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

Timur Almaev
psxta4@nottingham.ac.uk

Brais Martinez
brais.martinez@nottingham.ac.uk

Michel Valstar
michel.valstar@nottingham.ac.uk

This work was funded by the European Union’s Horizon 2020 research and innovation programme under grant agreement No 645378. 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.

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

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

Target AU - AU we want to detect.

Classifiers generalisability

Infrequency of occurrence of some AUs

Highly expensive manual AU annotation

Large impact of nuisance factors, e.g. identity

Problem

Proposed Solution

Train person-specific AU models

Learn all AU models jointly (Multi-Task)

Improve detection of rare AUs with frequent

Our Approach Overview

Latent Task Structure

Let matrix \( W \) hold \( T \) tasks that lie on a common linear subspace:

\[
W = LS,
\]

where \( L \) is the subspace generator and \( S \) is the matrix of linear coefficients.

Matrices \( L \) and \( S \) are learnt using a \( l_1 \)-based sparse dictionary learning approach.

Latent Tasks Structure Transfer

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

RLTS - Regularised Latent Task Structure

OverView

Domain

Example facial Action Units (AUs)

Problem

Classifiers generalisability

Proposed Solution

Train person-specific AU models

Learn all AU models jointly (Multi-Task)

Improve detection of rare AUs with frequent

Main Contribution: Latent Tasks Structure Transfer

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

\[
s_t = \arg \min_s E(Z,L,s) + \lambda \|s\|_1 + R(s),
\]

where

\[
R(s) = \tau \|s_{target} - s_{ref}\|_2^2,
\]

that vanishes when \( AU_{target} = AU_{ref} \).

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

DISFA dataset

McMaster dataset

This work was funded by the European Union’s Horizon 2020 research and innovation programme under grant agreement No 645378. 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.