

基于信号特征的房颤心电分类

一、小组成员

张渊 2520170002

刘雨薇 2120161018

二、问题描述

房颤是最常见的心律失常之一，是由心房主导折返环引起许多小折返环导致的房律紊乱。它几乎见于所有的器质性心脏病，在非器质性心脏病也可发生。可引起严重的并发症，如心力衰竭和动脉栓塞。临床上根据房颤的发作特点，将房颤分为阵发性房颤(房颤发生时间小于 7h，常小于 24h,可自行转复为窦性心律)、持续性房颤(发生时间大于 2 天，多需电转复或药物转复)，永久性房颤(不可能转为窦性心律)。

基于单导联的房颤心率识别是临床上的一大挑战，我们希望用数据挖掘的方法改进对心电信号特征的分类，提高预测的准确性。

三、目标与功能

基于心电信号的时域频域特征，用支持向量机对信号分类，实现区分正常信号，房颤信号，其他节律与噪声的功能。

工作由以下几部分组成：

(1) 选择数据集。ECG 数据由 AliveCor 提供，ECG 采集信号为 300Hz，采集设备已经进行过初步滤波。

(2) 数据预处理。训练数据长度不一，其中以 30 秒的居多，因此在训练模型的时候先不管短于 30s 的信号。超过 30 秒的信号截成两段。最后生成一个标准长度的训练集。

(3) 特征提取。按照文献中采用的特征，提取峰均功率比和对数熵。

(4) 特征分类，用支持向量机分类，把心电信号分为正常信号，房颤信号，其他节律与噪声四类。

四、数据与平台

数据：the PhysioNet/Computing in Cardiology Challenge 2017 官网上提供了训练集的数据。主办方的数据由 AliveCor 提供，ECG 采集信号为 300Hz，采集设备已经进行过初步滤波。所有数据以.mat 的格式存储，含有.hear 文件。

平台：生理信号处理使用 Matlab，其中用到滤波的包、SVM 的包。统计分析使用 R 语言。

五、理论基础

4.1 信号处理

人体心电是一种微弱的电信号，正常人心电信号的频率范围为 0.05 到 100Hz，能量主要集中的范围为 0.25-35Hz。它有如下特征：

(1) 微弱性：通过电子设备采集到的原始心电信号很微弱，其幅值一般不超过 5mV。

(2) 低频特性：它的能量主要集中的范围为 0.25-35Hz。属于低频信号。

(3) 易受干扰：心电信号很微弱，因此在采集信号的过程中会受到很多干扰，比如测试者的状态、环境因素等。

分析心电信号首先要进行滤波。心电信号的干扰主要有三种：

(1) 工频干扰

工频干扰是受测量环境中的电磁辐射的影响产生的，频率主要集中在 50Hz。它使心电波形上产生规律的细小波纹，影响我们对心电的观察。

(2) 肌电干扰

肌电噪声的形成是人的肌肉神经系统的活动，肌肉纤维的收缩引起物电变化，经过心电电极引导、放大就形成了肌电噪声。它属于高频噪声，主要能量集中在 30~300Hz 范围内，在波形上的表现为快速变化不规则波形。

(3) 基线漂移

一般由采集心电的电子设备引起，呼吸干扰和运动伪迹也可引起。基线漂移属于低频干扰，频率一般小于 5Hz，效果相当于在心电信号上加一个低频的正弦分量。

4.2 特征点的检测

基于时域波形的特征提取需要先定位主要波形的位置，如 QRS 波。R 波是心电信号中最突出的波形，因此可以首先对 R 波进行定位，再进一步定位 Q 波、

S 波。Pan-Tompkins 方法是一种很好的 R 波检测算法。理论基础比较复杂。我们基于该算法，加入了回检的步骤，提高了检测 R 波的准确性。

4.3 特征提取

心电信号的特征有时域特征、变换域特征。AF 病人的心电特点是 RR 间期绝对不规律，因此，RR 间期是十分重要的特征。另，变换域如小波变换域的特征也广泛地运用于心电信号分类中。

4.4 SVM 分类

六、项目进展

6.1 信号处理：

因此我们采用带通巴特沃斯滤波器，实现 2-50Hz 的滤波。

对心电信号进行分割，由于主办方提供的训练数据长度不一，其中以 30 秒的居多，因此在训练模型的时候先不管短于 30s 的信号。超过 30 秒的信号截成两段。最后生成一个标准长度的训练集。

6.2 特征提取

根据文献调研的结果，决定提取几个公认效果较好的特征：

(1) 峰均功率比：

不同层的小波系数的功率谱提供 ECG 信号不同尺度的频谱信息，因此可用于分析心房活动。已知心房活动通常发生在 4-9Hz 的频率范围内。因此，在 AF 发作期间，预期在该频率范围内 AF 信号的功率更集中。

$$\rho_{\ell}^D = \frac{\max_{f \in F} S_{\ell}^D(f)}{\int_{f \in F} S_{\ell}^D(f)},$$
$$\rho_{\ell}^C = \frac{\max_{f \in F} S_{\ell}^C(f)}{\int_{f \in F} S_{\ell}^C(f)},$$

提取步骤：小波变换--> 用每一层系数小波重构

--> 求功率谱

--> 在 4-9hz 的范围求峰均功率比

(2) 对数能熵：

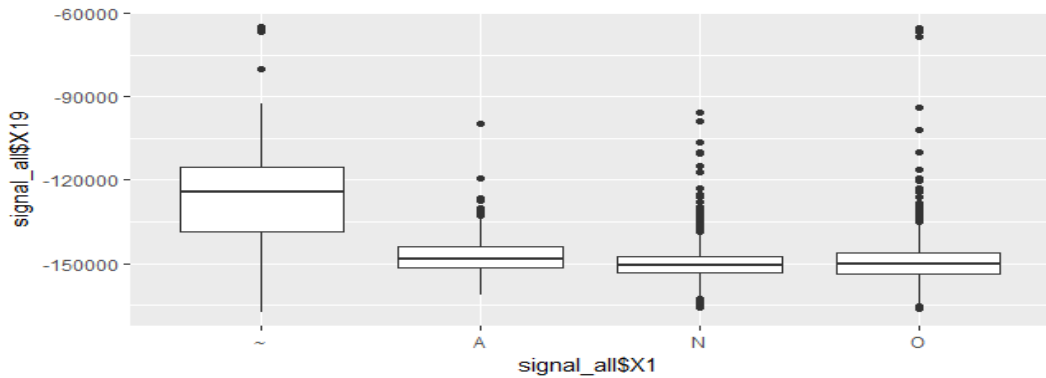
熵是衡量信号复杂度与不确定性的一种度量。小波熵能让我们度量不同信号频率的复杂度与无序程度。

提取步骤：小波变换--> 用每一层系数小波重构--> 求 log entropy

我们对心电信号进行 6 层小波分解，每层有一个高频系数一个低频系数，在对每一层尺度系数与细节系数重构后，每个系数对应一对特征。最终得到一个 24 维的特征矩阵。

6.3 对得到的特征进行统计分析

我们把不同组的特征最对比，希望粗略筛选出每个组区别较大的特征。如图所示，为某一层系数的峰均功率比在四组之间的盒图：



6.4 svm 模型训练及分类

(1) 采用 `libsvm` 包，进行多分类训练。以提取的特征作为输入。使用语句如下：`svmtrain(signal_label, signal_fea,cmd)`

(2) 训练后得到一个模型，`model`，格式为 `struct`。用 `model` 对训练集进行分类。使用语句如下：`svmpredict(test_label,test_fea,model)`。输出为四种类别。

用多分类 SVM，需要设置多个参数，我们采用网格搜索的方法来确定最优参数。最终多分类结果的准确率为 60%。

我们尝试用二分类 SVM 进行两两分类，然后对结果进行投票，产生一个多分类结果。但是效果不理想，投票的权重该如何确定是个问题。

七. 结果与讨论

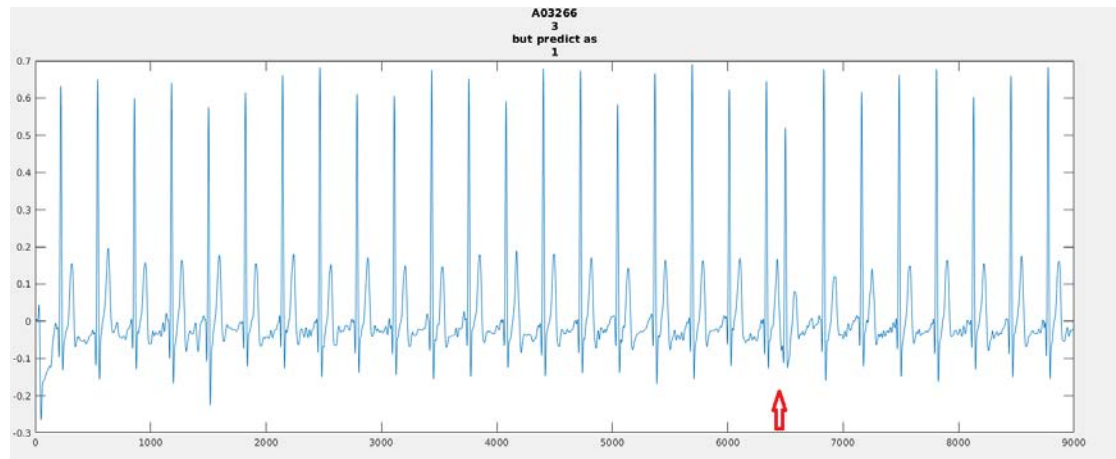
用 svm 进行两两分类的结果如下：

| | normal&AF | normal&other | normal&noisy | AF&other | AF&noisy | other&noisy |
|-----|-----------|--------------|--------------|----------|----------|-------------|
| 准确率 | 56.63% | 76.68% | 89.20% | 66.79% | 88.44% | 99.34% |

多分类的准确率为 61.2%

特征的选取非常重要，各组间的分类效果不好主要是：没有能明显区分四组的特征。每个组的特征值的分布都有重叠，导致分类有重合，效果不好。

查看未能正确分类的信号：



如图是一个其他节律信号被分类成 AF，即因为中间出现了一个“二联律”干扰了特征的识别。

因此，后续的工作主要是尝试不同的特征，找到区分度较好的特征。还可以尝试不同的分类方法。此外考虑是否可以用深度学习识别图像来完成房颤特征的识别？图像识别还没有运用到心拍分类中，可以进一步考虑是否可行。