

G-013

畳み込みニューラルネットワークを用いた舌画像による生化学データの異常判別 Abnormality discrimination of biochemical data by learning tongue image through convolution neural network

網島 克幸[†] 長谷川 豊[†] 森 康久仁[‡] 中口 俊哉[§] 須鎗 弘樹[‡]
Katsuyuki Tsunashima Yutaka Hasegawa Yasukuni Mori Toshiya Nakaguchi Hiroki Suyari

1. はじめに

東洋医学には、舌診という診断方法が存在する。舌診とは舌の色や形、または舌苔の厚さや色などを観ることで身体の状態の診断を行うものである。舌診の利点として、即時性と非侵襲性がある。しかし、その診断は医師の経験や主観に依存し、定性的である。そのため、舌診は定量的データを重んじる西洋医学では受け入れにくく、限られた分野でしか用いられていない。

この問題を解決するには、舌診を客観的に行うことが必要である。そのために、画像解析によって舌の特徴を定量化する研究[1]や、ベイジアンネットワークを用いて、舌の特徴から内臓疾患を判別する研究が行われている[2]。

また、近年画像解析において、畳み込みニューラルネットワークが活用されて、人間による判別が難しいとされる医療画像分野への応用が進んでいる。[3]では、畳み込みニューラルネットワークを用いて皮膚がんの識別を行ったところ、皮膚科の専門医と同程度の識別率が得られたと報告されている。

そこで、本研究ではこれらの背景から、客観的な舌診を行う研究の一環として、画像認識において活躍している畳み込みニューラルネットワークを用いることで、舌画像から内臓の異常を定量的に判別することを行う。畳み込みニューラルネットワークに舌画像を入力し、定量的なデータである生化学データの正常、異常をラベルとして学習を行う。

2. 舌診

東洋医学において、舌と内臓とは経路によって繋がっており、舌表面の各部位と各臓器はお互いに関連していると考えられている[4]。昔から舌と臓器との生理病理の関連については多くの文献に記載されている。図 1 は舌の表面の各部位と各臓器の対応を示した舌表面臓腑分布図である。一般的に舌表面は舌根、舌中、舌辺、舌尖の 4 部分に区分される。それぞれ、舌根は腎臓、舌中は脾臓と胃、舌辺は肝臓と胆嚢、舌尖は心臓と肺に対応し、その区域の変化が該当する内臓の変化を反映すると考えられている。

この考えのもと、舌の特定の部位の色や形、舌苔の厚さや色などを観ることで身体の状態を診断する方法が舌診である。

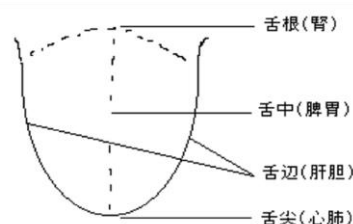


図 1 舌表面臓腑分布図[4]

3. 実験

3.1 実験概要

今回の実験では、畳み込みニューラルネットワークに舌画像を入力して、生化学データの正常、異常を判別する 2 クラス分類を行う。畳み込みニューラルネットワークのモデルとして ZFNet[5] と GoogLeNet[6] の 2 種類を用いる。それぞれのモデルで学習を行い、テストデータの判別率の計測を行うことを 2 回繰り返す。ここで、判別率とは、全データに対し、データを正しく判別できた割合のことである。

学習させる際の条件として、全データのうち 7 割を訓練データ、3 割をテストデータとし、試行毎に訓練データとテストデータの抽出を無作為に行う。

3.2 使用するデータ

今回の実験では学習データとして、千葉大学附属病院和漢診療科の外来患者と入院患者を対象に、事前に実験の説

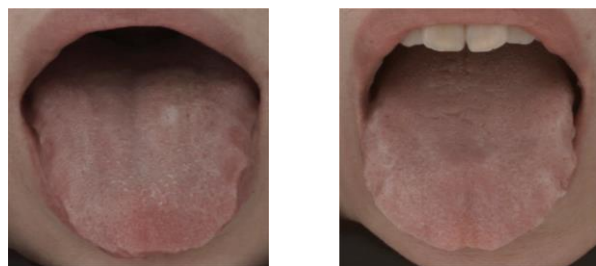


図 2 使用した舌画像例

[†] 千葉大学大学院融合理工学府, Graduate School of Science and Engineering, Chiba University

[‡] 千葉大学大学院工学研究院, Graduate School of Engineering, Chiba University

[§] 千葉大学フロンティア医工学センター, Center for Frontier Medical Engineering, Chiba University

明を行い同意が得られた 39 名の患者の舌の撮影画像と対応する生化学データを用いた。なお、これらのデータは千葉大学大学院医学研究院倫理審査承認番号 812 号の承認を得ている。図 2 に使用した舌画像の例を示す。

肝臓の機能の状態を示すといわれる 8 項目の生化学データに対し、それぞれ異常値を示した被験者の舌を「異常」クラス、正常値を示した被験者の舌を「正常」クラスとしてそれぞれ学習を行った。表 1 に使用する生化学データと該当する人数を示す。舌画像は 1 人の被験者につき 5 枚存在する。それぞれの生化学データごとに正常値の人数が異常値の人数の約 1.5 倍になるように選び学習させる。また、各舌画像に対し、ガンマ値を 19 段階に変化させた画像と輝度を 2 段階に変化した画像を追加しデータの拡張を行う。訓練データとテストデータの抽出の際には、同じ被験者の舌画像が、訓練データとテストデータの両方に含まれるようなことがないようにする。

表 1 判別する生化学データと使用データ数

生化学データ	データ数	
	正常(人)	異常(人)
ALP	7	5
CHE	7	4
G-GTP	18	12
GPT	7	4
LDH	8	5
T-BIL	10	7
T-CHO	8	5
PLT	6	3

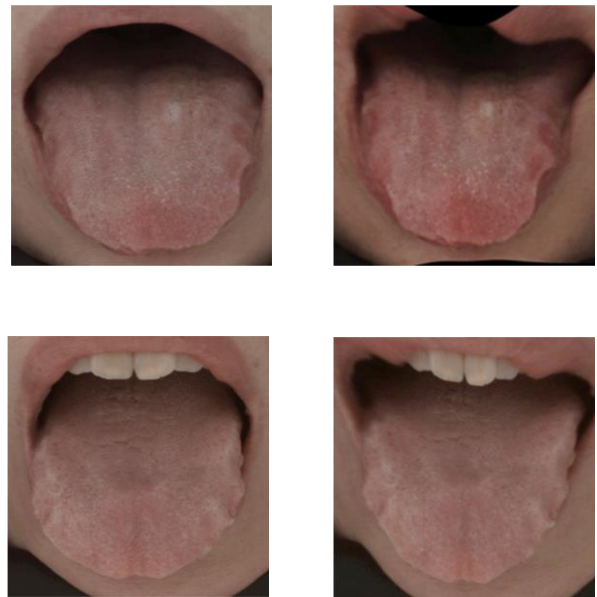


図 3 舌の形状を全て統一
左側が元の舌画像、右側が変形した舌画像

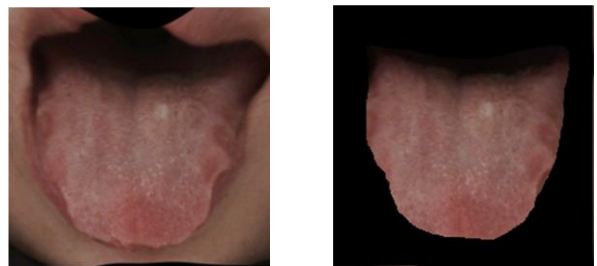


図 4 舌以外の部分を黒塗り

3.3 舌画像の前処理

舌の形状や色、大きさには個人差がある。学習の際には舌の表面の情報のみを用いるため、舌の形状による影響をなくす必要がある。そこで Thin plate spline(TPS)とアフィン変換を組み合わせた舌形状規格化手法を用いて、全ての舌画像を標準舌形状に統一するように変形する[7]。図 3 は複数の舌画像を舌形状規格化手法によって、同じ形になるように変形した様子を表している。左側の画像が元の舌画像であり、その画像をアフィン変換と TPS によって変形したのが右側の画像である。

また、唇や肌など舌以外の部分から情報を得て学習してしまうことを避けるため、舌画像の形状を統一した後、図 4 のように舌以外の部分を全て黒く塗りつぶした。

4. 結果

実験を行った結果を表 2 に示す。表 2 はそれぞれのモデル毎の肝臓に関する 8 種類の生化学データの正常、異常の判別率をまとめたものである。

表 2 から、生化学データによって判別率に差があるのが見て取れる。ZFNet を用いた場合には、CHE, GPT, T-CHO

表 2 各生化学データの判別率(%)

生化学データ	ZFNet		GoogLeNet	
	1回目	2回目	1回目	2回目
ALP	40	32	36	48
CHE	88	48	44	60
G-GTP	51	43	41	61
GPT	76	85	58	31
LDH	49	62	63	50
T-BIL	48	28	56	60
T-CHO	68	72	64	52
PLT	40	29	67	67

の判別において、判別率が7~8割程度の結果が得られるときがあった。また、それぞれの試行ごとに判別率の差が大きい場合があった。

そこで、一番判別率が高かった ZFNet による CHE の学習の様子を図 5 に示す。青色の線が訓練データの判別率、緑色の線がテストデータの判別率を示しており、1 エポック毎のそれぞれの判別率の推移を表している。テストデータの 1 エポックごとの判別率の差が大きいことが見てとれる。この原因は学習に使用したテストデータ数が少なかったためだと考えられる。また、試行ごとの判別率の差もテストデータに選ばれる画像に影響していると考えられる。

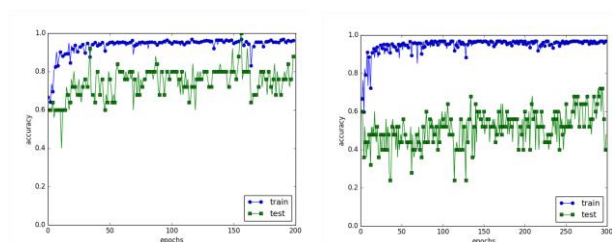


図 5 ZFNet による CHE の学習の様子
左が 1 回目、右が 2 回目

したがって、学習に使用するデータ数を増やすために、生化学データを組み合わせてもう一度学習を行うことにした。先ほどの実験で ZFNet での判別率が高かった上位 4 項目 CHE, GPT, LDH, T-CHO のうち 2 つを組みあわせる。どちらか一方でも異常があった被験者の舌を「異常」クラス、異常値を示す項目がなかった被験者の舌を「正常」クラスとして学習を行う。学習は先ほどと同じ条件で行いテストデータの判別率を計測した。表 3 に組み合わせた生化学データの組み合わせとそれぞれの人数を表す。

表 3 生化学データの組み合わせと使用データ数

生化学データ	データ数	
	正常(人)	異常(人)
CHE と GPT	12	7
CHE と LDH	13	7
CHE と T-CHO	13	8
GPT と LDH	14	9
GPT と T-CHO	12	8
LDH と T-CHO	14	9

表 4 に各生化学データの組み合わせごとに実験を行ったときの判別率を示す。

表 4 組み合わせた生化学データの判別率 (%)

生化学データ	ZFNet		GoogLeNet	
	1 回目	2 回目	1 回目	2 回目
CHE と GPT	71	50	73	60
CHE と LDH	43	83	59	67
CHE と T-CHO	57	85	43	44
GPT と LDH	50	56	50	69
GPT と T-CHO	57	40	43	44
LDH と T-CHO	66	71	56	60

先ほどの実験と同様に、ZFNet を用いた場合には CHE と GPT, CHE と LDH, CHE と T-CHO, LDH と T-CHO の判別において、7~8 割の判別率を得られることがあった。しかし、依然として試行ごとの判別率の差が大きいことが見てとれる。

一番判別率が高かった ZFNet による CHE と T-CHO の学習の様子を図 6 に示す。図 5 と比べると、テストデータの 1 エポックごとの判別率の差は先ほどの実験よりも小さくなった。しかし、どちらの試行でも訓練データの判別率の上昇に対しテストデータの判別率が上昇しておらず、過学習を起こしていることが見てとれる。これらの結果から、データ拡張だけでなく、もっと多くの舌画像の種類を用いて学習し、判別を行ってみる必要があると考えられる。

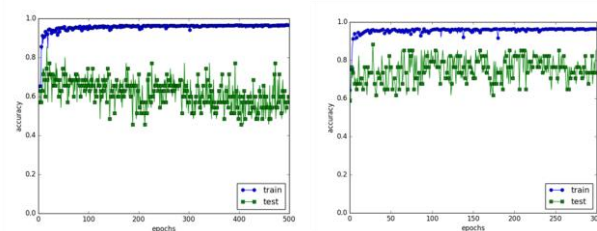


図 6 ZFNet による CHE と T-CHO の学習の様子
左が 1 回目、右が 2 回目

5. まとめと今後の展望

5.1 まとめ

本研究では、畳み込みニューラルネットワークを用いることで、舌画像から肝臓の状態に関係があるとされる生化学データの異常を定量的に判別する実験を行った。ZFNet を用いた学習により、生化学データによっては判別率が 7~8 割となる結果を示すことがあった。しかし、試行ごとの判別率の差が大きく、生化学データを組み合わせることで舌画像の種類を増やして学習を行ったが、改善は見られなかった。

5.2 今後の展望

今回の実験では、計 39 人分の舌画像という少ない種類の舌画像をガンマ変換と輝度変換によるデータ拡張によって大量に増やし学習を行った。過学習を抑制し、判別率を向上するためには、さらに舌画像の種類を増やして学習を行うことが考えられる。また、訓練データの判別率が 100% 近くに到達していることから、パラメータ数が多すぎる可能性がある。したがって、もっとパラメータ数の少ないモデルを用いた判別を行う必要もある。さらに、今回畳み込みニューラルネットワークに入力したのは舌画像の RGB の画素値であったため、それ以外の特徴量を用いて学習を行うことで判別率に影響があるかも確かめたい。

参考文献

- [1] Chuang-ChienChiu, "A novel approach based on computerized image analysis for traditional Chinese medical diagnosis of the tongue", Computer Methods and Programs in Biomedicine, vol. 61, pp77-89 (2000).
- [2] Bo Pang, David Zhang, Naimin Li and Kuanquan Wang, "Computerized tongue diagnosis based on Bayesian networks", IEEE Transactions on Biomedical Engineering, No10, vol.51, pp.1803-1810(2004).
- [3] Andre Esteva, Brett Kuprel, Rob Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau and Sebastian Thrun, "Dermatologist-level classification of skin cancer with deep neural networks", Nature, 542(7639):115-118(2017)
- [4] 張立也, "舌診の研究", 埼玉医科大学雑誌, 第 30 卷, 第 3 号, pp.T31-T32 (2003)
- [5] Matthew D. Zeiler, Rob Fergus, "Visualizing and Understanding Convolutional Networks", in Proc. of ECCV (2014).
- [6] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincen Vanhoucke, Andrew Rabinovich, "Going Deeper with Convolutions", in Proc. IEEE CVPR, pp.4-8(2015).
- [7] Kazunari Murai, Toshiya Nakaguchi, Akira Morita, Takao Namiki "Association Analysis of Tongue Color Spatial Distribution and Physiological Index based on Tongue Shape Normalization" Proceedings of 13th AIC Congress 2017, OS24-03(2017).