

令和 3 年 5 月 27 日現在

機関番号：32665

研究種目：若手研究

研究期間：2018～2020

課題番号：18K16600

研究課題名（和文）深層学習を用いた、頭部CT画像判定装置の開発

研究課題名（英文）Development of head CT image judgment device using deep learning

研究代表者

渡辺 充（WATANABE, Mitsuru）

日本大学・医学部・助教

研究者番号：00792673

交付決定額（研究期間全体）：（直接経費） 2,200,000円

研究成果の概要（和文）：急性期脳卒中や重症頭部外傷では迅速な診断および専門的加療が必要である。しかし全ての医療機関に専門家が配置されているわけではなく、適切な判断がされない可能性がある。近年深層学習と呼ばれる技術が脚光を浴びている。本研究では深層学習の手法を用いて頭部CT画像の判定装置を作成し、診断の補助として用いることを目的とした。実際の頭部CT画像を用いて学習を行い、9割を超える精度で緊急度・重症度の高い画像を判定することが可能であった。また、ワークステーションで解析した結果をコンパクトで量産されているデバイスに搭載しても数秒で結果が判定可能であり、実臨床で用いることが出来ると考えられた。

研究成果の学術的意義や社会的意義

急性期脳卒中や重症頭部外傷では迅速な診断および専門的加療が必要である。しかし全ての医療機関に専門家が配置されているわけではなく、適切な判断が行われない可能性がある。本研究では深層学習の手法を用いて頭部CT画像の判定装置を作成し、診断の補助として用いることを目的とした。同様の目的で遠隔診療が考えられる。しかし、対応する医師の負担が生じることや、個人情報である医用画像を施設外に送信する必要がありセキュリティ上の懸念が生じる。本研究で開発した画像判定装置を用いればこれらの問題が解決され、適切な医療を受ける機会が確保出来ると考えられた。

研究成果の概要（英文）：Acute stroke and severe head injury require prompt diagnosis and professional treatment. However, not all medical institutions have specialists. Therefore, the diagnostic imaging may be delayed.

In recent years, a technique called deep learning has been in the limelight. The purpose of this study was to create a head CT image judgment device using a deep learning method to assist in medical diagnosis.

For CT images with high urgency and severity, the program was able to judge with an accuracy of over 90%. Then, the results analyzed by the workstation were installed in a compact and mass-produced device. The device was able to judge the CT image in a few seconds. It had the judgment accuracy and speed that could be used clinically.

研究分野：脳神経外科学

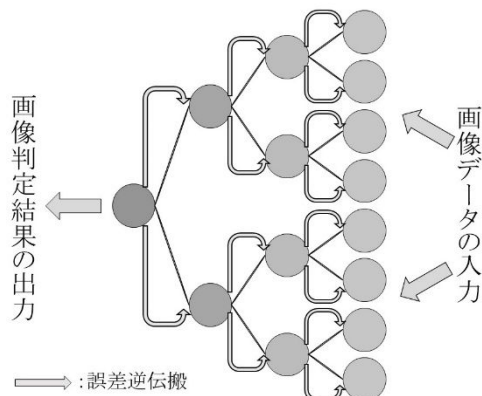
キーワード：深層学習 脳卒中 頭部外傷 画像診断 人工知能 画像認識

### 1. 研究開始当初の背景

近年「深層学習」と呼ばれる技術が脚光を浴びている。Image Net Large Scale Visual Recognition Challenge (ILSVRC) と呼ばれる画像認識のコンテストは、予め与えられた千種類、数十万枚の画像をコンピュータに学習させ、続いて未知の画像を判定させて正答率を競うというものである。人間は未知の画像を判別した場合の正答率が 95% 程度とされているが、2015 年には機械学習が 95% を僅かに超えた。

ここで用いられた画像判定の手法は、畳み込みニューラルネットワークと呼ばれる方法である。これは簡単に説明すると回帰式を求めるように、与えられたデータに対して最適な関数を求める方法である。この関数が多重に重なっているため「深層」と呼ばれる。また関数に代入した結果と実際の答えの誤差を比較して、関数の係数を調整し精度を高める誤差逆伝搬法を用いるため「学習」と呼ばれる(図1)。

図1 畳み込みニューラルネットワーク



従来このような計算をコンピュータに行かせた場合、非常に計算量が多いため高性能 CPU を用いても限られた時間で結果を得ることが難しかった。最近のこの分野の発展は、これまでコンピュータグラフィックスの表示に用いていたハードウェア Graphics Processing Unit (GPU) を流用することで、安価に並列計算が行えるようになったことによる。GPU は元々コンピュータディスプレイの各ドットに表示させる色を計算するため、数百から数千の単純計算を並列で行うことができる装置である。そのため、深層学習のような、計算量が多いが単純な課題を解くことに向いている。

また、多くの分野で注目されるようになり、GPU を搭載したカードサイズのデバイスが登場したことで、自動車の自動運転技術や産業ロボット等に組み込まれ普及を後押ししている。研究代表者は、このデバイスが脳画像診断に役立つことに着目した。

脳画像研究では Voxel Based Morphometry (VBM) に代表される研究手法が先行して用いられてきた。この方法は、画像の準備に制約(同一のモダリティで同一のシーケンスの MRI 画像を全例で撮像する必要がある)があり、CPU を用いなければ解くことが難しい計算が多く含まれる。従って、いつでも行えるわけではなく、緊急時に用いることには向かないため、実臨床とかけ離れていると言わざるを得ない。

本研究では広く普及しており、緊急時にも行うことが可能な頭部 CT 画像を基に、診療を補助するツールの開発を目指した。特に緊急度の高い頭部 CT 画像について、遠隔診療に頼らずとも、開発するデバイスが専門外の医師の判断をサポートすることができるようになることを目標とした。

### 2. 研究の目的

本研究は実際の臨床で用いられる頭部 CT 画像を用いて、緊急度・重症度の臨床的判断について、(1)深層学習の手法を用いて妥当な正答率で画像を分類することが可能であることを証明する。(2)量産されており安価でコンパクトなデバイスを用いて、臨床上許容できる時間で頭部 CT の判定が行えることを証明する。(3)臨床の現場に持ち込むことのできるサイズのデバイスを開発することを目的とした。

### 3. 研究の方法

#### 対象

日本大学医学部附属板橋病院脳神経外科で頭部 CT の検査が行われた、脳卒中および頭部外傷患者およびその疑いのある患者を対象とした。対象となる頭部 CT 画像は 2067 例であった。このうち、同一症例の短期間における繰り返しの撮影は、画像の差異に乏しく学習には不相当と考えられたため除外した。手術後の症例は画像の変化が顕著ではあるが回復過程を表していることが多く、また本研究の趣旨である、専門外の医師が画像の判定をする場面を想定しにくいと考え除外した。これにより対象症例数は 243 例であった。

## データの加工、分類

頭部 CT の画像データを災害時のトリアージタグに習い、「緑」専門的な加療の必要なし、「黄」緊急度は高くないが専門的な加療を要する、「赤」緊急度が高く直ちに専門的な加療を要する、「黒」加療を行っても転帰を改善させられる見込みがない」の 4 つに分類した。分類は脳神経外科専門医の判断と実際の経過を基に行った。

CT 装置より CD-R を介して出力した Digital Imaging and Communications in Medicine (DICOM) 画像を、DICOM サーバーである OsiriX に入力し管理した。全脳を撮影した画像のうち、等間隔に 8 枚を抽出し、256 × 256 pixel の png 形式の画像を用意した。一般的にサンプル数が少ない画像データを深層学習で扱う場合、data augmentation と呼ばれる手法で、画像データに対して、反転・回転・拡大縮小・ノイズ付加などを行い、十数倍のサンプルデータを作成して用いられる<sup>1)</sup>。これはサンプルに対して過学習してしまうことを避けるとともに、サンプル数の不足を補う方法である。本研究でも data augmentation を用いて 33048 枚の画像データを作成した。

これらの画像データから、深層学習を行うための学習用データ (80%)、調整用データ (10%)、性能評価用データ (10%) に分けて用意した。

## 解析方法

用いた機材は、学習用ワークステーションとして HP 製 Z820 (CPU: Intel Xeon E5-2630 2.60GHz × 2, memory 192GB, GPU: NVIDIA Tesla K20c × 2) を用いた。臨床の現場に持ち込むコンパクトな画像判定装置として、NVIDIA 製 Jetson AGX Xavier 開発者キットを用いた。

いずれも搭載した開発環境は OS: Ubuntu 18.04, CUDA 10.0.130, cuDNN 7.3.0.29, フレームワークは TensorFlow 1.11.0 を用いた。GUI 環境として NVIDIA DIGITS 6 を用いた。用いたネットワークは AlexNet を元に調整した。

## 頭部 CT 判定装置の開発

学習の結果を基に頭部 CT 判定装置の開発を行った (図 2)。NVIDIA Jetson AGX Xavier 開発者キットは GPU が搭載されたデバイスである (図 3)。これにワークステーションで学習した結果を搭載し、临床上現実的な時間内に妥当な判断が可能であることを検証した。

## 4. 研究成果

本研究では当初頭部 CT 画像を前述の 4 種類に分類し判定することを目指した。しかし症例数が不十分であったため、9 割以上の判定率とすることは困難であった。

今回の趣旨を考えると、特に緊急度の高い画像を漏らさずに判定できることが重要と考えられた。「緑・黄」の緊急度が低い画像、「赤」の緊急度が高い画像を確実に判別できるように分類を改めた。「黒」の専門的な加療を行っても改善が見込めない画像については、専門家に判断を仰ぐ必要があるという点では変わらないため、「緑・黄」と「赤・黒」を分類する課題に変更した。その結果、精度は 94.4% であった (図 4)。グラフの横軸は学習回数を示す。学習用データを繰り返し学習させることによって、性能評価データの正答率 (橙色) が徐々に上昇している。一方で本プログラムの回答と正解の誤差を示す損失関数 (青色: 学習データ、緑色: 性能評価データ) が低下していることが分かる。

このように学習した結果を画像判定装置に搭載した。画像判定を行ったところ、平均 2.8 秒で認識させることが出来た。これは临床上十分な時間であることが明らかとなった。

図2 CT画像の撮影から画像判定までの流れ

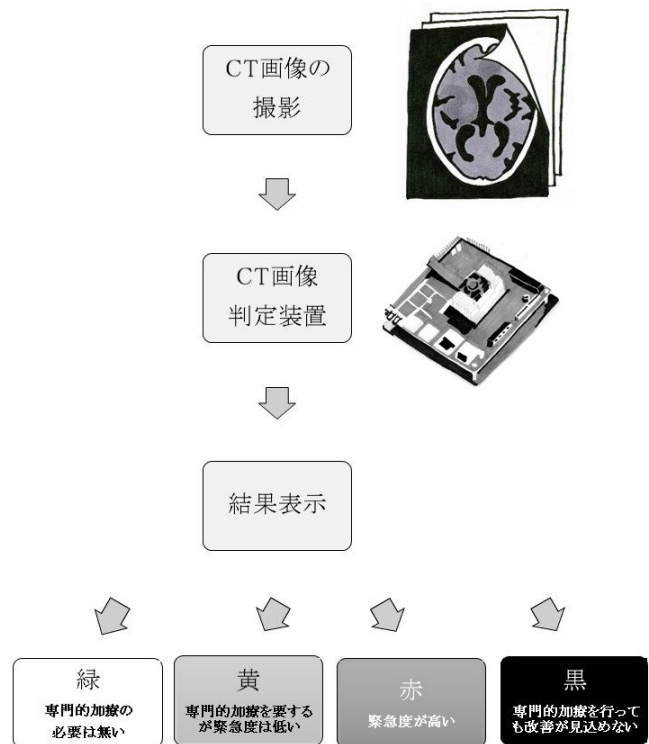


図3 手の上に乗る大きさの Jetson AGX Xavier

ここまでの研究では、画像を CT 装置から CD-R を介して、直接出力したものをオフラインで用いてきた。実臨床で用いることを想定すると、CT 装置に直接接続し画像を収集することが望ましい。しかし、セキュリティ上の懸念があるため、病院から許可を得ることは難しいと考えられた。そこで、画像判定装置にカメラを接続して、電子カルテや CT 装置のコンソール画面に表示された画像からデータを収集する方法を考えた（図 5）。

しかし画像判定精度は著しく低く、臨床で用いられる精度とはならなかった。

原因としてモニター画面への映り込みや、干渉縞といった特有のノイズが発生することが考えられた。Data augmentation で付与されるノイズと異なるため精度が上がらなかったと考えられた（図 6）。本件については、学習データの収集時点から、カメラを用いるなどの工夫が必要になると考えられた。

#### 今後の課題

現在新たな問題として人間ならば決して誤認することのないような画像が、深層学習でご認識してしまう問題が知られている。一例として自動車の自動運転技術で想定される、道路標識の認識において、道路標識の画像にわずかなノイズを加えることで「STOP」と「制限速度 80」を誤認してしまうという問題が報告されている<sup>2)</sup>。このような AI 騙しのようなエラーが発生するか否かは、臨床で使用して初めて明らかとなる予測困難な問題である。

また、AI が誤った判断をした場合、その責任の所在がどこにあるのかの議論もなされている。AI を作成したプログラマーの責任なのか、AI に判断を任せられた医師の責任なのか、責任の所在については 2018 年 12 月に厚生労働省が「人工知能 (AI) を用いた診断、治療等の支援を行うプログラムの利用と医師法第 17 条の規定との関係について」(医政医発 1219 第 1 号) という通達を出している。この通達では「人工知能 (AI) を用いた診断・治療支援を行うプログラムを利用して診療を行う場合についても、診断、治療等を行う主体は医師であり、医師はその最終的な判断の責任を負うこととなり、当該診療は医師法 (昭和 23 年法律第 201 号) 第 17 条の医業として行われる」とされている。現在のところあくまで AI は医師の補助をするのみという位置付けとされている。

#### 結語

画像判定装置として、量産されるコンパクトで安価なハードウェアを用いて、臨床に十分な精度と時間を有する装置を開発することが出来た。しかし当初予定した 4 分類で画像判定精度を確保するには、サンプル数の増加が必要であると考えられた。また、画像入力方法として、カメラを用いてデータを入力する方法を検討したが、ノイズの乗り方の違いから十分な精度を確保することが難しかった。

#### 文献

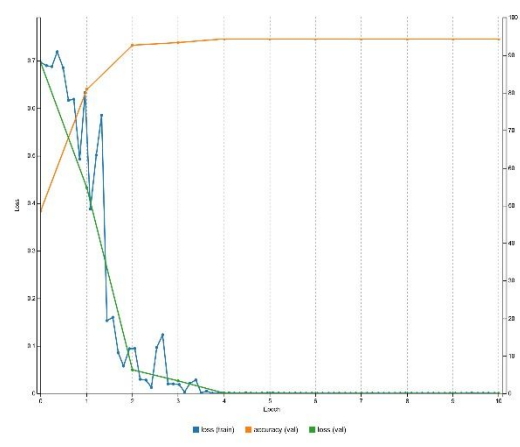


図 4 画像判定精度のグラフ

橙色：性能評価データの正答率（右軸）

青色：学習データの損失関数（左軸）

緑色：性能評価データの損失関数（左軸）

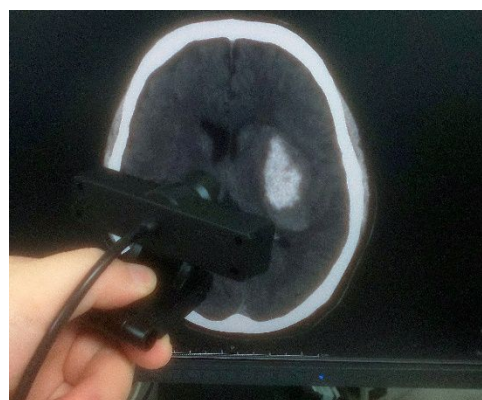


図 5 カメラでコンピュータモニターを撮影



図 6 コンピュータ画面を撮影し、映り込み・干渉縞のある画像

- 1) Connor S, et al., A survey on Image Data Augmentation for Deep Learning, Journal of Big Data, 2019
- 2) Kevin E, et al., Robust Physical-World Attacks on Deep Learning Visual Classification, Computer Vision and Pattern Recognition, 2018

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計2件（うち招待講演 1件 / うち国際学会 0件）

1. 発表者名 渡辺 充、青木 宏之、茂呂 修啓、塩川 諒治、前田 剛、大島 秀規、吉野 篤緒
2. 発表標題 頭部外傷後高次脳機能障害と大脳皮質の形態的变化（シンポジウム、神経外傷の画像診断）
3. 学会等名 第42日本脳神経外傷学会
4. 発表年 2019年

1. 発表者名 渡辺 充
2. 発表標題 DBS手術におけるMRI画像～DTIの応用～（ランチョンセミナー）
3. 学会等名 第43回日本脳神経CI学会（招待講演）
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------