

Deep Learning と時間分解近赤外分光鏡を用いた 予防医学システムの開発

大山 勝徳¹、胡 莉珍²、酒谷 薫²

¹日大工・情報、²日大工・電気電子

本研究は、NIRS による脳機能データ及び血液検査データを用いて適切な Deep Neural Network (DNN) を構築し、高齢者の認知機能障害を早期に発見するシステムを開発することを目的とする。本年度は、認知機能障害リスクの機械学習モデルを評価し、その有用性を考察した。

【1】MMSE による認知機能スクリーニングテスト

本研究では最初に、Mini Mental State Examination (MMSE) テストによって認知機能を計測する。このテストは認知症スクリーニングテストのために精神状態を体系的に評価することに有用であることが知られている[1]。MMSE スコアは、30 点満点の中で、24 点以上を異常なし、18 点から 23 点までを軽度認知障害 (MCI)、17 点以下を認知症 (severe cognitive impairment) と分類している。数ある指標の中で MMSE は最も信頼され、世界中で使用されている。ただし、MMSE による認知機能スクリーニングテストは、テストの内容に対する被験者の誤解がないことを前提として実施される必要がある点、時間を要する点、そして MMSE スコア 24 点から 26 点の間に相当する被験者の異常なしと MCI を識別するときに感度が比較的低いことがおもな課題であ。したがって本研究では、MMSE スコアの実測値を目的変数として、生体情報の分析結果から MMSE スコアに基づく効率的な認知機能の評価する方法を考察する。

【2】時間分解近赤外分光鏡 (TRS) による高齢者認知機能障害の評価法

今回の分析に用いた生体情報は、近赤外分光法の一つである時間分解分光法 (Time Resolved near-infrared Spectroscopy: 以下 TRS) による計測結果である。TRS は、ピコ秒パルス光と光拡散方程式を用いて、安静時のヘモグロビン (Hb) 濃度の絶対値を計測することが可能である[2]。

対象は、脳外科外来患者 202 名 (男性 87 名 女性 115 名、平均 73.4±13.0 歳)。MMSE テストを実施した後に、TNS 測定プローブを両側前額部に配置し、前頭前野の安静時の酸素化ヘモグロビン (HbO₂) 濃度 (μM)、脱酸素化ヘモグロビン (HbO) 濃度 (μM)、酸素飽和度 (SO₂)、さらには近赤外線光の発光部から受光部までの光路長を計測した。

【3】Deep Neural Network による MMSE スコアの推定

最初に、MMSE スコアを目的変数として、計測した TNIRS データそれぞれの項目の相関度を求めて、次に紹介する DNN による MMSE スコアの推定結果と比較する。今回の分析に用いた DNN は、Fig 1 のように、年齢を加え、2ch の TNIRS データ (SO₂、HbO、HbO₂、THb、PL1、PL2、PL3) を合わせた計 15 個の項目を説明変数としている。ここで、Photon path lengths (PLs) は光路長を表し、PL1 は 760nm 帯域の光路長、PL2 は 800nm 帯域の光路長、PL3 は 830nm 帯域の光路長を表す。今回使用した DNN のモデルは

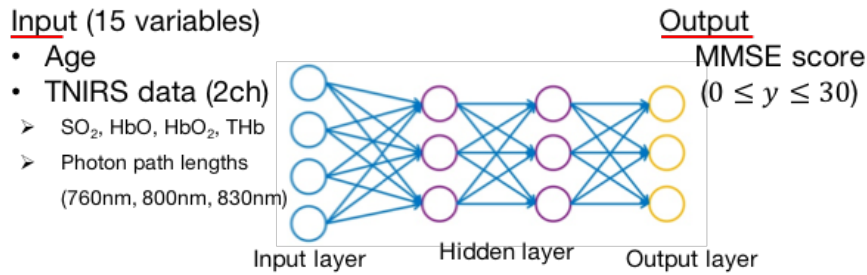


Fig 1. Structure of the deep neural network for the data analysis. Input vectors include age and TRS data on the left and right PFCs. The output vector is regression to estimate the MMSE score. The hidden layer contains no backward connections from downstream layers.

H2O プロジェクトの Deep Learning ライブラリ [3][4]を使用して実装したものである。隠れ層 (Hidden layer) は 2 層あり、各層に 400 個のニューロンユニットが配置されている。入力される説明変数から各ニューロンユニットに対する重み ($\alpha = \sum_{i=1}^n w_i x_i + b$) を付けて、次の層のニューロンユニットに対する出力 $f(\alpha)$ を決めている。すなわち、関数 f は非線形の活性化関数であり、バイアス b によりニューロンユニットの活性化の閾値を調整し、かつ、重み付けされた各入力 x_i の総和により $f(\alpha)$ が決まる。重み w_i を決定するために、教師データとなる目的変数と推定値から求める平均二乗誤差を損失関数 $L(W, B | j)$ とし、 L が最も低い重み w_i の組み合わせを求める。ここで W は j 番目の層における重み w_i の組み合わせを表す集合を表し、もう一方の B は j 番目の層におけるバイアスの組み合わせを表す。

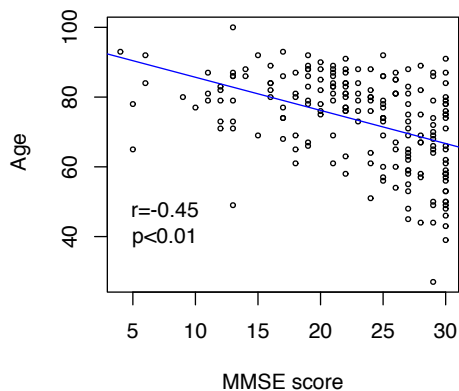


Fig 2. Relationship between MMSE score and age.

【4】MMSE スコアと各パラメータの相関

相関分析の結果、被験者の年齢と MMSE スコアとの間に有意な相関関係 ($r=-0.45$, $p < 0.01$) が認められた (Fig 2)。次に、TNIRS による酸素飽和と MMSE スコアの間にも有意な相関関係 ($r=0.38$, $p < 0.01$) を認められた。このことから、年齢や酸素飽和度は MMSE スコアの推定に最も貢献する可能性が高い。以下、相関度に関する表を Table 1 に示す。

Table 1. Comparison of TRS parameters between normal subjects and subjects with impaired cognitive function.

		Normal (MMSE ≥ 24)	Impaired (MMSE ≤ 24)	p value
Left	oxy-Hb	37.23	35.02	$p < 0.1$
	deoxy-Hb	17.94	19.22	$p < 0.05$
	total-Hb	55.17	54.24	-
	SO ₂	67.38	64.41	$p < 0.01$
	Optical path length (761 nm)	19.02	19.38	-
	Optical path length (791 nm)	19.31	19.51	-
	Optical path length (836 nm)	17.97	18.14	-
Right	oxy-Hb	39.41	35.80	$p < 0.01$
	deoxy-Hb	18.73	19.20	-
	total-Hb	58.15	54.99	$p < 0.1$
	SO ₂	67.69	64.90	$p < 0.01$
	Optical path length (761 nm)	19.03	19.00	-
	Optical path length (791 nm)	19.35	19.93	-
	Optical path length (836 nm)	18.25	18.79	-

【5】DNNによるクラス分類精度

今回使用したDNNと他の代表的な機械学習手法(ランダムフォレスト、サポートベクターマシン)による識別結果との比較を行った結果、Fig 3のように、MMSEスコア24未満のクラスとMMSEスコア24以上のクラスの2分類においてはDNNの結果が91.5%の識別精度(accuracy)で最も高い。このとき、202個のサンプルデータで数時間以内に計算可能、かつ、検証結果の偏りが少なくなる方法として、Leave-one-out交差検証を用いた。他の機械学習手法については基本的なハイパーパラメータで実行しており、調整することにより少し改善する余地を残すが、15個の説明変数からなるTRSのデータにおいてDNNは比較的適している可能性が高い。今回構成したDNNは各ニューロンユニットについて50%のドロップアウト率とL1正則を適用してあり、過学習は比較的少ない。ただし、サンプルデータが脳外科外来患者に限られる点では、それ以外のデータを合わせて今後も評価する必要がある。

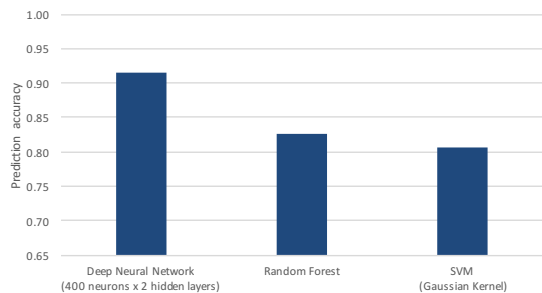


Fig 3. Comparison of classification accuracy by leave-one-out cross-validation (n=202). The DNN was tuned with 400 neurons in two hidden layers and a 50% dropout rate. The other machine learning algorithms were launched with basic hyper parameters, i.e., 500 trees and five variables tried at each split for the random forest and 0.1 of sigma 1 of cost C and for the SVM with Gaussian kernel function.

また特筆すべきこととして、PL1 から PL3 の光路長もまた MMSE スコアの推定に大きく貢献していることが分かった。前節の Table 1 に見られるように、MMSE スコアに対する相関度に関する有意差は認められない代わりに、MMSE ス

コアの計測値に基づく異常なし (Normal) と認知症リスクあり (Impaired) の群間で平均値に大きな差がある。このことは頭蓋骨から大脳新皮質の距離が影響するため、脳萎縮の程度に関する間接的な影響も示唆している。

【6】DNNによる推定スコア

次に、DNN の出力を回帰の結果として使用し、そのまま出力を推定値とすると、Fig 4 のように、MMSE テストによる計測結果と DNN による推定結果の間に正の相関 ($r=0.85$, $p<0.01$) が認められた。一方の RMS 誤差 (Root Mean Square error) は 3.02 であり、すなわち、推定値が標準分布に従うと仮定したとき ± 3.02 のバラつきがあることに留意する必要がある。特に計測値が 18 未満の被験者について偽陰性になるパターンがあるため、その条件について今後検証の余地がある。

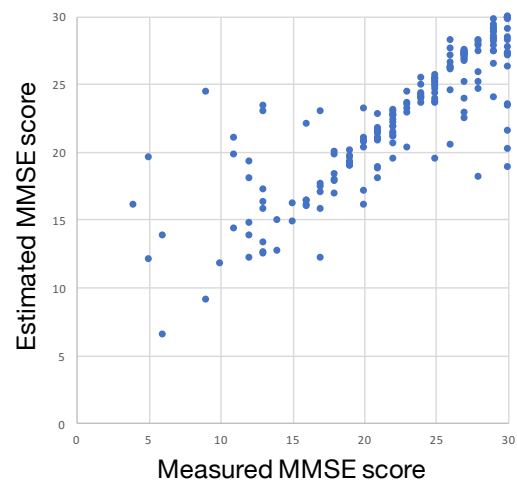


Fig 4. Correlation between measured and estimated MMSE scores ($r=0.85$, $p<0.01$).

【7】まとめ

本報告では、認知機能障害リスクを評価するための DNN モデルによるクラス分類精度や推定スコアの特徴を考察した。今後の課題として、MMSE スコアの計測値と推定値の残差が大きいグループに関する分析をさらに進める必要がある。また、サンプルデータを増やして DNN モデルのハイパーパラメータを調整することによる

実用性を向上させる一方で、DNN によるクラス分類結果や推定結果の可読性のたる根拠を現場の病院関係者へ提示する方法を今後求めていきたい。

参考文献

- [1] Tombaugh, TN, McIntyre, NJ (1992) The mini-mental state examination: a comprehensive review. *J Am Geriatr Soc* 40(9):922-935
- [2] 酒谷 薫 (2015) 次世代 NIRS – 時間分解 NIRS (TRS) による脳循環と脳機能計測. *Clinical Neuroscience* 33:716-718
- [3] Candel, A., Lanford, J., LeDell, et al. (2017) Deep Learning with H2O. https://h2o-release.s3.amazonaws.com/h2o/rel-slater/9/docs-website/h2o-docs/booklets/DeepLearning_Vignette.pdf. Accessed August 2017
- [4] Reddy, Y. B. (2017) Classification and automatic recognition of objects using H2O package. *Proceedings of SPIE* 10185, Cyber Sensing 101850F