HOKKAIDO UNIVERSITY

Title	FDG-PETを用いた脳ブドウ糖代謝率測定における人工知能を活用した非侵襲的測定方法に関する研究
Author(s)	川内, 敬介
Degree Grantor	北海道大学
Degree Name	博士(医理工学)
Dissertation Number	甲第15974号
Issue Date	2024-03-25
DOI	https://doi.org/10.14943/doctoral.k15974
Doc URL	https://hdl.handle.net/2115/92247
Type	doctoral thesis
File Information	Kawauchi_Keisuke.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

学　位　論　文

FDG-PET を用いた脳ブドウ糖代謝率測定における人工知能を活用した非侵襲的測定方法に関する研究

(Studies on non-invasive measurement method using artificial intelligence in measuring cerebral metabolic rate of glucose using FDG-PET)

2024 年 3 月

北　海　道　大　学

川内　敬介

Keisuke Kawauchi

学　位　論　文

FDG-PET を用いた脳ブドウ糖代謝率測定における人工知能を活用した非侵襲的測定方法に関する研究

(Studies on non-invasive measurement method using artificial intelligence in measuring cerebral metabolic rate of glucose using FDG-PET)

2024 年 3 月

北　海　道　大　学

川内　敬介

Keisuke Kawauchi

目次

1. 発表論文目録および学会発表目録 1
2. 緒言 2
2.1. FDG-PET 検査 2
2.2. 脳ブドウ糖代謝率 4
2.3. 人工知能 6
2.4. 説明可能な AI 8
2.5. 本研究の目的 9
3. 略語表 10
4. 方法 11
4.1. 対象 11
4.2. 使用した装置および設定 11
4.3. AI の構成 13
4.4. CNN への入力画像と入力方法 18
4.5. CNN の学習および検証 20
4.6. CNN の評価方法 21
5. 結果 22
5.1. CNN の学習結果 22
5.2. 実測の動脈血漿中の$\boldsymbol{Cpt}$と CNN によって得られた$\boldsymbol{Cpt}$との比較 23
5.3. 脳ブドウ糖代謝率の比較 25
5.4. 人工知能の判断根拠の可視化 27
6. 考察 29
7. 総括および結論 30
8. 謝辞 31
9. 参考文献 32

1. 発表論文目録および学会発表目録

本研究の一部は以下の論文に発表した。

1. Keisuke Kawauchi, Mui Saito, Kentaro Nishigami, Chietsugu Katoh

   A convolutional neural network-based system to estimate the arterial plasma radioactivity curve in ^{18}F-FDG dynamic brain PET study

   Nuclear Medicine Communications; 44(11): p.p.1029-1037, November 2023. DOI: 10.1097/MNM.0000000000001752

本研究の一部は以下の学会で発表した。

1 Keisuke Kawauchi, Chietsugu Katoh

   ^{18}F-FDG dynamic brain PET study estimating the arterial plasma radioactivity curve using a convolutional neural network (CNN)

   34th Annual Congress of the European Association of Nuclear Medicine – EANM'21, October 20 - 23, 2021

2. 緒言

2.1. FDG-PET 検査

^{18}F-fluorodeoxyglucose（^{18}F-FDG）はグルコースつまりブドウ糖の C-2 の位置の水酸基を陽電子放出核種である ^{18}F で置換したブドウ糖類似放射性薬剤である(Pacák et al., 1969)（図 1）。

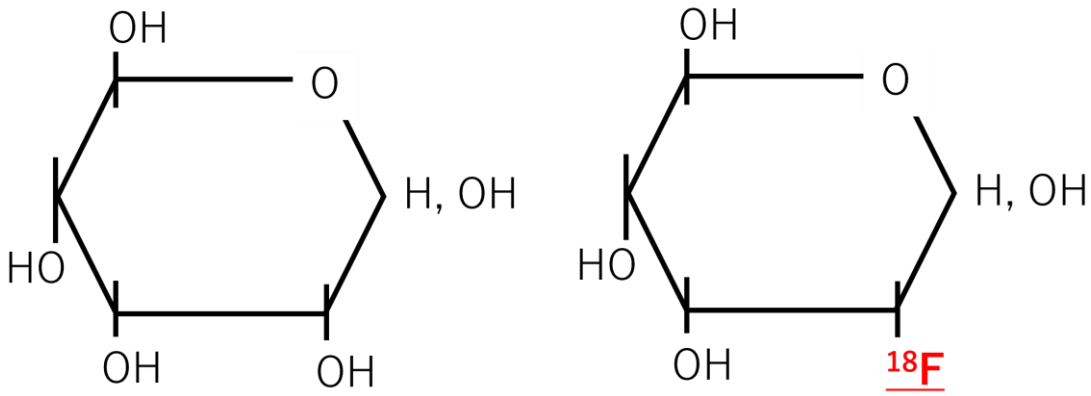

図 1. ブドウ糖（左）と FDG（右）の化学構造

投与された FDG は血液に運ばれて全身をめぐり、グルコーストランスポーターを介して細胞内に取り込まれ、ヘキソキナーゼによるリン酸化を受ける。ここまでの代謝は通常のブドウ糖と同様であるが、その後の代謝反応が進まず、リン酸化体として細胞内に滞留するという特徴を持っている。ブドウ糖類似物質である FDG が取り込まれるということは、組織にグルコースが取り込まれていることと同義であり、組織における糖代謝を表現することが出来る。このような特徴から FDG は糖代謝が盛んな細胞、例えば腫瘍細胞などに多く集積するため、Positron Emission Tomography (PET) 装置を用いることで病変部の可視化を行うことが可能となる(Som et al., 1980)。また、全身の腫瘍を検索できる利点があり、現在では主に悪性腫瘍の検出(Strauss & Conti, 1991)、良悪性の判定(Okada et al., 1991)、生物学的特性診断(Okada et al., 1992)、治療効果判定および予後予測(Chaise et al., 2007)などに使用されており、特に早期胃がんを除く他の検査や画像診断により病期診断、転移・再発の診断が確定できない場合に関しては国民健康保険の適用対象となっている。

FDG-PET を用いる利点としては、FDG がブドウ糖類似物質であるために拒絶反応がほとんどなく、半減期も約 110 分と比較的短いため、被検者への負担が少ないことがある。また、PET 装置は他の画像検査装置と異なりモーターで駆動するカメラではないことに加えて、感度も高いため、フレーム時間の短いダイナミック撮像（10 秒/1 フレーム以下でも撮像可能）による定量検査を実施できることなどの利点もある。

これらの利点を生かして悪性腫瘍以外では、脳機能の定量検査にも使用されている(Newberg et al., 2002)。例えばアルツハイマー病巣やてんかん病巣では神経細胞の活動が低下しており、糖代謝も低下しているため、通常よりも集積が低下している部位を評価することによって病巣を同定することが可能となる(Anchisi et al., 2005; Drzezga et al., 1999)。ダイナミック ^{18}F-FDG-PET による糖代謝画像の例を図 2 に示した。この例はスライス厚 3.4mm の頭部矢状断像が 42 スライス並べられた画像である。画像の色は、青が濃いほど糖代謝が低く、赤に近いほど糖代謝が高いことを表している。図中の赤矢印の部分に着目すると、周囲は黄緑に近い色であるのに対して、矢印の部分だけ青くなっており、左右の頭頂葉で糖代謝が低下していることが画像から把握できる。脳機能の検査に関しては、難治性部分てんかんで外科切除が必要な場合について国民健康保険の適用対象となっている。

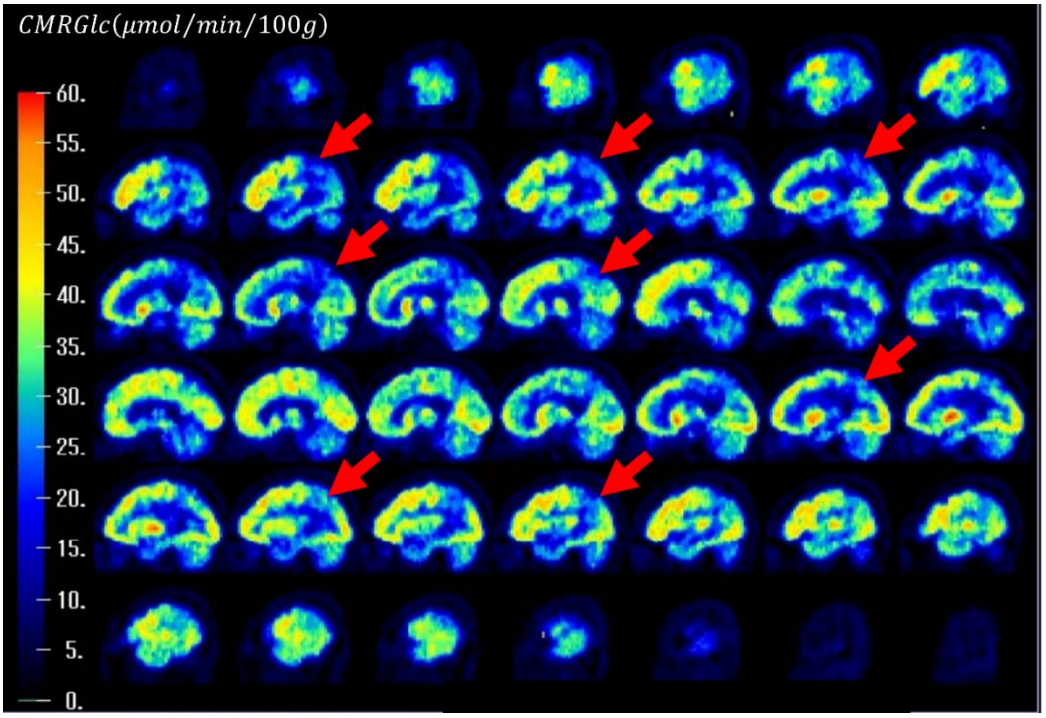

図 2. ダイナミック ^{18}F-FDG-PET による糖代謝画像の例

2.2. 脳ブドウ糖代謝率

図 2 で示した糖代謝画像の例は、PET 画像から脳ブドウ糖代謝率（Cerebral Metabolic Rate of Glucose、CMRGlc）を計算して可視化した画像である。ただし、^{18}F から放出された陽電子の消滅放射線を検出器でとらえただけの画像だけでは、単純に ^{18}F-FDG の体内分布とその位置を定量的に可視化しただけの画像に過ぎず、脳ブドウ糖代謝率を表す画像にはならない。PET を用いて脳ブドウ糖代謝率を測定するためには、^{18}F-FDG がどのように組織に取り込まれ、代謝されるのかという生理学的な動態を考慮する必要がある。そこで、^{18}F-FDG の生理学的な動態を考慮可能な「コンパートメントモデル」と呼ばれるモデルが考案された(Phelps et al., 1979; Sokoloff et al., 1977)。^{18}F-FDG の 3 コンパートメントモデルの模式図を図 3 に示した。

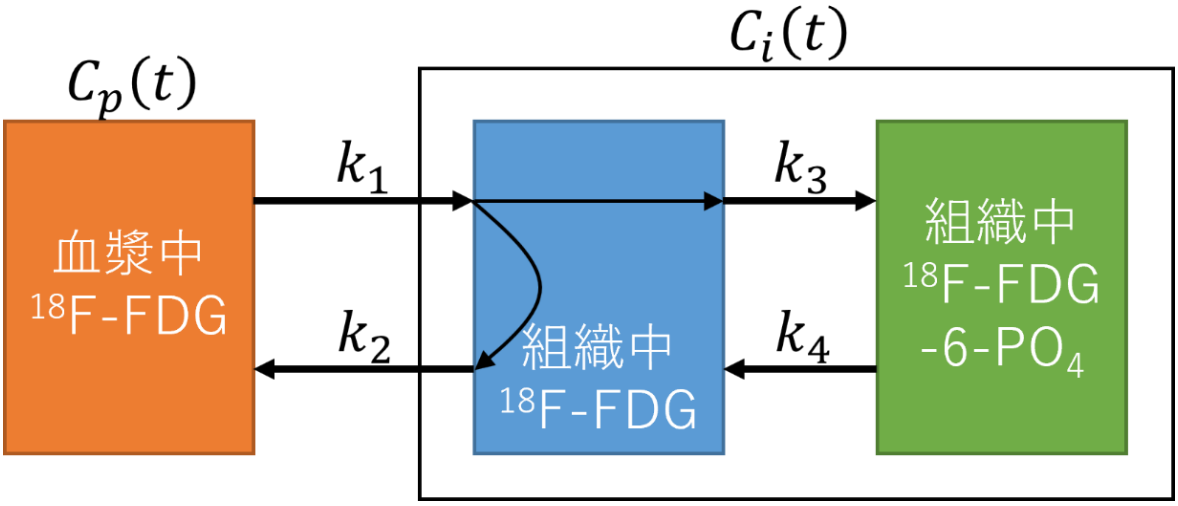

図 3. 3 コンパートメントモデル

図中のk_1~k_4は各コンパートメント間の ^{18}F-FDG の輸送および酵素反応の速度定数を表しており、$C_i(t)$は組織中の ^{18}F-FDG の放射能量、$C_p(t)$は動脈血漿中の ^{18}F-FDG の放射能量をそれぞれ表している。なお、$C_i(t)$および$C_p(t)$のtは、放射能を測定した時間を表す。この関係は次式で表すことが出来る。

$$C_i(t) = \frac{k_1}{\alpha_2 - \alpha_1}[(k_1 + k_4 - \alpha_1)e^{-\alpha_1 t} + (-k_3 - k_4 - \alpha_2)e^{-\alpha_2 t}] * C_p(t) \cdots (1)$$

$$\alpha_1 = \frac{k_2 + k_3 + k_4 \pm \sqrt{(k_2 + k_3 + k_4)^2 - 4k_2k_4}}{2} \cdots (2)$$

式(1)および(2)が成り立つとき、実際に得られた$C_i(t)$および$C_p(t)$を代入し、非線形最小 2 乗法によってk_1~k_4を決定する(Gallagher et al., 1978)。この時、CMRGlc は次式によって求められる。

$$\mathrm{CMRGlc} = \frac{GL}{LC} \cdot \frac{k_1 k_3}{k_2 + k_3} \cdots (3)$$

ここで、GL は血糖値で、LC は一括係数と呼ばれる ^{18}F-FDG とグルコースの差を補正する係数である。

^{18}F-FDG-PET を用いたコンパートメントモデルに基づく脳ブドウ糖代謝率測定の最大の利点は、ブドウ糖類似物質である ^{18}F-FDG を用いることで実際の糖代謝の様子を可視化できることにある。他の測定方法、例えば Magnetic Resonance Imaging (MRI)を用いた方法などでは糖代謝そのものを可視化することは困難であり、間接的に表現する事しかできない。このことからも、^{18}F-FDG-PET を用いた脳ブドウ糖代謝率測定が唯一無二の測定方法であるといえる。

ただし、この測定方法にも問題がある。コンパートメントモデル解析における式(1)の$C_p(t)$は前述のとおり、動脈血漿中の ^{18}F-FDG の放射能量である。つまり、検査中は動脈血の連続採血が必要となる。動脈採血では、動脈の周囲を取り巻く神経の影響で痛みを伴い、静脈採採血に比べて侵襲性が高い。さらに動脈は静脈と比較して深部に存在するために技術的難易度が高いことに加え、^{18}F-FDG-PET 検査においては意識がある状態で実施することから痛み等によって被検者が動いてしまう可能性があり、この動脈採血を連続して実施することはリスクが高い。以上のことから、連続動脈採血を実施せずに脳ブドウ糖代謝率測定を実施する方法が求められている。

2.3. 人工知能

ところで、近年では人工知能（Artificial Intelligence、 AI）を用いた画像の分析に関する研究が進められている。以下では、AIの作成方法である機械学習手法およびＡIの医用画像解析分野への応用について説明する。

2.3.1. 機械学習手法

機械学習とは、データをコンピュータに与え、データの持つ法則性や重要度等の特徴を分析・学習するアルゴリズムのことである(Thomas M. Mitchell, 1997)。現在の機械学習手法は大きく2つに分けられる。一つは「古典的機械学習」で、こちらは後述の深層学習に当てはまらない機械学習手法を指す。代表的な手法としては、決定木やサポートベクターマシン、ニューラルネットワーク等、様々な手法がある。特に人間の脳内の神経細胞の構造と機能を模した数理モデルであるニューラルネットワークは後述の深層学習の基となった手法である。もう一つは「深層学習」で、古典的なニューラルネットワークを３層以上に多層化した機械学習手法の一種である(Wang et al., 2018)。近年のコンピュータの性能向上等に伴い実用化に向けた研究が進められてきた。特に 2006 年に初めて実用的な深層学習(Hinton et al., 2006; Hinton & Salakhutdinov, 2006)が提唱されたことをきっかけに「第3次AIブーム」が始まり、現在に至るまで発展を続けている。さらに、画像の物体認識において初めて深層学習の概念を取り入れた AlexNet(Krizhevsky et al., 2017)は、「ImageNet Large Scale Visual Recognition Challenge」という画像認識の精度を競う競技会において、人手で特徴量を設計した他の手法を圧倒し、深層学習が画像の分析に有用であることを示した。これ以降、画像の分析においては深層学習を用いた AI が主流となっている。

2.3.2. ＡＩの医用画像解析分野への応用

深層学習を用いたAIによる画像分析に関する研究は医用画像解析の分野への応用も進んでいる。例えば、胸部 X 線写真から結核の病変部位を自動的に検出した報告(Lakhani & Sundaram, 2017)や、頭部 MRI 検査画像から脳腫瘍を抽出した報告(Li et al., 2017)、肝造影 Computed Tomography (CT) 検査画像における肝腫瘤の鑑別を実施した報告(Yasaka et al., 2018)、全身 FDG-PET 検査画像から腫瘍の有無を分類した報告(Kawauchi et al., 2020)などがある。

上記のような画像の分類や病変部位の特定などを目的とした報告の他に、画像から数値を得るAIに関する研究も報告されている。例えばKawauchiらは、全身FDG-PET検査画像から被検者の年齢と体重を回帰的に予測するAIについて報告しており、年齢については83.2%の確率で誤差5歳以内、体重については96.1%の確率で誤差5kg以内の予測が可能であったとしている(Kawauchi et al., 2019)。

以上を受けて、動脈採血が必要な式(1)の$C_p(t)$を、AI を用いて画像から推定することが出来れば侵襲性の高い動脈採血を実施せずに ^{18}F-FDG-PET を用いたコンパートメントモデルに基づく脳ブドウ糖代謝率測定が実施できる可能性が考えられた。ただし深層学習では内部で実行されている計算がブラックボックス化されているため、深層学習によって実現した AI では推定が精度よく実施できたとしても、推定の際の判断根拠を提示することが難しいという問題がある。

2.4. 説明可能な AI

深層学習により開発された AI は判断根拠の提示が難しいという問題を解決するために、Class Activation Mapping (CAM)と呼ばれる根拠の可視化手法が提案されている(Selvaraju et al., 2017)。CAM は深層学習を用いた画像分類や物体検出、セマンティックセグメンテーション等において AI が判断を実施した際に根拠とした部分を可視化することが出来る手法である。CAM の模式図を図 3 に示した。図は頭部 MRI 画像の模式図で、赤円は CAM が示した AI の判断根拠を表している。(A)は正しく病変部をとらえているが、(B)は頭蓋骨の一部を判断根拠にしてしまっている。これは AI の学習に用いた病変のある画像で(B)と同じように頭蓋骨が含まれているような場合に、病変そのものではなく他の構造物を根拠として学習してしまった例である。(B)のパターンが得られてしまった場合には、AI の推定結果を信頼することが難しいということを視覚的に判断することが出来るようになる。

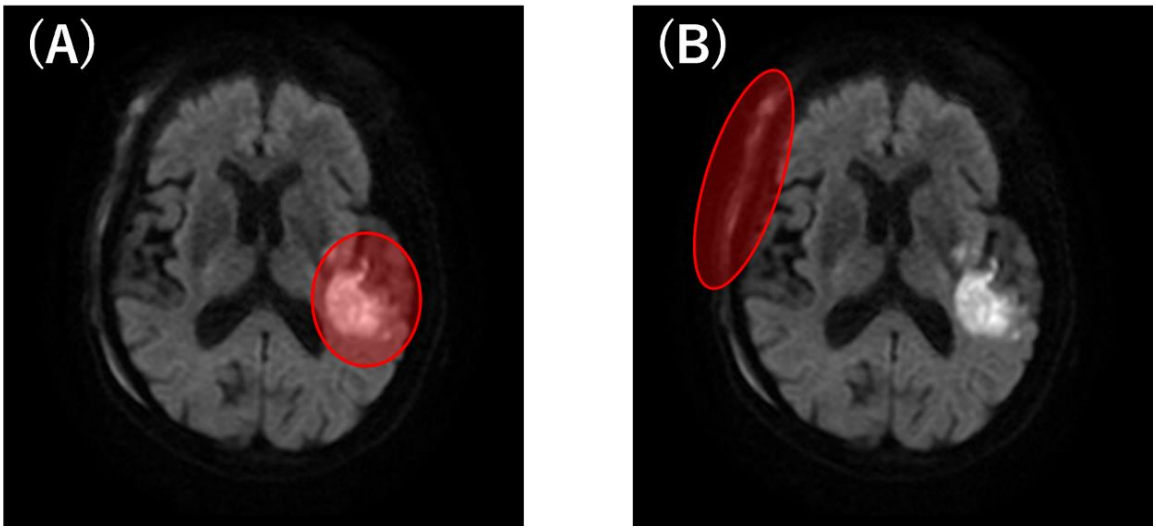

図 3. CAM の模式図

このように、CAM を用いても数学的根拠を示すことが困難であることには変わりないが、式(1)の$C_p(t)$を AI で推定した際の基準を確認することで、信頼可能な AI か判断することが出来るようになると考えられる。

2.5. 本研究の目的

^{18}F-FDG-PET を用いたコンパートメントモデルに基づく脳ブドウ糖代謝率測定は、脳のブドウ糖代謝を可視化する唯一無二の方法であるが、動脈採血が必要であり、侵襲性が高い。そこで本研究では、「侵襲性の高さ」を改善し患者の負担を低減するために、動脈採血が必要な式(1)の$C_p(t)$を、AI を用いて画像から推定する方法の開発を目的とした。さらに、$C_p(t)$推定の際の根拠となる部位を可視化することで人間が解釈可能な根拠に基づいて推定していることを提示し、確認できるようにすることを目的とした。

本研究の結果、AI を用いることで動脈採血を実施しない場合でも動脈採血を実施した場合と同等の式(1)の$C_p(t)$を得ることが可能なことを示した。また AI の判断根拠については、画像内の内頸動脈に基づいて推定していることを示した。内頸動脈は画像内で最も明確に撮像される動脈であり、これに着目して動脈血漿中の ^{18}F-FDG の放射能量である$C_p(t)$を推定していることは、人間の感覚にも反していない。これらの結果は AI を活用することで ^{18}F-FDG-PET による脳ブドウ糖代謝率測定の「侵襲性の高さ」を改善し、患者の負担を低減できる可能性を提示するものである。

3. 略語表

本文中および図中で用いた略語は以下のとおりである。

AI	Artificial Intelligence
CAM	Class Activation Mapping
CMRGlc	Cerebral Metabolic Rate of Glucose
CNN	Convolutional Neural Network
CT	Computed Tomography
FDG	FluoroDeoxyGlucose
FWHM	Full Width at Half Maximum
LOA	Limits Of Agreement
LOOCV	Leave-One-Out Cross-Validation
MRI	Magnetic Resonance Imaging
PET	Positron Emission Tomography
ROI	Region Of Interest

4. 方法

4.1. 対象

北海道大学病院にてダイナミック ^{18}F-FDG-PET による糖代謝検査を実施した神経疾患患者 29 名（37.3±13.2 歳、男性 11 名、女性 18 名）を対象とした。各患者に 185MBq の ^{18}F-FDG を静注し、ダイナミック ^{18}F-FDG-PET データの取得を実施した。ダイナミック ^{18}F-FDG-PET データの取得と同時に、前腕の動脈に挿入した留置針から 1ml の動脈血を合計 22 回採取し、直ちに遠心分離して動脈血漿を抽出した。

本研究は北海道大学病院の倫理審査委員会により承認されている（UMIN000018160）。また、本研究は後ろ向き研究であるため各患者からの書面によるインフォームドコンセントは省略した。

4.2. 使用した装置および設定

ダイナミック PET 検査にはシーメンス社製の PET 装置「EXACT HR +」を使用した。検査の際には ^{18}F-FDG 185MBq を肘静脈から静脈内に投与し、その直後から脳のダイナミック PET 画像を 60 分間収集した。ダイナミック撮像のフレーム数は22フレームで、収集のタイミングは 20 秒、50 秒、70 秒、100 秒、140 秒、180 秒、220 秒、270 秒、330 秒、390 秒、450 秒、570 秒、750 秒、930 秒、1110 秒、1350 秒、1650 秒、1950 秒、2250 秒、2550 秒、2850 秒、3150 秒とした。このデータについて、フィルタ補正逆投影法を用いて再構成した。生成した画像はすべて軸位画像で、画素数が128×128ピクセル、半値幅（full width at half maximum, FWHM）が 4.8mm、1 画素の実長が$2.5 \times 2.5 \times 6.3\ [mm]$であった。図 4 にダイナミック ^{18}F-FDG-PET による糖代謝検査の22フレーム分の画像のうち、内頸動脈が最も明瞭に撮影されるスライスの例を示した。画像の左上から右下に向かって時系列順に並んでおり、色が暗い（黒に近い）ほど FDG が少なく、色が明るい（白に近い）ほど FDG が多く集積している。時間が経過するにつれて FDG の集積が多くなり明るい部分が増加していることがわかる。

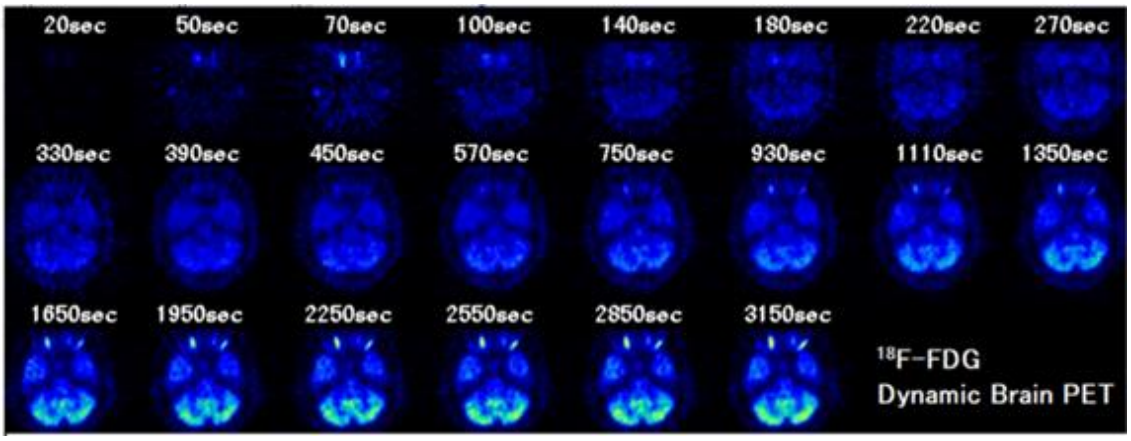

図 4. ダイナミック ^{18}F-FDG-PET による糖代謝検査の２２フレーム分の画像の例

動脈血中の ^{18}F-FDG の放射能量の測定は Aloka 社製のウェルカウンター「ARC-380CL」で測定し、動脈血中の ^{18}F-FDG の放射能量を得た。ただし、^{18}F-FDG は赤血球にも取り込まれるため、動脈血漿中の ^{18}F-FDG の放射能量$C_p(t)$を必要とするコンパートメントモデル解析のための正確な入力関数とはならない。そこで、脳組織に入らない赤血球の放射能を除去し、正確な動脈血漿中の時間放射能曲線を得た。図 5 に得られた動脈血漿中の ^{18}F-FDG の放射能量$C_p(t)$の例を示した。

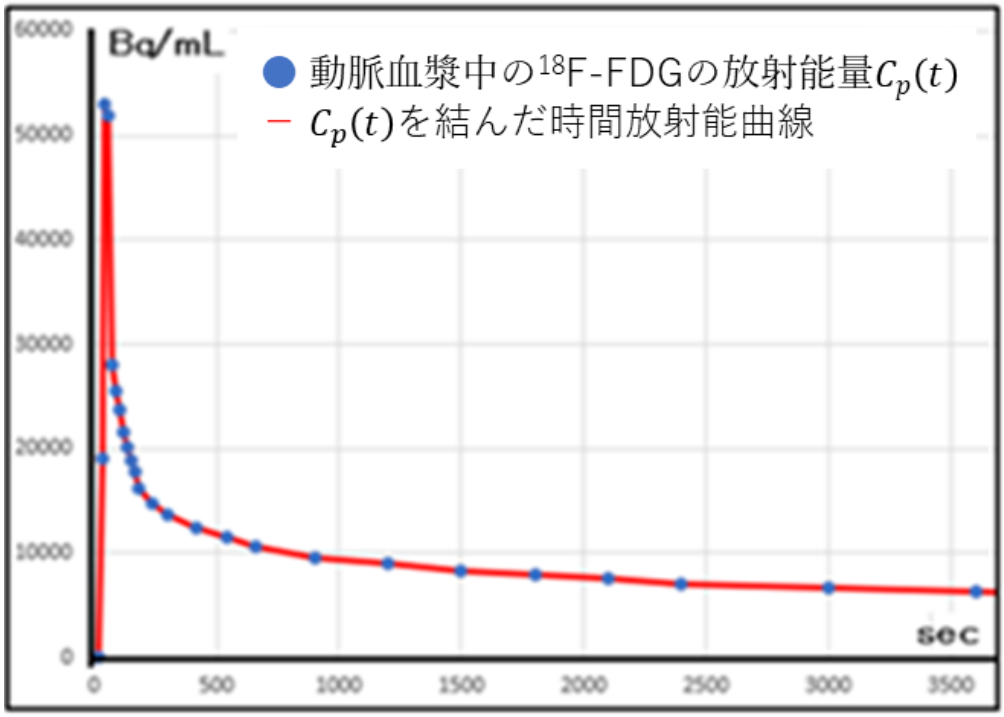

図 5. 動脈血漿中の ^{18}F-FDG の放射能量$C_p(t)$の例

4.3. AI の構成

AI の構成にはさまざまな種類があるが、本研究では深層学習の一種である畳み込みニューラルネットワーク（convolutional neural network, CNN）を用いた。以下では、深層学習および CNN と本研究で採用したCNNの構造について説明する。

4.3.1. 深層学習

深層学習は、人間の脳内の神経細胞の構造と機能を模したニューラルネットワークを用いた機械学習手法の一種であり、多層化したニューラルネットワークを用いることを特徴とする。図 6 に基本的なニューラルネットワークの模式図を示した。ニューラルネットワークは、入力層、隠れ層、出力層の 3 層で構成されている。この各層が生体でいうところのシナプス結合にあたるエッジで結ばれ、相互接続された生体でいうところのニューロン細胞にあたるノードで構成される。このようなニューラルネットワークにおいて、隠れ層に 2 層以上の複数の層を採用したニューラルネットワークを用いた機械学習手法を深層学習と呼ぶ。

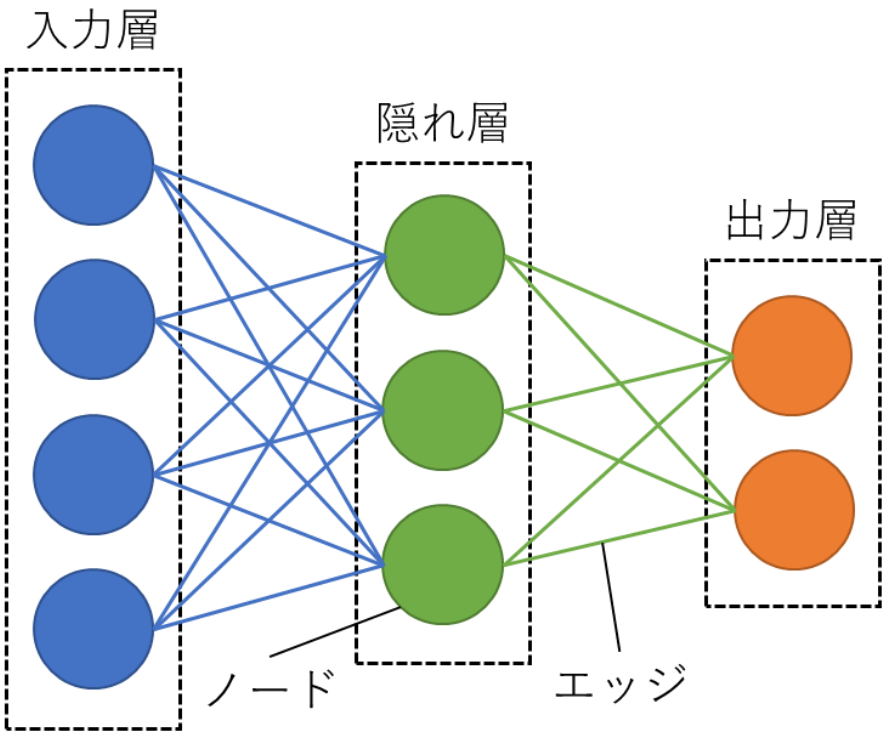

図 6. ニューラルネットワークの模式図

4.3.2. ＣＮＮ

畳み込みニューラルネットワーク（CNN）は深層学習の特殊なタイプであり、ニューラルネットワークの層構造の中に「畳み込み層」と呼ばれる層を持ち、層間で重みを共通化できることを特徴とする。畳み込み層は画像内の局所的な特徴量を抽出することが目的の層で、CNN は特に画像認識タスクに対して有効であることが報告されている(Krizhevsky et al., 2017)。PET 画像に対しても CNN を用いた報告が多くなされおり、例えば、治療前の FDG-PET に基づく CNN を用いた肺の悪性腫瘍の進行度分析と予後予測(Huang et al., 2022)や PET 画像の生成に CNN を用いた報告(Matsubara et al., 2022)、PET 画像のノイズ除去に CNN を用いた報告(Jaudet et al., 2021)などがある。そこで本研究では、動脈採血が必要な式(1)の$C_p(t)$を、AI を用いて画像から推定する方法として、CNN を用いることとした。

4.3.3. ネットワーク構造

CNN には様々な構成があるが、本研究では Xception(Chollet, 2016)と呼ばれるネットワークモデルに注目した。図 7 に Xception のネットワーク構造を示した。図中の Conv は畳み込み層、ReLU はランプ関数、stride は畳み込み時のフィルタの移動幅で特徴量を抽出する範囲にあたる値、SeparableConv は深さ方向に分割可能な畳み込み層、MaxPoolong はプーリング層（畳み込み層で抽出された特徴量のダウンサンプリングを実施し、特徴量を局所的な代表値に集約する役割の層）のうち代表値を最大値とするもの、GlobalAveragePooling はプーリング層のうち代表値を平均値とするものをそれぞれ表している。処理の流れはまず、入力層に画像を入力し、続いて隠れ層では図中の構造を8回繰り返した後、出力層で最終的な処理を実施する。Xception は、深さ方向に分離可能な畳み込み（SeparableConv）を利用することでパラメータ数を削減し、計算量の削減と汎用性向上を実現したモデルであり、これによって高性能なコンピュータを用いることなく実行可能な点で優れている。また、Xception は合計 71 層のネットワークモデルであるが、隠れ層は同じ構造を8回繰り返すという構成になっているため、プログラミング実装も比較的容易である。

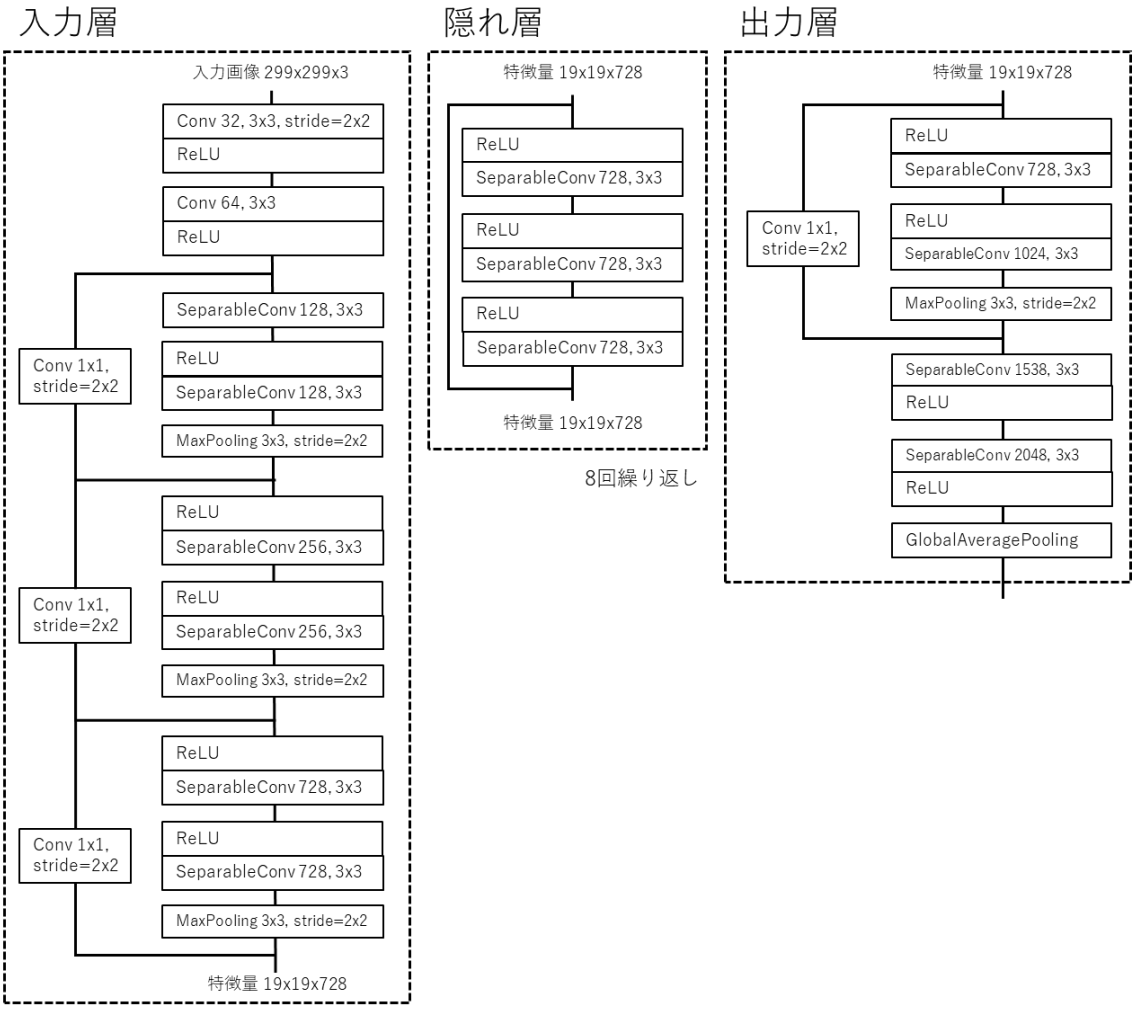

図 7. Xception のネットワーク構造

また、Xception を用いた医用画像解析に関する報告には、胃潰瘍の良性・悪性分類を行った報告(Liu et al., 2022)や急性期脳梗塞患者に対して hyperdense middle cerebral artery sign（主に塞栓性の超急性期梗塞で出現する治療方針に影響を与える重要な所見。頭部の単純 CT 検査において、中大脳動脈内の血栓および血栓の影響による末梢血管の高吸収構造が確認される。）の識別を行った報告(Shinohara et al., 2020)、胸部 X 線画像から COVID-19 の所見の識別を行った報告(Gülmez, 2022)などがあり、Xception の報告から約7年経過した現在でも新たな研究成果が報告されている。以上より本研究では、Xception と同等の構成のネットワークモデルを採用した。ただし、画像から数値を回帰予測することが目的であるため、出力層の最終出力のノード数を 1 とした。

4.3.4. CNN の学習

CNN を学習する際に頻出する問題に未学習と過学習がある。未学習とは、CNN が目的の達成に必要な情報を学習データから適切に学習することが出来ていない状態である。学習の開始時は当然未学習の状態であるが、学習ステップが進んでも未学習のまま学習が適切に進まない場合がある。これは目的に対して入力データに不適切なデータが含まれている場合やネットワーク構造が簡素すぎる場合などに生じる恐れがある。一方で過学習とは、CNNが入力されたデータのみを完璧に学習してしまう状態を言う。過学習の場合では、入力データ以外のデータに対する予測が適切に行えず、汎用性の低い状態になってしまう。これは、入力データが少なすぎる場合やネットワーク構造が複雑すぎる場合などに生じる恐れがある。

これらの現象は、学習ステップ（エポック数）に対して不正解の割合がプロットされた曲線（損失曲線）によって判断することが出来る。図 8 に損失曲線の模式図を示す。まず、適切に学習が進んだ場合は(A)のように学習データに対する損失曲線と検証データに対する損失曲線が同じように減少を続けていく。それに対して過学習の場合は(B)のように学習ステップが進むにつれて学習データの損失は 0 に近づくが、検証データの損失は大きくなっていく。また未学習の場合は(C)のように学習ステップが進んでも学習データ、検証データ共に不正解の割合が減少していかないという傾向を示す。ただし、未学習の場合も過学習の場合も明確な閾値は存在しない。そのため、実際の予測値の確認や前述のCAMを用いた判断根拠の可視化との組み合わせ評価が重要である。

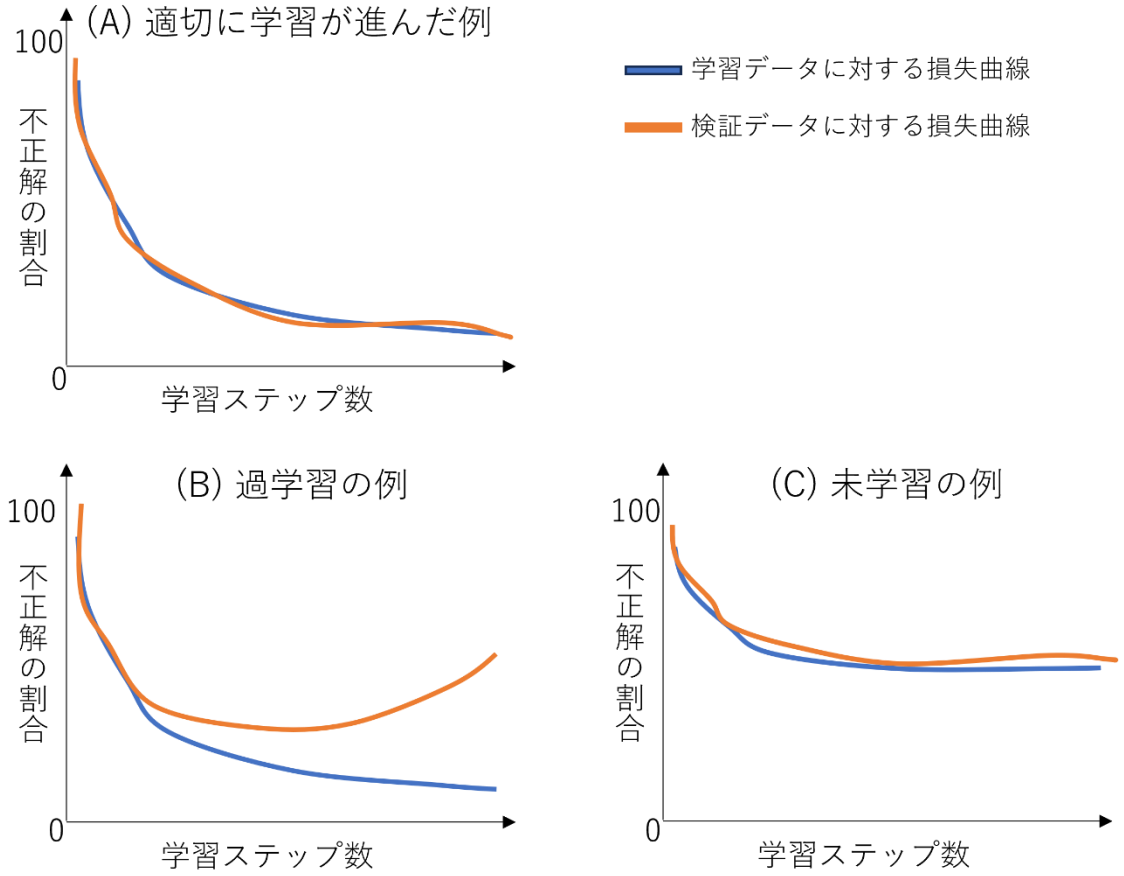

図 8. 損失曲線の模式図

4.3.5. CNN を実行したコンピュータの構成

本研究は下記の構成で実行した。

- OS：Windows 10 pro 64bit
- CPU：intel Core i9-10900K
- GPU：NVIDIA GeForce RTX 3090 24GB
- 深層学習フレームワーク：TensorFlow 2.5.0.
- プログラミング言語：Python 3.8.5

4.4. CNN への入力画像と入力方法

図 9 に CNN への入力画像の例を示した。画像はある症例におけるダイナミック ^{18}F-FDG-PET による糖代謝検査の第 10 フレーム（左）および第 20 フレーム（右）の画像である。この形式の画像が 1 症例あたり22フレーム×29 症例あるため、入力に利用可能な画像は最大で 638 枚ある。

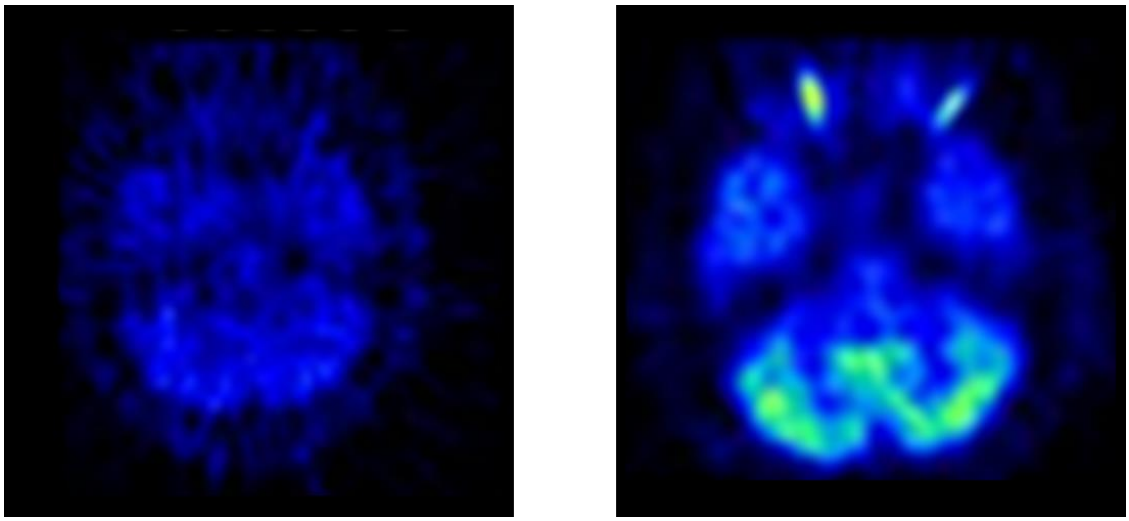

図 9. CNN への入力画像の例

この入力画像を学習フェーズと検証フェーズに分けて使用する。学習フェーズは、コンピュータに画像などのデータを入力し、入力画像の持つ法則性や重要度等の特徴を分析・学習させるフェーズであり、いわゆる AI を作成するフェーズである。
一方で検証フェーズは、作成した AI に学習フェーズで使用しなかった未知データを入力することによって AI の性能を検証フェーズである。このように検証フェーズでは未知のデータを入力する必要があるため、画像を学習フェーズ用と検証フェーズ用に分割した。本研究では画像および付随する動脈血漿中の ^{18}F-FDG の放射能量 $C_p(t)$を入力する際の分割手法に leave-one-out cross-validation（LOOCV）法を採用した(Molinaro et al., 2005)。模式図を図 10 に示した。図はデータが 4 つの場合の例で、A,B,C,D のように検証用の標本を入れ替えながら学習と評価を実施する。

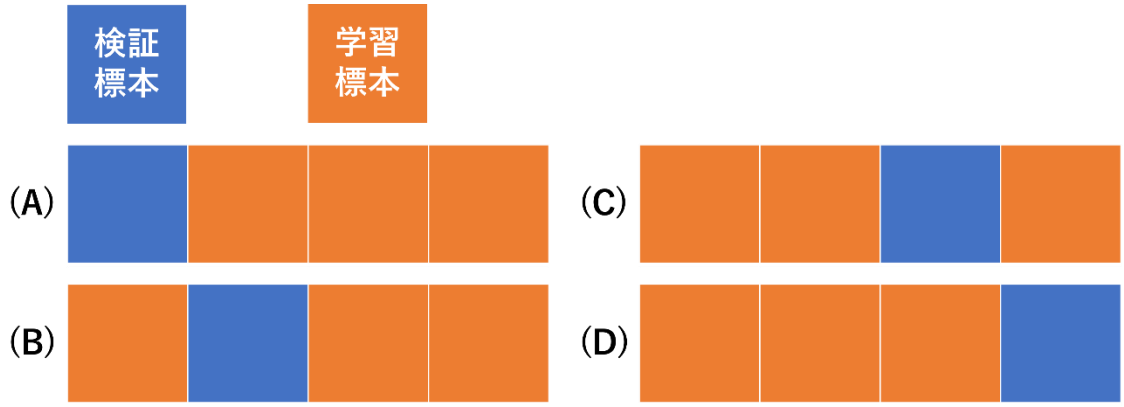

図 10. LOOCV 法の模式図

LOOCV 法とは「学習対象の標本群から検証に使用する 1 標本を抽出し、残りの標本群を用いて学習、検証用の 1 標本で検証を実施する」という流れを全標本で検証を実施するまで繰り返すという学習・検証方法である。LOOCV 法は症例数や画像枚数が多い場合には時間がかかりすぎるため不適と考えられており、1 標本ではなく複数標本を抽出する k-fold cross-validation 法を採用することが一般的である。しかし、本研究では 29 症例、638 枚と画像枚数が少ないため LOOCV 法を用いることが好適であると考えて採用した。

4.5. CNN の学習および検証

本研究の目的は、^{18}F-FDG-PET を用いたコンパートメントモデルに基づく脳ブドウ糖代謝率測定の問題点である「侵襲性の高さ」を改善し患者の負担を低減するために、動脈採血が必要な動脈血漿中の ^{18}F-FDG の放射能量$C_p(t)$を、CNN を用いて画像から回帰的に推定することであり、CNN の学習および検証は次の流れで実施した。

① LOOCV 法を用いて、対象 29 症例から検証フェーズ用の 1 症例と学習フェーズ用の 28 症例にデータを分割
② 学習フェーズ用の 28 症例、各 22 フレーム、合計 616 枚の画像および当該画像における動脈血漿中の ^{18}F-FDG の放射能量$C_p(t)$を AI に入力して、学習
③ 検証フェーズ用の 1 症例、22 フレーム、22 枚の画像を学習した AI に入力して、動脈血漿中の ^{18}F-FDG の放射能量$C_p(t)$を推定
④ ③で推定した動脈血漿中の ^{18}F-FDG の放射能量$C_p(t)$と実際に動脈血から取得した動脈血漿中の ^{18}F-FDG の放射能量$C_p(t)$を比較して、検証
⑤ ①～④を検証フェーズ用の 1 症例を入れ替えながら 29 回繰り返し試行

上記の①～⑤の流れで CNN を学習・評価を実施するが、そのままでは AI の判断の根拠は不明なままである。そこで本研究では、CNN が反応した部分を強調表示する CAM を用いて、4.5 の流れで学習した AI の動脈血漿中の ^{18}F-FDG の放射能量$C_p(t)$の推定の際の判断根拠の可視化を実施した。

4.6. CNN の評価方法

4.6.1. ブランドアルトマンプロットを用いた CMRGlc の評価

下記の手順で動脈採血から得られた実測値を用いた場合の CMRGlc と CNN による推定値を用いた場合の CMRGlc を算出し、評価する。

① 各患者に対して、図のように脳内に 100 個の関心領域（ROI, Region of Interest）を配置し、脳組織中の ^{18}F-FDG の放射能量を求める

② 3 コンパートメントモデルにより、動脈採血から得られた実測値を用いた場合の CMRGlc と CNN による推定値を用いた場合の CMRGlc を算出する

③ DIFF（動脈採血から得られた実測値を用いた CMRGlc と CNN による推定値を用いた CMRGlc の差分値）および MMEAN（動脈採血から得られた実測値を用いた CMRGlc と CNN による推定値を用いた CMRGlc の平均値）を算出し、ブランドアルトマンプロット(Martin Bland & Altman, 1986)を作成する

④ 平均 DIFF±1.96×DIFF の標準偏差の値である一致限界（LOA、Limits of Agreement）を計算し、DIFF の何％が LOA 内に含まれているか計算する

⑤ DIFF の 95％以上が LOA に含まれていれば、動脈採血から得られた実測値を用いた CMRGlc と CNN による推定値を用いた CMRGlc については同等であると解釈できる

4.6.2. CAM を用いた注目部位の可視化評価

CAM を用いて CNN が動脈血漿中の 18F-FDG の放射能量$C_p(t)$を推定した際の画像内の着目部分を可視化する。CNN に入力する画像において動脈の放射能量に伴って変化する要素としては内頸動脈がある。したがって学習された CNN は、画像内の内頸動脈周辺に着目することが望ましい結果であると考える。そこで可視化評価としては、CAM で注目部位の可視化を実施し、その注目部位に内頸動脈が含まれているか否かを目視により確認する。

5. 結果

5.1. CNN の学習結果

北海道大学病院にて 185MBq の ^{18}F-FDG を静注し、ダイナミック ^{18}F-FDG-PET データの取得および動脈血採取を実施した神経疾患患者 29 名（37.3±13.2 歳、男性 11 名、女性 18 名）を対象に、後ろ向き研究を実施した。CNN の学習は LOOCV の1セットあたり平均で 10 分程度を要し、学習後の推定および CAM による注目部位の可視化は 1 患者当たり 0.1 秒程度で完了した。

ここで、学習の際に未学習や過学習の問題が発生しなかったかを確認する。図 11 は、本研究で行われた 29 の試行のうちの 1 つの学習曲線を示している。この症例では、学習データに対する学習曲線と検証データに対する学習曲線が、学習ステップが進むにつれて徐々に小さく（予測精度が高く）なっており、2つの曲線が大きく乖離することもないということから、CNN の学習過程において、未学習でも、過学習もなく、適切な学習が実施されたことが確認できる。残りの 28 の症例でも同様の傾向が観察されたことから、未学習や過学習はなかった。

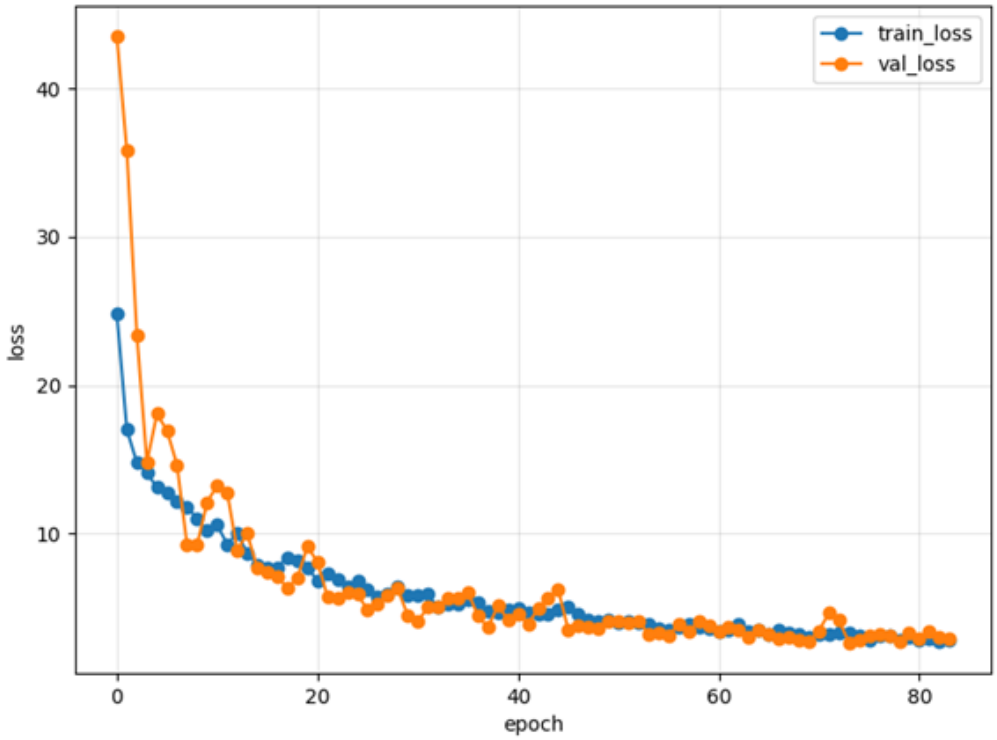

図 11. 損失曲線

5.2. 実測の動脈血漿中の$\boldsymbol{C_p(t)}$と CNN によって得られた$\boldsymbol{C_p(t)}$との比較

図 12 は、動脈採血を実施した実測の動脈血漿中の$\boldsymbol{C_p(t)}$と CNN が推定した$\boldsymbol{C_p(t)}$を示している。図の縦軸は$\boldsymbol{C_p(t)}$の放射能量 Bq/mL、横軸は経過時間を表しており、各点におけるバーは動脈採血を実施した実測の動脈血漿中の$\boldsymbol{C_p(t)}$と CNN が推定した$\boldsymbol{C_p(t)}$とのばらつきを表したエラーバーである。図より、CNN を用いた提案手法を用いることで、ばらつきはあるものの動脈血のサンプリングなしでも、動脈血のサンプリングを実施した場合の実測の動脈血漿中の$\boldsymbol{C_p(t)}$と同等の値が得られる例があることが確認できる。

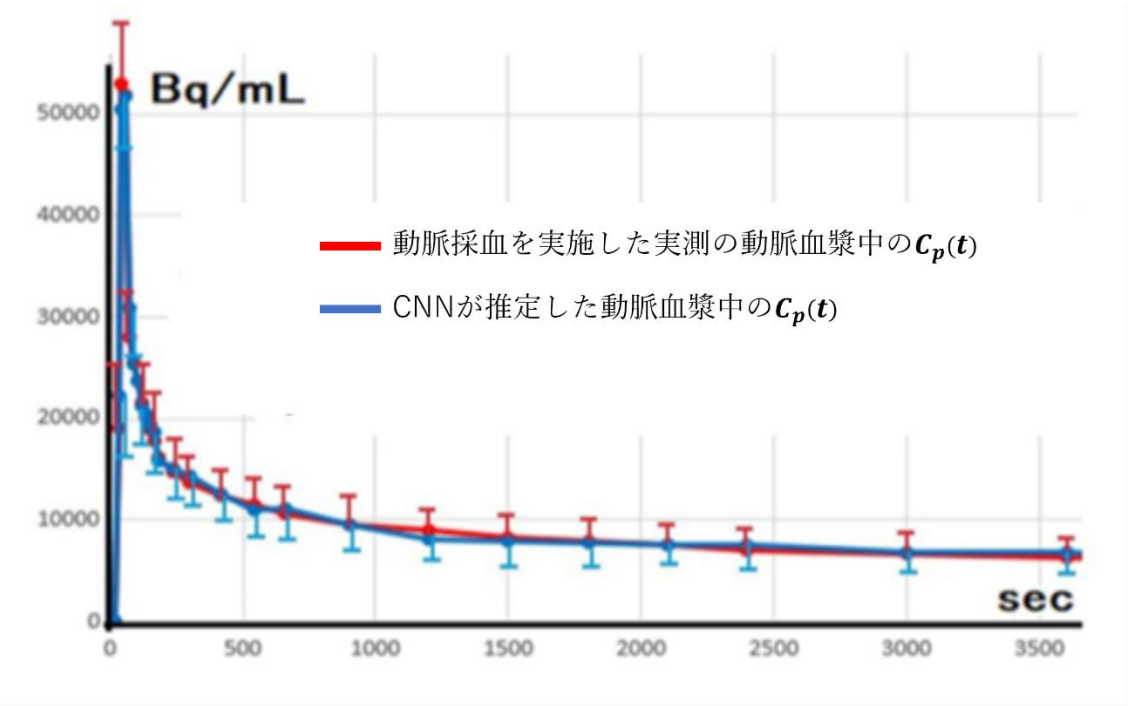

図 12. 動脈採血を実施した実測の動脈血漿中の$\boldsymbol{C_p(t)}$と CNN が推定した$\boldsymbol{C_p(t)}$

さらに、実測の動脈血漿中の$\boldsymbol{C_p(t)}$と CNN が推定した$\boldsymbol{C_p(t)}$とを比較した結果を表 1 に示した。29症例のうち 21 の症例では、22 フレーム中 18 フレーム以上が 10%以内の誤差を示した。一方、残りの 8 症例では予測誤差が 10%を超えるフレームが多数確認された。

表 1. 動脈採血を実施した実測の動脈血漿中の$C_p(t)$と CNN が推定した$C_p(t)$のスライス毎の平均誤差と誤差が 10%以内のスライスの割合

	平均誤差 %	誤差が 10%以内のスライスの割合
patient_01	3.9	100%
patient_02	6.1	86%
patient_03	1.6	100%
patient_04	5.0	86%
patient_05	4.2	95%
patient_06	1.7	100%
patient_07	3.3	100%
patient_08	4.3	95%
patient_09	3.3	100%
patient_10	17.3	32%
patient_11	3.3	100%
patient_12	5.5	91%
patient_13	2.5	100%
patient_14	17.6	41%
patient_15	9.0	64%
patient_16	5.1	91%
patient_17	12.6	64%
patient_18	3.5	95%
patient_19	8.4	77%
patient_20	3.4	100%
patient_21	3.0	100%
patient_22	2.9	100%
patient_23	2.7	100%
patient_24	33.8	18%
patient_25	4.2	100%
patient_26	47.3	14%
patient_27	30.2	18%
patient_28	34.5	9%
patient_29	2.7	95%

5.3.　脳ブドウ糖代謝率の比較

図 13 に示したように、各症例において、100 個の ROI を脳内に配置した。これらの ROI は、縦軸と横軸に沿って 2cm ごとに 1 ピクセルの領域として自動的に定義されている。ただし、脳室内および脳外は除外している。

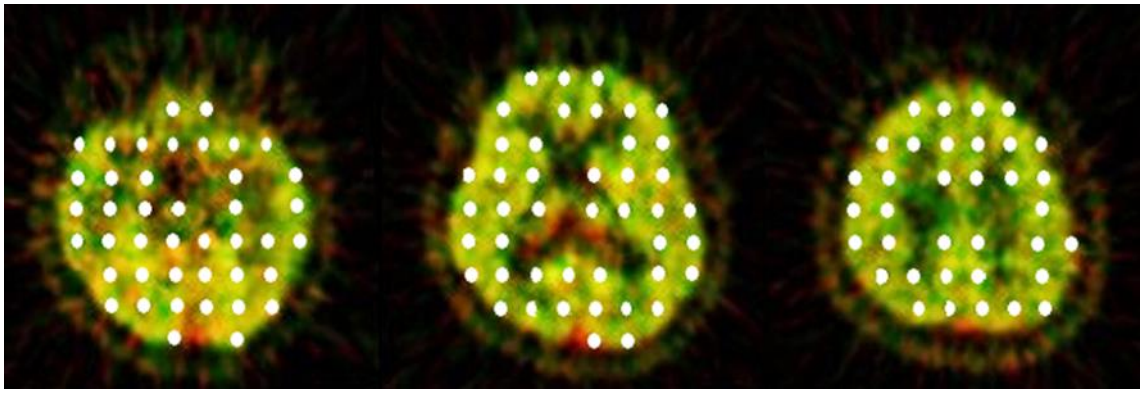

図 13. 画像内に設置した ROI の例

CMRGlc は 3 コンパートメントモデルを用いて計算された。結果を図 14 に示した。採血をして得られた実測の CMRGlc 値と CNN によって推定された CMRGlc 値との相関係数は 0.98 と有意に高い相関を示している。なお、ガイドライン(Cohen, 1988)によれば、0.5 以上の相関係数は強い相関を示すとされている。

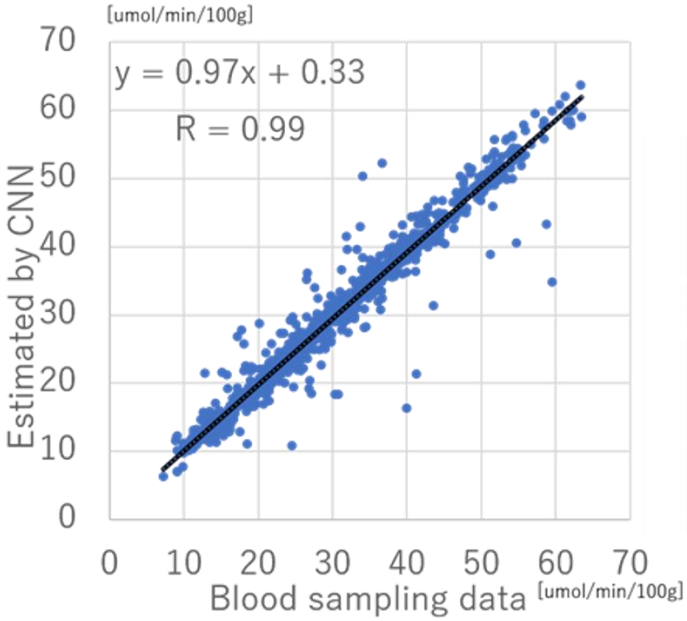

図 14. CMRGlc の散布図

図 15 はブランドアルトマンプロットを示しており、縦軸は採血から得られたCMRGlc 値と CNN によって推定された CMRGlc 値の差（DIFF）を表し、横軸はこれらの値の平均（MMEAN）を表している。破線は一致限界（LOA）を表し、"平均DIFF±1.96×DIFF の標準偏差 "として計算される。DIFF 値の 95%が LOA 内に収まっていれば、誤差が正規分布に従っていることを示し、採血から得られた CMRGlc 値とCNNによって推定されたCMRGlc値との間に一貫性があるとみなすことができる。本研究の結果は、DIFF 値の 97%が LOA 内に収まり一貫性を示した。

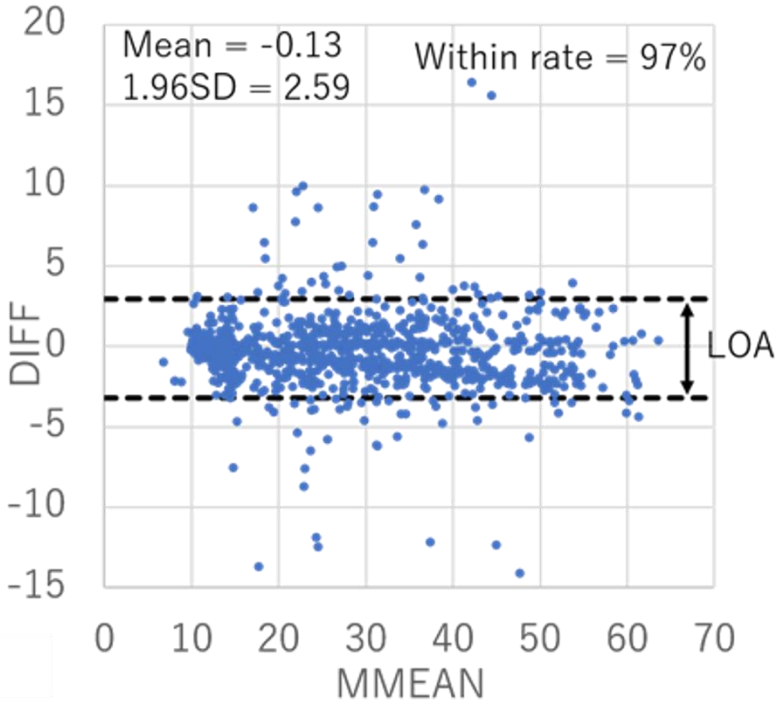

図 15. CMRGlc に関するブランドアルトマンプロット

5.4. 人工知能の判断根拠の可視化

CNN が情報を抽出する際に最も着目した画像内の部位（活性化領域）を調査するために、通常設定の Grad-CAM および注目部位をより強調するために調整を実施した独自の表示領域調整済み Grad-CAM による判断根拠の可視化を実施した。

5.4.1. 通常設定の Grad-CAM による可視化

通常設定の Grad-CAM の基本設定では注目度が低い領域を青色、高い領域を赤色としてその間を虹色に補間するグラデーションカラーを採用している。これによって重要度の変化が色としてわかりやすく表示できる。この分析の例を図 16 に示した。注目すべきは、すべての症例において少なくとも 1 つの画像で、内頸動脈近傍を含む活性化領域が明瞭に示されたことである。一方で図 17 は内頸動脈近傍以外の領域が強調された例である。具体的には、(A)は視神経周辺の強調領域、(B)は小脳に隣接する強調領域である。このように内頸動脈近傍以外を強調した画像では顕著な予測誤差を示す傾向を確認した。また、図 16、図 17 に共通するが、重要度の変化が激しい領域では色が重なりあってしまい注目領域の判断が難しくなる場合があることを確認した。そこで表示方法を工夫する取り組みを実施した。

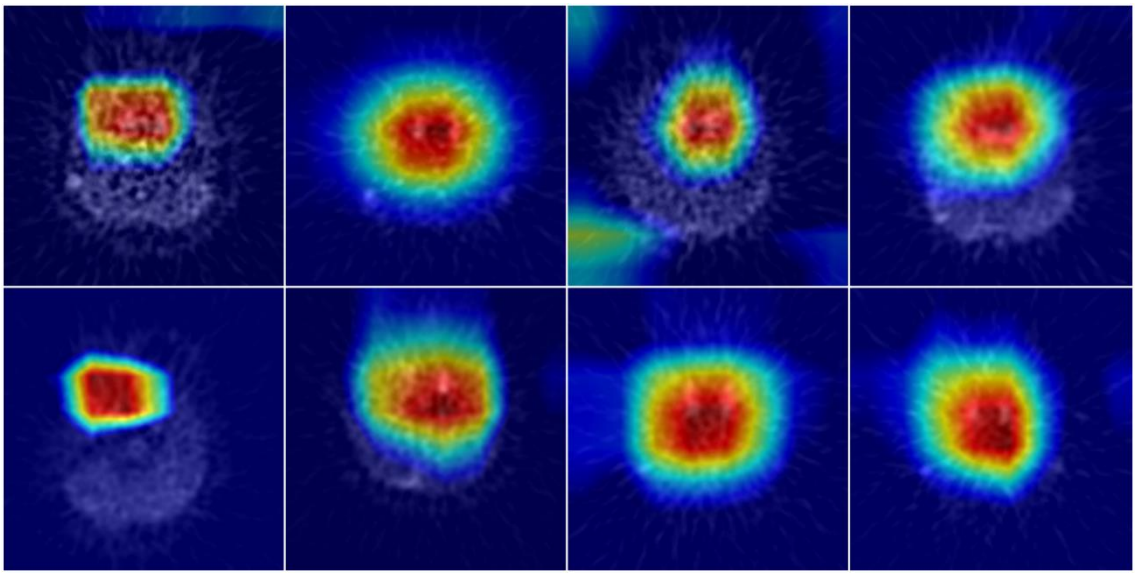

図 16. Grad-CAM の例

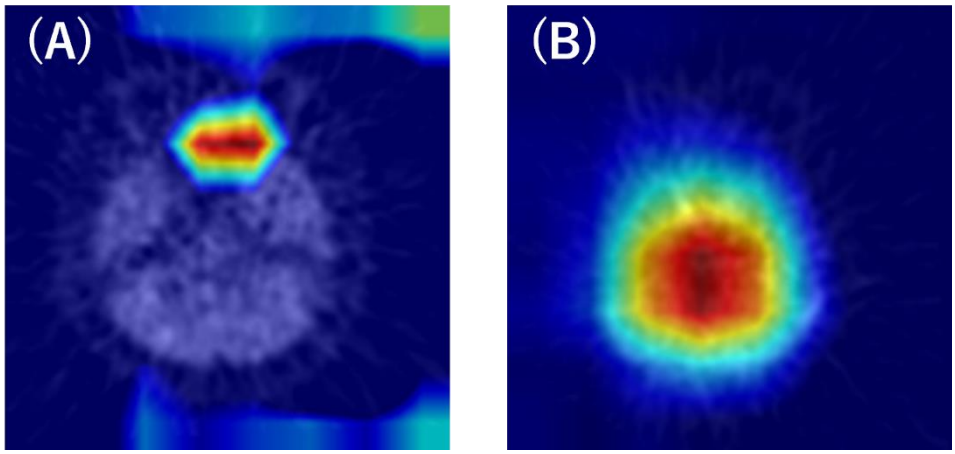

図 17. 内頸動脈近傍以外を強調した画像の例

5.4.2. 表示領域調整済み Grad-CAM による可視化

表示領域調整済み Grad-CAM では、通常の Grad-CAM のうち、注目度が 70%以上の領域を赤色、それ以下の部分を青色の 2 色に敢えて塗り分けることで、注目すべき領域をより強調して表示しようとする試みである。この分析の例を図 18 に示した。画像は図 16 と同様の患者の同様のスライスであり、結果は通常設定の Grad-CAM と同様であるが、重要度が 70%以上となった領域が内頸動脈近傍に広がっていることがより明確に判断できるようになったと考える。

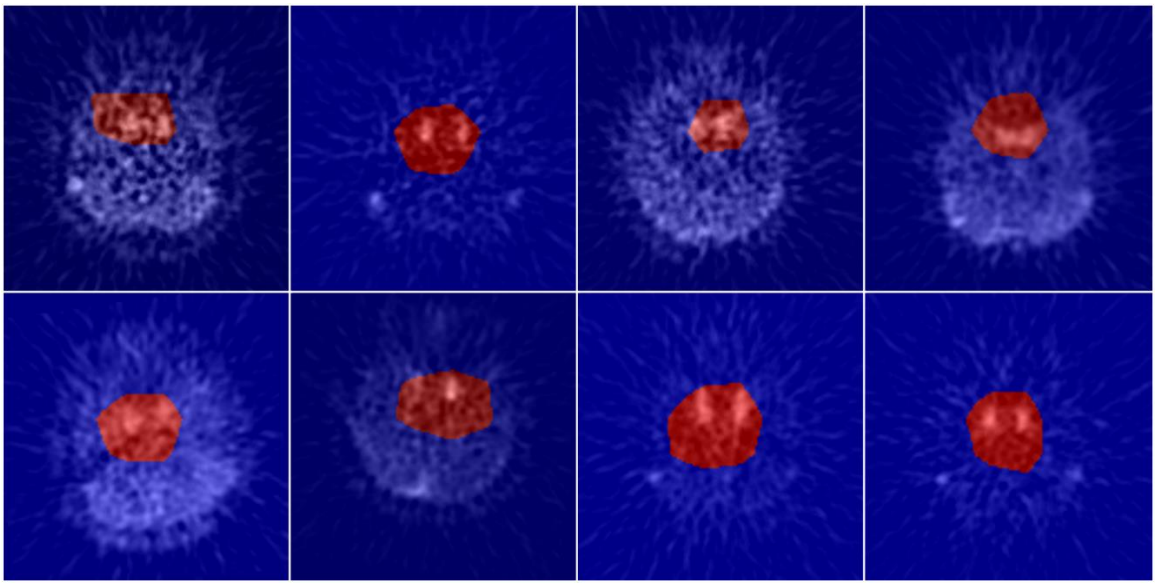

図 18. 表示調整済み Grad-CAM の例

6. 考察

本研究では、CNN を用いて動脈血漿中の ^{18}F-FDG の放射能量である$C_p(t)$の推定を実施した。その結果、21 の症例において、22 フレーム中 18 フレーム以上の予測誤差が 10%の範囲内に収まることが確認された。さらに、血液サンプルから得られた推定 CMRGlc と CNN から得られた推定 CMRGlc の相関係数は、0.98 という非常に有意な値を示した。これらの知見は、CNN が 18F-FDG 脳 PET データを用いて動脈血漿時間放射能曲線を正確に推定できる可能性を示唆しており、多様なダイナミックイメージングデータを非侵襲的に定量解析できる可能性につながると考えている。

画像データを用いて脳 PET の入力関数を推定する手法は従来から文献でいくつか報告されている(Bartlett et al., 2019; Zanotti-Fregonara et al., 2011)。しかしこれらの手法では、実際の検査のタイミングにおいても動脈血または静脈血のサンプリングが必要である。本提案手法でも CNN の学習時には動脈血のサンプリングは依然として必要である。本研究の結果は、十分に訓練された CNN を使用することで、実際の検査のタイミングでは採血の必要性がなくなる可能性を示唆している。

29症例のうち、8 つの症例において予測誤差が大きくなった明確な原因は特定できていない。これは CNN の判断根拠がブラックボックスになっていることが影響している。予測誤差の原因を特定するために本報告では CAM を用いて、CNN の判断根拠を可視化した。可視化の結果より、予測に失敗したケースでは内頸動脈の視認性が他のケースに比べて低下していることが観察された。さらに、図 17 のように、CAM が内頸動脈以外の領域に焦点を当てる傾向を示した。これらの結果から考えると、症例によって内頸動脈の見え方にばらつきがあり、内頸動脈が最も顕著に識別できるフレームだとしても、他の症例と比較すると内頸動脈が明瞭に描出されない症例が存在したことが、予測精度の低下につながった可能性があると推測している。この問題に対処するためには、見え方のばらつきに対応できるように訓練データに同様の傾向を持つ画像の数を増やすことが有効であると考える。

このように可視化技術である CAM を用いることで「CNN の注目領域」を可視化することができ、CNN のブラックボックス的なアプローチに対して説明可能な AI の開発を促し、ユーザーの信頼を高めることが期待される。

7. 総括および結論

本研究から得られた新知見は以下のとおりである。

- ^{18}F-FDG-PET による脳ブドウ糖代謝率測定において、CNN を用いることで動脈採血無しで血漿中の ^{18}F-FDG の放射能量を推定できる可能性を示した。
- CNN の判断根拠を可視化した結果、画像内の内頸動脈近傍に注目している例では予測精度が高い傾向を確認した。一方で、内頸動脈に着目できなかった例では予測精度が低い傾向を確認した。
- 予測精度が低い症例では、精度の高い症例と比較して内頸動脈近傍が明瞭に描出されていない傾向を確認した。

前述のとおり、脳 PET で使用する入力関数を画像データから推定する手法自体は報告されていたが、いずれの手法でも採血を避ける手法に関する報告は見られなかった。本研究の成果は、CNN を用いることによって ^{18}F-FDG-PET による脳ブドウ糖代謝率測定において、動脈採血無しに血漿中の ^{18}F-FDG の放射能量を推定できる可能性が示唆された。これによって検査における動脈採血という侵襲性の高い行為を実施する必要がなくなり、患者の負担軽減につながると考えられる。またこの結果は、多様なダイナミックイメージングデータの非侵襲的定量解析に応用できる可能性があり、今後の臨床現場において、種々の検査における侵襲性の低減へとつながっていくことが期待される。

ただし本研究においては、CNN の判断根拠がブラックボックスである影響を受けて一部の症例において、的確な予測が出来なかったことに関する原因究明が出来ていない点が limitation として挙げられる。現時点では可視化手法によって画像ごとに内頸動脈の明瞭さが異なることが影響していると考えている。原因究明のためには。学習データに内頸動脈の明瞭さが異なる種々のデータを追加し、学習データの量とバリエーションを増加させることで予測精度が改善するかなどに関する調査を行っていく必要があり、今後の研究課題であると考えている。

8. 謝辞

本論文を作成するにあたり、多大なご指導とご鞭撻を賜りました北海道大学大学院保健科学院・加藤知恵次教授に深く御礼申し上げます。また、研究にご協力いただいた研究室の皆様に厚く御礼申し上げます。最後に、本論文は多くの方々のご協力に支えられて完成したものであり、ここに全ての皆様へ心から感謝の意を表します。

9. 参考文献

Anchisi, D., Borroni, B., Franceschi, M., Kerrouche, N., Kalbe, E., Beuthien-Beumann, B., Cappa, S., Lenz, O., Ludecke, S., Marcone, A., Mielke, R., Ortelli, P., Padovani, A., Pelati, O., Pupi, A., Scarpini, E., Weisenbach, S., Herholz, K., Salmon, E., … Perani, D. (2005). Heterogeneity of brain glucose metabolism in mild cognitive impairment and clinical progression to Alzheimer disease. *Archives of Neurology*, *62*(11), 1728–1733. https://doi.org/10.1001/ARCHNEUR.62.11.1728

Bartlett, E. A., Ananth, M., Rossano, S., Zhang, M., Yang, J., Lin, S. fei, Nabulsi, N., Huang, Y., Zanderigo, F., Parsey, R. V., & DeLorenzo, C. (2019). Quantification of Positron Emission Tomography Data Using Simultaneous Estimation of the Input Function: Validation with Venous Blood and Replication of Clinical Studies. *Molecular Imaging and Biology*, *21*(5), 926–934. https://doi.org/10.1007/S11307-018-1300-1

Chaise, C., Itti, E., Petegnief, Y., Wirquin, E., Copie-Bergman, C., Farcet, J. P., Delfau-Larue, M. H., Meignan, M., Talbot, J. N., & Molinier-Frenkel, V. (2007). [F-18]-Fluoro-2-deoxy-D-glucose positron emission tomography as a tool for early detection of immunotherapy response in a murine B cell lymphoma model. *Cancer Immunology, Immunotherapy* , *56*(8), 1163. https://doi.org/10.1007/S00262-006-0265-0

Chollet, F. (2016). Xception: Deep Learning with Depthwise Separable Convolutions. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, *2017-January*, 1800–1807. https://doi.org/10.1109/CVPR.2017.195

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences Second Edition*. LAWRENCE ERLBAUM ASSOCIATES.

Drzezga, A., Arnold, S., Minoshima, S., Noachtar, S., Szecsi, J., Winkler, P., Römer, W., Tatsch, K., Weber, W., & Bartenstein, P. (1999). 18F-FDG PET Studies in Patients with Extratemporal

and Temporal Epilepsy: Evaluation of an Observer-Independent Analysis. *Journal of Nuclear Medicine*, *40*(5), 737–746. https://jnm.snmjournals.org/content/40/5/737
Gallagher, B. M., Fowler, J. S., Gutterson, N. I., MacGregor, R. R., Wan, C. N., & Wolf, A. P. (1978). Metabolic Trapping as a Principle of Radiopharmaceutical Design: Some Factors Responsible for the Biodistribution of [18F] 2-Deoxy-2-Fluoro-D-Glucose. *Journal of Nuclear Medicine*, *19*(10), 1154–1161. https://jnm.snmjournals.org/content/19/10/1154
Gülmez, B. (2022). A novel deep neural network model based Xception and genetic algorithm for detection of COVID-19 from X-ray images. *Annals of Operations Research*, 1. https://doi.org/10.1007/S10479-022-05151-Y
Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, *18*, 1527–1554.
Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, *313*(5786), 504–507. https://doi.org/10.1126/SCIENCE.1127647/SUPPL_FILE/HINTON.SOM.PDF
Huang, B., Sollee, J., Luo, Y. H., Reddy, A., Zhong, Z., Wu, J., Mammarappallil, J., Healey, T., Cheng, G., Azzoli, C., Korogodsky, D., Zhang, P., Feng, X., Li, J., Yang, L., Jiao, Z., & Bai, H. X. (2022). Prediction of lung malignancy progression and survival with machine learning based on pre-treatment FDG-PET/CT. *EBioMedicine*, *82*. https://doi.org/10.1016/J.EBIOM.2022.104127
Jaudet, C., Weyts, K., Lechervy, A., Batalla, A., Bardet, S., & Corroyer-Dulmont, A. (2021). The Impact of Artificial Intelligence CNN Based Denoising on FDG PET Radiomics. *Frontiers in Oncology*, *11*. https://doi.org/10.3389/FONC.2021.692973
Kawauchi, K., Furuya, S., Hirata, K., Katoh, C., Manabe, O., Kobayashi, K., Watanabe, S., & Shiga, T. (2020). A convolutional neural network-based system to classify patients

using FDG PET/CT examinations. *BMC Cancer*, *20*(1), 1–10. https://doi.org/10.1186/S12885-020-6694-X/FIGURES/4
Kawauchi, K., Hirata, K., Katoh, C., Ichikawa, S., Manabe, O., Kobayashi, K., Watanabe, S., Furuya, S., & Shiga, T. (2019). A convolutional neural network-based system to prevent patient misidentification in FDG-PET examinations. *Scientific Reports*, *9*(1). https://doi.org/10.1038/S41598-019-43656-Y
Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386
Lakhani, P., & Sundaram, B. (2017). Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. *Radiology*, *284*(2), 574–582. https://doi.org/10.1148/radiol.2017162326
Li, Z., Wang, Y., Yu, J., Guo, Y., & Cao, W. (2017). Deep Learning based Radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma. *Sci Rep*, *7*(1), 5467. https://doi.org/10.1038/s41598-017-05848-2
Liu, Y., Zhang, L., Hao, Z., Yang, Z., Wang, S., Zhou, X., & Chang, Q. (2022). An xception model based on residual attention mechanism for the classification of benign and malignant gastric ulcers. *Scientific Reports*, *12*(1). https://doi.org/10.1038/S41598-022-19639-X
Martin Bland, J., & Altman, D. G. (1986). STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. *The Lancet*, *327*(8476), 307–310. https://doi.org/10.1016/S0140-6736(86)90837-8
Matsubara, K., Ibaraki, M., Nemoto, M., Watabe, H., & Kimura, Y. (2022). A review on AI in PET imaging. *Annals of Nuclear Medicine*, *36*(2), 133–143. https://doi.org/10.1007/S12149-021-01710-8/METRICS
Molinaro, A. M., Simon, R., & Pfeiffer, R. M. (2005). Prediction error estimation: a comparison of resampling methods.

Bioinformatics, *21*(15), 3301–3307. https://doi.org/10.1093/BIOINFORMATICS/BTI499
Newberg, A., Alavi, A., & Reivich, M. (2002). Determination of regional cerebral function with FDG-PET imaging in neuropsychiatric disorders. *Seminars in Nuclear Medicine*, *32*(1), 13–34. https://doi.org/10.1053/SNUC.2002.29276
Okada, J., Yoshikawa, K., Imazeki, K., Minoshima, S., Uno, K., Itami, J., Kuyama, J., Maruno, H., & Arimizu, N. (1991). The Use of FDG-PET in the Detection and Management of Malignant Lymphoma: Correlation of Uptake with Prognosis. *Journal of Nuclear Medicine*, *32*(4).
Okada, J., Yoshikawa, K., Itami, M., Imaseki, K., Uno, K., Itami, J., Kuyama, J., Mikata, A., & Arimizu, N. (1992). Positron emission tomography using fluorine-18-fluorodeoxyglucose in malignant lymphoma: a comparison with proliferative activity. *Journal of Nuclear Medicine : Official Publication, Society of Nuclear Medicine*, *33*(3), 325–329. https://pubmed.ncbi.nlm.nih.gov/1740697/
Pacák, J., Toík, Z., & Erný, M. (1969). Synthesis of 2-deoxy-2-fluoro-D-glucose. *Journal of the Chemical Society D: Chemical Communications*, *2*, 77–77. https://doi.org/10.1039/C29690000077
Phelps, M. E., Huang, S. C., Hoffman, E. J., Selin, C., Sokoloff, L., & Kuhl, D. E. (1979). Tomographic measurement of local cerebral glucose metabolic rate in humans with (F-18)2-fluoro-2-deoxy-D-glucose: Validation of method. *Annals of Neurology*, *6*(5), 371–388. https://doi.org/10.1002/ANA.410060502
Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. *In ArXiv:1610.02391v3*.
Shinohara, Y., Takahashi, N., Lee, Y., Ohmura, T., & Kinoshita, T. (2020). Development of a deep learning model to identify hyperdense MCA sign in patients with acute ischemic stroke. *Japanese Journal of Radiology*, *38*(2), 112–117.

https://doi.org/10.1007/S11604-019-00894-4/METRICS
Sokoloff, L., Reivich, M., Kennedy, C., Rosiers, M. H. D., Patlak, C. S., Pettigrew, K. D., Sakurada, O., & Shinohara, M. (1977). THE [14C]DEOXYGLUCOSE METHOD FOR THE MEASUREMENT OF LOCAL CEREBRAL GLUCOSE UTILIZATION: THEORY, PROCEDURE, AND NORMAL VALUES IN THE CONSCIOUS AND ANESTHETIZED ALBINO RAT1. *Journal of Neurochemistry*, *28*(5), 897–916. https://doi.org/10.1111/J.1471-4159.1977.TB10649.X
Som, P., Atkins, H. L., Bandoypadhyay, D., Fowler, J. S., MacGregor, R. R., Matsui, K., Oster, Z. H., Sacker, D. F., Shiue, C. Y., Turner, H., Wan, C. N., Wolf, A. P., & Zabinski, S. V. (1980). A Fluorinated Glucose Analog, 2-fluoro-2-deoxy-D-glucose (F-18): Nontoxic Tracer for Rapid Tumor Detection. *Journal of Nuclear Medicine*, *21*(7), 670–675. https://jnm.snmjournals.org/content/21/7/670
Strauss, L. G., & Conti, P. S. (1991). The Applications of PET in Clinical Oncology. *Journal of Nuclear Medicine*, *32*(4).
Thomas M. Mitchell. (1997). *Machine Learning*. McGraw-Hill, Inc.
Wang, J., Wang, W. M., Tian, Z., & Li, Z. (2018). Classification of multiple affective attributes of customer reviews: Using classical machine learning and deep learning. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3207677.3277953
Yasaka, K., Akai, H., Abe, O., & Kiryu, S. (2018). Deep Learning with Convolutional Neural Network for Differentiation of Liver Masses at Dynamic Contrast-enhanced CT: A Preliminary Study. *Radiology*, *286*(3), 887–896. https://doi.org/10.1148/radiol.2017170706
Zanotti-Fregonara, P., Chen, K., Liow, J. S., Fujita, M., & Innis, R. B. (2011). Image-derived input function for brain PET studies: many challenges and few opportunities. *Journal of Cerebral Blood Flow & Metabolism*, *31*(10), 1986. https://doi.org/10.1038/JCBFM.2011.107