

# Facial Features Based Hybrid Methods for Emotion Recognition

Bharati Dixit, Arun Gaikwad

**Abstract:** Effective human machine interaction systems are need of the time so the work carried out deals with one of such significant HMI tasks- automatic emotion recognition. The experimentation carried out for this study is focused to facial expressions based emotion recognition. Two techniques of emotion recognition based on hybrid features are designed and experimented using JAFFE database. The first technique referred as "Hybrid Method1" is designed around feature descriptor obtained through local directional number & principal component analysis and feed forward neural network used as classifier. The second technique referred as "Hybrid Method 2" is designed around feature descriptor obtained through histogram of oriented gradients, local binary pattern and Gabor filters. PCA- principal component analysis is used for dimensionality reduction of feature descriptor and k-nearest neighbors as classifier. The average emotion recognition accuracy achieved through method 1 and method 2 is 85.24% and 93.86% respectively. Effectiveness of both the techniques is compared on the basis of performance parameters such as accuracy, false positive rate, false negative rate and emotion recognition time. Emotion recognition has wide application areas so the work carried out can be applied for suitable application development.

**Keywords :** Emotion Recognition, Facial expressions, Local directional number Histogram of oriented gradients, Hybrid Features.

## I. INTRODUCTION

Automatic human emotion recognition is much sought after and useful for effective human machine interaction from the perspective of many applications like cognitive assessment, pain monitoring, driver alert system, behaviour analysis etc. Emotions are integral part of human being [1] and impact social skills, interpersonal skills and learning capabilities & decision making capabilities to a large extent

Emotion recognition can be implemented through verbal approach and non-verbal approach. Scientists across the globe have experimented methods using single modality of input as well as multimodal inputs. The inputs like Gait, facial

Expressions, EEG signals, body posture and gestures etc. can be used to design emotion recognition system under category of non-verbal approach. The paper is based on the findings obtained through experimentation based on facial

expressions as input

To design any emotion recognition system, the feature extraction is very significant as it plays crucial role in modelling classifier [2]-[3] and hence in achieving better performance parameters for the designed system. The features can be derived from confined region (i.e. local regions) of interest or the complete region to be operated is considered as single one (i.e. global region) and features are extracted from this region. There are variety of methods available for feature extraction for local as well as global region and called as local features (model based methods) [4]-[5] and global features (appearance based methods)[6] respectively. Some of the popularly used methods for local feature extraction are LDN - Local directional number, LBP- local binary pattern, HOG - Histogram of gradients, geometric features etc. Some of the popularly used global feature extractors are Gabor Filters, wavelets, PCA - Principal Component Analysis, LDA - Linear discriminant analysis etc. Both local and global feature extractors have proved to provide good performance and are in practice. Combinations of local and global features provide benefits of both the local and global features and called as hybrid features [7] or hybrid method.

The study undertaken explores hybrid features based emotion recognition techniques for experimentation. Paper is further developed under different sections. Section II provides broader perspective of emotion and emotion recognition. Section III provides implementation methodology. Section IV elaborates on experimental results and performance analysis. Section V concludes the findings of experimentation.

## II. EMOTION AND EMOTION RECOGNITION - BROADER PERSPECTIVE

The Emotion is defined as a psychological state that arises spontaneously and is accompanied by physiological changes. Human emotions are revealed through speech, face, Gait, EEG signals, Body posture and gestures. The work undertaken for research is based on facial expressions. Face is a multi-signal and multi-message system. The face broadcasts messages about emotion, mood, attitude, character, intelligence, attractiveness, age, gender, race etc. Facial cues makes available signals such as static, slow and rapid [8]- [9].

**Revised Manuscript Received on September 25, 2019**

**Bharati Dixit**, Associate Professor, Department of Computer Engineering, MIT College of Engineering, Pune and Research Scholar, SCOE, Pune, India

**Dr. Arun Gaikwad**, Professor and former Principal, Department of Electronics and Telecommunications, ZCER, Pune, India.

Facial aspects and facial features gives information about static signals. The slow signals reflects variation in facial appearance with respect to time. Rapid signals are generated through movement of facial muscles. Correct estimation of emotions can be made from these rapid facial signals - the blueprints – which distinguish each of the primary emotions like anger, fear, disgust, sad, surprise, happy and neutral.

Emotional corpus is available in the form of naturalistic emotions, simulated emotions and induced emotions. Popularly used emotion recognition systems which are based on facial expressions can use local features, global features or hybrid features. The work undertaken explores hybrid features and titled as Hybrid Method1 and Hybrid Method 2 for further references in the paper. Emotion recognition has wide application range such as e-learning, driver alert system, cognitive assessment, behavior analysis etc. The implementation methodology for hybrid method 1 and hybrid method 2 are discussed further in section III.

III. IMPLENTATION METHODOLOGY

Emotion recognition is implemented using two different techniques referred as "Hybrid Method 1" & "Hybrid Method 2". Details of each method such as system block diagram, proposed algorithm, logical flow of implementation is discussed further in subsections 3.1 and 3.2.

3.1 Emotion recognition - hybrid method 1

As the method is based on hybrid features so LDN - local directional number is used as local feature extractor and PCA - principal component analysis functions as global feature extractor. The hybrid features are formed through fusion of LDN and PCA features. This hybrid feature descriptor is further fed to FFNN classifier which provides output in the form of six emotion classes. Process flow diagram is as shown in Figure 1.

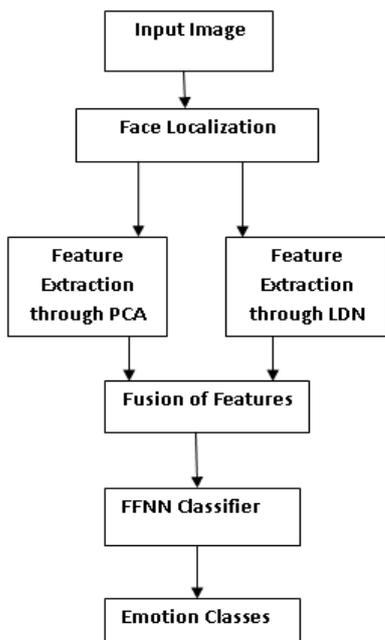


Figure 1: Hybrid Method 1 - flow diagram

Computation of feature descriptor through LDN [10] involves convolution of original image with the mask which

results in extraction of edge response image. The most negative and the most positive directions are used to encode the information. The generated six bit binary number is called LDN image. LDN is computed for each region of face image and Histogram of each region is also calculated. All these histograms are concatenated to create feature descriptor.

Principal component analysis can effectively represent images in low dimensional subspace. While doing so PCA does not require the knowledge of geometry and reflectance of face image. PCA [11] decomposes face image into Eigen faces which work as initial principal component of feature vector.

Neural network is very effective tool to perform many machine learning tasks [12]. Feature vector, obtained through fusion of LDN and PCA features, is provided as input to FFNN. Feed Forward neural network is used here for classification of images into six classes. Experimentation is performed on JAFFE database and sample screenshots and performance analysis is discussed in details in section 4.

3.2 Emotion recognition - hybrid method 2

As the method is designed using HOG and LBP as local feature descriptor and Gabor filter bank as global feature descriptor. The combination of HOG, LBP and Gabor filters provides feature descriptor which is hybrid in nature but at the same time quite high in dimension. PCA is used for dimensionality reduction and provides optimum feature vector which is further fed to kNN classifier for performing classification into six emotion classes. Implementation steps for Hybrid Method 2 are shown in the form of flow diagram in Figure 2.

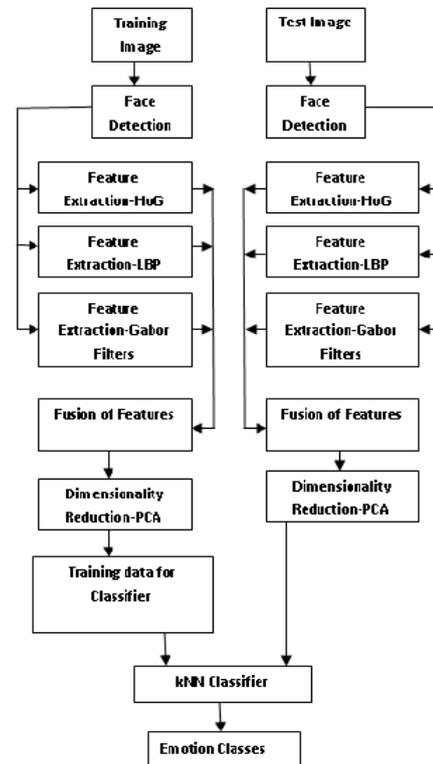


Figure 2: Hybrid Method 2- flow diagram

Details of some important steps is as follows: One of the local feature descriptors used here is HOG - Histogram of oriented gradients [13] which is popularly used in of pattern recognition. HOG is based on the idea that appearance and shape describes any image. The image is partitioned into cells and histogram is computed for pixel within the cells. Concatenation of all these histograms results in feature vector.

HOG has inbuilt advantage of illumination invariance as it operates on local cells which are invariant of photometric and geometric transformation. The steps involved in calculation of HOG descriptor is as represented in Figure 3.

Steps	Details of the steps
S1	Acquire Image to be treated as input image
S2	Normalize Image for Gamma and Color
S3	Compute Gradients
S4	Register weighted vote into spatial and oriented cells
S5	Perform contrast Normalization for all the blocks
S6	Perform histograms for all the block regions
S7	Concatenate Histograms of the block regions over detected window
S8	Formation of feature vector

Figure 3: Steps to obtain HOG feature vector [13]

Local Binary Pattern [14]-[16] is simple and effective operator used for real world applications as it offers robustness against gray scale changes and computational simplicity. Use of threshold function results in providing a binary number to each pixel. The neighbourhood pixels contribute in this process and popularly used neighborhood is 3x3.

The LBP operator labels the pixels of an image by thresholding the 3x3 neighborhood of each pixel with the value of the central pixel and concatenating the results binomially to form a number. The threshold function for the basic LBP is represented as Equation (1).

$$f(I(Z_0), I(Z_i)) = \begin{cases} 0 & \text{if } I(Z_i) - I(Z_0) \leq \text{threshold} \\ 1 & \text{if } I(Z_i) - I(Z_0) > \text{threshold} \end{cases} \quad (1)$$

Feature descriptor derived through LBP is used during fusion process while forming final feature vector. Gabor Filter bank [17]-[18] of 5 orientations and 8 scales are used for experimentation. Gabor filters are explained as in equation (2)

$$F(x, y) = \exp(-\frac{(X^2 + \gamma^2 Y^2)}{2\sigma^2}) * \cos(2\pi X/\lambda) \quad (2)$$

$$X = x \cos\theta + y \sin\theta, \quad Y = -x \sin\theta + y \cos\theta$$

Where  $\theta$  is orientation,  $\sigma$  is effective width;  $\lambda$  is wavelength and  $\gamma$  is aspect ratio.

Concatenation of features obtained through Gabor filter bank with features obtained through HOG and LBP is carried out. The feature vector obtained through fusion of features has curse of dimensionality so to reduce the dimensionality of feature vector PCA is applied. PCA is a multivariate technique available in various models and depending upon the requirement of application chosen accordingly. Simplification of the data set details and analysis of the variables is achieved with application of PCA [19]-[20].

PCA computes new variables called principal components which are obtained as linear combinations of the original variables. The values of these new variables are called factor scores which are interpreted geometrically as projections of the observations on to principal components. The first principal component is required to have the largest possible variance. The second component should be orthogonal to the first one and must have maximum possible inertia. Similarly other principal components are calculated.

k-nearest neighbors classifier [21]-[22] provides output as a class membership. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors. k is a positive integer, typically small.

Hybrid Method 2 provides output in the form of emotion profile which reveals degree of presence/absence of each emotion for the image under study. For the experiments carried out, the bar with lowest amplitude on the profile represents dominant emotion in the image.

This approach is very effective for naturalistic emotional corpus.

#### IV. EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS

Experimentation is carried out on publically available JAFFE database [23] for six universally accepted emotions. The experimental results in the form of snapshots of intermediate steps and performance analysis are discussed in details in subsections 4.1 and 4.2 respectively.

##### 4.1 Experimentation results

The hybrid methods designed for facial expressions based emotion recognition are experimented on JAFFE database for six universally accepted emotions. Experimentation is carried out on 183 images of JAFFE database. Some of the snapshots obtained as intermediate results during experimentation for Hybrid method 1 are as represented from Figure 4(a) to 4(d) for the detected emotion Sad.

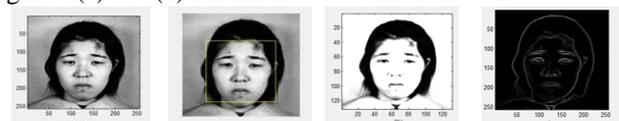


Figure 4(a) Original Image  
Figure 4(b) Localized Face Region  
Figure 4(c) Application of PCA  
Figure 4(d) Application of LDN

Figure 4: Snapshots of Hybrid method 1

## Facial Features Based Hybrid Methods for Emotion Recognition

The second method experimented on JAFFE database is referred as hybrid method 2 and the output is represented in the form of emotion profile and text box. Sample screen shots of input, intermediate outputs and final output is as represented from Figure 5(a) to 5(d) for the detected emotion anger.

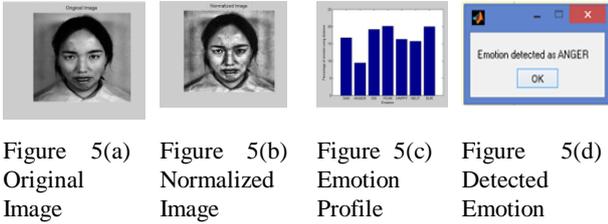


Figure 5(a) Original Image    Figure 5(b) Normalized Image    Figure 5(c) Emotion Profile    Figure 5(d) Detected Emotion

Figure 5: Snapshots of Hybrid method 2

Performance parameters of Hybrid method 1 and Hybrid method 2 are calculated in the form of TPR-true positive rate FPR/FAR -false acceptance rate and FNR/FRR -false rejection rate. Analysis of both the methods is conducted in following subsection

### 4.2 Performance analysis

The experimented methods are tested for accuracy, false acceptance rate, and false rejection rate and emotion recognition time. The database is divided into training dataset and testing dataset. 65% of the overall sample space is used as training data and rest 35% sample space is used as testing data. After testing the images the result is represented in the form of confusion matrix.

Hybrid method 1 and Hybrid method 2 are experimented on JAFFE database and results are represented in the form of confusion matrix as shown in Table 1 and Table 2 respectively.

Table 1. Confusion matrix for JAFFE Database for Hybrid Method 1

PC AC →	SA	AN	DI	FE	HA	SU
↓						
SA	25	1	2	3	0	0
AN	1	24	1	1	0	2
DI	1	1	28	0	0	0

Table 4. Average Values of Performance Parameters

Implementation Method	Average TPR (%)	Average FPR / FAR (%)	Average FNR / FRR (%)
Hybrid Method 1	85.24	3.34	14.75
Hybrid Method 2	93.86	1.025	5.92

Graphical representation of class wise true positive rate, false positive rate, false negative rate and average values of TPR, FPR and FNR for both the methods are shown from Figure 6 to Figure 9.

FE	2	2	1	27	0	0
HA	0	0	1	0	28	2
SU	1	0	1	1	3	24
<b>Total</b>	30	28	34	32	31	28
<b>Accuracy</b>	<b>80.6</b>	<b>82.</b>	<b>93.3</b>	<b>84.3</b>	<b>90.3</b>	<b>80</b>

SA-Sad, AN-Anger, DI-Disgust, FE-Fear, HA-Happy, SU-Surprise, PC-Predicted Emotion Class, AC- Actual Emotion Class

Average Accuracy =  $(80.65\% + 82.76\% + 93.34\% + 84.36\% + 90.32\% + 80.00\%) / 6 = 85.24\%$

Performance parameters such as class wise TPR and FNR are calculated for both Hybrid Method 1 and Hybrid Method 2 and tabulated in Table 3.

Table 3. Class wise TPR, FPR and FNR for Hybrid Method 1 and Hybrid Method 2

EM	TPR (%)		FPR/FAR (%)		FNR/FRR (%)	
	HM 1	HM 2	HM 1	HM 2	HM 1	HM 2
SA	80.6	93.54	3.67	0.69	19.34	3.22
AN	82.7	90.00	2.94	0.68	17.24	10.0
DI	93.3	93.10	4.47	2.02	6.66	3.44
FE	84.3	96.87	3.73	1.39	15.62	3.12
HA	90.3	93.54	2.29	1.37	9.67	6.45
SU	80.0	96.66	2.94	0.00	20.00	3.33

EM - Emotions, HM1 - Hybrid Method 1, HM2 - Hybrid Method 2

SA -Sad, AN - Anger, DI - Disgust, FE - Fear, HA -Happy, SU - Surprise

While calculating the performance parameters class wise images used for experimentation are 31 for class sad, 29 for class anger, 30 for class disgust, 32 for class fear, 31 for class happy and 30 for class surprise. Average values of true positive rate, false positive rate and false negative rate as represented in Table 4.

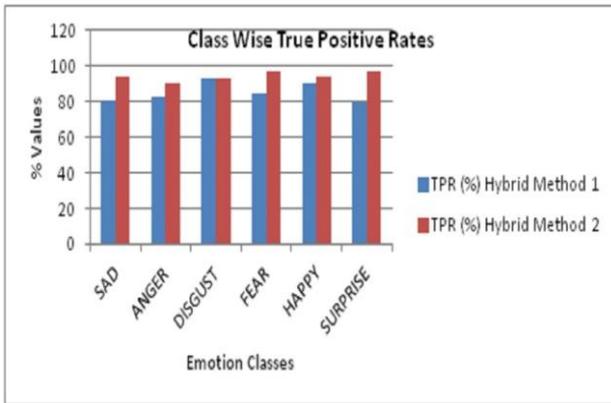


Figure 6: Emotion Class wise True Positive Rate

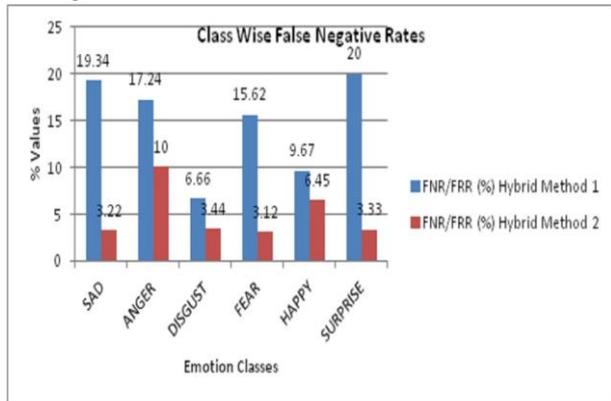


Figure 8: Emotion Class wise False Negative Rate

Emotion recognition time obtained through Hybrid method 1 is 1.02 Seconds and through Hybrid Method 2 is 1.108 seconds. Comparative analysis of both the hybrid methods results in the conclusion that hybrid method 2 has outperformed hybrid method 1 in terms of emotion recognition accuracy whereas both the methods are at par as far as of emotion recognition time is concerned.

## V. CONCLUSION

Emotion recognition methods designed around hybrid features are studied and experimented on JAFFE database and compared on the basis of emotion recognition accuracy and emotion recognition time. Both the methods provide emotion recognition accuracy more than 85% and emotion recognition time around 1 second. The 'Hybrid Method 2' designed using combination of HOG, LBP and Gabor features and KNN as classifier outperforms the 'Hybrid Method 1' designed using combination of features obtained through LDN and PCA and FFNN as classifier.

Classification accuracy is better for Hybrid method 2 as compare to Hybrid method 1 for all the classes except the emotion class disgust. For emotion class disgust both the methods provide results which are at par with each other. Any of the method can be implemented for the application such as cognitive assessment system, driver alert system, e-learning, patient pain monitoring, recommendation systems, healthcare products etc.

The experimentation can be extended for pose variation for image database and on image sequences as well due to growing demand of video analytics applications. The emotion recognition accuracy can be enhanced further by variation in feature fusion process and by modeling another

classifier.

## REFERENCES

- [1] K. Oatley, et al., "The Experience of Emotions in Everyday Life," Journal of Cognitive Emotions, vol. 8, pp. 369-381, 1994
- [2] A.N. Ratna et al., "Comparative Analysis Of Machine Learning KNN, SVM and Random Forests Algorithm For Facial Expression Classification," International Seminar On Application For Technology Of Information And Communication (Isemamtic), pp. 163-168, 2016.
- [3] F. Nacer et al., "Exemplar-Based Facial Expression Recognition," Elsevier Information Science, vol. 460, pp. 318-330, 2018.
- [4] B. Allaert, et al., "Impact Of Face Registration Techniques On Facial Expression Recognition," Research Article Signal Processing: Image Communication, vol. 61, pp. 44-53, 2018
- [5] Junkai Chen, et al., "Facial Expression Recognition in Video with Multiple Feature Fusion," IEEE Transactions on Affective Computing, vol. 1, pp.1-13, 2016.
- [6] D. Hong-Bo, et al., "A New Facial Expression Recognition Method Based On Local Gabor Filter Bank And PCA Plus LDA," International Journal Of Information Technology, vol. 11, Issue 11, pp. 86-96, 2005.
- [7] Othmane El Meslouhi, et al., "Unimodal Multi-Feature and One dimensional Hidden Markov Models for Low Resolution Face Recognition," International Journal of Electrical and Computer Engineering, vol. 7, Issue 4, pp. 1915-1922, 2017
- [8] A.T. Madhumita, et al., "Image Based Facial Macro-Expression Recognition Using Deep Learning," International Conference on Digital Image Computing: Techniques and Applications - DICTA on small Datasets, Australia, pp.1-7, 2017.
- [9] Paul Ekman, et al., "Unmasking the Face" published by Maylor Books in 2003.
- [10] Adin Ramirez, et al., "Local Directional Number Pattern for Face Analysis: Face and Expression Recognition," IEEE Transactions On Image Processing, vol. 22, Issue 5, pp. 1740-1752, May 2013.
- [11] YaxinSun et al., "Cognitive Facial Expression With Constrained Dimensionality Reduction," Journal of Neurocomputing, vol. 22, pp. 397-408, 2017.
- [12] M. Hongying, et al., "Time-Delay Neural Network for Continuous Emotional dimension Prediction From Facial Expression Sequences," IEEE Transactions On Cybernetics, vol. 46, Issue 4, 916-929, 2016.
- [13] N. Dalal, et al., "Histograms of oriented gradients for human detection," IEEE Computer Society Conference on pattern recognition and Computer Vision, 2005.
- [14] Z. Baochang, et al., "Local Derivative Pattern Versus Local Binary Pattern: Face Recognition with High-Order Local Pattern Descriptor," IEEE Transactions On Image Processing, vol.19, Issue 2, pp. 533-544, 2010.
- [15] International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 2, February 2012) 6 Detection of Misbehaving Nodes in Ad Hoc Routing Isha V. Hatware, Atul B. Kathole, Mahesh D. Bompilwar
- [16] International Journal of Scientific & Engineering Research, Volume 4, Issue 6, June-2013 39 ISSN 2229-5518 "SURVEY OF TOPOLOGY BASED REACTIVE ROUTING PROTOCOLS IN VANET" Atul B.Kathole, Yogadhar Pande.
- [17] Z. Ligang, et al., "Facial Expression Recognition Using Facial Movement Features," IEEE Transactions On Affective Computing, vol.2, Issue 4, pp. 219-229, 2011.
- [18] S. L. Happy, et al., "Automatic facial expression recognition using features of salient facial patches," IEEE transactions on Affective Computing, vol. 6, Issue 1, pp.1-12, 2015.
- [19] Y. Jian, et al., "Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition," IEEE Transactions On Pattern Analysis And Machine Intelligence, vol. 26, Issue 1, pp. 131-137, 2004.
- [20] S. Venkatramaphani kumar et al., "Face Recognition with Modular Two Dimensional PCA under Uncontrolled Illumination variations," International Journal of Electrical and Computer Engineering, vol. 6, Issue 4, pp. 1610-1616, 2016.
- [21] D. Coomans, et al., "Alternative K-Nearest Neighbour Rules In Supervised Pattern Recognition: Part 1. K-Nearest Neighbour Classification By Using Alternative Voting Rules" Analytica Chimica Acta., vol. 136, pp. 15-27, 1982.
- [22] B. S. Everitt, et al., "Miscellaneous Clustering Methods, in Cluster Analysis," 5th Edition, John Wiley & Sons, Ltd, Chichester, UK, 2011.
- [23] Standard Dataset Available: [http://www.kasrl.org/jaffe\\_download.html](http://www.kasrl.org/jaffe_download.html)