



(12)发明专利申请

(10)申请公布号 CN 110008674 A

(43)申请公布日 2019.07.12

(21)申请号 201910226702.4

(22)申请日 2019.03.25

(71)申请人 浙江大学

地址 310058 浙江省杭州市西湖区余杭塘路866号

(72)发明人 褚逸凡 沈海斌

(74)专利代理机构 杭州求是专利事务有限公司 33200

代理人 郑海峰

(51) Int. Cl.

G06F 21/32(2013.01)

A61B 5/00(2006.01)

A61B 5/0402(2006.01)

G06N 3/04(2006.01)

G06N 3/08(2006.01)

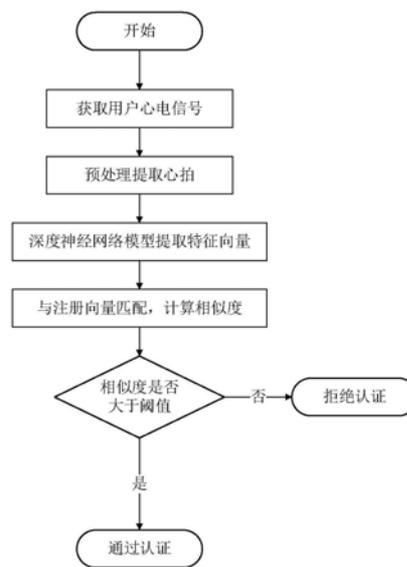
权利要求书2页 说明书5页 附图2页

(54)发明名称

一种高泛化性的心电信号身份认证方法

(57)摘要

本发明公开了一种高泛化性的心电信号身份认证方法,涉及生物信号身份认证领域。解决现有方法泛化性较差,且对于新用户,需进行重新训练的问题。方法采用了并行多尺度一维残差网络架构,使用了三种尺寸不同的卷积核进行并行特征提取,能够对不同长度的心电信号分别提取特征,有效地保留了心电信号中包含的时序性,提高了整体身份识别的效果和准确性。采用了中心目标函数与边界目标函数对神经网络进行训练,保证了所提取特征类间离散性和类内聚合性,从而使得同类特征聚集得更加紧密,异类特征分散得更加开阔,在匹配时能够提高匹配得准确性,同时使得训练得到得模型对训练数据的依赖程度降低,大大提高了模型的泛化性和鲁棒性。



1. 一种高泛化性的心电信号身份认证方法,其特征在于,所述方法包括以下几个步骤:

1) 获取用户的心电信号;

2) 对用户心电信号进行预处理,提取固定长度心拍;

3) 通过经由中心目标函数与边界目标函数训练的并行多尺度一维残差网络架构深度神经网络模型,从预处理后的固定长度心拍中提取心电信号特征向量;

4) 将生成的用户心电信号特征向量与用户预先注册的心电信号特征模板进行匹配,根据二者的相似度确定匹配结果,完成认证。

2. 如权利要求1所述的高泛化性的心电信号身份认证方法,其特征在于,所述步骤2)中,对用户心电信号进行预处理步骤具体包括:

1) 对所述短时间心电信号进行重采样至500Hz,并对重采样后的信号进行滤波,去除噪声;

2) 对预处理后的心电信号进行QRS检测,获取R波峰值点位置,取R波峰值点前99个采样点和R波峰值点后200个采样点,与R波峰值点一起组成固定长度为300个采样点的心拍。

3. 如权利要求1所述的高泛化性的心电信号身份认证方法,其特征在于,所述步骤3)中的深度神经网络模型具体包括:

1) 所述的神经网络模型采用并行多尺度一维残差网络架构,其中每一个尺度中残差块采用的卷积核的尺寸分别为 1×3 , 1×5 和 1×7 ;

2) 所述的神经网络采用softmax目标函数,中心目标函数与边界目标函数的加权之和作为总的目标函数进行训练,目标函数形式如下:

$$\mathcal{L} = \mathcal{L}_S + \lambda_C \cdot \mathcal{L}_C + \lambda_M \cdot \mathcal{L}_M$$

$$\mathcal{L}_S = -\sum_{i=1}^m \log \frac{e^{W_{y_i}^T f(x_i) + b_{y_i}}}{\sum_{j=1}^n e^{W_j^T f(x_i) + b_j}}$$

$$\mathcal{L}_C = \frac{1}{2} \sum_{i=1}^m \|f(x_i) - c_{y_i}\|_2^2$$

$$\mathcal{L}_M = \sum_{i=1}^m \sum_a^K \sum_b^K [D_{ia}^P - D_{ib}^N + \alpha]$$

$$D_{ia}^P = \|f(x_i) - f(x_{ia}^P)\|_2^2$$

$$D_{ib}^N = \|f(x_i) - f(x_{ib}^N)\|_2^2$$

其中, \mathcal{L} 为总的目标函数, \mathcal{L}_S 、 \mathcal{L}_C 和 \mathcal{L}_M 分别表示softmax目标函数、中心目标函数和边界目标函数; m 和 n 分别表示训练样本个数和训练样本类别数; x_i 和 y_i 分别表示第 i 个训练样本与第 i 个训练样本对应的类别; $f(x_i)$ 为第 i 个训练样本的心电信号特征向量; c_{y_i} 为所有类别为 y_i 的样本的平均值; K 为近邻搜索个数; x_{ia}^P 和 x_{ib}^N 分别表示与样本 x_i 类别相同的距离最远的第 a 个样本,和与样本 x_i 类别不同的距离最近的第 b 个样本; α 为距离常数; $W_{y_i}^T$ 表示softmax全连接层的权值矩阵的第 y_i 个列向量的转置, b_{y_i} 表示softmax全连接层的偏置向量的第 y_i 个值, W_j^T 表示softmax全连接层的权值矩阵的第 j 个列向量的转置, b_j 表示softmax全连接层的偏置向量的第 j 个值; λ_C 和 λ_M 分别表示中心目标函数和边界目标函数的权重。

4. 如权利要求1所述的高泛化性的心电信号身份认证方法,其特征在于,所述步骤4)中的匹配过程具体包括:

1) 指定的用户在认证前,已经预先注册过一段短时间心电信号并进行相同的预处理,采用训练完的神经网络模型,从这段心电信号中提取出心电信号特征向量,将其作为指定用户的已注册特征模板向量;

2) 计算获得的指定用户心电信号特征向量与已注册特征模板向量之间的欧式距离,将欧式距离作为二者的匹配相似度;

3) 将匹配相似度与预算设定的阈值进行对比,若相似度大于设定的阈值,则认为二者互相匹配,通过认证;反之拒绝通过认证;

4) 若从获取的用户一段短时间的心电信号中能够提取多个心电信号特征向量,则采用投票机制进行决定,若能够通过匹配的特征向量个数大于特征向量总数的50%,则认为通过认证;反之拒绝通过认证。

一种高泛化性的心电信号身份认证方法

技术领域

[0001] 本发明涉及模式识别、信号处理和身份认证领域,特别是涉及一种高泛化性的心电信号身份认证方法。

背景技术

[0002] 随着信息技术的高速发展,信息化与数字化在给人类带来了巨大的便利的同时,也放大了个人或组织数据安全受到危险的可能性,许多身份认证方法正是因此应运而生。近年来,随着各项技术和采集设备的成熟,许多身份认证方法,开始从传统的指纹、人脸等生物识别,转向更加新兴、可靠、安全的生物信号特征识别。心电信号具有唯一性、活体性、独特性、隐私性,同时又兼具了一定的稳定性与可采集性,是一种良好的生物信号身份识别特征。

[0003] 目前的基于心电信号的身份识别方法,通常需要对心电信号的多个特征点进行检索,并根据特征点的位置进行特征提取与组合。由于不少基准点的定位较为模糊,定位较为困难,并且不同设备采集的心电信号中通常基准点的位置会有较大的变化,因此这些方法往往具有较差的泛化性,在不同的数据集上表现差异较大,身份识别的准确性也会受到非常大的影响。此外,对于未注册的用户,这些方法通常需要对模型重新进行训练,无法满足身份认证的实时性要求。

[0004] 综上所述,若要将基于心电信号的身份认证方法广泛应用到生活中,必须满足以下条件:(1)对不同设备采集的心电信号,需要有较高的泛化能力,保证识别准确度处在较为稳定的范围内;(2)对于新用户,身份认证模型无需进行重新训练。

[0005] 因此,本发明提出了一种高泛化性的、基于并行多尺度一维残差网络的心电信号身份认证方法,并且可以投入日常生活的实际应用。

发明内容

[0006] 本发明的目的在于克服现有技术的不足,提出一种可以适用于不同设备采集的心电信号,并且具有高精度与高泛化性的身份识别方法。

[0007] 为了实现上述目的,本发明采用了以下技术方案:

[0008] 一种高泛化性的心电信号身份认证方法包括以下步骤:

[0009] 1) 获取用户心电信号;

[0010] 2) 对用户心电信号进行预处理,提取固定长度心拍;

[0011] 3) 通过经由中心目标函数与边界目标函数训练的并行多尺度一维残差网络架构深度神经网络模型,从预处理后的固定长度心拍中提取心电信号特征向量;

[0012] 4) 将生成的用户心电信号特征向量与用户预先注册的心电信号特征模板进行匹配,根据二者的相似度确定匹配结果,完成认证。

[0013] 优选的,所述步骤2)中的预处理步骤具体为:

[0014] 1) 对所述短时间心电信号进行重采样至500Hz,并对重采样后的信号进行滤波,去

除噪声；

[0015] 2) 对预处理后的心电信号进行QRS检测,获取R波峰值点位置,取R波峰值点前99个采样点和R波峰值点后200个采样点,与R波峰值点一起组成固定长度为300个采样点的心拍。

[0016] 优选的,所述步骤3)中的深度神经网络模型具体为:

[0017] 1) 所述的神经网络模型采用并行多尺度一维残差网络架构,其中每一个尺度中残差块采用的卷积核的尺寸分别为 1×3 , 1×5 和 1×7 ;

[0018] 2) 所述的神经网络采用softmax目标函数,中心目标函数与边界目标函数的加权之和作为总的目标函数进行训练,目标函数形式如下:

$$[0019] \quad \mathcal{L} = \mathcal{L}_S + \lambda_C \cdot \mathcal{L}_C + \lambda_M \cdot \mathcal{L}_M$$

$$[0020] \quad \mathcal{L}_S = - \sum_{i=1}^m \log \frac{e^{W_{y_i}^T f(x_i) + b_{y_i}}}{\sum_{j=1}^n e^{W_j^T f(x_i) + b_j}}$$

$$[0021] \quad \mathcal{L}_C = \frac{1}{2} \sum_{i=1}^m \|f(x_i) - c_{y_i}\|_2^2$$

$$[0022] \quad \mathcal{L}_M = \sum_{i=1}^m \sum_a^K \sum_b^K [D_{ia}^P - D_{ib}^N + \alpha]$$

$$[0023] \quad D_{ia}^P = \|f(x_i) - f(x_{ia}^P)\|_2^2$$

$$[0024] \quad D_{ib}^N = \|f(x_i) - f(x_{ib}^N)\|_2^2$$

[0025] 其中, \mathcal{L} 为总的目标函数, \mathcal{L}_S 、 \mathcal{L}_C 和 \mathcal{L}_M 分别表示softmax目标函数、中心目标函数和边界目标函数; m 和 n 分别表示训练样本个数和训练样本类别数; x_i 和 y_i 分别表示第 i 个训练样本与第 i 个训练样本对应的类别; $f(x_i)$ 为第 i 个训练样本的心电信号特征向量; c_{y_i} 为所有类别为 y_i 的样本的平均值; K 为近邻搜索个数; x_{ia}^P 和 x_{ib}^N 分别表示与样本 x_i 类别相同的距离最远的第 a 个样本,和与样本 x_i 类别不同的距离最近的第 b 个样本; α 为距离常数; $W_{y_i}^T$ 表示softmax全连接层的权值矩阵的第 y_i 个列向量的转置, b_{y_i} 表示softmax全连接层的偏置向量的第 y_i 个值, W_j^T 表示softmax全连接层的权值矩阵的第 j 个列向量的转置, b_j 表示softmax全连接层的偏置向量的第 j 个值; λ_C 和 λ_M 分别表示中心目标函数和边界目标函数的权重;

[0026] 3) 所述的神经网络模型在训练时采用数据量大,覆盖面广,并且与测试数据集完全不同的数据集作为训练样本数据集。

[0027] 优选的,所述步骤4)中的匹配过程具体为:

[0028] 1) 指定的用户在认证前,已经预先注册过一段短时间心电信号并进行相同的预处理,采用训练完的深度神经网络模型,从这段心电信号中提取出心电信号特征向量,将其作为指定用户的已注册特征模板向量;

[0029] 2) 计算获得的指定用户心电信号特征向量与已注册特征模板向量之间的欧式距离,将欧式距离作为二者的匹配相似度;

[0030] 3) 将匹配相似度与预算设定的阈值进行对比,若相似度大于设定的阈值,则认为

二者互相匹配,通过认证;反之拒绝通过认证;

[0031] 4)若从获取的用户一段短时间的心电信号中能够提取多个心电信号特征向量,则采用投票机制进行决定,若能够通过匹配的特征向量个数大于特征向量总数的50%,则认为通过认证;反之拒绝通过认证。

[0032] 本发明的高泛化性的心电信号身份认证方法的优点是:

[0033] 1.该方法对心电信号中的基准点依赖少,只需要检测R波峰值点的位置,无需确定其他较难检测的基准点的位置。

[0034] 2.该方法采用了深度神经网络模型进行特征提取,与传统方法相比,深度神经网络能够提取更加高层次的抽象特征,从而有效地提高特征的准确性。

[0035] 3.该方法采用了并行多尺度一维残差网络架构,使用了三种尺寸不同的卷积核进行并行特征提取,能够对不同长度的心电信号分别提取特征,有效地保留了心电信号中包含的时序性,从而使得所提取的特征中包含了更为完整的心电信号信息,提高了整体身份识别的效果和准确性。

[0036] 4.该方法采用了中心目标函数与边界目标函数对深度神经网络进行训练,这两个目标函数保证了所提取特征类间离散性和类内聚合性,从而使得同类特征聚集得更加紧密,异类特征分散得更加开阔,在匹配时能够提高匹配得准确性,同时使得训练得到模型对训练数据的依赖程度降低,大大提高了模型的泛化性和鲁棒性。

[0037] 5.该方法在进行模型训练时,选择了多种不同的大规模公开心电信号数据集作为训练样本,覆盖了多种采样频率,多种采集环境,多种健康状况下的心电信号,能够更好地发挥所述深度神经网络模型架构的性能,增大所述目标函数的训练效果,从而大大提高了该方法对不同设备采集的数据的泛化性和鲁棒性。

附图说明

[0038] 图1是本发明提供的基于心电信号的身份认证流程图;

[0039] 图2是本发明提供的深度神经网络模型的结构图;

具体实施方式

[0040] 下面结合技术方案和附图对本发明进行进一步的说明。

[0041] 如图1所示,本发明的一种高泛化性的心电信号身份认证方法包括以下步骤:

[0042] 1)获取用户的设定时间心电信号;

[0043] 2)对用户心电信号进行预处理,提取固定长度心拍;

[0044] 3)通过经由中心目标函数与边界目标函数训练的并行多尺度一维残差网络架构深度神经网络模型,从预处理后的固定长度心拍中提取心电信号特征向量;

[0045] 4)将生成的用户心电信号特征向量与用户预先注册的心电信号特征模板进行匹配,根据二者的相似度确定匹配结果,完成认证。

[0046] 在本实施例中,所述的步骤2)中预处理步骤具体为:

[0047] 1)对所述短时间心电信号进行重采样至500Hz,并对重采样后的信号进行滤波,去除噪声;

[0048] 2)对预处理后的心电信号进行QRS检测,获取R波峰值点位置,取R波峰值点前99个

采样点和R波峰值点后200个采样点,与R波峰值点一起组成固定长度为300个采样点的心拍。

[0049] 在本实施例中,所述的步骤3)中深度神经网络模型结构如图2所示,具体内容如下:

[0050] 1)所述的神经网络模型采用并行多尺度一维残差网络架构,其中每一个尺度中残差块采用的卷积核的尺寸分别为 1×3 , 1×5 和 1×7 ;

[0051] 2)所述的神经网络采用softmax目标函数,中心目标函数与边界目标函数的加权之和作为总的目标函数进行训练,目标函数形式如下:

$$[0052] \quad \mathcal{L} = \mathcal{L}_S + \lambda_C \cdot \mathcal{L}_C + \lambda_M \cdot \mathcal{L}_M$$

$$[0053] \quad \mathcal{L}_S = -\sum_{i=1}^m \log \frac{e^{W_{y_i}^T f(x_i) + b_{y_i}}}{\sum_{j=1}^n e^{W_j^T f(x_i) + b_j}}$$

$$[0054] \quad \mathcal{L}_C = \frac{1}{2} \sum_{i=1}^m \|f(x_i) - c_{y_i}\|_2^2$$

$$[0055] \quad \mathcal{L}_M = \sum_{i=1}^m \sum_a^K \sum_b^K [D_{ia}^P - D_{ib}^N + \alpha]$$

$$[0056] \quad D_{ia}^P = \|f(x_i) - f(x_{ia}^P)\|_2^2$$

$$[0057] \quad D_{ib}^N = \|f(x_i) - f(x_{ib}^N)\|_2^2$$

[0058] 其中, \mathcal{L} 为总的目标函数, \mathcal{L}_S 、 \mathcal{L}_C 和 \mathcal{L}_M 分别表示softmax目标函数、中心目标函数和边界目标函数; m 和 n 分别表示训练样本个数和训练样本类别数; x_i 和 y_i 分别表示第 i 个训练样本与第 i 个训练样本对应的类别; $f(x_i)$ 为第 i 个训练样本的心电信号特征向量; c_{y_i} 为所有类别为 y_i 的样本的平均值; K 为近邻搜索个数; x_{ia}^P 和 x_{ib}^N 分别表示与样本 x_i 类别相同的距离最远的第 a 个样本,和与样本 x_i 类别不同的距离最近的第 b 个样本; α 为距离常数; $W_{y_i}^T$ 表示softmax全连接层的权值矩阵的第 y_i 个列向量的转置, b_{y_i} 表示softmax全连接层的偏置向量的第 y_i 个值, W_j^T 表示softmax全连接层的权值矩阵的第 j 个列向量的转置, b_j 表示softmax全连接层的偏置向量的第 j 个值; λ_C 和 λ_M 分别表示中心目标函数和边界目标函数的权重;

[0059] 3)所述的神经网络模型在训练时采用数据量大,覆盖面广,并且与测试数据集完全不同的数据集作为训练样本数据集。

[0060] 在本发明的一个优选实例中,所述的步骤4)中的匹配过程具体包括:

[0061] 1)指定的用户在认证前,已经预先注册过一段短时间心电信号并进行相同的预处理,采用训练完的深度神经网络模型,从这段心电信号中提取出心电信号特征向量,将其作为指定用户的已注册特征模板向量;

[0062] 2)计算获得的指定用户心电信号特征向量与已注册特征模板向量之间的欧式距离,将欧式距离作为二者的匹配相似度;

[0063] 3)将匹配相似度与预算设定的阈值进行对比,若相似度大于设定的阈值,则认为二者互相匹配,通过认证;反之拒绝通过认证;

[0064] 4)若从获取的用户一段短时间的心电信号中能够提取多个心电信号特征向量,则

采用投票机制进行决定,若能够通过匹配的特征向量个数大于特征向量总数的50%,则认为通过认证;反之则拒绝通过认证。

[0065] 本发明首先采用 1×3 , 1×5 和 1×7 的卷积核来构建并行多尺度一维残差网络架构的神经网络模型,进而使用公开的心电信号数据集:ECG-ID,PTB和PhysioNet/Cinc数据集作为训练集。对训练集中的每条心电信号,本发明检测心电信号中的全部R波峰值点,并取每个R波峰值点之前的99个采样点和之后的200个采样点,提取一个长度300个采样点的心拍。将所有心电信号中提出的心拍作为神经网络的训练样本,采用softmax目标函数、中心目标函数和边界目标函数对所构建的神经网络模型进行训练,获得一个完整的可用模型。

[0066] 在获得了已训练的神经网络模型后,本发明对与训练数据集的采样设备、采样环境均不同的MIT-BIHArrhythmia数据集上进行测试,首先对MIT-BIHArrhythmia数据集中的每一个用户进行注册,采用相同的预处理方法获得心拍,并通过已训练的神经网络模型获取相应的特征向量,将每一个用户50%的特征向量的平均值作为该用户的注册模板向量。最后将每个用户剩下50%的特征向量作为测试特征向量,计算每个测试特征向量与该用户注册模板向量之间的相似度并与阈值进行对比,采用投票机制对每个对比结果进行决策,确定二者是否互相匹配。本发明在所述MIT-BIHArrhythmia数据集中获得了95.99%的识别准确率,较论文^[1]中的91.1%的识别准确率提高了4.89%,具有更加优越的泛化性和准确性。[1]Zhao Z,Zhang Y,Deng Y,et al.ECG authentication system design incorporating a convolutional neural network and generalized S-Transformation[J].Computers in biology and medicine,2018,102:168-179。

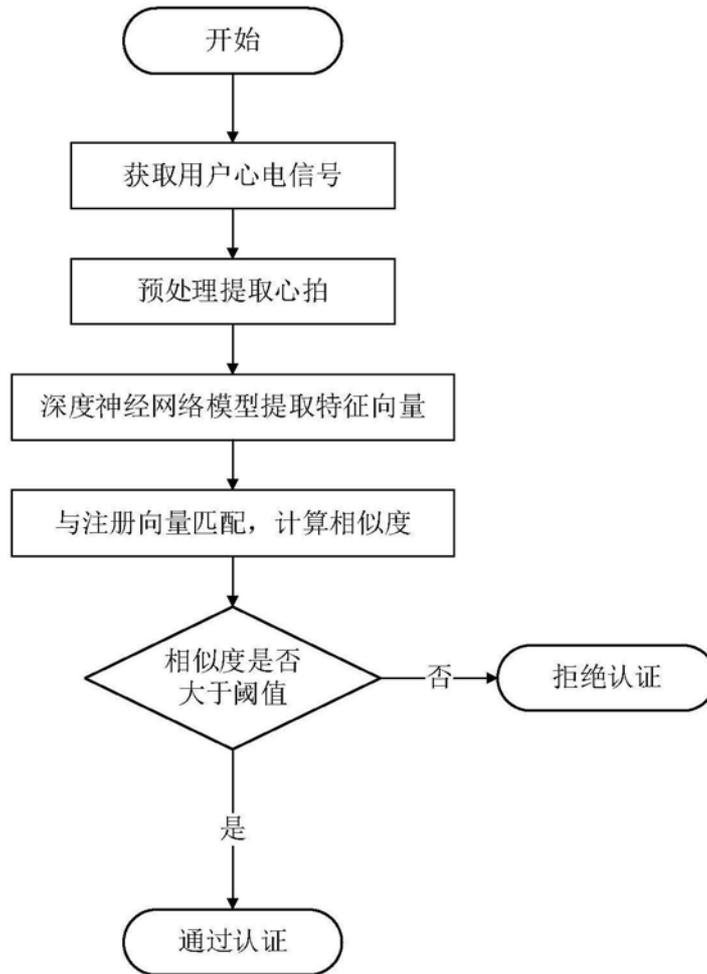


图1

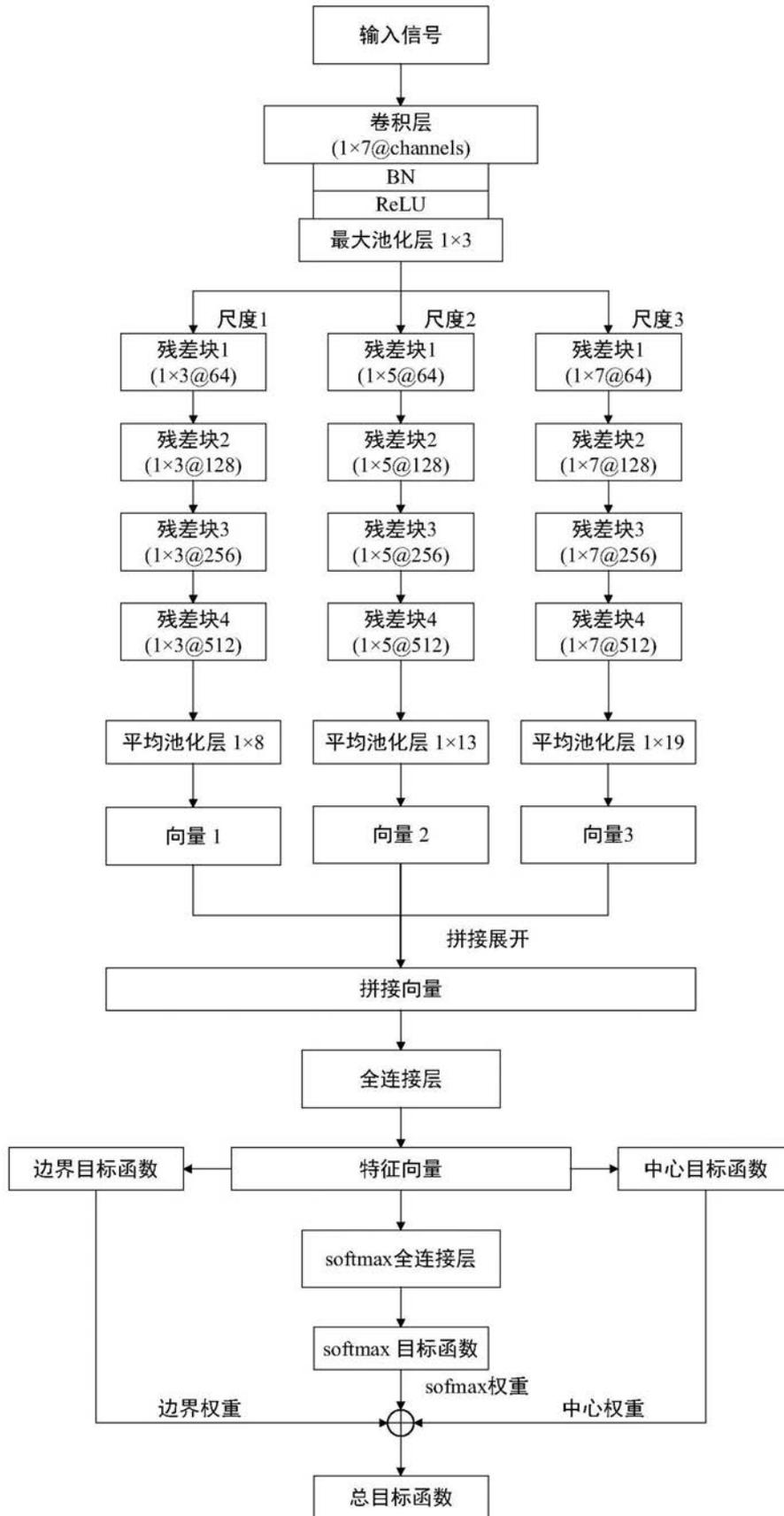


图2

| | | | |
|----------------|---|---------|------------|
| 专利名称(译) | 一种高泛化性的心电信号身份认证方法 | | |
| 公开(公告)号 | CN110008674A | 公开(公告)日 | 2019-07-12 |
| 申请号 | CN201910226702.4 | 申请日 | 2019-03-25 |
| [标]申请(专利权)人(译) | 浙江大学 | | |
| 申请(专利权)人(译) | 浙江大学 | | |
| 当前申请(专利权)人(译) | 浙江大学 | | |
| [标]发明人 | 褚逸凡 沈海斌 | | |
| 发明人 | 褚逸凡 沈海斌 | | |
| IPC分类号 | G06F21/32 A61B5/00 A61B5/0402 G06N3/04 G06N3/08 | | |
| CPC分类号 | A61B5/0402 A61B5/7225 G06F21/32 G06N3/0454 G06N3/08 | | |
| 代理人(译) | 郑海峰 | | |
| 外部链接 | Espacenet SIPO | | |

摘要(译)

本发明公开了一种高泛化性的心电信号身份认证方法，涉及生物信号身份认证领域。解决现有方法泛化性较差，且对于新用户，需进行重新训练的问题。方法采用了并行多尺度一维残差网络架构，使用了三种尺寸不同的卷积核进行并行特征提取，能够对不同长度的心电信号分别提取特征，有效地保留了心电信号中包含的时序性，提高了整体身份识别的效果和准确性。采用了中心目标函数与边界目标函数对神经网络进行训练，保证了所提取特征的类间离散性和类内聚合性，从而使得同类特征聚集得更加紧密，异类特征分散得更加开阔，在匹配时能够提高匹配得准确性，同时使得训练得到得模型对训练数据的依赖程度降低，大大提高了模型的泛化性和鲁棒性。

