Fusion of multi representation and multi descriptors for facial expression recognition

M Santosh¹ and A Sharma²

¹,² Computer Science and Engineering, M. M. Engineering College, Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala – 133207

¹manta.ahlawat2@gmail.com, ²asharma@mmumullana.org

Abstract. Facial Expression Recognition has become vital for efficient Human Computer Interaction. In this paper, we propose effective facial expression recognition approach for recognizing six basic facial expressions. Our approach consists of three main phases which are: (1) face detection and pre-processing, (2) features extraction and (3) facial expression classification. The face pre-processing phase is performed using the facial landmarks. After the face is aligned and cropped, facial regions of interest (eyes, nose and mouth) are detected. In the features extraction phase, we used Histogram of oriented gradients (HOG), Local Binary Pattern (LBP) and the fusion of the two features. For the last step, Support Vector Machine (SVM) is used to recognize the facial expression. To evaluate the performance of our approach, we used three popular datasets which are The Extended Cohn-Kanade (CK+), The Japanese Female Facial Expression (JAFFE) and Oulu-CASIA NIR-VIS dataset (CASIA). In addition, 10 folds cross-validation scheme is used to evaluate the performance of our approach. Our proposed fusion of multi representations and multi descriptors achieves better or competitive performance compared with the state-of-the-art methods. The accuracies of our approach are 99.18%, 95.77% and 99.09% for CK+, JAFFE and CASIA, respectively. The results prove the efficiency of our approach although the challenging conditions from one dataset to another.

1. Introduction

Facial Expression Recognition methodologies are categorized into two groups - deep learning methods and hand crafted methods. The hand crafted methods include Local Phase Quantization (LPQ), Local mean binary pattern (LMBP), Histograms of oriented gradients (HOG) and Boosted Local Binary pattern (BLBP) whereas Deep learning methods [1] include Convolution neural networks CNN and alexnet architecture.

Various researches have been done to develop methods for recognizing human facial expressions. Chao Qi, Min Li, Qiushi Wang and Huiquan Zhang in [2] proposed a cognitive and mapped binary pattern method for recognition of expression and evaluated their method on CK+ dataset. To extract the facial contours they used LBP operators and segmented face area by using pseudo 3D model into sub regions. Two classifiers - SVM and Softmax were used for classification with resulting accuracy of 87.8%. Biao Yang, Jinneng Cao, Rongrong Ni and Yuyu Zhang in [3] proposed weighted mixture deep neural network (WMDNN) for processing two channels of facial images i.e. greyscale images and their corresponding local binary pattern (LBP) and result was calculated by using SoftMax.
classification. Three datasets: CK+, JAFEE, and Oulu-CASIA provided an average accuracy of 97%, 92.2% and 92.3% respectively. Y. Ding, Q. Zhao, B. Li and X. Yuan in [4] proposed a method to classify facial expressions by using image sequence method from videos. They used Taylor Feature Pattern (TFP) which is based on LBP and Taylor expansion and obtained effective features for the two datasets i.e. JAFEE and CK. Different distance measures i.e. Chi-square distance, Modified G-statistic (MG) and histogram intersection (HI) were used in this experiment and provided highest recognition rate of 0.9484, 0.9249 and 0.9343 and 0.9186, 0.9286 and 0.9143 for JAFEE and CK datasets respectively.

S L Happy, Aurobinda Routray in [5] used facial patches to extract discriminative features which in turn are used for recognizing facial expressions. They used six facial expressions and classified using one-against-one classification technique. Moreover, they also proposed a detection technique named automatic learning free facial landmark and provided an overall accuracy of 91.8% and 94.14% on JAFEE and CK+ datasets. In [6], Bing-Fei Wu and Chun-Hsien Lin proposed a method to correct misclassified samples of the expressions and used Weighted Center Regression Adaptive Feature Mapping (W-CR-AFM). The datasets used in this research were Cohn-Kanade (CK+), Radbound Faces Dataset (RaFD), and Amsterdam Dynamic Facial Expression Set (ADFES). CNN model was used to classify the images after pre-processing was regarded to be a generic model. GM and W-CR-AFM together provided an accuracy of 89.84%, 96.27% and 92.7% for CK+, RaFD and ADFES dataset respectively.

M. Hassan, Md. Golam Rabiul Alam, M. Z. Uddin, M. Huda, Ahmad Almogren, G. Fortino in [7] used Zygomaticus Electromyography (zEMG), Photoplethysmogram (PPG) and Electro-Dermal Activity (EDA) signals obtained from sensors to extract facial features. They used fine gaussian SVM and deep belief network to correctly classify feature fusion vector of five emotions and provided an overall accuracy of 89.3% on DEAP dataset. In [8], Chen-chiung Hsieh, Mei Hua Hsih, Mengkai Jiang, Yunmaw Cheng, Enhui Liang proposed method based on directional gradient operators to acquire facial features gathered from CK+ dataset. SVM classifier was used for classification of six facial expressions and achieved an average recognition rate of 94.7%.

Yan Ouyang, N. Sang, Rui Huang in [9] proposed technique which used HOG and LBP features for extracting facial features. Sparse representation based classification (SRC) was used to classify images gathered from CK+ dataset and provided an overall recognition rate of 95.64% for six classes. In [10], Xiaoming Zhao, Xugan Shi & Shiqing Zhang have proposed an algorithm for recognizing facial expressions by using deep neural network techniques. This research was done on images from CK+, MUG and RaFD datasets. CNN model was used for classification of dataset and provided highest recognition rate of 99.3% for CK+ dataset.

Selective transfer machine is used for analysis of facial expressions in [11] by Wen-Sheng Chu, Fernando De la Torre, and Jeffrey F. Cohn. This approach provided an average recognition rate of 96.4% and 86.5% for CK+ and GEMEP-FERA datasets. An innovative technique employing recognition facial expressions is proposed in [12] by Matlovic, Tomas & Gaspar, Peter & Moro, Robert & Simko, Jakub & Bielikova, Maria, in which electro encephalopathy (EEG) signals are used for detection of human emotions. This technique provided an overall accuracy of 53%. With the help of image processing, we are able to do various experiments ranging from medical applications [13][14] to emotion recognition.

In [15], Kamlesh Mistry, Li Zhang, Siew Chin Neoh, Chee Peng Lim, and Ben Fielding proposed a method for recognition of facial expression which employs particle swarm optimization (PSO) for optimization of features along with modified LBP and micro genetic algorithm (mGA). Their approach provided an overall accuracy of 94.66% for CK+ dataset by training it on ensemble based SVM classifier.

In our previous work [16], we have used HOG features and used Canberra distance for classifying the facial expressions into emotions whereas in [17] we used fusion of LBP and HOG features with Canberra distance classifier. Table 1 summaries the comparison of all the previous papers.
Table 1. Summary of the comparison of all the previous papers.

| [ref] | Methodologies | Accuracy | Dataset |
|-------|---------------|----------|---------|
| [2]   | Pseudo 3D model, SBM, Softmax | 87.8%    | CK+     |
| [3]   | Weighted mixture deep neural network (WMDNN), LTP, Softmax | 97%, 92.2%, 92.3% | CK+, JAFFE, CASIA |
| [4]   | Taylor Feature Pattern (TFP), LBP, Chi-square distance, Modified G-statistic (MG) and histogram intersection (HI) | 0.9484, 0.9286 | JAFFE, CK |
| [5]   | One-against-one classification technique, facial landmark | 91.8%, 94.14% | JAFFE and CK+ |
| [6]   | Weighted Center Regression Adaptive Feature Mapping (W-CR-AFM), CNN | 89.84%, 96.27%, 92.7% | (CK+), (RaFD), (ADFES) |
| [7]   | Fine gaussian SVM and deep belief network, feature fusion | 89.3% | (EDA), (PPG) and (zEMG) |
| [8]   | Directional gradient operators, SVM | 94.7% | CK+ |
| [9]   | HOG, LBP and Sparse representation based classification (SRC) | 95.64% | CK+ |
| [10]  | CNN model | 99.3% | CK+, MUG and RaFD |
| [11]  | Selective transfer machine | 96.4%, 86.5% | CK+, GEMEP-FERA |
| [12]  | Machine learning | 53% | electro encephalopathy (EEG) signals |
| [15]  | Particle swarm optimization (PSO), LBP and micro genetic algorithm (mGA), ensemble | 94.66% | CK+ |

2. Datasets

It becomes important to use standard datasets for training and testing purpose because it aids in performance comparison of different algorithms. The methodology performance is evaluated on three different datasets which are as follows:

- Cohn-Kanade Dataset (CK+)
- Japanese Female Facial Expressions (JAFFE)
- Oulu-CASIA NIR-VIS dataset.

CK+ dataset [18] was released with additional 107 sequences and 26 subjects to original CK dataset (486 sequences and involved from 97 subjects). Total 593 images were collected from 123 persons which were Afro American and Euro American. Mostly images are gray scaled but some are colored. Resolution of some part of dataset is 640 × 490 pixels and other part is 640 × 480 pixels. It contains 8 different facial expressions six basic expressions and one contempt expression and neutral face. As CK+ is sequence based dataset, in our experiment, we have used peak frames (last three frames) for identification of expressions.
Table 2. Summary of Datasets.

| Dataset | Facial Expressions | Total Subjects | Total Images | Image Type          | Resolution      |
|---------|--------------------|----------------|--------------|---------------------|-----------------|
| CK+     | 8                  | 123            | 593          | Mostly gray         | 640* 490        |
| JAFFE   | 7                  | 10             | 213          | Gray                | 256* 256        |
| CASIA   | 6                  | 80             | 1440         | Color video sequences| 320*240        |

JAFFE dataset [19] was collected at Kyushu University. It is divided into seven facial expressions: surprise, neutral, sadness, fear, anger, happiness, and disgust. It was collected from ten female models of Japan. Each Facial Expressions consists of three images and in total 123 images in gray scale were captured with resolution of 256 × 256.

CASIA [20] dataset was collected from total 80 subjects. These subjects posed six different facial expressions: happiness, anger, surprise, sadness, disgust and fear. Images were collected in three conditions of illumination: weak, normal, and dark. Dataset was collected at a rate of 25 frames per second with resolution of 320×240 pixels. Table 2 presents the summary of expressions depicted in the datasets used. Similar to CK+ dataset, we have used peak frames for identification of expressions.

Table 3. Facial Expressions included in each dataset.

| Dataset | Surprise | Neutral | Fear | Anger | Happiness | Contempt | Disgust | Sadness |
|---------|----------|---------|------|-------|-----------|----------|---------|---------|
| CK+     | ✓        | ✓       | ✓    | ✓     | ✓         | ✓        | ✓       | ✓       |
| JAFFE   | ✓        | ✓       | ✓    | ✓     | ✓         | ✓        | ✓       | ✓       |
| CASIA   | ✓        | ✿      | ✓    | ✓     | ✓         | ✿        | ✓       | ✓       |

3. Proposed Approach

In the proposed approach, we have detected and aligned the face through landmark detection. Aligned facial landmarks are used to detect the face, eyes, mouth and nose regions. We have used multi-descriptors to extract the features which are Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP) and Fusion of HOG and LBP. Firstly HOG is used to extract the features of all the representations (whole face, eyes, mouth and nose regions) and then fused them alongside. Similarly, we concatenated all facial representations using LBP descriptors. Lastly, we concatenated fused LBP features and fused HOG features alongside each other. The features of each method are fed to 10-fold Support Vector Machine. Proposed approach is tested on three different datasets which are available online. Figure1 depicts the proposed approach.
3.1. Face Detection
For the face and facial component (eyes, noise, lips) of interest detection facial texture characteristics are mainly focused. To detect and retrieve the face region, Fusion transformed and Shallow features [21] (FTDS) approach is used.

First of all, the face is detected through 68 facial landmarks using Dlib library [22]. 2D location of eyes is determined to perform in-plane rotations. After rotation – three furthest points are selected as boundaries of the face in left, right and bottom.

\[ d_2 = 0.6 \times d_1 \]

Where
\( d_1 \) is distance from lower boundary to eyes
\( d_2 \) is distance from upper boundary to eyes

Then face is cropped using the four boundaries and resized to 240 *240 pixels.

After the face is detected, face region parts of the face are extracted.
Fig 3. Shows the facial regions for CASIA, JAFFE and CK+ Dataset from top to bottom.
3.2. **Feature Extraction**

Histogram of Oriented Gradients (HOG) and Local Binary pattern (LBP) are used for feature extraction. The HOG feature centres around the shape or structure of an image. It can give the edge course too. This is finished by obtaining the orientation and gradient. These orientations are determined in 'restricted' partitions. This implies that the total picture is separated into small portions and for every portion, the angles and direction is determined.
At last the HOG [23] produces a Histogram for every one of these portions independently. The histograms are made utilizing the orientations and gradients of the value of image pixel. It depends upon the cells which are the localized part and gradient depends upon the orientations. So HOG is dependent upon the gradient orientation in cells (localized regions). However, it is an easy and simple method to pass the irregular form of input object and is also good to a change that happens in the geometry and illumination.

Local Binary Pattern (LBP) [24] is used for classification purpose in computer vision as visual descriptor. It is a straightforward yet exceptionally productive texture based operator which performs the labelling of pixels of a picture by thresholding the area of every pixel and results is saved as binary number. Basically it divides the image into a window of different cells. Every pixel in each cell is compared to its neighbouring pixel. Then 0 and 1 is generated for each cell accordingly, following that histogram is computed.

**Figure 5.** Feature extraction procedure of segmented images using LBP

3.3. Support Vector Machine

Support vector (SVM) is a supervised learning algorithm, in which, we plot every point of data in space of n-dimension (The total features in feature vector is value of n) with the estimation of each element being the estimation of a specific arrangement. Cross-validation is a measurable method used to appraise the ability of machine learning models. It is regularly used and applied in AI to select and compare a structured framework for a given prescient issue since it is simple to execute, straightforward. It is a re-sampling methodology used to assess AI models on a chunk of information available. We make groups by finding the hyper-plane that separates the classes well overall. In this study binary SVM with 10 cross validations is used for classification of all the facial features using HOG, LBP and concatenation of HOG and LBP[18].

4. Results and Discussion

Our approach detects the facial features of different face datasets and evaluating the classification performance by conducting three experiments on each dataset separately. All the segmented images of lips, eyes and nose from JAFEE, CK+ and CASIA datasets are trained into HOG, LBP and feature fusion of LBP and HOG. Then each trained feature vector of face, eyes, mouth and nose are fed into SVM classifier and their respective accuracies are recorded. Classification accuracy of face (Acc_face) is higher for HOG feature as compare to LBP in case of all datasets evaluated. Classification accuracy
of eyes (Acc_eyes) is higher for HOG feature in case of JAFFE and CK+ dataset but LBP classification accuracies is higher for CASIA dataset. In case of mouth classification (Acc_mouth) accuracy of HOG feature is higher than LBP for JAFEE and CK+ datasets. Classification accuracy of nose (Acc_nose) is greater for HOG for CK+ and JAFEE dataset. But fusion of HOG and LBP features for the combination of all the segmented facial features (Acc_combined) reported highest accuracy of 95.77 % for JAFFE dataset, 99.18 % for CK+ dataset and 99.16% for CASIA dataset.

![Figure 6](image6.png)

**Figure 6.** Comparison of Accuracies of HOG, LBP for JAFEE, last column Acc_combined represents the accuracies of HOG and LBP fused.

![Figure 7](image7.png)

**Figure 7.** Comparison of Accuracies of HOG, LBP for CK+, last column Acc_combined represents the accuracies of HOG and LBP fused.
Figure 8. Comparison of Accuracies of HOG, LBP for CASIA, last column Acc_combined represents the accuracies of HOG and LBP fused.

5. Performance Comparison

In comparison to previous papers, our proposed method outperforms all of them, these papers proposed different algorithms based on Principal Component Analysis (PCA), HOG, Local Phase quantization (LPQ), Convolutional Neural Network (CNN). Comparison of past papers on JAFEE dataset with proposed method is presented in table 4, on CK+ dataset with proposed method is presented in table 5 and on CASIA dataset with proposed method is presented in table 6. Most methods are trained and tested on CK+, JAFEE or JAFFE, CASIA. Biago Yang covered all three datasets. If we compare SVM with neural networks, it is memory efficient as well reduces time complexity. All the recorded accuracies of the proposed method are greater than previous papers which clearly depict that fusion of LBP and HOG outperforms all other features.

Table 4. Comparison of Accuracies of Previous papers and proposed method for JAFFE dataset.

| Author (ref)                        | Methodology                        | Classification          | Accuracy  |
|-------------------------------------|------------------------------------|-------------------------|-----------|
| R. Ramnathan et al. (2009) [5]      | Feature extraction                 | SVM                     | 83.3%     |
|                                     | 2D Gabor Filters                   |                         |           |
| Frank Y. Shin et al. (2008) [25]    | DWT - Discrete Wavelet Transform   | SVM (RBF Kernel)        | 94.13%    |
| Shishir Bashyal et al. (2007) [26]  | Gabor Wavelets                     | LVQ-Linear Vector Quantization | 87.51%  |
| Zhengyou Zhang et al. (1998) [27]   | Geometry based features            | Multilayer Perceptron   | 73.3%     |
|                                     | Gabor Wavelets                     |                         | 92.2%     |
| Proposed Method                     | HOG, LBP, feature fusion           | Binary SVM              | 95.77%    |
### Table 5. Comparison of Accuracies of Previous papers and proposed method for CK+ dataset.

| Author (ref)               | Methodology                           | Accuracy |
|----------------------------|---------------------------------------|----------|
| Chao Qi et al. (2018) [2]  | Feature extraction LBP - Local Binary Pattern | 84%      |
| Kamlesh Mistry et al. (2017) [15] | MLBP - Modified Local Binary Pattern | 94.66% (Ensemble SVM 90.70%) |
| Bing-Fei Wu et al. (2018) [6] | Adaptive Feature mapping              | 87.78%   |
| Wenfei Gu et al. (2012) [28] | Radial encoding of local Gabor Features | 91.51%   |
| Proposed Method            | HOG , LBP , feature fusion            | 99.18%   |

### Table 6. Comparison of Accuracies of Previous papers and proposed method for CASIA dataset.

| Author (ref)               | Methodology                           | Accuracy |
|----------------------------|---------------------------------------|----------|
| Biao Yang (2017) [3]       | WMDNN - weighted mixture deep neural network | 92.3%    |
| F.Bougourzi (2020) [21]    | HOG , LPQ and BSIF                    | 89.67%   |
| Jadd haddad[2020][29]      | 3D convolutional neural networks (3D-CNN) | 97.56% 86% |
| Hepend zhang 2019[30]      | deep convolution neural networks (WMDCNN) |         |
| Proposed Method            | HOG , LBP , feature fusion            | 99.16%   |
6. Conclusion
In this paper, we proposed an effective approach for facial expression recognition using fusion of multi representation and multi descriptors. Three fusion scenarios are performed; first we fuse the extracted features of LBP using all presentations (whole face, eyes, mouth and nose regions) by concatenating them alongside each other. In the second fusion scenario, we concatenated the extracted features of all facial representation using HOG descriptor. The third fusion scenario is obtained by concatenating the fused LBP features and the fused HOG features, all alongside each other. We have evaluated our approach on three most popular datasets CK+, JAFFE and CASIA and got results 99.18%, 95.77% and 99.09% respectively.

7. References
[1] Rohini Goel, Avinash Sharma, and Rajiv Kapoor, "Object Recognition Using Deep Learning" published in Journal of Computational and Theoretical Nanoscience Vol. 16, 4044–4052, 2019.
[2] Qi, Chao & Li, Min & Wang, Qiushi & Zhang, Huiquan & Xing, Jinling & Gao, Zhifan & Zhang, Huailing. (2018). Facial Expressions Recognition Based on Cognition and Mapped Binary Patterns. IEEE Access. PP. 1-1. 10.1109/ACCESS.2018.2816044.
[3] B. Yang, J. Cao, R. Ni and Y. Zhang, "Facial Expression Recognition Using Weighted Mixture Deep Neural Network Based on Double-Channel Facial Images," in IEEE Access, vol. 6, pp. 4630-4640, 2018, doi: 10.1109/ACCESS.2017.2784096.
[4] Ding, Q. Zhao, B. Li and X. Yuan, "Facial Expression Recognition From Image Sequence Based on LBP and Taylor Expansion," in IEEE Access, vol. 5, pp. 19409-19419, 2017, doi: 10.1109/ACCESS.2017.2737821.
[5] S. L. Happy and A. Routray, "Automatic facial expression recognition using features of salient facial patches," in IEEE Transactions on Affective Computing, vol. 6, no. 1, pp. 1-12, 1 Jan.-March 2015, doi: 10.1109/TAFFC.2014.2386334.
[6] B. Wu and C. Lin, "Adaptive Feature Mapping for Customizing Deep Learning Based Facial Expression Recognition Model," in IEEE Access, vol. 6, pp. 12451-12461, 2018, doi: 10.1109/ACCESS.2018.2805861.
[7] Hassun, M., Alam, M.G., Uddin, M.Z., Huda, M., Almogren, A., & Fortino, G. (2019). Human emotion recognition using deep belief network architecture. Inf. Fusion, 51, 10-18.
[8] Chen-Chiung Hsieh, Mei-Hua Hsih, Meng-Kai Jiang, Yun-Maw Cheng, and En-Hui Liang. 2016. Effective semantic features for facial expressions recognition using SVM. Multimedia Tools Appl. 75, 11 (June 2016), 6663–6682. DOI:https://doi.org/10.1007/s11042-015-2598-1
[9] Ouyang, Yan et al. “Accurate and robust facial expressions recognition by fusing multiple sparse representation based classifiers.” Neurocomputing 149 (2015): 71-78.
[10] Xiaoming Zhao, Xugan Shi & Shiqing Zhang (2015) Facial Expression Recognition via Deep Learning, IETE Technical Review, 32:5, 347-355, DOI: 10.1080/02564602.2015.1017542
[11] Wen-Sheng Chu, De la Torre F, Cohn JF. Selective Transfer Machine for Personalized Facial Expression Analysis. IEEE Trans Pattern Anal Mach Intell. 2017;39(3):529-545. doi:10.1109/TPAMI.2016.2547397
[12] Matlovic, Tomas & Gaspar, Peter & Moro, Robert & Simko, Jakub & Bielikova, Maria. (2016). Emotions detection using facial expressions recognition and EEG. 18-23. 10.1109/SMAP.2016.7753378.
[13] Somasekar, J. & Sharma, A. & Reddy, N. & Reddy, Y.. (2020). Image analysis for automatic enumeration of rbc infected with plasmidium parasites-implications for malaria diagnosis. Advances in Mathematics: Scientific Journal. 9. 1221-1230. 10.37418/amsj.9.3.48.
[14] A. Sharmal And J. Somasekar, “Contrast Image Construction Technique for Medical Imaging”, published in Advances in Mathematics: Scientific Journal (Adv. Math., Sci. J.) vol-9-no-6-2020 (pp 3325–3329)
[15] Kamlesh Mistry, Li Zhang, Siew Chin Neoh, Chee Peng Lim, and Ben Fielding proposed a method for recognition of facial expression which employs particle swarm optimization (PSO) for optimization of features along with modified LBP and micro genetic algorithm (mGA). Their approach provided an overall accuracy of 94.66% for CK+ dataset by training it on ensemble based SVM classifier.

[16] Santosh, Mamta & Sharma, Avinash. (2019). A Proposed Framework for Emotion Recognition Using Canberra Distance Classifier. Journal of Computational and Theoretical Nanoscience. 16. 3778-3782.

[17] Mamta Santosh, Avinash Sharma, "Facial Expression Recognition using Fusion of LBP and HoG Features" published in International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-8 June, 2019

[18] 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," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, San Francisco, CA, 2010, pp. 94-101, doi: 10.1109/CVPRW.2010.5543262.

[19] Michael J. Lyons, Miyuki Kamachi, Jiro Gyoba. Coding Facial Expressions with Gabor Wavelets (IVC Special Issue)arXiv:2009

[20] Guoying Zhao, Xiaohua Huang, Matti Taini, Stan Z. Li, Matti Pietikäinen, Facial expression recognition from near-infrared videos, Image and Vision Computing, Volume 29, Issue 9, 2011, Pages 607-619, ISSN 0262-8856

[21] Bougourzi, Fares & Dornaika, Fadi & Mokrani, K. & taleb-ahmed, Abdelmalik & Ruichek, Yassine. (2020). Fusion Transformed Deep and Shallow features (FTDS) for Image-Based Facial Expression Recognition. Expert Systems with Applications. 156. 113459. 10.1016/j.eswa.2020.113459.

[22] King, D. E. (2009). Dlib-ml: A machine learning toolkit. Journal of Machine Learning Research, 10(Jul), 1755–1758

[23] Kong, X., et al., A HOG-SVM based fall detection IoT system for elderly persons using deep sensor. Procedia computer science, 2019. 147: p. 276-282.

[24] Xiao, B., et al., 2D-LBP: an enhanced local binary feature for texture image classification. IEEE Transactions on Circuits and Systems for Video Technology, 2018. 29(9): p. 2796-2808.

[25] Shih, F.Y., C.-F. Chuang, and P.S. Wang. Performance comparisons of facial expression recognition in JAFFE dataset. International Journal of Pattern Recognition and Artificial Intelligence, 2008. 22(03): p. 445-459.

[26] Bashyal, S. and G.K. Venayagamoorthy, Recognition of facial expressions using Gabor wavelets and learning vector quantization. Engineering Applications of Artificial Intelligence, 2008. 21(7): p. 1056-1064.

[27] Zhang, Z., et al. Comparison between geometry-based and gabor-wavelets-based facial expression recognition using multi-layer perceptron. in Proceedings Third IEEE International Conference on Automatic face and gesture recognition. 1998. IEEE.

[28] Gu, W., et al., Facial expression recognition using radial encoding of local Gabor features and classifier synthesis. Pattern recognition, 2012. 45(1): p. 80-91.

[29] Haddad, J., O. Lézoray, and P. Hamel. 3D-CNN for Facial Emotion Recognition in Videos. in International Symposium on Visual Computing. 2020.

[30] Ramanathan, R., et al. A support vector machines approach for efficient facial expression recognition. in 2009 International Conference on Advances in Recent Technologies in Communication and Computing. 2009. IEEE.