

CNNを用いた心電図の自動診断支援に関する研究

HONDA, Yoshiyuki / 本田, 祥之

(出版者 / Publisher)

法政大学大学院理工学・工学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 理工学・工学研究科編

(巻 / Volume)

62

(開始ページ / Start Page)

1

(終了ページ / End Page)

7

(発行年 / Year)

2021-03-24

(URL)

<https://doi.org/10.15002/00023982>

CNNを用いた心電図の自動診断支援に関する研究

RESEARCH ON CNN-BASED AUTOMATIC DIAGNOSIS SUPPORT FOR ELECTROCARDIOGRAM

本田祥之

Yoshiyuki HONDA

指導教員 藤井章博

法政大学大学院理工学研究科電気電子工学専攻博士後期課程

In this study, we proposed a method to determine normal and abnormal ECG waveforms and to classify them into multiple classes. In this method, the ECG is not treated as time-series waveform data, but as two-dimensional waveform image data, and classified. We used four classes from the MIT-BIH arrhythmia database: N class, VEB class, SVEB class, and F class, and classified them using CNN, VGG16, VGG19, and ResNet. The maximum accuracy was achieved with VGG16 and VGG19. 99.22% of sensitivity, 99.52% of specificity, and 99.3% of accuracy were obtained.

Keywords : ECG, CNN, MIT-BIH

1. 序論

(1) 研究背景

心電図検査は単に循環器疾患の診断にとどまらず、内科系・外科系を問わず疾患や病態の診断、経過の判定、術前・術後評価、薬効評価や副作用の検出、さらには健康診断の基本項目となっている。有病者のみならず就労者や学童・一般住民らの健康状態を量るバロメーターとしても利用され、その活用範囲と目的は広範にわたる。心電図の診断は、生化学的検査などの定量的な計測とは異なり、定性的な判断による部分が多い。即ち、数字で表しにくい波形に頼る判読であるといえる。このため、専門家により異なる診断や波形の判読、複雑波形や異常所見が重複した場合の自動診断には限界があると言われていた。さらに、この欠点が最も顕著なのが不整脈であり、同じ波形や一定の規則性を発揮するものは別にして、規則性のまったくない、またはほとんど見られない複雑不整脈、時間とともに変化する不整脈などにおける自動診断能力には、いまだ限界がある[1]。

心電図検査では多くの場合、心電計に装備されたコンピュータを用いた自動解析プログラムによる自動診断所見に基づいて臨床応用されている。自動解析には波形計測の処理から得られた所見をミネソタ分類、その他アルゴリズムの処理により診断を導く方式となっており、その方式は各社により異なる診断プログラムが使用されている。しかし、自動診断に用いられる診断基準や診断アルゴリズムは、対象となる患者の性別、年齢などによっても大きく左右されるため、統一されておらず、それによる様々な問題点が存在している。

(2) 研究目的

現在、医療分野では自動診断の診断制度を改善し、臨床における有効性をさらに高めるための研究が盛んに行われている。そして、近年ではルールベースによる手法に加え、機械学習を用いた手法も散見されるようになってきているが、既存の研究では限定された異常所見を対象とした分類が多い。また、心電図データを取り扱う場合ノイズを除去や前処理を行ってから分類を行う必要があるため、リアルタイム性のある診断支援を実現するのは難しいのではないかと考えられる。そこで本研究では、心電図波形の診断支援を目的とし、心電図波形が正常範囲内に含まれるかのような異常所見があるのかを推定し、多クラス分類を行う手法の提案を目的とする。

2. 心電図の自動診断

(1) 従来のアプローチ

従来の心電図の自動診断には、ミネソタコードによるルールベースの分類が用いられてきた。ミネソタコードとは、心電図所見を出来るだけ客観的に計測し、共通の尺度で心電図を分類しようとする試みである。運動負荷前、および後の12誘導心電図のQRS群やST-T波の高さ、幅や形および不整脈の有無などを決められた数字(コード)で表現するものである。ミネソタコードは最初、ミネソタ大学のBlackburnらが成人男子を対象にして疫学用、特に虚血性心疾患の疫学調査用に考慮された分類として1960年に報告された。その後も部分的な修正を重ね1968年にはWHOのモノグラフとして出版され、国際的に循環器疫学検診の心電図所見分類の基準としてその使用が

推奨されている。日本では、1967年の日本循環器管理研究協議会心電図小委員会でさらに一部改訂したミネソタコードを採用し、疫学用心電図分類に使用している。

このミネソタコードを用いた分類には先述した問題点が存在し、医療現場で利用するには限界があることを指摘されている。心電図自動分類の標準化のためにミネソタコードを適用する場合にも、その長所、短所をよく理解し、検査の対象、目的や方法の差により、その解釈や評価は慎重でなければならないとされている[2]。

(2) 機械学習を用いた手法

近年では、機械学習を用いた心電図分類手法も増えてきている。その多くは、心電図を時系列数値データとして取り扱い、再帰型ニューラルネットワーク(RNN)の一種である長短期記憶(LSTM)を用いた分類手法である。また、畳み込みニューラルネットワークを用いた画像分類による方法も提案されている。しかし、事前にフーリエ変換やウェーブレット変換等の前処理が必要であることや[3][4]、画像サイズを極端に大きくできないことによって一度に入力できる心電図の時間なども少なくなってしまうという問題点も存在する。また、特定の異常所見を限定した正常異常分類が多く、複数の異常所見を対象とした分類の研究事例も多くないのが現状である[5][6]。

3. 提案手法

(1) 手法概要

従来の手法では、心電図波形を時系列データとして取り扱い分類する手法が主流であったが、本研究では心電図波形を2次元波形画像として取り扱い、CNN及び、学習済みモデルを利用したファインチューニングによる深層学習を適用する。提案手法の概要図を図1に示す。分類用のデータには、MIT-BIH不整脈データベースを用いる。このデータベースは患者のホルター心電図からのデータについて医師から様々なラベリングがされており、分類性能を客観的に評価することができ、関連研究での利用例も多く性能の比較を行いやすいためである。データセットの作成方法については、後の項で記述する。

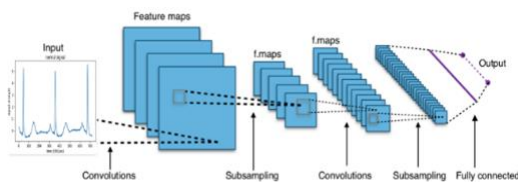


図1 提案手法の概要図

(2) MIT-BIH 不整脈データベース

ECGの公開データとして有名なMIT-BIH Arrhythmia Database (mitdb)を使用した。このデータセットはホルター心電図記録装置で記録された47名の患者から収集した48レコードからなり、各レコードファイルには約30分間の2誘導(II, V1)のシグナルデータが格納されて

いる。また、各R波のピーク位置に対してアノテーションが付与されている。シグナルデータのサンプルを以下の図2に示す。このデータは大きくN(正常)、VEB(心室性期外収縮)、SVEB(上室期外収縮)、F(融合収縮)、Q(不明)の5つのクラスに分類される。このデータのクラス表を図3に示す。本研究ではこの5つのクラスのうちN、VEB、SVEB、Fの4つのクラスのデータを使用した。実際に使用する際には、特定の病状のデータのみを使用するためにペースメーカーを使用している患者のデータについては除外した。

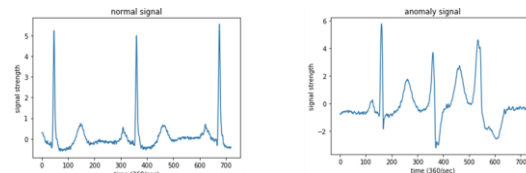


図2 MIT-BIH心電図波形の例(左:正常、右:異常所見あり)

AAMI Heartbeat Classes	N	SVEB	VEB	F	Q
Description	Any heartbeat not in S, V, F or Q class	Supraventricular ectopic beat	Ventricular ectopic beat	Fusion beat	Unknown beat
MIT-BIH Heartbeat types (annotation symbols)	Normal (N)	Atrial premature beat (A)	Premature ventricular contraction (V)	Fusion of ventricular & normal beat (F)	Paced beat (I)
	Left bundle branch block beat (L)	Aberrated atrial premature beat (a)	Ventricular escape beat (E)		Fusion of paced and normal beat (f)
	Right bundle branch block beat (R)	Nodal (Junctional) premature beat (J)			Unclassified beat (Q)
	Atrial escape beat (e)	Supraventricular ar premature beat (S)			
	Nodal (Junctional) beat (j)				

図3 MIT-BIHクラス一覧

(3) 画像データセット作成方法

画像作成の具体的な手順として、10拍分の心電図波形画像を図4に示す。赤い点で示すR波から次のR波までで区切り、10枚の心電図画像に分割する。R波を基準に分割した10枚の心電図波形画像を重ね合わせることで、図5に示す1枚の画像にまとめる。このような方法で画像を作成することで、取り扱う心電図の長さ依存せず固定サイズの画像を作成することができるので、深層学習における計算コストの問題を回避することができる。

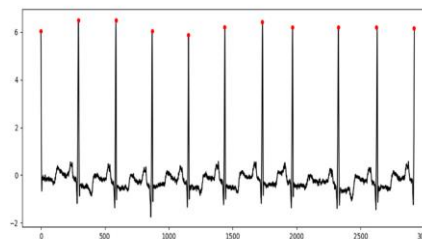


図4 アノテーションが付与された心電図波形

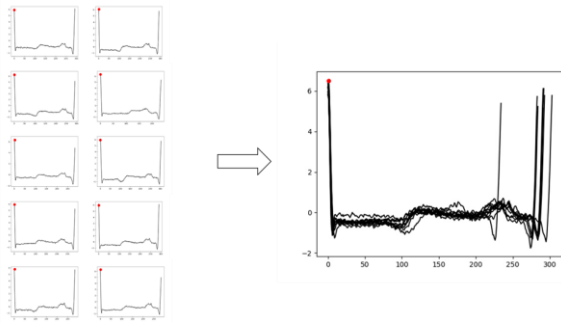


図5 画像の合成方法

これらの手順でNクラス, VEBクラス, SVEBクラス, Fクラスの画像をそれぞれ 958, 694, 269, 79 枚ずつ作成した. 作成した各クラスの画像の例を図6に示す. 上記の手順で作成したこれらの画像をデータセットとし, 学習, 及び分類に利用した.

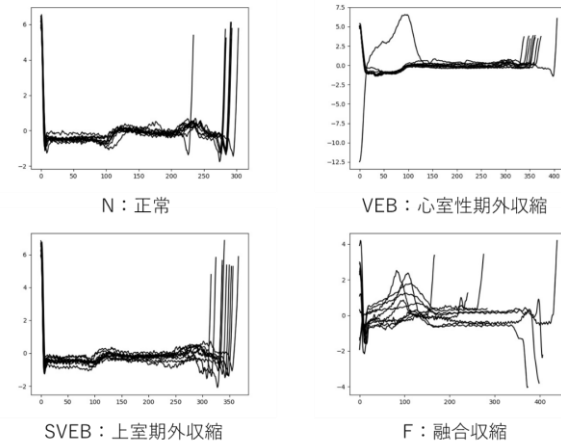


図6 作成した画像の例(左上:正常、右上:心室性期外収縮、左下:上室性期外収縮、右下:融合収縮)

(4) 分類におけるデータセットの分割

作成したデータセットを, train データ(60%), validation データ(20%), test データ(20%)に分割し, 学習及びテストを行った. 学習にはCNN, VGG16, VGG19, ResNet によるファインチューニングを用いた.

4. 結果

(1) 提案手法により得られた結果

この項では, CNN, VGG16, VGG19, ResNet の4つの分類器での分類精度について順に記述していく. 評価指標には accuracy(正解率), precision(適合率), recall(再現率)/sensitivity(感度), specificity(特異度), f-measure(F 値), support(データ数)を用いる.

a) CNN

CNN での学習結果を図7に示す. 最終的な正解率は以下の表1の通りになった.

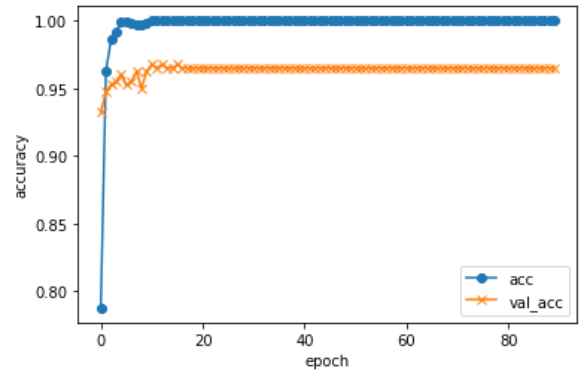


図7 CNNの学習結果

表1 CNNにおける分類正解率

	Train	Validation	Test
Accuracy	100%	96.5%	97.3%

また, 各ラベルごとの混同行列を図8に, そこから算出される評価を表2に示す.

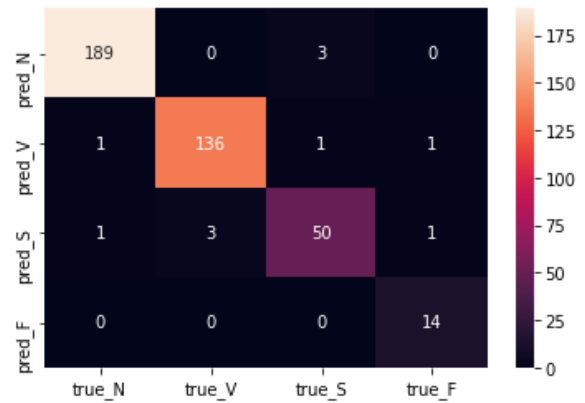


図8 CNNでの分類結果

表2 CNNにおける分類の性能評価

クラス	再現率	適合率 /感度	F 値	特異度	データ数
N	0.98	0.99	0.99	0.99	191
VEB	0.98	0.98	0.98	0.99	139
SVEB	0.91	0.93	0.92	0.99	54
F	1	0.88	0.93	1	16

b) VGG16

VGG16 での学習結果を図9に示す. 最終的な正解率は以下の表3の通りになった.

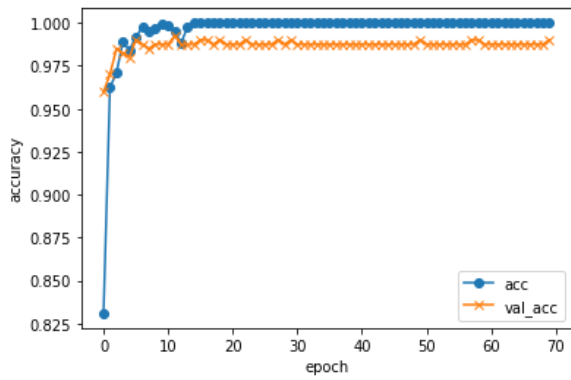


図9 VGG16の学習結果

表3 VGG16における分類正解率

	Train	Validation	Test
Accuracy	100%	99.0%	99.3%

また、各ラベルごとの混同行列を図10に、そこから算出される評価を表4に示す。

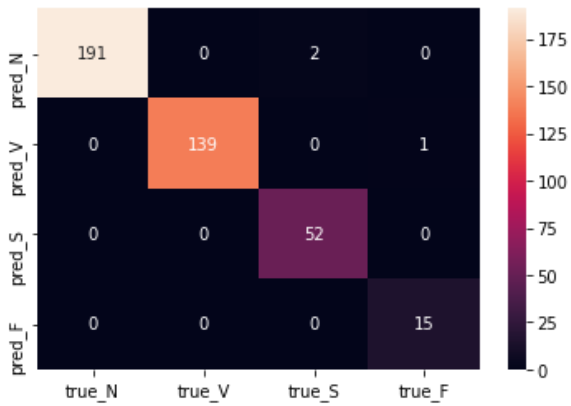


図10 VGG16での分類結果

表4 VGG16における分類の性能評価

クラス	再現率	適合率 /感度	F値	特異度	データ数
N	0.99	1	0.99	0.99	191
VEB	0.99	1	1	1	139
SVEB	1	0.96	0.98	1	54
F	1	0.94	0.97	1	16

c) VGG19

VGG16での学習結果を図11に示す。最終的な正解率は以下の表5の通りになった。

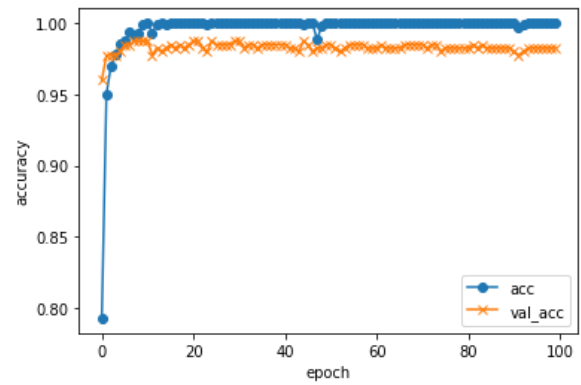


図11 VGG19の学習結果

表5 VGG19における分類正解率

	Train	Validation	Test
Accuracy	100%	99.0%	99.3%

また、各ラベルごとの混同行列を図12に、そこから算出される評価を表6に示す。

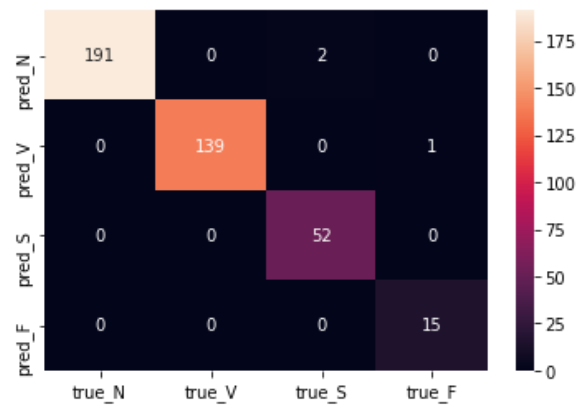


図12 VGG19での分類結果

表6 VGG19における分類の性能評価

クラス	再現率	適合率 /感度	F値	特異度	データ数
N	0.99	1	0.99	0.99	191
VEB	0.99	1	1	1	139
SVEB	1	0.96	0.98	1	54
F	1	0.94	0.97	1	16

d) ResNet

VGG16での学習結果を図13に示す。最終的な正解率は以下の表7の通りになった。

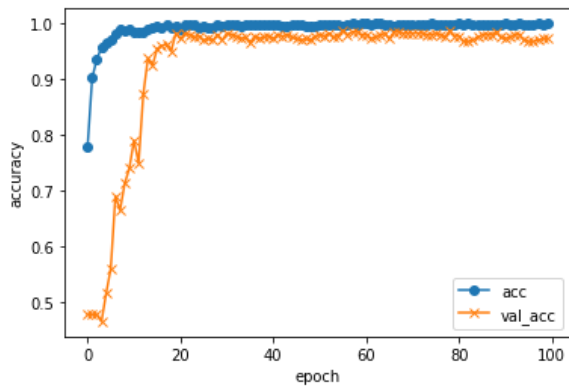


図 1 3 ResNet の学習結果

表 7 ResNet における分類正解率

	Train	Validation	Test
Accuracy	100%	99.0%	98.0%

また、各ラベルごとの混同行列を図 14 に、そこから算出される評価を表 8 に示す。

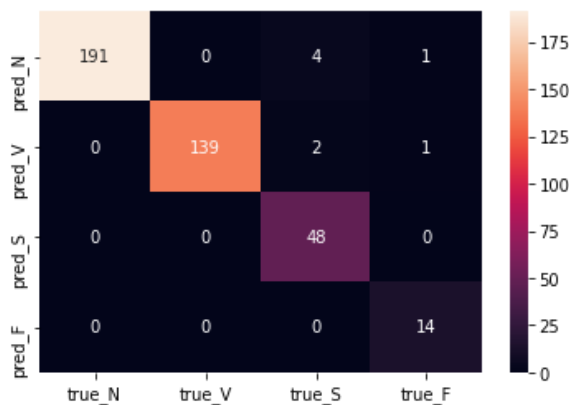


図 1 4 ResNet での分類結果

表 8 ResNet における分類の性能評価

クラス	再現率	適合率 / 感度	F 値	特異度	データ数
N	0.97	1	0.99	0.98	191
VEB	0.98	1	0.99	0.99	139
SVEB	1	0.89	0.94	1	54
F	1	0.88	0.93	1	16

(2) 先行研究及び関連分野との比較

提案手法及び関連研究の分類対象としたクラス、sensitivity, specificity, accuracy をまとめたものを以下の表 9 に示す。また提案手法の sensitivity, specificity は各クラスの結果をデータ数で重みづけ平均したものを記載した。

表 9 提案手法との性能比較

著者	分類対象のクラス	感度	特異度	正解率
Acharya et al.[5]	2	95.32%	91.04%	93.18%
Inan et al.[6]	2 (N, V)	—	—	95.16%
Isin et al.[7]	3 (Normal, R, PB)	—	—	92 %
Rai et al.[8]	3 (N, L, R)	99.01%	99.53%	99.07%
Jambukia et al.[9]	3 (N, L, R)	—	—	99.41%
Jambukia et al.[9]	4 (N, L, R, PB)	—	—	98.68%
Wu et al.[6]	4(NSR, AF, OTHER, NOISE)	97.56%	99.19%	97.56%
Oh et al.[10]	5 (N, L, R, V, A)	97.50%	98.70%	98.10%
Mahfuz et al.[5]	5(N, L, R, PB, V)	100%	100%	100%
提案手法	4(N,VEB,S,VEB,F)	99.22%	99.52%	99.3%

5. 考察

(1) 提案手法に対する考察

前章の結果から、CNN での分類では N クラス、VEB クラスに比べて SVEB クラス、F クラスの分類性能が低いことがわかる。これは、学習や検証に用いるデータの少なさが原因ではないかと考えられる。しかし、VGG16 や VGG19 では、ファインチューニングにより少ないデータからもしっかりと特徴を抽出し、データの少なさに関する問題を回避することができた。また、ResNet での結果も CNN の結果を上回ることはできたが、SVEB クラスや F クラスなどデータの少ないクラスの誤判定が目立つ結果となった。以上のことから提案手法で最も良い性能になったのは同値で VGG16, VGG19 であった。結果ではそれぞれ、N クラス、VEB クラスではすべてのテストデータに対して正答しており、SVEB クラスからは 54 個中 2 個、F クラスからは 16 個中 1 個の誤判定であった。学習に用いた SVEB クラスの画像は図 15 のような画像が多く、テストデータに含まれた 2 枚の誤判定画像(図 16)のような画像はあまり見られなかったため、判定がうまくいかなかったのだと考えられる。

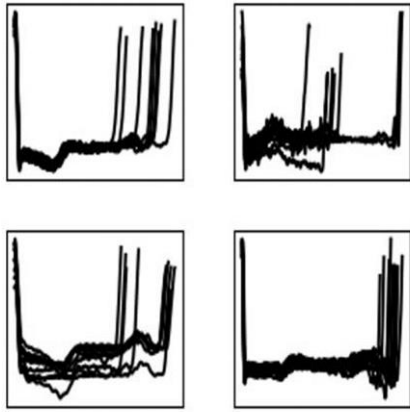


図 1 5 SVEB クラスの学習データ

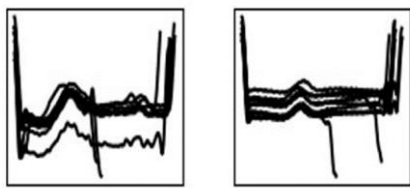


図 1 6 SVEB クラス誤判定画像

Fクラスも同様に、学習データは図 17 のような画像がほとんどで、誤判定画像(図 18)とは形状の異なっていたためであると考えられる。

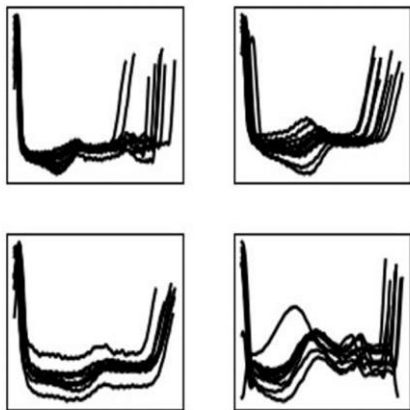


図 1 7 F クラスの学習データ



図 1 8 F クラス誤判定画像

(2) 関連研究との比較に対する考察

取り扱うクラスや分類するクラス数が異なるため単純

な比較を行うのは難しいが、表 9 より関連研究と比較したとき sensitivity では 2 番目, specificity では 3 番目, accuracy では 3 番目に高い数値であった. 同じ 4 クラス分類においては, Jambukia らの研究[9], Wu らの研究[6] よりも sensitivity, specificity, accuracy とともに最も高い性能を出すことができた. 提案手法よりも高い性能を出している研究もあるが, 提案手法では心電図データに前処理を加えておらず, 任意の時間長の心電図を扱うことができるという点から, 様々な応用が可能であると考えられる.

6. 結論

本研究では, 心電図の自動診断の限界に着目し, 心電図波形の診断支援を目的として, 心電図波形が正常範囲に含まれるかどのような異常所見があるのかを推定し, 多クラス分類を行う手法の提案を目的とした. 本手法は, 心電図を時系列波形データとして取り扱うのではなく, 2次元波形画像データとして取り扱い分類に応用する方法である. MIT-BIH 不整脈データベースの中から N クラス, VEB クラス, SVEB クラス, F クラスの 4 クラスを使用し, CNN, VGG16, VGG19, ResNet を使用し分類を行った. VGG16, VGG19 で最大の精度を出すことができ, sensitivity(感度)では 99.22%, specificity(特異度)では 99.52%, accuracy(正解率)では 99.3%の結果となった. 複雑な前処理が必要なく, 任意の時間長の心電図を扱えるという点から様々な応用の可能性が考えられる手法の提案を行うことができた.

謝辞

本研究を進めるにあたり, 熱心なご指導を頂きました法政大学応用情報工学科藤井章博教授に感謝いたします. 研究活動を支えてくださった法政大学大学院工学研究科平成 26 年度修士課程修了清水宏泰氏, 藤井研究室の皆様感謝いたします. また, 修士課程修了に至るまで, 不自由なく学業に励めるよう支えてくださった両親にも感謝いたします.

参考文献

- 1) 平岡昌和, “心電図自動診断の限界”, 日本不整脈心電学会, J-Stage, 心電図, 2015 年 35 巻 2 号 p. 149-155.
- 2) 岡本 登, 岩塚 徹, ”ミネソタコードによる心電図分類の問題点”, 第 3 回日循協総会誌, 33, 1968.
- 3) Md. Rashed-Al-Mahfuz, Mohammad Ali Moni, Pietro Li'o, Sheikh Mohammed Shariful Islam, Shlomo Berkovsky, Matloob Khushi, and Julian M.W. Quinn, “Deep Convolutional Neural Networks based Beats Classification to Identify ECG Characteristic Frequencies”, JOURNAL OF LATEX CLASS FILES, VOL. 14, NO. 8, AUGUST 2015

- 4) Ziqian Wu, Xujian Feng and Cuiwei Yang*, "A Deep Learning Method to Detect Atrial Fibrillation Based on Continuous Wavelet Transform", 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)
- 5) U. R. Acharya, H. Fujita, S. L. Oh, U. Raghavendra, J. H. Tan, M. Adam, A. Gertych, and Y. Hagiwara, "Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural network," *Future Generation Computer Systems*, vol. 79, pp. 952–959, 2018.
- 6) O. T. Inan, L. Giovangrandi, and G. T. Kovacs, "Robust neuralnetwork- based classification of premature ventricular contractions using wavelet transform and timing interval features," *IEEE transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2507–2515, 2006.
- 7) A. Isin and S. Ozdalili, "Cardiac arrhythmia detection using deep learning," *Procedia computer science*, vol. 120, pp. 268–275, 2017.
- 8) H. M. Rai and K. Chatterjee, "A novel adaptive feature extraction for detection of cardiac arrhythmias using hybrid technique mrdwt & mpnn classifier from ecg big data," *Big data research*, vol. 12, pp. 13–22, 2018.
- 9) S. H. Jambukia, V. K. Dabhi, and H. B. Prajapati, "Ecg beat classification using machine learning techniques," *International Journal of Biomedical Engineering and Technology*, vol. 26, no. 1, pp. 32–53, 2018.
- 10) S. L. Oh, E. Y. Ng, R. San Tan, and U. R. Acharya, "Automated diagnosis of arrhythmia using combination of cnn and lstm techniques with variable length heart beats," *Computers in biology and medicine*, vol. 102, pp. 278–287, 2018.