International Journal on Applications in Information and Communication Engineering
Volume 8 : Issue 1 : June 2022, pp 34 – 40 www.aetsjournal.com ISSN (Online) : 2394 - 6237

MULTIMODEL FUSION IMAGE FOR EMOTION DETECTION USING CNN IN DEEP LEARNING

K.VENKATASALAM , NANDHINI.M , SUVITHA.M , SWETHA.M , THAMARAISELVI.M

Abstract— Facial expressions convey non-verbal information between humans in face-to- face relations. Automatic facial expression recognition, which plays a vital part in mortal-machine interfaces, has attracted adding attention from experimenters since the early nineties. Classical machine learning approaches frequently bear a complex point birth process and produce poor results. In this paper, we apply recent advances in deep learning to propose effective deep Convolutional Neural Networks (CNNs) that can directly interpret semantic information available in faces in an automated manner without hand-designing of features descriptors. We also apply different loss functions and training tricks in order to learn CNNs with a strong bracket power. The experimental results show that our proposed networks outperform state-of-the- art styles on the well- known FERC-2013 dataset handed on the Kaggle facial expression recognition competition. In comparison to the winning model of this competition, the number of parameters in our proposed networks intensely decreases, that accelerates the overall performance speed and makes the proposed networks well suitable for real- time systems.

Keywords— facial emotion recognition; conventional FER; deep learning-based FER; convolution neural networks;

I. INTRODUCTION

Ever since computers were constructed, people have wanted to make artificially intelligent (AI) systems that are mentally and/ or physically original to humans. In the once decades, the increase of generally available computational power handed a helping hand for developing fast learning mach

Dr K.VENKATASALAM M.E., PH.D, Associate Professor , Department of Computer Science and Engineering , Mahendra Engineering College , Tamil Nadu , India . (Email : venkatasalamk@mahendra.info)
NANDHINI.M, Department of Computer Science and Engineering , Mahendra Engineering College , Tamil Nadu , India . (Email : nandhini5471@gmail.com)
SUVITHA.M, Department of Computer Science and Engineering , Mahendra Engineering College , Tamil Nadu , India . (Email : suvitham8@gmail.com)
SWETHA.M, Department of Computer Science and Engineering , Mahendra Engineering College , Tamil Nadu , India . (Email : shwetaamohan@gmail.com)
THAMARAISELVI.M Department of Computer Science and Engineering , Mahendra Engineering College , Tamil Nadu , India . (Email : thamaraiselvi17@gmail.com)

also fete mortal feelings provides a new dimension to mortal- machine relations, for case, smile sensor in marketable digital cameras or interactive announcements. Robots can also profit from automated facial expression recognition. However, they can reply upon this and have applicable actions, If robots can prognosticate mortal feelings. In this paper, we borrow deep learning fashion and propose effective infrastructures of Convolutional Neural Networks to break the problem of facial expression recognition. We also apply different loss functions associated with supervised learning and several training tricks in order to learn CNNs with a strong discrimination power. We show that Multiclass SVM loss works better than cross-entropy loss (combining with soft max function) in facial expression recognition. Either, the evaluation of the test sets using multiple crops with different scales and reels can yield an delicacy boost compared to single crop evaluation.

Facial expression recognition competition FERC-2013 (1) was held on 2013 by Kaggle. The winner is RBM platoon (14) with the delicacy of69.4 on public test set and71.2 on private test set. To the stylish of our knowledge, this is the stateof-the- art on the FERC-2013 dataset so far. The trials show that our proposed system achieves better delicacy than their results on both two test sets. The rest of the paper is organized as follow. In section II we compactly epitomize some affiliated work on facial expression recognition. In section III we describe our proposed CNN infrastructures. Our trials and evaluation are shown in section IV. The conclusion is section V with some discussion for the future work

1)DEEP LEARNING

Deep Learning is a type of machine learning and artificial intelligence (AI) that imitates the way humans gain certain types of knowledge. Deep learning is an important element of data wisdom, which includes statistics and prophetic modeling. It's extremely salutary to data scientists who are assigned with collecting, assaying and interpreting large quantities of data; deep learning makes this process briskly and lightly.

Deep learning vs. machine learning

Still, how do they differ? Deep learning distinguishes itself from classical machine learning by the type of data that it works with and the styles in which it learns, If deep learning is a subset of machine learning.

Machine learning algorithms influence structured, labeled data to make prognostications — meaning that specific features are defined from the input data for the model and organized into tables. This does n't inescapably mean that it does n't use unshaped data; it just means that if it does, it generally goes through somepre-processing to organize it into a structured format.

Deep Learning eliminates some of datapre-processing that's generally involved with machine learning. These algorithms can ingest and reuse unshaped data, like textbook and images, and it automates point birth, removing some of the reliance on mortal experts. For illustration, let's say that we had a set of prints of different faves, and we wanted to classify by " cat", " canine", " hamster", et cetera. Deep learning algorithms can determine which features (e.g. cognizance) are most important to distinguish each beast from another. In machine learning, this scale of features is established manually by a mortal expert.

Also, through the processes of grade descent and back propagation, the deep learning algorithm adjusts and fits itself for delicacy, allowing it to make prognostications about a new print of an beast with increased perfection.

Machine learning and deep learning models are able of different types of learning as well, which are generally distributed as supervised learning, unsupervised learning, and underpinning learning. Supervised learning utilizes labeled datasets to classify or make prognostications; this requires some kind of mortal intervention to marker input data rightly. In discrepancy, unsupervised learning does n't bear labeled datasets, and rather, it detects patterns in the data, clustering them by any identifying characteristics. Underpinning learning is a process in which a model learns to come more accurate for performing an action in an terrain grounded on feedback in order to maximize the price..

2)How deep learning works

Deep Learning neural networks, or artificial neural networks, attempts to mimic the mortal brain through a combination of data inputs, weights, and bias. These rudiments work together to directly fete, classify, and describe objects within the data.

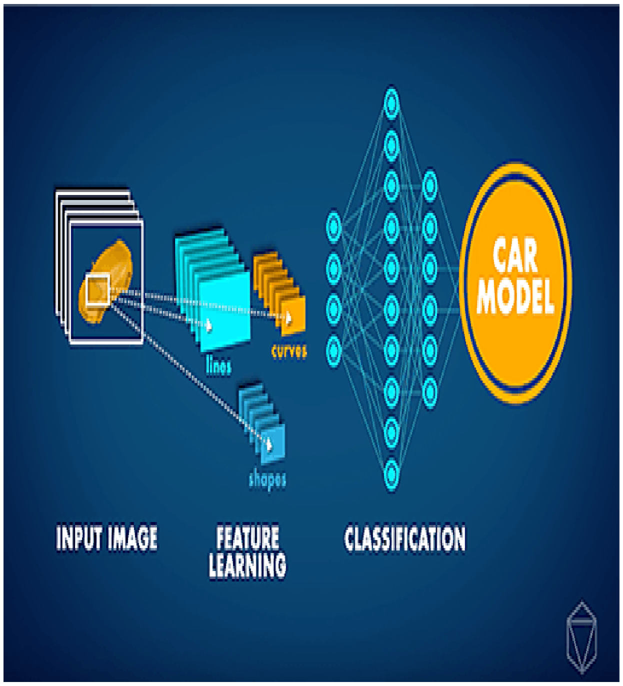
Deep neural networks correspond of multiple layers of connected bumps, each structure upon the former subcaste to upgrade and optimize the vaticination or categorization. This progression of calculations through the network is called forward propagation. The input and affair layers of a deep neural network are called visible layers. The input subcaste is where the deep learning model ingests the data for processing, and the affair subcaste is where the final vaticination or bracket is made.

Another process called back propagation uses algorithms, like grade descent, to calculate crimes in prognostications and also adjusts the weights and impulses of the function by moving backward through the layers in an trouble to train the model. Together, forward propagation and back propagation allow a neural network to make prognostications and correct for any crimes consequently. Over time, the algorithm becomes gradationally more accurate.

The below describes the simplest type of deep neural network in the simplest terms. Still, deep learning algorithms are incredibly complex, and there are different types of neural networks to address specific problems or datasets. For illustration,

Convolutional neural networks (CNNs), used primarily in computer vision and image bracket operations, can descry features and patterns within an image, enabling tasks, like object discovery or recognition. In 2015, a CNN bested a mortal in an object recognition challenge for the first time.

Recurrent neural network (RNNs) are typically used in natural language and speech recognition applications as it leverages sequential or times series data.



II. LITERATURE SURVEY

1)IM Revina, WRS Emmanuel

This paper describes the check of Face Expression Recognition (FER) ways which include the three major stages similar as preprocessing, point birth and bracket. This check explains the colorful types of FER ways with its major benefactions. The performance of colorful FER ways is compared grounded on the number of expressions honored and complexity of algorithms. Databases like JAFFE, CK, and some other variety of facial expression databases are bandied in this check. The study on classifiers gather from recent papers reveals a more important and dependable understanding of the peculiar characteristics of classifiers for exploration fellows.

2)A Qayyum, I Razzak

In this paper, we present progressive light residual literacy to classify robotic emotion recognition in children. Unlike earlier residual neural network, we reduce the skip connection at the earlier part of the network and increase gradationally as the network go deeper. The progressive light residual network can explore further point space due to limiting the skip connection locally, which makes the network more vulnerable to disquiet which help to deal with overfitting problem for lower data. Experimental results on standard children feelings dataset show that the proposed approach showed a considerable gain in performance compared to the state of the art styles

3)G Mattavelli, E Barvas, C Longo

This study aims at assessing emotion recognition and demarcation in PD. Recognition of six facial expression was studied in order to clarify its relationship with motor, cognitive and neuropsychiatric symptoms. Perceptivity in differencing happy and fearful faces was delved to address controversial findings on impairment in early stages of emotion processing. To do so, seventy PD cases were tested with the Ekman 60 Faces test and compared with 46 neurologically unimpaired actors. Cases' performances were identified with clinical scales and neuropsychological tests. A subsample of 25 PD cases and 25 control actors were also tested with a backward masking paradigm for perceptivity in happiness and fear demarcation. Results showed that PD cases were bloodied in facial emotion recognition, especially for fearful expressions. The performance identified with perceptual, administrative and general cognitive capacities, but facial expression recognition poverties were present indeed in cognitively unimpaired cases. In discrepancy, cases' perceptivity in backward masking tasks wasn't reduced as compared to controls. Taken together our data demonstrate that facial emotion recognition, and sweat expression in particular, is critically affected by neurodegeneration in PD and related to cognitive capacities; still, it appears before other cognitive impairments. Saved performances in differencing shortly presented facial expressions, suggest unimpaired early stages of emotion processing.

4)A Peña, A Morales, I Serna, J Fierrez

This work explores facial expression bias as a security vulnerability of face recognition systems.

Despite the great performance achieved by state-of-the- art face recognition systems, the algorithms are still sensitive to a large range of covariates. To present a comprehensive analysis of how facial expression bias impacts the performance of face recognition technologies. Our study analyzes i) facial expression impulses in the most popular face recognition databases; and ii) the impact of facial expression in face recognition performances. Our experimental frame includes two face sensors, three face recognition models, and three different databases. Our results demonstrate a huge facial expression bias in the most extensively used databases, as well as a affiliated impact of face expression in the performance of state-of-the- art algorithms. This work opens the door to new exploration lines concentrated on mollifying the observed vulnerability.

III. EXISTING SYSTEM

Based on the computer vision research, Haar wavelet is used to image feature detection for object recognition. The success of the real-time face recognition systems are limited by the varying quality of images due to unreliable environment conditions. Human is counted manually for attendance systems. Counting the humans results in an inaccurate result and there will be no database proof.

IV. PROPOSED SYSTEM

Our proposed method can be classified into two main part which are detection on facial reorganization and speech reorganization. In facial reorganization, the CNN algorithm has been used in order to detect the emotional expression. Same as, the CNN algorithm has been used in detection of emotion by using speech signals.

V. SYSTEM ARCHITECTURE

A system architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system.

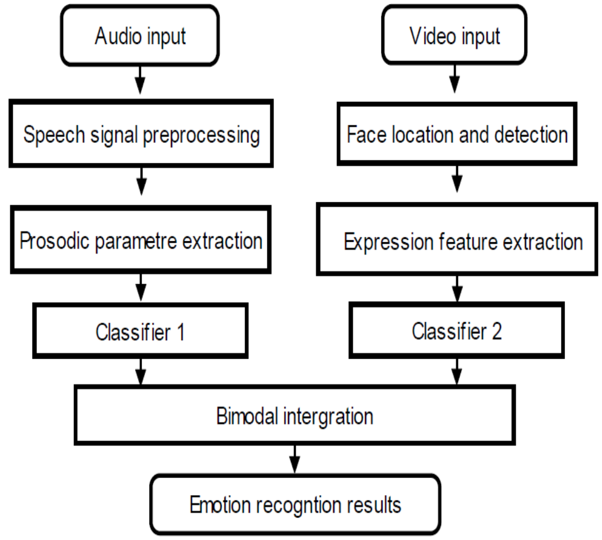


Figure 3.1.Conceptual model

VI. MODULE IMPLEMENTATION

1)MODULE LIST

- Training
- Dataset collection
- Preprocessing
- CNN layer Feature extraction
- Testing with result and analysis
- Signal collection
- Speech Preprocessing
- Feature extraction

2)MODULE DESCRIPTION

A. TRAINING

The augmentation of images are collected and features are extracted and trained into data based features are extracted created the data based data base. mat Testing of images we verified the person under the different testing.

B. DATA COLLECTION

The dataset is divided into two training set and test set. Each sample represents a business sign labeled as one of 2 classes. The shape of a business sign image is gauged to 256 × 256 pixels in 3 channel RGB representation. Below, there are a many arbitrary samples from the dataset vilo jones images are collected

C. PREPROCESSING

Variations that are inapplicable to facial expressions, similar as different backgrounds, illuminations and head acts, are fairly common in unconstrained scripts. Thus, before training the deep neural network to learn meaningful features,pre-processing is needed to align and homogenize the visual semantic information conveyed by the face Illumination and discrepancy can vary in different images indeed from the same person with the same expression, especially in unconstrained surroundings, which can affect in largeintra-class dissonances Image size

D. CNN LAYER FEATURE EXTRACTION

The conception of convolution neural networks. They're veritably successful in image recognition. The crucial part to understand, which distinguishes CNN from traditional neural networks, is the complication operation. Having an image at the input, CNN scans it numerous times to look for certain features. This scanning (complication) can be set with 2 main parameters stride and padding type. As we see on below picture, process of the first complication gives us a set of new frames, shown then in the alternate column (sub caste). Each frame contains information about one point and its presence in scrutinized image. Performing frame will have larger values in places where a point is explosively visible and lower values where there are no or little similar features. Latterly, the process is repeated for each of attained frames for a chosen number of times. In this design we chose a classic model which contains only two complication layers. The ultimate subcaste we're convolving, the further high- position features are being searched. It works also to mortal perception. To give an illustration, below is a veritably descriptive picture with features which are searched on different CNN layers. As you can see, the operation of this model is face recognition. You may ask how the model knows which features toseek.However, searched features are arbitrary, If you construct the CNN from the morning. Also, during training process, weights between neurons are being acclimated and sluggishly CNN starts to find similar features which enable to meet predefined thing, i.e. to fete successfully images from the training set. Different image point pattern.

E. TESTING WITH RESULT AND ANALYSIS

Live camera we take input image and preprocessing of image images. The preprocessed of images the person under video the person belongs of the person under expression mode.

F. SIGNAL COLLECTION

Human emotion signal is collected from the Kaggle website under different emotion signal.

Emotion recognition is the process of identifying human emotion. People vary widely in their accuracy at recognizing the emotions of others.

The datasets is collected from the Kaggle website for collection of different emotion recognition.

G. SPEECH PREPROCESSING

The preprocessing of signal the signal label range is resized into rows and columns of data.

The signal is continuous signal we converted into discrete signal .

H. FEATURE EXTRACTION

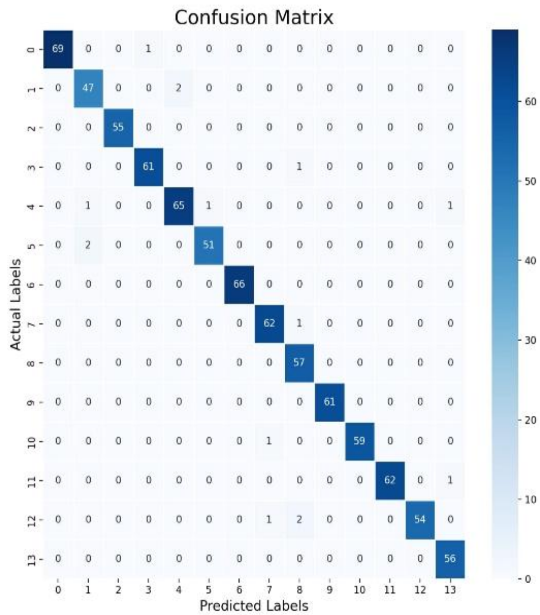
Extracting temporal or spectral features from the EEG signal's time or frequency domain, respectively, and finally, designing a multi-class classification strategy.

Feature quality dramatically increases the accuracy of the emotion classification strategy.

Cnn model is build model for feature extraction of training of signal for key features extraction.

VII. SCREENSHOT

Figure 4.1.DIFFERENT EMOTIONS



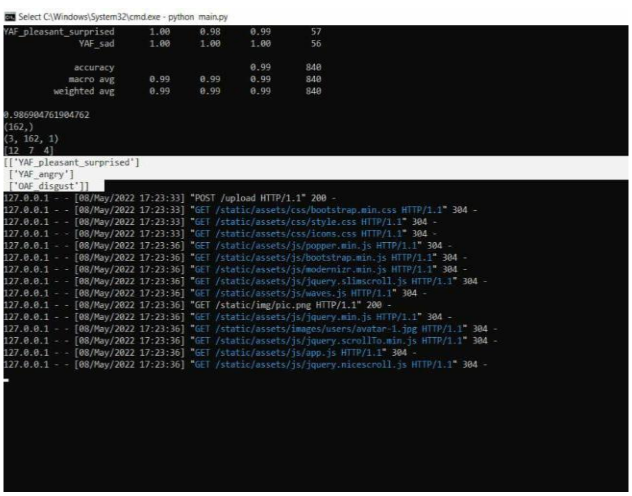
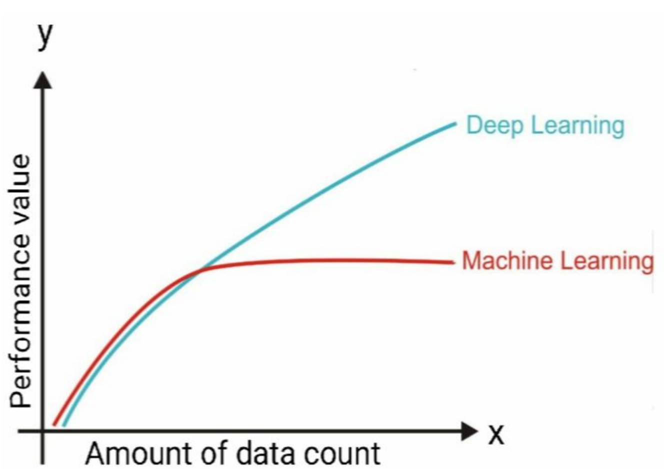
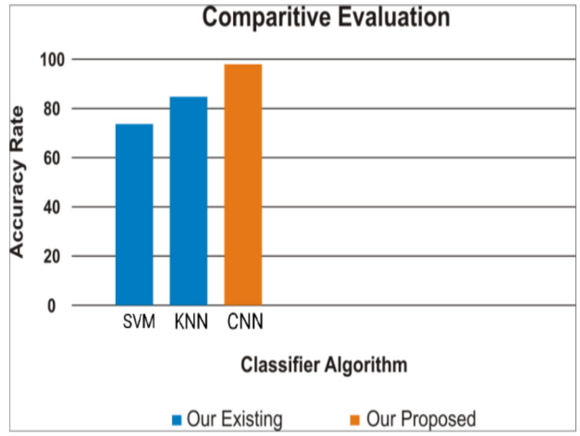


Figure 4.2. OUPUT OF SPEECH

VIII. RESULT AND COMPARISON





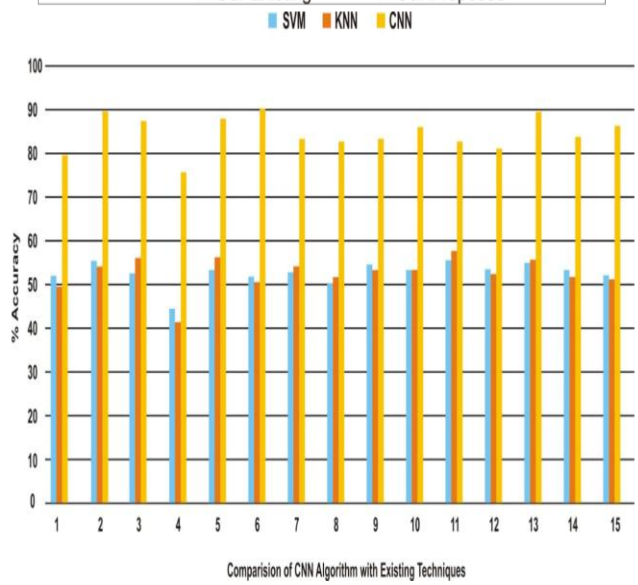


Figure 4.3.COMPARISON OF CNN

IX. CONCLUSION

We have presented a complete and fully automated approach for facial expression identification by simultaneously utilizing the face surface and face subsurface features. We presented a new algorithm for the face identification and recognition, which can more reliably extract the face features and achieve much higher accuracy than previously proposed facial identification approaches. The proposed approach present a very low degree of complexity, which makes it suitable for real-time applications. Depending upon the selected features and the measured region properties of the human face, the different expression of the human was further classified using CNN (convolution layers). The proposed method is superior compared with other state-of-the-art approaches and that the analysis of the general image quality of the face images reveals highly valuable information that may be very efficiently used in discriminate them from fake trait.

REFERENCES

[1] Paul Viola, Michael J. Jones, Robust Real-Time Face Detection, International Journal of Cumputer Vision 57(2), 2004.

[2] Ayushi Gupta, Ekta Sharma, NehaSachan and Neha Tiwari. Door Lock System through Face Recognition Using MATLAB. International Journal of Scientific Research in Computer Science and Engineering, Vol-1, Issue-3, 30 June 2013.

[3] Daniel Georgescu, A Real-Time Face Recognition System Using Eigenfaces, Journal of Mobile Embedded and Distributed Systems, vol. III, no. 4, 2011.

[4] Ole Helvig Jensen, Implementing the Viola-Jones Face Detection Algorithm, 2008.

[5] M. A. Turk and A. P. Pentland. Eigenfaces for Recognition. Journal of Cognitive Neuroscience, 3(1):71 - 86, 1991.

[6] M. A. Turk and A. P. Pentland. Face recognition using eigenfaces. IEEE Computer SocietyConference on Computer Vision and Pattern Recognition, CVPR 91, pages 586 - 591, 1991.

[7] I.Yugashini, S. Vidhyasri, K.Gayathri Devi, Design And Implementation Of Automated Door Accsessing System With Face Recognition, International Journal of Science and Modern Engineering(IJISME), Volume-1, Issue-12, November 2013.

[8] Liton Chandra Paul, Abdulla Al Sumam. Face Recognition Using Principal Component Analysis Method. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET),Volume 1, Issue 9, November 2012.

[9] A. C. Rafael and D. Sidney, “Affect detection: An interdisciplinary review of models, methods, and their applications,” *IEEE Trans. Affective Comput.*, vol. 1, no. 1, pp. 18–34, Jun. 2010.

[10] N. Sebe, H. Aghajan, T. Huang, N. M. Thalmann, and C. Shan, “Special issue on multimodal affective interaction,” *IEEE Trans. Multimedia*, vol. 12, no. 6, pp. 477–480, Oct. 2010.

[11] P. Ekman, T. Dalgleish, and M. E. Power, *Basic Emotions, Handbook of Cognition and Emotion*. Chichester, U.K.: Wiley, 1999.

[12] C. Darwin, *The Expression of Emotions in Man and Animals*. John Murray, 1872 (reprinted by the University of Chicago Press, Chicago, IL, 1965).

[13] J. F. Cohn, “Advances in behavioral science using automated facial image analysis and synthesis,” *IEEE Signal Process. Mag.*, vol. 27, no. 6, pp. 128–133, Nov. 2010.

[14] K. I. Chang, K. W. Bowyer, and P. J. Flynn, “Multiple nose region matching for 3-D face recognition under varying facial expression,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1695–1700, Oct. 2006.

[15] P. Ekman, W. V. Friesen, and J. C. Hager, *The Facial Action Coding System: A Technique for the Measurement of Facial Movement*. San Francisco, CA: Consulting Psychologist, 2002.

[16] M. J. Lyons, J. Budynek, A. Plante, and S. Akamatsu, “Classifying facial attributes using a 2-D Gabor wavelet representation and discriminant analysis,” in *Proc. 4th Int. Conf. Automatic Face Gesture Recognit.*, Mar. 2000, pp. 202–207.

[17] L. D. Silva and S. C. Hui, “Real-time facial feature extraction and emotion recognition,” in *Proc. 4th Int. Conf. Inform. Commun. Signal Process.*, vol. 3. Dec. 2003, pp. 1310–1314.

[18] I. Cohen, N. Sebe, Y. Sun, M. S. Lew, and T. S. Huang, “Evaluation of expression recognition techniques,” in *Proc. Int. Conf. Image Video Retrieval*, Jul. 2003, pp. 184–195.

[19] G. Guo and C. R. Dyer, “Learning from examples in the small sample case: Face expression recognition,” *IEEE Trans. Syst., Man, Cybern. B*, vol. 35, no. 3, pp. 477–488, Jun. 2005.

[20] M. Pantic and I. Patras, “Dynamics of facial expression: Recognition of facial actions and their temporal segments from face profile image sequences,” *IEEE Trans. Syst., Man, Cybern. B*, vol. 36, no. 2, pp. 433–449, Apr. 2006.

[21] K. Anderson and P. W. McOwan, “A real-time automated system for the recognition of human facial expressions,” *IEEE Trans. Syst., Man, Cybern. B*, vol. 36, no. 1, pp. 96–105, Feb. 2006.

[22] H. Gunes and M. Piccardi, “Automatic temporal segment detection and affect recognition from face and body display,” *IEEE Trans. Syst., Man, Cybern. B*, vol. 39, no. 1, pp. 64–84, Feb. 2009.

[23] Y. Wang and L. Guan, “Recognizing human emotional state from audiovisual signals,” *IEEE Trans. Multimedia*, vol. 10, no. 5, pp. 659–668, Aug. 2008.