

基于深度学习可分析标签相关性的心电信号分类

张诗雨¹, 孙占全²

¹上海理工大学理学院, 上海

²上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2024年4月1日; 录用日期: 2024年5月1日; 发布日期: 2024年5月7日

摘要

心血管疾病是全球主要死因之一, 及早发现心脏病的风险因素对于预防猝死至关重要, 基于心电图的心律失常自动检测是心脏病筛查的重要手段, 对降低心脏病死亡率至关重要。临床上同一患者可能同时出现多个心电异常类型, 因此, 解决心电图数据中多标签问题且考虑标签关系显得尤为关键。本文提出了一个可分析心电图多标签相关性的卷积神经网络分类模型, 该模型包含三个模块: 空间特征提取模块, 卷积神经网络作为基础架构, 用来提取心电图的特征; 标签与特征相关性嵌入学习模块, 用于探索标签之间的相关性, 并将标签与特征相关性嵌入信息融合到特征空间中; 解码器模块, 用于预测相应心律失常类别。为解决样本不均衡, 采用非对称损失函数平衡正负标签。通过对两个心电数据集进行实验, 实验结果显示, 本文提出的模型相较于现有心电图分类方法, 表现更优异。

关键词

心电信号分类, 多标签分类, 注意力机制, 深度学习, 数据不均衡

ECG Signals Classification by Analyzing Label Correlation Based on Deep Learning

Shiyu Zhang¹, Zhanquan Sun²

¹College of Science, University of Shanghai for Science and Technology, Shanghai

²School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 1st, 2024; accepted: May. 1st, 2024; published: May. 7th, 2024

Abstract

Cardiovascular disease is one of the leading causes of death in the world. Early detection of risk factors for heart disease is crucial for the prevention of sudden death. Clinically, multiple types of ECG abnormalities may occur in the same patient at the same time. Therefore, it is particularly important to solve the problem of multiple labels in ECG data and consider the label relationship. In this paper, a convolutional neural network classification model is proposed to analyze the correlation between ECG multi-labels. The model includes three modules. In the spatial feature extraction module, convolutional neural network is used as the infrastructure to extract ECG features. The label and feature correlation embedding learning module is used to explore the correlation between labels and fuse the label and feature correlation embedding information into the feature space. Decoder module, which predicts the probability of the corresponding arrhythmia class. Besides, to solve the data imbalance, asymmetric loss was used to balance the positive and negative labels. Experiments on two electrocardiogram datasets show that the proposed model has better performance than the existing ECG classification methods.

Keywords

Electrocardiogram Classification, Multi-Label Classification, Attention Mechanism, Deep Learning, Imbalanced Data

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

心电图(Electrocardiogram, ECG)是一种常用的非侵入性方法,用于评估心脏的电活动。通过分析心电图信号,医生可以诊断心脏疾病、监测心脏功能以及预测心血管事件的风险,然而,传统的心电图分类方法通常需要依赖专业医生对心电图进行解读。

近年来,深度学习技术的迅猛发展为心电图分类带来了新的机遇,深度学习模型,尤其是卷积神经网络(Convolutional Neural Networks, CNN)和循环神经网络(Recurrent Neural Networks, RNN),在图像和序列数据处理方面取得了显著的成功,这些模型可以自动学习特征表示,并通过大规模的心电数据集的训练来提高模型的性能。利用深度学习模型进行心电图分类具有许多潜在优势。首先,深度学习模型能够从原始的心电图信号中提取抽象的特征表示,无需手动设计特征工程,减轻了人工干预的负担。其次,深度学习模型可以通过端到端的训练方式直接从数据中学习分类规则,避免了繁琐的手动规则制定过程。此外,深度学习模型具有较强的泛化能力,可以处理不同心电图类型、噪声和干扰等情况,并具备适应新数据的能力。

基于深度学习的心电图分类模型都取得了不错的分类效果,Ge等人提出了一种算法,将残差卷积模块和挤压激励注意力机制模块结合起来,从原始心电图信号中提取特征[1]。Yao等人在其研究中提出了一种名为时间增量卷积神经网络(TimeIncremental Convolutional Neural Network, TI-CNN)的方法,旨在为心电图模型提供更大的输入长度灵活性[2]。后来,他们提出了一种名为基于注意力的时间增量卷积神经网络的新模型(Attention-based Time-Incremental Convolutional Neural Network, ATI-CNN),以TI-CNN作为参考,通过整合CNN、循环单元和注意力机制模块,实现了对心电信号信息的空间特征和时间特征的

融合[3]。He 等人提出了一种基于深度神经网络的心律失常自动分类方法, 该方法包括残差卷积模块和双向长短期记忆层(Bidirectional Long Short-Term Memory, BLSTM) [4]。这些研究表明深度学习模型可以有效诊断心律失常, 并且深度学习在未来为辅助医生进行疾病诊断和为患者提供护理服务方面具有巨大潜力。

尽管这些模型采用了不同的方法都取得了有效的成果, 但其中一些模型将心电图分类任务视为单标签分类任务。因此, 心电图自动分类仍面临许多挑战, 一方面, 一些心电图记录可能同时包含多个标签, 例如在 CPSC 2018 数据库中[5], 虽然大多数心电信号只表现出一种心律失常类型, 但有些心电信号可能有两个或两个以上的心律失常类型, 因为一个病人可能同时患有多种心脏疾病。许多现有的方法将多标签分类任务视为单标签分类任务, 即使该记录有两个或多个类型, 模型只给出一种分类结果, 从而会导致分析结果的不准确。此外, 目前心电分类方法未考虑到标签之间的相关性。因此, 多标签心电图分类仍然是心电图分析中具有挑战性和重要性的问题。为提高模型的分类型性能, 本文提出的模型利用标签相关性来更好地捕捉心电图不同心律失常类型之间的信息, 为模型进行分类提供指导。此外, 不同通道的特征在分类上具有不同水平的重要性, 为了解决这个问题, 我们在网络中使用通道注意力机制模块, 自适应地调整不同通道特征的权重, 以提高模型的分类型性能。本文提出了一个基于卷积神经网络和注意力机制的可以分析标签之间相关性的分类模型, 总结起来, 本研究有以下几点主要贡献:

(1) 本文提出的模型包含三个模块: (a) 空间特征提取模块, 卷积神经网络作为基础架构, 用来提取心电图的特征; (b) 标签与特征相关性嵌入学习模块, 用于探索标签间的相关性, 并将标签与特征相关性嵌入信息融合到特征空间中; (c) 解码器模块, 预测相应心律失常类型。

(2) 本文可以直接从心电数据标签与卷积神经网络的输出特征中学习标签嵌入, 并高效地分析标签之间的相关性, 从而有效地提高模型的分类型性能。此外, 模型结合了卷积神经网络、注意力机制和非对称损失函数, 通道注意力可以自适应地调整不同通道特征的权重, 使模型增强对空间特征的感知学习能力; 为了解决数据不均衡问题, 采用了非对称损失函数调整正负标签的权重, 进一步提高了模型的分类型结果。

(3) 在两个公开的心电数据集上进行实验, 实验结果表明, 本文所提出的模型在心律失常分类方面优于一些现有模型的分类型结果。

2. 仿真模型

2.1. 问题定义

在本文中, 多标签心律失常检测任务是对 ECG 序列中的不同心律失常进行诊断。本论文所提出的模型采用 12 导联心电图记录作为输入, 并输出一个或多个标签, 即相应的 ECG 心律失常类型。训练集可表示为 $S = (X, Y)$, 其中 X 是一个输入信号, 即长度为 L 的 12 导联 ECG 记录, Y 是数据集的真实标签的集合。在我们的模型中, 令 $Y = [Y_0, Y_1, Y_2, \dots, Y_k, \dots, Y_K]$, 其中 $Y_k \in \{0, 1\}$, $K = 0, 1, \dots, K$ 。这里, Y_k ($0 \leq k \leq K$) 对应于一个心律失常的真实标签。如果 ECG 记录 X_i 含有第 k 类异常标签, 则 Y_k 等于 1, 否则, Y_k 等于 0, 如下表达式所示:

$$Y_k = \begin{cases} 1, & \text{如果心电记录 } X_i \text{ 包含第 } k \text{ 类的心律失常标签,} \\ 0, & \text{如果心电记录 } X_i \text{ 不包含第 } k \text{ 类的心律失常标签.} \end{cases} \quad (1)$$

多标签深度学习模型的主要任务是从模型 S 中获得一个多标签分类函数 $F(X) \rightarrow P$, 该函数可以将每个输入信号映射到一组向量 P 。使用 X 作为输入, 我们的模型预测一条心电记录对应的每个类别的概率为 $P = [P_0, P_1, P_2, \dots, P_k, \dots, P_K]$, 其中 P_k 是模型预测的第 k 个类别的概率, K 是 ECG 数据集类别的数量。在训练期间, 模型的目的是最小化预测概率 P 相对于其真实标签 Y 的损失, 让模型预测的概率无限接近

其真实标签。

2.2. 仿真模型

本文提出了一种融合了标签与特征关系嵌入的模型，可以检测不同心电图心律失常的类型，并将标签嵌入到特征提取过程中，从而实现有效的多标签诊断。总体框架如图 1 所示，主要由三个模块组成：(1) 空间特征提取模块，卷积神经网络作为基础架构，用来提取心电图的特征；(2) 标签与特征相关性嵌入学习模块，用于探索标签之间的相关性，并将标签与特征相关性嵌入信息融合到特征空间中；(3) 解码器模块，预测相应心律失常的概率。

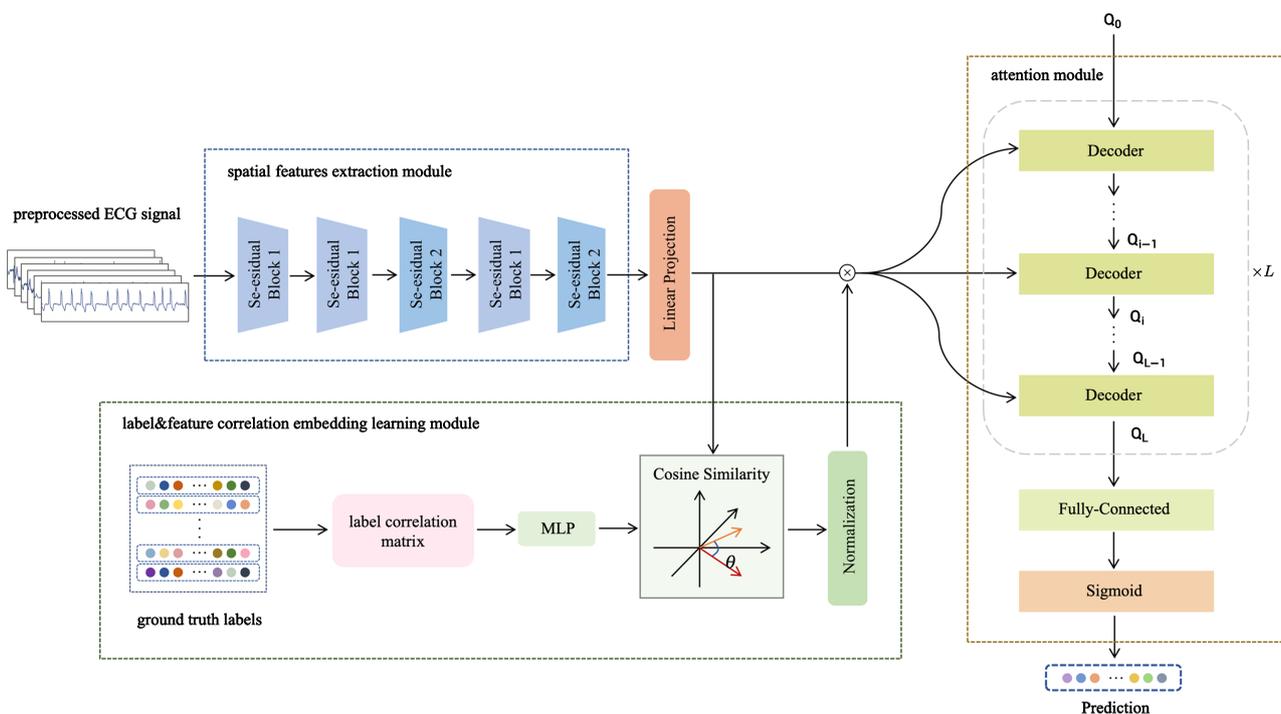


Figure 1. The framework of our proposed model

图 1. 本文提出模型的框架结构

2.2.1. SE 模块

挤压激励(Squeeze and Excitation, SE)注意力机制模块[6]通过对通道之间的相互依赖关系进行建模，自适应地重新校准特征通道的权重，其目标是通过明确卷积特征通道之间的相互依赖关系，增强网络的分类效果，使网络能够进行特征重新校准，从而学会利用全局信息选择性地强调包含心律失常信息丰富的特征通道，并抑制不相关的特征通道。对于包含重要信息的特征通道应予以重视并赋予较大的权重。因此本文提出将 SE 模块应用于心律失常分类任务，让模型有选择地强调与心律失常类别相关性大的特征，同时抑制其他不相关的特征。SE 模块的结构如图 2 所示，模块包含挤压模块和激励模块。对于任何输入 U ，其中 $U \in \mathbb{R}^{C \times H \times W}$ ， U 首先经过挤压操作，之后输入激励模块，从而为特征中的不同通道产生不同的权重。

$U = [u_1, u_2, \dots, u_c]$ 作为 Squeeze 模块的输入，Squeeze 模块中的操作是全局信息嵌入，将输入特征进行全局平均池化计算，生成与通道有关的 $1 \times 1 \times C$ 向量 z ，计算公式如下：

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (2)$$

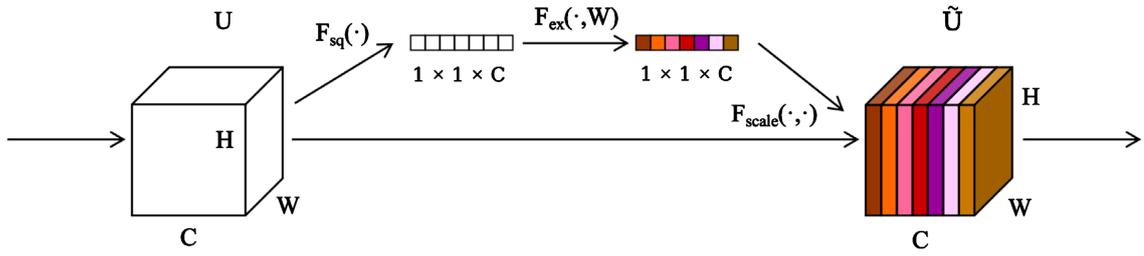


Figure 2. The internal architecture of SE block [6]
图 2. SE 模块的内部结构[6]

利用挤压操作聚合信息，之后进行激励操作，以完全捕获通道依赖性，激励模块用于自适应重新校准，主要是为了提升模型的泛化能力。激励模块包含了两个非线性的全连接层，第一次全连接操作将其通道数进行缩小，降维比为 r ，即将通道数降为 $\frac{C}{r}$ ，第二次全连接操作将通道数还原成 C ，最后将得到的向量中的元素映射为 0 至 1 范围内的小数，作为不同通道的权重，具体的计算公式如下：

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \quad (3)$$

其中 σ 表示修正线性单元层(Rectified Linear Unit, ReLU)激活函数， $W_1 \in \mathbb{R}^{\frac{C}{r} \times C}$ ， $W_2 \in \mathbb{R}^{C \times \frac{C}{r}}$ ，最终将权重与原来的特征对应相乘，得到重新设置通道权重后的特征为 $\tilde{U} = [\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_C]$ 。

$$\tilde{u}_c = F_{scale}(u_c, s_c) = s_c \times u_c \quad (4)$$

2.2.2. 空间特征提取模块

受到残差网络[7]的启发，采用卷积神经网络来提取局部特征，该主干网络包括两类模块，如图 3(a)所示，第一种卷积模块由三个卷积层、一个平均池化层、一个 SE 模块[6]、若干个归一化层(Batch Normalization, BN)和 ReLU 激活函数组成，与残差网络[7]有所不同，在跳跃连接处加入了最大池化层。与第一类模块相比，第二类模块缺少一个池化层，残差网络的连接方式采用了直接连接，可见图 3(b)。Conv1x3xn 表示卷积层的核大小为 1x3，核数为 n 。Conv3x3xn 表示卷积层的核大小为 3x3，核数为 n 。第一类卷积模块重复 3 次，唯一的区别是卷积层通道数的变化，分别为 C_1 、 C_2 和 C_3 ，第二类卷积一共有两个，除了卷积层的通道数不同，其余层结构一致。所有的平均池化层使用池化核大小为 5，从而使输入长度在每个基本卷积块中减小五倍。每个卷积层后面都跟随一个 BN 层和一个 ReLU 激活函数，BN 层有助于在训练中获得更稳定的参数更新[8]，ReLU 是一种广泛使用的激活函数，因其在深度神经网络中防止梯度消失问题方面的能力而备受认可[9]。在这一阶段，将输入的心电信号作为主干网络的输入，用于提取局部特征，给定输入的心电图记录 $X \in \mathbb{R}^{1 \times 12 \times L}$ ，通过主干网络提取到的特征 $\tilde{X} \in \mathbb{R}^{C_3 \times 12 \times 40}$ 作为线性投影层(Linear Projection Layer)的输入，该层将特征映射成二维特征 $F \in \mathbb{R}^{C_3 \times \text{hidden_dim}}$ 。

2.2.3. 特征与标签嵌入学习模块

本模板主要是探索标签与标签之间的相关性，将特征与标签的相关性嵌入融进主干网络的特征中，通过计算学习卷积后的特征与标签的余弦相似度，从而获得二者之间的相关性，指导模型更好地进行心律失常的分类。首先，统计各个心律失常出现的总次数，构建标签相关矩阵 E (label correlation matrix)，标签相关矩阵的每个元素可以用条件概率表示：

$$e_{i,j} = p_{i,j}(Y_i|Y_j) = \frac{t(Y_i, Y_j)}{t(Y_j)} \quad (5)$$

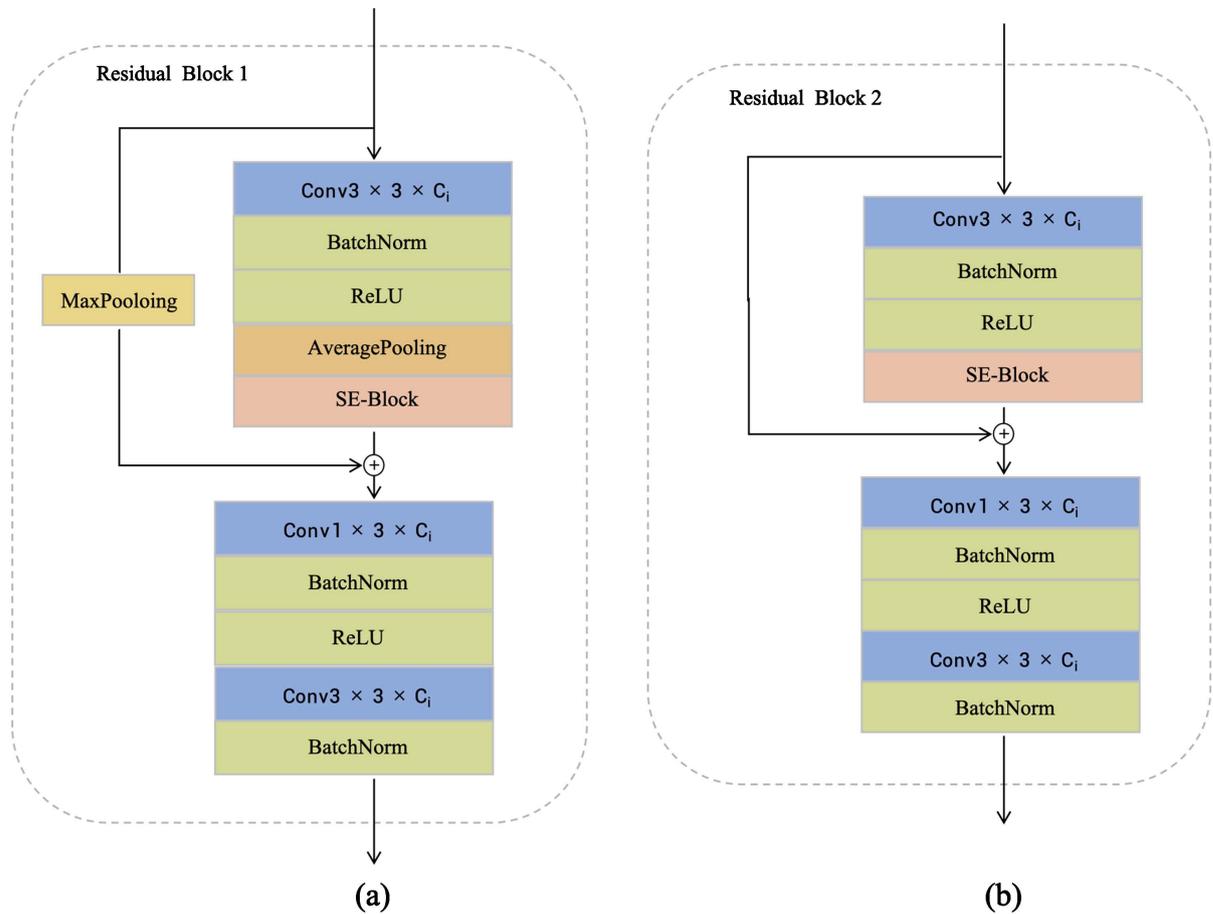


Figure 3. The internal architecture of residual block
图 3. 残差模块的内部结构

式中 $t(Y_i, Y_j)$ 为标签 Y_i 与 Y_j 一起出现的次数, $t(Y_j)$ 表示心律失常 Y_j 发生的次数, $p_{i,j}(Y_i|Y_j)$ 表示心律失常 Y_j 发生时发生 Y_i 的概率, 条件概率构成标签相关矩阵 E , $E = [E_0, E_1, \dots, E_k, \dots, E_K]^T$, $E \in \mathbb{R}^{K \times K}$, $E_k = [e_{k,0}, e_{k,1}, \dots, e_{k,k}, \dots, e_{k,K}]$ 。利用多层感知机(Multilayer Perceptron, MLP)将每个 ECG 心律失常类型的标签相关矩阵映射到语义嵌入空间, 生成标签相关性嵌入 $M = [M_0, M_1, \dots, M_K]^T$, 从而得到相关矩阵定义的标签依赖关系, 计算方式如下:

$$M_k = \phi(E_k) \quad (6)$$

其中, ϕ 表示 MLP, 由三个全连阶层以及一个 ReLU 组成, 得到了最佳的标签相关性嵌入。为了让标签与特征更好的融合, 采用余弦相似度求标签与特征之间的相关性:

$$s_c^k = \text{ReLU} \left(\frac{F_c \times M_k}{\|F_c\| \times \|M_k\|} \right) \quad (7)$$

其中 $F = [F_0, F_1, \dots, F_{C_3}]^T$, 之后将得到的与标签有关的通道注意力系数归一化:

$$\hat{s}_c^k = \frac{s_c^k}{\sum_{c=0}^{C_3} s_c^k} \quad (8)$$

对于特征向量中的每个点, 归一化之后的注意力系数越大, 则说明该特征与对应的标签嵌入相关性越大, 将归一化的注意力系数对每一类的特征向量进行加权叠加:

$$\widehat{F}_k = \sum_{c=0}^K s_c^k \widehat{F}_k \quad (9)$$

其中 \widehat{F}_k 是与第 k 个心律失常类别相关的加权特征, 最终得到该模块的输出 $\widehat{F} = [\widehat{F}_0, \widehat{F}_1, \dots, \widehat{F}_k, \dots, \widehat{F}_K]^T$ 。

2.2.4. 解码器模块

注意力模块由多个解码器层和一个线性投影层组成, 该模块可用于提取全局特征, 并预测每个类别的概率作为模型的输出。受到 ML_Decoder [10]的启发, 相比于 Transformer [11]的解码器, 结构可见图 4(a), 本文提出的模型解码器模块减少一个自注意力模块, 由于在经过注意力模块之前, 将特征输入一个投影层, 投影层可以将查询转换为任何想要的输出, 从而使自注意力模块成为冗余。因此, 与 Transformer 相比, 本文将自注意力层去除, 将查询直接输入多头注意力机制, 结构可见图 4(b)。

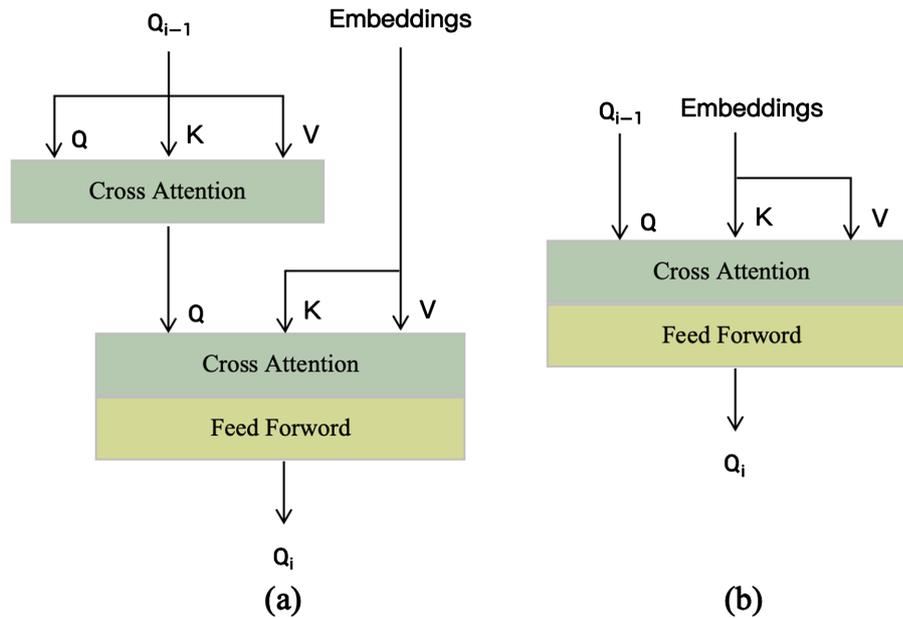


Figure 4. (a) The internal architecture of Decoder of Transformer; (b) The internal architecture of Decoder of our proposed model

图 4. (a) transformer 的解码器内部结构; (b) 本文的解码器内部结构

将特征与标签嵌入学习模块的输出特征 \widehat{F} 作为查询输入解码器模块, 第 i 层解码器根据其前一层的输出更新查询。每一个解码器模块包含一个交叉注意力层(cross-attention layer)和一个前馈神经网络层(Feed-Forward Networks layer, FFN), 具体如下:

$$CrossAttention: \widetilde{Q}_{i-1} = MultiHeadAttention(Q_{i-1}, \widehat{F}, \widehat{F}) \quad (10)$$

$$FeedForward: Q_i = FFN(\widetilde{Q}_{i-1}) \quad (11)$$

其中, 交叉注意力模块使用多头注意力函数实现, 多头注意力函数和 FFN 与标准的 Transformer 解码器[11]中的定义相同。在交叉注意力模块中, key 和 value 是特征与标签嵌入学习模块的输出特征, 交

又注意力模块可以分析标签与主干网络提取的特征之间的关系, 因此, 通过每个解码器层, 查询根据特征与提取到的空间特征之间的相关性逐层更新。最后, 将第 L 层解码器的输出进行全连接操作(Fully Connection, FC)以及 sigmoid 函数, 得到模型预测的每个心律失常类别的概率 P :

$$\text{Logits} : P = \text{sigmoid} \left(FC(Q_L) \right) \quad (12)$$

2.3. 损失函数

在心电数据集中, 一条记录包含少量的正标签和大量的负标签, 导致了样本之间的不均衡, 为了更有效地解决这个问题, 本论文采用了简化的非对称损失函数(Asymmetric-loss, ASL) [12], 该损失在正样本和负样本上的操作不同。与传统的交叉熵损失相比, 非对称损失减少了负样本的贡献, 并在训练过程中更加强调正标签。我们利用以下非对称损失函数来计算每个训练样本 X 的损失:

$$L = \frac{1}{K} \sum_{k=1}^K \begin{cases} (1 - P_k)^{\gamma^+} \log(P_k), Y_k = 1, \\ (P_k)^{\gamma^-} \log(1 - P_k), Y_k = 0. \end{cases} \quad (13)$$

其中, P_k 是模型预测的各个类别的概率, Y_k 是一个二进制标签, 表示心电记录 X 的实际标签中是否含有第 k 类标签。

3. 仿真结果

3.1. 实验环境

本文所提出的模型在一台配备 GeForce GTX GPU、8 GB 内存的服务器上进行训练和测试, 该模型使用 PyTorch1.7.0 和 CUDA 10.1 框架进行实现。

3.2. 数据集

3.2.1. CPSC2018 心电数据集

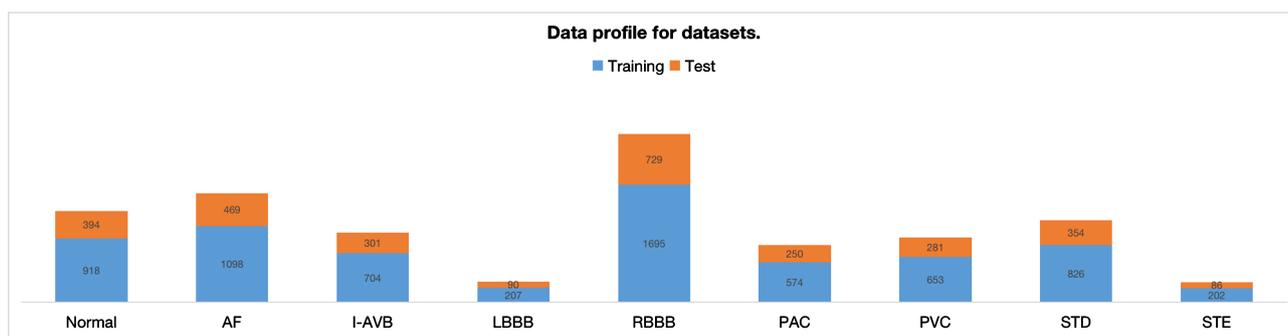


Figure 5. Data profile for CPSC 2018

图 5. CPSC 2018 心电数据集的数据概况

CPSC 2018 心电数据集[5]是中国生理信号挑战赛第一届比赛的数据集, 该数据集的信号来自 11 家医院, 包含了九种类型的心电图数据: 一种是正常的心电数据类型(Normal), 另外八种是心律失常类型分别为: (1) 心房颤动(Atrial Fibrillation, AF); (2) 一度房室传导阻滞(First-degree Atrioventricular Block, I-AVB); (3) 左束支传导阻滞(Left Bundle Branch Block, LBBB); (4) 右束支传导阻滞(Right Bundle Branch Block, RBBB); (5) 房性早搏(Premature Atrial Contraction, PAC); (6) 室性早搏(Premature Ventricular Contraction, PVC); (7) ST 段压低(ST-segment Depression, STD); (8) ST 段抬高(ST-segment Elevation, STE)。图 5 详细描述了数据集的

数据概况, 训练集一共包含 6877 个心电图记录, 测试集包含 2954 个心电图记录, 每个心电图记录有 12 个导联, 每一条记录的长度不同, 从 6 秒到 60 秒不等, 大多数记录有一个标签为正, 但有些记录可能有两个或两个以上的标签为正, 训练集和测试集分别有 476 个和 203 个多标签类型的实验数据。

3.2.2. PTB-XL 心电数据集

实验的另一个数据集是 PTB-XL 心电图数据库[13], 包含来自 18,885 名患者的 21,837 份记录。这些心电图记录为 12 导联, 持续时间为 10 秒, 采样频率为 500 Hz, 并且该数据集在性别方面保持平衡, 男性占 52%, 女性占 48%。此外, 心电图信号由两名经验丰富的心脏病专家进行注释, 并且, 该多标签心电数据集一共有 71 个不同陈述, 包含了 44 个不同的诊断陈述、19 个不同的形式陈述和 12 个不重叠的节律陈述, 其中 4 个陈述既属于诊断陈述又属于形式陈述。本论文主要对诊断陈述的数据集进行实验, PTB-XL 提供了超类和子类层次的分类情况, 可用于训练更广泛类别上的分类算法, 而不是原始的细粒度诊断标签, PTB-XL 心电数据集的详细信息见图 6。本文使用 PTB-XL 心电数据集推荐的训练集和测试集进行实验, 其中, 训练集数据 19,267 例, 测试集数据 2163 例, 推荐的测试集数据标签质量高, 数据分布与训练集数据相似。

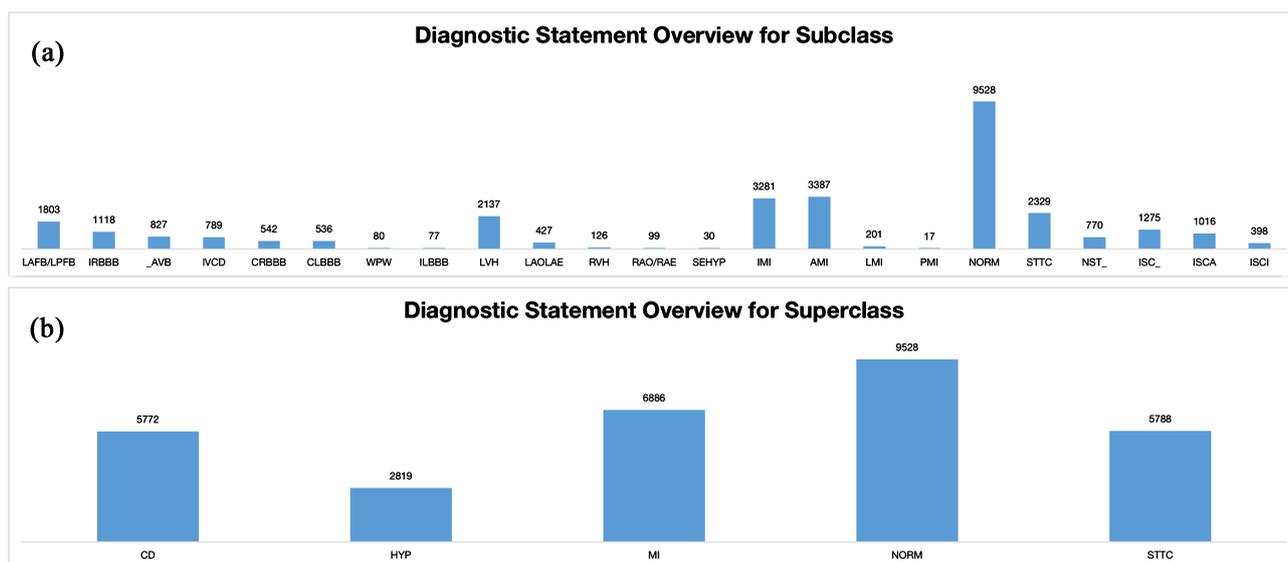


Figure 6. Diagnostic statement overview for subclass (a) and superclass (b) of PTB-XL

图 6. PTB-XL 心电数据集子类(a)及超类(b)诊断陈述概述

3.3. 数据预处理和训练设置

CPSC 2018 心电数据集包含了 6 秒到 60 秒不等的心电图记录, 卷积神经网络无法处理不同长度的输入, 因此需要调整原始数据的大小, 本文对降采样后的心电信号进行裁剪或填充, 使它们的长度相同。将数据集进一步分成训练集和测试集, 比例为 8:2, 后续的实验均用训练集进行模型的训练, 最终将训练好的模型在测试集上进行测试得到分类结果。通过对超参数的精细调整, 最终选择了 32 作为批处理大小。使用 Adam 优化器[23]对整个模型进行了 60 个时期的训练, 以迭代更新权重, 其中学习率在开始时设置为 0.0001, 此外, 使用早停策略来确定适当的迭代次数, 以避免过度拟合。当验证损失保持不变时, 说明网络已经收敛, 训练停止, 从而保证可以得到最佳的分类效果。

3.4. 评价指标

在这项研究中, 我们使用了典型的分类指标来验证模型的分类性能, 包括每个分类类型和子类的准确率、精确率、召回率以及所有类别的平均 F1 值。准确率是正确预测数量与总预测数量的比率, 精确率是正确的正类别预测数量与所有正类别预测数量的比率, 召回率是正确的正类别预测数量与真正的正类别和假负类别的总数的比率, F1 值是精确率和召回率的加权平均值。以下的方程式可用于计算各个类别和整体的精确度、召回率和 F1 值:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{TP}} \quad (14)$$

$$\text{Precision} = \frac{\text{TP}}{\text{FP} + \text{TP}} \quad (15)$$

$$\text{Recall} = \frac{\text{TP}}{\text{FN} + \text{TP}} \quad (16)$$

$$F1 = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FN} + \text{FP}} \quad (17)$$

其中 K 是数据集分类的数量, C 表示不同类别, TP 表示真实值和预测值都是正样本, 即真阳性; TN 表示真实值和预测值都是负样本, 即真阴性; FP 表示真实值为负样本, 预测值为正样本, 即假阳性; FN 表示真实值为正样本, 预测值为负样本, 即假阴性, 详细可见下表 1。

Table 1. Confusion matrix
表 1. 混淆矩阵

	预测值 = 1	预测值 = 0
真实值 = 1	TP	FN
真实值 = 0	FP	TN

3.5. 仿真结果

3.5.1. PTB-XL 心电数据集仿真结果

本文比较了我们提出的模型与近几年学者提出的方法在 PTB-XL 数据集上不同分类层次的性能, 如表 2 所示, 针对 PTB-XL 数据集, Smigiel 等人提出了三种神经网络架构: 基于卷积神经网络的神经网络架构、基于 SincNet 的神经网络架构以及基于卷积网络的增加了基于熵的特征的神经网络架构。在这项工作中, Smigiel 开发了一个深度神经网络用于原始心电图信号的自动分类[14]。另外, 为了探索和实现有效的 ECG 识别, Smigiel 等人提出了一种卷积神经网络, 通过添加基于熵的特征来执行单个 QRS 波的编码[15]。除此以外, Jyotishi 等人将 FSL (Few-Shot Learning) 网络的结果与基于 Softmax 的神经网络进行分类的评估评分进行比较, 提出的神经网络解释一组从 ECG 信号提取的 QRS 波群, 通过训练深度卷积神经网络来识别 PTB-XL 心电数据集中 2、5 和 20 种不同的心脏病类别[16]。Jyotishi 等人开发了一个专注的时空学习网络(Attentive spatio-temporal learning network, ASTLNet), 它可以通过利用多导联 ECG 信号的并发时空变化来学习更好的诊断表征。ASTLNet 由两个模块组成, 即时空表征学习(Spatio-temporal representation learning, STRL)模块和注意时空聚集(Attentive spatio-temporal aggregation, ASTA)模块。STRL 模块用于学习多尺度的时空特征, ASTA 模块用于聚合学习到的特征[17]。而本文所提出的模型主要是分析心电数据标签之间的相关性, 从而提升模型对多标签心电数据的分类效果。在 PTB-XL 心电数据集上进行了一系列实验, 与以上的方法进行了比较, 针对超类诊断(Super diagnosis)的 F1 值为 0.754, 准确率

为 0.864, 精确率为 0.783, 召回率为 0.732, 评价指标均高于比较的模型。此外, 在 PTB-XL 的子类诊断 (Sub diagnosis) 方面, 本文提出模型的 F1 值、精确率和准确率均高于现有模型的仿真结果。由这些实验结果表明, 该模型在 PTB-XL 心电数据集都取得了很好的结果, 这表明该模型在心电图心律失常分类方面是有效的。

Table 2. Comparison of results of the proposed model and those of the existing model on sub and super diagnosis of PTB-XL
表 2. 本文提出的模型与现有模型在 PTB-XL 心电数据集子类 and 超类诊断上的分类结果对比

模型	分类级别	Accuracy	Precision	Recall	F1
Smigiel <i>et al.</i> [14]		0.698	0.355	0.339	0.332
Smigiel <i>et al.</i> [15]		0.685	-	-	0.345
Palczynski [16]	sub diagnosis	0.671	-	-	0.324
Jyotishi [17]		0.541	-	-	0.460
Proposed model		0.727	0.479	0.580	0.482
Smigiel <i>et al.</i> [14]		0.765	0.714	0.662	0.680
Smigiel <i>et al.</i> [15]		0.763	-	-	0.683
Palczynski [16]	Super diagnosis	0.790	-	-	0.717
Jyotishi [17]		0.627	-	-	0.736
Proposed model		0.902	0.783	0.732	0.754

3.5.2. CPSC2018 心电数据集仿真结果

本文在 CPSC2018 心电数据集上对本文提出的网络和一些现有的分类网络进行了比较分析, 表 3 列出了不同模型方法的各项性能指标, 粗体显示的值表示相应指标的最佳值, 与 TI-CNN [2]、ATI-CNN [3] 以及 MLC-CNN[21]相比, 本文提出的模型分类结果 F1 分别提升了约 0.057、0.017 和 0.002。与文献[4]、文献[18]、文献[17]、文献[19]、文献[1]中模型分类结果相比, 本文提出的模型 F1 值较高。Chen 等人提出的模型 F1 达到了 0.837, 其模型整合了 130 个模型来构建一个复杂且庞大的集成模型[20], 由于这种架构复杂, 因此, 模型预测和预测需要更多的计算成本, 相比之下, 本文提出的模型取得了明显的改进, 更具竞争力。

Table 3. Comparison of results of the proposed model and those of the existing model on CPSC-2018
表 3. 本文提出的模型与现有模型在 CPSC2018 心电数据集上的分类结果对比

模型	Normal	AF	I-AVB	LBBB	RBBB	PAC	PVC	STD	STE	F1
TI-CNN [2]	0.750	0.900	0.920	0.760	0.640	0.830	0.870	0.810	0.460	0.772
He <i>et al.</i> [4]	0.755	0.846	0.780	0.790	0.748	0.829	0.869	0.870	0.655	0.806
ATI-CNN [3]	0.789	0.920	0.933	0.789	0.736	0.861	0.872	0.850	0.556	0.812
Zhang <i>et al.</i> [18]	0.805	0.919	0.864	0.866	0.926	0.735	0.851	0.814	0.535	0.813
Jyotishi <i>et al.</i> [17]	0.790	0.923	0.867	0.892	0.937	0.757	0.837	0.798	0.558	0.818
Geng <i>et al.</i> [19]	0.824	0.925	0.882	0.937	0.939	0.734	0.767	0.835	0.600	0.827
MLC-CNN [21]	0.792	0.936	0.889	0.779	0.744	0.866	0.876	0.876	0.683	0.827
Ge <i>et al.</i> [1]	0.790	0.920	0.870	0.870	0.930	0.780	0.860	0.810	0.590	0.828

Chen <i>et al.</i> [20]	-	-	-	-	-	-	-	-	-	-	0.837
Proposed model	0.798	0.913	0.867	0.826	0.916	0.850	0.849	0.846	0.596	0.829	

3.6. 消融实验

为了验证标签与特征相关性嵌入学习模块和解码器模块的有效性, 进行了一系列消融实验, 将我们的网络分别与没有标签与特征相关性嵌入学习模块和解码器模块的两种网络变体进行比较。该模型由三个模块组成: 空间特征提取模块、标签与特征相关性嵌入学习模块和解码器模块, 用于比较的三个模型的结构如下:

(1) 模型 1: 该模型仅使用空间特征提取模块进行特征提取, 我们设计这个模型是为了验证标签与特征相关性嵌入学习模块和解码器模块的作用。

(2) 模型 2: 该模型包含标签与特征相关性嵌入学习模块, 不包含解码器模块, 设计这个模型是为了验证解码器模块的效果。

(3) 模型 3: 该模型包含解码器模块, 不包含标签与特征相关性嵌入学习模块, 设计这个模型是为了验证标签与特征相关性嵌入学习模块的效果。

三个参考模型和本文提出模型在 CPSC 2018 心电数据集上进行实验, 得到 F1 值比较见表 4, 模型 2 的 F1 值比模型 1 的高出 0.191, 而模型 3 的 F1 值比模型 1 高 0.163, 而模型 2 包含标签与特征相关性嵌入学习模块, 不包含解码器模块, 而模型 3 包含解码器模块, 不含标签与特征相关性嵌入学习模块, 这表明, 标签与特征相关性嵌入学习模块比解码器模块的分类效果更好。相比于三个参考的对比模型, 本文提出的模型在所有模型中的 F1 值最高, 为 0.829, 并且在 Normal、AF、PAC、PVC、STD、STE 子类的 F1 值是所有模型中最高的。这些结果表明, 标签与特征相关性嵌入学习模块和解码器模块在特征提取和心律失常分类中都是有效的, 而两个模块的组合比单独使用一个模块分类效果更好, 两个模块的联合可以提高模型的性能。

Table 4. Comparison of results of the proposed model and those of the existing model on CPSC 2018
表 4. 本文提出的模型与现有模型在 CPSC 2018 心电数据集上的分类结果

模型	Normal	AF	I-AVB	LBBB	RBBB	PAC	PVC	STD	STE	F1
模型 1	0.532	0.754	0.614	0.832	0.823	0.315	0.645	0.564	0.373	0.606
模型 2	0.754	0.856	0.886	0.895	0.923	0.694	0.845	0.756	0.564	0.797
模型 3	0.714	0.843	0.854	0.875	0.903	0.634	0.776	0.746	0.576	0.769
Proposed model	0.798	0.913	0.867	0.826	0.916	0.850	0.849	0.846	0.596	0.829

4. 结束语

本文提出了一种用于提取标签之间相关性的卷积神经网络, 模型结合了卷积神经网络、注意力机制和非对称损失函数, 包含空间特征提取模块、标签与特征相关性嵌入学习模块以及解码器模块。首先, 空间特征提取模块使用卷积神经网络作为基础架构用来提取心电图的空间特征, 通过加入注意力机制, 自适应地调整不同通道的权重, 使模型增强对空间特征的感知学习能力, 学习到与不同类别相关的特征; 标签与特征相关性嵌入学习模块用于探索标签与卷积网络提取到的特征之间依赖关系, 并将标签之间相关性嵌入信息融合到特征空间中; 解码器模块用于进行相应心律失常分类, 得到每个类别的预测分数; 除此以外, 模型采用非对称损失函数来调整正负标签的权重。

与几种最新的方法相比, 本文提出的模型在 CPSC 2018 和 PTB-XL 两个公开的心电数据集上进行实验, 在 CPSC 2018 心电数据集上的 F1 值达 0.829, 在 PTB-XL 心电数据集上的 F1 值为 0.482 和 0.754, 其分类结果与近年来提出的模型的实验结果相比, 本文提出的模型 F1 值高于其他分类模型, 表现出卓越的分类性能。除此以外, 还对模型中的各个模块在 CPSC 2018 心电数据集上进行消融实验, 实验结果表明模型中的标签与特征相关性嵌入学习模块和解码器模块对心电分类都有着显著的作用, 并且两个模块的联合使用效果优于单个模块的使用。因此, 本文通过大量的实验有效验证了模型以及模型中的各个模块对多标签心电信号心律失常分类的有效性。

基金项目

本文由国防基础研究计划项目(JCKY2019413D001)、上海理工大学医学工程交叉项目(10-21-302-413)、上海市自然科学基金项目(19ZR1436000)部分资助。

参考文献

- [1] Ge, R., Shen, T., Zhou, Y., *et al.* (2021) Convolutional Squeeze-and-Excitation Network for ECG Arrhythmia Detection. *Artificial Intelligence in Medicine*, **121**, Article ID: 102181. <https://doi.org/10.1016/j.artmed.2021.102181>
- [2] Yao, Q., Fan, X., Cai, Y., *et al.* (2018) Time-Incremental Convolutional Neural Network for Arrhythmia Detection in Varied-Length Electrocardiogram. 2018 *IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*, Athens, 12-15 August 2018, 754-761. <https://doi.org/10.1109/DASC/PiCom/DataCom/CyberSciTec.2018.00131>
- [3] Yao, Q., Wang, R., Fan, X., *et al.* (2020) Multi-Class Arrhythmia Detection from 12-Lead Varied-Length ECG Using Attention-Based Time-Incremental Convolutional Neural Network. *Information Fusion*, **53**, 174-182. <https://doi.org/10.1016/j.inffus.2019.06.024>
- [4] He, R., Liu, Y., Wang, K., *et al.* (2019) Automatic Cardiac Arrhythmia Classification Using Combination of Deep Residual Network and Bidirectional LSTM. *IEEE Access*, **7**, 102119-102135. <https://doi.org/10.1109/ACCESS.2019.2931500>
- [5] Liu, F., Liu, C., Zhao, L., *et al.* (2018) An Open Access Database for Evaluating the Algorithms of Electrocardiogram Rhythm and Morphology Abnormality Detection. *Journal of Medical Imaging and Health Informatics*, **8**, 1368-1373. <https://doi.org/10.1166/jmih.2018.2442>
- [6] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/CVPR.2018.00745>
- [7] He, K., Zhang, X., Ren, S., *et al.* (2016) Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/CVPR.2016.90>
- [8] Ioffe, S. and Szegedy, C. (2015) Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *Proceedings of the 32nd International Conference on Machine Learning*, Lille, 6-11 July 2015, 448-456.
- [9] Nair, V. and Hinton, G.E. (2010) Rectified Linear Units Improve Restricted Boltzmann Machines. *Proceedings of the 27th International Conference on International Conference on Machine Learning*, Haifa, 21-24 June 2010, 807-814.
- [10] Ridnik, T., Sharir, G., Ben-Cohen, A., *et al.* (2023) ML-Decoder: Scalable and Versatile Classification Head. 2023 *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, 2-7 January 2023, 32-41. <https://doi.org/10.1109/WACV56688.2023.00012>
- [11] Vaswani, A., *et al.* (2017) Attention Is All You Need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 6000-6010.
- [12] Ridnik, T., Ben-Baruch, E., Zamir, N., *et al.* (2021) Asymmetric Loss for Multi-Label Classification. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, 10-17 October 2021, 82-91. <https://doi.org/10.1109/ICCV48922.2021.00015>
- [13] Wagner, P., Strodthoff, N., Boussetjot, R.-D., *et al.* (2020) PTB-XL, a Large Publicly Available Electrocardiography Dataset, *Scientific Data*, **7**, Article No. 154. <https://doi.org/10.1038/s41597-020-0495-6>
- [14] Śmigiel, S., Pałczyński, K. and Ledziński, D. (2021) ECG Signal Classification Using Deep Learning Techniques Based on the PTB-XL Dataset. *Entropy*, **23**, Article 1121. <https://doi.org/10.3390/e23091121>

-
- [15] Śmigiel, S., Pałczyński, K. and Ledziński, D. (2021) Deep Learning Techniques in the Classification of ECG Signals Using R-Peak Detection Based on the PTB-XL Dataset. *Sensors*, **21**, Article 8174. <https://doi.org/10.3390/s21248174>
- [16] Pałczyński, K., Śmigiel, S., Ledziński, D., *et al.* (2022) Study of the Few-Shot Learning for ECG Classification Based on the PTB-XL Dataset. *Sensors*, **22**, Article 904. <https://doi.org/10.3390/s22030904>
- [17] Jyotishi, D. and Dandapat, S. (2023) An Attentive Spatio-Temporal Learning-Based Network for Cardiovascular Disease Diagnosis. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, **53**, 4661-4671. <https://doi.org/10.1109/TSMC.2023.3257022>
- [18] Zhang, D., Yang, S., Yuan, X., *et al.* (2021) Interpretable Deep Learning for Automatic Diagnosis of 12-Lead Electrocardiogram. *iScience*, **24**, Article 102373. <https://doi.org/10.1016/j.isci.2021.102373>
- [19] Geng, Q., Liu, H., Gao, T., *et al.* (2023) An ECG Classification Method Based on Multi-Task Learning and CoT Attention Mechanism. *Healthcare*, **11**, Article 1000. <https://doi.org/10.3390/healthcare11071000>
- [20] Chen, T.-M., Huang, C.-H., Shih, E.S.C., *et al.* (2020) Detection and Classification of Cardiac Arrhythmias by a Challenge-Best Deep Learning Neural Network Model. *iScience*, **23**, Article 100886. <https://doi.org/10.1016/j.isci.2020.100886>
- [21] Ge, Z., Jiang, X., Tong, Z., *et al.* (2021) Multi-Label Correlation Guided Feature Fusion Network for Abnormal ECG Diagnosis. *Knowledge-Based Systems*, **233**, Article 107508. <https://doi.org/10.1016/j.knosys.2021.107508>