

基于 LSTM-Attention 网络的短期风电功率预测

钱勇生¹, 邵洁², 季欣欣², 李晓瑞², 莫晨², 程其玉²

[1. 上海电器科学研究所(集团)有限公司, 上海 200063;

2. 上海电力大学电子与信息工程学院, 上海 200090]

摘要: 提出一种基于 LSTM-Attention 网络的短期风电功率预测方法。首先, 使用 LSTM 网络对数值天气预报(NWP)数据的特征信息进行提取, 同时采用注意力机制有效分析了模型输入与输出的相关性, 从而获取了更多重要时间的整体特征; 其次, 使用卷积神经网络(CNN)提取 NWP 数据的局部特征, 并引入压缩和奖惩网络(SE)模块学习特征权重, 利用特征重新标定方式提高网络表示能力; 最后, 将局部特征和整体特征进行特征融合, 通过分类器输出分类结果。利用 NOAA 提供的美国加利福尼亚州某风电场的数据进行案例分析, 证明了所提方法的有效性。试验结果表明, 与 BP 神经网络、自回归积分滑动平均模型(ARIMA)模型和 LSTM 模型相比, LSTM-Attention 模型具有更高的预测精度, 证明了该方法的有效性。

关键词: 风电功率预测; LSTM; 卷积神经网络; 压缩和奖惩网络模块; 注意力机制

中图分类号: TM 715 **文献标志码:** A **文章编号:** 1673-6540(2019)09-0095-06

Short-Term Wind Power Forecasting Based on LSTM-Attention Network

QIAN Yongsheng¹, SHAO Jie², JI Xinxin², LI Xiaorui², MO Chen², CHENG Qiyu²

[1. Shanghai Electrical Apparatus Research Institute (Group) Co., Ltd., Shanghai 200063, China;

2. College of Electronics and Information Engineering, Shanghai University of Electric Power, Shanghai 200090, China]

Abstract: A short-term wind power forecasting method based on long short-term memory-attention (LSTM-Attention) network was presented. Firstly, the LSTM network was used to extract the feature information of numerical weather prediction (NWP) data, and the attention mechanism was used to effectively analyze the correlation between input and output of the model, so as to obtain more global features of important moments. Secondly, the convolutional neural network (CNN) was used to extract the local features of NWP data, squeeze-excitation (SE) blocks were introduced to learn the feature weights, and the feature re-calibration method was used to improve the network representation ability. Finally, local and global features were fused, and the classification results were output by classifier. A case study of a wind farm in California, American provided by National Oceanic and Atmospheric Administration (NOAA) was conducted to demonstrate the effectiveness of the proposed method. The experimental results showed that LSTM-Attention model had higher prediction accuracy than BP neural network, autoregressive integrated moving average (ARIMA) model and LSTM model, which proved the validity of the proposed method.

Key words: wind power forecasting; long short-term memory; convolutional neural network; squeeze-excitation blocks; attention mechanism

作者简介: 钱勇生(1992—), 男, 硕士研究生, 研究方向为深度学习、新能源等。

邵洁(1981—), 女, 博士, 副教授, 研究方向为计算机视觉、图像处理、机器学习。

季欣欣(1994—), 男, 硕士研究生, 研究方向为计算机视觉、深度学习。

0 引言

随着化石燃料开采枯竭和减少温室气体排放的必要性,可再生能源得到了广泛的开发和利用。风力发电是世界上增长最快的能源之一。据报告统计,2017 年底全球累计风电装机容量达到 539.58 GW^[1]。然而,风能具有随机性、波动性和间歇性等特点。风电的大规模并网将影响电力系统的安全性、稳定性以及电能质量^[2-4]。因此,准确的风力预测对风电在电力系统中的实时调整起着重要作用。风速预测可以调整电力分配、优化控制策略和降低运行成本^[5-7]。风电功率预测按照不同时间范围,可以分为 4 个时间段:超短期(几秒到几分钟)、短期(15 min 到 24 h)、中期(24 h 到 1 周)和长期(1 周到数年)^[8]。

风电预测主要基于风速预测,已有的风速预测技术主要分为物理方法、统计方法和混合方法^[9-11]。物理模型是基于数值天气预报(NWP)数据。依据大气层空气运动的情况,对风电场进行精确的物理描述,但需要大量的计算资源^[12]。传统统计方法通常利用历史序列(风电功率、历史风速、风向、气压、温度等)找到测量数据(风电功率或风速)的关系,例如时间序列分析法^[13]、人工神经网络(ANN)^[14]、支持向量机法(SVM)^[15]和卡尔曼滤波法^[16]等。混合方法是结合不同模型的有用信息,呈现出优越的性能。如混合神经网络和模糊逻辑^[17]、ANN 和遗传算法(GA)的组合^[18]、混合小波分解和 ANN^[19]以及多算法混合^[20]。在上述方法中,具有优越的非线性拟合和泛化能力的神经网络使用最广泛。但是,传统神经网络存在一些缺点:(1)输入变量和训练样本的严格要求,训练样本过多或过少均会影响训练效果;(2)当特征尺寸过大,难以有效提取高质量特征时,神经网络难以获得良好的效果。深度学习是 ANN 的发展,可以很好地解决以上问题。通过学习海量数据和强大的学习和泛化能力,能够确定哪些特征是必不可少的,哪些特征在没有人力的情况下可以省略,进而提高分类和预测等任务的准确性^[21]。

本文提出基于 LSTM-Attention 网络的短期风电功率预测方法。该网络由卷积神经网络、长期短期记忆(LSTM)及压缩和奖惩网络(SE)模块组

成。首先,通过对模型输入包括风速数据(NWP、本地测量数据或附近观测数据)进行数据标准化预处理;其次,利用历史数据之间的统计关系,对历史数据的训练建立 LSTM-Attention 网络预测模型,使用注意力机制的 LSTM 神经网络提取时间序列数据的整体特征以及使用嵌入 SE 块的卷积神经网络(CNN)提取局部特征,并将局部特征和整体特征进行融合;最后,利用风速和风力发电功率的关系模型得到风电功率数据。试验结果表明,LSTM-Attention 网络预测模型的预测精度比传统预测方法更高。

1 风电场输出功率的影响因素

风力涡轮机的输出功率取决于风速。风速在一段时间范围内变化,并取决于区域天气状况、季节变化和环境因素^[22]。穿过风力涡轮机的转子叶片总扫掠面积(A)的风机功率可以定义为

$$P = \frac{1}{2} C_p \rho(t) A v^3 \quad (1)$$

式中: C_p ——风机的功率系数,由叶尖角度、叶片设计以及风速和转子速度之间的关系决定;

$\rho(t)$ ——随时间变化的空气密度,取决于周围温度、湿度和大气压力;

v ——风速。

通常,测量的风速数据是接近地球表面的速度,因此修改风速以获得风力机轴高度的速度:

$$v(t) = v_H(t) \left(\frac{H_1}{H_0} \right)^\alpha \quad (2)$$

式中: $v_H(t)$ ——风塔测量的风速;

H_1, H_0 ——风轮机轴和风塔的高度;

α ——风切变指数,一般为 1/7。

2 注意力机制的 LSTM 网络模型及相关结构

2.1 注意力机制的 LSTM 网络

循环神经网络(RNN)是一种用于处理时间序列问题并获得良好性能的神经网络。更具体地说,当处理后续输入数据时,先前输入所携带的信息可以保持在网络内。RNN 在语音识别、自然语言和机器翻译等任务中得到广泛应用。因固有的

时间依赖性, RNN 适用于风电预测任务。但因具有在反向传播期间消失梯度问题的缺点^[23], 在长时间序列预测中, RNN 仍将陷入无法掌握长期规律和捕捉长期特征的两难境地, 从而限制了 RNN 处理相当长的时间序列信息的能力。LSTM 作为 RNN 的扩展, 能够通过用特殊设计的模块替换简单的激活节点来学习长期依赖性。每个 LSTM 单元由 4 个交互单元组成, 包括内部单元、输入门、遗忘门和输出门^[24]。LSTM 结构如图 1 所示。

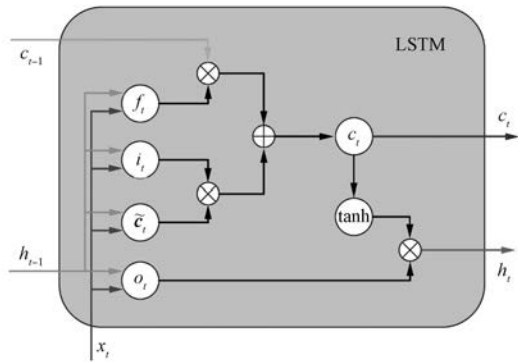


图 1 LSTM 结构

内部单元通过自我重复连接记忆前一时间步的单元状态。输入门控制输入激活流入内部单元状态。输出门控制输出激活流入 LSTM 单元输出。遗忘门缩放内部单元状态, 使 LSTM 单元能够自适应地忘记或重置单元的存储器。因此, LSTM 自然能够学习输入序列的长期和短期依赖性, 使 LSTM 适用于考虑长期背景的风电预测任务。假设时间步长为 \$t(t=1, \dots, n, n\$ 为时间步长的总数), 则 LSTM 单元的输入状态为 \$x_t\$, 遗忘门为 \$f_t\$, 输入门为 \$i_t\$, 输出门为 \$o_t\$, 单元状态记忆为 \$c_t\$, 以及隐藏状态输出为 \$h_t\$。在上一个时间步长 \$t-1\$, 将细胞状态记忆表示为 \$c_{t-1}\$, 并将隐藏状态输出为 \$h_{t-1}\$。变量之间的关系可以描述为

$$f_t = \text{sigmoid}(W_{f_n} h_{t-1} + W_{f_x} x_t + b_f) \quad (3)$$

$$i_t = \text{sigmoid}(W_{i_n} h_{t-1} + W_{i_x} x_t + b_i) \quad (4)$$

$$\tilde{c}_t = \tanh(W_{c_n} h_{t-1} + W_{c_x} x_t + b_c) \quad (5)$$

$$o_t = \text{sigmoid}(W_{o_n} h_{t-1} + W_{o_x} x_t + b_o) \quad (6)$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \tilde{c}_t \quad (7)$$

$$h_t = o_t \otimes \tanh(c_t) \quad (8)$$

式中: \$W, b\$——训练网络参数;

\$\tilde{c}_t\$——tanh 层创建的中间候选值的矢量。

人类的视觉系统始终关注不同数据的相关局部信息, 注意机制起着非常重要的作用。为了使模型能够更多地关注时间序列中的重要信息, 本文提出一种注意力机制的 LSTM 网络结构, 如图 2 所示。

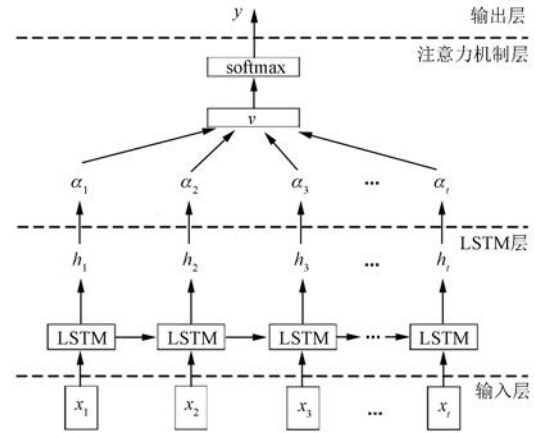


图 2 注意力机制的 LSTM 结构

图 2 中, 输入序列为一段时间的气象数据表示 \$x_1, x_2, \dots, x_t\$, 将每个输入传入 LSTM 单元, 得到对应隐藏层的输出 \$h_1, h_2, \dots, h_t\$。在隐藏层引入注意力机制, 计算各个输入分配的注意力概率分布值 \$\alpha_1, \alpha_2, \dots, \alpha_t\$, 注意力机制将会产生注意力权重矩阵 \$\alpha\$ 和特征表示 \$v\$。具体描述为

$$u_i = \tanh(W_s h_i + b_s) \quad (9)$$

$$\alpha_i = \frac{\exp(u_i^T u_s)}{\sum_{i=1}^t \exp(u_i^T u_s)}, \sum_{i=1}^t \alpha_i = 1 \quad (10)$$

$$v = \sum_{i=1}^t \alpha_i h_i \quad (11)$$

式中: \$W_s\$——权值矩阵;

\$b\$——偏置量;

\$u_s\$——随机初始化的时间序列注意力机制矩阵。

最后, 利用 softmax 函数得到预测类别为 \$y\$, 计算式如下:

$$y = \text{softmax}(W_v v + b_v) \quad (12)$$

2.2 SE_CNN 网络

SE_CNN 网络由 CNN 和 SE 模块级联而成。CNN 是一种广泛用于计算机视觉任务的前馈型神经网络。多层 CNN 已被应用于解决诸如图像分类、视频识别和自然语言处理等问题。通过利

用多个滤波器对输入宽度和高度进行卷积,从而在一个卷积层产生滤波器的激活图。可堆叠若干卷积层以产生输入的更高级抽象。因此,卷积层可以用作模型重要构建块,将包含风速和风力时间序列的原始输入数据转换为可以由更高层使用的特征图。由于直接从数据中学习滤波器的参数,在特征提取阶段期间几乎不需要先验知识。SE 设法计算输出卷积通道的权重,通过强调重要特征和抑制通道之间的无用特征使网络更有效^[25]。SE 模块示意图如图 3 所示。

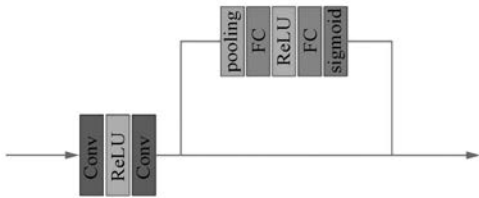


图 3 SE 模块示意图

首先使用全局平均池来生成通道向量,将全局空间信息压缩成一个通道描述符,这个描述符嵌入了通道特征响应的全局分布,使后面的网络层获得全局感受野的信息。压缩功能如式(13)所示:

$$z_c = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H u_c(i, j) \tag{13}$$

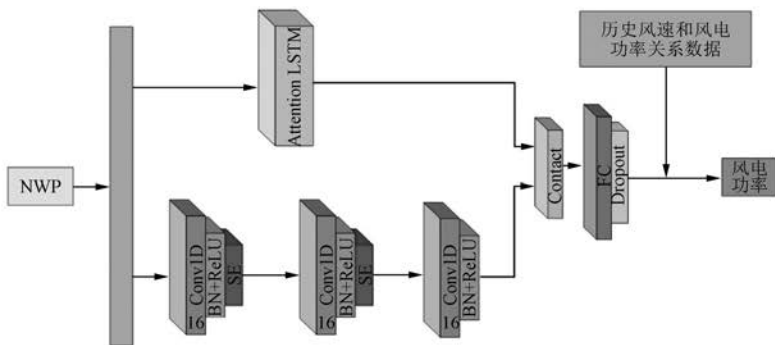


图 4 基于 LSTM-Attention 网络的风电功率预测模型

建立的网络模型共有 5 个超参数:注意力机制的 LSTM 输入层维数、隐藏层个数、SE_CNN 的 Conv1D 层卷积核数量、大小以及输出层维数。在多次训练网络后,最佳参数设置如下:输入层维数为 5,即输入变量个数,隐藏层个数设为 8,3 个 Conv1D 层均设为 16 组卷积核,卷积核大小分别设为 8×8、5×5 和 3×3,其后紧跟一个 BN (Batch

式中: H 、 W ——图像的高度和宽度;
 c ——压缩通道的第 c 个元素。
 对于输入特征映射大小 $H \times W \times C$,输出大小将为 $1 \times 1 \times C$ 。

为了充分利用前一阶段的通道聚合信息,获取各通道信息的依赖关系,通过基于通道依赖性的筛选机制为每个通道学习特定采样的激活,控制每个通道的激励。奖惩功能如下:

$$s = \sigma [W_2 \delta (W_1 z)] \tag{14}$$

式中: σ ——sigmoid 激活函数;
 W_1 、 W_2 ——2 个全连接层 (FC);
 δ ——ReLU 函数。

3 基于 LSTM-Attention 网络的风电功率预测模型

基于 LSTM-Attention 网络的风电功率预测模型如图 4 所示。预测模型分为 2 个部分:注意力机制的 LSTM 网络和 SE_CNN 网络。首先,在对原始 NWP 数据进行标准化处理后作为预测模型的输入;其次,利用注意力机制的 LSTM 网络和 SE_CNN 网络分别提取时间序列 NWP 数据的整体特征和局部特征后进行特征融合,并利用分类器预测风速;最后,利用历史风速和风电功率的关系得到风电功率。

normalization) 函数和一个 ReLU 激活函数,其中,前 2 个 Conv1D 层后又嵌入 SE 模块来提取有效特征。该预测任务是根据历史数据预测下一步的风速,因此输出维数设为 1。为了避免网络训练时的过度拟合问题,并确保网络的泛化能力足够好,FC 后增加设定值为 0.2 的 Dropout 层。

4 试验验证

4.1 预测结果误差评估标准

使用平均绝对误差 MAE 和均方根误差 RMSE 并研究预测精度。定义如下：

$$E_{MAE} = \frac{1}{N} \sum_{i=1}^N |Y_i - \hat{Y}_i| \tag{15}$$

$$E_{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - \hat{Y}_i)^2} \tag{16}$$

式中： N ——预测的数量；
 i ——预测点序列编号；
 Y_i ——当时的标准化实际观察值；
 $\hat{Y}_i$ ——同一时期的预测标准化值。

4.2 预测模型的结果

使用 NOAA 提供的美国加利福尼亚州某风电场的实测数据进行案例研究,以验证本文所提方法。获取的数据包含风力测量数据和风电场的 NWP 数据(100 m 的空气密度、压力、温度、风速和风向)。本文选取 2011 年 3 月 22 日至 2011 年 3 月 31 日的数据,采样时间间隔是 5 min,共计 2 880 个采样点,以 90% 数据作为训练集,10% 数据作为测试集。

基于 Python 平台的 Keras 深度学习框架构建 LSTM-Attention 预测模型。使用 Adam 算法确定权重参数的迭代更新方法以选择优化程序,计算梯度的一阶矩估计和二阶矩估计而为不同的参数设计独立的自适应性学习率,迭代次数为 100。风电功率预测结果如图 5 所示。表 1 所示为 LSTM 和 LSTM-Attention 的预测误差。

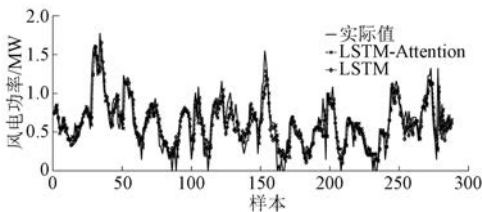


图 5 风电功率预测结果

表 1 LSTM 和 LSTM-Attention 的预测误差 %

方法	MAE	RMSE
LSTM-Attention	0.858	1.148
LSTM	0.863	1.154

可以看出,LSTM 和 LSTM-Attention 方法的预测结果均接近实际风电功率曲线,此外 LSTM-Attention 的预测精度高于基于原始 NWP 数据的 LSTM 模型。LSTM-Attention 模型还可以降低网络的复杂性并增强模型的泛化能力。

4.3 与不同模型的比较

将 LSTM-Attention 模型与 BP 神经网络和自回归积分滑动平均模型(ARIMA)模型进行比较。不同模型的预测结果如图 6 所示。表 2 所示为不同预测模型的误差结果。

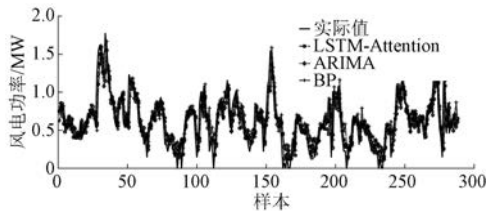


图 6 不同模型的预测结果

表 2 不同预测模型的误差结果 %

方法	MAE	RMSE
LSTM-Attention	0.858	1.148
ARIMA	0.943	1.249
BP	1.048	1.387

从图 6 和表 2 可以看到,对于相同的测试集中,LSTM-Attention 模型是接近实际风电功率曲线。与 BP 神经网络和 ARIMA 模型相比,LSTM-Attention 模型的误差分别减少了约 0.20% 和 0.09%。

5 结 语

对于大规模风电接入的电网,准确的风电预测模型在运行和控制中具有重要的意义。与传统学习方法相比,深度学习方法具有数据学习和泛化能力的优势。本文提出了基于 LSTM-Attention 网络的短期风电功率预测方法。为了获得更好地提取 NWP 的高阶特征,利用注意力机制的 LSTM 网络和 SE_CNN 网络分别提取整体特征和局部特征并进行融合,其中,注意力机制有效分析了模型输入与输出的相关性,从而获取了更多重要时间的整体特征。引入 SE 模块学习特征权重,利用特征重新标定方式提高网络表示能力。与 BP 神经网络和 ARIMA 模型相比,LSTM-Attention 的预

测精度更高,且具有泛化和高维函数逼近能力的优点。这验证了 LSTM-Attention 模型在风力预测领域的先进性和实用性。

【参考文献】

- [1] Global statistics [DB/OL]. [2018-04-06]. [http: // gwec.net/global-figures/graphs/](http://gwec.net/global-figures/graphs/).
- [2] 刘永前,韩爽,胡永生.风电场出力短期预报研究综述[J].现代电力,2007,24(5): 6.
- [3] 牛东晓,范磊磊.风电功率预测方法综述及发展研究[J].现代电力,2013,30(4): 24.
- [4] ZHU L, SHI H, AND DING M. A chaotic BP neural network used to wind power prediction [C] // 2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 2018.
- [5] 刘斯伟.并网双馈风电机组对电力系统暂态稳定性的影响机理研究 [D]. 北京: 华北电力大学, 2016.
- [6] 迟永宁,王伟胜,戴慧珠.改善基于双馈感应发电机的并网风电场暂态电压稳定性研究[J].中国电机工程学报,2007,27(25): 25.
- [7] 赵洪山,刘兴杰,李聪.基于机组动态风速信息的风电场有功控制策略[J].电力科学与工程,2013,29(1): 9.
- [8] 陶玉波,陈昊,秦晓辉,等.短期风电功率预测概念和模型与方法[J].电力工程技术,2018,37(5): 7.
- [9] 薛禹胜,郁琛,赵俊华,等.关于短期及超短期风电功率预测的评述[J].电力系统自动化,2015,39(6): 141.
- [10] 曾杰,张华.基于最小二乘支持向量机的风速预测模型[J].电网技术,2009,33(18): 144.
- [11] 谷兴凯,范高锋,王晓蓉,等.风电功率预测技术综述[J].电网技术,2007,31(增刊2): 335.
- [12] FOLEY A M, LEAHY P G, MARVUGLIA A, et al. Current methods and advances in forecasting of wind power generation [J]. Renewable Energy, 2012, 37(1): 1.
- [13] 岳莉莉.基于时间序列分析的风速短期预测方法研究[D].北京:华北电力大学,2012.
- [14] 范高锋,王伟胜,刘纯,等.基于人工神经网络的风电功率预测[J].中国电机工程学报,2008,28(34): 118.
- [15] 王晓兰,王明伟.基于小波分解和最小二乘支持向量机的短期风速预测[J].电网技术,2010,34(1): 179.
- [16] 修春波,任晓,李艳晴,等.基于卡尔曼滤波的风速序列短期预测方法[J].电工技术学报,2014,29(2): 253.
- [17] NEGNEVITSKY M, JOHNSON P, SANTOSO S. Short term wind power forecasting using hybrid intelligent systems [C] // 2007 IEEE Power Engineering Society General Meeting, 2007.
- [18] 韦姝.基于混合模型的风电场短期风速风功率预测研究[D].南京:东南大学,2015.
- [19] DOUCOURE B, AGBOSSOU K, CARDENAS A. Time series prediction using artificial wavelet neural network and multi-resolution analysis: Application to wind speed data[J].Renew Energy,2016,92: 202.
- [20] LIU H, CHEN C, TIAN H, et al. A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks [J]. Renewable Energy,2012,48: 545.
- [21] LECUN Y, BENGIO Y, HINTON G. Deep learning [J].Nature,2015,521: 436.
- [22] 马祎炜,俞俊杰,吴国祥,等.双馈风力发电系统最大功率点跟踪控制策略[J].电工技术学报,2009,24(4): 202.
- [23] BENGIO Y, SIMARD P, FRASCONI P. Learning long-term dependencies with gradient descent is difficult[J].IEEE Transactions on Neural Networks, 1994,5(2): 157.
- [24] GERS F, SCHRAUDOLPH N N, SCHMIDHUBER J. Learning precise timing with LSTM recurrent networks [J].Journal of Machine Learning Research, 2002, 3(1): 115.
- [25] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J].IEEE Transactions on Pattern Analysis and Machine Intelligence. doi:10.1109/TPAMI.2019.2913372.

收稿日期: 2019-06-17