

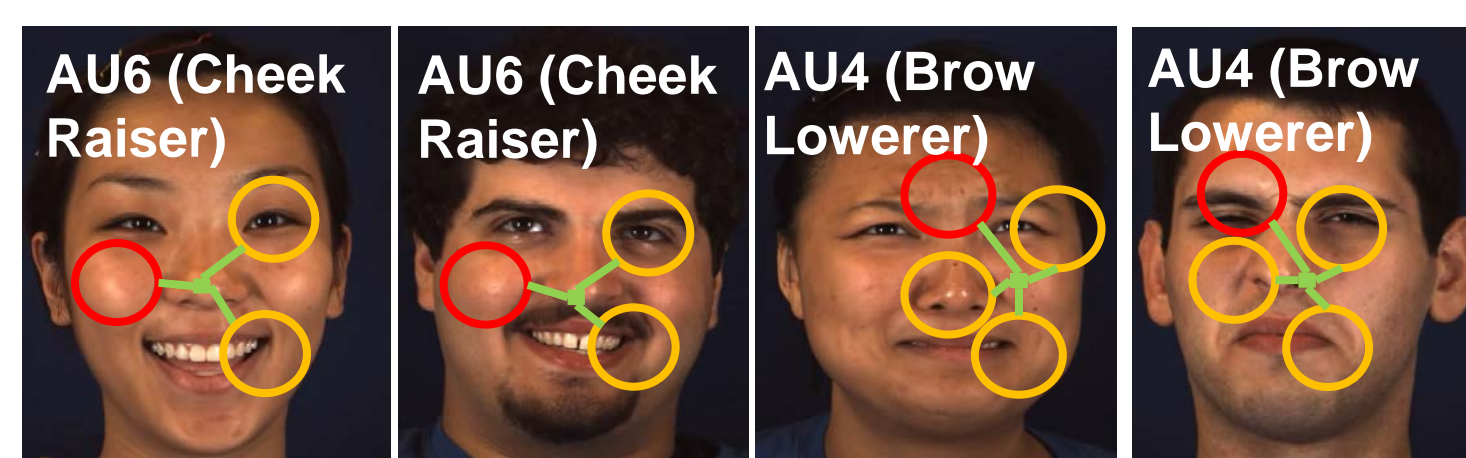
Motivation

What is facial action unit?

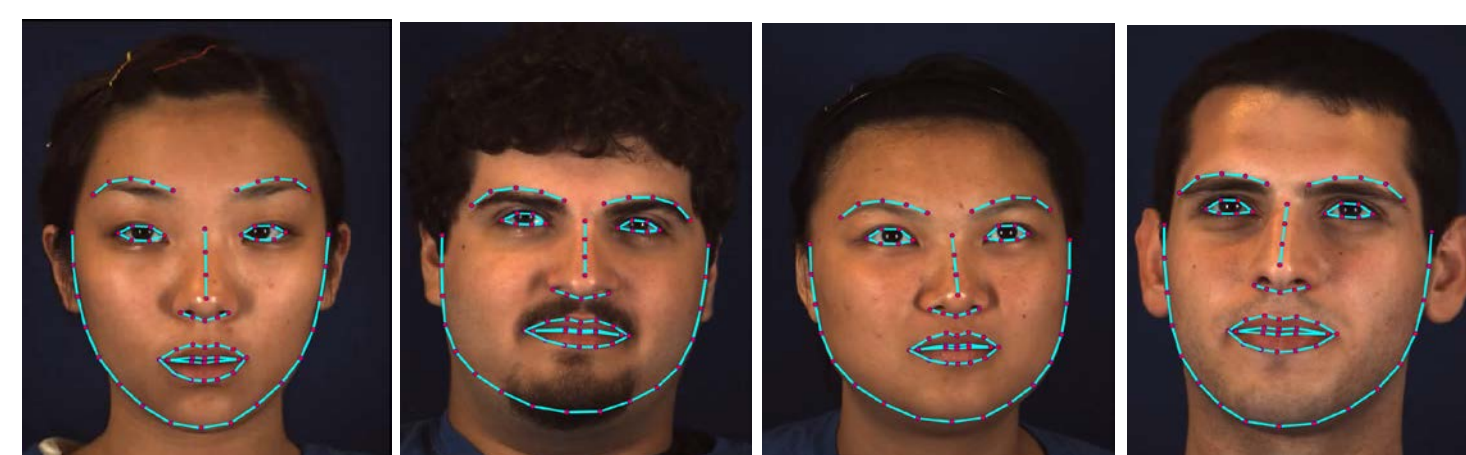
Facial action units refer to a set of facial muscle movements coded by their appearance on the face, which can be used for coding nearly any anatomically possible facial expression.



Taking into account the relationship of multiple related face regions can provide better robustness than using individual single local regions separately.

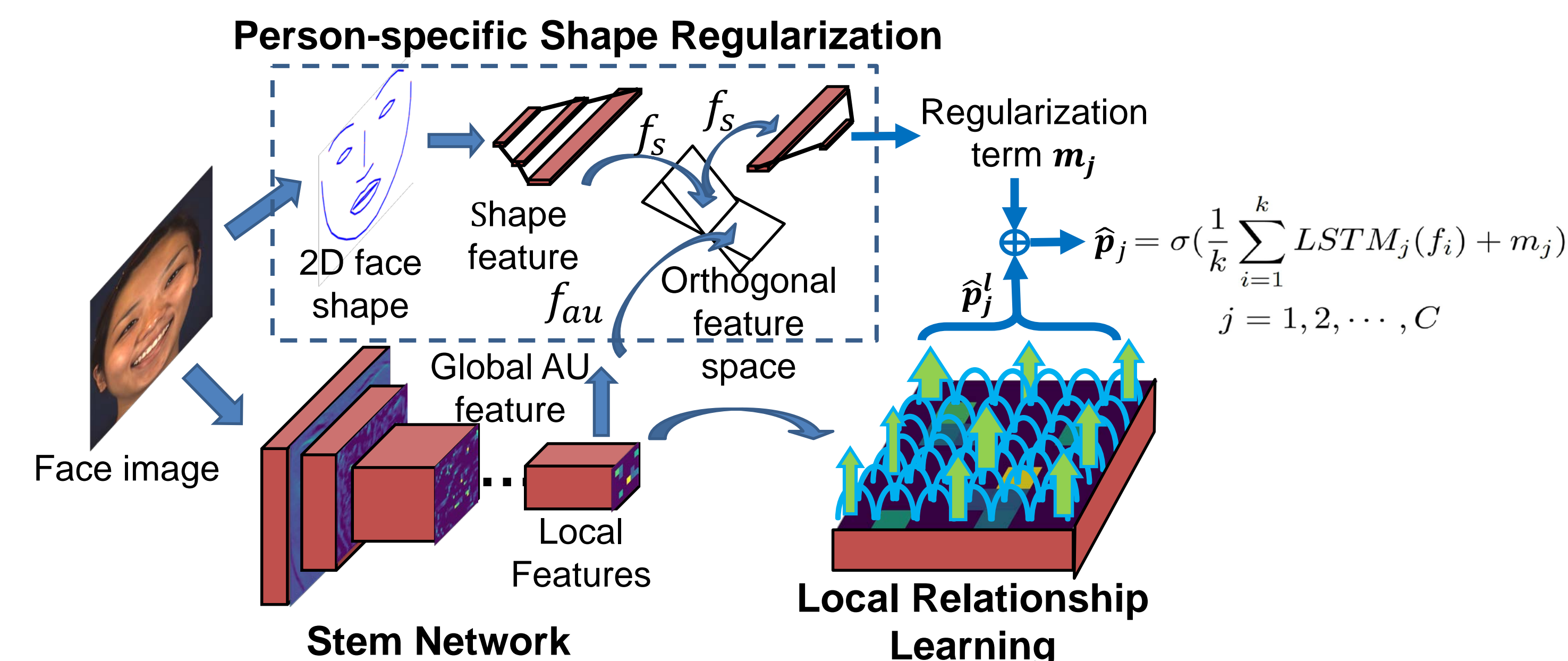


The influence of person-specific shape information can be reduced by regularization based on facial shape during feature learning.



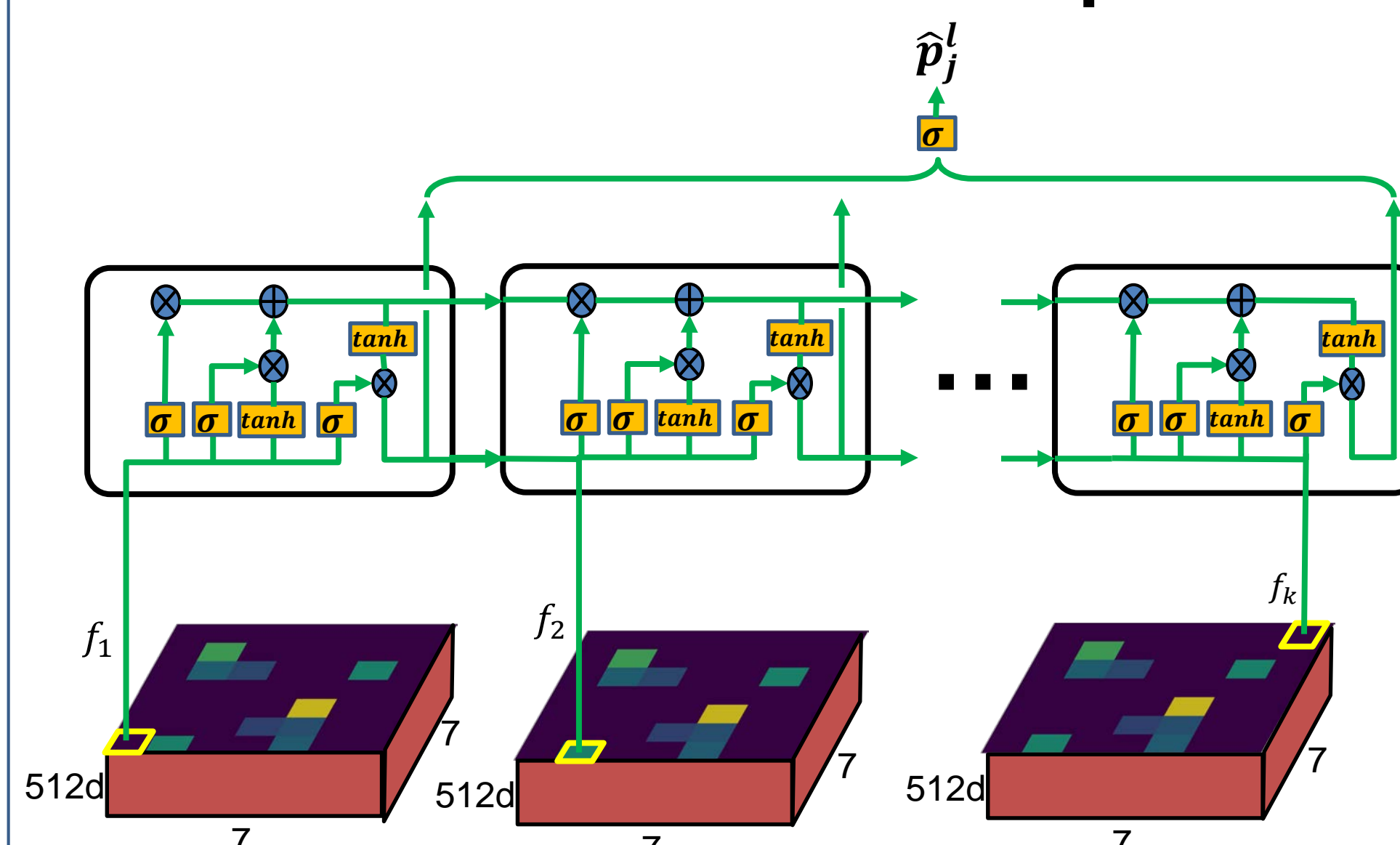
Overview

LP-Net for AU detection consists of a stem network, a local relationship learning module to model the related facial regions and a person-specific shape regularization module to reduce the shape bias.



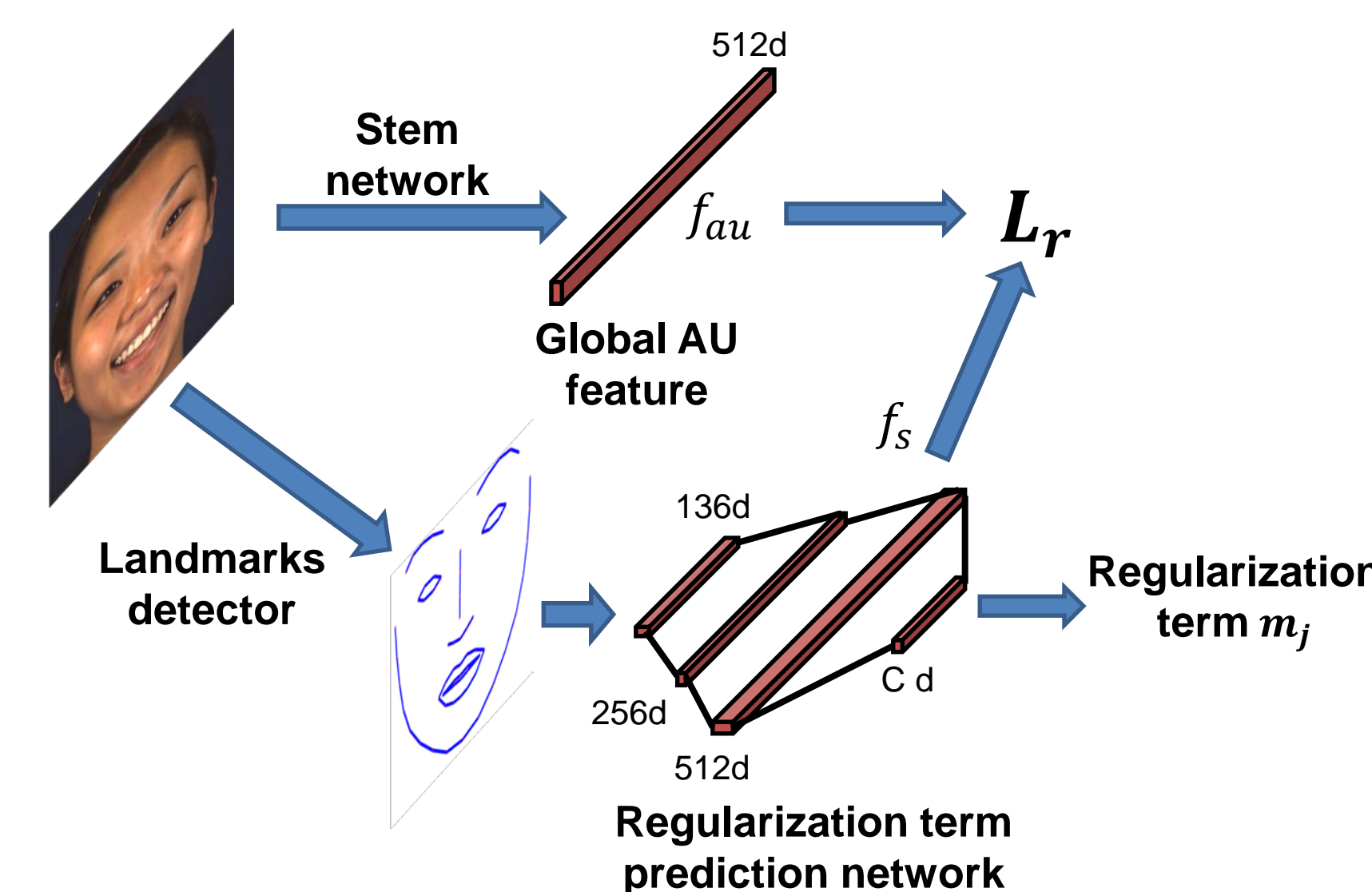
Detailed Structure

L-Net for Local Relationship Learning



- The output feature maps from the Stem-Net are used as the local features.
- LSTM is used for relationship learning.

P-Net for Person-specific Shape Regularization



- Facial landmark features are used for regularization term prediction.
- The landmark features are projected to an AU-independent feature space by using loss function:

$$L_r = |f_{au} \cdot f_s|$$

Loss functions

Loss function for AU prediction: weighted binary cross-entropy.

$$L_{au} = -\frac{1}{C} \sum_{j=1}^C w_c [p_j \log \hat{p}_j + (1 - p_j) \log(1 - \hat{p}_j)]$$

Overall loss function

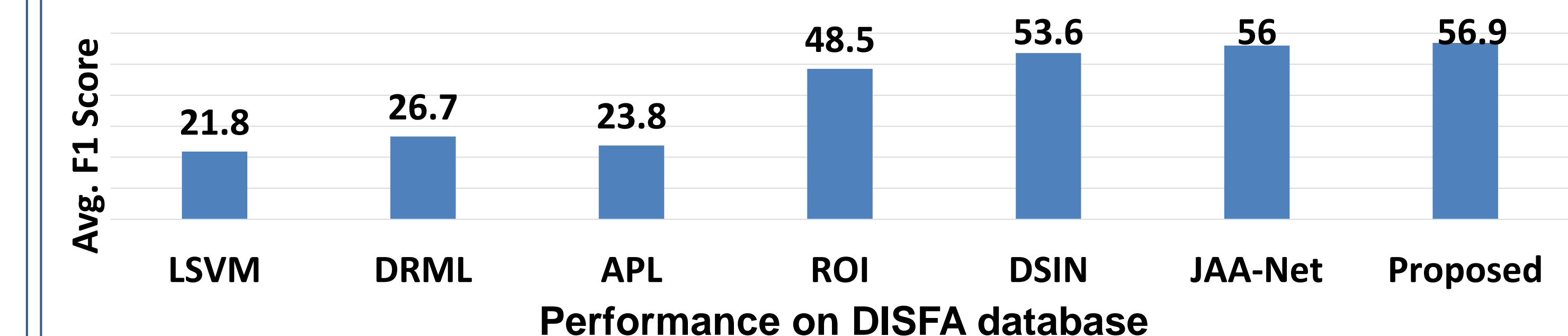
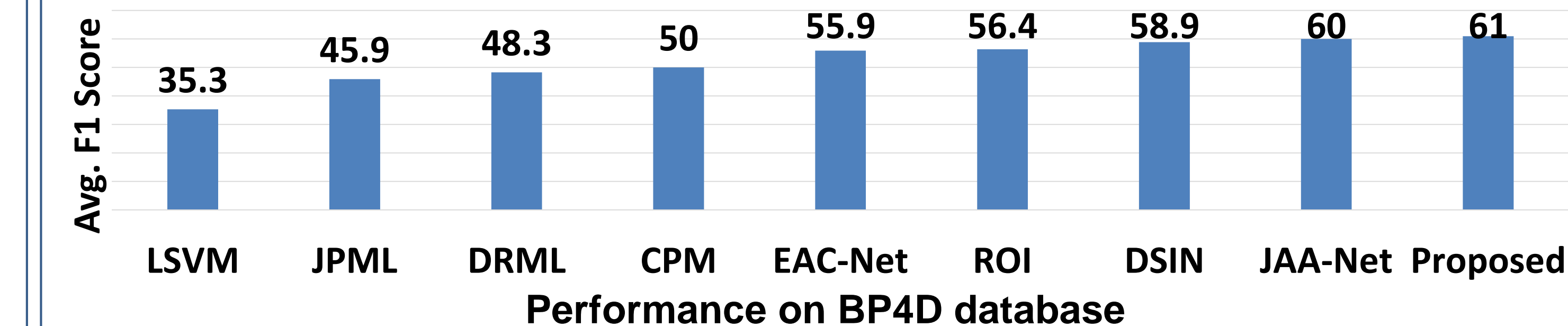
$$L_{all} = L_{au} + \lambda L_r$$

Experimental Results

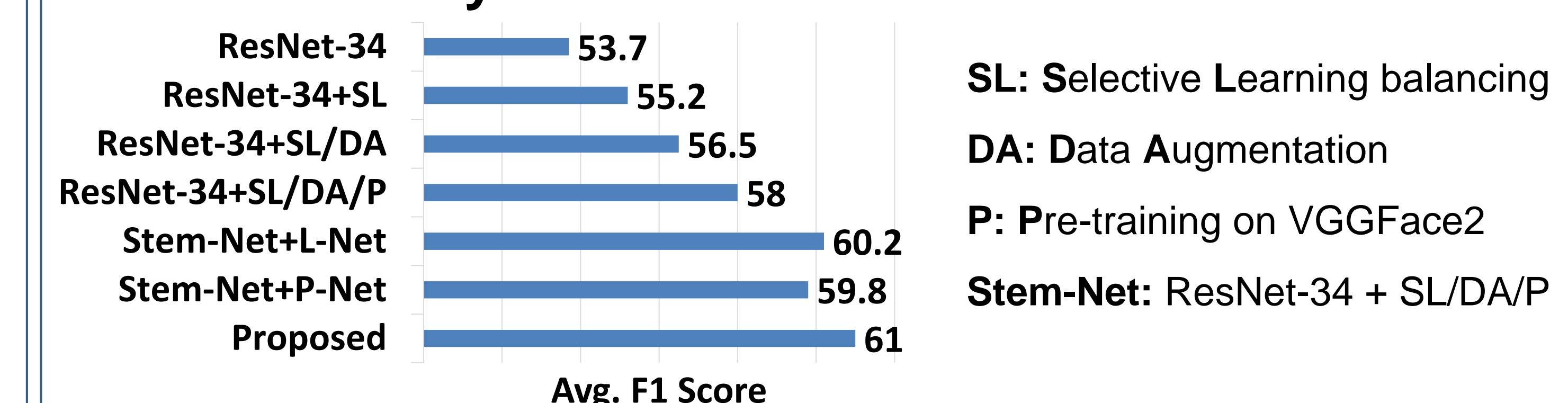
Databases

	No. Subj.	No. Vids.	No. Imgs.	No. AU	Protocol
BP4D	41	324	~140,000	12	three-fold
DISFA	27	27	~130,000	8	Fine-tune

Results



Ablation Study on BP4D Database



Qualitative Analysis

