

東洋医学における舌診と深層学習に関する研究
Study on Tongue Diagnosis and Deep Learning
Based on Traditional Chinese Medicine

安 振宇

Zhenyu An

山口大学大学院 東アジア研究科

The Graduate School of East Asian Studies

Yamaguchi University

指導教員：中田 充，葛 崎偉，鷹岡 亮

Supervisor: Mitsuru Nakata, Qi-Wei Ge, Ryo Takaoka

要 旨

学位論文題目：東洋医学における舌診と深層学習に関する研究

指導教員：中田 充, 葛 崎偉, 鷹岡 亮

申請者名：安 振宇 (アン シンウ)

2018年10月入学

山口大学大学院 東アジア研究科 博士課程

東洋医学とは古代中国を起源とし、日本や韓国を含む世界中の国々に広がり、発展してきた医学である。東洋医学における診察は、四診（望診、聞診、問診および切診）と呼ばれる4種類の診察法で構成されている。この中で望診は、患者の顔色、表情、皮膚、爪、頭髮、舌などを注意深く観察する診察法であるが、これらのうちでも舌を見る舌診が特に重要とされている。舌診とは、舌の色・乾燥度・舌の苔（舌苔）などを診ることにより、五臓六腑の状態を診断する手法である。なお、舌は直接的或いは間接的に多くの臓腑と繋がっている。一般的に舌表面は舌尖部、舌辺部（左と右）、舌中部、舌根部の5部位に別れ、それぞれ舌尖部は心肺、舌辺部は肝胆、舌中部は脾胃、舌根部は腎、と各臓腑に対応し、それらの部位による変化は内臓気血の盛衰が舌に影響し、病変が舌に現れた症状として、内臓の変化を反映したものである。

五臓六腑とは、東洋医学において人体の内臓全体を指す際に使用される表現であり、具体的には五臓は肝・心・脾・肺・腎を指し、六腑は胆・小腸・胃・大腸・膀胱・三焦を指す。経脈とは身体を縦に流れる脈であり、気と血の通路の役割を果たすものである。気・血・津液は生きるために必要なものであり、いわゆる代謝物質である。また、経穴は一般的に経脈に沿って配置されており、人体の表面に存在している。

鍼灸治療は、東洋医学の治療法の1つであり、鍼または灸で経穴を刺激することによって、その刺激が経脈に伝わり、体内の気・血・津液が改善され、その結果、五臓六腑の機能が活性化される。鍼灸治療は、副作用が少なく、病気の予防や難病の治療などにも有効であることから、WHOに認められ、今や世界中に広まっている。しかしながら、鍼灸治療は未だ経験的・

臨床的に行われることが多く、治療の仕組みや過程が科学的に解明されていない部分が多い。そのため、現代の西洋医学や科学技術的な観点も取入れた客観的・統一的な研究方法や評価基準・評価方法の確立が求められている。

このような東洋医学の現状を踏まえ、深層学習という Artificial Intelligence (AI) 技術を活用して、四診から鍼灸治療における治療方針の決定までを総合的に支援する鍼灸治療支援システムの開発を目指している。このシステムは主に、四診を支援してその結果を解析する機能（機能1）、治療用経穴の候補を抽出する機能（機能2）、治療用経穴の候補を評価する機能（機能3）の3つから構成されている。

本論文では、鍼灸治療支援システムの機能1として舌診を取り上げ、深層学習を活用した舌状態の自動認識による臓腑の健康状態を自動診断する5分割診断法と舌画像の識別による裂紋舌（れつもんぜつ）の自動診断法を提案した。

5分割診断法では、Mask R-CNN を用いて舌画像の認識と切り抜きを行い、その結果を5つの部位に分割した上で、各部位の画像における症状の有無を判定することで、各部位に対応する臓腑の健康状態を判定する。各部位の症状の有無の判定には、LeNet, ResNet50, ResNet101, DenseNet169 によるアンサンブル学習を用いる。

また、裂紋舌の自動診断法では、舌画像の認識と切り抜きから得られた舌のみの画像データを裂紋舌の自動認識学習に適用して、裂紋舌のみの画像データを抽出した上で、裂紋舌の裂紋状態を判定する。裂紋舌の自動認識と裂紋舌の裂紋状態の判定には、LeNet, ResNet50, ResNet101, DenseNet169, ConvNeXt-Tiny の5つの画像認識モデルによるアンサンブル学習を用いる。

提案した5分割診断法と裂紋舌の自動診断法を評価するために、各々を適用する評価実験を行った。実験結果より次のことが明らかになった。なお、それぞれの手法では比較的古いモデル LeNet と比較的新しいモデル ConvNeXt-Tiny も活用するがそれらの有用性も確認された。

(1) 舌の5部位に対応する臓腑の健康状態の判定については、アンサンブル学習を用いた場合、その正解率が80%以上で、5つのモデルを個別に用いて学習した時のいずれの最良の結果よりも高くなっている。よって、舌画像による臓腑健康状態の診断においては、アンサンブル学習を用いた5分割診断法は極めて効果的であると言える。

- (2) 裂紋舌の自動認識学習については、アンサンブル学習を用いた場合、設定した4種類の学習率において、いずれも正解率が90%であり、5つの個別の画像認識モデルの学習における最良のテスト結果より高くなっていることから、アンサンブル学習を用いた方が効果的である。
- (3) 裂紋舌の裂紋状態判定学習については、5つの個別の画像認識モデルの学習とアンサンブル学習を用いた場合、それらの最良のテスト結果として、各々の最高な正解率が80%と85%であることから、アンサンブル学習を利用した方がより効果的であると言える。

以上の(1)~(3)より、本論文で提案した5分割診断法と裂紋舌の自動診断法は、それぞれ臓腑健康状態の自動診断と裂紋舌の自動診断に極めて効果的である。

これらの診断法では、深層学習を舌診に活用することで、より客観的かつ効果的な診断支援が可能となり、健康管理や医療の向上にも繋がるものである。

謝 辞

本論文の作成にあたり、終始適切な助言を賜り、また丁寧に指導して下さった多くの方々に、この場をお借りして感謝の意を申し上げます。

山口大学大学院の葛 崎偉教授には、指導教員として情報科学分野で初学者である筆者を受け入れていただき、心より感謝を申し上げます。情報科学の基礎的知識に始まり、研究手法、学会発表など、研究に関するあらゆることを教えていただきました。博士論文の完成まで、何度となく障壁に直面したものの、最後までやり遂げることができたのは、ひとえに葛 崎偉教授のご指導・ご支援のおかげです。葛 崎偉教授には、さまざまな面で厳しくも温かく見守っていただき、ここに深謝の意を表します。

山口大学大学院東アジア研究科中田 充教授、鷹岡 亮教授には、指導教員として、学術誌論文・学位論文の細部にわたり貴重なご意見・ご指摘と、論文執筆についての激励をいただきましたことをここに記すとともに、心より感謝申し上げます。

本研究を遂行するにあたり、裂紋舌画像と裂紋状態の酷い画像の選出については中国山東中医薬大学の楊佃会教授・主任医師からご協力をいただいた。ここで謝意を表します。

また、5 年半の間共に励まし合い、研究をするということに対して、何度も相談に乗ってくださった呉 彪博士（2021 年 3 月修了）、甘 泉博士（2022 年 3 月修了）、楊 航氏、張 慶琪氏をはじめ、葛研究室の仲間には常に刺激的な議論を頂き、精神的にも支えられました。ここに感謝の意を申し上げます。

最後に、筆者の博士課程への進学希望を温かく応援してくれた家族・友人にも、心から感謝の気持ちを伝えておきたいです。特に、両親の絶え間ない精神的な励ましと信頼によって、私は最後まで諦めずに本研究を頑張ってきました。困難な時には様々な形で支えてくれたことに深く感謝します。

ここに記しきれない多くの方々のご支援・ご協力によって、本研究が成立していることに、改めて厚く御礼申し上げます。皆様、ありがとうございました。

目次

第1章	はじめに	1
1.1	研究背景と研究目的	1
1.2	本論文の構成と位置付け	4
第2章	研究に関する準備	7
2.1	東洋医学における舌診	7
2.1.1	舌診とは	7
2.1.2	舌表面と五臓六腑の関係	9
2.1.3	裂紋舌	13
2.2	深層学習	14
2.2.1	人工知能, 機械学習と深層学習の関係	14
2.2.2	深層学習を舌診に活用した先行研究	18
2.2.3	Mask R-CNN 手法と6つの画像認識モデル	20
2.2.4	アンサンブル学習と水増し手法	24
2.2.5	2クラス分類問題における評価指標	27
2.2.6	深層学習による舌診の効果および意義	29
第3章	深層学習による臓腑健康状態の診断	31
3.1	深層学習による舌の5分割診断法の提案	31
3.1.1	舌診支援システムにおける処理の流れ	31
3.1.2	舌画像の自動認識・切り抜き・5分割	33
3.1.3	臓腑状態判定のための学習	35

3.2	舌の5分割診断法による学習結果と評価	39
3.2.1	5つの画像認識モデルによる結果	39
3.2.2	アンサンブル学習を用いた実験	42
第4章	深層学習による裂紋舌画像の認識	45
4.1	深層学習による裂紋舌の自動診断法の提案	45
4.1.1	裂紋舌の自動診断の処理手順	45
4.1.2	裂紋舌の自動認識とその裂紋状態判定学習	46
4.1.3	ハイパーパラメータの設定	47
4.2	裂紋舌の自動診断法による学習結果と評価	49
4.2.1	5つの画像認識モデルによる結果	49
4.2.2	アンサンブル学習を用いた実験	51
第5章	おわりに	54

目次

2.1	五臓間の五行説と臓腑間の表裏関係	10
2.2	舌表面の部位と臓腑の対応	12
2.3	ニューラルネットワークの構造	17
2.4	人工知能, 機械学習と深層学習の関係	17
2.5	LeNet の構成 (文献 [47] の Fig.1 より引用)	21
2.6	Shortcut connection の仕組み (文献 [46] の Fig.2 より引用)	21
2.7	DenseNet の構成 (文献 [48] の Fig.1 より引用)	22
2.8	モデルのスケーリング (文献 [49] の Fig.2 より引用)	23
2.9	ResNet から ConvNeXt に変更する流れ (文献 [50] の Fig.2 より引用)	24
3.1	舌診支援システムにおける処理の流れ	32
3.2	舌の自動認識と切り抜き	32
3.3	舌画像の 5 分割画像	32
3.4	舌辺部画像の処理	36
3.5	臓腑の状態判定学習に用いる画像の例	38
4.1	深層学習を用いた裂紋舌の診断の流れ	46

表目次

2.1	混同行列	27
2.2	混同行列の指標	28
2.3	テスト結果の評価指標	28
3.1	Loss の種類とその意味	34
3.2	臓腑の状態判定学習に用いる画像の内訳	35
3.3	LeNet によるテスト結果	40
3.4	ResNet50 によるテスト結果	40
3.5	ResNet101 によるテスト結果	41
3.6	DenseNet169 によるテスト結果	41
3.7	EfficientNet-B0 によるテスト結果	41
3.8	5つのモデルによる最良のテスト結果	41
3.9	アンサンブル学習によるテスト結果	43
3.10	モデル組合せのアンサンブル学習の結果	44
4.1	裂紋舌の自動認識学習に用いる画像の内訳 (水増しなし)	47
4.2	裂紋舌の裂紋状態判定学習に用いる画像の内訳 (水増しあり)	47
4.3	5つのモデルによる裂紋舌の自動認識学習の結果 (水増しなし)	50
4.4	5つのモデルによる裂紋舌の裂紋状態判定学習の結果 (水増しあり)	50
4.5	アンサンブル学習による裂紋舌の自動認識テスト結果 (水増しなし)	53
4.6	アンサンブル学習による裂紋舌の裂紋状態判定テスト結果 (水増しあり)	53
4.7	4つの画像認識モデルを対象としたアンサンブル学習のテスト結果	53

第1章 はじめに

1.1 研究背景と研究目的

東洋医学とは古代中国を起源とし、日本や韓国を含む世界中の国々に広がり、発展してきた医学である。東洋医学における診察は、四診（視覚による望診、聴覚と嗅覚を用いる聞診、会話による問診、患者の体に触れる切診）と呼ばれる4種類の診察法で構成されている。この中で望診は、患者の顔色、表情、皮膚、爪、頭髪、舌などを注意深く観察する診察法であるが、これらのうちでも舌を見る舌診が特に重要とされている。舌診とは、舌の色・乾燥度・舌の苔（舌苔）などを診ることにより、五臓六腑の状態を診断する手法である [1]。なお、舌は直接的或いは間接的に多くの臓腑と繋がっている。一般的に舌表面は舌尖部、舌辺部（左と右）、舌中部、舌根部の5部位に別れ、それぞれ舌尖部は心肺、舌辺部は肝胆、舌中部は脾胃、舌根部は腎、と各臓腑に対応し、それらの部位による変化は内臓気血の盛衰が舌に影響し、病変が舌に現れた症状として、内臓の変化を反映したものである。

五臓六腑とは、東洋医学において人体の内臓全体を指す際に使用される表現であり、具体的には五臓は肝・心・脾・肺・腎を指し、六腑は胆・小腸・胃・大腸・膀胱・三焦を指す。経脈とは身体を縦に流れる脈であり、気と血の通路の役割を果たすものである。気・血・津液は生きるために必要なものであり、いわゆる代謝物質である。また、経穴は一般的に経脈に沿って配置されており、人体の表面に存在している [2,3]。

鍼灸治療は、東洋医学の治療法の1つであり、鍼または灸で経穴を刺激することによって、その刺激が経脈に伝わり、体内の気・血・津液が改善され、その結果、五臓六腑の機能が活性化される。鍼灸治療は、副作用が少なく、病気の予防や難病の治療などにも有効であることから、1989年にWorld Health Organization (WHO) に認められ、今や世界中に広まっている。

しかしながら、鍼灸治療は未だ経験的・臨床的に行われることが多く、治療の仕組みや過程が科学的に解明されていない部分が多い。そのため、現代の西洋医学や科学技術的な観点も取入れた客観的・統一的な研究方法や評価基準・評価方法の確立が求められている [4, 5].

このような東洋医学の現状を踏まえ、深層学習という Artificial Intelligence (AI) 技術を活用して、四診から鍼灸治療における治療方針の決定までを総合的に支援する鍼灸治療支援システムの開発を目指している。このシステムは主に、四診を支援してその結果を解析する機能（機能 1）、治療用経穴の候補を抽出する機能（機能 2）、治療用経穴の候補を評価する機能（機能 3）の 3 つから構成されている。

これまでに著者の所属している研究グループでは、カラーペトリネットを用いて人体における経穴・経絡のモデルを構築し、カラーペトリネット解析ツール (CPN-Tools) を用いた鍼灸治療シミュレーション実験を通じて、治療現場では観測できない臓腑状態の変化の過程を数値的に提示する仕組みを実現してきた [6-8]。このモデルは機能 3 の基盤となるものである。本研究では、鍼灸治療支援システムの機能 1 として舌診を取り上げ、深層学習を舌診に活用し、舌状態の自動認識による臓腑の健康状態を自動診断する 5 分割診断法と舌画像の識別による裂紋舌の自動診断法を提案する。なお、本研究においては、舌の 5 部位に対応する臓腑の健康状態を自動診断した後のステップとして、病気のある臓腑に適切な水分や潤いが欠如しているか、さらにその病気の度合いまでを判定するために、舌診の対象として裂紋舌を取り上げ、その自動認識と関連する裂紋状態の自動診断技術の実現を目指している。裂紋舌は舌質（舌の肉部分）の表面が割れている状態として定義されており、臓腑の潤いが足りていない病理状態（運動不足・栄養不足）を示唆するとされている。また、裂紋舌の亀裂にはその位置と深さによって病気のある臓器と度合いが異なる [9].

本研究で提案する舌の 5 分割診断法と裂紋舌の自動診断法を鍼灸治療支援システムの機能 1 に活用することで、臓腑の健康状態と病気のある臓腑に必要な水分や潤いが不足しているかどうかを判定し、さらにその病気の度合いまで評価して提供することが可能となる。これらの情報と聞診、問診、切診の結果に基づいて、鍼灸治療支援システムの機能 1 で「症状と臓腑状態の関係データベース」を作成し、臓腑状態の判定を実施する。次に鍼灸治療支援システムの機

能 2 では、臓腑状態に対して、鍼灸治療用の候補の経穴を抽出する。ここでは、409 個経穴の効能によって、「経穴と臓腑の関係データベース」を作成し、治療用の候補経穴を抽出する。また、鍼灸治療支援システムの機能 3 では、抽出した経穴を用いて、鍼灸治療のシミュレーションを行い、その結果により、最良の鍼灸治療案が得られる。最後には、得られた最良の鍼灸治療案を提案することが可能になる。

また、本研究で提案する舌の 5 分割診断法と裂紋舌の自動診断法は、舌診を行う際に、臓腑の健康状態と病気のある臓腑に必要な水分や潤いが不足しているかどうかを判定し、さらにその病気の度合いまでを自動的に評価して提供できるため、迅速かつ効率的に行われるようになる。これにより、医師や東洋医学の専門家は他の重要な診療活動や症例の分析に時間を割くことができ、患者の診断や治療案の立案がより迅速に進められる。なお、医師や専門家の主観的な見解や経験による要素を減少させ、客観的な評価が可能となる。これにより、異なる医師や専門家の間での一貫性が向上し、より客観性のある診断が期待できる。

1.2 本論文の構成と位置付け

本論文の構成は以下のようになっている。

第 2 章では、東洋医学における舌診が発展してきた背景と現状を紹介し、東洋医学における舌診とそれに関連する舌苔、舌象、舌質の概念を具体的に説明する。なお、舌表面と五臓六腑の関係および裂紋舌について紹介する。また、深層学習を舌診に活用した先行研究を踏まえつつ、本研究で用いる深層学習の手法と 2 クラス分類の評価指標を述べる。

第 3 章では、五臓六腑の自動診断技術の実現を目指して、深層学習を活用した舌の 5 分割診断法を提案する。まずは、舌診支援システムの構成と処理手順について概説した上で、5 分割診断法における舌の自動認識と切り抜き、舌画像の 5 分割、そして、5 つの画像認識モデルを用いた臓腑状態判定のための学習について説明する。次に、臓腑状態判定のための学習結果を評価するために、5 つの画像認識モデルを用いた時の舌診（舌画像分類）の性能を調べることで各モデルの評価を行う。最後に、そのうちの 4 つのモデルを融合させるアンサンブル学習を用いた時の学習結果を評価する。さらに、その結果を 5 つの個別の学習結果と比較を行い、アンサンブル学習の有効性を示す。

第 4 章では、裂紋舌の裂紋状態の自動診断技術の実現を目指して、深層学習を活用した裂紋舌の自動診断法を提案する。まずは、裂紋舌の自動診断手順を概説した上で、5 つの画像認識モデルを用いた裂紋舌画像の自動認識と裂紋舌に基づく裂紋状態判定の学習、および各々の学習で設定したハイパーパラメータについて説明する。次に、裂紋舌の自動診断法による裂紋舌の自動認識と裂紋舌の裂紋状態判定のための学習結果を評価するために、5 つの画像認識モデルを用いた時の裂紋舌画像の自動認識と裂紋舌に基づく裂紋状態判定の学習結果を調べることで各モデルの評価を行う。なお、裂紋舌の裂紋状態判定学習において、裂紋舌画像に対して水増し手法を活用することでその有効性を示す。最後に、5 つのモデルを融合させるアンサンブル学習を用いた時の学習結果を評価して、その結果を 5 つの個別の学習結果と比較を行い、アンサンブル学習の有効性を示す。

第 5 章では、本研究で得られた研究成果をまとめた上で、その研究成果の応用に関する今後の展望および研究をさらに発展させるための課題を述べる。

これまでに舌画像の自動認識に関連する研究として、舌を自動認識するための TongueNet の提案 [10]、舌の観察結果に基づいたファジィ処理の研究 [11]、舌領域からの舌苔領域のみの抽出に関する研究 [12]、舌画像の精度向上のためのシステム開発 [13] などが行われている。また深層学習を用いた舌診に関する研究としては、新型コロナウイルス感染症 (COVID-19) の診断支援のために Convolutional Neural Network (CNN) 技術を活用して舌苔の種類を自動認識する GreasyCoatNet モデルを開発した研究 [14] や、Faster R-CNN 手法を活用して、裂紋舌 (れつもんぜつ)、歯痕舌 (しこんぜつ)、お点舌 (おてんぜつ)、お斑舌 (おはんぜつ)、膩苔 (じたい)、剥落苔 (はくらくたい)、腐苔 (ふたい) の 7 種類の舌画像を分類し、それぞれの舌象 (舌の変化) の特徴と性別、年齢、代謝性疾患との関係を解明する研究 [15] および、舌画像解析における舌領域分割と舌特徴抽出の定量化および自動化のため、ソフトウェアを開発する研究 [16] などが試みられている。しかしこれらの研究は、東洋医学の舌診でいう五臓六腑の健康状態と裂紋舌の裂紋状態を深層学習を用いて自動診断するまでには至っていない。

本研究は、五臓六腑の自動診断技術と裂紋舌に基づく裂紋状態の自動診断技術の実現を目指して、画像に含まれる舌を自動で認識して切り抜き、舌画像を 5 つの部位に分割した上で各部位に対応する臓腑の健康状態を自動的に診断する 5 分割診断法および舌画像を裂紋舌画像と非裂紋舌画像の 2 種類に分類した上で裂紋舌画像から酷い裂紋状態と軽い裂紋状態を識別する自動診断法を提案するものである。

これらの診断法では、深層学習を舌診に活用することで、以下のような利点が期待される。

- (1) 高度なパターン認識：舌の状態や特徴を画像データとしてアンサンブル学習に用いることで、微細な変化や特徴をより正確に認識できるようになる。
- (2) 客観的な診断支援：人間の主観的な要素を減少させ、より客観的かつ統一的な舌診の実現が可能となる。

- (3) 大規模データの効果的な利用：舌診のデータは大量かつ多様であり、深層学習はこれらのデータから潜在的な傾向やパターンを発見しやすくするため、医学的知識の拡充や病気の理解が進む可能性がある。
- (4) 複数の特徴の統合：舌診は舌の色、形状、舌苔など、さまざまな特徴を総合的に評価する。一方、深層学習は複数の特徴を統合して分析する能力を持っており、これによってより包括的な診断が行えるようになる。
- (5) データベースとの連携：深層学習を用いた舌診の結果はデータベースと連携することができ、大規模な舌診データと統合して分析することが可能である。これにより、臓腑の健康状態や舌の種類を把握し、診断の精度向上や新たな知見の発見が期待される。
- (6) 医療人材のサポート：深層学習を活用した舌診は、医療従事者の診断業務をサポートする一方で、経験豊富な専門家の知識を補完する役割も果たすことができる。
- (7) 疾患の早期検出：舌の微細な変化や特徴の異常を検知し、病気の初期段階での注意喚起や診断が可能になる。これにより、日頃の臓腑健康状態を管理でき、治療を早い段階で実施することにより、病気を未然に防ぐことが可能となる。

第2章 研究に関する準備

2.1 東洋医学における舌診

2.1.1 舌診とは

舌診は、伝統的な東洋医学における四診の中の視覚による望診に属しており、舌の状態を観察することで、舌象(舌の色・つや・形態という形象を指す)を認識して、人体の健康状態や疾患の情報を診断する手法である。また、舌象は舌診の主要な観察対象であり、人体の肉眼で観察された舌表面のことを指し、舌象の観察は舌質(舌体とも呼び、舌の肉質を指す)、舌苔(舌の表面についている苔)と舌下絡脈の3つの部分に分けて順次に行われる。具体的に舌質の観察は舌神(臓腑機能の状態が舌の上に反映したものを指し、舌の柔軟性と動きは舌質のつやに現れる)、舌色(舌質の色)、舌形(舌の形状)、舌態(舌の外観や状態)を含み、人体の気血津液や臓腑機能の健康状態を表す。舌苔の観察は苔色(舌表面に付いている苔の色)と苔質(舌苔の質感や特性)に分けられ、病気の性質や具合を反映するとされており、舌下絡脈の観察は気血の循環状態を見分けることができる。なお、一般的に、舌色が淡紅色であり、異常な舌形(齒痕・光滑・裂紋・嫩・瘦・小・短・芒刺・斜・卷・胖・大・舌瘡・舌腫など)がなく、舌質の動きが自由で苔色が白く、かつ舌苔の形が薄い状態の舌を健康な人の舌診と言われている [17,18].

舌診の歴史は古く、その起源は今からおおよそ3000年程前の殷商時代に遡り、遺跡から発掘された甲骨文に舌診の関連文章が記載されていた。そして紀元前5世紀から紀元前4世紀にかけて、「中国古代四大名医」の1人である扁鵲が初めて舌診に関する記録を残した。その後、舌診は西周時代から春秋戦国時代と秦漢時代の時を経て、徐々に発展して行き、東洋医学の舌診基礎として固められた。特に春秋戦国時代では、舌診は中国の伝統的な医学体系の一環とし

て医学書の『黄帝内経』に言及され、東洋医学の舌診の形成と今後の発展のために堅実な基盤を築き上げた [19].

そして漢唐時代では、舌診は『黄帝内経』における理論に基づいて、臨床実践の中で繰り返し検証を行い、着実に発展して行った。中でも舌診の発展に大きな貢献を果たしたのは、張仲景や華佗（外科医）などを中心とした医学者達である。また、張仲景が執筆した『傷寒論』は、東洋医学の歴史において、最も影響力のある書籍の 1 つであり、舌診を東洋医学の真髄となる「弁証論治」の重要な部分として取り入れた。弁証論治は、四診により病人の各症状を把握し、総合的に分析して証候を判断し、治療原則に随って治療を行うことを指している。なお、張仲景は『黄帝内経』に書かれている大量な舌診の理論を臨床実践に応用し、舌診の理論と弁証論治を密接に関連付け、弁証の指標と論治の根拠の 1 つとさせた。張仲景は『傷寒論』の他に、『金匱要略』も執筆された。この医学書では、舌形、舌色、舌苔などの要素が疾患の診断にどのように関連しているかを詳細に説明されている [17, 19].

その後宋金元時代では、東洋医学が日々発展して行くのに伴い、舌診の理論と臨床実践がさらに多くの新たな進歩を達成した。特に『敖氏傷寒金鏡録』は敖氏が執筆した中国における初の舌診を研究対象とした専門書であり、舌診の発展史上の重要な節目と言える。この医学書は舌診学の基礎を固めたと同時に、更なる発展を促進し、そして明清時代の「温病学」発展に大きな影響を与えた。例えば、張登が執筆した『傷寒舌鑑』や、申斗垣が執筆した『傷寒觀舌心法』など多くの舌診専門書は『敖氏傷寒金鏡録』に基づいて、臨床実践を積み重ねた上で発展してきた医学書である。従って、明清時代から民国時代における舌診の発展は壮観であり、東洋医学の進歩に重要な寄与をしたと言える [19].

さらに近代から現代に入ると、東洋医学は科学の発展に伴い、舌診に対する研究がより深まり、舌診の客観的な評価や規範化に関する研究が行われ、多くの成果を収めた。

このように東洋医学は中国国内で数千年にわたり発展し続けてきたが、時間の経過とともに周辺のアジア諸国にも伝播していった。日本、韓国、ベトナムなどの国々では、東洋医学の原則を受け入れつつ、それぞれの文化や環境に合わせた独自の医学体系が形成されるようになり、舌診もこれらの国々で受け継がれ、東洋医学という包括的な概念の一部として認識されるよう

になった。

東洋医学の歴史を有するアジア諸国では、現在、西洋医学と東洋医学の二本立ての体制で、東洋医学も正規の医療として医療政策に取り込まれている。正規の医師として、近代西洋医学の医師と東洋医学の医師が共存する医療体系である。また、各国の政府機関の中に東洋医学の担当部局が設置されており、国立の研究機関や病院も設立され、伝統医学専門大学や欧米の大学や研究機関と連携して研究活動を行っている [5,20]。

総じていえば、舌診は東洋医学の一部として始まり、次第に国際的にも認識されるようになった。東洋医学の一環として、舌診は今日でも広く用いられ、患者の診断と治療に重要な役割を果たしている。

2.1.2 舌表面と五臓六腑の関係

東洋医学では、人体の内臓を五臓六腑として認識されている。五臓には肝・心・脾・肺・腎が含まれており、六腑には胆・小腸・胃・大腸・膀胱・三焦が含まれている。東洋医学における臓腑は、単に器官そのものを指すのではなく、それらの働きによって生じる様々な現象も含まれている。五臓は、飲食物の栄養成分から、人体の代謝物質である気・血・津液を生成し、それらを貯蔵する袋状の器官と考えられている。六腑は、飲食物が連続して送られる1本の消化管とされている [2,21]。

東洋医学において、五臓六腑は以下に示す各機能を有すると考えられている [1]。

- 肝は、血を貯蔵し、全身への配分を決定する。全身の気の働きを調節する。
- 心は、全身に血を送り、精神や意識もつかさどる。
- 脾は、消化・吸収をする。栄養分と水分を全身に送る。栄養分や臓腑を持ちあげる。
- 肺は、呼吸をする。全身の水分の巡りを調節する。気や津液などを押しあげる、押しさげる。
- 腎は、精を貯蔵し、必要に応じて配分する。水分の代謝を調節する。肺が吸い込んだ気を、腎まで下げる。
- 胆は、胆汁を貯蔵し、必要があれば排出して消化を助ける。

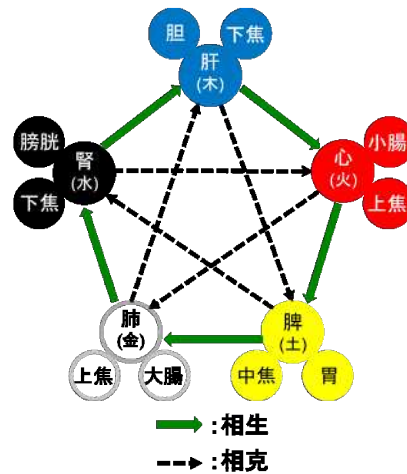


図 2.1: 五臓間の五行説と臓腑間の表裏関係

- 小腸は、胃から送られてきた飲食物を栄養分と不要物に分ける。
 - 胃は、飲食物をざっと消化する。脾の働きを助けて消化・吸収を調整する。
 - 大腸は、小腸から送られてきた不要物から水分を吸収して、大便を排泄する。
 - 膀胱は、腎から送られてきた水分（尿）を貯めておき、腎の決定に従って排泄する。
- 三焦は、特別な腑であり、詳細は後述する。

東洋医学では五臓の相互関係を五行説に当てはめる考え方がある。すなわち、五臓の肝・心・脾・肺・腎を、それぞれ木（もく）・火（か）・土（ど）・金（きん）・水（すい）の5つの要素に対応付けて表現する。また五臓間には相生と相克との関係がある [2,21]。各要素が五臓六腑に対応する理由は、以下のように考えられている [2]。

- 肝は、木が成長するように、気を体内に柔軟に循環させる。
- 心は、体を炎のようにぽかぽか温める。
- 脾は、万物を生み出す土のように、栄養分を生成する。
- 肺は、金属のような収斂効果で、気や津液を下方に引き寄せる。
- 腎は、水が低い場所に流れるように精気を蓄え、水分を調節する。

相生の関係とは相手を生み育てる母子関係のことであり、相克の関係とは相手を抑制する関係と考えられている [2,21]。具体的な相生相克の関係は以下の通りである [2]。

- 木生火は、木同士が擦れ合って火を起こし、さらに木を加えると火が勢いが強まる。
- 火生土は、火が燃えると灰と土が生成される。
- 土生金は、地下（土の中）には鉱物（金属）が含まれている。
- 金生水は、金属は水を滋養し、その表面に水滴を生じさせる。
- 水生木は、水が土地に湿潤さや栄養をもたらし、それによって木が成長する。
- 木克土は、木は根を張るために土を押し分け、土地を破壊する。
- 土克水は、土は水を吸収しやすく、水分を保持する性質があるが、同時に土が多すぎると水が滞り、水の流れが阻害される。
- 水克火は、水は火を消す性質がある。
- 火克金は、火は高温で金を溶かし、融かす性質がある。
- 金克木は、金は斧や鋸などの金属の道具を使って木を切り倒すことができる。

以上より、五行説における相生の力は、「木 → 火 → 土 → 金 → 水 → 木」の順で働きを促進し、相克の力は、「木 → 土 → 水 → 火 → 金 → 木」の順で働きを抑制すると考えられている [2,21].

東洋医学においては、相生相克の関係が均衡状態を維持している時、健康が保たれた状態であるとされる。従って、この関係は、どちらか一方が強すぎても弱すぎても健康は保たれない。五臓間の相互関係から、病気の原因や治療方法を考えることもある [21,22].

また、臓腑間には表裏関係が存在する。これは、経脈の流れによる臓と腑の密接な関係である。腑は表であり、臓は裏である。臓と腑は、互いに強く結びついているため、一方が変調を起こすと、もう一方も影響を受けやすくなる。具体的な結びつきは、肝と胆、心と小腸、脾と胃、肺と大腸、腎と膀胱である [21,23]. 図 2.1 は五臓間の五行説（相生相克関係）と臓腑間の表裏関係を表したものである。

なお、腑の1つである三焦は、具体的な臓腑ではなく、臓腑を含む胸腹部の一部である。古くから伝わる東洋医学の文献である『皇帝内径（こうていないけい）』（[24]- [28]）に初めてこの言葉が出現し、『難経』（[29]）では「有名無形（ゆうめいむけい）」（名のみで、実質がない）とされている。また『中蔵経』では、三焦は、すべての臓腑を総領し、各臓腑を協調させて、それぞれの正常な生理機能を発揮させることで、人の健康を保っていると記載されている [30].

つまり、東洋医学における三焦の役割は、五臓六腑のバランス状態を維持することで、これにより人体の健康が保たれると考えられている。三焦は上焦（じょうしょう）・中焦（ちゅうしょう）・下焦（げしょう）から構成される。上焦は横隔膜（おうかくまく）の上部であり、心と肺も含む。中焦は横隔膜の下から臍（へそ）までの腹部であり、脾を含む。下焦は臍より下の部分であり、肝と腎を含んでいる [31,32]。

東洋医学において、五臓六腑が健康であれば、人体は健康であると考えられている [33]。また、五行説に基づく五臓同士の関係が主要であり、表裏一体となる臓腑間の関係はそれを補うものであると考えられている [1,23]。

東洋医学では、人体の生命活動を維持するために4つの基本要素である気・血・津液・精が必要とされており、それぞれお互いに体内でバランスよく循環することにより、臓器や組織が正しく働き、病気にかかりにくくなるとされている。また、人体生命活動のエネルギー源となる気は血と巡り会い、形成した通路が経絡である。経絡を通して、舌と内臓が繋がり、舌表面の各部位と各臓腑が関連している。よって、東洋医学において、舌は直接的あるいは間接的に多くの臓腑と繋がっている。一般的に舌表面は舌尖部、舌辺部（左と右）、舌中部、舌根部の5つの部位に分けられ、それぞれ心肺、肝胆、脾胃、腎に対応する。舌の各部位の変化は、内臓の気血の盛衰が舌に影響して現れる症状と関連していると考えられている。よって、舌の表面の変化を捉えることで内臓の状態の変化を把握することができる [9]。図 2.2 は舌表面臓腑分布を示したものである。

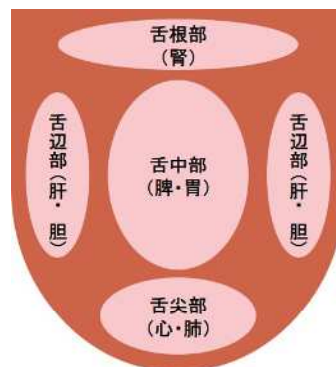


図 2.2: 舌表面の部位と臓腑の対応

2.1.3 裂紋舌

舌診では、裂紋舌を舌質（舌の肉部分）の表面が割れている状態として定義されており、臓腑の潤いが足りていない病理状態（運動不足・栄養不足）を示唆するとされている。また、裂紋舌の亀裂にはその深さと位置によって症状の度合いと病気のある臓器が異なる。例えば、裂紋舌の亀裂が左側の舌辺部から右側の舌辺部まで横一列に広がっている場合、舌辺部と舌中部に対応する肝・胆と脾・胃に病気が存在すると言える。なお、亀裂が深ければ深いほど、酷い症状を示し、浅い亀裂は軽い症状を示している [9]。

裂紋舌は、口腔内の乾燥や粘膜の状態によって引き起こされることがあるが、その原因は多岐にわたると考えられている。以下に一般的なものを示す [34]。

- 栄養不足：特定のビタミンやミネラルの不足（例：ビタミン B 群や鉄分）が口腔内の健康に影響を与え、裂紋舌の原因となる場合がある。
- ストレス：ストレスや精神的な負担が高まると、口腔内の状態に影響を与える場合があり、裂紋舌が現れることがある。
- 疾患や病気：特定の疾患や病気（例：口内炎や乾燥症候群など）が裂紋舌を引き起こす場合がある。
- 薬物の副作用：特定の薬物や医薬品の副作用として口腔内の異常が起こる場合がある。
- 加齢：加齢に伴って、口腔内の粘膜や唾液の分泌量が減少し、裂紋舌が現れることがある。

従って、裂紋舌は単一の要因によるものではなく、様々な要因が影響する可能性があると言える。そのため、裂紋舌が現れた際には、その原因を特定し、適切な対処法を見つけるために医師や歯科医に相談することが重要である。

2.2 深層学習

2.2.1 人工知能，機械学習と深層学習の関係

「人工知能 (AI: Artificial Intelligence)」という言葉は、1956 年に米国の計算機科学研究者である John McCarthy によって初めて命名された用語である [35]。AI を広い意味では、人間のような知能をコンピューターに持たせる技術のことを指しており、細かく分けると以下のレベル 1~レベル 4 の 4 つの段階に分けられる [36]。

- レベル 1: 単純な制御プログラム。これは単純な論理回路を含む制御工学・システム工学を指しており、昔からある歴史の長い分野である。
- レベル 2: 古典的な AI。これはデータの入出力が多彩であり、推論・探索などが内部で行われている。
- レベル 3: 機械学習を取り入れた AI。これはプログラム自体が作業 (タスク) を通して学習して行く仕組みである。
- レベル 4: 深層学習を取り入れた AI。これは機械学習の中で、データの特性や特徴となる特徴量を学習して行く最新の AI モデルである。

また、AI の歴史として、現在までに大きく 3 回の AI ブームが存在している。まず第一次 AI ブームが来たのは 1950 年代後半から 1960 年代である。この時期に「人工知能」という言葉が誕生したが、当時の AI は推論や探索といった今まですでにあつた手法を組み合わせた非常に簡単なものであり、現実社会の課題を解決するに至らず、ブームは終結した。次に第二次 AI ブームでは、1980 年代にインターネットが登場し、個人用の PC・ソフトウェアである Macintosh や Windows が出始めた。しかし、AI に必要な大規模データやテクノロジーがまだ不足している問題、コンピューターの性能の問題、コンピューターに学習させる手法の問題により、ブームは終了した。そして 30 年後、それらの問題は解決され、さらに、革新的なアイデアと共に起きたのが第三次 AI ブームである。この時期では、インターネット技術の発達と成熟により、データを保存するクラウドの登場やビックデータ (大規模なデータ) の活用が可能になり、さらにテクノロジーの進歩と、革新的なアイデアである深層学習の登場によって、実現不可能

と思われてきた自動運転、自然な自動翻訳、汎用的なタスク処理が可能になった [36].

機械学習は、コンピューター上で動作するアルゴリズムの研究分野であり、コンピューターにデータを読み込ませ、データ内に潜む有用なパターンや規則性などを学習させることで、未知のデータに対する予測や判断を行うためのアルゴリズムを獲得することを可能にするデータ解析技術である。また、機械学習における学習は、コンピューターが手元にあるデータから何らかのパターンや規則性を見つける作業を意味する。機械学習は人間の能力を超えるデータの処理が可能であり、人間の見つけることのできないパターンや規則性を見つけることができるため、近年では非常に注目されている技術である [37].

機械学習には、教師あり学習、教師なし学習、強化学習の 3 種類があると言われている。教師あり学習は、入力する学習データに推論の正解となる情報を付与した上で、学習させる方法である。教師なし学習は、学習データに正解情報を付けない状態で行う学習である。例えば、犬と猫の各画像に「犬」と「猫」の正解情報を付けた上で学習させるのが教師あり学習であり、犬と猫の各画像に正解情報を付けずに学習させるのが教師なし学習である。この例では、教師あり学習の結果、犬か猫かを予測するモデルを構築することができ、未知の画像に対して、それが犬なのか猫なのかを予測できるようになる。教師なし学習の場合は、与えられた画像について、特徴が似ている画像同士でグループ分けすることになる。ただし、この際にコンピューターは各グループに犬か猫かの名前を付けることはできないため、各グループの意味付けや結果の解釈は人間がする必要がある。また、教師あり学習も教師なし学習も、ただ単に学習データを与えるのではなく、機械学習のアルゴリズムも人間が指定する必要がある [37].

教師あり学習には、回帰とクラス分類があり、いずれも入力情報と正解情報を与えて、予測モデルを構築する必要がある。また、予測する答えが連続的な数値データの場合を回帰と呼び、非連続なカテゴリーカルデータの場合をクラス分類と呼ぶ。なお、回帰の代表的なアルゴリズムには、線形回帰や回帰木（決定木）などがある。線形回帰は単純な線形関数による回帰モデルであり、回帰木は非線形のより複雑な回帰モデルを構築することができる。一方、クラス分類には、2つのクラスの確率を予測する 2 クラス分類と、3つ以上のクラスの確率を予測する多クラス分類があり、通常は予測確率をもとに予測値のクラスを決定する。代表的なアルゴリズム

ムとして、ロジスティック回帰や分類木（決定木）などがある [37].

教師なし学習には、学習データの類似度を発見させ、その類似度に基づいて、グループ分けを行う「クラスタリング」や、学習データから重要な情報を抽出し、可能な限り情報を保存して、高次元のデータを低次元に変換する「次元削減」などがあり、いずれも学習データを与えて、何らかの規則性を発見させる仕組みとなっている。また、クラスタリングと次元削減の各々の代表的なアルゴリズムとして、k-means 法と主成分分析などがある [37].

強化学習は、報酬という成果を最大化することを目的として、試行錯誤の学習を繰り返す学習方法である。その実例としては、家庭用ゲームの AI 機能、車の自動運転、google のロボットアームなどで活用されている [37].

深層学習は、教師あり学習や教師なし学習や強化学習と同じく、機械学習に属する 1 種の手法であり、人間の脳神経回路を模倣した学習である。人間の脳神経回路は、ニューロンという神経細胞の集まりから構成されており、ニューロンは刺激を受けると興奮し、その刺激を他の細胞に伝播する仕組みとなっている。また、これらの刺激は電気信号としてやり取りされ、人間の記憶や思考となっている。そしてこれらの脳神経回路に基づいて、研究者たちが数理モデルとしてコンピューター上で再現したものをニューラルネットワークと呼び、その構成要素であるニューロンが数理モデル化されたものをパーセプトロンと呼ぶ。図 2.3 に示すように、ニューラルネットワークの構造には、入力層、中間層、出力層があり、中間層の数を増やすと、より複雑な問題を解くことができ、より賢い AI が作れるようになる [38].

図 2.4 に示すように、人工知能の定義が最も広い意味を含んでおり、その中の専門領域では、学習データに対して、予測や判断を行うアルゴリズムを獲得する技術として機械学習があり、さらに機械学習の 1 種の手法として深層学習が含まれている。総じて言えば、これらは「人工知能 ⊃ 機械学習 ⊃ 深層学習」という包含関係にある。

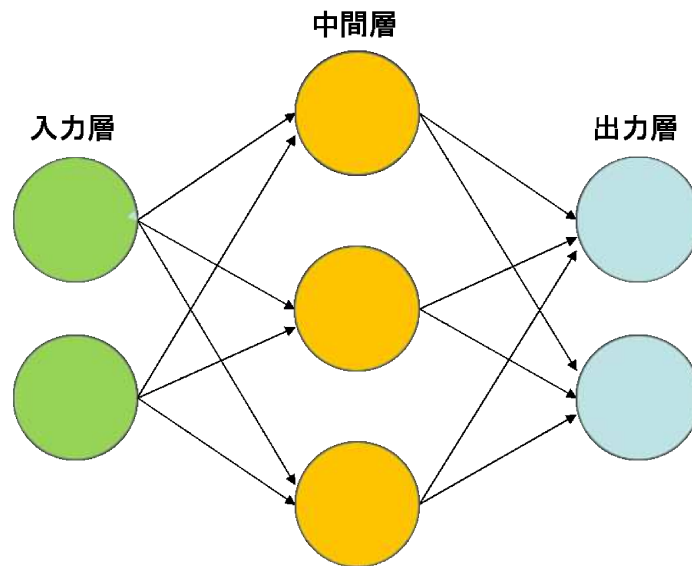


図 2.3: ニューラルネットワークの構造

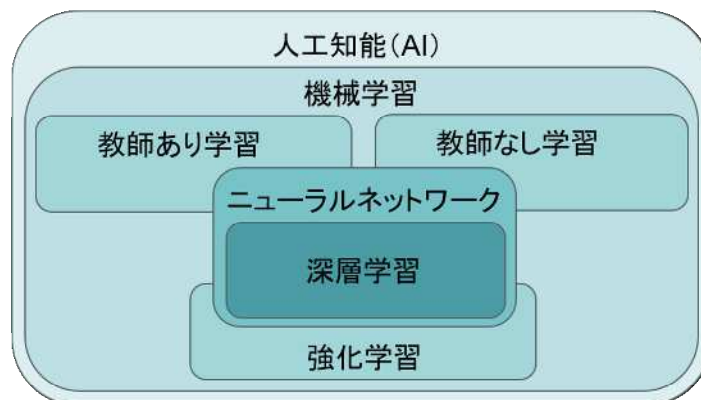


図 2.4: 人工知能, 機械学習と深層学習の関係

2.2.2 深層学習を舌診に活用した先行研究

人工知能の大いなる発展につれ、東洋医学における知能化も進化してきた。近年、深層学習アルゴリズムの代表となる畳み込みニューラルネットワーク（CNN：convolutional neural network）が複数の研究分野で重要な役割を果たしていると同時に、東洋医学の知能化発展にも多大な貢献をもたらした。

CNN は主に画像認識の分野で用いられる深層学習アルゴリズムであり、データを学習することにより入力画像から特徴量を抽出し、それらを区別することができるようになる。また、CNN はデータを入力する入力層から始まり、次にデータ内の特徴を抽出する畳み込み層と抽出された特徴を集約するプーリング層の 2 つがセットで繋がり、そして出力を 1 次元にする役割を持つ全結合層と出力層が続く構造となる。なお、畳み込みは、画像データをうまく取り入れる方法のことを指しており、画像は複数のピクセルによって表現されている。例えば、 28×28 のピクセルで表現された画像があり、それを 784 の一次元のデータとしてインプットする際には、本来持っているはずの重要な画像の情報を失う可能性がある。そこで畳み込みという方法を用いることで、画像データの情報をなるべく失うことなく取り込むことができる。CNN を利用した画像認識技術の応用例としては、無人レジ店舗、ドライブレコーダーを活用した物体検出、医療における画像診断などがある [38]。

また、舌診の客観化を目的とした研究は、舌の状態を観察・分析し、人体における五臓六腑の健康状態や疾病の診断・予防に活用する研究分野であり、東洋医学から数十年の発展を経て、現在では舌画像の特徴量を抽出して疾患の早期診断や予防に役立つモデルの開発などにおいて、多くの研究成果が得られている。舌診の客観化で最初に試みた研究では、主に舌象の分類を定量化する研究などが見られたが、近年の機械学習技術の発展に伴い、舌診の客観化研究は臨床においてより身近になり、研究者たちは舌診を通じて病気の診断と予防を直感的に理解することができるようになった。

例えば、Zhang らは、サポートベクターマシン (SVM：support vector machine) を活用した舌画像に基づく糖尿病の自動診断法を開発する研究を行い、その結果では 78.77% の正解率で糖

尿病を自動診断することができた [39]。この研究では、東洋医学における舌画像に基づく糖尿病の診断方法が極めて貴重であり、舌診の自動化を実現させる可能性を示している。舌形の自動認識に関する研究では、Li らは歯痕舌の自動認識を目的として、まず歯痕舌の画像 297 枚と非歯痕舌の画像 344 枚を収集する。次に、歯痕舌の画像を識別させるために、歯痕舌の凹型領域の情報を人工的に表示させる。その後、CNN を歯痕舌の特徴抽出に活用して、歯痕舌と非歯痕舌の 2 クラス分類を行う。その結果では 72.7% の正解率で歯痕舌を自動認識することができた [40]。舌苔の自動識別に関する研究では、肖らは CNN を活用して、舌苔を薄白苔、白厚苔、薄黄苔、黄厚苔、灰黒苔の 5 種類に多分類するモデルを提案した。その結果では 94.85% の正解率で舌苔を 5 種類に自動識別することができた [41]。舌色の自動識別研究については、Qi らはアンサンブル学習の 1 種であるランダムフォレスト手法 [42] を活用して、728 枚の舌画像を淡色、淡紅色、紅色、紫紅色、紫色の 5 種類の舌色に多分類するモデルを提案した。その結果では 78.81% の正解率で舌色を 5 種類に自動識別することができた [43]。舌象の自動識別研究については、胡らは東洋医学における 9 種類の舌象の内、陰虚体質と陽虚体質の 2 種類に対して、Python 上で画像処理を行うためのライブラリである PIL (Python Imaging Library) を用いて画像解像度の縮小を行い、次に CNN の軽量化・高速化のブレイクスルーの 1 つとなった MobileNets で 2 クラス分類するモデルを構築した。その結果では 70.5% の正解率で舌象を 2 種類に自動識別することができた [44]。なお、正解率については、分類問題における評価指標であり、詳細は後述する。

また、これまでに舌画像の自動認識に関連する研究として、舌を自動認識するための TongueNet の提案 [10]、舌の観察結果に基づいたファジィ処理の研究 [11]、舌領域からの舌苔領域のみの抽出に関する研究 [12]、舌画像の正解率向上のためのシステム開発 [13] などが行われている。また深層学習を用いた舌診に関する研究としては、新型コロナウイルス感染症 (COVID-19) の診断支援のために CNN 技術を活用して舌苔の種類を自動認識する GreasyCoatNet モデルを開発する研究 [14] や、Faster R-CNN 手法を活用して、裂紋舌 (れつもんぜつ)、歯痕舌 (しこんぜつ)、お点舌 (おてんぜつ)、お斑舌 (おはんぜつ)、膩苔 (じたい)、剥落苔 (はくらくたい)、腐苔 (ふたい) の 7 種類の舌画像を分類し、それぞれの舌象 (舌の変化) の特徴と性別、年齢、

代謝性疾患との関係を解明する研究 [15] および、舌画像解析における舌領域分割と舌特徴抽出の定量化および自動化のため、ソフトウェアを開発する研究 [16] などが試みられている。

しかしこれらの研究は、東洋医学の舌診でいう五臓六腑の健康状態と裂紋舌の裂紋状態を深層学習を用いて自動診断するまでには至っていない。

本研究は、五臓六腑の自動診断技術と裂紋舌に基づく裂紋状態の自動診断技術の実現を目指して、画像に含まれる舌を自動で認識して切り抜き、舌画像を 5 つの部位に分割した上で各部位に対応する臓腑の健康状態を自動的に診断する 5 分割診断法および舌画像を裂紋舌画像と非裂紋舌画像の 2 種類に分類した上で裂紋舌画像から酷い裂紋状態と軽い裂紋状態を識別する自動診断法を提案するものである。

2.2.3 Mask R-CNN 手法と 6 つの画像認識モデル

文献 [10] では、提案の TongueNet の有効性を定量的に評価するために、Precision, Dice, mIoU, FPR, FNR および ME の 6 つの評価指標を設定し、3 セットのテストデータを用いて FCN, U-Net, SegNet, Deep-Tongue, Mask R-CNN [45] および TongueNet という 6 つの方法の比較評価を行っている。実験結果より、3 つのテストセットにおいて、TongueNet は FPR 以外のすべての評価指標で最も高い性能を示している。Mask R-CNN については、Precision は次善であり、それぞれ 96.97%, 96.67%, 97.32% という高い正確さを示しており、Precision 以外の評価指標においても TongueNet に近い評価値を得ている。

しかし、TongueNet はソースコードが未公開のため、本研究では、舌の自動認識を進めるに当たり、性能も高く、ソースコードも公開されている機械学習の Mask R-CNN 手法を活用する。Mask-R-CNN とは Microsoft の研究グループが開発した深層学習のモデルであり、一般物体検出と Segmentation (画像の領域分割) を同時に行うマルチタスクモデルである。これにより、入力された画像に対して、決められた種類の物体を認識してその位置とクラス名を特定し、なおかつ画像中のピクセル単位で物体の写っている領域を特定することができる。なお、本研究では舌を含む画像に対して、Mask R-CNN 手法で画像認識モデル ResNet50 [46] による物体検出・領域分割を行う。

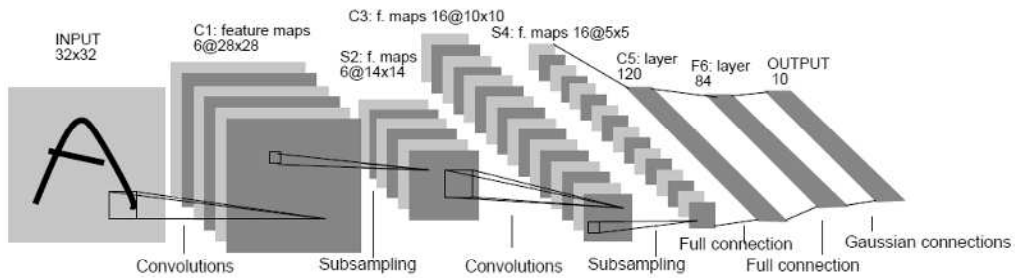


図 2.5: LeNet の構成 (文献 [47] の Fig.1 より引用)

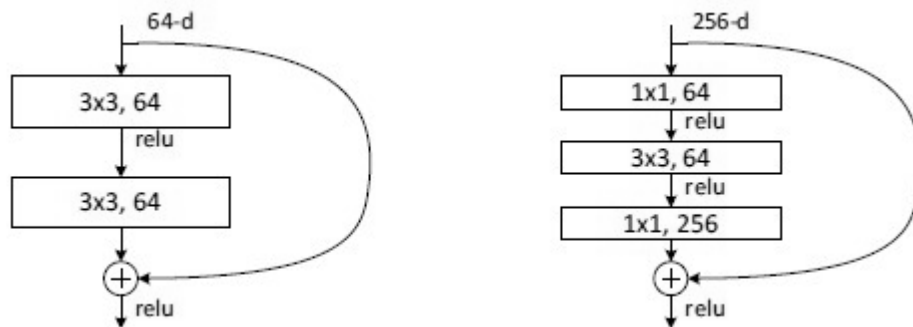


図 2.6: Shortcut connection の仕組み (文献 [46] の Fig.2 より引用)

本研究はまず五臓六腑の自動診断技術の実現を目指して、舌画像の自動認識・切り抜きによって舌の各部位に対応する全ての臓腑の健康状態を診断するという発想に基づいて実施したものであり、その第一歩として、古典的な LeNet [47]、現在でも広く用いられている ResNet や DenseNet [48]、および比較的新しく、かつ高性能な EfficientNet [49] による画像分類の深層学習を行う。次に、裂紋舌に基づく裂紋状態の自動診断技術の実現を目指して、裂紋舌の自動認識と裂紋状態を判定するために、画像認識モデルである LeNet, ResNet50, ResNet101, DenseNet169 および ConvNeXt-Tiny [50] による画像分類の機械学習を行う。

LeNet は CNN の元祖であり、LeCun らが手書き数字認識を行うネットワークとして 1998 年に提案したものである。図 2.5 に示すように、LeNet は畳み込み層とプーリング層を繰り返して行い、最後に全結合層を経て結果が出力される構造となっている [47]。

ResNet は画像認識分野で最も成功を収めたモデルの 1 つであり、2015 年に Microsoft Research から発表され、その年に画像認識モデルの精度を競う著名なコンペティションである ImageNet

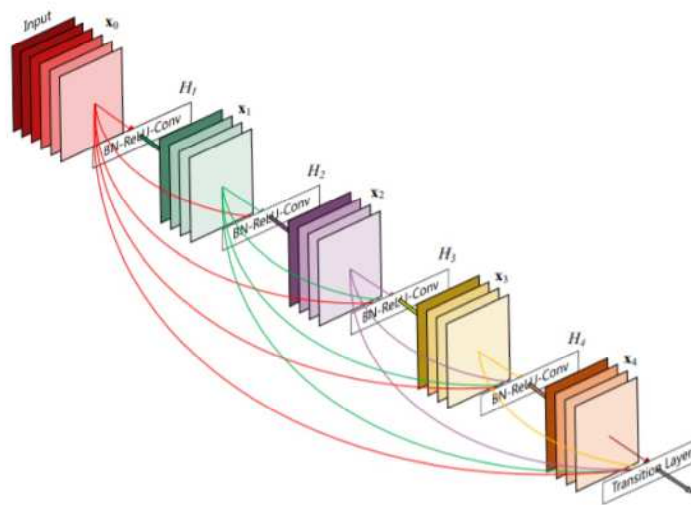


図 2.7: DenseNet の構成 (文献 [48] の Fig.1 より引用)

Large Scale Visual Recognition Challenge (ILSVRC) にて優勝している [46]. ResNet の特徴は, CNN 構造を多層化した際に学習が難しくなる問題 (勾配消失問題) を解決するための仕組みとして shortcut connection を導入したことであり, その結果 152 もの層を重ねて学習ができた. 図 2.6 に示すように, shortcut connection の仕組みは手前の層から来た入力をいくつかの層を飛ばして, その後に計算された値と足し合わせるものである. これによって, 飛ばされたいくつかの層は手前の層からの入力との誤差 (residual) のみを予測することで効率よい学習ができることから ResNet と命名されている [46].

DenseNet は ResNet を改良したモデルである. ResNet がある層に与えられた信号をそれよりも少し上位の層の出力に追加する shortcut connection により深いネットワークを学習できるようにしたのに対し, DenseNet では, 図 2.7 に示すように前方の各層からの出力すべてが後方の層への入力として用いられている. レイヤー間が密に (dense) 結合していることから DenseNet と呼ばれる. これにより, 勾配消失の削減, 特徴伝達の強化, 特徴の効率的な利用, パラメータ数の削減, 正則化効果の期待を実現している [48].

EfficientNet は 2019 年に Google が発表したモデルであり, ResNet よりも効率なパラメータ数で, 高い精度を目指したものである. 図 2.8 に示すように, EfficientNet はモデルの「深さ」

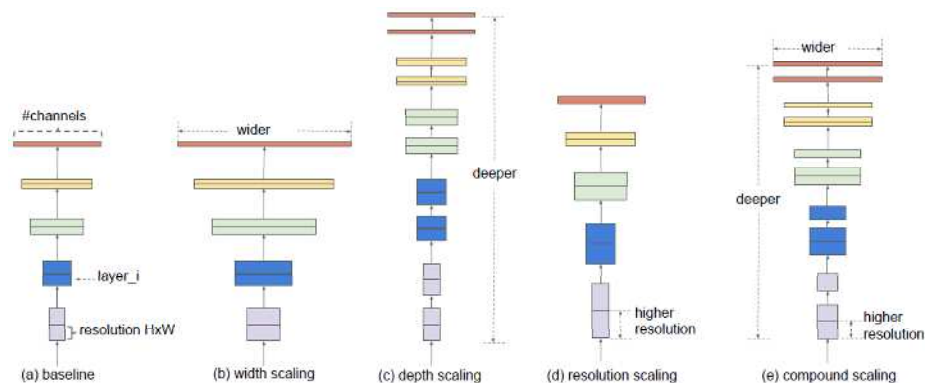


図 2.8: モデルのスケーリング (文献 [49] の Fig.2 より引用)

と「広さ」と「入力画像の解像度」の3つをバランスよく調整することで精度を向上させる構成となっている。なお、EfficientNet には B0 から B7 のモデルがあり、数字が上がるに連れてパラメータ数が大きくなり精度も向上する [49]。

ConvNeXt については、Swin Transformer など 2021 年代の ViT 系モデルの構造設計によって ResNet を書き換えたモデルとして 2022 年に提案されたものである。ConvNeXt のモデル構造は、必要に応じて容易に拡張や変更が行えるように設計されており、スケーラビリティが特徴的となっている。よって、提供されるデータが増えると正解率が向上する。なお、ConvNeXt-Tiny と ConvNeXt-Base はそれぞれ ResNet50 と ResNet200 をベースとして作られたモデルである。また、図 2.9 に示すように、ImageNet において、ConvNeXt-Tiny の正解率は 82.00% で、ResNet50 の 78.80% より大幅に改善されており、同等計算量の Swin Transformer Tiny の 81.30% よりも正解率が高いことが分かる。従って、ConvNeXt は画像分類において最適化され、より優れた結果を達成している [50]。

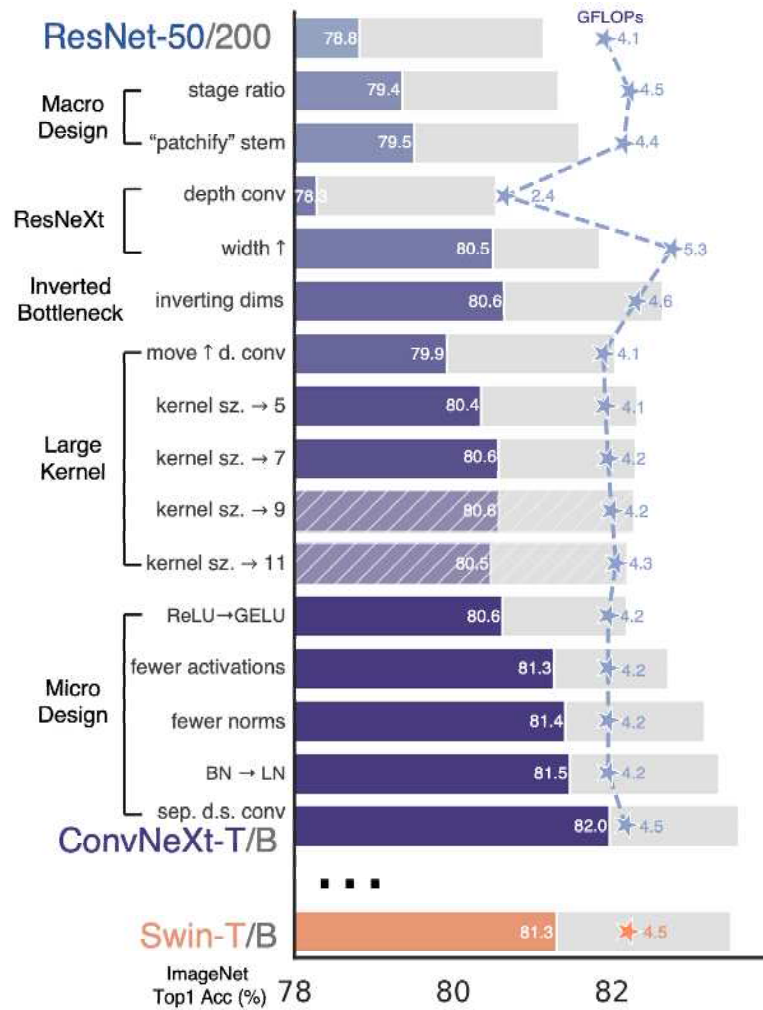


図 2.9: ResNet から ConvNeXt に変更する流れ (文献 [50] の Fig.2 より引用)

2.2.4 アンサンブル学習と水増し手法

本研究では、五臓六腑の自動診断における臓腑の健康状態を判定する正確さと、裂紋舌の自動診断における裂紋舌の自動認識と裂紋状態を判定する割合を高めるために、複数の画像認識モデルを融合したアンサンブル学習を行う。なお、裂紋舌の裂紋状態を判定する学習では、画像の枚数が極めて少ないため、今回はデータを拡張するという水増し手法を取り入れる学習を行う。

アンサンブル学習とは、「複数の機械学習モデルを組み合わせるタイプの機械学習ア

ルゴリズムのこと」[42]である。これにより、単独では精度の高くないモデルを多数用いることで高い精度をたたき出すことが可能となり、未学習のデータに対しての予測能力の向上が図られる。また、機械学習において、推定結果と実測値の差であるバイアス (Bias) と、推定結果の誤差であるバリエーション (Variance) はトレードオフの関係にある。例えば、モデルによる予測においてバイアスが大きい場合、そのモデルは入力と出力の関係性を正確に表現できていないと言える。いわゆる「学習不足」の状態である。一方、モデルの予測においてバリエーションが大きい場合、そのモデルは学習データのノイズまで学習しており、テストデータなど未知のデータでは正確に予測できないと言える。いわゆる「過学習」の状態である。そこでアンサンブル学習を活用することで、バイアスとバリエーションを適切なバランスに調整し、過学習を防ぎながら、高い予測結果を維持することができる。なお、アンサンブル学習の代表的なアルゴリズムとして、バギング (Bagging)、ブースティング (Boosting)、スタッキング (Stacking) がある [51]。

バギングはアンサンブル学習の中でも代表的かつ主流なアルゴリズムであり、並列的にモデルを構築し、それぞれのモデルの総合的な結果を用いるという手法である。バギングを活用するメリットとしては、学習時間を短縮できることや、バリエーションを抑えることを可能とするなどがある。また、クラス分類におけるバギングは、単に複数のモデルの予測確率の平均を取るのが一般的な方法であるが、複数作成したモデルの予測確率にばらつきがある場合、予測確率の高いモデルには大きめの重みを付ける加重平均のアンサンブル学習もある。クラス分類における加重平均をとる場合、複数のモデルの予測確率は 0 から 1 までの重みとともにそれぞれ乗算され、その後、各々で重み付けされた予測確率の合計値をもとに最終予測のクラスを決定する。なお、加重平均するアンサンブル学習において、以下のアプローチが考えられる [52]。

- 学習した複数のモデルを検証した各々の確率に基づいて、その確率が高いモデルには大きい重みを付け、低いモデルには小さい重みを付けることができる。
- 学習した複数のモデルに重み付けした予測確率の合計値が最も高くなるように最適化できる。

アンサンブル学習では、並列的に学習するバギングに対して、直列的に学習を行うブースティングがある。このアルゴリズムでは、最初のモデルでうまく推定できていない部分を補完する

ために、その部分の学習データに重みを付けて次のモデルで学習するという反復処理を行い、最終的には予測確率の高いモデルに重みを与えて、最終的なモデルを作成する。ブースティングの特徴としては、学習するモデルの数を増やすごとに、最終結果をより良いものに改善することができることや、バイアスを小さくすることが可能であるなどがある。また、一般的にブースティングの方がバギングより高い予測確率が見込めるが、最初から順番にモデルを処理していくため、バギングより学習時間がかかるとされている [42].

スタッキングの Stack は積み重ねるという意味であり、複数のモデルを積み重ねて、前のモデルの出力を特徴量 (1つの学習データ) として最終的なモデルを作成し予測する手法である。スタッキングのメリットとしては、モデルの多様性を確保できることや、それぞれのモデルの弱点を補うこと、安定した予測結果が得られることなどがある。しかし、計算コストがかかる点や採用するモデルによって予測性能が大きく変わってしまうことが難点とされている [51].

近年のニューラルネットワーク (特に CNN) による画像認識技術はとても進歩しており、開発されたモデルの中にはとても深い層をもつものもある。そのようなモデルの中には多くのパラメータを有するものもあり、学習のために大量のデータが必要とされる場合がある。しかし、データによっては必要な数を用意できない場合もある。特に医療系などのプライバシーに関わるようなデータは比較的集めにくい現状となっている。少ないデータ数でも学習を行えるようにするためにデータの拡張が必要である。しかし、ただ闇雲に増やせばよいというわけではなく、テストしたときにより良い予測確率を発揮するためにはどのような変化を加えるかも考慮する必要がある [38].

画像認識の機械学習において、データの拡張は、学習用の画像データに対して、変形させたり組み合わせたりと「変換」を施すことでデータの量と分布を増加させる手法である。この「変換」には、画像を上下左右に、指定された値の範囲で移動させる「上下平行移動」や「左右平行移動」、画像を大きくさせる「拡大」や小さくさせる「縮小」、画像を上下左右に反転させる「上下反転」や「左右反転」、画像を指定された角度に回す「回転」、他にも「平行移動」や「明るさの調整」、「背景の置換」、「ノイズの付与」など様々な種類がある。

本研究では裂紋舌の裂紋状態を判定する学習において、画像データに対して、画像処理ライ

ブラリである OpenCV の flip 関数による「上下反転」、「左右反転」および「上下反転+左右反転」の 3 種類の水増し手法を活用する。

2.2.5 2 クラス分類問題における評価指標

機械学習における評価指標は、学習させたモデルの性能やその予測値の良し悪しを測る指標である。本研究では、2 クラス分類問題において、学習させたモデルの性能を評価するにあたり、再現率 (Recall)、特異度 (Specificity)、正解率 (Accuracy) の 3 つの指標を取り扱い、評価を行う。また、2 クラス分類において、データを分類した際に、その正クラスと負クラスに属する数を整理しておく表のことを混同行列と呼び、混同行列から様々な評価指標を計算することができるため、分類問題には必須な要素である。この「正クラス」と「負クラス」には、それぞれ 2 パターンの意味がある。

例えば、動物の画像を見て、それがネコかどうか判断させるモデルを作ったとする。「正クラス」には「ネコの画像を正しくネコと判断できた」と、「ネコ以外の画像を正しくネコではないと判断できた」の 2 パターンがある。同様に、「負クラス」には、「ネコの画像をネコではないと判断した」と、「ネコ以外の画像をネコと判断した」の 2 パターンがある。このように、「正クラス」と「負クラス」には合計で 4 つのパターンがあり、これらを混同行列の指標とされている。なお、混同行列とその指標および本研究で取り扱う 3 つの評価指標をそれぞれ表 2.1、表 2.2 および表 2.3 に示す [37, 53]。

表 2.1: 混同行列

		モデルが予測したクラス	
		正 (Positive)	負 (Negative)
実際のクラス	正 (Positive)	TP(True Positive)	FN(False Negative)
	負 (Negative)	FP(False Positive)	TN(True Negative)

表 2.2: 混同行列の指標

指標	説明
TN(True Negative)	実際の負クラスを正しく負クラスと予測された数
FN(False Negative)	実際の正クラスを間違えて負クラスと予測された数
FP(False Positive)	実際の負クラスを間違えて正クラスと予測された数
TP(True Positive)	実際の正クラスを正しく正クラスと予測された数

表 2.3: テスト結果の評価指標

評価指標	意味	計算式
再現率 (Recall)	実際の正クラスの内, 正しく正クラスと分類出来た割合	$TP/(TP+FN)$
特異度 (Specificity)	実際の負クラスの内, 正しく負クラスと分類出来た割合	$TN/(FP+TN)$
正解率 (Accuracy)	予測クラス総数の内, 正しく分類出来たクラスの割合	$(TP+TN)/(TP+FP+FN+TN)$

2.2.6 深層学習による舌診の効果および意義

深層学習による舌診は、臓腑健康状態の自動診断や裂紋舌の自動診断にとどまらず、健康管理や医療の向上につながることを期待される。ここでは、その具体的な側面に注目し、深層学習による舌診の効果および意義を以下に示す。

- (1) 高度なパターン認識：深層学習を舌診に活用することで、豊富な舌の状態や特徴を画像データとして学習でき、これにより、舌の微細な特徴や病態のパターンをより正確に認識できるようになる。従って、深層学習による舌診は、医師が見逃す可能性のある細かい変化まで検知することが期待される。
- (2) 客観的な診断支援：医師による舌診は主観的な要素が含まれることがあるが、深層学習は客観的かつ標準化された評価を提供するため、医師の主観的な要素を削減し、より客観的かつ統一的な舌診の実現が期待される。
- (3) 大規模データの効果的な利用：舌診のデータは膨大かつ多様であり、深層学習はこれらのデータから潜在的な傾向やパターンを発見しやすくするため、医学的知識の拡充や病気の理解が進む可能性がある。
- (4) 複数の特徴の統合：舌診は舌色、形状、舌苔など、さまざまな特徴を総合的に評価する。一方、深層学習は複数の特徴を統合して分析する能力を持っており、これによってより包括的な診断が行えることが期待される。
- (5) データベースとの連携：深層学習による舌診の結果はデータベースと連携することができ、大規模な舌診データと統合して分析することが可能である。これにより、臓腑の健康状態や舌の種類を把握し、診断の正解率向上や新たな知見の発見が期待される。

- (6) 医療人材のサポート：深層学習による舌診は自動化されたプロセスであるため、医療従事者の診断業務をサポートし、作業負担を軽減させると同時に舌診の結果を素早く提供できる。従って、舌診はより効率的に活用され、患者への適切な対応が迅速に行われる。また、経験豊富な専門家の知識を補完する役割も果たすことが期待される。
- (7) 疾患の早期検出：舌の微細な変化や特徴の違いを検出し、疾患の早期段階での警告や診断が見込まれる。これにより、日頃の臓腑健康状態を効果的に管理でき、治療の早期介入や予防が期待される。

以上の(1)~(7)より、深層学習による舌診は、高度な技術力と大規模なデータ処理能力に基づいて、医師に比べてより客観的かつ効果的な医療診断が提供され、患者の健康増進に寄与することが期待される。

第3章 深層学習による臓腑健康状態の診断

本章では、五臓六腑の自動診断技術の実現を目指して、深層学習を活用した舌の5分割診断法について述べる。

まずは、舌診支援システムの構成と処理手順について概説した上で、5分割診断法における舌の自動認識と切り抜き、舌画像の5分割、そして、LeNet, ResNet50, ResNet101, DenseNet169, EfficientNet の5つの画像認識モデルを用いた臓腑状態判定のための学習について説明する。次に、臓腑状態判定のための学習のテスト結果を評価するために、5つの画像認識モデルを用いた時の舌診（舌画像分類）の性能を調べることで各モデルの評価を行う。さらに、そのうちの4つのモデルを融合させるアンサンブル学習を活用して舌画像を分類し、その結果を上記モデルの結果と比較することで、正解率の高い舌診支援システムの実現方法について考察する。

3.1 深層学習による舌の5分割診断法の提案

3.1.1 舌診支援システムにおける処理の流れ

本研究では、図3.1に示す舌診支援システムを構築することを目指している。このシステムは次の(1)～(3)の処理からなっている。(1)舌を含む写真が入力されると、図3.2の様に舌を認識して切り抜き、舌のみの画像データを生成する(図3.1②③)。(2)文献[18]で示された比率に従い、図3.3の様に切り抜かれた舌画像を5部位に分割し、各々の部位の画像データを生成する(図3.1⑤)。(3)5分割された舌の各部位の画像に対して提案する5分割診断法を適用して、対応する臓腑の健康状態を自動診断する(図3.1⑦⑧)。以降、3.1.2節および3.1.3節にて、(1)(2)の具体的な処理と(3)の臓腑状態を判定するための学習について述べる。

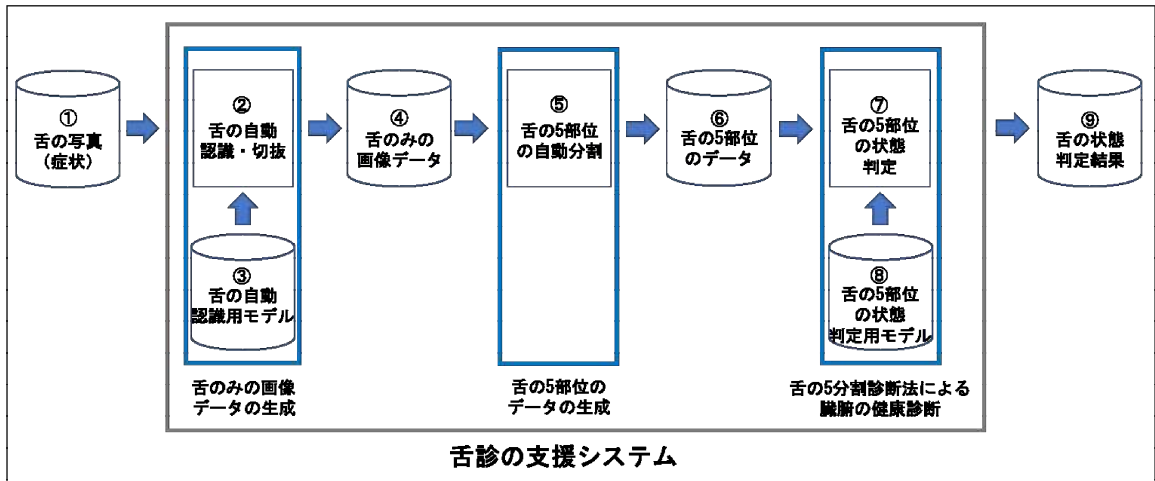


図 3.1: 舌診支援システムにおける処理の流れ

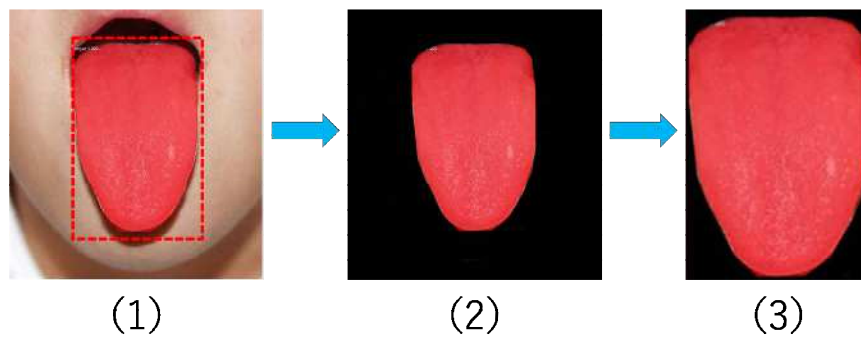


図 3.2: 舌の自動認識と切り抜き

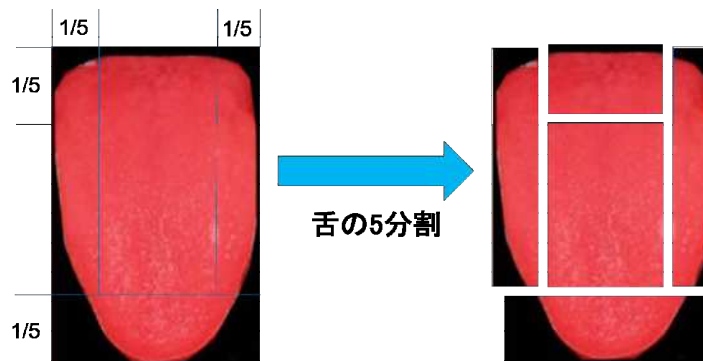


図 3.3: 舌画像の 5 分割画像

3.1.2 舌画像の自動認識・切り抜き・5 分割

学習データとテストデータとして、数多くの舌の画像が含まれる文献 [9, 18, 54–57] から舌の画像 1226 枚をスキャンして集めた。これらの画像は色深度が 32 かつ画像サイズが (幅 94 ピクセル×高さ 104 ピクセル) ~ (幅 1773 ピクセル×高さ 1404 ピクセル) となっている。学習データとテストデータの割合は 9:1 とし、それぞれの画像の枚数を 1104 枚と 122 枚にした。これらの画像に対し、アノテーションツールである LabelMe [58] を用いて舌の輪郭に沿って頂点を打つことで画像中の舌の領域を表す Polygons (多角形) を作成してラベル名は全て「tongue」とした。作成した舌のアノテーションデータは json ファイルとして保存しており、ファイルの中には Polygons の座標情報、画像データとラベル名が保存されている。

次に、Mask R-CNN 手法を利用して舌の自動認識のための学習とテストを行う。学習の際には自動認識に使用する「舌の自動認識用モデル」(学習の結果)の候補が複数生成されるが、その中で各 Loss の合計値 $Loss_{sum}$ が低いものを舌画像の認識に用いる。表 3.1 は各 Loss の値について文献 [59, 60] における説明をまとめたものである。

学習結果を評価するテストとして、Tensorflow のパッケージ Keras で提供されている事前学習済みモデルの ResNet50 [45] を用いて 1104 枚の舌画像を学習させた。その際、epoch 数 (一つの学習データを繰り返し学習する数 [61]) は “heads epoch 10, all epoch 30” と設定した。学習の結果として生成された 30 個のモデルのうち、最も低い $Loss_{sum}$ 値 (0.2120) を持つモデルを用いて、122 枚のテストデータを認識させて舌を正しく認識した割合を確認した。その結果、舌表面に「矢印」が写っている画像 7 枚、舌尖部の所に「楕円」が写っている画像 1 枚と舌根部の写り具合が暗い画像 1 枚 (合計 9 枚) を除く 113 枚の画像を正しく認識できた。従って、舌を自動認識できた割合は 92.62% であり、このモデルを「舌の自動認識用モデル」に用いることとする。なお、「矢印」や「楕円」は画像を取得した文献において舌の症状を示すために付けられたマークであった。

認識した舌画像を切り抜く方法は、まず「舌の自動認識用モデル」によって舌画像の中の tongue とラベル付けされた多角形の外側を黒く塗りつぶす (図 3.2(2))。その後、自動認識用

モデルを用いて検出された四角い境界ボックスに合わせて舌の部分を切り抜く。これらの作業により図 3.2(1) から (3) の画像が生成される。

なお、臓腑の状態判定学習を行う際には学習データとなる舌の 5 部位の画像サイズを統一しておく必要があるため、「舌の自動認識用モデル」の学習に用いた 1104 枚の舌の画像を全て図 3.2(3) のような画像となるように切り抜き、それらの画像サイズの高さと幅の平均値（今回は高さが 596 ピクセル、幅が 602 ピクセルであった）を求め、Python の画像処理ライブラリ Pillow を用いてその大きさにリサイズしておく。

次に、切り抜かれた舌のみの画像（例えば図 3.2 の (3)）を、文献 [18] に示されている比率に基づいて図 3.3 のように分割して舌の 5 部位を切り出す。具体的には、まず舌画像の上下左右の端から縦幅と横幅の 1/5 の位置を定める（図 3.3 の左）。次に画像の下端から 1/5 までを切り出して舌尖部とし、残った画像の左右 1/5 を切り出して左右それぞれの舌辺部とする。さらに残りの画像の上 1/5 を舌根部とし、最後まで残った画像が舌中部となる。切り出し作業にあたっては、Python と OpenCV を用いてトリミング処理を行い、さらに左右の舌辺部の画像は OpenCV を用いて横方向に結合して 1 つの舌辺部画像としている。これらの作業により、舌辺部（高さ 476 ピクセル×幅 242 ピクセル）、舌根部（高さ 120 ピクセル×幅 360 ピクセル）、舌尖部（高さ 120 ピクセル×幅 602 ピクセル）と舌中部（高さ 356 ピクセル×幅 360 ピクセル）が得られ、臓腑状態判定のモデルに入力するための画像が得られる。

表 3.1: Loss の種類とその意味

Loss の値	説明
rpn_class_loss	RPN の画素単位領域に対する損失関数の値
rpn_bbox_loss	RPN のバウンディングボックスに対する損失関数値
$mrcnn_class_loss$	Head のクラス名に対する損失関数の値
$mrcnn_bbox_loss$	Head のバウンディングボックスに対する損失関数値
$mrcnn_mask_loss$	画素単位領域に対する損失関数の値
$Loss_{sum}$	上記 Loss の値の合計値（合計損失値）

3.1.3 臓腑状態判定のための学習

臓腑の状態を判定するための学習に用いる画像データとして、前述の文献 [9, 18, 54–57] から得た 1226 枚の画像に加えて、新たに文献 [19, 34, 62] より 491 枚の画像をスキャンして合計画像を 1717 枚に増やした上で、それらから画像が示す症状等に基づいて 523 枚を選び出して用いることとした。これは、文献 [9, 18, 54–57] からの画像には、舌根部に腎の症状が現れている画像や舌辺部に肝および胆の症状が現れている画像が極めて少なく、良好な学習結果を得ることが困難であると想定されたためである。これら 523 枚の画像を、舌尖部（心・肺に対応）、舌辺部（肝・胆に対応）、舌中部（脾・胃に対応）、舌根部（腎に対応）に分割することで、最終的にそれぞれの部位ごとに学習に用いることができる画像は 140 枚、105 枚、212 枚、66 枚となった。

臓腑の状態を判定する学習は、各部位に対応する臓腑が病気か健康かを予測する 2 クラス分類問題に属している。そこで、各部位の画像を、対応する臓腑が病気であるクラスと健康であるクラスの 2 クラスに分けて学習を行う。すなわち、ある部位の画像に病気の症状があると判定された場合、その部位に対応する臓腑が病気であると分類される。なお、臓腑が健康であるクラスの各部位の画像は、1717 枚の画像のうち病気であるクラスに属していない画像から、症状が見られないと思われる画像を目視により選び出した。このような方法で作成した学習用、検証用、テスト用の各画像データの内訳を表 3.2 に示す。

表 3.2 に示したように、収集した舌根部の画像が最も少ないが、今回はデータを拡張する水増し手法を取り入れずに 66 枚のみの舌根部の画像を LeNet, ResNet50, ResNet101, DenseNet169,

表 3.2: 臓腑の状態判定学習に用いる画像の内訳

舌の 5 部位	舌尖部		舌辺部 (左右)		舌中部		舌根部	
	病気	健康	病気	健康	病気	健康	病気	健康
学習用画像 (枚)	117	117	86	86	180	180	51	51
検証用画像 (枚)	13	13	9	9	20	20	5	5
テスト用画像 (枚)	10	10	10	10	12	12	10	10

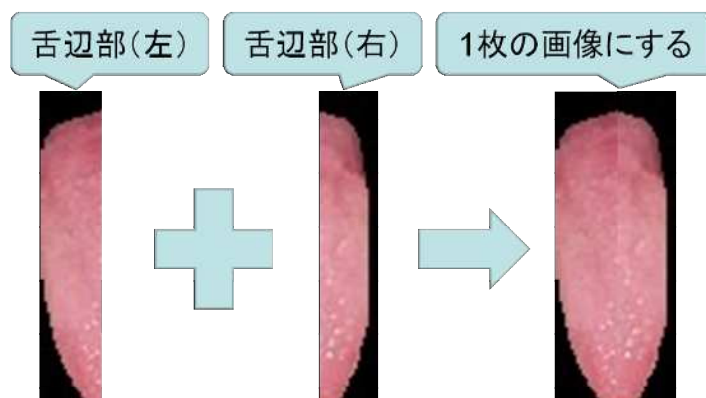


図 3.4: 舌辺部画像の処理

EfficientNet-B0 の 5 つの画像認識モデルを用いた学習およびアンサンブル学習に活用する。表中の「学習用画像」「検証用画像」「テスト画像」とは、それぞれ、学習においてモデルのトレーニング（学習）に用いる画像、トレーニングしたモデルを比較して採用するモデルを決定するための画像、そして、採用したモデルの性能を評価するための画像である。舌辺部については、左右の片側あるいは両側に症状があると判断された場合に、肝または胆が病気であると判定する。左右の舌辺部の画像を個別に学習するとこの判定が難しくなるため、図 3.4 のように舌辺部の左右の画像を 1 枚に結び合わせた画像を学習用、検証用、テスト用のデータとして用いる。図 3.5 に状態判定学習に用いる舌の各部位の画像の例を示す。なお状態判定学習では、舌の部位ごとに学習が行われる。

本研究では、LeNet, ResNet50, ResNet101, DenseNet169, EfficientNet-B0 の 5 つの画像認識モデルを用いて「舌の 5 部位の状態判定用モデル」(図 3.1 ⑧) を生成するための学習を行う。さらに、臓腑の健康状態を判定する正確さを高めるために、複数の画像認識モデルを融合したアンサンブル学習も行う。これらの学習においては、学習回数とパラメータ更新ごとに用いるサンプルの数 (Batch size) [63] をそれぞれ “epoch 100” と “4” と設定している。なお、過学習を防ぐ手段として、5 つの画像認識モデルを学習する際には、検証用画像を病気と健康に正しく分類できた割合 (Validation accuracy) (以降、検証正解率と呼ぶ) を確認し、最良となる epoch で学習を停止する Early Stopping [64] を適用する。一方で、アンサンブル学習を行

う際には、重み付けを 5000 回行い、そのうち最も高い検証正解率を持つモデルを出力結果として評価する。学習結果の中で、テスト用画像に対して正しく識別した割合の最も高いモデルが判定した結果を臓腑状態の最終判定結果として採用する。これら臓腑状態判定における全ての学習において、よく使われている機械学習の最適化アルゴリズムの 1 つである Adam [65] を学習率の値を 0.001 から 0.01 まで 0.003 刻みで増やしながら適用している。学習率は、機械学習の最適化において 1 回の学習で重みパラメータをどのくらい変化させるのかを定める指標であり、学習率が大きくなりすぎると Loss の値が発散し、小さくなりすぎると Loss の値の収束が遅くなる [66]。よって効率的なパラメータ探索処理の実現には最適な学習率を見つけ出すことが重要である。

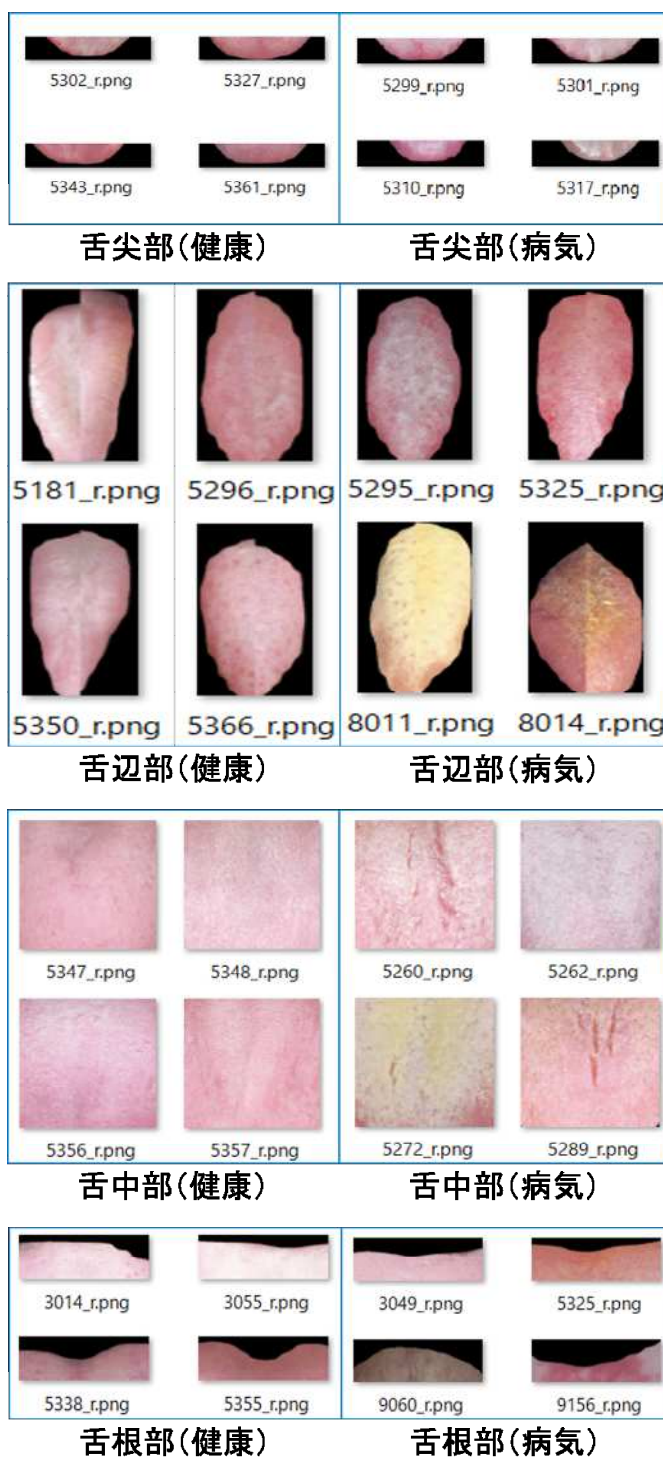


図 3.5: 臓腑の状態判定学習に用いる画像の例

3.2 舌の 5 分割診断法による学習結果と評価

これまで述べてきた舌診支援システムの臓腑状態判定のための学習結果を評価するために、5つの画像認識モデルそれぞれについて、テスト用画像を分類するテストを実施し、分類結果の正解率を比較評価する。さらに正解率の向上を目指して、複数のモデルを組み合わせたアンサンブル学習を行い、そのテスト結果を評価する。これらの結果で最良のモデルを、舌診支援システムにおける「舌の 5 部位の状態判定用モデル」(図 3.1 ⑧)として採用する。

実験に用いた画像認識モデルのうち、LeNet は文献 [47] で示された構造に基づいてモデルを構築・初期化してから利用しており、Resnet50, Resnet101, DenseNet169 および EfficientNet-B0 は、Tensorflow のパッケージである keras.applications からダウンロードした事前学習済みモデルを利用している。また、各々の事前学習済みモデルはファインチューニングにより、全結合層を 2 クラス分類用に変更し、再学習は後半の層から行うようにしている。

3.2.1 5つの画像認識モデルによる結果

ここでは、LeNet, ResNet50, ResNet101, DenseNet169, EfficientNet-B0 によるテスト用画像を分類する実験を行い、その結果を比較する。実験の手順は以下のとおりである。

【実験手順】5つの画像認識モデルに対して、舌の部位ごとに以下を繰り返す。

1. 4つの学習率ごとに、学習して得られたモデルに検証用画像を用いて検証した結果、検証正解率の値が最も高いモデルを選ぶ。
2. 選び出された4つのモデルから、表 3.2 のテスト用画像を用いた正解率が最も高いモデルを、その画像認識モデルをその部位に適用したモデルとして採用する。

5つの画像認識モデルについて、採用されたモデルによる臓腑の状態判定学習のテスト結果を表 3.3～表 3.7 に示す。例えば、LeNet を用いた場合の、舌尖部、舌辺部、舌中部および舌根部における正解率は、それぞれ 55%、60%、71%および 55%であった(表 3.3)。また、表 3.7 より、EfficientNet-B0 の正解率はいずれの部位でも 50%であるが、特異度が 0%であった。つま

り、「健康」の判定が出来なかった。これは、EfficientNet-B0 を用いた学習においては、学習用画像が少ないため、「健康」の舌画像の特徴量を抽出できないことが原因であると考えられる。

これらの 5 つのモデルによる学習において、舌の部位ごとに正解率が最も高いモデルの結果を表 3.8 にまとめて示す。部位ごとに最高の正解率が得られたのは、それぞれ ResNet50, ResNet101, DenseNet169 および ResNet50 を用いた時である。このうち、正解率が最も高いものは舌辺部の 80% であり、学習に用いたモデルと学習率は “ResNet101 と 0.001” である。しかし、舌尖部、舌中部、舌根部の正解率はそれぞれ 70%, 75%, 65% と比較的低い結果となっており、これらのモデルを個別に用いて学習した結果は全般的に正解率が十分とは言えない。そこで、複数の画像認識モデルを用いたアンサンブル学習を活用することで正解率の向上を図る。但し、EfficientNet-B0 は、いずれの部位の正解率も他の 4 つのモデルを超えるものはないため、アンサンブル学習には用いないこととする。

表 3.3: LeNet によるテスト結果

舌の 5 部位	舌尖部			舌辺部 (左右)			舌中部			舌根部		
	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率
テスト結果	70%	40%	55%	20%	100%	60%	67%	75%	71%	30%	80%	55%

表 3.4: ResNet50 によるテスト結果

舌の 5 部位	舌尖部			舌辺部 (左右)			舌中部			舌根部		
	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率
テスト結果	50%	90%	70%	90%	40%	65%	58%	83%	71%	80%	50%	65%

表 3.5: ResNet101 によるテスト結果

舌の 5 部位	舌尖部			舌辺部 (左右)			舌中部			舌根部		
	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率
テスト結果	70%	60%	65%	100%	60%	80%	42%	100%	71%	40%	80%	60%

表 3.6: DenseNet169 によるテスト結果

舌の 5 部位	舌尖部			舌辺部 (左右)			舌中部			舌根部		
	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率
テスト結果	60%	40%	50%	80%	70%	75%	50%	100%	75%	20%	80%	50%

表 3.7: EfficientNet-B0 によるテスト結果

舌の 5 部位	舌尖部			舌辺部 (左右)			舌中部			舌根部		
	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率
テスト結果	100%	0%	50%	100%	0%	50%	100%	0%	50%	100%	0%	50%

表 3.8: 5 つのモデルによる最良のテスト結果

舌の 5 部位	舌尖部			舌辺部 (左右)			舌中部			舌根部		
	ResNet50(0.007)			ResNet101(0.001)			DenseNet169(0.01)			ResNet50(0.01)		
テスト結果	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率
		50%	90%	70%	100%	60%	80%	50%	100%	75%	80%	50%

3.2.2 アンサンブル学習を用いた実験

臓腑状態判定の正解率を高めるために、3.2.1 節で述べた 4 つの画像認識モデルを用いてアンサンブル学習を行う。この処理では、(1) 4 つの画像認識モデルそれぞれについて、学習用画像と検証用画像を用いて学習および検証を行い、最も高い検証正解率をもつモデル M_1, M_2, M_3, M_4 を取得し、(2) 得られた 4 つのモデルから加重平均を用いたアンサンブル学習 [52] により新たなモデルを得る、という手順で実行される。手順 (2) では、重み付けされたモデルに対して、検証正解率が最も高いモデルを出力させるために、5000 回の探索を行う。その詳細は次の通りである。

- (a) M_1, \dots, M_4 に対して、重み $w_{M_1}, w_{M_2}, w_{M_3}, w_{M_4}$ の値を均等に割り振り、かつ合計値が 1 となるように設定し、アンサンブル学習で得られたモデルを h_0 とする。
- (b) i を 0 に初期化する。
- (c) M_1, \dots, M_4 に対して、新たな重み $w'_{M_1}, w'_{M_2}, w'_{M_3}, w'_{M_4}$ を合計値が 1 となるようにランダムに設定し、アンサンブル学習で得られたモデルを h_i とする。
- (d) もし h_i の検証正解率が h_0 の検証正解率より大きい場合、 h_i の検証正解率を h_0 に代入し、 i を 1 増やして (e) へ。そうでなければ、 i を 1 増やして (e) へ。
- (e) $i < 5000$ であれば (c) へ。そうでなければ h_0 をアンサンブル学習の結果として出力して終了。

表 3.9 は、LeNet, ResNet50, ResNet101 および DenseNet169 に対して上記のアンサンブル学習を適用した結果である。表には舌の部位ごとのアンサンブル学習において、学習率を 0.001 から 0.01 まで 0.003 刻みで増やして実験した中で、最も高い正解率とその時の学習率を示している。例えば、舌尖部の最も高い正解率は 80% であり、その時の学習率は 0.001 である。同様に、舌辺部の場合は学習率が 0.007 の時に最も高い正解率 90% が得られている。また、舌中部と舌根部の場合は、学習率がそれぞれ 0.001 と 0.007 の時に、それぞれの高い正解率 92% と 85% が得られている。

さらに、古典的と言われる LeNet, および、効果や特性が似通った ResNet50 と ResNet101

表 3.9: アンサンブル学習によるテスト結果

舌の 5 部位	舌尖部			舌辺部 (左右)			舌中部			舌根部		
学習率	0.001			0.007			0.001			0.007		
テスト結果	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率
		100%	60%	80%	100%	80%	90%	83%	100%	92%	80%	90%

とを併用していることによる効果を確認するために、3つのネットの組み合わせ (ResNet50, ResNet101, DenseNet169), (LeNet, ResNet101, DenseNet169) および (LeNet, ResNet50, DenseNet169) に対しても、前述と同様の学習を行った。表 3.10 に示したとおり、いずれの組み合わせにおいても、舌の全ての部位について、モデルを 4つ全て用いたときの結果よりも劣っている。従って、古典的な LeNet や、同特性の ResNet50 と ResNet101 の併用も効果的であることが言える。

加えて、表 3.3～表 3.7 および表 3.8 に示したモデルを個別に用いた際の正解率と表 3.9 の正解率とを比較すると、全ての部位において、アンサンブル学習の結果は個別のネットの結果を 10%以上上回っていることが分かる。よって、舌画像による臓腑健康状態の診断においては、アンサンブル学習は極めて効果的であると言える。

表 3.10: モデル組合せのアンサンブル学習の結果

対象のモデル	ResNet50, ResNet101, DenseNet169			LeNet, ResNet101, DenseNet169			LeNet, ResNet50, DenseNet169		
舌の 5 部位	舌尖部								
学習率	0.001			0.001			0.007		
テスト結果	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率
	90%	40%	65%	70%	80%	75%	70%	80%	75%
舌の 5 部位	舌辺部 (左右)								
学習率	0.004			0.01			0.007		
テスト結果	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率
	60%	90%	75%	90%	80%	85%	80%	90%	85%
舌の 5 部位	舌中部								
学習率	0.007			0.01			0.001		
テスト結果	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率
	75%	100%	88%	75%	100%	88%	75%	100%	88%
舌の 5 部位	舌根部								
学習率	0.01			0.01			0.01		
テスト結果	再現率	特異度	正解率	再現率	特異度	正解率	再現率	特異度	正解率
	90%	70%	80%	90%	60%	75%	80%	60%	70%

第4章 深層学習による裂紋舌画像の認識

本章では、裂紋舌の裂紋状態の自動診断技術の実現を目指して、深層学習を活用した舌表面状態の画像識別による裂紋舌の自動診断法について述べる。

まずは、裂紋舌の自動診断の処理手順について概説した上で、LeNet, ResNet50, ResNet101, DenseNet169, ConvNeXt-Tiny [50] の5つの画像認識モデルを用いた裂紋舌の自動認識と裂紋舌の裂紋状態判定学習を説明する。なお、各々の学習結果を評価するために、5つの画像認識モデルを用いた時の舌診（画像分類）の性能を調べることで各モデルの評価を行う。また、裂紋舌画像に対する水増し手法の活用、アンサンブル学習の活用、および学習結果の比較に基づき、裂紋舌の自動診断における正解率向上について考察する。

4.1 深層学習による裂紋舌の自動診断法の提案

4.1.1 裂紋舌の自動診断の処理手順

本研究では、図4.1に示すような裂紋舌の自動診断を目指している。この診断は、次の(1)～(3)の処理からなっている。(1) まず舌を含む写真が入力されると、図3.2に示すように、その舌を自動的に認識・切抜するように舌のみの画像データを生成する(図4.1 ②③)。(2) 次に、生成された舌画像において、「裂紋舌の自動認識用モデル」を適用して、裂紋舌の画像データと非裂紋舌の画像データを生成する(図4.1 ⑤⑥)。(3) 最後に、生成された裂紋舌のデータに対して、「裂紋舌の裂紋状態判定用モデル」を適用して、裂紋舌の裂紋状態を自動診断する(図4.1 ⑨⑩)。以降4.1.2節および4.1.3節にて、(2)(3)の学習とハイパーパラメータの設定について述べる。

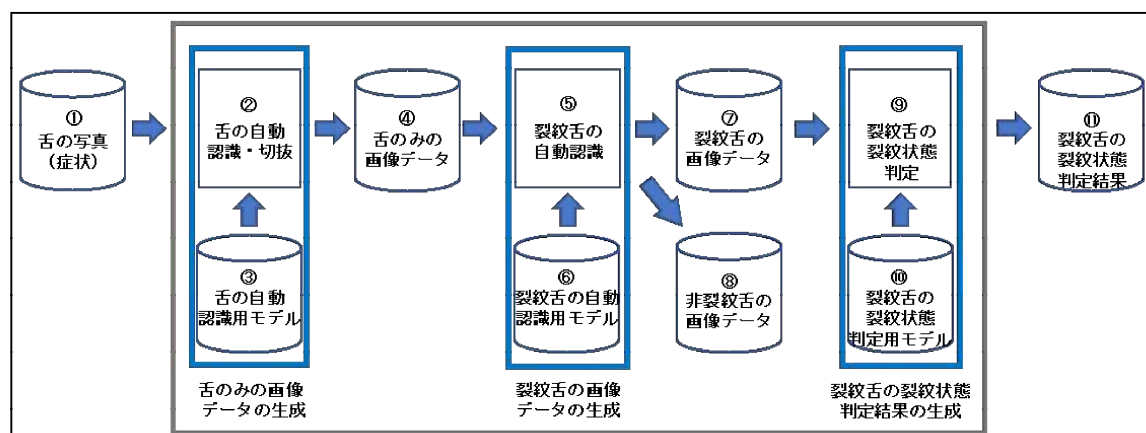


図 4.1: 深層学習を用いた裂紋舌の診断の流れ

4.1.2 裂紋舌の自動認識とその裂紋状態判定学習

裂紋舌の自動認識を行うために、画像データを次のように用意する。文献 [9, 18, 54-57] からの画像のみで裂紋舌を自動認識する実験を行うとすると、裂紋舌と記されている画像の枚数が少なく、優れた学習結果（高い正解率）を得ることが困難であると判断した。そこで、新たに文献 [19, 34, 62] から 491 枚の画像を収集して、合計 1717 枚の画像から 180 枚の裂紋舌画像を選び出して用いることにした。

学習の方法としては、裂紋舌の自動認識学習は裂紋舌か非裂紋舌かを予測する 2 クラス分類問題に属しているため、裂紋舌のクラスと非裂紋舌のクラスの 2 つのクラスに分けて行うこととする。具体的には、非裂紋舌のクラスの画像は、1717 枚の画像から裂紋舌であるクラスに属している画像を除き、非裂紋舌であると思われる画像を目視により選び出した。なお、この学習では、裂紋舌と非裂紋舌を区別する特徴を自動的に学習し、新たな舌画像に対しても適切な分類を行えるようになる。これにより、裂紋舌の自動認識において高い信頼性が確保され、さまざまな応用領域で有用な成果が期待される。

次に、裂紋舌の裂紋状態判定学習の方法としては、裂紋舌の裂紋状態を裂紋が酷いクラスと裂紋が軽いクラスの 2 つのクラスに分けて行うこととする。また、学習の準備として、裂紋舌の自動認識学習に用いた 180 枚の裂紋舌画像から裂紋状態が酷いものを 40 枚選び出し、それらを除いた 120 枚から裂紋状態が軽いと思われる 40 枚の画像を目視により選び出した。しか

しこれらの画像の枚数は極めて少ないため、今回はデータを拡張するという水増し手法を取り入れる学習を行う。具体的には、それぞれのクラスの画像に対して、画像処理ライブラリである OpenCV の flip 関数による「上下反転」、「左右反転」および「上下反転+左右反転」の 3 種類の変換を施すことでデータを水増しして、5 つの画像認識モデルを個別に用いた学習およびアンサンブル学習に活用する。この水増し手法により、多種多様な画像を学習用データとして用意することができ、それぞれの変換によって生じる画像の違いは、表現のバリエーションをカバーするのに有効的である。

以上のような学習方法に基づいて分類した学習用、検証用とテスト用画像の枚数の内訳をそれぞれ表 4.1 と表 4.2 に示す。

表 4.1: 裂紋舌の自動認識学習に用いる画像の内訳 (水増しなし)

画像データ クラス	学習用画像 (枚)	検証用画像 (枚)	テスト用画像 (枚)
裂紋舌	153	17	10
非裂紋舌	153	17	10

表 4.2: 裂紋舌の裂紋状態判定学習に用いる画像の内訳 (水増しあり)

画像データ クラス	学習用画像 (枚)	検証用画像 (枚)	テスト用画像 (枚)
裂紋状態が酷い	135	15	10
裂紋状態が軽い	135	15	10

4.1.3 ハイパーパラメータの設定

本研究では、LeNet, ResNet50, ResNet101, DenseNet169, ConvNeXt-Tiny の 5 つの画像認識モデルを用いて「裂紋舌の自動認識用モデル」(図 4.1 ⑥)と「裂紋舌の裂紋状態判定用モデル」(図 4.1 ⑩)を生成するための学習を行う。また、裂紋舌の自動認識と裂紋舌の裂紋状態を判定する正解率を高めるために、複数の画像認識モデルを融合したアンサンブルも行う。

これらの学習においては、epoch 数と Batch size をそれぞれ 100 と 4 と設定している。なお、過学習を防ぐ手段として、5 つの画像認識モデルを学習する際には、検証正解率を確認し、最良となる epoch で学習を停止する Early Stopping を適用する。一方で、アンサンブル学習を行う際には、重み付けを 5000 回行い、そのうち最も高い検証正解率を持つモデルを出力結果として評価する。学習結果の中で、テスト用画像に対して正しく識別した割合の最も高いモデルが判定した結果を裂紋舌の裂紋状態の最終判定結果として採用する。これら裂紋舌の裂紋状態判定における全ての学習において、Adam の値を 0.001 から 0.01 まで 0.003 刻みで増やしながらか適用している。

裂紋舌の自動認識学習と裂紋舌の裂紋状態判定学習は両者共に 2 クラス分類に属する問題である。前者では、正クラスと負クラスはそれぞれ裂紋舌の画像と非裂紋舌の画像を指しており、後者では、正クラスと負クラスはそれぞれ裂紋舌の裂紋状態が酷い画像と裂紋状態が軽い画像を表している。

4.2 裂紋舌の自動診断法による学習結果と評価

これまで述べてきた裂紋舌の自動認識と裂紋舌の裂紋状態判定のための学習結果を評価するために、5つの画像認識モデルそれぞれについて、テスト用画像を分類するテストを実施し、分類結果の正解率を比較評価する。さらに正解率の向上を目指して、複数のモデルを組み合わせたアンサンブル学習を行い、そのテスト結果を評価する。これらの結果で最良のモデルを、それぞれ「裂紋舌の自動認識用モデル」(図 4.1 ⑥)と「裂紋舌の裂紋状態判定用モデル」(図 4.1 ⑩)として採用する。

実験に用いた画像認識モデルのうち、LeNet は文献 [47] で示された構造に基づいてモデルを構築・初期化してから利用しており、Resnet50, Resnet101, DenseNet169 および ConvNeXt-Tiny は、Tensorflow のパッケージである keras.applications からダウンロードした事前学習済みモデルを利用している。また、各々の事前学習済みモデルはファインチューニングにより、全結合層を 2 クラス分類用に変更し、再学習は後半の層から行うようにしている。

4.2.1 5つの画像認識モデルによる結果

ここでは、LeNet, ResNet50, ResNet101, DenseNet169, ConvNeXt-Tiny による表 1 と表 2 のテスト用画像を分類する実験を行い、その結果を比較する。実験の手順は以下のとおりである。

【実験手順】 裂紋舌の自動認識学習と本研究で使用する 3 種類の水増し手法を適用した裂紋舌の裂紋状態判定学習に対して、5つの個別の画像認識モデルによる学習ごとに以下を繰り返す。

1. 設定した 4 つの学習率ごとに、学習して得られたモデルに検証用画像を用いて検証した結果、検証正解率の値が最も高いモデルを選ぶ。
2. 選び出された 4 つのモデルから、テスト用画像を用いた正解率が最も高いモデルを、その画像認識モデルによる学習に適用したモデルとして採用する。

表 4.3: 5つのモデルによる裂紋舌の自動認識学習の結果（水増しなし）

モデル (学習率)	テスト結果 再現率	特異度	正解率
LeNet (0.004)	80%	60%	70%
ResNet50 (0.007)	80%	80%	80%
ResNet101 (0.01)	80%	80%	80%
DenseNet169 (0.001)	70%	100%	85%
ConvNeXt-Tiny (0.007)	80%	90%	85%

表 4.4: 5つのモデルによる裂紋舌の裂紋状態判定学習の結果（水増しあり）

モデル (学習率)	テスト結果 再現率	特異度	正解率
LeNet (0.004)	40%	100%	70%
ResNet50 (0.004)	70%	90%	80%
ResNet101 (0.01)	60%	40%	50%
DenseNet169 (0.001)	60%	80%	70%
ConvNeXt-Tiny (0.001)	30%	100%	65%

5つの画像認識モデルについて、採用されたモデルによる裂紋舌の自動認識学習と本研究で使用する3種類の水増し手法を適用した裂紋舌の裂紋状態判定学習のテスト結果を表6と表7に示す。

表4.3より、裂紋舌の自動認識学習において、正解率が最も高いものは85%であり、学習に用いた画像認識モデルと学習率では“DenseNet169と0.001”と“ConvNeXt-Tinyと0.007”の2つが存在するが、後者の方の再現率と特異度の差が前者より小さいため、“ConvNeXt-Tiny”は裂紋舌の自動認識学習において、裂紋舌のクラスと非裂紋舌のクラスの両方に対してより均衡して高い性能を示していると考えられる。

また、表4.4に示すように、本研究で使用する3種類の水増し手法を適用した裂紋舌の裂紋状態判定学習において、正解率が最も高いものは80%であり、学習に用いた画像認識モデルと学習率は“ResNet50と0.004”であった。しかし、表7において、正解率が最も高いもの以外の画像認識モデルによる学習の正解率はすべて70%以下と比較的低い結果となっており、これ

らのモデルを個別に用いて学習した結果は全般的に正解率が十分とは言えない。そこで、複数の画像認識モデルを用いたアンサンブル学習を活用することで正解率の向上を図る。

4.2.2 アンサンブル学習を用いた実験

裂紋舌の自動認識と裂紋状態判定の正解率を高めるために、4.2.1 節で述べた 5 つの画像認識モデルを用いてアンサンブル学習を行う。この処理では、(1) 5 つの画像認識モデルそれぞれについて、学習用画像と検証用画像を用いて学習および検証を行い、最も高い検証正解率をもつモデル M_1, M_2, M_3, M_4, M_5 を取得し、(2) 得られた 5 つのモデルから加重平均を用いたアンサンブル学習により新たなモデルを得る、という手順で実行される。手順 (2) では、重み付けされたモデルに対して、検証正解率が最も高いモデルを出力させるために、5000 回の探索を行う。その詳細は次の通りである。

- (a) M_1, \dots, M_5 に対して、重み $w_{M_1}, w_{M_2}, w_{M_3}, w_{M_4}, w_{M_5}$ の値を均等に割り振り、かつ合計値が 1 となるように設定し、アンサンブル学習で得られたモデルを h_0 とする。
- (b) i を 0 に初期化する。
- (c) M_1, \dots, M_5 に対して、新たな重み $w'_{M_1}, w'_{M_2}, w'_{M_3}, w'_{M_4}, w'_{M_5}$ を合計値が 1 となるようにランダムに設定し、アンサンブル学習で得られたモデルを h_i とする。
- (d) もし h_i の検証正解率が h_0 の検証正解率より大きい場合、 h_i の検証正解率を h_0 に代入し、 i を 1 増やして (e) へ。そうでなければ、 i を 1 増やして (e) へ。
- (e) $i < 5000$ であれば (c) へ。そうでなければ h_0 をアンサンブル学習の結果として出力して終了。

表 4.5 と表 4.6 は、LeNet, ResNet50, ResNet101, DenseNet169 および ConvNeXt-Tiny に対して、上記のアンサンブル学習を裂紋舌の自動認識と 3 種類の水増し手法を取り入れた裂紋舌の裂紋状態判定に適用した各々のテスト結果である。これらの表には、学習率を 0.001 から 0.01 まで 0.003 刻みで増やして実験した中で、最も高い正解率とその時の学習率を示している。例えば、表 4.5 においては、本研究で設定した 4 つの学習率に対する正解率はどれも 90% で一

貫しており、同時にこれらのテスト結果は裂紋舌の自動認識における最も高い正解率となっている。同様に、表 4.6 のアンサンブル学習による 3 種類の水増し手法を取り入れた裂紋舌の裂紋状態判定の場合は、学習率が 0.007 の時に最も高い正解率 85% が得られている。

さらに、比較的新しい画像認識モデルである ConvNeXt-Tiny による効果を確認するために、LeNet, ResNet50, ResNet101 および DenseNet169 の 4 つの画像認識モデルに対しても、前述と同様の学習を行った。表 4.7 に示したとおり、4 つの画像認識モデルを対象としたアンサンブル学習によるテスト結果は、ConvNeXt-Tiny を含めた 5 つの画像認識モデルを対象としたアンサンブル学習の最も高い正解率とすべて同じであることが分かる。ただし、水増し手法を取り入れた裂紋舌の裂紋状態判定においては、表 4.7 に示すように、5 つの画像認識モデルを対象としたアンサンブル学習の方の再現率と特異度の差が 4 つの画像認識モデルを対象としたアンサンブル学習より 20% 小さいため、“ConvNeXt-Tiny” は水増し手法を活用した裂紋舌の裂紋状態判定において、裂紋舌の裂紋状態が酷いクラスと裂紋状態が軽いクラスの両者に対してより均衡して高い性能を示していると言える。従って、比較的新しい画像認識モデルである ConvNeXt-Tiny も効果があると言える。

加えて、表 4.3 と表 4.4 に示したモデルを個別に用いた際の正解率と表 4.5 と表 4.6 のアンサンブル学習による正解率とを比較すると、アンサンブル学習の正解率は LeNet, ResNet50, ResNet101, DenseNet169 および ConvNeXt-Tiny を個別に用いた際の正解率を上回ることが分かる。よって、アンサンブル学習は 5 つの個別の画像認識モデルによる学習より優れていると言える。

表 4.5: アンサンブル学習による裂紋舌の自動認識テスト結果（水増しなし）

学習率 \ テスト結果	再現率	特異度	正解率
0.001	80%	100%	90%
0.004	80%	100%	90%
0.007	80%	100%	90%
0.01	80%	100%	90%

表 4.6: アンサンブル学習による裂紋舌の裂紋状態判定テスト結果（水増しあり）

学習率 \ テスト結果	再現率	特異度	正解率
0.007	80%	90%	85%

表 4.7: 4つの画像認識モデルを対象としたアンサンブル学習のテスト結果

対象のモデル：LeNet, ResNet50, ResNet101, DenseNet169			
裂紋舌の自動認識テスト結果（水増しなし）			
学習率 \ テスト結果	再現率	特異度	正解率
0.001	80%	100%	90%
0.007	80%	100%	90%
0.01	80%	100%	90%
裂紋舌の症状度合い判定テスト結果（水増しあり）			
0.001	70%	100%	85%
0.004	70%	100%	85%
0.007	70%	100%	85%

第5章 おわりに

本論文では AI 技術の深層学習を活用して、舌の 5 部位に対応する臓腑の健康状態を自動診断する 5 分割診断法と、舌表面状態の画像識別による裂紋舌の自動診断法を提案した。提案した 5 分割診断法において、類似する先行研究はなく、臓腑の健康状態の診断を舌画像の自動認識によって行う先駆的な研究である。また左右 2 つに分かれている舌辺部に対して、統合した画像を用いて診断する点も特徴として挙げられる。

5 分割診断法では、Mask R-CNN を用いて舌画像の認識と切り抜きを行い、その結果を 5 つの部位に分割した上で、各部位の画像における症状の有無を判定することで、各部位に対応する臓腑の健康状態を判定する。各部位の症状の有無の判定には、LeNet, ResNet50, ResNet101, DenseNet169 によるアンサンブル学習を用いる。

提案した 5 分割診断法を評価するために、9 冊の書籍にある 1717 枚の画像から選び出した 523 枚を対象に、5 分割診断法を適用する評価実験を行った。実験結果より次のことが明らかになった。なお、本手法では比較的古いモデルも活用するがその有用性も確認された。

- 舌の 5 部位に対応する臓腑の健康状態の判定については、5 つのモデルを個別に用いて学習した時の最良の結果として、5 部位のいずれも正解率が 65% 以上であり、また、その時の再現率・特異度のどちらも 50% 以上である。
- 一方、アンサンブル学習を用いた場合、その正解率が 80% 以上で、5 つのモデルのいずれよりも高くなっている。よって、舌画像による臓腑健康状態の診断においては、アンサンブル学習を用いた 5 分割診断法は極めて効果的であると言える。

本研究の舌画像は文献より収集したため、舌根部を含む全ての部位が比較的きれいに映されている。従って、本研究の成果を実臨床で応用する場合は、舌根部まで確実に撮影するよう

に、患者が舌を出した状態で光源を舌根部まで届くように配置する必要がある。また、舌運動の障害や筋力低下、舌小帯の短縮など、本来舌診が困難な患者に対しては、医療用超小型カメラの活用が考えられる。

一方、舌表面状態の画像識別による裂紋舌の自動診断法では、Mask R-CNN を用いて舌画像の認識と切り抜きを行い、その結果を裂紋舌の自動認識学習に適用して、裂紋舌のみの画像データを抽出した上で、裂紋舌の裂紋状態を判定する。裂紋舌の自動認識と裂紋舌の裂紋状態の判定には、LeNet, ResNet50, ResNet101, DenseNet169, ConvNeXt-Tiny の 5 つの画像認識モデルによるアンサンブル学習を用いる。

提案した舌表面状態の画像識別による裂紋舌の自動診断法を評価するために、まず 9 冊の書籍にある 1717 枚の画像から選び出した 360 枚を対象に、裂紋舌の自動認識学習を行った。次に収集した 180 枚の裂紋舌画像から選び出した 80 枚を対象に、“上下反転”、“左右反転”および“上下反転+左右反転”の 3 種類の水増し手法で画像データを 4 倍に拡張して 5 つのモデルおよびこれらのアンサンブルの評価実験を行った。これらの実験結果より次のことが明らかになった。なお、アンサンブル学習では比較的新しい画像認識モデルである ConvNeXt-Tiny も取り入れて活用するがその効果と有用性も確認された。

- 裂紋舌の自動認識学習については、5 つの画像認識モデルの学習における最良のテスト結果として、いずれも正解率が 70% 以上である。また、アンサンブル学習を用いた場合、設定した 4 種類の学習率において、いずれも正解率が 90% であり、5 つの個別の画像認識モデルの学習における最良の正解率より 20% 高くなっていることから、アンサンブル学習を用いた方が効果的である。
- 裂紋舌の裂紋状態判定学習については、5 つの個別の画像認識モデルの学習とアンサンブル学習を用いた場合、それらの最良のテスト結果として、各々の最高な正解率が 80% と 85% であることから、アンサンブル学習を利用した方がより効果的である。

今後の展望として、これらの診断法では、深層学習を舌診に活用することで、次のような利点が見込まれる。まず、舌の微細な変化や特徴をより正確に捉えるための高度なパターン認識

が期待されること。次に、主観的な要素を減少させて、客観的かつ一貫性のある舌診を実現する診断支援が期待されること。また、大量かつ多様な舌診データから潜在的な傾向やパターンを発見しやすくし、医学的知識の拡充や病気の理解が進む可能性があること。さらに、舌色、形状、舌苔など多様な特徴を総合的に評価する複数の特徴の統合も期待されること。なお、深層学習を用いた舌診結果とデータベースの連携により臓腑の健康状態や舌の種類を把握し、診断の正解率向上や新たな知見の発見が期待され、医療従事者の診断業務をサポートすると同時に、経験豊富な専門家の知識を補完する役割も果たす可能性がある。

今後の課題として、(1) 学習効果を上げるために、更なる舌画像の収集や新たな水増し手法の適用を行うこと、(2) 舌診の正解率を向上させるために、EfficientNet-B* (*:1-7)、Lambda Networks, MLP-Mixer, ViT (Vision Transformer) などの最新の画像認識モデルも取り入れたアンサンブル学習を試みること、(3) 舌色や舌質・舌苔の状態等に基づいた更なる詳細な基準を設けて、画像を病気別により正解率の高い割り振りの方法を検討すること、などが挙げられる。

これらの課題を解決するために、実際に医療現場の医師と連携を取りながら、多くの患者の様々な舌象の変化を考察した上で、関連する舌画像を病気別に分類して、データベースを構築する必要がある。また、最新の画像認識モデルを取り入れたアンサンブル学習も探索して、より正解率の高い学習方法を探す。

参考文献

- [1] 兵頭 明, 『東洋医学のしくみ』, 新星出版社, 2012.
- [2] 兵頭 明, 『経絡・ツボの教科書』, 新星出版社, 2012.
- [3] 形井 秀一, “日本鍼灸の歴史”, 全日本鍼灸学会雑誌, vol. 62, no. 1, pp. 12-28, 2012.
- [4] 篠原 昭二, 坂口 俊二, 浦山 久嗣, 香取 俊光, 河原 保裕, 小林 健二, “第二次日本経穴委員会の提言”, 医道の日本, vol. 71, no. 5, pp. 142-154, 2012.
- [5] 山下 仁, 津嘉山 洋, “【総説】国際化する鍼灸：その動向と展望(1) 欧米における普及と臨床研究の進歩”, 日本補完代替医療学会誌, vol. 3, no. 3, pp. 77-81, 2006.
- [6] 甘 泉, 呉 靱, 中田 充, 葛 崎偉, “カラーペトリネットによる東洋医学の人体モデルの構築”, IEICE Trans. Inf.& Syst., vol. J104-D, no. 04, pp. 217-227, 2021.
- [7] Q. Gan, R. Wu, M. Nakata and Q.W. Ge, “Construction of a human body model for acupuncture and moxibustion treatment by colored petri nets”, BioSystems, vol. 210, no. 104526, pp. 1-11, 2021 (<https://doi.org/10.1016/j.biosystems.2021.104526>).
- [8] 甘 泉, 呉 靱, 葛 崎偉, “気・血・津液を考慮した東洋医学の人体モデルの構築”, 東アジア研究, no. 20, pp. 85-98, 2021.
- [9] 陳 勇, 『舌診論』, 合同会社花乱社, 2018.
- [10] C. Zhou, H. Fan and Z. Li, “Tonguenet : Accurate localization and segmentation for tongue images using deep neural networks”, IEEE Access, vol. 7, pp. 148779-148789, 2019.

- [11] 和辻 直, “ファジィ理論を応用した舌診断支援システムの構築と検証”, 明治鍼灸医学, no. 33, pp. 11-33, 2003.
- [12] G. Nambu, T. Namiki, T. Nakaguchi and T. Tanaka, “Extraction of tongue coating area from tongue image for automated tongue diagnosis”, Journal of Ergonomic Technology, vol .20, no .1, pp. 1-9, 2020.
- [13] 三宅 洋一, “伝統医学への期待”, Journal of Image & Information Management, no. 5, p. 36, 2018.
- [14] X. Wang et al., “Constructing tongue coating recognition model using deep transfer learning to assist syndrome diagnosis and its potential in noninvasive ethnopharmacological evaluation”, Journal of Ethnopharmacology, vol. 285, no. 114905, pp. 1-8, 2021 (<https://doi.org/10.1016/j.jep.2021.114905>).
- [15] T. Jiang et al., “Deep learning multi-label tongue image analysis and its application in a population undergoing routine medicalcheckup”, Hindawi, vol. 2022, no. 3384209, pp. 1-12, 2022.
- [16] 石川 堅也, 他, “舌診断支援のための深層学習を用いた舌特徴抽出ソフトウェアの開発”, IEICE Technical Report, vol. 122, no. MI-188, pp. 47-48(MI), 2022.
- [17] 辰巳 洋, 『一語でわかる中医用語辞典』, 源草社, 2009.
- [18] 許 家佗, 『中医舌診臨床図解』, 化学工業出版社, 2017. (in Chinese)
- [19] 王 憶勤, 『漢英対照中医舌診図譜』, 上海科学技術出版社, 2007. (in Chinese)
- [20] H. Yamashita, H. Tsukayama and C. Sugishita, “Popularity of complementary and alternative medicine in Japan: a telephone survey”, Complement Ther Med, vol. 10, pp. 84-93, 2002.

- [21] 孫 広仁, 鄭 洪新, 『中医基礎理論』, 中国中医薬出版社, 2012 (in Chinese).
- [22] 東 丈夫, 村上 光太郎, 『漢方薬の実際知識増補版』, 東洋経済新報社, 1991.
- [23] 天津中医薬大学, 学校法人後藤学園 編集責任, 劉 公望, 兵頭 明, 平馬 直樹, 路 京華 編集, 学校法人後藤学園中医学研究所 訳, 『針灸学 [基礎篇]』, 東洋学術出版社, 2007.
- [24] 東 丈夫, 『素問』, たにぐち書店, 2006.
- [25] 東 丈夫, 『靈枢』, たにぐち書店, 2006.
- [26] 河北医学院, 『素問経 校訳』, 人民衛生出版社, 2017 (in Chinese).
- [27] 山東中医学院, 河北医学院, 『皇帝内径靈枢経 校訳 上冊』, 人民衛生出版社, 2018 (in Chinese).
- [28] 山東中医学院, 河北医学院, 『皇帝内径靈枢経 校訳 下冊』, 人民衛生出版社, 2018 (in Chinese).
- [29] 凌 耀星, 『難経語訳』, 人民衛生出版社, 2013 (in Chinese).
- [30] 黄 作陣 編, 『中蔵経校注』, 学苑出版社, 2008 (in Chinese).
- [31] 宮川 浩也, “三焦新解 その1 三三焦論 張介賓『質疑録』”, 日本内径医学会, no. 36, pp. 16-21, 1991.
- [32] 宮川 浩也, “三焦新解 その8 分類”, 日本内径医学会, no. 43, pp. 17-18, 1992.
- [33] 薛 永東, 『黄帝内経五臓六腑養生法』, 北方婦児出版社, 2010 (in Chinese).
- [34] 費 兆馥, 顧 亦棣, 『舌診カラーガイド (新装版)』, エルゼビア・ジャパン, 2016.
- [35] John McCarthy, “人工知能の論理と哲学”, 『稻盛財団: 京都賞と助成金』, pp. 94-127, 1992.
- [36] 松尾 豊, 『人工知能は人間を超えるか ディープラーニングの先にあるもの』, KADOKAWA/中経出版, 2015.

- [37] アンドレアス・C・ミュラー, サラ・ガイド, 『Python ではじめる機械学習』, 株式会社オライリー・ジャパン, 2020.
- [38] 斎藤 康毅, 『ゼロから作る Deep Learning Python で学ぶディープラーニングの理論と実装』, 株式会社トップスタジオ, 2019.
- [39] J.F. Zhang, J.T. Xu et al., “*Diagnostic Method of Diabetes Based on Support Vector Machine and Tongue Images*”, Biomed Research International, pp. 1-9, 2017(DOI : 10.1155/2017/7961494).
- [40] X.Q. Li, Y. Zhang, Q. Cui, X.M. Yi, Y. Zhang, “*DTooth-Marked Tongue Recognition Using Multiple Instance Learning and CNN Features*”, IEEE Transactions on Cybernetics, vol. 49, no. 2, pp. 380-387, 2019(DOI : 10.1109/TCYB.2017.2772289).
- [41] Q.X. Xiao, Z. Zhang, H. Zhang, “*軽量畳み込みニューラルネットワークに基づく舌苔の色分類方法*”, 測定制御技術, vol. 38, no. 3, 2019(in Chinese).
- [42] 坂本 俊之, 『作ってわかる! アンサンブル学習アルゴリズム入門』, シーアンドアール研究所, 2019.
- [43] Qi Z, Tu LP, Chen JB, Hu XJ, Xu JT, Zhang ZF. “*The Classification of Tongue Colors with Standardized Acquisition and ICC Profile Correction in Traditional Chinese Medicine.*”, Biomed Research International ,no. 12, 2016(DOI : 10.1155/2016/3510807).
- [44] 胡 繼礼, カン 紅星, “*畳み込みニューラルネットワークに基づく舌象の分類*”, 安慶師範大学学报 (自然科学版) , vol. 24, no. 4, pp. 44-49, 2018(in Chinese).
- [45] K. He, G. Gkioxari, P. Dollar and R. Girshick, “*Mask r cnn*”, Proc.2017 IEEE ICCV, pp. 2980-2988, 2017.
- [46] K. He, X. Zhang, S. Ren and J. Sun, “*Deep residual learning for image recognition*”, Proc. ICLR, pp. 770-778, 2015.

- [47] Y., LeCun, P., Haffner, L., Bottou and Y., Bengio, “*Object recognition with gradient-based learning*”, Shape, Contour and Grouping in Computer Vision, Lecture Notes in Computer Science, vol. 1681, pp. 319-345, 1999.
- [48] G. Huang, Z. Liu, L. Van Der Maaten and K.Q. Weinberger, “*Densely connected convolutional networks*”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2261-2269, 2017.
- [49] M.Tan and Q.V. Le, “*EfficientNet: Rethinking model scaling for convolutional neural networks*”, Proc. PMLR, pp. 6105-6114, 2019.
- [50] Z. Liu et al. , “*A ConvNet for the 2020s*”, 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), New Orleans, LA, USA, pp. 11966, 2022 (<https://doi.org/10.1109/CVPR52688.2022.01167>).
- [51] 門脇 大輔, 『Kaggleで勝つデータ分析の技術』, 技術評論社, 2019.
- [52] Z.H. Zhou, 『アンサンブル法による機械学習：基礎とアルゴリズム』, 近代科学社, 2017.
- [53] 松岡 勝也, Mhd Irvan Irvan, 小林 良輔, 山口 利恵, “多要素認証におけるスコアフュージョンへのニューラルネットの適用”, Information Processing Society of JapanJapan, vol. 2019, pp. 734-739, 2019.
- [54] 王 彦暉, 『舌象図鑑-臨床実用』, 化学工業出版社, 2018. (in Chinese)
- [55] 周 幸来, 『舌診快速入門第2版』, 遼寧科学技術出版社, 2016. (in Chinese)
- [56] 黄 英儒, 黄 国東, 黄 晓, 『舌診-名老中医』, 化学工業出版社, 2018. (in Chinese)
- [57] 羅 大倫, 『図解舌診』, 江西科学技術出版社, 2015. (in Chinese)
- [58] B.C. Russell, A. Torralba, K.P. Murphy and W.T. Freeman, “*LabelMe: A database and web-based tool for image annotation*”, International Journal of Computer Vison, vol. 77, Issues 1-3, pp. 157-173, 2008.

- [59] S. Viswanath et al., “*Terrain surveillance system with drone and applied machine vision*” , Journal of Physics : Conference Series, vol. 2115, no. 1, p. 012019, 2021.
- [60] 平田 結愛, 笠松 雅史, 村上 幸一, 重田 和弘, “Mask R-CNN を用いた理科教育向け植物判定システムの検討”, 人工知能学会全国大会論文集, JSAI2020, p. 2H5GS1302, 2020.
- [61] F. Chollet, 『Python と Keras によるディープラーニング』, マイナビ出版社, 2018.
- [62] 藤本連風, 平田耕一, 山本哲斎, 『針灸舌診アトラス』, 緑書房, 2012.
- [63] 斎藤康毅, 『ゼロから作る Deep Learning-Python で学ぶディープラーニングの理論と実装』, 株式会社トップスタジオ, 2019.
- [64] 瀧雅人, 『機械学習スタートアップシリーズ これならわかる深層学習入門』, 講談社, 2017.
- [65] D.P. Kingma and J.L. Ba, “*ADAM: A method for stochastic optimization*” , Proc. ICLR 2015.
- [66] L.N. Smith, “*Cyclical learning rates for training neural networks*” , Proc. IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 464-472, 2017.