Supervised Contrastive Learning with Identity-Label Embeddings for Facial Action Units Recognition

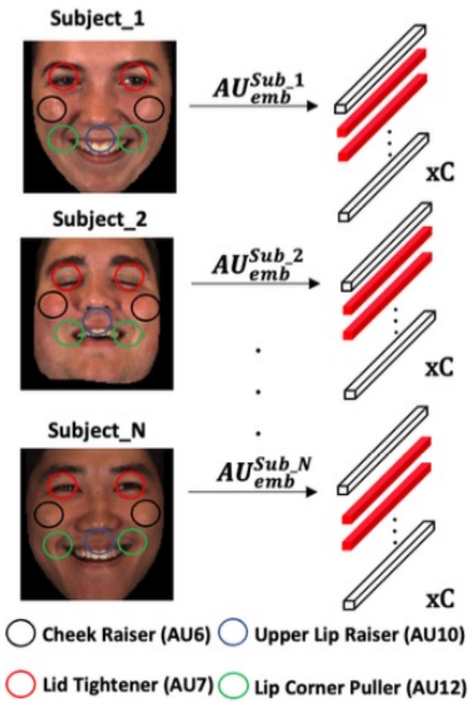
Tangzheng Lian[1], David A Adama[2], Pedro Machado[2], Doratha E Vinkemeier[1]

Nottingham Trent University

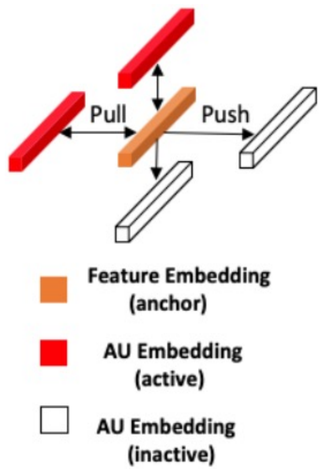
BMVC 2023

Research motivation

- The inherent differences in facial behaviour and appearance across individuals.

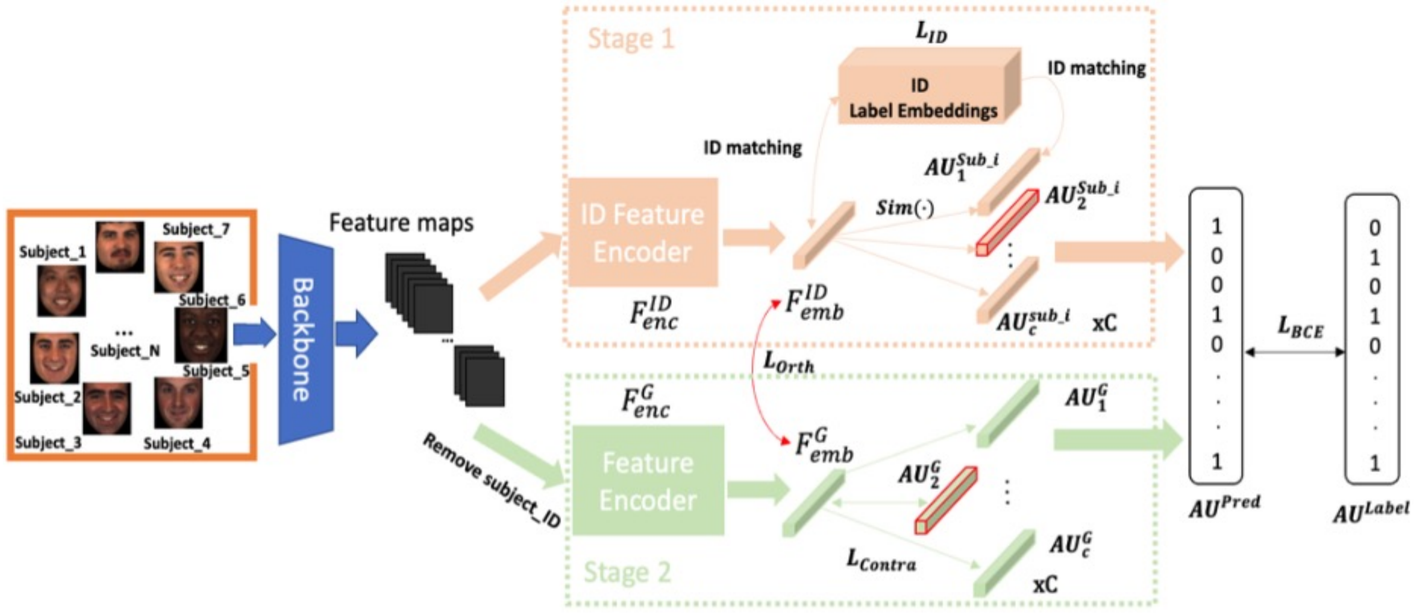


- The complex interactions among different AUs.



Method

- We propose a novel two-stage training framework, called CL-ILE. In the first stage of CL-ILE, we introduce identity-label embeddings (ILEs) to train an ID feature encoder capable of generating person-specific feature embeddings for input face image images.
- In the second stage, we present a data-driven method that implicitly models the relationships among AUs using contrastive loss in a supervised setting while eliminating the person-specific features generated by the first stage to enhance generalizability.



$$\mathcal{L}_{Contra} = \frac{1}{|B|} \sum_{(F_{emb}^{G}, y) \in B} \frac{1}{|P(y)|} \sum_{p \in P(y)} -\log \frac{exp(sim\left(F_{emb}^{G}, AU_{p}^{G}\right)/\tau)}{\sum_{c=1}^{C} exp(sim\left(F_{emb}^{G}, AU_{c}^{G}\right)/\tau)}$$

$$\mathcal{L}_{BCE} = -\frac{1}{C} \sum_{c=1}^{C} w_c \left[y_c \log s\left(F_{emb}^{ID} \cdot AU_c^{Sub_i}\right) + (1 - y_c) \log\left(1 - s\left(F_{emb}^{ID} \cdot AU_c^{Sub_i}\right)\right)\right]$$

Results

BP4D

Method/AUs	1	2	4	6	7	10	12	14	15	17	23	24	Avg.
DRML† [21]	36.4	41.8	43.0	55.0	67.0	66.3	65.8	54.1	33.2	48.0	31.7	30.0	48.3
ATF* [42]	39.2	35.2	45.9	71.6	71.9	79.0	83.7	65.5	33.8	60.0	37.3	41.8	55.4
LP-Net*† [31]	43.4	38.0	54.2	77.1	76.7	83.8	87.2	63.3	45.3	60.5	48.1	54.2	61.0
AU-GCN† [27]	46.8	38.5	60.1	**[80.1]**	[79.5]	84.8	88.0	67.3	52.0	63.2	40.9	52.8	62.8
SRERL† [24]	46.9	45.3	55.6	77.1	78.4	83.5	87.6	63.9	[52.2]	[63.9]	47.1	53.3	62.9
PIAP *† [37]	[54.2]	[47.1]	54.0	79.0	78.2	**[86.3]**	89.5	66.1	49.7	63.2	**[49.9]**	52.0	64.1
SupHCL*† [6]	52.8	45.7	**[61.6]**	79.5	79.3	84.7	86.9	[67.6]	51.4	62.5	48.6	52.3	64.4
ME-Graph† [28]	52.7	44.3	[60.9]	[79.9]	**[80.1]**	[85.3]	**[89.2]**	**[69.4]**	**[55.4]**	**[64.4]**	[49.8]	[55.1]	**[65.5]**
CL-ILE (Ours)	**[55.1]**	**[52.1]**	55.0	78.2	75.5	83.4	[88.1]	67.4	51.9	59.5	46.9	**[62.2]**	[64.6]

DISFA

Method/AUs	1	2	4	6	9	12	25	26	Avg.
DRML† [21]	17.3	17.7	37.4	29.0	10.7	37.7	38.5	20.1	26.7
ATF* [42]	45.2	39.7	47.1	48.6	32.0	55.0	86.4	39.2	49.2
AU-GCN† [27]	32.3	19.5	55.7	[57.9]	**[61.4]**	62.7	90.9	[60.0]	55.0
SRERL† [24]	45.7	47.8	59.6	47.1	45.6	73.5	84.3	43.6	55.9
LP-Net *† [31]	29.9	24.7	[72.7]	46.8	49.6	72.9	[93.8]	**[65.0]**	56.9
ME-Graph† [28]	[54.6]	47.1	**[72.9]**	54.0	[55.7]	[76.7]	91.1	53.0	[63.1]
PIAP† [37]	50.2	[51.8]	71.9	50.6	54.5	**[79.7]**	**[94.1]**	57.2	**[63.8]**
CL-ILE (Ours)	**[58.9]**	**[56.4]**	69.1	**[58.5]**	54.4	72.2	85.9	47.3	62.8

Correlation Matrix