



(12) 发明专利

(10) 授权公告号 CN 111310572 B

(45) 授权公告日 2023.05.05

(21) 申请号 202010052132.4

A61B 5/346 (2021.01)

(22) 申请日 2020.01.17

A61B 5/318 (2021.01)

G06F 18/24 (2023.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 111310572 A

(43) 申请公布日 2020.06.19

(73) 专利权人 上海乐普云智科技股份有限公司

地址 201612 上海市松江区莘砖公路668号  
A座16层

(72) 发明人 王斌 曹君

(74) 专利代理机构 北京慧诚智道知识产权代理

事务所(特殊普通合伙)  
11539

专利代理师 李楠

(51) Int. Cl.

G06F 18/2131 (2023.01)

A61B 5/00 (2006.01)

(56) 对比文件

CN 106725428 A, 2017.05.31

CN 107714023 A, 2018.02.23

CN 107832737 A, 2018.03.23

CN 107837082 A, 2018.03.27

CN 107951485 A, 2018.04.24

CN 107981858 A, 2018.05.04

CN 109871742 A, 2019.06.11

CN 110472229 A, 2019.11.19

US 2018350468 A1, 2018.12.06

张俊飞等.基于深度学习的ECG心拍数据分类设计.《自动化与仪器仪表》.2019,(第12期),第71-75页.

审查员 轩海珍

权利要求书2页 说明书7页 附图5页

(54) 发明名称

利用心搏时间序列生成心搏标签序列的处理方法和装置

(57) 摘要

本发明实施例涉及一种利用心搏时间序列生成心搏标签序列的处理方法和装置,方法包括:获取心搏时间序列;心搏时间序列包括多导联心搏数据;按照设定数据量对多导联心搏数据进行数据切割,得到多组心搏分析数据;将多组心搏分析数据进行数据组合,得到四维张量数据{B,H,W,C};对四维张量数据进行张量格式转换处理,将四维张量数据中的数据高度收缩为1,并对数据宽度进行压缩,输出为{B,1,W<sub>1</sub>,C<sub>1</sub>}的输出张量;对输出张量进行转换,得到特征张量{B,W<sub>1</sub>,C<sub>1</sub>};

将特征张量与随机初始化的权重矩阵相乘,输出嵌入特征张量{B,W<sub>1</sub>,d<sub>model</sub>};其中,d<sub>model</sub>为输入到Transformer模型的特征向量的维度;将嵌入特征张量输入到训练好的Transformer模型,输出心搏时间序列对应的心搏标签序列。

应的心搏标签序列。



CN 111310572 B

1. 一种利用心搏时间序列生成心搏标签序列的处理方法,其特征在于,所述处理方法包括:

获取心搏时间序列;所述心搏时间序列包括多导联心搏数据;

按照设定数据量对所述多导联心搏数据进行数据切割,得到多组心搏分析数据;

将所述多组心搏分析数据进行数据组合,得到四维张量数据;所述四维张量数据具有四个因子 $\{B,H,W,C\}$ ,其中因子B为批量数据、因子H为高度数据、因子W为宽度数据、因子C为通道数据;所述批量数据为所述多组心搏分析数据的组数;

对所述四维张量数据进行张量格式转换处理,将所述四维张量数据中的高度数据收缩为1,并对宽度数据进行压缩,输出为 $\{B,1,W_1,C_1\}$ 的输出张量;

对所述输出张量 $\{B,1,W_1,C_1\}$ 进行转换,得到特征张量 $\{B,W_1,C_1\}$ ;

将所述特征张量与随机初始化的权重矩阵 $\mathbf{E} \in \mathbb{R}^{W_1 \times d_{\text{model}} \times C_1}$ 相乘,输出嵌入特征张量 $\{B,W_1,d_{\text{model}}\}$ ;其中, $d_{\text{model}}$ 为输入到Transformer模型的特征向量的维度;

将所述嵌入特征张量输入到训练好的Transformer模型,输出所述心搏时间序列对应的心搏标签序列。

2. 根据权利要求1所述的处理方法,其特征在于,在所述将所述嵌入特征张量输入到训练好的Transformer模型之前,所述方法还包括:训练所述Transformer模型。

3. 根据权利要求2所述的处理方法,其特征在于,所述训练所述Transformer模型具体包括:

对作为训练样本的心搏时间序列进行心搏数据的数据标注;所述数据标注包括对心搏数据的心搏类型和心搏R点位置的标注;

按照设定采样频率和采样长度进行第一数据量的心搏片段提取;

在提取到的心搏片段中,根据所述数据标注确定所述心搏R点位置对应的心搏类型,得到神经网络机器翻译NMT标签序列;

对所述NMT标签序列进行整理,得到符合自然语言处理NLP模型语句要求的作为训练样本的心搏标签序列;

以作为训练样本的心搏时间序列和作为训练样本的心搏标签序列对Transformer模型进行训练。

4. 根据权利要求3所述的处理方法,其特征在于,所述对所述NMT标签序列进行整理具体包括:

确定所述心搏标签序列的字段长度;

在所述NMT标签序列的第一个字段之前添加标记“S”;

在所述NMT标签序列的最后一个字段之后添加标记“/S”;

根据所述字段长度,在所述标记“/S”之后的字段中填充标记“Pad”。

5. 根据权利要求3或4所述的处理方法,其特征在于,所述以作为训练样本的心搏时间序列和作为训练样本的心搏标签序列对Transformer模型进行训练具体包括:

得到所述作为训练样本的心搏时间序列的训练样本的嵌入特征张量 $\{B,W_1,d_{\text{model}}\}$ ;

将所述训练样本的嵌入特征张量 $\{B,W_1,d_{\text{model}}\}$ ,和,数据标注得到NMT标签序列作为训练样本输入数据,将所述整理得到的训练样本的心搏标签序列作为训练样本输出数据,对所述Transformer模型进行训练。

6. 根据权利要求1所述的处理方法,其特征在于,所述对所述四维张量数据进行张量格式转换处理,将所述四维张量数据中的高度数据收缩为1,并对宽度数据进行压缩,输出为 $\{B, 1, W_1, C_1\}$ 的输出张量具体为:

设定多导联心搏数据的导联数量为所述四维张量数据的高度数据;

按照设定步幅,对所述四维张量数据使用CNN卷积神经网络进行多层网络卷积计算,得到高度数据收缩为1且宽度数据被压缩的输出张量。

7. 根据权利要求1所述的处理方法,其特征在于,所述Transformer模型为基于注意力机制,采用了编码器-译码器架构的模型。

8. 一种设备,包括存储器和处理器,其特征在于,所述存储器用于存储程序,所述处理器用于执行权利要求1至7任一项所述的方法。

9. 一种计算机可读存储介质,包括指令,当所述指令在计算机上运行时,使所述计算机执行权利要求1至7任一项所述的方法。

## 利用心搏时间序列生成心搏标签序列的处理方法和装置

### 技术领域

[0001] 本发明涉及数据处理技术领域,尤其涉及一种利用心搏时间序列生成心搏标签序列的处理方法和装置。

### 背景技术

[0002] 心血管疾病是威胁人类健康的主要疾病之一,利用有效的手段对心血管疾病进行检测是目前全世界关注的重要课题。

[0003] 心电图(ECG)是现代医学中诊断心血管疾病的主要方法,利用ECG诊断各种心血管疾病,本质上就是提取ECG的特征数据对ECG进行分类的过程。专家医生在心电图的阅读分析过程中,都是需要同时比较各个导联(单导数据除外)的信号在时间顺序上的变化,导联之间的相关性(空间关系)和变异,然后才能够做出一个比较准确的判断。而这种依赖于医生经验的方式,准确率无法得到保障。

[0004] 随着科技的进步,利用计算机对ECG进行自动准确的分析已经得到了快速的发展。但是,虽然市场上大多数的心电图分析软件都可以对数据进行自动分析,但由于心电图信号本身的复杂与变异性,目前自动分析软件的准确率远远不够,无法达到临床分析使用的要求。

### 发明内容

[0005] 本发明的目的是针对现有技术的缺陷,提供一种利用心搏时间序列生成心搏标签序列的处理方法。本方法通过将心搏时间序列建模为自然语言中的“源语句”,将心搏时间序列的标签序列建模为“目标语句”,对Transformer模型进行改进训练,利用训练后的模型对基于心搏时间序列处理转换得到的嵌入特征张量进行处理,输出心搏标签序列。

[0006] 为实现上述目的,第一方面,本发明提供了一种利用心搏时间序列生成心搏标签序列的处理方法,包括:

[0007] 获取心搏时间序列;所述心搏时间序列包括多导联心搏数据;

[0008] 按照设定数据量对所述多导联心搏数据进行数据切割,得到多组心搏分析数据;

[0009] 将所述多组心搏分析数据进行数据组合,得到四维张量数据;所述四维张量数据具有四个因子{B,H,W,C},其中因子B为批量数据、因子H为高度数据、因子W为宽度数据、因子C为通道数据;所述批量数据为所述多组心搏分析数据的组数;

[0010] 对所述四维张量数据进行张量格式转换处理,将所述四维张量数据中的高度数据收缩为1,并对宽度数据进行压缩,输出为{B,1,W<sub>1</sub>,C<sub>1</sub>}的输出张量;

[0011] 对所述输出张量进行转换,得到特征张量{B,W<sub>1</sub>,C<sub>1</sub>};

[0012] 将所述特征张量与随机初始化的权重矩阵 $E \in R^{W_1 \times d_{model} \times C_1}$ 相乘,输出嵌入特征张量{B,W<sub>1</sub>,d<sub>model</sub>} ;其中,d<sub>model</sub>为输入到Transformer模型的特征向量的维度;

[0013] 将所述嵌入特征张量输入到训练好的Transformer模型,输出所述心搏时间序列对应的心搏标签序列。

[0014] 优选的,在所述将所述嵌入特征张量输入到训练好的Transformer模型之前,所述方法还包括:训练所述Transformer模型。

[0015] 进一步优选的,所述训练所述Transformer模型具体包括:

[0016] 对作为训练样本的心搏时间序列进行心搏数据的数据标注;所述数据标注包括对心搏数据的心搏类型和心搏R点位置的标注;

[0017] 按照设定采样频率和采样长度进行第一数据量的心搏片段提取;

[0018] 在提取到的心搏片段中,根据所述数据标注确定所述心搏R点位置对应的心搏类型,得到神经网络机器翻译(Neural Machine Translation,NMT)标签序列;

[0019] 对所述NMT标签序列进行整理,得到符合自然语言处理(Natural Language Processing,NLP)模型语句要求的作为训练样本的心搏标签序列;

[0020] 以作为训练样本的心搏时间序列和作为训练样本的心搏标签序列对Transformer模型进行训练。

[0021] 进一步优选的,所述对所述NMT标签序列进行整理具体包括:

[0022] 确定所述心搏标签序列的字段长度;

[0023] 在所述NMT标签序列的第一个字段之前添加标记“S”;

[0024] 在所述NMT标签序列的最后一个字段之后添加标记“/S”;

[0025] 根据所述字段长度,在所述标记“/S”之后的字段中填充标记“Pad”。

[0026] 进一步优选的,所述以作为训练样本的心搏时间序列和作为训练样本的心搏标签序列对Transformer模型进行训练具体包括:

[0027] 对所述作为训练样本的心搏时间序列按照上述权利要求1所述方法得到所述户作为训练样本的心搏时间序列的训练样本的嵌入特征张量 $\{B, W_1, d_{\text{model}}\}$ ;

[0028] 将所述训练样本的嵌入特征张量 $\{B, W_1, d_{\text{model}}\}$ ,和,数据标注得到NMT标签序列作为训练样本输入数据,将所述整理得到的训练样本的心搏标签序列作为训练样本输出数据,对所述Transformer模型进行训练。

[0029] 优选的,所述对所述四维张量数据进行张量格式转换处理,将所述四维张量数据中的高度数据收缩为1,并对宽度数据进行压缩,输出为 $\{B, 1, W_1, C_1\}$ 的输出张量具体为:

[0030] 设定多导联心搏数据的导联数量为所述四维张量数据的高度数据;

[0031] 按照设定步幅,对所述四维张量数据使用CNN卷积神经网络进行多层网络卷积计算,得到高度数据收缩为1且宽度数据被压缩的输出张量。

[0032] 优选的,所述Transformer模型为基于注意力机制,采用了编码器-译码器架构的模型。

[0033] 本发明实施例提供的利用心搏时间序列生成心搏标签序列的处理方法。本方法通过将心搏时间序列建模为自然语言中的“源语句”,将心搏时间序列的标签序列建模为“目标语句”,对Transformer模型进行改进训练,利用训练后的模型对基于心搏时间序列处理转换得到的嵌入特征张量进行处理,输出心搏标签序列。

[0034] 第二方面,本发明实施例提供了一种设备,该设备包括存储器和处理器,存储器用于存储程序,处理器用于执行第一方面及第一方面的各实现方式中的方法。

[0035] 第三方面,本发明实施例提供了一种包含指令的计算机程序产品,当计算机程序产品在计算机上运行时,使得计算机执行第一方面及第一方面的各实现方式中的方法。

[0036] 第四方面,本发明实施例提供了一种计算机可读存储介质,计算机可读存储介质上存储有计算机程序,计算机程序被处理器执行时实现第一方面及第一方面的各实现方式中的方法。

### 附图说明

[0037] 图1为本发明实施例提供的利用心搏时间序列生成心搏标签序列的数据处理系统结构示意图;

[0038] 图2为本发明实施例提供的利用心搏时间序列生成心搏标签序列的处理方法流程图;

[0039] 图3为本发明实施例提供的Transformer模型的训练方法流程图;

[0040] 图4为本发明实施例提供的初步特征提取CNN模块示例图;

[0041] 图5为本发明实施例提供的Transformer模型结构示意图;

[0042] 图6为本发明实施例提供的一种设备结构示意图。

### 具体实施方式

[0043] 下面通过附图和实施例,对本发明的技术方案做进一步的详细描述。

[0044] 本发明实施例提供的利用心搏时间序列生成心搏标签序列的处理方法,可以用于心搏标签序列的生成。心律失常往往是序列性的改变,虽然每个心搏的定位、定性是分析的基点,但从序列层面统揽全局才能做到全面,准确。诸如文氏现象、干扰性分离、并行收缩、传出阻滞等,绝非是基于少数几个心搏可以作出诊断的。因此形成心搏标签序列对于心电分析是非常有意义且必要的。

[0045] 图1为本发明实施例提供的本发明实施例提供的利用心搏时间序列生成心搏标签序列的数据处理系统结构示意图;本发明的处理方法通过图1所示的系统结构来实现。

[0046] 图1所示的系统结构中,输入数据为心搏时间序列,包括多导联心搏数据,通过心搏时间序列预处理模块进行数据切割、组合、得到四维张量数据,然后通过初步特征提取模块,得到高度数据收缩为1的输出张量;通过后处理模块得到嵌入特征张量,其中包括Transformer模型的特征向量的维度;最后通过Transformer模型,输出心搏时间序列对应的心搏标签序列。

[0047] 初步特征提取模块的作用在于进行数据隔离和格式转换,便于输入不同格式的数据,连接后续不同的模型,为后续模型统一接口的格式。

[0048] 图2为本发明实施例提供的利用心搏时间序列生成心搏标签序列的处理方法流程图,下面结合图2,对本发明实施例提供的利用心搏时间序列生成心搏标签序列的处理方法进行说明。

[0049] 根据图2本发明上述处理方法的主要步骤包括:

[0050] 步骤110,获取心搏时间序列;

[0051] 其中,心搏时间序列包括多导联心搏数据;

[0052] 具体的,导联心搏数据是指各个导联的心搏数据,各导联心搏数据的获取方法可以根据申请人在先申请的专利201711203259.6《基于人工智能自学习的心电图自动分析方法和装置》中步骤100-步骤120的方法获得。

[0053] 步骤120,按照设定数据量对多导联心搏数据进行数据切割,得到多组心搏分析数据;

[0054] 具体的,对心搏时间序列,以设定数据量对全部的导联心搏数据进行切割生成导联的心搏分析数据。切割得到每组心搏分析数据中都包括多个导联的数据。切割定长的心搏时间序列时,不需要将某个R波位于整个时间序列的中心。此步骤由心搏时间序列预处理模块执行。

[0055] 步骤130,将多组心搏分析数据进行数据组合,得到四维张量数据;

[0056] 具体的,四维张量数据具有四个因子 $\{B,H,W,C\}$ ,其中因子B为批量数据、因子H为高度数据、因子W为宽度数据、因子C为通道数据;批量数据为多组心搏分析数据的组数。此步骤由心搏时间序列预处理模块执行。

[0057] 步骤140,对四维张量数据进行张量格式转换处理,将四维张量数据中的高度数据收缩为1,并对宽度数据进行压缩,输出为 $\{B,1,W_1,C_1\}$ 的输出张量;

[0058] 具体的,此步骤由初步特征提取模块执行。初步特征提取模块中可以包含卷积运算,也可以使用傅里叶变换、小波变换等频域特征提取方法。初步特征提取模块能够进行初步的特征提取和输入张量的维度调整。维度调整具有两个作用:

[0059] (1)为了使得ECG分类网络能够支持多种输入张量数据,以及支持单导联数据和多导联数据,消除输入变化对后续模型的影响,经过调整之后,输出张量的格式为 $\{B,1,W_1,C_1\}$ ,高度数据压缩为1,保证张量可以与后续transformer网络匹配。

[0060] (2)通过初步特征提取模块可以缩短心搏时间序列的长度。通过缩短心搏时间序列数据长度,可以有效提高整个模型的性能。

[0061] 下面给出了初步特征提取模块的一种实现方式,卷积神经网络(Convolutional Neural Networks,CNN)方式。

[0062] 设定多导联心搏数据的导联数量为四维张量数据的高度数据;按照设定步幅,对四维张量数据使用CNN进行多层网络卷积计算,得到高度数据收缩为1且宽度数据被压缩的输出张量。

[0063] 在具体的执行过程中:

[0064] 将导联数量4作为高度数据,数据量大小是1000个心电图电压值,设输入数据张量尺寸 $\{B,H,W,C\}$ 为 $\{128,4,1000,1\}$ 。那么,初步特征提取模块可以设计为如图4所示的三层CNN模块结构。

[0065] 第一层网络,CNN卷积核大小为 $3 \times 3$ ,卷积核数量为16,步幅为 $[2,2]$ 。CNN之后连接批归一化和Relu模块。网络的输出为 $[128,2,500,16]$ 。

[0066] 第二层网络,CNN卷积核大小为 $3 \times 3$ ,卷积核数量为32,步幅为 $[1,1]$ 。CNN之后连接批归一化和Relu模块。网络的输出为 $[128,2,500,32]$ 。

[0067] 第三层网络,CNN卷积核大小为 $3 \times 3$ ,卷积核数量为32,步幅为 $[2,2]$ 。CNN之后连接批归一化和Relu模块。网络的输出为 $[128,1,250,32]$ 。

[0068] 其中,步幅为卷积核执行卷积运算时每次移动的数量。步幅为2的效果是卷积计算输出的高度和宽度均减半,从而达到维度调整的目的。

[0069] 经过初步特征提取CNN模块之后,高度数据压缩为1,保证了张量可以与后续transformer网络匹配。时间序列长度压缩为250,有利于网络训练性能的提高。

[0070] 步骤150,对输出张量  $\{B, 1, W_1, C_1\}$  进行转换,得到特征张量  $\{B, W_1, C_1\}$ ;

[0071] 在本步骤的转换过程中,将压缩为1的高度数据去除。

[0072] 步骤160,将特征张量与随机初始化的权重矩阵  $E \in R^{W_1 \times d_{model} \times C_1}$  相乘,输出嵌入特征张量  $\{B, W_1, d_{model}\}$ ;

[0073] 其中,  $d_{model}$  为输入到Transformer模型的特征向量的维度;

[0074] 上述步骤150和步骤160由后处理模块执行,将初步特征提取模块输出的  $\{B, 1, W_1, C_1\}$  输出张量,转为  $\{B, W_1, C_1\}$  特征张量,并与随机初始化的权重矩阵  $E \in R^{W_1 \times d_{model} \times C_1}$  相乘,其中,  $d_{model}$  为输入到Transformer模型的特征向量的维度。输出嵌入特征张量  $\{B, W_1, d_{model}\}$ 。

[0075] 步骤170,将嵌入特征张量输入到训练好的Transformer模型,输出心搏时间序列对应的心搏标签序列。

[0076] 具体的,Transformer模型基于注意力机制,采用了编码器-译码器架构的神经网络模型。如图5所示,图中左半部分框图为编码器(Encoder)模块,右半部分框图为解码器(Decoder)模块。基于注意力机制的神经网络模型主要有两个优势:(1)避免使用循环神经网络,从而使得训练得以并行化;(2)注意力机制,获得长距离的记忆能力。

[0077] 其中,编码器模块中包含多个相同的层重复堆叠,每层包含两个子层:多头自注意力子层(multi-head attention layer or self-attention layer)和一个位置前馈层(feed forward layer)。两层之间通过残差和层标准化(layer norm)连接。

[0078] 解码器模块使用与编码器类似的层架构。不同之处在于解码器层中每层包含两个注意力子层。除了多头自注意力子层,还包括多头编码器注意力子层。层与层之间通过残差和层标准化连接。

[0079] 在本发明的具体实现中,对于Transformer模型进行了改进。

[0080] 对于常规的Transformer模型,在编码器(Encoder)模块之前,需先对数据进行位置编码。Transformer模型中没有循环网络结构,为了提供序列的位置信息,需要使用位置编码保留每个“词”,对本专利而言是心搏标签,的位置信息。

[0081] 进行位置编码有多种方式,可以使用参数学习策略,也可以使用固定的参数,本发明使用了固定位置编码。使用不同频率的正弦、余弦函数来生成位置向量,公式如下:

$$[0082] \quad PE_{(pos, 2i)} = \sin(pos/10000^{2i/d_{model}})$$

$$[0083] \quad PE_{(pos, 2i+1)} = \cos(pos/10000^{2i/d_{model}})$$

[0084] 其中pos表示序列中词的位置,  $i$  表示位置向量中词语编码的维度;

[0085]  $PE_{(pos, 2i)}$  表示偶数位置的词,  $PE_{(pos, 2i+1)}$  表示奇数位置的词,通过将偶数位置和奇数位置的词分别用正弦函数和余弦函数编码,因此每个词语就带上了相对的位置信息。

[0086] 在本发明的具体实现中,与常规的Transformer模型不同,由于编码器的输入为多导联心搏数据的嵌入特征张量,本身就是时间序列含有位置信息,因此,在编码器之前不需进行位置编码。

[0087] 最后,使用集束搜索集束搜索(Beam Search)算法计算得到用于输入观测序列的心搏标签序列。

[0088] 本发明首次将Transformer模型用于心搏分类领域,也对Transformer模型进行了相应的改进,在应用改进的Transformer模型执行上述流程之前,首先进行Transformer模型训练,模型的训练方法步骤如图3所示,具体如下:

[0089] 步骤210,对作为训练样本的心搏时间序列进行心搏数据的数据标注;

[0090] 其中,心搏时间序列中的心搏数据的长度可以是1秒到60秒。数据标注包括对心搏数据的心搏类型和心搏R点位置的标注。

[0091] 步骤220,按照设定采样频率和采样长度进行第一数据量的心搏片段提取;

[0092] 步骤230,在提取到的心搏片段中,根据数据标注确定心搏R点位置对应的心搏类型,得到神经网络机器翻译(Neural Machine Translation,NMT)标签序列;

[0093] 步骤240,对NMT标签序列进行整理,得到符合自然语言处理(Natural Language Processing,NLP)模型语句要求的作为训练样本的心搏标签序列;

[0094] 具体的,对NMT标签序列进行整理具体包括:

[0095] 确定心搏标签序列的字段长度;

[0096] 在NMT标签序列的第一个字段之前添加标记“S”;

[0097] 在NMT标签序列的最后一个字段之后添加标记“/S”;

[0098] 根据字段长度,在标记“/S”之后的字段中填充标记“Pad”。

[0099] 步骤250,以作为训练样本的心搏时间序列和作为训练样本的心搏标签序列对Transformer模型进行训练。

[0100] 具体的,对作为训练样本的心搏时间序列按照上述步骤120-步骤160方法得到户作为训练样本的心搏时间序列的训练样本的嵌入特征张量 $\{B, W_1, d_{model}\}$ ;

[0101] 将训练样本的嵌入特征张量 $\{B, W_1, d_{model}\}$ ,和,数据标注得到NMT标签序列作为训练样本输入数据,将整理得到的训练样本的心搏标签序列作为训练样本输出数据,对Transformer模型进行训练。

[0102] 对于得到训练样本的嵌入特征张量 $\{B, W_1, d_{model}\}$ 的方法在上述步骤120-步骤160中已经说明,下面,以一个具体的例子说明怎样通过数据标注得到NMT标签序列,以及如何整理得到的训练样本的心搏标签序列作为训练样本输出数据。

[0103] 以采样率是200Hz,5s为采样长度,取得设定数据量大小是1000个的心电图电压值的一个片段。

[0104] 此时得到的心搏片段中的数据标注结果可以表示为:

[0105]

心搏类型	N	V	N	N	N
心搏R点位置	112	267	523	724	909

[0106] 其中,N为窦性心搏,V表示室性早搏。

[0107] 在NMT标签序列中,只保留类型信息,得到心搏标签序列如下:

[0108]

心搏类型	N	V	N	N	N
------	---	---	---	---	---

[0109] 该序列就是数据标注得到的作为训练样本的NMT标签序列。

[0110] 根据上述步骤240的规则对NMT标签序列进行整理,得到作为训练样本输出数据的心搏标签序列如下:

[0111]

S	N	V	N	N	N	/S	Pad								
---	---	---	---	---	---	----	-----	-----	-----	-----	-----	-----	-----	-----	-----

[0112] 本发明实施例提供的利用心搏时间序列生成心搏标签序列的处理方法。本方法通过将心搏时间序列建模为自然语言中的“源语句”，将心搏时间序列的标签序列建模为“目标语句”，对Transformer模型进行改进训练，利用训练后的模型对基于心搏时间序列处理转换得到的嵌入特征张量进行处理，输出心搏标签序列。

[0113] 图6为本发明实施例提供的一种设备结构示意图，该设备包括：处理器和存储器。存储器可通过总线与处理器连接。存储器可以是非易失存储器，例如硬盘驱动器和闪存，存储器中存储有软件程序和设备驱动程序。软件程序能够执行本发明实施例提供的上述方法的各种功能；设备驱动程序可以是网络和接口驱动程序。处理器用于执行软件程序，该软件程序被执行时，能够实现本发明实施例提供的方法。

[0114] 需要说明的是，本发明实施例还提供了一种计算机可读存储介质。该计算机可读存储介质上存储有计算机程序，该计算机程序被处理器执行时，能够实现本发明实施例提供的方法。

[0115] 本发明实施例还提供了一种包含指令的计算机程序产品。当该计算机程序产品在计算机上运行时，使得处理器执行上述方法。

[0116] 专业人员应该还可以进一步意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、计算机软件或者二者的结合来实现，为了清楚地说明硬件和软件的可互换性，在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本发明的范围。

[0117] 结合本文中所公开的实施例描述的方法或算法的步骤可以用硬件、处理器执行的软件模块，或者二者的结合来实施。软件模块可以置于随机存储器(RAM)、内存、只读存储器(ROM)、电可编程ROM、电可擦除可编程ROM、寄存器、硬盘、可移动磁盘、CD-ROM、或技术领域内所公知的任意其它形式的存储介质中。

[0118] 以上所述的具体实施方式，对本发明的目的、技术方案和有益效果进行了进一步详细说明，所应理解的是，以上所述仅为本发明的具体实施方式而已，并不用于限定本发明的保护范围，凡在本发明的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

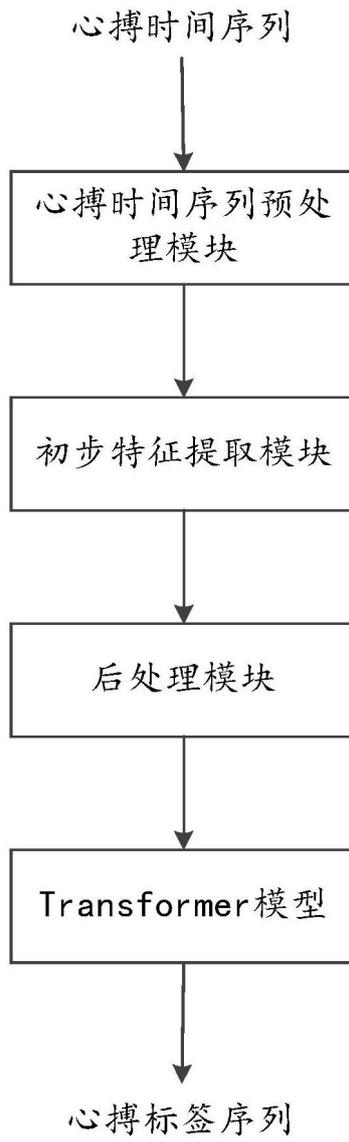


图1

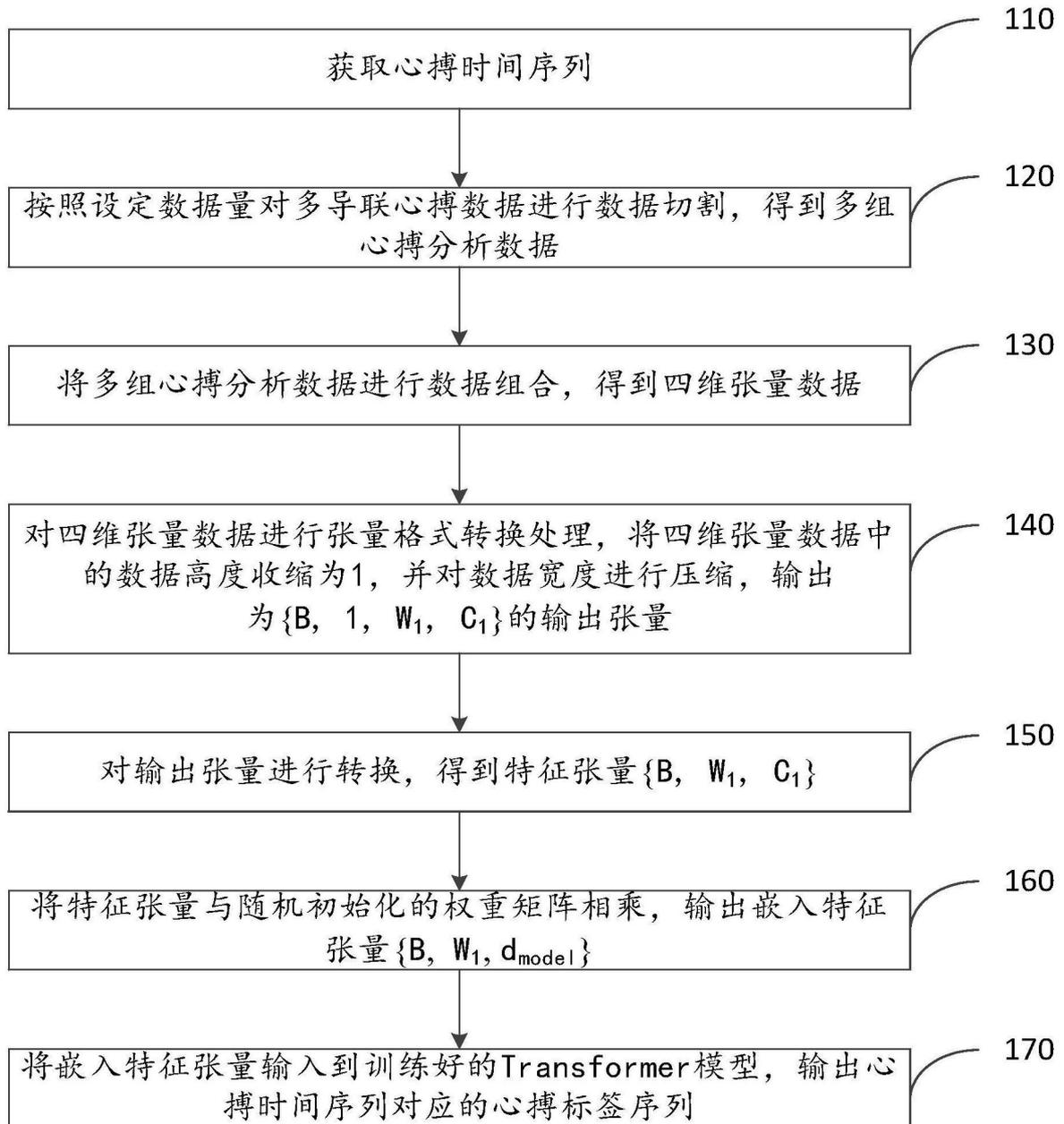


图2

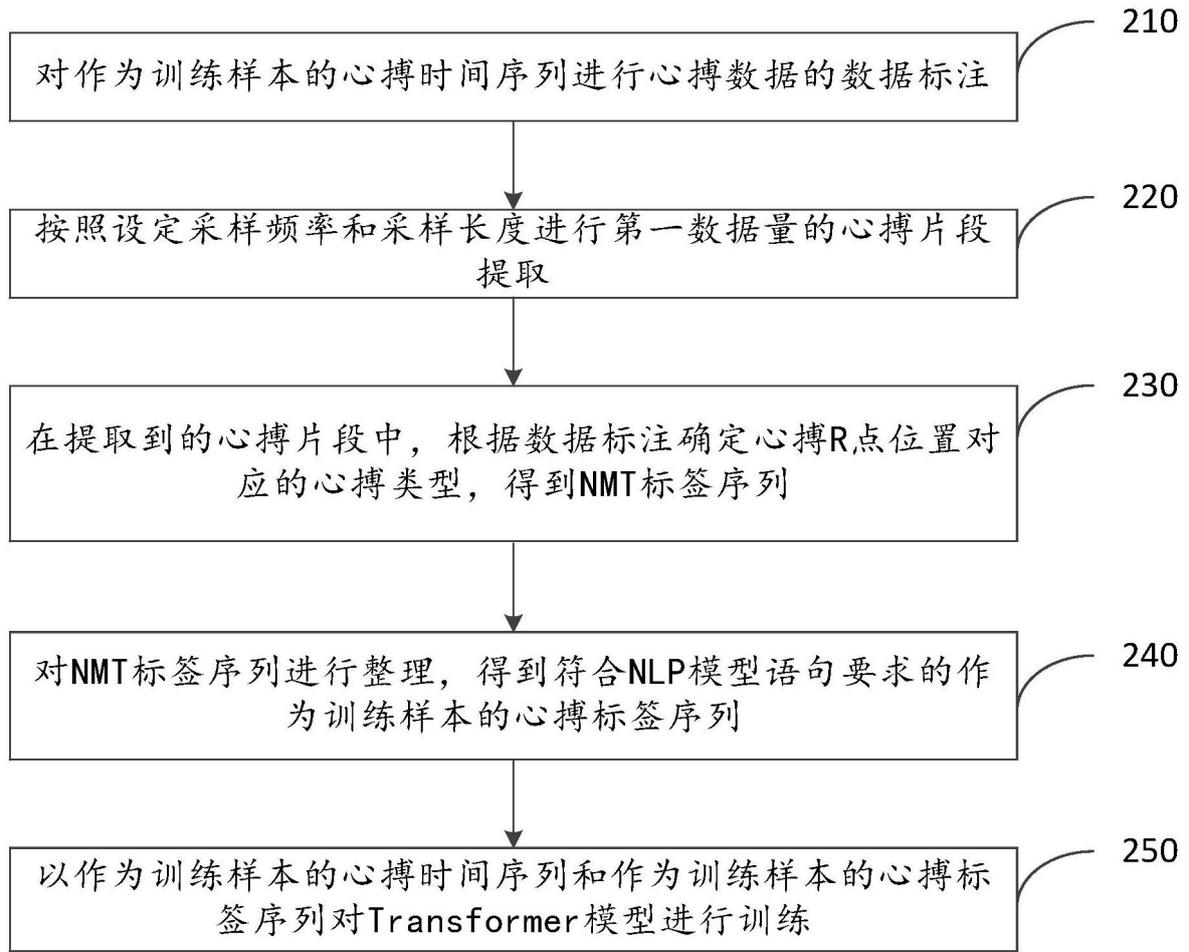


图3

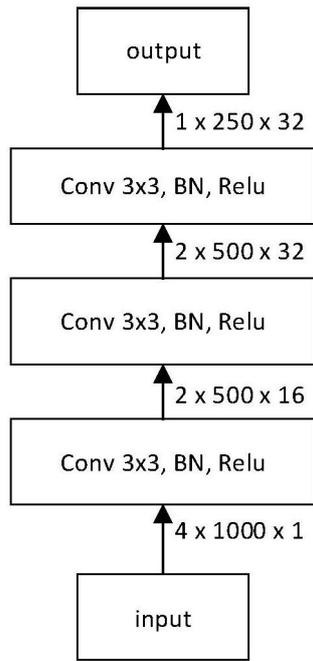


图4

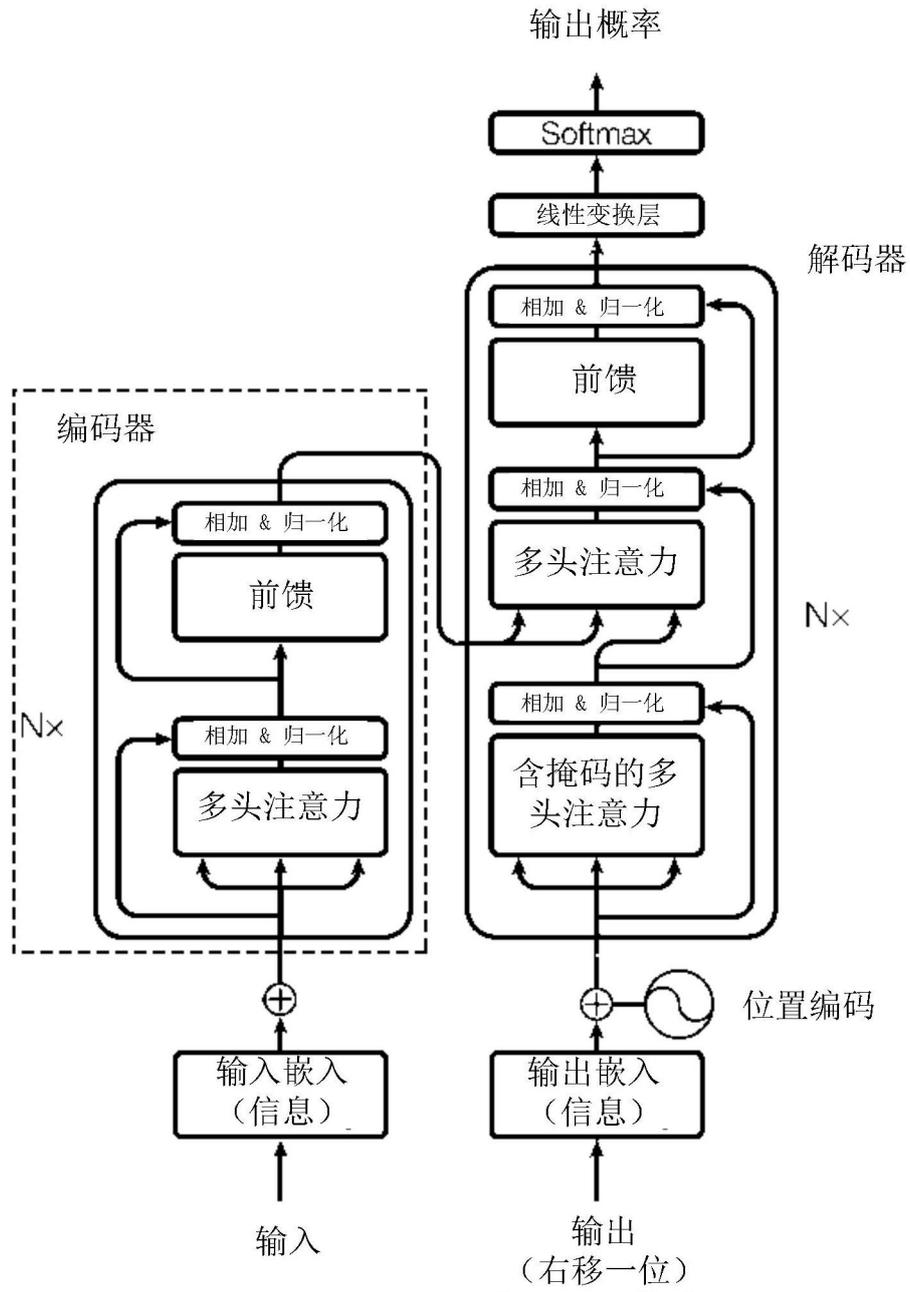


图5

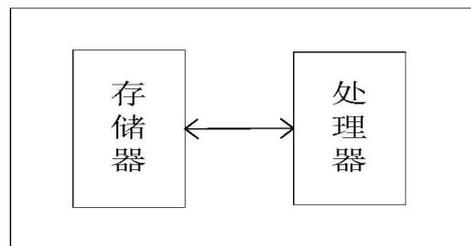


图6