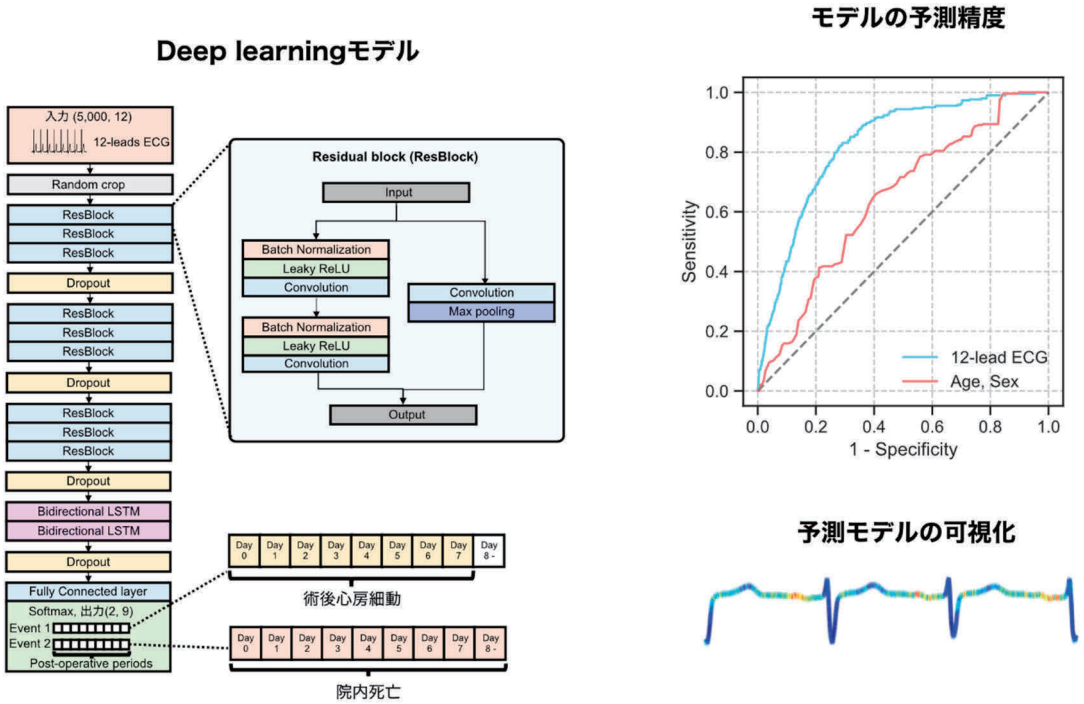
170　“心電図を診る目”をもつ人工知能モデルの開発	遠山　岳詩

【目的】近年、人工知能の技術革新はめざましく、心電図検査においても Deep learning の技術を応用した研究が盛んに行われている。大量のデータを用いて診断や予測を行う Deep learning モデルは、通常、数百万〜数千万の大量のデータを必要とすることから、開発対象が限定され、大量にデータを準備することが困難な疾患・病態は Deep learning 研究の対象にならない。そのような課題に対して、心電図情報を圧縮する Deep learning の基礎となる“心電図を見る目”を持つモデルを構築・活用することで、少数データでの Deep learning モデルの開発に繋げることが可能となる。そこで、今回は術後に発症する心房細動（POAF）を研究対象として、“心電図を見る目”を活用した Deep learning モデルを構築することを目的とした。POAF は、入院期間の延長、発作性心房細動の発症、後期脳梗塞のリスクを高めることが知られており、POAF のスクリーニングモデルを開発することは臨床的に重要である。

【方法】本研究では、九州大学病院にて外科的手術を受けた 18 歳以上の心房細動のない被験者 27,564 人を対象として、手術の 30 日前までに記録された非心房細動の心電図データ 43,980 件を解析した。院内死亡と早期退院による打ち切りや競合リスクを考慮し、Time-to-event 解析を扱うことができる深層学習モデルを構築した。本モデルは、事前に年齢・性別等を予測させる課題を与えた Deep learning モデル（心電図を見る目に相当）を構築し、それを Fine tuning することで、モデルの精度向上を図った。開発したモデルの性能は、時間依存性の評価指標を用いて評価した。

【結果】内部検証データセットにおける POAF の発生率は 3.6%であった。構築したモデルの POAF 予測精度は、時間依存性 c 統計量で 0.83（95%信頼区間 : 0.80〜0.85）であった。また、感度 79.9%、特異度 73.5%、陽性的中率 10.2%、陰性的中率 99.0%を達成した。顕著性マップによる解析では、特に aVF、V1、V2、V5、V6 の P 波と ST 領域に着目していることが明らかとなった。術前心電図を用いた深層学習モデルは、比較的小規模のデータベースをもとに構築されたものであるが、高い陰性的中率を持ち、術後に適切な心電図モニタリングが必要となるハイリスクの POAF 患者を特定するための実用的なモデルとなり得る。

術後心房細動のハイリスク患者をスクリーニングするための Deep learning モデル



170. “心電図を診る目”をもつ人工知能モデルの開発

遠山　岳詩

九州大学病院 ARO 次世代医療センター

Key words：Deep learning，心電図，術後心房細動，fine-tuning

緒　言

近年、人工知能の技術革新はめざましく、医療においても Deep learning の技術を応用した研究が盛んに行われている。循環器領域の最も基本的な検査である心電図検査においては、不整脈自動診断［1］、不整脈非発作時の心電図から潜在する不整脈の予測［2］、貧血の有無の判断［3］などが報告されている。大量のデータを用いて診断や予測を行うこれらの Deep learning モデルは、すでに循環器専門医の能力に匹敵、もしくはそれ以上の性能であると言える。しかしながら、通常高精度の Deep learning モデルには、数百万～数千万の大量のデータを必要とすることから、開発される対象が限定されてしまい、少数例のデータしかない疾患・病態はそもそも Deep learning の研究対象にならない。そのような課題に対して、心電図情報を圧縮する Deep learning の基礎となる“心電図を見る目”を活用することで少数データでの Deep learning モデルの開発を促進することが可能となる。

大量のデータを用意することが困難な課題として、本研究では、術後心房細動（POAF）の発症を術前の心電図から予測する Deep learning モデル構築を対象とした。POAF は術後のストレスによって誘発される心房細動である。POAF はこれまで良性で一過性の心房細動として認識されていたが［4］、近年の研究により、POAF は入院期間の延長、発作性心房細動の発症、後期における脳卒中の高リスクと関連することが明らになっている［5，6］。したがって、POAF の高リスク患者を同定する Deep learning モデルの開発は、臨床的意義のある重要なスクリーニングツールとなり得る。本研究は 2023 年 2 月、Circulation:Arrythmia and electrophysiology にて報告した［7］。

方　法

1. データソース

本研究は、九州大学病院の倫理審査委員会の承認を受け、ヘルシンキ宣言に基づき実施した。研究の対象は、2015 年から 2020 年の間に九州大学病院で外科手術を受け、手術の 30 日前までに記録された非心房細動の心電図を有する 18 歳以上の入院患者である。術前心電図のうち、心房細動、ノイズの多い心電図、電極が外れたもの、重複した心電図、右側心電図、心拍数 120 bpm 以上または 40 bpm 未満、心室頻拍または頻回の心室性期外収縮、評価不能または術前の心電図検査として不適切と判断されたものは解析から除外した。なお、洞調律、ペーシング調律は本研究の対象とした。Time-to-event データを扱える Deep learning モデルを構築するため、POAF の発生は、術後 7 日以内の 12 誘導心電図に心房細動が記録されたものと定義し、院内死亡は競合リスクとして扱った。図 1 のように、12 誘導心電図の 10 秒間の波形データを Deep learning モデルの入力として使用した。入力データをランダムにトリミングすることでデータ拡張を行った。心電図データは複数の残差ブロック（ResBlock）と時系列データを処理する LSTM からなるニューラルネットワークで処理された。モデルの出力は、イベント種類（POAF または院内死亡）および離散時間（0～7 日目、8 日目以降）の数に対応する 2×9 テンソルとした。推論時もモデルのランダム性を有効にして推論を 10 回繰り返すことで、その結果をアンサンブルした。

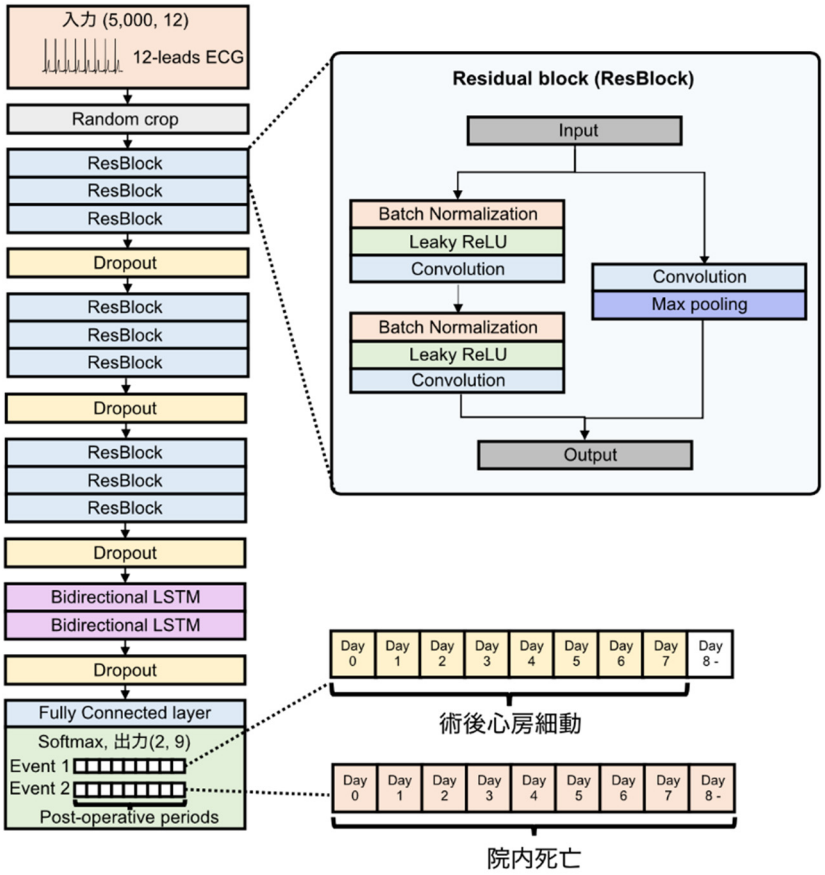


図 1. 術後心房細動のハイリスク患者を検出する Deep learning モデル構造

12 誘導心電図の時系列データをモデルの入力とし、術後心房細動と院内死亡のそれぞれのイベントに対する離散時間カテゴリを出力とする。本モデルの入力から隠れ層部分は、事前に年齢、性別を予測する Deep learning モデルを事前学習させたものを使用している。

まずグリッドサーチを行い、モデルの最適なハイパーパラメータを特定した。本研究は少数例でのモデル開発であり、POAF モデルの構築前に、非術前患者を含む 137,551 件の 12 誘導心電図から年齢・性別の予測を行うモデル（心電図を見る目に相当）の事前学習を行った上で、今回の POAF 予測のモデルとして Fine-tuning することで、少数例のトレーニングデータでの精度向上を図った。モデル構築の処理のコードは GitHub を参照されたい（https://github.com/TakeshiTohyama/DLmodel_POAF）。なお、Deep learning モデルが 12 誘導心電図のどこに注目して予測しているかを明らかにするために、SmoothGrad による顕著性マップを用いた。

結　果

1. 本研究の解析対象とモデルの構築

本研究では、POAF 予測モデル構築に先立って、137,551 件の 12 誘導心電図から年齢・性別を予測するモデルの事前学習を行い、”心電図を診る目”にあたる基本モデルの構築を行った。POAF 予測モデルの構築においては、心房細動のない患者 27,564 人の 43,980 件の術前心電図を、被験者レベルでトレーニング用、パラメータ決定用、内部検証用の 3 つのデータセットに 7：1：2 の割合で分け、これらのデータでモデルのパラメータの Fine-tuning と検証を行なった。なお、内部検証用のデータにおける、POAF の発生率は 3.6%であった。

2. モデル評価

図 2a 及び 2b は、構築された POAF 予測の Deep learning モデルの予測性能であり、7 日後の時間依存性 c 統計量は 0.83（95%信頼区間：0.80～0.85）であった。参考として年齢と性別を予測因子とした単純な予測モデルの c 統計量は 0.64 であった。術後 7 日目の臨床指標について、モデルは感度 79.9%、特異度 73.5%、陽性的中率 10.2%、陰性的中率 99.0%を達成した。

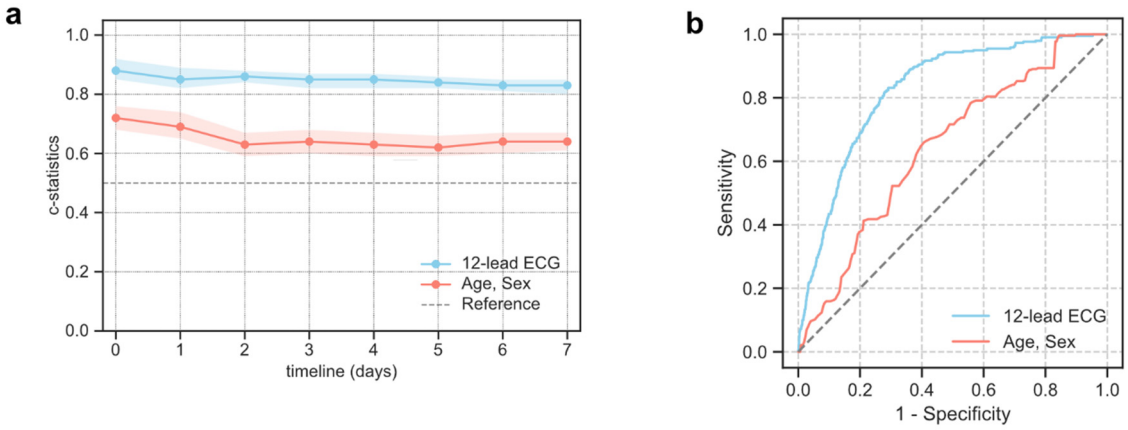


図 2. 術後心房細動の予測モデルの精度評価

a）予測モデルの時間依存性 c 統計量。x 軸は術後日数、y 軸は推定時間依存性 c 統計量を表す。青は 12 誘導心電図からの予測する Deep learning モデルであり、赤は年齢と性別からの予測モデルである。それぞれの斜線部分は 95%信頼区間を示す。

b）術後 7 日以内の心房細動の予測に対する受信者動作特性（ROC）曲線。

3. モデル解釈

予測性能に寄与した心電図領域について明らかにするため、SmoothGrad を用いた顕著性マップにて評価を行った。まず、図 3a は代表的な心電図波形であるが、P 波や ST の電圧値の低い領域に注目していた。これは、以前の発作性心房細動の発症予測［8］で報告されたものとほぼ同等の領域について注目している結果であった。また、誘導別では、aVF、V1、V2、V5、V6 を重要視していることが明らかとなった（図 3b）。

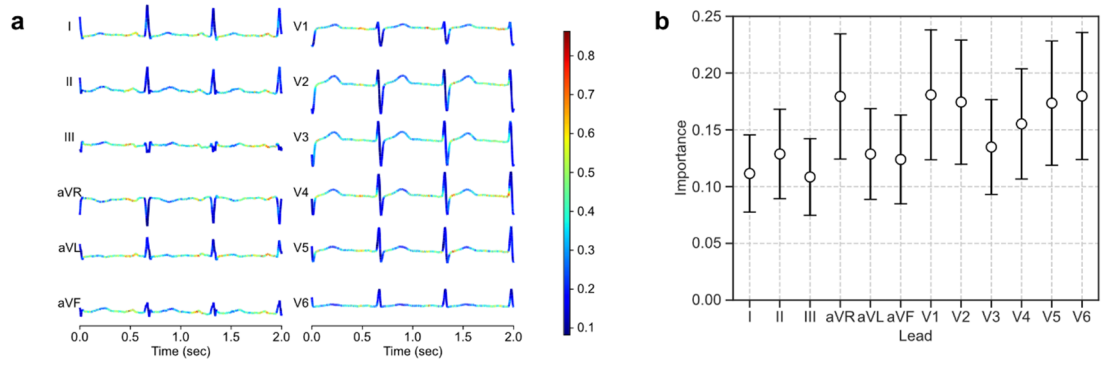


図 3. SmoothGrad による Deep learning モデルの可視化

a）術後心房細動の予測における重要視した波形領域を顕著性マップにて表現した。
波形領域においては、P波やST領域の電圧値の低い領域に注目している。

b）電極ごとの重要度（平均値±標準偏差）。特に、aVF、V1、V2、V5、V6を重要視している。

考　察

先行研究では、臨床変数に基づく POAF の予測モデルは 0.64 から 0.72 であったと報告されている［9，10］。本研究結果を、以前の報告と直接比較することは難しいが、今回の結果は、Deep learning モデルが高い精度で POAF を予測できることを示した。高い陰性的中率と低い陽性的中率は、POAF の発症率が低いことに起因している可能性があるが、本モデルによる予測は、ハイリスク患者を特定するための初期スクリーニングとして使用されるため、高い陰性的中率は臨床応用において最も重要な要件となる。本研究における限界として、後ろ向き研究として実施されたため、POAF の検出が不十分であった可能性があげられる。今後は、前向きに収集したデータを使ってもモデルの改良・検証を行っていく必要がある。今回開発した術前心電図を用いた深層学習モデルは、術後に適切な心電図モニタリングが必要となるハイリスクの POAF 患者を特定するための有効なスクリーニングツールとなり得る。

共同研究者・謝辞

本研究の共同研究者は、九州大学大学院医学研究院循環器内科学の井手友美、池田昌隆、永田拓也、坂本和生、筒井裕之、九州大学病院 ARO 次世代医療センターの廣瀬将行、船越公太、岸本淳司、戸高浩司、そして、九州大学病院メディカル・インフォメーションセンターの中島直樹である。

文　献

1） Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nature Medicine 2019; 25(1): 65–69. DOI: 10.1038/s41591-018-0268-3.

2） Attia ZI, Noseworthy PA, Lopez-Jimenez F, Asirvatham SJ, Deshmukh AJ, Gersh BJ, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. The Lancet 2019; 394(10201): 861–867. PMID: 31378392. DOI: 10.1016/S0140-6736(19)31721-0.

3） Kwon J myoung, Cho Y, Jeon KH, Cho S, Kim KH, Baek SD, et al. A deep learning algorithm to detect anaemia with ECGs: a retrospective, multicentre study. The Lancet Digital Health 2020; 2(7): e358–e367. PMID: 33328095. DOI: 10.1016/S2589-7500(20)30108-4.

4） Koshy AN, Hamilton G, Theuerle J, Teh AW, Han HC, Gow PJ, et al. Postoperative Atrial Fibrillation Following Noncardiac Surgery Increases Risk of Stroke. The American Journal of Medicine 2020; 133(3): 311-322.e5. PMID: 31473150. DOI: 10.1016/j.amjmed.2019.07.057.

5） Lin MH, Kamel H, Singer DE, Wu YL, Lee M, Ovbiagele B. Perioperative/Postoperative Atrial Fibrillation and Risk of Subsequent Stroke and/or Mortality: A Meta-Analysis. Stroke 2019; 50(6): 1364–1371. PMID: 31043148. DOI: 10.1161/STROKEAHA.118.023921.

6） Hravnak M, Hoffman LA, Saul MI, Zullo TG, Whitman GR. Resource utilization related to atrial fibrillation after coronary artery bypass grafting. American Journal of Critical Care : An Official Publication, American Association of Critical-Care Nurses 2002; 11(3): 228–238. PMID: 12022486.

7） Tohyama T, Ide T, Ikeda M, Nagata T, Tagawa K, Hirose M, et al. Deep Learning of ECG for the Prediction of Postoperative Atrial Fibrillation. Circulation: Arrhythmia and Electrophysiology 2023; 16(2): e000752. PMID: 24663335. DOI: 10.1161/CIRCEP.122.011579.

8) Khurshid S, Friedman S, Reeder C, Di Achille P, Diamant N, Singh P, et al. ECG-Based Deep Learning and Clinical Risk Factors to Predict Atrial Fibrillation. Circulation 2022; 145(2): 122–133. DOI: 10.1161/CIRCULATIONAHA.121.057480.
9) Mariscalco G, Biancari F, Zanobini M, Cottini M, Piffaretti G, Saccocci M, et al. Bedside tool for predicting the risk of postoperative atrial fibrillation after cardiac surgery: the POAF score. Journal of the American Heart Association 2014; 3(2): 1–9. PMID: 24663335. DOI: 10.1161/JAHA.113.000752.
10) Kolek MJ, Muehlschlegel JD, Bush WS, Parvez B, Murray KT, Stein CM, et al. Genetic and Clinical Risk Prediction Model for Postoperative Atrial Fibrillation. Circulation: Arrhythmia and Electrophysiology 2015; 8(1): 25–31. PMID: 25567478. DOI: 10.1161/CIRCEP.114.002300.