

# 深層学習を用いた救急車からの心電図における異常波形の分類

及川 遼<sup>†</sup> 土井 章男<sup>†</sup> チャクラボルティ バサビ<sup>†</sup> 伊藤 智範<sup>‡</sup>

岩手県立大学<sup>†</sup> 岩手医科大学<sup>‡</sup>

## 1. はじめに

近年、救急搬送時間の長い地域にて、12誘導心電図伝送システムが導入されている。12誘導心電図伝送システムとは、救急車両内にて、内部心筋梗塞が疑われる患者や事故等による外傷患者に対して、救急救命士が心電図の計測を行い、データを画像情報として救急病院等に伝送するものである。伝送された心電図を病院にいる医師が診断することで救急車の移動時間を有効活用し、早期の治療開始が可能となる。特に心筋梗塞などの虚血性心疾患は心不全やショックを合併する可能性が高くなるため早期の発見が必要であり、そのような症状に対して極めて有用なシステムである。伝送された心電図の解析を行う際には、異常波形の様々なパターンを把握しておくことが必要であり、循環器内科の医師でない限り短時間で正確な診断はできない。しかし、救急病院や医師不足の地域では循環器内科でない医師が診断を行う状況にあり、循環器内科医師と同等以上の性能をもつ自動診断システムが求められている。

本稿では第一段階として、救急車両内で計測した12誘導心電図を深層学習を用いて解析し、重症度の分類と評価を行った。

## 2. 関連研究

近年の心電図自動診断に関する研究では、特定の異常所見に対する診断が多く見られ、特に心筋梗塞や不整脈の分類・判別が比較的盛んに行われている [1, 2]。一方で、異常所見の種類によらない重症度の分類に関する研究は十分に行われていない [3]。

また、救急車両内で計測された心電図は、心筋梗塞等の兆候を明確に示している、患者の体動や電極の外れによるノイズが混入しやすい環境で計測されている、といった通常的心電図とは異なる特徴を有しているため、専用の分類器を作成することで既存の分類器よりも高い精度での判別が可能になると考えられる。

## 3. モデルの作成

### 3.1. データセット

学習に用いるデータセットは以下の流れで作成された。(1) 岩手県立二戸病院が保有している、救急車両内で計測した12誘導心電図の画像データ 88 症例に対し

て、12誘導分の波形それぞれを3分割し、1症例につき36枚の画像を切り出した。(2) 各画像に対して岩手医科大学の循環器専門医がラベル付けを行った。ラベル付けは正常、中等症以下、重症の3クラスで行った。(3) 各画像に対して細線化および二値化を行った。

最終的なデータ 3168枚のうち正常波形は1590枚、異常波形は1578枚であり、次節ではこれらの80%を学習データ、20%をテストデータとする。また、学習データの20%を検証データとして5分割交差検証を行う。

### 3.2. 学習

本研究では事前学習済みの EfficientnetB0 ネットワークを用いて分類器の作成を行う。入力画像のサイズは  $224 \times 224$  ピクセルと設定し、一致しない場合はリサイズを行う。また、学習データに対して  $\pm 30$  ピクセルの水平移動および垂直移動を行い、エポックごとにわずかに異なるデータセットを使用する。

最適化手法は Adam を使用し、ハイパーパラメータは  $\alpha = 0.0001$ ,  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$  とする。

また、ミニバッチサイズを 64, epoch を 100 とする。最良の分割時の学習曲線を図 1 に示す。

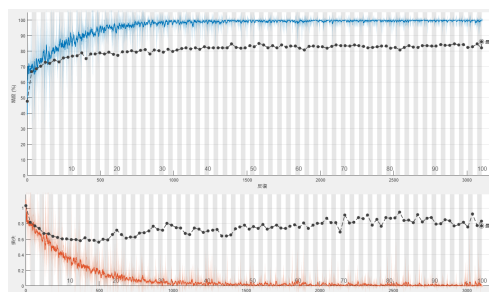


図 1: 学習曲線

## 4. 評価と考察

外部検証を行った結果得られた、各分割における混同行列を図 2 に示す。また、Accuracy および Kappa 係数を表 1 に示す。

		Confusion Matrix			
		0	1	2	3
Output Class	0	382 80.0%	24 5.0%	18 3.8%	91 19.2%
	1	25 4.1%	66 13.5%	3 0.6%	74 15.2%
	2	11 1.7%	5 0.9%	80 16.8%	81 16.7%
	3	91 18.0%	74 15.2%	70 14.3%	89 18.3%
		0	1	2	3

図 2: 混同行列

Classification of Abnormal Waveforms in Electrocardiograms from Ambulances Using Deep Learning

Ryo OIKAWA<sup>†</sup>, Akio DOI<sup>†</sup>, Basabi CHAKRABORTY<sup>†</sup>, Tomonori ITOH<sup>‡</sup>

<sup>†</sup>Iwate Prefectural University, <sup>‡</sup>Iwate Medical University



図 2: 混同行列 (続き)

表 1: Accuracy および Kappa 係数

Fold	Accuracy(%)	Kappa 係数
1	86.3	0.810
2	86.5	0.814
3	83.6	0.816
4	86.6	0.801
5	81.9	0.810
平均	85.0	0.810

なお、ここにおいて Accuracy とは正答率を意味し、分類結果と正解ラベルが一致している割合を表す。

Kappa 係数とは、ある現象を 2 人の観察者が観察した場合の結果がどの程度一致しているかを表す統計量である。結果として、最良の分割時、外部検証における Kappa 係数は 0.816 であり、循環器専門医との比較が、Landis らの目安 [4] において”ほぼ完璧な一致”と判断される 0.81-1.00 の範囲であった。このことから、本研究で作成したモデルは有効性があると考えられる。最悪の分割時では、外部検証における Kappa 係数は 0.801 であり、”かなりの一致”と判断される 0.61-0.80 の範囲であった。

一方で、一部のデータにおいて、周知のいずれのクラスとも類似していないと判断され、消去法で分類されるケースが散見された。波形の例とクラス活性化マッピングの結果を図 3 に示す。なお、クラス活性化マッピングとは、AI がどのような特徴を基に予測を行っているかを可視化する手法である。図 3.(a) のような凹凸の少ない滑らかな形状の波形において偽陽性となる事例については、心電図波形の二値分類を行う先行研究でも報告がされており [3]、当該形状の正常波形をデータセットに多く含めるなどを行う必要があると考えら

れる。図 3.(b) については学習の際のデータ拡張において垂直移動を行ったことで高さの特徴が失われ波形の ST 上昇や T 波増高といった所見の検出が難しくなってしまった可能性が考えられる。クラス活性化マッピングの結果からも、図 3.(b) の波形において所見の見られる部位に AI が十分に着目できていないことが確認できる。また、心筋梗塞等では経時的な波形の変化がみられ、様々な形状が存在するため、学習データの不足が分類失敗の原因として考えられる。結論として、これらの問題は学習データの増加およびデータ拡張方法の見直しで改善されるものと思われる。

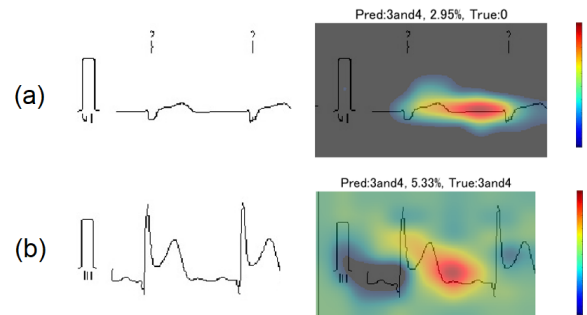


図 3: 消去法で分類されている波形の例

## 5. おわりに

本研究では救急車両内で計測した 12 誘導心電図を深層学習を用いて解析し、重症度の分類と評価を行った。その結果、外部検証にて平均 85.0% の分類精度および 0.810 の Kappa 係数を得た。今後の展望として学習データの増加およびデータ拡張方法の見直しを行った上でモデルの再学習を行うこと、および第二段階として症例単位での重症度分類を行うことを検討している。

## 参考文献

- 1) U. Rajendra Acharya, Hamido Fujita, Shu Lih Oh, Yuki Hagiwara, Jen Hong Tan, Muhammad Adam, Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals, Information Sciences, Vol.415-416, pp.190-198, 2017.
- 2) G. Sannino, G. De Pietro, A deep learning approach for ECG-based heartbeat classification for arrhythmia detection, Future Generation Computer Systems, Vol.86, pp.446-455, 2018.
- 3) S. Furubayashi, T. Imai, S. Ishihara, K. Fujiu, K. Ohe, A study on normal abnormal determination of electrocardiogram waveform using deep learning, JSAI Technical Report, Type 2 SIG, pp.05-01-05-05, 2018.
- 4) J. Richard Landis and Gary G. Koch, The Measurement of Observer Agreement for Categorical Data, Biometrics, Vol.33, No. 1, pp.159-174, 1977.