Towards Subject Agnostic Affective Emotion Recognition

Amit Kumar Jaiswal[1] Haiming Liu[2] Prayag Tiwari[3]

[1]University of Surrey, UK
[2]University of Southampton, UK
[3]Halmstad University, Sweden

October 22, 2023

Outline

Introduction & Problem

Challenges

- Occurrence of distributional shift from the non-stationary nature of EEG signals.
- Limited theoretical background for (affective) emotion recognition in the presence of domain shift in data.
- Large distributional shifts between domains limit generalisation to unseen domains.
- Typical ML approaches is difficult to utilise for subject-agnostic EEG-driven emotion recognition due to structural variability across different domains.

Problem

- Learning from Multi-source data
  - Multimodal EEG data of emotional video clips
  - We refer to each data source a domain or a subject
- Identify emotions during realistic data (EEG) shifts

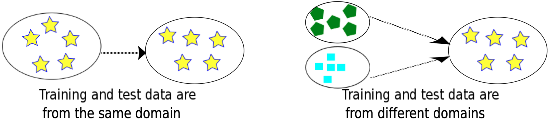
Introduction & Problem

Domain Adaptation & Generalisation

- Domain adaptation (DA) [Ganin 2015] and domain generalisation (DG) [Muandet 2013] are two classes of approaches to mitigate the impact of discrepancies among distributions of the training and testing data from varied sources.
- We refer the distribution of the training data drawn from the source domain, and the testing data is drawn from target domain.
    - DA: Leverage data from a single or multiple source domains including limited data from the target domain from training
    - DG: Leverage data from multiple source domains for training, where no training data from the target domain.

Background

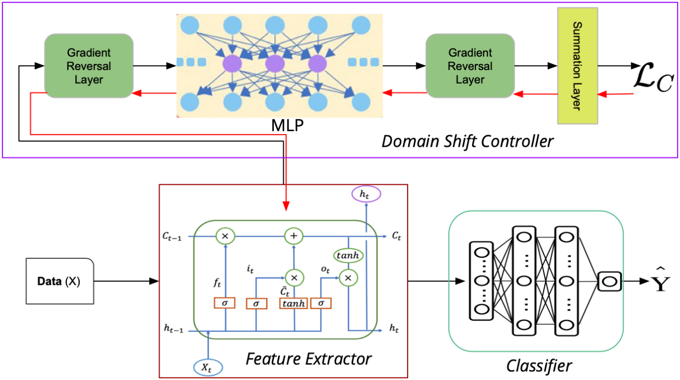
- **Domains** are modelled as a probability distributions over an input space



- **Domain adaptation** approaches to modelling scenarios, where a trained model on a source domain performing a certain task (classification, regression, etc.), but then applied to a different task in a different target domain.
  - **Input:** A lot of labelled data in the source domain and only unlabelled data in the target domain.
  - **Output:** A common representation between source domain data and target domain data and a model on the new representation for use in the target domain.
- **Distributional shift:** Representations learned in the source domain might not work well in the target domain

Our Approach - Graphical Overview

- We introduce, a **meta-learning based augmented domain adaptation (MeLaDA)** which includes
  - a sum-decomposable domain shift controller $D_C(\omega)$, which leverages the features extracted by $F(\theta)$ to assess the dissimilarity between the current domain and the shift-independent domain
  - a temporal multi-layer perceptron (MLP), which consists of a feature extractor $F(\theta)$ classifier $C(\phi)$



Our Approach - Theoretical Basis

- **Problem Settings:** The input space for EEG data as $\mathcal{E}_I$, and the output space as $\mathcal{E}_O$.
  - A domain $\mathcal{D}$ is a joint distribution $P_{\mathcal{E}_I \mathcal{E}_O}$ over the space $\mathcal{E}_I \times \mathcal{E}_O$.
  - Samples of a domain S_i refers to a set of $\{\mathcal{E}_{I\,i}, \mathcal{E}_{O\,i}\}$
  - A mapping function $\mathcal{Q}$ to transform one domain into another while minimising the divergence $(d(\cdot,\cdot))$ across the domains
  - Alignment technique: $\mathcal{Q}_{da} = \arg\min_{\mathcal{Q}} d(\mathcal{Q}(S_i), \mathcal{Q}(S_j))$

  Shift-independent Domain
  - Multi-source domain adaptation (or domain generalisation)

$$\text{Proposition:} \mathcal{Q}_{mda} = \arg\min_{\mathcal{Q}} \sum_{S_i, S_j \in S_{\text{shift}}} d(\mathcal{Q}(S_i), \mathcal{Q}(S_j))$$

- Domain shift in EEG patterns led to domain discrepancy, so we introduce the **sum-decomposable component** based on [Wagstaff 2019], a summation layer to effectively represent any permutation-invariant function

Our Approach - Domain Shift Controller $D_C(\omega)$

- **Adversarial elements:** We integrate two gradient reversal layers (GRL) [Ganin 2016] before and after a two-layer MLP, followed by an additional layer to further augment its performance
- **Rationale for Adversarial actors:** To represent the **domain discrepancy**, which is the largest difference between two domains to depict their divergence
  - **Pros:** In forward propagation, the GRL act as an identity map and reverses the direction of gradients in backword propagtion, and is overall an alternative of 'Discriminator' in GAN.
  - **Cons:** Two GRL layers exhibit instability during training
- We extension MND meansure [Wang 2018], as Maximum mean norm discrepancy (M_{MND})

$$M_{\text{MND}}(S_i, S_j) := \max_{g \in \mathcal{Q}} \left\| \mathbb{E}_{x \in S_i} g(x) - \mathbb{E}_{x \in S_j} g(x) \right\|_2$$

- Objective function of the $D_C(\omega)$

$$\mathcal{L}_c = \sum_{S_i \in S_{\text{shift}}} \max_{g \in \mathcal{Q}} \min_{\tau \in V_{\text{lspace}}} \left\| \mathbb{E}_{x \in S_i} g(x) - \tau \right\|_2$$

Contributions

- We proposed MeLaDA, an approach to develop a subject-agnostic EEG-based emotion recognition model.
- We setup theoretical aspects to mitigate the problem of domain shift, a way to incorporate any type of domain discrepancy by introducing a sum-decomposable network.
- We carried out experiments on the publicly available EEG-based aBCIs dataset, SEED[1]. The results indicate that our proposed approach outperforms domain generalisation methods. Moreover, our proposed approach, MeLaDA, exhibits comparable time and storage costs to domain generalisation methods.

[1] https://bcmi.sjtu.edu.cn/home/seed/seed.html

Model Training Strategy

- We utilise meta-learning approach to ensure the generalisation ability of the domain shift controller.
- Our training approach is of two-stage,
  - Training domain shift controller: Optimise $\mathcal{L}_c$ to minimise domain shift and couple via meta-learning

$$\mathcal{L}_{\text{meta}} = \sum_{(x_i, y_i) \in S_{\text{valid}}} \tanh \left(\ell(x^i, y^i; \theta^{'}) - (\ell(x^i, y^i; \theta) \right)$$

  - Overall loss for updating ω

$$\mathcal{L}_C(\theta, \phi, \omega; S_{\text{train}}) + \lambda \mathcal{L}_{\text{meta}}(\theta^{'}, \phi, \omega; S_{\text{valid}})$$

  - Training the overall network: We utilise model-agnostic meta learning approac [Finn 2017] to train both the domain shift controller and classifier as two distinct tasks alternatively.

Model Training Strategy - Algorithms

Algorithm 1 Training the Domain Shift Controller

Input: Given a domain $\mathscr{D}$ and θ, ϕ, ω are parameters.

Output: ω

1: $\mathscr{D} : (S_{\text{train}}, S_{\text{valid}}) \leftarrow \mathscr{D}$ ▷ Random partitioning

2: $\mathscr{L}_C(\theta, \omega; S_{\text{train}}) \leftarrow D_C(F(X))$ ▷ Meta-training stage

3: $\theta \leftarrow \theta' - \alpha_\theta \mathscr{L}_C(\theta, \omega; S_{\text{train}})$ ▷ Meta-training stage

4: $\mathscr{L}_{\text{meta}}(\theta, \theta', \phi; S_{\text{valid}}) \leftarrow$ ▷ Meta-validation stage

$\hookrightarrow \sum_{(x_i, y_i) \in S_{\text{valid}}} \tanh\left(\ell(x^i, y^i; \theta') - (\ell(x^i, y^i; \theta)\right)$

5: Update ω utilising $\mathscr{L}_C + \lambda \mathscr{L}_{\text{meta}}$ ▷ Optimisation

Algorithm 2 Training the MeLaDA framework

Input: Given a domain $\mathscr{D}$ and θ, ϕ, ω are parameters.

Output: θ, ϕ, ω

1: $\mathscr{D} : (S_{\text{train}}, S_{\text{valid}}) \leftarrow \mathscr{D}$ ▷ Random partitioning

2: $\mathscr{L}_C(\theta, \omega; S_{\text{train}}) \leftarrow D_C(F(X))$ ▷ Meta-training stage

3: $\mathscr{L}^{\text{train}}_{\text{classif}}(\theta, \phi) \leftarrow \ell(C(F(X_{\text{train}})), Y_{\text{train}})$ ▷ Meta-training stage

4: $\theta \leftarrow \theta' - \alpha_\theta \mathscr{L}_C(\theta, \omega; S_{\text{train}})$ ▷ Meta-training stage

5: $\mathscr{L}^{\text{valid}}_{\text{classif}}(\theta', \phi) \leftarrow \ell(C(F(X_{\text{valid}})), Y_{\text{valid}})$

$\hookrightarrow$ ▷ Meta-validation stage

6: Update θ, ϕ, ω utilising $\lambda \mathscr{L}_C + \mathscr{L}^{\text{valid}}_{\text{classif}} + \mathscr{L}^{\text{train}}_{\text{classif}}$

$\hookrightarrow$ ▷ Optimisation

Experiment - Dataset and Feature Extraction

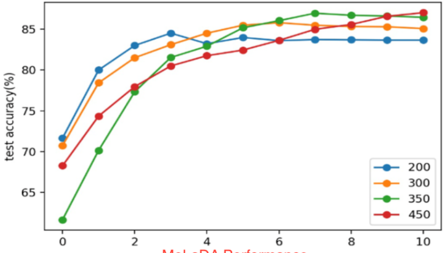
Dataset: We employ SEED dataset for emotion recognition and aBCIs using EEG signals. We use EEG signals data which has stimulation of 15 video clips (or 15 domains/subjects) from Chinese movies. It contains three distinct emotion classes which are happiness, neutrality, and sadness.

- Each subject is repeated 3 times in intervals of one week.
- **Extracted Features:** We use the existing differential entropy features, which corresponds to the logarithmic spectral energy of a fixed-length EEG sequence within a specific frequency band.
- We extracted the spectral energy for EEG patterns variability, where we apply the short-time Fourier transform using a non-overlapping Hanning window of 1 second to the EEG signal, provided five frequency bands.
- Given the input data is time-series, with each sample of dimension 310 (62 channels $\times$ 5 frequency bands), and we resample the feature with a time-step of 15 and a 1-second overlap.
- Following data preprocessing, we have 3184 samples per domain, and the overall data size is 15 x 3184 $=$ 47,760 samples

Result

Models	DA		Models	DG	
	MA	SD		MA	SD
SVM [46]	0.567	0.16	-	-	-
TCA [46]	0.640	0.15	MLDG [36]	0.795	0.12
TPT [46]	0.752	0.13	FC [38]	0.806	0.12
DAN [29]	0.838	0.08	DICA [33]	0.7	0.08
DANN [29]	0.792	0.13	DResNet [33]	0.853	0.08
WGAN-DA [31]	0.871	0.07	PnP [34]	0.854	0.07
MeLaDA (Ours)	0.864	0.09	MeLaDA	0.864	0.09

Table 1
The mean accuracy and standard deviation for both domain adaptation (DA) and domain generalisation (DG) are reported with comparative baseline methods on the SEED dataset.



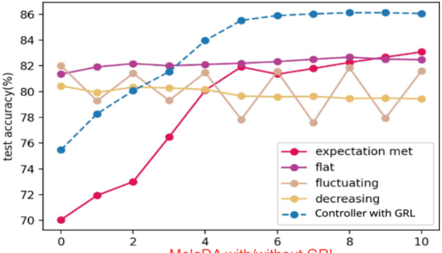


Figure: Fast Adaptation to Domains

Conclusion

- Our work introduces MeLaDA, an augmented domain adaptation approach using meta learning, for subject-agnostic EEG-based emotion recognition.
- MeLaDA allows for subject-agnostic model development without requiring source domain data during testing.
- The approach employs a sum-decomposable domain shift controller for enhanced domain adaptation, which is theoretically proven.
- MeLaDA combines adversarial actors - GRL and meta-learning techniques, facilitating generalisation to new domains with minimal self-adaptive iterations.
- Experimental results on the SEED dataset demonstrate superior performance of MeLaDA compared to traditional domain generalisation methods.
- MeLaDA's suitability for constructing subject-agnostic affective models is highlighted, surpassing conventional domain adaptation, domain generalisation, and ASFM (adaptive subspace feature matching) methods.

References I

Chelsea Finn, Pieter Abbeel et Sergey Levine.
Model-agnostic meta-learning for fast adaptation of deep networks.
In International conference on machine learning, pages 1126–1135. PMLR, 2017.

Yaroslav Ganin et Victor Lempitsky.
Unsupervised domain adaptation by backpropagation.
In International conference on machine learning, pages 1180–1189. PMLR, 2015.

Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand et Victor Lempitsky.
Domain-adversarial training of neural networks.
The journal of machine learning research, vol. 17, no. 1, pages 2096–2030, 2016.

References II

Krikamol Muandet, David Balduzzi et Bernhard Schölkopf.
Domain generalization via invariant feature representation.
In International conference on machine learning, pages 10–18. PMLR, 2013.

Edward Wagstaff, Fabian Fuchs, Martin Engelcke, Ingmar Posner et Michael A Osborne.
On the limitations of representing functions on sets.
In International Conference on Machine Learning, pages 6487–6494. PMLR, 2019.

Jindong Wang, Wenjie Feng, Yiqiang Chen, Han Yu, Meiyu Huang et Philip S Yu.
Visual domain adaptation with manifold embedded distribution alignment.
In Proceedings of the 26th ACM international conference on Multimedia, pages 402–410, 2018.

References III