

原 著

脳形態 MRI を用いたアルツハイマー病臨床診断の為の深層学習モデルの構築
と判断領域の可視化

武田 基秀

新潟大学大学院医歯学総合研究科 生体機能調整医学専攻 脳機能解析分野
(主任：五十嵐 博中 教授)

Construction of a deep learning model for clinical diagnosis of Alzheimer's disease using
morphological brain MRI and visualization of discriminative decision area

Motohide Takeda

Department of Integrated Neuroscience, Brain Research Institute,
Graduate School of Medical and Dental Science, Niigata University

Reprint requests to: Motohide Takeda

Center for Integrated Human Brain Science, Brain Research Institute, Niigata University,
1-757 Asahimachi-dori, Chuo-ku, Niigata City, 951-8585, Japan
Mail: tkd.institute@gmail.com

別刷請求先：〒951-8585 新潟市中央区旭町通り 1-757

新潟大学脳研究所統合脳機能研究センター 武田 基秀
メールアドレス：tkd.institute@gmail.com

要 旨

【緒言】

アルツハイマー病は徐々に進行する認知障害のため社会適応が困難となり、その介護の負担を中心として大きな社会問題となっている。一方、人工知能の医療への活用は進展しており大きな変革を生み出す途上にある。本研究は、最適化した深層学習を脳形態の磁気共鳴画像法（MRI）に用いることにより、高い判別性能を持った診断補助システムを構築するとともに、通常ブラックボックスである判断根拠を可視化することにより、臨床補助システムとして臨床診断に付加価値を付与することのできる画像判別システムを構築し判別性能の評価を行った。

【対象と方法】

対象は 2015/4/6 から 2017/6/13 までの期間に撮像された 68 例(正常例 (Control) 36 症例、AD 症例 32 症例)の MRI 形態画像を用いた。事前学習された VGG16 及び ConvNeXtTiny(CNXT)を最適化した深層学習モデルを使用して、鑑別診断能を検証した。さらに Grad-CAM、SmoothGrad を用い本研究の深層学習モデルが AD と判断した根拠について、重み付けの高い部位の可視化を行い判断根拠の明確化を行った。

【結果】

診断能の評価において、VGG16 を用いた場合、受信者動作特性 (ROC) 分析の総合評価指標である AUC 値は 0.990、正解率 0.943、再現率 0.900、適合率 1.000 であった。次に、CNXT を用いた場合、AUC 値は 0.986、正解率 0.986、再現率 0.983、適合率 0.992 であった。重み付けの高い部位の可視化では、海馬、側脳室といった AD において早期から萎縮をきたす部位の重みが高かった。

【考察】

本研究の結果について、最近の類似研究の結果、臨床医の判別精度と比較して、データが少ないにもかかわらず良好な性能を達成でき、臨床の補助システムとして医療に付加価値を与えうる潜在力があることの立証になったと考えた。CNXT と SmoothGrad を使用することにより、可視化において安定した良好な結果を得ることができ、臨床と本技術の活用の中で価値を生み出す潜在性を示唆していると考えられた。

課題として、実践において臨床的価値を付加するために、臨床の現場と共同して本技術の更なる改良が必要であること、並びに、画像の質に対する堅牢性を担保する必要性から、データを拡大した研究が必要であることがあげられる。

【結論】

深層学習モデルを使用した脳形態の MRI の画像診断を用いることにより、良好な診断性能を持った判別システムを構築することができた。さらに判断根拠を可視化してその情報を供給し、臨床に診断補完システムという付加価値を与えることが可能であることを示すことができた。

Key Words: Alzheimer's disease, Brain magnetic resonance imaging, Deep Learning, Convolutional Neural Networks, Explainable AI

緒 言

日本における認知症の患者数は2012年に462万人、その内アルツハイマー病 (Alzheimer's disease)¹⁾ (AD)は313万人と調査され、現在、認知症600万人、AD400万人と推計されている。²⁾ ADを代表疾患の一つとする認知症は、人間が知的な発達を遂げ獲得された知能・適応性などの能力が失われ日常生活に困難を来す疾患である。臨床症状として、主にエピソード記憶障害の他、失行、失認、実行機能障害などの認知障害を伴う。認知症の原因疾患として他にびまん性レビー小体病、前頭側頭型認知症、血管性認知症などがあるが、ADは認知症の67.7%と多数を占めると報告されている³⁾。ADは増加傾向にあり²⁾、徐々に進行する認知障害のため社会適応が困難となり、その介護の負担も含め大きな社会問題となっている。政府は、令和元年に、認知症施策推進大綱を定め、2025年までの活動計画を示している²⁾。

ADの臨床診断は、診断基準に基づき、記憶・認知機能のテスト、改訂版長谷川式簡易認知評価スケール、ミニ・メンタルステート試験、アルツハイマー評価スケール、ミニ・メンタルステート試験、アルツハイマー型認知評価スケール、ウェックス成人知能検査、高齢者うつスケールなどが用いられる。これらを用いた時のAD診断精度は感度81%、特異度70%と十分でない。⁴⁾ また、MRIによるADか否かの診断能はMCIの早期を含めても、感度84%、特異度74%と報告されている。^{5) 6)}そのため臨床徴候、画像診断、生物学的マーカーなどを組み合わせて診断することが推奨されているが、ADの確定診断は病理診断であること、また各認知症疾患に重複が認められることも少なからずあり日常臨床での診断が困難な症例も存在する。特に、アミロイド β (A β)-PET⁷⁾、Tau-PET⁸⁾、髄液A β 42および髄液tau、髄液リン酸化tau⁹⁾などの研究的診断法が実用化されてきた現在、これらを用いた診断と臨床診断には乖離が認められることが指摘されている。しかし、これらの診断法は、いずれも実用化の途上であること、費用の問題ないしは侵襲性より広く日常臨床に用いられるには種々の障壁がある。以上より日常臨床で用いることができる非侵襲的かつ客観的な指標を開発することが喫緊の課題である。

ADにおけるMRIは、現状では診断より脳梗塞などの他の認知障害をきたす疾患の除外の為にこなわれることが多い。一方、ADの病理はA β 42の凝集による老人斑とリン酸化tauの凝集による神経原繊維変化の出現、および神経細胞の脱落によって特徴つけられるが、この所見は海馬、海馬傍回を含む側頭葉内側部、および側頭頭頂連合野で病初期より認められ、かつその侵襲が強い。その結果として形態学的に同部位の脳実質の萎縮をきたす(図1)。したがって、これらの特徴的な形態学的変化を判断できる客観的な指標があれば臨床における診断の補助として用いることが可能である。^{10) 11) 12)}

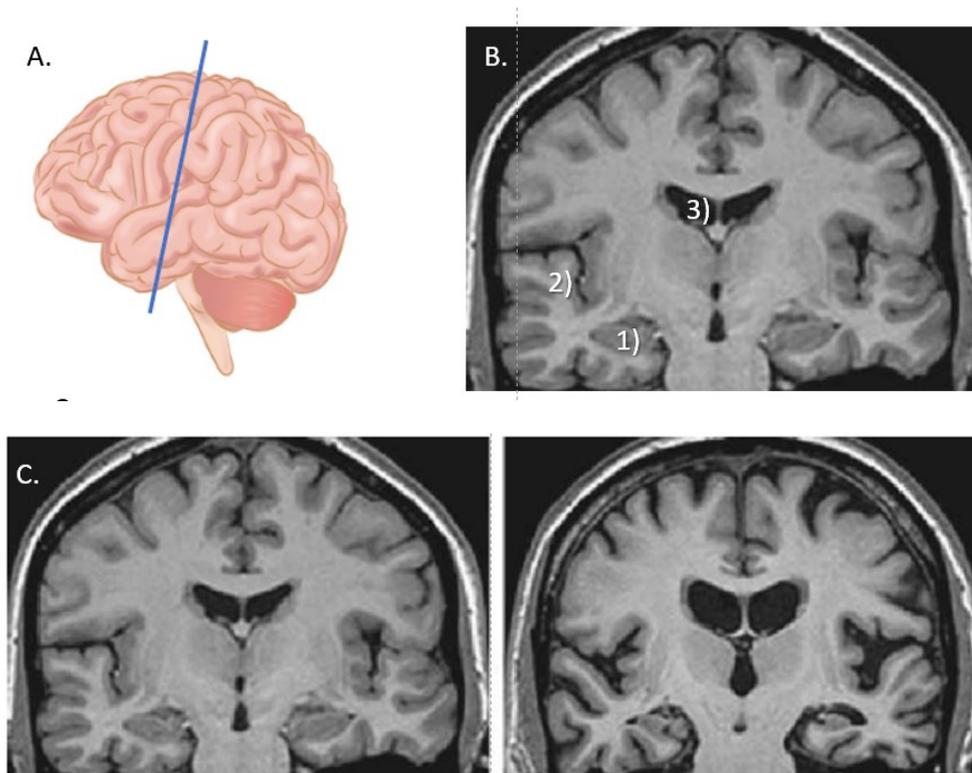


図 1 ADにおける代表的な形態変化と機械学習に用いた MRI 画像

A. 3D-MRI 形態画像より切り出した海馬を中心とした断面。B. 学習・判定に用いた断面の正常例。 C. 正常例および AD 例の代表的な画像。AD では B. で数字を振った 1) 海馬、2) 側頭葉上葉の萎縮及び、3) 側脳室の拡大が認められる。

AD の確定診断は病理所見によってなされ、磁気共鳴画像法 (MRI) は臨床診断に付加価値を与えるが、臨床診断において最も重視されるのは認知障害をはじめとした臨床徴候である。一方、臨床において人工知能の活用が進んでおり¹³⁾、臨床現場での研究開発と活用の不断の努力が医療の更なる進展による価値創出のための重要な基盤¹⁴⁾¹⁵⁾と考える。斯かる状況において、AD の臨床診断の補助として、人工知能の学習済み深層学習モデルを使用した脳形態の磁気共鳴画像法 (MRI) の画像診断を用いることが可能であるかを検証し、更に、通常ブラックボックスである深層学習モデルの判断根拠を可視化することにより、説明性を付加することが可能かを検討した。深層学習は、画像認識技術を基礎に、対象を認識、分類することに優れた人工知能の技術である。その中で、広く使われる畳み込みニューラルネットワーク (Convolutional Neural Networks (CNN)) は、訓練データによる学習課程で分類・判別能力を蓄積するという優位性を持つアルゴリズムである¹⁶⁾。実施済みのアルゴリズムを他の問題にオープンに活用し成果を上げていく転移学習 (Transfer Learning) の技術により、2014 年に発表され評価が定着している CNN の応用

手法である VGG16¹⁷⁾並びに現在活用が拡大している ConvNeXtTiny (CNXT)¹⁸⁾が広く用いられている。

本研究においては AD 及び年齢適合認知機能正常例(Control)の MRI 形態画像を用い、VGG16 および CNXT に学習させ、これら CNN を最適化し本研究に用いるために微調整し修正された modified CNN (mCNN) を構築し、これを用いて判別および判別性能の検証を行った。次に、いかなる特徴を判断根拠しているかについての可視化を、Grad-CAM++¹⁹⁾と SmoothGrad²⁰⁾ を実装することにより行い、mCNN が捉えている特徴を Heat Map (HM) で表示させ評価した。

対象 と 方法

本研究は北日本脳神経外科病院の臨床審査委員会の承認(「統計学的 MRI 画像処理および AI を用いた正常加齢と認知症の鑑別診断、承認番号 2017-01)および新潟大学倫理委員会の承認(「研究課題名:人工知能を用いた MRI による認知障害診断アルゴリズムの開発」、承認番号 2018-0334)を受けて施行された。MRI 形態画像は、2015 年 4 月 6 日から 2017 年 6 月 13 日までの期間に、北日本脳神経外科病院(現:五泉中央病院)にて撮像された 77 症例(男性 28 例、女性 49 例)の匿名化した MR 画像データベースから、MCI の 9 例を除外し、残りの 68 例を用い、その内訳は、正常例が 36 症例(男性 8 例:平均年齢 72.75±10.73、女性 28 例、平均年齢 73.35±5.09)並びに、AD 症例 が 32 症例(男性 15 例、平均年齢 80.93±5.24、女性 17 例、平均年齢 79.59±6.45)である。AD の臨床診断は米国国立老化研究所およびアルツハイマー協会(NIA-AA)が策定した dementia due to Alzheimer's disease の診断基準中、definite AD を満たすものとした²¹⁾。NIA-AA の definite AD 診断基準は、認知症があり、数ヶ月から年四に緩徐進行、認知機能低下の客観的病歴がある、次の 1 つ以上の項で病歴と検査に明らかに低下している: 健忘症状、非健忘症状:失語、視空間機能、遂行機能、次の所見がない場合:脳血管障害、Levy 小体型認知症、前頭型認知症、進行型失語症、他の内科・神経疾患の存在、薬剤性認知障害、とされている。正常例は脳ドック目的ないしは頭痛等で受診した認知機能正常例とした。

深層学習モデルの構築および基本訓練として、Ubuntu18.04.5LTS オペレーティングシステムと NVIDIA GeForce RTX2080Ti GPU(NVIDIA Corporation、カリフォルニア州、USA)を搭載した PC 上において、python 3.9.15、tensorflow 2.10.0、keras 2.10.0 を用いて実行した。VGG16 および CNXT の深層学習モデルの最適化手法の調整を行い、最適化した。さらに、最適化された深層学習モデルが画像中のどの部位に判別根拠の重み付を有するかについて可視化手法を用いて検証した。

深層学習モデルの構築および認識効率の評価のための、MRI 形態画像データの撮像条件および画像処理法は以下の如くである。撮像には Siemens 社製 3T MRI (Magnetom、Siemens-AG、Tubingen、Germany)を用い 3D-MPRAGE にて全脳形態画像を取得した。撮像条件は、

TR/TE=2300/2.96 ms、T_{eff} 900 ms、echo spacing time 7.1 ms、bandwidth 240 Hz/pixel、FOV220x220x150mm、256x256x128matrix、Voxel size 0.898 x 0.898 x 1.3 とした。得られたMPRAGE の 3D-dataset から、Matlab (Version R2022a, MathWorks, USA) を用い homemade のスクリプトを用いて、脚間窩を中心に 1mm スライス厚で 1 例あたり 9 枚から 15 枚の状断スライスを切り出した (図 2)。画像は、130*130pixel で切り出した。

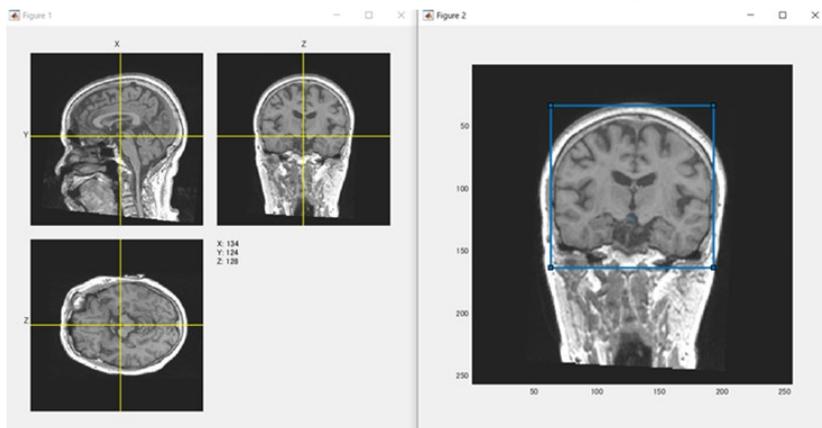


図 2 画像切り出し

さらに、頭蓋骨が判別並びに判別根拠の可視化に干渉することが指摘されており¹⁴⁾、頭蓋骨の影響を軽減するため頭蓋骨除去したデータセットを作成した。頭蓋骨除去には、数値計算ソフト Matlab (Version R2022a, MathWorks, USA) を用いて作成した homemade の Matlab スクリプトを使用し、画像を SPM12 (Wellcome Trust Centre for Neuroimaging, UK) の分割化 (Segmentation) アルゴリズムに基づき、灰白質、白質、脳脊髄液、頭蓋骨、軟部組織に分割し、画像から頭蓋骨をマスクすることにより行った。

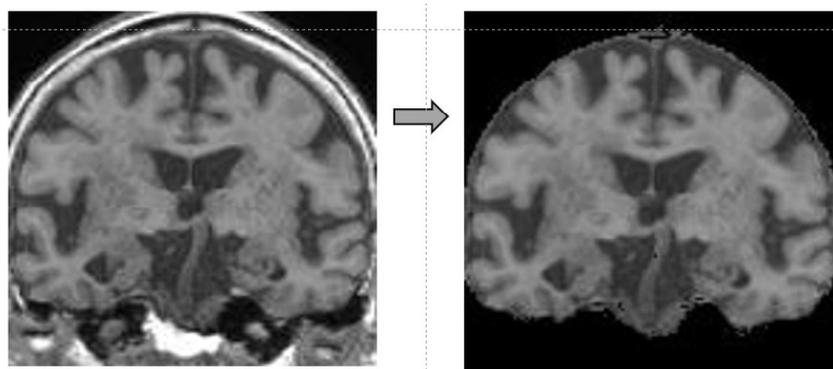


図 3 頭蓋骨除去：左図は除去前、右図は除去後である。

最終的に、アルツハイマー症例及び正常認知機能例、併せて 1020 枚の画像データセットを作成した。いずれのデータセットにおいても、全データセット中の 80% (810 枚) を訓練用に使用し、学習(training、AD:225、Control:285)と学習結果評価 (validation、AD:135、Control:165)、独立した残り 20% (210 枚) を最終的な性能検証(test、AD:120、Control:90)

に使用した。

CNN アルゴリズムの選択について、Case A は、VGG16 を使用した mCNN による検証とし、図 3 に示すモデル構造を作成して実験に使用した。Case B は、CNXT を使用した。深層学習の技術の進化において、100 万枚を超えるデータである Imagenet を使った判別精度を競うコンペティションが毎年行われ、VGG16 は 2014 年に 2 位となり、そのシンプルな構造から汎用使用されている。それ以後の進化の事例を図 4 に示すが、CNXT は Attention を使用した Encoder-Decoder モデルを導入することにより Vision transformer と同等の判別能を持つ CNN であり²²⁾、Case B に採用したものである。

Case A mCNN (VGG16)

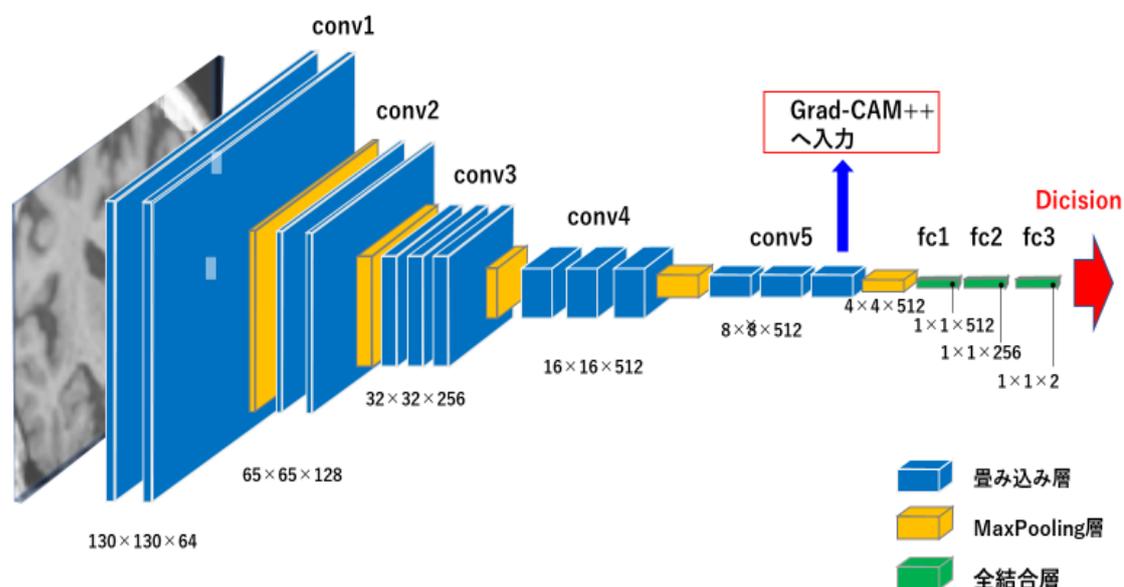
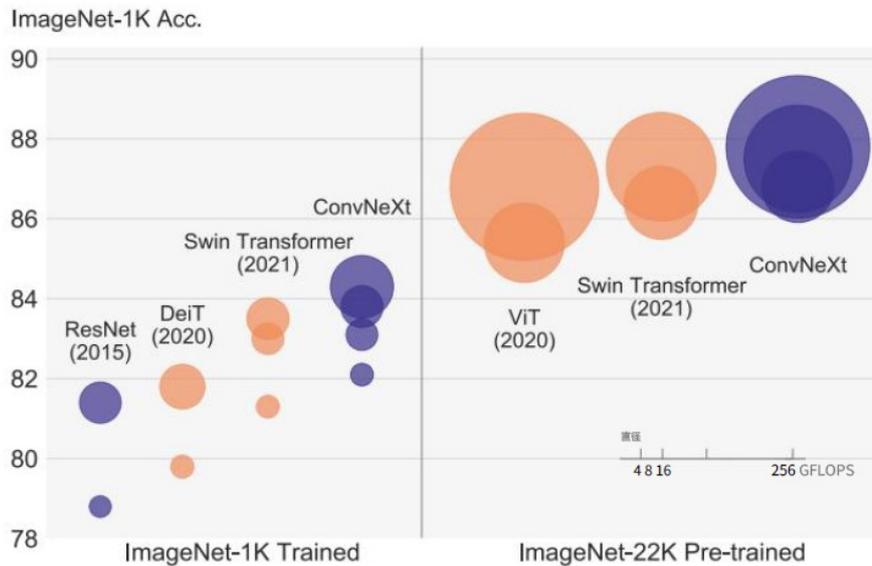


図 3 Case A: mCNN (VGG16) のモデル構造¹⁶⁾¹⁷⁾

畳み込み層 (Conv2) と MaxPooling 層の重ね合わせ構造をもち、conv5 までが VGG16 の基本構造で、それに自作の層を加えた。conv4 までは、VGG16 の学習済みの重みをそのまま固定して使い、それ以後は、今回の訓練での学習を反映させた。判断根拠の可視化において、Grad-CAM++では conv5 の最終層の情報を取り、SmoothGrad は Input と Output のデータを用いて感受性を分析することにより判別の重み付の高い領域を HM で表示した。



引用元：Zhuang Liu et al: A ConvNet for the 2020s, Figure 1¹⁸⁾

図4 Case-B CNXT 技術の位置づけ

mCNN の最適化について、プログラムの稼働での学習効率と性能を上げ、過学習を抑制するために、バッチ数、最適化関数(Adam、RMSprop、SGD の選択)¹⁶⁾、損失関数、学習率、エポック数、early stopping 導入等の設定と調整を行い最適なシステムを探索した。Case A のセッティングは、最適化関数は Adam、RMSprop、SGD のすべてを使い、損失関数は binary cross-entropy、batch size は 16、学習率は Adam および RMSprop では $1 \times e^{-7}$ 、SGD は $1 \times e^{-3}$ と小さくゆっくり遷移するようにし、最大 epoch 数 2000 とした。また、図 3 に記載の通り VGG16 の固有の重みは conv4 までこの学習で変化しないように固定し、最適化する層は、それ以降と増設部分とした。今回の場合、画像の画素数が小さいため、既存のデータから新しく作成された合成データを追加することによって、データの量を増やすために使用される手法である画像拡張を、回転(10 度)、平行移動(なし)、拡大・縮小(0.1)、左右反転(有り)として適用した。Case B のセッティングは、最適化関数は AdamW を使用し、損失関数は Sparse Categorical Crossentropy、batch size は 32、学習率は $1 \times e^{-4}$ 、最大 epoch 数 2000 とした。また、すべての層が学習する設定とした。画像拡張は回転(0)、平行移動(0.2)、拡大・縮小(0.1)、左右反転(0.5)を設定した。更に Cutout(0.25)で画像に欠損(dropout)を与え拡張と同様の効果も試した。

診断性能の評価については、Python のオープンソース scikit-learn を使い得られた ROC 曲線の AUC 値、並びに。図 5 に示す性能指標で評価した。

	AD判定	Control判定
AD画像	TP	FN
Control画像	FP	TN

図5 性能指標

正解率 (Accuracy) :

$$(TP+TN)/(TP+FP+FN+TN)$$

再現率(感度) (Recall) : $TP/(TP+FN)$

適合率(精度)(Precision): $TP/(TP+FP)$

F1-値: $(2 \times \text{適合率} \times \text{再現率}) / (\text{適合率} + \text{再現率})$

特異度 (Specificity) : $TN/(FP+TN)$

深層学習の判断根拠の可視化については、Grad-CAM++と、SmoothGrad の異なる2方法を使用した。訓練時の学習評価時の画像からランダムに選ばれた30枚、テスト時のAD画像、Control画像、加えて分類エラー（誤判断）を起こした画像を、それぞれランダムに30画像を選び、判別の重み付の高い領域をHMで表示し、MRI形態画像に重ね合わせ出力させたもので評価した。

結 果

診断性能の評価の結果について、Case-A(VGG16)は、図6に示す学習曲線で実行され、判別性能はROC分析のAUC値は0.990(図7)、正解率(accuracy): 0.943、再現率(recall) : 0.900、適合率(precision): 1.000、FI-measure値: 0.947の、ADとControlを判別する良好な結果を示した(表3)。次に、Case B (CNXT)は、ROC分析AUC値は0.986、正解率: 0.986、再現率: 0.983、適合率: 0.992、FI-measure値: 0.987と同様に良好な結果を得た(表3)。

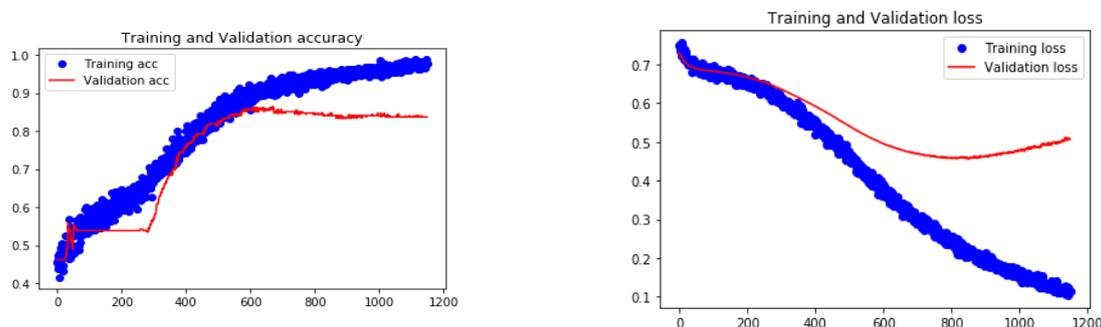


図6 CaseA 学習曲線
VGG16 (Adamの場合)

(左図: Training Accuracy、Validation Accuracy、右図: Training Loss、Validation Loss)

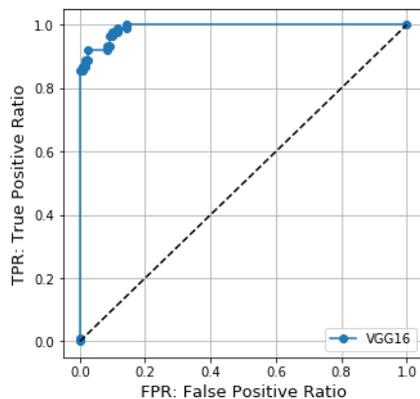


図 7 Case A ROC 曲線 (Adam の場合) : AUC 値: 0.990

表 3 診断性能の評価結果 (縦: 性能指標 横: 最適化関数)

最適化関数	Case A (VGG16)			Case B (CNXT)
	Adam	RMSprop	SGD	AdamW
ROC分析AUC値	0.990	0.989	0.991	0.986
正解率(accuracy)	0.943	0.929	0.929	0.986
再現率(recall)	0.900	0.883	0.883	0.983
適合率(precision)	1.000	0.991	0.991	0.992
適合率と再現率 の調和平均(FI- measure)	0.947	0.934	0.934	0.987

mCNN がいかなる特徴をその判断根拠しているかについての可視化した結果は、Case A では、Grad-CAM++を使用した結果の代表的事例について、図 8、図 9、図 10 に示す。図 11 は、SmoothGrad を使用した事例を示している。Case B では、CNXT と SmoothGrad を用い、AD 正解例を図 12、Control 正解例を図 13、誤判断例を図 14 に示す。いずれも、海馬や海馬傍回の萎縮縮、側副溝の拡大を捉えている。

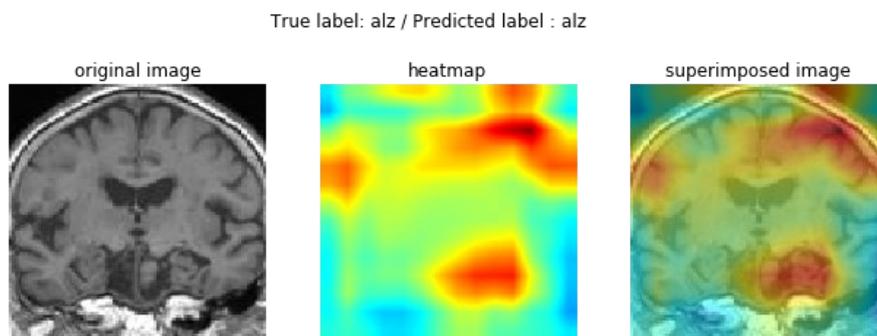


図 8 Case A(1) VGG16(Adam)の場合の Grad-CAM++ による HM
 海馬の萎縮、頭頂葉皮質の萎縮を検知し高いクラス活性化度を HM で示している。

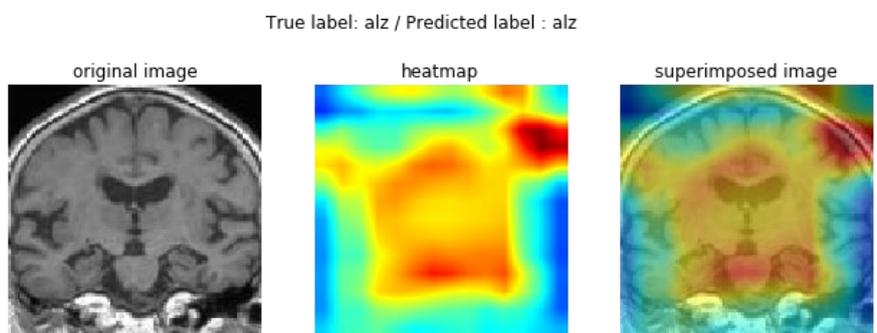


図 9 Case A(2) VGG16(RMSprop)の場合の Grad-CAM++ による HM
 海馬の萎縮、脳室拡大、頭頂葉皮質の萎縮を検知している。

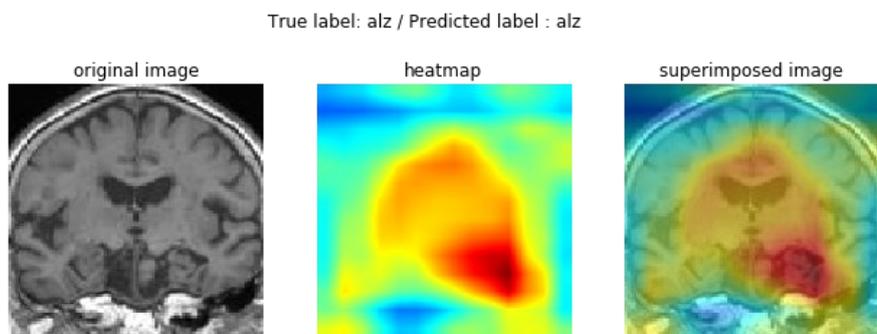


図 10 Case A(3) VGG16(SGD)の場合の Grad-CAM++ による HM
 海馬の萎縮を検知し、皮質は異常があるとはとらえていない。

True label: 0 / Predicted ALZ : 0.9257074 / Predicted NORM : 0.074292585

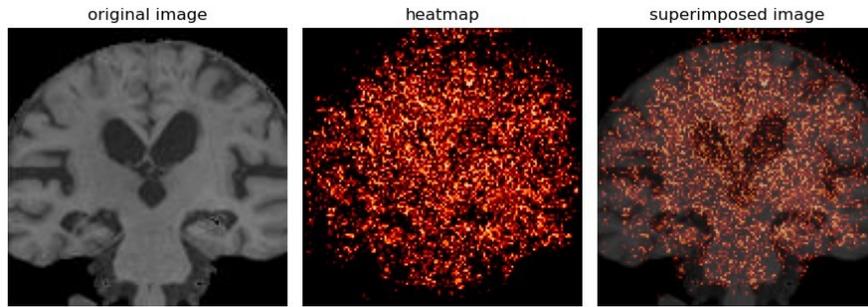


図 1.1 CaseA(4) VGG16(Adam)の場合のと SmoothGrad による HM
全体にクラス活性化を示す HM が見られ、特徴領域を特異的に示してはいない。

True label: 0 / Predicted ALZ : 0.9502011 / Predicted NORM : 0.04979894

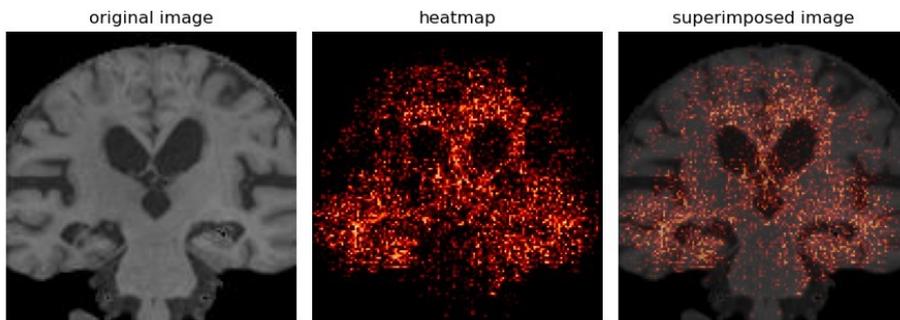


図 1.2 Case B(1) CNXT と SmoothGrad AD 正解例の HM
AD 画像を 95%の確度で AD と正解した事例で、HM は判別した根拠の特徴として AD で病初期より萎縮の認められる海馬、側頭葉上葉及び、側脳室を捉えている。

True label: 1 / Predicted ALZ : 0.018197618 / Predicted NORM : 0.9818024

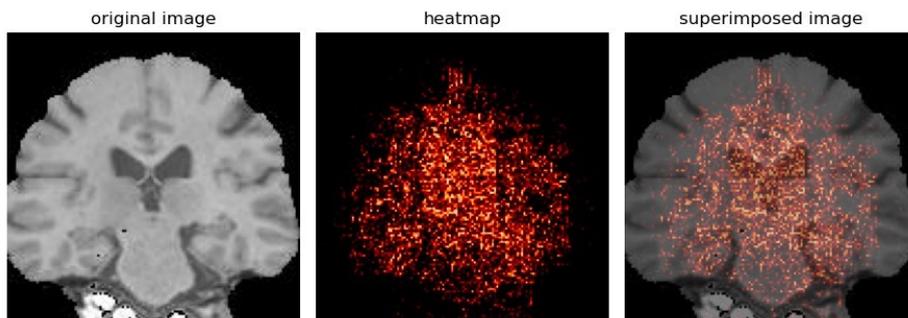


図 1.3 Case B(2) CNXT と SmoothGrad Control 正解例の HM
Control 画像を 98%の確度で Control と正解した事例。

True label: 0 / Predicted ALZ : 0.4624982 / Predicted NORM : 0.5375018

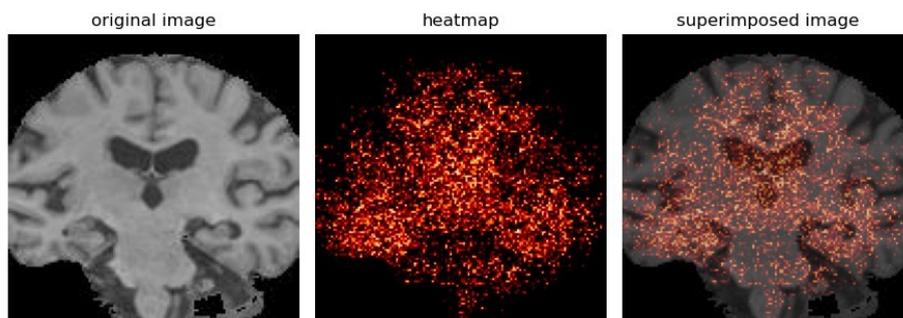


図 1 4 Case B(3) CNXT と SmoothGrad Control 誤判断例の HM
AD の臨床画像を 54%の確度で Control と誤判断した事例。

考 察

本研究の結果の判別性能と関連する研究の結果との比較について、Case-A (VGG16) において、正解率:0.943、再現率 (感度) 0.900、適合率(精度):1.000、特異度:1.000、ROC 分析における AUC 値:0.990 であり、CaseB (CNXT) において、正解率:0.986、再現率:0.983、適合率:0.992、特異度:0.989、ROC 分析 AUC 値:0.986 であった。最近の類似研究の結果を見ると、Modupe Odusami et al²³⁾は、Alzheimer's Disease Neuroimaging Initiative (ADNI)のデータを用い、ResNet18 をニューラルネットワークとして使用し、正解率 : 0.808, 感度 0.918、特異度 0.838 を得ている。Atif Mehmood et al²⁴⁾ は、ADNI のデータを用い、VGG を使用し、正解率 0.987、感度 0.982、特異度 0.991 を得ている。QiLi et al²⁵⁾ は、ADNI のデータを用い、VGGNet を使用し、正解率 0.950、感度 0.914、特異度 0.994、AUC 値 0.994、更に ResNet により、0.95/0.93/0.96/0.954 を得ている。Kirti Raj Bhatele et al²⁶⁾は、VGG16、VGG19 及び ResNet50 を使い、得られた VGG19:正解率 0.90 感度 0.70、特異度 1.00、F1score 0.83、VGG16: 0.88/0.69/1.00/0.81、ResNet50:0.81/0.63/0.64/0.64 の結果より、調整した VGG19 を良好な結果であるとしている。これら最近の類似研究の結果との比較において、本研究の結果は同等若しくは優位性を持つと考えられた。本研究のデータの少なさを考慮すると、少ないデータを用いても十分な判別能力を有すると考えられた。

一方、臨床医の判別精度に関しては、伊藤健吾ら⁵⁾は、MRI による AD と Control の判別においては側頭葉内側部に委縮の評価が主体となり診断能は(MCI を含む)感度 84%、特異度 74%を示したと報告している。日本神経学会診断ガイドライン認知症疾患診療ガイドライン 2017¹⁾では、AD と Control の判別性能を、20 研究を調査した結果として、感度 0.868(0.95CI 0.819~0.917)、特異度 0.787(0.95CI

0.703~0.871)とし、MRIを用いた12研究においてADとControlの鑑別は感度0.850、特異度0.88としている。また、2006年に著者が調査し得た医療における認知症の診断状況の資料では、認知症の診断はその診断基準に基づき各種テスト用いた時のAD診断精度は特異度70%、感度81%と聴取したことを示した⁴⁾。

以上の検討から、本研究の結果の判別性能は、データが少ないにもかかわらず良好な性能を達成でき、臨床の補助システムとして医療に付加価値を与える潜在力があることの立証になったと考えた。

次に、mCNNの最適化において、小さいデータ量という弱点への対処では、ここ10年間で実力を示したCNNアルゴリズムの選択では、VGG16を検討の基礎に置き、その後創出されたResnet、Efficientなどを試行しながら、最終的に最新の発展版であるConvNeXt(CNXT)を使い¹⁸⁾、これを用いることにより、小さいデータ量のため起こる過学習による性能の劣化を抑制して良好な結果となった。

深層学習の判断根拠の可視化と分析について、可視化技術毎の実験結果は以下の通りであった。Case AのGrad-CAM++は代表的な例での可視化を示した(図8、9、10)。但し、可視化は一部において結果を示すことはできたものの、HMが特定な場所を示さないケースもあり不安定であった。Case Bにおいて、CNXTとSmoothGradを使用することにより、可視化において安定した良好な結果を得ることができた。図11はADの画像を95%の確度でADと判定した事例である。HMは判別における重み付の高さに基づいた特徴的領域として、ADにおいて顕著に認められる海馬、側頭葉上葉の萎縮、及び側脳室の拡大が認められる場所(図1)を示している。図12は、Control例の画像を98%の確度でControlと判定した事例である。HMは特定の特徴を有する場所は示さず分散し図12と異なっている。図13は誤判定をした事例である。ADの臨床画像を、Controlの確度を54%、ADの確度を46%として、Controlと判定し誤判定となっている。HMは判定確度が拮抗している状態においては、図11のAD判定と図12のControl判定の中間の所見を呈していた。この結果は、判断指標の連続性を示していると考えられ、臨床と本技術の活用の中で価値を生み出す潜在性を示唆していると考えられた。

今後の課題として、本技術の比較優位性は検証出来たものの、実践において臨床的価値を付加するために本技術の更なる改良が必要である。例えば、Grad-CAM++とSmooth-Gradの画像の示すHMの情報が潜在価値を有することは検証できた一方、日常臨床における価値を創出する応用技術の形成には今後の研究が必要である。更に実証プロジェクトとして、研究、応用開発、事業開発に向けて活動することが重要である。次に、データについて、単施設で撮像された画像であり、今後画像の質に対する堅牢性を担保する必要性とデータ量が小さい問題があり、応用開発においては創出価値に大きく影響するので、データを拡大した継続的な研究が必要であると考えられた。

結 論

最適化した深層学習を脳形態 MRI に用いることにより臨床での使用を検討しうる性能の高い判別システムが可能であることを立証し、さらに、実装した判別システムの判断根拠を可視化してその情報を供給し、臨床に診断補完システムという付加価値を与えることが可能であることを示すことができた。

文 献

- 1) 日本神経学会: 認知症疾患診療ガイドライン 2017. 第 1 版, 医学書院, 東京 : p210-213, 2017. <https://www.neurology-jp.org/guidelinem/2022> 年 11 月 1 日閲覧
- 2) 厚生労働省老健局 : 認知症の将来推計について, 認知症施策の総合的な推進について(参考資料), pp3, 令和元年 6 月 20 日. (厚生労働省.ホームページ) <https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/> 2022 年 11 月 1 日閲覧
- 3) 厚生労働省 : 図表 1-1-15 人口高齢化を乗り越える社会モデルを考える、血管性及び詳細不明の認知症、アルツハイマー病の患者数の推移. 平成 28 年版厚生労働白書. pp33, 2016. <https://www.mhlw.go.jp/wp/hakusyo/kousei/16/backdata/01-01-01-15.html> 2022 年 11 月 1 日閲覧
- 4) 武田基秀他 : 医療における認知症の診断状況 (アルツハイマー型認知症診断の方向性).新潟大学統合脳機能研究センター・三菱商事共同研究報告書資料 エクゼクティブサマリー,pp3, 2006 年 4 月 10 日.(私信)
- 5) 伊藤健吾、木村泰之、加藤隆司 : 画像診断、認知症の予防とケア 第 3 章 各論 1 認知症の診断 2.画像診断, 公益財団法人長寿科学振興財団ホームページ: pp81-83, 2019. <https://www.tyojyu.or.jp/kankoubutsu/gyoseki/ninchisho-yobocare/h30-3-2.html> 2023 年 1 月 9 日閲覧
- 6) Bloudek L M, Spackman D E, Blankenburg M and Sullivan S D: Review and meta-analysis of biomarkers and diagnostic imaging in Alzheimer's disease. J Alzheimers Dis 26: 627-45, 2011.
- 7) Nordberg A, Rinne J O, Kadir A and Långström B : The use of PET in Alzheimer disease, NATURE REVIEWS NEUROLOGY , 2016
- 8) Scho M, Lockhart S N, Schonhaut D R, Schwimmer H D, Rabinovici G D and Jagust W J: PET Imaging of Tau Deposition in the Aging Human Brain, Neuron 89: 971-982, 2016.
- 9) Blennow K: A Review of Fluid Biomarkers for Alzheimer's Disease, Moving from

CSF to Blood , Neurol Ther6 (Suppl 1):S15–S24, 2017.

- 10) Sivera R , Delingette H, Lorenzi M, Pennec X, Ayache N, for the Alzheimer's Disease Neuroimaging Initiative: A model of brain morphological changes related to aging and Alzheimer's disease from cross-sectional assessments .Neuroimage 198: 255-270.2019
- 11) Vemuri P and Jack C R Jr: Role of structural MRI in Alzheimer's disease. Alzheimer's Research & Therapy: alzres.biomedcentral.com:1-3,2010.
<https://alzres.biomedcentral.com/articles/10.1186/alzrt47> 2023年1月9日閲覧
- 12) Apostolova L G, Green A E, Babakchanian S, Hwang K S, Chou Y-Y, Toga A W and Thompson P M: Hippocampal atrophy and ventricular enlargement in normal aging, mild cognitive impairment and Alzheimer's disease. Alzheimer Dis Assoc Disord ,26(1):17-19, 2012,
- 13) Myszczyńska M A, Ojiamies P N, Lacoste A M B, Saffari D N A, Mead R, Hautbergue G M, Holbrook J D and Ferraiuolo L: Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. Nature Reviews Neurology 16(8):440-456, 2020.
- 14) Ito Y, Fukuda M, Matsuzawa H, Masuda H, Kobayashi Y, Hasegawa N, Kitaura H, Kakita A and Fujii Y: Deep learning-based diagnosis of temporal lobe epilepsy associated with hippocampal sclerosis: An MRI study. Epilepsy Research 178: 106815, 2021.
- 15) 服部 高明、安田 永智、松澤 等: マルチモーダルな脳画像と脳機能データを用いたマルチモーダル機械学習、平成 31 年度新潟大学脳研究所「脳神経病理資源活用の疾患病態共同研究拠点」共同利用・共同研究報告書:pp1-2.2019.(私信) https://www.bri.niigata-u.ac.jp/docs/4-1.takaaki_hattori.pdf
2023年1月9日閲覧
- 16) 斎藤康毅著:ゼロから作る Deep Learning,初版,株式会社オラリー・ジャパン,東京, pp39-52 pp165-175, 2016.
- 17) Simonyan K. and Zisserman A : Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556, 2014.
<https://doi.org/10.48550/arXiv.1409.1556> 2023年1月9日閲覧
- 18) Liu Z, Mao H, Wu C-Y, Feichtenhofer C , Darrell T and Xie S: A ConvNet for the 2020s , IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR): 11966-11976, 2022. <https://arxiv.org/abs/2201.03545> 2023年1月9日閲覧
- 19) Chattopadhyay A , Sarkar A , Howlader P and Balasubramanian V N : Grad-CAM : Generalized Gradient-Based Visual Explanations for Deep Convolutional Networks, 2018 IEEE Winter Conference on Applications of Computer Vision (WACV): 839–847, 2018.

- 20) Smilkov D, Thorat N, Kim B , Viegas F and Wattenberg M: SmoothGrad: removing noise by adding noise. Google Inc.,2017
https://arxiv.org/pdf/1706.03825.pdf?source=post_page 2023 年 1 月 9 日閱覽
- 21) Jack C R Jr, Bennett D A, Blennow K, Carrillo M C, Dunn B, Haeberlein S B, Holtzman D M, Jagust W, Jessen F, Karlawish J, Liu E, Molinuevo J L, Montine T, Phelps C, Rankin K P, Rowe C C, Scheltens P, Siemers E, Snyder H M, Sperling R; Contributors. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement.* 2018 Apr;14(4):535-562. doi: 10.1016/j.jalz.2018.02.018. PMID: 29653606; PMCID: PMC5958625.
- 22) Dosovitskiy A et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale、 Proceeding of The International Conference on Learning Representations (ICLR) 2021, Vienna, Austria May 04 2021.
<https://openreview.net/forum?id=YicbFdNTTy> 2023 年 1 月 9 日閱覽
- 23) Odusami M, Maskeliūnas R, Damaswvicius R and Krilavius T: detection of Early Stage from Functional Brain Changes in Magnetic Resonance Images Using a Finetuned ResNet18. *Diagnostics* 2021, 11(6):1071.2021.
- 24) Mehmood A, Yang S, Feng Z, Wang M, Ahmad AL S, Khan R, Maqsood M and Yaqub M: A Transfer Learning Approach for Early Diagnosis of Alzheimer's Disease on MRI Images. *Neuroscience* 460: 43-52. 2021.
<https://doi.org/10.1016/j.neuroscience.2021.01.002> 2023 年 1 月 9 日閱覽
- 25) Li Q and Yang M Q: Comparison of machine learning approaches for enhancing Alzheimer's disease classification. *PeerJ* 9:e10549. 2021
<https://peerj.com/articles/10549/> 2023 年 1 月 9 日閱覽
- 26) Bhatele K R and Bhadauria S S: Classification of Neurodegenerative Diseases Based on VGG 19 Deep Transfer Learning Architecture: A Deep Learning Approach.
<http://dx.doi.org/10.7717/peerj.10549>, 2021. 2023 年 1 月 9 日閱覽

謝 辞

新潟大学脳研究所統合脳機能研究センター 中田力教授、五十嵐博中教授、松澤等博士、伊藤陽祐博士、大野健博士をはじめとした同センターの皆様に、長年にわたり、ご指導ご支援をいただいたことに深謝いたします。また解析データの提供を賜りました北日本脳神経外科病院（現五泉中央病院）佐藤光弥先生に深謝いたします。