息状态在当前状态的保存程度,重置门是当前状态和先前信息的结合程度。

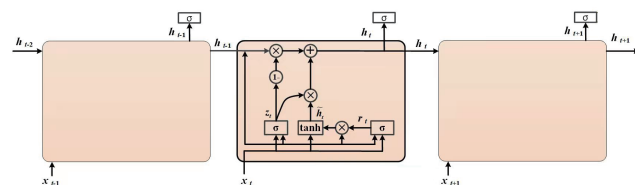


图 3 LSTM 结构图

若假设 x_t 为输入; h_t 为隐含层的输出; z_t 为更新门; r_t 为重置门,则 GRU 单元的 h_t 计算公式为:

$$r_t = \sigma(W_r x_t + U_r h_{t-1}) \quad (3)$$

$$z_t = \sigma(W_z x_t + U_z h_{t-1}) \quad (4)$$

$$\tilde{h}_t = \tanh(W_h x_t + r_t \odot h_{t-1}) \quad (5)$$

$$h_t = z_t h_{t-1} + (1 - z_t) \tilde{h}_t \quad (6)$$

式中： W_r 、 W_z 、 W 和 U_r 、 U_z 分别为权重向量和偏置向量； σ 和 $\tanh$ 分别为 sigmoid 激活函数和双曲线正切函数。

3 误差评价

对于预测误差的评价，文本选取了两类评估指标：常用的评价误差标准平均绝对误差率（MAPE）、常用于评价时间序列预测性能的均方根百分比误差（RMSPE）。

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \bar{y}_i|}{y_i} \times 100\% \quad (7)$$

$$\text{RMSPE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \bar{y}_i}{y_i} \right|^2} \cdot 100 \quad (8)$$

式中： n 为样本数目； y_i 与 $\bar{y}_i$ 分别为真实值和预测值。

4 实验分析与讨论

本次实验使用设备 CPU 配置为 Intel(R) Core(TM) i5-5200U CPU，内存为 12 GB，Python 为 3.7.7 版本，编辑器为 Jupyter notebook。本文选取数据预处理后的数据集，并按 8:2 的比例分为训练集和测试集，实验使用本文方法 ACGRU 模型进行预测，与 CNN、LSTM 等单一的预测模型以及 CNN-GRU 的混合神经网络模型进行对比预测，同时将预测误差以及运行时间作为评价指标。

4.1 实验结果及误差性能对比

实验对本文使用的预测模型 ACGRU 以及单一的 CNN、LSTM 神经网络模型以及混合神经网络模型 CNN-GRU 进行预测实验。实验结果如图 4-7 所示，分别为负荷预测以及真实值的对比图，由图可以得知实线为真实的负荷值，虚线为预测的真实值，可以看出，本文使用的方法预测值与真实值的重合程度较好，特别是在负荷峰值预测较其他预测方法更为贴近真实值的负荷功率，其他方法在峰值预测效果较差。在模型图的拟合中，可以看出本文的拟合性较好，在拐点还是其他地方的拟合性都比较好，其次是 CNN-GRU 模型除了在峰值和拐点处的拟合性较差外其他的真实值和预测值重合性较好，然后是 LSTM 的模型，CNN 的拟合性最差。

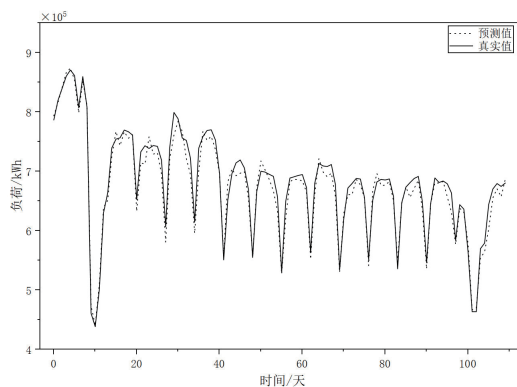


图 4 本文模型预测对比图

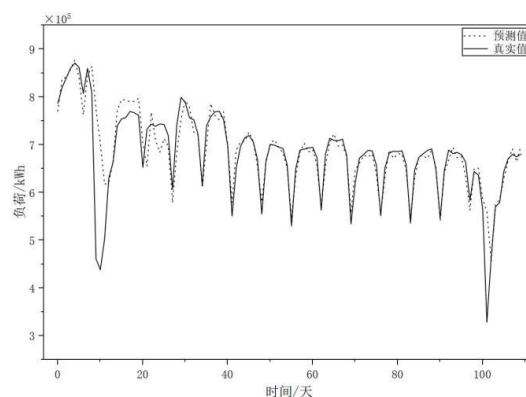


图 5 LSTM 预测对比图

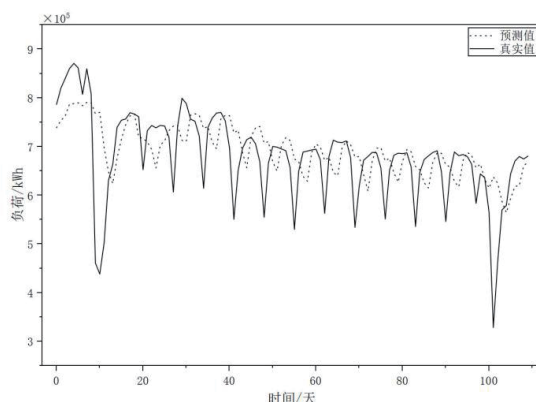


图 6 CNN 预测对比图

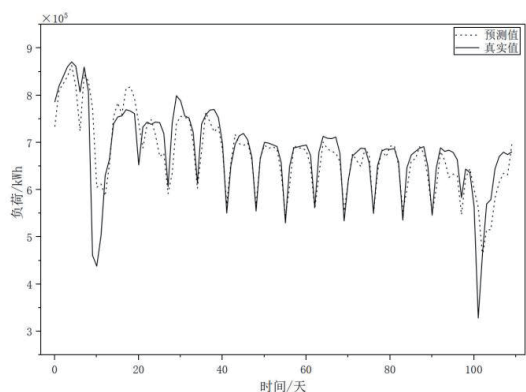


图 7 CNN-GRU 预测对比图

由表 1 可知，在 MAPE 的误差评价中，本文使用的模型 ACGRU 较好，CNN-GRU 模型次之，其次为 LSTM 模型，CNN 模型较差。则在 RMSPE 的评价中，CNN 的表现较好，LSTM 较差。同时在 MAPE 指标的比较下，本文使用的模型比 CNN-GRU 混合模型减少了 0.98% 的误差，并且比单一的模型 LSTM、CNN 分别减少了 3.22%、4.91% 的误差。在 RMSPE 的误差评价中，本文模型算法的误差比 CNN-GRU 混合神经网络模型减少了 0.87%，比 CNN、GRU 分别减少了 6.34%、6.97%。可以看出混合神经网络模型较单一的模型表现比较好，结合单一模型的优点，本文模型结合不

同模型的优点,注意力机制在混合模型中,提取关键的影响因素特征,提高了模型的预测精度,结合表2可知,不同模型的训练时间耗时间不同,本文模型使用的方法训练时间为42.910 s,训练时间比其他模型要短,但是测试时间相比较长也不超过2 s。综合考虑,本文使用的模型相比较而言有更好的预测效果,更具有优势。

表1 模型性能对比表

模型	RMSPE	MAPE
本文模型	0.105 7	0.045 3
LSTM	0.175 4	0.077 5
CNN	0.169 1	0.094 4
CNN-GRU	0.114 4	0.055 1

表2 模型训练时间

模型	训练时间/s	测试时间/s
本文模型	429.10	1.910
LSTM	517.43	1.939
CNN	434.22	1.371
CNN-GRU	443.40	1.379

5 结论

在负荷时间序列预测的基础上,本文提出了基于注意力机制的卷积门控混合神经网络模型(ACGRU),结合负荷数据相空间重构,使用CNN提出负荷系统相空间下的空间特征,使用注意力机制获取关键特征,从而得到更精确的预测精度。本文方法既考虑到负荷时间序列的短期负荷预测方法的优势,同时也考虑到实际的负荷是多种影响因素共同作用。实验结果表明,本文提出的模型比其他实验模型取得了更好的实验结果,预测精度也较高,在峰值预测效果更好。

参考文献:

- [1] KHWAJA A S, ZHANG X, ANPALAGAN A, et al. Boosted neural networks for improved short-term electric load forecasting[J]. Electric power systems research, 2017, 143: 431-437.
- [2] MASSAOUDI M, REFAAT S S, CHIHI I, et al. A novel stacked generalization ensemble-based hybrid LGBM-XGB-MLP model for short-term load forecasting[J]. Energy, 2021, 214: 118874.
- [3] LIANG Y, NIU D X, HONG W C. Short term load forecasting based on feature extraction and improved general regression neural network model[J]. Energy, 2019, 166: 653-663.
- [4] LU J, ZHANG Q, YANG Z, et al. Short-term load forecasting

method based on CNN-LSTM hybrid neural network model[J]. Automation of electric power systems, 43(8): 131-137.

- [5] LIU Y Z, XU N. An IGA-WLSSVR short-term load forecasting model based on chaotic time series[J]. Control engineering, 2021, 28(2): 245-250.
- [6] 吕金虎, 陆君安, 陈士华. 混沌时间序列分析及其应用[M]. 武汉: 武汉大学出版社, 2020.
- [7] 丁明, 虞海彪, 刘练, 等. 基于多变量相空间重构和RBF神经网络的光伏功率预测方法[J]. 电子测量与仪器学报, 2020, 34(8): 1-7.
- [8] 王育飞, 杨启星, 薛花. 考虑混沌特征的增强型大脑情绪神经网络光伏发电功率超短期预测模型[J]. 高电压技术, 2021, 47(4): 1165-1177.
- [9] 陈黎君. 基于改进的EMD-PSR单变量混沌时间序列短期预测研究[D]. 武汉: 武汉理工大学, 2020.
- [10] LU J X, ZHANG Q P, YANG Z H, et al. A hybrid model based on convolutional neural network and long short-term memory for short-term load forecasting[C/OL]. // 2019 IEEE Power & Energy Society General Meeting (PESGM). Piscataway: IEEE, 2019[2025-12-30]. <https://ieeexplore.ieee.org/document/8973549>. DOI: 10.1109/PESGM40551.2019.8973549.
- [11] 王晨阳, 段倩倩, 周凯, 等. 基于遗传算法优化卷积长短记忆混合神经网络模型的光伏发电功率预测[J]. 物理学报, 2020, 69(10): 149-155.
- [12] KOUHI S, KEYNIA F, RAVADANEGH S N. A new short-term load forecast method based on neuro-evolutionary algorithm and chaotic feature selection[J]. International journal of electrical power & energy systems, 2014, 62: 862-867.
- [13] FAN G F, PENG L L, HONG W C. Short term load forecasting based on phase space reconstruction algorithm and bi-square kernel regression model[J]. Applied energy, 2018, 224: 13-33.
- [14] DONG M, GRUMBACH L. A hybrid distribution feeder long-term load forecasting method based on sequence prediction[J]. IEEE transactions on smart grid, 2020, 11(1): 470-482.
- [15] SEHOVAC L, NESEN C, GROLINGER K. Forecasting building energy consumption with deep learning: a sequence to sequence approach[C]. // 2019 IEEE International Congress on Internet of Things (ICIOT). Piscataway: IEEE, 2019: 108-116.
- [16] AFRASIABI M, MOHAMMADI M, RASTEGAR M, et al. Probabilistic deep neural network price forecasting based on residential load and wind speed predictions[J]. IET renewable power generation, 2019, 13(11): 1840-1848.

基于卡警数据的路口排队长度计算模型研究

常宗智^{1,2}
CHANG Zongzhi

摘要

随着汽车保有量的快速增加，城市交通拥堵问题日益严重，交叉口排队长度是评估交通运行效率的关键指标。基于此，文章提出一种结合卡警抓拍数据与信号灯配时参数的路口排队长度近似计算方法。首先，通过卡警数据提取车辆到达时间、瞬时速度及停车次数，结合信号相位划分交通流周期；其次，基于冲击波理论构建排队形成与消散的动态模型，推导绿灯亮起后车流消散波传播速度与集结波回溯规律；最后，通过实际案例验证模型精度。结果表明，该方法能够有效减少传统检测器的部署成本，为动态信号优化提供理论支持。

关键词

卡警抓拍数据；信号灯配时；冲击波模型；排队长度

doi: 10.3969/j.issn.1672-9528.2026.02.043

0 引言

城市交叉口是交通网络的瓶颈，其通行效率直接影响整个路网的运行效率。随着车辆保有量的增加，交叉口排队溢出现象频发，导致交通延误和拥堵加剧，尤其是在早晚高峰时段。传统的排队长度检测方法依赖于线圈检测器或路口雷

达，但其易损坏且维护成本高。近年来，车牌识别系统（LPR，即电警、卡口数据）因其高精度和易获取性，逐渐成为交通状态监测的重要数据源。

目前，国内外学者对信号交叉口排队长度的研究，根据数据来源不同，研究方式主要分为3种：基于定点数据的研究、基于移动数据的研究和基于多源数据融合的研究。例如，贾利民等^[1]利用车辆行驶经过传感器的时刻与车尾时距估计车辆排队的长度。由于车辆传感器的覆盖范围有限，且

1. 东南大学 江苏南京 210096
2. 南京金智视讯技术有限公司 江苏南京 211100

[17]PACKARD N H, CRUTCHFIELD J P, FARMER J D, et al. Geometry from a time series [J]. Physical review letters, 1980, 45(9): 712-716.

[18]TAKENS F . Detecting strange attractors in turbulence[C]// Dynamical Systems and Turbulence, Warwick 1980. Berlin:Springer,2006:366–381.

[19]MARTINERIE J M , ALBANO A M , MEES A I , et al. Mutual information, strange attractors, and the optimal estimation of dimension[J]. Physical review A, 1992, 45(10):7058-7064.

[20]CAO L Y . Practical method for determining the minimum embedding dimension of a scalar time series[J]. Physica D nonlinear phenomena, 1997, 110(1/2):43-50.

[21]SHERSTINSKY A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network[J]. Physica D nonlinear phenomena, 2020, 404:132306.

[22] 黄伟建, 李永涛, 黄远. 基于混合神经网络和注意力机制

的混沌时间序列预测 [J]. 物理学报 ,2021,70(1):235-243.

[23] YAN R , LIAO J Q, YANG J , et al. Multi-hour and multi-site air quality index forecasting in Beijing using CNN, LSTM, CNN-LSTM, and spatiotemporal clustering[J]. Expert systems with applications, 2020, 169:114513.

[24] JIAO R H, ZHANG T M, JIANG Y Z, et al.Short-term non-residential load forecasting based on multiple sequences LSTM recurrent neural network[J].IEEE access,2019,6: 59438-59448.

【作者简介】

赵瑞晓（1996—），女，河南濮阳人，硕士研究生，助教，研究方向：人工智能、计算机应用。

顾航（1992—），男，江苏沭阳人，硕士研究生，助教，研究方向：人工智能、计算机应用。

（收稿日期：2025-07-03 修回日期：2026-02-04 ）