[1]Mehrabian A,Ferris SR.Inference of attitudes from nonverbal communication in two channels.[J].Journal of consulting psychology,1967,31(3): 248.

[2]Zhao Y, Xu J. A convolutional neural network for compound micro-expression recognition[J].

Sensors, 2019,19(24): 5553.

[3] Du S,Martinez A M.Compound facial expressions of emotion: from basic research to clinical

applications[J]. Dialogues in clinical neuroscience,2015,17(4):443.

[5]Hickson S, Dufour N, Sud A, Kwatra V, Essa I. Eyemotion: Classifying facial expressions in VR

using eye-racking cameras[A].l/ 2019 IEEE Winter Conference on Applications of ComputerVision (WACV)[C].2019: 1626-1635.

[4] Chen C H, Lee I J, Lin L Y.Augmented reality-based self-facial modeling to promote the emc

tional expression and social skills of adolescents with autism spectrum disorders[J]. Research indevelopmental disabilities, 2015,36: 396-403.

[6] Girard J M, Cohn J F, Jeni L A, Lucey S, De la Torre F.How much training data for facial action

unit detection?{[A]. l/ 2015 11th IEEE International Conference and Workshops on AutomaticFace and Gesture Recognition (FG)[C].2015:1-8.

[7]Friesen W V, Ekman P, others. EMFACS-7: Emotional facial action coding system[J]. Unpublished manuscript,University of California at San Francisco, 1983,2(36): 1.

[8] 何良华. 人脸表情识别中若干关键技术的研究[D]. 东南大学, 2005.

[9] 周书仁, 梁昔明, 朱灿,等. 基于ICA与HMM的表情识别[J]. 中国图象图形学报, 2008, 13(12):2321-2328.

[10] 周书仁. 人脸表情识别算法分析与研究[D]. 中南大学, 2009.

[11] 应自炉, 唐京海, 李景文,等. 支持向量鉴别分析及在人脸表情识别中的应用[J]. 电子学报, 2008, 36(4):725-730.

[12] Khorasani K. Facial expression recognition using constructiveneural networks[C]// Signal Processing, Sensor Fusion, andTarget Recognition X. Signal Processing, Sensor Fusion, andTarget Recognition X, 2001:1588 - 1595.

[13] Kyperountas M, Tefas A, Pitas I. Salient feature and reliableclassifier selection for facial expression classification[J]. PatternRecognition, 2010, 43(3):972-986.

[14] Zheng W, Zhou X, Zou C, et al. Facial expression recognitionusing kernel canonical correlation analysis (KCCA).[J]. IEEETransactions on Neural Networks, 2006, 17(1):233.

[15] 付晓峰. 基于二元模式的人脸识别与表情识别研究[D]. 浙江大学,2008.

[16] Yacoob Y, Davis L S. Recognizing Human Facial ExpressionsFrom Long Image Sequences Using Optical Flow[J]. PatternAnalysis & Machine Intelligence IEEE Transactions on, 1996,18(6):636-642.

[17] Tsalakanidou F, Malassiotis S. Real-time 2D+3D facial actionand expression recognition[J]. Pattern Recognition, 2010,43(5):1763-1775.

[18] Wang J, Yin L. Static topographic modeling for facialexpression recognition and analysis[J]. Computer Vision &Image Understanding, 2007, 108(1):19-34.

[19] Sung J, Kim D. Pose-Robust Facial Expression RecognitionUsing View-Based 2D $+$ 3D AAM[J]. IEEE Transactions onSystems, Man, and Cybernetics - Part A: Systems and Humans,2008, 38(4):852-866.

[20] Kotsia I, Pitas I. Facial Expression Recognition in ImageSequences Using Geometric Deformation Features and SupportVector Machines[J]. IEEE Transactions on Image Processing APublication of the IEEE Signal Processing Society, 2007,16(1):172.

[21]K. He,X. Zhang,S. Ren,and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Conference onComputer Vision and Pattern Recognition, 2016, pp.770-778.

[22] Y.Tang, "Deep learning using linear support vector machines," arXiv preprint arXio:1306.0239, 2013.

[23]S. Li and W. Deng, "Reliable crowdsourcing and deep locality preserving learning for unconstrained facial expression recogni-tion,"IEEE Transactions on Image Processing, vol.28, no. 1, pp.356370,2018.

[24] K. Wang, X.Peng,J. Yang, D. Meng, and Y.Qiao,"Region attention networks for pose and occlusion robust facial expression recogni-tion,"IEEE Transactions on Image Processing, vol.29, pp. 4057-4069,2020.

[25]Y.Li, J.Zeng, S. Shan,and X. Chen, "Occlusion aware facial expression recognition using cnn with attention mechanism,"IEEETransactions on Image Processing, vol.28, no.5, pp.2439-2450,2018.

[26] Y. Fan,V. Li, and J. C. Lam,"Facial expression recognition with deeply-supervised attention network," IEEE Transactions on Affec-tive Computing, 2020, doi: 10.1109/TAFFC.2020.2988264.

[27]T.Xu, J. White, S. Kalkan, and H. Gunes, "Investigating bias and

fairness in facial expression recognition," in European Conference onComputer Vision, 2020, pp.506-523.

[28] K. Wang, X. Peng, J. Yang, S. Lu, and Y.Qiao, "Suppressing uncer-

tainties for large-scale facial expression recognition," in Proceedingsof the IEEE Conference on Computer Vision and Pattern Recognition,2020, pp.6897-6906.

[29]A. Vaswani,N. Shazeer, N. Parmar, J. Uszkoreit,L. Jones,A. N.Gomez,L. Kaiser, and I. Polosukhin, " Attention is all you need,"arXiv preprint arXiv:1706.03762,2017.

[30]'A. Dosovitskiy, L. Beyer,A. Kolesnikov, D. Weissenborn, X.Zhai, T. Unterthiner,M. Dehghani,M. Minderer,G. Heigold, S.Gellyet al., "An image is worth 16x16 words: Transformers for imagerecognition at scale," arXio preprint arXiv:2010.11929,2020.

[31] Aouayeb，M.， Hamidouche, W.，Soladie,C., Kpalma，K., Seguier，R.: Learning

vision transformer with squeeze and excitation for facial expression recognition.arXiv preprint arXiv:2107.03107 (2021)4

[32] A. Dosovitskiy,L. Beyer,A. Kolesnikov, D. Weissenborn, X.Zhai,T.Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S.Gellyet al., "An image is worth 16x16 words: Transformers for imagerecognition at scale," arXiv preprint arXiv:2010.11929,2020.

[33] W. Wang, E.Xie, X. Li,D.-P. Fan, K.Song, D. Liang, T. Lu, P. Luo and L. Shao,"Pyramid vision transformer: A versatile back-bone for dense prediction without convolutions," arXio preprintarXio:2102.12122,2021.

[34] Z.Chen,L. Xie, J. Niu, X. Liu,L. Wei, and Q.Tian, "Visformer: The vision-friendly transformer," arXio preprint arXiv:2104.12533,2021.

[35] H.Wu,B.Xiao, N.Codella, M.Liu, X. Dai, L. Yuan, and L.Zhang,"Cvt: Introducing convolutions to vision transformers," arXivpreprint arXio:2103.15808,2021.

[36] Z.Liu, S.Luo,W.Li, J.Lu,Y. Wu, C. Li, and L.Yang, "Convtrans-former: A convolutional transformer network for video framesynthesis," arXiv preprint arXiv:2011.10185,2020.

[37] Zheng C , Mendieta M , Chen C . POSTER: A Pyramid Cross-Fusion Transformer Network for Facial Expression Recognition[J]. 2022.