

大規模心電図データを用いた波形区間検出誤差に基づく異常群分類モデルの構築

小西 拓己 【コミュニケーション&コラボレーション研究室】

1 はじめに

心電図は、心臓の電氣的活動を非侵襲的に計測できる検査であり、不整脈や心疾患の評価において広く用いられている。心電図から診断される疾患の一つに、QT延長症候群 (Long QT Syndrome: LQTS) がある。LQTSは致死性不整脈や突然死の原因となり得る心疾患であり、早期発見および適切なリスク評価が重要である。LQTSの診断においては、心電図上の補正QT間隔 (QTc) が基盤となる指標であるが、QT区間の判定は専門医にとっても困難な症例が存在する。特に、大規模データを抱えるスクリーニングにおいては、人手による確認には限界があり、自動化手法の構築が強く求められている。

従来の研究では、LQTSの分類精度の向上を目的として、アルゴリズムやモデル構造の改良に重点が置かれてきた。LQTSの診断に用いられる心電図には、ノイズや調律異常等により、診断に適さない心電図波形が含まれる場合があるが、入力波形の選定が及ぼす影響については、十分に議論されていない。

本研究では、計測信頼性の低い心電図波形を識別する方式を提案し、信号品質が良好かつ心調律が安定した波形を事前に選別する枠組みを示す。従来手法との比較を通じて、入力波形の選別の有効性を検証する。

2 方法

本研究では、2006年9月から2024年2月までの約54万件の12誘導心電図レコードを使用した。各レコードは500 Hzでサンプリングされた10秒間の波形データである。

波形品質スコアは、vital_sqiライブラリを用いて各誘導から36個の信号品質特徴量を抽出し、Isolation Forestにより異常度スコアを算出した。波形は異常度スコアの分位点に基づき、Good (下位70%: 重み1.0)、Border (下位70-90%: 重み0.5)、Bad (上位10%: 除外)の3群に分類した。

分類モデルにはSE-ResNet[1]とMulti-head Attentionを組み合わせた構成を採用し、入力として12誘導心電図波形および臨床変数(年齢、性別、身長、体重)を用い、出力としてLQTSの確率値を得た。LQTSの正解ラベルは、自動診断による $QTc \geq 500$ msを陽性(約4%)、 $QTc < 500$ msを陰性として定義した。5分割交差検証により評価を行い、バッチサイズ64、最大エポック数100 (Early stopping: patience 15)の設定で、波形選別を行ったデータを用いたモデルと選別を行わないモデルをそれぞれ学習させ、ROC-AUC, PR-AUC, 感度, 特異度, F1スコアにより分類性能を比較した。

3 結果

表1に分類性能の比較結果を示す。波形選別を導入したモデルは、選別を行わないベースラインモデルと比較して、全体的に分類性能の向上を示した。ROC-AUCは0.977から0.985に、PR-AUCは0.710から0.750に向上した。特に感度 (Sensitivity) は0.551から0.680へと大きく改善した。一方で、特異度 (Specificity) は0.993から0.994と高い水準を維持しており、偽陽性の増加を抑えたまま検出性能の向上が確認された。また、F1スコアは0.644から0.720へと上昇し、陽性クラスに対する識別性能の改善が示された。

表1 分類性能の比較

手法	ROC-AUC	PR-AUC	Sens.	Spec.	F1
ベースライン (波形選別無し)	.977	.710	.551	.993	.644
重み付け (波形選別有り)	.985	.750	.680	.994	.720

4 考察

従来手法に比べ、感度が大きく改善され、従来モデルでは見逃されていた異常例の一部を新たに検出できるようになったと考えられる。極端なQTc延長を示すレコードでは、ノイズや調律異常によりQT区間が不明瞭となる場合が多く、これらの波形が学習データに含まれ、モデルが異常群に共通する形態的特徴を十分に学習できなかった可能性がある。波形選別の導入により、このような不安定な波形の寄与が抑制され、異常群の検出率が向上したと考えられる。

5 まとめ

本研究では、波形区間検出誤差に着目し、波形品質評価を導入した異常群分類方式を構築した。その結果、入力波形の選別により分類性能の向上が確認され、計測信頼性の低い波形を事前に抽出する有効性が示された。本研究は入力品質に基づく新たな診断支援の枠組みを提示し、LQTS検出の信頼性向上に寄与するものである。本手法の限界として、心電計に搭載された自動診断機能が出力したQTc値を正解ラベルとして用いている点が挙げられる。その結果、専門医による臨床診断結果との乖離が生じ得る可能性がある。今後の展望として、専門医によるアノテーションや複数計測アルゴリズムの合意に基づくラベルの導入によって、より厳密な評価が可能になると考えられる。

参考文献

- [1] J Park, et al. Ecg-signal multi-classification model based on squeeze-and-excitation residual neural networks. *Applied Sciences*, Vol. 10, No. 18, p. 6495, 2020.