

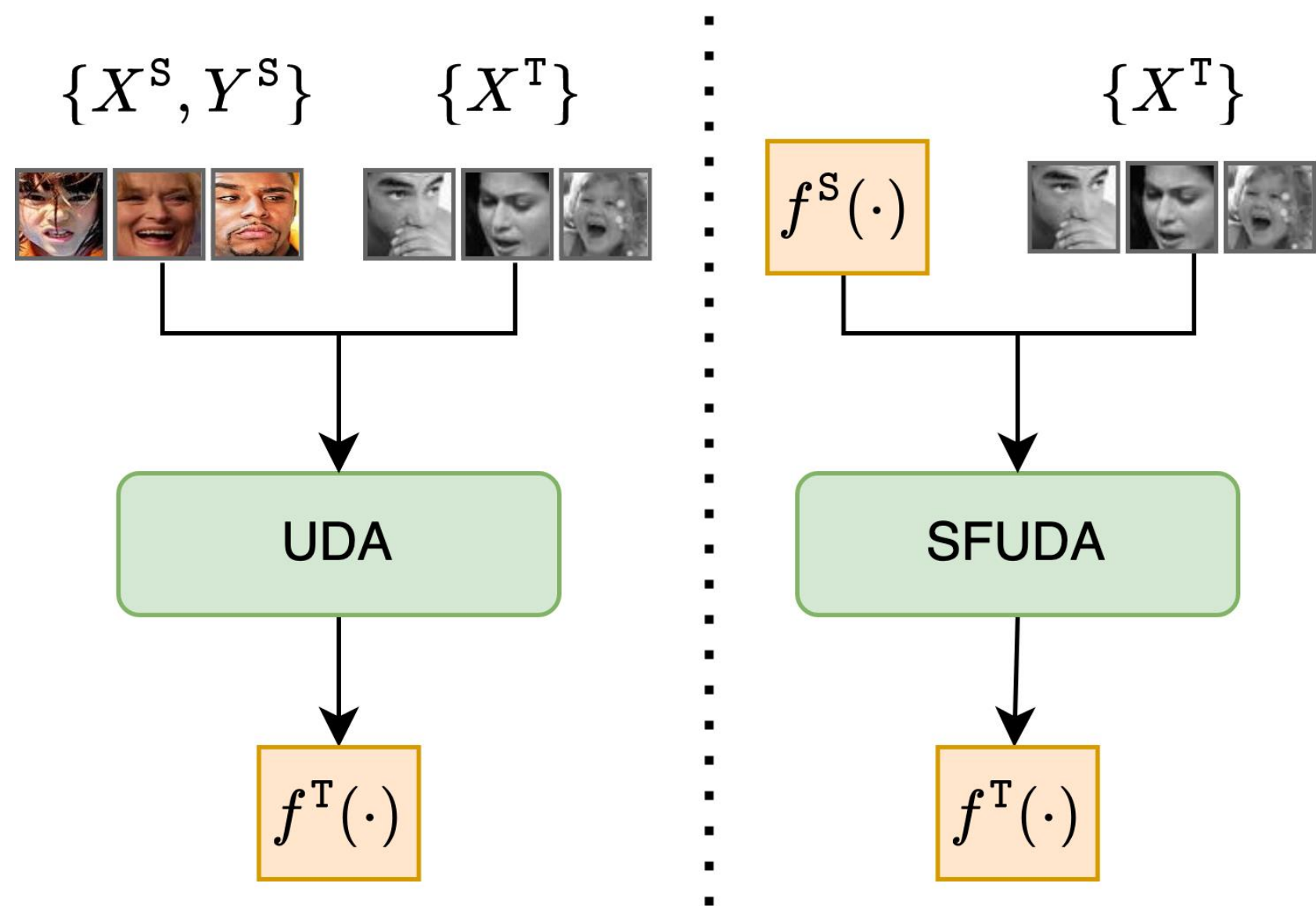
Cluster-level pseudo-labelling for source-free cross-domain facial expression recognition

Alessandro Conti, Paolo Rota, Yiming Wang, Elisa Ricci

Introduction

Problem

- Domain Adaptation applied on the task of **Facial Expression Recognition (FER)**.
- No access to target labels (**Unsupervised**)
- No access to source data (**Source-Free**).



Contributions

- Perform **self-supervised pretraining** of the target model
- Introduce a **novel pseudo-labelling technique based on clustering**.

Highlighted facts

- Address Source-Free Unsupervised Domain Adaptation
- Apply the problem on Facial Expression Recognition
- Propose a novel pseudo-labelling function based on cluster statistics
- Achieve competitive results with non-source-free methods

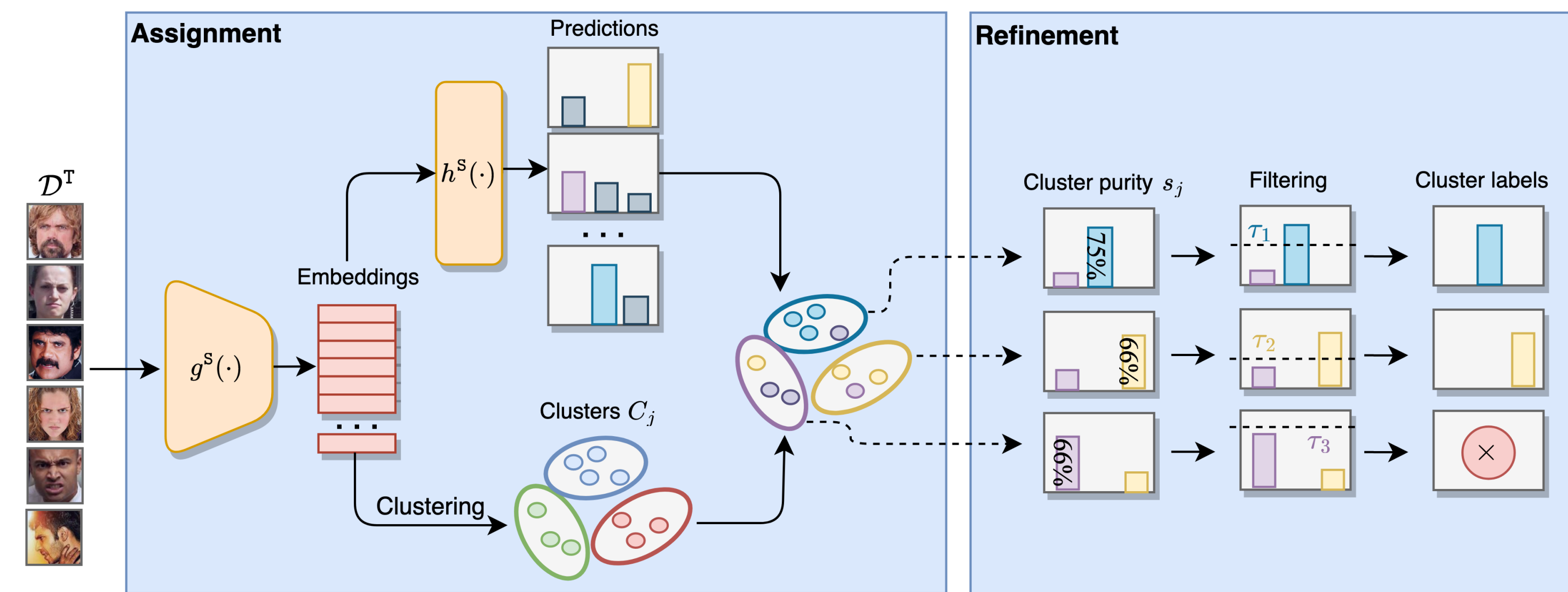
Contact: alessandro.conti-1@unitn.it
Paper: <https://arxiv.org/abs/2210.05246>



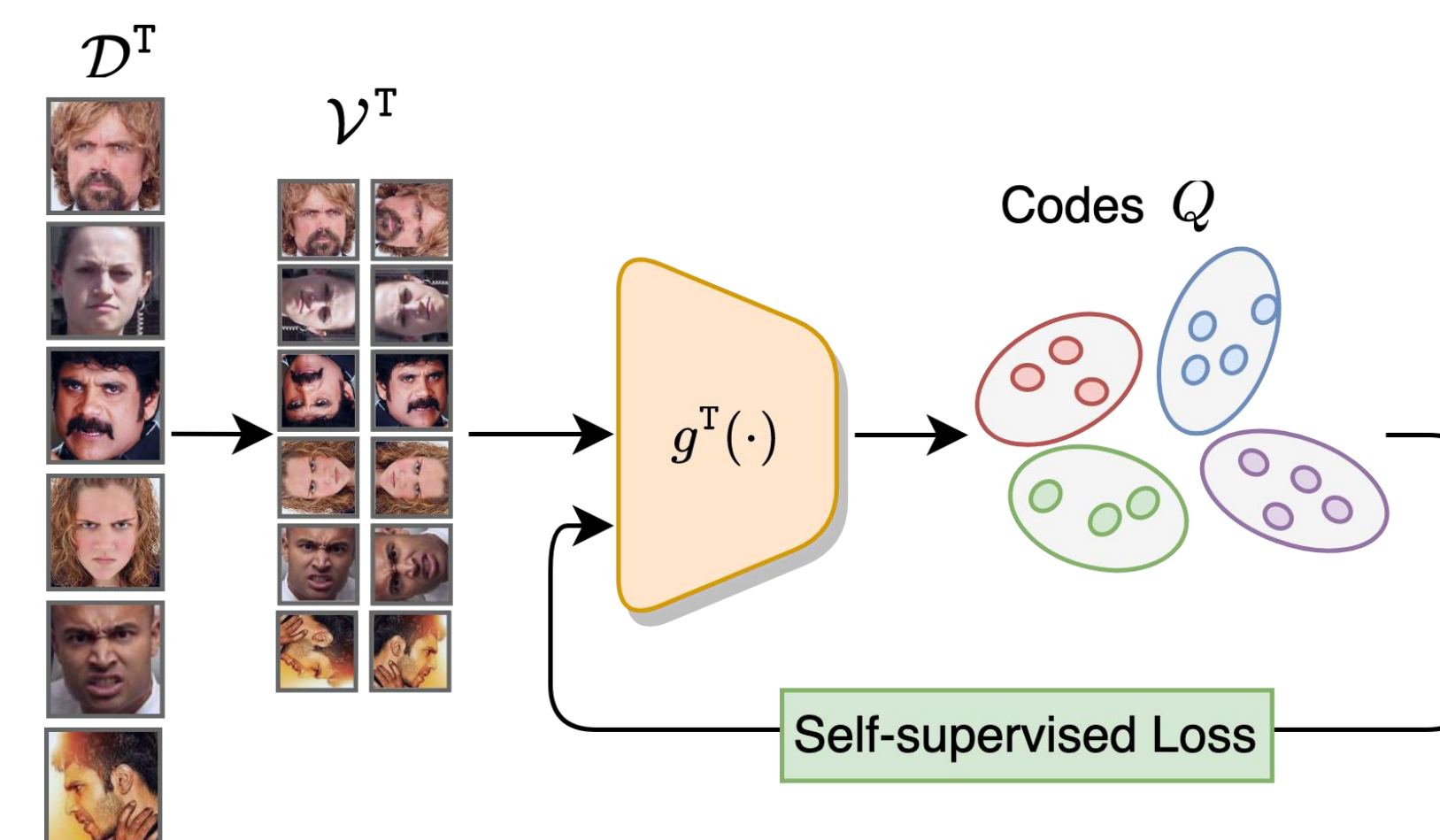
This work was supported by the EU JPI/CH SHIELD project, by the PRIN project PREVUE (Prot. 2017N2RK7K), the EU H2020 MARVEL (957337) project, the EU ISFP PROTECTOR (101034216) project, the EU H2020 SPRING project (871245), and by Fondazione VRT. It was carried out under the "Vision and Learning joint Laboratory" between FBK and UNITN.

Method

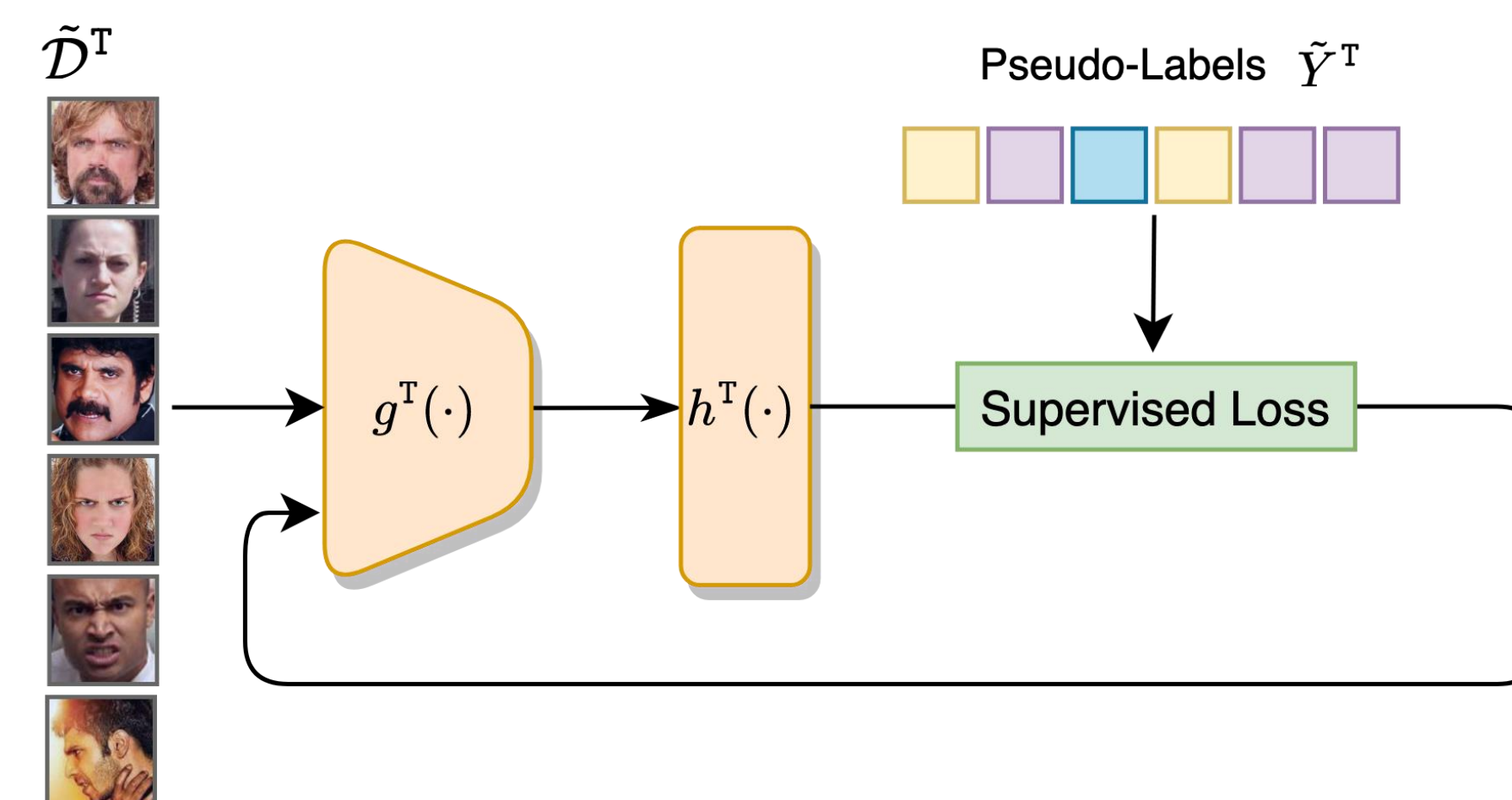
- Propose a **three-stage pipeline**



- Stage 1**: assign and refine pseudo-labels
- Pseudo-label samples and cluster features with the source model
- Count pseudo-labels in each cluster
- The **cluster label** is the major class in the cluster
- The % of the major class is the **cluster purity**
- Use a percentile class-wise threshold to ignore the least pure clusters



- Stage 2**: pretrain the target model with SwAV



- Stage 3**: finetune the network for FER with the pseudo-labelled subset and the cluster labels

Results

- Compared with UDA methods (with access to the source data) and SHOT (a source-free method)
- Achieve **competitive performance** in four adaptation scenarios
- Outperform all the UDA counterparts in AFE to FER2013.

Method	AFE \rightarrow ExpW	AFE \rightarrow FER2013	RAF-DB \rightarrow ExpW	RAF-DB \rightarrow FER2013
ICID [14]	54.85	46.44	68.52	53.00
DFA [43]	62.53	36.88	47.42	47.88
LPL [22]	54.51	49.82	68.35	53.61
DETNet [20]	58.41	45.39	43.92	42.01
FTDNN [38]	55.29	48.58	68.08	53.28
ECAN [21]	62.52	46.15	48.73	50.76
CADA [27]	58.50	48.61	63.74	54.71
SAFN [35]	55.17	50.07	68.32	53.31
SWD [18]	56.56	51.84	65.85	53.70
AGRA [6]	65.03	51.95	69.70	54.94
SHOT-IM [23]	53.52	49.51	53.13	49.44
SHOT [23]	54.12	49.44	53.51	49.36
CluP (DeepClusterV2)	62.56	50.47	65.43	53.83
CluP (SwAV)	65.00	52.51	66.60	53.71

- For class-wise accuracy, outperform the source model finetuned with SHOT for most of the classes.

Method	Surprise	Fear	Disgust	Happiness	Sadness	Anger	Neutral
SHOT-IM [23]	28.29	45.05	9.86	75.97	56.12	40.66	71.96
SHOT [23]	28.18	43.24	10.25	75.59	53.53	40.18	74.37
CluP (DeepClusterV2)	29.44	45.05	2.83	83.15	77.70	34.72	76.65
CluP (SwAV)	37.89	30.63	13.57	80.72	50.85	44.51	74.49

- Consistent improvements** from standard pseudo-labelling when ablating the backbone and the score function

Backbone	Score	AFE \rightarrow ExpW	AFE \rightarrow FER2013	RAF-DB \rightarrow ExpW	RAF-DB \rightarrow FER2013
Source	Conf.	56.43	48.36	59.79	50.47
Source	Purity	56.54	47.34	61.18	54.29
SwAV	Conf.	62.88	51.27	63.22	50.68
SwAV	Purity	65.00	52.51	66.60	53.71

Conclusions

- We propose a novel pseudo-labelling technique for Source-Free Unsupervised Domain Adaptation for Facial Expression Recognition
- Our method achieves competitive results against methods with access to the source data