

機械学習を用いた医療用画像診断支援システムの設計と開発 Design and Development of a Medical Imaging Diagnostic Support System Using Machine Learning

三瓶 栄治[‡] 孫 静涛[†] 山口 淳[‡]
Eiji Sambei Jingtao Sun Jun Yamaguchi

1. はじめに

近年クラウドや機械学習の技術進歩により、医療分野の様々な分野に AI 技術が浸透し、特に医療用画像診断支援に関する研究開発の注目が集まっている。日本国内においても、日経リサーチが実施した調査アンケートによると、AI による画像診断技術を求めている医師が圧倒的に多いことがわかった [1]。この背景には、日本において放射線科医が不足している問題や、業務負担が各国と比較して 4.3 倍にも及ぶ問題がある [2]。本研究では、こうした問題に対処する為に、機械学習による医用画像診断支援システムを提案し、開発時の課題と解決ソリューションを紹介する。

医用画像は X 線撮影のような、胸部の 2 次元画像を取得する方法や、CT・MRI のような断層を撮影する方法など、用途に応じて複数存在する。AI を医用画像処理へ適用する手法 [3,4] に関しても、用途に応じ複数存在している。しかし既存の研究では、特定のデータを対象とし、特定の基盤上に、限定的な学習モデルを構築できる手法を提案されているが、医師のニーズに応じて、一般化された診断データの活用や様々な最新の学習モデルを基盤上に適応できる医療診断支援システムに関する設計が少ない。

本研究では、撮影された医用画像の特性や、医師のニーズに合わせて適切な形で診断を支援するシステムを提案する。実装では、個別の医用画像に特化した AI を、コンテナ技術を用いて API 化することにより、様々な医用画像に対応する為の拡張性があるシステムを構築した。また、このシステム上で動作する、医師に対するサービスとして、主に 2 つのサービスを構築した。1 つ目が Grad-CAM [5] の手法で生成したヒートマップを Chat-GPT と組み合わせ、読影診断を補助するサービスである。2 つ目は U-Net によって出力されたマスク画像を元に、医用 3 次元オブジェクトを構築し、Web-UI 上で確認できるサービスである。このサービスは、医師が使用した医用画像の疾患位置確認がしやすくなり、利便性と仕事の効率性を向上できる。

本稿は下記の様に構成されている。2 節では、関連研究を紹介し、3 節においては、診断支援システムのシナリオと提案手法を紹介する。4 節では、システムアーキテクチャの全体像と、プロトタイプの実装手法を説明する。5 節では、実装した Web-UI と AI モデルの可視化結果と性能評価を紹介する。最後に 6 節で、今後の展望と課題に関して考察する。

2. 関連研究

Hicham Moujahid らが起こった研究 [3] では、Grad-CAM を用いて胸部 X 線画像から AI が病気であると診断した根拠を、可視化する実験を行っている。Grad-CAM はモデルの特徴マップを使用し、AI が画像のどの領域をみてクラス分類を行ったのかをヒートマップで可視化する手法である。この手法は、AI が診断した疾患の根拠を可視化する手法として代表的である。

また、代表的な手法として U-Net を用いた手法がある。s0mnaths 氏のソースコード [4] を初めとして、U-Net を使用した医用画像のセグメンテーションに関する研究、ソースコードは Kaggle 上などに複数存在する。

こういった代表的な個別の手法に関する研究や知見が蓄積されてきた一方で、医師に対する診断支援を前提としたシステム全体の統合に関して考慮された研究は限定的である。AI をパッケージ化し、医師に向けたサービスとして展開する総合的手法に関する研究が欠如している。本研究は、これら代表的な AI を用いた医用画像処理手法により出力された結果を、さらに医師に対する診断支援という目的で、医師の利用条件に応じて医療診断支援システムが適切に対応でき、医師の利便性や効率性の向上が期待できる。

3. 診断支援のシナリオと提案手法

本節では、医師が抱えている課題と、その解決ソリューションのシナリオを述べる。

3.1 診断支援のシナリオ

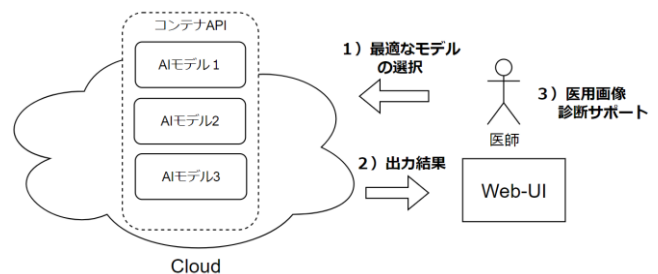


図 1 医療診断支援のシナリオ

図 1 に診断支援に関わるシナリオを示す。図 1 では、医師が遠隔地から医用画像を含んだリクエストを送信している。医師側のコンピュータには AI モデルは搭載されていない。医師がいるローカル環境には GPU を初めとして、

[‡] 東京工科大学 Tokyo University of Technology コンピュータサイエンス学部

[†] 日立製作所 研究開発グループ R&D, Hitachi, Ltd.

AI モデルを動かすための環境が整っていない。医師は様々なタイプの医用画像を元に診断を行う必要がある状況に置かれている。医師は、人手不足と業務過多の為、一人当たりには使える診断(医用画像の読影)時間が限られており、20代から 30 代の比較的若い医師を初めとして、多くの医師がチャット GPT 等の AI 活用による業務効率化を期待している[1]。提案手法により解決される課題は、大きく以下の 3 つである。

- 1) 人手不足と業務過多により、一人当たりにかかる診断時間が少ない。
- 2) AI モデルを動かせるローカル環境がなく、医用 AI のニーズがある[1]にも関わらず医用 AI を導入できない。
- 3) AI モデルの出力結果をそのまま使いこなせない。

課題 1 という大きな問題を解決する為には AI 導入による業務効率化が期待されるが、課題 2,3 がボトルネックになると考えられる。

3.2 提案手法

本節では 3.1 節のシナリオから抽出された課題を解決する為の提案手法について紹介する。本研究では、撮影された医用画像の特性や、医師のニーズに合わせて適切な形で診断を支援するシステムを提案する。また、本研究では各医用 AI モデルをコンテナ化しクラウド上に展開する事で、医師側のローカル環境の状況に関わらず、医用 AI の利用を可能にする。AI の出力結果をそのまま医師に伝えるのではなく診断支援という目的に沿う見やすい形式で提供する。

具体的には Grad-CAM による出力を元の入力された医用画像と合成し、さらに Chat-GPT によって簡易的に診断させることにより、医師が注目すべき領域と注目すべき理由を瞬時に提供する。これは医師の医用画像読影の時間を減らすだけでなく、見落としも防止する。また、U-Net からの出力から医用 3 次元オブジェクトを構築することで、医師は自分が注目している医用画像が、全体のどこに位置しているのかを把握しやすくなる。

4. プロトタイプシステムの設計と実装

この節ではプロトタイプシステムの設計と実装方法について述べる

4.1 アーキテクチャ設計

個別の AI モデルを管理しやすくするために、本研究で提案するシステムは、マイクロアーキテクチャを採用し、個別の AI モデルをシステムから分離しやすくする。図 2 にアーキテクチャの設計図を示す。

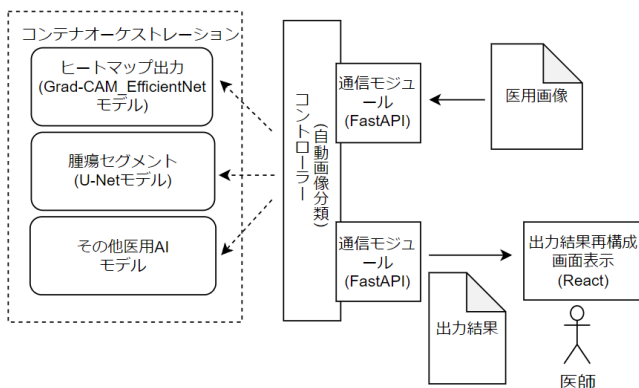


図 2 医療診断支援システムアーキテクチャ

本提案のそれぞれの AI モデルは、Docker を使用してコンテナ化されている。入力された医用画像のタイプに応じて最適な AI 処理モデルが選択できる。モデルの選択に関して機械学習による判別が考えられる。医師が通常関知しない、AI モデルと医用画像の相性を機械学習により判別する事で、医師のモデル選択に関する手間が解消される可能性がある。各コンテナは FastAPI により、API 化されており、医用画像を送信すると、結果をレスポンスする。レスポンスされた結果は、診断支援という目的に沿うようにさらに再構成され、最終的に医師が Web-UI にて確認できるようになる。Web-UI の構築には React を用いて、動的なサイトを実現する。

Docker によりコンテナ化された各 AI モデルは独立して管理が可能である。コンテナ化による管理は、医師のニーズに合わせて新たなモデルを追加したり、モデルを更新したりできるような柔軟な拡張性を提供する。

4.2 実装

各医用 AI モデルの出力結果を、Web-UI によって再構成できることを確認する為に、2 節の関連研究で述べた 2 つの代表的 AI モデルを実装する。

4.2.1 Grad-CAM を用いた医用画像読影実装

データセットは既存の公開されている幼児の胸部 X 線画像[6]を使用し、モデルが肺炎かどうかを 2 値分類する様に学習させる。その後、Grad-CAM[5]を学習させたモデルの最終の畳み込み層に適用していく。機械学習アーキテクチャには EfficientNet[7]を採用した。

Grad-CAM の計算は以下の過程で行われる。

1. 画像を入力し、対象クラスのスコアを出力する。
2. 対象クラスのスコアに対して、最後の畳み込み層の各特徴マップの勾配を計算する。
3. 勾配のグローバル平均を取ったものを、特徴マップのニューロンに対する重要度(1)として使用する。これが重み α_k^c となる。

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \underbrace{\frac{\partial y^c}{\partial A_{ij}^k}}_{\text{gradients via backprop}} \quad (1)$$

global average pooling

- ・ i と j は特徴マップの空間次元である。
- ・ Z は特徴マップ内の要素の数である。

4. これらの重みを用いて、重み付けされた特徴マップの線形結合を計算し、ReLU 関数を適用して最終的なヒートマップを生成する。このヒートマップ(2)が、クラス c の予測に最も影響を与えた画像の領域を示す。

$$L_{\text{Grad-CAM}}^c = \text{ReLU} \left(\underbrace{\sum_k \alpha_k^c A^k}_{\text{linear combination}} \right) \quad (2)$$

この手法を医用画像のクラス分類にも適用する事ができる。肺炎患者の CT 画像で、病気であるかどうかをバイナリ分類するように学習させる。こうして学習されたモデルは、最終層の畳み込みレイヤーの特徴マップが、疾患箇所に対応するようにチューニングされている。Grad-CAM は最終の畳み込み層以外にも適用することが出来るが、一般的には深い層ほど高度な視覚的特徴を捉える能力があるとされるため[8]、通常 Grad-CAM は最も深い畳み込み層に適用される。図 3 は Grad-CAM を用いたヒートマップの出力を示す。図 4 は Chat-GPT に入力したヒートマップと医用画像の合成画像を示す。

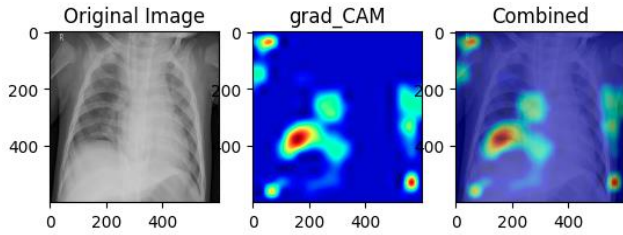


図 3 Grad-CAM を用いたヒートマップ出力

この方法により診断の根拠(疾患の場所)がヒートマップで出力することが可能となる。図 3 の左の画像(Original Image)は入力画像であり、真ん中の画像(grad-CAM)は出力されたヒートマップである。右の画像(Combined)は入力画像とヒートマップを合成した画像である。

これらの画像を Chat-GPT4 に入力し簡易的な診断をする様に指示をする(図 4)。なお、入力画像と共に Grad-CAM を使用したヒートマップであることを Chat-GPT4 に伝えた。

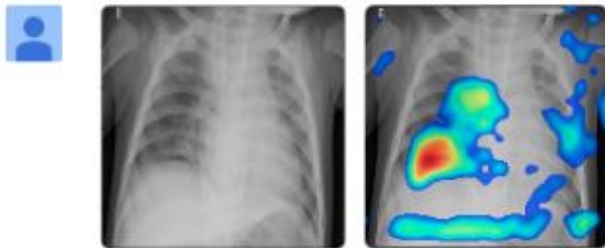


図 4 Chat-GPT に 2 つの画像を入力する

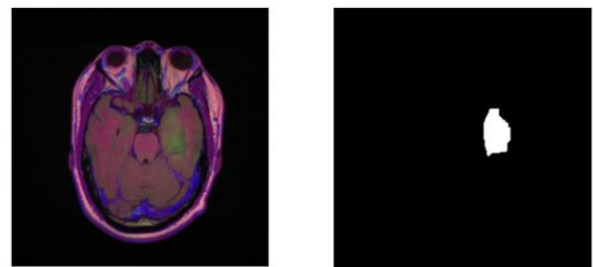
これらの入力画像に対して、医学的な説明を行うように指示したところ、Chat-GPT は診断の結果を出力した。具体的には、胸部 X 線画像上のこのヒートマップに示される領域は、肺の下葉や中葉に相当する部分に集中している。肺炎の場合、しばしば肺の下葉に症状が現れる事が多いため、この結果は肺炎の典型的な所見と一致している。

また、入力情報に、医学的な知見は含めなかった状態で、右肺にしかない中葉の位置を認識することや、肺炎の症状と関連付けて診断している事は注目に値する。厳密な診断に関して、精度の課題はまだあるが、大まかに診断で注目すべき領域を医師に提示できるため、業務効率化に貢献する。

4.2.2 U-Net を用いた医用 3 次元オブジェクト構築による診断補助

U-net[9]は、画像セグメンテーション（画像内のオブジェクトの領域を識別するタスク）に特化しており、医療画像のセグメンテーションに活用されている。

このセグメント技術を応用し、脳腫瘍の 3 次元オブジェクトを構築する手法を考案する。U-net で脳腫瘍をセグメントするように学習されたモデルで、MRI により断層的に撮影された脳画像を処理すると、層状に脳腫瘍をセグメントすることができる。図 5 は脳画像とマスクの出力結果を示す。図 5 の画像 a は、MRI で撮影された患者の脳の断面画像である。この断面画像を U-net で学習したモデルに入力すると、画像 b のような脳腫瘍の領域がセグメントされたマスク画像が生成される。今回は公開されている脳腫瘍のデータセット[10]を使用する。



(a) 入力画像

(b) マスク画像

図 5 脳画像とマスク

層状に撮影した医用画像なので、セグメントしたマスク画像を縦方向に連結していけば、3 次元オブジェクトが構築できる。今回作成した 3 次元オブジェクトは、腫瘍を赤色、脳全体を青色の点群データで表現されている(図 6)。

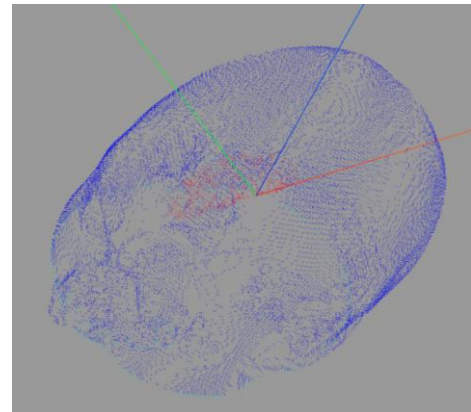


図 6 脳全体と腫瘍の 3 次元オブジェクト

5. 評価

この節では、実装した医療診断支援システムの可視化結果と実装した AI モデルの初歩的な性能評価を述べる。

図 7 は、4.2.1 節の Grad-CAM を用いた AI モデルの出力結果が最終的に Web-UI 側で再構成された例である。左側の医用画像に関する Chat-GPT による説明が、右の領域に表示されている。左下の操作バーはヒートマップの透過具合を自由に変更できる仕様になっている。

医師は診断で注目すべき領域を瞬時に見極めることができ、医用画像読影時間の短縮に貢献すると考えられる。



図 7 Grad-CAM と Chat-GPT を使用した診断支援 Web-UI の例

図 8 の画像は、4.2.2 節の U-Net を用いた AI モデルの出力結果が最終的に Web-UI[11]側で再構成された例である。左領域 a で医用画像を選択すると、右領域 b の医用 3 次元オブジェクトの対応するレイヤーが、シアン色に発光して脳全体のどこの断層を覗いているのかを把握できる。

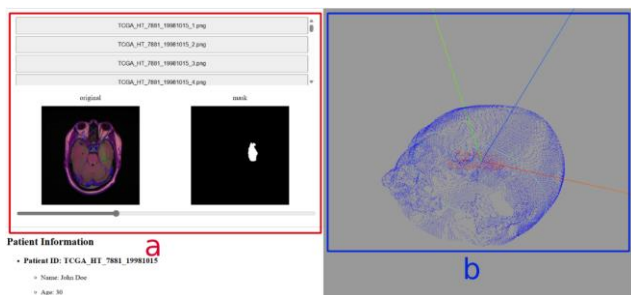


図 8 U-Net を用いた医用 3 次元オブジェクトによる診断支援 Web-UI の例

今回作成された 3 次元オブジェクトは、断層撮影の間隔が比較的広く、密度が薄いデータセットであったが、脳全体の中に占める腫瘍の密度が視覚的に表現されている。この表現は、医者の読影を支援すると考えられる。

表 1 は今回作成した、AI モデルの性能比較である。それぞれのモデルのアーキテクチャに応じて、評価指標が異なるので、単純に比較できない点は注意が必要である。作成された 2 つのモデルとも、学習データ量は大規模ではないものの、精度において高い指標を示した。

表 1 今回作成した AI モデル性能比較

AI モデル	学習データ	精度	損失関数
EfficientNet-Grad-CAM	約 5800 枚	0.89	2 値交差エントロピー
U-net	約 4000 枚	0.96	Dice 損失+BCELoss

今回 Docker によって、各 AI モデルをコンテナ化したが、AI を搭載したコンテナは起動に時間がかかる問題が生じた。機械学習ライブラリの立ち上げに時間がかかっていると考えられる。また、入力画像を API に送信してから、AI が処理した出力が返ってくるまで、ラグが生じた。今回は CPU による処理で実装した為、計算に時間を要したと考えられる。今後は GPU などと連携し、タイムラグを減少させ、実用に耐える設計にする必要がある

6. まとめと今後の展望

本研究では、機械学習を用いた医療診断支援システムを提案した。医用 AI モデルにより、医用画像読影の時間を短縮することが可能となる。このシステムはクラウドを効果的に利用しており、医師側のローカル環境に計算リソースを要求しない。プロトタイプで医用 AI モデルのコンテナ化を実装し、作成した API エンドポイントを通して、医用画像が適切に処理されたことを確認した。出力された結果をもとに、医師の診断を支援する目的で、効果的な Web-UI を実装した。

今後の課題としては、5 章でも述べたコンテナの処理時間の短縮に加え、学習データ量を増やして医用 AI モデルの汎化性能を高めることが必要である。また、研究環境の特性上、医療関係者と直接連携を取れなかったことは本研究の限界であり課題である。今後、医師が実際に求めているニーズを調査し、より診断の効率化に貢献できるように Web-UI の構成を改善していく必要がある。

謝辞

本研究にご協力いただいた、鈴木飛鳥氏、平野達也氏、山本泰生氏、羅婷氏に感謝申し上げます。

参考文献

- [1] 日経リサーチ, “期待高まる AI、それでも 8 割の医療機関は未導入。”, https://service.nikkei-r.co.jp/report/healthcare_id191, (accessed 2024-6-10).
- [2] 中島 康雄, 山田 恵, 今村 恵子, 小林 和子, “Radiologist supply and workload: international comparison”, *Japanese Journal of Radiology*, Vol.26, pp.455-465, (2008).
- [3] Hicham Moujahid, Bouchaib Cherradi, Oussama El Gannour, Wamda Nagmeldin, Abdelzahir Abdelmaboud, Mohammed Al-Sarem, Lhoussain Bahatti, Faisal Saeed, Mohammed Hadwan, “A Novel Explainable CNN Model for Screening COVID-19 on X-ray Images”, *The Computer Systems Science and Engineering journal*, Vol.46, No.2, pp.1789-1809, (2023).
- [4] s0mnaths, “BrainMRI-UNET-PyTorch”, <https://www.kaggle.com/code/s0mnaths/brain-mri-unet-pytorch>, (accessed 2024-6-10).
- [5] Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, Dhruv Batra, “Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization”, *International Journal of Computer Vision*, Vol.128, pp.336-359, (2019).
- [6] Daniel Kermany, Kang Zhang, Michael Goldbaum, “Large Dataset of Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images”, <https://data.mendeley.com/datasets/rschbjr9sj/3>, (accessed 2024-6-10).
- [7] Mingxing Tan, Quoc V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks”, *International Conference on Machine Learning*, Vol.97, pp.6105-6114, (2019).
- [8] Matthew D Zeiler, Rob Fergus, “Visualizing and Understanding Convolutional Networks”, *Computer Vision—ECCV*, pp.818-833, (2014).
- [9] Olaf Ronneberger, Philipp Fischer, Thomas Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation”, *MICCAI, Medical Image Computing and Computer-Assisted Intervention*, pp 234–241, (2015).
- [10] The Cancer Imaging Archive, “Brain MRI segmentation”, <https://www.kaggle.com/datasets/mateuszbudza/lgg-mri-segmentation>, (accessed 2024-6-10).
- [11] Eiji Sambei, “react-unet”, <https://github.com/medical-ai-project/react-unet>, (accessed 2024-6-10).