

12 誘導心電図の深層学習による致死性不整脈発症予測

Prediction of Fatal Arrhythmias by Deep Learning of 12-Lead ECG

赤田 翔¹⁾ 西原 大裕²⁾ 中村 一文³⁾ 高谷 陽一²⁾ 右田 剛史⁴⁾ 高橋 規一⁴⁾

Sho AKADA Takahiro NISHIHARA Kazufumi NAKAMURA Yoichi TAKAYA Tsuyoshi MIGITA Norikazu TAKAHASHI

1 はじめに

心疾患は悪性新生物に次ぐ日本人の死因第 2 位であり、令和 5 年度には死因全体の 14.7% を占めた [1]。高齢化による老年人口の増加に伴い、心疾患による死亡者数は今後増加していくと予想される。心疾患の一種である心筋症は、心臓の機能障害とともに心筋の異常を引き起こす病気であり、症状によって、拡張型心筋症、肥大型心筋症、拘束型心筋症、不整脈原性右室心筋変性症の 4 種類に分類される [2] (分類不能心筋症を加えて 5 種類とすることもある [3])。

肥大型心筋症 (HCM: Hypertrophic Cardiomyopathy) は、突然死の原因になり得るため、HCM 患者には長期にわたる慎重な経過観察が必要である。実際、HCM 患者の死因の約 40% は突然死であると報告されている [4]。したがって、突然死リスクの高い HCM 患者を事前に同定することは重要な課題である。この課題の解決に向けた取り組みとして、2011 年の ACCF/AHA ガイドライン [5] や 2014 年の ESC ガイドライン [6] の策定があるが、これらには限界があることが指摘されている。前者においては、突然死の陰性的中率は高いが陽性的中率は低い [7]。後者においては、高リスクと判定される HCM 患者の突然死発生率は高いものの、突然死またはそれに近い状況に陥る HCM 患者が事前に高リスクと判定される割合は半数以下と低い [8]。

本稿では、HCM 患者の心電図データを人工知能 (AI: Artificial Intelligence) の一種である畳み込みニューラルネットワーク (CNN: Convolutional Neural Network) に適切に学習させることにより、突然死の主要因子の 1 つである致死性不整脈の発症を約 90% の高い精度で予測できることを示す。本研究で用いる CNN モデルは、心房細動の予測に関する研究 [9] で用いられたモデルにわずかな変更を加えたものである。CNN の学習には最適化アルゴリズム Adam [10] とミニバッチ学習を用い、それらにベイズ最適化によるハイパーパラメータの自動探索と層化グループ K 分割交差検証を組み合わせる。学習およびテストに用いるデータは、岡山大学病院循環器内科において収集された、HCM 患者 179 人の致死性不整脈発症前の 12 誘導心電図データ 1,013 件である。文献 [9] では 180,922 人の心電図データ 649,931 件が用いられた

- 1) 岡山大学大学院環境生命自然科学研究科 Graduate School of Environmental, Life, Natural Science and Technology, Okayama University
- 2) 岡山大学大学院医歯薬学総合研究科 Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University
- 3) 岡山大学学術研究院医歯薬学域 Faculty of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University
- 4) 岡山大学学術研究院環境生命自然科学学域 Faculty of Environmental, Life, Natural Science and Technology, Okayama University

ことを考慮すると、本研究のデータは極めて少ない。また、文献 [9] では計算量削減のために III 誘導と単極子誘導を除いた 8 波形を学習に用いているが、本研究では 12 波形すべてを用いる。

2 CNN モデル

本研究の目的は、HCM 患者が将来に致死性不整脈を発症するか否かを、その患者の 12 誘導心電図を用いて判定する CNN を構築することである。12 誘導心電図は 12 種類の波形を 10 秒間記録したもので、心電計のサンプリング周波数が 500Hz (サンプリング間隔は 2 ミリ秒) であるから、各波形は 5,000 個の数値からなる時系列となる。上記の目的のためには、60,000 (= 5,000 × 12) 個の入力ノードと、発症と非発症に対応する 2 個の出力ノードを有する CNN モデルの構造を適切に設計し、致死性不整脈を発症していない HCM 患者の心電図データと、致死性不整脈を発症した HCM 患者の発症前の心電図データを学習させる必要がある。

本研究で用いる CNN モデルの全体構成を図 1 に示す。このモデルは心房細動の予測に関する研究 [9] で用いられた CNN を少し変更したものである。具体的には Classification Axis における全結合層を 2 層から 3 層に増やした。全結合層が 2 層では推論が不安定になることが予備実験において確認されたためである。図 1 のモデルは Temporal Axis, Lead Axis, Classification Axis の 3 レイヤーで構成される。以下では、これらのレイヤーと Temporal Axis に含まれる ResBlock (図 2 参照) の詳細について述べる。

Temporal Axis とは心電図における時間軸のことであり、波形ごとに時間軸に沿った特徴抽出を行う。ResBlock における畳み込みフィルタのサイズは 3 ブロックごとに (7, 1) → (5, 1) → (3, 1) と徐々に小さくなっており、層が深くなるにつれてデータの詳細な特徴に注目する。また、畳み込み層のチャンネル数は 3 ブロックごとに 16 → 32 → 64 と増加しており、層が深くなるにつれてより複雑な特徴を読み取ることができる。

ResBlock は残差接続を含む畳み込み層である。図 2 の左側には畳み込み層が 2 つ存在する。それらの層ではフィルタの形状が $k \times 1$ となっていることから、波形ごとに特徴を抽出する。畳み込み後にはバッチ正規化 (Batch Normalization) と活性化関数 (ReLU) による非線形変換を行う。図 2 の右側の残差接続では、左側の出力と次元数を合わせるために 1×1 の畳み込みおよび最大値プーリング (Max Pooling) を行う。ブロックの最終的な出力は左側と右側の和であり、これによって勾配消失を防いだり、過学習を抑制したりしながら層を深くすることを可能にしている。

Lead Axis では Temporal Axis で抽出された各波形の特徴から全電極の特徴を抽出してネットワークの発火を行う。Lead Axis の畳み込みフィルタの形状が 1×12 と

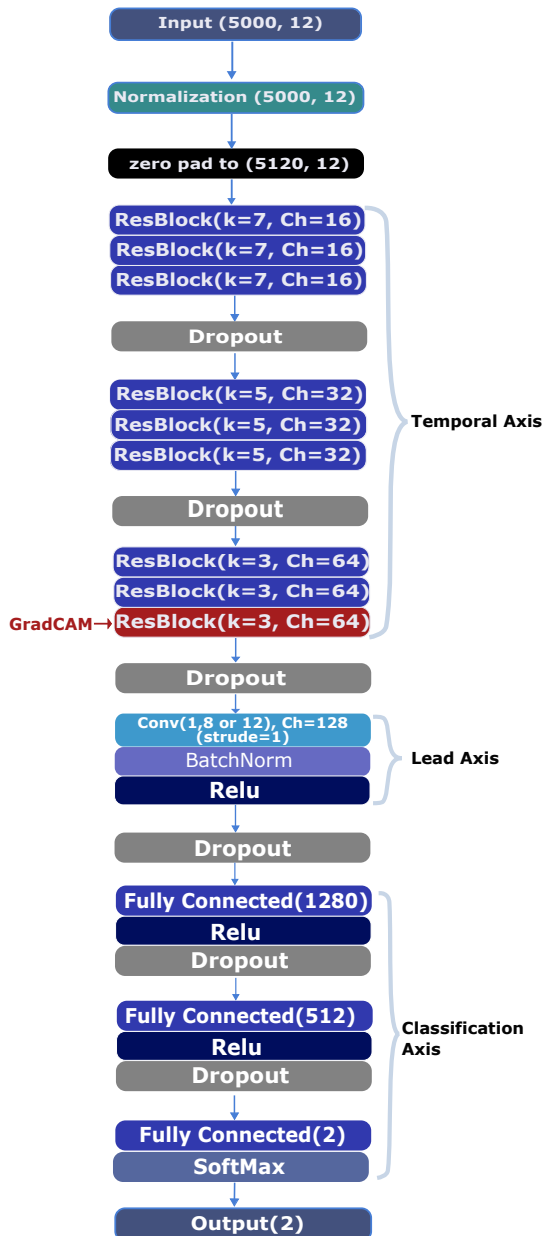


図 1 CNN モデルの全体構成

なっていることから、全電極の特徴を抽出していることが分かる。また、畳み込み後には ResBlock と同様にバッチ正規化 (Batch Normalization) と活性化関数 (ReLU) による非線形変換を行う。

Classification Axis では全結合層に接続して分類を行う。Lead Axis からの出力を平坦化 (1次元化) したときの全ノード (1,280 個) を 1,280 ノード, 512 ノード, 2 ノードの全結合層に順に接続する。最終出力層には Softmax 関数を用いており、致死性不整脈の発症予測を確率として出力する。

3 CNN の学習法とその実装

3.1 データセットの作成

CNN の学習には、岡山大学病院循環器内科で収集された、HCM 患者 179 人の致死性不整脈発症前の 12 誘導心電図データを用いる。各心電図データには、その患者が致死性不整脈を発症していればラベル 1 が、そうでな

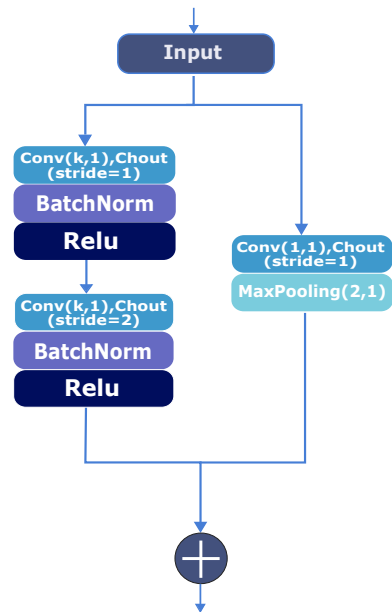


図 2 ResBlock の構成

ければラベル 0 が付与されている。これらのデータには同一患者の心電図が複数件含まれており、患者ごとのデータ件数は 1 から 150 と大きくばらついている。そのためすべてのデータを利用して CNN の学習を行うと、データ数の非常に多い患者の影響を強く受けてしまう恐れがある。そこで各患者のデータから最大 15 件 (16 件以上ある場合には最新のものから 15 件) を選んでデータセットを作成する。

3.2 交差検証のためのデータ分割

本研究では、図 3 に示すように、5 分割交差検証を用いて CNN の予測性能を評価する。まず、全データから 10% を選んでテスト用とする。次に、残りの 90% のデータを 5 分割して 4 つを訓練用, 1 つを検証用とする。検証用の選び方は 5 通りあるので、この処理によって 5 種類のデータセットが得られる。

データセットには同一患者の心電図が複数件含まれているため、ランダムに訓練用, 検証用, テスト用に分けると、訓練用とテスト用の両方に同一患者の心電図が存在する可能性がある。同一患者の心電図は特徴が類似しているため、この状況では CNN の予測性能を過大評価してしまう恐れがある。そこで層化グループ K 分割 (SGKF: Stratified Group K-Fold) 交差検証によって CNN の予測性能を評価する。SGKF は訓練用とテスト用の両方に同一患者の心電図が含まれないようにするデータ分割法である。SGKF の実装には sklearn.model_selection の StratifiedGroupKFold [11] を利用する。

3.3 CNN の学習とハイパーパラメータの探索

データ分割によって得られた 5 種類のデータセットの各々を用いて図 1 の CNN モデルのミニバッチ学習を行う。バッチサイズは 16 とし、最適化アルゴリズムには Adam [10] を用いる。また、エポック数が 500 に到達するか、検証用データに対する損失値が規定回数連続して減少しなければ学習を終了する。

CNN の学習において、学習率, ドロップアウト率, 重み減衰率などのハイパーパラメータは学習の収束性や

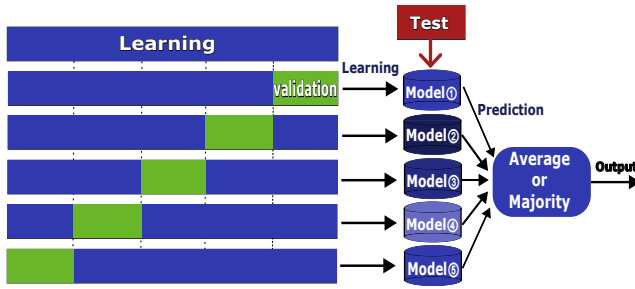


図3 学習とテストの流れ

表1 実験に用いたデータセットの内訳

発症群	非発症群	計
418 件 (102 人)	595 件 (77 人)	1,013 件 (179 人)

CNN の予測性能に大きな影響を及ぼすため、値を適切に設定する必要がある。しかし、それらの値を手動でチューニングするのは時間と労力の掛かる作業であり実用的とは言えない。そこで本研究では、ベイズ最適化 [12] を用いて探索を自動化し、検証用データに対する損失値を小さくするハイパーパラメータの値を求める。実装には skopt の `gp_minimize` [13] を利用する。

3.4 CNN の性能評価

5 種類のデータセットを用いて CNN の学習を行うことによって、パラメータ値の異なる 5 つの CNN が得られる。それらの各々にテスト用データを与えて致死性不整脈の発症を予測させ、精度、再現率、適合率、F 値によって性能を評価する。また、交差検証によっても CNN の性能を評価する。すなわち、5 個の CNN の出力の平均化や多数決によって予測を行い、精度、再現率、適合率、F 値によって性能を評価する。

3.5 CNN の判断根拠の可視化

学習済み CNN が HCM 患者の将来における致死性不整脈発症を高い確率で予測できたとしても、実際の医療現場では医師が診断を行う。CNN が心電図のどこに着目して予測しているかを可視化すれば、医師の診断を支援できる可能性がある。そこで本研究では、図 1 の Temporal Axis における最終残差ブロックに Grad-CAM++ [14] を適用して各波形の時間帯ごとの重要度を求め、さらに各波形の重要度の最大値を 12 波形で比較する。これによって、CNN がどの波形のどの時間帯に着目しているかを医師に提示できる。Grad-CAM++ の実装には `pytorch-gradcam` の `GradCAMpp` を使用する。

4 評価実験

4.1 実験概要

図 1 の CNN モデルと前述の学習法の有効性を検証するため、岡山大学病院循環器内科で収集された、HCM 患者 179 人の致死性不整脈発症前の心電図データを用いて評価実験を行う。データセットの内訳を表 1 に示す。「発症群」は、心電図測定後に致死性不整脈を発症した患者のデータを表し、「非発症群」は、心電図測定後も致死性不整脈を発症していない患者のデータを表す。また、実験におけるプログラム実行環境を表 2 に示す。

表2 プログラム実行環境

OS	Ubuntu 22.04.3 LST
CPU	Intel®Xeon Silver 4210R CPU@2.40GHz × 40
GPU	NVIDIA GeForce RTX 3090 × 4
プログラミング言語	Python 3.10.12
機械学習ライブラリ	PyTorch 2.0.1

表3 各モデルと交差検証の性能評価 (12 波形学習時)

評価対象	精度	再現率	適合率	F 値
モデル 1	0.84193	0.91665	0.76474	0.829
モデル 2	0.79547	0.95335	0.71962	0.813
モデル 3	0.83849	0.93415	0.78663	0.844
モデル 4	0.82150	0.92854	0.73495	0.817
モデル 5	0.85286	0.95585	0.76123	0.843
交差検証 (平均)	0.90662	0.91698	0.85246	0.905
交差検証 (多数決)	0.90154	0.97561	0.89854	0.899

4.2 CNN 単体と交差検証の性能評価

まず、データ分割によって得られた 5 種類のデータセットのそれぞれを図 1 の CNN モデルに学習させて 5 つの学習済みモデルを得た。次に、それらのモデルの各々にテスト用の心電図データを与えて致死性不整脈発症の有無を予測させ、精度、再現率、適合率、F 値によって CNN モデル単体の予測性能を評価した。さらに、同じテスト用データを 5 つの学習済みモデルに与え、それらの出力の平均化または多数決によって致死性不整脈発症の有無を予測させ、上と同じ指標で予測性能を評価した。結果を表 3 に示す。ただし、表 3 の数値は、データセットを 10 分割してテスト用データの選び方を 10 通りすべて試したときの平均である。

まず、精度について考察する。CNN モデル単体の場合が 80~85% で、交差検証 (平均) と交差検証 (多数決) の場合がともに約 90% であるため、交差検証によって精度が向上していることがわかる。

再現率については、CNN モデル単体の場合が 92~96% で、交差検証 (平均) の場合が約 92% であるため交差検証の効果は見られなかった。しかし、交差検証 (多数決) の場合は約 98% と高い値となり、モデル単体の場合よりも偽陰性率が大幅に減少している。

適合率については、モデル単体の場合が 72~76% で、交差検証 (平均) と交差検証 (多数決) の場合がそれぞれ約 85%、約 90% となり、交差検証によって適合率が大幅に上昇している、すなわち、偽陽性率が大幅に減少していることがわかる。

F 値に関しても同様に、モデル単体の場合が 82~84% で、交差検証の場合は平均、多数決ともに約 90% であることから、交差検証によって F 値が大幅に上昇していることがわかる。

4.3 学習に用いる波形の数による性能の違い

関連研究 [9] では、12 波形のうち III 誘導と 3 つの単極肢誘導を除いた 8 波形を学習に使用している。そこで、CNN モデルの学習に用いる波形を 12 種類から 8 種類に絞り、上と同じ性能評価実験を行った。結果を表 4 に示す。ただし、表 4 の数値は、表 3 の場合と同様に、データセットを 10 分割してテスト用データの選び方を

表 4 各モデルと交差検証の性能評価 (8 波形学習時)

評価対象	精度	再現率	適合率	F 値
モデル 1	0.83352	0.84939	0.77635	0.801
モデル 2	0.83146	0.93329	0.74432	0.825
モデル 3	0.81413	0.91439	0.74803	0.814
モデル 4	0.85405	0.91402	0.79552	0.844
モデル 5	0.84568	0.96329	0.77144	0.846
交差検証 (平均)	0.92383	0.98049	0.86530	0.917
交差検証 (多数決)	0.91776	0.98293	0.85480	0.912

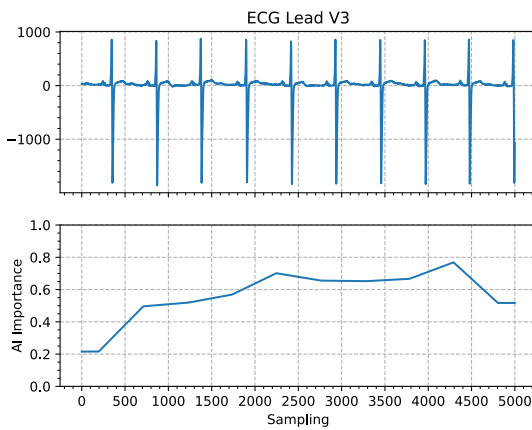


図 4 波形 V3 の時間帯ごとの重要度

10 通りすべて試したときの平均である。

表 3 と表 4 を比較すると、両者に顕著な差は見られないものの、8 波形の方が精度が 1~2 ポイント高いことが分かる。この理由について現時点では断言できないが、学習に使用した 8 波形から残りの 4 波形を導出できる [15] ため、冗長な情報が削除されて性能が向上したと推察される。

4.4 Grad-CAM++による判断根拠の可視化

ある患者の心電図における 12 波形の各々について、Grad-CAM++を用いて時間帯ごとの重要度を求めた。そのうちの 4 波形 (V3,V2,III,V1) の重要度を可視化した結果を図 4~7 に示す。各図において、上には心電図の波形が、下には時間帯ごとの重要度が描かれている。重要度は 0 から 1 までの値で与えられ、値が大きいほど重要度が高いことを意味する。重要度が高い時間帯を上部の心電図データと照らし合わせることで、CNN の判断根拠を容易に確認することができる。

12 波形の時間帯ごとの重要度の最大値を求めて可視化した結果を図 8 に示す。この図から、波形 V3,V2,III,V1 の順に重要度が高いことがわかる。最初に図 8 を見て重要な波形を抽出し、次に各波形の時間帯ごとの重要度 (図 4~7) を見ることによって、CNN の判断根拠を効率的に確認することができる。

5 おわりに

肥大型心筋症患者の心電図データを畳み込みニューラルネットワーク (CNN: Convolutional Neural Network) に学習させ、学習済み CNN によって致死性不整脈の発症予測を行った。関連研究 [9] の CNN モデルに修正を加

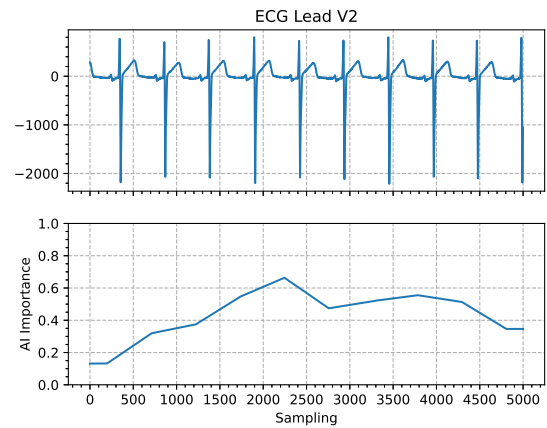


図 5 波形 V2 の時間帯ごとの重要度

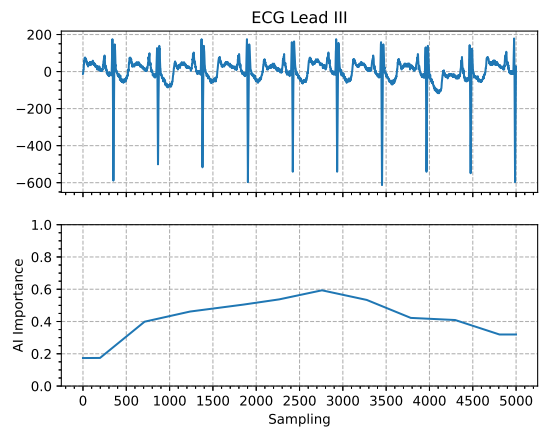


図 6 波形 III の時間帯ごとの重要度

え、交差検証等の手法を導入することにより、患者の将来の致死性不整脈の発症を約 90% の高い精度で予測できることを示した。また、説明可能な AI を目指し、Grad-CAM++によって CNN モデルの判断根拠を可視化した。どの波形のどの時間帯が重要であるかを視覚的に示すことができるので、医師の診断を効果的に支援できる可能性がある。

今後の展望として、肥大型心筋症以外の心筋症もつ患者の心電図データに対しても、本研究の CNN モデルや学習方法を適用して致死性不整脈の発症を予測できるか検証する。本研究の手法が有効であることが明らかになれば、より多くの心疾患に対して将来の致死性不整脈の発症を予測できると期待される。

謝辞

本研究は JSPS 科研費 JP23K21703 の助成を受けたものである。ここに謝意を表す。

倫理審査

本研究は岡山大学倫理審査委員会の承認を得て実施された (番号 2304-018)。

参考文献

- [1] “令和 5 年 (2023) 人口動態統計月報年計 (概数) の概況,” <https://www.mhlw.go.jp/toukei/saikin/hw/jinkou/geppo/nengai23/dl/gaikyouR5.pdf> (2024)

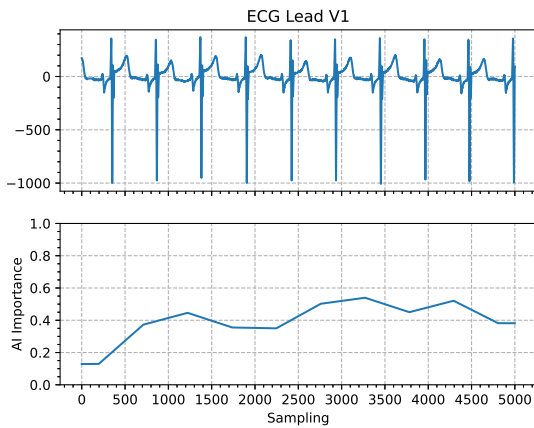


図 7 波形 V1 の時間帯ごとの重要度

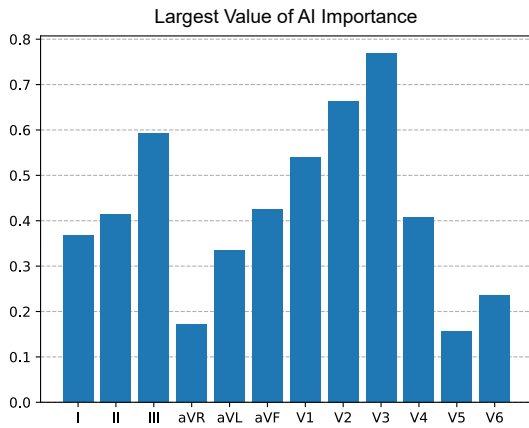


図 8 各波形の時間帯ごとの重要度の最大値

年 6 月 8 日アクセス).

- [2] “日本循環器学会／日本心不全学会合同ガイドライン：心筋症診療ガイドライン（2018 年改訂版）,” https://www.j-circ.or.jp/cms/wp-content/uploads/2018/08/JCS2018_tsutsui_kitaoka.pdf (2024 年 6 月 8 日アクセス).
- [3] 北風政史, “心筋症をどう考えるか,” 日本内科学会雑誌, vol.103, no.2, pp.273–276, 2014.
- [4] T. Kubo, T. Hirota, Y. Baba, Y. Ochi, A. Takahashi, N. Yamasaki, N. Hamashige, K. Yamamoto, F. Kondo, K. Bando, E. Yamada, T. Furuno, T. Yabe, L. Y. Doi, and H. Kitaoka, “Patients’ characteristics and clinical course of hypertrophic cardiomyopathy in a regional Japanese cohort—results from Kochi RYOMA study—,” *Circulation Journal*, vol.82, no.3, pp.824–830, 2018.
- [5] B. J. Gersh, B. J. Maron, R. O. Bonow, J. A. Dearani, M. A. Fifer, M. S. Link, S. S. Naidu, R. A. Nishimura, S. R. Ommen, H. Rakowski, C. E. Seidman, J. A. Towbin, J. E. Udelson, and C. W. Yancy, “2011 ACCF/AHA guideline for the diagnosis and treatment of hypertrophic cardiomyopathy: A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines,” *Circulation*, vol.124, no.24, pp.e783–e831, 2011.
- [6] P. M. Elliott, A. Anastakis, M. A. Borger, M. Borggrefe, F. Cecchi, P. Charron, A. A. Hagege, A. Lafont, G. Limongelli, H. Mahrholdt, W. J. McKenna, J. Mogensen, P. Nihoyannopoulos, S. Nistri, P. G. Pieper, B. Pieske, C. Rapezzi, F. H. Rutten, C. Tillmanns, H. Watkins, and C. O’ Mahony, “2014 ESC guide-

lines on diagnosis and management of hypertrophic cardiomyopathy,” *Russian Journal of Cardiology*, vol.5, pp.7–57, 2015.

- [7] C. O’Mahony, M. Tome-Esteban, P. D. Lambiase, A. Pantazis, S. Dickie, W. J. McKenna, and P. M. Elliott, “A validation study of the 2003 American College of Cardiology/European Society of Cardiology and 2011 American College of Cardiology Foundation/American Heart Association risk stratification and treatment algorithms for sudden cardiac death in patients with hypertrophic cardiomyopathy,” *Heart*, vol.99, no.8, pp.534–541, 2013.
- [8] B. J. Maron, S. A. Casey, R. H. Chan, R. F. Garberich, E. J. Rowin, and M. S. Maron, “Independent assessment of the European Society of Cardiology sudden death risk model for hypertrophic cardiomyopathy,” *The American Journal of Cardiology*, vol.116, no.5, pp.757–764, 2015.
- [9] Z. Attia, P. A. Noseworthy, F. Lopez-Jimenez, S. J. Asivatham, A. J. Deshmukh, B. J. Gersh, R. E. Carter, X. Yao, A. A. Rabinstein, B. J. Erickson, S. Kapa, and P. A. Friedman, “An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction,” *Lancet*, vol.394, no.10201, pp.861–867, 2019.
- [10] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” arXiv preprint arXiv:1412.6980, pp.1–15, 2014.
- [11] “StratifiedGroupKFold,” https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.StratifiedGroupKFold.html (2024 年 6 月 8 日アクセス).
- [12] J. Snoek, H. Larochelle and R. P. Adams, “Practical Bayesian optimization of machine learning algorithms,” *Advances in Neural Information Processing Systems*, eds. by F. Pereira, C. J. Burges, L. Bottou, and K. Q. Weinberger, vol.25, pp.1–9, 2012.
- [13] G. Louppe and M. Kumar, “Bayesian optimization with skopt,” https://scikit-optimize.github.io/stable/auto_examples/bayesian-optimization.html (2024 年 6 月 8 日アクセス).
- [14] A. Chattopadhyay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, “Grad-CAM++: Generalized gradient-based visual explanations for deep convolutional networks,” *Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision*, pp.839–847, 2018.
- [15] J. Malmivuo and R. Plonsey, *Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields*, pp.277–289, Oxford University Press, 1975.