

基于注意力机制与卷积神经网络的不均衡多标签心电图分类方法研究

张诗雨¹, 孙占全², 顾浩丞², 陈骏立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6月12日; 录用日期: 2024年7月5日; 发布日期: 2024年7月12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重要。本文提出了一个多标签心电图分类模型, 包括卷积神经网络、Transformer和注意力机制等模块, 其中卷积神经网络用于提取局部特征, Transformer用于提取全局特征, 采用非对称损失函数平衡正负标签。对两个心电数据集进行实验, 实验结果表明, 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, 注意力机制, Transformer, 深度学习

Attention-Based Convolutional Neural Network for Imbalanced Multi-Label Electrocardiogram

Shiyu Zhang¹, Zhanquan Sun², Haocheng Gu², Junli Chen²

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: Jun. 12th, 2024; accepted: Jul. 5th, 2024; published: Jul. 12th, 2024

文章引用: 张诗雨, 孙占全, 顾浩丞, 陈骏立. 基于注意力机制与卷积神经网络的不均衡多标签心电图分类方法研究[J]. 建模与仿真, 2024, 13(4): 4317-4334. DOI: 10.12677/mos.2024.134391

Abstract

Cardiovascular disease is one of the leading causes of death worldwide. Early detection of risk factors for heart disease is essential to prevent sudden death. The automatic detection of arrhythmia based on electrocardiogram is an important means of heart disease screening, which is crucial to reduce the mortality of heart disease. In clinical settings, patients may exhibit multiple abnormal categories rather than just one. Therefore, it becomes necessary to address the issue of multiple labels in electrocardiogram data and consider the relationships between these labels. We propose a multi-label classification model composed of convolutional neural networks, Transformers and attention mechanisms. The convolutional neural network is used to extract local features, Transformer is used to extract global features, and an asymmetric loss function is used to balance positive and negative labels. Experimental results on two electrocardiogram dataset show that the proposed model has better performance than the existing electrocardiogram classification methods.

Keywords

Electrocardiogram Classification, Multi-Label, Attention, Transformer, Deep Learning

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

心血管疾病是全球人口死亡的首要原因，占据了总死亡人数的31%以上[1]。统计数据表明，猝死在心脏病死亡病例中占比超过80%，其中一半以上的病例与心律失常有关[1] [2]。心电图(Electrocardiography, ECG)是一种记录心脏电活动的技术，能够准确反映心脏在每个周期中的电信号。心脏的电活动起源于窦房结，通过传导系统刺激心房和心室，从而引发心脏的收缩和泵血功能。ECG中的波形变化是由于电位的连续变化所引起的，因此，ECG提供了关于心脏节律和形态的信息，并且有助于识别心脏电信号异常。

诊断心律失常通常需要经过训练有素且具备临床经验的心脏病专家进行诊断，但自动监测心律失常和诊断系统在诊所、社区医疗中心和家庭健康护理计划中的需求不断增加。心电图包含了大量有价值的信息，能够直接反映心脏的生理状态，包括心脏电信号和结构变化的形态和时间特征[3]。

基于多个公开的心电数据，学者提出了许多基于机器学习的心律失常自动检测方法。例如，Beyli等人提出了一种利用支持向量机进行ECG检测的方法[4]。Pan等人开发了一种实时算法，用于检测ECG信号中的QRS复合体[5]。此外，Jambukia等人总结了多项使用机器学习方法的实验，发现人工神经网络在心电信号分类中越来越受欢迎[6]。

不同于传统的机器学习方法，近年来学者提出了许多使用深度学习模型进行心律失常自动识别的网络模型。Ge等人提出了一种算法，将残差卷积模块和挤压激励块结合起来，从原始心电图信号中提取特征[7]。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(Time-Incremental Convolutional Neural Network, TI-CNN)，旨在为心电图模型提供更大的输入长度灵活性[8]。后来，他们提出了一种名为基于注意力机制的时间增量卷积神经网络(Attention-based Time-Incremental Convolutional Neural Network, ATI-CNN)的新模型，以TI-CNN作为参考[9]，通过整合卷积神经网络(Convolutional Neural Network, CNN)、循环单元和注意力机制模块，实现对心电信号信息的空间和时间特征融合。Zhang等人提出了一

种空间-时间注意力的卷积循环神经网络,将空间-时间注意力机制融入卷积循环神经网络,以便在空间和时间两个维度上专注于代表性的特征[10]。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,该方法包含残差卷积模块和双向长短期记忆层模块[11]。这些研究均验证了基于卷积神经网络的模型在诊断心律失常中的有效性,并且也证明深度学习网络在未来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,并取得了有效的成果,但如何实现基于心电图的高精度心律失常识别仍是待解难题,特别是目前大部分研究都将心电图分类任务视为单标签分类任务,即一条心电图数据对应一种心电异常类型。而实际临床上,一个患者通常同时包含多种异常类型,即多个标签,例如在 CPSC 2018 心电数据集中[12],有些心电信号就同时包含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不准确。为了解决心电图分类任务的挑战,最近,一些学者提出了一些多标签分类的模型。例如,Prabhakararao 等人提出了一种基于注意力的时间卷积神经网络(Attention-based Temporal Convolutional Neural Network, ATCNN)模型,可以用于 12 导联心电图数据的多标签分类[13],Sheikh 等人提出了一个基于注意力机制模块的多标签分类模型[14]。此外,Chen 等人发现右束支传导阻滞通常伴随着早搏性心房收缩和早搏性室性收缩,除非是心房颤动[15]。但以上模型没有详细分析标签之间的相关性,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。

为提高模型的性能,本文提出的模型利用标签之间的相关性更好地捕捉心电图不同心律失常之间的信息。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,以提高分类性能。本论文提出了一个基于 Transformer 和卷积神经网络的多标签分类模型(Multi-Label Classification Model based on Transformer and Convolutional Neural Networks, MLCT-CNN),总结起来,本研究有以下几点主要贡献。

- 1) 本论文提出的深度学习模型结合了卷积神经网络、注意力机制和非对称损失函数等模块,通过卷积模型提取心电图局部特征,通过注意力集中提取有区别性特征,非对称损失函数有效处理数据不均衡问题。
- 2) 本研究使用卷积神经网络从心电数据中学习标签嵌入的特征,使用 Transformer 的解码器结构分析标签嵌入之间的相关性,从而有效提高多标签心电图分类性能。
- 3) 在两个公开数据集上进行实验,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现有的分类方法。

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 $Y=[Y_0,Y_1,Y_2,\dots,Y_g,\dots,Y_{G-1}]$,其中 $Y_g \in \{0,1\},g=0,1,\dots,G-1$ 。这里, $Y_g(0 \leq g \leq G-1)$ 对应于一个心律失常的真实标签。如果 ECG 记录 X_i 含有第 g 类异常标签,则 Y_g 等于 1,否则, Y_g 等于 0,如下表达式所示:

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

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 $P = [P_0, P_1, P_2, \dots, P_g, \dots, P_{G-1}]$ ，其中 P_g 是模型预测的第 g 个类别的概率， G 是 ECG 数据集类别的数量。在训练期间，模型的目标是最小化预测概率 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

本文所提出的模型 MLCT-CNN 如图 1 所示,包括卷积神经网络(Convolutional Neural Network, CNN)、CBAM (Convolutional Block Attention module) [16]和解码器分类块三个模块。该模型充分考虑了心律失常检测中的关键特征,包括局部波形特征和全局特征[17]。模型主要包括两个部分,第一部分卷积神经网络被用作提取局部特征的基础架构,第二部分使用注意力机制区分特征的不同通道与不同空间的特征的重要性,使用解码器分类块来提取全局特征,并分析标签嵌入特征之间的相关性,并隐式地捕捉不同标签之间的关系。

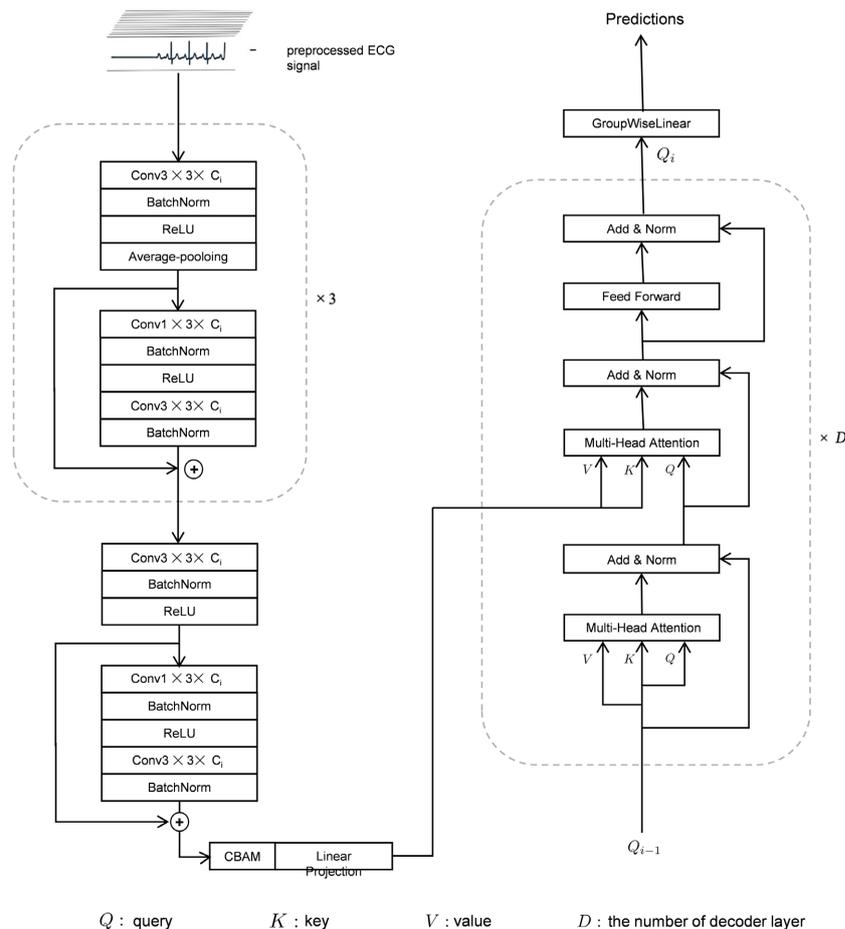


Figure 1. The framework of our proposed MLCT-CNN
 图 1. 本论文所提出的 MLCT-CNN 模型框架

2.2.1. 主干网络模块

受到残差网络[18]的启发,为了防止由于模型深度过大导致模型的分类型性能下降,我们采用卷积神经

网络来提取局部特征,如图2所示。该主干网络包括两类模块,第一种卷积模块由3个卷积层、一个平均池化层、若干个归一化层(Batch Normalization, BN)层和修正线性单元层(Rectified Linear Unit, ReLU)组成,与第一个模块相比,第二个模块缺少一个池化层。第一个卷积模块重复3次,通道数分别为 C_1 、 C_2 和 C_3 。通过两种不同卷积模块进一步提取心电图数据的特征。每个卷积层后面都跟着一个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[19]。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受到认可[20]。在主干网络中,将心电图信号作为主干网络的输入,用于提取局部特征。给定输入 $X \in \mathbb{R}^{1 \times 12 \times L}$,通过卷积网络提取特征为 $F \in \mathbb{R}^{C_3 \times 12 \times L}$ 。

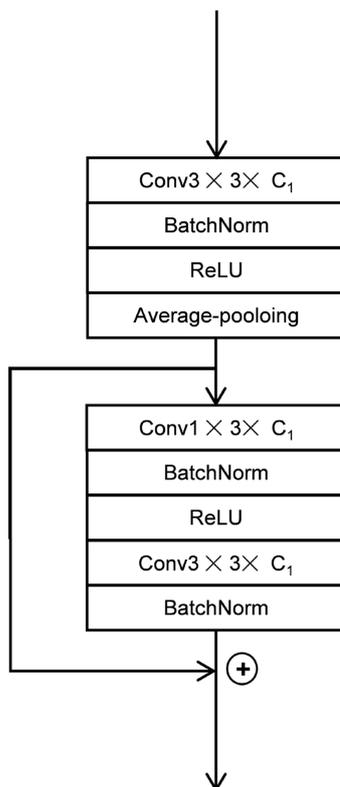


Figure 2. The framework of backbone network
图2. 主干网络结构示意图

2.2.2. 注意力机制模块

模型使用了多个不同的注意机制模块:通道注意机制、空间注意机制和Transformer的解码器模块。图1中的CBAM由通道注意力机制和空间注意力机制组成,图3和图4详细阐述了模型的结构。首先,通道注意力机制通过对特征通道不同程度的重要性进行不同的选择,得到通道注意力系数,而空间注意力机制是在特征中选择不同重要性的特征来获得空间注意力系数,输入特征 F 通过这两个模块得到详细特征 F'' 。这两步分别用式(2)和(3)表示:

$$F' = M_C(F) \otimes F \quad (2)$$

$$F'' = M_S(F') \otimes F' \quad (3)$$

其中, \otimes 表示对应元素相乘, $M_C(F)$ 表示通道注意力模块(Channel Attention Module)产生的系数,而 $M_S(F')$ 表示空间注意力产生模块(Spatial Attention Module)的系数。

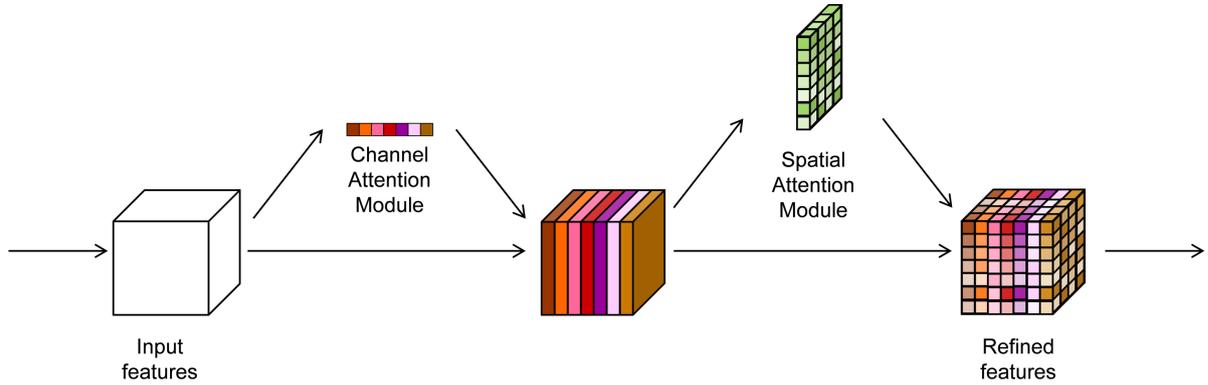


Figure 3. The framework of CBAM
图 3. CBAM 模块的内部结构

如图 4(a)所示, 通道注意力模块首先对输入特征进行通道维度的最大池化和平均池化操作, 得到两个一维向量: $F_{avg}^C \in C \times 1 \times 1$ 以及 $F_{max}^C \in C \times 1 \times 1$ 。这两个向量随后作为共享权重的多层感知机(Multi-Layer Perceptron, MLP)的输入。MLP 对这两个向量进行处理, 然后将处理后的结果相加, 并经过 Sigmoid 函数映射, 最终得到通道注意力权重。这个权重反映了每个通道在特征表达中的重要程度, 具体计算方法如式(4)所示。

$$M_C(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) = \sigma(W_1(W_0(F_{avg}^C)) + W_1(W_0(F_{max}^C))) \quad (4)$$

其中, σ 表示激活函数, $W_0 \in \mathbb{R}^{\frac{C}{r} \times C}$, $W_1 \in \mathbb{R}^{\frac{C}{r} \times C}$ 。

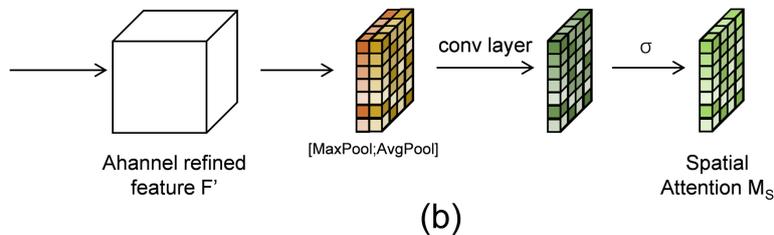
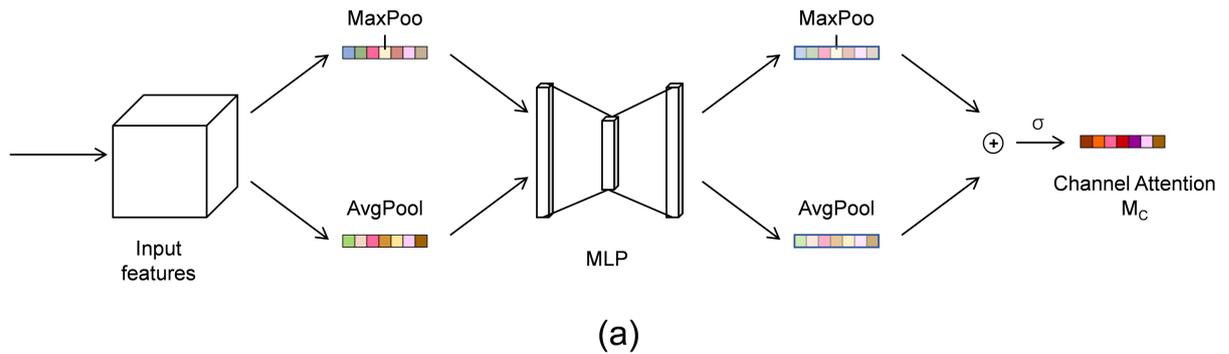


Figure 4. (a) The internal structure of channel attention mechanism; (b) The internal structure of spatial attention mechanism
图 4. (a) 通道注意力的内部结构; (b) 空间注意力的内部结构

如图 4(b)所示, 对于空间注意力模块, 输入特征同样进行空间维度的最大池化和平均池化操作, 生成两个矩阵: $F_{avg}^S \in 1 \times H \times W$ 以及 $F_{max}^S \in 1 \times H \times W$ 。这两个矩阵在通道维度上进行串联, 形成一个 $2 \times H \times W$

的特征映射。这个特征映射随后通过一个卷积层进行处理，卷积层的输出通道数为 1，从而得到一个空间注意力权重矩阵。这个矩阵的每个元素表示对应空间位置的重要程度，有助于模型聚焦于关键区域。整个空间注意力权重的计算过程可以用式(5)表示。

$$M_s(F) = \sigma\left(f^{7 \times 7}\left(\left[\text{AvgPool}(F); \text{MaxPool}(F)\right]\right)\right) = \sigma\left(f^{7 \times 7}\left(\left[F_{avg}^S; F_{max}^S\right]\right)\right) \quad (5)$$

其中， σ 表示激活函数， $f^{7 \times 7}$ 表示卷积核大小为 7×7 的卷积操作。通道注意力机制聚焦于对 ECG 分类更为关键的特征通道，而空间注意力主要聚焦于特征的空间特征。两种不同的注意力机制对 ECG 均有帮助，同时使用两种模型更能提高模型分类结果，通道注意力机制以及空间注意力机制不改变特征的形状。在通道空间注意力之后是一个线性投影层(Linear Projection Layer)。该层将特征投影到与下一阶段中解码器模块所需的查询维度匹配，将投影后的特征重新调整。

经过通道空间注意力得到标签嵌入特征，包含标签的相关信息。为自动提取全局特征并分析标签嵌入之间的关系，引入了解码器分类块[21]。如图 1 所示，将标签嵌入输入多个 Transformer 解码器层(见图 5)和一个线性投影层组成，该模块可用于提取全局特征，并分析标签嵌入之间的相关性，从而预测每个类别的概率作为模型的输出。

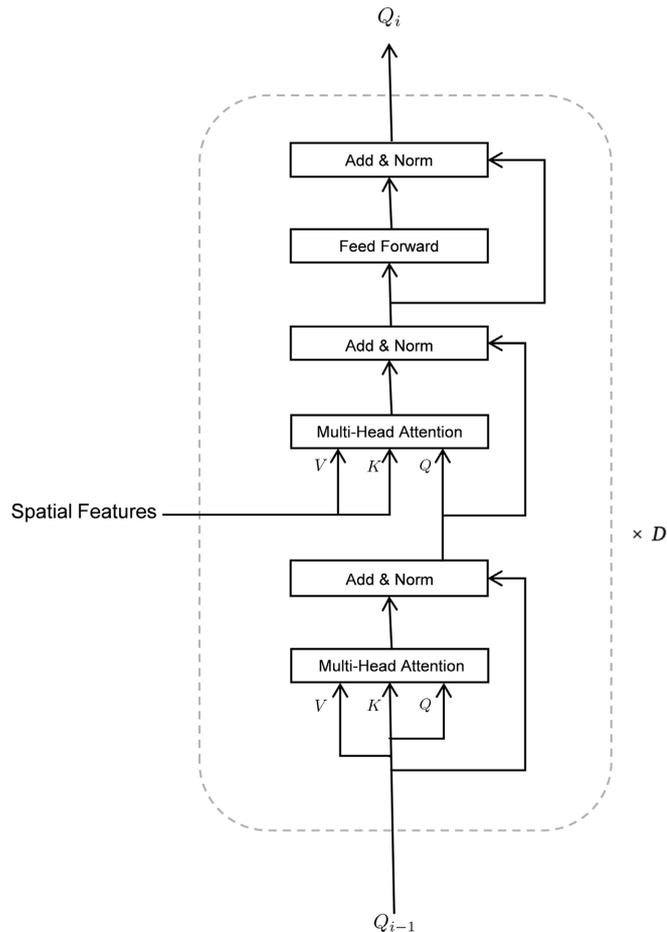


Figure 5. The internal architecture of MLCB

图 5. MLCB 的内部结构

在 Transformer 解码器中，使用标签嵌入特征作为查询 $Q_i \in \mathbb{R}^{C_3 \times \text{hidden_dim}}$ ，第 i 层 Transformer 解码器

根据其前一层的输出更新查询(query)。解码器分类块包含自注意力层(self-attention layer)、交叉注意力层(cross-attention layer)和前馈神经网络层(Feed-Forward Networks layer, FFN), 具体如下:

$$\text{self-attention} : Q_i^{(1)} = \text{MultiHead}(Q_{i-1}, Q_{i-1}, Q_{i-1}) \quad (6)$$

$$\text{cross-attention} : Q_i^{(2)} = \text{MultiHead}(Q_i^{(1)}, F'', F'') \quad (7)$$

$$\text{FFN} : Q_i = \text{FFN}(Q_i^{(2)}) \quad (8)$$

其中, 自注意力和交叉注意力模块都使用相同的多头注意力函数 **MultiHead** 实现, 多头注意力函数的计算如式(9), 将 h 个基础注意力机制模块的输出进行拼接得到多头注意力机制的输出。

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O \text{ where } \text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (9)$$

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (10)$$

式中, $W_i^Q \in \mathbb{R}^{d_{\text{model}} \times d_k}$, $W_i^K \in \mathbb{R}^{d_{\text{model}} \times d_k}$, $W_i^V \in \mathbb{R}^{d_{\text{model}} \times d_v}$, $W_i^O \in \mathbb{R}^{hd_k \times d_{\text{model}}}$, h 表示多头注意力机制中基础注意力机制模块的个数。注意力机制模块首先将输入的 Q 与 K 的转置进行相乘, 之后用 d_k 以及 **softmax** 操作进行归一化得到权重, 乘以 V , 得到单个注意力机制的输出, 如式(6)。在自注意力模块中, **query**、**key** 和 **value** 都来自标签嵌入 Q_{i-1} , 而在交叉注意力模块中, **key** 和 **value** 来自卷积网络的空间特征 F'' 。自注意力模块可以通过计算标签嵌入特征之间的相似度分析标签嵌入特征之间的依赖关系, 而交叉注意力机制通过计算先后生成的标签嵌入之间的相似度计算其相关性。

在 D 层 **Transformer** 解码器层之后是 **GroupWiseLinear** 层, 得到模型的预测结果 $P \in \mathbb{R}^{G}$, 表示模型预测到的 G 个心律失常类别的概率。

$$P_i = \sum_{j=1}^d (Q_{D(i,j)} \times w_j + b_i) \quad (i \in 0, 1, \dots, G-1) \quad (11)$$

其中 w_j 和 b_i 是模型的参数。

2.2.3. 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) [22], 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函数相比, 非对称损失函数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:

$$L = \frac{1}{G} \sum_{g=1}^G \begin{cases} (1 - P_g)^{\gamma^+} \log(P_g), Y_g = 1, \\ (P_g)^{\gamma^-} \log(1 - P_g), Y_g = 0. \end{cases} \quad (12)$$

其中, P_g 是预测类别的概率, Y_g 是一个二进制标签, 用于表示心电图记录 X 的实际标签中是否含有标签 g 。 γ^- 和 γ^+ 分别为负聚焦和正聚焦的参数, 通过对这两个参数的设置, 可以更好地控制正样本和负样本对损失函数的贡献, 并帮助模型从正样本中学习有意义的特征。

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**、**8GB** 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

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 6 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。

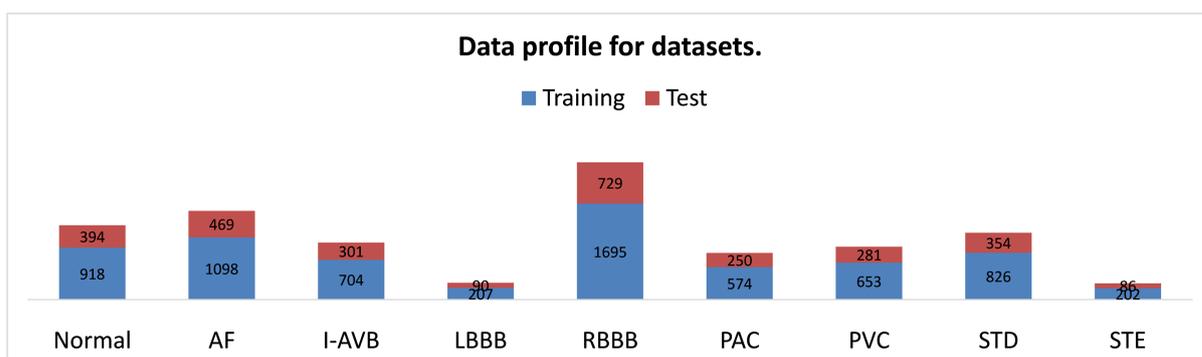


Figure 6. Data profile for CPSC 2018
图 6. CPSC 2018 心电数据集的数据概况

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，该多标签心电图数据集一共有 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 心电图数据集的详细信息见图 7。本文使用 PTB-XL 心电图数据集推荐的训练集和测试集进行实验，其中，训练集数据 19267 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。

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 CPSC 2018 心电数据集分成训练集和测试集，比例为 8:2，测试集和训练集数据不重合。而针对 PTB-XL 心电图数据集，本文使用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。

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[24]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。

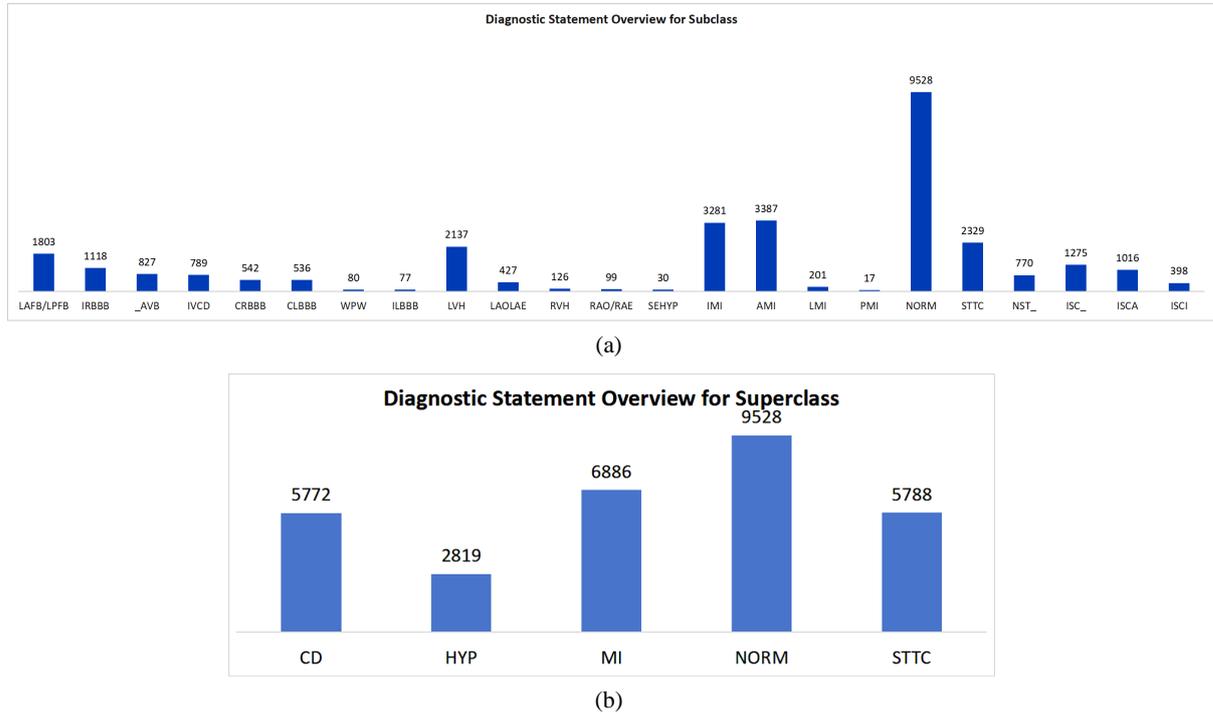


Figure 7. Diagnostic statement overview for subclass (a) and superclass (b) of PTB-XL
图 7. PTB-XL 心电图数据集子类(a)及超类(b)诊断陈述概述

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，包括准确率、精确度、召回率以及 F1 分数，分类指标都包含各个类别和整体的分类情况。

准确率是正确预测数量与总预测数量的比率，准确率代表了该算法检测出真实样本占有所有样本的比率，表示模型的正确预测，准确率越高，则该算法检测出样本真实情况的能力越强。可以使用以下公式来计算这个指标：

$$\text{Accuracy}_g = \frac{\text{TP}_g + \text{TP}_g}{\text{TP}_g + \text{FP}_g + \text{TP}_g + \text{TP}_g} \quad (13)$$

$$\text{Accuracy} = \frac{\sum_{g=0}^{G-1} \text{Accuracy}_g}{G} \quad (14)$$

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：

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

$$\text{Precision} = \frac{\sum_{g=0}^{G-1} \text{Precision}_g}{G} \quad (16)$$

召回率衡量的是被正确分类为正类的样本数量与总正类样本数量的比例。这个指标可以通过以下公式来计算，它可以反映模型在将正类样本分类的能力：

$$\text{Recall}_g = \frac{\text{TP}_g}{\text{TP}_g + \text{FN}_g} \quad (17)$$

$$\text{Recall} = \frac{\sum_{g=0}^{G-1} \text{Recall}_g}{G} \quad (18)$$

F1 值是精确率和召回率的加权平均值：

$$F1_g = \frac{2 \times \text{TP}_g}{2 \times \text{TP}_g + \text{FP}_g + \text{FN}_g} \quad (19)$$

$$F1 = \frac{\sum_{g=0}^{G-1} F1_g}{G} \quad (20)$$

其中， G 是数据集分类的数量， g 表示不同类别，TP (True Positive)表示真实值和预测值都是正样本，即真阳性；TN (True Negative)表示真实值和预测值都是负样本，即真阴性；FP (False Positive)表示真实值为负样本，预测值为正样本，即假阳性；FN (False Negative)表示真实值为正样本，预测值为负样本，即假阴性。

针对 CPSC 2018 心电数据集，一共有 9 个标签，1 个标签为正常，其余 8 个标签为心律失常类型，心律失常可分为四个大类：房颤(AF)、传导阻滞(Block)、早搏(PC)以及 ST 段改变，分类指标的计算分别如式(21)、(22)、(23)以及(24)。

$$F_{af} = \frac{2 \times \text{TP}_2}{N_{2x} + N_{x2}} \quad (21)$$

$$F_{block} = \frac{2 \times (\text{TP}_3 + \text{TP}_4 + \text{TP}_5)}{N_{3x} + N_{x3} + N_{4x} + N_{x4} + N_{5x} + N_{x5}} \quad (22)$$

$$F_{pc} = \frac{2 \times (\text{TP}_6 + \text{TP}_7)}{N_{6x} + N_{x6} + N_{7x} + N_{x7}} \quad (23)$$

$$F_{st} = \frac{2 \times (\text{TP}_8 + \text{TP}_9)}{N_{8x} + N_{x8} + N_{9x} + N_{x9}} \quad (24)$$

其中， F_{af} 是 AF 的 F1 值； F_{block} 是阻滞大类的 F1 值，包括 I-AVB、LBBB 和 RBBB 三种类型的心律失常； F_{pc} 为早搏大类的 F1 值，其中早搏包含 PAC 和 PVC； F_{st} 是 ST 段变化的 F1 值，其中包括 STD 和 STE。变量数 N_{ij} 的计数规则见表 1。

3.5. 仿真结果

3.5.1. CPSC 2018 心电数据集仿真结果

本模型的分类结果与 CPSC 2018 比赛排名前 10 的结果进行了比较，比较结果如表 2 所示，本文的模型分类结果在所有模型中排名第二，尤其在包含 I-AVB、LBBB 和 RBBB 三种心律失常类型的 block

子类别中，该模型子类 F1 分数最高，比第二名的 F1 分数高 0.013，而整体的 F1 值与最高 F1 值相比，本文提出的模型稍低 0.002。然而，Chen 等人整合了 130 个模型来构建一个复杂且庞大的集成模型[15]，由于这种复杂的架构，预测效率较低，并且需要更多的计算成本。

Table 1. Counting rules for the numbers of the variables

表 1. 变量数量的计数规则

		预测									
		Normal	AF	I-AVB	LBBB	RBBB	PAC	PVC	STD	STE	Total
标签	Normal	N_{11}	N_{12}	N_{13}	N_{14}	N_{15}	N_{16}	N_{17}	N_{18}	N_{19}	N_{1x}
	AF	N_{21}	N_{22}	N_{23}	N_{24}	N_{25}	N_{26}	N_{27}	N_{28}	N_{29}	N_{2x}
	I-AVB	N_{31}	N_{32}	N_{33}	N_{34}	N_{35}	N_{36}	N_{37}	N_{38}	N_{39}	N_{3x}
	LBBB	N_{41}	N_{42}	N_{43}	N_{44}	N_{45}	N_{46}	N_{47}	N_{48}	N_{49}	N_{4x}
	RBBB	N_{51}	N_{52}	N_{53}	N_{54}	N_{55}	N_{56}	N_{57}	N_{58}	N_{59}	N_{5x}
	PAC	N_{61}	N_{62}	N_{63}	N_{64}	N_{65}	N_{66}	N_{67}	N_{68}	N_{69}	N_{6x}
	PVC	N_{71}	N_{72}	N_{73}	N_{74}	N_{75}	N_{76}	N_{77}	N_{78}	N_{79}	N_{7x}
	STD	N_{81}	N_{82}	N_{83}	N_{84}	N_{85}	N_{86}	N_{87}	N_{88}	N_{89}	N_{8x}
	STE	N_{91}	N_{92}	N_{93}	N_{94}	N_{95}	N_{96}	N_{97}	N_{98}	N_{99}	N_{9x}
	Total	N_{x1}	N_{x2}	N_{x3}	N_{x4}	N_{x5}	N_{x6}	N_{x7}	N_{x8}	N_{x9}	

Table 2. Comparison of the proposed model and related work on CPSC 2018

表 2. 本论文提出的模型与比较模型在 CPSC 2018 心电数据集的实验结果比较

		F_{af}	F_{block}	F_{pc}	F_{st}	$F1$
Rank1	Chen	0.933	0.899	0.847	0.779	0.837
Rank2	Cai	0.931	0.912	0.817	0.761	0.830
Rank3	He	0.914	0.879	0.801	0.742	0.806
Rank4	Yu	0.918	0.890	0.789	0.718	0.802
Rank5	Yan	0.924	0.882	0.779	0.709	0.791
Rank6	Wang	0.905	0.902	0.722	0.708	0.783
Rank7	Lin	0.911	0.891	0.775	0.670	0.782
Rank8	Shen	0.921	0.858	0.797	0.676	0.778
Rank9	Mao	0.906	0.876	0.773	0.711	0.776
Rank10	Zhang	0.894	0.857	0.733	0.683	0.766
	本论文	0.907	0.925	0.801	0.712	0.835

本节还在 CPSC 2018 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 列出了不同方法的各项性能指标，粗体显示的值表示相应指标的最佳分数。与 TI-CNN[8]和 ATI-CNN [9]相

比, MLCT-CNN 分类结果 F1 值分别提升了约 0.063 和 0.023。与文献[11]、文献[25]、文献[26]、文献[27]、文献[28]、文献[7]中提出模型的分类结果相比, 本文提出的模型 F1 分别提高了约 0.029、0.022、0.017、0.008、0.008 和 0.007。特别是对于正常、LBBB 和 RBBB 类别, 本文提出的模型的 F1 值均高于这些模型。总之, 相比于之前提出的用于心电数据集分类的模型, MLCT-CNN 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 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
TI-CNN [8]	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> [11]	0.755	0.846	0.780	0.790	0.748	0.829	0.869	0.870	0.655	0.806
ATI-CNN [9]	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> [25]	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> [26]	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> [27]	0.824	0.925	0.882	0.937	0.939	0.734	0.767	0.835	0.600	0.827
MLC-CNN [28]	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> [7]	0.790	0.920	0.870	0.870	0.930	0.780	0.860	0.810	0.590	0.828
本论文提出的模型	0.808	0.907	0.892	0.937	0.945	0.754	0.848	0.805	0.619	0.835

3.5.2. PTB-XL 心电数据集仿真结果

本章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分类性能, 如表 4 所示。本文所提出的模型针对超类诊断(super diagnosis)的 F1 值为 0.762, 准确度为 0.898, 精确度为 0.833, 召回率为 0.722。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、准确率和召回率均高于现有的模型。由这些实验结果表明, 该模型在 CPSC 2018 心电数据集和 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异常分类方面是有效的。

Table 4. Comparison of results of the proposed model and those of the existing model on PTB-XL

表 4. 本文提出的模型与现有模型在 PTB-XL 数据集上的分类结果

模型	分类级别	Accuracy	Precision	Recall	F1
Smigiel <i>et al.</i> [29]		0.698	0.355	0.339	0.332
Smigiel <i>et al.</i> [30]		0.685	-	-	0.345
Palczynski <i>et al.</i> [31]	sub diagnosis	0.671	-	-	0.324
Jyotishi <i>et al.</i> [26]		0.541	-	-	0.460
本论文		0.967	0.544	0.464	0.473
Smigiel <i>et al.</i> [29]		0.765	0.714	0.662	0.680
Smigiel <i>et al.</i> [30]		0.763	-	-	0.683
Palczynski <i>et al.</i> [31]	super diagnosis	0.790	-	-	0.717
Jyotishi <i>et al.</i> [26]		0.627	-	-	0.736
本论文		0.898	0.833	0.722	0.762

3.6. 消融实验

a) 关于非对称损失函数参数的选择, 本文在两个不同的心电数据集上进行了详尽的实验。在 SPSC 2018 心电数据集上, 选取两个固定的 γ^- 值, 即 2 和 4, 实验结果如图 8(a)所示。分析表明, 当 γ^+ 增加时, F1 值呈现下降趋势, 然而, 当 γ^- 设置为 4, 且 γ^+ 从 0 增加到 1 时, F1 值却有所提升。当 $\gamma^+ = 1$ 且 $\gamma^- = 4$ 时, 模型在 SPSC 2018 心电数据集上的分类效果达到最优。

此外, 在 PTB-XL 心电数据集上, 本文对非对称损失的参数 γ^- 和 γ^+ 进行了对比实验, 结果如图 8(b)所示, 该模型在超级诊断方面, 当 $\gamma^- = 2$ 且 $\gamma^+ = 0$ 时, 模型达到了最佳的 F1 值 0.762。综上所述, 不同的心电数据集在非对称损失函数参数选择上存在差异。为了获得最佳的分类效果, 需要根据具体的数据集进行实验, 以确定最佳的参数设置。

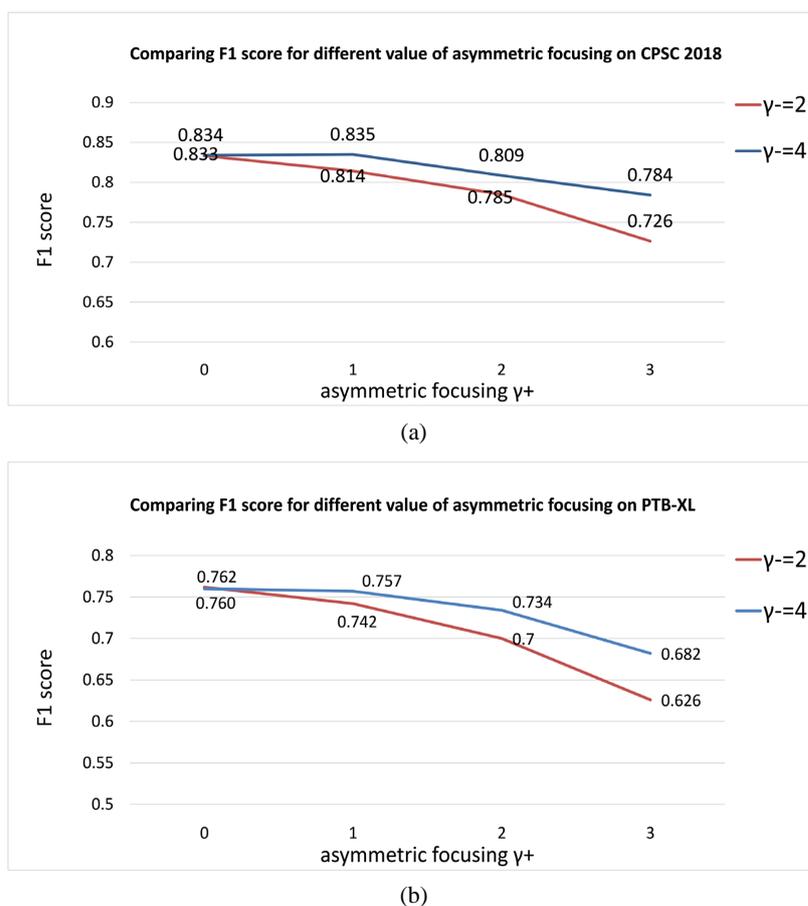


Figure 8. Comparing F1 score for different value of asymmetric focusing γ^+ , for $\gamma^- = 2$ and $\gamma^- = 4$ on (a) CPSC 2018 and (b) PTB-XL

图 8. 不同的正向聚焦因子 (γ^+) 值在 (a) CPSC2018 和 (b) PTB-XL 心电数据集上的 F1 对比, 其中 $\gamma^- = 2$ 和 $\gamma^- = 4$

b) 为了验证 CBAM 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 CBAM 和 MLCB 的两种网络变体进行比较。该模型由三个部分组成: 主干网络、CBAM 和解码器模块。用于比较的三个模型的结构如下:

模型 1: 该模型仅使用卷积神经网络作为主干网络进行特征提取, 设计这个模型是为了验证 CBAM 和解码器模块的作用。

模型 2: 该模型不包含解码器模块, 设计这个模型是为了验证 CBAM 的效果。

模型 3: 该模型不包含 CBAM。这个模型是为了验证解码器模块的效果。

表 5 对比了几个参考模型和本文提出模型的 F1 值, 模型 2 的 F1 值比模型 1 的高出 0.187, 而模型 3 的 F1 值比模型 1 高出 0.224。模型 2 包含 CBAM, 而模型 3 包含解码器。

Table 5. The F1 score of our proposed model and the reference models on CPSC 2018

表 5. 本文提出的模型和消融实验各个模型在 CPSC 2018 心电数据集上的 F1 分类结果

模型	Normal	AF	I-AVB	LBBB	RBBB	PAC	PVC	STD	STE	F1
模型 1	0.586	0.721	0.509	0.843	0.865	0.225	0.729	0.584	0.321	0.598
模型 2	0.705	0.884	0.826	0.889	0.925	0.354	0.759	0.723	0.526	0.785
模型 3	0.799	0.894	0.886	0.902	0.941	0.712	0.831	0.804	0.628	0.822
本论文提出的模型	0.808	0.907	0.892	0.937	0.945	0.754	0.848	0.805	0.619	0.835

3.7. 方法对比

在本节中, 我们对比了不同注意力机制模块的作用。除了 CBAM 模块[16], 我们对比了 CA (Coordinate Attention)模块[32], SE (Squeeze and Excitation)模块[33], SCSE (Concurrent Spatial and Channel Squeeze & Excitation)模块[34]以及 GMAP (Global Max Pooling and Global Average Pooling)模块[7]。与这几种注意力机制相比, 带有 CBAM 模块的模型最终的分分类效果更好, 如表 6 所示, CBAM 可以自适应地增强包含心律失常信息的通道, 抑制不相关的通道, 从而改善模型的分分类性能。

Table 6. Comparison of different attention mechanism on CPSC 2018

表 6. 不同的注意力机制在 CPSC2018 心电数据集上的实验结果比较

注意力机制	Normal	AF	I-AVB	LBBB	RBBB	PAC	PVC	STD	STE	F1
None	0.799	0.894	0.886	0.902	0.941	0.712	0.831	0.804	0.628	0.822
CA	0.788	0.887	0.892	0.902	0.935	0.707	0.850	0.785	0.585	0.815
SE	0.789	0.903	0.898	0.868	0.940	0.691	0.852	0.794	0.621	0.817
SCSE	0.790	0.904	0.886	0.902	0.934	0.724	0.832	0.820	0.575	0.819
本论文提出的模型	0.808	0.907	0.892	0.937	0.945	0.754	0.848	0.805	0.619	0.835

3.8. 仿真结果分析

本文提出了一个能够自动检测心律失常类别且无需预处理的模型, 通过以上的实验结果显示, 我们的方法在心电图记录分类任务中分类结果良好, 在 CPSC 2018 心电数据集中测试集的平均 F1 值达到 0.835, 在 PTB-XL 心电数据集中, 超类诊断(super diagnosis)的 F1 值为 0.762, 子类诊断(sub diagnosis)的 F1 值为 0.473, 均超过了一些现有模型的分分类效果, 证明本文提出的模型在多标签心电图分分类检测的有效性。通过消融实验表明, 本文提出的分分类网络中的每个模块都是必不可少的, 并且都能一定程度上增强模型的分分类效果。在提出的网络中, CBAM 比近几年提出的其他注意力机制模块更有效, 分分类效果更优, 并且 CBAM 模块和解码器多标签分分类模块都提高了模型的分分类效果, 两个模块的组合分分类性能更好。

正如图 9 所示，混淆矩阵是用于评估分类模型性能的一个矩阵，其中是心电数据集标签的种类数，在 CPSC 2018 心电数据集上的混淆矩阵则表明本文提出的模型预测结果与实际的标签存在高度一致性，从而表明模型对多标签心电数据集具有良好的分类效果。

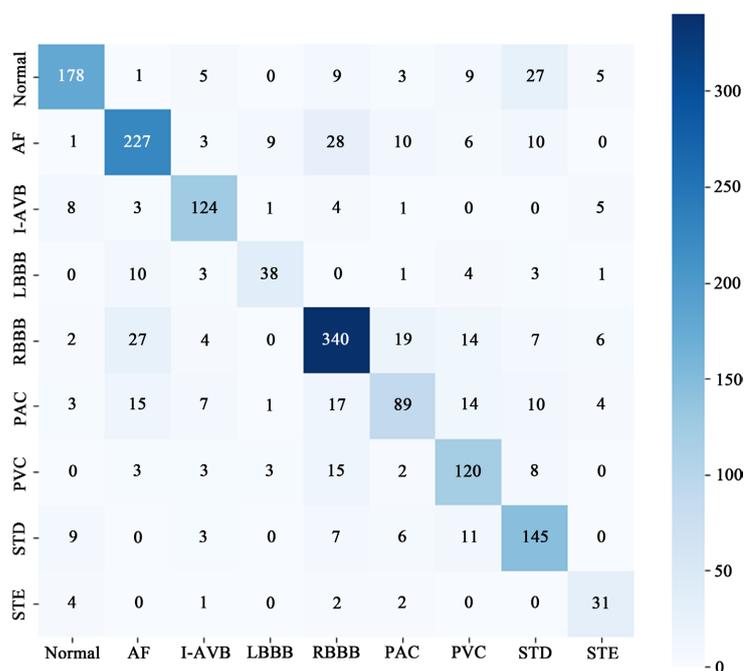


Figure 9. Confusion matrix of the proposed model on the test datasets of CPSC2018
图 9. 在 CPSC 2018 测试数据集上的混淆矩阵

4. 结束语

本文研究了一种深度神经网络架构 MLCT-CNN，该模型融合了卷积神经网络、Transformer、通道注意力机制以及空间注意力机制等先进技术，旨在提升多标签心律失常分类的准确性与效率。本文并未采用原始的 Transformer 结构，而是针对性地选取了 Transformer 的解码器块，专注于分析标签嵌入特征间的相关性，进而优化心电图的多标签分类效果。

在对比实验中，本文提出的 MLCT-CNN 模型在两个公开的心电数据集上均展现出了优越的性能，特别是在 LBBB 和 RBBB 两类心律失常的诊断上表现突出。具体而言，在 CPSC 2018 12 导联心电数据集上，MLCT-CNN 模型取得了 0.835 的 F1 值，充分证明了其在多标签分类任务中的有效性。此外，在 PTB-XL 数据集上，无论是子类诊断还是超类诊断，MLCT-CNN 均展现出了较高的性能，其子类诊断 F1 值为 0.473，超类诊断 F1 值为 0.762，均超过了近年来现有的分类方法，进一步验证了 MLCT-CNN 在多标签心律失常分类中的优越性。

通过对两个心电数据集上的实验结果进行综合分析，可以得出结论：本文提出的 MLCT-CNN 模型在 F1 分数上显著优于其他分类模型，显示出卓越的分类性能。这一结果不仅有效验证了 MLCT-CNN 对于多标签心律失常分类的有效性，同时也为未来的心电图分类研究提供了新的思路和方法。

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] Mendis, S., Puska, P., Norrving, B., *et al.* (2011) Global Atlas on Cardiovascular Disease Prevention and Control. World Health Organization.
- [2] Mehra, R. (2007) Global Public Health Problem of Sudden Cardiac Death. *Journal of Electrocardiology*, **40**, S118-S122. <https://doi.org/10.1016/j.jelectrocard.2007.06.023>
- [3] Hagiwara, Y., Fujita, H., Oh, S.L., Tan, J.H., Tan, R.S., Ciaccio, E.J., *et al.* (2018) Computer-Aided Diagnosis of Atrial Fibrillation Based on ECG Signals: A Review. *Information Sciences*, **467**, 99-114. <https://doi.org/10.1016/j.ins.2018.07.063>
- [4] Übeyli, E.D. (2007) ECG Beats Classification Using Multiclass Support Vector Machines with Error Correcting Output Codes. *Digital Signal Processing*, **17**, 675-684. <https://doi.org/10.1016/j.dsp.2006.11.009>
- [5] Pan, J. and Tompkins, W.J. (1985) A Real-Time QRS Detection Algorithm. *IEEE Transactions on Biomedical Engineering*, **32**, 230-236. <https://doi.org/10.1109/tbme.1985.325532>
- [6] Jambukia, S.H., Dabhi, V.K. and Prajapati, H.B. (2015) Classification of ECG Signals Using Machine Learning Techniques: A Survey. 2015 *International Conference on Advances in Computer Engineering and Applications*, Ghaziabad, 19-20 March 2015, 714-721. <https://doi.org/10.1109/icacea.2015.7164783>
- [7] Ge, R., Shen, T., Zhou, Y., Liu, C., Zhang, L., Yang, B., *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>
- [8] Yao, Q., Fan, X., Cai, Y., Wang, R., Yin, L. and Li, Y. (2018) Time-Incremental Convolutional Neural Network for Arrhythmia Detection in Varied-Length Electrocardiogram. 2018 *IEEE 16th Conference on Dependable, Autonomic and Secure Computing, 16th Conference on Pervasive Intelligence and Computing, 4th Conference 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/cyberscitech.2018.00131>
- [9] Yao, Q., Wang, R., Fan, X., Liu, J. and Li, Y. (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>
- [10] Zhang, J., Liu, A., Gao, M., Chen, X., Zhang, X. and Chen, X. (2020) ECG-Based Multi-Class Arrhythmia Detection Using Spatio-Temporal Attention-Based Convolutional Recurrent Neural Network. *Artificial Intelligence in Medicine*, **106**, Article ID: 101856. <https://doi.org/10.1016/j.artmed.2020.101856>
- [11] He, R., Liu, Y., Wang, K., Zhao, N., Yuan, Y., Li, Q., *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>
- [12] Liu, F., Liu, C., Zhao, L., Zhang, X., Wu, X., Xu, X., *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>
- [13] Prabhakararao, E. and Dandapt, S. (2023) Multi-Label ECG Classification Using Temporal Convolutional Neural Network.
- [14] Sheikh, D., Verma, H. and Chauhan, N. (2024) Reduced Lead ECG Multi-Label Classification with Higher Generalization Using 2D SEResnets with Self Attention. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-024-18116-z>
- [15] Chen, T., Huang, C., Shih, E.S.C., Hu, Y. and Hwang, M. (2020) Detection and Classification of Cardiac Arrhythmias by a Challenge-Best Deep Learning Neural Network Model. *iScience*, **23**, Article ID: 100886. <https://doi.org/10.1016/j.isci.2020.100886>
- [16] Woo, S., Park, J., Lee, J. and Kweon, I.S. (2018) CBAM: Convolutional Block Attention Module. In: *Computer Vision—ECCV 2018*, Springer, 3-19. https://doi.org/10.1007/978-3-030-01234-2_1
- [17] Tseng, K.K., Lee, D. and Chen, C. (2016) ECG Identification System Using Neural Network with Global and Local Features. *International Association for Development of the Information Society*.
- [18] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/cvpr.2016.90>
- [19] Ioffe, S. and Szegedy, C. (2015) Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *International Conference on Machine Learning*, Lille, 6-11 July 2015, 448-456.
- [20] Nair, V. and Hinton, G.E. (2010) Rectified Linear Units Improve Restricted Boltzmann Machines. In: *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, Omnipress, Haifa, 807-814.

-
- [21] Vaswani, A., Shazeer, N., Parmar, N., *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.
- [22] Ridnik, T., Ben-Baruch, E., Zamir, N., Noy, A., Friedman, I., Protter, M., *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>
- [23] Wagner, P., Strodthoff, N., Bousseljot, R., Kreiseler, D., Lunze, F.I., Samek, W., *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>
- [24] Da, K. (2014) A Method for Stochastic Optimization.
- [25] Zhang, D., Yang, S., Yuan, X. and Zhang, P. (2021) Interpretable Deep Learning for Automatic Diagnosis of 12-Lead Electrocardiogram. *iScience*, **24**, Article ID: 102373. <https://doi.org/10.1016/j.isci.2021.102373>
- [26] 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>
- [27] Geng, Q., Liu, H., Gao, T., Liu, R., Chen, C., Zhu, Q., *et al.* (2023) An ECG Classification Method Based on Multi-Task Learning and Cot Attention Mechanism. *Healthcare*, **11**, Article No. 1000. <https://doi.org/10.3390/healthcare11071000>
- [28] Ge, Z., Jiang, X., Tong, Z., Feng, P., Zhou, B., Xu, M., *et al.* (2021) Multi-Label Correlation Guided Feature Fusion Network for Abnormal ECG Diagnosis. *Knowledge-Based Systems*, **233**, Article ID: 107508. <https://doi.org/10.1016/j.knosys.2021.107508>
- [29] Śmigiel, S., Pałczyński, K. and Ledziński, D. (2021) ECG Signal Classification Using Deep Learning Techniques Based on the PTB-XL Dataset. *Entropy*, **23**, Article No. 1121. <https://doi.org/10.3390/e23091121>
- [30] 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 No. 8174. <https://doi.org/10.3390/s21248174>
- [31] Pałczyński, K., Śmigiel, S., Ledziński, D. and Bujnowski, S. (2022) Study of the Few-Shot Learning for ECG Classification Based on the PTB-XL Dataset. *Sensors*, **22**, Article No. 904. <https://doi.org/10.3390/s22030904>
- [32] Hou, Q., Zhou, D. and Feng, J. (2021). Coordinate Attention for Efficient Mobile Network Design. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 13708-13717. <https://doi.org/10.1109/cvpr46437.2021.01350>
- [33] 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>
- [34] Roy, A.G., Navab, N. and Wachinger, C. (2018) Concurrent Spatial and Channel “Squeeze & Excitation” in Fully Convolutional Networks. In: *Medical Image Computing and Computer Assisted Intervention—MICCAI 2018*, Springer International Publishing, 421-429. https://doi.org/10.1007/978-3-030-00928-1_48