A Methodological and Systematic Survey of AI-Driven Approaches and Applications for Facial Expression Recognition

M Bakiaraj[1] and B Subramani[2]
[1, 2]Department, of Computer Science,
Shri Nehru Maha Vidyalaya College of Arts and Science, Coimbatore, India

Abstract: Facial Expression Recognition (FER) is an emerging discipline that employs visual analysis with extensive learning to analyze and interpret human emotions using technology. This article provides a comprehensive and organized overview of FER, an area at the leading edge of automated vision and deep learning. FER is essential for accurately reading human emotions. This extensive study examines many sophisticated models, including Generative Adversarial Networks (GANs), Deep Belief Networks (DBN), and Convolutional Neural Networks (CNNs). The main emphasis is their capacity to reliably detect fundamental emotions such as happiness, sorrow, rage, and surprise. The research showcases substantial improvements in the accuracy and effectiveness of emotion recognition, achieved via methods such as saliency mapping and advanced feature extraction. This research emphasizes the shift towards AI-driven approaches in FER, which results from a collaborative and multidisciplinary effort including computational science, technology, and healthcare worldwide. It highlights the significant advantages of the sector in areas such as medical services, IoT, and privacy, emphasizing its contribution to enhancing user experience and engagement with intelligent systems. The poll evaluates different approaches and recognizes upcoming difficulties and directions, highlighting FER's significant contribution to comprehending and examining intricate emotional states.

Keywords:

Facial Expression, Convolutional Neural Network, Deep learning, Recurrent Neural Network, Generative Adversarial Networks

1. INTRODUCTION

Facial expressions serve as innate emotional cues, posing a challenge for robots yet readily understandable for humans. FER is widely used in several fields, such as cell phones, emotional computing, intelligent systems, and psychological research [1]. FER systems are created within the fields of computer vision and deep learning to identify primary emotions such as anger, disgust, fear, happiness, sorrow, surprise, and contempt [2].

Several research studies have used computer vision in facial expression recognition (FER) for various applications such as healthcare, the Internet of Things (IoT), security, and human-computer interactions [3]. Emotions are regarded as cognitive states or stages, and Facial Expression Recognition (FER) plays a pivotal role in improving usability and contentment across different digital platforms [4]. Various facial expression recognition (FER) research datasets are described, emphasizing the need for a well-balanced dataset for dependable outcomes. The publication delineates the disparities in datasets about children, adults, and seniors, highlighting the divergence in emotion categories and recording circumstances [5].

The study's objective is to conduct a comprehensive analysis of facial expression recognition (FER) research, compare various datasets, examine the difficulties and probable future directions in FER, and delineate potential practical applications. The text provides a comprehensive overview of many investigations, from dataset discussions to examining current methodologies and algorithms. It concludes by offering ideas for future study [6]. Moreover, the study examines research articles based on FER, emphasizing essential keywords, geographical factors, article citations, and publications accessible via the SCOPUS database. The content encompasses a historical viewpoint on the patterns of publishing in the field of FER [7, 8, 9, 10, and 11].

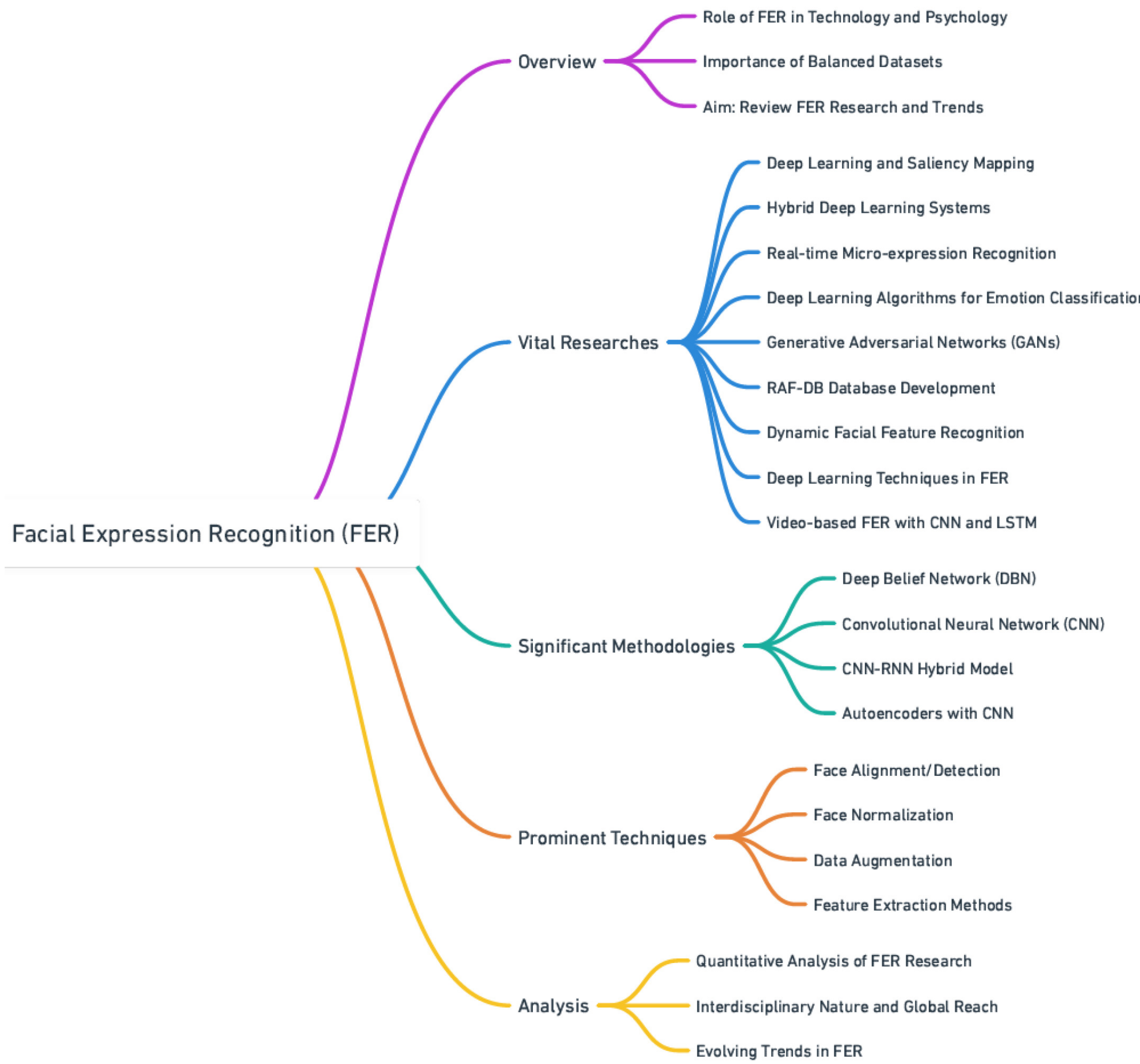

Figure 1. Overview of the Survey

2. Vital Researches

Kumari and Bhatia [12] developed a facial emotion detection method utilizing deep learning and saliency mapping. Their model preprocesses input data with a Generative Adversarial Network (GAN), contrast enhancement, and bilateral filtering. An emotion detection system employs a deep Convolutional Neural Network (CNN) that has been trained using the Nadam optimizer. Their technique was evaluated on FER+, CK+, and JAFFE datasets, yielding impressive accuracy rates: 84.2% on FER+, 97.7% on JAFFE, and 99.9% on CK+. Khan et al. [13] proposed an advanced hybrid deep learning system for facial expression recognition. This system, based on a Spatial Transformer

Network (STN), excels in feature extraction and classification. It overcomes challenges such as lighting and orientation issues and has been effectively tested on FER-2013, JAFFE, CK+, and FERG datasets, showing notable improvements in recognizing facial expressions.

Zheng and Blasch [14] proposed an innovative neural network model designed to recognize micro-expressions in real-time. This model combines Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Vision Transformer approaches to extract spatial and temporal features efficiently. After being trained on diverse datasets of facial micro-expressions, this hybrid model demonstrated notable improvements in accuracy of recognition, especially when integrating results from multiple models. Yaseen et al. [15] utilized deep learning algorithms for classifying human emotions from facial images. They experimented with various models including EfficientNetB0, DenseNet169, and a hybrid of VGG16+EfficientNetB0. Their methodology, tested on FER2013 and JAFFE datasets, resulted in high detection ratios of 90.6% and 95.6% respectively, demonstrating the effectiveness of their hybrid approaches.

S. Karthika & Durgadevi [16] focuses on Generative Adversarial Networks (GANs), detailing their fundamentals, variants, and applications. It provides an overview of GANs in various fields such as computer vision and language processing, discussing their potential future developments alongside existing challenges and limitations.Li and Deng [17] developed a novel database, RAF-DB, for facial expression recognition featuring diverse, naturally-posed images. They introduced a Deep Locality-Preserving Convolutional Neural Network (DLP-CNN) for recognizing complex, real-world facial expressions. This method, tested on multiple datasets, showed superior performance in recognizing expressions compared to existing techniques.

Kim et al. [18] proposed an approach for facial expression recognition that addresses the limitations of softmax function predictions. They introduced an autoencoder technique for generating neutral-emotion facial images, leading to improved recognition of dynamic facial features. Their method, validated on CK+ and JAFFE datasets, showed high accuracy and outperformed existing methods. Lee [19] review focuses on deep learning techniques for facial expression recognition, exploring their advantages, limitations, and potential advancements. The study highlights issues in the field and compares deep learning's effectiveness with traditional methods, suggesting future directions such as creating larger datasets and developing more robust and explainable models.

Deng et al. [20] presented a CNN-based method for video-based facial expression recognition. Their approach involved using a 3D CNN to extract spatiotemporal information from video frames, followed by an LSTM network to capture facial expression dynamics. This method demonstrated state-of-the-art results on multiple benchmark datasets, effectively utilizing expression-related muscular movements in videos.

3. Significant Methodologies

Several models exemplify the dynamic progression of AI-powered face emotion identification, each playing a role in advancing the field and enhancing the precision and complexity of emotion detection systems. The Deep Belief Network (DBN) is recognized as an initial approach in the development of face emotion detection models. DBN leverages its ability to identify complex patterns in inputs, which significantly contributes to accuracy in classification tasks. However, its usage has been limited due to the emergence of more efficient methods. For instance, a cited study using DBN achieved around 20% accuracy on a limited set of emotions. The study highlights that while DBN represented a substantial advancement in FER technology, it has been mostly replaced by more sophisticated neural network models such as Convolutional Neural Networks (CNN)[21].

A Convolutional Neural Network (CNN) is a kind of artificial neural network that is particularly effective for analyzing visual data. It is designed to automatically learn and extract features from images or other types of grid-like data by using convolutional layers. The Convolutional Neural Network (CNN) is generally acknowledged as a very effective and extensively used technique for face emotion recognition. Convolutional neural networks (CNNs) have continuously achieved superior performance compared to older models such as deep belief networks (DBNs) owing to their capacity to effectively process intricate visual input and extract pertinent characteristics for the purpose of emotion categorization. The survey discusses various customizations and enhancements of the CNN model, such as the integration of different activation functions and pooling filters. The earlier alterations have resulted in notable improvements in precision, showcasing the model's versatility and efficacy in the realm of Facial Expression Recognition (FER) [21].

The report also highlights the novel technique of integrating Convolutional Neural Networks (CNN) with Recurrent Neural Networks (RNN) to develop hybrid models for face emotion identification. These hybrid models combine the advantages of both CNNs, which are proficient in image processing, and RNNs, which excel in processing sequential data. An example cited in the document is the use of the Deformable Part Model for face detection and Dlib for facial feature extraction. This CNN-RNN hybrid model has been shown to yield the highest mean performance accuracy, indicating its efficacy in more complex FER scenarios where both spatial and temporal features are crucial.

The work from [21] explores the use of Autoencoders in conjunction with Convolutional Neural Networks (CNNs) to improve the accuracy of emotion identification. Autoencoders are a kind of neural networks that have the ability to rebuild their input data in a space with fewer dimensions. When used with CNNs, they enhance the model's capacity to reliably identify and categorize emotions. This approach is part of the ongoing evolution in FER technology, aiming to refine the accuracy and reliability of emotion detection models.

4. Prominent Techniques

Face Alignment/Detection: Face alignment and detection are crucial initial steps in facial emotion recognition. The survey highlights various techniques used for these purposes, such as the Viola-Jones Face Detector and Haar Classifiers. These methods are essential for identifying and isolating faces from images or video frames, ensuring that subsequent analysis focuses solely on facial features. The Viola-Jones method is noted for its reliability and computational efficiency in detecting frontal faces, while Haar Classifiers are used for their ability to identify facial features by reducing pixel size groups. These techniques form the foundation for accurate feature learning and alignment in FER systems[22].

Face Normalization: Face normalization involves standardizing facial data to account for variations in illumination and head poses, which can significantly impact the performance of FER models. The survey discusses various methods of normalization, including illumination normalization and histogram equalization. These techniques are designed to minimize internal differences in features due to varying lighting conditions or angles. Some approaches involve combining histogram equalization with other normalization methods to enhance global contrast in facial images, thereby improving the accuracy of emotion detection models[23].

Data Augmentation: In the context of FER, data augmentation is used to enhance the training datasets, particularly when there is a lack of sufficient images. This technique generates new data from existing images by applying various transformations while maintaining the integrity of facial data. The survey references studies that discuss different augmentation methods, including pose synthesis and illumination synthesis. Data augmentation is vital for training deep neural networks in FER, as it ensures a diverse range of data, leading to more robust and accurate emotion recognition models[24].

Feature Extraction Methods: The document covers several feature extraction techniques crucial for FER, including textual, edge-based, global and local, geometric, and patch-based features. Textural features, such as those extracted using Gabor filters or Local Binary Patterns (LBP), focus on the textural aspects of facial images[25]. Edge-based features, using methods like Histogram of Oriented Gradients (HOG), emphasize the edges and contours of facial features [26]. Global and local features, extracted through methods like Principal Component Analysis (PCA) and Independent Component Analysis (ICA), provide a comprehensive view of facial characteristics. Geometric features are about extracting discrete elements like lines and curves, while patch-based features involve segmenting the face into patches and analyzing each segment separately. These techniques collectively contribute to the detailed and nuanced analysis of facial expressions required for accurate emotion recognition [27]

5. Analysis

This quantitative analysis underscores the expanding field of FER research, highlighting its interdisciplinary nature, global reach, and the evolving trends influenced by technological and societal changes.The survey analyzed FER research papers based on crucial keywords. This analysis used the SCOPUS database, filtering journal articles and reviews from 2005 to 2021. A total of 558 documents were initially retrieved, which were then filtered down to 463 documents after excluding lecture notes, conferences, and workshops. Further removal of duplicates resulted in 318 research papers being selected for analysis. The analysis considered each document's publication title, publication year, journal/source title, and the number of citations [28].

The analysis revealed that most research on FER is conducted in the field of Computer Science, followed by Engineering and Medicine. This is indicative of the interdisciplinary nature of FER, combining aspects of technology and human psychology. The trend in Computer Science is particularly significant, aligning with FER's roots in computer vision problems[29].The survey found that India leads in publishing papers related to FER, followed by the United States, the United Kingdom, China, and Germany. This geographical analysis highlights the global interest in FER research and its relevance across diverse cultural and technological landscapes [30].

Table 1. Comprehensive Overview of Methods and Techniques in FER

Category	Method	Description
Global and Local Features[31]	Principal Component Analysis	Extracts global and low-dimensional features, widely used in FER.
	Independent Component Analysis	Extracts local characteristics using multichannel observations.
	Stepwise Linear Discriminant Analysis	Extracts localized features using regression models.
	Discrete Fourier Transform	Conventional method for extracting global features.
	Gabor Wavelet Transform	Used for extracting local features.
Geometric Features [32]	Corner Detection	Extracts corners of an object.
	Edge Detection	Defines boundaries of image regions.
	Blob Detection	Recognizes sections of images.
	Local Curvelet Transform	Extracts geometric features like mean entropy, standard deviation, and kurtosis using a three-stage steerable pyramid representation.
Patch-Based Feature [33]	Patch Extraction and Matching	Converts extracted patches into distance characteristics.
	K-Nearest Neighbor	Used after extracting features from patches.

	Gabor Features with Patch-Based Extractor	Combined with patch-based extractor to improve accuracy for small sample sizes.
Deep Belief Network (DBN) [34]	Deep Belief Network	Based on Restricted Boltzmann Machine (RBM), learns abstract facial image information.
Convolutional Neural Network (CNN) [35]	Convolutional Neural Network	Consists of convolutional layers, pooling layers, and fully connected layers. Efficient in image detection, scene segmentation, and FER.
Advanced CNN Models [36]	Dense Convolutional Network	Uses concatenation for passing feature maps between layers.
	Extreme Inception Network	Replaces inception modules with depthwise separable convolutions.
	Residual Networks with Parallel Towers	Adds scaling of parallel towers within each module.
	Shuffle Network	Designed for mobile devices, using pointwise convolutions.
	Mobile Network	Uses pointwise convolutions in depthwise separable convolutions.
	Efficient Network	Uses compound scaling method for balancing network dimensions.
Deep Autoencoder (DAE) [36]	Deep Autoencoder	Reproduces the input dataset at the output, encoding information into a representation that can be regenerated.
Recurrent Neural Network (RNN) [36]	Recurrent Neural Network	Incorporates temporal information, suitable for predicting sequential data. Uses Long Short-Term Memory units for video-based expression recognition.

6. CONCLUSION

The thorough examination of FER highlights its substantial development and increasing influence in diverse domains, propelled by computer vision and profound learning breakthroughs. Groundbreaking research has successfully used techniques such as deep learning, saliency mapping, and hybrid models that integrate CNNs with RNNs, resulting in significant advancements in the precise identification and analysis of fundamental emotions. The study showcases a worldwide, multidisciplinary endeavor, notably prominent in medicine, technology, and science, with significant contributions from nations such as the USA, India, and China. The survey's methodological analysis of different feature extraction methods and neural network models emphasizes the dynamic character of the area and its critical importance in an extensive spectrum of usage, including healthcare, security, and human-computer interaction. The dynamic field of FER not only presents exciting opportunities for future breakthroughs in technology and society but also provides a platform for tackling forthcoming issues in the identification and interpretation of emotions.

REFERENCES

[1] K. Chengeta and S. Viriri, ''A survey on facial recognition based on local directional and local binary patterns,'' in Proc. Conf. Inf. Commun. Technol. Soc. (ICTAS), Mar. 2018, pp. 1–6, doi: 10.1109/ICTAS.2018.8368757.

[2] G. Rajeswari and P. IthayaRani, ''Literature survey on facial expression recognition techniques,'' in Proc. 3rd Int. Conf. Commun. Electron. Syst. (ICCES), Oct. 2018, pp. 137–142, doi: 10.1109/CESYS.2018.8723953.

[3] I. M. Revina and W. R. S. Emmanuel, ''A survey on human face expression recognition techniques,'' J. King Saud Univ.- Comput. Inf. Sci., vol. 33, no. 6, pp. 619–628, Sep. 2018, doi: 10.1016/J.JKSUCI.2018.09.002.

[4] B. Martinez, M. F. Valstar, B. Jiang, and M. Pantic, ''Automatic analysis of facial actions: A survey,'' IEEE Trans. Affect. Comput., vol. 10, no. 3, pp. 325–347, Jul. 2019, doi: 10.1109/TAFFC.2017.2731763.

[5] V. Franzoni, G. Biondi, D. Perri, and O. Gervasi, ''Enhancing mouthbased emotion recognition using transfer learning,'' Sensors, vol. 20, no. 18, pp. 1–15, 2020, doi: 10.3390/s20185222

[6] B. Li and D. Lima, ''Facial expression recognition via ResNet-50,'' Int. J. Cognit. Comput. Eng., vol. 2, pp. 57–64, Jun. 2021, doi: 10.1016/J.IJCCE.2021.02.002

[7] M. Al-Shabi, W. P. Cheah, and T. Connie, ''Facial expression recognition using a hybrid CNN–SIFT aggregator,'' in Proc. Int. Workshop Multi-Disciplinary Trends Artif. Intell., in Lecture Notes in Computer Science: Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, vol. 10607, Aug. 2016, pp. 139–149, doi: 10.1007/978-3-319-69456-6_12.

[8] P. J. Bota, C. Wang, A. L. N. Fred, and H. P. Da Silva, ''A review, current challenges, and future possibilities on emotion recognition using machine learning and physiological signals,'' IEEE Access, vol. 7, pp. 140990–141020, 2019, doi: 10.1109/ACCESS.2019.2944001.

[9] T. Kundu and C. Saravanan, ''Advancements and recent trends in emotion recognition using facial image analysis and machine learning models,'' in Proc. Int. Conf. Electr., Electron., Commun., Comput., Optim. Techn. (ICEEECCOT), Dec. 2017, pp. 1–6, doi: 10.1109/ ICEEECCOT.2017.8284512.

[10] S. Bhattacharya and M. Gupta, ''A survey on: Facial emotion recognition invariant to pose, illumination and age,'' in Proc. 2nd Int. Conf. Adv. Comput. Commun. Paradigms (ICACCP), Feb. 2019, pp. 1–6, doi: 10.1109/ICACCP.2019.8883015.

[11] A. S. Vyas, H. B. Prajapati, and V. K. Dabhi, ''Survey on face expression recognition using CNN,'' in Proc. 5th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS), Mar. 2019, pp. 102–106, doi: 10.1109/ICACCS.2019.8728330.

[12]. Kumari, N., & Bhatia, R. (2023). Saliency map and deep learning based efficient facial emotion recognition technique for facial images. Multimedia Tools and Applications. https://doi.org/10.1007/s11042-023-16220-0

[13]. Khan, N., Singh, A. V., & Agrawal, R. (2022). Enhanced Deep Learning Hybrid Model of CNN Based on Spatial Transformer Network for Facial Expression Recognition. International Journal of Pattern Recognition and Artificial Intelligence, 36(14). https://doi.org/10.1142/s0218001422520280

[14]. Zheng, Y., & Blasch, E. (2023). Facial Micro-Expression Recognition Enhanced by Score Fusion and a Hybrid Model from Convolutional LSTM and Vision Transformer. Sensors, 23(12), 5650. https://doi.org/10.3390/s23125650

[15]. Yaseen, A. F., Shaukat, A., & Alam, M. (2022). Emotion Recognition from Facial Images using Hybrid Deep Learning Models. 2022 2nd International Conference on Digital Futures and Transformative Technologies(ICoDT2). https://doi.org/10.1109/icodt255437.2022.9787474

[16]. S, Karthika., & Durgadevi, M. (2021). Generative Adversarial Network (GAN): a general review on different variants of GAN and applications. 2021 6th International Conference on Communication and Electronics Systems (ICCES). https://doi.org/10.1109/icces51350.2021.9489160

[17]. Li, S., & Deng, W. (2019). Reliable Crowdsourcing and Deep Locality-Preserving Learning for Unconstrained Facial Expression Recognition. IEEE Transactions on Image Processing, 28(1), 356–370. https://doi.org/10.1109/tip.2018.2868382

[18]. Kim, J.-H., Kim, B.-G., Roy, P. P., & Jeong, D.-M. (2019). Efficient Facial Expression Recognition Algorithm Based on Hierarchical Deep Neural Network Structure. IEEE Access, 7, 41273–41285. https://doi.org/10.1109/access.2019.2907327

[19]. Lee, S. (2023). A Brief Review of Deep Learning for Facial Expression Recognition. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4318896

[20]. Deng, X., Zhang, L., Loy, C.C., Tang, X., & Cao, X. (2018). Residual attention network for video-based facial expression recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 5573-5582).

[21] C. Marechal, D. Mikołajewski, K. Tyburek, P. Prokopowicz, L. Bougueroua, C. Ancourt, and K. Wegrzyn-Wolska, ''Survey on AIbased multimodal methods for emotion detection,'' in High-Performance Modelling and Simulation for Big Data Applications (Lecture Notes in Computer Science: Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, vol. 11400, 2019, pp. 307–324, doi: 10.1007/978-3-030-16272-6_11.

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

[23] A. S. Meuwissen, J. E. Anderson, and P. D. Zelazo, ''The creation and validation of the developmental emotional faces stimulus set,'' Behav. Res. Methods, vol. 49, no. 3, p. 960, Jun. 2017, doi: 10.3758/ S13428-016-0756-7.

[24] V. LoBue and C. Thrasher, ''The child affective facial expression (CAFE) set: Validity and reliability from untrained adults,'' Frontiers Psychol., vol. 5, p. 1532, Jan. 2015, doi: 10.3389/FPSYG.2014.01532.

[25] B. Nojavanasghari, T. Baltrušaitis, C. E. Hughes, and L.-P. Morency, ''EmoReact: A multimodal approach and dataset for recognizing emotional responses in children,'' in Proc. 18th ACM Int. Conf. Multimodal Interact., Oct. 2016, pp. 137–144, doi: 10.1145/2993148.2993168.

[26] Radboud Faces Database—Behavioural Science Institute. Accessed: Aug. 29, 2021. [Online]. Available: https://www.ru.nl/ bsi/research/facilities/radboud-faces-database/

[27] 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 Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.-Workshops (CVPRW), Jun. 2010, pp. 94–101, doi: 10.1109/CVPRW.2010.5543262.

[28] JAFFE Dataset | Papers With Code. Accessed: Aug. 29, 2021. [Online]. Available: https://paperswithcode.com/dataset/jaffe

[29] X. Feng, B. Lv, Z. Li, and J. Zhang, ''Automatic facial expression recognition with AAM-based feature extraction and SVM classifier,'' in Proc. Mex. Int. Conf. Artif. Intell., in Lecture Notes in Computer Science: Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, vol. 4293, 2006, pp. 726–733, doi: 10.1007/11925231_69.

[30] A. Ghofrani, R. M. Toroghi, and S. Ghanbari, ''Realtime face-detection and emotion recognition using MTCNN and miniShuffleNet V2,'' in Proc. 5th Conf. Knowl. Based Eng. Innov. (KBEI), Feb. 2019, pp. 817–821, doi: 10.1109/KBEI.2019.8734924.

[31] W. Chen, M. J. Er, and S. Wu, ''Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain,'' IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 36, no. 2, pp. 458–466, Apr. 2006, doi: 10.1109/TSMCB.2005.857353.

[32] D. A. Pitaloka, A. Wulandari, T. Basaruddin, and D. Y. Liliana, ''Enhancing CNN with preprocessing stage in automatic emotion recognition,'' Proc. Comput. Sci., vol. 116, pp. 523–529, Jan. 2017, doi: 10.1016/J.PROCS.2017.10.038.

[33] P. Hu, D. Cai, S. Wang, A. Yao, and Y. Chen, ''Learning supervised scoring ensemble for emotion recognition in the wild,'' in Proc. 19th ACM Int. Conf. Multimodal Interact. (ICMI), Nov. 2017, pp. 553–560, doi: 10.1145/3136755.3143009.

[34] C. Shorten and T. M. Khoshgoftaar, ''A survey on image data augmentation for deep learning,'' J. Big Data, vol. 6, no. 1, pp. 1–48, Dec. 2019, doi: 10.1186/S40537-019-0197-0.

[35] Y. Ji and K. Idrissi, ''Automatic facial expression recognition based on spatiotemporal descriptors,'' Pattern Recognit. Lett., vol. 33, no. 10, pp. 1373–1380, Jul. 2012, doi: 10.1016/J.PATREC.2012.03.006.

[36] Z.-H. Feng, J. Kittler, W. Christmas, P. Huber, and X.-J. Wu, ''Dynamic attention-controlled cascaded shape regression exploiting training data augmentation and fuzzy-set sample weighting,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 3681–3686, doi: 10.1109/CVPR.2017.392.