

メタ学習による個人に適応した脳波からの感情予測

Meta-Learning for Personalized Emotion Prediction from EEG Signals

宮本 佳奈^{*1*2}

Kana Miyamoto

田中 宏季^{*1*2}

Hiroki Tanaka

中村 哲^{*1*2}

Satoshi Nakamura

^{*1}奈良先端科学技術大学院大学

Nara Institute of Science and Technology

^{*2}理化学研究所 革新智能統合研究センター

Center for Advanced Intelligence Project, RIKEN

We have developed an emotion induction system that predicts participants' emotions from EEG and provides personalized music. Although it is important to secure the amount of data for training emotion prediction models, it is a burden for the participants to record EEG data for a long time. In this study, we aim to investigate a training method for using a small amount of EEG data. We propose using meta-learning that trains a pre-training model that can be adapted easily to each participant. As a result of predicting valence and arousal from EEG, the method with meta-learning showed a significantly lower prediction error than the method without meta-learning ($p < .001$).

1. はじめに

日常生活において、感情のコントロールが求められる場面が存在する。例えば、プレゼンテーション中に感じる緊張は発表のパフォーマンスを低下させるおそれがあり、発表時には緊張の緩和が求められる [徳久 19]。また、感情状態はうつ病などの精神疾患にも関係することから、感情喚起はこれまで盛んに研究されてきた。音楽は、感情を変化させるために使用されてきた刺激の一つである [Kreutz 08, Schubert 13]。しかし、同じ音楽を聴いてもすべての人が同じ感情になるわけではなく、また同一人物であっても状況によって誘発される感情は異なる [Larsen 91]。このことから我々は、効果的な感情喚起には個人の状態に合わせて音楽を生成することが必要であると考えている。我々はこれまでに、脳波から参加者の感情を予測し、予測した感情を利用して音楽を生成するシステムを提案した [Miyamoto 22]。このシステムは3つの要素から成る。1つ目の要素は、音楽生成器である。音楽生成器は、入力された感情と似た感情を誘発させる効果のある音楽を作る。2つ目の要素は、脳波から感情を予測するモデルである。音楽生成器へ喚起したい感情を入力しても、音楽聴取時の個人の状態によって喚起したい感情を必ずしも参加者が感じるわけではない。このため、モデルは音楽を聴取して実際に誘発された感情を脳波から予測する。3つ目の要素は、音楽生成器のコントローラである。脳波から予測された感情を使用して、喚起したい感情に参加者の感情を近づけられるよう、音楽生成器に入力する感情を調整する。参加者の感情状態を反映して音楽をリアルタイムに生成することで、効果的な感情喚起が提供できると考えた。以上の3つの要素を備えたシステムの構築と検証を行なった結果、システムを使用して作られた音楽が感情喚起に有効であることが示された [Miyamoto 22]。

現在、我々は感情喚起システムをプレゼンテーション中や会話中などの実環境で使用することを目指している。これを実現するための課題の一つとして、感情予測モデルの学習のための、脳波の収録時間の短縮を挙げる。脳波は個人差が大きいことから、参加者ごとにモデルを学習することが望まし

い [Lin 17]。モデルの学習に使用するデータ量は重要であり、データ不足はモデルの性能の低下に繋がることが指摘されている [Wang 18]。このため、モデルを使用する前に可能な限り長く脳波収録を行うことが望ましい。しかし、脳波計を長時間装着することは参加者にとって負担となるため、モデルの性能を保ちながら脳波の収録時間を短縮することが求められている。そこで本稿では、少量の脳波データを用いた高精度な感情予測モデルの学習手法を検討する。

転移学習は少量のデータを用いたモデルの学習によく利用される。これは、あるドメインで学習した事前学習モデルを別のドメインへ少量データを用いて適応させる学習手法である。脳波を用いた予測にも転移学習は利用されており、大量の脳波データで学習された事前学習モデルを、目的のドメインの対象の参加者の脳波データで適応する試みがされてきた。しかし、事前学習モデルは複数人の脳波を混同したデータから学習された [Lan 18, Miyamoto 22]。学習に使用するサンプルが誰の脳波データであるか識別されていないため、従来の事前学習モデルは各個人の脳波の特徴を適切に考慮できていないと考えられる。つまり、一部の参加者には適応しやすいが、他の参加者には適応しにくい事前学習モデルを学習している可能性がある。少量の脳波データでモデルを学習するためには、どの参加者にも適応しやすい事前学習モデルの獲得が望まれる。この問題を解決するために、Model-Agnostic Meta-Learning (MAML) [Finn 17] を使用した学習が提案されている。MAMLは、さまざまなタスクに適応しやすいモデルを少量のデータで学習するメタ学習である。脳波を取り扱う研究において、1つのタスクを1名分の脳波データとすることで、誰にでも適応しやすい一般性のあるモデルを学習できると考えられている。MAMLは脳波からの睡眠レベルの分類 [Banluesombatkul 20]などに利用され、有効性が示されている。また、脳波を用いた感情予測にも使用されており、ミュージックビデオを刺激とした Dataset for Emotion Analysis using EEG, Physiological and Video Signals や映画を刺激とした SJTU Emotion EEG Dataset を使用して、MAMLの有効性の検証が行われている [Duan 20]。ただし、音楽聴取中の脳波からの感情予測にMAMLは利用されていない。そこで本稿では、MAMLが我々の感情喚起システムを構成する要素である音楽聴取中の脳波を用いた感情予測モデルの学習にも有効であるかを調査する。

連絡先: 宮本 佳奈, 奈良先端科学技術大学院大学, 奈良県生駒市高山町 8916-5, 0743-72-5265, miyamoto.kana.mk4@is.naist.jp

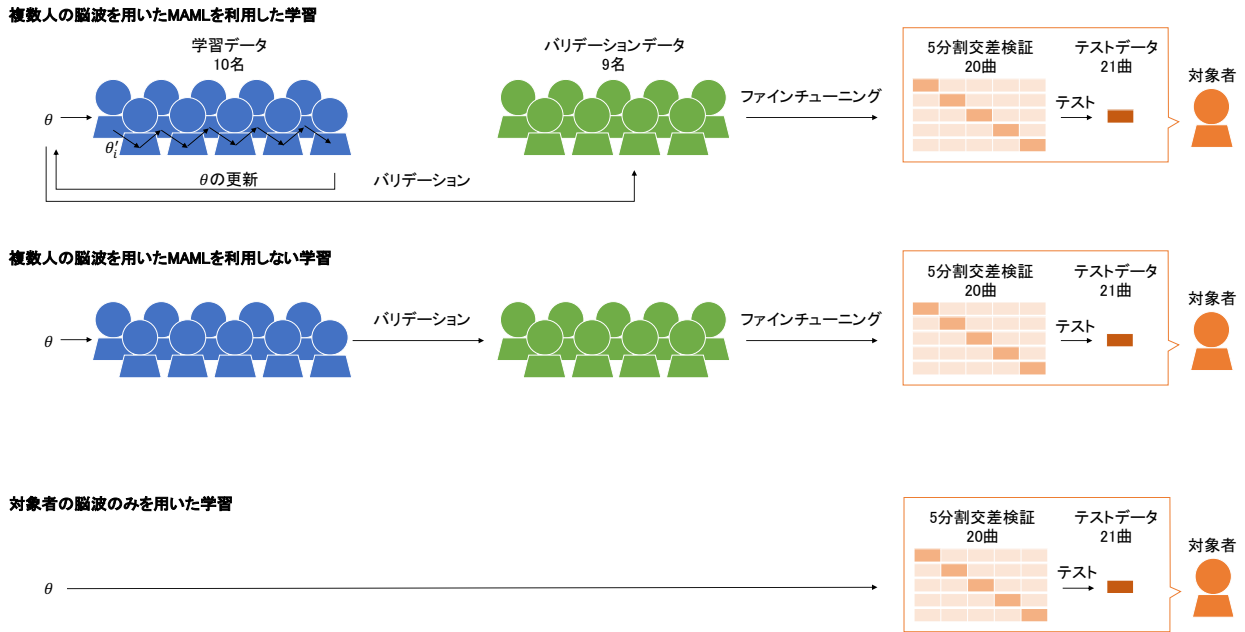


図 1: 少量の脳波データを用いた感情予測モデルの 3 種類の学習手法

我々は以下の 3 種類の学習手法を比較する。

- 提案手法
 - 複数人の脳波を用いた MAML を利用した学習
- ベースライン手法
 - 複数人の脳波を用いた MAML を利用しない学習 [Lan 18, Miyamoto 22]
 - 対象者の脳波のみを用いた学習 [Ehrlich 19, Miyamoto 22]

1 つ目は、複数人の脳波を用いて MAML を利用して事前学習モデルを学習し、対象者の少量の脳波データを使用してモデルをファインチューニングする提案手法である。2 つ目は、複数人の脳波を用いて MAML を利用せずに事前学習モデルを学習し、対象者の少量の脳波データを使用してモデルをファインチューニングするベースライン手法である。3 つ目は、対象者の少量の脳波データのみを使用してモデルを学習するベースライン手法である。我々はこの 3 手法を用いて、音楽聴取中の感情の valence と arousal を予測する。さらに予測値と真値の Root Mean Squared Error (RMSE) を計算し、3 手法を比較する。本稿の内容は、我々の先行研究の日本語版である [Miyamoto 21]。

2. 実験方法

MAML が音楽聴取中の少量の脳波からの高精度な感情予測に有効であるか調査するため、MAML を使用しない 2 つのベースライン手法と感情予測精度を比較する。本章では、実験に使用した脳波のデータセット、学習に利用した特徴量とモデル構造を示した後、モデルを学習する 3 手法について述べる。

2.1 データセット

本実験で使用したデータセットは、我々の先行研究で作成したものである [Miyamoto 22]。このデータセットは、音楽を

聴取している間の脳波と実験参加者が評価した自身の感情を収録したものである。この収録は、奈良先端科学技術大学院大学の研究倫理委員会の承認を受けて行われた。インフォームドコンセントに同意した健康な 20 名（男性 10 名、女性 10 名）が収録に参加した。参加者の年齢は、19 歳から 37 歳（平均 24.3 歳）であった。データ収集に使用された音楽は、音楽生成器を利用して作られた 20 秒間の 41 曲の音楽である。音楽の詳細な情報は、我々の先行研究に示している [Miyamoto 22]。また、音楽のサンプルは以下の URL で公開されている (<https://sites.google.com/view/music-generator>)。参加者は、音楽聴取によって誘発された valence と arousal を Self-Assessment Mannequin (SAM) [Bradley 94] を用いてそれぞれ 0 から 1 の範囲の 9 段階で評価した。1 曲ごとに評価された valence と arousal を音楽聴取中の参加者の感情の真値としている。我々は、この真値を脳波から予測することを目指す。また、脳波の計測には CGX 社製の Quick-30 を使用した。

2.2 特徴量とモデル構造

我々は感情予測のモデルとして Convolutional Neural Network (CNN) を採用した。CNN の入力は脳波計の電極の位置関係と脳波の周波数成分を考慮した行列である。行列の作成のため、まずデータセットの脳波を 1 秒ごとに切り分け、バンドパスフィルタを使い、theta (4-7 Hz), alpha (8-13 Hz), low beta (14-21 Hz), high beta (22-29 Hz), gamma (30-45 Hz) の 5 つの周波数帯に分けた。電極ごとに 5 つの周波数帯の脳波波形の分散の対数を計算した。このようにして得られた 1 曲あたり 20 サンプルの特徴量を脳波計の電極配置に従って行列に当てはめた。音楽聴取後に SAM で評価された参加者の感情の valence と arousal が 20 サンプルの真値として設定された。詳細な特徴量の計算と行列の作成方法は、我々の先行研究に示している [Miyamoto 21]。

CNN の構造は、畳み込み層（フィルタサイズ: 2×2 , スライド幅: 1, フィルタの数: 8), バッチ正規化層, ReLU 層, 畳み込み層 (2×2 , 1, 8), バッチ正規化層, ReLU 層, 畳み込み層 (2×2 , 1, 8), バッチ正規化層, ReLU 層, 全結合層

(出力次元: 2), 回帰出力層である. 最適化手法は Stochastic Gradient Descent を適用した.

2.3 少量の脳波データを用いた感情予測モデルの学習

データセットには 20 名分のデータが含まれている. 1 名ずつ対象者に設定し, 各手法で 20 回の学習を行なった. 学習時間の削減のため, 調整するハイパーパラメータの種類を限定している. また, テストデータや学習データは各参加者において 3 手法で統一している.

2.3.1 複数人の脳波を用いた MAML を利用した学習

この提案手法では, 複数人の脳波を用いて MAML を利用して事前学習モデルを学習し, 対象者の少量の脳波データを使用してモデルをファインチューニングする. モデルの学習手法を図 1 の上部に示した. 我々のデータセットの中からランダムに取り出した 10 名分のデータを学習データとし, 1 名分のデータを 1 つのタスク T_i とみなした. また, それぞれのタスクにおいてランダムに取り出した 20 曲分のデータをサポートセット, 残りの 21 曲分のデータをクエリセットとした. まず, タスクごとにサポートセットを使用して, 初期化したモデルパラメータ θ から学習率 α で θ'_i を学習した.

$$\theta'_i = \theta - \alpha \nabla_{\theta} \mathcal{L}_{T_i}(f_{\theta}) \quad (1)$$

パラメータの更新をした後, 10 タスクそれぞれでクエリセットを使用し, 学習したパラメータのロスを計算した. その後, 各タスクで計算されたロスの合計が小さくなるよう, パラメータ θ を学習率 $\beta = 10^{-1}$ で更新した.

$$\theta \leftarrow \theta - \beta \nabla_{\theta} \sum_i \mathcal{L}_{T_i}(f_{\theta'_i}) \quad (2)$$

上記の学習に使用されなかったデータセットの 9 名の全曲分のデータをバリデーションデータとし, 学習率 $\alpha \in \{10^{-1}, 10^{-2}\}$ のパラメータチューニングを行なった. またパラメータの更新は, バリデーションデータで計算されたロスが 5 回連続して下がらなくなるまで続けられた.

このようにして学習された事前学習モデルを使用し, 対象者のデータを用いてファインチューニングをした. 対象者のデータからランダムに 21 曲分のデータを取り出し, これをテストデータとした. また, 残りの 20 曲を使用し 5 分割交差検証を行なった. 15 曲分のデータを学習データ, 5 曲分のデータをバリデーションデータとし, 学習率 $\gamma \in \{10^{-1}, 10^{-2}\}$ のパラメータチューニングを行なった.

$$\theta \leftarrow \theta - \gamma \nabla_{\theta} \mathcal{L}(f_{\theta}) \quad (3)$$

またパラメータの更新は, バリデーションデータで計算されたロスが 5 回連続して下がらなくなるまで続けられた. 5 分割交差検証によりこれを 5 回繰り返し, テストデータを使用して予測値と真値の RMSE を計算した.

2.3.2 複数人の脳波を用いた MAML を利用しない学習

このベースライン手法では, 複数人の脳波を用いて MAML を利用せずに事前学習モデルを学習し, 対象者の少量の脳波データを使用してモデルをファインチューニングする. モデルの学習手法を図 1 の中部に示した. 我々のデータセットの中からランダムに取り出した 10 名の全曲分のデータを学習データとした. 学習データをシャッフルし, バッチサイズを 1024 とし, 初期化したモデルパラメータ θ を更新した. 学習に使

表 1: 3 手法による予測値と真値の RMSE

参加者	MAML あり		MAML なし		対象者のみ	
	val	aro	val	aro	val	aro
1	0.300	0.348	0.310	0.374	0.371	0.378
2	0.281	0.283	0.407	0.334	0.427	0.408
3	0.274	0.329	0.307	0.361	0.357	0.456
4	0.141	0.167	0.153	0.169	0.306	0.326
5	0.332	0.271	0.379	0.314	0.374	0.373
6	0.276	0.350	0.364	0.394	0.387	0.456
7	0.361	0.387	0.449	0.454	0.478	0.470
8	0.283	0.292	0.331	0.328	0.380	0.380
9	0.235	0.360	0.301	0.380	0.378	0.356
10	0.103	0.197	0.116	0.205	0.377	0.264
11	0.193	0.276	0.249	0.314	0.289	0.328
12	0.204	0.252	0.260	0.276	0.346	0.323
13	0.220	0.321	0.266	0.360	0.314	0.374
14	0.213	0.264	0.285	0.290	0.380	0.377
15	0.253	0.297	0.289	0.297	0.360	0.362
16	0.120	0.196	0.231	0.251	0.258	0.318
17	0.424	0.392	0.452	0.406	0.511	0.468
18	0.191	0.189	0.240	0.238	0.382	0.383
19	0.106	0.246	0.202	0.292	0.233	0.323
20	0.375	0.317	0.415	0.385	0.424	0.405
平均	0.244	0.287	0.300	0.321	0.367	0.377
標準偏差	0.090	0.066	0.092	0.071	0.066	0.056

用されなかった 9 名の全曲分のデータをバリデーションデータとし, 学習率 $\delta \in \{10^{-1}, 10^{-2}\}$ のパラメータチューニングを行なった.

$$\theta \leftarrow \theta - \delta \nabla_{\theta} \mathcal{L}(f_{\theta}) \quad (4)$$

またパラメータの更新は, バリデーションデータで計算されたロスが 5 回連続して下がらなくなるまで続けられた. ファインチューニングの方法は, MAML を利用する手法と同じである.

2.3.3 対象者の脳波のみを用いた学習

このベースライン手法では, 対象者以外の脳波は使用しない. したがって 1 名分のデータからモデルが学習された. モデルの学習手法を図 1 の下部に示した. 上記の複数人の脳波を用いた学習のファインチューニングと同じ手順で初期化したモデルパラメータ θ を更新した.

3. 実験結果と考察

20 名をそれぞれ対象者と設定し, 3 手法によって valence と arousal を予測した. 予測値と真値の RMSE を表 1 に示す. この表より, 全参加者の RMSE の平均が最も低くなったのは, valence と arousal 共に MAML を利用した学習手法であった. また, フリードマン検定により, valence と arousal のそれぞれで 3 手法間に有意差を認めた ($p < .001$). さらに, ボンフェローニ補正を適用したウィルコクソンの符号順位検定により, MAML を利用した学習と複数人の脳波を用いるが MAML を利用しない学習や, MAML を利用した学習と対象者の脳波のみを用いた学習の間に valence と arousal 共に有意差を認めた ($p < .001$). また, MAML を使用しないベースライン手法間でも valence と arousal 共に有意差を認めた ($p < .001$).

以上の結果より, 複数人の脳波を用いた学習は対象者の脳波のみを用いた学習よりも低い RMSE で感情を予測できることがわかった. 対象者の脳波のみを用いた学習は, 他の参加者の脳波特性の影響を受けず, 対象者に適応しやすいと考えられ

る。しかし、今回の実験設定の場合、対象者のみでは学習データ量が不足しており、複数人の脳波を用いた学習の方が高精度に感情を予測できたと考えられる。また、複数人の脳波を用いた場合は、MAMLを利用することでより低いRMSEで感情を予測できることがわかった。MAMLを使用して個人を認識した学習を行なったことで、一般性が高く、少量の脳波データでも対象者に適応しやすい事前学習モデルが学習できたと考えられる。

4. まとめ

我々はこれまでに脳波から感情を予測し、予測結果を利用して個人に適応した音楽を生成する感情喚起システムを構築した。本稿ではシステムの利便性の向上のため、脳波から感情を予測するモデルの学習のための脳波収録時間を短縮することを目指した。MAMLは少量データを用いた学習に効果的であることから、MAMLが少量の音楽聴取中の脳波データを用いた感情予測に対しても有効であるかを調査した。複数人の脳波を用いてMAMLを利用した手法、複数人の脳波を用いるがMAMLを利用しない手法、対象者の脳波のみを用いる手法の3手法を比較した。この結果、提案手法のMAMLを利用した手法が最も低いRMSEで感情を予測でき、ベースライン手法との間に有意差を認めた。したがって、MAMLは音楽聴取中の少量の脳波データからの感情予測にも有効であることが示された。

メタ学習の研究は盛んに行われており、MAMLを改良したアルゴリズムも提案されている[Raghu 20, Oh 21]。脳波の特性を利用してアルゴリズムを改良することをこれからの課題として挙げ、我々は今後も高精度な感情予測を実現する学習手法を調査する。

謝辞

本研究はJST, CREST, JPMJCR19A5の支援を受けたものである。また、本研究は理研の大学院生リサーチ・アソシエイト制度の下での成果である。

参考文献

[Banluesombatkul 20] Banluesombatkul, N., Oupphan, P., Leelaarporn, P., Lakhan, P., Chaitusaney, B., Jaimchariya, N., Chuangsuwanich, E., Chen, W., Phan, H., Dilokthanakul, N., et al.: Metasleep learner: A pilot study on fast adaptation of bio-signals-based sleep stage classifier to new individual subject using meta-learning, *IEEE Journal of Biomedical and Health Informatics* (2020)

[Bradley 94] Bradley, M. M. and Lang, P. J.: Measuring emotion: the self-assessment manikin and the semantic differential, *Journal of behavior therapy and experimental psychiatry*, Vol. 25, No. 1, pp. 49–59 (1994)

[Duan 20] Duan, T., Shaikh, M. A., Chauhan, M., Chu, J., Srihari, R. K., Pathak, A., and Srihari, S. N.: Meta learn on constrained transfer learning for low resource cross subject EEG classification, *IEEE Access*, Vol. 8, pp. 224791–224802 (2020)

[Ehrlich 19] Ehrlich, S. K., Agres, K. R., Guan, C., and Cheng, G.: A closed-loop, music-based brain-computer

interface for emotion mediation, *PLoS one*, Vol. 14, No. 3, pp. 1–24 (2019)

[Finn 17] Finn, C., Abbeel, P., and Levine, S.: Model-agnostic meta-learning for fast adaptation of deep networks, in *International conference on machine learning*, pp. 1126–1135 (2017)

[Kreutz 08] Kreutz, G., Ott, U., Teichmann, D., Osawa, P., and Vaitl, D.: Using music to induce emotions: Influences of musical preference and absorption, *Psychology of music*, Vol. 36, No. 1, pp. 101–126 (2008)

[Lan 18] Lan, Z., Sourina, O., Wang, L., Scherer, R., and Müller-Putz, G. R.: Domain adaptation techniques for EEG-based emotion recognition: a comparative study on two public datasets, *IEEE Transactions on Cognitive and Developmental Systems*, Vol. 11, No. 1, pp. 85–94 (2018)

[Larsen 91] Larsen, R. J. and Ketelaar, T.: Personality and susceptibility to positive and negative emotional states., *Journal of personality and social psychology*, Vol. 61, No. 1, p. 132 (1991)

[Lin 17] Lin, Y.-P. and Jung, T.-P.: Improving EEG-based emotion classification using conditional transfer learning, *Frontiers in human neuroscience*, Vol. 11, p. 334 (2017)

[Miyamoto 21] Miyamoto, K., Tanaka, H., and Nakamura, S.: Meta-Learning for Emotion Prediction from EEG while Listening to Music, in *Companion Publication of the 2021 International Conference on Multimodal Interaction*, pp. 324–328 (2021)

[Miyamoto 22] Miyamoto, K., Tanaka, H., and Nakamura, S.: Online EEG-based Emotion Prediction and Music Generation for Inducing Affective States, *IEICE Transactions on Information and Systems*, Vol. E105-D, No. 5 (2022)

[Oh 21] Oh, J., Yoo, H., Kim, C., and Yun, S.-Y.: {BOIL}: Towards Representation Change for Few-shot Learning, in *International Conference on Learning Representations* (2021)

[Raghu 20] Raghu, A., Raghu, M., Bengio, S., and Vinyals, O.: Rapid Learning or Feature Reuse? Towards Understanding the Effectiveness of MAML, in *International Conference on Learning Representations* (2020)

[Schubert 13] Schubert, E.: Emotion felt by the listener and expressed by the music: literature review and theoretical perspectives, *Frontiers in psychology*, Vol. 4, p. 837 (2013)

[Wang 18] Wang, F., Zhong, S.-h., Peng, J., Jiang, J., and Liu, Y.: Data augmentation for eeg-based emotion recognition with deep convolutional neural networks, in *International Conference on Multimedia Modeling*, pp. 82–93 (2018)

[徳久 19] 徳久弘樹, 大野直紀, 中村聡史: 自身のみが聴取可能な音楽によるプレゼンテーション支援手法の提案, Technical report, 電子情報通信学会 (2019)