



(12)发明专利申请

(10)申请公布号 CN 110558971 A

(43)申请公布日 2019.12.13

(21)申请号 201910713533.7

(22)申请日 2019.08.02

(71)申请人 苏州星空大海医疗科技有限公司

地址 215000 江苏省苏州市高新区珠江路
455号3号楼313-2

(72)发明人 杜晓薇

(74)专利代理机构 北京科家知识产权代理事务
所(普通合伙) 11427

代理人 张丽

(51)Int.Cl.

A61B 5/0402(2006.01)

A61B 5/00(2006.01)

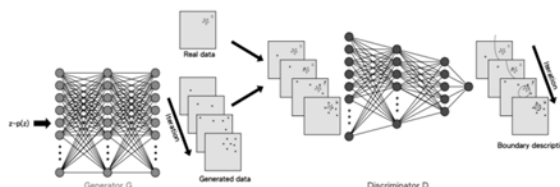
权利要求书2页 说明书6页 附图2页

(54)发明名称

基于单目标及多目标的生成对抗网络心电图异常检测方法

(57)摘要

本发明公开了一种基于单目标及多目标的生成对抗网络心电图异常检测方法,其将异常检测作为二元分类问题,不依赖于关于正常数据的假设并且需要较少的计算资源,在此基础上,本发明提出了一种新的异常检测算法SO-GAAL,它利用生成性广告-学习学习直接生成信息潜在异常值,解决了由“维数灾难”引起的信息缺失。本发明将SO-GAAL的网络架构从单个发生器扩展到具有不同目标的多个发生器(MO-GAAL),以防止单个发生器陷入模式崩溃问题。本发明为心电图异常检测提供了新的思路和解决方法,首次将对抗生成网络应用于心电图异常检测领域当中。



1.一种基于单目标的生成对抗网络心电图异常检测方法,其特征在于,所述方法包括如下步骤:

步骤1、数据预处理:

步骤1.1、对采集到的心电图信号数据集进行去噪处理:在去噪声处理中,普遍采用数字滤波技术实现去噪效果,即在算法处理部分对输入信号的噪声部分进行频率分析并去除这部分的信号,主要包括工频干扰、基线漂移、肌电干扰等干扰信号;

步骤1.2、将去噪处理后的图像用向量化表示,即将去噪后的心电信号图转化成神经网络可以处理的浮点数张量;

步骤1.3、将向量化后的数据进一步进行标准化处理,使浮点数据的值介于0-1范围内;

步骤2、生成器和判别器的构造:生成器和判别器均由多层神经网络构成,前者有多个输出值,后者仅有一个输出值;

步骤3、潜在异常点的生成:用以随机噪音pz为输入的生成器随机生成多个潜在异常点;

步骤4、优化判别器:要求判别器尽最大可能的将上一步骤获得的潜在异常点从真实数据中分别出来,即要求满足判别结果的损失函数最小:

$$\max_{\theta_d} V_d = \frac{1}{2n} \left[\sum_{i=1}^n \log(D(x^{(i)})) + \sum_{j=1}^k \sum_{i=1}^{n_j} \log(1 - D(G_j(z_j^{(i)}))) \right]$$

步骤5:重复步骤3至步骤5,直到判别器与生成器达到纳什均衡;

步骤6:判别器返回异常点的分数,即1-D(X)。

2.一种基于多目标的生成对抗网络心电图异常检测方法,其特征在于,所述方法包括如下步骤:

步骤1、数据预处理:

步骤1.1、对采集到的心电图信号数据集进行去噪处理:在去噪声处理中,普遍采用数字滤波技术实现去噪效果,即在算法处理部分对输入信号的噪声部分进行频率分析并去除这部分的信号,主要包括工频干扰、基线漂移、肌电干扰等干扰信号;

步骤1.2、将去噪处理后的图像用向量化表示,即将去噪后的心电信号图转化成神经网络可以处理的浮点数张量;

步骤1.3、将向量化后的数据进一步进行标准化处理,使浮点数据的值介于0-1范围内;

步骤2、生成器和判别器的构造:生成器和判别器均由多层神经网络构成,前者有多个输出值,后者仅有一个输出值;

步骤3、数据划分:将预处理后的数据集 $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{d \times n}$ 分割成k个子集 $X_{1:k}$,即将原始数据输入到判别器中,然后根据输出的判别结果,计算彼此间的相似性,最后将相似的用户划分在一起,将原数据平均分割成k份,满足 $\bigcup_{j=1}^k X_j = X$;

步骤4、潜在异常点的生成:用k个以随机噪音pz为输入的子生成器分别学习相应数据子集中数据的生成机制,即要求第j个子生成器 G_j 生成的异常点数据和原数据的第j个子集 X_j 的输出值相似;即满足:

$$\min_{\theta_j} V_{G_j} = -\frac{1}{n} \sum_{i=1}^n [T_j \log(D(G_j(z_j^{(i)}))) + (1 - T_j) \log(1 - D(G_j(z_j^{(i)})))]$$

步骤5、集成潜在异常点：将不同子生成器生成的数据潜在异常点集合集成在一起，为整个数据集提供一个合理的参考分布；

步骤6、优化判别器：要求判别器尽最大可能的将上一步骤获得的潜在异常点从真实数据中分别出来，即要求满足判别结果的损失函数最小；

$$\max_{\theta_d} V_d = \frac{1}{2n} \left[\sum_{i=1}^n \log(D(x^{(i)})) + \sum_{j=1}^k \sum_{i=1}^{n_j} \log(1 - D(G_j(z_j^{(i)}))) \right]$$

步骤7：重复步骤3至步骤6，直到判别器与生成器达到纳什均衡；

步骤8：判别器返回异常点的分数，即 $1-D(X)$ 。

基于单目标及多目标的生成对抗网络心电图异常检测方法

技术领域

[0001] 本发明涉及医疗心电图异常检测技术领域,具体为一种基于单目标及多目标的生成对抗网络心电图异常检测方法。

背景技术

[0002] 近年来,随着深度学习、模糊识别、神经网络等技术水平的提高,心电图的异常检测方法不断得到改进和发展。目前关于心电信号异常的识别分类方法包含滤波和阈值法、模板匹配法、小波变换法、数学形态学法、神经网络法等。然而这些方法都或多或少面临着许多的限制。譬如数学形态学法虽具有较高的精确性、严谨性和抗噪性,但它对数据的预处理有着较高的要求,并因为极大的计算量而很少在实际中应用。神经网络法则属于近年来新兴的应用于生物医学的技术手段,其超强的学习能力使得异常检测的精确率得到提高、判别时间缩短、抗噪性能也得到改进,但是在实际应用中,神经网络法仍处于探索阶段。

[0003] 关于异常点的检测方法,经典模型包括基于模型的检测算法和基于近邻的检测算法。其中基于模型的检测算法是直接为所有数据构造一个模型,然后基于数据与已构建的正常模式间的偏差计算离群值。基于近邻的检测算法则是度量数据点的稀有度,将远离其最近邻居的数据点判别为离群点。然而不同于以上两种经典模型,基于生成潜在离群点的检测算法则提出了新的检测思路。它首先通过生成潜在离群点来构造一个有标签的数据集,然后训练一个二分类器来对该数据集进行后续的检测。然而随着数据结构复杂度的增加,这些方法没办法为好的判别效果提供保障。

发明内容

[0004] 为了克服上述问题,本发明提供一种基于单目标及多目标的生成对抗网络心电图异常检测方法。

[0005] 本发明的技术方案是提供一种基于单目标的生成对抗网络心电图异常检测方法,其特征在于,所述方法包括如下步骤:

[0006] 步骤1、数据预处理:

[0007] 步骤1.1、对采集到的心电图信号数据集进行去噪处理:在去噪声处理中,普遍采用数字滤波技术实现去噪效果,即在算法处理部分对输入信号的噪声部分进行频率分析并去除这部分的信号,主要包括工频干扰、基线漂移、肌电干扰等干扰信号;

[0008] 步骤1.2、将去噪处理后的图像用向量化表示,即将去噪后的心电信号图转化成神经网络可以处理的浮点数张量;

[0009] 步骤1.3、将向量化后的数据进一步进行标准化处理,使浮点数据的值介于0-1范围内;

[0010] 步骤2、生成器和判别器的构造:生成器和判别器均由多层神经网络构成,前者有多个输出值,后者仅有一个输出值;

[0011] 步骤3、潜在异常点的生成:用以随机噪音 p_z 为输入的生成器随机生成多个潜在异

常点；

[0012] 步骤4、优化判别器：要求判别器尽最大可能的将上一步骤获得的潜在异常点从真实数据中分别出来，即要求满足判别结果的损失函数最小；

$$\max_{\theta_d} V_d = \frac{1}{2n} \left[\sum_{i=1}^n \log(D(x^{(i)})) + \sum_{j=1}^k \sum_{i=1}^{n_j} \log(1 - D(G_j(z_j^{(i)}))) \right]$$

[0013] 步骤5：重复步骤3至步骤5，直到判别器与生成器达到纳什均衡；

[0014] 步骤6：判别器返回异常点的分数，即1-D(X)。

[0015] 一种基于多目标的生成对抗网络心电图异常检测方法，其特征在于，所述方法包括如下步骤：

[0016] 步骤1、数据预处理：

[0017] 步骤1.1、对采集到的心电图信号数据集进行去噪处理：在去噪声处理中，普遍采用数字滤波技术实现去噪效果，即在算法处理部分对输入信号的噪声部分进行频率分析并去除这部分的信号，主要包括工频干扰、基线漂移、肌电干扰等干扰信号；

[0018] 步骤1.2、将去噪处理后的图像用向量化表示，即将去噪后的心电信号图转化成神经网络可以处理的浮点数张量；

[0019] 步骤1.3、将向量化后的数据进一步进行标准化处理，使浮点数据的值介于0-1范围内；

[0020] 步骤2、生成器和判别器的构造：生成器和判别器均由多层神经网络构成，前者有多个输出值，后者仅有一个输出值，结构如图3所示；

[0021] 步骤3、数据划分：将预处理后的数据集 $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{d \times n}$ 分割成k个子集 $X_{1:k}$ ，即将原始数据输入到判别器中，然后根据输出的判别结果，计算彼此间的相似性，最后将相似的用户划分在一起，将原数据平均分割成k份，满足 $\bigcup_{j=1}^k X_j = X$ ；

[0022] 步骤4、潜在异常点的生成：用k个以随机噪音 p_z 为输入的子生成器分别学习相应数据子集中数据的生成机制，即要求第j个子生成器 G_j 生成的异常点数据和原数据的第j个子集 X_j 的输出值相似；即满足：

$$\min_{\theta_j} V_{G_j} = -\frac{1}{n} \sum_{i=1}^n [T_j \log(D(G_j(z_j^{(i)}))) + (1 - T_j) \log(1 - D(G_j(z_j^{(i)})))]$$

[0023] 步骤5、集成潜在异常点：将不同子生成器生成的数据潜在异常点集合集成在一起，为整个数据集提供一个合理的参考分布；

[0024] 步骤6、优化判别器：要求判别器尽最大可能的将上一步骤获得的潜在异常点从真实数据中分别出来，即要求满足判别结果的损失函数最小；

$$\max_{\theta_d} V_d = \frac{1}{2n} \left[\sum_{i=1}^n \log(D(x^{(i)})) + \sum_{j=1}^k \sum_{i=1}^{n_j} \log(1 - D(G_j(z_j^{(i)}))) \right]$$

[0025] 步骤7：重复步骤3至步骤6，直到判别器与生成器达到纳什均衡；

[0026] 步骤8：判别器返回异常点的分数，即1-D(X)。

[0027] 心电图一直是临床上诊断各种心脏疾病的重要手段，并广泛应用于手术麻醉、危重病人抢救、医学观察等领域。对于具有潜在性的心律异常的病人实施心电监护，可以有效地降低心血管疾病的死亡率，然而如何借助大数据信息手段更方便、快捷、有效的识别出异

常心电图信号,对于患者来说有着重要的意义。

[0028] 在本发明中,我们首先提出了一种新颖的基于最近的生成对抗性学习框架的心电图异常检测方法,我们称之为单目标生成对抗主动学习(SO-GAAL)。具体来说,它是一个在生成器G与判别器D之间执行的最大最小博弈过程,这也可以被认为是我们模型中的主动学习过程。我们采用随机产生的噪声作为生成器的输入,并通过鉴别器的引导直接产生位于真实数据内部或附近的信息性潜在异常值。结果,SO-GAAL中的鉴别器可以通过刻画出将潜在异常值与实际数据分开的划分边界,从而识别出异常值。然而,SO-GAAL产生信息潜在异常值的最终目标是为整个数据集提供合理的参考分布。如果所有信息潜在异常值随着训练的进程而发生在部分实际数据内部或附近,这可以被识别为模式崩溃问题,则SO-GAAL可能获得错误的检测结果。因此,当潜在异常值提供足够的信息时,应确定训练的停止节点,这对于没有任何先验信息的SO-GAAL来说是极其困难的。为了克服SO-GAAL的这一缺点,我们又进一步提出了一种多目标生成对抗主动学习(MO-GAAL),它将网络架构从单个发生器扩展到具有不同目标的多个发生器。这通过为整个数据集生成多个参考分布的混合来迁移模式崩溃问题。

[0029] 本发明的有益效果是:本发明的一种基于单目标及多目标的生成对抗网络心电图异常检测方法具有如下有益效果:

[0030] 1、我们将异常检测作为二元分类问题,不依赖于关于正常数据的假设并且需要较少的计算资源。在此基础上,我们提出了一种新的异常检测算法SO-GAAL,它利用生成性广告-学习学习直接生成信息潜在异常值,解决了由“维数灾难”引起的信息缺失。

[0031] 2、我们将SO-GAAL的网络架构从单个发生器扩展到具有不同目标的多个发生器(MO-GAAL),以防止单个发生器陷入模式崩溃问题。

[0032] 本发明为心电图异常检测提供了新的思路和解决方法,也是首次将对抗生成网络应用于心电图异常检测领域当中。

附图说明

[0033] 图1为基于SO-GAAL的离群点检测机制的模型示意图;

[0034] 图2为MO-GAAL的优化过程示意图;

[0035] 图3为基于MO-GAAL的离群点检测机制的模型示意图;

[0036] 图4为MO-GAAL的优化过程示意图。

具体实施方式

[0037] 为了使本发明实现的技术手段、创作特征、达成目的与功效易于明白了解,下面结合具体实施例,进一步阐述本发明。

[0038] 如图1至图4所示,本发明的一种基于单目标的生成对抗网络心电图异常检测方法,其特征在于,所述方法包括如下步骤:

[0039] 步骤1、数据预处理:

[0040] 步骤1.1、对采集到的心电图信号数据集进行去噪处理:在去噪声处理中,普遍采用数字滤波技术实现去噪效果,即在算法处理部分对输入信号的噪声部分进行频率分析并去除这部分的信号,主要包括工频干扰、基线漂移、肌电干扰等干扰信号;

[0041] 步骤1.2、将去噪处理后的图像用向量化表示,即将去噪后的心电信号图转化成神经网络可以处理的浮点数张量;

[0042] 步骤1.3、将向量化后的数据进一步进行标准化处理,使浮点数据的值介于0-1范围内;

[0043] 步骤2、生成器和判别器的构造:生成器和判别器均由多层神经网络构成,前者有多个输出值,后者仅有一个输出值,结构如图1所示;

[0044] 步骤3、潜在异常点的生成:用以随机噪音pz为输入的生成器随机生成多个潜在异常点;

[0045] 步骤4、优化判别器:要求判别器尽最大可能的将上一步骤获得的潜在异常点从真实数据中分别出来,即要求满足判别结果的损失函数最小;

$$\max_{\theta_d} V_d = \frac{1}{2n} \left[\sum_{i=1}^n \log(D(x^{(i)})) + \sum_{j=1}^k \sum_{i=1}^{n_j} \log(1 - D(G_j(z_j^{(i)}))) \right]$$

[0046] 步骤5:重复步骤3至步骤5,直到判别器与生成器达到纳什均衡;

[0047] 步骤6:判别器返回异常点的分数,即1-D(X)。

[0048] 一种基于多目标的生成对抗网络心电图异常检测方法,其特征在于,所述方法包括如下步骤:

步骤1、数据预处理:

[0049] 步骤1.1、对采集到的心电图信号数据集进行去噪处理:在去噪声处理中,普遍采用数字滤波技术实现去噪效果,即在算法处理部分对输入信号的噪声部分进行频率分析并去除这部分的信号,主要包括工频干扰、基线漂移、肌电干扰等干扰信号;

[0050] 步骤1.2、将去噪处理后的图像用向量化表示,即将去噪后的心电信号图转化成神经网络可以处理的浮点数张量;

[0051] 步骤1.3、将向量化后的数据进一步进行标准化处理,使浮点数据的值介于0-1范围内;

[0052] 步骤2、生成器和判别器的构造:生成器和判别器均由多层神经网络构成,前者有多个输出值,后者仅有一个输出值,结构如图3所示;

[0053] 步骤3、数据划分:将预处理后的数据集 $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{d \times n}$ 分割成k个子集 $X_{1:k}$,即将原始数据输入到判别器中,然后根据输出的判别结果,计算彼此间的相似性,最后将相似的用户划分在一起,将原数据平均分割成k份,满足 $\bigcup_{j=1}^k X_j = X$;

[0054] 步骤4、潜在异常点的生成:用k个以随机噪音pz为输入的子生成器分别学习相应数据子集中数据的生成机制,即要求第j个子生成器 G_j 生成的异常点数据和原数据的第j个子集 X_j 的输出值相似;即满足:

$$\min_{\theta_j} V_{G_j} = -\frac{1}{n} \sum_{i=1}^n [T_j \log(D(G_j(z_j^{(i)}))) + (1 - T_j) \log(1 - D(G_j(z_j^{(i)})))]$$

[0055] 步骤5、集成潜在异常点:将不同子生成器生成的数据潜在异常点集合集成在一起,为整个数据集提供一个合理的参考分布;

[0056] 步骤6、优化判别器:要求判别器尽最大可能的将上一步骤获得的潜在异常点从真实数据中分别出来,即要求满足判别结果的损失函数最小;

$$\max_{\theta_d} V_d = \frac{1}{2n} \left[\sum_{i=1}^n \log(D(x^{(i)})) + \sum_{j=1}^k \sum_{i=1}^{n_j} \log(1 - D(G_j(z_j^{(i)}))) \right]$$

[0057] 步骤7:重复步骤3至步骤6,直到判别器与生成器达到纳什均衡;

[0058] 步骤8:判别器返回异常点的分数,即 $1-D(X)$ 。

[0059] 本发明提出的是一种基于单目标生成对抗网络的心电图异常检测方法模型,并在此基础上扩展成一种基于多目标生成对抗网络的心电图异常检测方法模型。不同于传统的心电异常检测的特征识别方法,本文提出的发明不仅是无监督的、可适用于高维数据的异常点检测方法,而且首次将生成潜在异常数据与判别潜在异常数据间的博弈应用到心电图异常检测领域。

[0060] 如图3所示,本实施例中,一种基于多目标生成对抗网络的心电图异常检测方法是按如下步骤进行:

[0061] 步骤1、数据预处理:

[0062] 步骤1.1、对采集到的心电图信号数据集进行去噪处理:在去噪声处理中,普遍采用数字滤波技术实现去噪效果,即在算法处理部分对输入信号的噪声部分进行频率分析并去除这部分的信号,主要包括工频干扰、基线漂移、肌电干扰等干扰信号;

[0063] 步骤1.2、将去噪处理后的图像用向量化表示:考虑到心电图数据是一种黑白波形图像,我们可以将去噪后的心电信号图转化成神经网络可以处理的二维向量;

[0064] 步骤1.3、对向量化后的数据进一步进行标准化处理,使其包含的浮点数据值介于0-1范围内;

[0065] 步骤2、生成器和判别器的构造:生成器和判别器均由多层神经网络构成,前者有多个输出值,后者仅有一个输出值,结构如图2所示;

[0066] 步骤2.1、生成器的构造:生成器的目的就是随机生成均匀的潜在异常点分布对原数据进行补充,并根据异常点的分布相对密度小于正常点的相对密度这一属性,更有效的识别出原有数据中的异常点。这里我们引入一个均匀的参考分布 μ ,并通过数据相对于参考分布的相对密度 $\rho(x)$ 来确定数据 x 的聚散度。

[0067] 步骤2.2、判别器的构造:判别器则是构造一个函数 $\xi(x) \in [0, 1]$ 表示数据点输出正确的标签,并最小化函数 $\xi(x)$ 的损失函数 L_ξ 。最优的函数 $\xi(x)$ 应该满足以下条件:

$$\xi(x | \rho(x) \geq \tau) \rightarrow 1$$

$$\xi(x | \rho(x) \leq \tau) \rightarrow 0$$

对于 $\forall x$:

$$\xi(x_i | \rho(x_i) \geq \tau) \geq \xi(x_i | \rho(x_i) \leq \tau)$$

即,当数据的相对密度 $\rho(x)$ 小于合理阈值 τ 时,数据有很大的概率为异常点,则函数 $\xi(x)$ 应为其输出一个趋近于0的值;反之亦然。损失函数 L_ξ 则如下所示:

$$L_\xi = -\frac{1}{n} \sum_{i=1}^n (c_n y_i \log(\xi(x_i)) + c_o (1 - y_i) \log(1 - \xi(x_i)))$$

其中, c_n 与 c_o 分别为正常数据和异常数据的误分类成本;

[0068] 步骤3、数据划分:将预处理后的数据集 $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{d \times n}$ 分割成 k 个子集 $X_{1:k}$,即将原始数据输入到判别器中,然后根据判别结果的相似性,将原数据分割成 k 份

$$\cup_{j=1}^k X_j = X ;$$

[0069] 步骤4、潜在异常点的生成：利用不同的子生成器分别对应学习不同数据子集中的数据生成机制，即要求第j个子生成器 G_j 生成的异常点数据和原数据的第j个子集 X_j 的判别式的输出值相似。即满足：

$$\min_{\theta_j} V_{Gj} = -\frac{1}{n} \sum_{i=1}^n [T_j \log(D(G_j(z_j^{(i)}))) + (1-T_j) \log(1-D(G_j(z_j^{(i)})))]$$

[0070] 步骤5、集成潜在异常点：将上述步骤获取到的k个子生成器生成的数据潜在异常点分布集成在一起，形成整个数据集的潜在异常点的参考分布；

[0071] 步骤6、优化判别器：要求判别器尽最大可能的将上一步骤获得的整个数据集的潜在异常点从真实数据中分别出来，即要求满足判别结果的损失函数最小；

$$\max_{\theta_d} V_d = \frac{1}{2n} [\sum_{i=1}^n \log(D(x^{(i)})) + \sum_{j=1}^k \sum_{i=1}^{n_j} \log(1-D(G_j(z_j^{(i)})))]$$

[0072] 步骤7：重复步骤3至步骤6，直到判别器与生成器达到纳什均衡；

[0073] 步骤8：判别器返回异常点的分数，即 $1-D(X)$ 。

[0074] 以上实施例仅为本发明其中的一种实施方式，其描述较为具体和详细，但并不能因此而理解为对本发明专利范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本发明构思的前提下，还可以做出若干变形和改进，这些都属于本发明的保护范围。因此，本发明的保护范围应以所附权利要求为准。

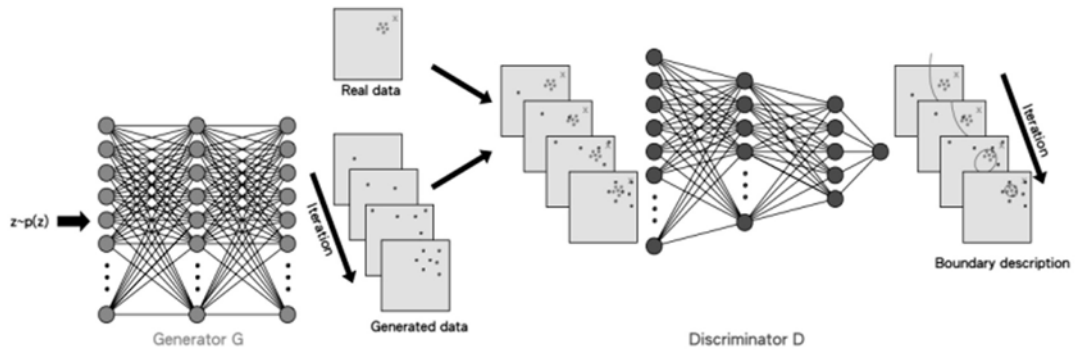


图1

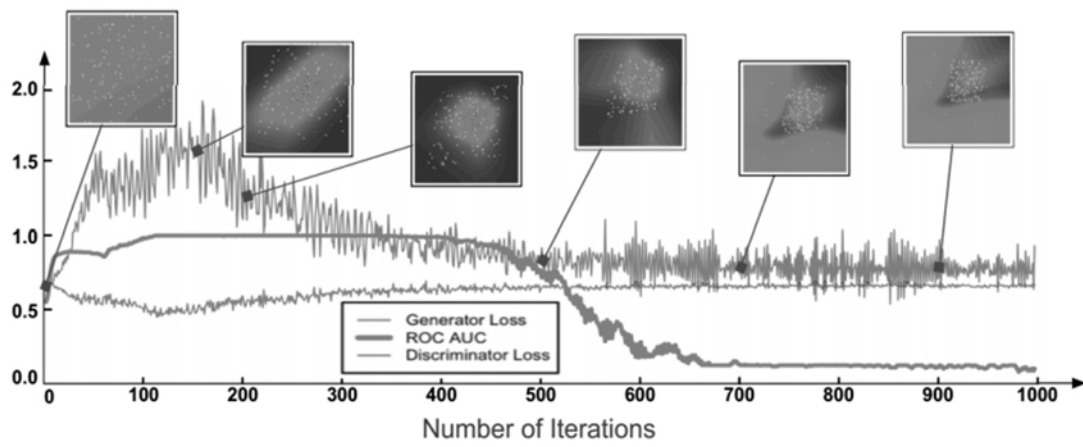


图2

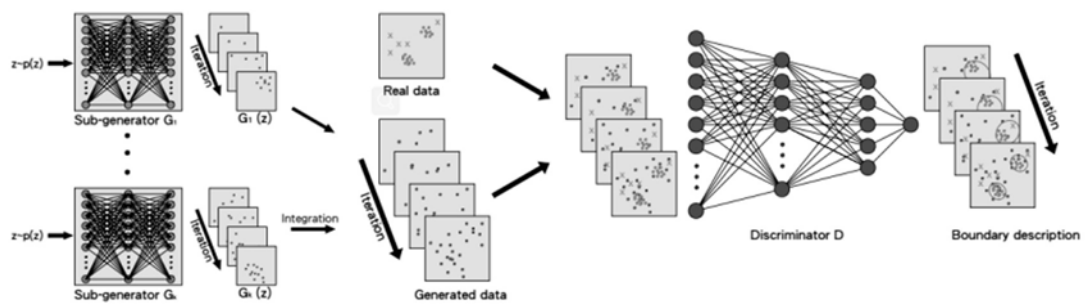


图3

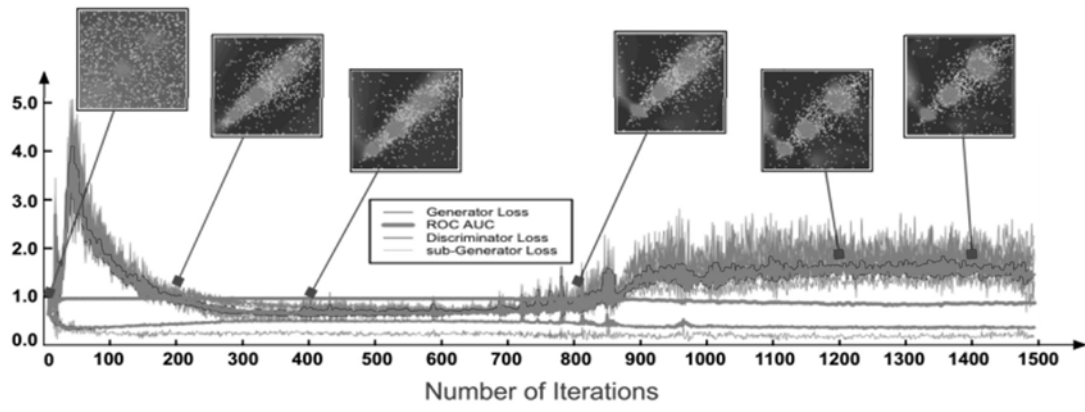


图4

专利名称(译)	基于单目标及多目标的生成对抗网络心电图异常检测方法		
公开(公告)号	CN110558971A	公开(公告)日	2019-12-13
申请号	CN201910713533.7	申请日	2019-08-02
[标]发明人	杜晓薇		
发明人	杜晓薇		
IPC分类号	A61B5/0402 A61B5/00		
CPC分类号	A61B5/0402 A61B5/7203 A61B5/7264 A61B5/7271		
代理人(译)	张丽		
外部链接	Espacenet SIPO		

摘要(译)

本发明公开了一种基于单目标及多目标的生成对抗网络心电图异常检测方法，其将异常检测作为二元分类问题，不依赖于关于正常数据的假设并且需要较少的计算资源，在此基础上，本发明提出了一种新的异常检测算法SO-GAAL，它利用生成性广告-学习学习直接生成信息潜在异常值，解决了由“维数灾难”引起的信息缺失。本发明将SO-GAAL的网络架构从单个发生器扩展到具有不同目标的多个发生器(MO-GAAL)，以防止单个发生器陷入模式崩溃问题。本发明为心电图异常检测提供了新的思路和解决方法，首次将对抗生成网络应用于心电图异常检测领域当中。

