Mohamed A. Saleh , Alan Ting Yong , N. Marbukhari , YM.Yussoff , N. Nabila Mohamed , Ali Abd Almisreb , Habibah Hashim

Research Article

Facial Expression Recognition: A New Dataset and a Review of the Literature

Mohamed A. Saleh [a*], Alan Ting Yong [b], N. Marbukhari [c], YM.Yussoff [d], N. Nabila Mohamed [e], Ali Abd Almisreb [f], Habibah Hashim [g]

[a*,b] School of computing and creative media, UCTS, Sarawak, Malaysia
[c,d,g] Faculty of Electrical Engineering, Universiti Teknology MARA (UiTM), Malaysia
[c] UIN Maulana Malik Ibrahim, Indonesia
[e] Faculty of Engineering & IT, Mahsa University, Selangor, Malaysia
[f] Faculty of Engineering & Natural Sciences, International University of Sarajevo, Bosnia & Herzegovina

*Corresponding author: yusna233@uitm.edu.my

Abstract

One of the biggest challenges in computer vision and deep learning is recognizing the emotion based on facial expression. Besides the challenge of model development in Deep Learning (DL), the dataset is considered one of the main factors that plays a crucial role in producing highly accurate deep learning frameworks for pattern classification, features extraction, and emotion recognition from images. Despite the existence of several dedicated datasets for the FER, the desired accuracy was not being achieved due to the enormous size of the available dataset, where some of such are unfeasible for regular computers or cloud computing services. In contrast, there is a lack of the other datasets' size, generalization, and quality. The objectives of this paper are two-fold. First, reviews the existing datasets. Second, presents a significant new dataset for facial expression called Emot-FE. This dataset contains 276,305 images of facial expressions formulated in a single file in the form of xlsx format. Emot-FE dataset has been filtered, pre-processed, labeled, and classified based on the seven labels (emotion's expressions). Thus, this dataset will be the largest in such format with high-quality images. This new dataset has been evaluated using VGGNet Convolutional Neural Networks where the recognition accuracy outperforms the previously achieved accuracies of the previous studies. This dataset is available for distribution to the researchers..

Keywords: face dataset; emotion recognition; CNN, Pattern classification, facial expression, Deep Learning

1. Introduction

Of late, automatic recognition for the expression of human emotion and feelings has gained momentum, especially with the concurrent revolution of Artificial Intelligence (AI). The trainable, feasible, and considerable dataset is an indispensable necessity for achieving high accuracy emotion recognition. Many studies have been conducted for collecting and annotating facial expression datasets to be utilized in CNN models. Thus, several facial expression datasets have been published, such as [1]–[7]. This study will focus on two goals: reviewing the currently used datasets and presenting a new dataset that can be utilized practically in the following proposed DL models.

The rest of this paper is arranged as follows. Existing facial expressions datasets were reviewed in Section II. An overview of a new dataset namely Emot-FE is presented in Section III. The dataset collection details of Emot-FE and the conclusion of this paper are discussed in Section IV and Section V, respectively..

2. Existing Databases

In this section, the publicly available datasets of emotional facial expressions are briefly reviewed. Therefore, the focus will be on the most recent and popular datasets that have been released recently and utilized effectively in the deep learning neural network training.

Before 2013 the popular dedicated datasets for recognizing emotional facial expressions were based on handcrafted features, while since 2013, the DL has been utilized instead, especially CNN models [8], [9].

JAFFE dataset [3] is one of the traditional facial emotion datasets that include 219 gray-scale images of facial expressions for Japanese ladies. Images were labeled into seven facial emotions (angry, sad, neutral, happy, surprise, fear, and disgust) for ten Japanese females, each of which is 256 by 256 pixels.

KDEF dataset [10] was collected by the Karolinska Institute and included 4,900 facial expression images. It was collected from 70 actors and labeled into seven basic emotions, similar to JAFEE dataset. The images were taken in five different angles (straight, full right/ left, and half right/ left). The size of KDEF images is 562 × 762 pixels.

M&M Initiative (MMI) dataset [2] [11] was provided as high-resolution images and video sequences of facial expressions and included more than 3,600 samples. Real color (24 bits) images were collected from 25 research staff members and students. MMI dataset was manually labeled into six basic facial expressions (Anger, Surprise, Disgust, Happy, Sadness, and Fear). The size of each image in MMI dataset is 640 × 480 pixels.

BU-3DFE database [12] was collected by Binghamton and contained 2,500 3D facial expression models. Fifty-six females and 44 males participated as subjects. BU-3DFE database labeled into seven (six basic expressions and normal) facial emotional expressions. The dataset's images are in colored 3D geometrical shapes with the size of 1040 × 1329 pixels.

 Extended CohnKanade (CK+) database [1] is considered one of the most utilized datasets. It contains 593 images of 123 subjects in 18 - 50 years, where 69% of them are female. Each images' sequences were recorded as frontal views and 30_degree views and formulated into 8-bit gray-scale arrays of size 640x490 or 640x480 pixels. Every image sequence contains 10 – 60 frames, where that sequence was labeled into seven (including contempt and excluding natural) emotional facial expressions.

TFD dataset, known as Toronto Face Database [13], contains 4,178 labeled images from the original 112,234 images. The annotated images were labeled like BU-3DFE dataset into six basic facial emotional expressions plus natural. The faces in the labeled images were resized to 48 × 48 pixels.

CMU-Multi-PIE dataset [14] was collected by Carnegie Mellon University (CMU) as a multi Profile Index of Emotion (PIE). The dataset comprises 755,370 facial images of the 337 participant subjects from different ethnicities, European American, 60%, Asian 35%, African American 3%, and 2%. The average age of the participated subjects was 27.9 years. The number of males was 235, while the females' number was 102. All of the annotated images were taken using 15 different viewpoints and 19 illumination effects. All CMU-Multi-PIE facial images were labeled into six-basic expressions: neutral, scream, surprise, smile, squint, and disgust. CMU-Multi-PIE dataset generally is utilized for multi-view facial expression analysis.

RaFD dataset [15] contains 8,040 Caucasian facial images of 49 participant subjects. Thirty-nine were adults (20 male and 19 female) and ten children (six females and four males). The images were obtained using five cameras from five different directions (at angles, 00, 450, 900, 1350, and 1800) and three gaze directions. The images were labeled into eight facial expressions (neutral, sadness, fear, contempt, surprise, disgust, happiness, and anger). Each image in RaFD dataset was resized to 1024 × 681 pixels.

The Oulu-CASIA database[16] included 2,880 images collected for 80 people (23-58 years old and 73.8% male). It was labeled into six basic facial emotional expressions: happiness, anger, sadness, disgust, surprise, and fear. All facial images in Oulu-CASIA dataset were captured in dark, weak, and normal illumination conditions. The image's resolution of Oulu-CASIA database is 320 × 240 pixels.

SFEW dataset [6] includes 1,635 images collected as static frames from the video. All images are for natural expressions in the wild environment. The images of SFEW dataset were annotated into seven emotional expressions, anger, disgust, fear, sadness, surprise, neutral, and happiness. The images in SFEW dataset are in

Mohamed A. Saleh , Alan Ting Yong , N. Marbukhari , YM.Yussoff , N. Nabila Mohamed , Ali Abd Almisreb , Habibah Hashim

RGB color mode and sizes 720 × 576 pixels.

DEAP dataset [17] was collected for comparative emotion analysis using physiological signals. Participated subjects' (n=32) reactions were recorded while watching 40 music videos, one-minute lengthy excerpts. In this dataset, the emotional states were annotated according to the arousal and valence dimensional plane. In other meaning, the range of emotional states of inactive to active (uninterested, bored, alert, excited) is described by arousal. In contrast, the range of unpleasant and pleasant (sad, stressed, happy, elated) is described by valence.

FER2013 dataset [4] was provided by Kaggle, contains 37,887 gray-scale images, and collected using Google image search API (Application Program Interface) for 184 facial emotional expressions' keywords. All images in FER2013 dataset were subjected to pre-process treatment such as cropping, resizing to 48×48 pixels, and gray scaling. Images of FER2013 were labeled as the basic six emotional expressions (happy, anger, disgust, sad, surprise, and fear) plus neutral.

DISFA dataset [18] was provided by the University of Denver, and its total collected images were 130,000. The images were recorded as video frames of 27 participated subjects (males and females, ages between 18 to 35 years) while watching videos. DISFA's images were annotated into 12 Action Units (AU) where all of that images were resized into 1024 × 768 pixels.

AM-FED [19] dataset comprises 168,359 facial frames extracted from 242 videos. All frames (images) in this dataset were labeled into 14 AUs. The frame's size in AM-FED dataset is 1024 × 768 pixels.

Compound Emotion (CE) dataset [20] contains 5,060 facial images of 230 subjects (100 males, 130 females, with 23 years as the average age). The images of the CE dataset were categorized into 22 basic and compound emotion modes. Different races and ethnicities from different continents were included. Facial images without glasses and non-covered foreheads were also included for male subjects. Only the shaved faces images were included. The facial image's size in CE dataset is 3000 × 4000 pixels.

BP4D-Spontaneous dataset [21] includes spontaneous facial expressions in the form of 3D video for 41 participated subjects (18 male, 23 female, and age 18 to 29 years). The subjects were from different ethnicities such as American, Hispanic, European, and Asian. -Americans. The images of BP4D-Spontaneous dataset were sized into 1040 × 1329 pixels.

EmotioNet dataset [7] contains one million facial emotional expression images obtained from the internet. EmotioNet is one of the most significant scale datasets, 5% (50,000) of that images were labeled manually while the rest (950,000) were annotated automatically. In EmotioNet, the images were categorized into 16 possible emotional expressions. Seven are basic emotional models (happy, sad, disgusted, surprised, angry, and fearful, and appalled). The others are the compound emotions (angrily-disgusted, sadly-disgusted, angrily-surprised, fearfully-surprised, fearfully-angry, happily-disgusted, sadly-angry, happily-surprised, and awed).

Aff-Wild Database [22] was dedicated to assisting the continuous effect of the valence-arousal-space in-the-wild. Aff-Wild contains more than 2,000 images, and 500 videos were collected from the spontaneous emotional response of 500 subjects while watching an emotional video, reacting to jokes, and doing some activities.

The FER-Wild dataset [23] includes 24,000 images of facial emotional expressions. All of the FER-Wild images were collected using emotional keywords from the internet. The images were labeled into seven emotional expressions (neutral, surprise, anger, fear, sadness, happiness, and disgust).

AffectNet dataset [24] comprises more than one million facial emotional expressions. Collecting images was using the search on the internet for 1,250 emotion keywords in six different languages. AffectNet has considered the most extensive dataset covering facial emotional expressions in both models (categorical and dimensional). 450,000 images were annotated manually into eleven discrete categories, eight emotions (Neutral, Sadness, Happiness, Surprised, Anger, Fear, Contempt, and Disgusted), plus No-object, Uncertain, and No-face. Images of the AffectNet's were sized into 425 × 425 with an STD of 349 × 349 pixels.

Extended Yale B face (B+) dataset [25] contains 16,128 facial images of ten subjects collected under 64 different illumination conditions for every post (total nine posts) to form 576 viewing conditions. B+ dataset images' size is 320 × 243 pixels.

Acted Facial Expressions in the Wild (AFEW 7) dataset [26]: was provided after its two predecessor versions EmotiW 2013 and SFEW. These two datasets were dedicated to recognizing the emotions in the Wild Challenge [26]. AFEW 7 comprises 1,809 videos collected from different movies for various conditions of spontaneous expressions, head poses, occlusions, and illuminations. This dataset's expressions were annotated into seven facial emotional expressions (neutral, happy, anger, surprise, disgust, fear, and sad). In the AFEW 7

dataset, the samples were grouped into three independent sets: 773 Training samples, 383 Validation samples, and 653 Testing samples.

RAF-DB [27] [28] dataset is a Real-world facial expression database that includes 29,672 real-world facial images collected from different internet resources. the images of RAF-DB dataset were manually labeled into seven basic facial expressions and 11 compound facial expressions. 15,339 images from that dataset were divided into 12,271 images for training and 3,068 images for testing.

The Expression in-the-Wild Database (ExpW) [29] contains 91,793 images labeled into seven basic facial emotional expressions (neutral, happy, disgust, surprise, angry, sad, and fear).

In one of the recent studies [30], LIRIS-CSE dataset was presented as a new dataset for children's spontaneous facial emotional expressions. LIRIS-CSE comprises 208 short video clips (5 seconds is the average length of every video), each video containing one specific expression. LIRIS-CSE was labeled into six basic spontaneous expressions for 12 children between 6 and 12 years old (7.3 years is their mean age) - five male and seven female.

EMOTIC dataset [31] presented in 2020 as 23,571 images for 34,320 people (males, females, children, teenagers, and adults) labeled based on their observed emotions. EMOTIC dataset was annotated into two facial emotional representations as 26 discrete categories and three continuous dimensions (Valence, Arousal, and Dominance) [31].

All these datasets of emotional facial expressions have been summarized and sorted chronically in Table 1 with providing the name of the dataset, its reference, year of release, number of samples (image or video), number of participated subjects, number of labels (facial emotional expressions), images'/frames' size, and the URL link for every dataset.

Table 1 A summary of the facial emotional expression datasets

Database	Year	Samples	Subject	Expression distribution	Image's size (pixels)	Access
JAFFE [3]	1998	219 images	10	7 facial expressions	562×762	http://www.kasrl.org/jaffe.html
KDEF[10]	1998	4,900 images	70	7 facial expressions	256×256	http://www.emotionlab.se/kdef/
MMI [2] [11]	2005/ 2010	740 images and 2,900 videos	25	6 facial expressions	640×480	https://mmifacedb.eu/
BU-3DFE [12]	2006	2,500 images	100	7 facial expressions	1040×1329	http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE Analysis.html
CK+ [1]	2010	593 image sequences	123	8 facial expressions	640×490	http://www.consortium.ri.cmu.edu/ckagree/
TFD [13]	2010	112,234 images	N/A	7 facial expressions	48×48	josh@mplab.ucsd.edu
Multi-PIE [14]	2010	755,370 images	337	7 facial expressions	N/A	http://www.flintbox.com/public/project/4742/
RaFD [15]	2010	8,040 images	39	8 facial expressions	1024×681	http://www.socsci.ru.nl:8180/RaFD2/RaFD
Oulu-CASIA [16]	2011	2,880 image sequences	80	6 basic expressions	320×240	http://www.cse.oulu.fi/CMV/Downloads/Oulu-CASIA
SFEW 2.0 [6]	2015	1,635 images	N/A	7 facial expressions	720×576	https://cs.anu.edu.au/few/emotiw2015.html
DEAP [17]	2012	40 one-minute long videos	32	4 facial expressions	N/A	http://www.eecs.qmul.ac.uk/mmv/datasets/deap/
FER-2013 [4]	2013	35,887 images	N/A	7 facial expressions	48×48	https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge
DISFA [18]	2013	130,000 stereo video frames	27	12 AUs	1024×68	https://computervisiononline.com/dataset/1105138646
AM-FED [19]	2013	168,359 facial video frames	N/A	14 AUs	1024×768	http://www.affectiva.com/facial-expression-dataset-am-fed/
CE [20]	2014	5,060 images	230	6 basic expressions	3000×4000	http://cbcsl.ece.ohio-state.edu/dbform_compound.html
BP4D-Spontanous [21]	2014	3D video database	41	N/A	1040×1329	http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE_Analysis.html
EmotioNet [7]	2016	1M images	N/A	7 basic expressions and 9 compound expressions	N/A	http://cbcsl.ece.ohio-state.edu/dbform emotionet.html

Mohamed A. Saleh , Alan Ting Yong , N. Marbukhari , YM.Yussoff , N. Nabila Mohamed , Ali Abd Almisreb , Habibah Hashim

Aff-Wild [22]	2016	500 videos from YouTube, 2,000 images	and500	Valence and arousal (continuous)	N/A	https://ibug.doc.ic.ac.uk/resources/first-affect-wild-challenge/
FER-Wild [23]	2016	24,000 images	24,000	7 facial expressions	N/A	https://cs.anu.edu.au/few/Group.htm
AffectNet[24]	2017	450,000 images	N/A	7 facial expressions	N/A	http://mohammadmahoor.com/databases-codes/
B+ [25]	2017	16,128 images	28	N/A	320×243	https://computervisiononline.com/dataset/1105138686
AFEW 7.0 [26]	2017	1,809 videos	N/A	7 facial expressions	N/A	https://sites.google.com/site/emotiwchallenge/
RAF-DB [27] [28]	2017/ 2018	29,672 images	N/A	7 facial expressions and 11 compound expressions	N/A	http://www.whdeng.cn/RAF/model1.html
ExpW [29]	2018	91,793 images	N/A	7 facial expressions	N/A	http://mmlab.ie.cuhk.edu.hk/projects/socialrelation/index.html
LIRIS-CSE [30]	2019	208 short videos	12	6 basic spontaneous expressions	N/A	https://childrenfacialexpression.projet.liris.cnrs.fr/site/requestnew
EMOTIC [31]	2020	23,571 images	34,320	26 discrete categories, and 3 continuous dimensions	N/A	http://sunai.uoc.edu/emotic/download.html

3. An Overview of Emot-FE Dataset

Emot-FE dataset comprises 276,305 images of facial expressions of different ethnicities, ages, and gender. All images were collected in different formats (png, gif, jpg, and jpeg), different poses, different backgrounds, and different lighting levels. The facial images in Emot-FE dataset are a variation of orientation, covered and non-covered head, wearing glass and non-glass, shaved and non-shaved beards. Figure 1 shows examples of some images of Emot-FE dataset. The images were annotated into seven facial expressions (coded from 0 to 6). Table 2 shows the number of images for each of such expressions. Emot-FE dataset was divided into three sets: Training, PrivateTest for validation, and PublicTest for testing. Training part includes 221,045, PrivateTest includes 27,630, and PublicTest 27,630.
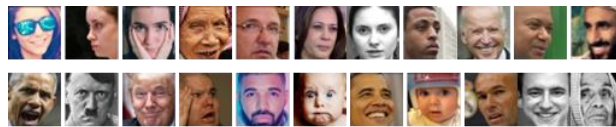

Figure 1: Sample Images of Emot-FE Dataset

Table 2 Number of images for annotated facial expressions

Code	Expression	Number of images
0	Angry	24245
1	disgust	3739
2	Scared	6048
3	Happy	131225
4	Sad	24746
5	Surprised	13620
6	Neutral	72682
Total		276305

4. Dataset Collection

Emot-FE dataset was created to be a practical facial expression dataset collected from the previous facial dataset and internet by focusing mainly on the seven basic facial expressions. The obtained images were filtered by excluding the duplicated images and the images that do not show facial expression. Using OpenCV, the collected images' faces were recognized, cropped, and resized with maintaining their aspect ratio. All cropped images have been filtered to exclude the non-facial images generated due to failing in the facial recognition process. Figure 2 shows sample examples of some pre-processed images of Emot-FE Dataset. The formatting was similar to the strategy that was proposed in [4] as follows: a) resizing all images to 48 × 48 pixels, b) converting all images into gray-scale, c) saving every image as one dimension array in a single excel file. As the results (will be elaborated in the following study) of training Emot-FE dataset on the R-CNN model, the achieved accuracy of facial emotional expressions recognition is 83.5%.

5. Dataset Format

Emot-FE dataset was designed in a way to ease its distribution, utilization, and downloading for the researcher. It was designed as a single xlsx file with a size of approximately 853MB. Thus three columns were used as the first column represents the emotion value as the emotion expression (label) of the corresponding image saved as an integer value (0, 1, 2, 3, 4, 5, 6). The second column represents the pixel's value, where every image is saved as a one-row array in a single cell. The third column represents every image's usage, where all images are split into three sets (Training, PublicTest, and PrivateTest).
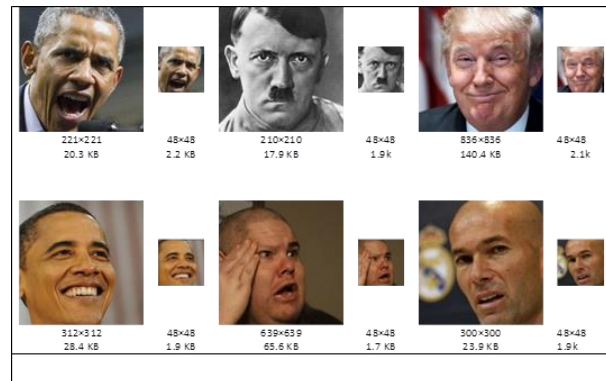
Figure 3: Some examples of original images associated with the pre-processed images of Emot-FE Dataset

6. Conclusion and future work

This paper presented the review for the state-of-the-art of facial expression datasets and presented a new dataset called Emot-FE. In Emot-FE dataset, the aim was to present an image database that includes an extensive collection of high-quality images for different formats (png, gif, jpg, and jpeg), different ethnic subjects, age, and genders. Emot-FE is considered the largest dataset in the form of a single xlsx file where it shows better emotion detection accuracy using R-CNN network. In the following study, we will show the results of this dataset's training performance on VGGNet model. Emot-FE dataset will be ready for distribution to the researchers, which can be provided via the author's email.

Acknowledgement

References

[1] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression," in IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, CVPRW 2010, 2010, pp. 94–101.

[2] M. Pantic, M. Valstar, R. Rademaker, and L. Maat, "Web-Based Database for Facial Expression Analysis," IEEE Int. Conf. Multimed. Expo, pp. 5–8, 2005.

[3] M. Lyons and S. Akamatsu, "Coding Facial Expressions with Gabor Wavelets Coding Facial Expressions with Gabor Wavelets," in third ieee international conference on automatic face and gesture recognition, 1998, no. May, pp. 2–8.

[4] I. J. Goodfellow et al., "Challenges in representation learning: A report on three machine learning contests," Neural Networks, vol. 64, pp. 59–63, 2013.

[5] A. Dhall, R. Goecke, S. Lucey, and T. Gedeon, "Collecting large, richly annotated facial-expression databases from movies," IEEE Multimed., 2012.

Mohamed A. Saleh , Alan Ting Yong , N. Marbukhari , YM.Yussoff , N. Nabila Mohamed , Ali Abd Almisreb , Habibah Hashim

[6]   A. Dhall, R. Goecke, S. Lucey, and T. Gedeon, "Static facial expression analysis in tough conditions: Data, evaluation protocol and benchmark," in Proceedings of the IEEE International Conference on Computer Vision, 2011.

[7]   J. E. Marti´nez Pe´rez and I. M. Marti´nez, "EmotioNet: An accurate, real-time algorithm for the automatic annotation of a million facial expressions in the wild," Hacienda Publica Esp., vol. 191, no. 4, pp. 73–93, 2016.

[8]   C. Mumenthaler, D. Sander, and A. S. R. Manstead, "Emotion recognition in simulated social interactions," IEEE Trans. Affect. Comput., 2020.

[9]   A. M. Proverbio, E. Camporeale, and A. Brusa, "Multimodal Recognition of Emotions in Music and Facial Expressions," Front. Hum. Neurosci., 2020.

[10]   A. Flykt, D. Lundqvist, A. Flykt, and A. Öhman, "The Karolinska directed emotional faces (KDEF)," CD ROM from Dep. Clin. Neurosci. Psychol. Sect. Karolinska Institutet, 1998.

[11]   M. Valstar and M. Pantic, "Induced Disgust, Happiness and surprise: an addition to the MMI Facial Expression Database," Proc Intl Conf Lang. Resour. Eval., pp. 65–70, 2010.

[12]   L. Yin, X. Wei, Y. Sun, J. Wang, and M. J. Rosato, "A 3D facial expression database for facial behavior research," in FGR 2006: Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition, 2006.

[13]   J. Susskind, a Anderson, and G. E. Hinton, "The Toronto face dataset," U. Toronto, Tech. Rep. UTML TR, 2010.

[14]   R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, "Multi-PIE," Image Vis. Comput., 2010.

[15]   O. Langner, R. Dotsch, G. Bijlstra, D. H. J. Wigboldus, S. T. Hawk, and A. van Knippenberg, "Presentation and validation of the radboud faces database," Cogn. Emot., 2010.

[16]   G. Zhao, X. Huang, M. Taini, S. Z. Li, and M. Pietikäinen, "Facial expression recognition from near-infrared videos," Image Vis. Comput., 2011.

[17]   S. Koelstra et al., "DEAP: A database for emotion analysis; Using physiological signals," IEEE Trans. Affect. Comput., 2012.

[18]   S. M. Mavadati, M. H. Mahoor, K. Bartlett, P. Trinh, and J. F. Cohn, "DISFA: A spontaneous facial action intensity database," IEEE Trans. Affect. Comput., 2013.

[19]   D. McDuff, R. El Kaliouby, T. Senechal, M. Amr, J. F. Cohn, and R. Picard, "Affectiva-mit facial expression dataset (AM-FED): Naturalistic and spontaneous facial expressions collected 'in-the-wild,'" in IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2013.

[20]   S. Du, Y. Tao, and A. M. Martinez, "Compound facial expressions of emotion," Proc. Natl. Acad. Sci., 2014.

[21]   X. Zhang et al., "BP4D-Spontaneous: A high-resolution spontaneous 3D dynamic facial expression database," Image Vis. Comput., 2014.

[22]   S. Zafeiriou, A. Papaioannou, I. Kotsia, M. Nicolaou, and G. Zhao, "Facial Affect 'In-the-Wild': A Survey and a New Database," in IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2016.

[23]   A. Mollahosseini, B. Hassani, M. J. Salvador, H. Abdollahi, D. Chan, and M. H. Mahoor, "Facial Expression Recognition from World Wild Web," in IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2016.

[24]   A. Mollahosseini, B. Hasani, and M. H. Mahoor, "AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild," IEEE Trans. Affect. Comput., pp. 1–18, 2017.

[25]   A. Sarafraz, "Yale Face Database B," Computer Vision Online. .

[26]   A. Dhall, J. Hoey, S. Ghosh, T. Gedeon, J. Joshi, and R. Goecke, "From individual to group-level emotion recognition: EmotiW 5.0," pp. 524–528, 2017.

[27]   S. Li, W. Deng, and J. P. Du, "Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild," in Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017.

[28] S. Li and W. Deng, "Reliable crowdsourcing and deep locality-preserving learning for unconstrained facial expression recognition," IEEE Trans. Image Process., 2019.

[29] Z. Z. • P. L. • C. C. L. • X. T. ArXiv:1609.06426v3, "From Facial Expression Recognition to Interpersonal Relation Prediction Zhanpeng," Int. J. Comput. Vis., 2018.

[30] R. A. Khan, A. Crenn, A. Meyer, and S. Bouakaz, "A novel database of children's spontaneous facial expressions (LIRIS-CSE)," Image Vis. Comput., vol. 83–84, pp. 61–69, 2019.

[31] R. Kosti, J. M. Alvarez, A. Recasens, and A. Lapedriza, "Using EMOTIC Dataset," IEEE Trans. Pattern Anal. Mach. Intell., vol. 42, no. 11, pp. 2755–2766, 2020.