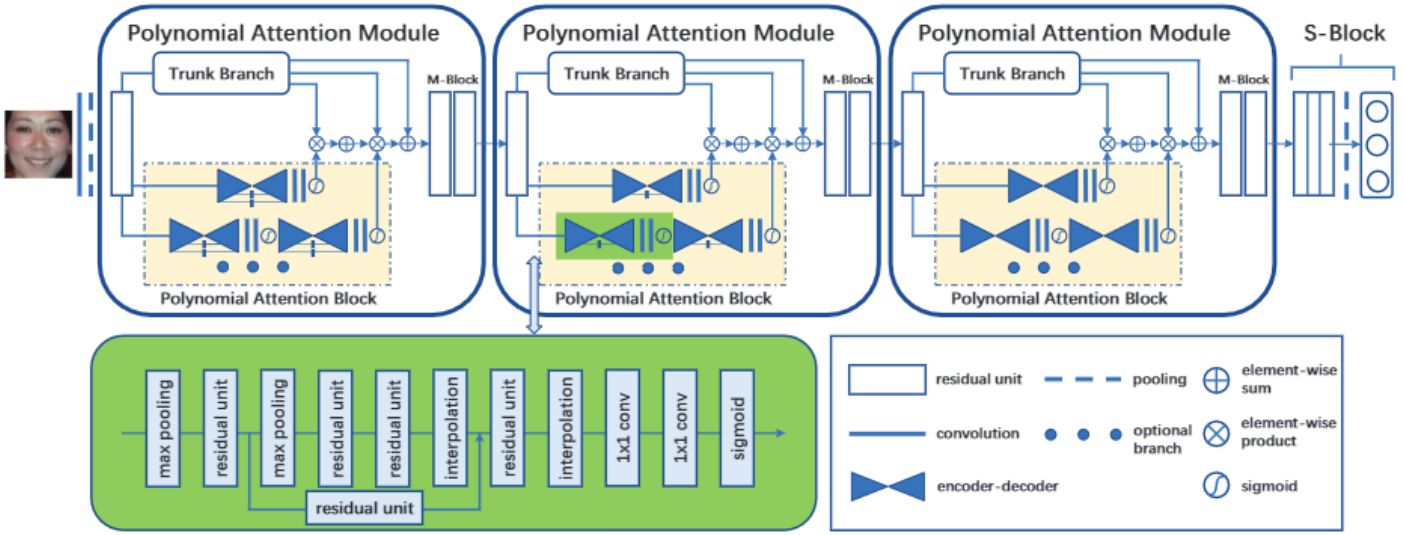
Polynomial Stacked-attention Network for Nationality Classification

Kunyan LI, Jie ZHANG, Shiguang SHAN

Frontiers of Computer Science, DOI: 10.1007/s11704-021-0455-0

Problems & Ideas

- Problems of Nationality Classification
  - It is hard to tell tiny differences on faces for countries on the same continent, such as East Asia
  - Few of methods consider explicitly excavating attention regions to tell these differences on faces
- Ideas: Propose a Polynomial Stacked-attention Network
  - Simultaneously learn features and attentions for nationality classification
  - explore more discriminative regions from the attention maps



Main Contributions

- Comparison on the Asian Face Dataset
  - Our method significantly outperforms Wang et al.[1] who employ Resnet50 to estimate nationality, with an improvement up to 6.2%
  - Our method also achieves better results than Attention-56[2], which demonstrates the effectiveness of carefully designing polynomial attention modules again

Table 1 The top-1 accuracy on the Asian Face Dataset.

Method	Wang et al. [1]	Att-56 [2]	Ours	Human
Acc.	0.750	0.776	0.812	0.389

[1] Yu W, Yang F, Haofu L, Jiebo L, and Xiangyang X, “Do they all look the same? deciphering chinese, japanese and koreans by fine-grained deep learning,” in IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), 2018, pp. 39–44

[2] Fei W, Mengqing J, Chen Q, Shuo Y, Cheng L, Honggang Z, Xiaogang W, and Xiaoou T, “Residual attention network for image classification,” in IEEE Conference on Computer Vision and Pattern Recogni-tion (CVPR), 2017, pp. 3156–3164.