Multi-View Facial Expression Recognition with AU-Guided Heatmaps

Hong-Quan Do[1,2], Hoang Van Thanh[3], Tu Minh Phuong[2,*]

[1] *FPT University, Hanoi, Vietnam*

[2] *Posts and Telecommunications Institute of Technology, Hanoi, Vietnam*

[3] *Quang Binh University, Quang Tri, Vietnam*

quandh13@fe.edu.vn, thanhhv@qbu.edu.vn, phuongtm@ptit.edu.vn

***Abstract*—Facial Expression Recognition (FER) is fundamental to affective computing and human-computer interaction, enabling systems to infer emotional states from facial cues. Current deep learning approaches rely solely on raw facial images and standard augmentation but lack explicit incorporation of facial anatomical knowledge. This forces models to discover expression-discriminative regions from scratch rather than leveraging established facial behavior research. To address this gap, we propose MV-FER, a multi-view framework that integrates spatial priors from the Facial Action Coding System (FACS). MV-FER employs three data representations: original facial images, horizontally flipped images with random erasing for view consistency, and AU-guided heatmaps generated by mapping landmarks to established Action Unit areas providing explicit spatial guidance toward expression-relevant facial regions. The three views are unified within a training paradigm that encourages learning of discriminative features while maintaining robustness against pose variations and partial occlusions. Benchmark experiments demonstrate that MV-FER achieves enhanced recognition accuracy over state-of-the-art approaches with similar architectures. These results confirm the effectiveness of our multi-view strategy, with progressive accuracy gains as augmented and AU-guided heatmap views are incorporated.**

***Keyword*— Action Unit, Facial Action Coding System, Facial Expression Recognition, Facial Landmark, Heatmap, Multi View**

Hong-Quan Do is a lecturer at FPT University, Hanoi, and a PhD candidate at the Posts and Telecommunications Institute of Technology. His research focuses on clustering, computer vision, face analysis, and recommender systems. He is also actively involved in various e-Government projects and e-Commerce recommendation applications across both the public and private sectors.

Hoang Van Thanh is a researcher at Quang Binh University. He received a Ph.D. degree in Computer Engineering from University of Ulsan, Republic of Korea, in 2020. His research interests include pattern recognition, machine learning, natural language processing, and computer vision.

Tu Minh Phuong is professor of computer science and Chairman of University Council at Posts and Telecommunications Institute of Technology. His research interests include machine learning, recommender systems, natural language processing, and computer vision. He received Ph.D. degree in control in technical systems from the National Academy of Sciences, Uzbekistant, in 1995.