

本文引用格式: 姚嘉伟,蔡延光.基于多核卷积和多头自注意力的心电图身份识别方法[J].自动化与信息工程,2023,44(5):32-37.

YAO Jiawei, CAI Yanguang. ECG identity recognition method based on multi-kernel convolution and multi-head self-attention[J]. Automation & Information Engineering, 2023,44(5):32-37.

基于多核卷积和多头自注意力的心电图身份识别方法*

姚嘉伟 蔡延光

(广东工业大学自动化学院, 广东 广州 510006)

摘要: 为提高心电图身份识别过程中神经网络的训练效率及识别正确率, 提出一种基于多核卷积和多头自注意力的心电图身份识别方法。首先, 利用多个大小不同的卷积核对预处理后的单个心拍进行特征提取; 然后, 采用多头自注意力模块加强卷积通道中全局特征的提取效果; 接着, 将特征通道一分为二进行半实例归一化, 使多头自注意力模块能够关注归一化前后的特征, 提升神经网络的收敛速度; 最后, 将多核多头自注意力模块进行 ResNet 残差连接。该方法在 QT 数据集上经过 20 个 epoch 训练, 实现了 94.92% 的识别正确率。利用 ResNet 进行对比实验的结果表明, 该方法能够有效地提升神经网络的训练效率及识别正确率。

关键词: 多核卷积; 多头自注意力机制; 半实例归一化; 心电图; 身份识别

中图分类号: TP391

文献标志码: A

文章编号: 1674-2605(2023)05-0005-06

DOI: 10.3969/j.issn.1674-2605.2023.05.005

ECG Identity Recognition Method Based on Multi-kernel Convolution and Multi-head Self-attention

YAO Jiawei CAI Yanguang

(College of Automation, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: To improve the training efficiency and recognition accuracy of neural networks in the process of electrocardiogram identity recognition, a electrocardiogram identity recognition method based on multi-kernel convolution and multi-head self-attention is proposed. Firstly, feature extraction is performed on a preprocessed single heartbeat using multiple convolution checks of different sizes; Then, a multi-head self-attention module is used to enhance the extraction effect of global features in the convolutional channel; Next, the feature channels are divided into two for semi instance normalization, enabling the multi-head self-attention module to focus on the features before and after normalization, thereby improving the convergence speed of the neural network; Finally, connect the multi-kernel and multi-head self-attention module to ResNet residual. This method achieved a recognition accuracy of 94.92% after 20 epochs of training on the QT dataset. The results of comparative experiments using ResNet show that this method can effectively improve the training efficiency and recognition accuracy of the neural network.

Keywords: multi-kernel convolution; multi-head self-attention mechanism; semi instance normalization; electrocardiogram; identity recognition

0 引言

心电图 (electrocardiogram, ECG) 是一种无创检查手段, 因价格便宜、操作简便, 是心脏疾病初筛的重要方式。随着人们对加密需求的不断提高, 生物认证的方式逐渐成为热点研究方向。利用 ECG 进行身份识别就是一种有效的生物认证方式。

近年来, 利用 ECG 进行身份识别引起了许多学者的关注。BIEL 等^[1]通过手动提取 ECG 的形态特征并采用相似分类法进行身份识别, 证明了利用 ECG 进行身份识别的可行性。LEE 等^[2]利用一种可扩展的极限学习机进行 ECG 身份识别, 采用增量学习的策略解决了由于新成员加入使模型参数发生更新而导致

* 基金项目: 广东省科技计划项目 (2016A050502060, 2020B1010010005); 广州市科技计划项目 (202206010011, SL2022B03J01092)。

的灾难性遗忘问题。DALAL 等^[3]从 ECG 中提取多累积量特征, 利用内核极限学习机 (kernel extreme learning machine, KELM) 对基于多累积量特征的 ECG 进行身份识别, 虽然取得了较高的识别正确率, 但该方法不是端对端的, 且需要先对 ECG 进行多阶统计量的特征计算, 影响识别效率。SRIVASTVA 等^[4]将残差网络 (ResNet) 和 DenseNet 的预训练模型进行堆叠, 利用 ECG 对堆叠后的模型进行微调, 构成身份识别模型, 虽能获得较好的识别效果, 但训练开销巨大。LI 等^[5]将极端随机树 (extra tree, ET) 和统一流形逼近与投影 (uniform manifold approximation and projection, UMAP) 两种方法叠加, 在 ECG 中提取集成特征, 再利用极限梯度提升器 (XGBoost) 对集成特征进行身份识别, 虽然能够达到 96.88% 的识别正确率, 但需要复杂的特征工程。JYOTISHI 等^[6]提出基于分层的长短期记忆模型和注意力机制的 ECG 身份识别系统, 在 250 个 epoch 训练后, 取得较好的识别效果。TUNCER 等^[7]提出一种基于 Tietze 模式和邻域成分分析 (neighborhood component analysis, NCA) 的身份识别方法, 虽然表现出色, 但处理过程十分繁琐。杨向林等^[8]提出一种小波分解和特征融合的 ECG 身份识别方法, 通过将波形特征和小波特征融合, 利用相关系数阈值法对融合特征进行身份识别, 达到 95% 的识别正确率。虽然过往利用 ECG 进行身份识别的研究已经提出了许多有效的方法, 但这些方法仅关注识别正确率, 没有考虑模型深度和训练效率的问题。

为减少模型深度, 提高训练效率, 并保证模型的识别正确率, 本文提出基于多核卷积和多头自注意力的多核多头自注意力 (multi-kernel multi-head self-attention, MKMSA) 模块, 并利用 ResNet 对 MKMSA 模块进行连接, 构造并训练能够准确进行身份识别的方法。

1 数据处理

1.1 数据预处理

ECG 中一个完整的 P、QRS、T 波的过程称为一个心拍。在不同的采样频率下, ECG 心拍的波形会出

现不同步的现象, 容易导致神经网络将其预测为 2 个不同的心拍; 同时, 神经网络无法对不同长度的数据进行处理。为此, 需对 ECG 进行预处理, 过程如下:

- 1) 将 ECG 进行重采样, 采样频率为 360 Hz;
- 2) 根据专家标注, 截取 P 波前 40 ms 到 T 波结束的片段作为一个心拍数据;
- 3) 将每个心拍的数据长度调整为 512 个样本点, 在长度不足的心拍数据后用 0 填充。

正常的心拍数据长度一般为 0.8~1.2 s, 预处理后的心拍数据长度为 1.42 s, 既能保证心拍的完整性, 又能满足神经网络的输入要求。

1.2 数据划分

对不同采样者的 ECG 进行独热编码标注, 并将心拍数量的 70% 作为训练集, 30% 作为测试集, 保证了每个身份标签对应的数据在训练集和测试集的分布相同。

2 识别方法

2.1 多核卷积通道

在神经网络训练的过程中, 卷积核大小是一个难以确定和调整的超参数。本文采用 Inception 模块的思想^[9], 并联卷积核大小不一的卷积神经网络。ECG 的信号较长, 为了尽可能地扩大卷积核, 同时减少网络参数, 采用空洞卷积层作为多核卷积通道的构成部分。通过这种方式, 神经网络在反向传播的过程中, 能自动地从最适合的通道快速拟合, 提升网络的训练速度。

2.2 特征权重调整

为提高特征的提取能力, 采用注意力机制^[10]广泛关注卷积通道中的整体特征, 并对权重进行调整。自注意力机制的计算过程如下:

- 1) 输入向量 x 被重新编码为特征向量 a , 即 $a = w^a x$;
- 2) 根据输入向量 x 进行位置编码, 生成位置向量 e , 即 $e = w^e x$;
- 3) 根据特征向量 a 与位置向量 e 的和, 分别计算匹配项 q 、被匹配项 k 、特征项 v , 即 $q = w^q (a + e)$ 、 $k = w^k (a + e)$ 、 $v = w^v (a + e)$;

4) 将 q 与 k 相乘, 得到匹配项与被匹配项的相似度 $k^T q$;

5) 将 $k^T q$ 与 v 相乘得到自注意力特征 b , 即

$$b = vk^T q \quad (1)$$

其中, w^a 、 w^e 、 w^q 、 w^k 、 w^v 均为神经网络训练过程中需要学习的参数矩阵。

将 q 、 k 、 v 整体看作为自注意力机制的一个特征提取头, 那么多头自注意力机制就是在自注意力机制的基础上, 采用多组 q 、 k 、 v 构成的多个特征提取头对输入向量 x 进行特征计算与拼接, 再通过线性转换 w^o 得到输出:

$$MultiHead(Q, K, V) = Concat(b_1 \cdots b_h) w^o \quad (2)$$

式中: $b_1 \cdots b_h$ 是公式(1)每一个特征提取头中产生的自注意力特征 b , 将其拼接后经过线性转换 w^o 得到多头自注意力的特征 M 。

2.3 实例归一化

实例归一化是由 ULYANOV 等^[11]提出的, 用于替代批归一化 (batch normalization, BN)^[12], 可提高对抗生成网络对图像风格迁移时神经网络的收敛速度。

假设 x_{ijk} 是第 $tijk$ 个元素, 其中 t 为批次中的索引, i 为特征通道, j 和 k 为跨越空间维度, 那么实例归一化先通过公式(4)、(5)分别计算整个批次实例的平均值 μ_{ii} 和方差 σ_{ii}^2 , 再由公式(3)计算归一化后的特征向量 y_{ijk} 。

$$y_{ijk} = \frac{x_{ijk} - \mu_{ii}}{\sqrt{\sigma_{ii}^2 + \varepsilon}} \quad (3)$$

$$\mu_{ii} = \frac{1}{HW} \sum_{l=1}^W \sum_{m=1}^H x_{ilm} \quad (4)$$

$$\sigma_{ii}^2 = \frac{1}{HW} \sum_{l=1}^W \sum_{m=1}^H (x_{ilm} - \mu_{ii})^2 \quad (5)$$

式中: ε 为一个很小的正数, 避免了除数等于 0 的情况, H 、 W 分别为第 t 批次中第 i 通道特征的高和宽。

2.4 MKMSA 模块

MKMSA 模块主要包括多核卷积通道、多头自注意力机制和半实例归一化^[13], 结构如图 1 所示。

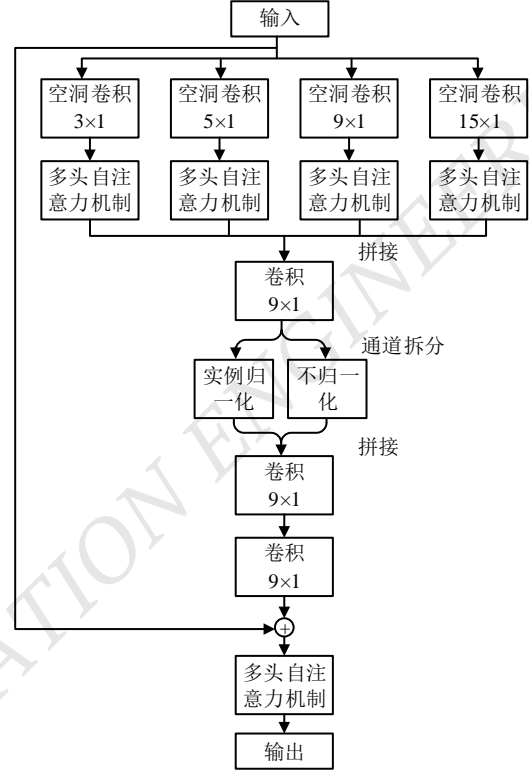


图 1 MKMSA 模块结构图

MKMSA 模块的具体计算过程如下:

1) 通过多个大小不同的空洞卷积核对输入 x 进行特征提取, 产生卷积特征 x_1 、 x_2 、 x_3 和 x_4 ;

2) 利用多头自注意力机制对 x_1 、 x_2 、 x_3 和 x_4 的特征权重进行调整, 生成多头自注意力特征, 并将多头自注意力特征进行合并, 生成融合多通道多头自注意力特征 M , 如公式(2)所示;

3) 对 M 进行卷积操作生成卷积特征 x_5 ; 利用半实例归一化^[12]对 x_5 的特征通道一分为二, 生成 x_{5-1} 和 x_{5-2} ; 利用公式(3)对 x_{5-1} 进行实例归一化生成 x_{5-1-n} , x_{5-2} 保持不变; 将 x_{5-1-n} 和 x_{5-2} 的通道进行拼接及卷积操作, 生成 x_6 , 半实例归一化能够增强 x_{5-1-n} 通道的独立性, 并使数据分布更加平滑稳定, 有利于梯度计算; 同时由于 x_{5-2} 没有实例归一化, 神经网络在学习过程中能够充分提取原始数据的特征,

从而加速神经网络的收敛；

4) 将卷积特征 x_6 和输入 \mathbf{x} 进行残差相加生成网络特征向量 \mathbf{x}_7 ；

5) 利用公式(2)对 \mathbf{x}_7 进行特征调整，生成模块输出 y 。

2.5 神经网络结构设计

采用 ResNet 的残差连接对 MKMSA 模块进行连接，本文方法的神经网络结构如图 2 所示。

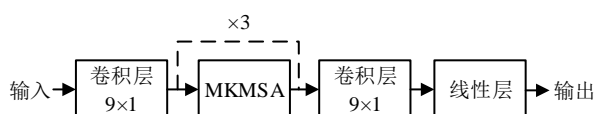


图 2 本文方法的神经网络结构图

图 2 中，2 个卷积层主要用于调整输入和输出 MKMSA 模块的特征通道数量；线性层整合具有类别区分性的特征，用于身份识别。

2.6 算法流程

基于多核卷积和多头自注意力的心电图身份识别方法的训练流程、识别流程分别如算法 1、算法 2 所示。

算法 1：训练流程

输入：心电图 X ，迭代次数 N ，身份标签 l_c

1. 将心电图 X 预处理成单个心拍 x_i ；
2. 初始化 $n = 0$ ；
3. 如果 $n < N$ ，将心拍 x_i 和身份标签 l_c 输入神经网络，计算预测值 y_i ；
4. 如果 $n > N$ ，保存神经网络参数并退出；
5. 利用 $Loss = -\frac{1}{N} \sum_i \sum_{c=1}^M y_{ic} \log(p_{ic})$ 计算每个预测值 y_i 和对应的标签类别 c 的损失，更新神经网络参数；
6. $n = n + 1$ ；
7. 输出：训练好的神经网络参数

算法 2：识别流程

输入：心电图 X ，训练好的神经网络参数

1. 将心电图 X 预处理成单个心拍 x_i ；
2. 加载训练好的神经网络参数；
3. 将心拍 x_i 输入网络模型，计算预测值 y_i ；
4. 输出：身份标签 y_i

3 实验

3.1 实验环境

本文方法采用 Pytorch 框架实现，实验采用的编程语言为 Python，处理器为 Intel I7-12700KF，内存为 32 GB，显卡为 Nvidia RTX 3060Ti。

3.2 实验数据

实验数据采用 Physionet^[14]的公开数据库 QT^[15]。QT 数据库包括 105 段 15 min 采样频率为 250 Hz 的 ECG，且含有专家标注的 P、QRS、T 波边界和峰值。

3.3 参数设置

实验采用交叉熵损失函数、选择随机梯度下降算法作为优化器，学习率为 0.002，batch size 为 256，训练了 20 个 epoch，本文方法的神经网络参数如表 1 所示。

表 1 本文方法的神经网络参数

网络类型	输出形状	数量
一维卷积	(256,1,512)	1
MKMSA 模块	(256,64,512)	3
一维卷积	(256,1,512)	1
全连接层	(256,105)	1

3.4 评估指标

实验采用正确率 (Accuracy) 对本文方法进行评价，计算公式为

$$A_{\text{accuracy}} = \frac{TP}{TP + FP} \quad (6)$$

式中： TP 为真阳性的数量，表示在预测的身份结果中与真实身份相匹配的人数， FP 为假阳性的数量，表示在预测的身份结果中与真实身份不匹配的人数。

3.5 实验结果与分析

对本文方法的神经网络进行训练，训练曲线如图 3 所示。

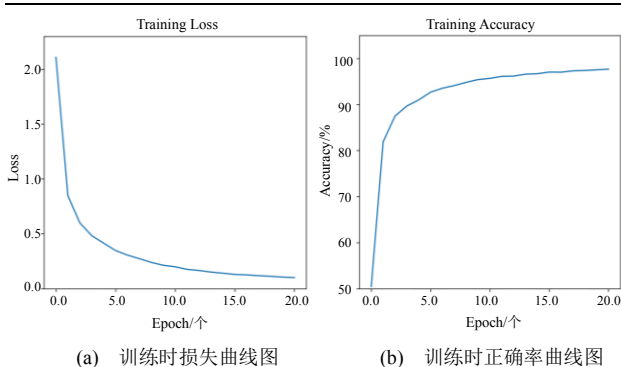


图3 本文方法的训练曲线

由图3可知，本文提出的神经网络在第15个epoch开始趋近于收敛状态，说明本文方法有效地提高了神经网络的训练效率。

本文方法在20个epoch上的训练结果如图4所示。

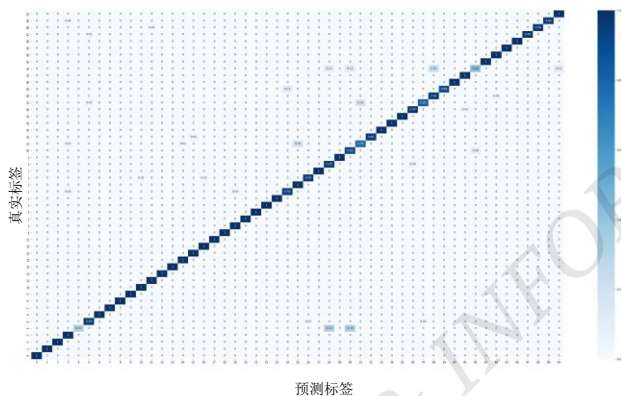


图4 本文方法的实验结果

由图4可知：本文方法能够正确地识别大部分心电图；同时由公式(6)对实验数据进行统计，本文方法在20个epoch上获得94.92%的识别正确率。

为进行对照实验，设计一个参数数量与本文神经网络相当的ResNet，并对其训练，训练曲线如图5所示。

由图5可知，ResNet在接近80个epoch时，训练集的正确率才超过90%；对比本文提出的神经网络，其收敛速度较慢。

在对照实验中，采用相同的优化器和损失函数在ResNet上训练了200个epoch，结果如图6所示。

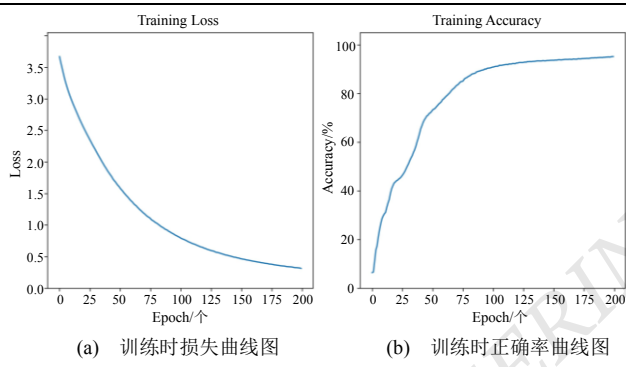


图5 ResNet上的训练曲线

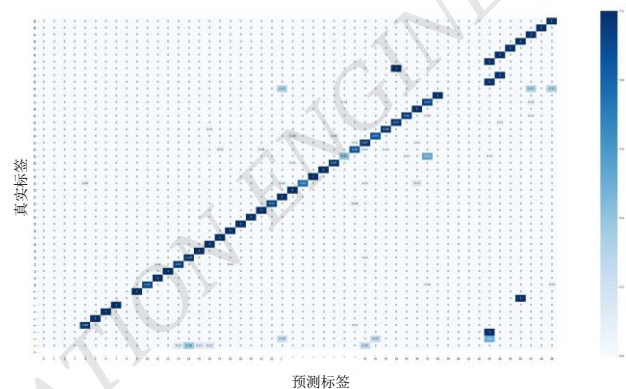


图6 ResNet上的实验结果

由图6可知：ResNet中出现了部分心电图无法正确识别的情况，同时由公式(6)对实验数据进行统计，其正确率为89.45%；相比本文方法，在保证正确率的情况下提高了训练效率，加快了神经网络的收敛。

4 结论

为了减少身份识别中神经网络的深度，提高神经网络的训练效率，本文设计了一种MKMSA模块并构建了基于多核卷积和多头自注意力的心电图身份识别方法，在公开数据集QT数据集上获得的识别正确率为94.92%，且有效地提升了训练效率。在未来的研究中，随着联邦机器学习的发展能够给心电图提供更好的隐私保护，使基于心电图的生物识别方法能够应用到更多现实的场景。

参考文献

[1] BIEL L, PETERSSON O, PHILIPSON L, et al. ECG analysis: a new approach in human identification[J]. IEEE Transactions on Instrumentation and Measurement, 2001,50(3):808-812.
 [2] LEE C L, CHEN Y T, WU A Y. A scalable extreme learning

- machine (S-ELM) for class-incremental ECG-based user identification[C]//2021 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, 2021: 1-5.
- [3] DALAL S, VISHWAKARMA V P. Classification of ECG signals using multi-cumulants based evolutionary hybrid classifier[J]. Scientific Reports, 2021, 11(1): 15092.
- [4] SRIVASTVA R, SINGH Y N, SINGH A. Statistical independence of ECG for biometric authentication[J]. Pattern Recognition, 2022, 127: 108640.
- [5] LI M, SI Y, YANG W, et al. ET-UMAP integration feature for ECG biometrics using stacking[J]. Biomedical Signal Processing and Control, 2022, 71: 103159.
- [6] JYOTISHI D, DANDAPAT S. An ECG biometric system using hierarchical lstm with attention mechanism[J]. IEEE Sensors Journal, 2021, 22(6): 6052-6061.
- [7] TUNCER T, AYDEMIR E, DOGAN S, et al. New human identification method using Tietze graph-based feature generation[J]. Soft Computing, 2021, 25(21): 13437-13449.
- [8] 杨向林, 严洪, 李延军, 等. 基于小波分解和数据融合方法的 ECG 身份识别[J]. 航天医学与医学工程, 2009, 22(4): 296-301.
- [9] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015:1-9.
- [10] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30.
- [11] ULYANOV D, VEDALDI A, LEMPITSKY V. Instance normalization: The missing ingredient for fast stylization[J]. arXiv preprint arXiv:1607.08022, 2016.
- [12] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]//International Conference on Machine Learning, 2015: 448-456.
- [13] CHEN L, LU X, ZHANG J, et al. Hinet: Half instance normalization network for image restoration[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021:182-192.
- [14] GOLDBERGER A L, AMARAL L A N, GLASS L, et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals[J]. Circulation, 2000, 101(23): e215-e220.
- [15] LAGUNA P, MARK R G, GOLDBERG A, et al. A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG[C]//Computers in Cardiology 1997. IEEE, 1997:673-676.

作者简介:

姚嘉伟, 男, 1996 年生, 硕士研究生, 主要研究方向: 机器学习。E-mail: yaojw199@163.com

蔡延光, 男, 1963 年生, 博士, 教授/博士生导师, 主要研究方向: 网络控制与优化、组合优化、智能优化、智能交通系统等。E-mail: caiyg99@163.com

《自动化与信息工程》建立作者交流微信群

亲爱的作者朋友们:

《自动化与信息工程》(双月刊)原名《广东自动化与信息工程》《广州自动化》,创刊于 1980 年。40 余年来,本刊一直坚持传播学术信息,发表学术研究成果,体现了自身的价值。沐风栉雨一步步脚踏实地地走到如今,更有赖于一众不离不弃的作者。我们的作者群体一直在壮大,不断有新生力量补充加入。为便捷地开展各种沟通与交流,应广大作者的要求,本刊建立作者交流微信群,欢迎各位学术新人和大神们投稿并加群,交流学术成果、分享学习心得。

本刊长期接受稿件,投稿邮箱是: aie@giiim.ac.cn。群二维码将以邮件回复您,欢迎您推介亲朋好友一起来投稿。若符合条件,您还可以申请担任我们的审稿专家或者学术通讯员。我们的每一点进步,都需要您的参与和支持,您的信任和配合是支撑我们继续前进的强大动力。

《自动化与信息工程》编辑部