

Stacked convolutional denoising autoencoders を用いた2誘導心電図からの特徴抽出および不整脈分類

高橋 柊^{1,a)} 落合 桂一¹ 深澤 佑介¹

概要：近年,さまざまなモノおよびコトをインターネットに接続することで,ログ収集および相互制御を行う Internet of Things(IoT) に注目が集まっている。IoT の活用により,今までセンシングが困難であった情報がリアルタイムに取得可能となることが期待されている。心電図(ECG) をリアルタイムに解析することができれば, IoT デバイスの活用によりリアルタイムに不整脈などを検出することが可能となる。本研究では, Stacked Convolutional Denoising Autoencoders(SCDAE) を用いた, ECG 波形からの高レベルな特徴抽出について検討する。また, 事前学習した SCDAE の構造および重みを抽出し, 全結合層を追加した分類器を再学習する不整脈分類手法を提案する。未知の ECG 波形からの不整脈分類において, 提案手法が既存手法 (Accuracy:92.7%) に対し高精度 (Accuracy:95.1%) であることを示す。

Feature extraction and arrhythmia classification from 2-lead ECG using stacked convolutional denoising autoencoders

SHU TAKAHASHI^{1,a)} KEIICHI OCHIAI¹ YUSUKE FUKAZAWA¹

1. はじめに

近年,さまざまなモノおよびコトをインターネットに接続することで,ログ収集および相互制御を行う Internet of Things(以下, IoT) に注目が集まっている [1]. IoT の活用により, 今までセンシングが困難であった情報がリアルタイムに取得可能となることが期待されている。高河原ら [2] は服型ウェアラブルデバイス「hitoe」を用いることで, 装着者の心電波形を検出する手法を提案している。服型デバイスおよび IoT の活用により, リアルタイムに心電波形を検出することが可能となる。

本研究は, 心電図 (ECG; Electrocardiogramis) からの不整脈自動検出を目的とする。ECG をリアルタイムに分析することができれば, IoT の活用によりリアルタイムなストレス度合い, 不整脈検出が可能となる [3]. ECG からの不整脈検出には以下のような問題が存在する。

観測対象による波形変化

ECG 波形は観測対象者や観測時間によって変化する。そのため, 観測対象に対しロバストな不整脈の検出を行うことは困難である。

心拍変動による波形変化

ECG 波形は観測対象者のストレス状態や興奮, 運動などの生理学および精神的変化によって変化する。そのため, 測定状態に対しロバストな不整脈の検出を行うことは困難である。

観測ノイズ

ECG 波形には観測ノイズが含まれている。そのため, 分析前には適切なノイズ除去が必要である。

観測対象および心拍変動による波形変動や観測ノイズに対しロバストな不整脈検出が可能となれば, 人や観測状態を選ばずリアルタイムに不整脈検出が可能となる。本研究では, Stacked Convolutional Denoising Autoencoders(以下, SCDAE) を用いた ECG からの不整脈分類手法を提案する。観測対象および心拍変動による波形変動に対して, SCDAE の畳み込み層が高レベルの特徴抽出を行い, プー

¹ NTT ドコモ
NTT DOCOMO, INC., Yokosuka, Kanagawa 239-8536,
Japan

^{a)} syuu.takahashi.us@nttdocomo.com

リング層が特徴の位置普遍性を獲得することが期待できる。また、SCDAEを事前学習(pre-training)をしたのち、SCDAEのエンコーダ部分に全結合のニューラルネットワークを追加し再学習(fine-tuning)することで、観測ノイズに対し頑強な分類モデルの構築が期待できる。

本研究の貢献は以下の通りである。

- ECGからの不整脈分類において問題となる心拍変化および観測ノイズ問題に対し、SCDAEを用いた高レベルの特徴抽出について検討する。
- 事前学習されたSCDAEに全結合層を追加することで、未知の波形について高精度な分類能力をもつ分類器を提案する。

本稿では、次章にて関連研究について述べ本研究との差分について説明する。次に、3章にて提案手法の詳細を説明する。4章では実データを用いた実験を行い、既存手法と提案手法の精度について検証する。最後に5章にて本研究のまとめおよび今後の課題について述べる。

2. 関連研究

ECGからの不整脈検出については、多くの手法が提案されている。図1に2誘導ECGの概要を示す。ECGは心臓の電気的な活動を記録している。2誘導ECGでは、一般にMLIIおよびV1の2種類の波形が観測され、各波形の相互変化を観察することで、不整脈の診断が可能となる。ECG波形は、心拍間隔(以下、RRI)を基準に分割され、window毎に不整脈であるか否かが決定される。そのため、ECG分類には各windowから有効な特徴を抽出することが求められる。

ECG波形からの特徴抽出にはウェーブレット変換をもちいた研究がある。Srivastavaら[4]は離散ウェーブレット変換(DWT; Discrete wavelet transform)を用い、ECG波形

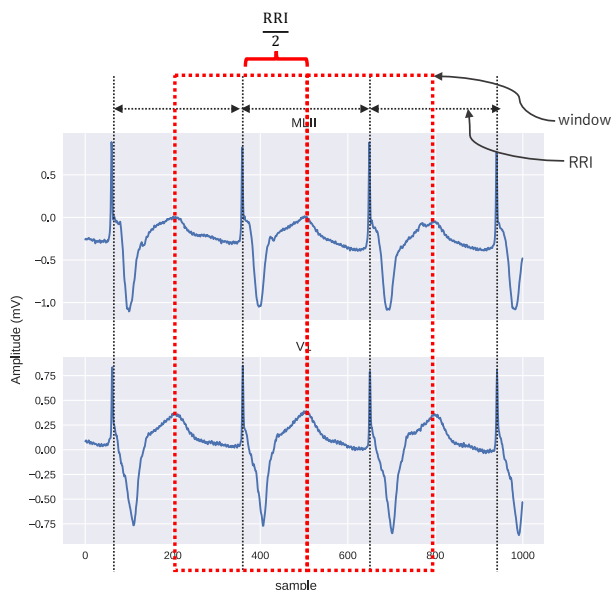


図1 ECGの概要

の周波数特性を抽出したのち、fuzzy hybrid neural networkにより波形を分類する手法を提案している。DWTを用いたスペクトル解析により圧縮された特徴が、ECG波形分類に有効であることを示している。Tangら[5]はECG波形よりピーク点(P, Q, R, SおよびT点)を検出し、各ピーク点間の距離およびRSにより次元削減した特徴を用い quantum neural networks(QNN)により波形を分類する手法について提案している。しかし、これらの手法はウェーブレット変換により抽出された特徴を用いるため、観測対象による波形変化が存在するというECGの特性上、未知の観測対象から得られたECGデータを分類することは困難である。

Osowski[6]らは観測対象によって変化するECG波形からラベル毎に共通した特徴を抽出するため、higher order statistics(HOS)を用いた特徴抽出について検討している。HOSを用いることで観測対象が異なるECG波形についても、ラベルごとに共通した特徴が抽出可能であることを示している。Zubair[7]らは、Convolutional Neural Networks(以下、CNN)を用いることで、未知の観測対象におけるECGからの不整脈検出について検討している。CNNを用いることで、新たに観測されたECG波形について高精度に不整脈検出が可能であることを示している。

本研究ではSCDAEを用いたECGからの特徴抽出および不整脈分類について検討する。既存研究においても、CNNを用いたECG波形の分類手法が提案されているが、本研究ではCNNにDenoising Autoencoderを組み合わせたSCDAEを用いることで、未知のデータやノイズに対しより頑強な特徴を抽出する。また、学習された重みを用いた分類器を構築することで不整脈分類を行う。関連研究と比較し、未知のデータに対する不整脈分類精度が高精度であることを示す。

3. 提案手法

本研究では、SCDAEを用いたECGからの特徴抽出および不整脈分類について検討する。提案手法は以下の2つのステップから構成される。

i) SCDAEによる特徴抽出(pre-training)

各波形からノイズを除去するSCDAEの事前学習(pre-training)を行う。SCDAEにより、未知のデータおよびノイズに対して頑強な特徴抽出が期待できる。

ii) 分類器の構築(fine-tuning)

SCDAEの学習により得られたモデルより、エンコーダ部分を抽出し全結合層を追加した分類器を構築する。SCDAEの事前学習により得られた重みを利用し分類器のfine-tuningを行うことで、高精度な分類器の構築が期待できる。

3.1 SCDAE による特徴抽出 (pre-training)

Autoencoder(以下, AE) はニューラルネットワークを使用した次元圧縮アルゴリズムである [8]. 入力データと出力データが同じようになるようニューラルネットワークを学習させることで, 多量の特徴量から入力データの説明に有効な圧縮された特徴量の抽出が期待できる. 入力 $x \in \mathbb{R}$ に対し, エンコーダ $y = f_{\theta}(x)$ とデコーダ $g_{\theta'}(y)$, 学習すべき重み θ, θ' および損失関数を L とした時, AE は式 1 の最適化問題として定義される.

$$\min_{\theta, \theta'} \frac{1}{2} \sum_{i=1}^n L(x_i, g(f(x_i))) \quad (1)$$

ただし, 活性化関数を σ , 学習により求められる重みを W , バイアスを b , $\theta = \{W, b\}$ としたとき, $f_{\theta}(x)$ は式 2, $g_{\theta'}(y)$ は式 3 として定義される.

$$f_{\theta}(x) = \sigma(W_x + b) \quad (2)$$

$$g_{\theta'}(y) = \sigma(W'_x + b') \quad (3)$$

Vincent ら [9] は AE の入力ベクトルにノイズを加えることで, より AE と比較しより汎化された特徴の抽出を行う Denoising Autoencoder(以下, DAE) について提案している. 図 2 に DAE の例を示す. DAE では入力 $X = x$ についての破壊プロセス (corruption process) $C(\bar{X}|X)$ から \bar{X} をサンプリングすることで, 入力ベクトル x を \bar{x} に破損させる. 破損させた \bar{x} を入力とし, AE と同様に x を復元するようなニューラルネットワークを学習させることで, AE と比較しより有効な特徴量の抽出を試みる. 入力データ X を生成する分布を $P(X)$ とした時, DAE は式 4 の期待値最適化問題として定義される.

$$\min_{\theta, \theta'} \mathbb{E}_{X \sim P(X), \bar{X} \sim C(\bar{X}|X)} L(X, g(f(\bar{X}))) \quad (4)$$

近年, 画像認識, 言語処理など様々な分野において Convolutional Neural Network(以下, CNN) が高い認識精度を

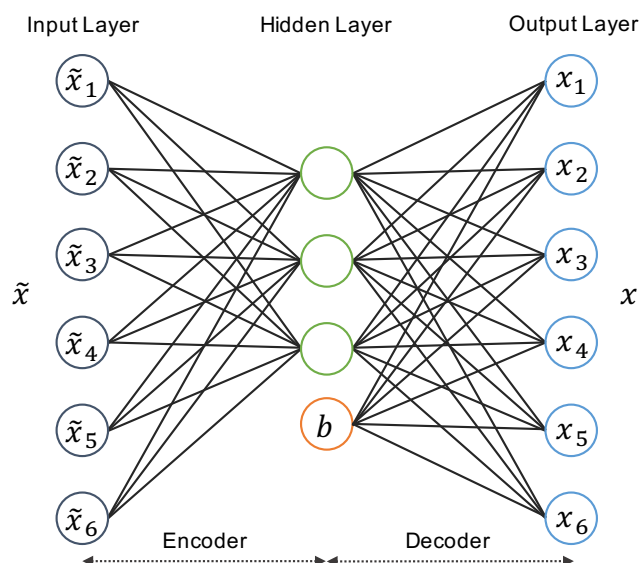


図 2 DAE の例

示している [10][11][12]. CNN では, 畳み込み層 (Convolutional layer) およびプーリング層 (Pooling layer) から構成されるニューラルネットワークである. 畳み込み層を用いることで入力データに出現するパターンを認識し, プーリング層により入力データにおけるパターン出現位置について頑健なモデルを生成することが可能となる [13]. Masci ら [14] は AE のエンコーダ層および中間層に CNN を利用した Convolutional Auto Encoder(以下, CAE) を提案している. CAE より学習された重みを用いた分類器が CNN と比較しよりよい分類精度であることを示している. Du ら [15] は CAE を多層化し, 入力データにノイズを付加した SCDAE を提案している. 既存の特徴抽出手法と比較し, SCDAE がより汎化した特徴を抽出することを示している.

本研究では, DAE のエンコーダ層とデコーダ層に畳み込み層およびプーリング層を持つ SCDAE により ECG から特徴抽出を行う. ECG 波形からの不整脈検出には, 観測対象および心拍変動による波形変化と観測ノイズ問題が存在する. 畳み込み層により高レベルの特徴認識を行い, プーリング層により特徴の位置不変性を獲得することができれば, 観測対象および心拍変動に対し頑強な特徴を抽出可能であると考えられる. また, 入力データにノイズを加えることで, 観測ノイズに対して頑強な特徴抽出が期待できる.

SCDAE は式 4 の期待値最適化問題と共通した性質を持つが, エンコーダ f_{θ} およびデコーダ $g_{\theta'}$ が全結合のニューラルネットワークと異なる. SCDAE におけるエンコーダを $convf_{\theta}$, デコーダを $convg_{\theta'}$ とすると, $k \in H$ 番目のフィルタ $convf_{\theta}(\bar{x})^k$ は式 5, $convg_{\theta'}(y)$ は式 6 により定義される. ただし, $*$ は 2 次元の畳み込み処理とする.

$$convf_{\theta}(\bar{x})^k = \sigma(x * W^k + b^k) \quad (5)$$

$$convg_{\theta'}(y) = \sigma\left(\sum_{k \in H} convf_{\theta}(\bar{x})^k * W'^k + b'\right) \quad (6)$$

SCDAE のエンコーダ層における畳み込み処理の前後にはプーリング層を挿入する. Pooling 層はエンコーダでは MaxPooling を行い, デコーダ層では, Upsampling を行う. MaxPooling は畳み込み層で計算されたフィルタの非最大値以外を削除し, Upsampling では入力の値と同様の値を任意の回数繰り返す.

SCDAE の損失関数 L_{SCDAE} には二乗誤差 (MSE) を用いる. 入力をノイズを加えた \bar{x} , 出力を x とした時, 損失関数 L_{SCDAE} は式 7 で定義される.

$$L_{SCDAE}(\theta) = \frac{1}{2n} \sum_{i=1}^n (x_i - \bar{x}_i)^2 \quad (7)$$

3.2 分類器の構築 (fine-tuning)

SCDAE は分類機能をもたない. 不整脈の検出を行うた

めに, 3.1 節で学習した SCDAE のエンコーダに全結合のニューラルネットワークを追加する. 図 3 に分類器構築の概念図を示す. 事前に学習した SCDAE の重みおよびエンコーダの構造を抽出し, 出力に全結合層 (Fully connected layer) を追加する. 全結合層の出力数を分類したい不整脈のラベルと揃えることで分類器を構築し, 入力 \bar{x} に対し, 出力を与えられたラベルとして分類器を学習する.

分類器における出力層の活性化関数には, softmax 関数を用いる. softmax 関数を用いることで, 各ユニットの出力を各ラベルへの所属確率として扱うことが可能となる. ここで, 出力層におけるユニット数を N , 入力を x , ユニット i の出力を x_i とした時, ユニット i の出力 $p(i)$ は式 8 で定義される.

$$p(i) = \frac{e^{x_i}}{\sum_{j=1}^N e^{x_j}} \quad (8)$$

分類器の損失関数 L_{CLF} には cross-entropy を用いる. サンプル数を n , 分類クラス数を m , サンプル i におけるクラス j の分類器出力を p_{ij} としたとき, 損失関数 L_{CLF} は式 9 で定義される.

$$L_{CLF}(\theta) = -\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m y_{ij} \log(p_{ij}) \quad (9)$$

4. 実験

4.1 実験条件

実験データには, 既存研究でひろく使われている 2 誘導 ECG データである MIT-BIH*1 を用いた. MIT-BIH には 48 人の 30 分間における ECG データが存在する. 各レコードには RRI およびラベルが存在し, ラベルは AAMI*2 の推奨する 5 分類に再分類可能となっている. 本研究では, 先行研究と同様の条件で実験するため, MIT-BIH から 4 レコード (102, 104, 107, 217) を除外し, 44 レコードに含まれる 100,389 の window を対象とする.

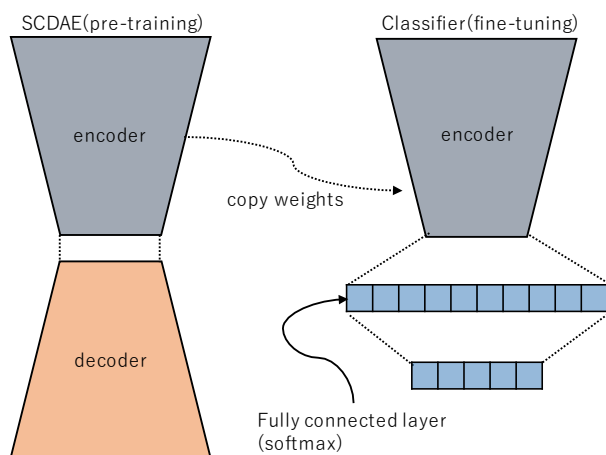


図 3 分類器の構築

*1 <https://physionet.org/physiobank/database/mitdb/>

*2 <http://www.aami.org>

ECG 波形は window 毎に RRI の差異が存在する. また, 観測対象者毎に振幅に際があるため, 正規化をする. window 毎に RRI が統一されたデータを生成するため, フーリエ変換を用いデータを固定レートにリサンプリングする. 次に, window 毎に異なる振幅を 0 - 1 間隔に正規化する. 本実験では, 各 window を 160 サンプルの固定レートにリサンプリングしたのち, 正規化を行った.

MIT-BIH のうち, 最初の 20 人 (100 - 124) の全区間におけるデータおよび, 24 人 (200 - 232) の最初 5 分間のデータをトレーニングデータとして利用する. トレーニングデータから除かれた 24 人 (200 - 232) 後半 25 分のデータをテストデータとして用いる. また, テストデータのうちランダムにサンプリングした 5000 件をバリデーションデータとして利用する. 表 1 に AAMI の推奨する不整脈の 5 分類と, ラベル数を示す.

4.2 実験結果

表 2 に実験で利用した SCDAE の構造を示す. SCDAE では, 3 層の畳み込み層およびプーリング層を利用した. また, 活性化関数には Relu[16] を用い, 最適化アルゴリズムには nadam[17] を利用した. また, 破壊分布として平均 0, 標準偏差 0.01 のガウス分布を用いた.

図 4 に SCDAE の学習曲線を示す. テストデータにおける損失 (loss) およびバリデーションデータに対する loss (val-loss) が学習回数ごとに少なくなっていることがわかる. このことから, SCDAE がトレーニングデータからの特徴抽

表 1 不整脈の 5 分類

Label	Description	Count	
		Train	Test
N	Normal beat	48,288	37,629
S	Supraventricular ectopic beat	442	2,097
V	Ventricular ectopic beat	2,205	4,327
F	Fusion beat	192	539
Q	Unknown beat	7	7

表 2 SCDAE の構造

	Layer	Type	Shape	Activation
Encoder	0	Noise	(160, 2)	
	1	Conv2D	(160, 16)	Relu
	2	MaxPooling	(80, 16)	
	3	Conv2D	(80, 8)	Relu
	4	MaxPooling	(40, 8)	
	5	Conv2D	(40, 4)	Relu
Decoder	6	MaxPooling	(20, 4)	
	7	UpSampling	(40, 4)	
	8	Conv2D	(40, 8)	Relu
	9	UpSampling	(80, 8)	
	10	Conv2D	(80, 16)	Relu
	11	UpSampling	(160, 16)	
	12	Conv2D	(160, 2)	Relu

出だけでなく、未知のデータに対しても同様に特徴を抽出可能であることがわかる。図5にテストデータの一部をSCDAEにより復元した例を示す。復元されたデータが入力波形のノイズを取り除いているとともに、特徴については引き続き保持していることがわかる。

学習したSCDAEよりエンコーダ部分の構造および学習された重みを抽出し、全結合層を追加することで分類器を構築する。表3に構築された分類器の構造を示す。エンコーダ層に32次元および5次元の全結合層を追加した。また、損失関数にはcross-entropy, 最適化アルゴリズムにはnadamを利用した。

図6に分類器の学習曲線, 図7に学習回数毎のAccuracyを示す。テストデータにおける損失(loss)およびバリデー

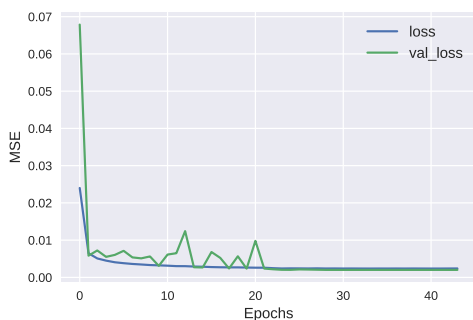


図4 SCDAEの学習曲線

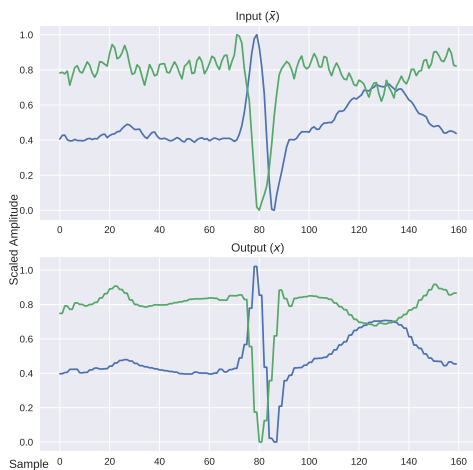


図5 SCDAEの復元例

表3 分類器の構造

Layer	Type	Shape	Activation	
0	Noise	(160, 2)		
1	Conv2D	(160, 16)	Relu	
2	MaxPooling	(80, 16)		
Encoder	3	Conv2D	(80, 8)	Relu
	4	MaxPooling	(40, 8)	
	5	Conv2D	(40, 4)	Relu
	6	MaxPooling	(20, 4)	
Classifier	7	Dence	80	Relu
	8	Dence	5	SoftMax

ションデータに対するloss(val-loss)が学習回数ごとに少なくなっていることがわかる。Accuracyについても学習を重ねる毎に、トレーニングデータおよびテストデータにて増加していることから、分類器が高精度であるとともに汎化能力を持っていることがわかる。

学習した分類器でテストデータを分類した際の分類精度を表4, 分類結果を表5に示す。十分な観測データが存在するN, VおよびFについてはprecision/recallともに高い精度で分類できていることがわかる。ラベルSについては, precisionが0.98と高い一方で, recallが0.51と低い。これは, 表1に示したようにラベルSのトレーニングデータが少ないためであると考えられる。ラベルQ(Unknown beat)については分類ができていない。これは, ラベルQのデータ数が極端に少ないため, 分類器が十分な学習を行えないためであると考えられる。また, Unknown beatは波形にラベルをつけることが困難なデータであるため, ラベリングが適切でない可能性がある。

提案手法と既存研究のAccuracyの比較を表6に示す。提案手法における未観測データの分類精度(Acuraccy)が, 既存研究と比較し高いことがわかる。このことから, 提案手法は既存研究と比較し, 未観測データからの不整脈検出

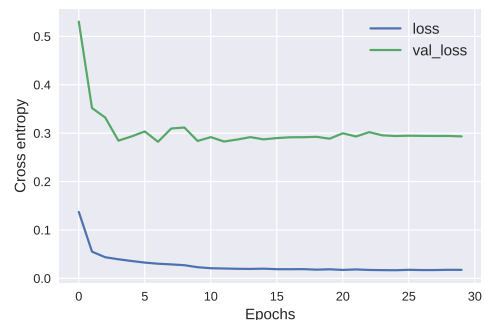


図6 分類器の学習曲線

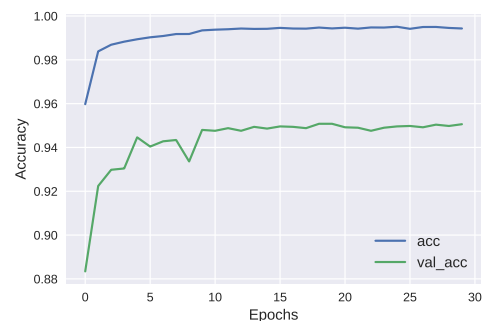


図7 分類器のAccuracy

表4 テストデータの分類精度

Label	Precision	Recall	F1-score	Support
N	0.96	0.99	0.98	37,629
S	0.98	0.51	0.67	2,097
V	0.86	0.82	0.84	4,327
F	0.88	0.73	0.80	539
Q	0.00	0.00	0.00	7

表 5 テストデータの分類結果

Label	Classification result				
	N	S	V	F	Q
N	37,387	18	201	23	0
S	732	1,063	302	0	0
V	756	3	3,538	30	0
F	85	0	63	391	0
Q	6	0	1	0	0

表 6 既存研究との精度比較

	Osowski '04[6]	Srivastava '13[4]	Tang '14[5]	Zubair '16[7]	Proposed
Acc	86%	85%	91.7%	92.7%	95.1%

に有効であるといえる。

5. まとめ

本研究では、SCDAE を用いた ECG からの不整脈検出手法を提案した。ECG からの不整脈検出で問題となる観測対象および心拍変動による波形変化については、CAE による高レベルかつ位置普遍性をもつ特徴抽出に着目し、観測ノイズについては DAE によるノイズ除去に着目することで、SCDAE による特徴抽出が有効であることを示した。また、事前学習した SCDAE の構造および重みを利用した分類器を構築し再学習することで、既存研究と比較し未知の波形分類が高精度で行えることを示した。

今後の課題として 2 誘導 ECG だけではなく、臨床分野で活用されている 12 誘導 ECG について SCDAE を適応することが挙げられる。12 誘導 ECG は、2 誘導 ECG と比較し多量かつ空間的な心電情報を観測することが可能である一方、情報が多量であるため目視での検査には限界がある。そのため提案手法を適応することで、効率的に不整脈の分類が可能となることが期待できる。

参考文献

[1] Gubbi, J., Buyya, R., Marusic, S. and Palaniswami, M.: Internet of Things (IoT): A vision, architectural elements, and future directions, *Future generation computer systems*, Vol. 29, No. 7, pp. 1645–1660 (2013).

[2] 高河原和彦, 小野一善: 心拍計測ウェアの技術と応用例, *電気学会誌*, Vol. 136, No. 3, pp. 139–142 (2016).

[3] Zhao, C., Zhao, M., Liu, J. and Zheng, C.: Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator, *Accident Analysis & Prevention*, Vol. 45, pp. 83–90 (2012).

[4] Srivastava, V. and Prasad, D.: Dwt-Based Feature Extraction from ecg Signal, *American J. of Eng. Research (AJER)*, Vol. 2, No. 3, pp. 44–50 (2013).

[5] Tang, X. and Shu, L.: Classification of electrocardiogram signals with RS and quantum neural networks, *International Journal of Multimedia and Ubiquitous Engineering*, Vol. 9, No. 2, pp. 363–372 (2014).

[6] Osowski, S., Hoai, L. T. and Markiewicz, T.: Support vector machine-based expert system for reliable heart-beat recognition, *IEEE transactions on biomedical en-*

gineering, Vol. 51, No. 4, pp. 582–589 (2004).

[7] Zubair, M., Kim, J. and Yoon, C.: An Automated ECG Beat Classification System Using Convolutional Neural Networks, *IT Convergence and Security (ICITCS), 2016 6th International Conference on*, IEEE, pp. 1–5 (2016).

[8] Hinton, G. E. and Salakhutdinov, R. R.: Reducing the dimensionality of data with neural networks, *science*, Vol. 313, No. 5786, pp. 504–507 (2006).

[9] Vincent, P., Larochelle, H., Bengio, Y. and Manzagol, P.-A.: Extracting and composing robust features with denoising autoencoders, *Proceedings of the 25th international conference on Machine learning*, ACM, pp. 1096–1103 (2008).

[10] Krizhevsky, A., Sutskever, I. and Hinton, G. E.: Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems*, pp. 1097–1105 (2012).

[11] Kim, Y.: Convolutional neural networks for sentence classification, *arXiv preprint arXiv:1408.5882* (2014).

[12] Zhang, X., Zhao, J. and LeCun, Y.: Character-level convolutional networks for text classification, *Advances in neural information processing systems*, pp. 649–657 (2015).

[13] Scherer, D., Müller, A. and Behnke, S.: Evaluation of pooling operations in convolutional architectures for object recognition, *Artificial Neural Networks-ICANN 2010*, pp. 92–101 (2010).

[14] Masci, J., Meier, U., Cireşan, D. and Schmidhuber, J.: Stacked convolutional auto-encoders for hierarchical feature extraction, *Artificial Neural Networks and Machine Learning-ICANN 2011*, pp. 52–59 (2011).

[15] Du, B., Xiong, W., Wu, J., Zhang, L., Zhang, L. and Tao, D.: Stacked Convolutional Denoising Auto-Encoders for Feature Representation, *IEEE Transactions on Cybernetics*, Vol. 47, No. 4, pp. 1017–1027 (2017).

[16] Glorot, X., Bordes, A. and Bengio, Y.: Deep Sparse Rectifier Neural Networks., *Aistats*, Vol. 15, No. 106, p. 275 (2011).

[17] Dozat, T.: Incorporating Nesterov momentum into Adam, Technical report, Stanford University, Tech. Rep., 2015.[Online]. Available: http://cs229.stanford.edu/proj2015/054_report.pdf (2015).