

DOI: 10.13382/j.jemi.B2407542

# 面向 ECG 彩虹码的双输入改进 VIT 识别研究

陈波<sup>1,2</sup> 孙辉<sup>2</sup> 储昭碧<sup>2</sup> 李育玲<sup>2</sup> 魏嘉乐<sup>2</sup>

(1. 功率半导体封装与可靠性安徽省重点实验室(合肥工业大学) 合肥 230009;

2. 合肥工业大学电气与自动化工程学院 合肥 230009)

**摘要:** 基于海量 ECG 数据, 辅助医生进行有效数据分析与诊断, 提高效率并减少医疗资源消耗, 实现 ECG 智能识别是当前一个重要研究方向。针对 ECG 智能识别单一图像、单一深度学习算法性能有限性问题, 提出了一种面向 ECG 彩虹码的双输入改进 VIT 识别方法。首先, 提出数学模型预测获取 ECG 标准周期, 并以抽频方法挖掘 ECG 潜在特征, 生成 ECG 彩虹码; 然后, 以卷积神经网络构建双输入特征提取模块, 提取多种 ECG 图像局部特征进行融合, 实现多维度 ECG 特征表示与融合, 采用 VIT 编码模块对融合特征进行全局关注, 实现基于多特征图像为输入的 ECG 识别。采用 MIT-BIH 数据库中的 ECG 进行实验, 所提 ECG 识别方法获得 99.41% 的平均准确率, 在现场采集的 N 类 ECG 中获得 100% 的准确率。实验结果表明, 提出的图像变换方法能够有效可视化 ECG 特征, 提出的识别方法能够有效实现 ECG 识别, 与其他同类型方法相比获得了更优的性能。

**关键词:** 心电信号; ECG 彩虹码; 图像变换; 双输入特征提取模块; 改进 VIT

**中图分类号:** TP274; TN911.7 **文献标识码:** A **国家标准学科分类代码:** 510.4099

## Research on two-input improved VIT recognition for ECG rainbow codes

Chen Bo<sup>1,2</sup> Sun Hui<sup>2</sup> Chu Zhaobi<sup>2</sup> Li Yuling<sup>2</sup> Wei Jiale<sup>2</sup>

(1. Anhui Province Key Laboratory of Semiconductor Packaging and Reliability (Hefei University of Technology), Hefei 230009, China; 2. School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China)

**Abstract:** Leveraging extensive ECG data, intelligent ECG recognition represents a pivotal research focus aimed at supporting physicians in conducting thorough data analysis and diagnosis, thereby enhancing efficiency and mitigating medical resource consumption. In order to solve the problem of feature loss and limited performance of single image and single deep learning algorithm in ECG intelligent recognition, a two-input improved VIT recognition method for ECG rainbow code is proposed. Firstly, a mathematical model is proposed to predict the standard period of ECG, and the potential features of ECG are mined by pumping method to generate ECG rainbow code. Then, a dual input feature extraction module is constructed with convolutional neural network to extract local features of multiple ECG images for fusion to achieve multi-dimensional ECG feature representation and fusion. A VIT coding module is used to pay global attention to fusion features to realize ECG recognition based on multi-feature images as input. The ECG recognition method in MIT-BIH database is used for experiments, and the average accuracy of the proposed ECG recognition method is 99.41%, and the accuracy of the N-type ECG collected in the field is 100%. The experimental results show that the proposed image transformation method can effectively visualize ECG features, and the effect is better than the traditional method. The proposed recognition method can realize ECG recognition effectively and has better performance than other similar methods.

**Keywords:** electrocardiography; ECG rainbow code; image transformation; dual input feature extraction module; improved VIT

## 0 引言

依据世界卫生组织发布的相关报告, 因心血管疾病

(cardiovascular disease, CVD) 造成的死亡人数在全球范围内占据首位<sup>[1-2]</sup>。心电图信号 (electrocardiography, ECG) 是反映心血管生命活动信息的最基本生理信号, 是医生对 CVD 诊断的基础性依据之一。基于海量 ECG 数

据研究 ECG 智能识别有效方法并辅助医生进行诊断,有利于提高诊断效率并缓解医疗资源压力<sup>[3-4]</sup>。

深度学习技术近些年飞速发展,学者们设计了诸多深度学习模型算法<sup>[5-7]</sup>。在医学领域将已知优秀算法进行应用逐渐变成学者们研究的热点,为 ECG 识别提供了新的思路。采用深度学习算法进行 ECG 识别,无需进行复杂的人工特征提取,其自身能够较为准确的提取输入数据特征<sup>[8-11]</sup>。通过深度学习算法实现 ECG 识别,可分为一维时间序列输入与二维图像输入的两种方法。Ritu 等<sup>[12]</sup>提出一种基于 SVM 和 Bi-LSTM 的心电图分类方法,利用一维时间序列数据在机器学习和深度学习联合方法实现 ECG 识别。Serkan 等<sup>[13]</sup>提出了一种自适应的一维卷积神经网络(convolutional neural networks, CNN),使用普通和患者特异性相结合的数据,使得模型更有针对性地进行学习训练,实现 ECG 识别。Mathunjwa 等<sup>[14]</sup>提出了一种结合递归图和两阶段深度学习的 ECG 分类方法,将 ECG 转换为递归图提供给 CNN 模型进行识别,最终在 ECG 识别中获得 98.41% 的准确率。Ullah 等<sup>[15]</sup>提出使用短时傅里叶变换将一维时间序列转换为二维光谱图,通过 CNN 模型对 8 种 ECG 进行分类,获得了较高的准确率。将一维时间序列变换为二维图像,能够更好反应原始数据的特征与模式,以图像为类深度学习算法输入数据进行 ECG 识别,其性能优于采用一维时间序列为输入数据的方法<sup>[16]</sup>。但是传统图像变换方法仍然存在特征损失问题,且采用单一图像与单一深度学习算法在 ECG 识别任务上存在性能有限性问题<sup>[17]</sup>。

针对上述问题,研究并提出一种 ECG 彩虹码(EGC rainbow code, ERC)的图像变换方法,以数学模型预测获得 ECG 标准周期,通过抽频方法挖掘 ECG 潜在特征,生成 ERC。根据 ECG 特性结合 CNN 局部特征提取的优势与视觉 transformer(vision transformers, VIT)全局特征关注的的能力,设计一种基于双输入改进 VIT 的混合模型,联合 ERC 图与短时傅里叶变换(short time fourier transform, STFT)图,即传统时频图,共同实现 ECG 识别,提升最终识别性能。

## 1 ECG 彩虹码图像变换方法

### 1.1 数学模型预测的 ECG 截取

假设心电采集设备测量的 ECG 可表示为  $x(t)$ , 形式表示为序列  $x(n)$ , 采用窗函数进行截取可表示为:

$$wfm(n) = \begin{cases} 1, & l_a \leq n \leq l_b \\ 0, & \text{其他} \end{cases} \quad (1)$$

其中,  $n = 0, 1, \dots, l-1$  表示采样量,  $n = t \times f$ ,  $f$  为采样频率,  $l_a$  与  $l_b$  为窗函数起点和终点。

截取后 ECG 可表示为  $x_a(n) = x(n) \times wfm(n)$ ,  $l_a$  与

$l_b$  可表示为:

$$l_a = \text{floor}\left(j_a + \frac{3(j_b - j_a)}{5}\right) \quad (2)$$

$$l_b = \text{floor}\left(j_c - \frac{2(j_c - j_b)}{5}\right) \quad (3)$$

其中,  $j_b$  为  $x(n)$  中第  $i$  个 R 峰对应序列号, 同理,  $j_a$ 、 $j_c$  为第  $i-1$ 、 $i+1$  个 R 峰对应序列号。

为保证截取后 ECG 长度一致性, 设置长度为 270 个数据点, 包括 R 峰后 160 个数据点, R 峰与 R 峰前 110 个数据点。当  $\text{floor}(2(j_b - j_a)/5) + 1 < 110$  时, 采用数学模型进行数据向前预测, 数学模型可表示为:

$$f_z(n) = a_0 + \sum_{i=1}^{l_z} a_i e^{-\left(\frac{(n-b_i)^2}{2\sigma_i^2}\right)} + r_a(n) \quad (4)$$

其中,  $r_a(n)$  表示数学模型修正分量。

通过  $f_z(n)$  对  $x_a(n)$  进行曲线拟合, 以斑马算法<sup>[18]</sup>进行参数辨识, 得到函数  $f_a(n)$ , 即  $f_a(n) \approx x_a(n)$ 。令  $m = n - 109 + \text{floor}(2(j_b - j_a)/5)$ , 则向前预测数据点  $f_a(m)$  可表示为:

$$f_a(m) = a_0 + \sum_{i=1}^{l_z} a_i e^{-\left(\frac{\left(m + 109 - \text{floor}\left(\frac{2(j_b - j_a)}{5}\right) - b_i\right)^2}{2\sigma_i^2}\right)} + r_a\left(m + 109 - \text{floor}\left(\frac{2(j_b - j_a)}{5}\right)\right) \quad (5)$$

其中,  $m = 0, 1, \dots, (108 - \text{floor}(2(j_b - j_a)/5))$ 。

当  $\text{floor}(3(j_c - j_b)/5) < 160$  时, 进行向后预测, 则向后预测数据点  $f_a(h)$  可表示为:

$$f_a(h) = a_0 + \sum_{i=1}^{l_z} a_i e^{-\left(\frac{\left(h + 160 - \text{floor}\left(\frac{3(j_c - j_b)}{5}\right) - b_i\right)^2}{2\sigma_i^2}\right)} + r_a\left(h + 160 - \text{floor}\left(\frac{3(j_c - j_b)}{5}\right)\right) \quad (6)$$

其中,  $h = 0, 1, \dots, (159 - \text{floor}(3(j_c - j_b)/5))$ 。

由式(5)和(6)可以得到预测后 ECG, 可表示为:

$$x_b(l) = \begin{cases} [f_a(m), x_a(n), f_a(h)], & g_a \geq 0, g_b \geq 0 \\ [x_a(g_c - 109), \dots, x_a(n), f_a(h)], & g_a < 0, g_b \geq 0 \\ [f_a(m), x_a(0), \dots, x_a(g_c + 160)], & g_a \geq 0, g_b < 0 \\ [x_a(g_c - 109), \dots, x_a(g_c + 160)], & g_a < 0, g_b < 0 \end{cases} \quad (7)$$

其中,  $g_c$  为  $x_a(n)$  中 R 峰对应序列号,  $g_a$ 、 $g_b$  值为  $108 - \text{floor}(2(j_b - j_a)/5)$ 、 $159 - \text{floor}(3(j_c - j_b)/5)$ ,  $x_b(l)$  数据点长度为 270, 其 R 峰对应序列号为 110。

### 1.2 ECG 彩虹码构建

对预测后的 ECG 序列  $x_c(l)$  进行快速傅里叶变换获得频率信号  $x_c(w)$ , 将  $x_c(w)$  等效为  $1 \times 270$  的一维矩阵,

分别采用大小为  $1 \times 2$ 、步长为 2 的卷积核与大小为  $1 \times 3$ 、步长为 3 的卷积核进行  $x_c(w)$  进行抽频处理,可表示为:

$$x_{b1}(w) = \frac{1}{g_d} \sum_{w=B_1} x_c(w) \quad (8)$$

$$x_{b2}(w) = \frac{1}{g_e} \sum_{w=B_2} x_c(w) \quad (9)$$

其中,  $x_{b1}(w)$  表示二阶频率信号、 $x_{b2}(w)$  表示三阶频率信号,大小分别为  $1 \times 135$ 、 $1 \times 90$ ,  $B_1$ 、 $B_2$  表示卷积核大小,  $g_d$ 、 $g_e$  为卷积核乘积值。

将二阶频率信号  $x_{b1}(w)$  与三阶频率信号  $x_{b2}(w)$  进行快速傅里叶反变换,得到二阶时域信号  $x_{c1}(l)$  与三阶时域信号  $x_{c2}(l)$ 。对  $x_c(l)$ 、 $x_c(w)$  进行  $g_h$  倍行扩展,得到时域信号矩阵  $A_{c1}$ 、频率信号矩阵  $A_{e1}$ ;对  $x_{c1}(l)$ 、 $x_{b1}(w)$  进行 2 倍列扩展与  $g_h$  倍行扩展,得到二阶时域信号矩阵  $A_{c2}$ 、二阶频率信号矩阵  $A_{e2}$ ;对  $x_{c2}(l)$ 、 $x_{b2}(w)$  进行 2 倍列扩展与  $g_h$  倍行扩展,得到三阶时域信号矩阵  $A_{c3}$ 、三阶频率信号矩阵  $A_{e3}$ 。由上述 6 个信号矩阵构建 ERC 矩阵,可表示为:

$$A_c = [A_{c1}^T, A_{c2}^T, A_{c3}^T]^T \quad (10)$$

$$A_e = [A_{e1}^T, A_{e2}^T, A_{e3}^T]^T \quad (11)$$

$$A = \frac{[A_c^T, A_e^T]^T - \min[A_c^T, A_e^T]^T}{\max[A_c^T, A_e^T]^T - \min[A_c^T, A_e^T]^T} \quad (12)$$

其中,  $A$  表示 ERC 矩阵,大小为  $270 \times 270$ ,将矩阵  $A$  的直线性映射为伪彩色图可以得到 ERC 图。

## 2 基于双输入改进 VIT 的混合模型

基于双输入改进 VIT 的混合模型以 CNN 与 VIT 实现 ECG 分类,采用 ECG 双通道特征图像输入,模型主要包括双输入特征提取模块、维度转换模块和 VIT 编码模块,双输入改进 VIT 的混合模型结构如图 1 所示。

1) 双输入特征提取模块:该模块主要用于提取 ECG 特征图的局部特征信息,位于双输入改进 VIT 的混合模型最前端,采用双通道对输入的两种不同 ECG 特征图进行局部特征提取并进行特征融合,由 CNN 构成。

2) 维度转换模块:该模块实现对双输入特征提取模块输出数据的维度变换与深度变换,并且将图像数据转换为 VIT 编码模块能够识别的特定数据格式。

3) VIT 编码模块:该模块实现对双输入特征提取模块提取到的局部特征进行全局关系的提取,实现局部特征与全局特征的联合,最终由该模块输出 ECG 识别结果。

双输入特征提取模块由 3 个连续串联的卷积层、归一化层、激活函数层与池化层共同提取 ECG 特征图低层特征,该模块可表示为:

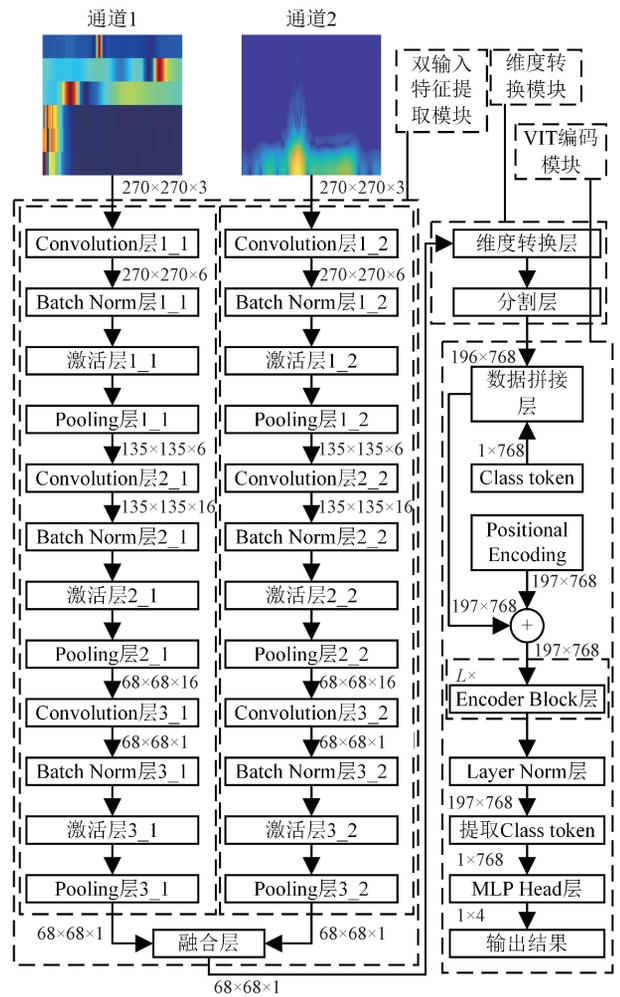


图 1 双输入改进 VIT 的混合模型结构图

Fig. 1 Hybrid model structure diagram of improved VIT with two inputs

$$s_i = \text{relu}(\text{bn}(\text{con}(k_{i-1}))) \quad (13)$$

$$k_i = \text{pooling}(s_i) \quad (14)$$

其中,  $\text{con}$  表示进行卷积运算,  $\text{bn}$  表示进行归一化运算,  $\text{relu}$  表示进行激活运算,  $\text{pooling}$  表示进行池化运算,  $k_0$  表示输入的特征图。

双输入特征提取模块各层配置如表 1 所示,表 1 中可学习属性是指特征提取模块内各网络层对应的可学习参数数量,“-”表示此位置无可学习参数。双输入特征提取模块,通道 1 与 2 配置相同,均采用表 1 所示配置,故共有可学习参数 2 386 个。

图 1 中的维度转换模块,由维度转换层和分割层构成。维度转换层通过双线性插值法将大小为  $68 \times 68 \times 1$  的图像转换为大小为  $224 \times 224 \times 3$  的图像,分割层则是将维度变换后的图像进行矩阵切割,可以避免采用大卷积核造成网络模型对优化器敏感<sup>[19]</sup>,可表示为:

表 1 特征提取模块配置

Table 1 Feature extractor module configuration

层名称	输入尺寸	卷积核大小	步长	深度	可学习属性
卷积层 1	270×270×3	3×3	1	6	权重 = 162、偏置 = 6
池化层 1	270×270×6	2×2	2	-	-
卷积层 2	135×135×16	3×3	1	16	权重 = 864、偏置 = 16
池化层 2	135×135×16	2×2	2	-	-
卷积层 3	68×86×16	3×3	1	1	权重 = 144、偏置 = 1
池化层 3	68×86×16	2×2	1	-	-

$$\mathbf{Z}_{ij} = \mathbf{G}_i \cdot \mathbf{H} \cdot \mathbf{I}_j \quad (15)$$

$$\mathbf{G}_i = [\mathbf{E}_i, \mathbf{E}_0, \dots, \mathbf{E}_0] \quad (16)$$

$$\mathbf{I}_j = [\mathbf{E}_j, \mathbf{E}_0, \dots, \mathbf{E}_0]^T \quad (17)$$

其中,  $\mathbf{H}$  表示维度转换层输出的图像数据,  $\mathbf{Z}_{ij}$  表示由  $\mathbf{H}$  切割后第  $(i-1) \times 14 + j$  个图像块,  $\mathbf{E}_i, \mathbf{E}_j$  为单位矩阵,  $\mathbf{E}_0$  为零元素矩阵。

VIT 编码模块去除了 VIT 中的大型卷积核进行图像切割的部分,保留了 VIT 数据嵌入功能和 VIT 串联编码器功能<sup>[20]</sup> VIT 中原始的切割功能由维度转换模块实现,去除了大型卷积核。

### 3 ECG 识别方法与实验方法设计

#### 3.1 ECG 识别方法

ECG 是 CVD 诊断的重要基准之一,实现 ECG 识别可以辅助医生进行相关疾病的诊断。研究并提出一种面向 ECG 彩虹码的双输入改进 VIT 识别,通过 ERC 图和 STFT 图作为输入数据,联合 CNN 与 VIT 的优势构建了双输入改进 VIT 的混合模型进行 ECG 识别。面向 ECG 彩虹码的双输入改进 VIT 识别方法如图 2 所示,其具体步骤如下:

##### 1) ECG 图像变换

通过第 1 节所提图像变换方法,获得 ERC 图,并以短时傅里叶变换获得 STFT 图。采用 ERC 图与 STFT 图,从不同维度进行特征可视化,作为 ECG 识别的输入数据。其中,STFT 图由短时傅里叶变换获得需要进行去除坐标轴和尺寸变换等操作<sup>[21]</sup>,确保图像大小为  $270 \times 270 \times 3$ 。

##### 2) 双输入改进 VIT 识别方法

采用第 2 节所提双输入改进 VIT 的混合模型进行 ECG 识别,以图 2 中的 ERC 图与 STFT 图为输入数据,通过 CNN 构建的双输入特征提取模块获取输入图像的低层局部特征并进行融合,再由维度转换模块进行融合特征的维度、深度变换处理,并将处理后数据传递给 VIT 编码模块进行全局特征提取,最终实现 ECG 识别任务。

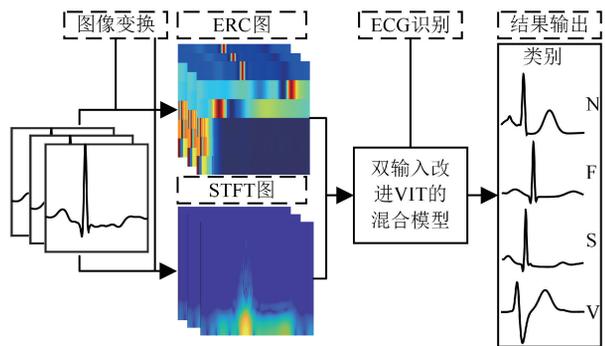


图 2 ECG 识别方法

Fig. 2 ECG recognition method

#### 3.2 实验方法设计

为使后续实验结果更为科学的横向对比,实验数据选用 MIT-BIH 心电数据库。该数据库 48 组 ECG 记录,采样频率为 360 Hz,采样幅值为 10 mV 电压信号,采样时长均达到半小时,每组数据中包含两个导联数据。选用数据量较多的 MLI 导联数据进行实验,其中 48 组 ECG 记录中有 4 组没有 MLI 导联,故最终选取 44 组 ECG 进行试验。

按照 AAMI 分类方法将 MIT-BIH 数据库 ECG 划分为 5 种类别,即正常心跳(N)、室上异位搏动(S)、室性早搏(V)、融合心跳(F)、未知心跳(U)。按照第 1.1 节方法对 ECG 进行截取,获得 N 类 ECG 数量为 90 006, S 类 ECG 数量为 2 648, V 类 ECG 数量为 7 005, F 类 ECG 数量为 779, U 类 ECG 数量为 15。优于 U 类 ECG 数量远远小于 N 类 ECG,故只针对 F、N、S、V 类 ECG 进行识别。

为验证所提 ECG 彩虹码图像变换方法的有效性与优势,以及面向 ECG 彩虹码的双输入改进 VIT 识别方法的可行性与优越性,设计 3 组验证实验,实验方法如下:

##### 1) ECG 彩虹码图像变换方法有效性与优势实验

将 ERC 图与 STFT 图、格拉姆角和场 (gramian angular summation field, GASF) 图、格拉姆角差场 (gramian angular difference field, GADF) 图、马尔可夫迁移场 (markov transition field, MTF) 图进行对比,以上述图像为输入数据,分别输入到 LetNet5 网络模型进行识别,通过该模型快速得到 ECG 识别结果,评价各图像的有效性 with 优势。由于 MIT-BIH 数据库中各类别 ECG 数量不均衡,会导致模型识别结果偏向数量多的 ECG 类别,故对 F、N、S、V 类 ECG 进行数据集均衡处理。为快速得到图像评价结果,采用小数据集进行实验,数据集如表 2 所示。

##### 2) 可行性实验与消融实验

采用 3.1 节 ECG 识别方法,将 ERC 图与 STFT 图输入到双输入改进 VIT 的混合模型进行识别,验证该方法的可行性。同时,进行消融实验,对双输入改进 VIT 的混

表 2 小样本数据集

Table 2 Small sample data set

ECG 类别标签	均衡处理前数量	均衡处理后数量
F	779	779
N	90 006	800
S	2 648	800
V	7 005	800

合模型的结构进行裁剪,将双输入特征提取模块裁剪为单输入特征提取模块,通过 ERC 图、STFT 图分别进行识别;将双输入特征提取模块与维度转换模块剔除,使网络模型恢复为 VIT,采用 ERC 图进行识别。通过消融实验验证各模块的有效性。为较好的评价各网络模型的性能,对 ECG 数据集进行均衡处理,获得标准数据集进行实验,数据集如表 3 所示。

表 3 标准本数据集

Table 3 Standard text data set

ECG 类别标签	均衡处理前数量	均衡处理后数量
F	779	2 337
N	90 006	2 500
S	2 648	2 648
V	7 005	2 500

### 3) 所提方法对比实验

为进一步验证所提方法的优势,与其他学者所发表文献中有关 ECG 识别的结果进行对比分析。为确保严格的横向对比,将选取同样采用 MIT-BIH 数据库进行 ECG 研究的文献进行分析。

## 4 实验结果与分析

实验采用基于 Python 的深度学习库 Pytorch 完成网络模型构建、训练与测试,实验在云服务器中完成,配有 2 个 RTX-4090 的 GPU。模型超参数设置为:迭代次数 40 次,采用批量大小 16,训练集、验证集与测试集按照 6:2:2 比例分配。各网络模型采用准确率、精确率、召回率、F1 值等 4 个评价指标进行对比。

按照 3.2 节设置的 ECG 彩虹码图像变换方法有效性与优势实验,将 4 类 ECG 进行图像变换,获得 ERC、STFT、GASF、GADF 与 MTF 图,如图 3 所示。通过将图 3(a)~(d) 进行组内对比,ERC 图具备有效可视化 ECG 特征的能力。ERC 图与图 3 中 STFT 图、GADF 图、MTF 图进行组间对比,ERC 图可从多角度进行 ECG 可视化,具有明显特征。

将图 3 所示图像的训练集与验证集对 LetNet5 网络模型进行训练,并用测试集进行 ECG 识别结果评估,各特征图像识别结果如图 4 所示。图 4 中 Pre 表示精确

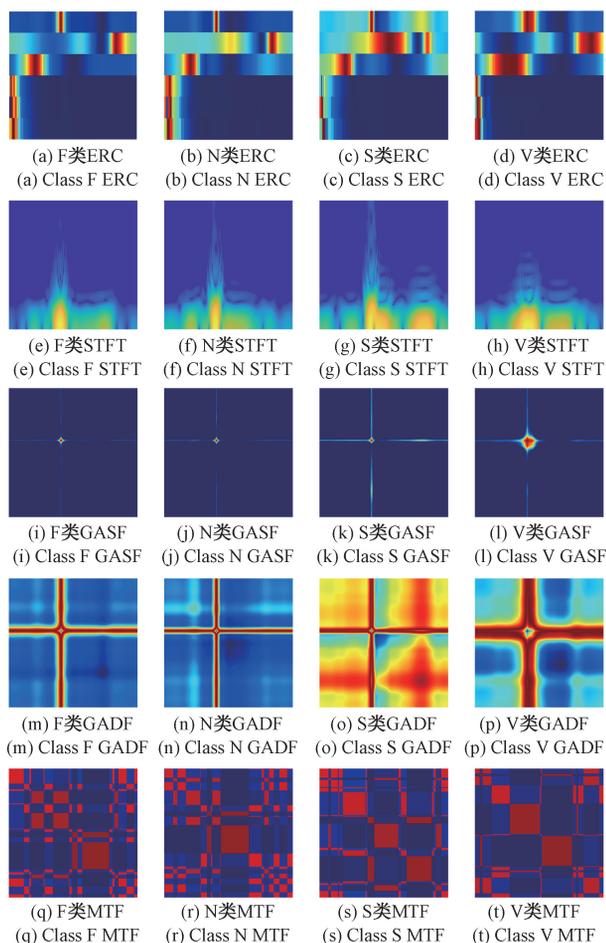


图 3 特征图像对比

Fig. 3 Feature image comparison

率,单位为%,Rec 表示召回率,单位为%。由图 4(a)可知,采用 ERC 图进行 ECG 识别的准确率为 98.7%,F1 值为 98.8%。相较于图 4(b)所示 STFT 图识别结果准确率提升了 0.1%,F1 值提升了 0.2%;相较于图 4(c)所示 GASF 图识别结果准确率提升了 0.3%,F1 值提升了 0.3%;相较于图 4(d)所示 GADF 图识别结果准确率提升了 0.4%,F1 值提升了 0.5%;相较于图 4(e)所示 MTF 图识别结果准确率提升了 0.9%,F1 值提升了 1%。在上述识别结果当中,召回率最高的为 ERC 图、STFT 图与 MTF 图,在 N 类 ECG 中获得 100%召回率。精确率最高的为 GADF 图,在 F 类 ECG 中获得 100%精确率。

为避免上述识别结果具有偶然性,对上述方法分别进行 5 次重复实验,取 5 次识别结果准确率和 F1 值的平均值进行对比分析,各方法的均值结果如图 5 所示。从图 5 中可以看出 ERC 图进行 ECG 识别的准确率为 98.71%,F1 值为 98.72%,在不同特征图像中效果最优,与图 4 结果保持一致。

综上所述,采用 ECG 彩虹码图像变换方法生成的特

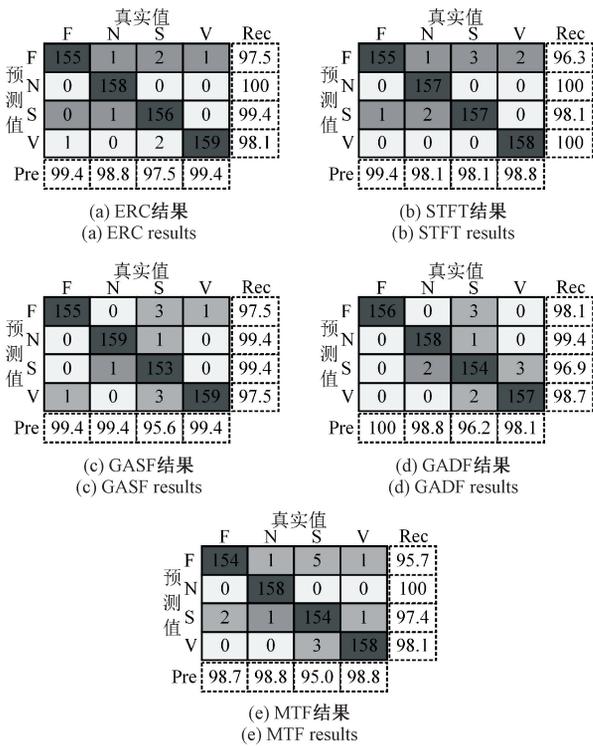


图 4 混淆矩阵对比

Fig. 4 Confusion matrix contrast

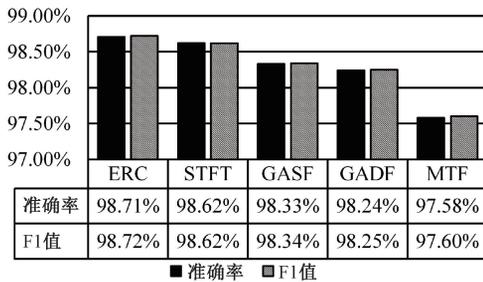


图 5 结果均值对比

Fig. 5 Result mean comparison

征图像可以用于 ECG 识别,且在不同变换方法生成的特征图像中获得了最优的识别效果,证明了该图像变换方法的有效性优势。

按照 3.2 节设置的可行性实验与消融实验进行分析,获得裁剪与未裁剪网络模型训练准确率曲线与损失率曲线如图 6 所示,方法 1 表示以 ERC 图与 STFT 图为输入,通过双输入改进 VIT 的混合模型进行 ECG 识别;方法 2 表示采用单输入特征提取模块,通过 ERC 图进行 ECG 识别;方法 3 表示采用单输入特征提取模块,通过 STFT 图进行 ECG 识别;方法 4 表示将双输入改进 VIT 的混合模型恢复为 VIT,通过 ERC 图进行 ECG 识别。从图 6 中可以看出,不同方法的准确率曲线和损失率曲线收

敛速度迅速,均在 30 次迭代后趋于稳定。其中,方法 1 的准确率曲线整体位于其他方法曲线上方,损失率曲线整体位于其他方法曲线下方。

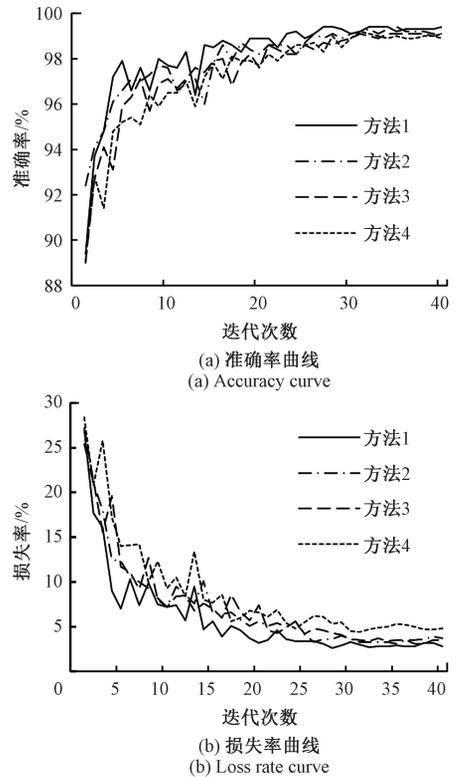


图 6 网络模型训练过程曲线

Fig. 6 Network model training process curve

由测试集获得 ECG 识别结果如图 7 所示。由图 7 可知,以 ERC 图与 STFT 图为输入数据,通过方法一进行 ECG 识别的准确率为 99.5%,F1 值为 99.5%。方法 1 与去除部分模块的方法 2 和 3 相比,识别准确率上升了 0.1%与 0.2%,F1 值同样上升了 0.1%与 0.2%。方法 1 与去除全部模块的方法 4 相比,识别准确率和 F1 值均上升了 0.4%。其中,召回率与精确率最高的均为方法 1 获得,在方法 1 中 F 类 ECG 精确率为 100%,V 类 ECG 召回率为 100%。

对上述方法分别进行 5 次重复实验,取平均值进行对比分析,各方法的均值结果如图 8 所示。从图中 8 中可以看出,平均值结果比图 7 所示的一次结果略有下降。在平均值结果中方法一获得的准确率为 99.41%,F1 值为 99.43%,准确率与其他 3 种方法相比较提升了 0.14%、0.22% 和 0.44%,F1 值提升了 0.16%、0.24% 和 0.46%。

综上所述,可以采用 ERC 图与 STFT 图为输入数据,通过双输入改进 VIT 的混合模型进行 ECG 识别,能够获得较好的识别效果,验证了所提方法的可行性。消融实

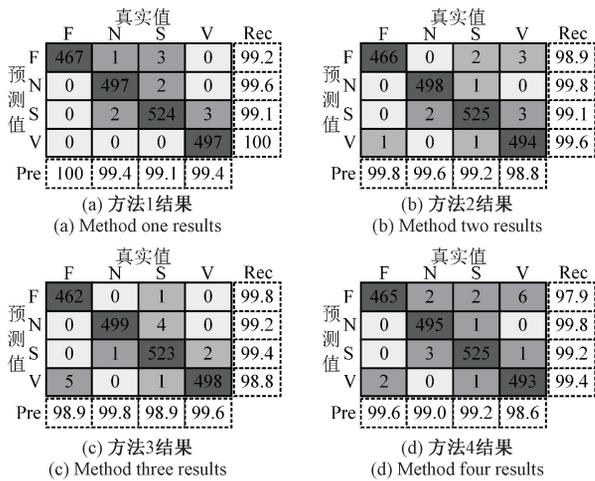


图 7 不同网络模型混淆矩阵对比

Fig. 7 Comparison of confusion matrix of different network models

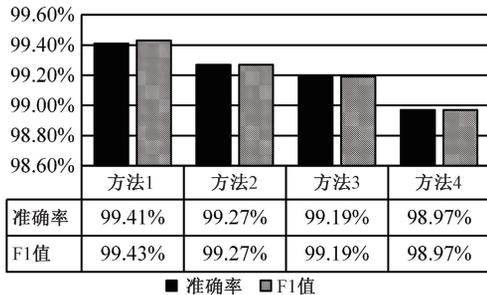


图 8 不同网络模型结果均值对比

Fig. 8 Comparison of the results of different network models

验结果表明,双输入特征提取模块与维度转换模块能够有效改善网络模型性能,在保证网络模型训练收敛速度不变的前提下提升模型的识别准确率,验证了所提模块的有效性。

按照 3.2 节设置的所提方法对比实验,选取了近 4 年内其他学者所发表文献中有关 ECG 识别的结果进行对比分析,如表 4 所示,其中混合方法是指面向 ECG 彩虹码的双输入改进 VIT 识别方法。

表 4 ECG 识别研究方法对比

Table 4 Comparison of ECG identification methods

方法	年份	准确率/%	F1 值/%	精确率/%	召回率/%
文献[22]方法	2020	97.20	93.00	93.16	-
文献[23]方法	2020	99.06	-	98.69	93.21
文献[24]方法	2021	99.11	-	95.46	92.33
文献[25]方法	2021	98.45	98.00	98.00	98.00
文献[26]方法	2021	99.11	-	93.55	96.72
文献[27]方法	2023	98.60	-	92.90	92.40
文献[28]方法	2023	99.30	99.50	96.40	98.00
混合方法	-	99.41	99.43	99.42	99.43

从表 4 中可以看出,针对 ECG 识别研究均采用了深度学习算法,学者们在各类深度学习算法上进行改进,使网络模型更加符合 ECG 固有特性,从而在 ECG 识别过程中获得更优的结果。表 4 所示方法识别准确率集中在 97.00%~99.50% 区间内,混合方法与其他文献方法相比在识别准确率、F1 值、精确率与召回率上均有较大优势。文献[28]通过 CNN-LSTM Attention 实现 ECG 识别,其性能与混合方法较为接近,混合方法在准确率指标下优于文献[28]方法,有 0.11% 的提升。虽然混合方法的 F1 值下降了 0.07%,但是精确率和召回率提升较大,分别上升了 3.02% 与 1.43%。混合方法与文献[22]、文献[23]、文献[24]、文献[25]、文献[26]、文献[27]等方法相比,其准确率的提升区间在 0.3%~2.21%,F1 值的提升区间在 1.4%~6.43%。

综上所述,面向 ECG 彩虹码的双输入改进 VIT 识别方法能够较好的实现 ECG 识别,且能够保持较高的识别准确率,在四种平均指标下,性能优于现有的多数方法,证明了所提方法的优势。

进一步验证效果,基于该方法设计 ECG 硬件采集平台及识别软件。在 LabVIEW 开发环境<sup>[29]</sup>下调用所提方法生成的动态库函数实现 ECG 识别。然后,通过 ADS1292R 芯片与 Arduino 开发板实现 ECG 采集,以无线串口模块发送数据至 ECG 识别软件进行识别。其中,ECG 采集设备如图 9 所示。

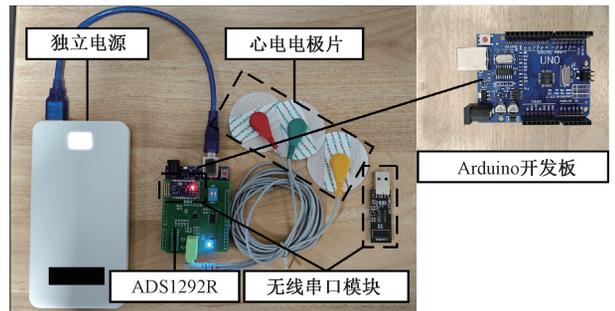


图 9 ECG 采集设备

Fig. 9 ECG acquisition equipment

由于非医学院,本实验室缺少 CVD 患者,故通过 ECG 采集设备仅获得了 N 类 ECG,在 ECG 识别软件中展现了一段采集的 ECG,如图 10 所示。从图 10 中可以看出,ECG 可以在设计的识别软件中有效可视化。

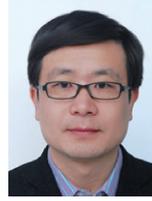
现场采集 ECG 的采样频率为 125 Hz,采集时长均大于 10 s,共采集 10 组数据,共生成 N 类 ECG 的数量为 130 条。通过 ECG 识别软件获得最终识别结果如图 11 所示。从图 11 中可知,130 条 N 类 ECG 均被正确识别,准确率为 100%,这是因为 N 类 ECG 特征较为明显,数据为同一对象的 ECG 样本。实验结果表明:通过所提方



- [ 8 ] 陈韬文, 宋家骏, 彭湘安, 等. 心电图 ECG 信号自动检测特征提取方法研究进展[J]. 电子测量技术, 2022, 45(19): 106-112.  
CHEN T W, SONG J J, PENG X AN, et al. Overview of automatic detection method of ECG signal based on feature extraction [ J ]. Electronic Measurement Technology, 2022, 45(19): 106-112.
- [ 9 ] HONG S D, ZHOU Y X, WU M, et al. Combining deep neural networks and engineered features for cardiac arrhythmia detection from ECG recordings [ J ]. Physiological Measurement, 2019, 40(5): 054009.
- [ 10 ] 刘光达, 周葛, 董梦坤, 等. 基于 FFNN 和 1D-CNN 的实时心律失常诊断系统与算法[J]. 电子测量与仪器学报, 2021, 35(3): 35-42.  
LIU G D, ZHOU G, DONG M K, et al. Real-time arrhythmia diagnosis system and algorithm based on FFNN and 1D-CNN [ J ]. Journal of Electronic Measurement and Instrumentation, 2021, 35 ( 3 ): 35-42.
- [ 11 ] JIANG M F, GU J Y, LI Y, et al. HADLN: Hybrid attention-based deep learning network for automated arrhythmia classification [ J ]. Frontiers in Physiology, 2021, 12: 683025.
- [ 12 ] RITU S, NAVIN R, RAJESH M. An empiric analysis of wavelet-based feature extraction on deep learning and machine learning algorithms for arrhythmia classification [ J ]. International Journal of Interactive Multimedia and Artificial Intelligence, 2021, 6(6): 25-34.
- [ 13 ] SERKAN K, TURKER I, MONCEF G. Real-time patient-specific ECG classification by 1-d convolutional neural networks [ J ]. IEEE transactions on bio-medical engineering, 2016, 63(3): 664-675.
- [ 14 ] MATHUNJWA B M, LIN Y T, LIN C H, et al. ECG arrhythmia classification by using a recurrence plot and convolutional neural network [ J ]. Biomedical Signal Processing and Control, 2021, 64: 102262.
- [ 15 ] ULLAH A, ANWAR M S, BILAL M, et al. Classification of arrhythmia by using deep learning with 2-decg spectral image representation [ J ]. Remote Sensing, 2020, 12(10): 1685.
- [ 16 ] 韩亮, 罗统军, 蒲秀娟, 等. 使用多尺度递归图和视觉转换器的心律失常分类[J]. 仪器仪表学报, 2022, 43(12): 149-157.  
HAN L, LUO T J, PU X J, et al. Arrhythmia classification using multi-scale recurrence plot and vision transformer [ J ]. Chinese Journal of Scientific Instrument, 2022, 43(12): 149-157.
- [ 17 ] 马广龙. 基于多模态神经网络的心电信号自动识别算法研究 [ D ]. 北京: 北京邮电大学, 2019.  
MA G L. Research on automatic recognition of ecg based on multimodal neural network [ D ]. Beijing: Beijing University of Posts and Telecommunications, 2019.
- [ 18 ] TROJOVSKÁ E, DEGHANI M, TROJOVSKY P. Zebra optimization algorithm: A new bio-inspired optimization algorithm for solving optimization algorithm [ J ]. IEEE Access, 2022, 10: 49445-49473.
- [ 19 ] 陈奔. 结合 Transformer 和 CNN 的白细胞自动分类与检测方法研究 [ D ]. 浙江: 杭州电子科技大学, 2023.  
CHEN B. Research on automatic leukocyte classification and detection method combined with Transformer and CNN [ D ]. Zhejiang: Hangzhou Dianzi University, 2023.
- [ 20 ] ALEXEY D. An image is worth  $16 \times 16$  words: Transformers for image recognition at scale [ J ]. ArXiv preprint arXiv: 2010.11929, 2020.
- [ 21 ] 沈道贤. 心脏电生理信号处理及心拍识别研究 [ D ]. 安徽: 合肥工业大学, 2022.  
SHEN D X. Research on cardiac electrophysiological signal processing and beat recognition [ D ]. Hefei: Hefei University of Technology, 2022.
- [ 22 ] 潘辉, 郑威, 张莹莹. 基于改进残差网络对心电信号的识别 [ J ]. 数据采集与处理, 2020, 35(4): 682-692.  
PAN H, ZHENG W, ZHANG Y Y. Recognition of ECG signal based on modified residual network [ J ]. Journal of Data Acquisition and Processing, 2020, 35 ( 4 ): 682-692.
- [ 23 ] LI ZH, ZHOU D SH, WAN L, et al. Heartbeat classification using deep residual convolutional neural network from 2-lead electrocardiogram [ J ]. Journal of Electrocardiology, 2020, 58: 105-112.
- [ 24 ] 李兴秀, 唐建军, 华晶. 结合 CNN 与双向 LSTM 的心律失常分类 [ J ]. 计算机科学与探索, 2021, 15(12): 2353-2361.  
LI X X, TANG J J, HUA J. Arrhythmia classification based on CNN and bidirectional LSTM [ J ]. Journal of Frontiers of Computer Science and Technology, 2021, 15(12): 2353-2361.
- [ 25 ] MAGHAWRY E, GHARIB T F, ISMAIL R, et al. An efficient heartbeats classifier based on optimizing convolutional neural network model [ J ]. IEEE Access, 2021, 9: 153266-153275.
- [ 26 ] BOUY L E, KHALIL M, ADIB A. ECG heartbeat classification based on multi-scale wavelet convolutional neural networks [ C ]. ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal

- Processing (ICASSP). IEEE, 2020: 3212-3216.
- [27] KHAN F, YU X J, YUAN ZH H, et al. ECG classification using 1-D convolutional deep residual neural network[J]. Plos One, 2023, 18(4): e0284791.
- [28] KHONDOWE E. 基于注意力机制的 CNN-LSTM 模型的心电信号分类[D]. 金华: 浙江师范大学, 2023.  
KHONDOWE E. ECG signal classification based on attention mechanism for CNN-LSTM [D]. Jinhua: Zhejiang Normal University, 2023.
- [29] 张安莉, 石卓勇, 范明邦, 等. 基于 LabVIEW 的无线运动传感器节点监测系统设计[J]. 国外电子测量技术, 2022, 41(1): 81-86.  
ZHANG AN L, SHI ZH Y, FAN M B, et al. Design of wireless motion sensor node monitoring system based on LabVIEW [J]. Foreign Electronic Measurement Technology, 2022, 41(1): 81-86.

## 作者简介



陈波(通信作者), 1999 年于合肥工业大学获得学士学位, 2005 年于同济大学获得硕士学位, 2011 年于中国科学技术大学获得博士学位, 现为合肥工业大学电气与自动化工程学院副教授, 主要研究方向为智能控制、机器视觉和生物医学工程。

E-mail: chenbo@hfut.edu.cn

**Chen Bo** (Corresponding author) received his B. Sc. degree from Hefei University of Technology in 1999, M. Sc. degree from Tongji University in 2005 and Ph. D. degree from University of Science and Technology of China in 2011, respectively. He is now an associate professor in School of Electrical Engineering and Automation, Hefei University of Technology. His main research interests include intelligent control, machine vision and biomedical engineering.