

研究论文

DOI: 10.11949/j.issn.0438-1157.20171534

基于 LSTM-RNN 模型的铁水硅含量预测

李泽龙, 杨春节, 刘文辉, 周恒, 李宇轩

(浙江大学控制科学与工程学院, 浙江 杭州 310027)

摘要: 针对高炉炼铁是一个动态过程, 具有大延迟, 工况复杂的特性。采用 LSTM-RNN 模型进行硅含量预测, 充分发挥了其处理时间序列时挖掘前后关联信息的优势。首先根据时间序列趋势及相关系数选择自变量, 并采用复杂工况的实际生产数据进行验证。然后用程序自动求解最优参数进行硅含量预测。最后将 LSTM-RNN 模型与 PLS 模型及 RNN 模型的结果进行对比, 验证该方法的优点。研究发现 LSTM-RNN 模型预测误差稳定, 预测精度较高, 比传统的统计学及神经网络方法取得了更好的预测精度。

关键词: 预测; 动态建模; 神经网络; 高炉炼铁; 硅含量

中图分类号: TQ 028.8

文献标志码: A

文章编号: 0438—1157 (2018) 03—0992—06

Research on hot metal Si-content prediction based on LSTM-RNN

LI Zelong, YANG Chunjie, LIU Wenhui, ZHOU Heng, LI Yuxuan

(College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, Zhejiang, China)

Abstract: The ironmaking in blast furnace, with large delay and complex conditions, is a dynamic process. The traditional methods for prediction of silicon content in hot metal are mostly based on the statistics or the simple neural networks, leading to lower accuracy. However, a model based on the long short-term memory-recurrent neural network (LSTM-RNN) is proposed to exploit the characteristics of the mutual information before and after the time series in this paper. The independent variables are selected according to the time series trend and the correlation coefficient. After that, the silicon content is predicted according to the input variables by optimizing the parameters automatically. In order to verify the constructed model, the extremely complex production data is used to compare the LSTM-RNN and simple RNN models. Remarkably, the result shows that the prediction error of LSTM-RNN model is stable and the prediction accuracy is high.

Key words: prediction; dynamic modelling; neural network; ironmaking; silicon content

引言

钢铁产业是工业生产的重中之重, 而高炉炼铁是炼钢的前提, 所以提高高炉炼铁效率具有重大的经济效益。硅含量是炼铁过程重要的参考指标。在稳定工况下, 当硅含量低于 0.4% 时, 炉温一般低于 1500℃, 无法正常炼铁; 硅含量高于 0.6% 时, 温度

过高, 焦炭的消耗过高。为了炼铁过程的稳定进行, 一般需要将硅含量控制在 0.4%~0.6% 之间。如果可以提前预知硅含量, 操作工将可以提前进行必要的操作来控制高炉平稳运行^[1]。现有的硅含量检测方法是出炉铁水送至化验室化验获得。这样得到的结果有严重的滞后, 对生产的参考性大大降低。只有及时准确地预测出高炉硅含量, 才能真正发挥这

2017-09-25 收到初稿, 2017-11-13 收到修改稿。

联系人: 杨春节。第一作者: 李泽龙 (1993—), 男, 硕士研究生。

基金项目: 国家自然科学基金项目 (61290321)。

Received date: 2017-09-25.

Corresponding author: Prof. YANG Chunjie, cjyang@ipc.zju.edu.cn

Foundation item: supported by the National Nature Science Foundation of China (61290321).

一项指标的价值。因此,对高炉铁水硅含量的及时预测具有非常重要的实际意义。

目前对于硅含量预测的模型主要有两大类:机理模型和数据驱动模型。机理模型主要是利用物理学中的热平衡特性和化学中的物料守恒原则进行分析,该类模型根据高炉冶炼中产生的化学反应,对炉身不同层级的物料反应进行建模^[2-5],然而由于高炉内部的反应极其复杂,而且模型中需要用到的许多参数无法测量到,所以预测效果较差。数据驱动模型^[6]主要使用了统计学方法及机器学习方法。所用的方法主要有:回归分析、偏最小二乘法(partial least squares regression, PLS)、主成分分析、自回归分析^[7]、神经网络方法^[8]、支持向量机^[9]。其中,回归分析、偏最小二乘法、主成分分析、自回归分析是采用统计学的方法,他们各自采取不同的统计学指标作为优化目标,效果也各不相同。神经网络模型与支持向量机则属于机器学习方法,他们以预测精度作为更加直观的优化目标,经过大量的训练学习,可以取得更好的效果。

计算机计算能力的增强为深度学习的应用创造了条件。神经网络由于处理非线性问题能力强,容错学习能力好,非常适合于硅含量的预测。曾燕飞等^[10]将BP网络用于硅含量的预测,但预测结果非常粗糙,偏差很大。宋菁华等^[11]将改进型的EMD-Elman神经网络用于硅含量预测,得到了较好的预测结果;蒋朝晖等^[12]采用Bootstrap方法用于硅含量预测,尝试给出预测结果的可信度,但是预测效果并不理想;宋贺达等^[13]采用控制理论中传统的非线性子空间模型进行硅含量预测,虽然降低了计算复杂度,但是效果也不理想。Zhou等^[14]采用模糊神经网络取得了更好的预测效果。但是,以上文章采用的是较为平稳工况的运行数据,如果工况波动比较大,则预测效果不好。针对复杂冶炼工况,本文相较前人的方法有了较大改进,可以针对复杂工况取得比较好的结果。

由于高炉炼铁是一个动态的时间序列,所以使用递归神经网络(recurrent neural network, RNN)更加合适。除此之外,高炉反应的过程是渐变的,也就是说当前的炉况与历史炉况相互关联。这就要求递归神经网络能够动态记忆历史信息,在学习新信息的同时保持历史信息留存持久化,所以本文引入了将RNN的隐含层添加长短时记忆网络结构(long short-term memory, LSTM)的方法进行研究^[15-18]。

LSTM在分析时间序列方面有着天然的优势,然而在工业数据中的应用较少,本文将此方法用于硅含量分析,以期取得较好的效果。

1 LSTM-RNN 神经网络模型

本节描述LSTM-RNN神经网络模型^[19]的结构、建模方法。

1.1 RNN 神经网络模型

RNN的精髓在于对隐藏层之间进行了连接,跨越了时间点。换句话说,隐含层的信息不仅传递到了输出层,还传递到了下一时刻的隐含层。图1表示了两个时间点网络的相互关联。

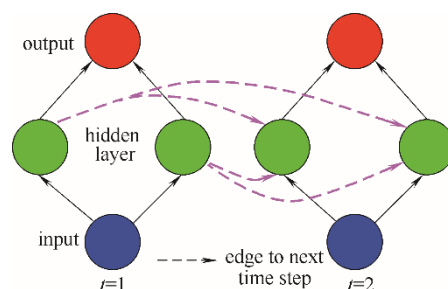


图1 RNN节点间关联

Fig.1 Structure of RNN

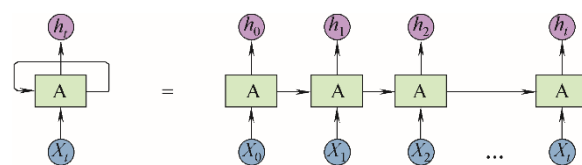


图2 RNN节点展开

Fig.2 Unfold of RNN

把网络展开的示意图如图2所示。展开的网络可以采用反向传播的方式进行训练,其本质就是BP误差反向传播,不同的是这种传播是跨时间步的,即参数是共享的。也就是每一步误差的反向传递不仅依赖当前网络状态,还依赖前步网络状态。

然而,传统的RNN存在一个梯度消失的问题^[20],即如果当前的输出与非常久远的时间序列有关,RNN将无法学习到。所以需要引入LSTMs^[21-29]。

1.2 LSTM对RNN的改进

LSTM模型由Hochreiter等提出^[30]。在常规的RNN中,隐含层一般是一个非常简单的节点,比如Tanh, LSTM则是将隐含层的简单节点改进为存储单元。存储单元的基本结构如图3所示。

存储单元由输入门 i 、输出门 o 、遗忘门 f 和记忆细胞 c 组成。前向传播时,输入门确定什么时候让激活传入存储单元,而输出门确定什么时候让激活

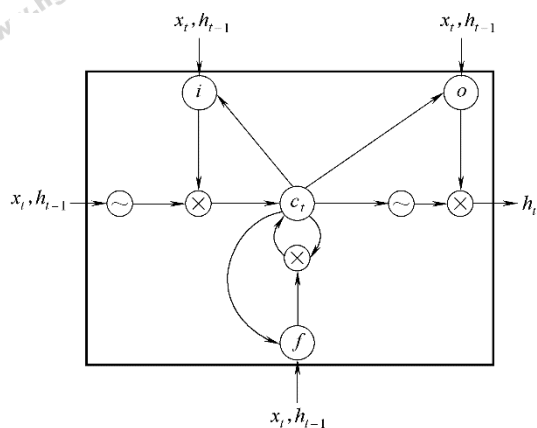


图 3 存储单元基本结构

Fig.3 Structure of memory cell

传出存储单元；反向传播时，输出门确定什么时候让错误流入存储单元，而输入门确定什么时候让它流出存储单元。输入门、输出门、遗忘门是控制信息流的关键。存储单元的运作原理可以用式（1）～式（5）表示^[31-33]

$$i_t = \delta(W_{ix}x_t + U_{ih}h_{t-1} + V_{ic}c_{t-1} + b_i) \quad (1)$$

$$f_t = \delta(W_{fx}x_t + U_{fh}h_{t-1} + V_{fc}c_{t-1} + b_f) \quad (2)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tanh(W_{cx}x_t + U_{ch}h_{t-1} + b_c) \quad (3)$$

$$o_t = \delta(W_{ox}x_t + U_{oh}h_{t-1} + V_{oc}c_t + b_o) \quad (4)$$

$$h_t = o_t \cdot \tanh(c_t) \quad (5)$$

其中， c_t 表示 t 时刻记忆细胞的计算方法； h_t 为 t 时间点 LSTM 单元的所有输出； W, U, V, b 分别是系数组成的矩阵和偏置的向量， δ 是激活函数 sigmoid； \cdot 是点乘运算； i_t, f_t, o_t 分别是 t 时间点输入门、遗忘门和输出门的计算方法。由图 3 可以看出输入门、遗忘门和输出门 3 个控制门的输出各自连接到了一个乘法元件上，从而分别控制信息流的输入、输出以及细胞单元的状态。

2 实例分析

本节结合实际，通过对某钢铁厂一座容积为 2600 m³ 的高炉数据进行分析，进行实验验证。介绍了变量确定的方法、模型评价指标的选择、模型参数的设定和对比结果分析。

2.1 数据的选择

由于实际生产中可能会出现硅含量的剧烈变动，为了表征实际的工况，本文选取了具有代表性的一段数据，数据中每隔一段时间都会出现非常剧烈的变动。

影响硅含量的因素有很多，鼓风的特性、入炉原料的性质、炉内物理状况都会影响硅含量。常用的

用于硅含量预测的变量有炉顶压力、透气性、喷煤、富氧率、炉顶温度、压差、料批、热封压力、热风温度、热风流量、鼓风湿度、前一炉硅含量。本文结合走势曲线以及与因变量的相关系数这一指标确定了 7 个与硅含量预测的相关度最高的变量，见表 1。并且将这 7 个变量作为 LSTM-RNN 的输入变量。

表 1 模型输入变量

Table 1 Input variable of model

Symbol	Variable	Unit	Correlation coefficient
x_1	air permeability	%	0.1101
x_2	oxygen-rich rate	%	-0.1031
x_3	top temperature	K	-0.1337
x_4	differential pressure	kPa	0.1741
x_5	batch of material	—	0.0293
x_6	hot blast flow	m ² ·h ⁻¹	0.0314
x_7	previous Si-content	%	0.5891

为防止不同自变量数量级的巨大差别造成奇异解，对数据进行归一化处理，映射到[-1,1]之间。

2.2 模型评价指标的选择

本文采用的评价指标是命中率和均方误差。

命中率是预测误差的绝对值小于或等于 0.1 的样本数占总样本数的比率，它体现预测模型的准确程度，其定义公式为

$$\text{ratio} = \frac{1}{n} \left(\sum_{i=1}^n \text{hit}_i \right) \times 100\% \quad (6)$$

$$\text{hit}_i = \begin{cases} 1, & |e_i| \leq 0.1 \\ 0, & \text{others} \end{cases} \quad (i = 1, 2, \dots, n) \quad (7)$$

其中，ratio 为命中率； n 为样本数目；hit 为命中； e 为预测误差。

均方误差指的是误差平方的平均值，它体现预测效果的稳定程度，其定义公式为

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n e_i^2 \quad (8)$$

2.3 模型参数设定

关于 LSTM-RNN 模型的参数设定，目前没有确定的标准，由于硅含量数据规模不大，而且计算机的计算能力已经足够强大，所以本文采用程序寻优的方法来确定最优参数。需要调节的主要参数有存储单元数目、结构和训练集时间窗口大小。对每种情况训练 10 次，监测其命中率。

监测结果说明模型效果波动非常小，每种参数设定下命中率的偏差小于 15%。当训练集时间窗为 21，存储单元数目为 20 时效果较好，后续的结果分析选择了该种设定下比较稳定的一组结果。最优参

数结果也满足隐含层数适中的先验知识——当节点数过小时,数据特征提取不足;当节点数过大时,数据则会出现过拟合。

依据程序寻优的结果,过于庞大的数据集对预测准确率呈反作用,因此对每一个时间点的预测,使用之前的 21 个时间点数据作为模型训练集。

将逐个时间点的数据输入网络,预测下一时间点的硅含量。然后将下一时间点的真实硅含量值当做已知输入网络,继续预测,在线滚动。

由于模型效果稳定性较高,对于不同高炉的硅含量预测,虽然内部物理化学特性不同,都可以采用此方法,通过程序自动寻优的方法建模,而不需要人工调试。

2.4 预测结果分析

为了说明 LSTM-RNN 模型的性能,本文将 PLS、典型 RNN 与 LSTM-DNN 进行性能比较,由于机器学习算法存在一定的波动,为了进一步说明其稳定性,本文采用 10 次训练的结果,取平均值以展示算法的平均水平。

如表 2 所示, LSTM-RNN 比 PLS 命中率提升 22%,比典型 RNN 提升了 12%。在均方误差方面, LSTM-RNN 模型比另外两种方法均有显著提升。

表 2 模型性能对比

Table 2 Performance comparison of models

Model	RATE/%	MSE
PLS	64	5.93×10^{-3}
RNN	74	4.26×10^{-3}
LSTM-RNN	86	2.03×10^{-3}

图 4、图 5、图 6 分别为 PLS 模型、典型 RNN 模型、LSTM-RNN 模型预测 500 个时间点的预测结果。由图可以看出,采用 PLS 模型的硅含量预测整体偏差比较大,采用 RNN 模型的硅含量预测出现了很多偏差比较严重的值,说明其预测效果均不是很理想。采用 LSTM-RNN 模型的硅含量预测,数据跟踪效果更好,其走势基本一致,且预测值更加接近,取得了更好的预测效果。

由于测试数据来源于波动比较大的炉况,因而 85% 的命中率不是十分理想。但是对于运行非常平稳的炉况,预测效果将会提升更多。

在某些应用场景下,操作工只需要知道硅含量的变化趋势,不需要知道硅含量的真实值。由于本模型对趋势的预测非常准确,非常适合于此类问题建模。

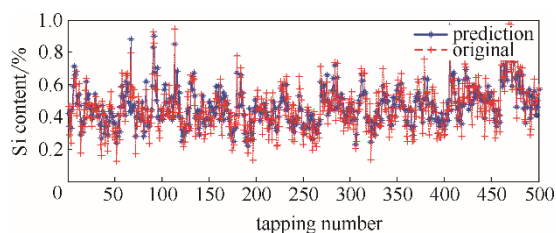


图 4 PLS 模型预测结果

Fig.4 Prediction of PLS

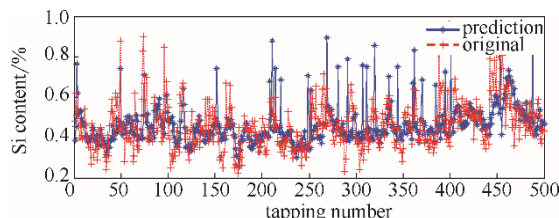


图 5 RNN 模型预测结果

Fig.5 Prediction of RNN

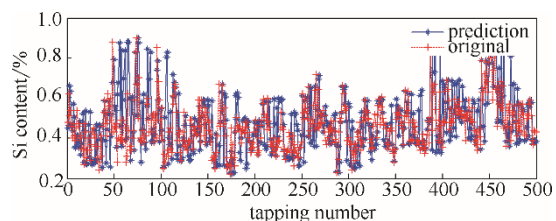


图 6 LSTM-RNN 模型预测结果

Fig.6 Prediction of LSTM-RNN

通过以上结果分析,发现本文采用的 LSTM-RNN 模型对铁水硅含量的预测效果较好,而且预测结果稳健,在实际应用中有一定的参考价值。

3 结 论

本文针对高炉炼铁的大滞后、动态工况特征,采取了适用于此类问题的 LSTM-RNN 模型;采用非常有代表性的变化复杂的现场数据对模型进行了验证,比前人的实验数据更加接近真实工况;采用自动求解最优参数的方法设定网络参数,充分发挥出了模型的特点,相较于前人的工作取得了较好的模型预测精度。而且预测结果稳定性好,为硅含量预测的研究开辟了新的思路,也对深度学习在工业现场的应用具有一定的实际意义。

符 号 说 明

- A —— 隐令层
- c_t —— t 时刻记忆细胞的计算方法
- e —— 预测误差
- f_t —— t 时间点遗忘门的计算方法
- hit —— 命中

i_t —— t 时间点输入门的计算方法
MSE —— 均方误差
 n —— 样本数目
 o_t —— t 时间点输出门的计算方法
ratio —— 命中率
 W, U, V, b —— 系数组成的矩阵和偏置的向量
 δ —— 激活函数 sigmoid
 \cdot —— 点乘运算

References

- [1] 李界家, 杨志宇, 曹阳. 铁水硅含量的集成模糊神经网络预测方法[J]. 计算机与应用化学, 2013, **30**(10): 1113-1116.
LI J, YANG Z Y, CAO Y. Research on hot metal silicon content prediction based on integrated neural network[J]. Computers and Applied Chemistry, 2013, **30** (10): 1113-1116.
- [2] TAHASHI H, KAWAI H, KOBAYASHI M, *et al.* Two dimensional cold model study on unstable solid descending motion and control in blast furnace operation with low reducing agent rate[J]. ISIJ International, 2005, **45**(10): 1386-1395.
- [3] CHU M S, YANG X F, SHEN F M. Numerical simulation of innovative operation of blast furnace based on multi-fluid model[J]. Journal of Iron and Steel Research, International, 2006, **13**(6): 8-15.
- [4] 赵敏. 高炉冶炼过程的复杂机理及其预测研究[D]. 杭州: 浙江大学, 2008.
ZHAO M. Complexity mechanism and predictive research for BF ironmaking process[D]. Hangzhou: Zhejiang University, 2008.
- [5] NOGAMI H, CHU M S, YAGI J. Multi-dimensional transient mathematical simulator of blast furnace process based on multi-fluid and kinetic theories[J]. Computers & Chemical Engineering, 2005, **29**(11): 2438-2448.
- [6] 郇传厚, 渐令, 陈积明, 等. 复杂高炉炼铁过程的数据驱动建模及预测算法[J]. 自动化学报, 2009, **35**(6): 725-730.
GAO C H, JIAN L, CHEN J M, *et al.* Data-driven modeling and predictive algorithm for complex blast furnace iron-making process[J]. Acta Automatica Sinica, 2009, **35**(6): 725-730.
- [7] SAXEN H. Short-term prediction of silicon content in pig iron[J]. Canadian Metallurgical Quarterly, 1994, **33**(4): 319-326.
- [8] LUO S H, GAO C H, ZENG J S, *et al.* Blast furnace system modeling by multivariate phase space reconstruction and neural networks[J]. Asian Journal of Control, 2013, **15**(2): 553-561.
- [9] 安剑奇, 陈易斐, 吴敏. 基于改进支持向量机的高炉一氧化碳利用率预测方法[J]. 化工学报, 2015, **66**(1): 206-214.
AN J Q, CHEN Y F, WU M. A prediction method for carbon monoxide utilization ratio of blast furnace based on improved support vector regression[J]. CIESC Journal, 2015, **66**(1): 206-214.
- [10] 曾燕飞, 李小伟. 基于 BP 神经网络的高炉铁水硅含量预测模型研究[J]. 控制与测量, 2006, **22**(19): 291-293.
ZENG Y F, LI X W. Model prediction model of blast furnace hot metal Si-content based on BP neural network's study[J]. Control and Measurement, 2006, **22**(19): 291-293.
- [11] 宋菁华, 杨春节, 周哲, 等. 改进型 EMD-Elman 神经网络在铁水硅含量预测中的应用[J]. 化工学报, 2016, **67**(3): 729-735.
SONG J H, YANG C J, ZHOU Z, *et al.* Application of improved EMD-Elman neural network to predict silicon content in hot metal[J]. CIESC Journal, 2016, **67**(3): 729-735.
- [12] 蒋朝辉, 董梦林, 桂卫华, 等. 基于 Bootstrap 的高炉铁水硅含量二维预报[J]. 自动化学报, 2016, **42**(5): 715-723.
JIANG Z H, DONG M L, GUI W H, *et al.* Two-dimensional prediction for silicon content of hot metal of blast furnace based on Bootstrap[J]. Acta Automatica Sinica, 2016, **42**(5): 715-723.
- [13] 宋贺达, 周平, 王宏, 等. 高炉炼铁过程多元铁水质量非线性子空间建模及应用[J]. 自动化学报, 2016, **42**(21): 1664-1679.
SONG H D, ZHOU P, WANG H, *et al.* Non-linear subspace modeling of multivariate molten iron quality in blast furnace ironmaking and its application[J]. Acta Automatica Sinica, 2016, **42**(21): 1664-1679.
- [14] ZHOU H, YANG C J, LIU W H, *et al.* A sliding-window T-S fuzzy neural network model for prediction of silicon content in hot metal[C]// 20th IFAC World Congress. IFAC-PapersOnLine, 2017, **50**(1): 14988-14991.
- [15] SENIOR A. Context dependent phone models for LSTM RNN acoustic modeling[J]. IEEE Int. Conf. Acoust, Speech Signal Process, 2015, (1): 4585-4589.
- [16] LIU C J, WANG Y Q, KSHITIZ K, *et al.* Investigations on speaker adaptation of LSTM-RNN models for speech recognition[C]//IEEE International Conference on ICASSP, 2016: 5020-5024.
- [17] GRAVES A, JAITLY N, MOHAMED A. Hybrid speech recognition with deep bi-directional LSTM[C]//Proceedings of the IEEE Workshop on Automatic Speech Recognition and Understanding, 2013.
- [18] SAK H, SENIOR A, BEAUFAYS F. Long short-term memory recurrent neural network architectures for large scale acoustic modeling[C]//Annual Conference of the International Speech Communication Association (Interspeech), 2014: 338-342.
- [19] ZACHARY C LIPTON. Learning to diagnose with LSTM recurrent neural networks[C]//ICLR 2016. Computer Science, 2017.
- [20] GRAVES A, SCHMIDHUBER J. Frame wise phoneme classification with bidirectional LSTM and other neural network architectures[J]. Neural Networks, 2005, **18**(5/6): 602-610.
- [21] SONG E, KANG H G. Multi-class learning algorithm for deep neural network-based statistical parametric speech synthesis[C]//Proc. 24th Eur. Signal Process Conf., 2016, 1951-1955.
- [22] ACHANTA S, GODAMBE T, GANGASHETTY S V. An investigation of recurrent neural network architectures for statistical parametric speech synthesis[C]//Proc. Interspeech, 2015: 859-863.
- [23] ZEN H, SAK H. Unidirectional long short-term memory recurrent neural network with recurrent output layer for low-latency speech synthesis[C]//IEEE Speech Signal Process, 2015: 4470-4474.
- [24] BYEON W, LIWICKI M, BREUEL T. Texture classification using 2D LSTM networks[C]//Pattern Recognition International Conference, 2014: 1144-1149.
- [25] BYEON W, BREUEL T M. Supervised texture segmentation using 2D LSTM networks[C]//Processing 2014 IEEE International Conference, 2014: 4373-4377.
- [26] PINHEIRO P, COLLOBERT R, JEBARAT, *et al.* Recurrent convolutional neural networks for scene labeling[C]//Proceedings of the 31 st International Conference on Machine Learning (ICML-14), 2014: 82-90.
- [27] SOCHER R, HUVAL B, BATH B, *et al.* Convolutional recursive deep learning for 3D object classification[C]// Advances in Neural Information Processing Systems, 2012: 665-673.
- [28] GONZALEZ-DOMINGUEZ J, LOPEZ-MORENO I, MORENO P J,

- et al.* Frame by frame language identification in short utterances using deep neural networks[J]. *Neural Networks*, 2015, **64**(C): 49-58.
- [29] GERS F A, SCHMIDHUBER J, CUMMINS F. Learning to forget: continual prediction with LSTM[J]. *Neural Computation*, 2000, **12**(10): 2451-2471.
- [30] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. *Neural Computation*, 1997, **9**(8): 1735-1780.
- [31] BENGIO Y, SIMARD P, FRASCONI P. Learning long-term dependencies with gradient descent is difficult[J]. *IEEE Transactions on Neural Networks*, 1994, **5**(2): 157-166.
- [32] GREFF K, SRIVASTAVA R K, KOUTNÍK J, *et al.* LSTM: a search space Odyssey[J]. *IEEE Transactions on Neural Networks & Learning Systems*, 2015, **28** (10): 2222.
- [33] GERS F A, SCHRAUDOLPH N N, SCHMIDHUBER J. Learning precise timing with LSTM recurrent networks[J]. *Journal of Machine Learning Research*, 2003, **3**(1): 115-143.