

基于深度学习的多标签不均衡心电信号分类

顾浩丞, 孙占全

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1月9日; 录用日期: 2024年3月21日; 发布日期: 2024年3月28日

摘要

心电图(ECG)自动分类技术已经成为医学诊断的热门研究领域。在临床上, 同一个患者通常伴随多种心律失常。为了能够精准分类出多种心律失常, 本文提出了一种涵盖多种注意力机制的一维卷积神经网络模型。相比传统的机器学习或者深度学习方法提取特征的能力有限, 引入的注意力机制能够引导模型学习心电信号中不同通道和空间位置之间的差异信息的同时结合全局稀疏自注意力, 提高分类仿真模型对与心律失常有关特征的关注程度; 其次, 通过重采样技术以及在损失函数的设计上融入了标签相关性策略来应对多标签带来的类别不平衡问题。在两个公开的多标签心电数据集CPSC2018和PTB-XL上评估该网络的性能, macro-F1分别达到84.13%和73.6%。实验结果表明, 该模型具有良好的分类性能。

关键词

心电图, 深度学习, 多标签分类, 卷积神经网络, 注意力机制, 数据不平衡

Multi-Label Imbalanced ECG Signal Classification Based on Deep Learning

Haocheng Gu, Zhanquan Sun

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

Received: Jan. 9th, 2024; accepted: Mar. 21st, 2024; published: Mar. 28th, 2024

Abstract

The technology of automatic Electrocardiogram (ECG) classification has emerged as a hot area of research in medical diagnostics. In clinical practice, a single patient often exhibits multiple types of cardiac arrhythmias. To accurately classify these diverse arrhythmias, this paper introduces a one-dimensional Convolutional Neural Network model that incorporates multiple attention me-

chanisms. Compared with the limited ability of traditional machine learning or deep learning methods to extract features, the introduced attention mechanism can guide the model to learn the difference information between different channels and spatial locations in ECG signals, and combine with global sparse self-attention to improve the classification simulation model's attention to arrhythmia-related features. We address the class imbalance problem brought about by multi-labels through resampling techniques and the integration of label correlation strategies into the design of the loss function. The performance of this network is evaluated on two public multi-label ECG datasets, CPSC2018 and PTB-XL, with macro-F1 reaching 84.13% and 73.6%, respectively. Experimental results indicate that the model demonstrates excellent classification performance.

Keywords

Electrocardiogram, Deep Learning, Multi-Label Classification, Convolutional Neural Network, Attention Mechanism, 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(ECG)是一种有效的无创、低成本的心血管体外监测工具,它可以通过检测和记录心脏电信号活动来诊断一系列的心脏疾病。然而,由于心电信号的形态和时间特征在不同身体状况下的不同患者中存在明显的差异,因此心电信号的分类是一个难题。此外,两种不同的疾病在心电图信号上可能具有大致相同的性质。同时,在整个 ECG 分析过程中,由于医生的疲劳和心电信号诊断需要深入的知识,可能会出现一些不可避免的个人错误,从而给患者带来无法预料的误诊后果。因此,ECG 自动分类已经被广泛应用于临床医学实践中。

ECG 自动分类可以将 ECG 信号分类到不同的心律失常类别中。在临床实践中,已经提出了许多机器学习方法,包括决策树[1]、随机森林[2]和支持向量机[3]。在这些方法中,需要数据预处理和手动特征提取来从原始信号中获得各种特征。近年来,随着深度学习技术的发展和广泛应用,卷积神经网络(CNN)和循环神经网络(RNN)等端到端的深度学习模型已被用于从原始信号中提取高水平特征来实现 ECG 诊断。Rajpurkar 等人[4]设计了一种深度卷积神经网络使用 91,232 个单导联心电图进行 12 种心律失常的分类。Xiong 等人[5]通过使用短时傅里叶变换(STFT)来转换 ECG 信号,将其表示为频谱图,并结合 CNN 模型实现准确的心律失常分类。Gao 等人[6]使用长短期记忆(LSTM)网络在非平衡的 ECG 数据集上检测心律失常,利用 LSTM 捕捉心电信号的长期依赖关系,模型能够准确地识别心律失常的存在并进行分类。

然而,在以前的大部分工作中,ECG 诊断被视为单一标签分类,它们忽略了在临床实践中同一个患者多种心律失常并发的可能性。这意味着一个心电信号记录中可能出现多个心律失常,这使得心律失常的自动诊断具有挑战性。例如,“心房颤动”和“右束支传导阻滞”通常同时出现,但“正常”和“右束支传导阻滞”很少同时出现[7]。Ran 等人[7]认为不同类型的心律失常之间可能存在相关性,采用标签关联性嵌入技术,引导网络在训练过程中学习到更准确的标签表示。Ge 等人[8]引入了包含通道注意力的 SE 模块。该模块通过学习特征通道之间的相关性来增强对重要特征的关注,并减少对不相关特征的依赖。Natarajan 等人[9]提出了一种用于多导联心电图分类的无卷积的纯 Transformer 方法,由多个自注意力模块和全连接层组成,通过学习波形之间的关联和重要特征,实现 ECG 分类任务。为了能够解决多标签分

类任务带来的类别不均衡挑战, 我们提出了一种基于多种注意力机制的一维卷积神经网络仿真模型, 本文的主要贡献如下:

1) 对于多标签任务的长尾分布现象, 对原始多标签样本适当地进行 Synthetic minority over-sampling technique (SMOTE) [10]重采样技术以平衡数据集。

2) 采用通道 - 空间交叉注意力机制(CSCA), 结合 Transformer 中的全局稀疏多头自注意力, 充分利用心电信号不同通道和空间位置之间的相关性的同时融入全局信息, 提高分类仿真模型对与心律失常有关特征的关注程度。

3) 在计算不同类别的加权损失时嵌入标签相关性信息矩阵, 提高模型在误分类情况下的惩罚权重。

2. 数据集与仿真模型

2.1. 数据集

我们选择了两个目前心律失常分类任务中最常用的多标签数据集来验证我们的模型能否在一条心电记录上识别多种心律失常类型。CPSC2018 [11]是 2018 年中国生理信号挑战赛提供的官方数据集。该数据集由 11 家医院收集而来, 官方公开的训练集中包含了 6877 条记录, 其中男性 3699 条, 女性 3178 条。信号的采样频率为 500 Hz, 每条记录包含标准 12 导联。由于各条记录不等长, 最短的有 6 s, 最长的达到了 144 s。通过简单的数据处理, 将不足 10 秒的记录复制该记录扩充到 10 秒, 超出 10 秒的记录截取前 10 秒, 以此得到相同维度的实验数据。PTB-XL [12]心电图数据集是来自 18,885 名患者的 21,837 个临床 12 导联心电图的大型数据集, 其中 52%为男性, 48%为女性。采样频率为 500 Hz, 长度为 10 秒, 每条记录包含标准 12 导联。两个数据集均按照 8:1:1 的比例随机划分为训练集、验证集和测试集。

由于多标签数据集中普遍存在类别样本数不均衡现象, 常见的几个多数类别高频出现, 而更多的低频类别罕见于样本中。这将导致模型倾向于将低频类别误分类为高频类别, 进而影响模型性能。为了缓解数据集长尾分布带来的消极影响, 我们在上述两个 ECG 多标签数据集上采用 SMOTE 重采样技术, 对含有低频类别的样本进行过采样, 对含有高频类别的样本进行欠采样。

图 1 展示了以 CPSC2018 数据集为例, 重采样前后第一个标签对应的 9 种不同心律失常类型的样本数。临床上, 一条 ECG 数据可能同时包含低频和高频类别的心律失常, 因此经过重采样之后的数据集类别分布只能达到一个相对平衡的状态。

2.2. 损失函数

上述的两个公开数据集经过简单的预处理用于作为网络的输入, 通过导联采集到的 ECG 信号可视为时间序列, 因此在训练过程中模型的输入可以表示为 $X_i = \{x_1, x_2, \dots, x_i, \dots, x_n\}, 1 \leq i \leq n$, 其中 n 表示训练集包含 n 条 ECG 记录, x_i 表示训练集中的第 i 条 ECG 记录。输入还包含用于表示心律失常信息的标签矩阵 $Y_i = \{y_1, y_2, \dots, y_i, \dots, y_n\}, 1 \leq i \leq n$, y_i 表示训练集中的第 i 条 ECG 记录对应的标签信息, 其维度为 $q \times 1$, 向量中每个数值均为 1 或 0, 这将引导模型学习不同心律失常类型的特征。在模型训练的过程中, 通过最小化二分类交叉熵损失函数的值来矫正模型, 使预测结果逐步趋向于标签。通过调整损失函数中不同类别的权重来平衡样本不同类别的贡献, 调整权重之后的损失函数如公式(1)所示:

$$L_{\text{weighted-BCE}} = \frac{1}{n} \sum_{i=1}^n \sum_{c=1}^q \omega_c \left[-y_c \log \hat{y}_c - (1 - y_c) \log (1 - \hat{y}_c) \right] \quad (1)$$

其中, n 代表 ECG 记录数量, q 为类别数量, y_c 为类别 c 的标签, \hat{y}_c 为类别 c 的模型预测概率。 w_c 为类别 c 的权重, 计算方法如公式(2)所示:

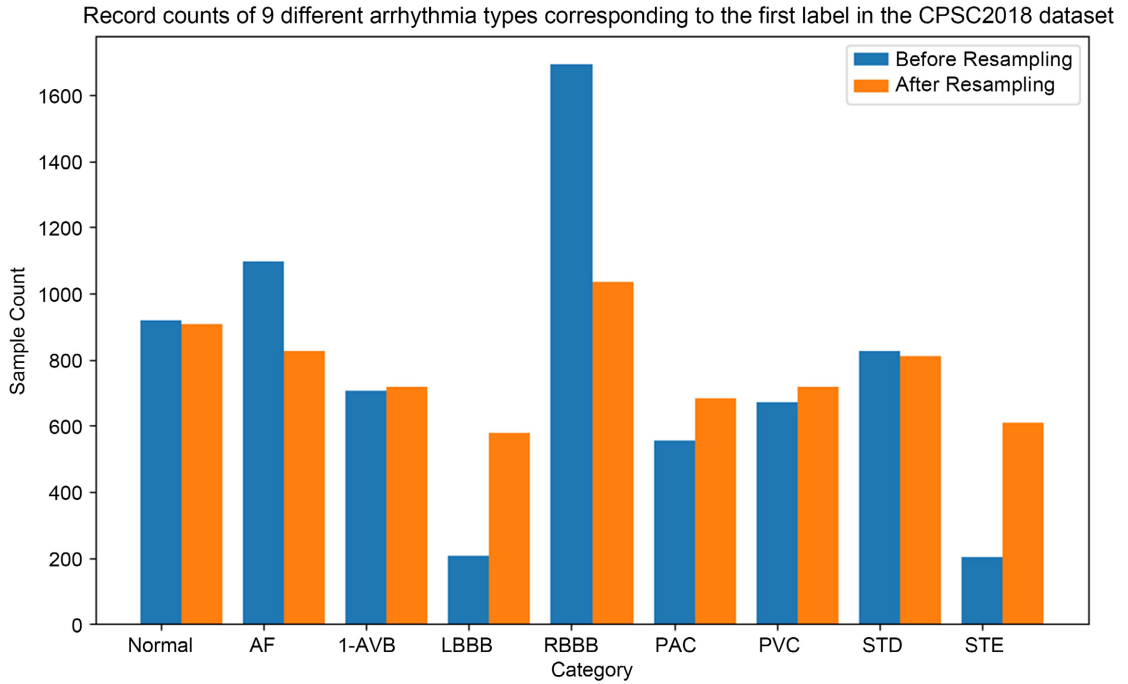


Figure 1. Class distribution before and after resampling

图 1. 重采样前后的数据集类别分布

$$\omega_c = \frac{1}{\log n_c} \quad (2)$$

在此基础上, 考虑到临床上对于不同类别的心律失常错误分类的代价不同。由于人为给定误分类代价矩阵需要较高的临床经验, 本文设计了一种误分类代价计算方法。在计算模型对每个样本全部类别的分类损失基础上, 增加了一项误分类惩罚系数 λ , 使模型提高了误分类代价较高时的分类性能。该计算方法先将包含每个类别 one-hot 形式的标签矩阵 A 乘以可训练的高维嵌入矩阵 B , 得到包含每个类别高维向量矩阵标签 C , 用余弦相似度来衡量这些向量之间的相似性。计算方法如公式(3)所示:

$$\text{Cos_similarity}_{ij} = \frac{C_i \cdot C_j}{\|C_i\| \cdot \|C_j\|} \quad (3)$$

其中, C_i 和 C_j 分别代表矩阵 C 的第 i 行和第 j 行向量, \cdot 表示向量点积, $\| \cdot \|$ 表示向量的范数。

定义误分类惩罚系数的计算方式如公式(4)所示:

$$\lambda = \sum_{n=1}^N \alpha \cdot (1 - \text{Cos_similarity}_{ij}) \quad (4)$$

其中, N 表示预测结果与真实标签不一致的标签对数量, α 为可训练的权重矩阵。

总损失函数的计算方式如公式(5)所示:

$$\text{Loss} = L_{\text{weighted-BCE}} + \lambda \quad (5)$$

2.3. 仿真网络模型

2.3.1. 网络整体结构

本文提出了一种涵盖多种注意力机制的一维卷积神经网络, 如图 2 所示, 网络主体包含四个阶段 (S0~S3)。

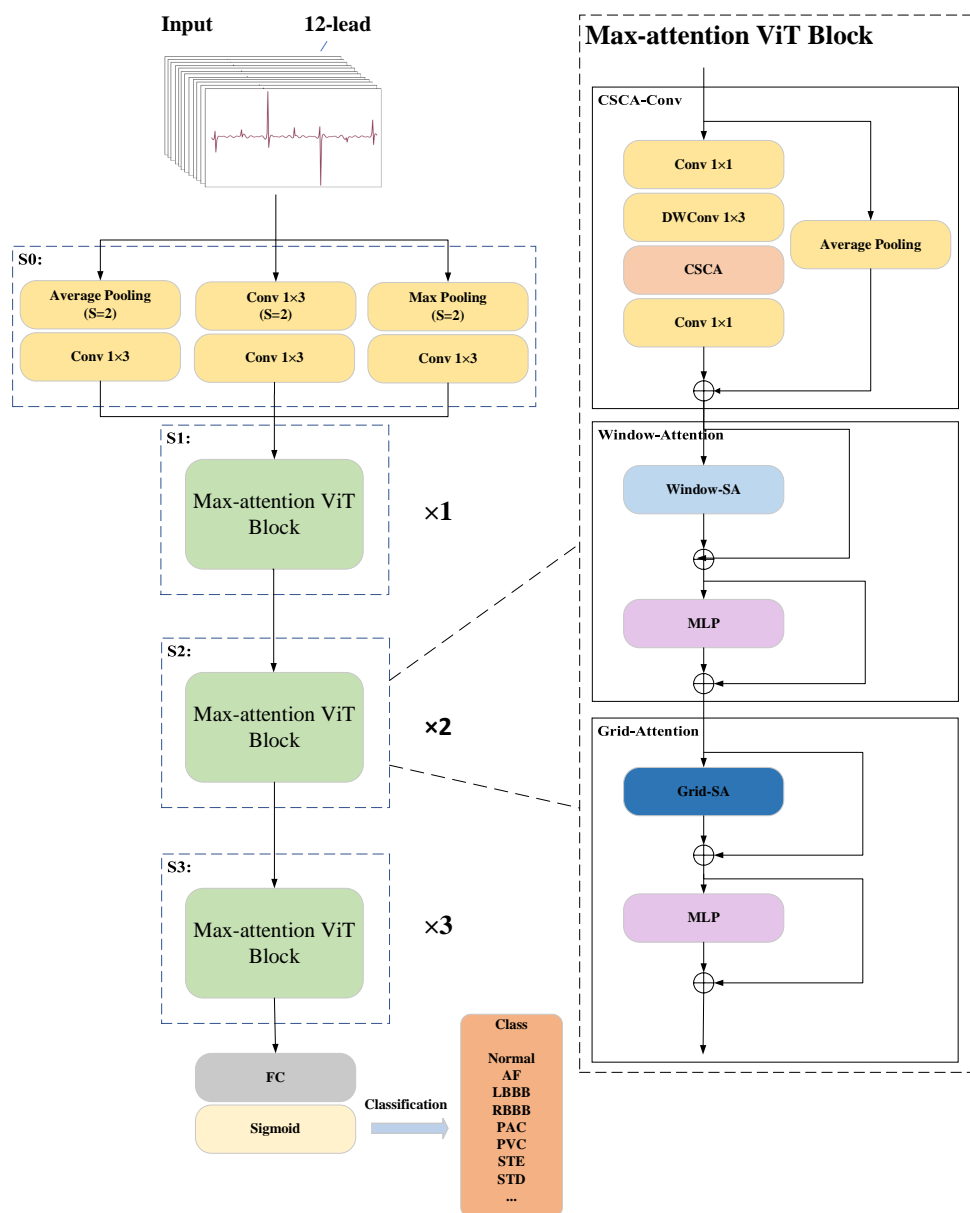


Figure 2. Overall structure of the network

图 2. 网络整体结构图

12 导联的 ECG 信号首先在 S0 阶段的三个分支分别使用 1×3 的一维卷积、平均池化和最大池化进行下采样为原始特征的一半，并使用 1×3 的一维卷积在仅改变通道数的基础上进一步提取特征；S1~S3 阶段使用相同结构的 Max-attention ViT 模块，堆叠的次数分别为 1、1、2，包含 CSCA-Conv、Window-Attention 和 Grid-Attention 三个模块。假设 x 是 CSCA-Conv 模块的输入特征，为简单起见，省略了归一化层和激活层，计算公式如公式(6)所示：

$$x \leftarrow \text{Conv}(\text{AvgPool}(x)) + \text{Conv}\left(\text{CSCA}\left(\text{DWConv}\left(\text{Conv}(x)\right)\right)\right) \quad (6)$$

其中， Avgpool 和 DWConv 表示在两条支路上分别使用一维平均池化和深度可分离卷积下采样一半，并使用 1×1 的一维卷积进行通道数翻倍， CSCA 表示通道 - 空间交叉注意力机制。

S3 阶段输出的特征展平后输入全连接层, 最终经过 Sigmoid 函数输出每个类别的分类结果。

2.3.2. 注意力模块

受到 Tu 等人[13]的启发, 本文提出的 Max-attention ViT 模块涵盖三种注意力机制。在 CSCA-Conv 模块中, 采用通道 - 空间交叉注意力机制(CSCA), 如图 3 所示, 空间注意力[14]和通道注意力[15]作为一种局部的应用, 可以分别在空间维度和通道维度上学习输入图像中不同位置以及不同通道之间的重要性。

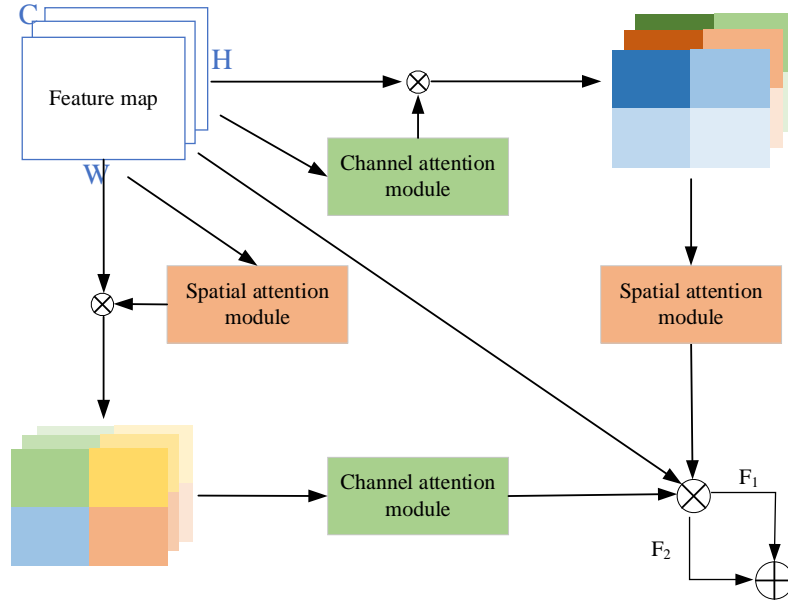


Figure 3. Overall structure of the network
图 3. 网络整体结构图

其中, \otimes 表示向量按元素乘积, \oplus 表示向量按元素相加, 将输入 CSCA 模块的特征图设置为 $F \in R^{12 \times W \times C}$, 经过上面的分支之后变为 $F_1 \in R^{12 \times W \times C}$, 如公式(7)所示:

$$F_1 = F \otimes SpatialAttention(F \otimes ChannelAttention(F)) \quad (7)$$

相应的, 输入的特征图经过下面的分支之后变为 $F_2 \in R^{12 \times W \times C}$, 如公式(8)所示:

$$F_2 = F \otimes ChannelAttention(F \otimes SpatialAttention(F)) \quad (8)$$

其中, 通道注意力的通道放缩比例为 12:1。输入特征在经过两个分支后得到的 F_1 和 F_2 在对应位置上进行元素相加得到进行通道注意力和空间注意力计算的高级特征, 这将有利于模型学习 ECG 记录中的不同心律失常类别的有效信息。

借鉴 Transformer 结构的多头自注意力机制, 在 Window-Attention 模块中, 使用含有局部多头自注意力的 Window-SA 模块来降低计算复杂度的同时允许模型捕捉不同位置之间的长距离依赖关系。如图 4 所示, 窗口自注意力将 CSCA 模块输出的特征图在空间维度上划分为多个不重叠的窗口, 计算每个窗口中特征向量之间的局部相关性, 提取更深层的局部特征信息, 然后将特征图恢复为原始维度。

将输入 Window-Attention 模块的特征图设置为 $x \in R^{12 \times W \times C}$, 计算公式如公式(9)所示:

$$\begin{aligned} x &\leftarrow x + Unwindow(Attention(Window(x))) \\ x &\leftarrow x + MLP(x) \end{aligned} \quad (9)$$

其中, MLP 为包含线性层、激活层和 $Dropout$ 层的多层感知器, $Attention$ 的多头自注意力计算如公式(10)所示:

$$Attention(Q, K, V) = \text{soft max}\left(Q \cdot K^T / \sqrt{d} + B\right) \cdot V \quad (10)$$

其中, Q, K, V 分别为查询矩阵、键矩阵和值矩阵, d 为输入向量的维度, B 为相对位置偏置矩阵。

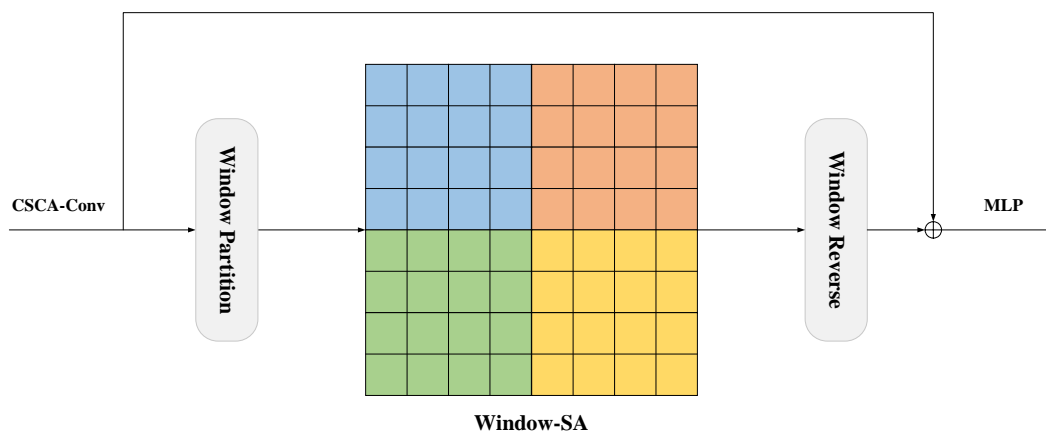


Figure 4. Window-Attention module structure diagram

图 4. Window-Attention 模块结构图

不同于 Window-Attention 的局部多头自注意力, 在 Grid-Attention 模块中, 采用全局稀疏自注意力的策略, 使用 Grid-SA 模块在降低计算复杂度的同时允许模型捕捉全局不同位置之间的依赖关系。如图 5 所示, 网格注意力在划分为多个不重叠的窗口的基础上, 将不同窗口的特征向量合并到同一个网格中计算特征向量之间的全局相关性, 提取更深层的全局特征信息, 然后将特征图恢复为原始维度。

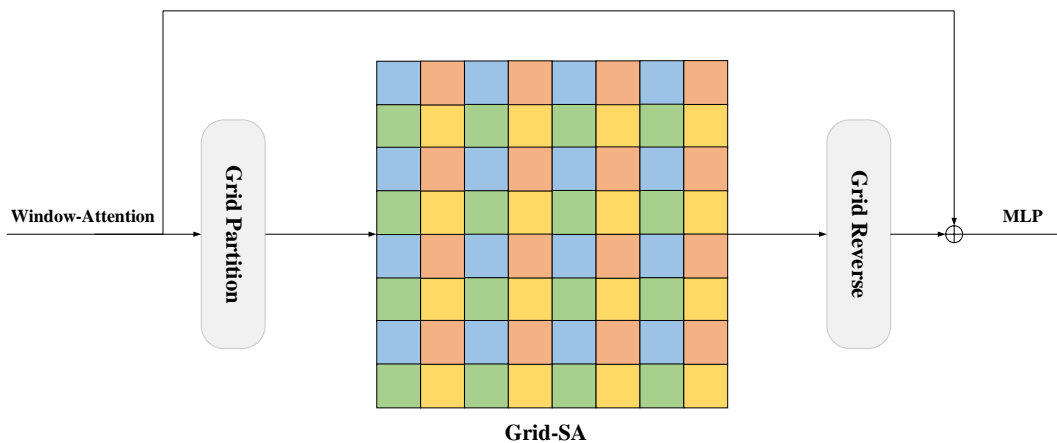


Figure 5. Grid-Attention module structure diagram

图 5. Grid-Attention 模块结构图

将输入 Grid-Attention 模块的特征图设置为 $x \in R^{12 \times W \times C}$, 计算公式如公式(11)所示:

$$\begin{aligned} x &\leftarrow x + \text{Ungrid}\left(\text{Attention}\left(\text{Grid}(x)\right)\right) \\ x &\leftarrow x + \text{MLP}(x) \end{aligned} \quad (11)$$

3. 实验结果与分析

3.1. 实验设置

本文实验是在一台配备 6G 内存的 NVIDIA GeForce 3060LaptopGPU 显卡上进行的, 该模型使用 Pytorch 框架搭建。在训练过程中, 采用随机梯度下降(SGD)算法来优化模型, 初始动量为 0.9, epoch 设置为 150, batch size 设置为 8, 随着训练的轮数增加, 学习率从初始的 0.1 逐渐降低到 0.001, 将网络中全连接层之前的 dropout 层的数值设置为 0.5。通过上述超参数的组合, 模型达到了最佳分类性能。

3.2. 评估指标

本文实验使用 macro-F1、准确率(Accuracy)、精确率(Precision)、召回率(Recall)和汉明损失(Hamming Loss)来作为评估仿真模型在多标签 ECG 分类任务中的评估指标, 计算公式如下所示:

$$\text{Hamming loss} = \frac{1}{(N \times L)} \sum_{i=1}^N \sum_{j=1}^L (y_{ij} \neq \hat{y}_{ij}) \quad (12)$$

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

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

$$\text{Recall} = \frac{TP}{TP + FN} \quad (15)$$

$$\text{Macro-F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (16)$$

其中, N 和 L 分别代表样本总数和每个样本的类别数, y_{ij} 和 \hat{y}_{ij} 分别表示第 i 个样本中第 j 个类别的真实值和预测值, TP 代表真阳性, FP 代表假阳性, TN 代表真阴性, FN 代表假阴性。除此以外, 与参考文献[16]一致, AUC(ROC 曲线下的面积)也被用于评估仿真模型在 PTB-XL 数据集上的性能。

3.3. CPSC2018 数据集实验结果

本文在公开的 CPSC2018 多标签数据集上评估我们的模型分类性能, 使用训练好的模型在测试集上进行分类并计算各项评估指标。使用上述 5 个评估指标对比了近几年在该数据集上多标签分类性能排名靠前的一些研究方法, 具体内容如表 1 所示。其中部分研究中仅有以 macro-F1 为主的个别指标, macro-F1 指标是衡量 ECG 多标签分类的主要指标, 因此本文主要以此为标准进行对比。参考文献[7]汇总了包括其提出的 LCEGNet 在内的 7 种网络结构的性能指标, 相比之下, 本文提出的涵盖多种注意力机制的一维卷积神经网络模型在 macro-F1 指标上提升了 6.46%~15.63%。Ge *et al.* [8]的网络结构中同样使用注意力模块, 但所使用的数据集在预处理过程中, 不同于大部分研究使用的是 10 秒的心电记录, 其使用 60 秒的心电序列作为模型的输入, 这意味着模型能够学习的特征更多, 为此我们使用其方法在 10 秒输入数据集上进行实验, macro-F1 指标最终为 0.813, 下降了 1.5%。Chen *et al.* [17]作为 CPSC2018 比赛中的第一名, 分别针对九类心律失常类型训练出最优模型, 通过结合多个基本模型的预测结果, 以实现更好的性能表现。相比之下, 我们的模型在简单高效的基础上 macro-F1 提升了 0.43%, 其余指标也比表格中的研究方法在不同程度上有所提高, 证明了我们的方法更适合用于 ECG 的多标签分类任务。

本文提出的仿真模型在 CPSC2018 测试集上计算并绘制了多标签混淆矩阵, 如图 6 所示, 展示了模型在不同类别上的正确分类和误分类情况, 颜色越深表示程度越高。

Table 1. The comparison of different models on the CPSC2018
表 1. CPSC2018 数据集上不同模型的性能对比

Model	Hamming loss	Accuracy	Precision	Recall	Macro-F1
ResNet [7]	0.0555	0.6807	0.7393	0.6610	0.6850
xECGNet [7]	0.0440	0.7304	0.8103	0.7313	0.7605
ML-GCN [7]	0.0428	0.7177	0.8236	0.7085	0.7535
MLC-CNN [7]	0.0472	0.6946	0.8034	0.6873	0.7337
P-GCN [7]	0.0455	0.7016	0.8162	0.6970	0.7443
KGGR [7]	0.0451	0.7132	0.8119	0.7062	0.7477
LCEGNet [7]	0.0407	0.7500	0.8326	0.7375	0.7767
Ge <i>et al.</i> [8]	/	/	0.830	0.827	0.828
Chen <i>et al.</i> [17]	/	/	/	/	0.837
Ours	0.0327	0.8139	0.8327	0.8508	0.8413

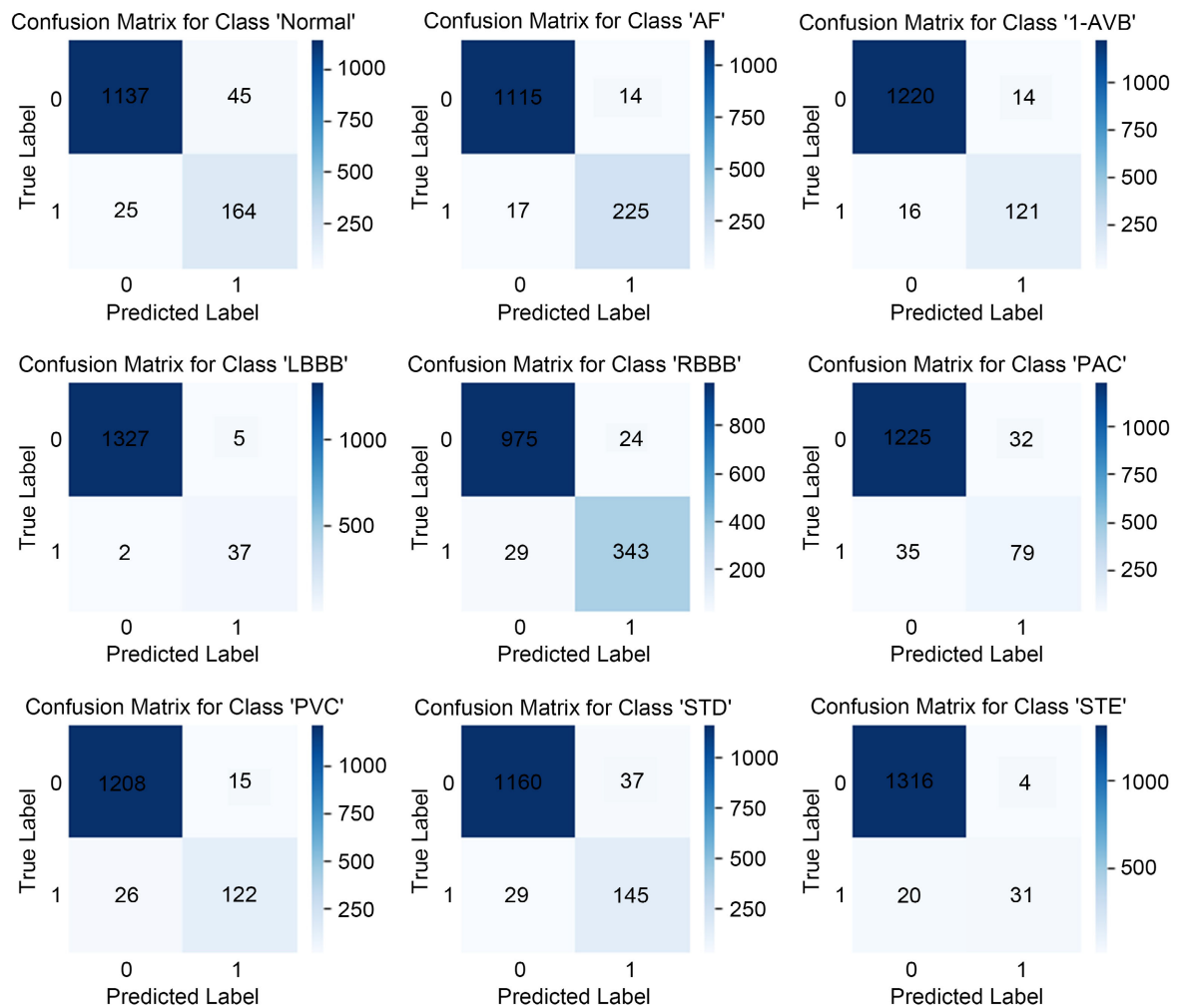


Figure 6. The confusion matrix on the test set of the CPSC2018
图 6. CPSC2018 测试集上的混淆矩阵

3.4. PTB-XL 数据集实验结果

为了验证模型的泛化能力, 我们在第二个公开的数据集 PTB-XL 上评估模型的分性能。数据集具有 5 个超类、23 个子类和 44 个诊断类, 由于在 23 个子类和 44 个诊断类分类标准下, 少部分类别的样本数极少, 现有的不均衡处理技术效果不显著。为了减轻类别不均衡的影响, 本次实验仅在 5 个超类上进行模型的训练和测试性能, 并将其与现有研究进行对比, 如表 2 所示。除 Precision 以外, 在其余几个主要指标上均高于现有研究方法, 充分说明本文提出的方法在多个数据集上具有良好的泛化性和分类性能。

Table 2. The comparison of different models on the PTB-XL

表 2. PTB-XL 数据集上不同模型的性能对比

Model	Accuracy	Precision	Recall	Macro-F1	AUC
Smigiel <i>et al.</i> [18]	/	0.714	0.662	0.680	0.910
DNNs [19]	/	/	/	0.683	0.907
FSL [20]	/	/	/	0.717	0.938
MRF-CNN [16]	/	0.730	0.710	0.720	0.930
Ours	0.644	0.703	0.787	0.736	0.941

3.5. 消融实验

为了验证仿真模型中 S1~S3 阶段使用的多种注意力机制的有效性, 以 CPSC2018 数据集为例, 进行了一系列消融实验, 具体内容如表 3 所示。首先, 将 Max-attention ViT 模块中三种注意力机制均替换为提出的 CSCA, 将其与现有研究中的几种注意力机制在同一模型结构下进行实验对比; 其次, 将 Max-attention ViT 模块中最后两个注意力机制分别全部替换为 Window-Attention 和 Grid-Attention; 最后, 将 Max-attention ViT 模块中最后两个注意力机制分别依次替换为 Window-Attention 和 Grid-Attention, 即图 2 所示的模型结构, 达到了最佳分类性能。

Table 3. Ablation experiments on different attention mechanisms

表 3. 不同注意力机制的消融实验

Model	Hamming loss	Accuracy	Precision	Recall	Macro-F1
ECA [21]	0.0732	0.7833	0.7928	0.8257	0.8092
CBAM [22]	0.0686	0.7882	0.7975	0.8276	0.8143
SE [15]	0.0657	0.8003	0.8079	0.8301	0.8166
SCSE [23]	0.0573	0.7891	0.8122	0.8334	0.8193
CSCA (Ours)	0.0524	0.8045	0.8187	0.8367	0.8221
CSCA + WA	0.0466	0.8072	0.8256	0.8484	0.8368
CSCA + GA	0.0437	0.8083	0.8263	0.8371	0.8273
CSCA + WA + GA	0.0392	0.8116	0.8358	0.8469	0.8413

4. 结束语

为了能够在多标签 ECG 自动分类任务中精准分类出多种心律失常, 本文提出了一种涵盖多种注意力

机制的一维卷积神经网络仿真模型。使用到的数据集包括 CPSC2018 和 PTB-XL 两个公开的多标签数据集, 针对多标签任务普遍存在的数据集长尾分布现象, 在数据处理层面, 对原始数据集的样本适当地进行 SMOTE 重采样技术以平衡数据集; 在算法优化层面, 在计算不同类别的加权损失时嵌入标签相关性信息矩阵, 提高模型在误分类情况下的惩罚权重, 引导模型正确分类。除此以外, 在网络结构中添加了三种不同的注意力机制, 使得模型同时加强对局部信息和全局信息的感知学习能力, 能够学习到与不同类别相关的高级特征, 达到精准分类的目的。未来的研究工作将考虑调整网络结构和参数设置, 有望进一步提升心律失常识别领域的性能, 并为实践提供更多有针对性的方法和指导。

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] Zhang, L., Peng, H. and Yu, C. (2010) An Approach for ECG Classification Based on Wavelet Feature Extraction and Decision Tree. 2010 *International Conference on Wireless Communications & Signal Processing (WCSP2010)*, Suzhou, 21-23 October 2010, 1-4. <https://doi.org/10.1109/WCSP.2010.5633782>
- [2] Kropf, M., Hayn, D. and Schreier, G. (2017) ECG Classification Based on Time and Frequency Domain Features Using Random Forests. 2017 *Computing in Cardiology (CinC)*, Rennes, 24-27 September 2017, 1-4. <https://doi.org/10.22489/CinC.2017.168-168>
- [3] Mehta, S.S. and Lingayat, N.S. (2008) Development of SVM Based Classification Techniques for the Delineation of Wave Components in 12-Lead Electrocardiogram. *Biomedical Signal Processing and Control*, **3**, 341-349. <https://doi.org/10.1016/j.bspc.2008.04.002>
- [4] Hannun, A.Y., Rajpurkar, P., Haghpanahi, M., Tison, G.H., Bourn, C., Turakhia, M.P. and Ng, A.Y. (2019) Cardiologist-Level Arrhythmia Detection and Classification in Ambulatory Electrocardiograms Using a Deep Neural Network. *Nature Medicine*, **25**, 65-69. <https://doi.org/10.1038/s41591-018-0268-3>
- [5] Huang, J.S., Chen, B.Q., Yao, B. and He, W.P. (2019) ECG Arrhythmia Classification Using STFT-Based Spectrogram and Convolutional Neural Network. *Biomedical Signal Processing and Control*, **7**, 92871-92880. <https://doi.org/10.1109/ACCESS.2019.2928017>
- [6] Gao, J.L., Zhang, H.P., Lu, P. and Wang, Z.M. (2019) An Effective LSTM Recurrent Network to Detect Arrhythmia on Imbalanced ECG Dataset. *Journal of Healthcare Engineering*, **2019**, Article ID: 6320651. <https://doi.org/10.1155/2019/6320651>
- [7] Ran, S.L., Li, X., Zhao, B.Z., Jiang, Y.N., Yang, X.Y. and Cheng, C. (2023) Label Correlation Embedding Guided Network for Multi-Label ECG Arrhythmia Diagnosis. *Knowledge-Based Systems*, **270**, Article ID: 110545. <https://doi.org/10.1016/j.knosys.2023.110545>
- [8] Ge, R.J., Shen, T.F., Zhou, Y., Liu, C.Y., Zhang, L.B., Yang, B.Q., *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>
- [9] Natarajan, A., Boverman, G., Chang, Y.L., Antonescu, C. and Rubin, J. (2021) Convolution-Free Waveform Transformers for Multi-Lead ECG Classification. *Conference on Computing in Cardiology (CinC)*, Brno, 13-15 September 2021, 1-4. <https://doi.org/10.23919/CinC53138.2021.9662697>
- [10] Chawla, N.V., Bowyer, K.W., Hall, L.O. and Kegelmeyer, W.P. (2002) SMOTE: Synthetic Minority Over-Sampling technique. *Journal of Artificial Intelligence Research*, **16**, 321-357. <https://doi.org/10.1613/jair.953>
- [11] Liu, F.F., Liu, C.Y., Zhao, L.N., Zhang, X.Y., Wu, X.L., Xu, X.Y., *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>
- [12] Wagner, P., Strodthoff, N., Bousseljot, R.D., Kreiseler, D., Lunze, F.I., Samek, W. and Schaeffter, T. (2020) PTB-XL, a Large Publicly Available Electrocardiography Dataset. *Scientific Data*, **7**, Article No. 154. <https://doi.org/10.1038/s41597-020-0495-6>
- [13] Tu, Z.Z., Talebi, H., Zhang, H., Yang, F., Milanfar, P., Bovik, A. and Li, Y.X. (2022) MaxViT: Multi-Axis Vision Transformer. *Computer Vision, ECCV 2022*, Vol. 13684, 459-479. https://doi.org/10.1007/978-3-031-20053-3_27

-
- [14] Jaderberg, M., Simonyan, K., Zisserman, A. and Kavukcuoglu, K. (2015) Spatial Transformer Networks. *Proceedings of the 28th Neural Information Processing Systems Conference (NIPS)*, Montreal, 8-13 December 2014, 2017-2025.
- [15] Hu, J., Shen, L., Albanie, S., Sun, G. and Wu, E.H. (2020) Squeeze-and-Excitation Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **42**, 2011 -2023. <https://doi.org/10.1109/TPAMI.2019.2913372>
- [16] Feyisa, D.W., Debelee, T.G., Ayano, Y.M., Kebede, S.R. and Assore, T.F. (2022) Lightweight Multireceptive Field CNN for 12-Lead ECG Signal Classification. *Computational Intelligence and Neuroscience*, **2022**, Article ID: 8413294. <https://doi.org/10.1155/2022/8413294>
- [17] Chen, T.M., Huang, C.H., Shih, E.S.C., Hu, Y.F. and Hwang, M.J. (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>
- [18] Smigiel, S., Palczynski, K. and Ledzinski, 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>
- [19] Smigiel, S., Palczynski, K. and Ledzinski, 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>
- [20] Palczynski, K., Smigiel, S., Ledzinski, 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>
- [21] Wang, Q.L., Wu, B.G., Zhu, P.F., Li, P.H., Zuo, W.M. and Hu, Q.H. (2020) ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 11531-11539. <https://doi.org/10.1109/CVPR42600.2020.01155>
- [22] Woo, S.H., Park, J., Lee, J.Y., Kweon, I.S., Ferrari, V., Hebert, M., *et al.* (2018) CBAM: Convolutional Block Attention Module. *15th European Conference on Computer Vision (ECCV)*, Munich, 8-14 September 2018, 3-19. https://doi.org/10.1007/978-3-030-01234-2_1
- [23] Roy, A.G., Navab, N. and Wachinger, C. (2018) Concurrent Spatial and Channel “Squeeze & Excitation” in Fully Convolutional Networks. *21st International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Vol. 11070, 421-429. https://doi.org/10.1007/978-3-030-00928-1_48