Survey on Facial expression recognition from images using deep learning architecture

Mrs. Nalini Jagtap
Assistant Professor, Department of Computer Engineering,
Dr. D. Y. Patil Institute of Engineering, Management and Research, Akurdi, Pune, India
Email: yadav.nalini@gmail.com

Mrs. Pratiksha Shevatekar
Assistant Professor, Department of Computer Engineering,
Dr. D. Y. Patil Institute of Engineering, Management and Research, Akurdi, Pune, India
Email: p_shevatekar@rediffmail.com

Ms. Rachana Mudholkar
Assistant Professor, Department of Computer Engineering,
Dr. D. Y. Patil Institute of Engineering, Management and Research, Akurdi, Pune, India
Email: mudholkar.rrm@gmail.com

Abstract: *Emotion detection and recognition is an interdisciplinary branch of computing technology, which is used to capture, store, process, and interpret human emotions and generate results for better decision-making in real time. It includes analyzing various emotions through facial expressions, postures, gestures, speech, text, and temperature changes in the human body.*

Keyword: *Convolutional deep belief network; Deep learning; Emotion detection; Gabor Filter; Neural networks; Wavelet Transform*

1. INTRODUCTION

Emotion detection and recognition solutions are one of the latest trends in the global IT market. These advanced solution suites capture and analyze all dimensions of human emotions, gestures, and other physiological changes in the subjects/people under observation. This technology could play a vital role in security surveillance, decision-making, and formulating strategies on the basis of time series analysis. Additionally, technology platforms like deep learning, structured on the basis of robust data repository and comprising extensive analysis of the behavioral patterns of people, help in providing high-precision emotion analytics.

Strong focus on strategies that reduce human effort has given a major push to automation technologies like emotion detection. Emotion detection aids in quicker decision-making process, which not only eliminates the biased human perceptions regarding people, but also helps in better understanding the customer's needs, thus improving the marketing and consulting services provided to the customers.

The rising demand for mapping human emotions, especially in the marketing and advertising sector; increasing concerns for security; and growing need for applications for support during emergency situation have bolstered the growth of emotion recognition and detection systems. Verticals such as banking, defense, and commercial security are showing significant growth in the adoption of biometrics. Currently, due to the increasing rate of cybercrimes and illegal trespassing, biometric authentication and deception detection systems have gained pace in terms of growth. These developments have cumulatively created an environment for emotion sensing technologies to be embedded within computing systems for enhanced computing experience, thereby resulting in better decision-making.

1. Applications

With the ongoing research in the field of robotics, especially in the field of humanoid robots, it becomes interesting to integrate these capabilities into machines allowing for a more diverse and natural way of communication**.** Some important applications of Facial expression recognition systems are as follow.

1. Surveillance of restricted area, Surveillance of huge crowd
2. Advertising
3. Data driven animations
4. Interactive games
5. Sociable robotics
6. Driver safety
7. Psychological disorder understanding
8. Counselling, interrogating, reward games

2. Existing Work

Development of an affective facial recognition system still remains a challenging task. Facial images and videos are affected by illumination conditions, human age, and variations in how the emotion is expressed. A first attempt was made in

1978 by Suwa et al. [1] to develop automatic facial expression system by analyzing twenty known areas of image structure. Since then many attempts have been made to develop automatic emotion recognition system that can narrate affective human feelings. Automatic facial expression recognition system is divided into three main stages, that is, (1) face detection and tracking, (2) feature extraction, and (3) emotion classification [2, 3]. For facial expression recognition systems, extracted facial features have been generally based on geometry [4] or appearance. In [5], facial element and muscle movement are used as features for facial expression recognition, which are obtained from patch based 3D Gabor filters. In [6], entropy, standard deviation, and mean of curvelet transform coefficients of each region are used as features for facial expression recognition. The extracted features are passed to online sequential extreme learning machine with radial basis function. SWT [7] is applied on the detected face, feature vector length is further reduced by applying DCT and selecting a few number of DCT coefficients. For classification of facial expression, an artificial neural network is trained on extracted features from the images.

Deep Learning

In recent times, deep learning has become very useful for many computer vision applications [8][9].In recent years by emerging powerful parallel processing hardware, Deep Neural Networks (DNN) become a hot topic in pattern recognition and machine learning science. Deep Learning (DL) scheme is based on the consecutive layers of signal processing units in order to mix and re-orient the input data to their most representative order correspond to a specific application. Facial expression classification has taken advantage of DNN classifiers in recent years. In [10] an Action Unit inspired Deep Network has been proposed which uses the DNN to extract the most representative features. [11] Achieved facial expression recognition based on a deep CNN. Face detection is achieved [11] by using Haar-like features and histogram equalization and then four-layer CNN architecture is constructed, including two convolution layers and two sub sampling layers (C-S-C-S) with Softmax classifier is used for multi-classification. [12] Reported a test accuracy of 92% based on the combination of 2D and 3D features trained on DCNNs from scratch. The fusion approach [13] has been proposed and the well established training and evaluation protocol for the BU-3DFE dataset (i.e. a smaller training data) is implemented, reached a competitive test accuracy of 89.31%.

Literature Review

Sr No	Literature	Features
1	Multimodal Emotion Recognition using Deep Learning Architectures, 2016[14]	• Multimodal Datasets – emoFBVP • four deep belief network (DBN) models generate robust multimodal features for emotion classification in an unsupervised manner • Convolutional deep belief network (CDBN) models that learn salient multimodal features of expressions of emotions
2	DeXpression: Deep Convolutional Neural Network for Expression Recognition, 2016[15]	• Datasets : MMI Dataset, CKP Dataset • The proposed four parts deep Convolutional Neural Network architecture. • Begins with Convolution 1, which applies 64 different filters. The next layer is Pooling 1, which down-samples the images and then they are normalized by LRN 1. • The next steps are the two FeatEx (Parallel Feature Extraction Block) block. • The features extracted by theses blocks are forwarded to a fully connected layer, which uses them to classify the input into the different emotions.
3	Deep Learning for Facial Expression Recognition: A step closer to a Smart Phone that Knows your Moods, 2017[16]	• Datasets **:** RaFD, CK+, JAFFE • Flipping images and rotating them by [-3,-2,-1,1,3,3] degrees for enlarging database for training DNN adding a wide range of samples drawn from different conditions and properties to DNN training sets, leads to a more reliable network
4	Facial Expression Recognition via Joint Deep Learning of RGB-Depth Map Latent Representations, 2017[13]	• Datasets : BU-3DFE • Fusion of depth map and RGB modalities result in a more robust classifier for FER. • A pipeline based architecture on transfer learning and deep convolutional neural network (DCNN)
5	A Method of Facial Expression Recognition Based on Gabor and NMF, 2016[17]	• Datasets : JAFFE facial expressions database • Expressional features are extracted by Gabor wavelet transformation and the high dimensional data are reduced by nonnegative matrix factorization (NMF). • Two layer classifier (TLC) is designed for expression recognition.

6	Deep Learning for Facial Expression Recognition:A step closer to a SmartPhone that Knows your Moods, 2017[18]	• Datasets : FER 2013 dataset, SFEW and AFEW • Improved face descriptors based on 2D and 3D Convolutional • Neural Networks, fusion of temporal and multimodal, including a novel hierarchical method combining features and scores are proposed with adapted CNN network achieved the accuracy of 58.8 %.
7	A Classification of Emotion and Gender Using Approximation Image Gabor Local Binary Pattern, 2017[19]	• Database: FERET, INDIAN FACE and AR FACE databases • Classification of gender and Emotion using the AIGLBP method, which includes the properties of wavelet decomposition and Gabor filter. Cross table classification rate reached 50+%. AIGLBP (Approximation image Gabor local binary pattern) is applied for feature extraction and SVM is used for classification.
8	Audio-Visual Emotion Recognition in Video Clips, 2016[20]	• Database : SAVEE, eNTERFACE'05, and RML databases • Audio features included prosodic features, MFCCs and FBEs. Visual features were computed from estimated key-frames representing each video content in terms of representative facial expressions. Visual data was described both by using geometric features and by means of a CNNbased model. Four types of classification methods were used, i.e. multiclass SVM and RF with and without applying PCA • improvements compared to previous state-of-the-art results
9	Transfer of Multimodal Emotion Features in Deep Belief Networks, 2017[11]	• Database : emoFBVP, Mind Reading, MMI, Cohn Kanade • 6-layer DBN trained network is proposed with transfer of features in layer by layer manner. • Proposed approach reduces training time and exhibits performance boost in accuracy
10	Automatic facial expression recognition based on a deep convolutional-neural-network structure, 2017[21]	• Database : Japanese Female Facial Expression Database(JAFFE) and the Extended Cohn-Kanade Dataset(CK+) • deep convolutional neural network (CNN) to devise a facial expression recognition system, which is capable to discover deeper feature representation of facial expression to achieve automatic recognition
11	Facial Expression Recognition Using Enhanced Deep 3D Convolutional Neural Networks, 2017[22]	• Database : MMI, CK+, FERA, DISFA • Proposed system consists of 3D Inception-ResNet layers followed by an LSTM unit that together extracts the spatial relations within facial images as well as the temporal relations between different frames in the video
12	Automatic emotion detection model from facial expression, 2017[23]	• Database : JAFFE • Live streaming and processed it using Grabor feature extraction and neural network • To detect the emotion facial attributes extraction by principal component analysis is used and a clusterization of different facial expression with respective emotions • the processed feature vector is channeled through the already learned pattern classifiers and facial expression is determined.

3. Challenges

Extensive literature survey has put down multiple challenges. Emotion recognition system from live streaming is characterized with following challenges which are considered in the scope of the research

1. Presence of multimodal features
2. Feature extraction in RGB color space
3. Recognition of dynamic expressions
4. Uneven illumination, blurred situations in videos
5. Absence of one or more modalities in live streaming
6. Distinguishing between more emotions such as Nervousness and Panic
7. Complex emotions : For example frustration can be accompanied by anger, therefore not only showing one emotion, but also the reason
8. Besides distinguishing between different emotions, also understanding the strength of an emotion
9. Multiple way of expressing emotions across various

people

4. Proposed System

The process of facial expression recognition mainly includes feature extraction and expression classification. The general procedure of facial expression recognition is shown by Fig. 1. The preprocessing is to eliminate some interference factors which have influences on the feature extraction and classification. The purpose of feature extraction is to obtain the attribute information which can characterize the class of input image. Classification is a differential process according to the characteristics of input image. Commonly used methods of feature extraction are based on image or gray values.

Figure 1. Facial Expression Recognition System

The process of facial expression recognition system in proposed research is shown by Fig. 2.

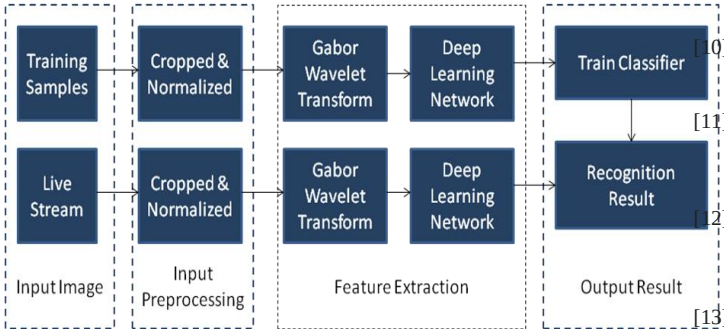

Figure 2. Proposed Facial Expression Recognition System

Facial expression is recognized by combining the Gabor wavelet transform with Deep Neural Network in proposed research. Nearest neighbour classifier and Voting Statistics is adopted to complete the two layer classification.

5. CONCLUSION

The emotion detection and recognition technology holds tremendous potential for diverse applications across various industries. The technology follows basic steps that include collecting, processing, and interpreting human emotions with the help of a computing system. The formation of emotion detection and recognition systems requires the combination of Deep learning with Neural Networks. It involves registration and execution of videos for emotion detection and identification

REFERENCES

[1] M. Suwa, N. Sugie, and K. Fujimora, "A preliminary note on pattern recognition of human emotional expression," in Proceedings of the 4th International Joint Conference on Pattern Recognition, pp. 408–410, Kyoto, Japan, November 1978.

[2] P. Ekman and W. V. Friesen, Facial Action Coding System, Consulting Psychologists Press, Stanford University, Palo Alto, Calif, USA, 1977.

[3] J. C. Hagerand, P. Ekman, and W. V. Friesen, Facial Action Coding System,AHuman Face, Salt Lake City,Utah, USA, 2002.

[4] M. H. Siddiqi, R. Ali, A. M. Khan, E. S. Kim, G. J. Kim, and S. Lee, "Facial expression recognition using active contour-based face detection, facial movement-based feature extraction, and non-linear feature selection," Multimedia Systems, vol. 21, no. 6, pp. 541–555, 2015.

[5] L. Zhang and D. Tjondronegoro, "Facial expression recognition using facial movement features," IEEE Transactions on Affective Computing, vol. 2, no. 4, pp. 219–229, 2011.

[6] Uc¸ar, Y. Demir, and C. G¨uzelis¸, "A new facial expression recognition based on curvelet transform and online sequential extreme learningmachine initialized with spherical clustering," Neural Computing and Applications, vol. 27, no. 1, pp. 131–142, 2016.

[7] Huma Qayyum, MuhammadMajid, Syed Muhammad Anwar, and Bilal Khan, "Facial Expression Recognition Using Stationary Wavelet Transform Features" , Hindawi Mathematical Problems in Engineering Volume 2017, Article ID 9854050, 9 pages.

[8] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," Neurocomputing, vol. 187, pp. 27–48, 2016.

[9] S. E. Kahou, X. Bouthillier, P. Lamblin, C. Gulcehre, V. Michalski, K. Konda, S. Jean, P. Froumenty, Y. Dauphin, N. Boulanger-Lewandowski et al., "Emonets: Multimodal deep learning approaches for emotion recognition in video," Journal on Multimodal User Interfaces, vol. 10, no. 2, pp. 99–111, 2016.

[10] M. Liu, S. Li, S. Shan, and X. Chen, "AU-inspired Deep Networks for Facial Expression Feature Learning," Neurocomputing, vol. 159, no. 1, pp. 126–136, 2015.

[11] Ke Shan, Junqi Guo*, Wenwan You, Di Lu, Rongfang Bie, "Automatic Facial Expression Recognition Based on a Deep Convolutional-Neural-Network Structure", IEEE SERA 2017, 978-1-5090-5756-6, June 7-9, 2017, London, UK

[12] X.-P. Huynh, T.-D. Tran, and Y.-G. Kim, "Convolutional neural network models for facial expression recognition using bu-3dfe database," in Information Science and Applications (ICISA) 2016. Springer Singapore, 2016, pp. 441–450.

[13] Oyebade K. Oyedotun, Girum Demisse, Abd El Rahman Shabayek, Djamila Aouada, Bjorn Ottersten, "Facial Expression Recognition via Joint Deep Learning of RGB-Depth Map Latent Representations", 2017 IEEE International Conference on Computer Vision Workshop (ICCVW), pp. 3161-3168, 2017

[14] Hiranmayi Ranganathan, Shayok Chakraborty and Sethuraman Panchanathan. "Multimodal Emotion Recognition using Deep Learning Architectures, Application of Computer Vision (WACV) IEEE conference, 2016

[15] Peter Burkert, Felix Trier, Muhammad Zeshan Afzal, Andreas Dengel, Marcus Liwicki, "DeXpression: Deep Convolutional Neural Network for Expression Recognition", arXiv.org, 2016

[16] Shabab Bazrafkan, Tudor Nedelcu, Pawel Filipczuk, Peter Corcoran, "Deep Learning for Facial Expression Recognition:A step closer to a SmartPhone that Knows your Moods" , 2017 IEEE International Conference on Consumer Electronics (ICCE)

[17] Jun Zhou, Sue Zhang, Hongyan Mei, and Dawei Wang, "A Method of Facial Expression Recognition Based on Gabor and NMF", Pattern Recognition and Image Analysis, 2016, Vol. 26, No. 1, pp. 119–124

[18] Valentin Vielzeuf, Stéphane Pateux, Frédéric Jurie, "Temporal Multimodal Fusion for Video Emotion Classification in the Wild", arXiv:1709.07200v1, 2017

[19] Kamaljeet Singh Kalsi, Dr. Preeti Rai, "A Classification of Emotion and Gender Using Approximation Image Gabor Local Binary Pattern", Cloud Computing, Data Science & Engineering - Confluence, 2017 7th International Conference on, 12-13 Jan. 2017

[20] Fatemeh Noroozi ; Marina Marjanovic ; Angelina Njegus ; Sergio Escalera ; Gholamreza Anbarjafari, "Audio-Visual Emotion Recognition in Video Clips", IEEE Transactions on Affective Computing (Volume: PP, Issue: 99), June 2017

[21] Hiranmayi Ranganathan, Shayok Chakraborty, Sethuraman Panchanathan, "Transfer of Multimodal Emotion Features in Deep Belief Networks", Signals, Systems and Computers, 2016 50th Asilomar Conference on, March 2017

[22] Behzad Hasani, Mohammad H. Mahoor, "Facial Expression Recognition Using Enhanced Deep 3D Convolutional Neural Networks", Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on, August 2017
[23] Debishree Dagar ; Abir Hudait ; H. K. Tripathy ; M. N. Das, "Automatic emotion detection model from facial expression", Advanced Communication Control and Computing Technologies (ICACCCT), 2016 International Conference on, January 2017

Authors Biography

Mrs. Nalini Jagtap, is a Research Scholar of Department of CSE in Savitribai Phule University. She completed her M.E. in department at PCCOE, SPPU. Her research interests are image processing, image mining, emotion detection and deep leaning.

Cite this paper:

Mrs. Nalini Jagtap, Mrs. Pratiksha Shevatekar, Ms. Rachana Mudholkar "Survey on Facial expression recognition from images using deep learning architecture", International Journal of Advances in Computer and Electronics Engineering, Vol. xx, No. xx, pp. xx-xx, June 2017.