

人工知能を利用した脳血流 SPECT 統計解析画像の識別

山本 泰司¹⁾ 上部 星雄²⁾ 矢田 伸広²⁾
 宮原 善徳²⁾ 勝部 敬¹⁾ 吉廻 毅¹⁾
 北垣 一¹⁾

1) 島根大学医学部放射線医学講座

2) 島根大学医学部附属病院放射線部

Classification of the Statistical Analysis Images of Cerebral Blood Flow SPECT by Using the Deep Learning.

Yasushi YAMAMOTO, Hoshio UWABE, Nobuhiro YADA,
 Yoshinori MIYAHARA, Takashi KATSUBE, Takeshi YOSHIZAKO
 and Hajime KITAGAKI

1) Department of Radiology, Faculty of Medicine, Shimane University.

2) Department of Radiology, Shimane University Hospital.

(article received : May 22 2020)

Key words : Deep learning, 3D-SSP, SPECT, CNN, DLB,

1. 緒言 (はじめに)

近年, 放射線画像分野における人工知能 (Artificial intelligence : AI) の研究が急速に進み, 臨床応用に向け様々な論文報告や学会発表が盛んである。例えば, 頭部 CT での脳出血検出¹⁾ や MRA 画像からの未破裂動脈瘤検出²⁾, CT や MRI でのノイズ低減技術³⁾ や再構成技術⁴⁾ への応用など, 多岐にわたりその注目度も大きい。核医学分野に目を向けると, PET での認知症領域における病態識別など診断領域への利用に関する報告がされている^{5,6)}。今回, 我々は脳血流 SPECT における統計解析画像に注目し, Deep learning (DL) を用いた認知症領域での疾患分類を試みた。

2. 方 法

Deep learning とは人間が行うタスクをコンピュータに学習させる機械学習で, 人間のニューロンをモデルとしている。基本となる畳み込みニューラルネットワーク Convolution neural network (CNN) を多層化したディープニューラルネットワーク Deep neural network (DNN) を使

用することでデータの特徴量を段階的に, より深く学習することを可能とする手法である。ベースとなる CNN もいくつかの層で構成され, 畳み込み層 (Convolution layer) では, 特定領域のニューロンを束ねた後に画像フィルタ処理で特徴量抽出器としての役割を担う。プーリング層 (Pooling layer) では束ねた層の最大値や平均値を出力して平行移動等によるロバスト性に対応することが可能となる。活性化関数 (Rectified linear unit : ReLU) は入力を次のニューロンに伝える際の選別器としての役割があり, 特に ReLU は計算コストが低いことから最も多用される活性化関数である。使用する画像は, 脳血流 Single photon emission computed tomography (SPECT) から解析で得られる統計解析画像, Three dimensional-stereotactic surface projections (3D-SSP)⁷⁻⁸⁾, Z-score map の jpg カラー画像とした。理由として, SPECT データに比較してパターン分類が容易, 解剖学的標準化により脳形態の個人差がなく DL の識別研究に適していると予測したことによる。識別は Alzheimer's disease (AD), Dementia with Lewy bodies (DLB), 異常所見なし (Not

finding：NF) の3クラスとし、経験年数8年以上の放射線診断医の放射線レポートを正解として評価した。準備するデータセットは、トレーニングセット (Training set：Tr.set), トレーニングデータの精度検証に使用するバリデーションセット (Validation set：Va.set), 分類精度を評価するテストセット (Test set：Te.set) の3種類とした。各データの詳細について下記に記載していく。

本研究は島根大学医の倫理委員会承認を得て行っている。

2-1. Z-score map の作成

SPECT データは、 ^{123}I -IMP 167 MBq 脳血流検査から得た。収集装置は Discovery NM/CT 670 pro, 画像再構成は 3D-OSEM 法 (subsets 10, iterations 6), 散乱線補正と減弱補正は, Dual energy window 法, CT based Attenuation correction 法を用いた。3D-SSP 解析のデータベースは自施設ノーマルデータベースを使用した。3D-SSP は Ant, Post, R-lat, L-lat, R-med, L-med の8方向表示 890×660 ピクセルの, 文字, スケール領域をトリミングにより 700×365 ピクセルにリサイズして用いた。

2-2. 各データの作成

^{123}I -IMP 検査の放射線レポートにて AD, DLB, NF の所見と Z-score map の低下領域が一致する記載があるものを2019年10月から遡って各症例30例選択し, Tr.set として各25例, Te.set として各5例を無作為に選別した。3D-SSP 解析からは全

脳, 視床, 小脳, 橋の4パターンの Z-score map が算出されるため, Te.set は20例 (5症例×4パターン), Tr.set については100例 (25症例×4パターン) とした。しかし, DLB にて認知症診断を専門とする医師の判断から視床, 小脳, 橋での参照画像から診断を迷う10例を除いた90例を採用した。AD と NF も症例数を揃えるために, 同じ医師の判断にて診断を迷うものから10例ずつを除外した。解析は20回行うが, Vr.set は90例から毎回ランダムに18例を使用した。患者データ背景は AD 症例 81.7 ± 6.5 歳 (male：female = 17例：13例), DLB 74.8 ± 8.8 歳 (m：f = 18：12), NF 66.4 ± 14.0 歳 (m：f = 10：20) である。

2-3. 解析環境

解析に使用したパーソナルコンピュータは CPU Intel core i7 3.6 GHz, メモリ 16 GB, OS は Windows10, ソフトは MATLAB2019b (Mathworks 社), Deep learning toolbox を使用し入力層から出力層までを **Table 1** に示すように10層で独自に作成した。エポック数などパラメータについては研究前の基礎実験として行い最適と思われる設定とした。

2-4. 解析の評価

評価は解析を20回繰り返して算出した, AD, DLB, NF 全体での正解率：Accuracy (正解した数 / 全体の数), 陽性であると識別した総数の何割が当たっていたかを示す適合率：Precision (真陽性 / (真陽性 + 偽陽性)), 本来は陽性と判断

Table 1 使用した DL アルゴリズムの詳細を示す。

No	Processing	Matrix × Channel	Activation	Set value
1	Input layer	365 × 700 × 3		RGB
2	Covolution layer (1)	365 × 700 × 8	Relu	Stride (1), Padding (1), 3 × 3 × 3 の畳み込み
	Batch normalization (1)	365 × 700 × 8		Offset 1 × 1 × 8, Scale 1 × 1 × 8
3	Max pooling Layer (1)	182 × 350 × 8	Relu	Stride (2), Padding (0), 2 × 2 の最大プーリング
4	Covolution layer (2)	182 × 350 × 16		Stride (1), Padding (1), 3 × 3 × 8 の畳み込み
	Batch normalization (2)	182 × 350 × 16		Offset 1 × 1 × 16, Scale 1 × 1 × 16
5	Max pooling layer (2)	91 × 175 × 16	Relu	Stride (2), Padding (0), 2 × 2 の最大プーリング
6	Covolution layer (3)	91 × 175 × 32		Stride (1), Padding (1), 3 × 3 × 16 の畳み込み
	Batch normalization (3)	91 × 175 × 32		Offset 1 × 1 × 32, Scale 1 × 1 × 32
7	Max pooling Layer (3)	45 × 87 × 32	Relu	Stride (2), Padding (0), 2 × 2 の最大プーリング
8	Covolution layer (4)	45 × 87 × 64		Stride (1), Padding (1), 3 × 3 × 32 の畳み込み
	Batch normalization (4)	45 × 87 × 64		Offset 1 × 1 × 64, Scale 1 × 1 × 64
9	Full connected layer	1 × 1 × 3	Softmax	
10	Output layer	1 × 1 × 3		

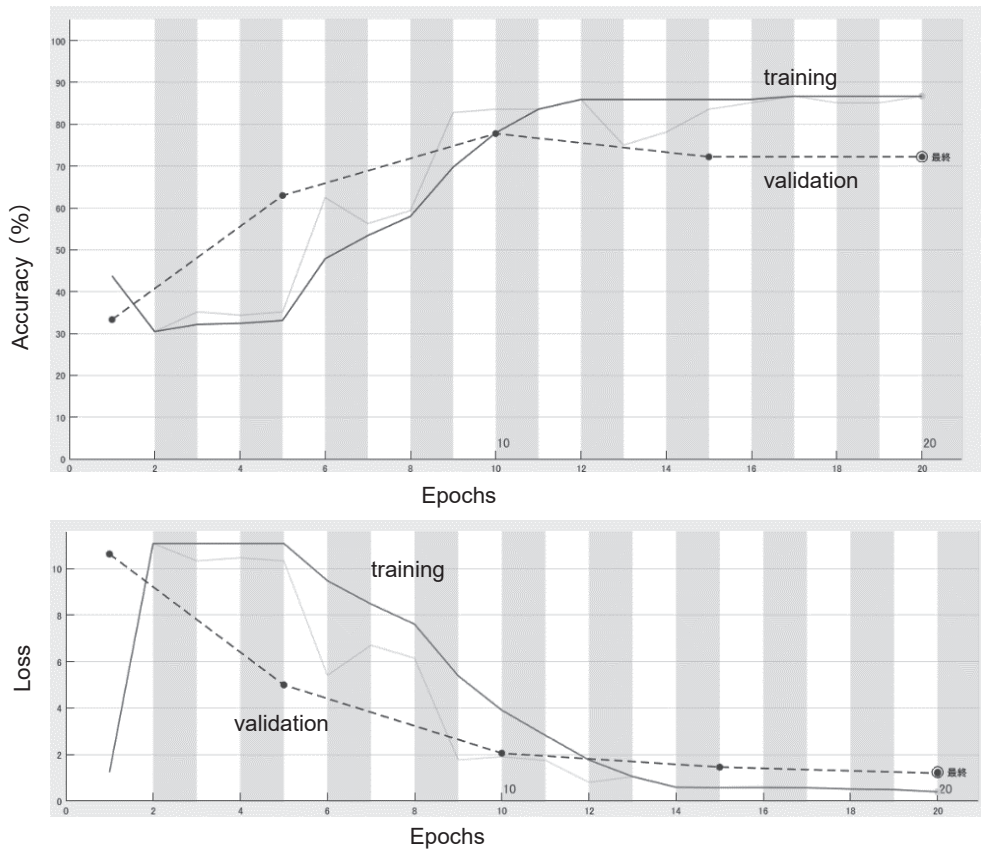


Fig. 1 グラフはトレーニングデータの精度検証の一例を示す。
解析毎に学習データの精度検証を確認した。
上段は学習データの精度、下段は学習データ内での損出データ数を示す。

すべきうちの、何割を陽性と判定できたかを示す再現率：Recall（真陽性/(真陽性+偽陰性)）、適合率と再現率のバランスを評価するF値：F-measure（ $2 \times \text{適合率} \times \text{再現率} / (\text{適合率} + \text{再現率})$ ）の平均値と標準偏差から算出した。

3. 結果

Fig. 1 は Tr.set の精度検証結果, **Fig. 2** は結果出力 Confusion matrix の一例を示し, 斜めの列が正解の数を示す。Tr.set の精度は $73.79 \pm 8.20\%$, 損出（信頼性の低いトレーニングデータ数）は 1.37 ± 0.54 , 処理時間は 4.57 ± 0.26 min であった。**Fig. 3** から Tr.set, Va.set, Te.set の正解率は, $81.1 \pm 5.9\%$, $74.5 \pm 6.8\%$, $78.5 \pm 8.9\%$, 適合率は, 82.6 ± 9.2 , 76.0 ± 11.2 , 80.9 ± 14.9 , 再現率は, 81.6 ± 13.3 , 74.5 ± 16.5 , 78.5 ± 19.6 , F 値は, 80.8

Confusion Matrix			
Output Class	AD	DLB	NC
	18 30.0%	0 0.0%	4 6.7%
	1 1.7%	20 33.3%	3 5.0%
	1 1.7%	0 0.0%	13 21.7%
Target Class			
Precision			
Recall			
Accuracy			
AD			
DLB			
NC			

Fig. 2 解析結果を示す Confusion Matrix。
縦の列から Recall を算出, 横の列から Precision を算出, 斜めの列から精度を算出。

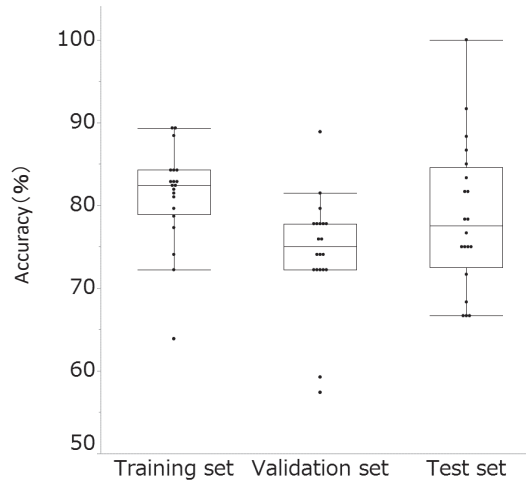


Fig. 3 20回の解析による各データの識別精度を示す。

Table 2 表は、Deep learning 解析から算出した各データセットの指標値を AD, DLB, NF で示す。

Data Type	Evaluation	AD	DLB	NC
Training set	method			
	Precision	0.794 ± 0.068	0.867 ± 0.085	0.818 ± 0.108
	Recall	0.721 ± 0.122	0.830 ± 0.111	0.882 ± 0.064
	F- measure	0.748 ± 0.079	0.841 ± 0.071	0.836 ± 0.046
Validation set	Precision	0.694 ± 0.101	0.773 ± 0.092	0.812 ± 0.112
	Recall	0.619 ± 0.152	0.808 ± 0.090	0.808 ± 0.169
	F- measure	0.642 ± 0.120	0.783 ± 0.059	0.791 ± 0.109
Test set	Precision	0.707 ± 0.109	0.883 ± 0.131	0.834 ± 0.150
	Recall	0.730 ± 0.229	0.917 ± 0.074	0.707 ± 0.185
	F- measure	0.696 ± 0.148	0.892 ± 0.078	0.746 ± 0.125

± 8.3 , 73.9 ± 12.0 , 77.8 ± 14.9 となり, Tr.set では全ての平均評価で 80% 以上であった。

Table 2 から AD, DLB, NF に分けて評価した結果, Te.set の適合率は $DLB > NF > AD$, 再現率は $DLB > AD > NF$, F 値は $DLB > NF > AD$ となり Tr.set, Va.set とともに同じ傾向にあった。Fig. 4(a) は放射線レポートでは AD であったが NF と判断, (b) は AD を DLB, (c) は DLB を AD, (d) は NF を AD と判断したケースを示す。

4. 考 察

Te.set において AD の適合率, 再現率が NF や DLB より低く, 画像パターンの特徴量が識別できないことが予測できる。例えば Fig. 4(a) は AD であり後部帯状回の血流が低下しているが,

他の領域の血流低下がなく NF へ分類された。AD の Tr.set 個々の画像を確認すると, (a) に近い画像が存在しなかった。(b) の AD 症例では後部帯状回の Z-score が高いものの, DLB の特異的低下領域である後頭葉にも低 Z-score の領域が広く存在し DLB と判断したことが予測される。また, AD や DLB の特異的低下領域の識別が容易な内側画像 (Rt.med, Lt.med) 以外, 前頭, 側頭葉領域の低下が結果に与える影響については明らかにできていない。今回, DLB が高い精度で識別できたのは, AD や NF の Tr.set 内に後頭葉低下の画像がなく, DLB 症例での特異的血流低下パターンであることを示している。(c) は DLB を AD と判断したデータ例であるが, AD と DLB 混合型の可能性もあり難しい症例であった。(d)

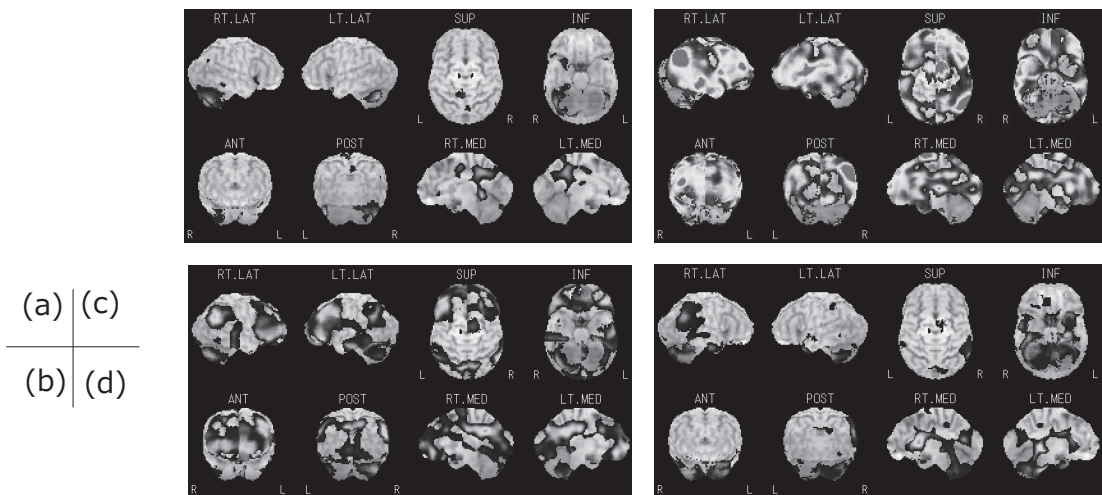


Fig. 4 識別を間違えた画像を示す。

(a) AD を NF と識別 (b) AD を DLB と識別 (c) DLB を AD と識別 (d) NF を AD と識別

は NF 症例であるが、AD に脳室周囲の Z-score が描出されているトレーニングデータが存在したためと考える。

以上から、DL のクラス分類は Tr.set と Te.set の特徴量の比較から判断されることから、できるだけ多様な AD、DLB の特異的血流パターンを学習させておくことが重要である。

解析の評価について、例えば DLB の Precision が高いのは AD や NF が DLB に識別される割合が低いことを示し、Recall は AD や NF に DLB が識別される割合が低いことを示す。これらは、F-measure が高い値を示すことになり、DLB の場合には血流低下として描出される領域の大きさや Z-score に左右されない特異的血流低下領域の認知症であることを示している可能性もある。

最近の論文では Grad-Cam⁹⁾ による CNN の判断根拠の可視化が可能になってきている。今後は Grad-Cam を用いて、画像の何処に注目してクラス分類しているのかを知ると同時に、症例数を更に増やすことで DL の精度向上に向けて研究を進める。

5. 結 論

全10層で構築した DL アルゴリズムを使用して、AD、DLB、NF の統計解析画像の識別を行い 78.5% の精度にとどまった。精度向上には、各疾患に特徴的なより多くの症例を追加する必要がある。

今後は Grad-Cam など DL の視覚化にも取り組み、精度向上に向けた研究を進める。

6. 引用文献 (references)

- 1) Chilamkurthy S, Ghosh R, Tanamala S, et al : Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. *Lancet*, **392** (10162): 2388-96, 2018
- 2) Ueda D, Yamamoto A, Nishimori M, et al : Deep Learning for MR Angiography : Automated Detection of Cerebral Aneurysms. *Radiology*, **290** (1): 187-94, 2019
- 3) Tatsugami F, Higaki T, Nakamura Y, et al : Deep learning-based image restoration algorithm for coronary CT angiography. *Eur Radiol*, **290** (1): 187-194, 2019
- 4) Ryu K, Nam Y, Gho SM, et al : Data-driven synthetic MRI FLAIR artifact correction via deep neural network. *J Magn Reson Imaging*, **50** (5): 1413-1423, 2019
- 5) Ding Y, Sohn JH, Kawczynski MG, et al : A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using ¹⁸F-FDG PET of the Brain. *Radiology*, **290** (2): 456-64, 2019
- 6) Liu M, Cheng D, Wang K, et al : Multi-Modality Cascaded Convolutional Neural Networks for Alzheimer's Disease Diagnosis. *Neuroinformatics*, **16** (3-4): 295-308, 2018
- 7) Minoshima S, Robert A, Koeppe M.A, et al : Anatomical standardization; Linear scaling and nonlinear

- warping of functional brain images. J Nucl Med, **23** (11): 1296-1309, 1993
- 8) 山本智朗, 菊池 敬, 對馬博之, 他: 核医学における臨床解析ソフトウェアの基礎と応用. 日本放射線技術学会核医学分科会 放射線医療技術学叢書 (28): p9-15, 日本放射線技術学会, 京都, 2011.
- 9) Iizuka T, Fukasawa M, Kameyama M: Deep-learning-based imaging-classification identified cingulate island sign in dementia with Lewy bodies. Sci Rep, **9** (1): 8944, 2019