

基于时域卷积网络的多尺度双线性天气预测模型

孔 震¹, 张华鲁^{1,2}, 岳圣凯³, 袁明磊³, 路 通³

(1. 南瑞集团有限公司, 江苏 南京 211106;
2. 东北电力大学电气工程系, 吉林 吉林 132012;
3. 南京大学计算机软件新技术国家重点实验室, 江苏 南京 210023)

摘 要: 针对时域卷积网络(TCN)提取能力受卷积层感受野限制, 难以对天气数据中的季节性信息和长时信息进行有效提取与分析的问题, 提出了一个新的基于 TCN 的多尺度双线性天气预测模型。该模型由 TCN 层和双线性汇合层 2 部分组成, 时域卷积层包含双路 TCN, 每个 TCN 利用历史观测数据独立提取特征, 除卷积核尺度之外, 其他网络参数均保持一致。多尺度的网络组合可以更深入挖掘数据中潜在关联信息; 时域卷积层的输出作为双线性汇合层的输入进行双线性融合, 规范化后得到最终输出, 即对未来天气的预测值, 进一步提升模型的特征表示能力。在公开的天气预测数据集上与 5 个基准方法进行对比, 实验结果表明所提方法的预测结果准确率更高; 此外, 对比 TCN, 多尺度双线性天气预测模型面对长时数据信息时表现更加稳定。

关 键 词: 时域卷积网络; 多尺度融合; 双线性汇合; 数据挖掘; 天气预测
中图分类号: TP 391 **DOI:** 10.11996/JGj.2095-302X.2020050764
文献标识码: A **文章编号:** 2095-302X(2020)05-0764-07

Multi-scale and bilinear models based on temporal convolutional network

KONG Zhen¹, ZHANG Hua-lu^{1,2}, YUE Sheng-kai³, YUAN Ming-lei³, LU Tong³

(1. Nari Group Corporation, Nanjing Jiangsu 211106, China;
2. Department of Electrical Engineering, Northeast Electric Power University, Jilin Jilin 132012, China;
3. National Key Lab for Novel Software Technology, Nanjing University, Nanjing Jiangsu 210023, China)

Abstract: The learning capability of such convolutional network models as temporal convolutional network (TCN) is limited to the scale of receptive field, and sometimes it is difficult to mine data from seasonal or long-term information, such as weather data for deep convolutional networks. In order to address this problem, a novel multi-scale and bilinear convolutional neural network was proposed, which was composed of the TCNs layer and bilinear layer. The TCNs layer contained two temporal convolutional networks, each of which extracted useful features independently from the

收稿日期: 2020-03-03; 定稿日期: 2020-06-16

Received: 3 March, 2020; Finalized: 16 June, 2020

基金项目: 国家电网有限公司总部科技项目

Foundation items: Scientific Foundation of State Grid Corporation of China

第一作者: 孔 震(1979-), 男, 江苏仪征人, 高级工程师, 硕士。主要研究方向为大数据、信息通信、人工智能及其应用等。

E-mail: kongzhen@sgepri.sgcc.com.cn

First author: KONG Zhen (1979-), male, senior engineer, master. His main research interests cover information and communication, artificial intelligence and applications. E-mail: kongzhen@sgepri.sgcc.com.cn

通信作者: 路 通(1976-), 男, 江苏盐城人, 教授, 博士, 博士生导师。主要研究方向为视频与图像数据处理、模式识别与计算机图形学等。

E-mail: lutong@nju.edu.cn

Corresponding author: LU Tong (1976-), male, professor, Ph.D. His main research interests cover video and image data processing, pattern recognition and computer graphics. E-mail: lutong@nju.edu.cn

historical data. All TCNs had the same parameters except for the kernel size, thereby further mining the underlying relevant information in data. The output of the TCNs layer was employed as the input of the bilinear layer, and after normalization, the final output could be eventually calculated, which enhanced feature representation capabilities of the model. The comparisons with five methods on the public datasets of weather forecast show that the proposed method can effectively improve the accuracy and performance in terms of the long-time data much better than the TCN.

Keywords: temporal convolutional network; multi-scale convolutional; bilinear fusion; data mining; weather forecast

天气预测在电力系统的电力规划、防灾减灾、实时监测等诸多领域都有着非常重要的作用,如光伏发电系统的输出功率很大程度上取决于光伏面板所能接收到的太阳辐射量,且与多种天气因素直接相关,光伏输出功率的预测则离不开天气的预测^[1]。再如风力发电系统,由风速随机变化引起的风电功率波动直接影响电力系统的稳定和控制^[2]。在防灾减灾方面,随着全球气候变暖和环流异常,极端天气的出现愈加频繁,近几年我国南方发生的冰灾事故造成了电网大面积的倒塔、杆塔损坏、断线、绝缘子脱落等故障,导致大面积、长时间停电,严重影响正常生产和生活,带来了沉重的电网修复负担。为了能够提早预测冰灾类天气、及时部署预防措施,对气温和风速等天气信息进行及时准确的预测非常必要。

传统的天气预测主要是通过卫星云图、统计学或动力-统计方法来进行的。近年来随着人工智能技术的飞速发展,出现了许多与其相关技术进行天气预测的研究,在智能化程度和精准度等方面有了一定提升。如深度网络中的循环神经网络(recurrent neural network, RNN)、长短时记忆网络(long short-term memory network, LSTM)^[3],作为时序预测研究方法和手段得到了广泛关注,在天气预测的研究中也取得了良好的效果。

目前的研究已经取得了很多进展,但是由于天气预测问题本身的复杂性和困难性,直今仍旧是一个没有完全解决的问题。为了进一步提高天气预测的精度,本文基于深度学习中的时域卷积网络(temporal convolutional network, TCN)对天气预测问题进行了研究,发现 TCN 在很多时序预测任务中有着比 RNN 更好的表现^[4]。进一步地,本文提出一种新的基于 TCN 的多尺度双线性天气预测模型,可利用历史天气观测数据进行学习,进而输出对未来天气的预测值。为了验证该模型的

有效性,还在大型公开天气数据集上与多种常用天气预测模型进行了对比。实验结果表明,与现有常见天气预测模型相比,该模型能够在多项天气指标的预测精度上达到或超过已有的现有方法。达到更加准确预测冰风灾害的目的,以最大限度的减少冰风灾害对输电线路带来的各种损失,具有非常重要的现实意义。

从已知的历史气象信息推测未来时段的天气信息,是一种典型的时间序列分析-预测问题。深度学习方法中使用最为广泛的序列处理方法则是基于 RNN 及其改进模型,如 LSTM、门控循环单元(gated recurrent unit, GRU)^[5]及 HUANG 等^[6]将递归注意模块与 CNN 结合的新方法等。LSTM 和 GRU 都在序列预测任务中取得了良好的效果。然而这类模型存在时序上的计算连接、无法进行并行优化、需要消耗大量的计算资源的问题。

BAI 等^[4]提出使用 TCN 对时序任务进行建模,其好处在于 TCN 可以使用并行计算的方法运行,比起 RNN 更加高效,对于长输入序列,训练需要的内存更少;YUAN 等^[7]提出了一种基于双因素注意力机制的天气预测模型,并得到了比 LSTM 等传统时序预测方法更好的结果;LEA 等^[8]在 TCN 的基础上提出了一种编解码模型,在计算机视觉检测和分割问题上取得了比 RNN 更好的效果;ZENG 等^[9]通过实验证明了 TCN 应用于碱基检测实际任务中在准确性和速度方面具有巨大的潜力;同时,BAI 等^[4]在 11 个不同的问题上将 TCN 与 LSTM, RNN, GRU 等方法进行了比较,结论是 TCN 不仅速度更快,而且有更高的精度。但由于 TCN 是一维的卷积结构,因此特征表征能力受卷积感受野的限制,在长时数据上,预测精度的表现不尽人意。

针对上述问题,本文进一步结合天气数据的特征,设计了一种新的基于 TCN 的多尺度双线性

模型,通过加强网络的特征表达能力,提高了TCN的长时间预测能力,并在长时间预测的多项指标上达到了本文所知的最好结果。

1 基于TCN的多尺度双线性天气预测模型

1.1 天气预测模型的输入输出定义

1.1.1 输入

输入 Y 包括2部分:① Y_1 是过去一段时间的历史气象信息;② Y_2 是超级计算机的预测值。历史中的 T 个连续时刻 $(1,2,\dots,T)$,与其对应的天气指标观测值记为 $Y=(y^1,y^2,\dots,y^T)$,对应的超级计算机预测值记为 $M=(\tilde{m}^1,\tilde{m}^2,\dots,\tilde{m}^T)$,其中某个时刻 t 所对应的气象信息数据 y^t 包含了 k 个天气指标值,其可表示为 $y^t=(y_1^t,y_2^t,\dots,y_k^t)$ 。

1.1.2 输出

输出是未来一段时间的气象信息,包括未来的 P 个连续时刻对应的天气指标值,可表示为 $\hat{Y}=(\hat{y}^{T+1},\hat{y}^{T+2},\dots,\hat{y}^{T+P})$,其中每个时刻的数据 $\hat{y}^{T+t'}$ 同样包含 k 个天气指标值。

1.1.3 天气预测模型

定义了输入输出后,天气预测模型 F 可以表示为

$$F(Y_1,\dots,Y_T;M_1,\dots,M_{T+p})=\hat{Y}_{T+1},\dots,\hat{Y}_{T+p} \quad (1)$$

1.2 时域卷积网络

在基于神经网络模型的时序问题研究中,通常采用RNN结构,如经典的LSTM,GRU网络等,另外一种经典网络结构——卷积神经网络则通常被用于图像数据的深度学习。最新的研究结果表明,在某些任务上,如音频合成、机器翻译等,卷积网络结构可以达到甚至优于RNN的效果,用于时序问题的卷积网络叫做TCN。

典型的TCN网络的基本结构如图1所示。

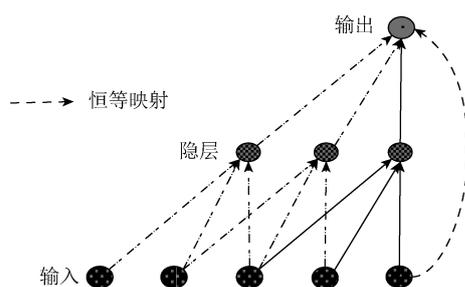


图1 时域卷积网络模型

Fig. 1 Temporal convolutional network

其中,TCN中包含了3种基本的结构:因果卷积^[10]、扩张卷积^[11]和恒等映射^[12]。

1.2.1 因果卷积

为了保证没有未来信息的泄露发生,TCN中严格按照时间先后顺序进行卷积操作,即时刻 t 的卷积操作仅发生在前一层的时刻 $t-1$ 和 $t-1$ 之前的数据上,不妨令卷积核为 $F=(f_1,f_2,\dots,f_k)$,其中 K 为卷积核的大小,输入序列为 $X=(x_1,x_2,\dots,x_T)$,则 x_T 处的因果卷积可以表示为

$$F(x_T)=\sum_{k=1}^K f_k x_{T-K+k} \quad (2)$$

1.2.2 扩张卷积

由于简单因果卷积网络的感受野与网络的深度呈线性关系,学习长时间时序依赖关系时网络会变得非常深,引入扩张卷积可以减少简单因果卷积的深度,引入扩张卷积后的因果卷积操作为

$$F_d(x_T)=\sum_{k=1}^K f_k x_{T-(K-k)d} \quad (3)$$

其中, d 为扩张卷积的系数。

为了能让网络学习到更长的时序依赖,需让天气预测模型能够利用更早的气象数据,可以通过扩大TCN中卷积的感受野来实现,也就是增加卷积核的大小 K 或者增大扩张系数增加卷积核的大小 d 。本文所提方法通过增大扩张系数来扩大卷积感受野,从而实现多尺度。扩张系数可选参数分别为1,2,4,6,8。

1.2.3 恒等映射

为了学习到长时间依赖,网络的深度变大,再加上扩张卷积,网络的稳定性就变得很重要。为了增加网络稳定性,还要加入恒等映射,加上恒等映射后网络的输出结果为

$$F(x_T)=F_d(x_T)+x_T \quad (4)$$

1.3 双线性池化

LIN等^[13]提出一种适用于二维图像卷积的双线性池化方法,其在一阶信息相同时,利用了二阶信息的不同进行分类的优化,在细粒度图像分类领域取得了当时最好的效果。LIN等^[14-15]在后续的实验证明,由于引入高阶统计信息的同时会带来计算复杂度成倍的提升、特征矩阵更加病态难以优化等问题,综合训练效率与测试性能,双线性的方法要优于三线性、四线性等。因此,为了合理地将多尺度TCN提取到的特征进行有效融合,同时引入高阶信息以增强方法性能,本文提出一种适用于一维卷积的双线性池化方法,令模

型中一路 TCN 卷积层输出为 $A_1=(a_1, \dots, a_n)$, 另一路输出为 $A_2=(a'_1, \dots, a'_n)$, 双线性层的融合输出 G' 为

$$G' = A_1^T A_2 = \begin{bmatrix} a_1^T \\ a_2^T \\ \dots \\ a_n^T \end{bmatrix} [a'_1, a'_2, \dots, a'_n]$$

$$= \begin{bmatrix} a_1^T a'_1 & a_1^T a'_2 & \dots & a_1^T a'_n \\ a_2^T a'_1 & a_2^T a'_2 & \dots & a_2^T a'_n \\ \dots & \dots & \dots & \dots \\ a_n^T a'_1 & a_n^T a'_2 & \dots & a_n^T a'_n \end{bmatrix} \quad (5)$$

双线性的输出层是格拉姆矩阵, 实际可以看作特征之间的偏心协方差矩阵, 每个元素由特定滤波器在特定位置卷积得到, 格拉姆计算实际上是计算两两特征之间的相关性, 而对角线元素度量各个维度自己的特性以及各维度之间的关系。取 G' 的对角线元素并按式(6)~(8)进行规范化, 得到双线性层的最终输出 G , 即

$$G = \text{diag}(G') \quad (6)$$

$$G = \text{sign}(G) \sqrt{|G|} \quad (7)$$

$$G = \frac{G}{\|G\|} \quad (8)$$

其中, $\text{diag}(\cdot)$ 为取矩阵的对角线元素作为输出向量。

1.4 基于时域卷积网络的多尺度双线性模型

本文提出一种基于 TCN 的多尺度双线性模型来进行天气预测, 该模型的框架如图 2 所示。

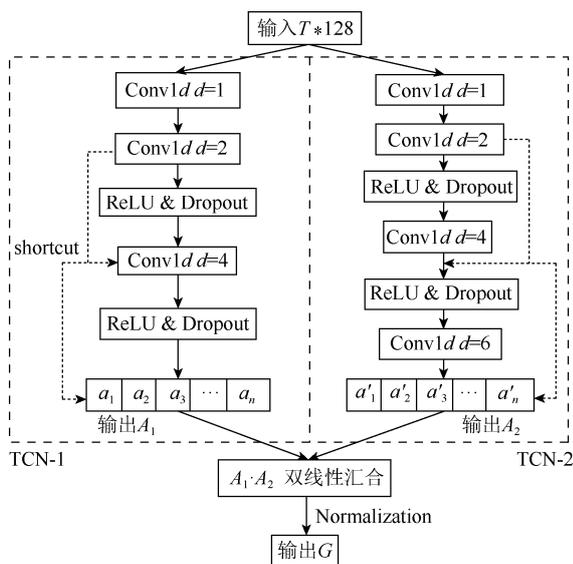


图 2 基于 TCN 的多尺度双线性天气预测模型
Fig. 2 TCN-based multi-scale bilinear weather prediction model

整个模型包含双路 TCN 网络, 每个 TCN 网络由扩张(因果)卷积层构成, 每个卷积层应用修正线性单元(rectified linear unit, ReLU)作为激活函数, 另外为了缓解过拟合现象, 加入了随机丢弃层(Dropout), 为了增强网络的稳定性, 还加入了恒等映射层, 2 个不同尺度网络(在右分支网络加入了 $d=6$ 的卷积层后, 两路输出可视为不同网络得到)的输出在双线性层进行双线性融合, 在规范化操作后, 最终得到网络的输出。除去扩张卷积带来的卷积核尺度变化之外, 其余网络参数均保持一致。

最后, 本模型训练过程中采用均方差函数(mean square error, MSE)作为损失函数, 其计算方法如下

$$MSE(Y, \hat{Y}) = \frac{\sum_{j=1}^q \sum_{i=1}^p (\hat{y}_i^j - y_i^j)^2}{p \cdot q} \quad (9)$$

其中, p 为天气预测的时间长度; q 为每个时刻所预测的天气指标的个数; \hat{y}_i^j 为预测值; y_i^j 为实际观测值。

2 实验结果与分析

2.1 数据集描述

本文的实验是在天气预报数据集 2018 上进行, 其是一个由全球 AI 挑战赛所提供的基准数据集^[16]。WFD2018 的数据采集是通过多站点的、多维度、长序列且高时间密度的气象观测站得到的。该数据集包含了北京 10 个气象站点, 共 3 年多的逐小时历史观测和气象预测数据组成, 由中国气象局北京城市气象研究所公开发布, 时间跨度长、密度高、缺失少。数据集由两部分组成: 一部分由气象观测站通过气象仪器实时监测获得的 9 个地面气象要素组成; 另一部分由中国北京气象局获得的 29 个气象预测要素组成, 这个 29 个气象预测要素是气象要素估计值。

本文所采用的训练数据包括了从为 2015 年 3 月 1 日至 2018 年 5 月 31 日的样本, 共 1 188 天。而 validation 集则为 2018 年 6 月 1 日至 2018 年 8 月 18 日的样本, 共 89 天; test 集 A 覆盖从 2018 年 8 月 29 日至 2018 年 9 月 24 日的样本, 共 27 天。

为了评估基于 TCN 的天气预测模型, 本文在实验中主要考察了该模型对多个气象站的 3 项天气指标的预测结果, 包括了地面上 2 m 高度处

温度(单位: °C)、地面上 2 m 高度处相对湿度(单位: %)和地面上 10 m 高度处风速(单位: m/s)。

2.2 评价标准

这里采用 3 个评价指标来衡量预测效果, 分别是均方根误差(root mean square error, RMSE)、平均绝对百分比误差(mean absolute percentage error, MAPE)和确定系数(R^2)。

其中, RMSE 与所评估数据一致的数值大小有关, 而 MAPE 和确定系数与数值大小无关, 只与重合度有关。特别是确定系数, 可衡量变量之间的线性相关性。RMSE 和 MAPE 越接近 0, 则意味着预测值与真实值越接近。当确定系数接近 1 时, 说明预测值和实际值之间有着很强的相关性。这 4 个评估指标的数学定义如下:

(1) 均方根误差

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (10)$$

其中, \hat{y}_i 为第 i 个样本的预测值; y_i 为第 i 个样本的真实值, 也为观测值。

(2) 平均绝对百分比误差

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|\hat{y}_i - y_i|}{|y_i|} \quad (11)$$

(3) 确定系数 R^2

$$R^2 = 1 - \frac{(\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (12)$$

2.3 对比方法

WRF (the weather research and forecasting model^[17]): 是一种传统的气象预测模式, 不仅可以用于真实天气的个案模拟, 也可以用其包含的模块组作为基本物理过程探讨的理论根据。

GBR(gradient boosting regressor^[18]): 是一种嵌入式学习方法, 适用于回归预测任务, 有着很强的泛化性能。在本文的对比试验中, GBR 模型采用了 100 个深度为 5 的基础估算子。

LED(LSTM encode decode^[19]): 基于 LSTM 编码-解码模型的方法, 适用于序列预测人物。将输入转换为一个固定长度的向量, 而解码过程将该向量解码为一个输出序列。

LEDA (LSTM encode decode with attention^[20]): 加入了时序注意力的 LSTM 编码-解码模型。注意力机制可以克服将输入序列转化为固定长度向量的问题。

2.4 实验结果分析

为了对本文所提出的基于 TCN 的多尺度双线性天气预测模型的实际效果进行评估, 在上文提到的 WFD 数据集上进行了天气预测实验, 实验采用一次处理数据长度为 128, 预测长度为 24, 得到实验结果见表 1~3。

表 1 地面上 2 m 相对湿度预测结果对比

Table 1 Comparison of prediction results of 2 m relative humidity on the ground

方法	RMSE	MAPE	R^2
WRF	18.39	22.08	0.33
GBR	33.03	64.58	-1.18
LED	17.52	32.72	0.34
LEDA	16.37	32.57	0.46
TCN	13.98	24.65	0.61
Proposed	13.47	24.64	0.61

表 2 地面上 2 m 温度预测结果对比

Table 2 Comparison of prediction results of the temperature of 2 m above ground

方法	RMSE	MAPE	R^2
WRF	2.63	14.29	0.66
GBR	7.06	39.56	-1.09
LED	6.16	36.84	-0.80
LEDA	5.78	34.58	-0.45
TCN	2.59	12.62	0.69
Proposed	2.60	13.01	0.70

表 3 地面上 10 m 高度风速预测结果对比

Table 3 Comparison of prediction results of wind speed at 10 m height on the ground

方法	RMSE	MAPE	R^2
WRF	1.63	98.58	-1.13
GBR	1.62	84.39	-0.58
LED	1.27	72.26	0.03
LEDA	1.20	63.99	0.13
TCN	1.42	80.43	-0.21
Proposed	1.12	85.10	0.14

从表 1~3 的数据可看出, 本文提出方法在 3 类 9 项指标中, 5 项指标优于 TCN, 1 项与 TCN 持平, 3 项略逊于 TCN, 并在 4 项预测指标上达到了本文所知的最好效果, 因此, 本文所提出的多尺度、双线性方法在 WFD 数据集上经过验证是有一定效果的。为了证明本文方法在长时数据上可以改善 TCN 捕获时序能力较差的问题, 在 WFD 数据集上, 分别在预测用历史数据长度参数为 32, 64, 128, 其预测长度分别为 6, 12, 24 的情况下, 与 TCN 进行对比试验, 实验结果分别如图 3~5 所示。图中显示, 本文方法在面对长时数据时, 相较于 TCN 更加的稳定, 部分预测指标更加精确。

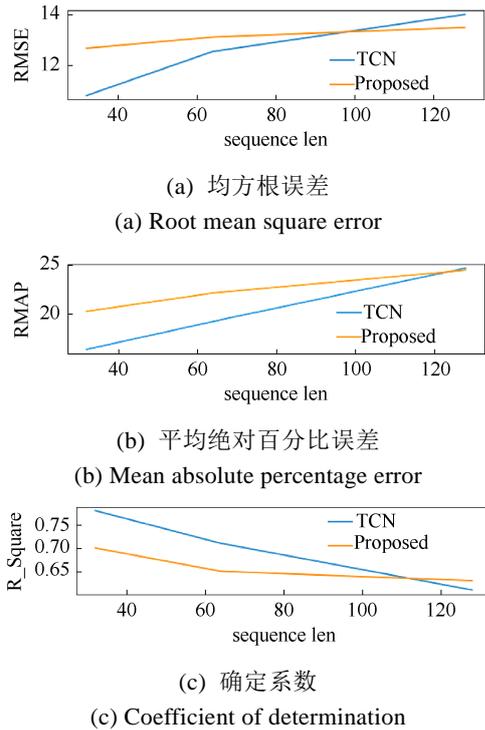


图 3 地面上 2 m 高度处湿度对比

Fig. 3 Comparison of humidity at a height of 2 m above the ground

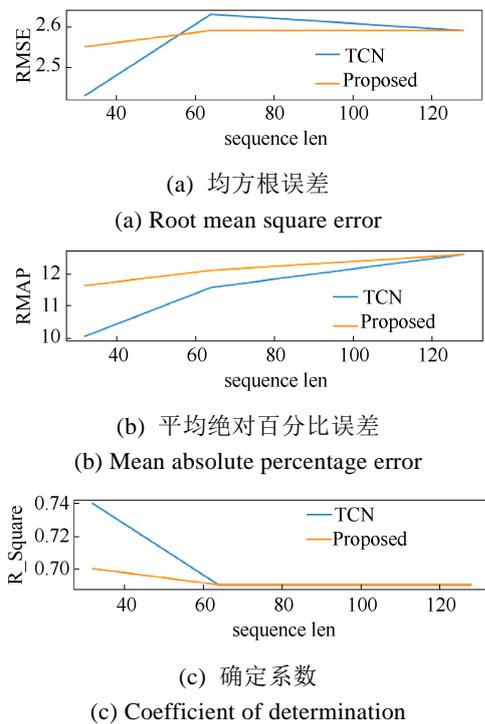


图 4 地面上 2 m 高度处温度对比

Fig. 4 Temperature comparison at a height of 2 m above the ground

为了评估网络参数 dilate 对实验效果的影响, 以地面上 2 m 温度(°C)RMSE 为评价指标进行对比试验, 一次处理数据长度为 128, 预测长度为 24, 其中 TCN-1 参数保持不变, 逐渐增加 TCN-2 中的

不同尺度 d 的卷积层。实验结果如图 6 所示, 综合 3 个指标, 本方法在 $d=6$ 时表现最佳。

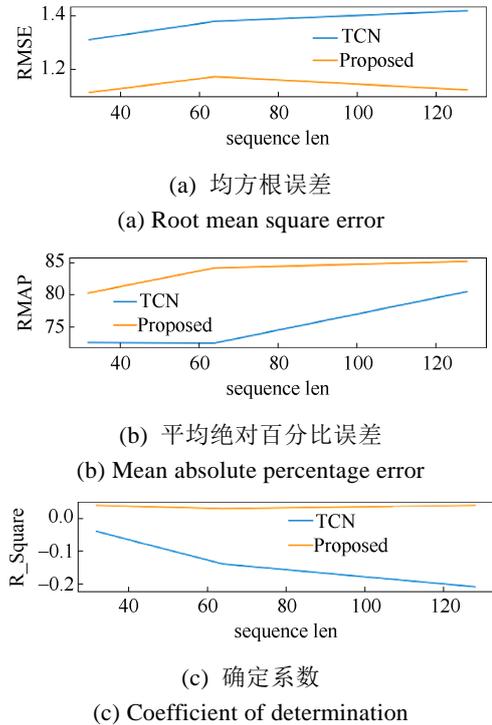


图 5 地面上 10 m 高度处风速对比

Fig. 5 Comparison of wind speed at a height of 10 m above the ground

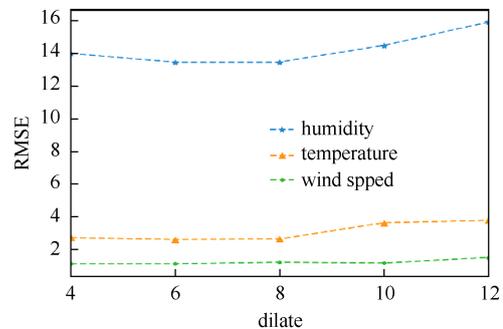


图 6 不同尺度对模型性能影响

Fig. 6 The impact of different scales on model performance

3 结束语

针对 TCN 面对长时数据预测效果不尽人意的问題, 本文提出了一种基于 TCN 的多尺度双线性的天气预测方法。多尺度卷积可以提升模型卷积的感受野, 在处理像天气这种季节性数据时, 可以增强模型对局部信息特征变化的处理能力, 学习到更好的时序特征。双线性池化操作可以在一阶统计信息相同的情况下利用模型的二阶统计信息, 在多尺度特征中融合各个尺度所特有的维度关系, 从而使多尺度的方法更容易训练和拟合。经过实验验证, 本文提出的天气预测模型在 WFD

数据集上有一定的优势。

参考文献

References

- [1] CHEN W D, CUI F, ZHOU H, et al. Impacts of different radiation schemes on the prediction of solar radiation and photovoltaic power[J]. *Atmospheric and Oceanic Science Letters*, 2017, 10(6): 446-451.
- [2] 唐西胜, 孙玉树, 齐智平. 基于 HHT 的风电功率波动及其对电力系统低频振荡的影响分析[J]. *电网技术*, 2015, 39(8): 2115-2121.
TANG X S, SUN Y S, QI Z P. Analysis of HHT-based wind power fluctuation and its impact on low-frequency oscillation of power system[J]. *Power System Technology*, 2015, 39(8): 2115-2121 (in Chinese).
- [3] ZIA T, ZAHID U. Long short-term memory recurrent neural network architectures for Urdu acoustic modeling[J]. *International Journal of Speech Technology*, 2019, 22(1): 21-30.
- [4] BAI S, KOLTER J, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[EB/OL]. [2020-02-01]. https://xueshu.baidu.com/usercenter/paper/show?paperid=f5d97f021345c683fa26222a20136b0c&site=xueshu_se&hitarticle=1.
- [5] CHO K, VAN MERRIENBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]// *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Stroudsburg: Association for Computational Linguistics, 2014: 1724-1734.
- [6] HUANG J B, ZHUO E H, LI H J, et al. Achieving accurate segmentation of nasopharyngeal carcinoma in MR images through recurrent attention[J]. *Medical Image Computing and Computer Assisted Intervention - MICCAI 2019*, 2019: 494-502.
- [7] YUAN M L, JI X Z, LU T, et al. A novel two-factor attention encoder-decoder network through combining temporal and prior knowledge for weather forecasting[C]// *2019 International Joint Conference on Neural Networks (IJCNN)*. New York: IEEE Press, 2019: 1-8.
- [8] LEA C, VIDAL R, REITER A, et al. Temporal convolutional networks: A unified approach to action segmentation[M]// *Lecture Notes in Computer Science*. Heidelberg: Springer, 2016: 47-54.
- [9] ZENG J W, CAI H M, PENG H, et al. Causalcall: nanopore basecalling using a temporal convolutional network[J]. *Frontiers in Genetics*, 2020, 10: 1332.
- [10] OORD A V D, DIELEMAN S, ZEN H, et al. WaveNet: a generative model for raw audio[EB/OL]. [2020-02-01]. https://xueshu.baidu.com/usercenter/paper/show?paperid=3400c5b533ad792d2b51b44dc91087f1&site=xueshu_se.
- [11] FISHER Y, VLADLEN K. Multi-scale context aggregation by dilated convolutions[EB/OL]. [2020-02-01]. https://xueshu.baidu.com/usercenter/paper/show?paperid=906431f7e4f3b6679f265fe39c229fd8&site=xueshu_se&hitarticle=1.
- [12] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]// *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. New York: IEEE Press, 2016: 770-778.
- [13] LIN T Y, ROYCHOWDHURY A, MAJI S. Bilinear CNN models for fine-grained visual recognition[C]// *2015 IEEE International Conference on Computer Vision (ICCV)*. New York: IEEE Press, 2015: 1449-1457.
- [14] LIN T Y, MAJI S. Improved bilinear pooling with CNNs[C]// *Proceedings of the British Machine Vision Conference 2017*. London: BMVA Press, 2017: 1-12.
- [15] LIN T Y, ROYCHOWDHURY A, MAJI S. Bilinear convolutional neural networks for fine-grained visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(6): 1309-1322.
- [16] Sinovation Ventures. Ai challenge 2018 weather forecasting data set [EB/OL]. [2020-01-03]. <https://challenger.ai/dataset/wfd2018>.
- [17] SKAMAROCK W C, KLEMP J B, DUDHIA J, et al. A description of the advanced research WRF version 3[M]. Colorado: National Center for Atmospheric Research, 2008:13-18.
- [18] YE J, CHOW J H, CHEN J, et al. Stochastic gradient boosted distributed decision trees[C]// *Proceeding of the 18th ACM Conference on Information and Knowledge Management - CIKM '09*. New York: ACM Press, 2009: 2061-2064.
- [19] SUTSKEVER I, VINYALS O, LE Q V. Sequence to sequence learning with neural networks[C]// *Proceedings of the 27th International Conference on Neural Information Processing Systems*. New York: ACM Press, 2014: 3104-3112.
- [20] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate[EB/OL]. [2020-02-01]. https://xueshu.baidu.com/usercenter/paper/show?paperid=3242f94dbfc892f63c9f51acd8ef8ce&site=xueshu_se&hitarticle=1.