

a. Shen Y., Aliamiri A., Voisin M., Avati A., Hannun A., and Ng A.:

Ambulatory Atrial Fibrillation Monitoring Using Wearable

Photoplethysmography with Deep Learning. arXiv:1811.07774v2. Nov 2018

- i. 吴恩达和三星合作，和Rajpurkar 2017的ECG一样，也是做心脏病分类（AF，心房颤动），但是采集方式从Single-lead ECG转为用可穿戴PPG，即[Photoplethysmography 光体积变化描记图法](#)：随着心跳会有一压力波通过血管进行传递，这个波会稍微改变血管的直径，PPG就是利用这个变化，是一种无需测量生物电信号就能获得心脏功能信息的光学技术。PPG主要用于测量血氧饱和度 (SpO2)，但也可不进行生物电信号测量就提供心脏功能信息。通过佩戴在智能手表上的PPG，其持续采集的数据可以用来定量AF Burden（心房颤动负荷），从而确定心脏病。
- ii. 本论文的工作是PPG数据上的二分类。其输入是30秒、20Hz的PPG输入，通过一个50层的1D CNN ResNeXt，输出是二分类标签（是否AF）。Loss是用二值交叉熵（Binary Cross Entropy）。前处理是用一个Finite Impulse Response (FIR) low-pass filter 低通滤波器，把数据下采样到5Hz。
- iii. 训练数据是从81个病人身上采集的8000小时PPG数据。这些数据是在ambulatory free-living 状态（即在家休养状态）下采集的，而不是在医院、受控的情况下采集，因此数据中存在大量的Motion Noise（运动噪声）。这些运动噪声可以通过手表的加速度感应器发现，并用于测试集中。数据集分两类：一类是29个病人的402段8小时记录，其中13人有AF，2人正常，14人有其他心脏病。另外一类则是53个健康人的341段3小时记录。病人的标注其实还是在同期录制的ECG上标注的，而且标注的时候还是多分类，和[Rajpurkar 2017](#)一样
- iv. 评估则是和[Hannun 2019](#)一样，加入了ROC、AUC等医学用统计计算。其结果是AUC在0.948，而且基本上不受Montion影响，而对照的基线(一个传统方法Inter-Beat Interval (IBI))则是最好AUC在0.86，而随着运动增加，可以一直下降到0.62左右。
- v. 比[Rajpurkar 2017](#)等做的更好的是，他们分析了一下网络中不同层次的特征图，与目前已知医学分析知识之间的相关性关系。并且他们也进行 saliency mapping分析，来确认输入中的哪些PPG节拍和最终网络输出的结果最相关。另外，他们也用t-SNE来可视化高层特征，如标签分布、用Motion的相关性等
- vi. Note : systolic and diastolic 收缩和舒张

b. Hannun A., Rajpurkar P., Haghighpanahi M., Tison G., Bourn C., Turakhia M., and Ng. A.: Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nature Medicine Vol 25, Jan 2019

- i. 灌水文：把2017年的论文，加了些医学术语，多攒了些数据，就从机器学习论文转为投到了Nature的医学论文
- ii. 比[Rajpurkar 2017](#)增加了接近一倍的数据（5.3万病人，9.1万记录），号称比MIT-BIH（也就是后来的PhysioNet）数据更强大。但是输出的分类变成了9个有病、加上sinus rhythm和noise，把AFIB和AFL合并成一类，去掉CHB。输出不再是1秒一次，而是256采样点（200Hz下1.28s）一次。对于数据集的医学描述也比[Rajpurkar 2017](#)更“医学”。评估除了F1，也用上了ROC。
- iii. 但是网络还是用[Rajpurkar 2017](#)的33层ResNet，F1也只是大概从0.77变成了0.80，并没有本质变化。评估也是在1秒的Sequence级别和30秒的set级别上进行，也是委员会做Ground Truth、6个个人医生做对比。另外，还把他们的2017年去PhysioNet上打比赛的成绩也掏出来做事。
- iv. 有趣的是，他们觉得LSTM和双向RNN都没有用，CNN就够了。
- v. 另外一个有趣的是他们对于在2017年之后大量类似论文的总结，如他们所描述的：Much of the previous work to employ DNNs toward ECG interpretation has focused on single aspects of the ECG processing pipeline, such as noise reduction^{13,14}, or feature extraction^{15,16}, or has approached limited diagnostic tasks, detecting only a handful of heartbeat types (normal, ventricular or supraventricular ectopic, fusion, and so on)^{17–20} or rhythm diagnoses (most commonly atrial fibrillation or ventricular tachycardia)^{21–25}

c. Rajpurkar P., Hannun A., Haghighpanahi M., Bourn C., Ng A., Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks. arXiv:1707.01836v1. Jul 2017

- i. 号称打败人类心脏学家、但是被骂的狗血淋头的心律失常检测方法，用34层的CNN来给ECG序列做序列分类，14个类别是sinus rhythm (SINUS 窦性心律)、noise、以及包含Atrial Fibrillation (AFIB 心房颤动)在内得12种病情等。
- ii. 简化成一个Seq2Seq问题，Loss就是标签序列和分类结果序列的Cross-entropy，网络是一个很规矩的简化ResNet：33层的BN-ReLu-Dropout-Conv，前后用max pool跳接。最后一层是全连接层，加上一个softmax输出多分类标签。输入因此是一个30*200维的向量，输出

是30维的标签，都不是很大。Filter“长度”是16，数量是64k（k从1开始，每4个Residual block加1）。

- iii. 数据是采样率为200Hz的30秒ECG信号及每秒一次的标注数据，用Zio Patch监护仪从2.6万个病人采集而来对6.4万记录，尤其是那些展示了心跳问题的病人。正常心率病人大概在2.2万，noise在1万，其余病人情况在1000到6000之间。标签是专家用Web工具标注的，但是是基于规则、没有校验的标注。Training:Validation是9:1。测试集是单独从328个病人取得336段记录，ground truth标签是3个医生一致出来得，对比是随便找了6个医生，单独标注。
- iv. Noise作为一个分类，其判定标准是：“The noise label is assigned when the device is disconnected from the skin or when the base line noise in the ECG makes identification of the underlying rhythm impossible”
- v. 评价方法是基于1秒的Sequence和基于30秒的Set，后者因此不会受到时间错位问题的影响。结果方面还是有很多心律失常与窦性心律混淆

d. Porumb et al., A Convolutional Neural Network approach to detect congestive heart failure. Biomedical Signal Processing and Control, Jun 2019

- i. 在PhysioNet的原始ECG数据上用CNN预测Congestive Heart Failure (CHF)，在自己的测试集上得到了100%。但是只是做了一个简单的二分类：正常还是CHF，并没有去触碰更复杂的relapse rate（复发率）的问题。
- ii. 他们对比的基线是以前用的Heart Rate Variability (HRV) - 即如何从ECG中计算心跳之间、一段时间上的variation。而这种检测Variation的模型往往需要长时间信号（比如24H）
- iii. 其工作主要包括：
 - 从PhysioNet数据中挑选18个正常、15个有病的人，每个人大约20h的双通道ECG数据，采样率为250Hz或者128Hz（归一化到128Hz）
 - 前处理是在R peak之前用235ms的时间窗，之后用390ms的时间窗来计算一个heart beat。同时为了去掉同样的beats，随机进行5选1来挑选beats，最后得到了大约27万正常beat、21万有病beat
 - 输入的时候，一个病人的数据要么在测试集、要么在训练集，数据随机分配了10次进行训练和评估，评估结果平均出来。所谓结果就是对单个beat进行分类，或者对于5分钟的beat进行一个投票。评价方法是accuracy, precision, sensitivity, specificity和AUC
 - 网络方面，使用3层1D conv来提取特征，然后用FC来分类。中间加入了一个GRAD-CAM，用于对特定输入的sequence进行其分类的可视化分析

- 由于是在5分钟的beat分类结果上来处理，结果很少。少数错误情况集中在特定3、4个病人的某几段数据上，其错误更多是来源于数据质量本身：噪声、movement artifacts等。最后是用GRAD-CAM做了一个热度图
- 最后总结，由于没有使用HRV，所以无需手动加工数据特征；网络很浅；单个人数据只用于训练或者只用于测试；可视化结论。因此其结果能够帮助所谓Clinical Decision Support System。
- 但是其问题在于数据分布比较偏颇，数据集太小。如果有更多数据集，可以用于Neuro-Behavior的研究（因为之前有研究使用HRV分析来做的）