

BMIのための入力脳波長の異なるCNNのブレンドによる 変化に鋭敏な実時間分類

Real-time EEG Classification Sensitive to Changes by Blending CNNs with Different Signal Lengths

太田貴士¹⁾, 舟橋健司²⁾

Takashi OTA and Kenji Funahashi

1) 名古屋工業大学 情報工学科 (〒 465-8555 名古屋市昭和区御器所町, t.ota.787@stn.nitech.ac.jp)

2) 名古屋工業大学 情報工学科 (〒 465-8555 名古屋市昭和区御器所町, kenji@nitech.ac.jp)

概要: 脳波で機械を操作するためのインターフェースの研究が多く行われており, ニューラルネットワークを用いた高い精度での認識手法が提案されている。しかし, これらの分類モデルは精度向上のために時間軸方向に入力が長く, 入力脳波の変化への反応が遅くてリアルタイム性に乏しい。本研究では, 入力信号長の異なる複数の CNN をブレンドすることでリアルタイム性を向上させる手法を提案する。

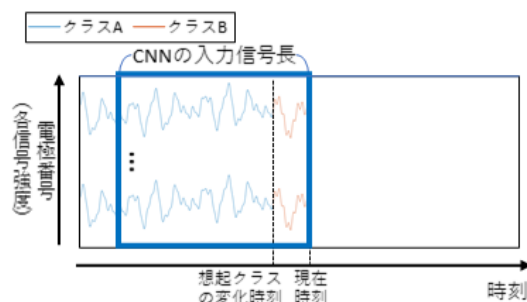
キーワード: ブレインマシンインターフェース, 脳波, ニューラルネットワーク

1. はじめに

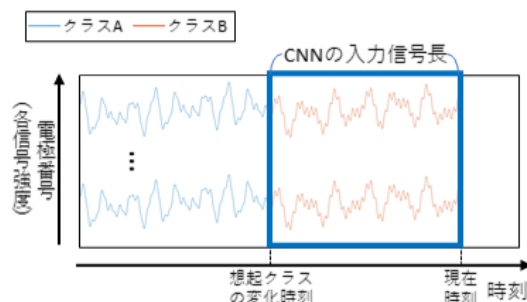
ブレインマシンインターフェース (BMI) は, 脳波などの脳情報を利用して脳と機械をつなぐ技術である。脳波は脳内で生じる電気活動であり, 頭部に設置した電極で記録される。近年は, 小型で可搬性に優れた, かつ時間的分解能が高い脳波計測機器が開発されていることから, 脳波 BMI の研究が盛んに行われている。例えば, サルの脳波でロボットアームを操作する研究 [1] や人間の脳波で車いすを操作する研究 [2] などが行われてきた。また, Schirmeister らはニューラルネットワークの一種である CNN でも従来の脳波解読技術である FBSCP [3] と同等の高い精度で脳波を分類できることを示した [4]。これにより, 高い精度で脳波分類が可能であり, かつ自動的に特徴選択が行えるニューラルネットワークを利用した脳波分類の手法が注目を集めるようになった。

CNN による脳波分類は, 横軸を時間軸, 縦軸を電極番号とした 2 次元平面に対する脳波強度データを CNN に入力して, 最終的に得られた予測確率が最大のクラスの出力を得るというものである。ところで, リアルタイムで脳波を分類するためには, CNN に対する入力信号が時間軸方向に長い場合にはユーザの想起クラスの変化, すなわち脳波の変化への追従が遅れてしまうというデメリットがある。すなわち, 図 1 (a) のように, ユーザの想起クラスが変化した直後は CNN への入力に変化後の想起クラスはほとんど含まれないため, 変化前の想起クラスの予測確率が高くなってしまい, 図 1 (b) のように変化後の想起クラスの脳波が CNN への入力の大部分を占めるようになると, ようやく変化後の想起クラスの予測確率が高くなる。Schirmeister らの研究 [4] で使われた CNN のうちの 1 つのアーキテクチャは入力データが 250Hz サンプリングの脳波に対して時間方向に 522 サンプルであった。もし, この CNN をそのままリアルタイム

対話システムに利用する場合, 図 1 (b) のように, 初めて入力信号が全て変化後の想起クラスの脳波となるのは, ユーザの想起クラスが変化してから, $522/250 \approx 2.1$ 秒後のことである。一方, この遅延を緩和するために単純に入力信号長を短くしてしまうと, 今度は分類性能が下がってしまう。筆者らが行った予備実験では, 2 つのクラスに対して各 750 個の学習データで, 入力信号長が 128 サンプルの CNN と入力信号長が 256 サンプルの CNN を学習させ, 両者の分類性能を



(a) 想起クラスが変化した直後



(b) 入力信号が全て変化後の想起クラスになったとき

図 1: 想起クラス変化時の入力信号の変化

比較したところ、前者は 64.1%、後者は 71.6%となった。このように、想起コマンドの変化への反応速度と分類性能がトレードオフの関係になっていて、ニューラルネットワークによりリアルタイム脳波分類を行うことは困難である。本研究では、時間軸方向の入力信号長の異なる CNN の出力結果をブレンド [5] することによって、この遅延を緩和する手法を提案する。

2. 提案手法

時間軸方向の入力信号長の異なる CNN の出力結果をブレンドして、その結果を得る手法を提案する。ブレンドは複数の弱分類器を加重平均することで、全体としての分類性能を向上させる方法である。図 2 は本研究で提案する手法を表す模式図である。入力信号長の長い CNN_{LONG} はまだ入力の半分以上を想起クラス変化前の脳波データが占めているため、現在時刻において既に想起クラスは A から B に変化しているが、クラス A である確率を 0.8 と示している。しかし、入力信号長の短い CNN_{SHORT} は既に入力の全体を想起クラス変化後の脳波データが占めているため、想起クラスの変化に対応できており、クラス B である確率を 0.95 と示している。これらの CNN の出力の加重平均を算出することで、全体として想起クラスの変化への対応が期待できる。このように時間軸方向の入力信号長の異なる CNN の出力結果をブレンドすることで、入力信号長の長い CNN だけを使う場合に比べて想起クラスの変化に早く反応できるようになり、かつ入力信号長の短い CNN に比べて分類性能が高くなると予想される。

3. 実験

3.1 概要

上記の提案手法の性能を実証するために、以下の手順で実験を行った。

1. 単一のクラスを想起させた学習用の脳波と途中で想起クラスを変化させたテスト用の脳波を計測する。
2. 学習データで入力信号長の異なる複数の CNN を学習させる。
3. テストデータに対して提案手法を適用し、提案手法の性能を評価する。

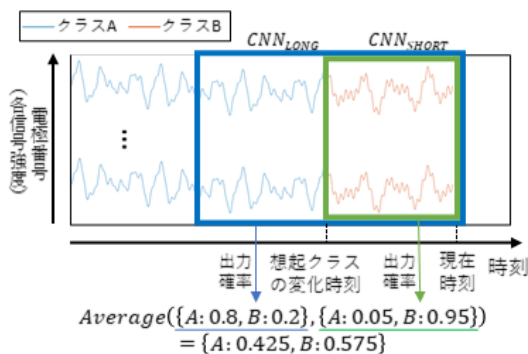


図 2: 提案手法

表 1: クロップのステップ幅と学習データ数

CNN の入力サイズ	ステップ幅	学習データ数
256	80	1152
384	96	928
512	136	632

3.2 データセット

本実験では、20 代男性の被験者一名の脳波を計測した。脳波の計測には、Emotiv EPOC X を利用した。電極数が 14、内部サンプリングレートは 2048Hz で、256Hz にダウンサンプリングされた信号を取得することができる。学習用脳波データは、右手開閉動作の想起と左手開閉動作の想起の 2 クラスそれぞれについて、16 秒間の計測を 1 試行として、24 試行計測したものである。テスト用脳波データ計測中は、計測の途中で被験者に想起クラスを変化させる指示を画面上に表示する。このような想起クラスの変化を含む 8 秒間の計測を 1 試行として、右手の動作から左手の動作に変化するパターンとその逆のパターンの 2 種類それぞれ 12 試行計測した。なお、学習用脳波データ、テスト用脳波データともに、計測タスクに対する慣れを防ぐために、計測タスクをランダムにシャッフルして計測した。

3.3 CNN の学習

CNN の学習にあたり、学習用脳波データを各 CNN の入力サイズ (サンプリングレート 256Hz のデータに対するサンプル数) にクロップした。クロップとは CNN の入力サイズよりも大きいデータから CNN の入力サイズに切り出すことであり、切り出す時刻を変えることで複数のデータとして扱う。画像認識の分野でよく使われているデータの水増しの手法であり、Schirrneister らの研究 [4] で脳波データに対してもクロップによる水増しが学習に有効であることが確認されている。なお、あるトライアルデータの特定部分が多数のクロップデータに含まれていると過学習になる恐れがあるので、CNN の時間軸方向の入力サイズに応じてクロップ時のステップ幅を変えた (表 1)。

各入力信号長における CNN の構成は、Schirrneister らの研究 [4] を参考にした。ただし、畳み込み窓のサイズを入力信号のサイズに合わせて変更し、ドロップアウトの割合とフィルター数を学習データの量に合わせて変更した。それ以外は Schirrneister らの研究 [4] の Deep ConvNet と同じ構成である。

3.4 提案手法の性能評価

予備実験より、CNN の入力サイズは 256, 384, 512 の 3 通りに設定した。また学習結果である分類器の精度はそれほど高くないためシードの違いにより異なる弱学習器が得られた。そのため同じ入力サイズの異なる分類器をブレンドするだけでも精度が向上するので、異なる入力サイズでの

ブレンドの比較対象を、同じ入力サイズでの同数のブレンドとした。まず、各入力サイズにつき6通りの乱数シードとともに学習データで学習させることで計18の分類器を得たうえで、以下の通り時間変化に伴う出力、すなわち各クラス確率の遷移を計算した。

(1) 8秒間のテストデータに対して16サンプル(約0.06秒)ステップで対応する長さのデータを各分類器に入力した。

(2) 続いて以下の組み合わせで、(1)の結果を均等な重みでブレンドした。なお、以降では括弧内の略称により表記する。

(a) 入力サイズ256の分類器6個(256×6)

(b) 入力サイズ384の分類器6個(384×6)

(c) 入力サイズ512の分類器6個(512×6)

(d) 入力サイズ256の分類器2個、入力サイズ384の分類器2個、入力サイズ512の分類器2個(256×2+2+2)

(e) 入力サイズ256の分類器3個、入力サイズ384の分類器2個、入力サイズ512の分類器1個(256×3+2+1)

(3) 最後に、各ブレンドに対して、(2)の結果のうち通常時の分類性能と想起クラス変化時の反応速度を調べた。すなわち、8秒間のテストデータを序盤の想起クラス変化前のフェーズ1、中盤の想起クラス遷移中のフェーズ2、終盤の想起クラス変化後のフェーズ3の3つに分割して、それぞれのフェーズでの正答率を計算した。

想起クラスの変化に鋭敏な分類器は、フェーズ2での精度が高くなると予想される。また、通常時の分類性能が高い分類器は、フェーズ1とフェーズ3での精度が高くなると予想される。なお、脳科学分野では、視覚刺激に追従するボタン押し操作実験などにおいて反応時間が200-300ms程度であると言われている。そこで想起クラス変化の指示が提示されてから被験者が変化後のクラスを実際に想起し始めるまでの遅延を考慮して、フェーズ2の開始時刻は、画面に想起クラスが提示されてから約0.25秒(64サンプル)後とした。また、遅延は各データによりばらついていることも考慮して、フェーズ2の継続時間は約2.5秒(640サンプル)間と十分に長く設定した。

4. 実験結果と考察

(1) 各分類器単体での出力例を、図3の(x)、(y)、(z)に示す。左手開閉動作に対応するクラスLeft、右手開閉動作に対応するクラスRightと分類される確率は、1節に記した通り、入力信号長が長い方が定常時の確率は比較的安定して高いが、短い方が想起クラスの変化に早く追従していることが分かる。

(2) 各ブレンドの結果のうち、あるテストデータに対する出力を図2の(a)-(e)に示す。まず、(x)と(a)、(z)と(c)を比較すると、ブレンドした方が比較的安定して高い確率を示していることが分かる。また(a)-(c)に注目すると、やはり、入力サイズの小さい分類器ほど早く想起クラスの変化に追従していることが分かる。加えて、(a)-(c)と異なる入力サイズの分類器をブレンドした(d)、(e)と比較すると、比較的安定して高い確率を示しつつ脳波の変化への対応が早くなる

ことが確認できる。

(3) 各ブレンド、各フェーズ毎の全てのテストデータに対する平均正答率を表2に示す。まず、フェーズ2の正解率は、(a)256×6のブレンド分類器では0.60、(b)384×6のブレンド分類器では0.54、512×6のブレンド分類器では0.52であり、入力サイズの小さい分類器ほどフェーズ2の正解率が高い、すなわち脳波の変化への対応が早いことが確認できる。また、(d)256×2+2+2と(e)256×3+2+1のブレンド分類器のフェーズ2の正解率はそれぞれ0.56と0.62である。入力サイズの異なる分類器の出力をブレンドすることで、比較的安定して高い確率を示しつつ脳波の変化への対応が早くなることが確認できる。

それぞれの結果について考察する。脳科学分野では反応時間は200-300ms程度であると言われているが、本入力データ採取時には「可能な限り迅速に想起クラスを変化させる」ことを意識していなかった。自動車の急制動時反応時間(空走時間)は意識的な実験でもペダルの踏み替えまでに0.75秒必要と言われている。そこで想起クラスの変更指示が画面に提示されてから被験者が実際に想起クラスを変えるまでにおよそ1秒要していたと仮定し、実際に脳波変化からの遅延は(c)512×6において1.1秒であったとすると、提案手法の(e)256×3+2+2では0.7秒と、大幅に短縮できている(図3)。さらに、(a)256×6の2.5秒から3.5秒付近では推定確率が50%は超えているものの100%には程遠く、(c)512×6では100%に近い値を示しているのに対して、(d)、(e)では比較的高い値を示している。すなわち、提案手法は変化への反応が早い上に入力サイズの小さい分類器に比べて通常時の分類性能が高いと言える(図2)。また、(e)は(d)に比べてフェーズ2の正解率が高いが、フェーズ1と3では低い(表2)。ブレンドの割合が高い入力サイズCNNの特性がブレンド結果により反映されているのだろう。ブレンド割合やブレンド時の重みも重要なパラメータであると考えられる。ところで、(a)-(c)3つの分類器に対するフェーズ1と3の全データに対する総合的な正答率に大きな差は無かった。入力サイズの大きい分類器ほど想起クラスの変化を含まない脳波に対する分類性能が高いという傾向は見られなかった。これは、入力サイズ512のCNNは、学習データが少なく(表1)、分類器の性能を十分に高められなかったためであろう。2クラス分類では推定確率50%をしきい値にクラス分類を行うため、例えば推定確率が55%と高くない場合でも正答と判断されてしまうが、複数クラス分類を想定した場合にはより高い性能が期待される。

5. むすび

本研究では、脳波の変化に鋭敏に反応することができるBMIを目指して、入力サイズの異なる分類器をブレンドする手法を提案した。また、ブレンドの割合を変えることで分類性能が変わることも確認できた。しかし、入力サイズ512サンプルのCNNが学習データ不足により、上手く学習できていなかった可能性があるため、提案手法の通常時の脳波分

類の精度の高さを検証することはできなかった。今後の展望としては、敵対生成ネットワークなどでデータの水増しを試みるなどして入力サイズ 512 サンプルの CNN の学習データ不足を解消して再度実験を行ったり、より適切な入力サイズの組み合わせを調べたりしたい。また、より定常時の分類性能を高く、脳波変化時の反応速度を早くするために、脳波の変化を検知して動的にブレンドの割合を変化させることも試みたい。

謝辞 本研究の一部は JSPS 科研費 JP20K11918 の助成を受けたものです。

参考文献

- [1] Johan Wessberg, et al.: Real-time Prediction of Hand Trajectory by Ensembles of Cortical Neurons in Primates, *Nature*, Vol. 408, No. 1, pp. 361–365, 2000.
- [2] 田中一男: 脳波指令で動く車いすの開発, *精密工学会誌*, Vol. 78, No. 8, pp. 662–665, 2012.
- [3] Ang KK, et al.: Filter Bank Common Spatial Pattern (FBCSP) in Brain-Computer Interface, *IEEE International Joint Conference on Neural Networks 2008, IJCNN*, pp. 2390–2397, 2008.
- [4] Schirrneister RT, et al.: Deep Learning With Convolutional Neural Networks for EEG Decoding and Visualization, *Human Brain Mapping*, Vol. 38, No. 11, pp. 5391–5420, 2017.
- [5] Andreas Töschler and Michael Jahrer: The BigChaos Solution to the Netflix Grand Prize, https://www.netflixprize.com/assets/GrandPrize2009_BPC_BigChaos.pdf.

表 2: 各組み合わせのフェーズ毎の平均正答率

ブレンド分類器	フェーズ		
	1	2	3
(a) 256×6	0.64	0.60	0.88
(b) 384×6	0.63	0.54	0.88
(c) 512×6	0.62	0.52	0.87
(d) 256×2+2+2	0.63	0.56	0.91
(e) 256×3+2+1	0.58	0.64	0.90

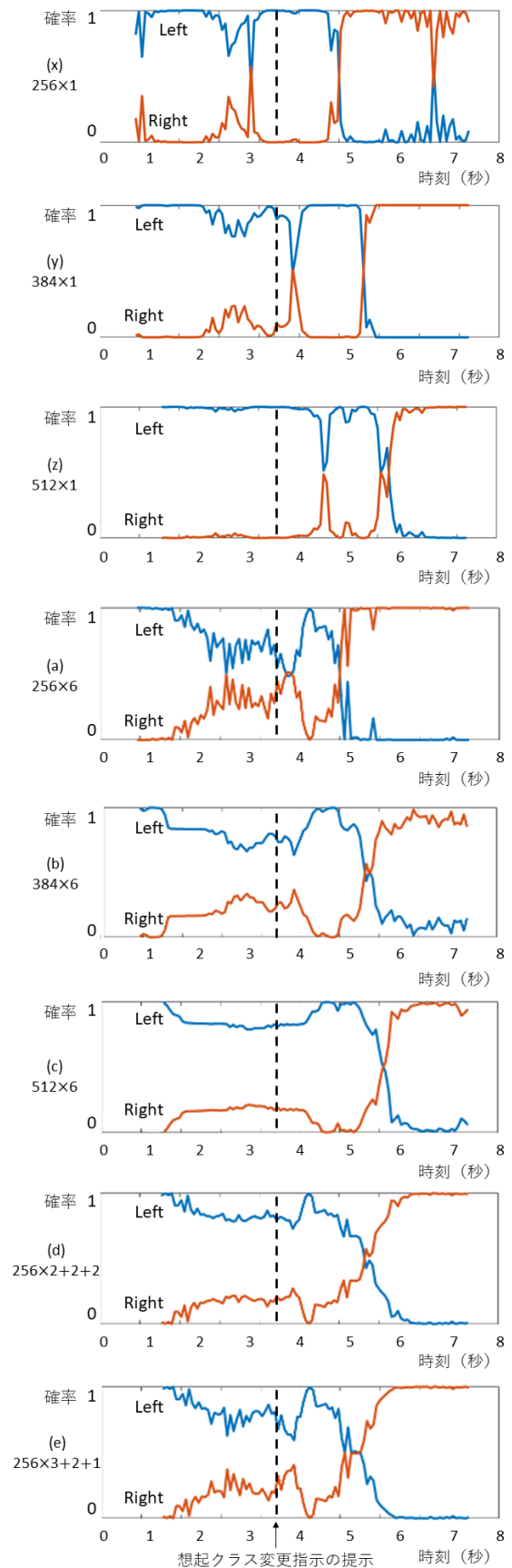


図 3: 出力確率の遷移