

博士論文 2025 年度（令和 7 年度）

生体信号識別に適した
機械学習フレームワークの構築

慶應義塾大学大学院 政策・メディア研究科

星野 貴行

生体信号識別に適した 機械学習フレームワークの構築

論文要旨

生体信号を用いたヒューマン・コンピュータ・インタフェース（生体信号型 HCI）は、身体接触や視覚・音声に制約がある状況でも利用可能な次世代インタフェースとして期待されており、機械学習による高精度な識別技術がその中核を担う。しかし、生体信号は計測負担が大きく大規模データの収集が困難であるうえ、被験者間の個人差が極めて大きく、従来の機械学習モデルは新規ユーザーに対して性能が不安定になるという問題が生じていた。この問題に対し、転移学習は個人差の影響を軽減しつつ事前計測被験者（ソース被験者）の情報を活用する有力な手法とされてきたが、その効果は多数のソース被験者が確保できる場合に限定されるという制約を抱えており、データ収集の困難性という本質的な課題は未解決であった。本研究は、少数のソース被験者データからでも個人差に適応し高精度な識別を実現する機械学習フレームワークを構築することを目的とし、三段階の手順で検討を行った。第一に、実用性の観点から機械学習モデルと転移学習手法の組合せを体系的に比較評価し、深層モデルとパラメータ空間型転移の組合せが最も高精度かつデータ効率に優れる第一選択であることを実証した。第二に、ソース被験者の最適な適用戦略を検証し、被験者ごとにモデルを学習しアンサンブル統合するマルチソース方式が、データセットの性質に依らず個人差に対して一貫して高い性能を示すことを明らかにした。第三に、上記で最適化された戦略基盤に、仮想的なソース被験者を生成して擬似的にソース数を拡張する Virtual Source Assistance (VSA) を導入し、少数ソース環境の本質的制約を克服し得ることを示した。以上より本研究は、生体信号型 HCI における機械学習・転移学習手法の設計・運用指針を体系化すると同時に、VSA によりデータ収集のボトルネックを解消する実用的なフレームワークを確立した。本成果は、生体信号型 HCI のユーザー負担と開発コストを大幅に低減し、生体信号型 HCI 技術の社会実装を加速させるものである。

キーワード

ヒューマンコンピュータインタフェース, 生体信号識別, 機械学習, 転移学習, アンサンブル学習

慶應義塾大学大学院 政策・メディア研究科
星野 貴行

Development of a Machine Learning Framework for Biological Signal Classification

Summary

Biosignal-based human-computer interfaces (HCI) are expected to serve as next-generation interaction modalities that remain usable even when physical contact, vision, or speech is constrained, with machine-learning-based recognition playing a central role in their functionality. However, biosignals are costly and burdensome to acquire, making large-scale data collection difficult, and they exhibit substantial inter-subject variability, causing conventional machine learning models to generalize poorly to new users. Transfer learning has been regarded as a promising approach for mitigating individual differences by leveraging data from previously recorded source subjects, yet its effectiveness is fundamentally limited to scenarios in which a sufficiently large number of source subjects can be secured. Thus, the essential challenge of data scarcity remains unresolved. This study aims to construct a machine learning framework capable of achieving high-accuracy recognition while adapting to individual differences, even when only a small number of source subjects are available. First, we systematically compared combinations of machine learning models and transfer learning methods from the perspective of practical applicability and demonstrated that deep neural networks combined with parameter-space transfer provide the most accurate and data-efficient option. Second, we examined strategies for utilizing source subjects and found that the multisource approach—training individual models for each source subject and integrating them via ensemble methods—consistently outperformed alternatives across datasets with varying characteristics. Third, we introduced Virtual Source Assistance (VSA), a method for generating virtual source subjects to artificially expand the source pool, thereby overcoming the fundamental limitations of small-source scenarios. Overall, this study systematizes design and operational guidelines for machine learning and transfer learning in biosignal-based HCI and establishes a practical framework that alleviates the data collection bottleneck through VSA. The proposed approach substantially reduces both user burden and development cost and is expected to accelerate the real-world deployment of biosignal-based HCI technologies.

Keywords

Human-Computer Interface, Biosignal Classification, Machine Learning, Transfer Learning, Ensemble Learning

Graduate School of Media and Governance, Keio University
Takayuki Hoshino

目 次

第 1 章	序論	1
1.1	研究背景	1
1.1.1	Human-Computer Interface (HCI) の進化と課題	1
1.1.2	生体信号型 HCI の可能性	2
1.1.3	生体信号型 HCI の基本構成	3
1.1.4	生体信号型 HCI における機械学習の役割	4
1.2	従来研究の課題	5
1.2.1	生体信号データ収集の困難性	5
1.2.2	生体信号における個人差の影響	6
1.2.3	転移学習による解決アプローチとその限界	7
1.3	研究目的	8
1.3.1	研究目標 1：実用的な生体信号識別に向けた機械学習と転移学習の比較検討	8
1.3.2	研究目標 2：複数被験者データの統合戦略に関する比較検討	9
1.3.3	研究目標 3：少数ソース被験者問題に対する仮想ソース被験者生成	9
1.3.4	本研究の学術的・社会的意義	9
1.4	論文構成	10
第 2 章	生体信号と機械学習の基礎	11
2.1	緒言	11
2.2	生体信号の発生原理と個人差	11
2.2.1	表面筋電位 (sEMG)	11
2.2.2	脳波 (EEG)	13
2.3	機械学習	15
2.3.1	機械学習の基本枠組み	16
2.3.2	生体信号識別における機械学習	17
2.4	転移学習	19
2.4.1	転移学習の定義と問題設定	20
2.4.2	転移学習の分類	21
2.4.3	生体信号識別における転移学習	22
2.5	本研究における問題設定	25
2.6	結言	26

第 3 章	実用的な生体信号識別に向けた機械学習と転移学習の比較検討	27
3.1	緒言	27
3.2	データセットおよび解析手法	29
3.2.1	データセット	29
3.2.2	前処理	31
3.2.3	識別	32
3.2.4	転移学習	33
3.2.5	アンサンブル処理	35
3.3	評価方法	35
3.3.1	SVM/ANN	37
3.3.2	STM-SVM/STM-ANN	39
3.3.3	FT-ANN	39
3.3.4	個別 SVM/ANN (i-SVM/i-ANN)	40
3.4	結果	40
3.4.1	全ソース被験者を用いた場合の識別精度	40
3.4.2	特定のソース被験者数における識別精度	42
3.4.3	転移学習手法と識別器の追加組み合わせの比較	45
3.5	考察	48
3.5.1	識別性能	48
3.5.2	ソース識別器の数	50
3.5.3	データ空間ベースとパラメータ空間ベースの比較	50
3.5.4	研究の限界	51
3.5.5	今後の課題	51
3.6	結言	52
第 4 章	複数被験者データの統合戦略に関する比較検討	53
4.1	緒言	53
4.2	データセット, 解析手法, および評価方法	55
4.2.1	データセット	55
4.2.2	区間分割と特徴抽出	59
4.2.3	バックボーンモデル	60
4.2.4	ファインチューニング (FT)	62
4.2.5	アンサンブル戦略	63
4.2.6	評価方法	64

4.3	結果	65
4.3.1	全体の認識精度	65
4.3.2	特徴量分布の可視化	67
4.3.3	アンサンブル戦略への trgANN の追加	68
4.4	考察	68
4.4.1	被験者内識別モデル (trgANN) と従来研究で使用されたモデルとの比較	68
4.4.2	シングルソースおよびマルチソースアプローチの評価	71
4.4.3	アンサンブル戦略の分析	72
4.4.4	sEMG ベース HCI への影響および適用可能性	73
4.4.5	本章の限界と今後の課題	74
4.5	結言	75
第 5 章	少数ソース被験者問題に対する仮想ソース被験者生成	76
5.1	緒言	76
5.2	仮想ソース支援 (VSA) の概念	77
5.3	データ準備	79
5.3.1	データセットおよび前処理	79
5.3.2	データの分割	81
5.4	手法	81
5.4.1	識別フレームワークの概要	81
5.4.2	識別器	82
5.4.3	仮想ソースの生成	83
5.4.4	転移学習およびアンサンブル戦略	85
5.5	評価	85
5.6	結果	86
5.6.1	被験者内識別精度	86
5.6.2	VSA の導入効果	87
5.6.3	被験者内モデルの統合効果	90
5.7	考察	91
5.7.1	識別器のアーキテクチャと被験者内識別精度	91
5.7.2	VSA の有効性と特徴	92
5.7.3	仮想ソース生成に関する考察	93
5.7.4	VSA の応用可能性とインパクト	94

5.7.5	限界	94
5.8	結言	94
5.9	章末付録: 異なる生体信号モダリティ (MEG) への拡張性評価	95
第 6 章	総合考察	97
6.1	提案フレームワークの意義と位置づけ	97
6.2	開発フェーズとデータ規模に応じた適用戦略	98
6.3	実世界応用への課題と展望	99
第 7 章	結論	102
	謝辞	105
	参考文献	107
付 録 A	機械学習の基礎	127
A.1	緒言	127
A.2	機械学習の枠組み	127
A.2.1	教師あり学習 (Supervised Learning)	127
A.2.2	教師なし学習 (Unsupervised Learning)	128
A.3	モデル選択と汎化性	128
A.3.1	線形性と非線形性	128
A.3.2	汎化性能と過学習	130
A.3.3	正則化	131
A.4	本研究で採用した機械学習モデルとその構成要素	133
A.4.1	サポートベクトルマシン	133
A.4.2	深層学習	135
付 録 B	カーネル法に基づくスタイル転移マッピング	139
B.1	緒言	139
B.2	データセットと前処理	142
B.2.1	データセット	142
B.2.2	MyoDataset	142
B.2.3	NinaPro DB5	142
B.2.4	OpenBMI	143
B.2.5	BCI Competition IV-2a (BCICIV2a)	143

B.2.6	データの分割	144
B.3	手法	144
B.3.1	バックボーン識別器	144
B.3.2	Style Transfer Mapping	146
B.3.3	Kernel Style Transfer Mapping	147
B.3.4	アンサンブル学習	148
B.4	評価	149
B.5	結果	149
B.5.1	追加分析	150
B.6	考察	153
B.7	結言	155
	研究業績	156

目 次

1.1	従来の HCI と生体信号型 HCI の一例. (※本図の一部は生成 AI モデル Google Gemini により作成)	2
1.2	生体信号型 HCI の処理の流れの概要. (※本図の一部は生成 AI モデル Google Gemini により作成)	4
1.3	被験者間変動によるデータ分布の不一致が引き起こす誤分類の模式図. (※本図の一部は生成 AI モデル Google Gemini により作成)	5
2.1	骨格筋における sEMG 発生機序の概念図. 中枢からの運動指令が脊髄を経由して α 運動ニューロンに伝達され, 神経筋接合部で筋線維の活動電位を誘発する過程を示す. 文献 [1, 2, 3] を参考に作成. (※本図の一部は生成 AI モデル Google Gemini により作成)	12
2.2	筋線維上を伝播する活動電位と, その容積伝導を介して皮膚表面で観測される sEMG 信号の関係. 多数の MUAP が空間的・時間的に加算されることで, 最終的な sEMG 波形が形成される. 文献 [1, 2, 3] を参考に作成. (※本図の一部は生成 AI モデル Google Gemini により作成)	12
2.3	興奮性シナプス後電位 (EPSP) の発生機序と活動電位 (AP) の時間的特性の比較. AP の持続時間は約 1 ms と短いのに対し, EPSP は約 10 ms かけて緩やかに変化するため, 時間的な加算が生じやすく, 頭皮上脳波の主要な信号源となる. 文献 [4, 5, 6] を参考に作成. (※本図の一部は生成 AI モデル Google Gemini により作成)	13
2.4	大脳皮質の褶曲構造 (脳回・脳溝) と電流双極子の配向性が EEG 計測に与える影響. 錐体細胞は皮質表面に対して垂直に配列するため, 脳回では双極子が頭皮電極に向かう形で形成され観測されやすいが, 脳溝では双極子が頭皮に対して水平, あるいは深部に位置するため観測されづらい. 文献 [4, 5, 6, 7] を参考に作成. (※本図の一部は生成 AI モデル Google Gemini により作成)	14
2.5	MNIST データセットで学習したモデルにおける, 同一ドメイン (MNIST) および異ドメイン (MNIST-M) に対する認識精度の比較. 文献 [8] を参考に作成. (※本図の一部は生成 AI モデル Google Gemini により作成)	17
3.1	本章で使用する転移学習の概念図. 各図はそれぞれ, (a) 転移前 (b) データ空間ベース転移学習 (c) パラメータ空間ベース転移学習を表す.	28
3.2	MyoDataset で使用されている計測機器 MyoArmband の装着の様子. 文献 [9] から引用.	30
3.3	8 種類の単一自由度動作. 文献 [10] より転載.	31

3.4	ターゲット被験者が被験者7のときのSVMとANNにおける(A)学習フェーズと(B)推論フェーズのフローチャート. CLFはClassifierの略でSVMもしくはANNを指す. また, Weightは学習フェーズで得られたターゲット被験者のキャリブレーションデータに対するソース識別器の識別精度(Accuracy)を指し, 0から1の値である. 文献[10]より転載.	37
3.5	ターゲット被験者が被験者7のときのSTM-SVMとSTM-ANNにおける(A)学習フェーズと(B)推論フェーズのフローチャート. CLFはSVMもしくはANNを指す. また, Weightは図3.4(A)の学習フェーズで得られたターゲット被験者のキャリブレーションデータに対するソース識別器の識別精度を指し, 0から1の値である. 文献[10]より転載.	38
3.6	ターゲット被験者が被験者7のときのFT-ANNにおける(A)学習フェーズと(B)推論フェーズのフローチャート. また, Weightは図3.4(A)の学習フェーズで得られたターゲット被験者のキャリブレーションデータに対するソース識別器の識別精度を指し, 0から1の値である. 文献[10]より転載.	39
3.7	ターゲット被験者が被験者7のときのi-SVMとi-ANNにおける学習フェーズと推論フェーズのフローチャート. CLFはi-SVMもしくはi-ANNを指す. 文献[10]より転載.	40
3.8	転移学習手法を適用した5条件および参照条件(i-SVMおよびi-ANN)の25名のターゲット被験者における平均識別精度. 文献[10]より転載.	41
3.9	ソース被験者数を7名とした場合の各手法の平均識別精度. 文献[10]より転載.	43
3.10	8クラス動作識別における混同行列. 縦軸が真のラベル, 横軸が予測ラベルを意味している. 上段3つがSTM-ANN, 下段3つがFT-ANNの結果. 左が25名分の総合的な結果, 中央が最良の被験者(被験者5)の結果, 右が最悪の被験者(被験者22)の結果. 文献[10]より転載.	44
3.11	被験者25名分の各クラスにおける試行ごと(3, 4, 5番目の試行)のSTM-ANNおよびFT-ANNの平均識別精度. 青いバーがSTM-ANN, 赤いバーがFT-ANNを意味する. 文献[10]より転載.	45
3.12	3回のテスト試行における真のラベル系列と, STM-ANN(上)およびFT-ANN(下)によって予測されたラベル系列. 最良(左)および最悪(右)の結果は, 図3.10と同じ被験者から得られた. 文献[10]より転載.	46
3.13	追加の転移学習と識別器を含めた平均識別精度. 文献[10]より転載.	47
4.1	シングルソースアプローチとマルチソースアプローチの概念図. (a) シングルソースアプローチでは, 複数のソース被験者のデータを事前に統合し, それを学習したニューラルネットワーク(NN)をターゲット被験者に対して微調整(FT)する. (b) マルチソースアプローチでは, ソース被験者ごとNNを学習し, 個別にFTし, それらをアンサンブル学習により統合する.	54

4.2	NinaPro DB2 と DB5 の実行課題一覧. データ配布サイト [11] より転載.	57
4.3	NinaPro DB で使用されている計測機器の装着の様子. (b) NinaPro DB5 (d) NinaPro DB2. 文献 [12] から引用.	58
4.4	ターゲットユーザ特化型モデル (trgANN) の学習および評価手順. 角が丸いブロック”Cal.”, ”Val.”, および”Test”は, それぞれキャリブレーション (調整), 検証, およびテスト用のデータを意味する. 文献 [13] より転載.	60
4.5	シングルソースモデル (genANN および FT-genANN) の学習および評価手順 ((A) genANN の事前学習手順, (B) genANN の評価手順, および (C) FT-genANN の評価手順). 文献 [13] より転載.	61
4.6	マルチソースモデルの学習および評価手順. (A) srcANN の事前学習手順. ターゲット被験者に対して, ソース被験者の数だけ srcANN が準備される. (B) HV, SV, WV, および VW の評価手順の一部. ”Prob.”と記されたブロックは, 動作クラスの数と同じ次元を持つ出力確率を表す. (C) FT-HV, FT-SV, FT-WV, FT-SVW の評価手順の一部. (D) 複数の確率を統合するアンサンブル手順と評価手順. ここで, N_{src} はソース被験者の数である. 文献 [13] より転載.	62
4.7	3 つのデータセットに対する 11 種類のモデルの精度を示す箱ひげ図. 破線の灰色の線は, trgANN の平均精度を表す. 文献 [13] より転載.	66
4.8	MyoDataset の特徴量を t-SNE により 2 次元に次元圧縮した結果. いずれの図でも同一のデータを表示しているが, 特定のクラスのみを赤で強調し, その他のクラスは灰色で示している. 文献 [13] より転載.	68
4.9	図 4.8 から抽出した (A) 安静クラス特徴および (B) 前腕回内クラス特徴における個人差. 各図では, 特定の被験者の特徴を赤で強調し, その他の被験者の特徴は灰色で表示している. 文献 [13] より転載.	69
4.10	3 種類のデータセットにおける trgANN を用いた場合および用いない場合の識別精度を示す. 青の実線は trgANN を用いない場合の精度を, オレンジの実線は trgANN を用いた場合の精度を, 灰色の破線は trgANN の平均精度をそれぞれ表している. 文献 [13] より転載.	71
5.1	生体信号識別における転移学習性能を改善するための仮想ソース支援 (Virtual Source Assistance: VSA) の概念図. (a) 実際のソース被験者データを改変して生成された仮想ソース被験者. (b) 従来手法と提案手法の比較. 従来の枠組みでは実ソースデータのみを使用するのに対し, VSA では実ソースと仮想ソースの両方のデータを用いてニューラルネットワーク (NN) を事前学習し, ターゲット被験者のキャリブレーションデータで FT を行い, さらにアンサンブル戦略により統合を行う. 文献 [14] より転載.	78

5.2	ニューラルネットワーク (NN) の構成. (a) 多層パーセプトロン (MLP), (b) 1 次元畳み込みニューラルネットワーク (1D-CNN), (c) 2 次元畳み込みニューラルネットワーク (2D-CNN). 「*」の記号は, FT の過程で更新されたパラメータを示す. 文献 [14] より転載.	82
5.3	(a) MyoDataset および (b) OpenBMI Dataset における識別精度を示す. 識別精度は, 識別に使用された実ソース被験者の数に応じて示されている. 各マーカーは, 全ターゲット被験者に対して 20 回の試行を平均した結果を表す. 文献 [14] より転載.	88
5.4	(a) MyoDataset および (b) OpenBMI Dataset における識別精度. 識別精度は, 実ソースおよび仮想ソースを組み合わせたソース被験者数に応じて示されている. 各マーカーは, 全ターゲット被験者に対する 20 回の試行の平均を表す. 星印は, 実ソース被験者が 10 名の場合の結果を示している. 文献 [14] より転載.	90
5.5	(a) MyoDataset および (b) OpenBMI Dataset における分類精度を示す. 記号「+」は, アンサンブル戦略に被験者内 NN を含めていることを表す. 文献 [14] より転載.	91
5.6	MEG-BCI データセットにおける分類精度. (a) アンサンブル戦略に被験者内 NN を用いない場合, (b) 用いた場合. 文献 [14] より転載.	96
A.1	k-means 法によるクラスタリングの概念図. (左) クラスタリング前のデータ分布, (中央) 3 つのクラスタへの分割例 ($k = 3$), (右) 4 つのクラスタへの分割例 ($k = 4$). 図中の $\times$ 印は, 算出された各クラスタの重心 (セントロイド) を示す.	128
A.2	線形モデルと非線形モデルの挙動の比較. 上段は識別問題における決定境界を示しており, 左上が線形カーネルによる線形識別, 右上が RBF カーネルによる非線形識別. 下段は回帰問題における回帰線を示しており, 左下が単回帰による線形回帰, 右下が多項式回帰による非線形回帰.	129
A.3	多項式回帰の次数を徐々に増やした場合の例.	130
A.4	多項式回帰に対する正則化の働き.	132
A.5	学習データ数の増加による過学習の抑制効果.	133
A.6	線形分離可能なデータに対する識別境界の比較. (左) 2 クラスを分離する超平面は複数存在し得る例. (右) SVM はサポートベクター (白抜き円) により決定される決定境界 (実線) を, マージン (破線で示す両側の境界間隔) が最大となるように選択.	134
A.7	カーネル法による高次元写像と非線形分離の概念図.	135
A.8	MLP の模式図. 本図の作成には NN-SVG [15] を使用した.	136

A.9	CNN の模式図. s はストライドの大きさ. 本図の作成には NN-SVG [15] を使用した.	137
B.1	STM の概要. ソースユーザのデータを o , ターゲットユーザのデータを $\times$ で示し, 背景色はソースユーザのデータに基づいて学習された決定領域を表している. 被験者間変動の影響により, ターゲットユーザのデータには分布のシフトや不整合が生じており, ソースデータで学習された識別器では正確な識別が困難な状態にある. STM は, 矢印で示される方向へターゲットデータを写像することでこの問題に対処する. これにより, ターゲットデータをソースデータの分布へと適合させ, ソース識別器を用いた高精度な識別を可能にする. 文献 [16] より転載.	140
B.2	(a) 線形および (b) 非線形の関係下における STM (スタイル変換マッピング) の性能を示す例. 左の列は, ソースデータおよびターゲットデータの元の分布を示している. 中央の列は従来の STM を適用した場合を描写しており, 右の列は RBF (動径基底関数) ベースの k -STM を示している. 円マーカー (o) は元のターゲットサンプルを表しており, これらは星印 ($\star$) で示されるソースドメインのプロトタイプに向かってマッピングされる. バツ印 ($\times$) は, マッピング後のターゲットデータの位置を示している. 従来の STM は非線形なマッピングを効果的に捉えることができないのに対し, RBF ベースの k -STM は線形および非線形関係の両方を適切に処理している. 文献 [16] より転載.	141
B.3	4 つのデータセットにおける STM および k -STM の識別精度 (ソース被験者数はデータセットに応じ最大 10 名まで変化). 各行はデータセット, 各列はバックボーン識別器に対応する. なお, 第 5 列は全識別器における平均精度を表す. 文献 [16] より転載.	150
B.4	各データセットにおける検証精度に対するプロトタイプ数 (k -means クラスタ数) の影響. 各パネルは, 4 つのバックボーン識別器における平均結果を示す. 文献 [16] より転載.	153

表 目 次

4.1	先行研究と本章の比較概要. アンサンブル戦略の列の略字 (HV, SV, WV, および SWV) の詳細は 4.2.5 項で説明される. 文献 [13] より転載して一部改変.	54
4.2	3 種類のデータセットにおける 11 種類のモデルの識別精度 (平均値 $\pm$ 標準偏差%). 太字の数値は, 各データセットにおいて最も高い精度を示したモデルの結果である. 文献 [13] より転載して一部改変.	65
4.3	3 種類のデータセットにおける, trgANN を用いたマルチソース手法の識別精度 (平均値 $\pm$ 標準偏差 %) および, trgANN の有無による精度差を示す. 太字の数値は, 各データセットで最も高い精度を示したモデルの結果である. 「+」の記号は, そのモデルのアンサンブル戦略に trgANN が含まれていることを示す. 文献 [13] より転載.	70
5.1	生成された仮想被験者数. 文献 [14] より転載.	84
5.2	被験者内識別精度. 文献 [14] より転載.	87
5.3	実ソース被験者数を変化させた場合の MyoDataset における精度. 文献 [14] より転載.	89
5.4	実ソース被験者数を変化させた場合の OpenBMI Dataset における精度. 文献 [14] より転載.	89
B.1	機械学習モデル (LR, SVM, RF, MLP) のハイパーパラメータの探索範囲. 文献 [16] より転載.	145
B.2	転移学習 (STM, k-STM) のハイパーパラメータの探索範囲. 文献 [16] より転載.	145
B.3	各データセットにおける三元配置反復測定分散分析の結果. 要因は, <i>transfer</i> (転移学習手法: STM, RBF k-STM, 線形 k-STM, ハイブリッド k-STM), <i>n_src</i> (ソース被験者数), および <i>model</i> (バックボーン識別器: LR, SVM, RF, MLP) である. 交互作用項 (例: <i>transfer:n_src</i>) は, 要因間の複合的な効果を表す. 文献 [16] より転載.	151
B.4	ボンフェローニ補正を伴う対応有りの <i>t</i> 検定の結果. 文献 [16] より転載.	152

第1章 序論

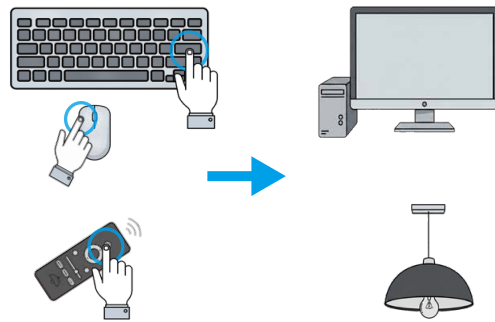
1.1 研究背景

1.1.1 Human-Computer Interface (HCI) の進化と課題

現代社会は、スマートフォンから産業用ロボットに至るまで、人間と機械が効果的に協調するための技術、すなわちヒューマン・コンピュータ・インタフェース (Human-Computer Interface: HCI) によって支えられている。HCI の発展を振り返ると、その歴史は一貫して「より直感的で自然なインタラクション」を志向してきたことが明らかである [17]。コンピュータが登場した 1940 年代の黎明期においては、巨大な配線盤やスイッチを操作することで回路を物理的に変更し、プログラムを実行する方式が主流であった。1960 年代に入ると、依然として紙テープやパンチカードといった原始的な入出力装置が用いられていたが、同時にライトペンによるタッチ操作などの萌芽も見られるようになった [18]。1970 年代には、キーボードによるコマンドライン・インタフェースが普及し、文字ベースで計算機と直接対話することが可能となった [19]。しかし、その利用には高度な専門知識を必要としたため、一般利用者にとっては依然として利用のハードルが高かった。転換点となったのは 1980 年代である。マウス操作とグラフィカル・ユーザ・インタフェース (Graphical User Interface: GUI) が確立され、視覚的な操作を可能にし、コンピュータの大衆化が大きく加速した [20]。さらに 2000 年代後半には、低コストなマルチタッチ方式の提案 [21] を契機に、マルチタッチを備えた端末が普及し、ユーザは「直接触れる」という人間本来の身体的操作を活用できるようになった。この革新はスマートフォンという新たなコンピューティング・パラダイムを生み出し、情報技術の利用形態を大きく変化させた [22]。このように、HCI の進化は人間と機械の間に存在する認知的・物理的距離を着実に縮め、技術の専門家のみならず広範な人々がその恩恵を享受できる社会を実現してきたと言える。

しかしながら、こうした革新的技術が登場している一方で、キーボードやタッチスクリーンといった物理的接触を前提とする従来型の HCI には、依然として本質的な制約が存在する。たとえば、四肢に障害を抱える人々にとって、キーボードの入力、マウスの精緻な操作、あるいはタッチスクリーン上での正確なタップといった行為は大きな困難を伴う。とりわけ筋萎縮性側索硬化症 (Amyotrophic Lateral Sclerosis: ALS) のような進行性神経疾患では、症状の進行に伴って動かせる筋肉が徐々に減少していくため、利用可能なインタフェースが次々と制限されてしまう。また、健常者であっても、育児や買い物で両手が塞がっている状況、料理中で手が汚れている状況、さらには運転中など、身体的接触を前提とする従来型の HCI が適さない状況は日常的に存在する。こうした課題に対して、視線や音声といった非

Conventional HCI



Biological signal-based HCI

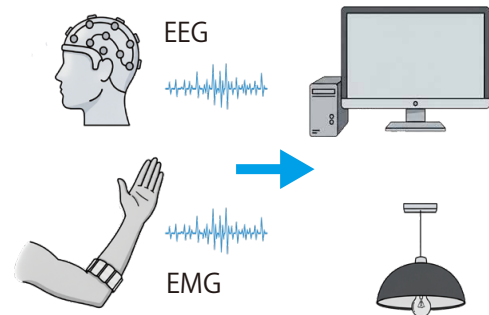


図 1.1: 従来の HCI と生体信号型 HCI の一例.

(※本図の一部は生成 AI モデル Google Gemini により作成)

接触型 HCI [23, 24] が有望視されているが、これらもまた環境的制約を免れない．たとえば、視線操作は運転中のような視線を大きく操作できない場面では危険を伴い得るし、音声操作は静寂が求められる場所やプライバシー保護が必要な空間、あるいは騒音の多い環境では実用性に乏しい．したがって、ユーザの身体的特性や利用環境の多様性を踏まえると、より多様な HCI の構築が、生活の質の向上に資するために必須である．

1.1.2 生体信号型 HCI の可能性

こうした背景のもと、物理的接触や視覚・音声に依存しない新たなインタフェース技術として、生体信号を用いた HCI (生体信号型 HCI) が注目を集めている (図 1.1)．生体信号とは、生体の神経・筋・循環・代謝・呼吸などの生理機構に伴う電氣的または物理・化学的变化を、身体表面あるいは光学的・機械的手法などを介して非侵襲的に計測した信号の総称である．代表的な電氣的信号としては、脳の神経活動に伴う電位変化を頭皮上で計測する脳波 (Electroencephalogram: EEG) [25]、筋線維の活動電位を皮膚表面で捉える筋電図 (Electromyogram: EMG) [26]、心筋の興奮伝導に伴う電位変化を記録する心電図 (Electrocardiogram: ECG) [27]、および眼球の静止電位を利用して眼球運動を推定する眼電位 (Electrooculogram: EOG) [28] が挙げられる．また、電氣的活動に限らず、近赤外光を用いて脳内の局所的な血液酸素化変化を測定する近赤外分光法 (functional Near-infrared Spectroscopy: fNIRS) [29] や、末梢血管の容積変化を光学的に検出する光電式容積脈波 (Photoplethysmogram: PPG) [30] など、光学的・循環系由来の信号も含まれる．これらの生体信号は、多様な生理機構に関する情報を含み、人間の行動意図や内部状態を推定するための有力な情報源であり、これらの信号をセンシングし、目的に応じた解析技術により意味ある情報として表現・変換することで、ユーザの意図や状態に応じた操作・支援が提案されている．

生体信号型 HCI の代表的な入力モダリティとしては、表面筋電位 (surface EMG: sEMG)

および EEG が挙げられる [26, 25]. これらの信号は、心拍数や血圧といった自律神経系の信号と比べ、ユーザの意図を比較的容易に反映できる点で、HCI への応用において優位性を持つ。たとえば、心拍や血圧などの自律神経系信号は随意的に制御することが困難であるため、明確な操作指令を割り当てるインタフェースとしては現実的でない。一方で、特定の筋や脳の活動は意識的に制御することが可能であり、それぞれの活動パターンを操作指令に対応づけることで、ユーザの意図を反映したインタラクションを実現できる。具体的には、EMG や EEG を用いて、コンピュータやスマートフォンの操作 [31, 32, 33]、電動車椅子の制御 [34, 35]、さらには家電製品の遠隔操作 [36] などが提案されている。また、これらの EMG や EEG はユーザの意図を反映させるだけでなく、ユーザの内部状態 (疲労 [37, 38]・ストレス [39, 40]・注意や覚醒度 [41]・情動 [42]) も反映するため、ユーザの状態に応じて動的に応答を変化させるインタラクションへの発展も期待される。

生体信号型 HCI が社会に与える影響は多岐にわたり、その応用可能性は急速に拡大している。生体信号型 HCI は当初、主に医療・リハビリテーション分野での応用が目指されていた。特に、筋萎縮性側索硬化症 (ALS) やロックイン患者のための EEG を用いた意思疎通支援システム [43, 44, 45, 46]、脊髄損傷患者や切断患者の機能回復 [47]、筋電義手 [48] などが代表例である。これに対し、近年ではウェアラブルな計測機器が普及し、非医療分野での応用も急速に進展している。たとえば、EEG では運転者の覚醒度 [49] やオンライン会議における疲労度 [50] の検出が行われており、EMG では手話の認識 [51] や Virtual Reality (VR) ゲーム操作 [52] などが提案されている。さらに近年では、ビッグテック企業が生体信号型 HCI 技術に巨額の投資を行い、産業応用の可能性が急速に拡大している。たとえば、Apple 社はイヤホン型デバイスへの EEG センサ搭載に関する特許を出願しており [53]、Meta 社はリストバンド型の sEMG デバイスを開発し、Augmented Reality (AR) や VR の直感的制御への応用を進めている [54]。このように、生体信号型 HCI は従来の物理的制約を乗り越え、人間の多様な状況とニーズに対応可能な次世代インタフェース技術として、学术界と産業界の両面から注目を集める重要な研究分野となっている。

1.1.3 生体信号型 HCI の基本構成

生体信号型 HCI のシステムは、一般に図 1.2 のように計測、信号処理、機械学習による識別・予測、およびデバイス実行という一連の処理フローで構成される。第一段階の計測において、生体信号の取得手法は電極を体内に留置または埋植する侵襲的手法と、体表面から計測する非侵襲的手法に大別される。侵襲的手法は、信号対雑音比が非常に良好で優れた信号取得が期待できる一方、外科的処置を必要とするためユーザへの身体的負担が大きい。対照的に、非侵襲的手法は安全性および運用上の簡便性に秀でており、HCI への展開可能性は大きく、本研究もこの計測方法に立脚する。また、本研究が主な対象とする EMG や EEG では、活動が想定される部位の皮膚表面上に電極を装着することで、微弱な生体電気信号を非侵襲的に取得する。第二段階の信号処理では、計測信号に対し、フィルタ処理によるノイズ低減や計測課題の特徴を強調する特徴量抽出が行われる。生の計測波形は一般に様々なノイズを含んでおり、加えて、時系列の高次元データであるため、これらの信号処理は、後段

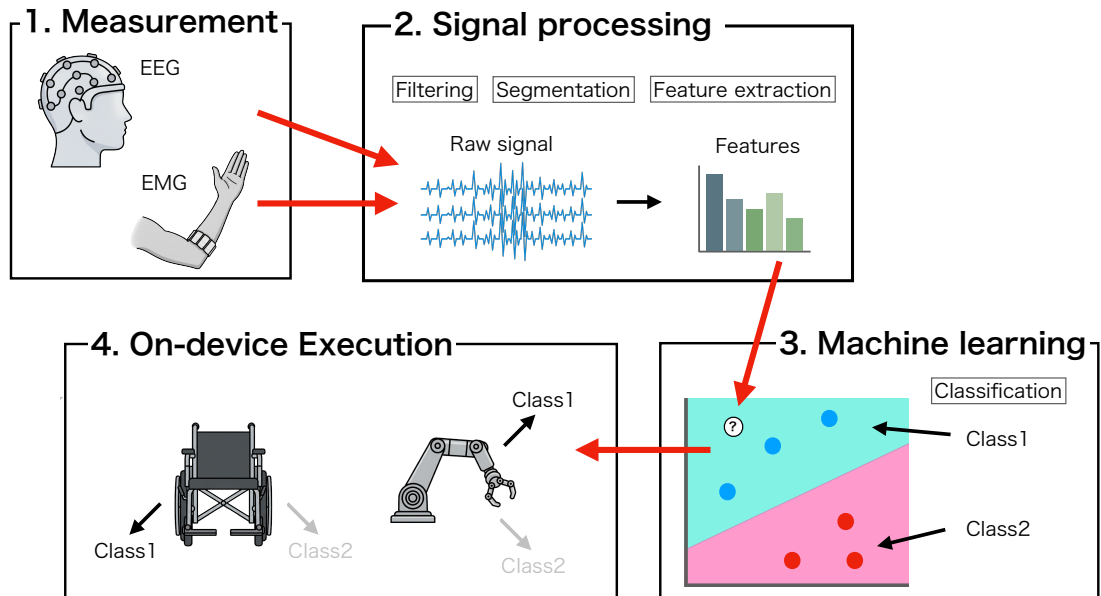


図 1.2: 生体信号型 HCI の処理の流れの概要.

(※本図の一部は生成 AI モデル Google Gemini により作成)

の機械学習モデルの識別精度や計算効率を担保する上で重要な処理である。第三段階の機械学習では、抽出された特徴量を事前に学習された機械学習モデルを用いて、ユーザの意図や状態を推定する。この機械学習による推定結果が、最終段階であるデバイス実行における具体的な操作コマンドと直結しており、推定結果に応じて外部デバイス（コンピュータや義手など）が制御される。したがって、これら一連の処理フローにおいて、機械学習の推定結果がデバイス実行の最終決定を行うため、第三段階の機械学習モデルの識別精度は、HCI システム全体の性能や実用性を直接左右する極めて重要な要素である。

1.1.4 生体信号型 HCI における機械学習の役割

機械学習は、データに潜むパターンや法則性をコンピュータ自らが自動的に抽出し学習する技術の総称である [55]。この技術は、事前に与えられたデータセットのパターンを自動で学習できるように設計されており、人間が明示的なルールを設計することが困難な複雑な現象に対して、データ駆動型で高精度な識別や予測を実現することが可能となる。この特性は、ノイズが多く複雑な時系列データである生体信号を扱う上で重要な特性である。例えば、EMG 信号を用いた制御の分野では、初期の研究において、計測信号の振幅が事前に設定した閾値を超えたかどうかを基準として、単純な ON/OFF の制御を行う方式が提案されていた [56]。このような閾値ベースの手法は実装容易性に優れる一方、開発者が課題ごとに閾値を設計する必要があるため、個人差や経時変動を考慮した基準を見いだすことは多大な労力を要するため、本質的な限界があった。これに対し、機械学習は生体信号に潜むパターンを高速かつ自動で学習できるため、今日では HCI の中核的基盤技術と位置づけられている。

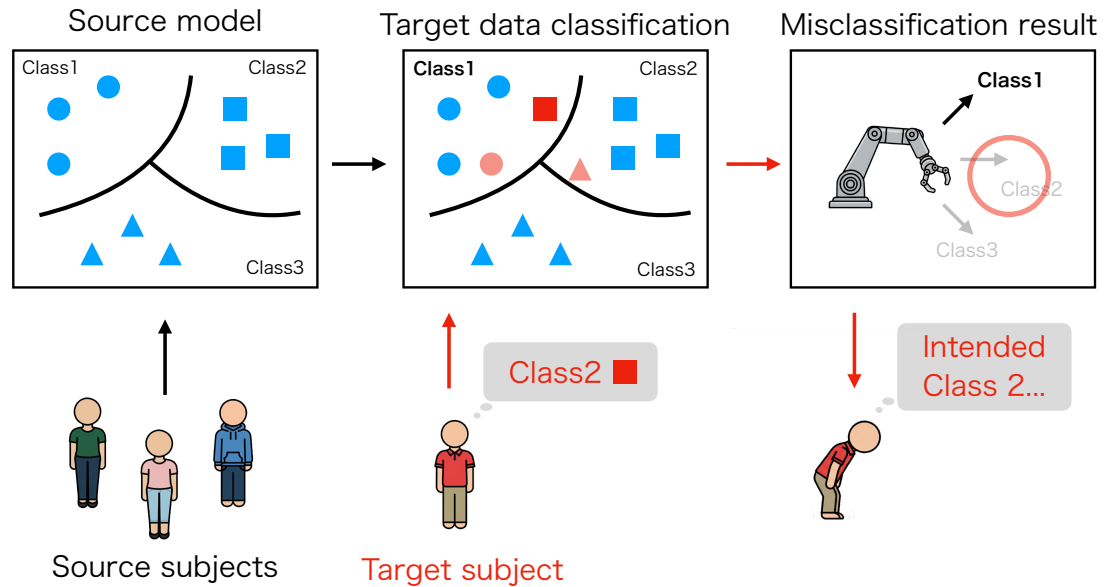


図 1.3: 被験者間変動によるデータ分布の不一致が引き起こす誤分類の模式図。

(※本図の一部は生成 AI モデル Google Gemini により作成)

1.2 従来研究の課題

1.2.1 生体信号データ収集の困難性

図 1.2 で示した通り，機械学習の精度がシステム全体の有用性を決定づけるため，その精度向上は，生体信号 HCI の実用化を達成するための最重要課題である．しかし，その高精度な機械学習モデルの学習に必要な大量のデータ収集は極めて困難であり，これが実用化を阻む最も深刻な障壁の一つとなっている．一般に，機械学習は，学習に利用したデータが多ければ多いほど性能が向上・安定することが知られており，大規模な訓練データの確保が極めて重要である [57, 58, 59]．たとえば，すでに大きな成果を上げている画像認識・音声認識・自然言語処理の分野では，数百万から数十億サンプル数を持つオープンデータセットが利用されており，機械学習モデルはこれらを元に多様なパターンを学習し，高い認識能力を獲得することができる [60, 61, 62]．このような大規模なデータの用意が可能になっている背景には，インターネットやスマートフォンの普及により，日常的に大量の画像・音声・テキストデータが自然に生成・蓄積される環境が整ったことが大きな要因としてある．これに対し，生体信号は大量のデータを手軽にアクセスできる環境が未だに整備されていない．これは大きく分けて，計測が困難である点とデータ共有が困難である点が原因である．前者の計測の困難さについては，一人の被験者から数百回のデータを計測するのに，sEMG では約 1 時間 [63]，EEG では 2-3 時間ほどの時間を要し [64]，大規模なデータ計測は被計測者に多大な身体的・精神的負担を強いる．また，良質なデータを計測するためには，高額な専用の計測機器や計測者の技術的スキルが必要であり，これも大規模な計測に対する障壁となっている．後者のデータ共有に関しては，生体信号は個人の身体的特性や生理的状态を大

きく反映するため、計測データは一般に個人情報とみなされ、プライバシー保護の観点から安易な共有や公開は行えないことに起因している。このように、生体信号分野は機械学習が大きな成果を上げている他分野と比較して、データ収集に大きな制約があり、少数のデータセットで可能な限り高い性能を発揮させる必要がある。

大規模なデータ計測が困難であることは、オープンデータセットに含まれている人数を見ても明らかである。近年、研究の再現性の担保や、開発されたアルゴリズムの性能を公平に比較・評価することを目的として、生体信号分野においてもオープンデータセットの構築と公開が積極的に進められている [65, 66]。しかし、これらの公開データセットに含まれる被験者数に着目すると、その規模は依然として限定的であり、前述した計測コストの高さが如実に反映されていることがわかる。たとえば、Tosin らは、上肢・手の動作認識を主目的とした sEMG のオープンデータセットを調査しており、9 個のデータセットの平均が 17.9 人 (最小 4 人, 最大 45 人) であることを報告している [65]。同様に運動想起課題における EEG においては、Altaheri らによれば、18 個のデータセットの平均が 18.3 人 (最小 1 人, 最大 109 人) と報告されている [66]。したがって、100 人を超えるような大規模な計測は稀である。また、仮に 100 名から 1 人あたり 1000 サンプルを計測できたとしても総数は 10 万サンプルに留まる。これは、先述の画像認識や自然言語処理分野のデータセットとは対照的であり、桁違いの規模の差が存在する。すなわち、生体信号処理分野におけるデータ環境は、オープンサイエンスの潮流の中にあいながらも、依然としてスモールデータの制約下であり、このデータの絶対的な不足が、未知のユーザや環境に頑健な機械学習モデルの構築を困難にする根本的な要因となっている。

1.2.2 生体信号における個人差の影響

生体信号分野における訓練データ不足の問題をさらに深刻化させる要因が、生体信号に本質的に内在する個人差である (図 1.3)。個々の被験者から大規模なデータを収集することが困難であるならば、次善の策として、複数の被験者からデータを収集し、それらを集約・活用することで訓練データを補強するアプローチが考えられる。しかし、生体信号分野においてこの戦略は、期待通りの成果を上げることが極めて難しい。これは、生体信号は被験者間の変動が非常に大きく、たとえ同一の課題であっても、観測される信号パターンは被験者ごとに大きく異なるためである。この個人差の要因は、大きく解剖学的・生理学的な側面と、機能的・認知的な側面に大別できる。前者としては、EMG における筋の厚み、皮下脂肪量、筋線維組成や、EEG における脳の皮質形状、頭蓋骨の厚みといった解剖学的・生理学的な差異が挙げられる。後者としては、運動実行における無意識の癖や、運動想起で用いるイメージ戦略 (運動感覚的か視覚的か) といった、機能的・認知的な差異が挙げられる。このように様々な要因が絡み合っ個人差を生み出しているため、複数の被験者のデータを単純に集約して汎用的なモデルを学習させようとしても、モデルは個々人の精緻なパターンを捉えきれず、被験者集団の平均的な特徴しか学習できない。これは、訓練データに含まれていない新しい被験者に対してモデルを適用する際に、識別精度が著しく低下する被験者間汎化の失敗に繋がり、実際、多くの研究でこの識別精度の低下が報告されている [67, 68, 69, 70]。

Zhu ら [70] は、複数被験者データで学習した機械学習モデルが有効なのは、新規被験者から得られるデータがごく少量の初期段階に限られると報告している。さらに、新規被験者のデータが約 15 分に達すると、その被験者のデータのみで学習した個人内モデルの方が、複数被験者モデルよりも高い精度を示すことを報告している。そのため、生体信号識別では、被験者ごとに限られたデータから個人ごとの機械学習モデルを構築することが、高い精度を得るための一般的な手法となっている [71, 72, 73, 74]。以上のように、生体信号の個人差は、被験者間でのデータ共有や汎用モデルの構築を本質的に困難にしており、少量のデータで高い性能を達成しなければならないという課題を一層際立たせている。

1.2.3 転移学習による解決アプローチとその限界

前述の個人差に関する問題を解決する有望なアプローチとして、転移学習 (Transfer Learning) が注目されている [75, 76]。これは、あるタスク (ソースドメイン) で事前に学習した知識を、類似するが異なるタスク (ターゲットドメイン) に応用する機械学習手法である。この転移学習は、ピアノを弾くことができる人は、初めて楽器に触れる人よりも他の楽器を比較的早く弾くことができることに喩えられる。このアプローチが有効である背景には、ソースドメインとターゲットドメインは一見異なっているとしても、両者に共通する基礎的な構造や特徴量、あるいは知識が存在するという仮定がある。一般に、機械学習は学習されたデータセットの範囲内でしか効果を発揮しないが、転移学習は、ソースドメインで得たこの共通知識を活用することで、その制約を部分的に克服するアプローチである。この特性は、ターゲットドメインで利用可能なデータが極めて少ない場合に特に有力であり、生体信号型 HCI の文脈においても、これは極めて重要な意味を持つ。すなわち、これを多数の事前計測ユーザ (ソース被験者) のデータ (ソースドメイン) から、新規ユーザ (ターゲット被験者) の少数のキャリブレーションデータ (ターゲットドメイン) への転移問題として扱うことで、個人差の問題を克服しつつ、新規ユーザのデータ収集負担を最小限に抑えることが可能になっている [77, 78]。

転移学習は、ターゲットユーザのデータ収集負担を軽減する一方で、依然として実用化を阻む重大な限界を抱えている。それは、既存の転移学習手法は、その有効性を担保するために多数のソース被験者のデータが利用可能であることを暗黙的に前提としている点である。EMG や EEG の分野においては様々な転移学習手法が提案されており、異なる転移学習手法であっても利用したソース被験者数が増えるほどに性能が向上し、精度が安定することが報告されている [79, 80, 9, 81]。しかし、これは裏を返せば、ソース被験者数が限定的である場合には、どの転移学習手法を採用したとしても十分な精度の改善が得られないということを意味する。この多数のソース被験者を必要とするという前提は、計測に多大な労力が必要とされる生体信号分野において極めて深刻な障壁となる。このため、従来の転移学習アプローチは確かにターゲットユーザのデータ収集負担を緩和するものの、依然として多くの人数からの事前計測が必要となっており、データ収集に関する課題の本質的解決には至っていない。したがって、少数ソース被験者であっても、高い性能を引き出す転移学習手法を開発し、データ収集の制約を根本から緩和できる新しいアプローチが必要とされている。

1.3 研究目的

ここまで論じたように、生体信号型 HCI の実用化を阻む課題として、データ収集の困難性、個人差、そして既存の転移学習の限界を指摘した。特に、従来の転移学習は多数のソース被験者を前提としており、少数のソース被験者からでは十分な性能が得られないという、実用化における深刻な課題が残されている。本研究は、この課題を解決することを目的とする。すなわち、ごく少数のソース被験者データからでも高精度な識別を実現する新たな転移学習を基盤とする機械学習フレームワークを構築し、データ収集の制約を緩和することを目指す。

このフレームワークの構築を達成するためには、段階的なアプローチが必要となる。まず、構築するフレームワークの土台として、数ある機械学習と転移学習の手法の中から最適な組合せを特定する必要がある。次に、その基盤の上で、限りあるソース被験者群のデータを最も効率的に扱うための戦略を確立しなければならない。そして最後に、これらによって最適化されたベースラインを用いてもなお残存する本質的な課題、すなわち少数ソース被験者問題を、本研究の核となる提案手法によって根本的に解決を目指す。したがって本研究では、この論理的な順序に基づき、以下の 3 つの段階的な研究目標を設定し、順次解明していく。

1.3.1 研究目標 1 : 実用的な生体信号識別に向けた機械学習と転移学習の比較検討

第一の研究目標は、本研究で構築するフレームワークの根幹をなす機械学習モデルと転移学習手法の最適な組み合わせを選定し、その技術的基盤を確立することである。この基盤の選定が不可欠な理由は、後続する研究目標の妥当性と性能が、この土台となる技術の選択に大きく依存するためである。現在、生体信号型 HCI においては、デバイス操作のコマンドを決定する機械学習モデルが選択され、その上で、個人差の問題を解決するために転移学習が適用されるため、様々な組合せが考えられる。機械学習においては、計算効率に優れ少量データでも機能する伝統的な機械学習から、表現力の高い深層学習まで多様な選択肢が存在する。同様に、転移学習においても、個人差の補正の方式に多様な手法が提案されている。実用化の観点からは、単に識別精度が高い組合せを見つけるだけでなく、ユーザ側の限られた計算資源でも動作する計算効率も同時に考慮する必要がある。しかしながら、これら機械学習と転移手法の多様な組み合わせを体系的に比較・評価した研究は十分ではなかった。そこで本目標では、まずこれらの技術的選択肢を、信号雑音比が優れ結果の解釈が比較的容易である EMG の識別タスクにおいて系統的に比較評価し、最も実用的かつ高精度な組合せの方針を特定する。sEMG を初期段階の評価に用いる理由は、信号対雑音比が優れている点に加え、タスク実行の客観的な確認が容易である点にある。EEG を用いた運動想起タスクでは、被験者が正しく想起を行えているか（正解ラベルの信頼性）を実験者が外部から確認することは困難であるが、sEMG を用いた運動実行タスクであれば、実験者が外部から正しく動作を行えていることを容易に確認でき、正解ラベルの信頼性が担保される。したがって、まずは不確実性の低い sEMG データを用いて提案フレームワークの技術的妥当性を検証する。

1.3.2 研究目標 2：複数被験者データの統合戦略に関する比較検討

第二の研究目標は、第一段階で特定した最適な技術基盤に対して、ソース被験者群のデータを効率的に扱うための適用戦略を確立することである。具体的には、ソース被験者群のデータを扱う際、(1) 全被験者のデータを統合して単一の汎用モデルを学習するシングルソースアプローチと(2) 被験者ごとに個別のモデルを学習し、それらの出力を後から統合するマルチソースアプローチ、この2つの方式が存在する。これらの方式は、対象となるデータセットの質と量に応じて、最適な選択が異なる可能性がある。具体的には、ソース被験者数が多く個人差が比較的小さいデータセットでは前者が、被験者数が少なく個人差が大きいデータセットでは後者が適していると予想される。しかし、これら2つのアプローチはこれまで体系的に比較されておらず、特に本研究の目的である少数のソース被験者という条件下において、どちらがより効果的であるかは明確でない。さらに、後者を採用する場合には、各モデルの出力をどのように統合するかというアンサンブル戦略の選択も性能を左右するが、その最適解も不明確である。そこで本段階では、これらの適用戦略を体系的に比較検証し、本研究の最終目的に取り組むための、最も効率的かつ高精度な適用戦略を特定する。これらの適用戦略は、前段階と同様に sEMG を用いて体系的に比較検証する。

1.3.3 研究目標 3：少数ソース被験者問題に対する仮想ソース被験者生成

研究目標 1 と 2 で最適化された既存戦略を用いたとしても、依然として少数ソース被験者問題は残存する。すなわち、如何に最適化された既存戦略を用いたとしても、転移学習はその有効性を担保するために多数のソース被験者のデータを暗黙的に前提としており、より少数のソース被験者しか利用できない状況では性能が著しく低下してしまう。この実用上の最大の制約を根本的に解決するため、本研究では第三の研究目標として、新手法 Virtual Source Assistance (VSA) を提案する。VSA は、実在する少数のソース被験者のデータを元に、新たな仮想被験者を生成することで転移学習に利用するソース被験者数を擬似的に増やすことを目的とする手法である。また、仮想被験者の生成に当たっては、生成元となった実被験者とは異なる性質を持つ多様な仮想被験者を生成することが精度向上に寄与すると仮説を立て、従来の保守的なデータ生成法に加え、より積極的なデータ生成法との比較を行う。これにより、少数の実ソース被験者からでも高精度な識別を可能とする手法を確立し、少数ソース被験者問題の解決を目指す。本段階における検証は、これまでの段階で用いてきた sEMG データに加え、より難易度の高い EEG データも評価に加える。動作実行の確認が容易な sEMG で手法の基礎的な有効性を確認した上で、内部状態の推定というより不確実性の高い EEG タスクへと適用範囲を拡張し、提案手法の汎用性と堅牢性を実証する。

1.3.4 本研究の学術的・社会的意義

本研究の完遂は、生体信号におけるデータ収集の困難性に端を発し、個人差や既存の転移学習の限界によって深刻化する一連の課題に対し、解決策を提示するものである。学術的

な意義としては、生体信号型 HCI における統合的な機械学習フレームワークを体系的に構築・提案した点にある。このフレームワークは、以下の 3 つの段階的な貢献によって構成される。第一に、これまで不明確であった機械学習モデルと転移学習手法の最適な組み合わせを、実用化の観点から体系的に評価し、その技術的基盤を特定した点。第二に、その基盤の上で、限られたソース被験者を効果的に扱うための適用戦略を比較検証し、その実用的なガイドラインを確立した点。そして第三に、本研究の核となる VSA が、仮想被験者という概念を導入することで、従来手法が前提としていた多数ソースを要する条件を緩和し、実用上の制約を大きく軽減した点である。社会的・工学的な意義としては、提案フレームワークによって少数のソース被験者からでも高精度な HCI を構築可能にすることで、開発初期のプロトタイピングや、被験者の確保が困難な希少疾患患者向けシステムの開発を大幅に加速させることが期待できる。また、本研究が提案するフレームワークは、データ収集に関わる開発者および被験者・ユーザ双方の身体的・時間的コストを根本から低減するものであり、生体信号型 HCI の普及を促進し、多様なユーザがその技術の恩恵を享受できる社会の実現に寄与するものである。

1.4 論文構成

本論文は、上記の研究目的を段階的に達成するため、以下の全 7 章で構成される。第 1 章では、生体信号型 HCI の現状と課題を概観し、データ収集の困難性、個人差、既存の転移学習の限界という課題を指摘した。それらを受け、本研究の 3 つの段階的な研究目標と、本研究が持つ学術的・社会的意義を定義した。第 2 章では、生体信号の発生原理と個人差を生じさせる原因、機械学習と転移学習の基礎について体系的に整理する。第 3 章では、研究目標 1「最適な手法選択の特定」に取り組む。第 2 章で整理した分類に基づき、生体信号型 HCI における機械学習モデルと転移学習手法の組合せを、実用化の側面から体系的に比較評価し、後続する研究の基礎となる最も実用的なベースラインを特定する。第 4 章では、研究目標 2「転移学習の適用戦略の最適化」に取り組む。第 3 章で特定した基盤を用い、ソース被験者群のデータをどう扱うかという戦略およびアンサンブル戦略を体系的に比較検証し、その実用的なガイドラインを確立する。第 5 章では、研究目標 3「少数ソース被験者問題の根本的解決」に取り組む。第 3、4 章で最適化された既存戦略を用いてもなお残存する本質的な課題を解決するため、本研究の核となる提案手法 VSA を提案し、その有効性を実証する。第 6 章では、これら一連の成果を俯瞰した総合考察を行う。提案手法が計測機器や計測モダリティを超えて機能するメカニズムや、個人差を除去すべきノイズではなく活用すべき基底と捉え直す学術的意義、さらには大規模データ時代における本手法の実用的な立ち位置について論じる。第 7 章では、本研究を総括することで包括的知見を明確にし、結論と今後の展望を示す。

第2章 生体信号と機械学習の基礎

2.1 緒言

実用的な生体信号型 HCI の実現において、最も大きな障壁となるのが、ユーザごとに異なる信号特性，すなわち個人差の存在である．従来の HCI 研究において，特定のユーザ環境下では高い操作性が報告されているシステムであっても，ひとたび対象者が変わると性能が著しく低下するという問題は，この個人差に起因するところが大きい．したがって，頑健なインタフェースを構築するためには，まず生体信号がどのようなメカニズムで生成され，なぜ個人差が生じるのかという生理学的な背景を深く理解する必要がある．その上で，変動する信号から普遍的な意図を抽出するための数理的な枠組み，すなわち機械学習および転移学習の理論体系を整理することが不可欠である．

2.2 生体信号の発生原理と個人差

2.2.1 表面筋電位 (sEMG)

sEMG は，骨格筋の収縮に伴って発生する微弱な電気信号を皮膚表面から計測したものである [1, 2, 3]．その発生機序は，中枢神経系からの指令に端を発する．図 2.1 に示すように，脳からの運動指令は脊髄を経由し， α 運動ニューロンへと伝達される．1つの α 運動ニューロンとそれに支配される複数の筋線維群は運動単位 (Motor Unit: MU) と呼ばれ，筋収縮の機能的な最小単位を構成する．神経インパルスが神経筋接合部に到達すると，筋線維膜上で脱分極が生じ，活動電位が発生する．図 2.2(a) に示すように，この電位変化はイオンの流入出を伴いながら筋線維上を伝播していく．この活動電位が筋線維上を伝播する速さは筋線維伝導速度 (Muscle Fiber Conduction Velocity: MFCV) と呼ばれる．個々の筋線維の活動電位が，同一 MU 内で同期して加算されたものが運動単位活動電位 (Motor Unit Action Potential: MUAP) である．最終的に観測される sEMG は，図 2.2(b) に示すように，活動する多数の MU からの MUAP が空間的・時間的に重畳され，さらに筋組織，筋膜，皮下脂肪，皮膚という一連の生体組織を介した容積伝導の影響を受けて，減衰・フィルタリングされた総和信号として記録される．

こうして計測される sEMG 信号には，解剖学的・生理学的な要因から大きな個人差が現れる．特に，大きな影響を受ける要因として，皮下脂肪の厚さや筋線維の構成が挙げられる．皮下脂肪は導電性が低いため電気信号を減衰・フィルタリングする作用があり，皮下脂肪が厚いほどに sEMG 振幅が低下し周波数成分も低くなる傾向が報告されている [82]．ま

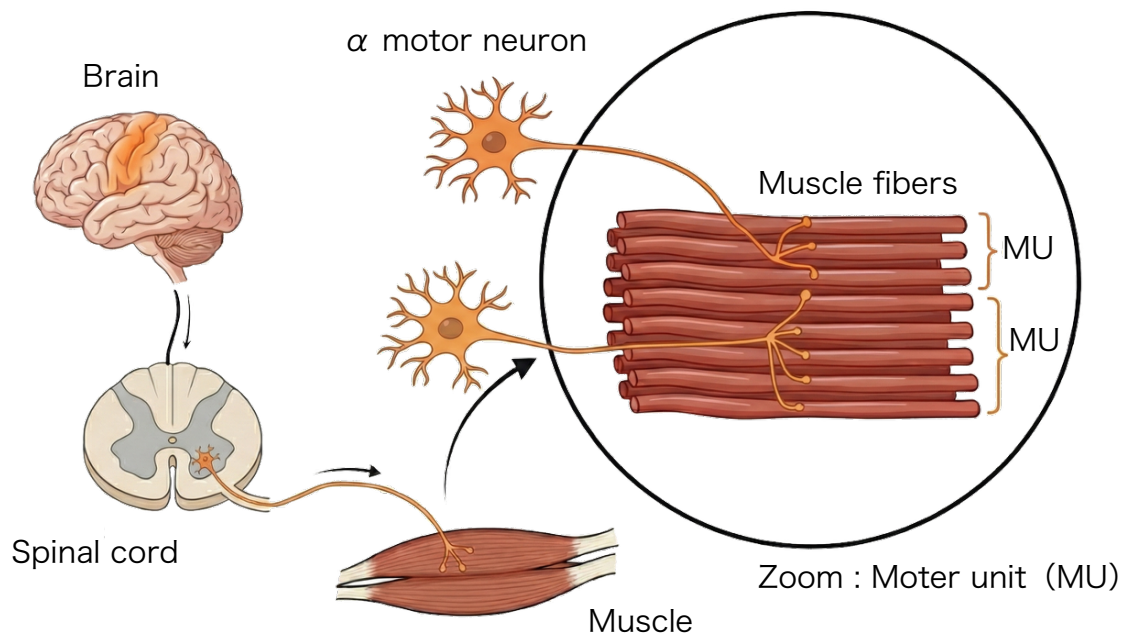


図 2.1: 骨格筋における sEMG 発生機序の概念図. 中枢からの運動指令が脊髄を経由して α 運動ニューロンに伝達され, 神経筋接合部で筋線維の活動電位を誘発する過程を示す. 文献 [1, 2, 3] を参考に作成.

(※本図の一部は生成 AI モデル Google Gemini により作成)

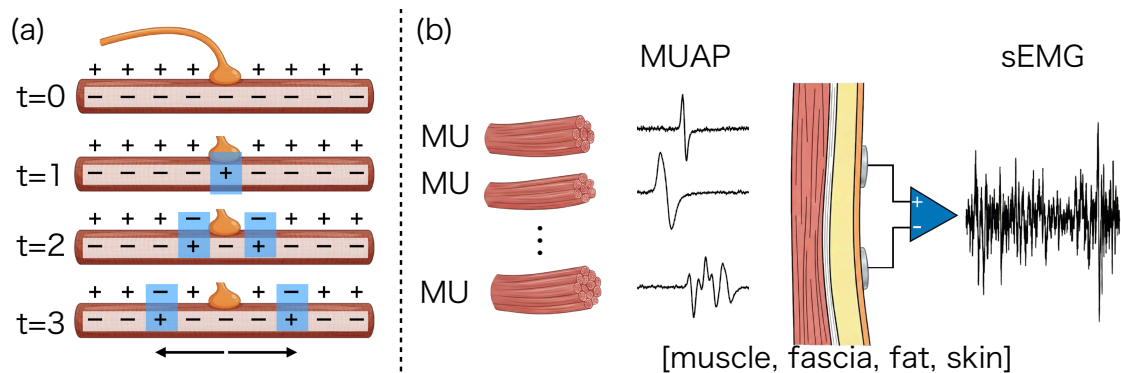


図 2.2: 筋線維上を伝播する活動電位と, その容積伝導を介して皮膚表面で観測される sEMG 信号の関係. 多数の MUAP が空間的・時間的に加算されることで, 最終的な sEMG 波形が形成される. 文献 [1, 2, 3] を参考に作成.

(※本図の一部は生成 AI モデル Google Gemini により作成)

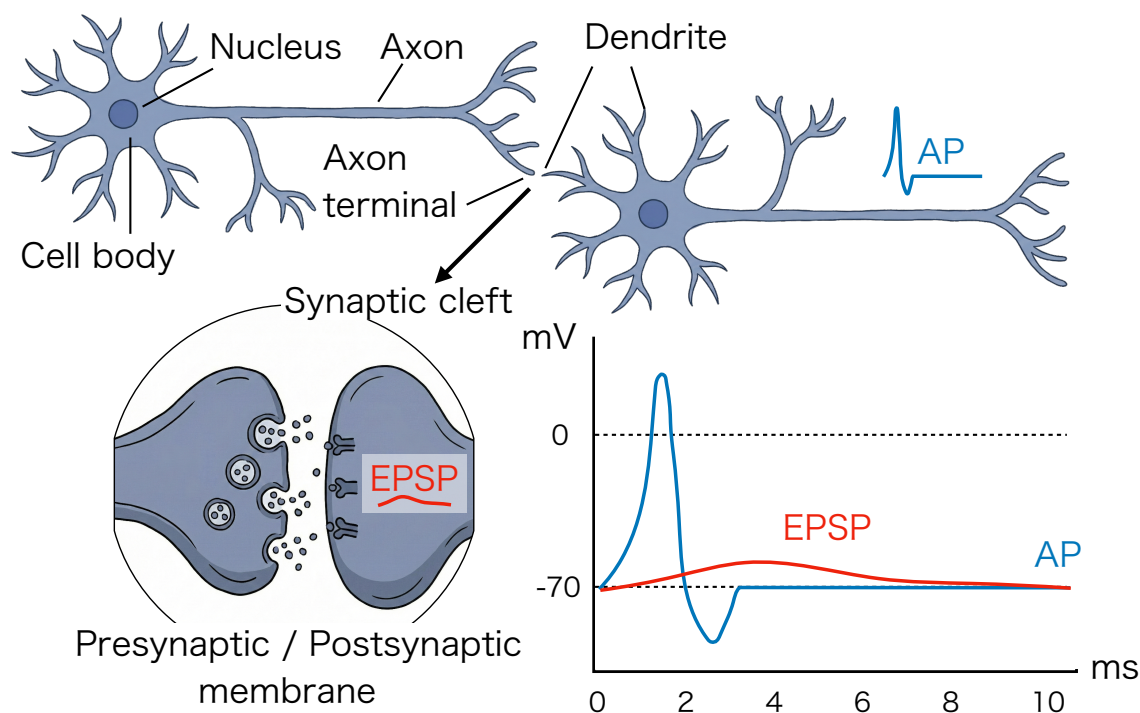


図 2.3: 興奮性シナプス後電位（EPSP）の発生機序と活動電位（AP）の時間的特性の比較。AP の持続時間は約 1 ms と短いに対し、EPSP は約 10 ms かけて緩やかに変化するため、時間的な加算が生じやすく、頭皮上脳波の主要な信号源となる。文献 [4, 5, 6] を参考に作成。

（※本図の一部は生成 AI モデル Google Gemini により作成）

た、皮下脂肪が厚い場合、電位は体表に到達するまでに広がりやすく、隣接した筋の活動が混ざって観測される傾向も報告されている [82]。筋線維の構成も信号に影響し、速筋線維の割合が高い場合は、MFCV が速く、高周波成分を多く含む信号が観測される [83]。この構成比に加え、筋線維径も信号に影響を与え、結果として男女間の差 [84] や加齢による差も生み出している [85]。この他にも、課題実行時の筋の温度 [86] や、疲労度 [87] によっても信号は変化し、これらが複合的に作用することで、個人ごとに性質が異なる信号が形成される。したがって、sEMG 解析においては、これら個人差を正確に捉えた上で解析を行う必要がある。

2.2.2 脳波 (EEG)

EEG は、大脳皮質に存在する多数の神経細胞が同期的に活動することで生じる電気信号を、頭皮上から非侵襲的に計測した信号である [4, 5, 6]。ただし、観測している信号源は sEMG とは大きく異なる。sEMG は筋線維から生じる活動電位を反映するのに対し、EEG では活動電位そのものではなく、主として大脳皮質の錐体細胞群におけるシナプス後電位 (Postsynaptic Potential: PSP) が主要な信号源となる。脳においても神経細胞の活性化に

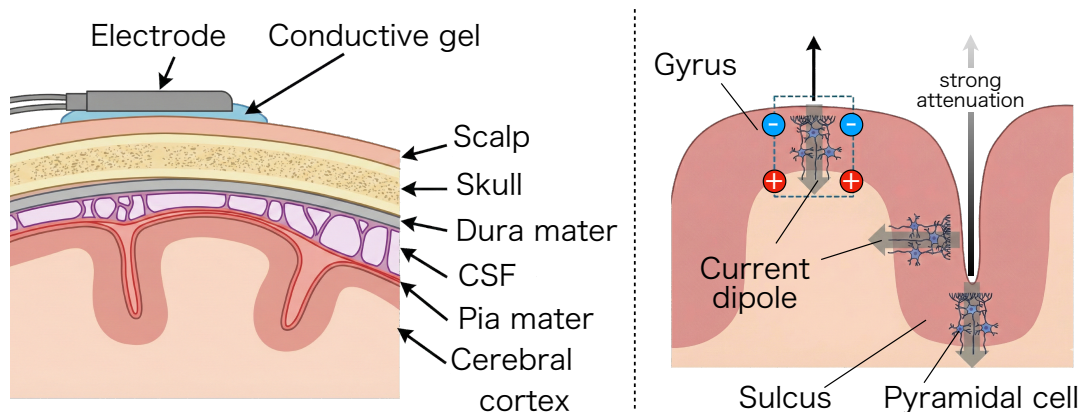


図 2.4: 大脳皮質の褶曲構造（脳回・脳溝）と電流双極子の配向性が EEG 計測に与える影響。錐体細胞は皮質表面に対して垂直に配列するため、脳回では双極子が頭皮電極に向かう形で形成され観測されやすいが、脳溝では双極子が頭皮に対して水平、あるいは深部に位置するため観測されづらい。文献 [4, 5, 6, 7] を参考に作成。

（※本図の一部は生成 AI モデル Google Gemini により作成）

応じて電位が生じるが、筋線維では MU ごとに同期的に活動電位が生じるのに対し、脳神経細胞の活動電位には、MU のような構造的要因による同期性は生じにくい。加えて、活動電位は急峻な二相性（プラスとマイナス）の波形を持つため、互いに相殺されてしまい、頭皮上ではほとんど計測されない。また、筋線維に比べて脳神経細胞は細胞の体積が小さく、発生する電流が微弱である点や、信号源と電極の間に高抵抗な頭蓋骨が介在することで信号が大きく減衰する点も、活動電位の観測を困難にしている要因である。

脳では活動電位が観測されづらいため、頭皮上から観測される電気信号は、より持続時間が長く同期性が高い PSP が支配的である (図 2.3)。PSP とは、情報を伝えるシナプス前細胞から放出された神経伝達物質が、シナプス間隙を介して、情報を受け取るシナプス後細胞の樹状突起にある受容体に結合することで生じる、局所的な膜電位の変化である。この PSP には興奮性と抑制性の双方が存在するが、頭皮上脳波の主成分となるのは興奮性シナプス後電位 (Excitatory Postsynaptic Potential: EPSP) である。活動電位は急速なイオンチャネルの開閉によるものであるため、パルスのような 1 ms 程度の短い持続時間であるのに対して、この EPSP は 10 ms 程度の時間を掛けて徐々に変化するため、時間的な加算が行われやすい。さらに、大脳皮質のシナプス後細胞の多くを占める錐体細胞は、皮質表面に対して垂直に整然と配列しているため、個々の細胞で生じた電位変化が効率的に空間的にも加算されやすく、EPSP は頭皮上においても計測可能な大きさとなるのである。特に、脳には脳回と脳溝という複雑なひだ構造が存在し、錐体細胞はこの皮質表面に対して常に垂直に配列している (図 2.4)。そのため、脳回 (凸部) では電流双極子が頭皮に対して垂直に形成され、EEG の主な信号源となる。一方で、脳溝 (凹部) の沈み込んでいく場所では、電極に直交した向きに電流双極子が形成されることになるため、EEG として観測されにくい。また、脳溝の底には脳回と同様に電極に対して垂直な電流双極子が形成されるが、電極までの距離が長いため、信号は大きく減衰してしまい、やはり EEG として観測されづらい。

sEMG と同様に、頭皮上で観測される EEG 信号にも、解剖学的な要因に起因する個人差が存在する。まず大きな要因として、信号源から電極に至るまでの生体組織の層構造（大脳皮質、脳脊髄液、硬膜、頭蓋骨、そして頭皮）の差が影響を与える。これらの組織の中で、脳脊髄液は生理食塩水に近い組成を持ち導電率が極めて高く電流を側方へ拡散させるのに対し、頭蓋骨は水分含有量が少なく導電率が非常に低いため信号を著しく減衰させる。この拡散させる層と減衰させる層の厚さや形状のバランスは個人ごとに異なるため、信号の強度や空間分布に複雑な個体差をもたらす要因となっている。さらに、信号源である大脳皮質の形状も、無視できない個人差の要因である。大脳皮質は複雑な褶曲構造（脳回と脳溝）を形成しており、その形状は個人ごとの神経回路の接続パターンや張力に基づく発生過程に依存して決定されると考えられている [88]。前述の通り、錐体細胞は皮質表面に垂直に配列するため、この褶曲パターンの違いは、信号源となる電気双極子の向きが個人ごとに異なることを意味する。

解剖学的な要因に加え、EEG には様々な生理学的な要因に起因する個人差も存在し、本研究が対象とする運動想起のタスクにおいては、次のような様々な個人差が影響する。まず運動想起とは、外的な筋活動や関節運動を伴うことなく、特定の運動行為を内的にシミュレーションし、その運動表象を再現する認知過程である [89]。例えば、右手の運動を想起した場合には左半球の一次運動野や運動前野において μ 帯域（8–13 Hz）や β 帯域（13–30 Hz）の事象関連脱同期（event-related desynchronization: ERD）が観測され、同時に同側半球では相対的な同期化（event-related synchronization: ERS）が生じる [90, 91]。ただし、周波数成分や ERD・ERS のパターンには個人差がある [90, 91]。この個人差は、先述の解剖学的な要因に依るところもあるが、その他の要因も多く指摘されている。例えば、自分の動きの感覚を思い出すキネステティックなイメージを得意とする者では、一次運動野や補足運動野の賦活が強く ERD が顕著となる一方、視覚的イメージに依存する者では同程度の脳活動が得られないことが報告されている [92]。また、スポーツ熟達者は運動想起時の ERD が強く出やすいことが報告されており [93, 94]、特に自身が高度に熟達した動作の想起においては、非熟達者と比較して顕著な ERD 増強が報告されている [94]。また、運動想起実行時の注意度が高いと ERD が強くなる一方、注意が散漫である場合には ERD が弱くなることも報告されている [95]。すなわち、EEG における個人差は、解剖学的構造・電気物性による伝達の違いだけでなく、脳がどのように運動を内部表現し処理するかという機能的・心理的要因にも大きく依存しているのである。

2.3 機械学習

生体信号型 HCI では、計測された信号からユーザの意図や状態を推定し、それに応じたデバイス制御を行う。この推定の心臓部となるのが機械学習であり、限られた計測データから高精度かつ頑健な識別器を構築できるかどうか、システム全体の実用性を左右する。本節では、生体信号型 HCI において最も広く採用されている教師あり学習を用いた識別の基本枠組みについて概説し、その後に生体信号における機械学習の活用についても概説する。また、本分野を専門としない読者に向けて、機械学習に関する基礎理論および用語の定義は

付録 A にまとめているため、必要に応じて参照されたい。

2.3.1 機械学習の基本枠組み

機械学習とは、データからパターンや規則性を自動的に抽出し、未知のデータに対しても汎化性能を持つ予測器を構成するための技術である [55]。一般に、入力空間 $\mathcal{X}$ 上の特徴ベクトル $\mathbf{x} \in \mathcal{X}$ と、そのラベル空間 $\mathcal{Y}$ 上の正解ラベル $y \in \mathcal{Y}$ の組 $(\mathbf{x}_n, y_n)$ を $n = 1, \dots, N$ 個集めた訓練データ集合

$$\mathcal{D}_{\text{train}} = \{(\mathbf{x}_n, y_n)\}_{n=1}^N \quad (2.1)$$

が与えられたとき、機械学習の目的は、入力 $\mathbf{x}$ からラベル y を予測する関数 $f_\theta: \mathcal{X} \rightarrow \mathcal{Y}$ を学習することである。ここで θ はモデルのパラメータであり、線形モデルであれば係数ベクトル、ニューラルネットワークであれば各層の重み行列やバイアス項などが含まれる。この際、本質的に重要となるのは、単に手元の $\mathcal{D}_{\text{train}}$ を丸暗記することではなく、学習には使用していない未知の入力 $\mathbf{x}$ に対しても正しい予測 y を出力できること、すなわち汎化性能を獲得することにある。

生体信号の識別においては、特徴ベクトル $\mathbf{x}$ は計測された生体信号から抽出された特徴量を表し、ラベル y はユーザの意図や状態を示すカテゴリ変数であることが多い。ここでは、例としてユーザが「グー」「チョキ」「パー」の3つの手形状を行っている状態を sEMG から識別する多クラス分類問題を考える。この際、学習データの収集は、ユーザに対して所定の動作タスクを行うよう指示を与え、その実行中の生体信号を計測することで行われる。一般的な実験パラダイムとしては、ディスプレイ上に「グー」「チョキ」「パー」のいずれかを示す視覚的指示をランダムに提示し、ユーザはその指示に従って直ちに動作を行う。システムは、各試行において視覚キューが表示されたタイミングと、その指示内容がグーなら $y = 1$ 、チョキなら $y = 2$ 、パーなら $y = 3$ を記録する。分析時には、そのタイミングを基準にして、各試行に対応する信号の一部区間を切り出す。このような試行を多数回繰り返し、入力信号と正解ラベルの組 $(\mathbf{x}_n, y_n)$ を蓄積することで、訓練データ集合 $\mathcal{D}_{\text{train}}$ が構築される。

学習データの構成、すなわちデータセット $\mathcal{D}_{\text{train}}$ に誰のデータを含めるかという戦略は、システムの運用形態を決定づける重要な要素である。生体信号識別の文脈では、これは主に被験者内識別 (within-subject classification) と被験者間識別 (cross-subject classification) の2つに大別される。被験者内識別は、対象となるユーザ自身のデータのみを用いてモデルパラメータ θ を学習させる方式である。この方式では、個人の信号特性に特化できるため比較的高い識別精度が期待できる反面、利用開始前にユーザごとにデータ収集と学習を行う負担が生じる。一方、被験者間識別は、対象ユーザ以外の他者のデータ群を用いて学習を行う方式である。これは、新規ユーザからのデータ収集をスキップする目的で構築されることが多く、複数人分のデータを事前に収集・学習させる。これにより、ある程度個人差を考慮したモデルを作成することができるが、被験者内識別に比べると大きく精度が落ちる問題点がある [67, 68, 69, 70]。したがって、HCI としての実用性を高めるうえでは、使用まで

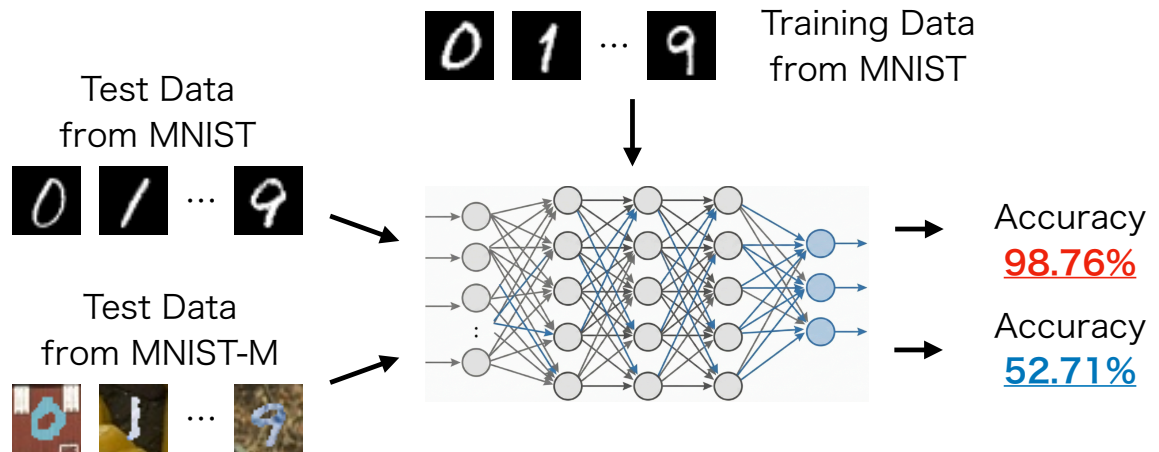


図 2.5: MNIST データセットで学習したモデルにおける、同ドメイン (MNIST) および異ドメイン (MNIST-M) に対する認識精度の比較. 文献 [8] を参考に作成.

(※本図の一部は生成 AI モデル Google Gemini により作成)

の負担を下げることも重要であるが、それ以上に識別精度を高めることが重要であるため、被験者内識別が主流となっている。

訓練データと未知データとの間のデータの性質の差に起因する識別精度の低下は、生体信号に限った話ではなく、機械学習の分野において「ドメインシフト」として知られる普遍的な課題である [96]. 図 2.5 は、Ganin ら [8] の実験を参考に作成したものであり、手書き数字認識におけるドメインシフトの影響を示している. ここでは、まずはじめに MNIST [97] と呼ばれる 0 から 9 までの手書き数字のデータセットを用いてニューラルネットワークを学習した. 続いて、同じ MNIST から、訓練データに用いていないテストデータを用いて評価した場合、この識別精度は 98.76% であった. 一方で、今度は同じ学習済みモデルに対して、先程の MNIST に対し、BSDS500 [98] と呼ばれる自然画像データセットのパッチを背景として合成した MNIST-M を用いて評価したところ、精度は 52.71% にまで低下した. この大幅な識別精度の低下は、いかに機械学習が訓練されたデータセットの影響を強く受けるかを物語っている. 人間であれば、このように背景のスタイルが変化したとしても、そこにある数字を正しく認識することは容易であるが、機械学習モデルにとっては、このような訓練データと未知のデータのドメイン間の分布の乖離は、認識能力を崩壊させる致命的な要因となる. 生体信号における被験者間の個人差もこれと等価な問題の原因であり、適切に個人差を補正しない限り被験者間でのデータの流用は極めて難しいことを意味している.

2.3.2 生体信号識別における機械学習

Parker ら [99] は、2000 年代までの EMG 識別の技術的変遷を概観している. 彼らによれば、1960 年以前の研究では、観測信号の振幅が事前に設定した閾値を超えたかどうかで ON/OFF を切り替えるといった、2-state の制御方式が一般的であった. 1960 年代から 1970 年代にかけては、筋収縮の強度や速度に対して複数の閾値を設けることで、3-state 以上の

識別が可能になったとされる。このような閾値ベースの手法は、2-state のように状態数が少ない場合には、直感的であり精度も高いと報告されている。一方で、状態数を増やすとユーザの操作負荷が増大するという課題も指摘されている。具体的には、3-state では直感的な操作が困難となり、4-state 以上では誤操作率が許容範囲を超えると指摘をしている。これらの限界を克服するため、1970 年後半以降はパターン認識を用いた機械学習手法が多クラス識別における標準的アプローチとして採用され始めた。機械学習を採用することで、今までは 1 次元的な閾値でしか判別ができなかったのに対し、多次元ベクトルを特徴として扱うことができ、高精度での識別が可能になった。このような流れを受け、現在では EMG の識別においては機械学習が標準的な枠組みとして定着している。

EEG においても EMG と類似した技術的発展の経緯が存在することが、Pfurtscheller ら [100] によって概観されている。彼らの報告によれば、ユーザの運動想起に伴う EEG の識別においても、閾値的な判別から機械学習へと変遷していったことが示されている。ここで特筆すべきは、sEMG とは異なり、EEG では比較的早い段階から機械学習が活用されていた点である。実際、Pfurtscheller らの解説において閾値による識別研究の紹介は少数にとどまり、その大多数は機械学習を中心として論じられている。この背景には、EEG 信号特有の複雑性といった要因に加え、研究が本格化した時期の技術的背景も強く影響していると推察される。ユーザの運動想起に伴う EEG 活動を用いて外部デバイス进行操作した最初期の研究として、1991 年の Wolpaw ら [101] によるカーソル移動の実験が広く知られている。この研究では、被験者の μ リズムの振幅を算出し、その値が事前に設定された 5 段階のレンジ（閾値）のいずれに属するかによって、カーソルを上下に移動させるルールベースの制御系が構築された。興味深いことに、この電圧レンジは固定ではなく、被験者ごとに、かつその訓練の進行度合いに応じて、オペレーターが逐次的に調整していたことが報告されている。このような人的介入の必要性は、EEG 信号における個人差や非定常性の扱いが、固定的な閾値処理では困難であることを浮き彫りにしている。こうした初歩的な手法の限界と、同時代における計算機能力の向上やアルゴリズムの成熟が重なったことで、EEG 信号処理の主流は、必然的に機械学習を用いたデータ駆動的なアプローチへと急速にシフトしていったものと考えられる。

EMG や EEG に適用されてきた機械学習手法は、その決定境界の形状に基づき、大きく線形識別器と非線形識別器に分類することができる。線形識別器とは、特徴空間上でクラスを分ける境界が直線・平面として表されるモデルであり、計算が軽量で学習データ数が少量であっても安定しやすい特徴がある。代表的な線形識別器として、線形判別分析 (Linear Discriminant Analysis: LDA) [102, 103] やロジスティック回帰 (Logistic Regression: LR) [104, 105] が挙げられる。LDA は、クラス内分散を最小化しつつクラス間分散を最大化するように写像軸を決定する手法であり、その実装の容易さと計算コストの低さが特徴である。LR は、未知のデータがどのクラスに属するのか確率を直接推定できるため、識別の確信度を評価したい場合や、どの特徴が識別に寄与しているのかを確認したい場合に有効な手法である。一方で、非線形識別器は、特徴空間においてクラス分布が複雑に入り組んでいる場合や、線形分離が困難なデータに対して有効なアプローチである。代表的な非線形識別器として、k 近傍法 (k-Nearest Neighbor: kNN) [106, 107] やサポートベクターマシン (Support

Vector Machine: SVM) [108, 109] が挙げられる。kNN は、特徴空間上の距離に基づき、近傍サンプルの多数決によってクラスを決定する直感的な手法であり、少数のデータからでもデータの局所的な分布を反映した複雑な決定境界を形成できる。SVM は、カーネル法を用いることでデータを高次元空間へ写像し、非線形な決定境界を構築する手法である。一般に非線形モデルは学習データが不足すると汎化性能を低下させる過学習を招きやすいが、SVM はマージン最大化という学習基準によってこの問題を抑制しており、EMG や EEG の識別ではスタンダードな識別器として用いられてきた [110, 111]。この他にも、単一の識別器ではなく、複数の識別器を組み合わせる予測を行うアンサンブル学習のアプローチも、生体信号処理において高い有効性を示している。代表的な手法として、ランダムフォレスト (Random Forest: RF) [112, 113] や AdaBoost [114, 115] が挙げられる。RF は、複数の決定木を並列して作成し、それらの多数決を取ることでモデルの分散を抑え、ノイズの多いデータに対しても安定した識別性能を実現する。AdaBoost は、前段で構築した識別器で誤識別されたデータに重みを与え、次段の識別器を構築するアプローチであり、逐次的にモデルを更新する手法である。以上は、EMG や EEG でよく使われる識別器の一例であり、この他にも様々な工夫が凝らされた識別器が採用されている。

近年では、ニューラルネットワーク (Neural Network: NN) に基づく手法も大きな注目を集めている。NN の起源は、1958 年に Rosenblatt らによって提案されたパーセプトロン (Perceptron) [116] にまで遡る。当初のパーセプトロンは単純な線形識別器であったが、その後、Widrow らによる ADALINE (Adaptive Linear Neuron) [117] において、誤差を最小化する学習則が導入され、数理的な基礎が固められた。しかし、これら単層のネットワークは線形分離不可能な問題を解けないという致命的な制約 (XOR 問題など) を抱えていた。この課題を解決するために登場したのが、多層パーセプトロン (Multi-Layer Perceptron: MLP) [118] であり、単層から多層に拡張したことで非線形分離が可能なモデルに進化した。この MLP は、文献においては人工ニューラルネットワーク (Artificial Neural Network: ANN) と呼ばれることもあり [119, 120, 121, 122]、特に EMG の識別において標準的な識別器として利用されてきた [123]。この MLP は NN の基礎的なモデルであるが、MLP を組み合わせることで最新の NN にも匹敵する MLP-Mixer [124] という手法が提案されており、その重要性は現在でも失われていない。また、時系列情報に特化させたりカレントニューラルネットワーク (Recurrent Neural Network: RNN) [125, 126, 127] や、空間的な情報に特化させた畳み込みニューラルネットワーク (Convolutional Neural Network: CNN) [97, 128, 129] などのアーキテクチャも採用されている。さらに、これらのモデルに続く新たな潮流として、自己注意機構 (Self-Attention Mechanism) を組み込んだ Transformer [130, 131, 132] の採用も進んでいる。

2.4 転移学習

転移学習 (transfer learning) とは、あるドメイン (ソースドメイン) で学習された知識を、別のドメイン (ターゲットドメイン) の学習に再利用することで、ターゲットドメインにおける性能向上や学習効率の改善を図る手法の総称である [75, 76]。しかし、転移学習という用語

は非常に多岐にわたる文脈で使用されており、転移学習のレビューを行った Weiss ら (2016) [76] はこの状況を "Throughout the literature on transfer learning, there are a number of terminology." と述べているように、分野で用語の使われ方が統一されていない。また、Weiss らは、ラベル付きデータとラベルなしデータの有無に基づいて転移学習の過程をどのように特徴づけるかについても、文献間で用語の使い方が一貫していないことも指摘している [76]。このような背景としては、Pan ら (2009) [75] は転移学習という概念が確立される以前から、類似したアイデアが異なる分野と文脈で研究されてきたことを指摘している。Pan らによれば、1995 年頃から "learning to learn", "life-long learning", "knowledge transfer", "inductive transfer", "multitask learning", "knowledge consolidation", "context-sensitive learning", "knowledge-based inductive bias", "metalearning", and "incremental/cumulative learning", など様々な名称で呼ばれてきたとされている [75]。そのため本研究では、転移学習の定義と問題設定について、分野で標準的で引用数の多い、Pan ら (2009) [75] と Weiss ら (2016) [76] の定義に基づき、以下のように形式的に定義し、議論を進めることとする。

2.4.1 転移学習の定義と問題設定

Pan ら [75] によれば、転移学習は「ドメイン (Domain)」と「タスク (Task)」という 2 つの基本概念によって構成される。

まず、ドメイン $\mathcal{D}$ は、特徴空間 (feature space) $\mathcal{X}$ とその周辺確率分布 (marginal probability distribution) $P(\mathbf{x})$ の 2 要素からなり、 $\mathcal{D} = \{\mathcal{X}, P(\mathbf{x})\}$ と定義される。ここで、 $\mathbf{x} = \{x_1, \dots, x_n\} \in \mathcal{X}$ は入力データベクトルを表す。生体信号識別の文脈においては、 $\mathcal{X}$ は計測された信号から抽出された特徴ベクトル空間に相当し、 $P(\mathbf{x})$ はある被験者から得られる生体信号データの生成分布を意味する。したがって、2 つのドメインが異なる ($\mathcal{D}_S \neq \mathcal{D}_T$) という状態は、特徴空間が異なる ($\mathcal{X}_S \neq \mathcal{X}_T$)、あるいは特徴空間は同一だが周辺確率分布が異なる ($P(\mathbf{x}_S) \neq P(\mathbf{x}_T)$) ことのいずれか、または両方を指す。生体信号型 HCI において、 $\mathcal{D}$ の構成要素による違いは以下のように解釈できる。まず、特徴空間 $\mathcal{X}$ については、ソースドメインでは 8 ch の電極を使用しているが、ターゲットドメインでは 4 ch の電極を使用しているような場合には、ドメイン間で計測されるデータの次元数が異なるため、 $\mathcal{X}_S \neq \mathcal{X}_T$ となる。次に、周辺確率分布 $P(\mathbf{x})$ については、同じ計測機器を用いていても ($\mathcal{X}_S = \mathcal{X}_T$)、例えば筋肉質である被験者 A と非筋肉質である被験者 B では、ある強度の信号の出やすさは個人に依存するため、分布が異なり $P_S(\mathbf{x}) \neq P_T(\mathbf{x})$ となる。

次に、タスク $\mathcal{T}$ は、ラベル空間 (label space) $\mathcal{Y}$ と予測関数 $f(\cdot)$ の 2 要素からなり、 $\mathcal{T} = \{\mathcal{Y}, f(\cdot)\}$ と定義される。予測関数 $f(\cdot)$ は、学習データを用いたトレーニングによって、条件付き確率分布 $P(y|\mathbf{x})$ を推定することで得られる関数であり、未観測の入力 $\mathbf{x}$ に対して対応するラベル $y \in \mathcal{Y}$ を予測するために用いられる。生体信号型 HCI において、 $\mathcal{T}$ の構成要素による違いは以下のように解釈できる。まず、ラベル空間 $\mathcal{Y}$ は識別対象となる動作クラス (例: 手の動作, 運動想起の種類など) に相当する。ソースドメインとターゲットドメインで識別したい動作セットが異なる場合 (例: ソースでは「グー・パー」、ターゲットでは「サムアップ・ピース」など)、ラベル空間が異なるため $\mathcal{Y}_S \neq \mathcal{Y}_T$ となる。また、予

測関数に関しては、たとえ識別動作が同一 ($\mathcal{Y}_S = \mathcal{Y}_T$) であっても、被験者ごとの神経活動パターンや生体組織の構造差などによって、入力信号 $\mathbf{x}$ に対するクラス y の条件付き確率 $P_S(y|\mathbf{x}) \neq P_T(y|\mathbf{x})$ に当たる。

これらの定義に基づき、転移学習は次のように定義される。ソースドメイン $\mathcal{D}_S$ とソースタスク $\mathcal{T}_S$ 、およびターゲットドメイン $\mathcal{D}_T$ とターゲットタスク $\mathcal{T}_T$ が与えられたとき、転移学習の目的は、 $\mathcal{D}_S$ と $\mathcal{T}_S$ から得られた知識 (Knowledge) を活用することで、 $\mathcal{D}_T$ におけるターゲット予測関数 $f_T(\cdot)$ の学習効率や予測性能を向上させることである。一般的な機械学習の枠組みではこれらが一致していることを想定しているが、転移学習では $\mathcal{D}_S \neq \mathcal{D}_T$ または $\mathcal{T}_S \neq \mathcal{T}_T$ のいずれか、あるいは両方が成立している状況を前提とする。

2.4.2 転移学習の分類

転移学習の分類については、先述の通り研究者や文脈によって様々な切り口が存在するが、本節では Pan ら [75] による「ソースとターゲットのドメインおよびタスクの関係性」に基づく分類と、Weiss ら [76] による「特徴空間の異同」に基づく分類の2つの主要な枠組みについて述べる。

(1) ドメインとタスクの関係性に基づく分類 (Pan らの定義)

Pan ら [75] は、ソースドメイン $\mathcal{D}_S$ とターゲットドメイン $\mathcal{D}_T$ 、およびソースタスク $\mathcal{T}_S$ とターゲットタスク $\mathcal{T}_T$ の関係性、さらにはラベル付きデータの可用性に基づいて、転移学習を以下の3つの主要な設定 (Settings) に分類している。

1. 帰納的転移学習 (Inductive Transfer Learning)

帰納的転移学習は、ソースタスクとターゲットタスクが異なる場合 ($\mathcal{T}_S \neq \mathcal{T}_T$) の設定である。ドメインの異同 ($\mathcal{D}_S = \mathcal{D}_T$ か否か) は問わない。この設定では、ターゲットドメインにおいて予測モデル $f_T(\cdot)$ を構築するために、ターゲットドメインのラベル付きデータが利用可能であることが前提となる。さらに、ソースドメインにラベル付きデータが存在するか否かによって、マルチタスク学習 (Multitask Learning) や自己教示学習 (Self-taught Learning) といったアプローチの設定に相当する。生体信号識別の文脈では、ある動作 (タスク A) の識別モデルを学習する際に、別の動作 (タスク B) のデータを補助的に用いるケース ($\mathcal{Y}_S \neq \mathcal{Y}_T$) などがこれに該当する。また、被験者ごとの生体特性の違いにより、たとえ同一の信号パターン $\mathbf{x}$ が観測されたとしても、それが示唆する課題 y は個人によって異なり得るケース ($P_S(y|\mathbf{x}) \neq P_T(y|\mathbf{x})$) などがこれに該当する。

2. トランスダクティブ転移学習 (Transductive Transfer Learning)

トランスダクティブ転移学習は、ソースタスクとターゲットタスクは同一 ($\mathcal{T}_S = \mathcal{T}_T$) であるが、ドメインが異なる ($\mathcal{D}_S \neq \mathcal{D}_T$) 場合の設定である。この設定では、ソース

ドメインには豊富なラベル付きデータが存在する一方、ターゲットドメインにはラベル付きデータが存在しないが、ラベルなしであればデータは利用可能であるという状況を想定する。これはドメイン適応 (Domain Adaptation) と呼ばれる問題設定と密接に関連しており、生体信号 HCI において、あるユーザ (ソース) のデータで学習したモデルを、キャリブレーションなしで別のユーザ (ターゲット) に適用する場合などがこれに該当する。

3. 教師なし転移学習 (Unsupervised Transfer Learning)

教師なし転移学習は、ターゲットタスクがソースタスクと異なるが関連しており ($\mathcal{T}_S \neq \mathcal{T}_T$)、かつ両ドメインにおいてラベル付きデータが存在しない設定である。主にクラスタリングや次元削減、密度推定などの教師なし学習課題を解く際に用いられるが、Pan らがまとめた当時ではほとんど研究が行われていない分野であるとされている。生体信号型 HCI においては、識別を行うためにはラベル付きデータが必須であるため、教師ありの機械学習を主目的とする本研究の対象範囲外となる。

(2) 特徴空間の異同に基づく分類 (Weiss らの定義)

Weiss ら [76] は、Pan ら [75] の定義を踏襲しつつ、近年ビッグデータ環境下で重要性を増している「特徴空間 $\mathcal{X}$ の異同」に着目し、転移学習を以下の 2 つに大別している。

1. 同種転移学習 (Homogeneous Transfer Learning)

同種転移学習は、ソースドメインとターゲットドメインの特徴空間が同一 ($\mathcal{X}_S = \mathcal{X}_T$) である場合を指す。この設定における主な課題は、特徴空間は同じであっても、周辺確率分布が異なる ($P(\mathbf{x}_S) \neq P(\mathbf{x}_T)$) という分布シフトの問題を解決することにある。生体信号型 HCI の文脈においては、同一の計測機器・電極配置を用いつつ被験者が異なるという状況を扱う場合、この同種転移学習の枠組みに対応する。

2. 異種転移学習 (Heterogeneous Transfer Learning)

異種転移学習は、ソースドメインとターゲットドメインの特徴空間が異なる ($\mathcal{X}_S \neq \mathcal{X}_T$) 場合を指す。例えば、テキストデータから画像データへ知識を転移する場合などがこれに該当し、分布の違いを補正する前に、異なる特徴空間を共通の潜在空間へ写像するなどの対処が必要となる。生体信号型 HCI の文脈においては、多数の電極を有する高密度計測システム (ソース) から得られた知識を、電極数の少ないウェアラブルデバイス (ターゲット) へ転移させる場合、この異種転移学習の枠組みに対応する。

2.4.3 生体信号識別における転移学習

前節で概説した通り、転移学習はソースとターゲットのドメインやタスクの関係性、および特徴空間の異同に応じて多様な問題設定として定式化される。しかし、本研究は、生体信号型 HCI においてはこれらに加えて、実運用上の制約を強く意識した設計が必要であると

考える．第一に考慮すべきは，ターゲットユーザ側（エッジ側）における計算・メモリ資源の制約である．生体信号インタフェースは，多くの場合，ウェアラブルデバイスやスマートフォン，あるいは義手内の組み込みシステムといった，計算能力やバッテリー容量に厳しい制限がある環境での動作が想定される．そのため，ユーザ側での機械学習モデルの学習や複雑な最適化計算を伴う転移学習手法は，高精度であったとしても，こうしたデバイス上で実行することは困難である．したがって，可能な限り軽量であり，ユーザ側のデバイスに過度な負荷をかけないアルゴリズムであることが求められる．第二に考慮すべきは，データ共有とストレージに関する制約である．生体信号は個人の健康状態や身体的特徴を反映する極めて機微なプライバシー情報を含んでいるため，他者との無制限な共有はセキュリティおよび倫理的な観点から好ましくない場合がある．また，転移学習のために多数のソース被験者のデータをターゲットユーザのデバイスに全て保存・展開することは，ストレージ容量の観点からも非効率的である．したがって，システムの実用性を高めるためには，ユーザを超えたデータの共有は最小限に留める，あるいはモデルパラメータや統計量といった抽象化された知識のみを共有する設計が望ましい．

この観点に立つと，現在機械学習分野で活発に研究されている手法であっても，本研究の想定する環境には必ずしも適さない．転移成分分析 (Transfer Component Analysis: TCA) [133] は，ソースデータとターゲットデータの分布の差を最小にするように，共通の特徴量空間に写像するカーネル法ベースの転移学習である．TCA は，データ分布の差を最小にしつつ主成分分析のように次元削減を行うことができる利点があるものの，TCA の実行には全ソースデータの共有が必須である．また，写像されたソースデータとターゲットデータを用いて機械学習モデルを新たに構築する必要があり，本研究の想定した環境には合わない．類似した手法である相関アライメント (CORrelation ALignment: CORAL) [134] は，ソースデータをターゲットデータの分布（共分散行列）に一致させるよう写像する手法であるが，これも同様にソースデータの共有と機械学習の再構築を必要とする．次に，深層学習ベースのアプローチである Deep CORAL [135] やドメイン敵対的ニューラルネットワーク (Domain-Adversarial Neural Network: DANN) [8] について検討する．これらは特徴抽出と識別の学習を統合できる強力な手法であるが，学習プロセスにおいてソースデータとターゲットデータの同時入力が前提となるため，依然としてデータ共有の制約を回避できない．一方で，漸進型ニューラルネットワーク (Progressive Neural Network: PNN) [136] は，ソースデータで学習された NN のパラメータを凍結して転用するため，ソースデータの共有を必ずしも前提としない．しかし，ターゲットタスク用に新たなニューラルネットワーク列を追加し，そのパラメータを学習する必要があるため，ユーザ側のデバイス上で深層学習モデルの構築および学習負荷が発生し，計算資源の制約を満たすことが困難である．以上のように，必ずしも全ての転移学習が，生体信号型 HCI の実運用における制約に適合するわけではない．

したがって，本研究の要件に合致するのは，ソースドメインの知識を効率的に利用し，少ない計算リソースでターゲットドメインへの適応を実現する手法である．これを満たす転移学習手法として，以下の手法が挙げられる．

- **スタイル転移マッピング (Style Transfer Mapping: STM) [137]**

STM は TCA や CORAL と類似した、データの写像に基づく転移学習手法である。しかし、STM は TCA や CORAL よりもソースデータに対する依存性が低い特徴がある。TCA や CORAL は、ソースとターゲットの分布の一致を目的としているため、分布の推定を安定させるために多量のソースデータが必要である。一方で、STM は分布の一致を行うのではなく、ソースデータの代表点に向かって写像を行う。特に、この代表点はプロトタイプと呼ばれ、k-means で計算された重心データなどが採用される。一般に、k-means の重心は実際のデータそのものではなく、データ群を要約した統計的な代表点であるため、STM では適応時に全ソースデータを共有する必要がなく、事前に計算された少数のプロトタイプのみを共有すれば足りる。また、STM ではターゲットデータをソースデータのように写像するため、ソースデータで学習された機械学習モデルを一切の調整なしに再利用することが可能であり、ユーザ側での再学習の負担がかからず、本研究の要件に合致する。

- **共変量シフト適応 (Covariate Shift Adaptation: CSA) [138]**

CSA は、入力分布が変化しても条件付き確率は変化しない ($P_S(y|\mathbf{x}) = P_T(y|\mathbf{x})$) という仮定のもと、分布のズレ（共変量シフト）を補正するアプローチである。一般的な CSA では、ソースデータの重要度を再重み付けすることで分布を一致させるが、本研究の文脈においては、統計的なパラメータを適応させる手法として位置づけられる。例えば、LDA をベースとする場合、ソースドメインで計算された平均ベクトルや共分散行列といった統計量を「事前知識」として利用し、ターゲットデータから得られた統計量と、重み付けを通じて統合することでモデルパラメータを更新する。このプロセスは、生データそのものではなく、抽象化された統計量のみを操作対象とするため、データ共有の必要がない。さらに、パラメータの更新は低コストな計算で完了するため、計算負荷を抑えることができ、実運用に適した手法である。

- **ファインチューニング (Fine-Tuning: FT) [139]**

FT は、NN において最も標準的な転移学習手法であり、ソースドメインで学習済みのネットワークの重みを初期値として利用し、ターゲットデータを用いてパラメータを微修正するアプローチである。類似した手法に Deep CORAL や DANN が存在するが、これらはドメイン間の分布の差を吸収するようにネットワークを学習させるために、ソースデータの同時入力が必要とされる。これに対し、FT の目的はあくまで事前学習された NN をターゲットドメインで再最適化するものであるため、ソースデータは不要である。また、FT は必ずしもネットワーク全体を微調整する必要はなく、出力に近い僅かな層のみを調整するのでも十分な効果があり、計算コストを低く抑えられる。PNN も同様に事前学習済み NN を利用するが、PNN ではそれと並列する形でターゲット専用の新しい NN を追加する手法であるため、計算コストが大きい。したがって、FT は、ソースデータを必要とせず、計算コストも比較的低いという点で、実運用環境において扱いやすい転移学習手法であると言える。

2.5 本研究における問題設定

前述の定義を本研究が扱う生体信号の被験者間転移（cross-subject transfer）の問題に当てはめると、本研究の問題設定は次のように定式化できる。

まず、ドメイン $\mathcal{D}$ とタスク $\mathcal{T}$ の構成要素のうち、両方で共有される要素について述べる。本研究において、ソース被験者（既存ユーザ）とターゲット被験者（新規ユーザ）は、同一の計測デバイスとプロトコルに従ってデータを取得するため、観測されるデータの物理的な次元や性質は等しい。したがって、特徴空間は共通であるとみなせる ($\mathcal{X}_S = \mathcal{X}_T$)。また、識別対象となる動作セットも共通であるため、ラベル空間も同一である ($\mathcal{Y}_S = \mathcal{Y}_T$)。

しかしながら、生体信号には解剖学的・生理学的な要因による不可避な個人差が存在し、これが確率分布の乖離を生じさせる。これは、同一の課題を行っても観測される信号の強度や周波数分布が異なることを意味し、ドメイン間の周辺確率分布が一致しないこと ($P_S(\mathbf{x}) \neq P_T(\mathbf{x})$) を意味する。したがって、 $\mathcal{X}_S = \mathcal{X}_T$ であるものの、周辺確率分布がソースとターゲット間で異なるため、 $\mathcal{D}_S \neq \mathcal{D}_T$ である。さらに、被験者ごとの生体特性の違いにより、たとえ同一の信号パターン $\mathbf{x}$ が観測されたとしても、それが示唆する課題 y は個人によって異なり得る。これは数学的には、入力から出力への対応規則、すなわち条件付き確率分布が被験者間で一致しないこと ($P_S(y|\mathbf{x}) \neq P_T(y|\mathbf{x})$) を意味する。したがって、 $\mathcal{Y}_S = \mathcal{Y}_T$ であるものの、条件付き確率分布がソースとターゲット間で異なるため、 $\mathcal{T}_S \neq \mathcal{T}_T$ である。

本研究では、以上の問題設定の元で、ターゲットとなる被験者から少数のラベル付きデータを取得し、これを元に転移学習を行う。この場合、Pan ら [75] の分類に従えば、ターゲット側にラベル付きデータが存在し、かつ条件付き分布が被験者間で一致しない ($\mathcal{T}_S \neq \mathcal{T}_T$) ため、帰納的転移学習（inductive transfer learning）に位置づけられる。Weiss ら [76] の定義に従う場合、 $\mathcal{X}_S = \mathcal{X}_T$ であるため、同種転移学習に位置づけられる。

以上より、本章では、本研究が対象とする生体信号型 HCI の基盤となる要素技術を整理した。まず sEMG と EEG を中心に、生体信号がどのような生理学的機構に由来し、どの段階で解剖学的・生理学的・機能的要因に起因する個人差が生じるかを概説した。次に、これらの信号からユーザの意図や状態を推定する枠組みとして教師あり機械学習を位置づけ、被験者内識別と被験者間識別の違い、およびドメインシフトの存在が識別性能に与える影響を整理した。さらに、こうしたドメインシフトを克服し少量データ条件下でも高い性能を実現するための枠組みとして転移学習を取り上げ、ドメイン／タスクの関係性および特徴空間の異同に基づく分類を踏まえつつ、本研究が扱う被験者間転移の問題設定を、帰納的かつ同種の転移学習として位置づけた。本章で整理した生体信号の特性、機械学習の枠組み、そして転移学習の定式化は、第3章以降で提案する具体的な手法設計および評価結果を理解するための理論的土台となる。

2.6 結言

本章では、生体信号型 HCI の構築に不可欠な基礎理論として、信号の発生機序、機械学習の枠組み、および転移学習の体系について整理を行った。sEMG および EEG は、ユーザの運動意図を反映する有用な情報源である一方で、皮下脂肪厚や筋線維組成、あるいは脳の褶曲構造や想起の方法の違いといった解剖学的・生理学的な個人差の影響を強く受ける。この個人差は、機械学習の文脈においては、訓練データ（ソースドメイン）とテストデータ（ターゲットドメイン）間の分布の乖離、すなわちドメインシフトとして顕在化し、他者のデータで学習したモデルの汎化性能を著しく低下させる要因となることを確認した。この課題に対し、本章では転移学習の理論的枠組みを導入し、生体信号の個人差問題との関連を議論した。特に、HCI の実運用環境を考慮した場合、データ共有や計算リソースの制約の観点から、無制限なデータの共有や重い再学習を伴う手法は適さないことを指摘した。その上で、本研究では STM, CSA, FT が実用的な転移学習手法であることを示した。また、本研究が取り組む課題を、特徴空間とラベル空間は共有されているが、周辺確率分布および条件付き確率分布が異なる条件下での、帰納的かつ同種の転移学習として定式化した。以上の議論により、生体信号における個人差の要因と、それを解決するための数理的なアプローチが確認された。次章以降では、本章で概観した基礎をもとに、生体信号型 HCI の実用化に向けた具体的な検証を行う。

第3章 実用的な生体信号識別に向けた機械学習と転移学習の比較検討

3.1 緒言

前章までに述べたように、生体信号 HCI では機械学習が重要な役割を持っているが、個人差の影響を強く受けるという弱点があった。そのため、転移学習と組み合わせることで、個人差による悪影響を軽減する必要がある。しかし、機械学習と転移学習にはそれぞれ多様な手法が存在するため、これらを適切に選択し利用する必要がある。本章では、数ある機械学習や転移学習の手法の中から有力な候補を選定し、それらの組み合わせを比較検討し、最適な構成を検討する。

第 2.3.2 項で概観したように、機械学習モデルは、線形モデルと非線形モデルに大別される。線形モデルは計算コストが低く、学習データ量が少ない場合にも比較的安定した識別が可能であるが、生体信号は本質的には非線形な性質を持つため、非線形モデルの方がより複雑な信号パターンを捉えるのに適していると考えられる。そのため、本章においては非線形モデルを採用する。また、機械学習モデルは伝統的な機械学習と、NN に基づく深層学習に大別される。そこで本章では、伝統的な機械学習の代表として SVM を、深層学習の代表として MLP (本章では以降、ANN と呼称) をそれぞれ選定する。SVM は EMG や EEG の識別において長らく標準的な識別器として利用されてきた実績がある [110, 111]。一方、ANN は基礎的な NN でありつつも高い表現能力を持ち、特に EMG 分野においては最も頻繁に使用されている識別器である [123]。以後、これら 2 つのアプローチの対比を明確にするため、SVM を浅層識別器 (Shallow Classifier)、ANN を深層識別器 (Deep Classifier) と呼称し、これらを転移学習のバックボーンモデルとして使用する。

次に、これらの識別器と組み合わせる転移学習手法の選定をする。第 2.4.3 項で論じたように、生体信号型 HCI の実運用においては、全ソースデータの共有や大規模な再学習を必要とする手法は適さない。この要件を満たす手法としては、STM と FT が最適である。STM は EEG [140] や sEMG [141] の識別において、従来手法である TCA や CSA などよりも優れた性能を示すことが報告されている。同様に FT も sEMG [142] や EEG [143] の識別においてその有効性が確認されている。ここで、両者のアプローチの違いを図 3.1 に示す。図 3.1(a) は転移学習適用前の状態を表しており、ソースドメイン (青色) とターゲットドメイン (赤色) のデータ分布が乖離しているため、ソースデータで学習された決定境界 (黒線) ではターゲットデータを正しく識別できないことを示している。これに対し、図 3.1(b) は STM の概念を表している。このアプローチでは、決定境界を固定したまま、ターゲットデータをソースデータの分布に合わせて写像 (移動) することで、既存の識別器での識別を

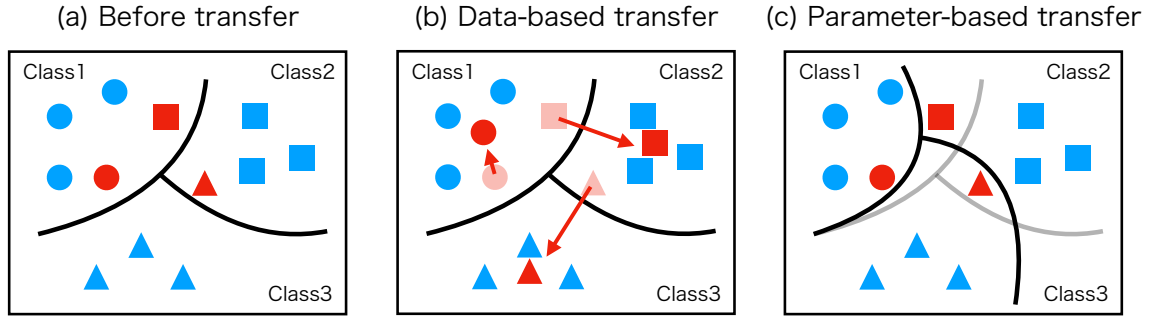


図 3.1: 本章で使用する転移学習の概念図. 各図はそれぞれ, (a) 転移前 (b) データ空間ベース転移学習 (c) パラメータ空間ベース転移学習を表す.

可能にする. 一方, 図 3.1(c) は FT の概念を表している. このアプローチでは, データ分布を固定したまま, ターゲットデータに合わせて識別器のパラメータ (決定境界) を更新・変形させることで, 適切な識別を実現する. 本章では, 前者のように特徴空間上のデータ分布を操作するアプローチをデータ空間ベースの転移学習, 後者のように識別器のモデルパラメータを操作するアプローチをパラメータ空間ベースの転移学習と呼称する.

これまでの研究では, 既存の同種手法との比較や転移学習の有無の比較に留まることが多く, また, 識別器の深浅を跨いだクロスオーバーな比較は十分にされていなかった [138, 140, 141, 144, 145]. したがって, これらの手法の組み合わせに関しては明確な基準が存在せず, 開発者はケースごとに構成を判断せざるを得ないのが現状である. そこで本章では, 識別器のタイプ (浅層・深層) と転移学習のアプローチ (データ空間ベース・パラメータ空間ベース) の最適な組み合わせを明らかにするための比較実験を行う. 具体的には, 転移学習を用いない場合をベースラインとし, データ空間ベースの転移学習 (STM) を適用した場合, およびパラメータ空間ベースの転移学習 (FT) を適用した場合の識別精度を定量的に評価する. 評価に際しては, STM は特徴空間自体を操作する手法であるため, 浅層識別器 (SVM) および深層識別器 (ANN) の双方と組み合わせで評価を行う. 一方, FT は事前学習済みの NN に対して実行する手法であるため, 深層識別器 (ANN) のみと組み合わせで評価を行う. これにより, 高い実用性を保ちつつも高精度な識別を可能にする構成を明らかにし, 本分野におけるベンチマークを確立することを目的とする.

検証には, 25 名の健常被験者から取得された前腕 8 動作を含む sEMG オープンデータセットを用いた. このデータセットに対し, 前述の組み合わせを適用して被験者間転移学習の精度を比較した. その結果, 深層識別器は浅層識別器よりも高い基礎性能を示し, さらに転移学習の適用によっていずれも大幅な精度向上が確認された. 中でも, パラメータ空間ベースの転移学習と深層識別器の組み合わせは, データ空間ベースの手法を含む他の全ての構成を上回り, 最も高い識別性能を達成した. 以上の結果に基づき, 本章では, パラメータ空間ベースの転移学習と深層識別器の組み合わせを標準的なアプローチとして位置づける.

3.2 データセットおよび解析手法

3.2.1 データセット

本章では、以下の公開データセットを使用した。このデータセットには、25名の健常者を対象として計測された8種類の前腕動作のsEMGデータが収録されている[9]。収集方法の詳細を以下に示す。

(1) 被験者

本実験には、18名の男性と7名の女性（年齢: 20–31歳，平均 ± 標準偏差 [SD]: 24.4 ± 3.0歳）から取得されたデータを用いた。なお，データ収集に関する実験は産業技術総合研究所の倫理委員会による承認を受けて実施されており（承認番号: HF2017-784），すべての被験者は参加前にインフォームドコンセントを提供した。

(2) 計測機器

sEMGデータの計測には、コンシューマー向けモバイル計測機器であるMyoジェスチャーコントロールアームバンド（Myo Gesture Control Armband, Thalmic Labs, Kitchener, Canada）を使用した。この機器は8チャンネル分の電極を持ち、それらを伸縮性のバンドで腕周りに固定することが可能であり、図3.2のように被験者の右前腕部に装着された。Bluetooth通信ポートを備えた4番目のチャンネルは腕橈骨筋の筋腹部から遠位方向に約1cmの位置に配置され、7番目のチャンネルは肘頭から遠位方向に6cm離れた位置に配置された。その他の6チャンネルは前腕部を囲むように等間隔に配置された。信号は200Hzのサンプリング周波数で収集され、5次バターワースフィルタを用いて15Hzのハイパスフィルタリングが施された。

本章で用いた計測機器は、前腕周囲に電極を周方向に配置するバンド型デバイスであり、これには明確な利点と欠点が存在する。まず、本デバイスの最大の利点は、その装着の容易さにある。この特徴は、一般ユーザが日常生活の中で手軽に利用できるHCIを実現する上で極めて重要である。一方で欠点として、電極の配置が解剖学的な知見に基づいたものではないため、電極の位置によっては複数の近傍筋からの信号が混入してしまうことが挙げられる。したがって、各チャンネルの信号を解剖学的な単一の筋へと一意に分離・同定することは困難であり、個別の筋活動に基づいた詳細な運動解析には不向きである。しかし、複数のジェスチャを安定して正しく認識するために、対象筋ごとに電極を毎回正確に配置することは、専門知識と時間を要し、日常利用を想定したHCIとしては運用上の負担が大きい。そのため、本章では個別の筋の活動を事前知識として仮定するのではなく、前腕周囲に配置された多チャンネルsEMGに含まれるパターン情報に基づいてジェスチャを識別する。この方針により、装着ずれや解剖学的個人差を含む実環境下でも運用可能な、実用性の高いHCIの実現に資することを目指す。

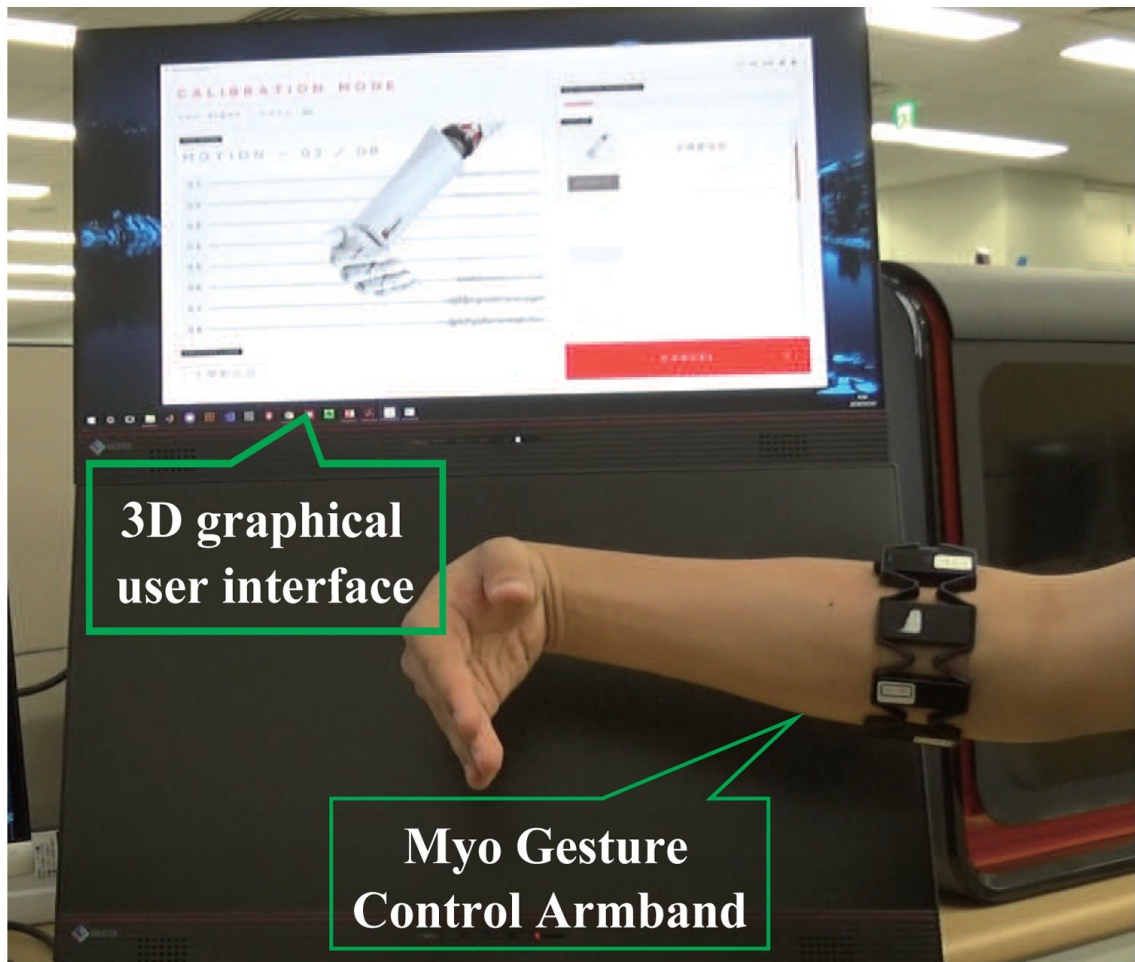


図 3.2: MyoDataset で使用されている計測機器 MyoArmband の装着の様子. 文献 [9] から引用.

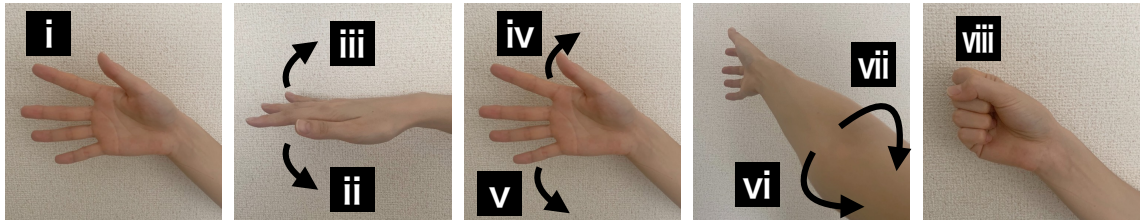


図 3.3: 8 種類の単一自由度動作. 文献 [10] より転載.

(3) 実験内容

被験者は以下の 8 種類の単一自由度動作を実施した (図 3.3 を参照):

- (i) 手開き (Hand Opening: HO)
- (ii) 手首の屈曲 (Wrist Flexion: WF)
- (iii) 手首の伸展 (Wrist Extension: WE)
- (iv) 手首の橈屈 (Radial Deviation: RD)
- (v) 手首の尺屈 (Ulnar Deviation: UD)
- (vi) 前腕の回内 (Forearm Pronation: FP)
- (vii) 前腕の回外 (Forearm Supination: FS)
- (viii) 手握り (Hand Closing: HC)

各動作をそれぞれ 5 回繰り返し, 合計 40 試行 (8 動作 $\times$ 5 回) を行った. 各試行は 6s のシーケンスで構成されている. 試行開始後 1 s 間, ノートパソコンの画面に提示される動画像によって, 被験者は実施する動作の指示を受ける. 被験者は指示に従って動作を実行し, 動作が完了した状態を 2 s 間維持する. その後, 被験者は 1s ほどかけて元の状態に戻し, 残り時間は待機する. 被験者はノートパソコンの前に立ち, 動作を行う際の腿と手の接触を避けるため, 両腕を手のひらが向かい合う状態で 10 cm 程度前進させ, その状態で実験を行った.

3.2.2 前処理

各 6s の試行には, 対象とする動作を実行していない非活動の時間区間が含まれている. このため, 動作の開始 (オンセット) を検出し, 収集されたデータから動作の実行中の活動区間を抽出する必要がある. 動作のオンセットはマルチスケールサンプルエントロピー (multiscale Sample Entropy: mSampEn) [146] を用いて検出した. 先行研究 [147] に基づいて, mSampEn アルゴリズムに必要な 4 つのパラメータ (N_w, N_s, m, r) をそれぞれ 40,

5, 2, および N_w の標準偏差の 0.25 倍に設定した。これら 4 つのパラメータはそれぞれ、エントロピーを計算する時間窓の長さを表すウィンドウ長 N_w , ウィンドウを時間方向に移動させる刻み幅を表すシフト幅 N_s , 比較する部分系列の長さを表すテンプレート次元 m , そして類似度判定に用いられる許容誤差 (Sigmoid 関数の傾きを制御するパラメータ) r である。チャンネルごとに mSampEn の値は得られ、それぞれ最小値と最大値を $[0, 1]$ に正規化した後、0.4 の閾値を設定し、全体のうち最初に閾値を超えた時点を動作のオンセットとして特定した。最終的に、特定されたオンセットを基準として各試行から活動区間として 1.5 s 間分のデータを抽出した。

抽出された 1.5 s 間の活動区間は、さらに解析窓というより短いデータ区間に分割された。これは、sEMG ベース HCI を実際に適用する場合、出力がこの解析窓の長さに依存することになり、基本的にインタフェースの 1 出力は 300 ms 未満であることが推奨されているためである [148]。本章では 250 ms を解析窓の長さとし、この窓を 50 ms ずつずらしながら (つまり 200 ms のオーバーラップを許しながら) 切り出した。各解析窓に対して、以下の時間領域および自己回帰モデルに基づく 11 次元の特徴量を算出した [149] :

- (i) 平均絶対値 (mean absolute value)
- (ii) ゼロ交差回数 (zero crossing)
- (iii) 傾き符号変化回数 (slope sign changes)
- (iv) 波形長 (waveform length)
- (v) 二乗平均平方根 (root-mean square)
- (vi) 6 次の自己回帰係数 (sixth-order autoregressive coefficients)

これらは、sEMG ベース動作認識において標準的な特徴量として広く用いられているものである。最終的に、250 ms の解析窓ごとに 88 次元の特徴量ベクトルが生成された。これらの特徴量は次元ごとに標準化 (平均 0 標準偏差 1) に変換した。

3.2.3 識別

本章では、STM と FT の先行研究 [140, 142] に倣い、被験者ごとに個別の識別器を構築した。

(1) 浅層識別器: Support Vector Machine

浅層識別器として、放射基底関数 (Radial Basis Function: RBF) カーネルを用いた SVM を採用し、8 クラスの sEMG データを識別した。SVM は、過去 20 年間にわたり sEMG データの識別に伝統的に使用されてきた識別器であり [110], ニューラルネットワークのような多層のアーキテクチャを持たないことから、浅層モデルとみなされる。浅層識別器とし

ては他にも線形判別分析 (LDA) などの選択肢があるが、本章では非線形でありながら訓練データ数が少ない場合でも比較的高い精度を示すことが知られている SVM を選択した。また、各動作クラスにつき 1 試行分のみを訓練データとして使用した (詳細は 3.3 節で後述)。SVM の実装には、Python で最も一般的に利用されている機械学習ライブラリである `scikit-learn` を用いた。RBF カーネルを用いた SVM のハイパーパラメータである C と γ については、 $10^{-3}, 10^{-2}, \dots, 10^3$ の範囲内で最適化を行った。この際、転移学習の有無に応じて異なる最適化手順を採用した (詳細は 3.3 節で後述)。また、本章で使用した SVM 識別器は離散的なクラスラベルを直接出力するのではなく、各クラスに対する確率値を出力し、その確率値を用いて更なる解析を実施した (詳細は 3.2.5 項で後述)。

(2) 深層識別器: Artificial Neural Network

深層識別器として、全結合層のみで構成された 3 つの隠れ層をもつ ANN を採用した。ANN は、sEMG によるジェスチャー識別において最も一般的かつシンプルな深層識別器である [123]。CNN や RNN など他の DNN アーキテクチャと比較して、ANN は構造が単純なため、転移学習の性能そのものに焦点を当てた解析が可能となる。本章における ANN の実装は、Python で最も広く使われている深層学習ライブラリである `Pytorch` を用いて行った。ANN には数多くのハイパーパラメータが存在し、それら全てを最適化するには非常に長い計算時間を要する。計算コストを抑えるため、いくつかのハイパーパラメータ群 (隠れ層の数、学習率、最適化手法、活性化関数) をそれぞれ 3, 0.01, Adam, Rectified Linear Unit (ReLU) に固定した。これらの事前設定されたハイパーパラメータの中でも特に、隠れ層の数は研究者の注目を集める傾向にある。隠れ層の数は識別性能に大きく影響し、層が少なすぎると過小学習 (underfitting) を引き起こし、一方で層が多すぎると過学習 (overfitting) につながる。このため、本章ではモデルの表現能力と汎化性能のバランスを考慮し、隠れ層の数を 3 層とした。次に、事前設定されていない他のハイパーパラメータについて最適化を行った。具体的には、各隠れ層のニューロン数を [22, 44, 88, 176, 352]、重み減衰 (weight decay) の値を $[10^{-1}, 10^{-2}, \dots, 10^{-5}]$ の範囲で最適化した。これらのパラメータの最適化にあたっては、転移学習の有無に応じて異なる最適化手順を採用した (詳細は 3.3 節で後述)。また、ANN 識別器はソフトマックス層から各クラスに対する確率値を出力し、この確率値を用いてさらなる解析を実施した (詳細は 3.2.5 項で後述)。

3.2.4 転移学習

(1) データ空間ベースの転移学習: Style Transfer Mapping

STM はデータ空間ベースの転移学習手法であり、一般的に浅層識別器と組み合わせて使用される。この手法の目的は、少量のターゲットデータをソースデータの代表点付近に写像し、写像されたターゲットデータとソースデータの代表点間の二乗誤差を最小化することで

ある．STMの実行は次の2つのステップからなる．最初のステップでは代表点の選定を行い，次のステップでは変換行列とバイアスの学習を行う．

第1ステップの代表点の選定においては，ソースデータセットに対してk-means クラスタリングアルゴリズムを適用した [150]．事前処理として，クラスタリングの性能に悪影響を及ぼす可能性のある外れ値的なソースデータを除去するためにSVMを使用した．まず，各ソース被験者のデータに対してSVMを学習させ，サポートベクター（SV）を特定した．その後，判別境界上に位置するSV（すなわち，一般的なパターンを代表しないデータ）を除去した．SVを除去した後，各ソース被験者の各クラスのデータに対してk-means クラスタリング（ $k = 15$ ）[140]を実施し，得られた各クラスのクラスタ中心ベクトルをSTMにおける「プロトタイプ」とした．キャリブレーションデータ（詳細は3.3節で後述）の i 番目のデータ $\mathbf{x}_{\text{cal}_i}$ に対しては，最も近いプロトタイプを代表点とし，これを $\mathbf{d}_i \in \mathbb{R}^{N_d}$ と表記した [151]． N_d は特徴量の次元数であり，本章では88である．

第2ステップでは，以下のコスト関数を最小化することにより，変換行列 $\mathbf{A} \in \mathbb{R}^{N_d \times N_d}$ とバイアス $\mathbf{b} \in \mathbb{R}^{N_d}$ を学習した．

$$\underset{\mathbf{A}, \mathbf{b}}{\operatorname{argmin}} \sum_{i=1}^{N_{\text{cal}}} \|\mathbf{A}\mathbf{x}_{\text{cal}_i} + \mathbf{b} - \mathbf{d}_i\|_2^2 + \beta \|\mathbf{A} - \mathbf{I}\|_F^2 + \gamma \|\mathbf{b}\|_2^2, \quad (3.1)$$

ここで， $\mathbf{x}_{\text{cal}_i} \in \mathbb{R}^{N_d}$ はキャリブレーションデータ， N_{cal} はターゲットデータの数， $\|\cdot\|_2^2$ は L_2 ノルム， $\|\cdot\|_F^2$ はフロベニウスノルム， $\mathbf{I} \in \mathbb{R}^{N_d \times N_d}$ は単位行列である．正則化パラメータ β および γ は，転移の程度を制御し，以下のように定義される．

$$\beta = \tilde{\beta} \frac{1}{N_d} \operatorname{tr} \left(\sum_{i=1}^{N_{\text{cal}}} \mathbf{x}_{\text{cal}_i} \mathbf{x}_{\text{cal}_i}^T \right), \gamma = \tilde{\gamma} N_{\text{cal}}, \quad (3.2)$$

ここで， $\operatorname{tr}(\cdot)$ は行列のトレースである．ハイパーパラメータ $\tilde{\beta}$ および $\tilde{\gamma}$ は，グリッドサーチにより最適化され，その探索範囲は0から3まで0.2刻みで設定された [140]．最終的に， $\mathbf{A}$ と $\mathbf{b}$ は以下のように計算された．

$$\mathbf{A} = \mathbf{Q}\mathbf{P}^{-1}, \mathbf{b} = \frac{1}{\hat{f}}(\hat{\mathbf{d}} - \mathbf{A}\hat{\mathbf{x}}), \quad (3.3)$$

ただし，

$$\mathbf{Q} = \sum_{i=1}^{N_{\text{cal}}} \mathbf{d}_i \mathbf{x}_{\text{cal}_i}^T - \frac{1}{\hat{f}} \hat{\mathbf{d}} \hat{\mathbf{x}}^T + \beta \mathbf{I}, \quad (3.4)$$

$$\mathbf{P} = \sum_{i=1}^{N_{\text{cal}}} \mathbf{x}_{\text{cal}_i} \mathbf{x}_{\text{cal}_i}^T - \frac{1}{\hat{f}} \hat{\mathbf{x}} \hat{\mathbf{x}}^T + \beta \mathbf{I}, \quad (3.5)$$

$$\hat{\mathbf{x}} = \sum_{i=1}^{N_{\text{cal}}} \mathbf{x}_{\text{cal}_i}, \hat{\mathbf{d}} = \sum_{i=1}^{N_{\text{cal}}} \mathbf{d}_i, \hat{f} = N_{\text{cal}} + \gamma. \quad (3.6)$$

ラベル付けされていない新規データは，学習された変換行列 $\mathbf{A}$ およびバイアス $\mathbf{b}$ を用いたアフィン変換によって線形写像され，その後，事前に学習済みのSVMまたはANN識別器によって識別された．

(2) パラメータ空間ベースの転移学習: FT

FT は、深層識別器と組み合わせて使用されるパラメータ空間ベースの転移学習手法である。FT は、事前学習済みの ANN の出力層付近の一部の層のみをターゲットユーザのキャリブレーションデータを用いて再学習する手法である。したがって、FT を適用するには、ソースデータで事前学習済みの ANN が必要となる。一般的な DNN では、入力層に近い層は特徴抽出に特化し、出力層に近い層は抽出された特徴の識別に特化していると考えられている [152]。FT では、ソースデータを用いて十分に事前学習されたネットワークの重み（パラメータ）を適切な初期値として利用するため、少量のキャリブレーションデータでも効率的な再学習が可能となり、結果として高い識別精度が得られる [139]。

FT を実施するために、本章では ANN の第 3 隠れ層と出力層の間のパラメータのみを再学習し、それ以外は固定した。また、キャリブレーションデータに対する過学習を防ぐために、ハイパーパラメータである L2 正則化項の強さを ANN の最適化時と同じ範囲 ($10^{-1}, 10^{-2}, \dots, 10^{-5}$) で再最適化した。FT は、ニューラルネットワークの隠れ層における表現を保持したまま、判別層（出力層付近）のパラメータのみを調整する方法であるため、SVM には適用できないことに留意されたい。

3.2.5 アンサンブル処理

個人ごとに識別器が構築された場合、マルチソース識別 (multisource classification) として知られるアンサンブル戦略を用いることが効果的となる [140, 153]。EEG の処理においては、ターゲットユーザと類似した 7 つのソースユーザの識別器 (ソース識別器) を用いたアンサンブル戦略が良好な性能を示すことが報告されている [79, 140]。しかし、sEMG 処理におけるソース識別器の最適な数については明らかになっていない。そこで本章では、使用するデータセットに含まれる第 1 ソース被験者から第 24 ソース被験者（データセットは 25 人分だが、1 人はターゲットとなるため）の各ケースにおいて、アンサンブル処理を実施した。まず、キャリブレーションデータに対する識別精度に基づいてソース識別器の順位付けを行い、上位 N ($N = 1, \dots, 24$) 個の識別器を選択した。次に、新規データに対して選択された識別器がそれぞれ各クラスの確率値を出力した。この際、キャリブレーションデータに対する識別精度を重みとして、この出力された確率値に掛け合わせ、クラスごとに N 個の確率値の和を取った。つまり、順位が高い (ターゲット被験者のキャリブレーションデータに良好な性能を示すソース) 識別器の出力値ほど識別に強く反映されるようになっている。

3.3 評価方法

本章では、データセットの 25 名の被験者のうち 1 名をターゲット被験者とし、残りの 24 名をソース被験者として用いた。各被験者を順番にターゲットとするため、評価は合計 25 回繰り返された。各ターゲット被験者に対して、ソース被験者の数を 1 名から 24 名まで順次増やしていき、識別精度がどのように変化するかを調査した。なお、ソース被験者の追

加は、3.2.5 項で説明したランキング（キャリブレーションデータに対する識別精度の降順）に従って行った．具体例を挙げると、図 3.4(A) に示されるように、ターゲット被験者が被験者 7 の場合、まずソース候補者（被験者 7 を除く被験者 1 から被験者 25）それぞれについて、個別の識別器を計 24 個作成した．次に、被験者 7 のキャリブレーションデータをこれらの識別器で識別し、24 個の識別精度を計算した．この中で最も高い精度を示したソース被験者（例えば、被験者 2）を最初のソース被験者として選択し、このソース被験者（被験者 2）を用いて、後述の 3.3.1 項、3.3.2 項、および 3.3.3 項に記載される各手法の識別性能を評価した．その後、他のソース被験者を「識別精度が高い順に」1 名ずつ（例えば被験者 9、被験者 15、…のように）追加していき、そのたびに性能を評価した．この結果、ターゲット被験者 7 に対しては合計 24 通りのソース被験者の組み合わせによる識別精度が得られた．この評価手順を、被験者 1 から被験者 25 までの全てのターゲット被験者について繰り返したため、合計 600 通り（25 名のターゲット被験者× 24 パターンのソース被験者追加方法）の評価結果を得た．

ターゲット被験者のデータに関して、最初の試行をキャリブレーションデータ、2 番目の試行を検証用データ、最後の 3 試行をテストデータとして設定した．識別器および転移学習手法の性能は、このテストデータ（3 試行）における 8 クラスの識別精度を用いて評価した．

転移学習手法と識別器の組み合わせにより、以下の 5 つの被験者間識別問題の条件を設定した．それぞれの評価手順を図 3.4、3.5、および 3.6 に示す．また、個別に学習した識別器（被験者内識別問題）による精度を調査するために、ターゲットデータのみで学習・テストを行う 2 つの参照条件（i-SVM および i-ANN）も設定した．この評価手順を図 3.7 に示す．したがって、最終的に以下に示す 7 つの条件を評価した．

1. SVM（転移学習なしで被験者間識別）
2. ANN（転移学習なしで被験者間識別）
3. STM-SVM（STM を用いた SVM で被験者間識別）
4. STM-ANN（STM を用いた ANN で被験者間識別）
5. FT-ANN（FT を用いた ANN で被験者間識別）
6. i-SVM（SVM で被験者内識別）
7. i-ANN（ANN で被験者内識別）

なお、上記の条件のうち最初の 5 条件では、アンサンブル戦略を採用した．識別器と転移学習の組み合わせのうち、どの組み合わせが最も高い識別性能を示すかを明らかにするために統計的な検定を実施した．本章では各要因の水準数に偏りがあるため、一般的な二元配置分散分析 [154] を直接適用するのではなく、段階的な統計検定を行った．まず初めに、4 つの条件（SVM, ANN, STM-SVM, STM-ANN）に対して二元配置分散分析を行い、STM が識別器に与える影響、および識別器の違いが識別精度に与える影響とその交互作用を評価した．次に、その 4 条件の中で最も良い精度を示した組み合わせと FT-ANN の比較を行った．こ

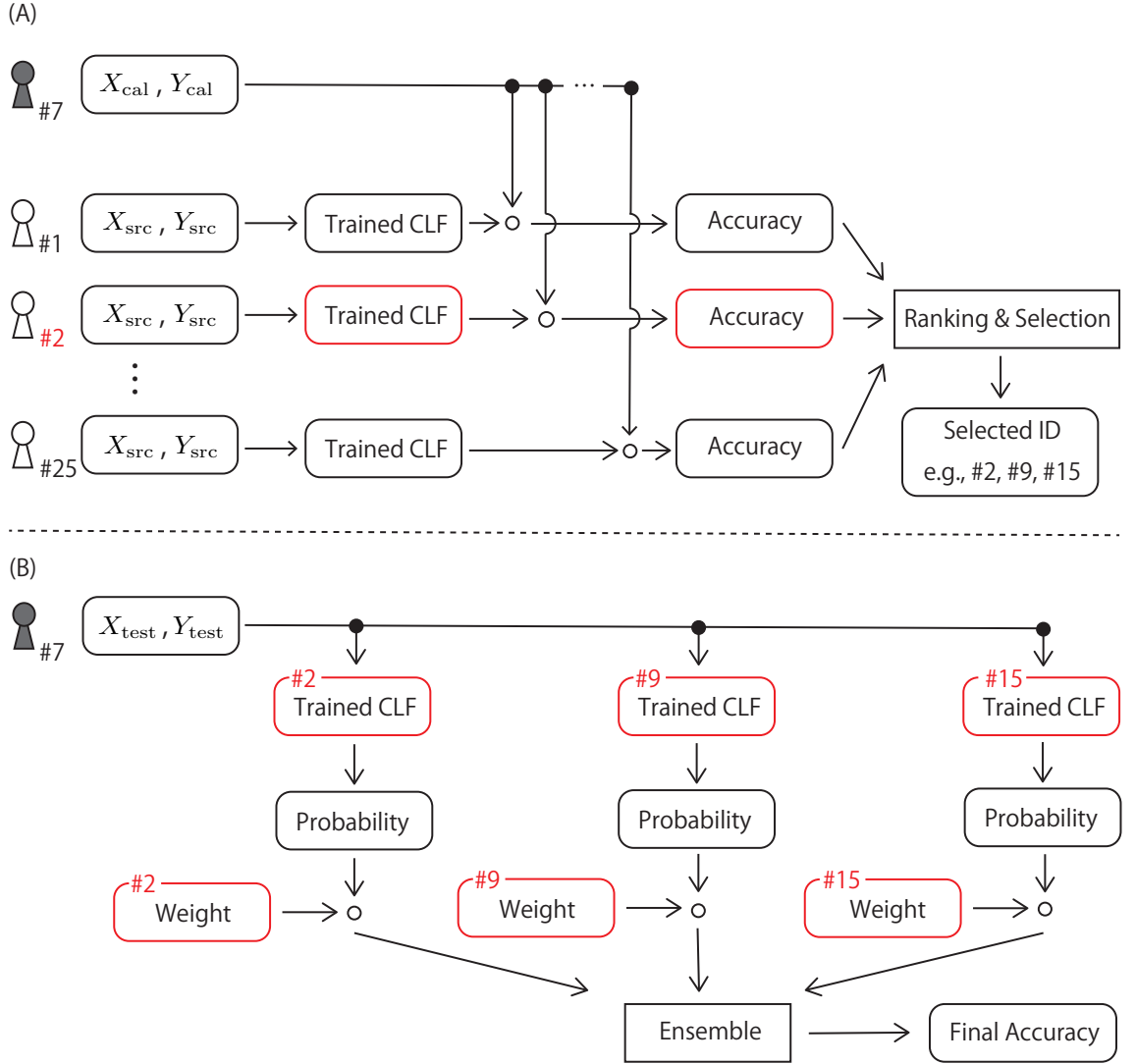


図 3.4: ターゲット被験者が被験者 7 のときの SVM と ANN における (A) 学習フェーズと (B) 推論フェーズのフローチャート. CLF は Classifier の略で SVM もしくは ANN を指す. また, Weight は学習フェーズで得られたターゲット被験者のキャリブレーションデータに対するソース識別器の識別精度 (Accuracy) を指し, 0 から 1 の値である. 文献 [10] より転載.

の 2 段階の統計検定を通じて, sEMG ベース動作認識に最も適した組み合わせを明らかにした.

3.3.1 SVM/ANN

図 3.4(A) および (B) に, SVM および ANN 条件における学習・テストのフローチャートを示す. これらの条件は, 転移学習を用いない場合の識別性能を示すために実施され, 被験者間識別のベースライン (基準) として用いられる. 最初にソース被験者ごとに個別の識別

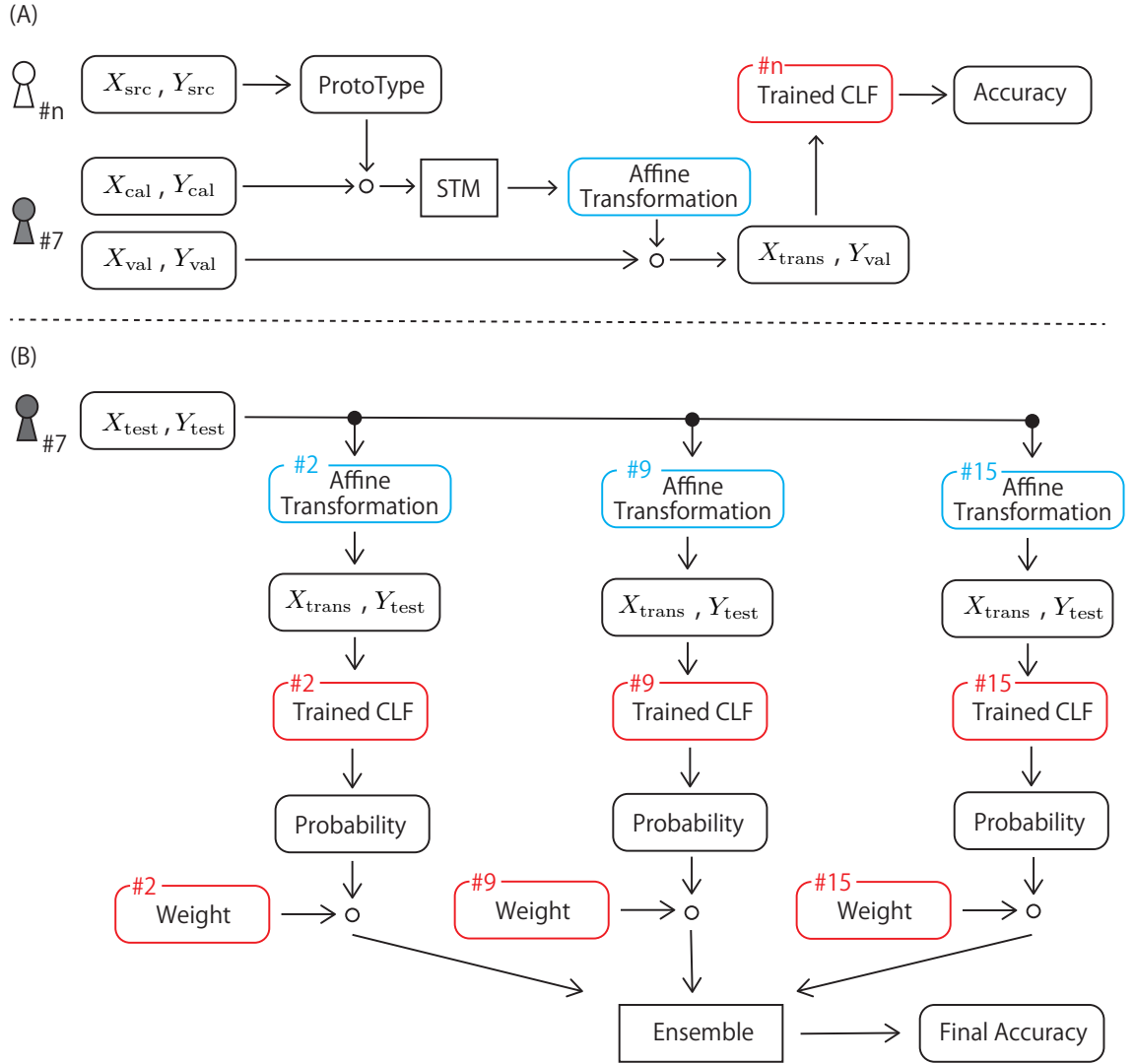


図 3.5: ターゲット被験者が被験者 7 のときの STM-SVM と STM-ANN における (A) 学習フェーズと (B) 推論フェーズのフローチャート. CLF は SVM もしくは ANN を指す. また, Weight は図 3.4(A) の学習フェーズで得られたターゲット被験者のキャリブレーションデータに対するソース識別器の識別精度を指し, 0 から 1 の値である. 文献 [10] より転載.

器 (SVM および ANN) を学習し, 5 分割交差検証を用いてハイパーパラメータの最適化を行った. 最適化後のソース識別器を用いてターゲット被験者のキャリブレーションデータを識別し, その結果の識別精度の高い順にソース識別器を選択した. その後, 選択されたソース識別器を用いてテストデータを識別し, アンサンブル処理による重み付けを行って動作ラベルを予測した.

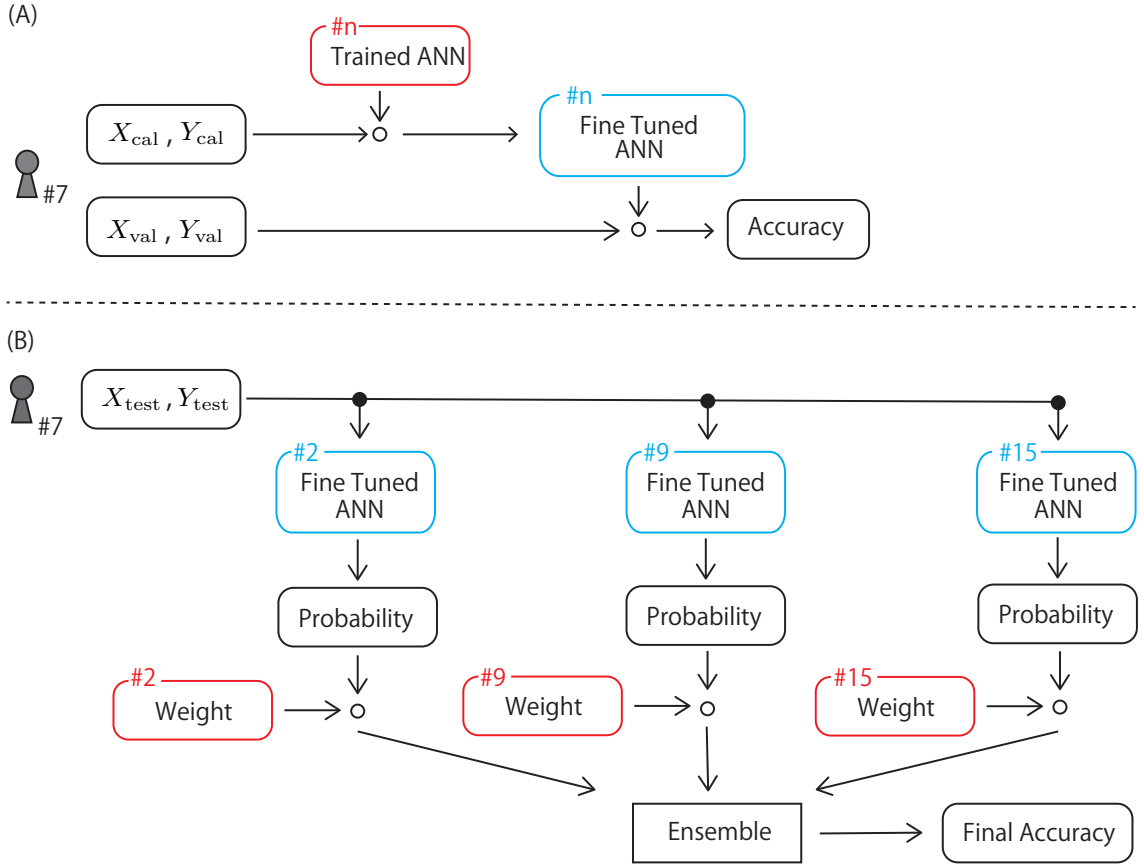


図 3.6: ターゲット被験者が被験者 7 のときの FT-ANN における (A) 学習フェーズと (B) 推論フェーズのフローチャート. また, Weight は図 3.4(A) の学習フェーズで得られたターゲット被験者のキャリブレーションデータに対するソース識別器の識別精度を指し, 0 から 1 の値である. 文献 [10] より転載.

3.3.2 STM-SVM/STM-ANN

図 3.5(A) および (B) に, STM-SVM および STM-ANN 条件における学習・テストのフローチャートを示す. SVM および ANN 条件と同様に, 個別のソース識別器を最適化し, ターゲット被験者用に上位 N 個の識別器を選択した. 次に, 選択した識別器ごとに STM のハイパーパラメータをターゲット被験者のキャリブレーションデータと検証データを用いて最適化した. テストデータは STM によりアフィン変換された. 最後に, 写像されたテストデータを選択した識別器で識別し, アンサンブル処理により重み付けを行って識別精度を評価した.

3.3.3 FT-ANN

図 3.6(A) および (B) に, FT-ANN 条件における学習・テストのフローチャートを示す. FT は深層学習特有の転移学習手法であるため, FT-SVM という条件は存在しない. SVM

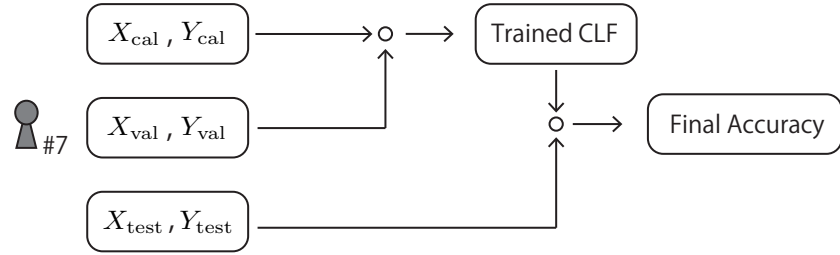


図 3.7: ターゲット被験者が被験者 7 のときの i-SVM と i-ANN における学習フェーズと推論フェーズのフローチャート. CLF は i-SVM もしくは i-ANN を指す. 文献 [10] より転載.

および ANN 条件と同様に, 個別の ANN を最適化し, ターゲット被験者用に上位 N 個の ANN を選択した. 選択された ANN の一部のパラメータをキャリブレーションデータと検証データで再学習する際, ハイパーパラメータである重み減衰を再度最適化した. 最終的に FT された ANN を用いてテストデータを識別し, アンサンブル処理による重み付けを行って動作ラベルを予測した.

3.3.4 個別 SVM/ANN (i-SVM/i-ANN)

図 3.7 に, 参照条件である i-SVM および i-ANN の学習・テストのフローチャートを示す. これらの条件では, 被験者内識別の一般的な手順に従い, ターゲット被験者のキャリブレーションデータと検証データのみを用いて識別器を学習・最適化した. その後, 最適化された識別器を用いてテストデータを識別した. 転移学習を行ったことでかえって精度が下がる, 負の転移が発生していない限り, 転移学習を用いた各条件の識別精度は, この個別学習条件の識別精度を上回ることが期待される.

3.4 結果

3.4.1 全ソース被験者を用いた場合の識別精度

図 3.8 は, 転移学習手法を適用した 5 条件および参照条件 (i-SVM および i-ANN) の, 25 名のターゲット被験者における平均識別精度を示している. 縦軸は 8 クラスの識別精度, 横軸はソース被験者数である. この図の実線上の各マーカーは, ターゲット被験者を 25 回繰り返し選択した際の平均識別精度を示している. 一方, 破線は参照条件 (i-SVM または i-ANN) における平均識別精度を示し, ターゲット被験者のキャリブレーションデータのみから得られたものであるためソース被験者数に依存せず一定の値となっている. 転移学習と組み合わせた識別器の精度が同色の破線を下回った場合, 転移学習手法が負の転移を引き起こしたことを意味する [75]. ソースデータのみを使用した従来の識別器 (すなわち, SVM および ANN) は, いずれも破線を下回っていた. STM と組み合わせた識別器 (STM-SVM および STM-ANN) については, ソースが 1 名の場合には i-SVM および i-ANN を上回ること

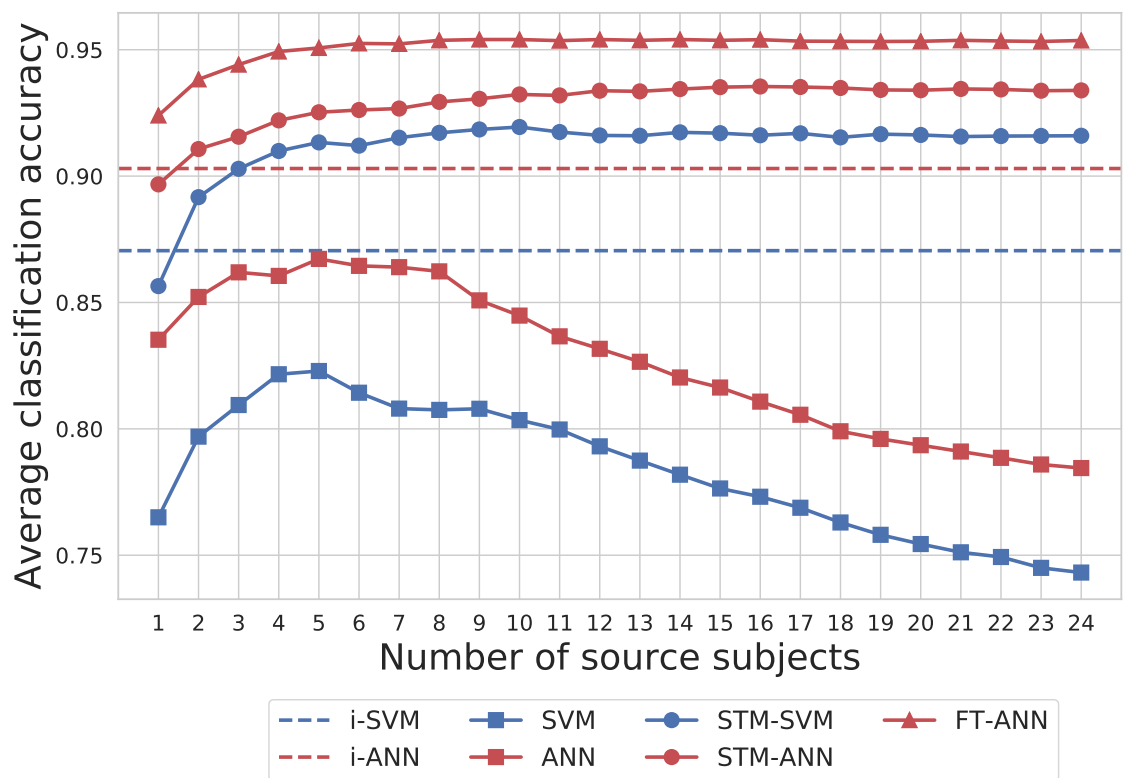


図 3.8: 転移学習手法を適用した5条件および参照条件 (i-SVM および i-ANN) の25名のターゲット被験者における平均識別精度. 文献 [10] より転載.

ができなかったが、ソース数が2名以上の場合には破線を上回る高い識別精度を示した。FTとANNを組み合わせた識別器（FT-ANN）もSTMと同様、i-ANNより高い識別精度を示したが、STMと異なり、ソース数が1名の場合から既に破線を上回っていた。識別精度が最大となったソース被験者数は、SVMでは5名（平均精度：0.8228）、ANNでは5名（平均精度：0.8672）、STM-SVMでは10名（平均精度：0.9194）、STM-ANNでは16名（平均精度：0.9354）、FT-ANNでは9名（平均精度：0.9540）であった。

3.3節で示した考え方にに基づき、どの識別器と転移学習手法の組み合わせが最も有効であるかを段階的な統計検定を用いて調査した。まず最初に、二元配置分散分析を行い、転移学習手法の有無（転移学習の適用あり・なし）および識別器（SVM・ANN）が識別精度に与える主効果および交互作用効果を検証した。その結果、転移学習手法群の主効果（ $F(1, 92) = 713.83$, $p < 0.001$ ）および識別器群の主効果（ $F(1, 92) = 49.65$, $p < 0.001$ ）は共に有意であった。また、交互作用効果についても有意な結果が得られた（ $F(1, 92) = 9.53$, $p = 0.0027$ ）。さらに、Tukeyの多重比較検定による事後解析の結果、STMを組み合わせた識別器は、転移学習を用いない識別器よりも有意に優れており（ $p < 0.001$ ）、ANNはSVMより有意に優れていることが明らかとなった（ $p < 0.001$ ）。また、STM-ANNは他の組み合わせ（SVM, ANN, および STM-SVM）よりも有意に高い性能を示した（SVMに対しては $p < 0.001$, ANNに対しては $p = 0.0312$, STM-SVMに対しては $p < 0.001$ ）。最後に、データ空間ベースの転移学習（STM）とパラメータ空間ベースの転移学習（FT）のどちらが優れているかを調べるために、STM-ANNとFT-ANNの識別精度について t 検定を実施した。その結果、FT-ANNがSTM-ANNよりも有意に優れていることが明らかとなった（ $p < 0.001$ ）。以上の結果から、本章で扱ったsEMGデータセットにおいては、FT-ANNが最も優れた手法であると結論付けられた。

3.4.2 特定のソース被験者数における識別精度

さらなる分析として、ソース被験者数を7名に設定した。7名という値は、本章と同一データセットを用いた先行研究 [9] や、異なるデータセットを用いた研究 [140, 79] においても良好な性能が報告されているため決定した。図 3.9 に、ソース被験者数を7名とした場合の各手法の平均識別精度を示す（SVM: 0.8080, ANN: 0.8639, STM-SVM: 0.9151, STM-ANN: 0.9267, FT-ANN: 0.9523）。各棒グラフは識別手法に対応し、破線は参照条件（i-SVM および i-ANN）の識別精度を示している。

次に、STM-ANNとFT-ANNに注目して、動作クラスごとの識別精度を調査した。図 3.10 にそれぞれの混同行列を示す。図中の上側の3つの混同行列はSTM-ANNに対応し、下側の3つはFT-ANNに対応している。左右の各列において、左側の上下の混同行列は全被験者（25名）の累積結果を示し、中央の混同行列は最良の被験者の結果、右側の混同行列は最も識別精度が低かった被験者の結果を示している。この分析においては、参照条件（i-ANN）の識別精度が最も高かった被験者（精度 0.9887）を最良、最も低かった被験者（精度 0.6121）を最悪と定義した。両手法において、最良結果は被験者 5、最悪結果は被験者 22 であった。左側の累積混同行列を見ると、STM-ANN、FT-ANNのいずれの場合も、クラス 1~5（手

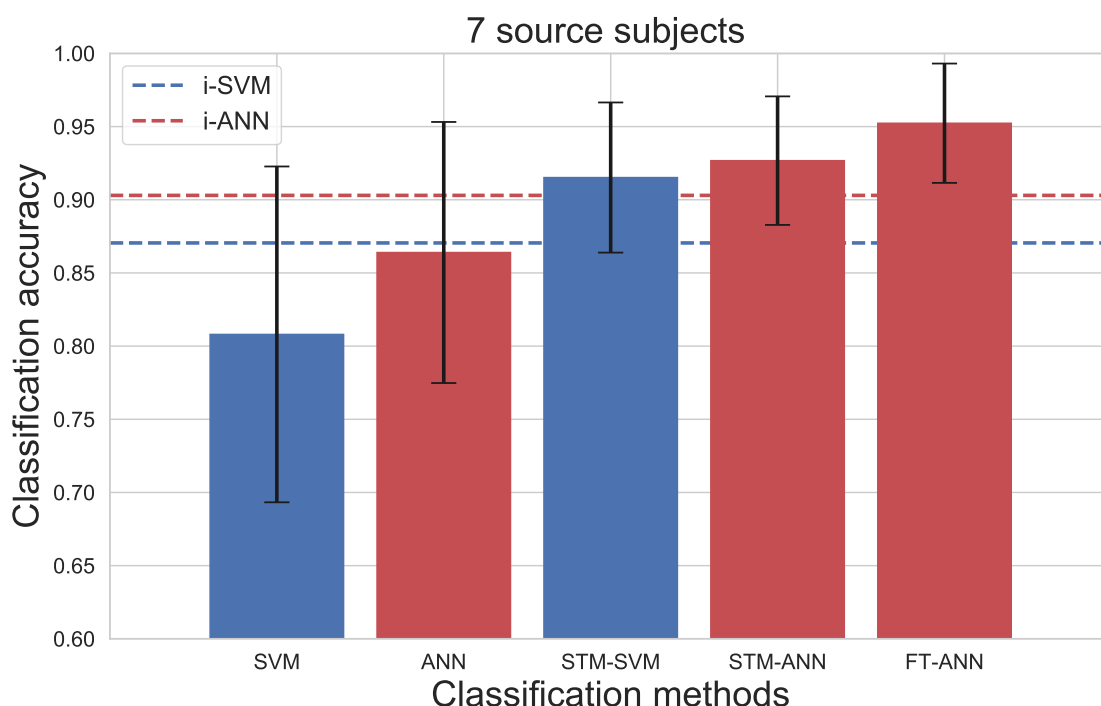


図 3.9: ソース被験者数を 7 名とした場合の各手法の平均識別精度。文献 [10] より転載。

開き、手首の屈曲、手首の伸展、手首の橈屈、および手首の尺屈）についてはほぼすべて正答であった。一方で、クラス 6 およびクラス 7（前腕の回内および前腕の回外）は比較的推定が難しく、特に対となる逆方向の動作が誤認されることが多かった。また、クラス 8（手握り）も一部異なるクラスとして誤認識されるケースが見られた。クラス 6 から 8 の識別精度については、STM-ANN と比較して FT-ANN の方がより高い精度を示していた。中央の混同行列に示した最良被験者の結果を見ると、両手法ともに非常に高い識別精度を達成していた。STM-ANN は完全に正しく識別されたのに対し、FT-ANN ではクラス 6 にわずかな誤りが存在した。右側の混同行列に示した最悪被験者の結果においては、どちらの転移学習手法も識別精度を約 20 % 改善したが、それでもクラス 6 から 8 には依然として誤認識が多く残っていた。

混同行列（図 3.10）では、各試行の方向に関する情報が集約されているが、本章で用いた sEMG データセットでは各動作が連続的に繰り返されている（例：クラス 1-1-1-1-2-2-2-...）。これは、同一クラスのデータであっても、後の試行ではデータ品質が初期の試行とは異なる可能性を示唆している。この影響を評価するために、各動作の試行ごとの精度を調査した。図 3.11 は、被験者 25 名分の各クラスにおける試行ごとの STM-ANN および FT-ANN の識別精度を示したものである。一部のクラスでは試行の影響を受けているように見受けられたが、すべての課題を通じて一貫した傾向はなく、転移学習手法（STM か FT か）による差異も確認されなかった。図 3.12 は、最良被験者と最悪被験者に焦点を当てた詳細な識別結果を示している。この結果からも、試行の経過による誤認識の増加傾向は確認されず、STM と FT の間でも同様の傾向が観察された。以上のことから、本実験において、STM および

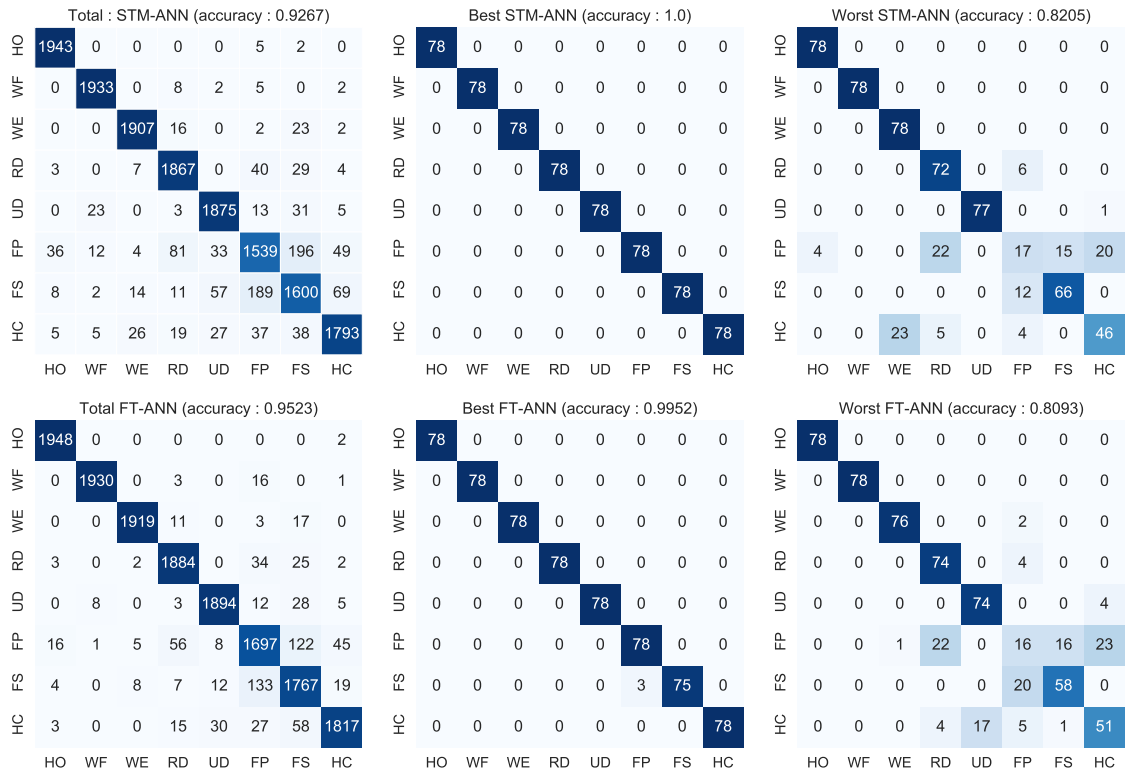


図 3.10: 8 クラス動作識別における混同行列. 縦軸が真のラベル, 横軸が予測ラベルを意味している. 上段 3 つが STM-ANN, 下段 3 つが FT-ANN の結果. 左が 25 名分の総合的な結果, 中央が最良の被験者 (被験者 5) の結果, 右が最悪の被験者 (被験者 22) の結果. 文献 [10] より転載.

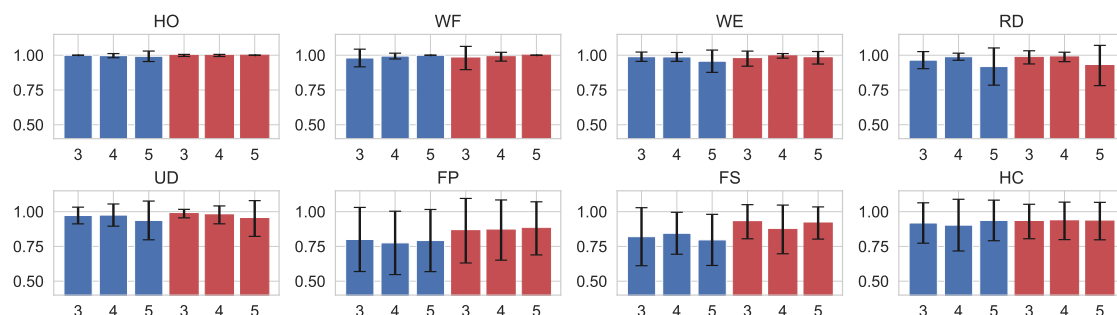


図 3.11: 被験者 25 名分の各クラスにおける試行ごと (3, 4, 5 番目の試行) の STM-ANN および FT-ANN の平均識別精度. 青いバーが STM-ANN, 赤いバーが FT-ANN を意味する. 文献 [10] より転載.

FT はいずれも同日内の短期間試行間でのデータ品質の変化に影響されなかったと結論づけられる.

3.4.3 転移学習手法と識別器の追加組み合わせの比較

本章では, 代表的な識別器として SVM と ANN を, 転移学習手法として STM と FT を選択し, その比較検証を行った. これは, これらが sEMG や EEG などの生体信号の識別において実用的であり, 比較において妥当性が高いと判断したためである [110, 111, 123, 141, 140, 144, 142]. しかし, 識別器と転移学習手法の組み合わせは他にも多数考えられ, それらによる性能差を検討することには, 選択した手法の正当性を確認する上で学術的意義がある. そこで本節では, 追加の比較として 1 種類の識別器 (LDA) と 3 種類の転移学習手法 (CSA, TCA, CORAL) を導入し, 以下の 8 通りの組み合わせと 1 つの参照条件を追加の評価対象とした.

1. CSA-LDA
2. STM-LDA
3. TCA-LDA
4. CORAL-LDA
5. TCA-SVM
6. CORAL-SVM
7. TCA-ANN
8. CORAL-ANN
9. i-LDA

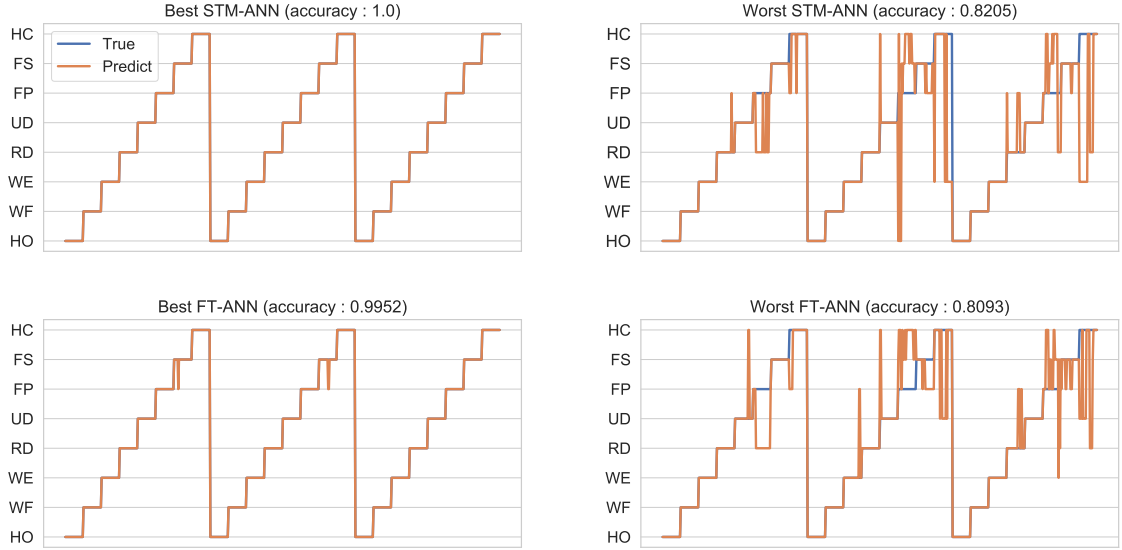


図 3.12: 3 回のテスト試行における真のラベル系列と，STM-ANN(上) および FT-ANN(下) によって予測されたラベル系列．最良 (左) および最悪 (右) の結果は，図 3.10 と同じ被験者から得られた．文献 [10] より転載．

これらの手法を選定した理由は次のとおりである．LDA は sEMG と EEG の両分野で識別で広く利用される線形識別器の一つである [155, 156]．この LDA を追加することで非線形な識別器に焦点を当てた本章の主要な分析に対し，線形モデルとしてのベースラインを提供する．TCA と CORAL は，STM と同様にデータ空間ベース転移学習手法であるが，写像のアプローチがそれぞれ異なる手法である．STM では，ターゲットをソースから計算されたプロトタイプに向かって写像させる．対して TCA は，ソースとターゲットの分布間距離を最小化するように，両者を共通の潜在部分空間へ写像する [133]．一方，CORAL は，ソースとターゲットの統計的性質の差異を最小化するように，ソースをターゲットの分布に向かって写像する [134]．ただし，TCA や CORAL は，ターゲットデータに合わせて特徴空間そのものを変換するため，変換後のデータを用いて識別器を学習する必要がある．そのため，ソース被験者の事前学習済みモデルをそのまま適用することはできないが，STM と同様のアプローチをとる既存手法として重要であるため，これらを性能評価のベースラインとして採用した．一方，CSA はパラメータ空間に基づく転移学習手法であり，ソースとターゲットのそれぞれのドメインの統計的なパラメータを重み付けをして統合することで，ターゲットドメインに適合した新たな識別モデル（決定境界）を構築する [138]．具体的には，平均 μ と共分散 Σ のみでモデルが記述される LDA のパラメトリックな性質を利用し，ソースとターゲットの統計量を線形的に統合することでパラメータ適応を実現する．これにより，勾配法によるパラメータの再最適化を行う FT に対し，統計量の単純な線形結合というより簡便なアプローチが，生体信号の個人差補正においてどの程度有効であるかを比較検証する．なお，TCA における RBF カーネルのハイパーパラメータ μ および σ は， 10^{-3} ， 10^{-2} ， $\dots$ ， 10^3 の範囲で最適化した．また，CSA のハイパーパラメータ τ および λ は，0.0，

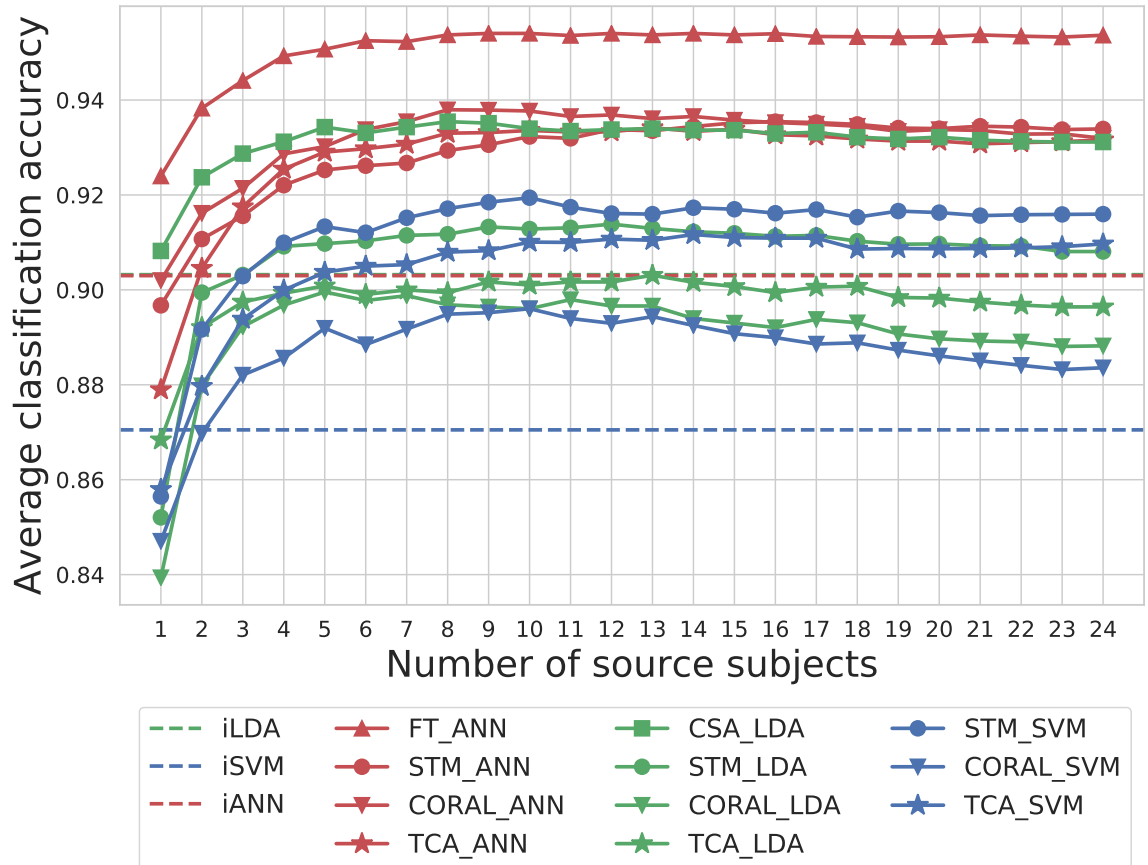


図 3.13: 追加の転移学習と識別器を含めた平均識別精度. 文献 [10] より転載.

0.1, ..., 1.0 の範囲で最適化した.

図 3.13 は, 3 種類の識別器 (LDA, SVM, ANN) および 5 種類の転移学習手法 (FT, STM, CORAL, TCA, CSA) を用いた全体的な識別結果を示している. 同図において, i-LDA の精度曲線は i-ANN の曲線とほぼ重なっており, その精度は 0.9032 であった. この値は i-SVM および i-ANN (それぞれ 0.8705 および 0.9030) の精度よりも高い. LDA を用いた手法群では, パラメータ空間に基づく転移学習手法である CSA-LDA が最も高い識別精度を示した. データ空間に基づく転移学習手法を LDA と組み合わせた場合, 精度は STM-LDA, TCA-LDA, CORAL-LDA の順に高かった. SVM を用いた手法群では, 組み合わせ可能な有効なパラメータ空間ベースの転移学習手法が存在しないため, データ空間ベースの手法のみを比較した. 精度は STM-SVM, TCA-SVM, CORAL-SVM の順に高かった. なお, TCA-SVM を適用する際には, TCA と RBF カーネル SVM の組み合わせが非常に相性が悪く (平均精度は約 0.30), 本研究では線形カーネル SVM を採用した点に注意が必要である. ANN を用いた手法群では, パラメータ空間ベースの転移学習手法である FT-ANN が, 識別器と転移学習手法の全組み合わせの中で最も高い識別精度を示した. ANN とデータ空間ベースの転移学習手法を組み合わせた場合, ソース被験者数が約 15 までは CORAL-ANN, STM-ANN, TCA-ANN の順に高い精度を示した. しかし, ソース被験者数が 15 を超える

条件では、これらの手法間で精度の差はほとんど認められなかった。以上の追加検証の結果からも、パラメータ空間に基づく転移学習手法と深層識別器の組み合わせ（FT-ANN）が最も優れているという本章の主要な分析結果が支持された。

3.5 考察

3.5.1 識別性能

i-SVM と i-ANN の識別精度はそれぞれ 0.8705 および 0.9030(図 3.8 参照) であり、これらは過去の研究で報告されている識別精度よりも低いものであった。本章では、時間領域および自己回帰モデルの特徴量に基づき SVM および ANN を構築し、8 つの電極ペアを用いて 25 名の参加者の 8 クラス前腕動作（手開き、手首の屈曲、手首の伸展、手首の橈屈、手首の尺屈、前腕の回内、前腕の回外、および手握り）の認識を行った。SVM を用いた研究としては、Lucas ら [157] は時間-周波数領域の特徴量を用いて 6 クラスの手指動作（手首の屈曲、手首の伸展、前腕の回外、前腕の回内、手開き、および手握り）の識別を行い、平均精度 0.953 を報告している。また、Alkan ら [158] は時間領域特徴量を用いて 4 クラスの前腕動作（肘の屈曲、肘の伸展、前腕の回内、および前腕の回外）の識別で平均精度 0.99 を報告している。ANN を用いた研究では、Amatanon ら [159] は時間領域特徴量を用いて 10 クラスのタイ語手話識別を行い、平均精度 0.979 を達成しており、Zhang ら [160] は時間領域特徴量を用いて 5 クラスの手指動作（握りこぶし、内側振り、外側振り、指開き、およびダブルタップ）の識別で平均精度 0.987 を達成している。これらの先行研究をまとめると、本章で扱った動作タスクについても 0.95 を超える識別精度が期待される。しかし、本章で得られた識別精度はその期待値よりも低かった。その主な原因は、基準となる識別器の訓練データ数の違いに起因すると考えられる。上記の過去の研究では、一般に多くの訓練データ（100 試行以上）を使用しているのに対し、本章では識別器の構築および最適化に用いたデータが各クラスにつき 1 試行分の訓練データおよび 1 試行分の検証用データのみであった。このようなデータ数の違いによる精度の差は想定されるものであるが、本章で示した精度（0.8705 および 0.9030）が、ウェアラブルセンサを用いて収録されたとく限られた（1 試行分の）トレーニング・検証データのみでも達成可能であることには意義がある。

また、本章の sEMG データセットには個人間の大きな差異が存在していることが確認された（図 3.8 の左端の青および赤の四角マーカーの結果を参照）。もし個人差が小さければ、ソースデータとターゲットデータは共通のデータ分布を有しており、単一ソースを用いた SVM および ANN の精度（図の左端の青および赤の四角マーカー）が破線の精度に近づくはずである。しかし、実際にはこれらの精度は破線の精度を下回った（図 3.8 の四角マーカー線と破線の組み合わせを参照）。この精度の低下は、sEMG データにおける個人差の大きさを裏付けるものであり [161, 162]、ターゲット被験者に類似したソース識別器を追加することで改善された。アンサンブル戦略は、このような個人差の影響を軽減することに寄与した。また、興味深いことに、ソース識別器数が過度に増加すると、アンサンブル戦略の効果は減少し、逆に精度が低下する場合があった。この現象は、ターゲット被験者と異なるデータを

持つソース識別器からの誤った投票が増加するために起こったと考えられる。このため、転移学習手法を用いずに過剰な数のソース識別器を使用することは負の転移を引き起こし、識別性能を悪化させることにつながると考えられる。

転移学習は、各ソース識別器の利点を最大限に活用できるように、個人間の変動を減少させることに貢献した。その結果、識別性能が向上し、同時にソース識別器数の増加に伴う精度低下を回避できた。STM と FT いずれの場合も、ソース識別器数が増える初期段階では SVM および ANN 同様に識別精度が向上し、その後も安定した性能を示した。これはターゲットとソースの間に共通した関係性が存在するという仮定に基づいている。STM と FT はターゲットとソースの間で情報を伝達するために異なるアプローチを用いている。STM アルゴリズムは線形写像を用いているため、その性能はターゲットとソースデータ間の相関に依存する。この相関の度合いは、被験者間変動の抑制や sEMG ベース動作認識の精度に影響を与えることが報告されている [163]。本章で用いたデータセットは単一自由度の単純な動作から構成されている。各被験者の動作強度 (例えば 1 回目は 50% の力で動作を発現し、2 回目は 70% で発現している、など) の変動は不明であるが、ソース候補者の動作のバリエーション全体を考えると、データ間の相関が比較的高くなることが予想される。この状況において STM は良好に機能し、ソース識別器数が増加しても安定した性能を示した。一方、FT の目的は、ソースデータによって学習済みのパラメータを用いて第 3 隠れ層から出力層間のパラメータを再学習することである。この方法は、たとえデータに大きな個人差があったとしても、特徴空間におけるクラス間のデータ構造は同一であり、ターゲットとソースの間で構造的な情報を共有できるという考えに基づいている。sEMG データの場合、筋肉の特性やその他の身体的要因によりデータに大きな個人差があるものの、各動作クラスにおける筋収縮パターン自体は類似しているはずである [164]。実際、本章において FT-ANN は最も高い識別精度を示した。上述した過去の研究 [157, 158, 159] と本章の結果を直接比較することは、使用されている特徴量、識別器、データセットなどが異なるため難しい。しかし重要な点は、本章での FT-ANN はターゲットユーザからわずか 1 試行分のデータしか用いていないにもかかわらず、8 クラス識別において 0.95 を超える精度を達成できることである。さらに FT-ANN の性能は、より高度な深層識別器を利用することで、さらに向上する可能性がある [165, 166]。Duan ら [165] は周波数領域特徴量を用いた CNN を提案し、10 クラスの手指ジェスチャー (親指を上げる、手を閉じる、手を開く、手を伸ばす、手を曲げる、人差し指と親指でリングを作る、中指と親指でリングを作る、薬指と親指でリングを作る、小指と親指でリングを作る、およびすべての指をまとめる) 識別で平均精度 0.94 を報告している。また Oh ら [166] は、スケール平均ウェーブレット変換に基づく新たな周波数領域特徴量を用いた CNN を提案し、6 クラスの手指ジェスチャー (つかむ、物をつまむ、カードを持つ、OK サイン、V サイン、および親指を上げる) 識別で平均精度 0.939 を報告している。本章では、手作業で設計された特徴量を用いる ANN によって良好な結果が得られたが、CNN によって得られる学習済み特徴量に基づいた識別器を用いることも有効な手段であると考えられる。

3.5.2 ソース識別器の数

アンサンブル戦略で使用するソース識別器の数は、ロバストな動作認識の性能に影響を与えるハイパーパラメータである。本章の結果から、転移学習手法を用いることで、ソース識別器の数を最大限に増やしても安定した識別精度が得られることが明らかになった。しかしながら、ソース識別器の数が増えると計算資源が多く必要となる。ここで、識別器の数を7個程度（またはそれ以上）に設定することは、シンプルかつ効果的なアプローチとなる。STM および FT いずれにおいても、ソース識別器の数が7個以上になると、識別精度にわずかな差しか見られなかった（図 3.8 参照）。最適なソース数を用いた場合と7個のソースを用いた場合の識別精度を列挙すると、それぞれ STM-SVM では 0.9194 と 0.9151, STM-ANN では 0.9354 と 0.9267, FT-ANN では 0.9540 と 0.9523 であった。これらの結果に加え、EEG 識別においても、Wei ら [79] や Li ら [140] が同様の結果を報告しており、ソースデータの数を多く設定すると識別精度がより安定することが示されている。ソース識別器数が「7」という数値は、EEG および sEMG データにおいて転換点である可能性がある。ただし、上記の研究結果は限られたデータセットに基づくものであるため、この数値はあくまで最低限の人数として考えることが望ましい。

3.5.3 データ空間ベースとパラメータ空間ベースの比較

二元配置分散分析の結果から示されたように、データ空間ベースの転移学習を深層識別器と組み合わせた場合（STM-ANN）は、浅層識別器と組み合わせた場合（STM-SVM）よりも識別性能が改善された。したがって、データ空間ベースおよびパラメータ空間ベースのどちらの転移学習手法においても、浅層識別器よりも深層識別器の方が適していると考えられる。

また、このパラメータ空間ベースの転移学習の優位性は、本章で主に取り上げた SVM や ANN だけでなく、追加の比較検証（3.4.3 節）で実施した LDA を用いた結果においても確認された。図 3.13 に示した通り、LDA 分類器においても、パラメータ空間ベースの手法である CSA は、データ空間ベースの手法（STM, TCA, CORAL）よりも一貫して高い精度を示した。SVM 分類器に対してパラメータ空間ベースの転移学習手法を適用した比較は、筋電位解析において当該手法が存在しないため実施できなかったが、文献 [167] のような手法を導入することで、同様の傾向が得られる可能性がある。i-ANN と i-LDA の識別精度はほぼ同等（それぞれ 0.9030 および 0.9032）であったものの、転移学習手法と組み合わせた場合には、ANN の方が LDA よりも強力なバックボーン識別器として機能した。これは、深層識別器の方が浅層識別器よりも転移学習との組み合わせに適していることを支持するものである。

sEMG データに対する転移学習の本質は、ターゲットとソースのデータ間に存在する非線形性を橋渡しする能力である。3.5.1 項で述べた通り、STM と FT はターゲットとソース間の情報を橋渡しする際に異なる戦略を採用している。ただし、STM と FT はともに線形変換を用いている。STM はシンプルな線形写像を実行し、FT も特徴空間へのマッピングを除

いて、線形変換のみを行っている。本章では、FT アルゴリズムが最後のパラメータのみを再学習しているため、再学習されたパラメータを使用する場合、中間表現を線形に転移していることになる。STM と FT は線形変換を共通して利用しているが、非線形の特徴抽出の恩恵を得ている FT が STM よりも優れていることが明らかとなった。このため、STM の性能は、非線形表現を獲得する機能を追加することでさらに改善される可能性がある [168]。本研究では、この課題に対し、非線形性を導入した拡張型 STM を独自に開発した。その数理的詳細および実験の結果については、付録 B に記述する。また、行列因子分解や特徴抽出を利用する方法 [169] を適用するなど、STM の性能を改善するための別のアプローチも考えられる。一方で FT についても、ANN の代わりに CNN [170] や RNN [171] など、より強力な深層特徴抽出器を採用することで、転移性能のさらなる向上が期待できる。

3.5.4 研究の限界

本章では、転移学習手法と識別器の関係を検討する際に、単一のデータセットのみを使用した。転移学習を採用した動機は、ウェアラブルセンサを用いた計測環境において、被験者間の変動性を軽減することである。多数の被験者からデータを収集することで、実験データセットには大きな被験者間変動が含まれており、転移学習手法の性能を検証する上で適した環境が整っている。本章で用いたデータセットは、他の公開データセットと比較して、比較的多い被験者数（25 名）を含んでいるという利点を持つ。例えば、sEMG ベース動作認識で広く使われる NinaPro データベースの DB5 [12] は、健常者 10 名のデータのみを使用している。しかしながら、オープンなデータセットを用いて性能を検証することは、研究結果の妥当性や標準化を促進する上で重要である。また、sEMG は被験者グループの種類や年齢（例えば、健常者や切断者、若年者や高齢者など）によって大きく変動することが知られているため [172, 173]、これらの多様な被験者群を網羅的にカバーするデータ収集またはデータセットの選択を行うことが望ましい。さらに、もう一つの限界として、本章では各被験者から収集されたデータが 1 日のみである点が挙げられる。各被験者が複数日にわたって提供したデータに対しても転移学習が有効に機能するかを検証することも有益だろう。

3.5.5 今後の課題

近年、ANN 以外の深層学習アーキテクチャ（例えば CNN や RNN）を用いた転移学習手法がいくつか提案されている [174]。本章では、浅層識別器と深層識別器の比較に焦点を置くことが重要であったため、これらの高度な深層アーキテクチャについては考慮しなかった。しかし、将来的には CNN や RNN などのアーキテクチャを含め、転移学習手法を併せて比較検討することを計画している。

また、これまでの研究と同様に、本章ではデータ空間ベースの転移学習手法およびパラメータ空間ベースの転移学習手法をそれぞれ独立に適用し、どちらが sEMG ベース動作認識において優れているかを議論した。しかし、これらを組み合わせることでさらに性能を向上させる可能性がある。例えば、まず STM によってターゲットデータをソース識別器に合

うように変換した後、変換したデータを正確に識別するために識別器の FT を行う方法が考えられる。このようなデータ空間ベースおよびパラメータ空間ベースのハイブリッド型の転移学習アプローチは、ソースとターゲット間のギャップをより効果的に埋める可能性がある。今後は、このようなハイブリッドアプローチが sEMG 識別問題に及ぼす影響を検討していく予定である。さらに、STM における線形変換関数を非線形関数に置き換えることで、ターゲットデータをソースデータにより類似した形式に変換できる可能性も検討する。なお、この非線形化への第一歩として、カーネル法を用いた非線形 STM (kernel-based STM: k-STM) を試作し、その有効性を予備的に検証した。この手法の数理的定式化および初期評価の結果については、付録 B に詳述する。

3.6 結言

本章では、8 種類の前腕動作を対象とした sEMG の被験者間識別問題において、浅層および深層の識別器それぞれに、データ空間ベースおよびパラメータ空間ベースの転移学習を適用し、識別精度の相対的な比較検討を行った。25 名の被験者から得られたデータを用いて個別の識別器を学習し、アンサンブル戦略を用いて被験者間識別を実施した。その結果、浅層識別器 (SVM) および深層識別器 (ANN) いずれにおいても、転移学習を利用することにより識別精度が大幅に向上し、さらにソース識別器の数を増やすことで識別の頑健性が得られることが確認された。特に、パラメータ空間ベースの転移学習 (FT) と深層識別器 (ANN) の組み合わせが最も高い精度を示した。これらの結果から、パラメータ空間ベースの転移学習と深層識別器を組み合わせる方法が、個人差に起因する複雑な分布変動に対して最も適応能力が高く、かつ実用的なアプローチであると考えられる。したがって、本章ではこの「深層識別器とパラメータ空間ベース転移学習」の構成を、標準的な技術基盤として結論づける。

第4章 複数被験者データの統合戦略に関する比較検討

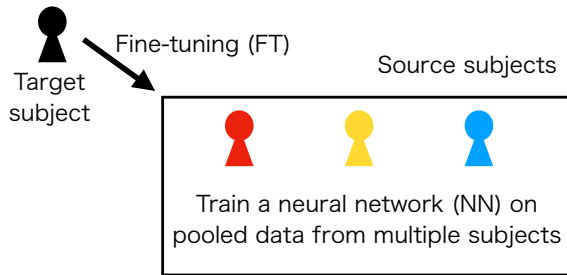
4.1 緒言

前章では、生体信号の個人差問題を解決するための技術的基盤として、sEMG のデータセットを利用して機械学習モデルと転移学習手法の最適な組み合わせを探索した。その結果、深層識別器 (ANN) に対し、パラメータ空間ベースの転移学習 (FT) を適用する手法が、他の組み合わせと比較して最も有効であることを示した。したがって、今後の開発においてはこの組み合わせが、実用化に向けた第一選択肢として位置づけられる。この組み合わせに関しては、既に多くの実践例が存在する [175, 176, 177, 178, 179, 180, 142, 10, 181]。しかし、事前に計測されている複数のソースユーザのデータを、FT のプロセスにおいてどのように統合・活用するのかという方法論に関しては、未だ統一された見解が得られておらず、大きく分けて二つの対照的なアプローチが混在しているのが現状である。

これら二つのアプローチの差異は、図 4.1 のように、本質的には「複数被験者の情報をどの段階で統合するか」という統合タイミングの相違に帰着する。第一のアプローチは、全ソースユーザのデータを一つの大規模なデータセットとして一元的に統合し、単一の汎用的なバックボーンモデルを学習させた後に、ターゲットデータで FT を行うアプローチである [175, 176, 177, 178, 179, 180]。本章では、これを「シングルソースアプローチ」と呼称する。第二のアプローチは、ソースユーザごとに個別のバックボーンモデルを学習し、それぞれに FT を適用した後に、それらの出力をアンサンブル学習によって統合するアプローチである [142, 10, 181]。本章では、これを「マルチソースアプローチ」と呼称する。これら二つのアプローチは、それぞれの研究において有効性が報告されているが、表 4.1 に示すようにそれぞれ独立して研究が進められ、体系的な比較がされていない。

現状、これらのアプローチの選択は、明確な理論的根拠に基づくというよりは、研究者や開発者の経験則や慣習に依拠して暗黙的に決定される傾向にある。しかし、両者はそれぞれ構造的に異なる利点と欠点を持つと考えられ、どちらのアプローチが適切かはデータセットの特性に依存する可能性がある。シングルソースアプローチの主な利点は、あらゆる被験者に対して適用可能な汎用的なバックボーンモデルが得られることである。その反面、訓練データセットの分散が大きくデータ量が不十分な場合には、表面的な平均的な特徴しか学習できず、本来は識別に有効な信号特性の学習が進まない可能性がある。一方、マルチソースアプローチの主な利点は、各ソース-ターゲットペアに対して個別に FT を行うことで、個人差を効果的に緩和できる点にある。その反面、同一のターゲット被験者のデータを用いて FT を行うため（すなわち、FT 後のモデルが互いに類似するため）、アンサンブル戦略に

(a) Single-source approach



(b) Multi-source approach

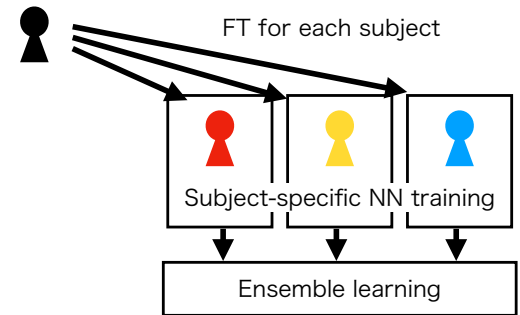


図 4.1: シングルソースアプローチとマルチソースアプローチの概念図. (a) シングルソースアプローチでは, 複数のソース被験者のデータを事前に統合し, それを学習したニューラルネットワーク (NN) をターゲット被験者に対して微調整 (FT) する. (b) マルチソースアプローチでは, ソース被験者ごと NN を学習し, 個別に FT し, それらをアンサンブル学習により統合する.

表 4.1: 先行研究と本章の比較概要. アンサンブル戦略の列の略字 (HV, SV, WV, および SWV) の詳細は 4.2.5 項で説明される. 文献 [13] より転載して一部改変.

文献	発行年	FT アプローチ	アンサンブル戦略	評価に使用されたデータセット数
[175]	2020	シングル	N/A	4
[176]	2021	シングル	N/A	1
[177]	2022	シングル	N/A	3
[178]	2023	シングル	N/A	1
[179]	2023	シングル	N/A	1
[180]	2023	シングル	N/A	3
[142]	2019	マルチ	HV	2
[10]	2022	マルチ	WV	1
[181]	2022	マルチ	HV	4
本章	2024	シングル, マルチ	HV, SV, WV, SWV	3

よるメリットが低下する可能性がある。こうした性質を考慮すると、どちらのアプローチが適切かは、対象とするデータセットの特性に強く依存すると考えられる。具体的には、被験者数が多く個人差が少ないデータセットに対してはシングルソースアプローチが適し、逆に被験者数が少なく個人差が大きい場合にはマルチソースアプローチが優位になると推察される。本章では、この仮説を検証するため、被験者数や計測機器特性、タスク難易度といった特性の異なる複数のデータセットを用いて、アプローチ間の相違を体系的に調査する。

そこで本章では、第3章で確立した FT-ANN を共通の技術基盤とし、複数被験者データの統合戦略として、シングルソースアプローチとマルチソースアプローチのどちらが適しているかを体系的に比較検討する。検証にあたっては、結果の汎用性を担保するため、特性が異なる3つのオープンデータセット（MyoDataset [9], NinaPro DB5 [12], NinaPro DB2 [63]）を使用する。また、マルチソースアプローチにおける複数の FT 済みモデルの統合手法（アンサンブル戦略）の選択が性能に与える影響を調査するため、4種類の異なるアンサンブル戦略を用意した。本章の検討を通じ、データセットの特性に応じた最適なデータ統合ガイドラインを導出することを目的とした。その結果、データセットの特性に依らずマルチソースアプローチはシングルソースアプローチよりも高い性能を示すことが明らかとなった。また、適切なアンサンブル戦略を選択することでより高い性能を実現でき、中でも Soft Voting という手法が実装が容易ながらも安定して高い性能を示すことが明らかとなった。よって本章では、パラメータ空間ベースの転移学習と深層識別器に対しては、マルチソースアプローチを基本戦略とし、その統合手法として Soft Voting を採用する構成を、最も実用的かつ高精度なデータ統合ガイドラインとして提案する。

4.2 データセット、解析手法、および評価方法

4.2.1 データセット

本章では、シングルソースおよびマルチソースアプローチの認識性能および FT の影響を調査するために、公開されている3つの sEMG データセットである MyoDataset [9], NinaPro DB5 [12], および NinaPro DB2 [63] を使用した。MyoDataset は第3章で用いたデータセットである。これらのデータセットはすべて、日常生活環境で容易に使用可能なウェアラブルかつワイヤレスの sEMG デバイスによって収録されたものである。各データセットに対して、元論文の設定 [9, 12, 63] を再現した信号の前処理および特徴抽出を行った。本章でこれらのデータセットを選定した理由は次の通りである。(1) いずれもウェアラブルかつワイヤレスのデバイスで記録されていること、(2) サンプリングレートや認識難易度などの特性がそれぞれ異なっていることである。MyoDataset と NinaPro DB5 は同一デバイスを使用している。第1の要素は、実環境における HCI の開発において重要である。第2の要素は、さまざまな計測条件におけるシングルソース・マルチソースアプローチの強みと弱みを明確化するのに役立つ。

(1) MyoDataset

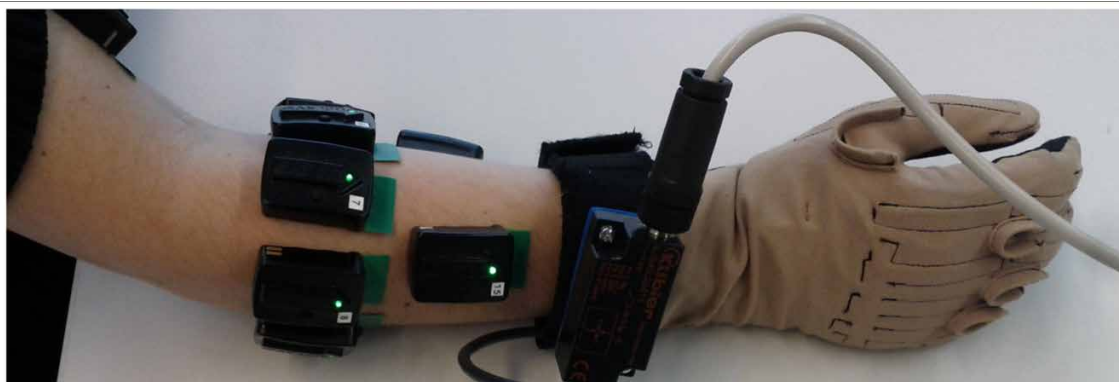
基本的な内容は3.2.1項で記述済みである。25名分の健常者を対象とし、単純な前腕動作8種類を5回ずつ計測したsEMGデータを収録している。MyoDatasetは単純な前腕動作課題で構成されているため、このデータセットのパターン認識の難易度は比較的低いと考えられる。また、収集された信号のサンプリングレートは比較的低く、被験者数は比較的多い。このため、多くの被験者を含み、個人差が比較的小さいシングルソースデータセットを構築可能であることから、本データセットにおいてはシングルソースアプローチの方が高い性能を示すことが予想された。一方で、マルチソースアプローチはFTされた各モデル間の多様性が失われるため、性能が低下すると考えられた。

(2) NinaPro DB5

このデータセットはNinawebチームによって公開されたもので、22歳から34歳の健常者10名（男性8名、女性2名）による前腕の運動およびジェスチャ52クラスのsEMGデータから構成されている[12]。各被験者は、エクササイズと呼ばれる前腕の運動およびジェスチャのまとまりであるエクササイズA・B・Cの3つから成る52種類の動作およびジェスチャを実施したが、本章では先行研究[12]の設定に従って、エクササイズBおよびCのみを認識課題として用いた(図4.2)。エクササイズBは8種類の等尺性および等張性の手のジェスチャと、9種類の基本的な手首の動作から構成されている。エクササイズCは23種類の把持および機能的動作から構成されている。したがって、本章のパターン認識課題には計40種類の動作およびジェスチャを用いた。各動作は各被験者によって5s間ずつ6回試行された。

sEMG信号の計測にはMyoジェスチャーコントロールアームバンドが右前腕の肘付近に2つ装着され、16チャンネルのsEMG信号が200Hzのサンプリング周波数で収集された。まず上側の計測機器が、第1チャンネルが腕橈関節付近に位置するように配置された。その後、下側の計測機器は、上側すぐ下に装着し、22.5°回転させて配置することで、下側が上側の隙間を埋めるように配置された(図4.3(d)参照)。本データセットは、第3.2.1項で述べたMyoDatasetと同じバンド型の計測機器を使用しており、解剖学的な知見に基づいた配置ではない点に注意されたい。信号に対してフィルタリングは施されていない。また、記録された5s間の信号から、データ提供元が定義するrestimulus変数に基づいて、活動動作を含む数s間のセグメントを抽出した。

NinaPro DB5は複雑な手の動作課題から成っているため、このデータセットのパターン認識の難易度は比較的高いと考えられる。また、信号のサンプリング周波数は比較的低く、被験者数も比較的小さい。このため、シングルソースアプローチにおいては単一のデータセットが過度に複雑になり、性能の低下が予想された。一方、マルチソースアプローチは各ターゲットとソースのペアごとに個人差を緩和するため、このデータセットではシングルソースアプローチよりも優れた性能を示すことが期待された。



(b) Delsys Trigno setup



(d) Double Myo setup

図 4.3: NinaPro DB で使用されている計測機器の装着の様子. (b) NinaPro DB5 (d) NinaPro DB2. 文献 [12] から引用.

(3) NinaPro DB2

このデータセットは NinaPro DB5 と同様に Ninaweb チームによって公開されたものであり、23 歳から 45 歳の健常者 40 名（男性 28 名，女性 12 名）による前腕動作およびジェスチャ 49 クラスの sEMG データから構成されている [63]。各被験者は、エクササイズ B・C・D から成る 49 種類の動作およびジェスチャを実施したが、本章では先行研究 [12] の設定に従い、エクササイズ B および C のみを認識課題として用いた。このため、NinaPro DB5 と同様に、本章のパターン認識には計 40 種類の動作およびジェスチャを用いた。各動作は各被験者によって 5s 間ずつ 6 回試行された。

Trigno ワイヤレス電極 (Delsys Inc, 米国) 12 チャンネルを被験者の右前腕に専用の粘着バンドで固定し、12 チャンネルの sEMG 信号が 2000 Hz のサンプリング周波数で収集された。12 チャンネルのうち 8 チャンネルは、橈骨上腕関節と同じ高さの前腕周囲に、リング状に等間隔で装着された。残りの 4 チャンネルは、解剖学的な知見を元に精密に装着された。まず、2 チャンネルが浅指屈筋と指伸筋の主要な活動点に装着され、さらに残りの 2 チャンネルが上腕二頭筋と上腕三頭筋の主要な活動点に装着された。なお、活動点は触診によって特定した。記録された信号には、5 次バターワースフィルターにより 15~500 Hz の帯域通過フィルタリングが施された。また、記録された 5s 間のデータから、データ提供元が定義する `restimulus` 変数に基づいて、活動動作を含む数 s 間のセグメントを抽出した。

NinaPro DB2 は複雑な手の動作課題で構成されているため、このデータセットのパターン認識の難易度は比較的高いと考えられる。この点は、マルチソースアプローチが適しているように思われる。一方、信号は高いサンプリング周波数で記録され、かつ比較的多くの被験者から収録されているため、被験者数の観点からはシングルソースアプローチも適しているように思われる。このため、シングルソース・マルチソースの両アプローチは、本データセットにおいてほぼ同等の認識性能を示す可能性があると考えられた。

4.2.2 区間分割と特徴抽出

それぞれ抽出された動作の実行中の活動区間を解析窓に分割した。MyoDataset に対しては、窓長を 250 ms，スライド幅を 50 ms とした。NinaPro DB5 および DB2 については、窓長を 200 ms，スライド幅を 100 ms とした。

各データセットにおける結果を従来研究と比較可能にするため、データセットで異なる手動設計の特徴量を使用した [9, 12, 63]。MyoDataset では、チャンネルごとに以下の 11 個の特徴量を抽出した。これらは Huang らが提案した時間領域 (TD) および自己回帰特徴量であり、(1) 平均絶対値、(2) ゼロ交差回数、(3) 傾き符号変化回数、(4) 波形長、(5) 二乗平均平方根、および (6–11) 6 次の自己回帰係数から成る [149]。NinaPro DB5 および DB2 では、チャンネルごとに以下の 4 つの特徴量を抽出した。これらは Hudgins らが提案した時間領域 (TD) 特徴量であり、(1) 平均絶対値、(2) ゼロ交差回数、(3) 傾き符号変化回数、および (4) 波形長から構成される [182]。これらの特徴量に基づき、MyoDataset, NinaPro DB5, および NinaPro DB2 における特徴次元数 D は、それぞれ 88, 64, 48 となった。

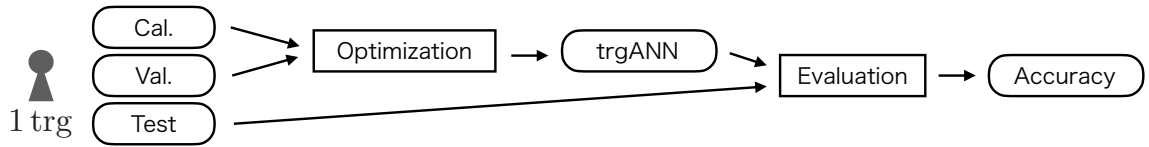


図 4.4: ターゲットユーザ特化型モデル (trgANN) の学習および評価手順. 角が丸いブロック”Cal.”, ”Val.”, および”Test”は, それぞれキャリブレーション (調整), 検証, およびテスト用のデータを意味する. 文献 [13] より転載.

4.2.3 バックボーンモデル

バックボーンモデルとして, 全結合層のみで構成された 3 つの隠れ層をもつ ANN を採用した. ANN は全結合層のみで構成される最も単純なニューラルネットワークであるため, シングルソースアプローチとマルチソースアプローチの性能差に焦点を当てるのに適している. また, ANN は他の伝統的な識別器と比較して, sEMG パターン認識の転移学習において優れた性能を示すことが報告されている [10]. 第 3 章で述べたように ANN にはいくつかのハイパーパラメータが存在するが, 計算コストを抑えるため, 次の 2 種類のハイパーパラメータのみを最適化の対象とした. 各隠れ層のニューロン数を $[D/4, D/2, D, 2D, 4D]$, 重み減衰の値を $[0.00001, 0.0001, \dots, 0.1]$ の範囲で最適化した. その他のハイパーパラメータについては以下のように統一して設定した. 各隠れ層には活性化関数として ReLU を適用し, バッチ正規化 (Batch Normalization) を施した. 出力層にはソフトマックス関数を使用し, 学習プロセスは学習率 0.01 の Adam を用いた.

シングルソースおよびマルチソースアプローチの性能を調査するため, 以下の 3 つのモデルを訓練した.

1. ターゲットユーザ特化型 ANN (trgANN)
2. ソースユーザ汎用 ANN (genANN)
3. ソースユーザ特化型 ANN (srcANN)

これらのモデルについて以下で詳細を述べる.

(1) ターゲットユーザ特化型 ANN: trgANN

このモデルはシングルソースアプローチを代表するものであり, ターゲットユーザのデータのみを用いて訓練された (図 4.4 参照). このモデルを用いた評価は本質的に被験者内識別 (within-subject classification) である. ハイパーパラメータ (ニューロン数および重み減衰値) は検証用データを用いて最適化し, すべての被験者で最も頻繁に選ばれたハイパーパラメータセットを一律に適用した.

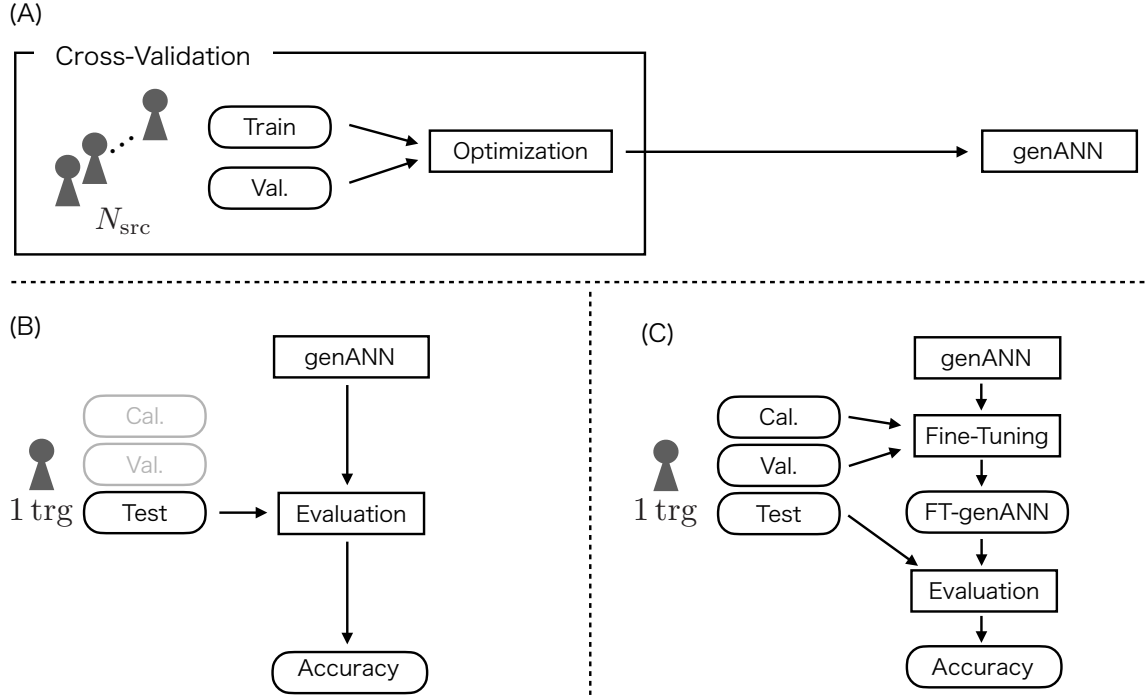


図 4.5: シングルソースモデル (genANN および FT-genANN) の学習および評価手順 ((A) genANN の事前学習手順, (B) genANN の評価手順, および (C) FT-genANN の評価手順). 文献 [13] より転載.

(2) ソースユーザ汎用 ANN: genANN

このモデルはシングルソースアプローチを代表するものであり, すべてのソースユーザのデータを単一の大規模データセットとして統合し, 訓練に使用した (図 4.5 参照). このモデルを用いた評価は本質的に被験者間識別 (cross-subject classification) である. MyoDataset および NinaPro DB5 に対しては, ハイパーパラメータを leave-one-source-subject-out 交差検証 (すなわち, 1 名を除く全ソース被験者を訓練データ, 除いた 1 名を検証データとして利用) を用いて最適化し, 任意のソース被験者を識別可能な ANN を構築した. NinaPro DB2 は被験者数が多すぎるため, 39 名のソース被験者をランダムに 10 グループに分割した 10-fold 交差検証を用いてハイパーパラメータを最適化した. 最終的に, 全体で最も頻繁に選ばれたハイパーパラメータセットを一律に適用した.

(3) ソースユーザ特化型 ANN: srcANN

このモデルはマルチソースアプローチを代表するものであり, ソース被験者ごとに個別に訓練された (図 4.6 参照). このモデルを用いた評価も本質的に被験者間識別 (cross-subject classification) である. 複数の srcANN を統合するためには, 4.2.5 節で後述するアンサンブル戦略を用いた. ハイパーパラメータは leave-one-trial-out 交差検証 (すなわち, ソース被験者ごとに 1 試行を検証データとして除外し, 残りの試行を訓練データとして使用) によ

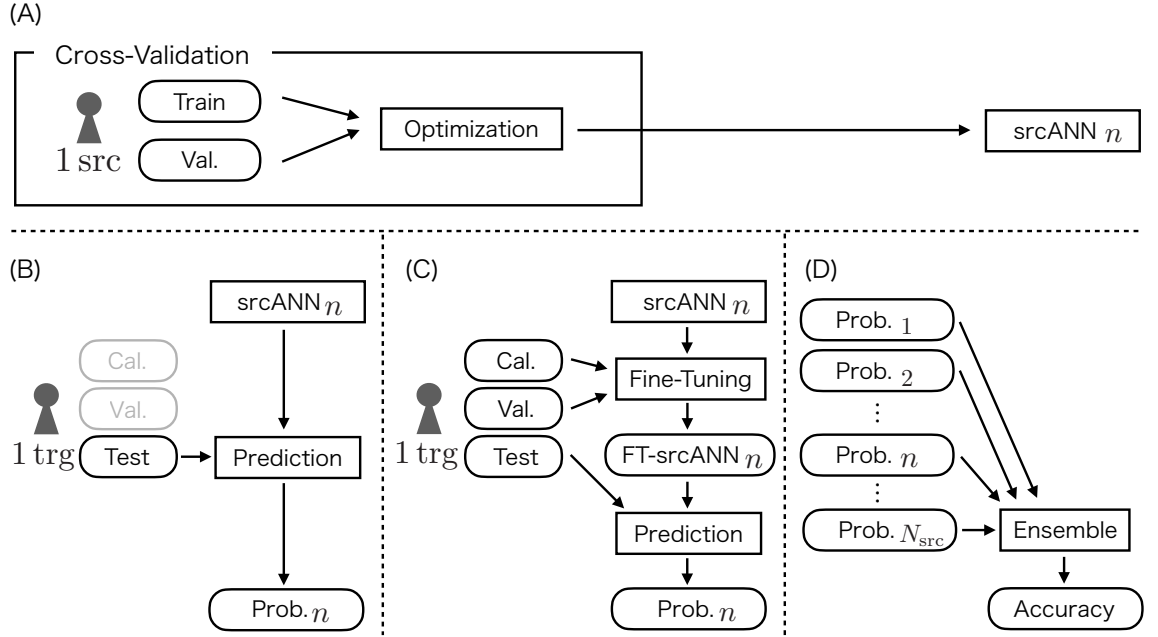


図 4.6: マルチソースモデルの学習および評価手順. (A) srcANN の事前学習手順. ターゲット被験者に対して, ソース被験者の数だけ srcANN が準備される. (B) HV, SV, WV, および VW の評価手順の一部. "Prob." と記されたブロックは, 動作クラスの数と同じ次元を持つ出力確率を表す. (C) FT-HV, FT-SV, FT-WV, FT-SVW の評価手順の一部. (D) 複数の確率を統合するアンサンブル手順と評価手順. ここで, N_{src} はソース被験者の数である. 文献 [13] より転載.

て最適化を行い, その被験者を効果的に識別できる srcANN を構築した. 最終的に, すべてのソース被験者で最も頻繁に選ばれたハイパーパラメータセットを一律に適用した.

4.2.4 ファインチューニング (FT)

FT は, 事前学習された ANN の後半の層をターゲットデータで再訓練する転移学習手法である. すべての層ではなく, 出力層に近い層のみを再学習するのが一般的である. 再学習する層数が多くなるほど, 訓練すべきパラメータ数が増加するため, 必要となるデータ量も増加することになる [139]. 本章ではターゲットのキャリブレーションデータ数が限られていることを考慮し, 最後の隠れ層と出力層の間のパラメータのみを FT の対象とした. また, 過学習を避けるために重み減衰による正則化を用い, その値は検証データを用いて最適化した. FT を施した genANN を FT-genANN, 同様に FT を施した srcANN を FT-srcANN と呼称する (図 4.5 および図 4.6 参照).

4.2.5 アンサンブル戦略

マルチソースアプローチにおいてアンサンブル戦略が認識精度に与える影響を調査するために、本章では以下の4種類のアンサンブル戦略を採用した。

(1) ハード投票法 (Hard voting: HV)

この戦略は、複数の srcANN/FT-srcANN によって予測されたクラスラベルの中から最も頻繁に現れるラベルを選択する方法である。例えば、2つのモデルがラベル A を予測し、3つのモデルがラベル B を予測した場合、HV はラベル B を出力ラベルとして選択する。複数のクラスが同数票を得た場合は、その中からランダムに1つを選択した。

(2) ソフト投票法 (Soft voting: SV)

この戦略は、すべての srcANN/FT-srcANN の出力確率を合算し、合計確率が最も大きいクラスラベルを選択する。HV と比較して柔軟性が高い戦略である。例えば、あるモデルが2つのクラスラベルにほぼ同じ確率を与えた場合でも、HV は強制的に1つのラベルを選択する必要があるが、SV では両方の可能性を考慮できる。

(3) 重み付き投票法 (Weighted voting: WV)

この戦略はSVの派生型であり、ターゲット被験者に類似した srcANN/FT-srcANN に対して高い重みを割り当てる方法である。HV および SV は、すべてのモデルが同等に有効であると仮定している。しかし実際には、ターゲットに対して類似したソースと類似しないソースが存在する可能性が高い。この点に対応するため、まずターゲットとソース被験者ペアごとの信頼度 (confidence) を算出し、それに基づいて srcANN/FT-srcANN の出力に重みを付与した。本章では、ターゲット被験者があるソース被験者に類似している場合、そのソース被験者の srcANN (FT 無しでも) がターゲットをうまく識別できるという仮定のもとで、その srcANN の認識精度を重みとして用いた (srcANN と FT-srcANN では同じ重みを用いたことに注意)。その後、各 srcANN/FT-srcANN の出力確率に対応する重みを掛け、SV と同様に重み付き確率の総和が最大となるクラスを出力ラベルとして選択した。

(4) 教師あり重み付き投票法 (Supervised weighted voting: SWV)

この戦略は WV とは別の重み付け方法であり、WV のようにモデルごとに一律の重みを与えるのではなく、教師あり学習を用いて各モデルの各クラスに対して詳細な重みを決定する。ここで、 N_s を srcANN/FT-srcANN の数、 N_c を識別対象となる動作の数とする。WV では各 srcANN/FT-srcANN に対して重みが設定されるため、重みの総数は N_s 個である。また WV は、モデルの識別能力がすべてのクラスに対して一律であることを暗黙的に仮定し

ている．しかし実際には，特定の動作に適合する（または適合しない）モデルが存在する可能性がある．SWV はこのような状況に対応するために，各 srcANN/FT-srcANN の各クラス出力に個別の重みを割り当てる（つまり，合計で $N_s \times N_c$ 個の重みを設定する）．重みの学習には以下の3つのステップを用いた．キャリブレーションデータの各サンプルに対して，各 srcANN/FT-srcANN を用いて識別し，その確率を取得する．得られた確率を各サンプルごとに連結し，ベクトル化する（ベクトルの長さは $N_s \times N_c$ ）．すべてのキャリブレーションデータのベクトルを用いて，メタ識別器（meta-classifier）を訓練する．本章ではロジスティック回帰をメタ識別器として採用し，その L2 正則化のハイパーパラメータは検証データを用いて $[0.0001, 0.001, \dots, 1000]$ の範囲で最適化した．テストデータを用いた推論時には，各モデルの出力を連結し，それを訓練済みのロジスティック回帰モデルの入力として最終的なクラスラベルを決定した．

4.2.6 評価方法

本章では，データセット中の1名をターゲット被験者，それ以外の被験者をソース被験者として設定し，これを全被験者について一度ずつ繰り返した．ある被験者をターゲットとして使用する際，その被験者のデータはキャリブレーションデータ，検証データ，およびテストデータに分割した．MyoDataset では，第1試行をキャリブレーション用，第2試行を検証用，残りの3試行をテスト用とした．NinaPro DB5 および DB2 では，第1，第3，および第4試行をキャリブレーション用，第6試行を検証用，残りの2試行（第2および第5試行）をテスト用とした．また，キャリブレーションデータを基準として，キャリブレーション，検証，テストデータの標準化を行った．これら3種類のデータは，すべての評価手順で共通して用いられた（図 4.4, 図 4.5, および図 4.6 参照）．一方，ある被験者をソース被験者として用いる際のデータ使用法は，評価手順によって異なった．genANN においては，交差検証のためソース被験者群を訓練グループと検証グループに分割した（図 4.5 参照）．訓練データを基準として，訓練データとテストデータを標準化した．srcANN においては，あるソース被験者のデータのうち1試行をテスト用として除外し，残りの試行を訓練データとして交差検証を実施した（図 4.6 参照）．こちらも訓練データを基準として，訓練データとテストデータを標準化した．

シングルソースおよびマルチソースアプローチの性能を，FT の有無と共に検証するため，本章では以下の11種類のモデル (1) trgANN, (2) genANN, (3) HV, (4) SV, (5) WV, (6) SWV, (7) FT-genANN, (8) FT-HV, (9) FT-SV, (10) FT-WV, および (11) FT-SWV の認識精度を評価した．なお，trgANN のみがソース被験者データを一切使用しない．したがって，trgANN は自身のデータのみを用いた自己ソース型（self-source）アプローチである．被験者内識別と被験者間識別の性能差は，trgANN とそれ以外のモデルの比較により評価できる．また，シングルソースアプローチとマルチソースアプローチの性能差は，genANN ベースのモデル群とアンサンブルベースのモデル群の比較により評価できる．具体的には，genANN はシングルソースアプローチに分類される一方で，HV, SV, WV, SWV はマルチソースアプローチに分類される．さらに，モデル名に FT が付与されているモデルは，FT

表 4.2: 3 種類のデータセットにおける 11 種類のモデルの識別精度（平均値 $\pm$ 標準偏差%）。太字の数値は、各データセットにおいて最も高い精度を示したモデルの結果である。文献 [13] より転載して一部改変。

Model	Transferred	Type	# of source	MyoDataset	NinaPro DB5	NinaPro DB2
trgANN	False	Within	Self	90.30 $\pm$ 7.64	56.35 $\pm$ 4.09	64.86 $\pm$ 7.25
genANN	False	Cross	Single	71.42 $\pm$ 15.62	17.28 $\pm$ 4.82	15.11 $\pm$ 2.99
HV	False	Cross	Multi	75.31 $\pm$ 18.70	18.14 $\pm$ 4.69	17.84 $\pm$ 3.40
SV	False	Cross	Multi	75.61 $\pm$ 18.93	19.40 $\pm$ 5.21	17.30 $\pm$ 3.18
WV	False	Cross	Multi	78.24 $\pm$ 17.08	20.24 $\pm$ 5.09	19.12 $\pm$ 3.46
SWV	False	Cross	Multi	91.96 $\pm$ 4.27	46.79 $\pm$ 4.34	62.69 $\pm$ 7.40
FT-genANN	True	Cross	Single	92.08 $\pm$ 5.22	46.24 $\pm$ 2.77	55.76 $\pm$ 7.69
FT-HV	True	Cross	Multi	95.26 $\pm$ 4.29	60.84 $\pm$ 4.68	67.78 $\pm$ 6.82
FT-SV	True	Cross	Multi	95.26 $\pm$ 4.29	61.72 $\pm$ 4.68	67.67 $\pm$ 6.85
FT-WV	True	Cross	Multi	95.33 $\pm$ 4.21	61.61 $\pm$ 4.56	67.79 $\pm$ 6.78
FT-SWV	True	Cross	Multi	94.68 $\pm$ 4.53	61.54 $\pm$ 5.01	68.86 $\pm$ 7.24

を施したシングルソース（FT-genANN）およびマルチソース（FT-HV, FT-SV, FT-WV, FT-SWV）アプローチである。したがって、FT の効果は FT の有無によるモデル間の比較によって評価可能である。

sEMG パターン認識に対して適切なアプローチを決定するため、各データセットにおいてシングルソースアプローチの認識精度と、マルチソースアプローチのうち最も保守的（最低精度）な認識精度との間で、Wilcoxon 符号付き順位検定（Wilcoxon signed-rank test）を行った。ここで、最悪の認識精度を示したマルチソースアプローチは、異なる 4 つのアンサンブル戦略を用いたマルチソースアプローチのうち、最低の精度を示したものを選択した。さらに、マルチソースアプローチの 4 つのアンサンブル戦略間に認識精度の差が存在するかを確認するため、Friedman 検定を実施した。有意差が確認された場合は、どのグループ間に差が存在するかを調べるため、Tukey の多重比較検定を実施した。すべての統計解析において、有意水準は 0.05 とした。

4.3 結果

4.3.1 全体の認識精度

表 4.2 に 3 つのデータセットにおける 11 種類のモデルの平均認識精度を示し、図 4.7 にその結果を箱ひげ図で示す。シングルソースアプローチとマルチソースアプローチ間の性能差について、以下のような傾向が観察された。FT を適用しない場合、両アプローチは全般的に被験者内識別である trgANN より精度が低かったものの、マルチソースアプローチはシングルソースアプローチよりも高い精度を示した。一方、FT を適用した場合、両アプローチとも FT を適用しない場合よりも認識精度が改善された。ただし、FT を適用したシングルソースモデル（FT-genANN）は NinaPro DB5 と DB2 において trgANN よりも低い精度に留まった。これに対して、FT を適用したマルチソースモデル（FT-HV, FT-SV, FT-WV,

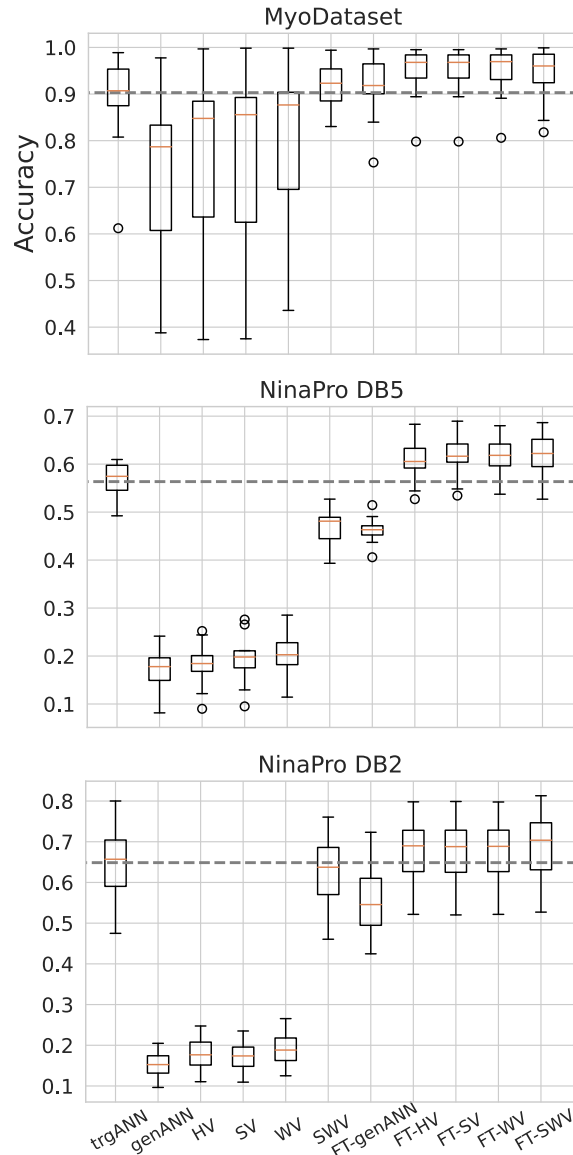


図 4.7: 3 つのデータセットに対する 11 種類のモデルの精度を示す箱ひげ図。破線の灰色の線は、trgANN の平均精度を表す。文献 [13] より転載。

および FT-SWV) は、シングルソースモデルおよび trgANN を常に上回る性能を示した。上記の観察結果から、FT の有無にかかわらずマルチソースアプローチはシングルソースアプローチよりも優れた結果を示し、特に FT を適用する場合、マルチソースアプローチは trgANN よりも高い認識精度を達成することが明らかとなった。

観察されたマルチソースアプローチの優位性に関する傾向は、統計的にも検証された。Wilcoxon 符号付き順位検定の結果、選択された FT 適用マルチソースアプローチの精度は、すべてのデータセットにおいて FT 適用シングルソースアプローチの精度よりも有意に高かった (MyoDataset: $Z = 2.630$, $p = 0.009$, NinaPro DB5: $Z = 3.742$, $p = 0.0002$, NinaPro DB2: $Z = 5.5859$, $p < 0.0001$)。さらに、Friedman 検定により、アンサンブル戦略間での有

有意な差が NinaPro DB5 ($\chi^2(3) = 14.4687, p = 0.0023$) と NinaPro DB2 ($\chi^2(3) = 49.8854, p < 0.0001$) では確認されたが, MyoDataset ($\chi^2(3) = 3.839, p = 0.2793$) では確認されなかった. 加えて, どのアンサンブル戦略間で差があるかを明らかにするため, Tukey の多重比較検定を行った. その結果, NinaPro DB5 では FT-HV と FT-SV 間に有意差が見られ, NinaPro DB2 では FT-HV と FT-SWV, FT-SV と FT-SWV, FT-WV と FT-SWV 間にそれぞれ有意差が認められた.

4.3.2 特徴量分布の可視化

すべてのデータセットにおいてシングルソースアプローチがマルチソースアプローチよりも劣っていた理由を明らかにするために, 4.2.2 節で述べた手動設計の特徴量を t 分布型確率的近傍埋め込み法 (t-SNE) を用いて可視化した. t-SNE は, 高次元データを低次元空間に視覚化するために広く用いられる機械学習アルゴリズムであり, 非線形かつ教師なしの次元削減手法として知られている [183]. 本章では, MyoDataset の全被験者のデータを単一のデータセットとして統合し, このデータセットに対して t-SNE を適用することで, 88 次元の特徴量を 2 次元へと圧縮した. t-SNE の実装には, Python の代表的な機械学習ライブラリである scikit-learn (バージョン 1.0.2) の `sklearn.manifold.TSNE` を使用し, その際のハイパーパラメータは scikit-learn ライブラリのデフォルト値を用いた.

図 4.8 に MyoDataset に対する t-SNE 可視化の結果を示す. それぞれの図には同一のデータが表示されている一方, 図の上部に記述されているクラスのデータを赤色で強調し, その他のクラスのデータを灰色で示している. 教師なしのアルゴリズムを使用しているにもかかわらず, 特徴量はおおむねクラスごとにまとまっていることがわかる. 特に安静 (resting), 手首屈曲 (wrist flexion), 手首伸展 (wrist extension) に対応する特徴量 (クラス 1, 2, および 3) は非常に明瞭に分離されている. ただし, それ以外のクラスの特徴量については, 同じ動作タスクであっても比較的ばらつきが大きかった.

図 4.8 をベースに, 個人差の有無を調査するため, 被験者単位でのハイライト表示を安静クラス (resting class) と前腕回内クラス (forearm pronation class) に対して行ったものを図 4.9 として示す. これらのクラスを選択した理由は, 安静クラスが t-SNE 空間で最も分散が小さく, 前腕回内クラスが最も分散が大きかったためである. 図 4.9(A) に図 4.8 から抽出した安静クラスの特徴量を示し, それぞれの図が特定の被験者を赤色でハイライトし, その他の被験者を灰色で表示している. 同様に, 図 4.9(B) は前腕回内クラスを同じ方法で示している. 図 4.9(A) に示した安静クラスでは, 特徴量は基本的に被験者ごとにまとまってクラスタを形成している. 一方, 図 4.9(B) に示した前腕回内クラスでは, 一部の被験者はよくまとまっているが, 全体として安静クラスに比べて特徴量がより散在しているように見える.

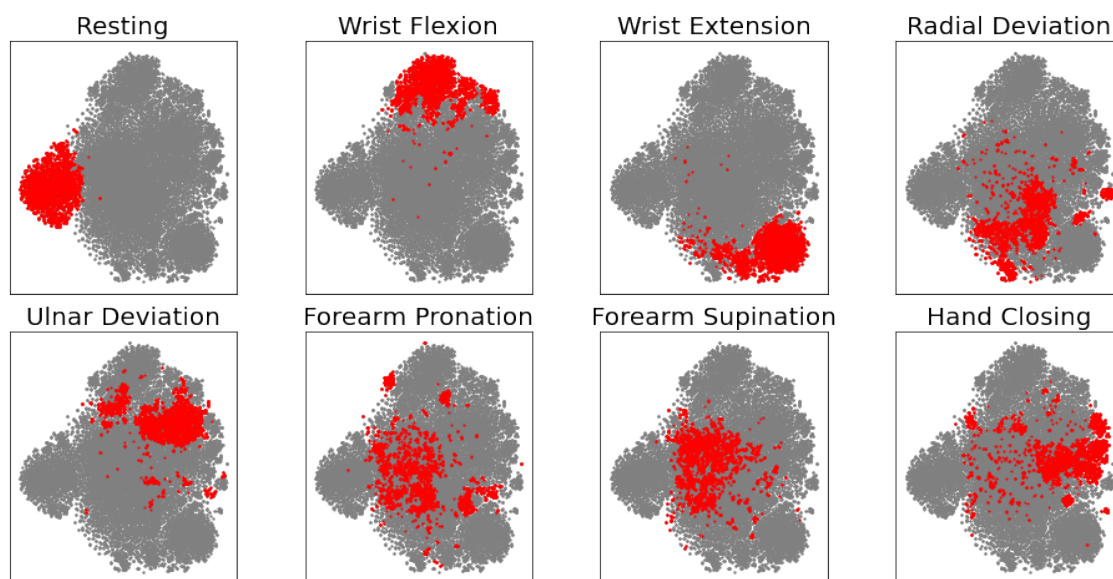


図 4.8: MyoDataset の特徴量を t-SNE により 2 次元に次元圧縮した結果. いずれの図でも同一のデータを表示しているが, 特定のクラスのみを赤で強調し, その他のクラスは灰色で示している. 文献 [13] より転載.

4.3.3 アンサンブル戦略への trgANN の追加

これまでのアンサンブル戦略では srcANN のみを用いたが, trgANN もアンサンブル戦略に含める (つまり, ターゲットユーザのデータのみで学習させた識別器をアンサンブル戦略に加える) ことが可能である. 表 4.3 に trgANN をアンサンブル戦略に追加した場合の精度を示し, 図 4.10 に trgANN 追加前後の平均精度を示す. trgANN を追加することにより, すべてのデータセットにおいて平均認識精度が全体的に向上した (表 4.3 の「Diff.」列参照). 特に, FT を適用していないモデル (HV, SV, WV, および SWV) における精度向上は, FT を適用したモデル (FT-HV, FT-SV, FT-WV, および FT-SWV) における精度向上よりも顕著であった. FT なしのモデルの中では, SWV が最も高い精度を示し, その精度は FT を適用したモデルの精度に近づいている. また, WV の性能は NinaPro DB5 および DB2 において大幅に向上した. 一方, FT を適用したモデルにおいては, trgANN を追加することによる性能改善は限定的であった. さらに, NinaPro DB5 では, trgANN を追加したことで FT-WV および FT-SWV の認識精度がわずかに低下した.

4.4 考察

4.4.1 被験者内識別モデル (trgANN) と従来研究で使用されたモデルとの比較

sEMG データは個人差が大きいため, 被験者内識別が sEMG パターン認識においてよく用いられている [26, 184, 185]. 本章で用いたデータセットに対しても, データ提供元のグ

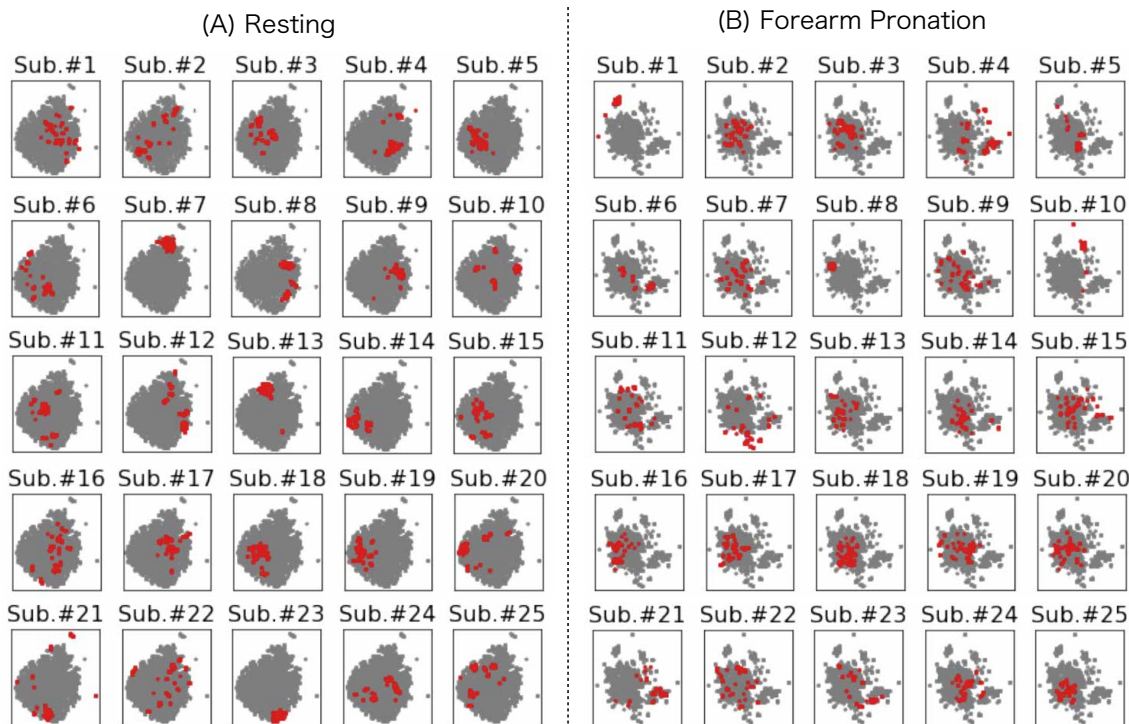


図 4.9: 図 4.8 から抽出した (A) 安静クラス特徴および (B) 前腕回内クラス特徴における個人差. 各図では, 特定の被験者の特徴を赤で強調し, その他の被験者の特徴は灰色で表示している. 文献 [13] より転載.

ループによって被験者内識別の評価がすでに行われている [12, 10]. そこで本章では, シングルソースおよびマルチソースアプローチを FT あり・なしの条件で評価するためのベースラインとして, まずは各データセットにおける被験者内識別モデル (trgANN) の性能を再検証した.

MyoDataset における trgANN の認識精度は, 表 4.2 に示されている通り 90.30%であった. MyoDataset は, 単純な 8 種類の単一自由度の前腕動作 (すなわち, 手開き, 手首の屈曲, 手首の伸展, 手首の橈屈, 手首の尺屈, 前腕の回内, 前腕の回外, および手握り) から構成されている. このような比較的難易度が低い動作を対象としたデータセットであるため, ターゲット被験者から得られた 1 試行のみを訓練に用いたとしても高い精度が得られることは妥当な結果である. また, 今回得られた 90.30%という精度は過去の研究 [10] において報告された精度と一致している.

NinaPro DB5 における trgANN の認識精度は, 表 4.2 に示されている通り 56.35%であった. 本章では, 指および手首の基本動作, 把持動作および機能的動作からなる合計 40 クラス (Exercise B および C) の認識を行った. このデータセットは比較的多くの動作・ジェスチャクラスを含んでおり, パターン認識の難易度は高いと考えられる. 本章で得られた認識精度は, データセット提供者の報告値よりも低い結果となった. 提供元の研究 [12] では, Hudgins の TD 特徴量とランダムフォレスト (RF) 識別器を用いて約 68%の精度が報告さ

表 4.3: 3 種類のデータセットにおける, trgANN を用いたマルチソース手法の識別精度 (平均値 $\pm$ 標準偏差 %) および, trgANN の有無による精度差を示す. 太字の数値は, 各データセットで最も高い精度を示したモデルの結果である. 「+」の記号は, そのモデルのアンサンブル戦略に trgANN が含まれていることを示す. 文献 [13] より転載.

Model	MyoDataset	Diff.	NinaPro DB5	Diff.	NinaPro DB2	Diff.
HV+	76.78 $\pm$ 18.12	1.46	27.81 $\pm$ 5.30	9.67	23.30 $\pm$ 3.91	5.46
SV+	77.24 $\pm$ 18.13	1.63	29.75 $\pm$ 5.38	10.35	22.44 $\pm$ 3.74	5.14
WV+	81.05 $\pm$ 14.87	2.81	55.34 $\pm$ 4.24	35.10	62.24 $\pm$ 8.36	43.12
SWV+	93.82 $\pm$ 4.59	1.86	56.96 $\pm$ 4.20	10.17	66.15 $\pm$ 7.33	3.46
FT-HV+	95.28 $\pm$ 4.32	0.01	61.77 $\pm$ 4.55	0.93	68.12 $\pm$ 6.82	0.33
FT-SV+	95.35 $\pm$ 4.18	0.08	62.41 $\pm$ 4.51	0.69	67.97 $\pm$ 6.82	0.30
FT-WV+	95.37 $\pm$ 4.17	0.04	61.24 $\pm$ 4.83	-0.37	69.14 $\pm$ 6.92	1.35
FT-SWV+	95.94 $\pm$ 2.82	1.26	61.47 $\pm$ 4.70	-0.06	69.34 $\pm$ 7.42	0.47

れている. このように本章と従来研究の間に生じた精度の差は, 主に以下の 2 つの要因に起因すると考えられる. 第一に, 従来研究で用いられた RF 識別器が, 本章で用いた ANN よりも本データセットに適していた可能性が挙げられる. 実際に, Pizzolato ら [12] は RF に加えて SVM を用いた評価も行っており, 識別器の種類によって精度が異なることを報告している. 第二に, 従来研究では 40 クラスに加えて, 識別が容易な安静クラス (resting class) が追加されていた点が挙げられる. 本章では安静クラスを追加していない. 従来研究では動作が始まる前のセクションから安静クラスのデータを抽出しているが, そのデータ数については明記されていないため推測の域を出ないものの, もしすべての動作課題ごとに安静クラスのデータが生成されていたとすると, 全体の認識精度が本章よりも大幅に高くなった可能性がある.

NinaPro DB2 における trgANN の認識精度は, 表 4.2 に示されているように 64.86% であった. NinaPro DB5 と同様の 40 クラスの識別を行ったが, NinaPro DB2 では DB5 よりもチャンネル数が少ないにもかかわらず, より高い認識精度が得られた. この精度差の主な原因として, 両データセット間の計測デバイスやサンプリング周波数の違いが挙げられる. Kim らの報告 [186] にもあるように, sEMG 動作認識においてサンプリング周波数は重要な要因である. 特に sEMG データは有効周波数が 500 Hz 程度までであるため, 高周波寄りの情報の有無が識別精度の差に影響を与えた可能性がある. データセット提供元も, NinaPro DB2 が DB5 よりも高い認識精度を示すことを報告している (RF 使用時の精度: 約 74% vs. 約 68%) [12]. この結果と比較すると本章でも同程度 (約 8%) の性能向上が起きていた. 一方, 本章においてデータ提供元の報告よりも低い認識精度となった原因については, 前述の NinaPro DB5 のケースと同様の理由 (識別器の違いおよび安静クラスの有無) が考えられる.

以上の結果から, trgANN による被験者内識別の認識精度は, 従来の研究で報告された精度とは若干異なるものの, 許容範囲内であることが示された. このため, これらの認識精度を, シングルソースおよびマルチソースアプローチの性能を転移学習 (FT) の有無と共に比較する際のベースラインとして使用することが妥当であると考えられる.

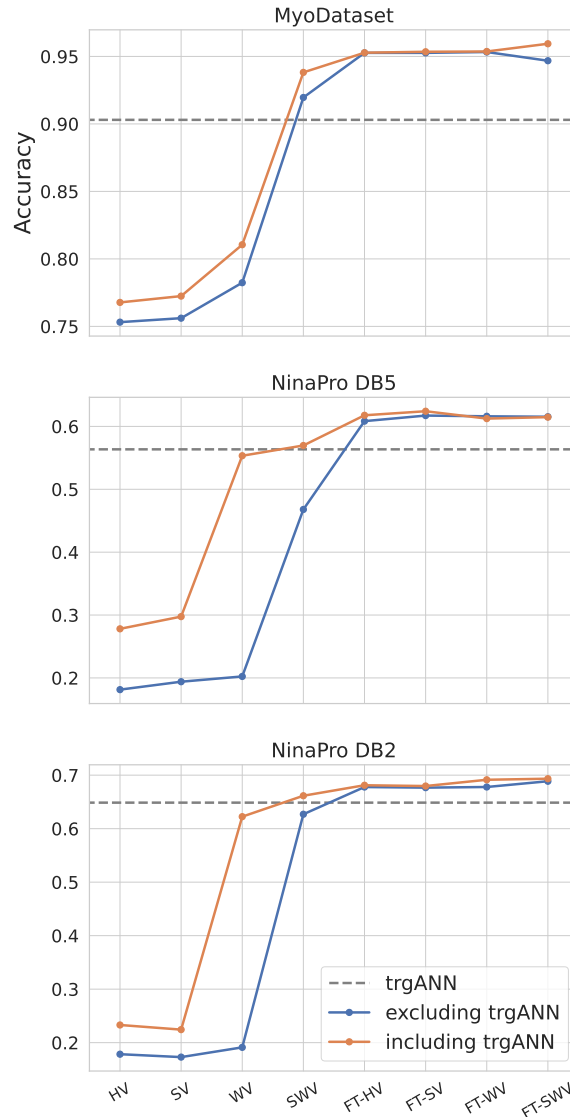


図 4.10: 3 種類のデータセットにおける trgANN を用いた場合および用いない場合の識別精度を示す。青の実線は trgANN を用いない場合の精度を，オレンジの実線は trgANN を用いた場合の精度を，灰色の破線は trgANN の平均精度をそれぞれ表している。文献 [13] より転載。

4.4.2 シングルソースおよびマルチソースアプローチの評価

表 4.2 に示した FT 適用モデルと非適用モデル間の精度差および図 4.9 の t-SNE による可視化結果から，被験者間に個人差が存在していることが明らかである。これらは同時に，被験者間識別において個人差による悪影響が起きる可能性を示しており，転移学習（FT）の必要性を強く示唆している。これに対し，本章では FT が個人差を効果的に緩和し，特にマルチソースアプローチと相性が良いことが確認された。また，統計的検定の結果から明らかのように，FT を適用する場合，シングルソースアプローチよりもマルチソースアプローチの方が適していることが示された。

FTは全体的なモデル性能を向上させたが、その向上後の性能は被験者内識別 (trgANN) を上回ることが理想である一方、シングルスソースアプローチと組み合わせた場合、MyoDatasetにおいてのみ trgANN の性能を上回った。データセットごとの結果の違いを詳細に検討するため、各データセットで最適化された ANN のニューロン数に着目した。入力特徴次元が 88 の MyoDataset では、trgANN と genANN において、それぞれ最適化されたニューロン数が 1 層目から 3 層目にかけて {352, 44, 44}, {352, 352, 176} であった。genANN が trgANN よりもニューロンの総数が多いのは、genANN がより大量のデータを処理する必要があるため、直感的にも理解できる。しかし、入力特徴次元が 64 の NinaPro DB5 では、trgANN と genANN の最適化ニューロン数はそれぞれ {256, 128, 64} と {32, 16, 16} であった。また、入力特徴次元が 48 の NinaPro DB2 では、trgANN と genANN の最適化ニューロン数はそれぞれ {192, 192, 96} と {96, 24, 24} であった。本来、genANN が大量のデータを処理するにはモデルの複雑性が増す必要があるはずだが、この場合は逆の結果が得られており直感に反するものである。これは、ハイパーパラメータ探索において、過学習のリスクを回避するために少ないニューロン数が選択された可能性がある。言い換えると、NinaPro DB5 および DB2 には動作の難易度の多さから個人差が大きく存在するため、genANN が十分に学習するのが困難であった可能性がある。

本章ではシングルスソース、マルチソース両方のアプローチの特性が明らかとなった。まず、シングルスソースアプローチはデータセットの複雑性の影響を強く受け、過学習に陥りやすいことが分かった。過学習の主な原因は訓練データ量の不足にあると考えられ、このデータ不足によって genANN が特定のパターンやノイズに過度に適合してしまったと推測される。このため、このアプローチの精度向上には何らかの追加的な正則化が必要であり、データ拡張 (data augmentation) が適切な手法であると考えられる。データ拡張とは、既存のデータから新たなデータを生成することで訓練データ量を増やす機械学習手法であり、効果的な正則化法として知られている [187, 188]。したがって、データ拡張を取り入れることで、データサンプルの不足に起因する過学習の低減が期待される。一方、マルチソースアプローチは追加的な正則化を取り入れなくとも、様々なデータセット条件で良好な性能を発揮した。また、興味深い傾向も観測された。それはソース数と性能の関係である。当初、ソース数が多くなると FT されたモデル間の多様性が失われ、性能が低下することを懸念していた。しかし、実際は認識性能に悪影響を与えることはなかった。

4.4.3 アンサンブル戦略の分析

本章において、データセットの性質に応じた適切なアンサンブル戦略の選択が重要であることが多くのケースで確認された。FT が適用されない場合には、個人差を効果的に軽減できるアンサンブル戦略が望ましいと考えられる。実際、全てのデータセットにおいて SWV が最も高い性能を示し、WV がそれに次ぐ性能を示した。単純な多数決型のアンサンブル戦略 (HV および SV) では、ターゲット被験者と類似しないソース被験者からの誤った投票によって性能が低下する可能性がある。一方、重み付けアンサンブル戦略 (WV および SWV) では、誤った投票の影響を軽減でき、比較的高い精度が得られた。特に、SWV は教師あり

学習に基づくため、より適切に重みを調整できる。このため、FT が適用されない条件下では、教師ありの重み付け型アンサンブル戦略が最も適していることが明らかとなった。これに対して、FT を適用した場合、全体を通して常に最適となるアンサンブル戦略は存在しなかった。ただし、NinaPro DB5 および DB2 では、わずかながら統計的に有意な差がアンサンブル戦略間で確認された。したがって、個人差が大きい場合には特に、アンサンブル戦略の選択が重要であることが本章の結果から示唆された。

trgANN をアンサンブル戦略に追加することは一般的に精度向上に寄与し、特に FT を適用しない場合にその恩恵が顕著であった（表 4.3 参照）。これは trgANN が srcANN よりもターゲット被験者にとって信頼性が高く、投票プロセスに好影響を与えるためである。特に NinaPro DB5 および DB2 において、WV+戦略で精度の大きな向上が見られたが、これは trgANN に対して投票の重みが偏った結果である可能性が高い。実際、NinaPro DB2 の WV+戦略で用いられた平均重みは、srcANN が 0.07 に対して trgANN は 0.56 であり、同様に NinaPro DB5 でも srcANN が 0.11、trgANN が 0.50 であった。このことから、trgANN が重み付き投票手法において支配的な影響を与えていることが推察される。しかし、1つのモデルに過度に依存すると、モデル間の多様性が失われ、アンサンブル戦略の利点が損なわれてしまう。このことは、表 4.3 に示した通り、FT-WV+および FT-SWV+に trgANN を追加した際に認識性能が低下した原因であると考えられる。したがって、trgANN をアンサンブル戦略に追加できる場合もアンサンブル戦略の選択は重要である。

4.4.4 sEMG ベース HCI への影響および適用可能性

本章の主な貢献は、FT を用いて複数のソース被験者データを統合するための新たなガイドラインを提供したことである。従来、sEMG ベースの HCI 開発者は明確な指針がないまま 2 種類のアプローチのいずれかを選択せざるを得ず、このプロセスには多大な人的関与と時間が必要だった。このため、本章の知見は、特に sEMG ベース HCI システムの開発者にとって有用である。本章の結果から、FT を適用したマルチソースアプローチが sEMG パターン認識に適していることが明らかになった。データセットの特徴に応じて適切なアンサンブル戦略を選択する必要があるが、その中でも SV は特に有力な候補である。FT-SV は trgANN を含めても全てのデータセットにおいて安定した性能を示している。また、SV は計算処理がシンプルで実装が容易なため、HCI システム開発者にとって取り入れやすい手法である。本章で提案したガイドラインに従うことで、開発者はモデル構築や最適化に伴う余計な労力を回避できる。その結果、ユーザのニーズに焦点を当てることが可能となり、sEMG を用いた HCI コミュニティに新たな知見をもたらすと期待される。

本章では FT に対する 2 つのアプローチの性能を調査したが、本章で得られた知見は sEMG 以外のモダリティにも応用可能である。本章で得られた本質的な知見は、「FT を用いたマルチソースアプローチはソースとターゲット間のデータ分布の差異を効果的に低減できる」という点である。このため、ソースドメインごとにデータが蓄積される他のモダリティにおいても、本章の知見が応用できる可能性が高い。特に、多種多様な生体信号との相性が良いと考えられる。生体信号の記録では、通常、信号は被験者ごとに計測され、そのまま被験者ご

とのデータとして保存されることが多い．実際に公開されている EEG のデータセットも、多くが被験者ごとに記録されている [189, 190]．また別の観点から見ると、本章は「データを取得した際のソースドメイン情報を記録する重要性」を示したとも言える．過去にはソースドメインが特に考慮されてこなかったデータであっても、ドメイン情報を活用することで高性能な FT が可能になる可能性がある．

4.4.5 本章の限界と今後の課題

本章で得られた知見は、健康な被験者を対象としており、被験者それぞれから収録された sEMG データは 1 日といった短期間のデータセットに基づいている．これまでの多くの研究でも健康な被験者が用いられてきたが、実際の sEMG ベースの HCI システムの重要なユーザ層として、四肢切断者が存在する．切断者は切断時期や切断の程度などが個々に異なるため、健常者よりもさらに個人差が大きいことが知られている [191]．そのため、今後は、個人差がより大きくなる可能性がある切断者を対象として、本章で提案した FT を用いたマルチソースアプローチの適用可能性を検証し、それに合わせてモデルを拡張する必要がある．また、sEMG データは同一の個人であっても測定デバイスの取り付け位置のずれ等により日々変化することが報告されている [192]．本章で用いたデータセットではすべて 1 日内で測定されているため、この影響は排除されている．しかし、日常的に使用される HCI システムでは、日ごとの信号変動に対するロバスト性が求められるため、今後この点を検証する必要がある．

本章では、シングルソースおよびマルチソースアプローチにおいて、すべての訓練データを使用した．しかしながら、訓練に用いるデータを賢く選択することにより、認識性能をさらに向上できる可能性がある．最近、機械学習分野ではデータ中心型人工知能（data-centric artificial intelligence; DCAI）と呼ばれるアプローチが大きな注目を集めている [193]．DCAI とは、ノイズを含む全てのデータを使用するのではなく、品質の高いデータの部分集合を選択的に用いて学習することによって、認識性能を改善する手法である．本章の t-SNE 可視化結果から、一部の被験者が他の大多数の被験者とは異なる特徴を示していることが明らかになった（例えば、図 4.9(B) における被験者 #1, #10, #21）．こうした被験者の外れ値データを訓練時に含めることは、認識性能に負の影響を与える可能性がある．このため、今後の研究課題として、データ選択が FT を適用したマルチソースアプローチに与える影響を評価していく必要がある．

マルチソースアプローチが FT に適していることを本章で確認したが、FT はターゲットのラベル情報を必要とする（つまり FT は教師あり手法に分類される）ため、HCI アプリケーションの運用を通じて徐々に蓄積される新しいラベル無しデータを活用できないという課題が残る．この課題に対して、教師なし深層ドメイン適応（unsupervised deep domain adaptation）の手法が有効であると考えられる．特に、深層相関アラインメント（Deep CORAL）[135] やドメイン敵対的ニューラルネットワーク（DANN）[8] は、既存の学習済みモデルを再利用可能である点から適している．これらの手法は、計算コストが高くなりがちであるが、損失関数の追加やネットワークの一部改変のみでラベルなしのターゲットデータを学習に利用

可能にするため、長期的な HCI アプリケーションの実現に向けて、これらの深層ドメイン適応手法の統合を検討する価値がある。本章はターゲットデータのラベルが存在する状況下での基礎的な知見を提供しているが、教師なし手法の検討は、よりロバストでユーザフレンドリーな HCI アプリケーション開発に向けた次のステップとして重要である。

4.5 結言

本章では、前章で確立された深層識別器とパラメータ空間ベースの転移学習という技術的基盤の上に立ち、複数のソース被験者のデータを最も効果的に活用するための統合戦略について比較検討を行った。特性の異なる 3 つの sEMG データセット (MyoDataset, NinaPro DB5, NinaPro DB2) を用いた網羅的な検証の結果、全ソースデータを一元的に統合して学習するシングルソースアプローチと比較して、被験者ごとに個別のモデルを構築し、それらをアンサンブル学習によって統合するマルチソースアプローチが、転移学習の性能を有意に向上させることを明らかにした。特に、データセットの複雑さや個人差の大小に関わらず、マルチソースアプローチは一貫して高い頑健性を示した。さらに、マルチソースアプローチにおける統合手法として 4 つのアンサンブル戦略を比較した結果、実装の容易さと高い識別性能を兼ね備えたソフト投票法が、実用上最も推奨される選択肢であることを示した。以上の結果より、本章では、深層識別器を用いたパラメータ空間ベースの転移学習においては、「マルチソースアプローチ」を基本戦略とし、その統合手法として「ソフト投票法」を採用する構成を、最も実用的かつ高精度なデータ統合ガイドラインとして結論づける。

第5章 少数ソース被験者問題に対する仮想 ソース被験者生成

5.1 緒言

本研究ではこれまで、生体信号型 HCI の実用化を阻む個人差の問題に対し、機械学習と転移学習の観点から段階的に検討を深化させてきた。第 3 章では、深層識別器とパラメータ空間ベースの転移学習の組み合わせが、生体信号型 HCI の普及の観点から実用的であり、かつ個人差によるデータ分布の変動に対して最も高い適応能力を持つことを示した。続く第 4 章では、この技術基盤の上で複数のソース被験者データを扱う戦略について検討し、被験者ごとに個別のモデルを構築して統合するマルチソースアプローチが、様々な状況において優れた識別精度と頑健性を示すことを明らかにした。これらの検討により、既存のソース被験者データを最大限に活用し、高精度な識別を実現するための実用的なガイドラインは確立されたと言える。

しかしながら、第 4 章までに最適化された転移学習戦略を用いたとしても、実用化においては依然として重大な制約が残されている。それは、転移学習の性能が、利用可能なソース被験者の物理的な数に強く依存するという点である。第 3 章の結果や、多くの先行研究 [79, 140, 194] が示唆するように、転移学習を用いた識別精度はソース被験者数が増加するほど向上し、安定する傾向にある。確かに、現状の転移学習はターゲットユーザのデータ収集負担を大幅に軽減することに成功している。しかし、その有効性を担保するために多数のソース被験者からのデータ収集を必要とするならば、それはターゲットユーザに課されていた負担を、単に多数のソース被験者へと分散させたに過ぎず、データ収集に関するコストを根本的に解決したとは言い難い。そのため、開発初期のプロトタイプ段階や、被験者の確保が困難な希少疾患患者を対象とする場合など、十分な数のソース被験者を確保できない状況下 (少数ソース被験者環境) では、いかに第 4 章までの最適化戦略を用いても、十分な識別精度を達成することは極めて困難となる。したがって、生体信号型 HCI の導入障壁を低減し、その普及を加速させるためには、少数のソース被験者しか利用できない状況下でも高い性能を発揮する転移学習フレームワークの開発が不可欠である。

この「少数ソース被験者問題」を解決するためには、物理的な被験者収集に頼ることなく、アルゴリズムによってデータの多様性を補完する新たなアプローチが必要となる。そこで本章では、実在する少数のソース被験者データから、新たな仮想の被験者データを生成する手法「Virtual Source Assistance (VSA)」を提案する。VSA は、既存データの改変によってデータを増やすデータ拡張 (Data Augmentation: DA) [195, 187, 196, 197] から着想を得ているが、その目的は従来の DA とは本質的に異なる。従来の生体信号識別における DA

は、主にターゲット被験者内のデータ量を増加させることで、被験者内識別における過学習を防ぐ正則化としての役割を果たしてきた [188]。これに対し、本研究が提案する VSA は、ソース被験者の「数」自体を擬似的に拡張することにより、転移学習フレームワークの多様性を高め、被験者間識別の性能を向上させることを主眼としている。すなわち、VSA を導入することで、実際に収集する必要のある物理的なソース被験者数を削減しつつ、転移学習の効果を最大限に引き出すことが期待される。

本章では、VSA の有効性を検証するため、仮想ソースの生成手法として線形補間 (Linear Interpolation) およびガウシアンノイズ (Gaussian Noise) の 2 種類を採用し、比較検討を行う。評価フレームワークには、第 4 章でその有効性が確認された、深層識別器とパラメータ空間ベースの転移学習を用いたマルチソースアプローチを採用する。識別器のアーキテクチャには、前章まで使用してきた ANN (本章以降、他のニューラルネットワークと区別するため MLP と呼称する) に加え、1 次元および 2 次元畳み込みニューラルネットワーク (1D-CNN, 2D-CNN) の 3 種類を用いることで、提案手法のモデル非依存性を評価する。検証データセットには、これまで用いてきた sEMG データ (MyoDataset) に加え、より信号の複雑性が高く個人差が顕著な EEG データ (OpenBMI Dataset [190]) を使用する。動作の客観的な確認が容易な sEMG と、より難易度の高い EEG という性質の異なるデータセットの双方において検証を行うことで、提案手法が特定の生体信号モダリティに依存しない汎用的な解決策であることを実証する。実験の結果、VSA はデータセットや識別器のアーキテクチャに依存せず、実ソースデータのみで訓練した場合と比較して高い識別精度を達成することを示す。特に、生成手法として線形補間を用いた場合において顕著な性能向上が確認された。本手法の確立により、物理的なデータ収集に伴う被験者の身体的・精神的負担を大幅に軽減し、少数のデータからでも高精度な生体信号識別システムの構築が可能であることを実証する。

5.2 仮想ソース支援 (VSA) の概念

VSA の基本コンセプトである「ソース被験者数の増加」は、アンサンブル学習における投票戦略からも着想を得ている [198]。アンサンブル学習は、理論的には 18 世紀に提唱されたコンドルセの陪審定理 (Condorcet's jury theorem) に基づいている [198, 199]。この定理は、個々の投票者が偶然を上回る確率で正しい判断を下す場合、投票者数を増やすことで集団としての判断精度が向上することを数学的に証明したものである。アンサンブル学習においては、各モデル (投票者) がランダムな推測を上回る性能を有する限り、それらを組み合わせることにより単一モデルよりも精度が向上するという原理として利用されている。この原理を踏襲し、VSA では仮想ソース被験者を生成して人工的に被験者数を増やし、実際のソース被験者数が限られている状況でも識別精度の向上を図っている。

コンドルセの定理を実際の問題に適用するには、以下の 3 つの重要な条件を満たす必要があるが、これらを実現するためには転移学習は重要な役割を果たす。第 1 に、各モデルが偶然を上回る精度を持つ必要がある。性能の悪いモデルは集団としての識別精度を低下させて

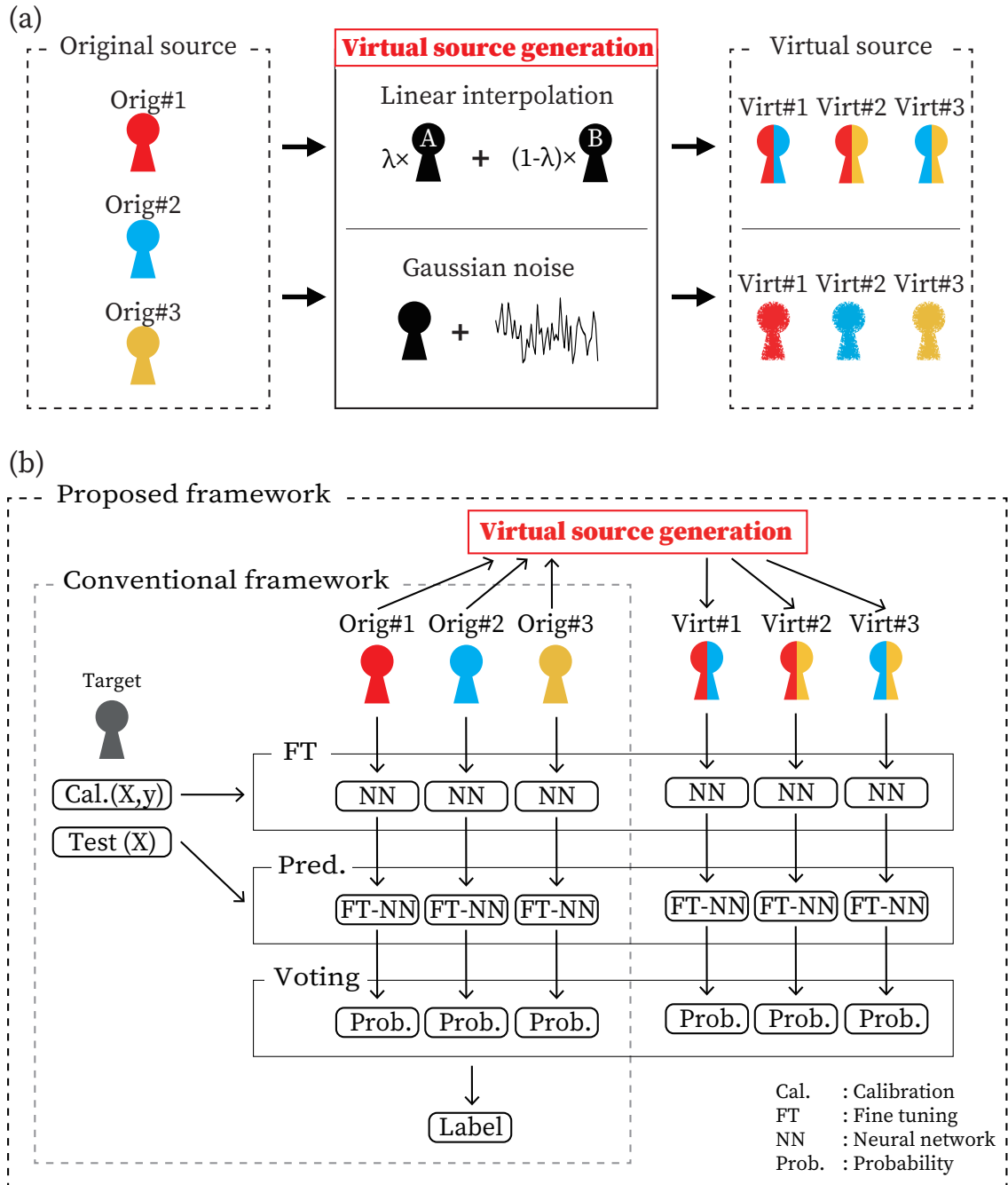


図 5.1: 生体信号識別における転移学習性能を改善するための仮想ソース支援 (Virtual Source Assistance: VSA) の概念図. (a) 実際のソース被験者データを改変して生成された仮想ソース被験者. (b) 従来手法と提案手法の比較. 従来の枠組みでは実ソースデータのみを使用するのに対し, VSA では実ソースと仮想ソースの両方のデータを用いてニューラルネットワーク (NN) を事前学習し, ターゲット被験者のキャリブレーションデータで FT を行い, さらにアンサンブル戦略により統合を行う. 文献 [14] より転載.

しまう。生体信号の識別において、ターゲット被験者以外の被験者から得た事前学習済みモデルをそのまま使用すると、被験者間の差異が大きいため十分な精度が得られないことが多い。転移学習はこの問題を緩和し、各モデルをターゲット被験者に適応させることで識別精度を高める。第2に、各モデルが多様な予測結果を生成する必要がある。複数のモデルが全く同じ予測を行う場合、投票によるメリットは消失する。FTは、高精度を達成しつつモデル間の多様性を保持するための効果的な戦略である。浅い層のパラメータを固定しておくことで、各ソース被験者に固有の特徴抽出能力を維持しつつ、深い層のパラメータを更新することでターゲット被験者に適応できる。第3に、安定した高い識別精度を得るためには被験者数の増加が不可欠である。しかし、実際の被験者からのデータ収集には多大な時間とコストがかかり、被験者に肉体的・精神的負担を与えてしまう。このため、仮想被験者を生成することは従来のデータ収集手法に代わる現実的な選択肢となる。生成されたデータは実際の人間の特徴を再現できていない可能性もあるが、転移学習を用いて統合することにより、その影響を低減し、仮想ソース被験者が効果的に投票を行えるようになると考えられる。

以上の投票を成功させるための3条件のうち、特にVSAにおいて重要となるのは2番目の条件であるソース間の「多様性の確保」である。仮想ソース被験者は実際の被験者データに由来するため、そのデータ特性が元のソース被験者に類似しやすく、多様性を確保することが根本的な課題となる。従来のDA手法は、元データの特性をできるだけ保持することを目的としているため[195, 187, 196, 197]、VSAに必要な仮想ソース被験者の生成には適さない。このため、本章では、十分な多様性を持つ仮想ソースを作成するためには、従来よりも積極的なデータ生成手法が必要になるという仮説を立てた。

本章では、仮想ソース被験者生成の手法として、2人の実際の被験者データ間の「線形補間」を採用した。従来、生体信号を対象とするDAでは、被験者固有の信号特性を変化させてしまうため、このような線形補間を避ける傾向があった。しかし、VSAにおいては、個々の被験者の特性を保つことよりも、複数の被験者間の特徴を融合してソースの多様性を増すことが重要である。この積極的なデータ生成手法の有効性を評価するため、比較対象として従来のマイルドな変換手法である「ガウシアンノイズ」を追加する手法と比較検討した。

5.3 データ準備

5.3.1 データセットおよび前処理

(1) MyoDataset

基本的な内容は3.2.1項で記述済みである。健常者25名を対象とし、単純な前腕動作8種類を5回ずつ計測したsEMGデータを収録している。

前処理は元の手法[9]に従い、高域通過フィルターの周波数をわずかに調整して行った。具体的には、生のsEMGデータに対して、元の前処理で使用された15 Hzではなく、識別に重要な振幅成分を過度に減衰させないために、10 Hzの5次バターワース高域通過フィルターを適用した。動作開始点の検出には、mSampEnを用い、ハイパーパラメータ N_w , N_s ,

m, r は文献 [146, 147] に従って設定した．動作開始点から 1.5 s 間の信号セグメントを 6 s 間の信号から抽出し，その後，250 ms の解析窓を 50 ms ずつオーバーラップさせてスライドさせながら分割した．これにより，各試行はスライディングウィンドウによって 26 セグメントに分割された．

識別器の訓練は被験者ごとに個別に行われ，モデルによって入力構成が異なる．1D-CNN および 2D-CNN の入力データは，「5 試行 $\times$ 26 セグメント，8 チャンネル，50 サンプル点」の形式とした．一方，MLP については，sEMG ベース動作認識で広く用いられている Huang の時間領域および自己回帰特徴量（TDAR 特徴量）を計算して入力とした [149]．具体的には，各チャンネルごとに以下の 11 種類の特徴量を抽出した：(1) 平均絶対値，(2) ゼロ交差回数，(3) 傾き符号変化回数，(4) 波形長，(5) 二乗平均平方根，および (6–11) 6 次の自己回帰係数である．8 チャンネルを用いているため，各セグメントは 88 次元の特徴量ベクトルとなり，MLP の入力形式は「5 試行 $\times$ 26 セグメント，88 特徴量」となった．

(2) OpenBMI Dataset

OpenBMI は，54 名の健康な被験者（年齢：24–35 歳，女性 25 名）から収集されたオープンな EEG データセットである [190]．各被験者について，運動想起（Motor Imagery: MI），事象関連電位（Event-Related Potential: ERP），定常状態視覚誘発電位（Steady-State Visual-Evoked Potential: SSVEP）という 3 つの主要なパラダイムでデータ収集が行われている．本章では，運動想起（MI）に焦点を当てた．各被験者は，(1) 右手の把持想起および (2) 左手の把持想起という 2 種類の運動想起タスクを実施した．タスクの実施順序はランダムであり，各タスクあたり 50 試行，合計 100 試行で 1 ブロックが構成されている．各試行において，62 チャンネルの測定装置（BrainAmp, Brain Products, Germany）を用いて，4 s 間の EEG データが 1000 Hz でサンプリングされた．

データ収集の実験は 4 ブロックから構成されており，第 1 および第 2 ブロックはそれぞれ識別器の訓練とオンラインテストが行われた．また，テスト中，被験者は識別器の出力に基づく視覚的フィードバックを受け取った．第 3 および第 4 ブロックは，それぞれ第 1 および第 2 ブロックと同様の構造であるが，別日に実施された．本章では第 1 ブロックおよび第 2 ブロックのデータのみを使用した．また，一部の被験者は運動想起タスクを効果的に遂行できないことが知られているため，被験者の選定を行った [200]．実際にデータ提供元により，54 名中 33 名が運動想起タスクで十分なパフォーマンスを示さなかったことが報告されている [190]．このため，本章では運動想起タスクを正常に遂行できた 21 名の被験者のみを対象とした．

データの前処理は，元の論文 [190] に詳細に記載された手順に従った．元の 62 チャンネルから運動野を中心に，国際 10-20 法に拡張した電極位置に従い，FC5/3/1/2/4/6, C5/3/1/z/2/4/6, CP5/3/1/z/2/4/6 の 20 チャンネルを選択し，さらなる解析に用いた．元のサンプリング周波数である 1000 Hz から計算負荷を軽減するため，MI 識別に必要な十分な時間分解能を維持した 100 Hz にダウンサンプリングされた [190]．選択された 20 チャンネルの EEG データは，5 次バターワースフィルタを用いて 8–30 Hz のバンドパスフィルタリングを施した．その後，

キュー提示後 1000–3500 ms のセグメントを抽出した。

識別器の訓練は被験者ごとに実施され、モデルごとに入力形式が異なる。1D-CNN および 2D-CNN については、入力データの次元を「2 ブロック× 100 試行, 20 チャンネル, 250 サンプル点」とした。MLP では、MI 識別で広く用いられる Filter Bank Common Spatial Pattern (FBCSP) 特徴量を用いた [201]。フィルタバンクは 6 つの周波数帯域 (8–12, 12–16, 16–20, 20–24, 24–28, 28–30 Hz) で構成し、各周波数帯域においてプロジェクション行列の最大および最小の 2 つずつ、計 4 つのフィルタから対数分散特徴量を抽出した。このプロセスにより、6 つの周波数帯域と各帯域 4 つのフィルタから計 24 個の特徴量が得られた。結果として MLP の入力形式は、「2 ブロック × 100 試行, 24 特徴量」となった。

5.3.2 データの分割

ある被験者がターゲットとして選択された場合、そのターゲット被験者のデータはキャリブレーションデータとテストデータに分割される (図 5.1(b) 参照)。MyoDataset では、第 1 試行がキャリブレーションデータとして使用され、第 2–5 試行がテストデータとして使用された。OpenBMI Dataset では、第 1 ブロックがキャリブレーションデータ、第 2 ブロックがテストデータとして使用された。キャリブレーションデータはバックボーン NN の FT に使用され、テストデータは識別精度の評価に用いられた。一方、ある被験者がソースとして指定された場合、そのソース被験者の全データが分割されずに使用された。ソースデータはバックボーン NN の事前訓練および仮想ソース生成に使用された。

5.4 手法

5.4.1 識別フレームワークの概要

図 5.1(b) に、従来型および VSA によって拡張した識別フレームワークの処理手順を示している。従来のフレームワークでは、バックボーン NN がソース被験者ごとに個別に訓練される。例えば、ソース被験者が 3 名の場合、3 つの異なるバックボーン NN が作成される。このように事前学習されたバックボーン NN は、ターゲット被験者のキャリブレーションデータを用いて FT される。FT 済みの各 NN はターゲットのテストデータを処理し、それぞれのモデルから予測確率が算出される。最後に、これらの確率が集約され、多数決投票によりクラスラベルが決定される。

VSA の手順は従来のフレームワークと類似しているが、重要な違いは「仮想ソース被験者」の追加である。VSA では、あらかじめ実際のソース被験者から仮想ソース被験者が生成される。例えば図 5.1(b) に示すように、3 名の仮想ソース被験者が VSA により生成される。仮想ソースのデータは実際のソースデータと組み合わせられ、計 6 名 (実際の 3 名 + 仮想 3 名) のソースを用いてクラスラベルが推定される。詳細な手順は後述する。

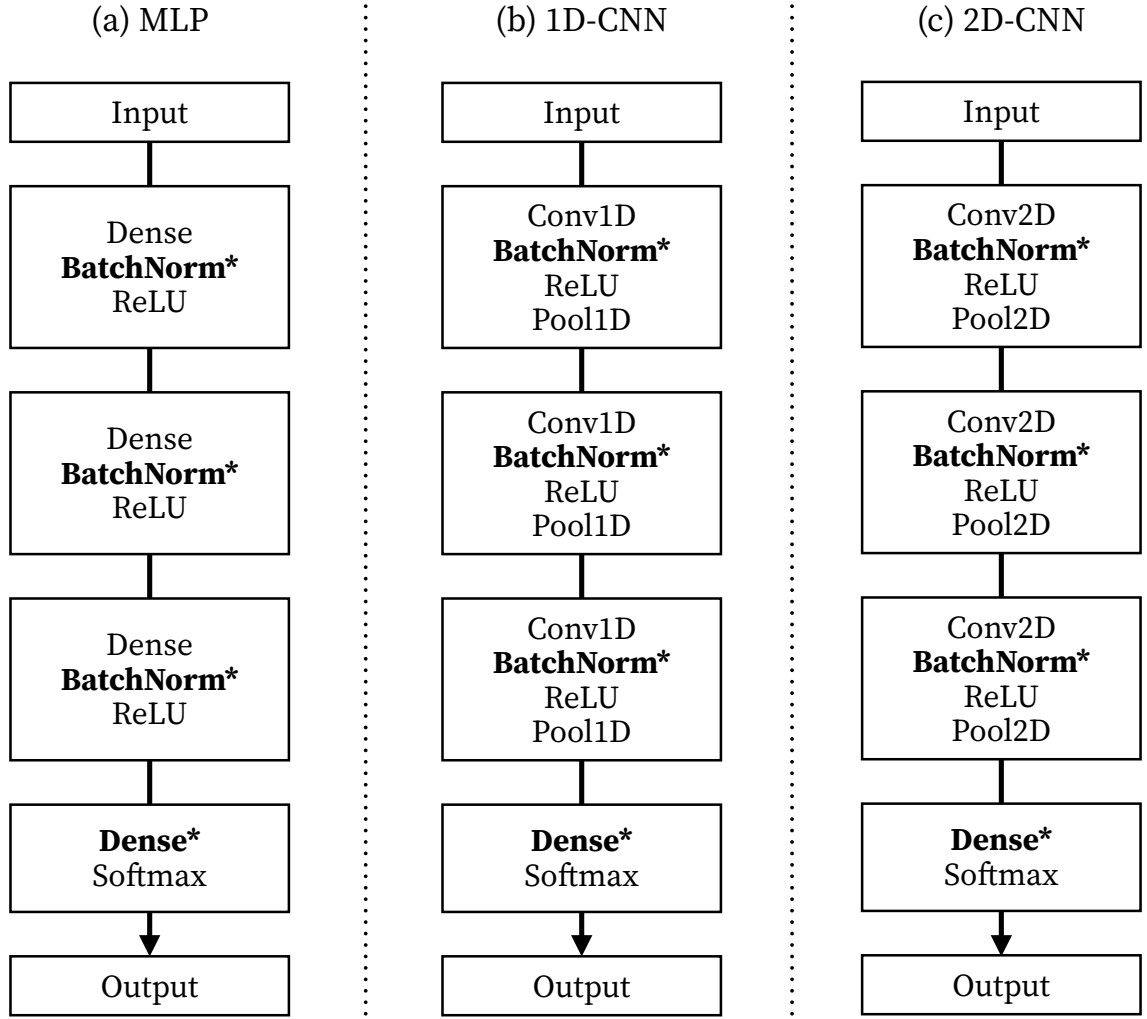


図 5.2: ニューラルネットワーク (NN) の構成. (a) 多層パーセプトロン (MLP), (b) 1 次元畳み込みニューラルネットワーク (1D-CNN), (c) 2 次元畳み込みニューラルネットワーク (2D-CNN). 「*」の記号は, FT の過程で更新されたパラメータを示す. 文献 [14] より転載.

5.4.2 識別器

本章では, 事前学習のためのバックボーン NN として, MLP, 1D-CNN, および 2D-CNN の 3 種類の NN を用いた. 結果の汎用性を高めるため, 図 5.2 に示すようなシンプルなアーキテクチャを採用した. MLP は各データセットごとに抽出した特徴量 (sEMG では TDAR 特徴量, EEG では FBCSP 特徴量) を入力として処理した. 一方, CNN は前処理済みの信号データを直接処理した. 特に指定がない限り, 各モデルの実装は PyTorch を用い, デフォルトのハイパーパラメータを使用した. 比較的小規模な訓練データセットであるため, すべてのモデルは Adam オプティマイザを用いたフルバッチ学習を 1000 イテレーション実行して最適化した. また, 過学習を防止するために, 重み減衰 $w = 0.00001$ を適用した.

(1) MLP

図 5.2(a) に MLP の構造を示す。この NN は 3 層の隠れ層からなり、各層のニューロン数は入力特徴量の次元数を D として $[2D, D, D/2]$ と設定した。各隠れ層では、全結合層 (Dense) 後にバッチ正規化と ReLU 活性化関数を順に適用した。最終層は識別のためのソフトマックス活性化を持つ全結合層であり、ニューロン数は識別対象となるクラス数に対応する (MyoDataset では 8, OpenBMI Dataset では 2)。

(2) 1D-CNN

図 5.2(b) に 1D-CNN の構造を示す。このモデルは、時間情報を処理するために 1 次元畳み込み (1D convolution) を適用する 3 層の畳み込み層からなる。各畳み込み層のフィルタ数はそれぞれ $[16, 32, 64]$ と設定した。1D 畳み込み層のカーネルサイズは、sEMG 信号では 3, EEG 信号では 4 に設定されており、これは EEG 信号 (250 時点) が sEMG 信号 (50 時点) に比べて長い時間的長さを持つことに対応するためである。各畳み込み層の後には、バッチ正規化, ReLU 活性化, 1D 平均プーリングを順番に適用した。1D 平均プーリング層のカーネルサイズは、sEMG 信号では 2, EEG 信号では 3 と設定した。最終層は識別のためのソフトマックス活性化を備えた全結合層であり、ニューロン数は識別クラス数に等しい。

(3) 2D-CNN

図 5.2(c) に 2D-CNN の構造を示す。このモデルは、時間的・空間的な情報を同時に処理するために 2 次元畳み込み (2D convolution) を適用する 3 層の畳み込み層からなる。各畳み込み層のフィルタ数はそれぞれ $[16, 32, 64]$ と設定した。2D 畳み込み層のカーネルサイズは、sEMG 信号では 3×3 , EEG 信号では 4×4 と設定した。各畳み込み層の後には、バッチ正規化, ReLU 活性化, 2D 平均プーリングを順番に適用した。平均プーリング層のカーネルサイズは sEMG 信号では 1×2 , EEG 信号では 1×3 である。チャンネル数は時点数に比べて小さいため、チャンネル方向にはプーリングを行っていない。最終層は識別のためのソフトマックス活性化を持つ全結合層であり、ニューロン数は識別クラス数に対応している。

5.4.3 仮想ソースの生成

実際のデータから仮想ソース被験者を生成するために、本章では「線形補間」と「ガウシアンノイズの付加」という 2 つの手法を検討する。両手法とも新たな仮想ソースの生成を目的としているが、線形補間はガウシアンノイズの付加よりも多様性に富んだ仮想ソースを生成することが期待される。

表 5.1: 生成された仮想被験者数. 文献 [14] より転載.

No. real subjects	No. linear interpolation virtual subjects	No. Gaussian noise virtual subjects
2	1 (3)	2 (4)
3	3 (6)	3 (6)
4	6 (10)	4 (8)
5	10 (15)	5 (10)
6	15 (21)	6 (12)
7	21 (28)	7 (14)
8	28 (36)	8 (16)
9	36 (45)	9 (18)
10	45 (55)	10 (20)

括弧内の数値は実被験者と仮想被験者の合計数.

(1) 線形補間

仮想ソース被験者を生成するため, 実際の被験者データのペア (被験者 A と被験者 B) に対して線形補間を適用する. ある被験者の前処理済み信号を $x_{i,j}$ とし, ここで i は i 番目のチャンネル, j は j 番目の時点を表す. 新たな仮想ソースデータ $x_{i,j}^*$ は以下のように生成される:

$$x_{i,j}^* = \lambda x_{i,j}^{A,c} + (1 - \lambda) x_{i,j}^{B,c} \quad (5.1)$$

ここで, c はクラスラベル, λ は 2 つの信号の混合比を調整するハイパーパラメータである. 本章では両被験者の特徴を均等に反映させるため, λ を 0.5 に設定した. 両データセットとも各被験者は同一数の試行を行っているため, 同一クラス内の対応する試行データ同士で補間が行われた. 例えば, 被験者 A のクラス 1 における第 1 試行データと, 被験者 B のクラス 1 における第 1 試行データが結合される. このように生成された信号は, 1D-CNN および 2D-CNN の事前学習に用いられた. MLP の場合は, 生成されたデータを特徴量抽出した後に事前学習を行った. この方法により, すべての実被験者ペアの組み合わせに対して仮想ソースを作成することが可能となる (表 5.1 を参照). 本章では, この線形補間手法を LI-VSA と表記する.

(2) ガウシアンノイズ

実際のソース被験者の前処理済み信号に弱いガウシアンノイズを付加することで, 仮想ソースデータを生成する. 仮想ソースデータ $x_{i,j}^*$ は, 以下のようにガウシアンノイズ $n_{i,j}$ を加えることで生成された:

$$\begin{aligned}x_{i,j}^* &= x_{i,j} + n_{i,j}, \\ n_{i,j} &\sim \mathcal{N}(0, \sigma^2).\end{aligned}\tag{5.2}$$

ここで、 $\mathcal{N}$ はガウス分布を示し、 σ^2 は分布の分散を表す。sEMG [187, 195] および EEG [196, 197] に関する先行研究を考慮して、モダリティごとに異なる分散の値を用いた。具体的には、sEMG 信号については $\sigma^2 = x_{i,j}^2/30$ 、EEG 信号については $\sigma^2 = 0.16$ と設定した。

生成された信号は 1D-CNN および 2D-CNN の事前学習に用いられた。MLP の場合は、生成データを手作り特徴量に変換した後、事前学習を実施した。従来の DA 手法は生成データを元データと混合して使用するが、本章の VSA 手法では生成データを独立した被験者由来のデータとして扱う。この方法においては、各実際の被験者に対してそれぞれガウシアンノイズを付加することで、実際の被験者数と同数の仮想ソースが生成された（表 5.1 参照）。本章ではこのガウシアンノイズに基づく手法を GN-VSA と表記する。

5.4.4 転移学習およびアンサンブル戦略

(1) ファインチューニング (FT)

FT は広く用いられている転移学習手法であり、事前に学習した NN のパラメータをターゲット被験者のデータを用いて再学習する [139, 142]。本章ではキャリブレーションデータの量が限られているため、再学習において NN のパラメータの一部のみを更新した（図 5.2 のアスタリスクで示した部分）。それ以外の重み減衰（weight decay）やオプティマイザなどのハイパーパラメータは事前学習時と同じ設定を維持した。

(2) アンサンブル戦略

複数の FT 済み NN の予測を統合するため、多数決投票（Majority Voting）を適用した。各 FT 済み NN は、入力データが各クラスに属する確率を表す予測確率を算出する。全ての NN から得られた全クラスの予測確率を合算し、その合計確率が最も高いクラスを識別結果として選択した。この多数決投票は第 3 の SV である。

5.5 評価

本章では、VSA が限られた数のソースで高い性能を示すことを仮説として設定した。そのため、利用可能な実ソース被験者数を 2 名から 10 名まで変化させた 9 つのシナリオにおいて性能評価を行った。識別実験では、各被験者をターゲットとし、それ以外の被験者をソースとして扱った。ソース被験者の組み合わせは膨大であるため、候補者の中からソース被験者をランダムに選択した。このランダムな選択により、特定のターゲット被験者に対して相性の良い、または悪いソースの組み合わせが生じる可能性がある。このランダム性の影響を

抑えるため、重複しない 20 通りのソース被験者セットを生成した。例えば、MyoDataset には 25 名の被験者データがあり、1 名をターゲットとして選択した場合、残りの 24 名がソース被験者候補となる。ソース被験者数を 3 名に設定した場合、24 名からランダムに 3 名を選び、これを 20 通り作成した。このように、ターゲット被験者の選択は 25 通りあり、ソース被験者数は 2 名から 10 名までの 9 段階、それぞれを 20 回繰り返すことで、合計 $25 \times 9 \times 20$ 回の評価を実施した。OpenBMI Dataset についても同様の方法を適用し、合計 $21 \times 9 \times 20$ 回の評価を行った。以上の評価方法をとることで、誰がソースになるかわからないという実環境の不確実性を含んだ検証が可能になる。

ターゲットおよびソース被験者を用いて、従来のフレームワークと VSA を適用したフレームワークの識別精度を比較した。従来のフレームワークでは実ソース被験者のみを用いたが、VSA フレームワークでは実ソースおよび仮想ソース被験者を併用した。また、仮想ソース生成手法の特性を明らかにするために、線形補間を用いた LI-VSA とガウシアンノイズを用いた GN-VSA の識別精度も評価した。公平な比較を行うため、特に実ソースが 3 名の場合を対象として比較した。実ソース 3 名の場合、LI-VSA と GN-VSA が生成する仮想ソースの数が等しくなるため、仮想ソース数の違いによる影響を排除した状態で、仮想ソース生成手法そのものの性能を直接的に評価することが可能となる（表 5.1 参照）。

データセットごとに、二要因反復測定分散分析を実施した。要因は、仮想ソース追加手法（従来手法、LI-VSA、GN-VSA）および NN モデル（MLP、1D-CNN、2D-CNN）の 2 つである。特に本章では VSA 手法の有効性の差異に焦点を当てているため、主に仮想ソース追加手法の違いを評価した。仮想ソース追加手法に有意な差が認められた場合、Bonferroni 補正を用いた対応のある t 検定を実施し、各フレームワーク間の比較を行った。統計検定における検出力の過大評価を防ぐため、各ターゲット被験者における 20 回の繰り返し結果を平均化し、1 つの平均値として分析した。統計検定における有意水準は 0.05 とした。

5.6 結果

本節では、提案手法である VSA の有効性を評価する。まず、本研究における実験設定および実装した識別器の妥当性を確認するため、基準となる被験者内識別の結果について述べる。次に、本研究の主要な結果である VSA を用いた転移学習の識別精度を示し、最後に被験者内モデルをアンサンブルに組み込んだ際の効果について述べる。

5.6.1 被験者内識別精度

各ターゲット被験者のデータのみを用いてモデルを訓練・評価する被験者内識別を実施した結果を表 5.2 に示す。MyoDataset においては、MLP が $89.6 \pm 5.8\%$ 、1D-CNN が $91.6 \pm 5.9\%$ 、2D-CNN が $92.2 \pm 5.5\%$ の精度を達成した。第 3 章における MLP（TDAR 特徴量）の精度 90.3% と比較しても同等の水準であり、CNN モデルはそれを上回る結果を示した。一方、OpenBMI Dataset においては、MLP が $87.3 \pm 11.2\%$ と最も高い精度を示

表 5.2: 被験者内識別精度. 文献 [14] より転載.

Model	MyoDataset	OpenBMI Dataset
MLP	89.6 ± 5.8	87.3 ± 11.2
1D-CNN	91.6 ± 5.9	77.7 ± 13.3
2D-CNN	92.2 ± 5.5	72.7 ± 13.9

各値は, 分類精度 (%) と標準偏差の組で表記した.

し, 1D-CNN は $77.7 \pm 13.3\%$, 2D-CNN は $72.7 \pm 13.9\%$ であった. 先行研究 [190] で報告されている FBCSP と LDA を用いた精度 (80.0%) と比較し, 本実験の MLP はこれを上回る結果を得た.

5.6.2 VSA の導入効果

図 5.3 に MyoDataset および OpenBMI Dataset における識別精度を示し, それぞれの精度指標の詳細を表 5.3 および 5.4 にまとめた. 図中の各マーカーおよび表中の各セルは, すべてのターゲット被験者について 20 回の繰り返しを行った平均精度を示す. 例えば, MyoDataset においては 25 名の各被験者について 20 回の繰り返しを行い, 合計で 500 回の識別結果を用いて平均値を算出している. MyoDataset においては MLP および 2D-CNN が高い性能を示したが, 1D-CNN の性能は最も低かった (図 5.3(a)). 一方, OpenBMI Dataset においては MLP が最も高い性能を示し, 2D-CNN の性能は最も低かった (図 5.3(b)). いずれのデータセットにおいても, 実ソース被験者数を増やすことで識別精度は向上した.

図 5.3 に示すように, VSA は従来のフレームワークよりも高い性能を示し, 中でも LI-VSA は評価した手法の中で最高の識別精度を達成した. 従来のフレームワークは, いずれの識別器およびデータセットでも 10 名の実被験者データを使用したときに最高の精度に達したが, VSA はそれよりも少ない被験者数でより高い精度を実現した. 具体的には, MyoDataset において, VSA は平均 5.3 名 (MLP で 5 名, 1D-CNN で 6 名, 2D-CNN で 5 名) の実被験者データで, 従来の 10 名の実被験者を用いた場合の精度を超えた. また, OpenBMI Dataset でも, 全 NN モデルにおいて平均 4.0 名の実被験者データで, 従来の 10 名の実被験者を用いた場合の精度を上回った. これらの結果は, LI-VSA が実被験者数を減らしつつも, 実被験者を多く用いた場合と同等以上の精度を達成できることを示している. このことから, VSA は実被験者データが限られている状況を効果的に補い, 高い識別精度を維持できることが示唆された.

両データセットに対して二元配置反復測定分散分析を実施した. MyoDataset では, 仮想ソース追加手法 ($F(2, 48) = 74.339, p < 0.001$), NN モデル ($F(2, 48) = 24.958, p < 0.001$) の主効果およびその交互作用 ($F(4, 96) = 6.198, p < 0.001$) において有意差が認められた. また OpenBMI Dataset でも, 仮想ソース追加手法 ($F(2, 40) = 37.7986, p < 0.001$), NN モデル ($F(2, 40) = 44.044, p < 0.001$) の主効果およびその交互作用 ($F(4, 80) = 18.085, p < 0.001$) において有意差が認められた.

さらに, 仮想ソース追加手法間の差を明らかにするため, Bonferroni 補正を用いた対応のあ

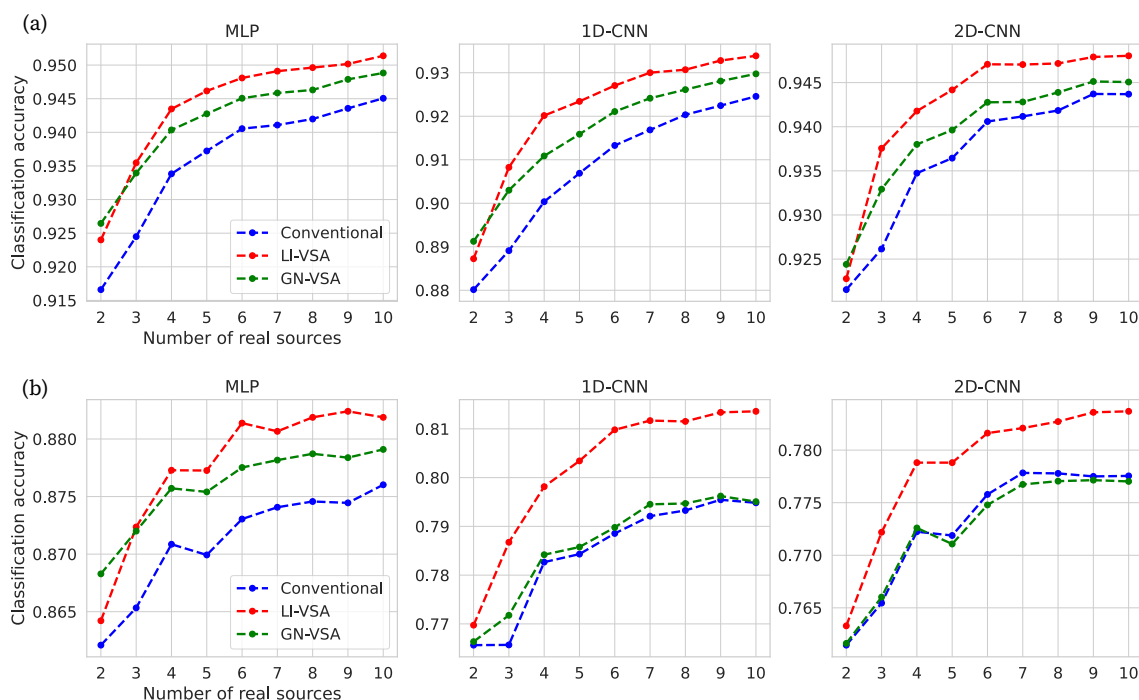


図 5.3: (a) MyoDataset および (b) OpenBMI Dataset における識別精度を示す。識別精度は、識別に使用された実ソース被験者の数に応じて示されている。各マーカーは、全ターゲット被験者に対して 20 回の試行を平均した結果を表す。文献 [14] より転載。

る t 検定を実施した。その結果、MyoDataset においては、従来手法と GN-VSA 間 ($t(74) = -9.061$, 補正後 $p < 0.001$), 従来手法と LI-VSA 間 ($t(74) = -13.201$, 補正後 $p < 0.001$), GN-VSA と LI-VSA 間 ($t(74) = -4.613$, 補正後 $p < 0.001$) の全ての組み合わせで有意差が確認された。同様に、OpenBMI Dataset においても、従来手法と GN-VSA 間 ($t(62) = -5.298$, 補正後 $p < 0.001$), 従来手法と LI-VSA 間 ($t(62) = -7.706$, 補正後 $p < 0.001$), GN-VSA と LI-VSA 間 ($t(62) = -5.365$, 補正後 $p < 0.001$) の全ての組み合わせで有意差が示された。これらの統計解析により、VSA は従来のフレームワークよりも有意に優れており、特に LI-VSA が仮想ソース生成手法として非常に優れていることが明らかになった。

仮想被験者の追加が識別性能の向上にどの程度貢献したかを評価するために、仮想と実際の被験者数を合計した総ソース数と識別精度との関係进行分析した (図 5.4)。その結果、データセットおよび NN モデルにかかわらず、ソース被験者数が増加するほど識別精度が向上するという明確な傾向が確認された。MyoDataset においては、従来のフレームワークが最も急激な精度改善を示した。これは、従来のフレームワークが実際のソース被験者のみを用いているため、当然の結果と考えられる。一方、LI-VSA においては、総ソース数が 10 名の場合でも、表 5.1 に示すように実際のソース被験者は 4 名のみである。しかしながら、LI-VSA が示す精度向上の曲線は従来のフレームワークに極めて近いものとなった。OpenBMI Dataset では、総ソース数が 2 から 10 名の範囲において、従来手法と LI-VSA の精度向上の曲線がより近い傾向を示した。これらの結果は、LI-VSA で生成した仮想ソースが、実際のソース

表 5.3: 実ソース被験者数を変化させた場合の MyoDataset における精度. 文献 [14] より転載.

No. real sources	MLP			1D-CNN			2D-CNN		
	Conventional	LI-VSA	GN-VSA	Conventional	LI-VSA	GN-VSA	Conventional	LI-VSA	GN-VSA
2	91.7 $\pm$ 5.4	92.4 $\pm$ 5.1	92.6 $\pm$ 5.0	88.0 $\pm$ 7.0	88.7 $\pm$ 6.9	89.1 $\pm$ 6.6	92.2 $\pm$ 5.9	92.3 $\pm$ 5.8	92.4 $\pm$ 5.7
3	92.4 $\pm$ 5.2	93.5 $\pm$ 4.9	93.4 $\pm$ 4.8	88.9 $\pm$ 6.7	90.8 $\pm$ 6.2	90.3 $\pm$ 6.2	92.6 $\pm$ 5.5	93.8 $\pm$ 4.9	93.3 $\pm$ 5.1
4	93.4 $\pm$ 4.8	94.3 $\pm$ 4.6	94.0 $\pm$ 4.6	90.0 $\pm$ 6.2	92.0 $\pm$ 5.4	91.1 $\pm$ 5.8	93.5 $\pm$ 4.9	94.2 $\pm$ 4.6	93.8 $\pm$ 4.7
5	93.7 $\pm$ 4.8	94.3 $\pm$ 4.6	94.6 $\pm$ 4.6	90.7 $\pm$ 5.8	92.3 $\pm$ 5.4	91.6 $\pm$ 5.6	93.6 $\pm$ 4.8	94.4 $\pm$ 4.4	94.0 $\pm$ 4.7
6	94.1 $\pm$ 4.8	94.8 $\pm$ 4.6	94.5 $\pm$ 4.6	91.3 $\pm$ 5.6	92.7 $\pm$ 5.2	92.1 $\pm$ 5.4	94.1 $\pm$ 4.5	94.7 $\pm$ 4.1	94.3 $\pm$ 4.4
7	94.1 $\pm$ 4.7	94.9 $\pm$ 4.6	94.6 $\pm$ 4.6	91.7 $\pm$ 5.5	93.0 $\pm$ 5.1	92.4 $\pm$ 5.3	94.1 $\pm$ 4.4	94.7 $\pm$ 4.1	94.3 $\pm$ 4.3
8	94.2 $\pm$ 4.8	95.0 $\pm$ 4.6	94.6 $\pm$ 4.6	92.0 $\pm$ 5.4	93.1 $\pm$ 5.1	92.6 $\pm$ 5.3	94.2 $\pm$ 4.4	94.7 $\pm$ 4.1	94.4 $\pm$ 4.4
9	94.4 $\pm$ 4.7	95.0 $\pm$ 4.6	94.8 $\pm$ 4.5	92.2 $\pm$ 5.4	93.3 $\pm$ 5.1	92.8 $\pm$ 5.3	94.4 $\pm$ 4.3	94.8 $\pm$ 4.1	94.5 $\pm$ 4.3
10	94.5 $\pm$ 4.7	95.1 $\pm$ 4.5	94.9 $\pm$ 4.5	92.5 $\pm$ 5.3	93.4 $\pm$ 5.1	93.0 $\pm$ 5.2	94.4 $\pm$ 4.2	94.8 $\pm$ 4.1	94.5 $\pm$ 4.2

3 種類のニューラルネットワーク (MLP, 1D-CNN, 2D-CNN) における従来手法, LI-VSA, および GN-VSA の結果. 数値は分類精度 (%) とその標準偏差.

表 5.4: 実ソース被験者数を変化させた場合の OpenBMI Dataset における精度. 文献 [14] より転載.

No. real sources	MLP			1D-CNN			2D-CNN		
	Conventional	LI-VSA	GN-VSA	Conventional	LI-VSA	GN-VSA	Conventional	LI-VSA	GN-VSA
2	86.2 $\pm$ 10.8	86.4 $\pm$ 10.7	86.8 $\pm$ 10.5	76.6 $\pm$ 10.8	77.0 $\pm$ 11.0	76.6 $\pm$ 10.8	76.1 $\pm$ 11.9	76.3 $\pm$ 11.7	76.2 $\pm$ 11.9
3	86.5 $\pm$ 10.4	87.2 $\pm$ 10.4	87.2 $\pm$ 10.2	76.6 $\pm$ 11.1	78.7 $\pm$ 11.1	77.2 $\pm$ 10.9	76.5 $\pm$ 11.6	77.2 $\pm$ 11.6	76.6 $\pm$ 11.6
4	87.1 $\pm$ 10.4	87.7 $\pm$ 10.4	87.6 $\pm$ 10.2	78.3 $\pm$ 11.0	79.8 $\pm$ 10.8	78.4 $\pm$ 11.0	77.2 $\pm$ 11.6	77.9 $\pm$ 11.3	77.3 $\pm$ 11.6
5	87.0 $\pm$ 10.4	87.7 $\pm$ 10.2	87.5 $\pm$ 10.2	78.4 $\pm$ 10.9	80.3 $\pm$ 10.9	78.6 $\pm$ 11.0	77.2 $\pm$ 11.8	77.9 $\pm$ 11.4	77.1 $\pm$ 11.8
6	87.3 $\pm$ 10.2	88.1 $\pm$ 10.2	87.8 $\pm$ 10.0	78.9 $\pm$ 10.9	81.0 $\pm$ 10.6	79.0 $\pm$ 10.6	77.6 $\pm$ 11.5	78.2 $\pm$ 11.4	77.5 $\pm$ 11.5
7	87.4 $\pm$ 10.3	88.1 $\pm$ 10.3	87.8 $\pm$ 10.1	79.2 $\pm$ 10.4	81.2 $\pm$ 10.7	79.5 $\pm$ 10.4	77.8 $\pm$ 11.3	78.2 $\pm$ 11.3	77.7 $\pm$ 11.4
8	87.5 $\pm$ 10.4	88.2 $\pm$ 10.3	87.9 $\pm$ 10.2	79.3 $\pm$ 10.5	81.2 $\pm$ 10.8	79.5 $\pm$ 10.5	77.8 $\pm$ 11.5	78.3 $\pm$ 11.3	77.7 $\pm$ 11.5
9	87.4 $\pm$ 10.3	88.2 $\pm$ 10.2	87.8 $\pm$ 10.1	79.5 $\pm$ 10.4	81.3 $\pm$ 10.6	79.6 $\pm$ 10.4	77.8 $\pm$ 11.4	78.4 $\pm$ 11.2	77.7 $\pm$ 11.5
10	87.6 $\pm$ 10.3	88.2 $\pm$ 10.3	87.9 $\pm$ 10.1	79.5 $\pm$ 10.7	81.4 $\pm$ 10.8	79.5 $\pm$ 10.6	77.8 $\pm$ 11.6	78.4 $\pm$ 11.3	77.7 $\pm$ 11.7

3 種類のニューラルネットワーク (MLP, 1D-CNN, 2D-CNN) における従来手法, LI-VSA, および GN-VSA の結果. 数値は分類精度 (%) とその標準偏差.

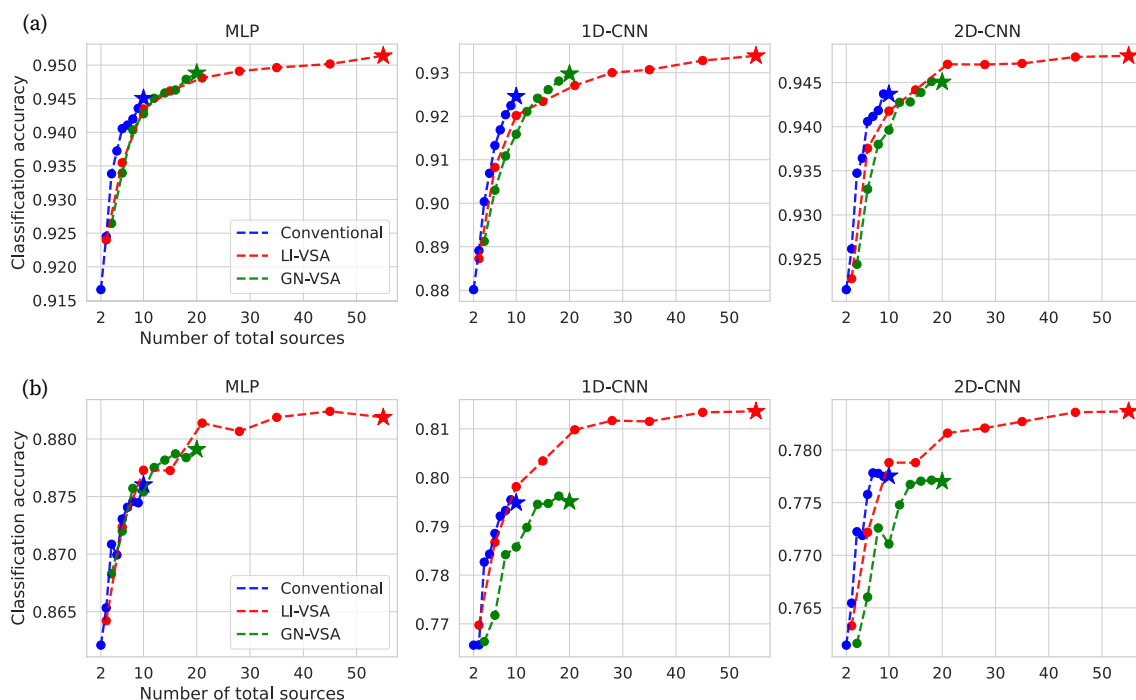


図 5.4: (a) MyoDataset および (b) OpenBMI Dataset における識別精度．識別精度は，実ソースおよび仮想ソースを組み合わせたソース被験者数に応じて示されている．各マーカーは，全ターゲット被験者に対する 20 回の試行の平均を表す．星印は，実ソース被験者が 10 名の場合の結果を示している．文献 [14] より転載．

被験者とほぼ同等の有効性で識別性能の向上に貢献していることを示している．

5.6.3 被験者内モデルの統合効果

前項までの結果では，FT モデルのみのアンサンブル評価を行った．しかし，表 5.3, 5.4 と表 5.2 を比較すると，ソース被験者数が極端に少ない場合など，状況によっては FT よりも被験者内識別器の方が高い精度を示すケースが確認された．そこで，FT モデルのアンサンブルに被験者内モデルを投票者として組み込んだ場合の精度を検証した．

図 5.5 にその結果を示す．被験者内 NN を導入することで，特にソース被験者数が少ない領域において分類精度の向上が確認された．一方で，ソース被験者数が増加するにつれて，被験者内 NN を組み込むことによる利得は減少した．例えば，ソース被験者数が 10 名の場合では，被験者内 NN の有無による精度差はわずか 0.1% 程度に留まった．

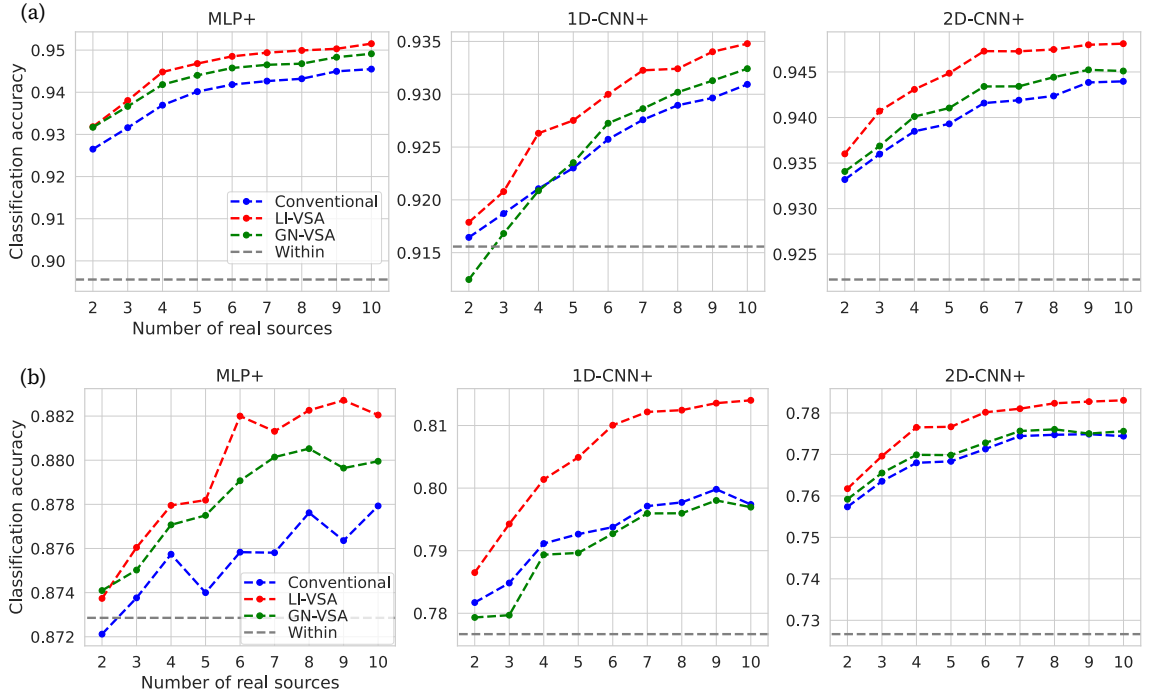


図 5.5: (a) MyoDataset および (b) OpenBMI Dataset における分類精度を示す．記号「+」は、アンサンブル戦略に被験者内 NN を含めていることを表す．文献 [14] より転載．

5.7 考察

5.7.1 識別器のアーキテクチャと被験者内識別精度

本章の結果 (Section 5.6.1) において、データセットによって識別器の性能傾向が異なる現象が確認された．各データセットごとに、その性能差が何に起因するのかを考察する．

MyoDataset においては、NN のアーキテクチャ毎に OpenBMI Dataset ほど大きな差はなかったが CNN との相性が良く、特に 2D-CNN が最高精度を達成した．この要因として、1D-CNN が主に信号の時間的特徴に焦点を当てて学習するのに対し、2D-CNN は畳み込み処理を通じて時間的特徴と空間的特徴の双方を同時に抽出できた点が挙げられる．本研究で用いた MyoDataset はリング状の多電極センサにより取得されており、各チャンネルの信号は独立しているのではなく、電極配置に基づいた空間的な相関を有している．2D-CNN は、その畳み込みにおいて、時間方向のみならず空間方向の畳み込みが可能であり、この電極間の空間的配置を含んだ情報を効果的に学習できたため、より高い識別性能を実現したと推察される．

一方、OpenBMI Dataset (EEG) においては、逆に MLP が最も高い精度を示し、CNN の性能は低調であった．これは、EEG 信号が本来もつ非線形的な特性を MLP がより適切に捉えることができるためであり、先行研究 [190] で採用されている FBCSP と LDA の組み合わせでは捉えきれない情報を学習できた可能性がある．一方で、本研究で実装した CNN は非線形分類器であるにもかかわらず、先行研究のモデルよりも低い精度となった．この

予想外の結果は、本研究で用いた CNN アーキテクチャが比較的単純であることに起因する可能性が高い。実際、Lawhern ら [202] および Liu ら [203] は EEG 信号処理に対して、より高度な CNN アーキテクチャを採用している。Lawhern ら [202] は depthwise ならびに separable convolution [204] を組み込み、Liu ら [203] は分岐型ネットワーク、ショートカット構造、DA など高度な構造を導入している。こうしたアーキテクチャ上の工夫は、複雑な EEG 信号を直接処理する上で極めて重要であると考えられる。また、特に 2D-CNN が 1D-CNN よりも悪い結果となったのは、今回のデータ配列に起因する可能性がある。本データセットの電極の配置は、国際的に認知された拡張 10-20 法に基づくものであるが、電極の物理的な位置関係を保存するようなマッピングを行わず、配布されたままのチャンネルインデックス順に配列した。そのため、2D-CNN の畳み込みカーネルが、物理的には離れた位置にある電極間の相関を誤って「局所的な特徴」として学習してしまい、これがノイズとして機能したため精度が低下したと推察される。

以上の通り、データセットの特性とモデル構造の適合性によって、ベースラインとなる識別精度には差異が認められた。しかしながら、本研究の主眼は個々の識別器の構造最適化ではなく、提案手法である VSA による転移学習の有効性検証にある点に留意する必要がある。実際に第 5.6 節の結果が示す通り、VSA はいずれのデータセット、いずれのアーキテクチャ (MLP, 1D/2D-CNN) においても一貫して精度の向上をもたらした。特に、OpenBMI における 2D-CNN のように、ベースライン精度が十分に高くない条件下においてさえも VSA による改善効果が確認された事実は、むしろ提案手法の汎用性と堅牢性を裏付けるものである。したがって、ここで議論した識別器アーキテクチャへの依存性は、本研究が提案する VSA フレームワークの有効性を損なうものではない。

5.7.2 VSA の有効性と特徴

仮想ソースからのデータを組み込むことで識別精度が著しく向上し、その効果は実際のソース被験者を追加することと同程度であった。図 5.3 に示したように、統計的検定によっても確認された通り、LI-VSA は限られた実ソース被験者数の条件下でも従来手法より有意に優れた性能を示した。特に、LI-VSA は平均 4.7 名 (sEMG で 5.3 名、EEG で 4.0 名) の実ソース被験者を使用するだけで、従来手法が 10 名の実ソース被験者を用いた場合の性能を超えた。表 5.1 に示すように、LI-VSA は 4 名の実ソース被験者から 6 名の仮想ソースを生成し、合計 10 名のソースを構築した。また、図 5.4 に示した通り、従来手法と LI-VSA の性能曲線が類似していることから、仮想ソースを追加することは実ソースを追加するのと同等の効果を持つことが示唆される。このことから、VSA は実際のデータ収集の代替手段となり、実ソース被験者からデータ収集を行う際の身体的・精神的負担を大きく軽減できると考えられる。

仮想ソースを生成する際は、それらが実際のソース被験者とは異なる多様な特性を持つことが重要である。本章の統計分析により、LI-VSA が GN-VSA を上回る性能を示したことから、実被験者とは異なる特性を持つ多様な仮想ソース生成の重要性が示された。この多様性への着目は VSA の基本的かつ特徴的なポイントである。従来の DA 手法は、元データに極

めて近いデータを生成することを目的としているのに対し [187, 195, 196, 197], VSA は多様性に富む仮想ソース被験者を生成する。今後は、より革新的な方法を検討し、従来の DA を超える方法で多様性と性能をさらに最大化することが課題となる。

5.7.3 仮想ソース生成に関する考察

本章で良好な結果を示した線形補間の特筆すべき点として、ソースデータ間の時間的な不整合に対する頑健性が挙げられる。本実験における線形補間では、被験者間の波形の厳密な時間的位置合わせは実施していない。前処理の段階で動作開始点に基づいたセグメンテーションは行われているものの、生体信号には本質的なばらつきが存在する。例えば OpenBMI Dataset においては、キュー提示後 1000–3500 ms の区間を一律に抽出して解析に使用しているが、実際の運動想起が開始されるタイミングや脳活動のピーク潜時には、被験者や試行ごとに不可避なズレが生じる。それにもかかわらず、LI-VSA が従来手法を有意に上回る性能を示したという事実は、本手法がこうした信号の時間的な微細なズレに対して過敏ではなく、高いロバスト性を有していることを示唆している。厳密なアライメント処理を必要とせず、大まかなセグメンテーションのみで機能するというこの特性は、実環境におけるノイズや変動に対する許容性を示しており、VSA の実用性を裏付ける重要な要素であると言える。したがって、線形補間は、良質な仮想ソース被験者を生成するための重要なベースラインであると位置づけられる。

質の高い仮想ソース被験者を生成することに加え、その数を増やすことも同様に重要である。図 5.3 に示した通り、実ソースが 2 名の場合には、GN-VSA が LI-VSA より一貫して高い精度を示した。この結果は、各手法が生成する仮想ソースの数に起因しており、表 5.1 に示されるように、2 名の実ソース被験者では GN-VSA の方が LI-VSA よりも多数の仮想ソースを生成できるためである。このことから、たとえ多少質が落ちたとしても、仮想ソースの数を増やすことが少数の高品質仮想ソースを生成するより有効である可能性が示唆される。今後の VSA の改善では、仮想ソースの質と量のバランスを最適化することが重要になる。

本章では、2 種類の単純な仮想ソース生成手法を用いたが、仮想ソース数を増加させるためにはさらなる方法が考えられる。LI-VSA においては、補間係数 λ を 0.5 に固定し、2 名の実ソースを等しく混合したが、 λ を変化させることによって、同じ実ソースのペアからさらに多様な仮想ソースを生成できる。また、複数被験者のデータを同時に補間したり、線形外挿を利用したりすることで、仮想ソースの多様性と量を大幅に増加させられる。さらに、従来の DA 手法を改良し、より大きな変動を持つ仮想データを生成することも検討すべきである。ただし、過度に強いガウシアンノイズは本来の信号が持つ重要な特徴を消去してしまう可能性があるため、生成敵対ネットワーク (GAN) などの、より柔軟で強力な DA 手法を導入することが考えられる [205]。これらの戦略を探求することは、将来の研究において重要な目標となる。

5.7.4 VSA の応用可能性とインパクト

VSA はドメイン固有の知識に依存せず動作するため、他の生体信号への拡張が容易である。例えば、脳磁場を測定し、リハビリテーション分野の研究に用いられる脳磁図 (MEG) は、VSA の有望な適用対象である [206]。MEG は大型機器を必要とし維持費用も高いため、多人数からデータを収集することが非常に困難である。予備的評価として、公開 MEG データセットに VSA を適用したところ、付録 5.9 に示す通り小規模な実験ながら有望な結果が得られた。このように VSA は複数種類の信号で有効であるため、特定のモダリティを超えて機能する堅牢性を備え、モデルの汎化性能向上にも寄与すると考えられる。したがって、VSA は多様な生体信号を用いたアプリケーション開発の基礎的手法となり得る。

また、VSA は、生体信号ベースの HCI 開発において、膨大なデータ収集に依存するという根本的課題を解決する上で大きな前進となり得る。重要な点として、VSA は長時間の記録セッションを減らすことで参加者の負担を軽減し、データ収集をより人道的かつ持続可能なものとする。また、参加者の募集や運営に伴う多大な負担を軽減することで、HCI 開発を加速できる。

さらに、VSA は患者固有のパターンを分析する医療分野の機械学習応用にも大きな可能性を秘めている。特に、まれな疾患に関わるシナリオでは被験者数の確保が困難であり、少数の実被験者データで効率よくモデルを訓練できる VSA が重要な役割を果たすだろう。最終的に、蓄積した知識を基に仮想ソースを生成することが可能になれば、事前収集したソースデータへの依存を大幅に低減、あるいは排除できる可能性がある。

5.7.5 限界

本章のデータセットは健康な被験者のみを対象としているが、生体信号ベースの HCI は患者にも適用されると期待される [102, 72]。患者群では疾患の進行度や損傷の重症度により、個人間変動が大きくなる可能性がある。例えば、切断の程度により使用可能な電極数が変わることがあり [191]、患者を対象としたさらなる検証が必要である。

5.8 結言

本章では、転移学習における利用可能なソース被験者数を増やす VSA を提案した。2つの仮想ソース生成手法 (LI-VSA, GN-VSA) を用い、MyoDataset および OpenBMI Dataset で評価した。その結果、VSA はデータセットや NN アーキテクチャを問わず従来手法を上回る性能を示し、特に LI-VSA は有意に高い識別精度を達成した。また、仮想ソース追加は実ソース追加と同様に効果的であり、VSA の潜在力が示された。VSA はドメイン知識を必要とせず柔軟に適用可能であるため、生体信号ベースの正確で使いやすい HCI の開発に新たな道を開くことが期待される。

5.9 章末付録: 異なる生体信号モダリティ (MEG) への拡張性評価

MEG データに対する VSA の有効性を評価するため、本研究では公開 MEG データセットである MEG-BCI dataset [207] を用いた。この評価においては、線形補間によって仮想ソースを生成し、バックボーンとしては手作り特徴量を入力とする MLP を採用した。本付録において、この組み合わせは sEMG および EEG のいずれのデータセットにおいても一貫して高い性能を示しており、MEG データに対する VSA の評価手法としても信頼性が高いと判断した。MEG-BCI dataset には、17 名の健常参加者（年齢 22–40 歳、女性 3 名）のデータが含まれている [207]。実験では、各参加者が手のイメージ、足のイメージ、引き算のイメージ、単語生成のイメージという 4 種類の課題を実施した。参加者は 2 日間にわたり 2 セッションを行い、各セッションでは課題ごとに 50 試行が行われた。各試行では、306 チャンネルの Elekta Neuromag システム (Elekta Oy, Finland) を用いて 7 s の MEG 信号が記録された。このシステムは 102 個の磁力計および 204 個の平面型グラジオメータで構成され、サンプリング周波数は 1000 Hz である。

VSA の実験では、Brain-computer Interface 研究で一般的に用いられる手・足の運動イメージ課題を対象とした [64]。さらに、先行研究 [207] において FBCSP により 55% を超える分類精度を達成した 7 名の参加者を選定した。前処理手順は文献 [207] に従った。まず、磁力計はノイズの影響を受けやすいため、204 チャンネルのグラジオメータから得られた MEG データのみを使用した。次に、キュー提示後 500–3500 ms の区間を抽出した。抽出した区間に対して FBCSP 法を適用し、フィルタバンクの構築には 10 種類の周波数帯 (8–12, 10–14, ..., 26–30 Hz) を用いた。各周波数帯では、投影行列の上位 3 フィルタと下位 3 フィルタから対数分散特徴量を抽出した。その結果、MLP の入力 は 2 日間 $\times$ 100 試行、および 60 特徴量となった。ターゲット被験者とした場合には、1 日目のデータをキャリブレーション用、2 日目のデータを評価用に用いた。一方、ソース被験者とした場合には、データ分割を行わず 2 日間のすべてのデータを学習に利用した。

図 5.6 に MEG-BCI データセットに対する実験結果を示す。VSA は従来フレームワークよりも高い分類精度を達成した。ソース被験者数を 6 名とした場合、VSA は被験者内分類精度 68.4% を上回り、被験者内分類器の有無にかかわらず 70.9% の最高精度を示した。興味深いことに、MEG を用いた実験では、実データのみを用いる従来フレームワークは、ソース被験者数が増加しても分類精度の向上がほとんど見られなかった。一方、LI-VSA は、ソース被験者を追加するたびに分類精度が着実に向上するという一貫した傾向を示した。本実験は比較的少数の被験者を対象としているものの、これらの結果は VSA が MEG データに対して有効であることを示唆しており、生体信号の他の種類にも適用可能であることを示すものである。

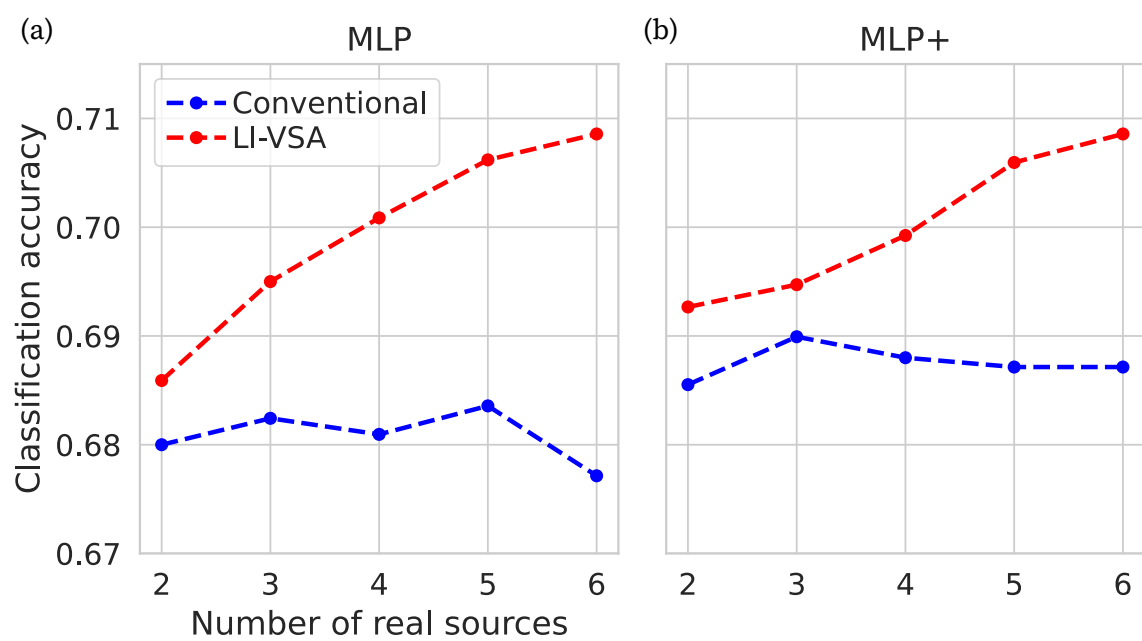


図 5.6: MEG-BCI データセットにおける分類精度. (a) アンサンブル戦略に被験者内 NN を用いない場合, (b) 用いた場合. 文献 [14] より転載.

第6章 総合考察

本研究では、生体信号型 HCI の実用化を阻む最大の障壁である「データ収集の困難性」と「個人差」という 2 つの課題に対し、機械学習と転移学習を統合したフレームワークを構築することで解決策を提示してきた。第 3 章では、システムの核となるバックボーンモデルと転移学習手法の選定を行い、深層識別器とパラメータ空間ベースの転移学習の組み合わせが、個人差によるデータ分布の変動に対して最も高い適応能力を持つことを明らかにした。続く第 4 章では、この基盤手法の上で複数のソース被験者データを扱うための統合戦略を検討した。その結果、全データを一元化するシングルソースアプローチよりも、被験者ごとに個別のモデルを構築して統合するマルチソースアプローチが、データの多様性を維持しつつ高精度な識別を実現できることを実証した。そして第 5 章では、物理的なソース被験者数が限られるという根本的な制約に対し、実データから仮想的なソース被験者を生成する VSA を提案した。これにより、わずかな実データからでも転移学習の効果を最大化し、少数ソース環境下における識別精度を飛躍的に向上させることに成功した。本章では、これらの成果を俯瞰し、本研究が提案するフレームワークが持つ学術的意義および実用的な価値について、データの不均質性、モデルの選択理由、そして実世界応用への展望という観点から総合的に考察する。

6.1 提案フレームワークの意義と位置づけ

本研究の特筆すべき成果の一つは、提案フレームワークが計測条件や生体信号モダリティの差異を超えて、一貫した有効性を示した点にある。本論文で扱った MyoDataset, NinaPro DB5, NinaPro DB2, OpenBMI, MEG-BCI といったデータセットは、それぞれ全く異なる特性を持っている。例えば sEMG において、MyoDataset や NinaPro DB5 はサンプリング周波数が 200 Hz と低く、電極配置も解剖学的な位置を考慮しない等間隔のリング型配置である。これに対し、NinaPro DB2 はサンプリング周波数が 2000 Hz と高く、等間隔なリング型の配置に加えて、解剖学的に特定された筋に対してピンポイントな電極も配置されている。一般に、解剖学的に特定された単一筋肉の計測と、複数の筋肉活動が混在する計測は、信号の物理的な意味合いや情報量が根本的に異なるため適切なアプローチが異なる [208]。しかし、本研究の結果は、提案フレームワークがこのようなデータの素性の違いを吸収し、転移学習による性能向上をもたらすことを示した。さらに、提案フレームワークが複数の計測モダリティ (sEMG, EEG, MEG) においても同様の有効性が確認された点は、より一層の意義を持つ。これらの 3 つの生体信号は、その物理的な発生機序が異なるだけでなく、観測される周波数帯域や空間分解能といった時空間特性も大きく異なる。これは、提案手法が

特定のハードウェア仕様や計測プロトコルに過適合しているのではなく、生体信号に内在する普遍的なパターンを捉える高いロバスト性を有していることを強く示唆するものである。よって、どのようなマルチチャネルデータであっても、そこに被計測者の意図や状態の情報が含まれている限り、提案フレームワークはその潜在的な構造を抽出・活用できるという点は、多様なデバイスが混在する実社会への展開において極めて重要な知見である。

従来の研究においては、個人差は普遍的なモデル構築を妨げる「ノイズ」として扱われ、いかにしてそれを除去し万人に通じる共通項を見出すかに主眼が置かれてきたものが多い。たとえば、第4章で比較対象としたシングルソースアプローチは、全被験者のデータを統合することで個性をならし、平均的な共通項のみを抽出しようとする試みであったと解釈できる。また、既存の転移学習手法においても、この共通項の抽出を志向するアプローチは主流である。たとえば、TCA や DANN といった転移学習は、特徴空間上でソースとターゲットの分布間距離を最小化するようにデータを変形させることで、明示的に共通項の抽出を行っている。同様に、EEG 分野でよく用いられる Euclidean Alignment は、各被験者の共分散行列を白色化することで、統計的な個人差を強制的に潰す働きをする [209]。これら従来手法の根底にあるのは、データを規格化し、情報を削ぎ落とすことで共通解を得ようとするアプローチである。

一方で、本研究が見出したフレームワークは、むしろ個々のソース被験者が持つ多様な個人差を積極的に保存・活用する戦略である。提案法のマルチソース FT は、ソースとターゲットを無理に単一の共通表現へ収束させることではなく、既存のソースモデルを出発点とし、ターゲットとの間に存在する個人の差分のみを事後的に補正するアプローチをとる。これは見方を変えれば、個々のソース被験者を未知の被験者を表現するための「基底」として使用していると解釈できる。数学的なベクトル空間においては、基底ベクトルが十分に多様であれば、任意のベクトルをそれらの線形結合として表現できる。このアナロジーを踏まえれば、生体信号においてソース被験者のバリエーション（基底）を十分に確保・保存しておくことは、未知のターゲットユーザをその近傍、あるいは組合せとして近似する数理的な操作に対応すると考えられる。したがって、VSA による多様な仮想被験者の生成は、新たな基底ベクトルを擬似的に追加し、ターゲットユーザを表現可能な空間を拡張・補完する操作として位置づけられる。これはすなわち、個人差を除去すべきノイズとしてではなく、フレームワークの汎化性能を支える不可欠な構成要素として再定義するものである。

6.2 開発フェーズとデータ規模に応じた適用戦略

本研究では、識別器として MLP や CNN といった、今日においては比較的古典的とも言える枯れた技術を主として採用した。近年では Transformer をはじめとする大規模モデルが主流となりつつあるが、本研究があえてベーシックなモデルを採用した背景には、明確な意図がある。第一に、Transformer のような巨大なモデルは、学習に膨大なデータを必要とし、本研究が対象としたような少数データ環境下では学習が不安定になりやすく、結果の再現性を担保することが難しい。第二に、本研究の主眼は、新しいネットワークアーキテク

チャの提案ではなく、転移学習や VSA といった「データ操作の枠組み」の有効性を検証することにある。そのため、誰もが特性を理解している標準的なモデルを用いることで、得られた性能向上がモデルの特殊性によるものではなく、提案フレームワークそのものの効果であることを客観的かつ説得力を持って示すことができると考えたためである。しかし、本フレームワークはモデルに依存しないため、内部の識別器を最新のアーキテクチャにアップデートすることは、極めて自然かつ推奨される発展形と言える。したがって、より多くの実被験者データや大規模な公開データセットが利用可能なフェーズにおいては、基盤モデルを Transformer 等の最新アーキテクチャに刷新し、提案フレームワークが到達する性能の上限や汎化性能を改めて検証することが、本研究の発展として極めて重要である。

データセットの規模の観点から言えば、近年 Meta 社が 6000 人超の規模で sEMG 研究を推進するなど、ビッグデータ化の波は生体信号分野にも押し寄せている [54]。もし、そのような理想的な大規模データセットが手元にあるならば、画像認識や自然言語処理分野と同様に、巨大なモデルを用いた王道的な機械学習アプローチを採用することが最適解となるであろう。現に、この Meta の研究では、スケーリング則 (モデルサイズとデータ量を増やすと誤差がべき乗則で下がる現象) に一致する挙動を報告しており、王道的なアプローチが生体信号でも成立することを裏付けている。しかし、現実のシステム開発、特に新規プロダクトの開発や、特定の患者層に向けた医療機器の開発現場において、初期段階からそのような大規模データが利用できるケースは極めて稀である。どのようなプロジェクトであっても、まずは少数のデータでプロトタイプを作成し、その有用性を評価して開発を継続するか否かを判断する PoC (概念実証) のフェーズが必ず存在する。本研究が提案するフレームワークは、まさにこの「データは少ないが、可能性を検証したい」という開発の初期段階において最大の価値を発揮する。VSA を用いることで、少数の被験者データからでも信頼性の高い評価モデルを構築でき、開発の意思決定を強力に支援することができる。すなわち、本研究はビッグデータ時代の到来を否定するものではなく、そこに至るまでのコールドスタート問題を埋めるための架け橋となる技術としての価値も大きい。

6.3 実世界応用への課題と展望

提案フレームワークは、生体信号型 HCI の社会実装を加速させる基盤技術として位置づけられ、その初期の開発を支えるものである。本研究では、主に sEMG と EEG を対象に離散値の識別課題で評価を行ってきたが、この離散的な意図抽出技術だけでも、工学的および医学的に多大な貢献が期待できる。本節では、本研究で確立した技術基盤が、実際のアプリケーションとしてどのように社会に展開されていくのか、その展望を論じる。

まず、日常生活環境における工学的応用の文脈において、sEMG は最も早期に社会実装へ移行可能なモダリティである。生体信号型 HCI の実用性は、計測の容易さと識別精度に大きく依存するが、sEMG は EEG 等と比較して装着の負担が少なく、かつ高い信号対雑音比を有する点で優位性を持つ。この領域における先駆的な事例として、本研究でも使用した Thalmic Labs 社の Myo が挙げられる。同製品は 200 ドル程度という安価な価格設

定に加え、プレゼンテーションのスライド操作などを sEMG で行う公式アプリケーションを提供し、一般向けの筋電インタフェースの可能性を広く示した。しかし、技術系メディア CIO.com の記事 [210] においては、その将来性は認められつつも、実用的価値は限定的であると評されている。その主な要因として、意図した動作が正しく認識されない精度の不十分さや、既存の入力インタフェースに対する明確な優位性が確立できていない点が指摘されている。指摘されている精度の課題については、本研究が提案するフレームワークによって改善が見込まれる。加えて、既存の入力インタフェースに対する sEMG の優位性が発揮される具体的な場面として、近年急速に普及が進む AR や VR を含む XR (Cross Reality) 空間が挙げられる。XR 空間においては必ずしも従来型のキーボードやマウスの入力が良いとは限らず、カメラによるハンドトラッキングが盛んに試されている。しかし、カメラによるトラッキングでは死角（オクルージョン）の影響を受け、入力可能な空間が限定される。これに対し、筋電を用いたジェスチャ操作は、このオクルージョン問題を回避することができる。また、将来的には基本ジェスチャ認識は、デバイスベンダーから標準 API として提供されると予想される。しかし、アプリケーション固有の動作や、独自のショートカットなど、標準外のカスタムジェスチャを追加したいという需要は常に存在する。このような場面において、少数のデータ計測を行うだけで、新規ジェスチャを高精度に学習・追加できる本提案フレームワークの特性は、システムの実用性と拡張性を担保する上で決定的な役割を果たすだろう。

医学・福祉分野への展開においても、本研究が対象とした離散的識別技術は重要な貢献をもたらす。まず sEMG については、神経損傷に対する機能回復 [47] や筋電義手 [48] などが、長らく代表的な応用例として知られている。一方、EEG は sEMG に比べて信号強度は微弱であり識別精度が低くなりやすいが、随意的な中枢神経活動を直接計測できるという代替不可能な利点を持つ。この特性から、脳卒中後のニューロリハビリテーションとしての応用が進められており、従来のリハビリテーションと組み合わせることで生活動作の改善がメタアナリシスによって報告されている [211]。これらのシステムでは、ユーザーの意図を正しく識別しフィードバックすることが重要であり、高い識別精度が求められる。しかし、病態や損傷レベルの違いにより健常者以上に個人差が顕著となるため、他者のデータを流用することは容易ではない。本研究の評価は健常者に限定されているものの、提案フレームワークが示した強力な個人差補正能力は、このような患者固有の信号特性への適応においても有効に機能する可能性が高い。加えて、易疲労性を持つ患者を対象とした大規模なデータセット構築は臨床的に困難が伴うが、少数の実データから仮想ソースを生成し、データの多様性を確保できる VSA は、このデータ希薄性の問題を解決する上でも極めて有用であると考えられる。

上述したような実応用をさらに深化させ、より複雑な支援を実現するためには、本研究で確立した基盤の上に立ち、提案法を以下の観点から拡張していくことが重要である。まず、計測された波形データからより本質的かつ詳細な情報を抽出・活用することである。本研究では、第 5 章において波形データを CNN で処理する方法も試みたものの、基本的には時系列データから特徴量を抽出して入力を構成するアプローチに基づいている。このプロセスはノイズ耐性を高め、学習を安定させる効果がある一方で、波形そのものが持つ微細な形状

情報や、チャンネル間の発火タイミングの時間差といった情報が縮退されている側面は否めない。また、現状の評価は離散的な識別に留まっているが、実際の機器操作や身体拡張においては、より動的で連続的な制御が不可欠である。今後は、本研究で構築した少数データ学習の枠組みを維持しつつ、特徴量抽出に頼らず生波形を直接扱うモデルへと拡張し、離散的なクラス分類から、連続的な推定へと拡張することは次なるステップとなる。

最後に、システムの在り方そのものに関する展望を述べる。本研究では一貫して、生体信号を外部デバイス进行操作するための入力として見做し、HCI を Human-Computer **Interface** の側面から探求してきた。しかし、工作機械の操作や身体拡張といった高度な実世界応用においては、対象物との接触に伴う反力や環境からのフィードバックが存在し、それに応じた動的な相互作用が不可欠となる。すなわち、システムは単なる一方通行の入力 (Interface) ではなく、環境とのループを含む Human-Computer **Interaction** へと昇華されなければならない。高度な Interaction を実現するためには、その構成要素である Interface が、どのような状況下でも使用者の意図を正確に伝達できる信頼性を持つことが大前提となる。本研究は、システム全体のボトルネックとなりうる「入力の信頼性」を、少数データ学習というアプローチによって確立したものである。この強固な Interface 基盤の上に、環境からのフィードバック情報を統合し、より有機的な Interaction を構築していくことこそが、本研究が目指す生体信号型 HCI の将来像である。

第7章 結論

本論文は、生体信号型 HCI の実用性を高めることを目的とし、特に、少ないデータで、個人差の大きい状況でも高精度を維持するためのフレームワークを提示した。生体信号型 HCI は医療・福祉・産業まで応用が広がっており、アプリケーションの信頼性と機能性は機械学習の識別精度に大きく依存する。しかし、高精度な識別の実現には、長時間の計測を要するデータ収集負担と、被験者間で計測信号が著しく異なる個人差が立ちはだかり、実用化の大きな障壁となっている。そこで本研究は、転移学習に基づくフレームワークを提案し、少数データ下であっても、個人差に堅牢で高精度な識別を実現することを目指した。そのために、まず、第3章で実用化の観点から最適な機械学習・転移学習手法の選択指針を示し、第4章でソース被験者データの活用に関する運用指針を示した。さらに第5章では、従来の転移学習が多人数のソース被験者を前提とするという制約に対し、少数のソース被験者数でも性能を向上させる新規手法 Virtual Source Assistance (VSA) を提案した。以上により、本論文は生体信号識別に適した機械学習の設計指針と、少数ソース環境を補完する VSA を統合し、現実的制約下でも高精度を持続するための実用フレームワークの実現可能性を検証した。

第3章では、実用化の観点から機械識別器と転移学習手法の組合せを横断的に評価し、どれを第一選択とすべきかを検証した。生体信号分野では様々な転移学習手法が提案されてきたが、実運用の観点からは、限られたラベルを用いて短時間かつ低計算コストでキャリブレーションしつつ、高精度な識別を達成できることが望ましい。この要件を満たす有力な枠組みとして、データ空間型とパラメータ空間型の二系統が挙げられるものの、両者を同一条件下で比較検証した研究は十分ではなく、どの機械識別器と組み合わせるべきかについても明確な指針が乏しかった。この状況では、研究者や開発者は最適な組合せを早期に選ぶことができず、導入コストやユーザ負担が増大することに繋がってしまうため、早期に指針を示す必要があった。そこで本章では、データ空間型の代表として STM、パラメータ空間型の代表として FT を選定し、浅層識別器の代表 SVM と深層識別器の代表 ANN に組み合わせ、公開データセットを用いた統一プロトコルで横断比較を行った。その結果、いずれの転移法も非転移より高精度を示し、とくに個人差の大きいソース被験者を追加した場合には、非転移では精度低下が生じ得る一方、転移学習の導入によりこの劣化を抑制できることも確認した。とりわけ ANN と FT の組み合わせが最も高い識別精度と良好なデータ効率を示し、この背景として、生体信号における転移の鍵がターゲットとソース間の非線形性ギャップの橋渡しにあり、深層特徴抽出の恩恵を受けられる FT が線形写像に依存する STM を上回ることが示唆された。以上より、本章は深層モデルとパラメータ空間型転移学習の組合せを、生体信号 HCI における転移学習適用の第一選択として推奨する指針を示した。

第 4 章では、ソース被験者の効率的な活用方法に関する選択指針を提示した。第 3 章において、深層モデルとパラメータ空間型転移学習の組合せが最適であることを示したが、この枠組みには二つの実装方式が存在する。一つは、複数人のデータを単一のデータセットに統合して単一の深層モデルを学習し、ターゲット被験者に合わせて転移学習を行うシングルソースアプローチである。もう一つは、各ソース被験者ごとに個別の深層モデルを学習し、それぞれをターゲット被験者へ転移学習で適応させた後、アンサンブル統合するマルチソース方式である。これら二つはいずれも既存研究で用いられているが、実運用の観点からどちらがより高い安定性と汎用性をもたらすかについては、体系的な比較検証が行われていなかった。本章では、データセットに含まれる人数やデータ量、課題の複雑さ、計測機器の違いといった条件によって、最適な方式が異なると仮説を立てた。この仮説を検証するため、条件の異なる三つの公開データセットを対象に統一プロトコルを設計し、各方式における ANN の FT 精度を比較検証した。その結果、仮説とは異なり、条件の異なるすべてのデータセットにおいてマルチソースアプローチがシングルソースアプローチを有意に上回った。とりわけ、複雑な動作課題を含むデータセットではマルチソースアプローチの優位性が顕著であり、シングルソースアプローチでは転移学習を適用しても十分な精度が得られないことが明らかになった。このような結果となった要因として、マルチソースアプローチでは各ソース被験者に対して個別に学習したモデルを転移させることで被験者ごとの特徴差を局所的に補正できたのに対し、シングルソースアプローチでは事前学習の段階で被験者間のばらつきを十分に捉えられず、転移学習の効果が限定的であったことが示唆された。さらに、4 種類のアンサンブル戦略の影響も評価した結果、課題の複雑性に依りて適切な手法が異なることが示されたが、SV が実装の容易さと安定性の観点から実用的な選択肢であることが確認された。以上の結果から、実運用においてはマルチソースアプローチを基本構成とし、アンサンブル手法として SV を採用することが推奨される。本章の成果により、個人差の大きい環境下においても高精度を維持できる転移学習の実装指針が明確化され、生体信号 HCI の実用化に向けた堅牢な運用フレームワークの基盤が確立された。

第 5 章では、転移学習は十分な人数の事前計測が欠かせないという制約を乗り越えるために、実被験者データを基に仮想被験者を生成し、擬似的にソース数を拡張する VSA を提案した。これまでの研究で生体信号識別における機械学習の有効な設計指針は整理されつつある一方、転移学習が十分に力を発揮するには多人数の事前計測を要するという前提が残っており、依然として実装上のボトルネックとなっていた。多人数の事前計測という要件は、多数の参加者に負担を強いるだけでなく計測側にも負担であり、特に試作初期や希少疾患領域では実装上の大きな障壁となる。この問題に対処するため、本章では実被験者データから仮想被験者データを生成することで、擬似的に人数を増やし、実人数が少数であっても高精度な識別を可能にすることを目指した。この仮想被験者生成においては、コンドルセの陪審定理に着想を得て、生成元の被験者と類似した性質を持つ仮想被験者を生成するよりも、より多様な被験者が含まれることが重要であると仮説を立てた。この仮説の下、多様な被験者を生成しやすい LI-VSA と、元データに近い被験者を生成しやすい GN-VSA の二つの生成法を試み、仮想被験者を用いない従来法との比較検証を行った。実験には、EMG および EEG のオープンデータセットを対象に統一プロトコルを設計し、仮想被験者の有無による

三つの NN の FT 精度を比較検証した。実験の結果、仮想被験者の導入により、従来法よりも有意に高精度な識別が達成されることが確認された。特に、多様な被験者を生成する LI-VSA が優れた性能を示し、従来法が実被験者 10 名で到達する精度を、平均 4.7 名の実被験者で凌駕した。また、仮想被験者の追加は実被験者の追加と同様に精度向上に寄与する傾向が見られ、仮想被験者の生成と追加が実被験者の計測に代わる有効な手段となり得ることが示された。以上より、本章の知見は、従来は欠かすことができなかった実計測への依存度を軽減し、参加者の身体的・精神的負担とデータ取得コストを抑えつつ高精度を維持できるフレームワークの実現可能性を示した。

以上を総括すると、本研究は、生体信号型 HCI の実用化を阻むデータ収集負担と個人差という二大障壁に対し、二段階の貢献によって実用的な解決の枠組みを提示した。第一に、第 3 章および第 4 章を通じて、これまで体系的な指針が乏しかった機械学習および転移学習手法の選定と運用方式について、深層モデルとパラメータ空間型転移の組合せ、ならびにマルチソースアプローチの採用が最適であることを実証し、高精度と堅牢性を両立する設計・運用指針を確立した。第二に、第 5 章において、従来の転移学習が前提としていた多人数の事前計測という最大のボトルネックを解消するため、少数の実データから仮想的なソースを生成する VSA を新たに提案し、データ収集コストとユーザ負担を大幅に削減しつつ高精度を達成する道筋を示した。この設計指針の体系化と、少数ソース環境を補完する VSA の提案を統合することで、本論文は、限られたデータという現実的制約下でも高精度な識別を可能にする、生体信号 HCI のための実用的な生体信号識別フレームワークを確立した。

本研究が示したフレームワークは、単なる生体信号識別の精度の改善にとどまらず、生体信号型 HCI を誰もが、どこでも、すぐに使える基盤技術へと押し上げるための道筋を具体化するものである。少数データでも個人差に適応できる汎化性能の高い機械学習を実現できることは、長時間の事前計測や専門家依存を減らし、医療やリハビリ、産業支援、教育やスポーツ、エンターテインメントに至るまで、様々な応用分野へシームレスに拡がることを意味する。今後、より優れた機械学習モデルや転移学習手法が登場しても、本研究の枠組みは基盤技術として依然として有益であり、とりわけ仮想被験者という概念は重要さを増していく。仮想被験者の導入は、希少疾患や小規模開発チームのようにデータ獲得が難しい領域にも公平な技術アクセスをもたらし、より包摂的で持続可能なイノベーションの循環を生む。また、多様性と品質を同時に高める仮想被験者生成技術が確立されれば、ヒトへの負担をさらに軽減しながら、高信頼かつ実用的な識別をいっそう現実的に実現できる。ひいては、実計測に依存せずシミュレーションのみで仮想被験者を生成することが実現されれば、生体信号型 HCI の開発速度と普及範囲は飛躍的に拡大する。本論文が切り拓いた方法論が、研究室の外へ、社会の現場へと接続され、誰もが自らの能力を拡張できる未来の礎となることを切に願う。

謝辞

本研究を完遂するにあたり、多くの方々からのご指導とご支援を賜りました。この場をお借りして、心より御礼申し上げます。

まず、主査であり指導教員である慶應義塾大学 環境情報学部 青山敦 教授に心より御礼申し上げます。青山教授には、私が学部生の頃より長きにわたり、専門知識や技能はもとより、研究者としての姿勢を一からご指導賜りました。特に、本研究の核心である「仮想被験者」の概念につきましては、研究の初期段階において先生にその可能性を見出していただき、私の挑戦を強く後押ししていただきました。このご支援こそが、本研究を完遂する上での最大の原動力となりました。学生の自由な発想を最大限に尊重し、学問に対して常に真摯に向き合われる先生の姿勢に、私は研究者・教育者としての在るべき姿を学びました。先生より頂いた薫陶を胸に、今後も研究活動に邁進し、社会に貢献できる研究者・教育者となるべく精進してまいります。

また、本論文の副査をお引き受けいただき、多角的な視点から貴重なご意見を賜りました、慶應義塾大学 環境情報学部 牛山潤一 教授、仰木裕嗣 教授、国立研究開発法人 産業技術総合研究所 人工知能研究センター 叶賀卓 主任研究員に深く感謝申し上げます。異なる専門分野の第一線で活躍される先生方より、それぞれの知見に基づいた俯瞰的かつ建設的なご助言を賜りましたことは、著者の視野を広げ、本論文の学術的基盤を強固なものとする上で、得難い機会となりました。牛山教授には、生体信号の専門的見地から本手法の議論を深めていただき、本手法の本質的な働きや意義について多くの気づきを得るに至る、貴重なご指導を賜りました。仰木教授には、現実世界でのインタラクションを見据えた物理的な観点から、本論文における考察に深みを与えていただくとともに、私の研究の進むべき道筋を鮮やかに照らしていただきました。叶賀先生には、生体信号分野の機械学習の専門的見地から、本研究の根幹を成す手法の開発において不可欠なご指導を仰ぎ、研究の遂行を技術面のみならず精神面からも強力に支えていただきました。先生方より賜りました多角的なご指導は、本論文の質を高めるに留まらず、私自身の人間的な成長や視野の広がりをもたらす、得難い経験となりました。ここに記して、心よりの感謝を捧げます。

私の機械学習研究の礎を築いてくださった、国立研究開発法人 産業技術総合研究所の兼村厚範先生に深く感謝申し上げます。学部生から修士課程にかけて、研究の道を志した当初に兼村先生より賜ったご指導は、私の研究者としての原点となっております。とりわけ、基礎理論から応用に至るまで、時間を惜しまず、紙とペンを用いて一つ一つ数式を書き下しながら、丁寧に原理を説いてくださった時間は、私にとって何物にも代えがたい貴重な経験でした。あの時の先生との対話を通じて得られた数理への深い理解がなければ、本研究を成し遂げることは到底叶いませんでした。

本研究の一部は、慶應義塾大学 森泰吉郎記念研究振興基金、および慶應義塾大学 潮田記念基金の助成を受けて行われました。これらのご支援のおかげで、本研究を円滑に遂行することができました。ここに深く感謝いたします。

日頃より研究生生活を共にし、切磋琢磨した慶應義塾大学 青山敦研究室の諸氏に感謝いたします。多様なバックグラウンドを持つメンバーと過ごした日々は常に知的な刺激に満ちており、日常の何気ない会話や議論の数々が、本研究における重要な着想や飛躍の契機となりました。互いに高め合える素晴らしい仲間に出会えたことに、心より感謝申し上げます。

長きにわたる研究生生活において、私は多くの良き友人に恵まれました。その中でもとりわけ、中村佳祐 氏、小畠大平 氏、そして 故 Bishop Merrill W. B. Adamson に感謝いたします。大学入学当初よりの親友である中村氏の、学問を探究する喜びを全身で体現するその姿は、私に常に多大な刺激を与えてくれました。彼との出会いこそが、私の魂に学問や研究への火を灯した、人生の決定的な転機でした。また、小畠氏には、私の大学生活における頼れる兄貴分として、研究活動の枠を超え、生活や人生のあらゆる局面で親身に寄り添い、支えていただきました。共に携わった仕事の場においても、私の至らぬ点を補い、導いてくださった彼の存在に、心より感謝します。そして、留学中に私を家族のように迎え入れてくれた Adamson 氏に、特別な感謝を捧げます。毎夜、夜が更けるまで学問や人生について語り明かしたあの日々は、私のその後の生き方に計り知れない影響を与え、一生の財産となりました。本論文の完成を誰よりも楽しみに待っていてくれたであろう彼に、ここに謹んで完成の報告をするとともに、万感の思いを込めて感謝を捧げます。

最後に、いかなる時も私の味方であり続け、心の拠り所となってくれた家族に深く感謝いたします。私の選んだ道を尊重し、常に背中を押し続けてくれた、父 星野伸司、母 星野由美子、姉 星野佐由理、そして生前、限りない愛情を注いでくれた祖母 星野ヤス江 に心より感謝します。また、私を温かく迎え入れ、長きにわたる研究生生活を優しく見守り、多大なご支援をくださった、義父 松永隆治、義母 松永浩美、義姉 松永美紀 に深く感謝申し上げます。皆様の温かな支えがあったからこそ、私は安心して研究に打ち込むことができました。そして、誰よりも近くで私を信じ、苦楽を共にしながら支え続けてくれた妻の結花に、心からの愛と感謝を捧げます。

参考文献

- [1] L. McManus, G. De Vito, and M. M. Lowery, “Analysis and biophysics of surface emg for physiotherapists and kinesiologists: Toward a common language with rehabilitation engineers,” *Frontiers in neurology*, vol. 11, p. 576729, 2020.
- [2] 株式会社 ALTs, “表面筋電位 (emg) とは,” <https://alts-device.com/technology/myoelectric-emg/>, 2025, アクセス日: 2025-11-21.
- [3] 中村英夫, “いまさら聞けない生体計測 1 (筋電),” *生体医工学*, vol. 57, no. 4-5, pp. 149–157, 2019.
- [4] E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, and A. J. Hudspeth, *カandel神経科学*, 金澤一郎 and 宮下保司, Eds. メディカル・サイエンス・インターナショナル, 2014.
- [5] S. Sanei and J. A. Chambers, *EEG signal processing*. John Wiley & Sons, 2013.
- [6] 青山敦, 上田一貴, 川崎真弘, 渡部宏樹, 辰岡鉄郎, 杉剛直, 船井正太郎, 笹井俊太郎, 中村哲, 田中宏季, 宮本佳奈, 岩城達也, 宮内英里, 兒玉隆之, 加納慎一郎, 夏目季代久, 平田雅之, 西下慧, and 間瀬英之, *脳波計測・解析の実用ハンドブック ～計測の留意点、解析データの取り扱い、最新応用事例～*. R&D 支援センター, 2024.
- [7] R. R. Daru, M. M. Rabby, T. Ko, Y. Shinglot, R. Raihan, and A. Adnan, “Electrically equivalent head tissue materials for electroencephalogram study on head surrogates,” *Applied Sciences*, vol. 14, no. 6, p. 2495, 2024.
- [8] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, “Domain-adversarial training of neural networks,” *Journal of Machine Learning Research*, vol. 17, no. 1, pp. 2030–2096, 2016.
- [9] S. Kanoga, T. Hoshino, and H. Asoh, “Semi-supervised style transfer mapping-based framework for semg-based pattern recognition with 1-or 2-dof forearm motions,” *Biomedical Signal Processing and Control*, vol. 68, p. 102817, 2021.
- [10] T. Hoshino, S. Kanoga, M. Tsubaki, and A. Aoyama, “Comparing subject-to-subject transfer learning methods in surface electromyogram-based motion recognition with shallow and deep classifiers,” *Neurocomputing*, vol. 489, pp. 599–612, 2022.

- [11] The Ninapro team, “Db2 data,” <https://ninapro.hevs.ch/instructions/DB2.html>, 2014, accessed: 2025-12-01.
- [12] S. Pizzolato, L. Tagliapietra, M. Cognolato, M. Reggiani, H. Müller, and M. Atzori, “Comparison of six electromyography acquisition setups on hand movement classification tasks,” *PloS one*, vol. 12, no. 10, p. e0186132, 2017.
- [13] T. Hoshino, S. Kanoga, M. Tsubaki, and A. Aoyama, “Comparison of fine-tuned single-source and multi-source approaches to surface electromyogram pattern recognition,” *Biomedical Signal Processing and Control*, vol. 94, 2024.
- [14] T. Hoshino, S. Kanoga, T. Takase, and A. Aoyama, “Virtual source assistance: Improving transfer learning performance for biological signal classification with data from few source subjects,” *Available at SSRN 5247869*, 2025.
- [15] A. LeNail, “Nn-svg: Publication-ready neural network architecture schematics.” *Journal of Open Source Software*, vol. 4, no. 33, p. 747, 2019.
- [16] T. Hoshino, S. Kanoga, and A. Aoyama, “Kernel-based style transfer mapping for cross-subject biological signal classification,” *IEEE Access*, vol. 14, pp. 4496–4507, 2025.
- [17] P. E. Ceruzzi, *A history of modern computing*. MIT press, 2003.
- [18] I. E. Sutherland, “Sketch pad a man-machine graphical communication system,” in *Proceedings of the SHARE design automation workshop*. ACM, 1964, pp. 6.329–6.346.
- [19] D. M. Ritchie, S. C. Johnson, M. Lesk, B. Kernighan *et al.*, “Unix time-sharing system: The c programming language,” *Bell System Technical Journal*, vol. 57, no. 6, pp. 1991–2019, 1978.
- [20] D. C. Smith, C. Irby, R. Kimball, B. Verplank, and E. Harslem, “Designing the star user interface,” *Byte*, vol. 7, no. 1982, pp. 242–282, 1982.
- [21] J. Y. Han, “Low-cost multi-touch sensing through frustrated total internal reflection,” in *Proceedings of the 18th annual ACM symposium on User interface software and technology*, 2005, pp. 115–118.
- [22] J. West and M. Mace, “Browsing as the killer app: Explaining the rapid success of apple’s iphone,” *Telecommunications Policy*, vol. 34, no. 5-6, pp. 270–286, 2010.
- [23] A. T. Duchowski, “A breadth-first survey of eye-tracking applications,” *Behavior Research Methods, Instruments, & Computers*, vol. 34, no. 4, pp. 455–470, 2002.

- [24] M. B. Hoy, “Alexa, siri, cortana, and more: an introduction to voice assistants,” *Medical reference services quarterly*, vol. 37, no. 1, pp. 81–88, 2018.
- [25] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. V. Vaughan, “Brain–computer interface technology: A review of the first international meeting,” *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 164–173, 2000.
- [26] M. A. Oskoei and H. Hu, “Myoelectric control systems—a survey,” *Biomedical Signal Processing and Control*, vol. 2, no. 4, pp. 275–294, 2007.
- [27] T. Stracina, M. Ronzhina, R. Redina, and M. Novakova, “Golden standard or obsolete method? review of ecg applications in clinical and experimental context,” *Frontiers in Physiology*, vol. 13, p. 867033, 2022.
- [28] C. Belkhiria, A. Boudir, C. Hurter, and V. Peysakhovich, “Eog-based human–computer interface: 2000–2020 review,” *Sensors*, vol. 22, no. 13, p. 4914, 2022.
- [29] N. Naseer and K.-S. Hong, “fnirs-based brain-computer interfaces: a review,” *Frontiers in human neuroscience*, vol. 9, p. 3, 2015.
- [30] J. Park, H. S. Seok, S.-S. Kim, and H. Shin, “Photoplethysmogram analysis and applications: an integrative review,” *Frontiers in physiology*, vol. 12, p. 808451, 2022.
- [31] Z. Fu, H. Li, Z. Ouyang, X. Liu, and J. Niu, “Typing everywhere with an emg keyboard: A novel myo armband-based hci tool,” in *International Conference on Algorithms and Architectures for Parallel Processing*. Springer, 2020, pp. 247–261.
- [32] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.-P. Jung, and S. Gao, “High-speed spelling with a noninvasive brain–computer interface,” *Proceedings of the national academy of sciences*, vol. 112, no. 44, pp. E6058–E6067, 2015.
- [33] A. Campbell, T. Choudhury, S. Hu, H. Lu, M. K. Mukerjee, M. Rabbi, and R. D. Raizada, “Neurophone: brain-mobile phone interface using a wireless eeg headset,” in *Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds*, 2010, pp. 3–8.
- [34] B. Rebsamen, C. Guan, H. Zhang, C. Wang, C. Teo, M. H. Ang, and E. Burdet, “A brain controlled wheelchair to navigate in familiar environments,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 6, pp. 590–598, 2010.
- [35] A. C. Manero, S. L. McLinden, J. Sparkman, and B. Oskarsson, “Evaluating surface emg control of motorized wheelchairs for amyotrophic lateral sclerosis patients,” *Journal of neuroengineering and rehabilitation*, vol. 19, no. 1, p. 88, 2022.

- [36] C.-T. Lin, B.-S. Lin, F.-C. Lin, and C.-J. Chang, "Brain computer interface-based smart living environmental auto-adjustment control system in upnp home networking," *IEEE Systems Journal*, vol. 8, no. 2, pp. 363–370, 2014.
- [37] S. C. Gandevia, "Spinal and supraspinal factors in human muscle fatigue," *Physiological reviews*, vol. 81, no. 4, pp. 1725–1789, 2001.
- [38] E. Wascher, B. Rasch, J. Sanger, S. Hoffmann, D. Schneider, G. Rinkenauer, H. Heuer, and I. Gutberlet, "Frontal theta activity reflects distinct aspects of mental fatigue," *Biological psychology*, vol. 96, pp. 57–65, 2014.
- [39] J. Wijsman, B. Grundlehner, J. Penders, and H. Hermens, "Trapezius muscle emg as predictor of mental stress," *ACM transactions on embedded computing systems (TECS)*, vol. 12, no. 4, pp. 1–20, 2013.
- [40] X. Hou, Y. Liu, O. Sourina, Y. R. E. Tan, L. Wang, and W. Mueller-Wittig, "Eeg based stress monitoring," in *2015 IEEE international conference on systems, man, and cybernetics*. IEEE, 2015, pp. 3110–3115.
- [41] S. Makeig, T.-P. Jung, and T. J. Sejnowski, "Awareness during drowsiness: dynamics and electrophysiological correlates." *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie experimentale*, vol. 54, no. 4, pp. 266–273, 2000.
- [42] R. J. Davidson, "Affective style and affective disorders: Perspectives from affective neuroscience," *Cognition & emotion*, vol. 12, no. 3, pp. 307–330, 1998.
- [43] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988.
- [44] E. Donchin, K. M. Spencer, and R. Wijesinghe, "The mental prosthesis: assessing the speed of a p300-based brain-computer interface," *IEEE transactions on rehabilitation engineering*, vol. 8, no. 2, pp. 174–179, 2000.
- [45] M. Middendorf, G. McMillan, G. Calhoun, and K. S. Jones, "Brain-computer interfaces based on the steady-state visual-evoked response," *IEEE transactions on rehabilitation engineering*, vol. 8, no. 2, pp. 211–214, 2000.
- [46] M. Cheng, X. Gao, S. Gao, and D. Xu, "Design and implementation of a brain-computer interface with high transfer rates," *IEEE transactions on biomedical engineering*, vol. 49, no. 10, pp. 1181–1186, 2002.
- [47] P. H. Peckham and J. S. Knutson, "Functional electrical stimulation for neuromuscular applications," *Annual Review of Biomedical Engineering*, vol. 7, no. 1, pp. 327–360, 2005.

- [48] B. Hudgins, P. Parker, and R. N. Scott, “A new strategy for multifunction myoelectric control,” *IEEE transactions on biomedical engineering*, vol. 40, no. 1, pp. 82–94, 1993.
- [49] C.-T. Lin, C.-H. Chuang, C.-S. Huang, S.-F. Tsai, S.-W. Lu, Y.-H. Chen, and L.-W. Ko, “Wireless and wearable eeg system for evaluating driver vigilance,” *IEEE Transactions on biomedical circuits and systems*, vol. 8, no. 2, pp. 165–176, 2014.
- [50] R. Riedl, K. Kostoglou, S. C. Wriessnegger, and G. R. Müller-Putz, “Videoconference fatigue from a neurophysiological perspective: Experimental evidence based on electroencephalography (eeg) and electrocardiography (ecg),” *Scientific reports*, vol. 13, no. 1, p. 18371, 2023.
- [51] J. G. Abreu, J. M. Teixeira, L. S. Figueiredo, and V. Teichrieb, “Evaluating sign language recognition using the myo armband,” in *2016 XVIII symposium on virtual and augmented reality (SVR)*. IEEE, 2016, pp. 64–70.
- [52] S. Rawat, S. Vats, and P. Kumar, “Evaluating and exploring the myo armband,” in *2016 International Conference System Modeling & Advancement in Research Trends (SMART)*. IEEE, 2016, pp. 115–120.
- [53] “Biosignal sensing device using dynamic selection of electrodes,” Patent US20 230 225 659A1, 7 20, 2023. [Online]. Available: <https://patents.google.com/patent/US20230225659A1/en>
- [54] P. Kaifosh and T. R. Reardon, “A generic non-invasive neuromotor interface for human-computer interaction,” *Nature*, pp. 1–10, 2025.
- [55] C. M. Bishop, *Pattern recognition and machine learning*. Springer, 2006, vol. 4, no. 4.
- [56] A. Fougner, Ø. Stavadahl, P. J. Kyberd, Y. G. Losier, and P. A. Parker, “Control of upper limb prostheses: Terminology and proportional myoelectric control—a review,” *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 20, no. 5, pp. 663–677, 2012.
- [57] M. Banko and E. Brill, “Scaling to very very large corpora for natural language disambiguation,” in *Proceedings of the 39th annual meeting of the Association for Computational Linguistics*, 2001, pp. 26–33.
- [58] A. Halevy, P. Norvig, and F. Pereira, “The unreasonable effectiveness of data,” *IEEE intelligent systems*, vol. 24, no. 2, pp. 8–12, 2009.
- [59] J. Hestness, S. Narang, N. Ardalani, G. Diamos, H. Jun, H. Kianinejad, M. M. A. Patwary, Y. Yang, and Y. Zhou, “Deep learning scaling is predictable, empirically,” *arXiv preprint arXiv:1712.00409*, 2017.

- [60] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein *et al.*, “Imagenet large scale visual recognition challenge,” *International journal of computer vision*, vol. 115, no. 3, pp. 211–252, 2015.
- [61] R. Ardila, M. Branson, K. Davis, M. Henretty, M. Kohler, J. Meyer, R. Morais, L. Saunders, F. M. Tyers, and G. Weber, “Common voice: A massively-multilingual speech corpus,” *arXiv preprint arXiv:1912.06670*, 2019.
- [62] M. Imran, U. Qazi, and F. Offi, “Tbcov: two billion multilingual covid-19 tweets with sentiment, entity, geo, and gender labels,” *Data*, vol. 7, no. 1, p. 8, 2022.
- [63] M. Atzori, A. Gijsberts, C. Castellini, B. Caputo, A.-G. M. Hager, S. Elsig, G. Giatlidis, F. Bassetto, and H. Müller, “Electromyography data for non-invasive naturally-controlled robotic hand prostheses,” *Scientific data*, vol. 1, no. 1, pp. 1–13, 2014.
- [64] C. Brunner, R. Leeb, G. Müller-Putz, A. Schlögl, and G. Pfurtscheller, “BCI competition 2008–graz data set A,” *Institute for knowledge discovery (laboratory of brain-computer interfaces), Graz University of Technology*, vol. 16, no. 1-6, p. 1, 2008.
- [65] M. C. Tosin, J. C. Machado, and A. Balbinot, “Semi-supervised upper limb movement classifier: Current scenario and upcoming challenges,” *Journal of Artificial Intelligence Research*, vol. 75, pp. 83–127, 2022.
- [66] H. Altaheri, G. Muhammad, M. Alsulaiman, S. U. Amin, G. A. Altuwaijri, W. Abdul, M. A. B., and M. Faisal, “Deep learning techniques for classification of electroencephalogram (eeg) motor imagery (mi) signals: A review,” *Neural Computing and Applications*, vol. 35, no. 20, pp. 14 681–14 722, 2023.
- [67] C. Castellini, A. E. Fiorilla, and G. Sandini, “Multi-subject/Daily-life activity EMG-based control of mechanical hands,” *Journal of NeuroEngineering and Rehabilitation*, vol. 6, no. 1, pp. 1–11, 2009.
- [68] M. Kołodziej, A. Majkowski, and M. Jurczak, “Acquisition and analysis of facial electromyographic signals for emotion recognition,” *Sensors*, vol. 24, no. 15, p. 4785, 2024.
- [69] B. Mirkovic, S. Debener, M. Jaeger, and M. De Vos, “Decoding the attended speech stream with multi-channel eeg: implications for online, daily-life applications,” *Journal of neural engineering*, vol. 12, no. 4, p. 046007, 2015.
- [70] J.-Y. Zhu, W.-L. Zheng, and B.-L. Lu, “Cross-subject and cross-gender emotion classification from EEG,” in *Proceedings of the International Union for Physical and Engineering Sciences in Medicine*, 2015, pp. 1188–1191.

- [71] S. Micera, J. Carpaneto, and S. Raspopovic, “Control of hand prostheses using peripheral information,” *IEEE Reviews in Biomedical Engineering*, vol. 3, pp. 48–68, 2010.
- [72] G. Onose, C. Grozea, A. Anghelescu, C. Daia, C. J. Sinescu, A. V. Ciurea, T. Spircu, A. Mirea, I. Andone, A. Spânu, C. Popescu, A.-S. Mihăescu, S. Fazli, M. Danóczy, and F. Popescu, “On the feasibility of using motor imagery EEG-based brain–computer interface in chronic tetraplegics for assistive robotic arm control: A clinical test and long-term post-trial follow-up,” *Spinal Cord*, vol. 50, no. 8, pp. 599–608, 2012.
- [73] K. Choi and A. Cichocki, “Control of a wheelchair by motor imagery in real time,” in *International conference on intelligent data engineering and automated learning*. Springer, 2008, pp. 330–337.
- [74] X. Xu, Y. Zhang, Y. Luo, and D. Chen, “Robust bio-signal based control of an intelligent wheelchair,” *Robotics*, vol. 2, no. 4, pp. 187–197, 2013.
- [75] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.
- [76] K. Weiss, T. M. Khoshgoftaar, and D. Wang, “A survey of transfer learning,” *Journal of Big Data*, vol. 3, no. 1, pp. 1–40, 2016.
- [77] T. Bao, S. Q. Xie, P. Yang, P. Zhou, and Z.-Q. Zhang, “Toward robust, adaptive and reliable upper-limb motion estimation using machine learning and deep learning—A survey in myoelectric control,” *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 8, pp. 3822–3835, 2022.
- [78] Z. Wan, R. Yang, M. Huang, N. Zeng, and X. Liu, “A review on transfer learning in EEG signal analysis,” *Neurocomputing*, vol. 421, pp. 1–14, 2021.
- [79] C.-S. Wei, Y.-P. Lin, Y.-T. Wang, C.-T. Lin, and T.-P. Jung, “A subject-transfer framework for obviating inter-and intra-subject variability in EEG-based drowsiness detection,” *NeuroImage*, vol. 174, pp. 407–419, 2018.
- [80] Z. Lan, O. Sourina, L. Wang, R. Scherer, and G. R. Müller-Putz, “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.
- [81] D. Liu, J. Zhang, H. Wu, S. Liu, and J. Long, “Multi-source transfer learning for EEG classification based on domain adversarial neural network,” *IEEE Transactions on Rehabilitation Engineering*, 2022.

- [82] T. A. Kuiken, M. M. Lowery, and N. S. Stoykov, "The effect of subcutaneous fat on myoelectric signal amplitude and cross-talk," *Prosthetics and orthotics international*, vol. 27, no. 1, pp. 48–54, 2003.
- [83] E. Kupa, S. Roy, S. Kandarian, and C. De Luca, "Effects of muscle fiber type and size on emg median frequency and conduction velocity," *Journal of applied physiology*, vol. 79, no. 1, pp. 23–32, 1995.
- [84] W. Troni, R. Cantello, and I. Rainero, "Conduction velocity along human muscle fibers in situ," *Neurology*, vol. 33, no. 11, pp. 1453–1459, 1983.
- [85] K. Mase, H. Kamimura, S. Imura, and K. Kitagawa, "Effect of age and gender on muscle function-analysis by muscle fiber conduction velocity," *Journal of Physical Therapy Science*, vol. 18, no. 1, pp. 81–87, 2006.
- [86] D. Farina, L. Arendt-Nielsen, and T. Graven-Nielsen, "Effect of temperature on spike-triggered average torque and electrophysiological properties of low-threshold motor units," *Journal of Applied Physiology*, vol. 99, no. 1, pp. 197–203, 2005.
- [87] C. J. De Luca, "Myoelectrical manifestations of localized muscular fatigue in humans." *Critical reviews in biomedical engineering*, vol. 11, no. 4, pp. 251–279, 1984.
- [88] D. C. Van Essen, "A tension-based theory of morphogenesis and compact wiring in the central nervous system," *Nature*, vol. 385, no. 6614, pp. 313–318, 1997.
- [89] J. Decety and J. Grèzes, "Neural mechanisms subserving the perception of human actions," *Trends in cognitive sciences*, vol. 3, no. 5, pp. 172–178, 1999.
- [90] G. Pfurtscheller, C. Neuper, D. Flotzinger, and M. Pregenzer, "Eeg-based discrimination between imagination of right and left hand movement," *Electroencephalography and clinical Neurophysiology*, vol. 103, no. 6, pp. 642–651, 1997.
- [91] G. Pfurtscheller, C. Brunner, A. Schlögl, and F. L. Da Silva, "Mu rhythm (de) synchronization and eeg single-trial classification of different motor imagery tasks," *NeuroImage*, vol. 31, no. 1, pp. 153–159, 2006.
- [92] C. Neuper, R. Scherer, M. Reiner, and G. Pfurtscheller, "Imagery of motor actions: Differential effects of kinesthetic and visual-motor mode of imagery in single-trial eeg," *Cognitive brain research*, vol. 25, no. 3, pp. 668–677, 2005.
- [93] S. Wolf, E. Brölz, D. Scholz, A. Ramos-Murguialday, P. M. Keune, M. Hautzinger, N. Birbaumer, and U. Strehl, "Winning the game: brain processes in expert, young elite and amateur table tennis players," *Frontiers in behavioral neuroscience*, vol. 8, p. 370, 2014.

- [94] H. Sugino and J. Ushiyama, “Gymnasts’ ability to modulate sensorimotor rhythms during kinesthetic motor imagery of sports non-specific movements superior to non-gymnasts,” *Frontiers in Sports and Active Living*, vol. 3, p. 757308, 2021.
- [95] P. Li, D. Yu, L. Cheng, and K. Wang, “Influence of attentional state on eeg-based motor imagery of lower limb,” *Frontiers in Human Neuroscience*, vol. 19, p. 1545492, 2025.
- [96] S. Ben-David, J. Blitzer, K. Crammer, A. Kulesza, F. Pereira, and J. W. Vaughan, “A theory of learning from different domains,” *Machine learning*, vol. 79, no. 1, pp. 151–175, 2010.
- [97] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [98] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, “Contour detection and hierarchical image segmentation,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 5, pp. 898–916, 2011.
- [99] P. Parker, K. Englehart, and B. Hudgins, “Myoelectric signal processing for control of powered limb prostheses,” *Journal of electromyography and kinesiology*, vol. 16, no. 6, pp. 541–548, 2006.
- [100] G. Pfurtscheller and C. Neuper, “Motor imagery and direct brain-computer communication,” *Proceedings of the IEEE*, vol. 89, no. 7, pp. 1123–1134, 2001.
- [101] J. R. Wolpaw, D. J. McFarland, G. W. Neat, and C. A. Forneris, “An eeg-based brain-computer interface for cursor control,” *Electroencephalography and clinical neurophysiology*, vol. 78, no. 3, pp. 252–259, 1991.
- [102] A. H. Al-Timemy, G. Bugmann, J. Escudero, and N. Outram, “Classification of finger movements for the dexterous hand prosthesis control with surface electromyography,” *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 3, pp. 608–618, 2013.
- [103] W.-Y. Hsu, “Eeg-based motor imagery classification using enhanced active segment selection and adaptive classifier,” *Computers in biology and medicine*, vol. 41, no. 8, pp. 633–639, 2011.
- [104] W. Li, Y.-c. Liu, C.-f. Zheng, J. Miao, H. Chen, H.-y. Quan, S.-h. Yan, and K. Zhang, “Diagnosis of compressed nerve root in lumbar disc herniation patients by surface electromyography,” *Orthopaedic surgery*, vol. 10, no. 1, pp. 47–55, 2018.

- [105] R. Tomioka, K. Aihara, and K.-R. Müller, “Logistic regression for single trial eeg classification,” *Advances in neural information processing systems*, vol. 19, 2006.
- [106] F. AlOmari and G. Liu, “Analysis of extracted forearm semg signal using lda, qda, k-nn classification algorithms,” *Open Automation and Control Systems Journal*, vol. 6, no. 1, pp. 108–116, 2014.
- [107] A. Yazdani, T. Ebrahimi, and U. Hoffmann, “Classification of eeg signals using dempster shafer theory and a k-nearest neighbor classifier,” in *2009 4th International IEEE/EMBS Conference on Neural Engineering*. IEEE, 2009, pp. 327–330.
- [108] Z. O. Khokhar, Z. G. Xiao, and C. Menon, “Surface EMG pattern recognition for real-time control of a wrist exoskeleton,” *BioMedical Engineering OnLine*, vol. 9, no. 1, 2010.
- [109] W. Yi, S. Qiu, H. Qi, L. Zhang, B. Wan, and D. Ming, “Eeg feature comparison and classification of simple and compound limb motor imagery,” *Journal of neuroengineering and rehabilitation*, vol. 10, no. 1, p. 106, 2013.
- [110] D. C. Toledo-Pérez, J. Rodríguez-Reséndiz, R. A. Gómez-Loenzo, and J. Jauregui-Correa, “Support vector machine-based emg signal classification techniques: A review,” *Applied Sciences*, vol. 9, no. 20, p. 4402, 2019.
- [111] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, “A review of classification algorithms for eeg-based brain–computer interfaces,” *Journal of neural engineering*, vol. 4, no. 2, p. R1, 2007.
- [112] E. Yaman and A. Subasi, “Comparison of bagging and boosting ensemble machine learning methods for automated emg signal classification,” *BioMed research international*, vol. 2019, no. 1, p. 9152506, 2019.
- [113] M. Bentlemsan, E.-T. Zemouri, D. Bouchaffra, B. Yahya-Zoubir, and K. Ferroudji, “Random forest and filter bank common spatial patterns for eeg-based motor imagery classification,” in *2014 5th International conference on intelligent systems, modelling and simulation*. IEEE, 2014, pp. 235–238.
- [114] S. Liang, L. Wang, L. Zhang, and Y. Wu, “Research on recognition of nine kinds of fine gestures based on adaptive adaboost algorithm and multi-feature combination,” *Ieee Access*, vol. 7, pp. 3235–3246, 2018.
- [115] X. Li, L. Wang, and E. Sung, “Adaboost with svm-based component classifiers,” *Engineering Applications of Artificial Intelligence*, vol. 21, no. 5, pp. 785–795, 2008.
- [116] F. Rosenblatt, “The perceptron: a probabilistic model for information storage and organization in the brain.” *Psychological review*, vol. 65, no. 6, pp. 386–408, 1958.

- [117] B. Widrow and M. E. Hoff, “Adaptive switching circuits,” in *Neurocomputing: foundations of research*, 1988, pp. 123–134.
- [118] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.
- [119] J.-U. Chu, I. Moon, Y.-J. Lee, S.-K. Kim, and M.-S. Mun, “A supervised feature-projection-based real-time EMG pattern recognition for multifunction myoelectric hand control,” *IEEE/ASME Transactions on Mechatronics*, vol. 12, no. 3, pp. 282–290, 2007.
- [120] M. R. Ahsan, M. I. Ibrahimy, and O. O. Khalifa, “Electromyography (emg) signal based hand gesture recognition using artificial neural network (ann),” in *2011 4th international conference on mechatronics (ICOM)*. IEEE, 2011, pp. 1–6.
- [121] R. Chatterjee and T. Bandyopadhyay, “Eeg based motor imagery classification using svm and mlp,” in *2016 2nd International Conference on Computational Intelligence and Networks (CINE)*. IEEE, 2016, pp. 84–89.
- [122] Y. Mohan, S. S. Chee, D. K. P. Xin, and L. P. Foong, “Artificial neural network for classification of depressive and normal in eeg,” in *2016 IEEE EMBS conference on biomedical engineering and sciences (IECBES)*. IEEE, 2016, pp. 286–290.
- [123] M. Yasen and S. Jusoh, “A systematic review on hand gesture recognition techniques, challenges and applications,” *PeerJ Computer Science*, vol. 5, p. e218, 2019.
- [124] I. O. Tolstikhin, N. Houlsby, A. Kolesnikov, L. Beyer, X. Zhai, T. Unterthiner, J. Yung, A. Steiner, D. Keysers, J. Uszkoreit *et al.*, “Mlp-mixer: An all-mlp architecture for vision,” *Advances in neural information processing systems*, vol. 34, pp. 24 261–24 272, 2021.
- [125] J. L. Elman, “Finding structure in time,” *Cognitive science*, vol. 14, no. 2, pp. 179–211, 1990.
- [126] M. Simão, P. Neto, and O. Gibaru, “Emg-based online classification of gestures with recurrent neural networks,” *Pattern Recognition Letters*, vol. 128, pp. 45–51, 2019.
- [127] T.-j. Luo, C.-l. Zhou, and F. Chao, “Exploring spatial-frequency-sequential relationships for motor imagery classification with recurrent neural network,” *BMC bioinformatics*, vol. 19, no. 1, p. 344, 2018.
- [128] W. Wei, Y. Wong, Y. Du, Y. Hu, M. Kankanhalli, and W. Geng, “A multi-stream convolutional neural network for semg-based gesture recognition in muscle-computer interface,” *Pattern Recognition Letters*, vol. 119, pp. 131–138, 2019.

- [129] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggenberger, M. Tangemann, F. Hutter, W. Burgard, and T. Ball, “Deep learning with convolutional neural networks for eeg decoding and visualization,” *Human brain mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.
- [130] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” *Advances in neural information processing systems*, vol. 30, 2017.
- [131] Y. Liu, X. Li, L. Yang, and H. Yu, “A transformer-based gesture prediction model via semg sensor for human–robot interaction,” *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–15, 2024.
- [132] J. Zhang, K. Li, B. Yang, and X. Han, “Local and global convolutional transformer-based motor imagery eeg classification,” *Frontiers in Neuroscience*, vol. 17, p. 1219988, 2023.
- [133] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, “Domain adaptation via transfer component analysis,” *IEEE Transactions on Neural Networks*, vol. 22, no. 2, pp. 199–210, 2010.
- [134] B. Sun, J. Feng, and K. Saenko, “Return of frustratingly easy domain adaptation,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 30, no. 1, 2016.
- [135] B. Sun and K. Saenko, “Deep coral: Correlation alignment for deep domain adaptation,” in *Computer Vision – ECCV 2016 Workshops*. Springer, 2016, pp. 443–450.
- [136] A. A. Rusu, N. C. Rabinowitz, G. Desjardins, H. Soyer, J. Kirkpatrick, K. Kavukcuoglu, R. Pascanu, and R. Hadsell, “Progressive neural networks,” *arXiv preprint arXiv:1606.04671*, 2016.
- [137] X.-Y. Zhang and C.-L. Liu, “Writer adaptation with style transfer mapping,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 7, pp. 1773–1787, 2012.
- [138] M. M.-C. Vidovic, H.-J. Hwang, S. Amsüss, J. M. Hahne, D. Farina, and K.-R. Müller, “Improving the robustness of myoelectric pattern recognition for upper limb prostheses by covariate shift adaptation,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 9, pp. 961–970, 2015.
- [139] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, “How transferable are features in deep neural networks?” in *Advances in Neural Information Processing Systems*, 2014, pp. 3320–3328.

- [140] J. Li, S. Qiu, Y.-Y. Shen, C.-L. Liu, and H. He, “Multisource transfer learning for cross-subject EEG emotion recognition,” *IEEE Transactions on Cybernetics*, vol. 50, no. 7, pp. 3281–3293, 2019.
- [141] S. Kanoga, T. Hoshino, and H. Asoh, “Subject transfer framework based on source selection and semi-supervised style transfer mapping for sEMG pattern recognition,” in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 1349–1353.
- [142] K.-T. Kim, C. Guan, and S.-W. Lee, “A subject-transfer framework based on single-trial EMG analysis using convolutional neural networks,” *IEEE Transactions on Rehabilitation Engineering*, vol. 28, no. 1, pp. 94–103, 2019.
- [143] K. Zhang, N. Robinson, S.-W. Lee, and C. Guan, “Adaptive transfer learning for eeg motor imagery classification with deep convolutional neural network,” *Neural Networks*, vol. 136, pp. 1–10, 2021.
- [144] F. Demir, V. Bajaj, M. C. Ince, S. Taran, and A. Şengür, “Surface EMG signals and deep transfer learning-based physical action classification,” *Neural Computing and Applications*, vol. 31, no. 12, pp. 8455–8462, 2019.
- [145] S. Qi, X. Wu, W. Chen, J. Liu, J. Zhang, and J. Wang, “sEMG-based recognition of composite motion with convolutional neural network,” *Sensors and Actuators A: Physical*, p. 112046, 2020.
- [146] H.-B. Xie, W.-X. He, and H. Liu, “Measuring time series regularity using nonlinear similarity-based sample entropy,” *Physics Letters A*, vol. 372, no. 48, pp. 7140–7146, 2008.
- [147] S. Kanoga and A. Kanemura, “Assessing the effect of transfer learning on myoelectric control systems with three electrode positions,” in *2018 IEEE International Conference on Industrial Technology (ICIT)*. IEEE, 2018, pp. 1478–1483.
- [148] K. Englehart and B. Hudgins, “A robust, real-time control scheme for multifunction myoelectric control,” *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 7, pp. 848–854, 2003.
- [149] Y. Huang, K. B. Englehart, B. Hudgins, and A. D. C. Chan, “A Gaussian mixture model based classification scheme for myoelectric control of powered upper limb prostheses,” *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 11, pp. 1801–1811, 2005.
- [150] D. Arthur and S. Vassilvitskii, “k-means++: The advantages of careful seeding,” Stanford, Tech. Rep., 2006.

- [151] J. C. Bezdek and L. I. Kuncheva, “Nearest prototype classifier designs: An experimental study,” *International Journal of Intelligent Systems*, vol. 16, no. 12, pp. 1445–1473, 2001.
- [152] H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, “Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations,” in *Proceedings of the 26th annual international conference on machine learning (ICML)*, 2009, pp. 609–616.
- [153] S. Fazli, F. Popescu, M. Danóczy, B. Blankertz, K.-R. Müller, and C. Grozea, “Subject-independent mental state classification in single trials,” *Neural Networks*, vol. 22, no. 9, pp. 1305–1312, 2009.
- [154] G. A. Seber and A. J. Lee, *Linear regression analysis*. John Wiley & Sons, 2012, vol. 329.
- [155] C. Igual, L. A. Pardo Jr, J. M. Hahne, and J. Igual, “Myoelectric control for upper limb prostheses,” *Electronics*, vol. 8, no. 11, p. 1244, 2019.
- [156] P. Wierzgała, D. Zapała, G. M. Wojcik, and J. Masiak, “Most popular signal processing methods in motor-imagery bci: a review and meta-analysis,” *Frontiers in neuroinformatics*, vol. 12, p. 78, 2018.
- [157] M.-F. Lucas, A. Gaufriau, S. Pascual, C. Doncarli, and D. Farina, “Multi-channel surface EMG classification using support vector machines and signal-based wavelet optimization,” *Biomedical Signal Processing and Control*, vol. 3, no. 2, pp. 169–174, 2008.
- [158] A. Alkan and M. Günay, “Identification of EMG signals using discriminant analysis and SVM classifier,” *Expert Systems with Applications*, vol. 39, no. 1, pp. 44–47, 2012.
- [159] V. Amatanon, S. Chanhang, P. Naiyanetr, and S. Thongpang, “Sign language-Thai alphabet conversion based on electromyogram (EMG),” in *The 7th 2014 Biomedical Engineering International Conference*. IEEE, 2014, pp. 1–4.
- [160] Z. Zhang, K. Yang, J. Qian, and L. Zhang, “Real-time surface EMG pattern recognition for hand gestures based on an artificial neural network,” *Sensors*, vol. 19, no. 14, 2019.
- [161] R. N. Khushaba, “Correlation analysis of electromyogram signals for multiuser myoelectric interfaces,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 745–755, 2014.

- [162] T. Matsubara and J. Morimoto, “Bilinear modeling of EMG signals to extract user-independent features for multiuser myoelectric interface,” *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 8, pp. 2205–2213, 2013.
- [163] T. Hoshino, S. Kanoga, M. Tsubaki, and A. Aoyama, “Analysis and usage: Subject-to-subject linear domain adaptation in sEMG classification,” in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2020, pp. 674–677.
- [164] H. Han and S. Jo, “Supervised hierarchical Bayesian model-based electromyographic control and analysis,” *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 4, pp. 1214–1224, 2013.
- [165] N. Duan, L.-Z. Liu, X.-J. Yu, Q. Li, and S.-C. Yeh, “Classification of multichannel surface-electromyography signals based on convolutional neural networks,” *Journal of Industrial Information Integration*, vol. 15, pp. 201–206, 2019.
- [166] D.-C. Oh and Y.-U. Jo, “Classification of hand gestures based on multi-channel emg by scale average wavelet transform and convolutional neural network,” *International Journal of Control, Automation and Systems*, vol. 19, no. 3, pp. 1443–1450, 2021.
- [167] T. De Cooman, K. Vandecasteele, C. Varon, B. Hunyadi, E. Cleeren, W. Van Paesschen, and S. Van Huffel, “Personalizing heart rate-based seizure detection using supervised svm transfer learning,” *Frontiers in neurology*, vol. 11, p. 145, 2020.
- [168] J. Feng, L. Peng, and F. Lebourgeois, “Gaussian process style transfer mapping for historical chinese character recognition,” in *Document Recognition and Retrieval XXII*, vol. 9402. International Society for Optics and Photonics, 2015, p. 94020D.
- [169] M. C. Tresch, V. C. K. Cheung, and A. d’Avella, “Matrix factorization algorithms for the identification of muscle synergies: evaluation on simulated and experimental data sets,” *Journal of Neurophysiology*, vol. 95, no. 4, pp. 2199–2212, 2006.
- [170] U. C. Allard, F. Nougrou, C. L. Fall, P. Giguère, C. Gosselin, F. Laviolette, and B. Gosselin, “A convolutional neural network for robotic arm guidance using sEMG based frequency-features,” in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 2464–2470.
- [171] Z. Zhang, C. He, and K. Yang, “A novel surface electromyographic signal-based hand gesture prediction using a recurrent neural network,” *Sensors*, vol. 20, no. 14, p. 3994, 2020.
- [172] S. Muceli and D. Farina, “Simultaneous and proportional estimation of hand kinematics from EMG during mirrored movements at multiple degrees-of-freedom,” *IEEE Transactions on Rehabilitation Engineering*, vol. 20, no. 3, pp. 371–378, 2011.

- [173] A. Phinyomark, G. Chujit, P. Phukpattaranont, C. Limsakul, and H. Hu, “A preliminary study assessing time-domain emg features of classifying exercises in preventing falls in the elderly,” in *2012 9th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*. IEEE, 2012, pp. 1–4.
- [174] I. Sosin, D. Kudenko, and A. Shpilman, “Continuous gesture recognition from semg sensor data with recurrent neural networks and adversarial domain adaptation,” in *2018 15Th international conference on control, automation, robotics and vision (ICARCV)*. IEEE, 2018, pp. 1436–1441.
- [175] X. Chen, Y. Li, R. Hu, X. Zhang, and X. Chen, “Hand gesture recognition based on surface electromyography using convolutional neural network with transfer learning method,” *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 4, pp. 1292–1304, 2020.
- [176] T. S. Tavares, A. F. Osorio, I. H. Fantini, and R. R. Alves, “Leveraging sEMG gesture recognition training on edge devices,” in *17th International Symposium on Medical Information Processing and Analysis*, vol. 12088. SPIE, 2021, pp. 342–350.
- [177] S. J. Lehmler, M. Saif-ur-Rehman, G. Tobias, and I. Iossifidis, “Deep transfer learning compared to subject-specific models for sEMG decoders,” *Journal of Neural Engineering*, vol. 19, no. 5, 2022.
- [178] A. Zhang, Q. Li, Z. Li, and J. Li, “Multimodal fusion convolutional neural network based on sEMG and accelerometer signals for inter-subject upper limb movement classification,” *IEEE Sensors Journal*, vol. 23, no. 11, pp. 12 334–12 345, 2023.
- [179] G. Li, B. Wan, K. Su, J. Huo, C. Jiang, and F. Wang, “sEMG and IMU data-based hand gesture recognition method using multi-stream CNN with a fine-tuning transfer framework,” *IEEE Sensors Journal*, vol. 23, no. 24, pp. 31 414–31 424, 2023.
- [180] Y. Long, Y. Geng, C. Dai, and G. Li, “A transfer learning based cross-subject generic model for continuous estimation of finger joint angles from a new user,” *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 4, pp. 1914–1925, 2023.
- [181] K. Wang, Y. Chen, Y. Zhang, X. Yang, and C. Hu, “Multi-source integration based transfer learning method for cross-user sEMG gesture recognition,” in *2022 International Joint Conference on Neural Networks*. IEEE, 2022, pp. 1–8.
- [182] B. Hudgins, P. Parker, and R. N. Scott, “A new strategy for multifunction myoelectric control,” *IEEE Transactions on Biomedical Engineering*, vol. 40, no. 1, pp. 82–94, 1993.

- [183] L. Van der Maaten and G. Hinton, “Visualizing data using t-SNE.” *Journal of Machine Learning Research*, vol. 9, no. 11, pp. 2579–2605, 2008.
- [184] J. Wu, X. Li, W. Liu, and Z. J. Wang, “sEMG signal processing methods: A review,” in *Journal of Physics: Conference Series*, vol. 1237, no. 3, 2019.
- [185] S. Shen, X. Wang, F. Mao, L. Sun, and M. Gu, “Movements classification through sEMG with convolutional vision transformer and stacking ensemble learning,” *IEEE Sensors Journal*, vol. 22, no. 13, pp. 13 318–13 325, 2022.
- [186] T. Kim, J. Kim, B. Koo, H. Jung, Y. Nam, Y. Chang, S. Park, and Y. Kim, “Effects of sampling rate and window length on motion recognition using sEMG armband module,” *International Journal of Precision Engineering and Manufacturing*, vol. 22, no. 8, pp. 1401–1411, 2021.
- [187] P. Tsinganos, B. Cornelis, J. Cornelis, B. Jansen, and A. Skodras, “Data augmentation of surface electromyography for hand gesture recognition,” *Sensors*, vol. 20, no. 17, 2020.
- [188] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, “Understanding deep learning (still) requires rethinking generalization,” *Communications of the ACM*, vol. 64, no. 3, pp. 107–115, 2021.
- [189] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, “DEAP: A database for emotion analysis; using physiological signals,” *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2011.
- [190] M.-H. Lee, O.-Y. Kwon, Y.-J. Kim, H.-K. Kim, Y.-E. Lee, J. Williamson, S. Fazli, and S.-W. Lee, “EEG dataset and OpenBMI toolbox for three BCI paradigms: An investigation into BCI illiteracy,” *GigaScience*, vol. 8, no. 5, 2019.
- [191] A. H. Al-Timemy, G. Bugmann, and J. Escudero, “Adaptive windowing framework for surface electromyogram-based pattern recognition system for transradial amputees,” *Sensors*, vol. 18, no. 8, 2018.
- [192] H.-J. Hwang, J. M. Hahne, and K.-R. Müller, “Real-time robustness evaluation of regression based myoelectric control against arm position change and donning/doffing,” *PLoS One*, vol. 12, no. 11, 2017.
- [193] M. Motamedi, N. Sakharlykh, and T. Kaldewey, “A data-centric approach for training deep neural networks with less data,” *arXiv preprint arXiv:2110.03613*, 2021.
- [194] C. Sartzetaki, P. Antoniadis, N. Antonopoulos, I. Gkinis, A. Krasoulis, S. Perdakis, and V. Pitsikalis, “Beyond within-subject performance: A multi-dataset study of

- fine-tuning in the EEG domain,” in *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2023, pp. 4429–4435.
- [195] M. Atzori, M. Cognolato, and H. Müller, “Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands,” *Frontiers in Neurorobotics*, vol. 10, 2016.
 - [196] C. Rommel, J. Paillard, T. Moreau, and A. Gramfort, “Data augmentation for learning predictive models on EEG: A systematic comparison,” *Journal of Neural Engineering*, vol. 19, no. 6, 2022.
 - [197] F. Wang, S.-h. Zhong, J. Peng, J. Jiang, and Y. Liu, “Data augmentation for EEG-based emotion recognition with deep convolutional neural networks,” in *MultiMedia Modeling: 24th International Conference*. Springer, 2018, pp. 82–93.
 - [198] M. Re and G. Valentini, *Ensemble methods: A review*, 01 2012, pp. 563–594.
 - [199] N. de Condorcet, *Essai sur l’application de l’analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris: Imprimerie Royale, 1785.
 - [200] M. Ahn, H. Cho, S. Ahn, and S. C. Jun, “High theta and low alpha powers may be indicative of BCI-illiteracy in motor imagery,” *PLoS One*, vol. 8, no. 11, 2013.
 - [201] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, “Filter bank common spatial pattern (FBCSP) in brain–computer interface,” in *Proceedings of the IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*. IEEE, 2008, pp. 2390–2397.
 - [202] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, “EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces,” *Journal of Neural Engineering*, vol. 15, no. 5, 2018.
 - [203] X. Liu, S. Xiong, X. Wang, T. Liang, H. Wang, and X. Liu, “A compact multi-branch 1D convolutional neural network for EEG-based motor imagery classification,” *Biomedical Signal Processing and Control*, vol. 81, 2023.
 - [204] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2017, pp. 1251–1258.
 - [205] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” *Advances in neural information processing systems*, vol. 27, 2014.

- [206] S. T. Foldes, D. J. Weber, and J. L. Collinger, “MEG-based neurofeedback for hand rehabilitation,” *Journal of NeuroEngineering and Rehabilitation*, vol. 12, pp. 1–9, 2015.
- [207] D. Rathee, H. Raza, S. Roy, and G. Prasad, “A magnetoencephalography dataset for motor and cognitive imagery-based brain–computer interface,” *Scientific Data*, vol. 8, no. 1, 2021.
- [208] S. A. Raurale, J. McAllister, and J. M. del Rincon, “Real-time embedded emg signal analysis for wrist-hand pose identification,” *IEEE Transactions on Signal Processing*, vol. 68, pp. 2713–2723, 2020.
- [209] H. He and D. Wu, “Transfer learning for brain–computer interfaces: A euclidean space data alignment approach,” *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 2, pp. 399–410, 2019.
- [210] A. Sacco, “Hands on with thalamic labs’ myo gesture control armband,” *CIO*, Jun. 2015, accessed: 2026-02-03. [Online]. Available: <https://www.cio.com/article/247234/thalamic-labs-myo-review-gesture-control-armband.html>
- [211] Y. Lin, Y. Yuan, J. Chen, and X. Lin, “Motor imagery combined with brain–computer interface for stroke patients: a meta-analysis,” *Frontiers in Neurology*, vol. 17, p. 1672882, 2026.
- [212] S. Kanoga, T. Hoshino, and M. Tada, “A style transfer mapping and fine-tuning subject transfer framework using convolutional neural networks for surface electromyogram pattern recognition,” in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2022, pp. 1171–1175.
- [213] X. Niu, X. Zhang, and L. Yang, “MI-EEG recognition based on euclidean alignment and style transfer mapping,” in *2022 2nd International Conference on Computational Modeling, Simulation and Data Analysis (CMSDA)*. IEEE, 2022, pp. 184–188.
- [214] S. Ran, W. Zhong, D. Duan, L. Ye, and Q. Zhang, “SSTM-IS: Simplified STM method based on instance selection for real-time EEG emotion recognition,” *Frontiers in Human Neuroscience*, vol. 17, 2023.
- [215] L. Chen, Q. She, M. Meng, Q. Zhang, and J. Zhang, “Similarity constraint style transfer mapping for emotion recognition,” *Biomedical Signal Processing and Control*, vol. 80, 2023.
- [216] W. Huang, Z. Guan, K. Li, Y. Zhou, and Y. Li, “An affective brain–computer interface based on a transfer learning method,” *IEEE Transactions on Affective Computing*, 2023.

- [217] K. Englehart, B. Hudgin, and P. A. Parker, “A wavelet-based continuous classification scheme for multifunction myoelectric control,” *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 3, pp. 302–311, 2001.
- [218] K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, “Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b,” *Frontiers in Neuroscience*, vol. 6, p. 39, 2012.
- [219] T. Briggs, T. Oates *et al.*, “Discovering domain-specific composite kernels,” in *Proceedings of the Twentieth National Conference on Artificial Intelligence and the Seventeenth Annual Conference on Innovative Applications of Artificial intelligence*, 2005, pp. 732–739.
- [220] D. H. Wolpert and W. G. Macready, “No free lunch theorems for optimization,” *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 67–82, 1997.

付 録 A 機械学習の基礎

A.1 緒言

本研究では、生体信号からユーザの意図や状態を推定するために、機械学習の技術を多用している。機械学習は近年急速な発展を遂げている分野であり、多岐にわたるアルゴリズムや専門用語が存在する。本分野を専門としない読者にとって、本論文の第 3 章以降で展開される手法の選択理由や、数理的な挙動を直感的に理解することが難しい場合がある。そのため、本付録では、本論文を読み解く上で前提となる機械学習の基本概念、および頻出する専門用語の定義を整理する。

本付録の構成は以下の通りである。まず、機械学習の基礎となる学習の枠組みについて概説する。次に、生体信号のような複雑なデータを扱う上で重要となる、モデルの線形性と非線形性について述べる。さらに、限られた生体信号データから頑健なモデルを構築するために避けられない問題である過学習と、それを抑制する正則化について解説する。最後に、本研究で採用している機械学習に関する用語を略説する。

A.2 機械学習の枠組み

機械学習は、学習データに付与される情報の性質（正解ラベルの有無）や学習の目的に応じて大きく、教師あり学習・教師なし学習・強化学習の 3 つに大別される。さらに近年では、データの収集コストという現実的な制約を克服するため、半教師あり学習や自己教師あり学習といった発展的な枠組みも重要視されている。本研究においては、教師あり学習と教師なし学習をそれぞれ使用しているため、これらを中心に各枠組みの定義と本研究における位置づけを整理する。

A.2.1 教師あり学習 (Supervised Learning)

教師あり学習とは、入力データ（特徴量ベクトル） $\boldsymbol{x}$ と、それに対する正解ラベル（教師信号） y のペアの集合から、入力と出力の関係性を学習する枠組みである。さらに、出力が離散値の場合は識別、連続値の場合は回帰と、さらに細分化されている。生体信号型 HCI の文脈においては、ユーザがどの動作（例：手首の屈曲、伸展など）を行ったかをカテゴリとして予測する場合は識別となる。一方、関節の曲げ角度（ $0^\circ - 90^\circ$ ）や発揮している筋力（ $0 - 100\%$ ）といった連続値を推定する場合は回帰として定式化される。本研究においては、

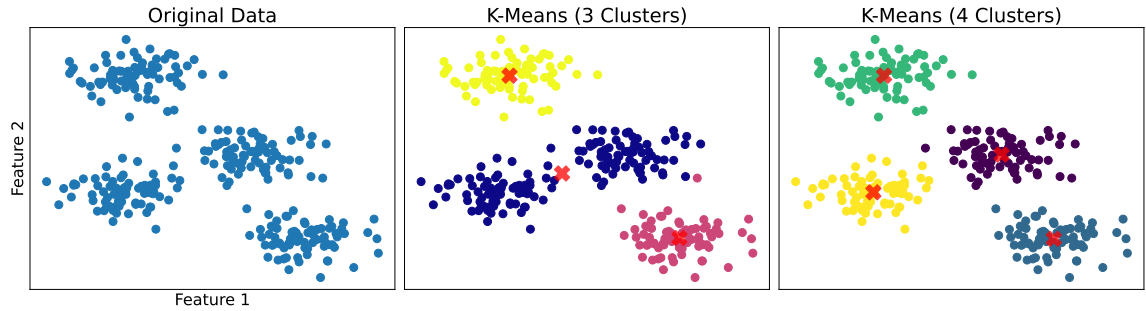


図 A.1: k-means 法によるクラスタリングの概念図. (左) クラスタリング前のデータ分布, (中央) 3つのクラスタへの分割例 ($k = 3$), (右) 4つのクラスタへの分割例 ($k = 4$). 図中の $\times$ 印は, 算出された各クラスタの重心 (セントロイド) を示す.

前者の識別の枠組みを主に使用し, ユーザが行った運動課題や, ユーザが想起した課題の予測を行う.

A.2.2 教師なし学習 (Unsupervised Learning)

教師なし学習とは, 教師あり学習とは異なり, 正解ラベル y を一切使用せず, 入力データ x のみを用いて, データに内在する構造やパターンを抽出する枠組みである. 代表的な応用として, 類似したデータをグループ化するクラスタリングや, データの次元削減などが挙げられる. 第 3 章では, STM における代表点 (プロトタイプ) の選定にクラスタリングが利用されており, 具体的には k-means という手法が採用されている. これは, 図 A.1 に示すように, 入力されたデータを事前に設定した個数のグループ (クラスタ) に自動的に分割する手法である. 本研究では, 分割された各クラスタの中心 (重心) を算出し, これを代表的なデータ点としてみなし, STM の写像先として利用している. また, 第 4 章では, 特徴量の次元削減に t-SNE という手法を採用している. 特徴量を可視化して分析する場合, データが 2 次元や 3 次元であれば可視化は容易であるが, より高次元の場合には視覚的に理解しやすい可視化は困難である. これに対し, t-SNE は, 元々の高次元の空間において互いに近い関係にあるデータ点同士が, 圧縮された低次元空間においても近くなるように写像する非線形次元削減手法である. これにより, 第 4 章では, 88 次元のデータを 2 次元に圧縮し可視化することで, 視覚的な分析を行っている.

A.3 モデル選択と汎化性

A.3.1 線形性と非線形性

機械学習モデルは, 入力データと出力データの関係性をどのような関数で表現するかによって, 線形モデルと非線形モデルに分類される. これらの差は, 識別問題においてはデータを分ける決定境界の性質として現れ, 回帰問題においてはデータを近似する回帰線の性質

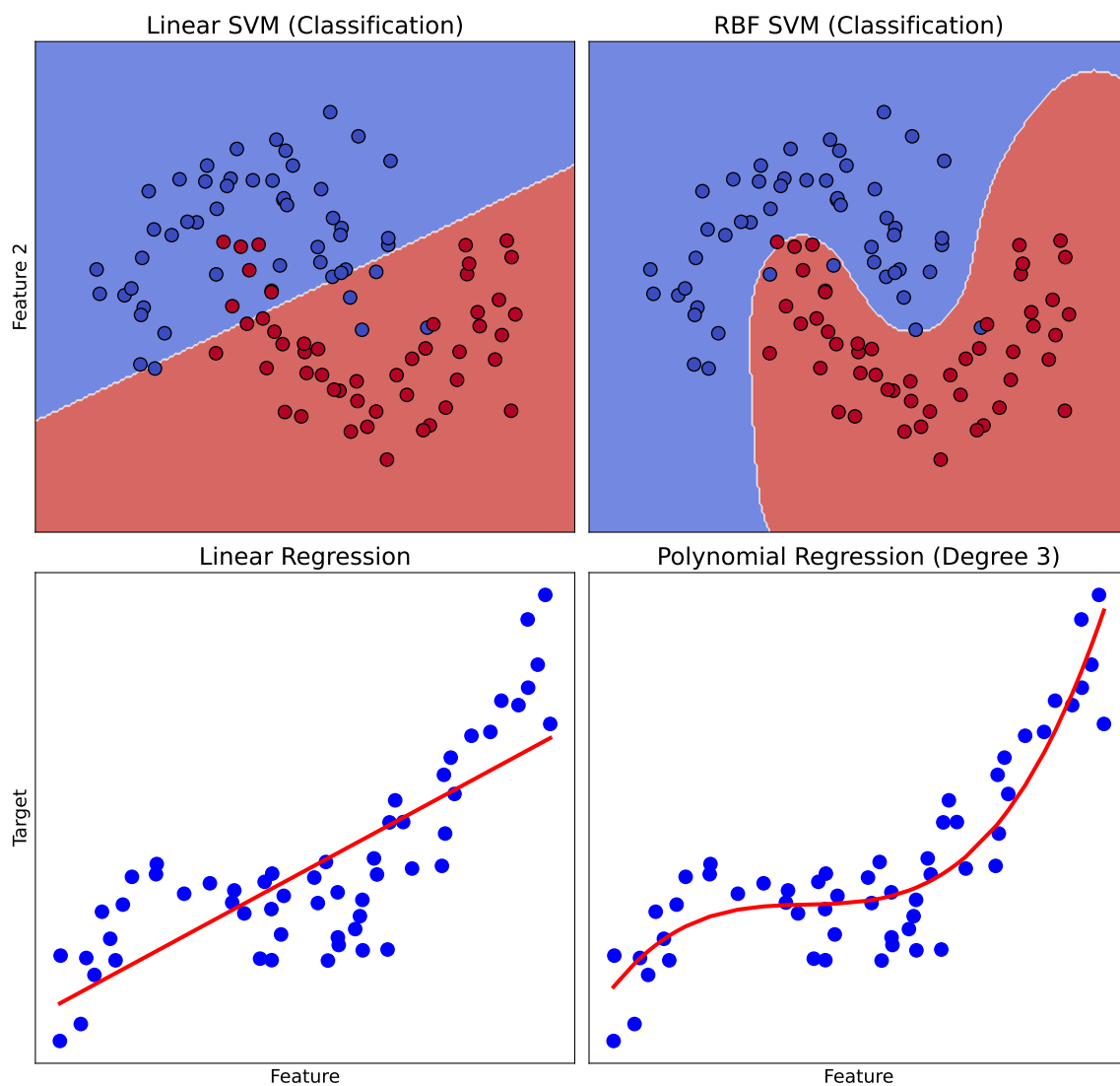


図 A.2: 線形モデルと非線形モデルの挙動の比較. 上段は識別問題における決定境界を示しており, 左上が線形カーネルによる線形識別, 右上が RBF カーネルによる非線形識別. 下段は回帰問題における回帰線を示しており, 左下が単回帰による線形回帰, 右下が多項式回帰による非線形回帰.

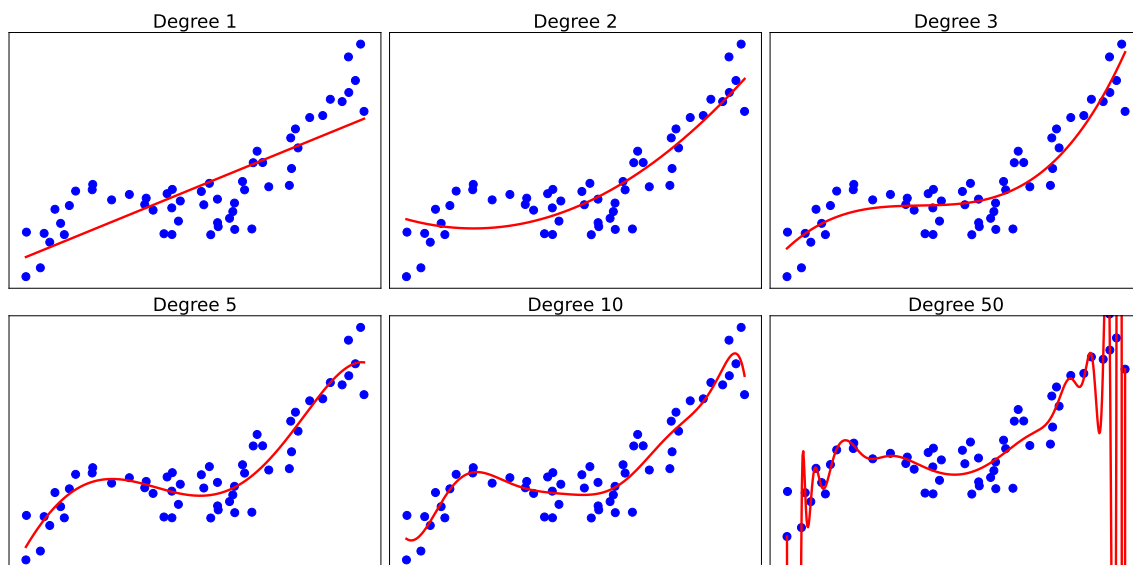


図 A.3: 多項式回帰の次数を徐々に増やした場合の例。

として現れる．図 A.2 は，識別と回帰における線形と非線形の違いを描いたものである．線形モデルは計算コストが低く，後述の過学習を起こしにくいという利点があるが，複雑なデータ分布を表現することには限界がある．一方で非線形モデルは，複雑なデータの構造を捉えることができる利点があるが，計算コストが高くなりがちであり，過学習を引き起こしやすい欠点がある．したがって，学習するデータの特徴に合わせて適切なモデルを選択することが重要である．

A.3.2 汎化性能と過学習

前項で線形モデルと非線形モデルの違いを概観したが，一見すると常に非線形なモデルを選択することが望ましいように思える．しかし，機械学習においてモデルを高度に非線形化させ表現力を高めることは必ずしも成功せず，むしろ失敗につながることが多い．これは，機械学習の究極の目的は，手元にある学習データを完璧に再現することではなく，まだ見ぬ未知のデータに対して正しい予測を行う能力，すなわち汎化性能を獲得することにあるためである．本項では，汎化性能を巡る問題と，モデルの複雑さを制御する技術について概説する．

汎化性能の確認に関しては，識別よりも回帰の方が直感的にわかりやすいため，回帰を中心に概説する．ここでは，図 A.2 でも用いた，多項式回帰を例にとり，モデルの複雑さ（多項式の次数）を変化させた場合の挙動を見ていく．多項式回帰とは，入力 x と出力 y の関係を d 次の多項式関数で近似する手法であり，以下のように表される．

$$y = \sum_{i=0}^d w_i x^i = w_0 + w_1 x + w_2 x^2 + \cdots + w_d x^d$$

ここで、 w_i は学習によって最適化される重みパラメータであり、 d はモデルの複雑さを決定する次数である。次数 d が 1 の場合は単純な線形モデルとなるが、 d が大きくなるにつれて曲線が自由に曲がれるようになり、高度な非線形モデルへと変化していく。

図 A.3 は d を徐々に大きくさせた場合の結果を示している。次数が 1 や 2 の場合、回帰線はデータの大きな傾向は掴んでいるものの、当てはまりが不十分となっている。次数が徐々に上がるにつれてデータによくフィットした曲線が得られるが、次数が 50 になった場合には学習データの各点を無理に通ろうとするあまり、データ本来の傾向からは逸脱した、不自然に波打つ曲線となっている。さらに次数を高めると、学習データの全てを通るような極めて複雑な曲線を学習することもでき、学習データに対する誤差をほぼゼロにすることも可能である。しかし、そのような曲線は学習データに含まれるノイズまで正確に暗記しているだけであり、未知のデータが入力された際には本来の傾向から大きく乖離し、全く見当外れの予測をしてしまう危険性を孕んでいる。機械学習においては、このようにモデルが学習データの細部に過度に適合してしまい、未知のデータに対する汎化性能を失ってしまった状態を過学習と呼ぶ。生体信号はノイズを多く含み、さらにデータ収集の困難さからデータ数が限られることが多いため、非線形モデルを用いた場合には常にこの過学習が深刻な問題となる。

A.3.3 正則化

モデルを必要以上に複雑にさせると過学習を引き起こしてしまうが、この過学習を抑制し汎化性能を維持するための技術の総称を正則化と呼ぶ。正則化手法として最も代表的なものは、学習時の損失関数に対して、モデルの重みの大きさに応じたペナルティを加算する手法である。ここでは、引き続き d 次の多項式回帰を例に挙げる。まず、通常多項式回帰では、予測値と実際のデータとの誤差の二乗和を最小化する以下の損失関数 $L(\mathbf{w})$ が用いられる。

$$L(\mathbf{w}) = \sum_{n=1}^N \left(y_n - \sum_{i=0}^d w_i x_n^i \right)^2$$

これに対して、L2 正則化を適用した場合には、 $L(\mathbf{w})$ に加えて、重みベクトルの大きさを示す L2 ノルムの二乗をペナルティ項として加算した、以下の目的関数 $J(\mathbf{w})$ を最小化する。

$$J(\mathbf{w}) = L(\mathbf{w}) + \lambda \|\mathbf{w}\|_2^2 = L(\mathbf{w}) + \lambda \sum_{i=1}^d w_i^2$$

ここで、 λ は正則化の強さを制御するためのハイパーパラメータ（後述）であり、ここでは 0 以上の値を取る。このペナルティ項が加わることで、学習は「誤差を小さくしつつ、かつ重みの大きさもできるだけゼロに近づける」という妥協点を探るようになる。

図 A.4 は正則化の強度 λ を変化させた場合の違いを表している。ここでは極端な過学習の例を見るために、20 個データに対して 19 次の多項式回帰を行っている。この場合、回帰曲線は全てのデータ点を通ることができ学習データに対する誤差は 0 になるが、強く過学習を引き起こしていることが見て取れる。これに対し、画像中央のように λ を適切に設定

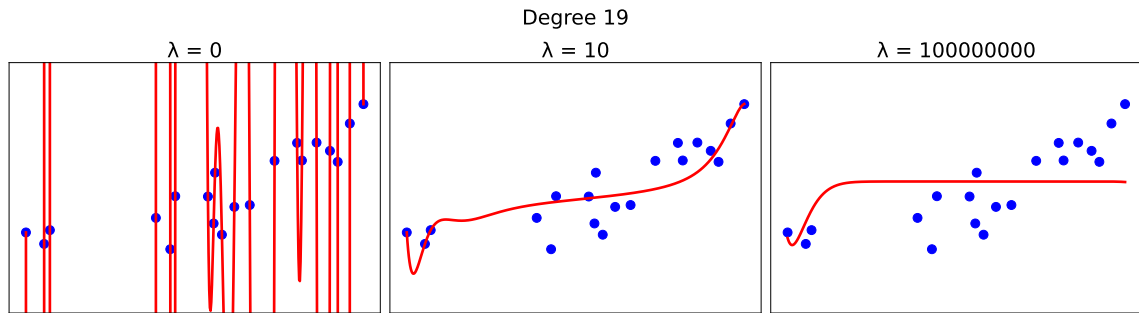


図 A.4: 多項式回帰に対する正則化の働き。

することでモデルの複雑性を適度に抑制し、データ本来の滑らかな傾向を捉えることができる。しかし、画像右のように、極端に λ を大きくしてしまうと、本来必要な複雑性までもが失われ、かえって精度を落とす原因となってしまう。以上のように、正則化は過学習を抑制する強力な手法であるが、その強さはデータセットに応じて適切に設定する必要がある。

機械学習において、モデルがデータから自動的に学習する重みパラメータとは異なり、 λ のように人間が事前に設定して学習プロセスそのものを制御する変数は、通常のパラメータの上位概念という意味で「ハイパーパラメータ」と呼ばれる。また、正則化の強さ λ だけにとどまらず、そもそも何次式でデータを近似するかという次数や、ニューラルネットワークにおける層の深さなど、モデルの構造自体を決定づける要素も同様にハイパーパラメータに分類される。したがって、未知のデータに対して高い汎化性能を獲得するためには、モデルの構造的複雑さと、正則化による制約の強さの両面から、データセットに適したハイパーパラメータの組み合わせを探索・決定することが不可欠である。

ハイパーパラメータの最適化には、手元にあるデータを学習用と評価用に分けて、未知のデータに対する汎化性能を疑似的に測定する必要がある。もし手元のデータを全て学習に使ってしまうと、学習データに対する誤差を小さくすることはできても、未知のデータに対して過学習を起こしているかどうかを判断できない。そこで、データセットを、モデルの学習に用いる学習データと、ハイパーパラメータの調整にのみ用いる検証データに分割する手法がとられる。さらに、特定のデータの偏りに依存せず、より信頼性の高い汎化性能を推定するための枠組みとして交差検証 (cross-validation) が広く用いられる。最も代表的な K-分割交差検証では、データセットを K 個のグループに分割する。そして、そのうちの 1 つを評価用データ、残りの $K - 1$ 個を学習用データとしてモデルの評価を行うプロセスを、評価用データのグループを入れ替えながら K 回繰り返す。最終的に得られた K 回の評価結果の平均をとることで、データセット全体に対する安定した性能評価が可能となる。このように、ハイパーパラメータによって、疑似的に未知のデータに対する汎化性能がどう変化するかを観察することで、最適なハイパーパラメータの選択を行う。

本項では、モデルの制約や評価手法の工夫によって過学習を抑制する方法を述べてきた。しかし、過学習を防ぎ汎化性能を高めるための最も本質的かつ強力なアプローチは「学習

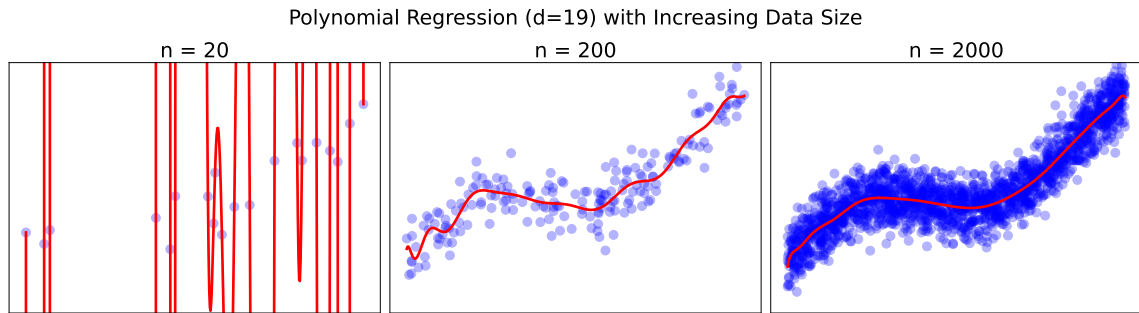


図 A.5: 学習データ数の増加による過学習の抑制効果.

データそのものの数を増やす」ことである。図 A.5 は、 $d = 19$ かつ $\lambda = 0$ の多項式回帰において、学習データの数のみを増加させた場合の挙動を示している。使用しているハイパーパラメータは同じにも関わらず、データ数が増えるにつれて、データの本質を捉えたような曲線が得られる様子が見て取れる。これは、データ数が増えるほど、個々のデータに含まれるノイズの影響は相対的に小さくなり、データが本来持つ真の傾向が顕在化するためである。しかしながら、本研究が対象とする生体信号においては、データの収集には被験者への電極装着や長時間の運動タスクの遂行など、多大な身体的・時間的コストが伴う。このように十分なデータを集めることが物理的に困難であるという生体信号特有の制約こそが、本分野における最大の課題であり、適切なハイパーパラメータの設定の重要性を物語っている。

A.4 本研究で採用した機械学習モデルとその構成要素

本研究の第 3 章以降では、生体信号の識別において、伝統的な機械学習の代表としてサポートベクターマシン (SVM)、深層学習の代表として多層パーセプトロン (MLP; 第 3, 4 章では ANN と呼称) および畳み込みニューラルネットワーク (CNN) を採用している。本節では、これらのモデルがどのように非線形な識別を実現し、またどのような要素技術によって学習の安定化や効率化を図っているのか、その構成要素について概説する。

A.4.1 サポートベクトルマシン

図 A.6 は、線形分離可能なデータに対して決定境界がどのように引かれるかを示している。単にデータを分離するという目的に対しては、左図のように無数の決定境界が考えられる。これに対し、SVM は右図のように各クラスのデータ点から最も離れた位置に決定境界を配置する、マージン最大化を行う。このマージンの最大化は、学習データに対する分類性能を確保するだけでなく、境界付近に十分な余裕 (マージン) を設けることで、データに含まれるノイズや変動に対する堅牢性を高めることを狙ったものである。

しかし、生体信号のような複雑なデータは、元の特徴空間において線形に分離できることは稀である。そこで SVM では、カーネル法と呼ばれる技術を用いて非線形な識別を実現

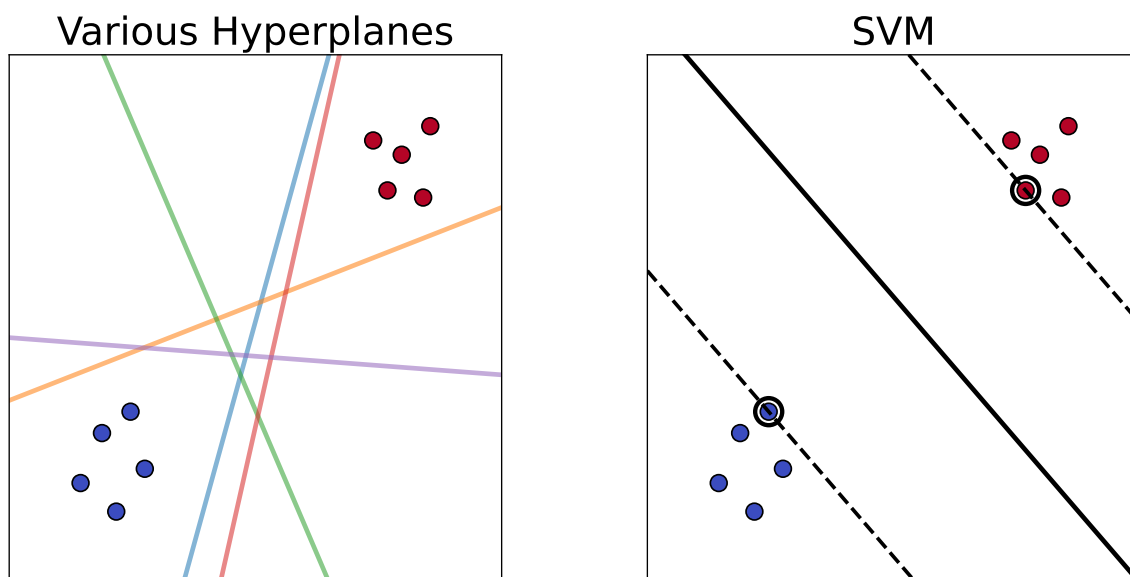


図 A.6: 線形分離可能なデータに対する識別境界の比較. (左) 2 クラスを分離する超平面は複数存在し得る例. (右) SVM はサポートベクター (白抜き円) により決定される決定境界 (実線) を, マージン (破線で示す両側の境界間隔) が最大となるように選択.

する. カーネル法とは, 元の低次元の特徴空間では線形分離不可能なデータを, 非線形な写像関数によってより高次元の空間へと変換し, その高次元空間において線形分離を行う手法である. ただし, この高次元空間への写像を明示的に計算すると計算量が膨大になる. そこで SVM では, カーネルトリックという数学的な性質を利用し, 計算量を大幅に削減する. カーネルトリックとは, 高次元空間での内積計算を, 元の低次元空間の関数計算に置き換える数学的な手法である. 具体的には, 非線形写像 Φ によってデータを高次元空間へ飛ばした際, その空間上での 2 点間の内積 $\Phi(\mathbf{x}) \cdot \Phi(\mathbf{y})$ の値が, 元の空間の変数を用いた特定の関数 $K(\mathbf{x}, \mathbf{y})$ の計算結果と一致することを利用するものである.

ここでは, 図 A.7 に示した高次元写像の具体的な仕組みを, 理解が容易な多項式カーネル (次数 2) を例に挙げて説明する. 元の 2 次元空間におけるデータベクトルを $\mathbf{x} = (x_1, x_2)$, $\mathbf{y} = (y_1, y_2)$ とする. また, 高次元空間への写像関数を $\Phi(\mathbf{x}) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)$ と定義する. このとき, 写像後の 3 次元空間における 2 点間の内積は次のように計算される.

$$\Phi(\mathbf{x}) \cdot \Phi(\mathbf{y}) = x_1^2 y_1^2 + 2x_1 x_2 y_1 y_2 + x_2^2 y_2^2$$

一方で, 元の 2 次元空間における内積 $\mathbf{x} \cdot \mathbf{y}$ を用い, カーネル関数を $K(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y})^2$ と定義して展開すると以下ようになる.

$$K(\mathbf{x}, \mathbf{y}) = (x_1 y_1 + x_2 y_2)^2 = x_1^2 y_1^2 + 2x_1 x_2 y_1 y_2 + x_2^2 y_2^2$$

両者の結果は完全に一致しており, これは高次元空間へ実際に写像して内積を計算することと, 元の空間でカーネル関数を計算することが等価であることを示している. このように, カーネルトリックを用いることで, 写像関数 Φ を具体的に構成 (明示的な座標計算) する

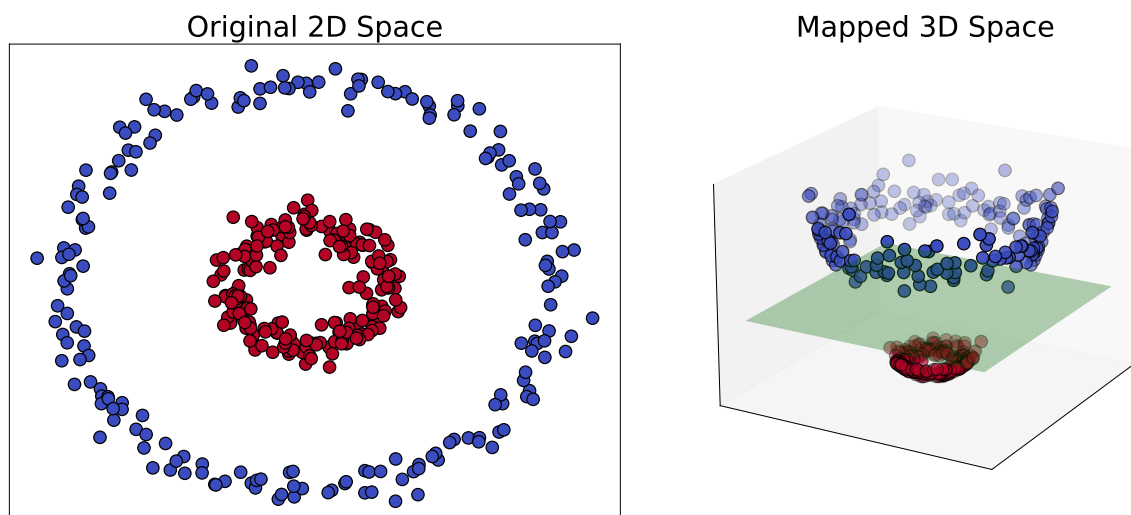


図 A.7: カーネル法による高次元写像と非線形分離の概念図.

ことなく，図 A.7（右）に示されるような高次元空間における線形識別を，元の低次元空間の演算のみで効率的に遂行できるのである．

第 3 章では，Radial Basis Function (RBF) カーネルと呼ばれる非線形カーネルを使用した．前述した多項式カーネルが有限次元への写像を行うのに対し，RBF カーネルは数学的に無限次元の特徴空間への写像を内包している．その関数形式は，2 点間の距離のガウス分布として $K(\mathbf{x}, \mathbf{y}) = \exp(-\gamma \|\mathbf{x} - \mathbf{y}\|^2)$ と定義される．この性質により，RBF カーネルはデータ間の局所的な類似性を極めて柔軟に捉えることが可能であり，複雑な分布を持つ生体信号の識別において非常に高い性能を発揮する．RBF カーネルを使用する SVM においては，誤分類の許容度を制御するペナルティ C と，RBF カーネルの関数の広がり方を決定する γ の 2 つがハイパーパラメータとなる．

A.4.2 深層学習

(1) 多層パーセプトロン

本研究の第 3 章から第 5 章において，主たる識別器として採用されているのが MLP である．MLP は，深層学習モデルの中で最も基礎的な構造を持ち，図 A.8 のように，入力層，複数の隠れ層，および出力層から構成される．最大の特徴は，隣接する層の全てのニューロンが互いに結合した全結合層を基本単位としている点である．丸で表される各ニューロンは，前層のニューロンからの出力にそれぞれ重みを掛け合わせ，さらにバイアスを足し合わせるといった線形変換を行う．この重みとバイアスが，学習データから最適化される重みパラメータである．MLP においては，隠れ層の数やニューロンの数がハイパーパラメータである．

線形変換を何層重ねたとしても，ネットワーク全体としては単一の線形変換と等価になってしまい，生体信号のような複雑なデータを表現することはできない．そこで，各ニューロ

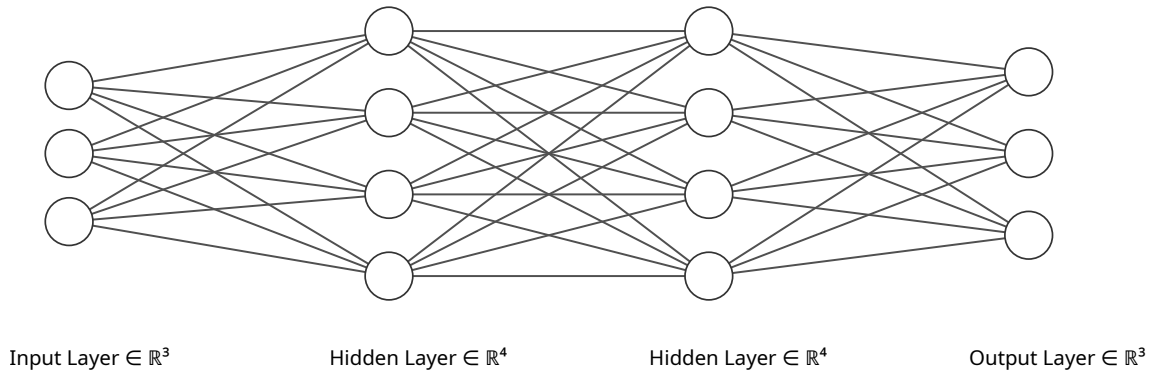


図 A.8: MLP の模式図. 本図の作成には NN-SVG [15] を使用した.

ンの出力に対して非線形な活性化関数を適用することで、ネットワークに非線形な表現力を与える. この活性化関数として最も代表的なものが Rectified Linear Unit (ReLU) 関数である. ReLU は、入力 x に対して以下の式で定義される単純な非線形関数である.

$$f(x) = \max(0, x)$$

この関数は、入力が 0 以下の場合は 0 を、0 より大きい場合はそのままの値を出力する. 従来のニューラルネットワークで主流であったシグモイド関数などは、層が深くなるにつれて学習に必要な勾配情報が減衰して消滅してしまう勾配消失問題を引き起こしやすかった. これに対し、ReLU は正の領域で勾配が常に 1 となるためこの問題を防ぐことができ、計算効率の良さも相まって、現在の深層学習において標準的に用いられている.

第 3 章以降の実験において、限られた生体信号データから頑健な MLP を学習させるため、いくつかの重要な技術を導入している. まず、深層学習モデルはパラメータ数が膨大であり、そのままでは前節で述べた過学習のリスクが極めて高い. 本研究では、これを防ぐために L2 正則化（深層学習の文脈では重み減衰 (weight decay)」と呼ばれる) を適用し、パラメータの極端な増大を抑制している. さらに、ネットワークの各隠れ層に入力されるデータの分布を標準化（平均 0, 分散 1 に変換）するバッチ正規化 (batch normalization) を導入している. これにより、学習の進行に伴って層への入力分布が変化してしまう内部共変量シフトと呼ばれる現象を抑制し、学習を高速化・安定化させることができる. また、パラメータの更新（最適化）には、パラメータごとに学習率を自動で適応的に調整する最適化アルゴリズムである Adam (Adaptive Moment Estimation) を採用している. これらの構成要素を組み合わせることで、ノイズの多い生体信号データにおいても、安定して高い汎化性能を発揮するネットワークの構築を実現している.

(2) 畳み込みニューラルネットワーク

本研究の第 5 章では、MLP に加えて、画像や時系列データのような局所的な相関関係を持つデータの処理に特化した、CNN を採用している. 前述の MLP における全結合層は、入力データの各要素をそれぞれ独立した情報として扱うため、波形の時間的な連続性や電極

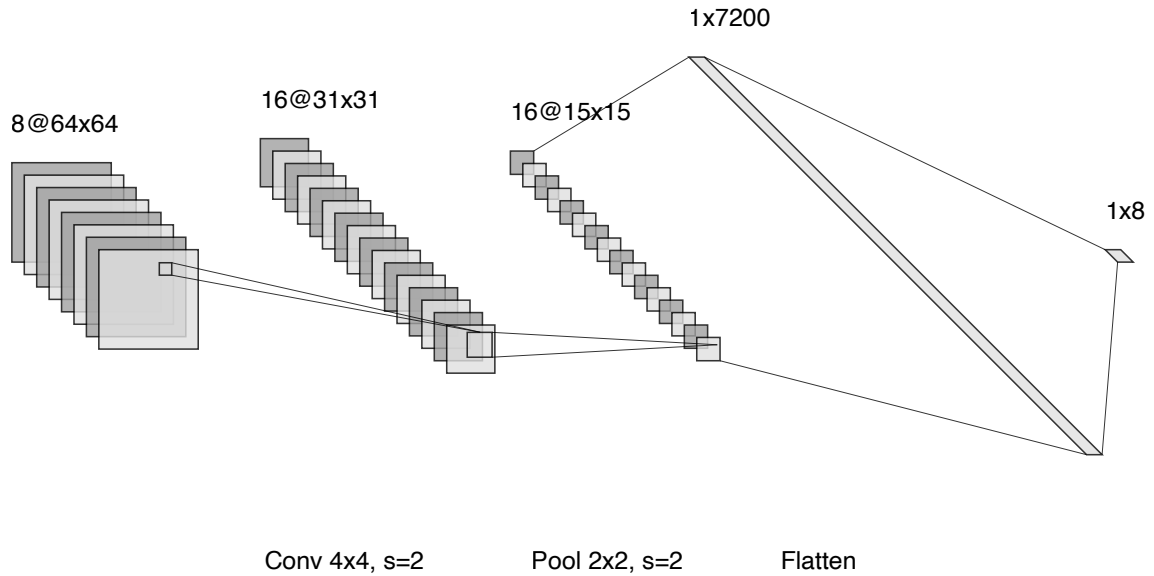


図 A.9: CNN の模式図. s はストライドの大きさ. 本図の作成には NN-SVG [15] を使用した.

間の空間的な配置関係といった、生体信号が本来持つ構造的な情報を直接的に利用することが難しい. これに対し, CNN は図 A.9 のように, 畳み込み層とプーリング層という独自の層を導入することで, データに内在する時空間的な構造を効果的に学習する.

畳み込み層では, フィルタと呼ばれる小さな重み行列を入力データ上でスライドさせながら積和演算を行う. これにより, フィルタのパターンと一致する局所的な特徴 (波形の立ち上がり, 特定の周波数成分, あるいは隣接電極間の電位差など) を抽出する. 本研究の第 5 章では, データセットの特性に応じて 2 種類の畳み込み層を使い分けている. 一つは時間方向の連続性のみを捉える 1 次元畳み込み (1D-CNN) であり, もう一つは時間的な変動に加えて, 多電極の空間的な配置関係も同時に考慮可能な 2 次元畳み込み (2D-CNN) である. CNN においては, 層をどのように設計するか, カーネルのサイズや枚数などがハイパーパラメータに当たる.

畳み込み層に続いて適用されるプーリング層は, 局所的な領域の情報を集約する役割を果たす. 本研究では, 外れ値に強く情報を広く拾うことが可能な平均プーリングを採用している. この処理により, 入力データに含まれる微小な位置ずれ, すなわち筋活動の開始時間のわずかなズレや, 電極の装着位置のズレに対する頑健性を獲得することができる. さらに, プーリング処理によってデータの空間サイズが縮小されるため, ネットワーク全体のパラメータ数を削減し, 過学習を抑制しつつ計算効率を高める効果もある. これらの構造的特性により, CNN は MLP とは異なるアプローチで生体信号の複雑なパターンを捉えることが可能となる.

なお, CNN においても MLP と同様に, ネットワークに非線形な表現力を与えるため, 畳み込み層の直後には活性化関数として ReLU が適用される. また, 学習の安定化と過学習の抑制という目的も共通しているため, CNN の各層にも 重み減衰とバッチ正規化が導入

され、最適化アルゴリズムには Adam が採用されている。これらの要素技術と CNN 独自の構造的特性（畳み込みとプーリング）を統合することで、CNN は MLP とは異なるアプローチで生体信号の複雑なパターンを捉える。

付 録 B カーネル法に基づくスタイル転移 マッピング

B.1 緒言

第 3 章では、学習済みの識別器を再学習なしで新規ユーザに適応可能な手法として STM を採用し、その有効性を検証した。STM は、図 B.1 に示すように、ターゲット被験者のデータをソース被験者のデータ分布上へと射影することで個人差を補正する。また、STM の学習は閉形式で解析的に求まるため高速であり、任意の識別モデルを再学習なしに再利用できるという利点がある。こうした実用性の高さから、STM は近年大きな注目を集めており、その性能や適用範囲を拡張する様々なアプローチが提案されている。Kanoga らは、畳み込みニューラルネットワークのファインチューニングの前に STM を適用する手法を導入し、後続の転移学習手法の前処理ステップとしての有効性を実証した [212]。Niu らは、STM とユークリッドアライメントという別の転移学習との組み合わせを提案し、STM が効果的な後処理ステップとしても機能することを示した [213]。Ran らは、選択されたサンプルのみに STM を適用するインスタンス選択アプローチを開発した [214]。Chen らは、限られたデータでも STM が良い性能を発揮できるようにする新しい正則化技術を導入した [215]。Huang らは、STM を用いて複数のソースユーザーからのデータを少数の代表的なドメインセットに融合する戦略を提案した [216]。以上のような多様な先行研究の存在は、STM の汎用性と強力な可能性を裏付けるものである。

しかしながら、第 3 章で議論したように STM には、写像関数が線形に限定されているという本質的な限界が存在する。図 B.2(a) に示すように、ソースとターゲットの関係が線形であれば STM は有効であるが、図 B.2(b) のような非線形な関係においては、分布の整合をとることができず、アンダーフィッティングを起こし識別精度の低下の要因となる。生体信号である sEMG や EEG は、第 2 章で述べたように、解剖学的・生理学的により強い非線形性が生じる。したがって、線形写像のみに依存する従来の STM は、これらの生体信号に内在する複雑な個人差を表現するには不十分である。

そこで本付録では、従来の STM をカーネル法によって拡張し、非線形写像を可能にした Kernel-based STM (k-STM) を提案する。k-STM は、データを高次元特徴空間へ写像することで、図 B.2(b) 右列に示すように、非線形な分布のズレを効果的に補正することを可能にする。本手法の有効性を検証するため、4 つの公開データセット (sEMG: MyoDataset [9], NinaProDB5 [12], EEG: OpenBMI [190], BCICIV2a [64]) を用いて評価を行った。実装においては、RBF カーネル、線形カーネル、およびその二つを組み合わせたハイブリッドカーネルを検討した。また、STM の識別器に依存しないという利点を活かし、ロジスティック

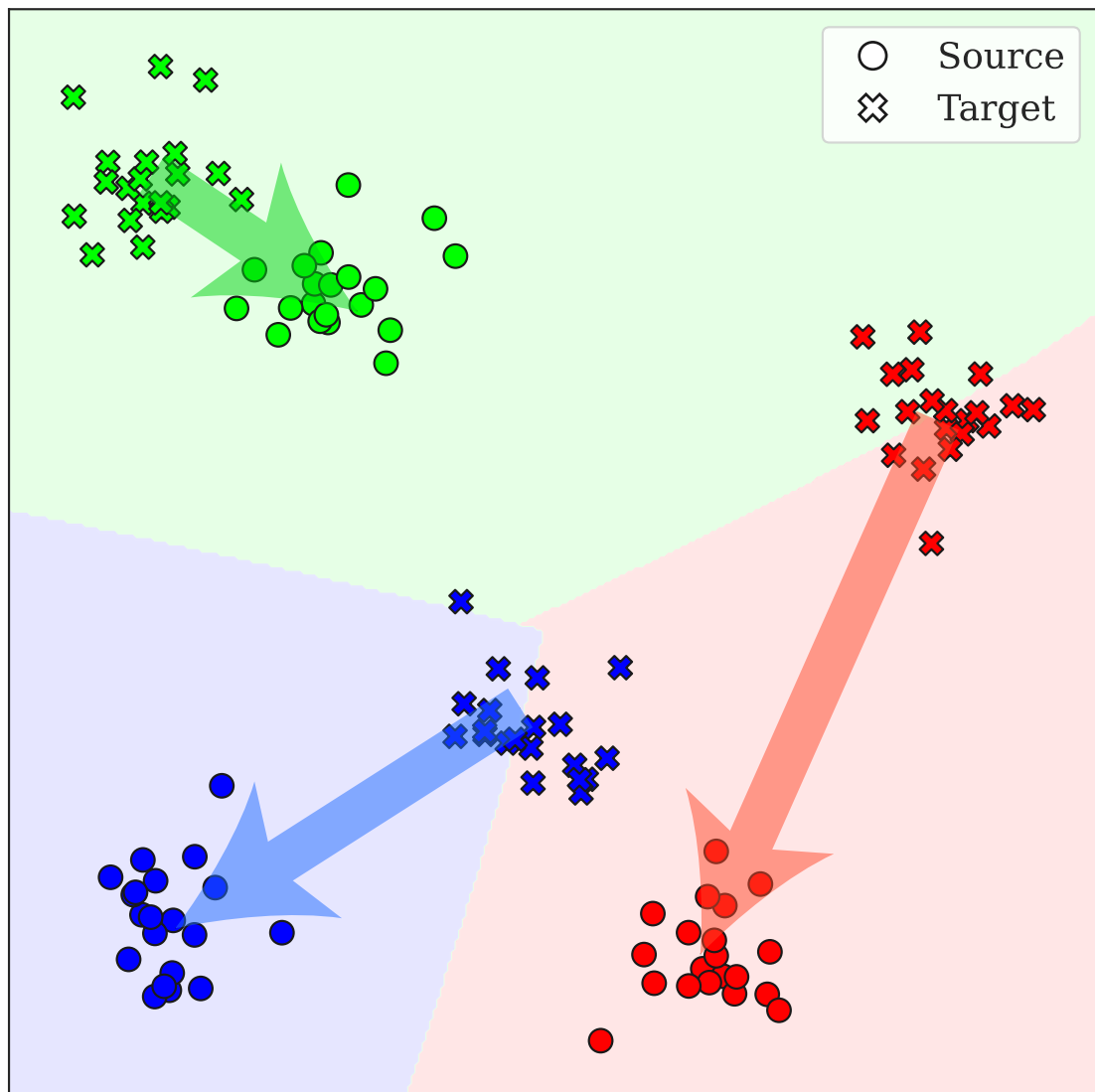


図 B.1: STM の概要. ソースユーザのデータを $\circ$, ターゲットユーザのデータを $\times$ で示し, 背景色はソースユーザのデータに基づいて学習された決定領域を表している. 被験者間変動の影響により, ターゲットユーザのデータには分布のシフトや不整合が生じており, ソースデータで学習された識別器では正確な識別が困難な状態にある. STM は, 矢印で示される方向へターゲットデータを写像することでこの問題に対処する. これにより, ターゲットデータをソースデータの分布へと適合させ, ソース識別器を用いた高精度な識別を可能にする. 文献 [16] より転載.

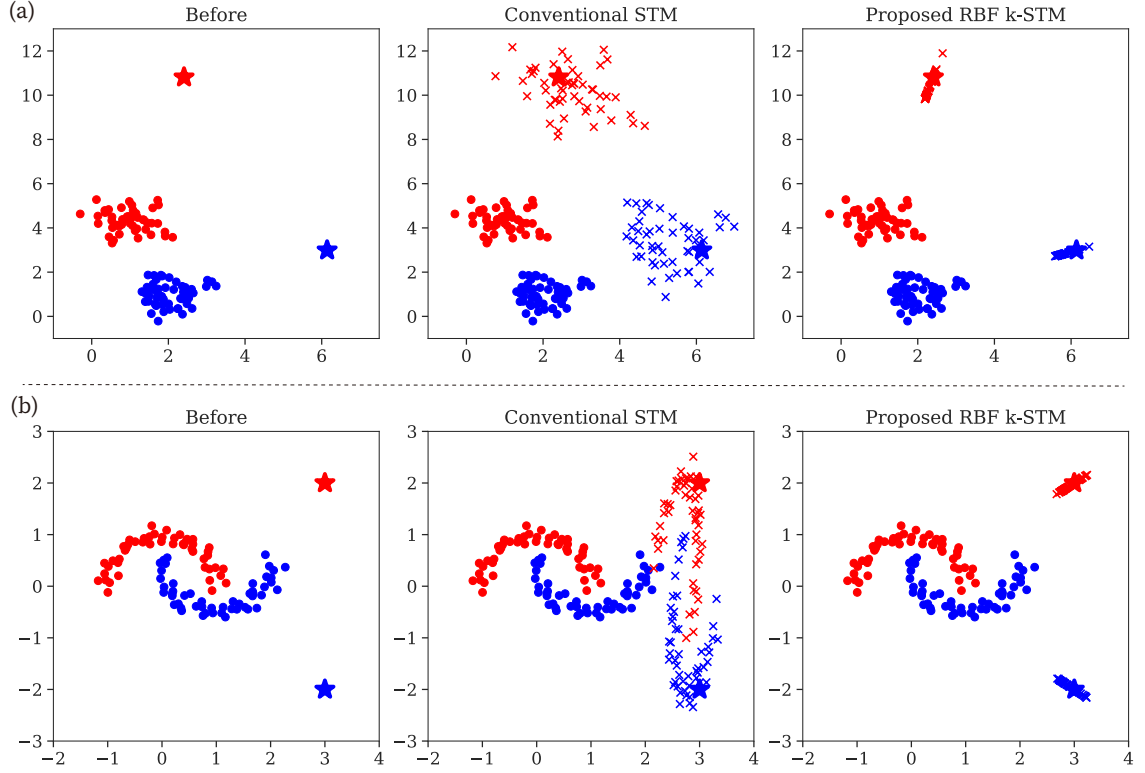


図 B.2: (a) 線形および (b) 非線形の関係下における STM (スタイル変換マッピング) の性能を示す例. 左の列は, ソースデータおよびターゲットデータの元の分布を示している. 中央の列は従来の STM を適用した場合を描写しており, 右の列は RBF (動径基底関数) ベースの k-STM を示している. 円マーカー (o) は元のターゲットサンプルを表しており, これらは星印 (★) で示されるソースドメインのプロトタイプに向かってマッピングされる. バツ印 (×) は, マッピング後のターゲットデータの位置を示している. 従来の STM は非線形なマッピングを効果的に捉えることができないのに対し, RBF ベースの k-STM は線形および非線形関係の両方を適切に処理している. 文献 [16] より転載.

回帰 (LR), サポートベクターマシン (SVM), ランダムフォレスト (RF), および多層パーセプトロン (MLP) という構造の異なる 4 種類の識別器を用いて評価を行った. 実験の結果, 非線形カーネルを使用した k-STM は, 従来の線形 STM よりも有意に高い識別精度を達成した. 本付録の結果は, 生体信号の個人差に含まれる非線形性を考慮することの重要性を示唆しており, k-STM がより表現力の高い転移学習手法として有効であることを示すものである.

B.2 データセットと前処理

B.2.1 データセット

本付録では, MyoDataset [9], NinaProDB5 [12], OpenBMI [190], および BCI Competition IV-2a (BCICIV2a) [64] の 4 つの公開データセットを用いた. 前者の 2 つは sEMG 信号, 後者の 2 つは EEG 信号により構成されている. 前処理および特徴抽出の手法については, 主にそれぞれの原著論文 [9, 12, 190, 64] に記載された方法に準拠した. なお, MyoDataset は第 3 以降使用しているデータセットであり, OpneBMI は第 5 で使用していたデータセットである.

B.2.2 MyoDataset

基本的な内容は 3.2.1 項で記述済みである. 健常者 25 名を対象とし, 単純な前腕動作 8 種類を 5 回ずつ計測した sEMG データを収録している.

前処理および特徴抽出は, 以下の手順で実施した. まず, 生の sEMG 信号に対し, カットオフ周波数 10 Hz の 5 次バターワースハイパスフィルタを適用した. 次に, マルチスケールサンプルエントロピーを用いて動作のオンセットを検出した. なお, ハイパーパラメータ N_w , N_s , m , r は先行研究 [146, 147] の設定を採用した. 各 6 秒間の試行記録から, 検出されたオンセットを始点として 1.5 秒間の区間を抽出した. 各区間は, 50 ms ずつシフトさせながら 250 ms の解析窓に分割され, これにより 1 試行あたり 26 個のセグメントが得られた [148]. 識別には, Huang らの時間領域および自己回帰 (TDAR) 特徴量を用いた [149]. その結果, 特徴ベクトル N_d の次元数は 88 (8 チャンネル $\times$ 11 特徴量) となった.

B.2.3 NinaPro DB5

基本的な内容は (2) 項で記述済みである. 健常者 10 名を対象とし, 単純な前腕動作 52 種類を 6 回ずつ計測した sEMG データを収録している. 本データセットでは, 各参加者は 3 つの実験タスクに参加しており, 本付録では 12 種類の指動作タスクで構成される Exercise A のデータのみを解析対象とした.

前処理および特徴抽出は、以下の手順で実施した。MyoDataset とは異なり、信号へのフィルタリングは行わなかった。動作のオンセットは、データセット提供者により提供された “restimulus” 変数 [12] を用いて特定し、これらのオンセットを始点として活動区間を抽出した。各区間は、100 ms ずつシフトさせながら 200 ms の解析窓に分割した [217]。識別には、Hudgins の時間領域特徴量 [48] を用いた。その結果、特徴ベクトル N_d の次元数は 64 (16 チャンネル $\times$ 4 特徴量) となった。

B.2.4 OpenBMI

基本的な内容は (2) 項で記述済みである。健常者 54 名を対象とし、2 種類の運動想起課題 (右手と左手) を 100 回ずつを 2 つの実験ブロックにわたって、計測した EEG データを収録している。なお、データ公開元の報告に基づき [200]、本付録では運動想起課題との相性が良かった 21 名の被験者を選定した。

前処理および特徴抽出は、以下の手順で実施した。運動野上に位置する 20 個の電極 (FC5/3/1/2/4/6, C5/3/1/z/2/4/6, CP5/3/1/z/2/4/6) からの EEG 信号を選択した。元のサンプリング周波数 1000 Hz を 100 Hz にダウンサンプリングし、5 次バターワースフィルタを用いて 8~30 Hz のバンドパスフィルタリングを施した。視覚オンセット後、1 秒から 3.5 秒までの 2.5 秒間の区間を抽出した。識別には、FBCSP [201] を用いた。フィルタバンクは、6 つの周波数帯域 (8–12, 12–16, ..., 24–28, および 28–30 Hz) を用いて構築された。各周波数帯域について、対応する空間フィルタ行列の上位 2 つおよび下位 2 つのフィルタに射影された信号から、対数分散特徴量を抽出した。その結果、特徴ベクトル N_d の次元数は 24 (6 帯域 $\times$ 4 フィルタ) となった。

B.2.5 BCI Competition IV-2a (BCICIV2a)

BCICIV2a は、9 名の健常者から収集された公開 EEG データセットである。各被験者は、左手、右手、足、および舌の運動想起という 4 つの運動想起課題を行った。データセットには被験者ごとに 2 つのセッションが含まれ、各セッションは 288 試行 (各タスク 72 試行) で構成されている。各試行は固視点と警告音の提示で開始し、その 2 秒後に運動想起タスクを指示する視覚キューが提示された。視覚キューは 1.25 秒間表示され、被験者は試行開始から 6 秒後に固視点が消滅するまで、対応する動作を想起するよう指示された。

データセット提供者による前処理手順の指定がなかったため、Ang ら [218] が記述した前処理手順を採用した。信号に対し、5 次バターワースフィルタを用いて 4~40 Hz のバンドパスフィルタリングを施した。次に、視覚キューのオンセット後 0.5 秒から 2.5 秒までの 2.0 秒間の区間を抽出した。識別には、9 つの周波数帯域 (4–8, 8–12, ..., および 36–40 Hz) を用いて FBCSP を抽出した。FBCSP は本来 2 クラス分類のために開発されたものであるため、多クラス分類問題に対処するために One-Versus-Rest (OVR) 戦略を採用した [218]。各周波数帯域について、対応する空間フィルタ行列の上位 2 つおよび下位 2 つのフィルタ

に射影された信号から，対数分散特徴量を抽出した．その結果，特徴ベクトル N_d の次元数は 144 (9 帯域 $\times$ 4 フィルタ $\times$ 4 OVR) となった．

B.2.6 データの分割

ある被験者がソース被験者として選択された場合，その全データをバックボーンとなる機械学習モデルの事前学習に使用した．一方，ある被験者がターゲット被験者として選択された場合，そのデータはキャリブレーションデータ，検証データ，およびテストデータに分割した．キャリブレーションおよび検証データは転移学習に用い，テストデータは転移学習の性能評価用として確保した．データ分割の方法は，以下に述べるようにデータセットごとに異なった．

- **MyoDataset:** 第 1 試行をキャリブレーション，第 2 試行を検証，残りの試行をテストに使用した．
- **NinaProDB5:** 第 1, 3, 6 試行をキャリブレーション，第 4 試行を検証，第 2, 5 試行をテストに使用した．
- **OpenBMI:** 第 1 ブロックを 80:20 の比率でランダムにキャリブレーションおよび検証セットに分割し，第 2 ブロックをすべてテストに使用した．
- **BCICIV2a:** 第 1 セッションを 80:20 の比率でランダムにキャリブレーションおよび検証セットに分割し，第 2 セッションをすべてテストに使用した．

B.3 手法

B.3.1 バックボーン識別器

バックボーンとなる機械学習モデルとして，ロジスティック回帰 (Logistic Regression: LR)，サポートベクターマシン (Support Vector Machine: SVM)，ランダムフォレスト (Random Forest: RF)，および多層パーセプトロン (Multilayer Perceptron: MLP) の 4 つの識別器を採用した．LR, SVM, および RF モデルの実装には scikit-learn を用い，MLP の実装には PyTorch を用いた．各識別器のハイパーパラメータは，ハイパーパラメータ自動最適化フレームワークである Optuna を用いて最適化し，モデルごとに 200 試行を実施した．ハイパーパラメータの探索範囲を表 B.1 にまとめる．なお，表に明記されていないハイパーパラメータについては，各ライブラリのデフォルト値を設定した．ハイパーパラメータの調整は，各ソース被験者のデータを用いた 3 分割交差検証によって行った．

表 B.1: 機械学習モデル (LR, SVM, RF, MLP) のハイパーパラメータの探索範囲. 文献 [16] より転載.

Classifier	Hyperparameter	Search Range
LR	C	Log-uniform distribution: $[10^{-4}, 10^4]$
	Penalty	{L1, L2}
	Solver	{liblinear (for L1), lbfgs (for L2)}
SVM	C	Log-uniform distribution: $[10^{-4}, 10^4]$
	Kernel	{linear, poly, RBF}
	Gamma	{scale, auto} (only for poly and RBF kernels)
RF	n_estimators	Integer values: [2, 300]
	max_depth	Integer values: [2, 100]
	min_samples_split	Integer values: [2, 20]
	min_samples_leaf	Integer values: [1, 20]
MLP	Learning rate	Log-uniform distribution: $[10^{-5}, 10^{-1}]$
	Weight decay	Log-uniform distribution: $[10^{-5}, 10^{-3}]$
	Number of layers	Integer values: [1, 5]
	Neurons per layer	Integer values: [2, 256], (each layer $\leq$ previous layer)

表 B.2: 転移学習 (STM, k-STM) のハイパーパラメータの探索範囲. 文献 [16] より転載.

Method	Hyperparameter	Search Range
STM	$\tilde{\beta}$	Uniform distribution: [0, 10]
	$\tilde{\gamma}$	Uniform distribution: [0, 10]
RBF k-STM	σ	Log-uniform distribution: $[10^{-5}, 10^5]$
	λ	Log-uniform distribution: $[10^{-5}, 10^5]$
Linear k-STM	λ	Log-uniform distribution: $[10^{-5}, 10^5]$
Hybrid k-STM	σ	Log-uniform distribution: $[10^{-5}, 10^5]$
	λ	Log-uniform distribution: $[10^{-5}, 10^5]$
	α	Uniform distribution: [0, 1]

B.3.2 Style Transfer Mapping

STM 手法の目的は、ターゲットユーザのデータ分布をソースユーザのそれに適合させることである。写像関数を学習するために、まず代表的なソースデータ点を、各キャリブレーションサンプルの写像先として設定した。具体的には、ソースデータ内の各クラスに対して個別に k -means クラスタリングを適用し、得られたクラスタ重心をクラス固有のプロトタイプとして用いた。次に、ターゲットユーザの各キャリブレーションサンプルを、同一クラス内に存在する最も近いプロトタイプに割り当てることで、対応する写像先を決定した。形式的に、 i 番目のキャリブレーションサンプルを $\mathbf{x}_{\text{cal}_i} \in \mathbb{R}^{N_d}$ 、それに対応する写像先プロトタイプを $\mathbf{d}_i \in \mathbb{R}^{N_d}$ と表記する。先行研究では一般に 15 個のクラスタが採用されてきたが [140, 141, 212, 216], 本付録ではクラスタ数を調整可能なハイパーパラメータとして扱い、 $\{1, 5, 15, 30\}$ の範囲で探索を行った。

次に、写像された各キャリブレーションサンプルとそれに対応するプロトタイプとの間の二乗誤差を最小化することにより、変換行列 $\mathbf{A} \in \mathbb{R}^{N_d \times N_d}$ およびバイアスペクトル $\mathbf{b} \in \mathbb{R}^{N_d}$ を推定した。この最適化問題は以下のように定式化される。

$$\underset{\mathbf{A}, \mathbf{b}}{\operatorname{argmin}} \sum_{i=1}^{N_{\text{cal}}} f_i \|\mathbf{A} \mathbf{x}_{\text{cal}_i} + \mathbf{b} - \mathbf{d}_i\|_2^2 + \beta \|\mathbf{A} - \mathbf{I}\|_F^2 + \gamma \|\mathbf{b}\|_2^2, \quad (\text{B.1})$$

ここで、 f_i は $[0, 1]$ の範囲をとる確信度、 N_{cal} はキャリブレーションサンプル数、 $\|\cdot\|_2^2$ および $\|\cdot\|_F^2$ はそれぞれ L_2 ノルムおよびフロベニウスノルム、 $\mathbf{I} \in \mathbb{R}^{N_d \times N_d}$ は単位行列を表す。確信度 f_i は半教師あり学習において一般的に用いられるパラメータであるが、本付録ではすべてのキャリブレーションデータに対して $f_i = 1$ を設定した。

正則化ハイパーパラメータ β および γ は、変換の程度を制御するものであり、以下のよう定義される。

$$\beta = \tilde{\beta} \frac{1}{N_d} \operatorname{tr} \left(\sum_{i=1}^{N_{\text{cal}}} \mathbf{x}_{\text{cal}_i} \mathbf{x}_{\text{cal}_i}^T \right), \quad \gamma = \tilde{\gamma} N_{\text{cal}}, \quad (\text{B.2})$$

ここで、 $\operatorname{tr}(\cdot)$ は行列のトレースを表す。先行研究では、 $\tilde{\beta}$ および $\tilde{\gamma}$ の探索範囲を $[0, 3]$ とするのが一般的であったが [140, 141, 214, 213, 212], より広範な探索範囲が性能向上に寄与する可能性があるため [215], 本付録ではこれを $[0, 10]$ まで拡張した（表 B.2 を参照）。ハイパーパラメータ $\tilde{\beta}$ および $\tilde{\gamma}$ は、検証データの識別精度を最大化するように、Optuna を用いて 200 試行の探索を行い最適化した。最終的に、 $\mathbf{A}$ および $\mathbf{b}$ は以下の閉形式解を用いて算出された。

$$\mathbf{A} = \mathbf{Q} \mathbf{P}^{-1}, \quad \mathbf{b} = \frac{1}{\tilde{f}} (\hat{\mathbf{d}} - \mathbf{A} \hat{\mathbf{x}}), \quad (\text{B.3})$$

ただし、

$$\mathbf{Q} = \sum_{i=1}^{N_{\text{cal}}} \mathbf{d}_i \mathbf{x}_{\text{cal}_i}^T - \frac{1}{\tilde{f}} \hat{\mathbf{d}} \hat{\mathbf{x}}^T + \beta \mathbf{I}, \quad (\text{B.4})$$

$$\mathbf{P} = \sum_{i=1}^{N_{\text{cal}}} \mathbf{x}_{\text{cal}_i} \mathbf{x}_{\text{cal}_i}^T - \frac{1}{\tilde{f}} \hat{\mathbf{x}} \hat{\mathbf{x}}^T + \beta \mathbf{I}, \quad (\text{B.5})$$

$$\hat{\mathbf{x}} = \sum_{i=1}^{N_{\text{cal}}} \mathbf{x}_{\text{cal}_i}, \quad \hat{\mathbf{d}} = \sum_{i=1}^{N_{\text{cal}}} \mathbf{d}_i, \quad \hat{f} = N_{\text{cal}} + \gamma. \quad (\text{B.6})$$

最適な写像パラメータ $\mathbf{A}$ および $\mathbf{b}$ が得られた後、新たな未知データ $\mathbf{x}_{\text{new}} \in \mathbb{R}^{N_d}$ は、以下のアフィン変換によってソース分布上へと写像される。

$$\mathbf{x}_{\text{trans}} = \mathbf{A}\mathbf{x}_{\text{new}} + \mathbf{b}. \quad (\text{B.7})$$

変換されたデータ $\mathbf{x}_{\text{trans}}$ は、ソースユーザの事前学習済み識別器を用いて識別される。

B.3.3 Kernel Style Transfer Mapping

提案手法である k-STM は、従来の STM における線形写像を非線形写像へと拡張するものであり、その基本的な処理手順はオリジナルの STM 手法に準拠している。k-STM の実行に先立ち、従来の STM と同様に、キャリブレーションサンプルとそれに対応するプロトタイプのパラを作成した。ここで、キャリブレーションサンプル行列を $\mathbf{X} \in \mathbb{R}^{N_{\text{cal}} \times N_d}$ 、対応するプロトタイプから構成される写像先行列を $\mathbf{D} \in \mathbb{R}^{N_{\text{cal}} \times N_d}$ と定義する。

k-STM における非線形写像は、データを高次元特徴空間へ陰に射影するカーネル関数を採用することで実現される。具体的には、カーネル行列 $\mathbf{K} \in \mathbb{R}^{N_{\text{cal}} \times N_{\text{cal}}}$ を構築する。ここで、各要素は $K_{i,j} = k(\mathbf{x}_{\text{cal}_i}, \mathbf{x}_{\text{cal}_j})$ として計算され、 $k(\cdot)$ は選択されたカーネル関数を表す。本付録では、RBF、線形、およびハイブリッドの 3 種類のカーネルを採用した。RBF カーネルは以下のように定義される。

$$k_{\text{R}}(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_2^2}{2\sigma^2}\right), \quad (\text{B.8})$$

ここで、 σ はカーネル帯域幅を表す。この RBF カーネルは最も代表的な非線形カーネルの一つであり、データ点同士が近接している場合にのみ強い影響を及ぼす。この特性により、ターゲット被験者固有の非線形特性を鋭敏に捉え、対応するプロトタイプへの効率的な写像の基盤となることが期待される。線形カーネルは、標準的な内積によって以下のように与えられる。

$$k_{\text{L}}(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^\top \mathbf{x}_j. \quad (\text{B.9})$$

線形カーネルは、従来の線形 STM と非線形 k-STM との公平な比較を行うためのベースラインとして選択された。これらのカーネル表現を用い、k-STM は以下の最適化問題を解くことで重み行列 $\mathbf{W} \in \mathbb{R}^{N_{\text{cal}} \times N_d}$ を学習する。

$$\underset{\mathbf{W}}{\text{argmin}} \|\mathbf{D} - \mathbf{KW}\|_F^2 + \lambda \|\mathbf{W}\|_F^2, \quad (\text{B.10})$$

ここで、 λ は適合精度とモデルの複雑さの間のトレードオフを制御する正則化ハイパーパラメータである。この問題に対する閉形式解は以下のように与えられる。

$$\mathbf{W} = (\mathbf{K}^\top \mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{K}^\top \mathbf{D}. \quad (\text{B.11})$$

最適な $\mathbf{W}$ が得られることで、k-STM はキャリブレーションデータとそれに対応するプロトタイプ間の複雑な非線形関係を効果的に捉えることが可能となる。

新たな未知データ点 $\mathbf{x}_{\text{new}}$ をソースユーザの分布へと写像するために、まずカーネルベクトル $\mathbf{k}_{\text{new}} \in \mathbb{R}^{N_{\text{cal}}}$ を計算する。ここで、 i 番目の要素は以下のように与えられる。

$$k_{\text{new},i} = k(\mathbf{x}_{\text{new}}, \mathbf{x}_{\text{cal}_i}). \quad (\text{B.12})$$

続いて、学習された重み行列 $\mathbf{W}$ を適用することで、変換後のデータが得られる。

$$\mathbf{x}_{\text{trans}} = \mathbf{k}_{\text{new}}^{\top} \mathbf{W}. \quad (\text{B.13})$$

最後に、変換されたデータ $\mathbf{x}_{\text{trans}}$ は、ソースユーザの事前学習済み識別器を用いて識別される。

加えて、RBF カーネルと線形カーネルの加重和として、ハイブリッドカーネルを導入する。

$$k_{\text{H}}(\mathbf{x}_i, \mathbf{x}_j) = \alpha \cdot k_{\text{R}}(\mathbf{x}_i, \mathbf{x}_j) + (1 - \alpha) \cdot k_{\text{L}}(\mathbf{x}_i, \mathbf{x}_j), \quad (\text{B.14})$$

ここで、 $\alpha \in [0, 1]$ は RBF 成分と線形成分の均衡を制御するパラメータである。このハイブリッドな定式化により、データセットの特性に応じて、線形写像と非線形写像の間で柔軟に適応することが可能となる。ハイパーパラメータ σ , λ , および α は、Optuna を用いて 200 試行にわたり最適化した。各ハイパーパラメータの探索範囲を表 B.2 にまとめる。本稿では以下、 $\alpha = 1$ の場合を RBF k-STM, $\alpha = 0$ の場合を線形 k-STM, それ以外の場合をハイブリッド k-STM と呼称する。

STM および k-STM の双方において、計算上の主要なボトルネックとなるのは、学習段階で実行される行列の逆行列演算である。両手法の計算量は共に $O(n^3)$ であるが、 n の定義は両者で異なる。STM において n は特徴量の次元数 N_d に相当するのに対し、k-STM において n はキャリブレーションサンプル数 N_{cal} に相当する。転移学習の文脈においては、ユーザの負担を最小限に抑えるため、利用可能なターゲットデータはごく少量であると仮定されるのが一般的である。したがって、k-STM の $O(N_{\text{cal}}^3)$ という計算量は実用上極めて効率的であり、迅速な被験者間適応に適している。

B.3.4 アンサンブル学習

STM および k-STM は、各ソース被験者に対して個別に実行されるため、その出力を統合する処理が必要となる。ターゲットユーザの写像後のデータは、各ソースユーザの事前学習済み識別器を用いて識別される。これら複数の識別器による予測結果を集約するため、本付録では予測確率の平均値に基づく多数決を採用した。この統合戦略は、第 4 章での SV であり、単純ではあるが、個人差を伴う転移学習タスクにおいて有効であることが示されている。

B.4 評価

各データセットにおいて、1名の被験者をターゲット被験者とし、ソース被験者数を1名から最大10名まで段階的に増加させながら識別精度を評価した。例えば MyoDataset の場合、ある1名がターゲット被験者として選択されると、残る24名がソース被験者の候補となる。設定されたソース被験者数（例：3名）に対し、候補群からランダムに被験者を選択し、STM および k-STM の性能評価を行った。ここで、全ての可能な組み合わせを評価することは計算コストの観点から困難であるため、各ソース被験者数において最大30通りの組み合わせをランダムに抽出し、その平均識別精度を算出した。この一連の手順を、すべての被験者がターゲット被験者となるまで繰り返した。

提案手法の有効性を検証するため、各データセットに対して以下の統計解析を実施した。まず、(1) 転移学習手法 (STM, RBF k-STM, 線形 k-STM, およびハイブリッド k-STM), (2) ソース被験者数 (1名から10名), および (3) バックボーン識別器 (LR, SVM, RF, および MLP) を要因とする三元配置反復測定分散分析 (ANOVA) を行った。本付録の主目的は提案手法の有効性を評価することにあるため、転移学習手法の要因において有意な主効果が認められた場合、多重比較を実施した。多重比較には、ボンフェローニ補正を伴う対応有りの t 検定を用いた。なお、すべての検定において有意水準は5%に設定した。

B.5 結果

図 B.3 に、全データセットにおける各手法の識別精度を示す。バックボーン識別器やデータセットの種類に依存せず、STM および k-STM は共に、ソース被験者数の増加に伴い識別精度が緩やかに向上する傾向を示した。すべてのデータセットにおいて、RBF k-STM およびハイブリッド k-STM は、概して従来の STM 手法よりも優れた性能を示した。各データセットにおける最高精度の構成は以下の通りである。MyoDataset では MLP とハイブリッド k-STM (ソース10名) を用いて精度 0.924, NinaProDB5 では LR とハイブリッド k-STM (ソース9名) を用いて精度 0.711, OpenBMI では SVM と RBF k-STM (ソース7名) を用いて精度 0.865, そして BCICIV2a では LR と線形 k-STM (ソース7名) を用いて精度 0.670 を達成した。

各データセットに対し、三元配置反復測定分散分析を実施した結果を表 B.3 に示す。同表の通り、すべてのデータセットにおいて、全主効果が統計的に有意であった。さらに、2次の交互作用についてもすべてのデータセットで有意であり、3次の交互作用については OpenBMI を除くすべてのデータセットで有意差が認められた。

すべてのデータセットにおいて転移学習の有意な主効果が認められたため、ボンフェローニ補正を伴う対応有りの t 検定を実施した。その結果を表 B.4 にまとめる。概して、RBF k-STM およびハイブリッド k-STM は、すべてのデータセットにおいて STM を有意に上回る性能を示した。MyoDataset および NinaProDB5 においては、ハイブリッド k-STM が最も高い性能を達成し、他のすべての手法を有意に上回った。OpenBMI においては、すべ

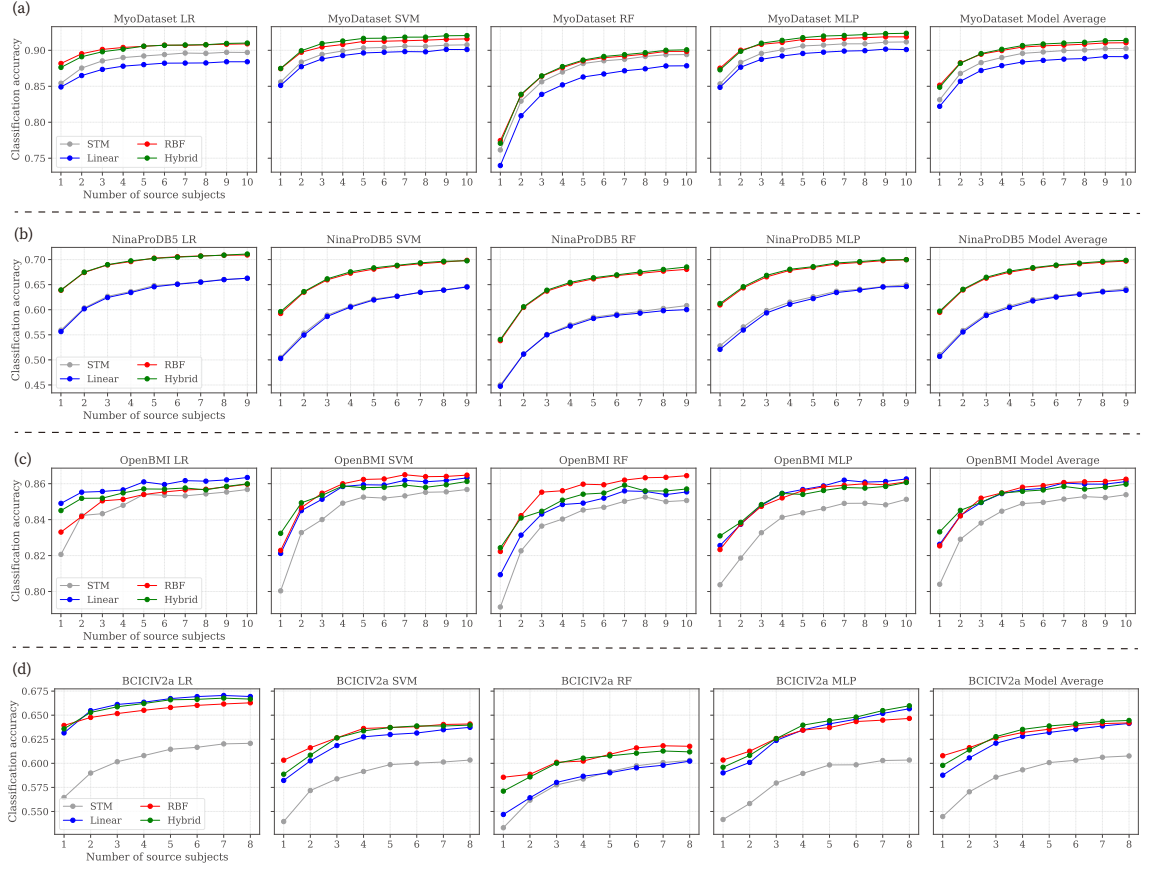


図 B.3: 4 つのデータセットにおける STM および k-STM の識別精度（ソース被験者数はデータセットに応じ最大 10 名まで変化）． 各行はデータセット，各列はバックボーン識別器に対応する． なお，第 5 列は全識別器における平均精度を表す． 文献 [16] より転載．

ての k-STM が STM を有意に上回ったものの，カーネル間には有意差が認められなかった． BCICIV2a においては，RBF k-STM およびハイブリッド k-STM が他の手法を上回ったが，両者の間に有意差は認められなかった． 以上の結果を総合すると，k-STM は，特に非線形カーネル（RBF やハイブリッドなど）を用いた場合において，従来の STM と比較して識別精度を有意に向上させることが統計的に裏付けられたと言える．

B.5.1 追加分析

本付録では，STM および k-STM における最適な k -means クラスタ（プロトタイプ）数について調査を行った． これは，先行研究 [140, 141, 212, 216] において，一般に最適化を行うことなくこの値を 15 に固定していたためである． プロトタイプ数は写像の質に影響を及ぼし，ひいては識別精度を左右する可能性があるため，性能を最大化するためにはその最適値を決定することが極めて重要である． そこで本付録では，1, 5, 15, および 30 クラスタという 4 つの設定を検証し，各ターゲットおよびソースのペアに対する検証精度に基づいて最適値を選択した． 検証データセットにおいて，クラスタ数を変化させた際の識別精度の

表 B.3: 各データセットにおける三元配置反復測定分散分析の結果. 要因は, *transfer* (転移学習手法: STM, RBF k-STM, 線形 k-STM, ハイブリッド k-STM), *n_src* (ソース被験者数), および *model* (バックボーン識別器: LR, SVM, RF, MLP) である. 交互作用項 (例: transfer:n_src) は, 要因間の複合的な効果を表す. 文献 [16] より転載.

Dataset	Effect	F-value	Num DF	Den DF	p-value
MyoDataset	transfer	23.976	3	72	< 0.001
	n_src	357.459	9	216	< 0.001
	model	59.533	3	72	< 0.001
	transfer:n_src	16.432	27	648	< 0.001
	transfer:model	3.828	9	216	0.002
	n_src:model	185.284	27	648	< 0.001
	transfer:n_src:model	2.426	81	1944	< 0.001
NinaProDB5	transfer	265.041	3	27	< 0.001
	n_src	492.284	8	72	< 0.001
	model	64.818	3	27	< 0.001
	transfer:n_src	26.754	24	216	< 0.001
	transfer:model	4.737	9	81	< 0.001
	n_src:model	44.192	24	216	< 0.001
	transfer:n_src:model	2.952	72	648	< 0.001
OpenBMI	transfer	4.980	3	60	0.004
	n_src	104.122	9	180	< 0.001
	model	2.868	3	60	0.044
	transfer:n_src	6.726	27	540	< 0.001
	transfer:model	3.711	9	180	< 0.001
	n_src:model	10.495	27	540	< 0.001
	transfer:n_src:model	1.186	81	1620	0.128
BCICIV2a	transfer	30.758	3	24	< 0.001
	n_src	61.568	7	56	< 0.001
	model	19.214	3	24	< 0.001
	transfer:n_src	11.220	21	168	< 0.001
	transfer:model	4.888	9	72	< 0.001
	n_src:model	4.337	21	168	< 0.001
	transfer:n_src:model	1.810	63	504	< 0.001

表 B.4: ボンフェローニ補正を伴う対応有りの t 検定の結果. 文献 [16] より転載.

Dataset	Comparison	t-value	modified p-value
MyoDataset	STM vs RBF	-14.865	< 0.001
	STM vs Linear	20.841	< 0.001
	STM vs Hybrid	-21.644	< 0.001
	RBF vs Linear	32.103	< 0.001
	RBF vs Hybrid	-3.427	0.004
	Linear vs Hybrid	-48.877	< 0.001
NinaProDB5	STM vs RBF	-55.706	< 0.001
	STM vs Linear	5.103	< 0.001
	STM vs Hybrid	-56.693	< 0.001
	RBF vs Linear	57.170	< 0.001
	RBF vs Hybrid	-4.262	< 0.001
	Linear vs Hybrid	-55.507	< 0.001
OpenBMI	STM vs RBF	-11.832	< 0.001
	STM vs Linear	-15.136	< 0.001
	STM vs Hybrid	-14.854	< 0.001
	RBF vs Linear	1.133	1.000
	RBF vs Hybrid	1.254	1.000
	Linear vs Hybrid	-0.129	1.000
BCICIV2a	STM vs RBF	-21.358	< 0.001
	STM vs Linear	-19.166	< 0.001
	STM vs Hybrid	-26.032	< 0.001
	RBF vs Linear	3.939	0.062
	RBF vs Hybrid	-0.190	1.000
	Linear vs Hybrid	-8.639	< 0.001

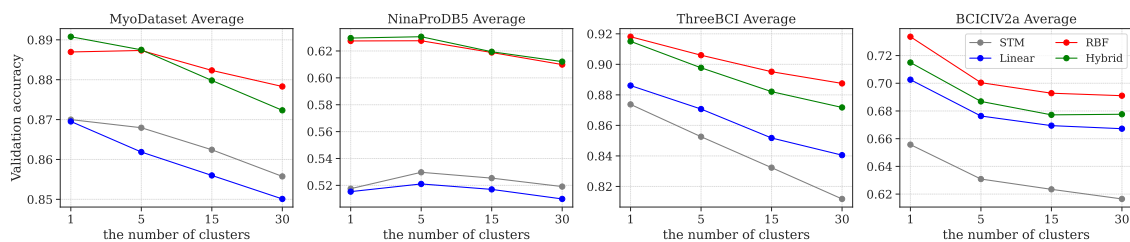


図 B.4: 各データセットにおける検証精度に対するプロトタイプ数 (k -means クラスタ数) の影響. 各パネルは, 4 つのバックボーン識別器における平均結果を示す. 文献 [16] より転載.

推移を図 B.4 に示す. 全体的な傾向として, STM および k -STM の双方において, クラスタ数が少ないほど一貫して高い精度が得られ, 単一のプロトタイプを用いた場合に最も高い性能が確認された. これらの結果は, 従来推奨されていた数よりも少ないクラスタ数を用いることで, 識別精度が向上する可能性があることを示唆している.

B.6 考察

統計解析の結果, 非線形写像を導入した転移学習手法は, sEMG および EEG 信号の識別において有意に有効であることが示された. 本付録では, 生体信号が非線形な個人差を有するという仮説に基づき, 従来の STM の枠組みを k -STM と呼称する非線形手法へと拡張した. 提案手法である k -STM は, RBF カーネルやハイブリッドカーネルといった非線形カーネルを導入することで, すべてのデータセットにおいて標準的な STM を一貫して上回る性能を示した. さらに, 非線形カーネル (RBF およびハイブリッド) に加え, 線形カーネルについても評価を行った. その結果, OpenBMI データセットを除くすべてのデータセットにおいて, 非線形カーネルが高い識別精度を達成した. これらの結果は, 観測された性能向上が単にカーネル法を用いたことによるものではなく, 非線形カーネルを導入したことに特異的に起因することを示唆している. したがって, 本付録における STM の非線形拡張は, 個人差を効果的に低減し, 識別性能を向上させるものであると結論付けられる.

以上の結果は, 線形写像と非線形写像の均衡を保つハイブリッドカーネルが, k -STM における有力な選択肢であるべきことを示している. 概して, 統計解析の結果は, ハイブリッドカーネルが最も高い識別精度を達成したことを示している. しかしながら, より詳細な検討を行ったところ, OpenBMI および BCICIV2a データセットにおける LR 識別器を用いた場合など, 特定の条件下では線形カーネルの方が優れた性能を示すことが明らかになった. これらの知見は, 最適なカーネルの選択がデータセットおよび識別器双方の特性に依存すること, また, 純粋な線形写像あるいは非線形写像のみを仮定することは過度な制約となり得ることを示唆している. ハイブリッドカーネルは, 線形成分と非線形成分を適応的に均衡させることでこの課題に対処し, それにより, より柔軟かつ現実的な解を提供するものである. 特筆すべき点として, 識別性能を向上させるために複数のカーネルを組み合わせるこ

との有効性は、生体信号識別の領域を超えて先行研究でも実証されており [219]、今後の研究における有望な方向性であることが示唆される。

各クラスにつき 15 個のプロトタイプを用いるという先行研究の推奨 [140, 141, 212, 216] とは対照的に、本付録では、より少ないプロトタイプ、特に単一のプロトタイプを用いた場合に、識別精度が向上することを見出した。複数の候補の中から最近傍のプロトタイプに各ターゲットサンプルを割り当てることは、より精緻な写像をもたらすという点で直感的には妥当に思えるかもしれないが、本付録の結果はそれとは逆の傾向を示している（図 B.4 を参照）。この乖離は、限られたキャリブレーションデータを用いて複数のプロトタイプを使用した場合、過度に詳細な写像が行われることによって生じる過学習のリスクに起因している可能性がある。さらに、多数のプロトタイプを使用することは、最適な写像先の決定に曖昧さを生じさせ、その結果、安定性および全体的な精度を低下させる要因となり得る。対照的に、より少ないプロトタイプを用いることで、ターゲットデータをより汎化された代表点へと写像することが可能となり、個人差の安定した補正と識別性能の向上が実現される。したがって、STM および k-STM の実用的な実装においては、より少ないプロトタイプを採用することが推奨される。本付録の知見に基づき、基本的な指針として、まず $k = 1$ を設定することを推奨する。この設定は、キャリブレーションデータが限られている場合において、過学習のリスクを効果的に低減し、より安定した写像先を提供するものである。

k-STM の主要な利点の一つは、様々な識別器との統合を容易にするその高い汎用性にある。ノーフリーランチ定理 [220] で示唆されているように、あらゆる問題設定において一貫して最適な性能を発揮できる単一の識別器は存在しない。したがって、特定の識別器に依存する転移学習手法 [138, 139, 8, 135] は、ある種のタスクにおいては良好な性能を示すかもしれないが、異なる条件下では精度が著しく低下する可能性がある。対照的に、k-STM は識別器に依存しない転移学習アプローチであり、線形モデルから非線形モデル、さらにはパラメトリック手法からノンパラメトリック手法に至るまで、広範な識別アルゴリズムとの柔軟な統合が可能である。この汎用性は本付録において実証的に検証されており、k-STM は LR, SVM, RF, および MLP といった根本的に異なるアーキテクチャを有する識別器のいずれに対しても、一貫して高い性能を発揮した。

さらに、k-STM はこれまでの STM 研究において蓄積された知見を直接的に活用可能である。例えば、Niu らおよび Kanoga らは、他の手法と組み合わせた際の前処理または後処理として STM が有効であることを実証しており [213, 212]、k-STM においても同様の統合が可能である。Huang らは、複数のソースドメインを単一の代表ソースへと統合するために STM を採用したが [216]、このプロセスに k-STM を導入することで、統合ソースの質をさらに向上させ得る。Ran らは、STM に適した最適なソースサンプルの選択手法を提案しているが [214]、このアプローチは有害なサンプルを除去することにより、k-STM における負の転移を抑制する上でも有用である。Chen らは、写像前後におけるターゲットサンプル間の関係性を保存するための新たな正則化技術を導入したが [215]、これは k-STM にとりわけ有益である。図 B.1 に示すように、k-STM はその強力な非線形写像能力ゆえに、写像されたデータをプロトタイプ周辺に過度に凝集させる傾向がある。Chen らの正則化を適用することで、写像後のデータ分布をより効果的に制御することが可能となり、それによ

り過学習のリスクを低減できると考えられる。総じて、k-STM はこれらの先行研究の成果を継承しつつ拡張するものであり、ロバストかつ汎用的な手法としての潜在能力を裏付けるものである。

k-STM の特性は、そのカーネルベースの定式化によって本質的に規定されており、これには明確な利点と不可避な制約の双方が伴う。生体信号の転移学習における典型的な文脈では、ターゲット被験者から収集されるデータ量は限られているのが一般的である。こうした状況下において、k-STM は学習および推論の双方の段階で極めて効率的である。この高速な処理は、速度と計算効率が優先される実世界への応用において重要な利点となる。しかしながら、ターゲット被験者のデータ量が極めて膨大になった場合、提案手法を適用することは計算上重大な困難を伴うこととなる。そのような場合には、カーネル法に依存しない他の非線形拡張が必要となる。例えば、非線形写像能力を維持しつつより大規模なデータに対処するためには、ニューラルネットワークベースの STM の開発が有効な将来の方向性となり得る。このようなスケーラブルなモデルの開発は、今後の研究課題である。

B.7 結言

本付録では、被験者間の生体信号識別に特化した、発展的な非線形転移学習手法である k-STM を提案する。線形写像に依存する従来の STM は、EMG や EEG といった生体信号において一般的に観察される、複雑かつ非線形な個人差を捉えることが本質的に困難であった。RBF カーネルやハイブリッドカーネルを含む非線形カーネルを導入することで、k-STM は非線形関係を効果的にモデル化し、識別精度の有意な向上をもたらした。4 つのバックボーン識別器 (LR, SVM, RF, および MLP) を用いた 4 つの公開データセット (MyoDataset, NinaProDB5, OpenBMI, および BCICIV2a) における評価により、従来の STM に対する k-STM の優位性が確認され、特に非線形カーネルの有効性が顕著に示された。k-STM の識別器に依存しない設計は、広範な機械学習モデルや既存の STM ベースの手法とのシームレスな統合を可能にし、実世界への応用における実用性と適応性を際立たせている。総じて、k-STM はロバストかつ汎用的な転移学習フレームワークであり、生体信号に基づく実用的なヒューマンコンピュータインタフェースの実現に向けた重要な進歩となるものである。

研究業績

主論文に関する論文

査読付き国際原著論文

- [1] **T. Hoshino**, S. Kanoga, A. Aoyama, “Kernel-Based Style Transfer Mapping for Cross-Subject Biological Signal Classification,” IEEE Access, vol. 14, pp. 4496-4507, 2025. doi:10.1109/ACCESS.2025.3649556
- [2] **T. Hoshino**, S. Kanoga, M. Tsubaki, A. Aoyama, “Comparison of Fine-Tuned Single-Source and Multi-Source Approaches to Surface Electromyogram Pattern Recognition,” Biomedical Signal Processing and Control, vol. 94, Art. no. 106261, 2024. doi:10.1016/j.bspc.2024.106261
- [3] **T. Hoshino**, S. Kanoga, M. Tsubaki, A. Aoyama, “Comparing subject-to-subject transfer learning methods in surface electromyogram-based motion recognition with shallow and deep classifiers,” Neurocomputing, vol. 489, pp.599-612, 2022. doi:10.1016/j.neucom.2021.12.081

プレプリント

- [1] **T. Hoshino**, S. Kanoga, T. Takase, A. Aoyama, “Virtual Source Assistance: Improving Transfer Learning Performance for Biological Signal Classification with Data from Few Source Subjects,” Available at SSRN, 2025. doi:10.2139/ssrn.5247869

その他発表済み成果

査読付き国際原著論文

- [1] M. Kubota, S. Bun, K. Takahata, S. Kurose, Y. Momota, Y. Iwabuchi, T. Tezuka, H. Tabuchi, M. Seki, Y. Yamamoto, R. Shikimoto, Y. Mimura, **T. Hoshino**, S. Shimohama, N. Suzuki, A. Morimoto, A. Oosumi, Y. Hoshino, K. Tai, H. Aoyagi, Y. Sato, J. Kuromitsu, J. Nakahara, M. Mimura, D. Ito “Plasma Biomarkers for Early Detection of Alzheimer ’ s Disease: A Cross-sectional Study in a Japanese

- Cohort ,” *Alzheimer ’ s Research & Therapy*, vol. 17, Art. no. 131, 2025. doi: 10.1186/s13195-025-01778-8
- [2] Y. Okawa, S. Kanoga, **T. Hoshino**, S. Ishikawa, “Evaluating the Impact of Replay-Based Continual Learning on Long-Term sEMG Pattern Recognition in Instance-Incremental Learning,” *IEEE Access*, vol. 12, pp. 182469-182482, 2024. doi: 10.1109/ACCESS.2024.3508799
- [3] Y. Date, S. Bun, K. Takahata, M. Kubota, Y. Momota, Y. Iwabuchi, T. Tezuka, H. Tabuchi, M. Seki, Y. Yamamoto, R. Shikimoto, Y. Mimura, **T. Hoshino**, S. Kurose, S. Shimohama, N. Suzuki, A. Morimoto, A. Oosumi, Y. Hoshino, M. Jinzaki, M. Mimura, D. Ito, “Can the clinical sign “head-turning sign” and simple questions in “Neucop-Q” predict amyloid β pathology?,” *Alzheimer ’ s Research & Therapy*, vol. 16, Art. no. 250, 2024. doi:10.1186/s13195-024-01605-6
- [4] Y. Okawa, S. Kanoga, **T. Hoshino**, S. Ishikawa, “Effect of regularization-based continual learning in instance-incremental learning scenario on long-term surface electromyogram-based pattern recognition,” *Advanced Biomedical Engineering*, vol. 13, pp. 363-373, 2024. doi:10.14326/abe.13.363
- [5] S. Bun, D. Ito, T. Tezuka, M. Kubota, R. Ueda, K. Takahata, S. Moriguchi, S. Kurose, Y. Momota, N. Suzuki, A. Morimoto, Y. Hoshino, N. Seki, Y. Mimura, R. Shikimoto, Y. Yamamoto, **T. Hoshino**, Y. Sato, H. Tabuchi, M. Mimura, “Performance of plasma A β 42/40, measured using a fully automated immunoassay, across a broad patient population in identifying amyloid status,” *Alzheimer ’ s Research & Therapy*, vol.15, Art. no. 149, 2023. doi:10.1186/s13195-023-01296-5
- [6] S. Kanoga, **T. Hoshino**, S. Kamei, T. Kobayashi, T. Ohmori, M. Uchiyama, M. Tada, “Comparison of seven shallow and deep regressors in continuous blood pressure and heart rate estimation using single-channel photoplethysmograms under three evaluation cases,” *Biomedical Signal Processing and Control*, vol.85, Art. no. 105029, 2023. doi:10.1016/j.bspc.2023.105029
- [7] S. Kanoga, **T. Hoshino**, H. Asoh, “Subject-transfer framework with unlabeled data based on multiple distance measures for surface electromyogram pattern recognition,” *Biomedical Signal Processing and Control*, vol. 74, Art. no. 103522, 2022. doi:10.1016/j.bspc.2022.103522
- [8] S. Kanoga, **T. Hoshino**, H. Asoh, “Semi-supervised style transfer mapping-based framework for sEMG-based pattern recognition with 1- or 2-DoF forearm motions ,” *Biomedical Signal Processing and Control*, vol. 68, Art. no. 102817, 2021. doi:10.1016/j.bspc.2021.102817

- [9] S. Kanoga, **T. Hoshino**, H. Asoh, “Independent Low-Rank Matrix Analysis-Based Automatic Artifact Reduction Technique Applied to Three BCI Paradigms,” *Frontiers in Human Neuroscience*, vol 14, Art. no. 173, 2020. doi:10.3389/fnhum.2020.00173

査読付き国際会議論文

- [1] **T. Hoshino**, S. Kanoga, A. Aoyama, “Channel- and Label-Flip Data Augmentation for Motor Imagery-Based Brain-Computer Interfaces,” 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC ’ 24), pp. 1-4, Florida, USA, 2024.07. doi:10.1109/EMBC53108.2024.10782028
- [2] S. Kanoga, R. Karakida, **T. Hoshino**, Y. Okawa. M. Tada, “Deep Generative Replay-Based Class-Incremental Continual Learning in sEMG-Based Pattern Recognition,” 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC ’ 24), pp. 1-4, Florida, USA, 2024.07. doi:10.1109/EMBC53108.2024.10781686
- [3] S. Kanoga, T. Nomaguchi, Y. Kobayashi, **T. Hoshino**, C. Tsutsumi, R. Sugimoto, T. Kouyama, R. Nakamura, “A Recurrent Deep Learning-based Monthly Deforestation Prediction Model in Eight Areas for Brazilian Amazon: A Pilot Study,” 2024 IEEE International Geoscience and Remote Sensing Symposium (IGARSS ’ 24), pp. 5208-5212, Athens, Greece, 2024.07. doi:10.1109/IGARSS53475.2024.10642628
- [4] Y. Okawa, S. Kanoga, **T. Hoshino**, T. Nitta, “Sequential Learning on sEMGs in Short- and Long-term Situations via Self-training Semi-supervised Support Vector Machine,” 44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC ’ 22), pp. 3232-3235, UK, 2022.07. doi:10.1109/EMBC48229.2022.9871311
- [5] S. Kanoga, **T. Hoshino**, M. Tada, “A Style Transfer Mapping and Fine-tuning Subject Transfer Framework Using Convolutional Neural Networks for Surface Electromyogram Pattern Recognition,” 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP ’ 22), pp.1171-1175, virtual, 2022.05. doi:10.1109/ICASSP43922.2022.9747657
- [6] S. Kanoga, T. Takase, **T. Hoshino**, H. Asoh, “Time-domain Mixup Source Data Augmentation of sEMGs for Motion Recognition towards Efficient Style Transfer Mapping,” 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC ’ 21), pp.35-38, virtual, 2021.11. doi:10.1109/EMBC46164.2021.9631048

- [7] M. Kato, S. Kanoga, **T. Hoshino**, T. Fukami, “Motor Imagery Classification of Finger Motions Using Multiclass CSP,” 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC ’ 20), pp. 2991-2994, Montreal, Canada, 2020.07. doi:10.1109/EMBC44109.2020.9176612
- [8] **T. Hoshino**, S. Kanoga, M. Tsubaki, A. Aoyama, “Analysis and Usage: Subject-to-subject Linear Domain Adaptation in sEMG Classification,” 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC ’ 20), pp. 674-677, Montreal, Canada, 2020.07. doi:10.1109/EMBC44109.2020.9175755
- [9] S. Kanoga, **T. Hoshino**, H. Asoh, “Subject Transfer Framework Based on Source Selection and Semi-Supervised Style Transfer Mapping for sEMG Pattern Recognition,” 45th International Conference on Acoustics, Speech, and Signal Processing (ICASSP ’ 20), pp. 1349-1353, Barcelona, Spain, 2020.05. doi:10.1109/ICASSP40776.2020.9054070
- [10] **T. Hoshino**, S. Kanoga, A. Kanemura, and T. Ogawa, “A No-Reference Metric of Cerebral Blood Flow Extraction for fNIRS Data”, IEEE International Conference on Systems, Man, and Cybernetics (SMC ’ 18), pp. 83-89, Miyazaki, Japan, 2018.10. doi:10.1109/SMC.2018.00024

その他国際会議発表

- [1] T. Kaku, S. Bun, K. Takahashi, Y. Mimura, T. Tezuka, M. Kubota, Y. Momota, **T. Hoshino**, H. Tabuchi, K. Takahata, S. Nakajima, H. Uchida, M. Mimura, D. Ito, “Evaluation of the ALPS Index as a Glymphatic System Biomarker Across the Alzheimer’s Disease Continuum,” Alzheimer’s Association International Conference 2025 (AAIC ’ 25), Toronto, Canada, 2025.06.
- [2] **T. Hoshino**, S. Kanoga, A. Aoyama, “Reducing Data Collection Demands in MEG-Based Classification Through Virtual Participants,” The 23rd International Conference on Biomagnetism (BIOMAG ’ 24), Sydney, Australia, 2024.08.
- [3] A. Aoyama, **T. Hoshino**, “Neural oscillations associated with adaptation to left-right reversed audition,” The 22nd International Conference on Biomagnetism (BIOMAG ’ 22), Birmingham, UK, 2022.08.
- [4] A. Aoyama, **T. Hoshino**, “Adaptation-related changes in neural oscillations during exposure to left-right reversed audition,” The 5th International Conference on Basic and Clinical Multimodal Imaging (BaCI ’ 21), Naples, Italy, 2021.10.

- [5] A. Aoyama, **T. Hoshino**, “Explicit and implicit auditory-motor spatial coordination during exposure to left-right reversed audition: an MEG study,” 4th International Conference on Basic and Clinical Multimodal Imaging (BaCI ’ 19), Chengdu, China, 2019.07.
- [6] **T. Hoshino**, A. Aoyama, “Estimation of MEG Sensor Signals on Bad Channels Using a Supervised Algorithm,” The 21st International Conference on Biomagnetism (BIOMAG ’ 18), Philadelphia, USA, 2018.08.
- [7] **T. Hoshino**, S. Kanoga, A. Kanemura, T. Ogawa, “Selecting artifactual independent components from fNIRS based on decoding analysis,” Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC ’ 18), Honolulu, USA, 2018.07.
- [8] **T. Hoshino**, S. Kanoga, A. Aoyama, “Dictionary learning of fNIRS signals to examine scalp blood flow,” Organization for Human Brain Mapping (OHBM ’ 18), Singapore, 2018.06.

国内発表

- [1] **星野貴行**, 叶賀卓, 青山敦, “仮想被験者を用いた生体信号の識別,” 日本生体工学会 専門別研究会 第 39 回実社会におけるマルチモーダル脳情報応用技術研究会, オーラル, オンライン, 2024.02.
- [2] 加藤将, 叶賀卓, 長島和馬, **星野貴行**, 深見忠典, “CNN を用いた指の運動想起脳波の識別,” 第 39 回日本医用画像工学会大会 (JAMIT ’ 20), オーラル, 山形, 2020.09.
- [3] 後藤優仁, **星野貴行**, 青山敦, “経頭蓋ランダムノイズ刺激による腹話術効果の抑制,” 第 59 回日本生体医工学会大会, 2020.05.
- [4] 後藤優仁, **星野貴行**, 青山敦, “経頭蓋ランダムノイズ刺激による腹話術効果の変調,” 日本生体医工学会第 27 回実社会におけるマルチモーダル脳情報応用技術研究会, 東京, 2024.02.
- [5] 後藤優仁, **星野貴行**, 青山敦, “経頭蓋電気刺激による腹話術効果の阻害に関する検討,” 第 49 回日本臨床神経生理学会, 福島, 2019.11.
- [6] 後藤優仁, **星野貴行**, 青山敦, “経頭蓋電気刺激による腹話術効果の阻害,” 生理学研究 所研究会「脳神経ダイナミクスの可視化と制御」, 愛知, 2019.07.
- [7] **星野貴行**, 叶賀卓, 兼村厚範, 小川剛史, 青山敦, “頭部運動に依存した脳血流と頭皮血流が fNIRS デコーディングに与える影響の解析,” 第 58 回日本生体医工学会大会, ポスター, 沖縄, 2019.06.

- [8] 星野貴行, 兼村厚範, 小川剛史, “近赤外分光画像法で計測された脳活動の時空間解析”, 情報論的学習理論ワークショップ (IBIS), 京都, 2016.11.
- [9] 野沢健人, 星野貴行, 福田拓也, 兼村厚範, “脳波データへの深層学習の適用”, 情報論的学習理論ワークショップ (IBIS), 京都, 2016.11.