Learning to transfer: transferring latent task structures and its application to person-specific facial action unit detection

Computer Vision Laboratory
cvl.cs.nott.ac.uk

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Timur Almaev
psxta4@nottingham.ac.uk

Brais Martinez
brais.martinez@nottingham.ac.uk

Michel Valstar
michel.valstar@nottingham.ac.uk

OVERVIEW

Domain

Example facial Action Units (AUs)

Problem

- Large impact of nuisance factors, e.g. identity
- Highly expensive manual AU annotation
- Infrequence of occurrence of some AUs

→ Classifiers generalisability

Proposed Solution

- Train person-specific AU models
- Learn all AU models jointly (Multi-Task)
- Improve detection of rare AUs with frequent

Task - a binary classification problem of a specific AU and a specific subject.

Target AU - AU we want to detect.

Reference AU - AU used to improve detection of the target AU.

METHODOLOGY

Our Approach Overview

Ref. AU data

Target AU data

Sufficient amount of labelled data is available for all subjects

Amount of labelled data for some subjects is limited or non-existent

Task Relatedness Estimation (Latent Structure)

Task Relatedness Estimation (Latent Structure)

Latent Tasks Structure Transfer

Latent Task Structures of reference and target AUs are constrained to be similar

RLTS - Regularised Latent Task Structure

Latent Task Structure

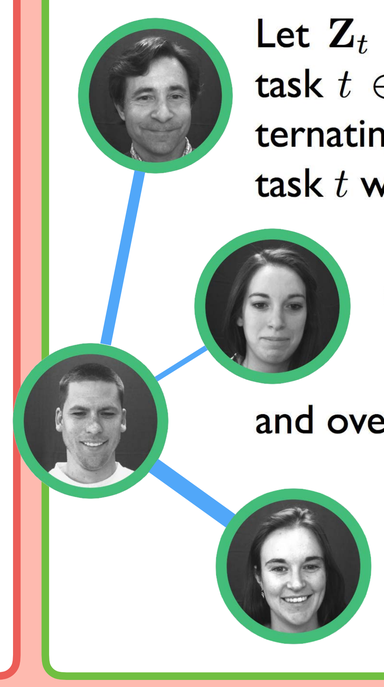
Let matrix $\mathbf{W}$ hold T tasks that lie on a common linear subspace:

$$\mathbf{W} = \mathbf{LS}, \qquad (1)$$

where $\mathbf{L}$ is the subspace generator and $\mathbf{S}$ is the matrix of linear coefficients.

Matrices $\mathbf{L}$ and $\mathbf{S}$ are learnt using a l_1-based sparse dictionary learning approach.

Latent Tasks Structure Learning



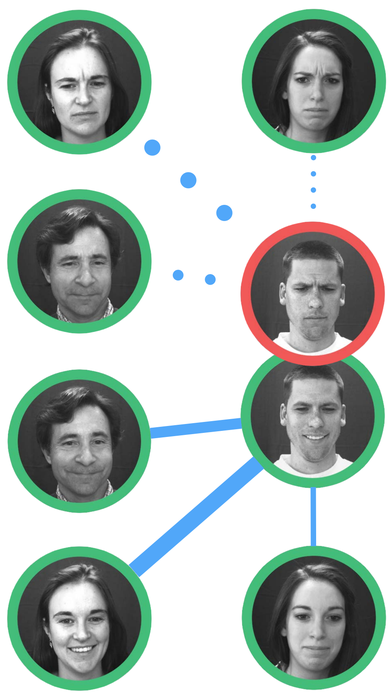
Let $\mathbf{Z}_t = \{(\mathbf{x}_t^i, y_t^i)\}_{i=1,\dots,N_t}$ be the dataset of task $t \in \{1...T\}$. $\mathbf{L}$ and $\mathbf{S}$ are computed by alternating between minimisation over $\mathbf{s}_t$ for every task t with fixed $\mathbf{L}$:

$$\mathbf{s}_t = \underset{\mathbf{s}}{\operatorname{argmin}}\ \mathcal{E}(\mathbf{Z}, \mathbf{L}, \mathbf{s}) + \lambda\|\mathbf{s}\|_1, \qquad (2)$$

and over $\mathbf{L}$ while fixing $\mathbf{S}$:

$$\underset{\mathbf{L}}{\operatorname{argmin}}\ \mathcal{E}(\mathbf{Z}, \mathbf{L}, \mathbf{S}) + \mu\|\mathbf{L}\|_F^2 \qquad (3)$$

Main Contribution: Latent Tasks Structure Transfer



Transfer is performed by alternating between Eq. 3 and Eq. 2 extended with an additional regularisation term $\mathcal{R}(\mathbf{s})$:

$$\mathbf{s}_t = \underset{\mathbf{s}}{\operatorname{argmin}}\ \mathcal{E}(\mathbf{Z}, \mathbf{L}, \mathbf{s}) + \lambda\|\mathbf{s}\|_1 + \mathcal{R}(\mathbf{s}), \qquad (4)$$

where

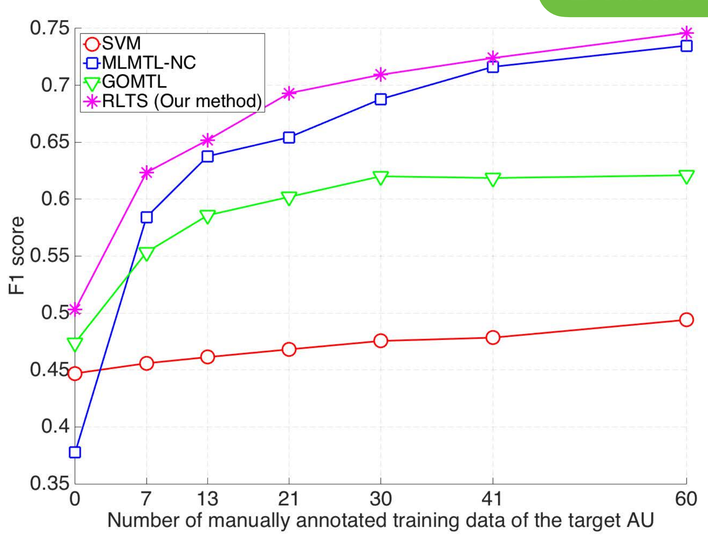
$$\boxed{\mathcal{R}(\mathbf{s}) = \tau\|\mathbf{s}_{target} - \mathbf{s}_{ref}\|_2^2,} \qquad (5)$$

that vanishes when $AU_{target} = AU_{ref}$.

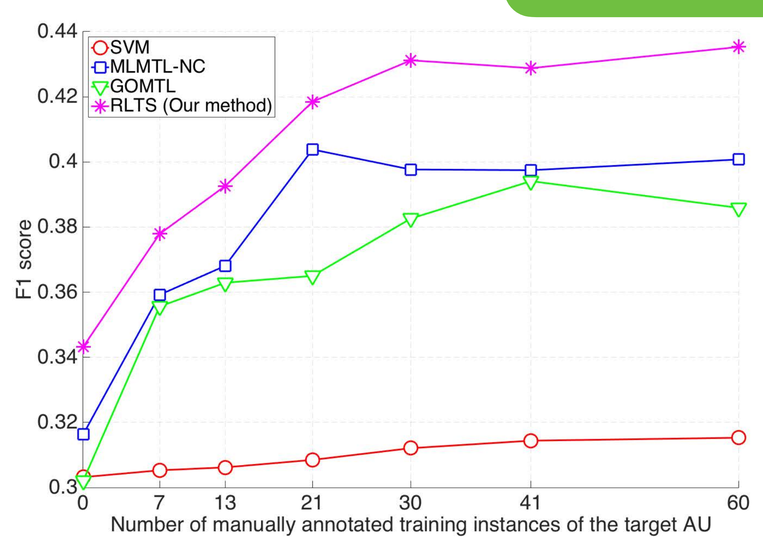
Given sufficient amount of AU_{ref} data, $\mathcal{R}(\mathbf{s})$ enables AU_{target} modelling *in absence or limited availability of* AU_{target} *data.*

RESULTS

DISFA dataset



McMaster dataset



This work was funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No645378. We are also very grateful for the access to the University of Nottingham High Performance Computing Facility. We would like to thank British Machine Vision Association (BMVA) for providing funds for attending the conference.