2019

Micro Expression Classification Accuracy Assessment

Pratikshya Sharma
University of Ulster

Sonya Coleman
University of Ulster

Pratheepan Yogarajah
University of Ulster

See next page for additional authors

Recommended Citation

Sharma, P., Coleman, S., Yogarajah, P. & Laurenc, T. (2019). Micro expression classification accuracy assessment. *IMVIP 2019: Irish Machine Vision & Image Processing*, Technological University Dublin, Dublin, Ireland, August 28-30. doi:10.21427/kbny-0a41

Authors
Pratikshya Sharma, Sonya Coleman, Pratheepan Yogarajah, and Laurenc Taggart

Micro Expression Classification Accuracy Assessment

Pratikshya Sharma[1*], Sonya Coleman[1], Pratheepan Yogarajah[1], Taggart Laurence[2]

School of Computing, Engineering & Intelligent Systems [1]
School of Nursing & Health Research [2]
University of Ulster, Northern Ireland
sharma-p3@ulster.ac.uk[*]

Abstract

The ability to identify and draw appropriate implications from non-verbal cues is a challenging task in facial expression recognition and has been investigated by various disciplines particularly social science, medical science, psychology and technological sciences beyond three decades. Non-verbal cues often last a few seconds and are obvious (macro) whereas others are very short and difficult to interpret (micro). This research is based on the area of micro expression recognition with the main focus laid on understanding and exploring the combined effect of various existing feature extraction techniques and one of the most renowned machine learning algorithms identified as Support Vector Machine (SVM). Experiments are conducted on spatiotemporal descriptors extracted from the CASME II dataset using LBP-TOP, LBP-SIP, LPQ-TOP, HOG-TOP, HIGO-TOP and STLBP-IP. We have considered two different cases for the CASME II dataset where the first case measures performance for five class i.e. happiness, disgust, surprise, repression and others and the second case considers three classes namely positive, negative and surprise. LPQ-TOP with SVM produced highest accuracy against rest of the approaches in this work.

Keywords: Micro-expression, feature extraction, feature classification, SVM.

1 Introduction

The common non-verbal cues that have attracted attention from research scholars of diverse backgrounds include body language, facial expressions, facial emotion, hand gestures, eye movements and eye contact. These cues are critically important since they are sent from the emotional brain [Patti, 2012] and are hallmarks of the most honest expressions. As the common proverb says “face is the index of mind”- facial expressions can contribute in understanding one’s facial emotion as well as psychological status. Facial expressions are broadly classified into macro and micro expressions where extensive research with high end results has already been obtained for macro expressions and research still continues. However, it has been argued that macro expressions may not be the perfect measure to determine a person’s real emotion and psychological status since these are voluntary as well as conscious expressions. In other words, a person exhibiting these expressions is aware of what they are doing and can regulate their facial emotions, hence may attempt to portray an expression contrary to their actual emotion. In such a scenario macro expressions cannot be used as a true cue. Micro expressions appear on one’s face in a spontaneous manner and a person has exceptionally less control over them hence, they have minimal chance of being regulated. It is believed that these expressions are leaked on one’s face as a result of conscious or unconscious suppression of true emotion [P. Ekman, 1969] [Yuan Zong et al., 2018]. Due to these factors micro expressions are often seen as more natural and genuine than macro expressions; unveiling useful cues for true emotional status. These expressions mainly differ in terms of intensity, duration and number of facial parts contributing to their formation [Sze-Teng Liong et al., 2017]. However, it has been debated that duration is the sole ground on which expressions can be categorized as macro or micro, ruling out intensity as another factor for classification [Wu Q. et al., 2011]. Additionally, expressions that have duration which extends beyond micro expression (0.04 to 0.2secs [Sze-Teng Liong et al., 2017]) duration ultimately fall under the category of macro expressions [Wu Q. et al., 2011] [P. Ekman, 1969] [P. Ekman, 2009]. Previous work [Xiaohong Li et al., 2016] has shown that micro expressions have comparably very low intensity than macro expressions making it tougher for recognition. It is said that more facial segments contribute in forming a macro expression whereas significantly fewer facial segments contribute to a micro expression. Micro expression recognition systems can be devised using both manual as well as automated approaches. The manual approach consists of observations made by highly trained individuals and have achieved a modest accuracy of 47% [M.G. Frank et al., 2009]. A psychological experiment [J. Endres et al., 2009] revealed an average recognition rate of 50% on the micro-expression training dataset METT [P. Ekman, 2003]. However recently, automated approaches have paced towards achieving an accuracy as high as 75.3% [Yandan Wang et al., 2017]. Some of the eminent application areas that have been associated with micro expression recognition include autism, schizophrenia, mental diseases, criminal investigation, lie detection and business negotiations [P. Ekman, 2009] [Yandan Wang et al.,2017]. In this paper we aim to assess classification accuracy of facial micro expressions by experimenting on CASME II [Yan W-J et al., 2014] database and compare it with the existing results of SMIC [Xiaobai Li et al., 2013] database. We have considered two different cases for CASME II database where in the first case we consider five class i.e. happiness, disgust, surprise, repression and others. In the second case we consider three classes namely positive, negative and surprise. The motive behind choosing three class expressions is because the key focus of this research lies in the application of micro expression recognition system on autistic faces. This type of faces tend to manifest very minimal expressions on their face therefore the concept of recognizing micro expression into three classes as positive, negative and surprise seemed more realizable. Also in five class category fear and sad expressions were not used for baseline evaluation however the author wanted to explore these two expressions for autistic faces along with other expressions. The remainder of the paper contains three sections where Section 2 discusses the methodology for micro expression recognition, Section 3 presents the results obtained followed by concluding remarks in Section 4.

2 Methodology

The bare bone structure of the methodology generally adopted for a micro expression recognition system (figure 1) consists of face detection, pre-processing, feature extraction followed by classification as its constituent components [M. Takalkar et al., 2018] [Kam Meng Goh et al., 2018].

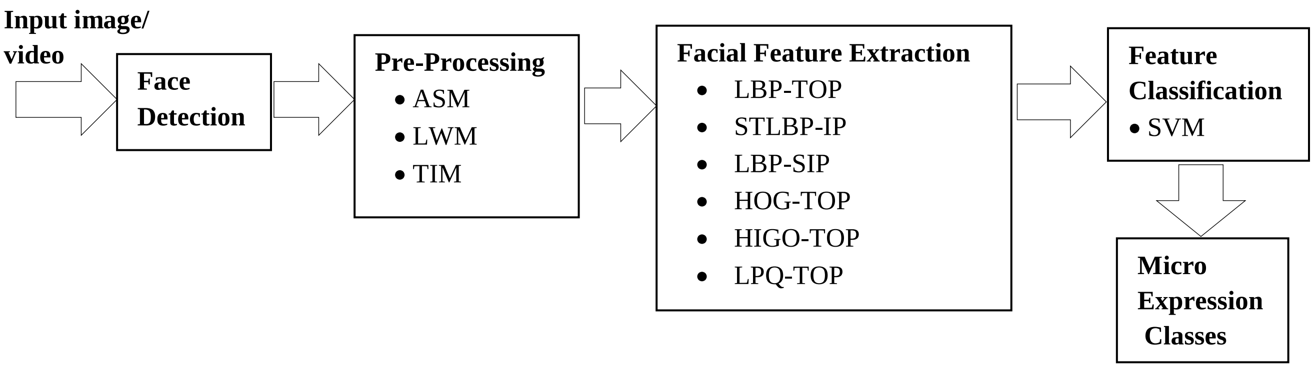

Figure 1 Bare bone structure for micro expression recognition system [M. Takalkar et al., 2018]

2.1 Face Detection and Pre Processing

As the starting point, the face is detected from an image or video and selected as the object of interest for further examination. One of the popular techniques for face detection is the Viola-Jones method [Viola P et al., 2004]. We have chosen the CASME II database [Yan W-J et al., 2014] which contains cropped faces and hence no face detection method was required. The next stage is known as pre-processing and its objective is to enrich features of interest and subdue the features that are of less importance [M. Takalkar et al., 2018]. For five class emotion, the inputs to the proposed system are face frontal view from CASME II images with face size of approximately 280 x 340 pixels. These images are pre-processed using Active Shape Modelling (ASM) [Cootes TF et al., 1995] and

Light Weighted Mean (LWM) [Goshtasby A, 1988]. In three class emotion, the CASME II micro expression database requires to be normalized using a Temporal Interpolation Model (TIM) [Z. Zhou et al., 2011] to contain 16 frames per video. By doing so it helps to meet appropriate parameter setting requirements and enable spatiotemporal descriptor extraction [Yuan Zong et al., 2018].

2.2 Facial Feature Extraction

The pre-processing stage is followed by feature extraction which can be broadly classified into appearance based and geometry based. These extracted features form the foundation for recognition of facial expressions into their respective classes. Here Local Binary Pattern with Three Orthogonal Plane (LBP-TOP) [Zhao G et al., 2007] was chosen since this method is one of the most popular extraction techniques for micro expression and was used by [Yan W-J et al., 2014] as baseline evaluation, for five class emotion. For calculating LBP a set of frames is required representing a video sequence with *T* time length. The LBP code is constructed for three planes namely XY, XT and YT. Next, we construct histograms for these planes individually and eventually concatenate them into a block of 5 x 5. These features along with class labels were fed as input to SVM for classification. Another method for extracting features for five class emotion is Spatiotemporal Local Binary Pattern with Integral Projection (STBLP-IP) [Xiaohua Huang et al., 2015]. The main reason for choosing this extraction technique was due to its ability to retain the attribute information for the shape in images along with more discriminative capability for categorizing micro expressions into various class. Initially the integral projection in two directions, horizontal and vertical is calculated. Then, LBP is calculated for extracting appearance information from the difference images produced in the above step and finally the resulting binary vector element is multiplied by a weight. The 1D LBP obtained for each frame is represented in a histogram. In the final step, the histograms obtained from horizontal and vertical projections are represented into a single vector by concatenation and classified using SVM. For micro expression with three class the first feature extraction method considered was LBP-TOP which follows same procedure as described earlier. The second feature extraction method is known as Local Binary Pattern with Six Intersection Point (LBP-SIP) [Y. Wang et al., 2014]. For extracting features from the input images this method uses only six unique points emanated from the intersection of three orthogonal planes. The final feature vector resulting from concatenation of local binary patterns of three planes was used for classification in our experiment. Due to its ability to reduce redundancy, LBP-SIP provides a condensed representation of features and hence was selected in this research work. The third technique is Histogram of Gradient with Three Orthogonal Plane (HOG-TOP) [Xiaobai Li, et. al., 2018] which calculates local gradient direction and gradient magnitude denoted by symbols ϴ and *m*. This method computes these values for every pixel on the XY plane. It also constructs a mapping between the directional and magnitude values calculated in previous step. This technique is drawn out from the existing HOG method which was originally developed for detecting human [N. Dalal et al., 2005]. Finally a histogram is constructed along with a weighted vote approach and the overall procedure is further extended for the XT and YT planes then concatenated into a vector. The feature vector obtained from HOG-TOP has been used in our work for classification. The fifth method is Histogram of Image Gradient with Three Orthogonal Plane (HIGO-TOP) and follows the similar procedure as HOG-TOP but uses simple vote approach [Xiaobai Li, et. al., 2018]. For colour videos both HOG and HIGO methods surpassed the performance in comparison to the LBP method [Xiaobai Li, et al., 2018] hence was chosen in this micro expression classification experiment.

In addition to these feature extraction techniques that are commonly used for micro-expressions, we propose to use Local Phase Quantization with Three Orthogonal Plane (LPQ-TOP) [Päivärinta J. et al., 2011] [V. Ojansivu et al., 2008] which has not previously been used for three class micro-expression classification with SVM on CASME II dataset. In this method the phase information can be derived using two dimensional Discrete Fourier transform with four coefficients having values of type integer (range of 0 to 255). These integer values can be formed using the binary coding technique. The local phase quantized information collected for all the neighbour pixels for a given image is modelled into a histogram. The parameters used include a local window of [5, 5, 5] with a correlation model of [0, 0] [Yuan Zong et al., 2018]. This notion is further extended on three orthogonal planes and ultimately concatenated into feature vector. This feature vector has been used in our work for classifying them into three emotion types. The phase quantization method is not affected by the changes in the illumination that have uniform nature, hence this method has been considered for examination in this research.

2.3 Feature Classification

To classify features for both five class and three class emotion on CASME II micro expressions database, the SVM technique [Vapnik P. et al., 1995] using LIBSVM [Chih-Chung et al., 2011] was used with three kernel functions namely linear, polynomial and radial basis function (RBF). For five class emotion the CASME II database with emotions labelled as happy, disgust, surprise, repression and others was considered similar to [Yan W-J et al., 2014]. Among the three kernels used in SVM for classification purpose the polynomial kernel reported the best performance for LBP-TOP. In case of STLBP-IP linear kernel was able to give best performance compared to RBF and polynomial kernel. In the second case only three classes of emotions represented as positive, negative and surprise were used. For three class emotion the micro expression images of CASME II database originally labelled as happy is considered as positive micro expression in our work. Similarly disgust, sad, and fear are considered as negative micro-expression and surprise sample is considered same as the original micro expression. The features extracted from CASME II dataset with relabelling for 130 samples has been adapted from [Yuan Zong et al., 2018]. The RBF kernel was able to produce best performance when used with HOG-TOP, LBP-SIP and LPQ-TOP independently. For LBP-TOP polynomial kernel reported the best performance whereas for HIGO-TOP linear kernel performed the best.

3 Results

Table 1 presents the classification results obtained for five class micro-emotion on CASME II database. For LBP-TOP the value of radius parameter on the *x* and *y* axes was set to 1, radius of T was set to 4, and number of neighbouring pixels *P* was set to 8. For LBP-TOP the highest accuracy of 62.98% (Table 1) was achieved in combination with SVM implemented using polynomial kernel and cross validation with leave-one-subject-out. Using STLBP-IP and SVM classification with leave-one-subject-out cross validation the highest accuracy achieved was 54.11% with linear kernel compared to RBF and polynomial kernel. In this case for constructing 1D LBP the mask parameter was set to 9.

Feature Extraction Method	**Recognition rate obtained (%)**	**SVM Kernel Used**	**Comparison**		**Five Emotion Class**
			Accuracy	**Reference**	
LBP-TOP	**62.98**	Polynomial	63.41	[Yan W-J et al., 2014]	Happy, Disgust, Surprise, Repression & Others
STLBP-IP	54.11	Linear	59.51	[Xiaohua Huang et al., 2015]	

Table 1: Classification accuracy comparison for five class emotion on CASME II.

Table 2 depicts results obtained for CASME II representing three class i.e. positive, negative and surprise for micro expression. The LBP-TOP with radius *R*=3 and neighbour *P*=8 achieved average accuracy of 55.20% using polynomial kernel for three class emotion. This is however lesser than the accuracy obtained for five class emotion i.e. 62.98% but comparable with the state of the art accuracy reported for three class emotion on CASME II. LBP-SIP was able to secure average accuracy of 56.08% using *R*=3 and a RBF kernel. Both HOG-TOP and HIGO-TOP

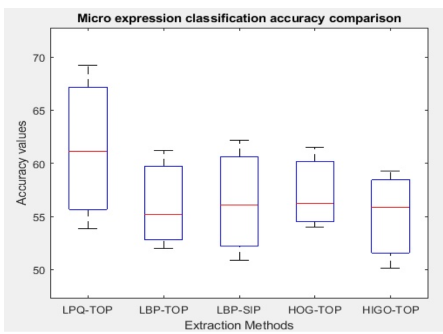

Figure 2 Micro expression classification accuracy comparison

techniques with 8 neighbour has obtained average accuracy of 56.22% with RBF and 55.87% with linear kernel respectively. Among the five methods LPQ-TOP was able to produce a maximum accuracy of 69.23% and average accuracy of 61.16%. This is by far the highest accuracy obtained in our experiment compared to all other methods for three class emotion for micro expression classification. Research work using SVM for classification of features extracted for micro expressions using LPQ-TOP could not be found as of now due to which appropriate comparison could not be made. Figure 2 demonstrates the lowest, highest and average accuracy obtained for three class emotion on CASME II dataset using the five different feature extraction methods with SVM.

Experiment results obtained for 3 class emotion			**Comparison with 3 class emotion**			**Emotion Used**
Feature Extraction Method	**Recognition rate obtained (%)**	**SVM Kernel Used**	**Accuracy**	**Reference**	**Class Label**	
LPQ-TOP	**61.16**	RBF	-	-	3	Happy, Sad, Fear, Disgust, Surprise
LBP-SIP	56.08	RBF	-	-		
LBP-TOP	55.20	Polynomial	55.87	[Xiaobai Li, et al., 2018]	positive, negative, surprise	
HOG-TOP	56.22	RBF	57.49			
HIGO -TOP	55.87	Linear	55.87			

Table 2: Classification accuracy Assessment and comparison for three class emotion on CASME II.

Feature Extraction Method	**Our Accuracy (CASME II)**	**Comparison**			**Class Label**	**Emotion Used**
		Accuracy	**Reference**	**Reference Database**		
LPQ-TOP	**61.16**	-	-	-	3	Happy, Sad, Fear, Disgust, Surprise
LBP-SIP	56.08	64.02	[Y. Wang et al., 2014]	SMIC	positive, negative, surprise	
LBP-TOP	55.20	76.06	[Xiaobai Li, et al., 2018]	SMICsub-HS		
HOG-TOP	56.22	80.28	[Xiaobai Li, et al., 2018]	SMICsub-HS		
HIGO-TOP	55.87	80.28				

Table 3: Classification Accuracy comparison for three class emotion between CASME II and SMIC.

4 Conclusions

On the basis of micro expression classification accuracy obtained it is seen that LBP-TOP with SVM was successful in achieving the accuracy comparable to the baseline accuracy for five class emotion (refer table 1). However applying same technique on CASME II dataset for classifying it into three class emotion produced ~8% lesser accuracy (refer table 2). This may be due to fact that fear and sad expression are difficult to classify. However, as these are preliminary results, they are positive and there is scope for improving these in future by parameter tuning and additional pre-processing. In case of STLBP-IP method, the maximum accuracy achieved was ~4% below the standard value reported. Few similar works could be found for micro expression hence tuning the method was challenging and time consuming. Using LBP-SIP technique for three class emotion the recognition value achieved (refer table 3) was satisfactory; the value attained was ~8% below that of SMIC dataset. HOG-TOP achieved fair accuracy value in the range of mid fifty on CASME II but is comparatively lesser than the value reported for SMIC (refer table 3). Similarly a satisfactory accuracy value was achieved using HIGO-TOP but is far below that of SMIC (refer table 3). Most of the current research work on CASME II dataset have focused on classifying micro expression into more than three class emotion. Very less work have been reported on CASME II dataset for classifying micro expression into three class emotion hence comparison of our accuracy value for three class emotion has been established wherever applicable. We have also compared our accuracy value obtained for CASME II dataset with existing accuracy values on SMIC dataset for micro expressions with three emotion class. The LPQ-TOP feature extraction method on CASME II micro expression database has been used by [Yuan Zong et al., 2018]. The work focuses on cross-database for micro expression recognition so suitable classification accuracy value with the combination of SVM and LPQ-TOP could not be found. Nevertheless in our experiment the features extracted using LPQ-TOP was classified using SVM and has produced highest classification accuracy compared to other extraction methods. Hence this technique seems to have some potential to be considered for further examination. Previous work on SMIC database has achieved good success rate for three class emotion thus improvising the accuracy for CASME II dataset for three class emotion will be the focus in future in addition to five class emotion for micro expression.

References

[Y. Wang et al., 2014] Wang, Yandan, John See, Raphael C.-W. Phan and Yee-Hui Oh, *LBP with Six Intersection Points: Reducing Redundant Information in LBP-TOP for Micro-expression Recognition*, Asian Conference on Computer Vision. Springer, 2014, pp. 525–537.

[Yuan Zong et al., 2018] Yuan Zong, Wenming Zheng, Xiaopeng Hong, Chuangao Tang, Zhen Cui, Guoying Zhao, *Cross-Database Micro-Expression Recognition: A Benchmark*. CoRR abs/1812.07742 (2018).

[Chih-Chung et al., 2011] Chih-Chung Chang and Chih-Jen Lin, *LIBSVM : a library for support vector machines*. ACM Transactions on Intelligent Systems and Technology, vol. 2, no. 3, p. 27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[Xiaobai Li, et al., 2018] Xiaobai Li, Xiaopeng Hong, Antti Moilanen, Xiaohua Huang , Tomas Pfister, Guoying Zhao , and Matti Pietikainen , *Towards Reading Hidden Emotions:A Comparative Study of Spontaneous Micro-Expression Spotting and Recognition Methods*, IEEE Transactions On Affective Computing, Vol. 9, No. 4, October-December 2018.

[Xiaohua Huang et al., 2015] Xiaohua Huang, Su-Jing Wang, Guoying Zhao, Matti Pietikainen, *Facial Micro-Expression Recognition using Spatiotemporal Local Binary Pattern with Integral Projection*, 2015 IEEE International Conference on Computer Vision Workshop (ICCVW).

[Yan W-J et al., 2014] Yan W-J, Li X, Wang S-J, Zhao G, Liu Y-J, et al. (2014), *CASME II: An Improved Spontaneous Micro-Expression Database and the Baseline Evaluation*, PLoS ONE 9(1): e86041. doi:10.1371/journal.pone.0086041.

[Zhao G et al., 2007] Zhao G, Pietikainen M (2007), *Dynamic texture recognition using local binary patterns with an application to facial expressions*, IEEE Trans Pattern Anal Mach Intell 29: 915–928.

[Patti, 2012] Patti Wood, *Snap: Making the Most of First Impressions, Body Language, and Charisma*, Paperback – October 9, 2012, 13 edition.

[Xiaohong Li et al., 2016] Xiaohong Li, Jun Yu,Shu Zhan, *Spontaneous facial micro-expression detection based on deep learning* , 978-1-5090-1345-6/16/,2016 IEEE.

[M.G. Frank et al., 2009] M. G. Frank, M. Herbasz, K. Sinuk, A. Keller, and C. Nolan., *I see how you feel: Training laypeople and professionals to recognize fleeting emotions*, International Communication Association, Sheraton New York City, 2009. 1, 2, 7.

[Yandan Wang et al., 2017] Yandan Wang, John See, Yee-Hui Oh, Raphael C.-W. Phan, Yogachandran Rahulamathavan, Huo-Chong Ling5, Su-Wei Tan, Xujie Li, *Effective recognition of facial micro-expressions with video motion magnification*, Multimed Tools Appl (2017) 76:21665–21690 , DOI 10.1007/s11042-016-4079-6.

[J. Endres et al., 2009] J. Endres and A. Laidlaw, *Micro-expression recognition training in medical students: a pilot study*, BMC medical education, vol. 9, no. 1,p. 47, 2009.

[P. Ekman, 2009] P. Ekman, *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage*, (Revised Edition), WW Norton & Company, 2009.

[P. Ekman, 1969] Ekman P,Friesen WV,(1969), *Nonverbal leakage and clues to deception*.Psychiatry 32: 88–106.

[Wu Q. et al., 2011] Wu Q., Shen X., Fu X. (2011) *The Machine Knows What You Are Hiding: An Automatic Micro-expression Recognition System*, In: D'Mello S., Graesser A., Schuller B., Martin JC. (eds) Affective Computing and Intelligent Interaction. ACII 2011. Lecture Notes in Computer Science, vol 6975. Springer, Berlin, Heidelberg.

[M. Takalkar et al., 2018], *A survey: facial micro-expression recognition*, Madhumita Takalkar1 & Min Xu1 & Qiang Wu1 & Zenon Chaczko, Multimed Tools Appl (2018) 77:19301–19325.

[Viola P et al., 2004] Viola P, Jones MJ (2004), *Robust real-time face detection*, Int J Comput Vis 57:137–154

[Kam Meng Goh et al., 2018] Kam Meng Goh, Chee How Ng, Li Li Lim, U. U. Sheikh, 2018, *Micro-expression recognition: an updated review of current trends, challenges and solutions*, The Visual Computer https://doi.org/10.1007/s00371-018-1607-6

[Päivärinta J et al., 2011] Päivärinta J., Rahtu E., Heikkilä J. (2011), *Volume Local Phase Quantization for Blur-Insensitive Dynamic Texture Classification*. In: Heyden A., Kahl F. (eds) Image Analysis. SCIA 2011. Lecture Notes in Computer Science, vol 6688. Springer, Berlin, Heidelberg, Springer, 2011, pp. 360–369.

[V. Ojansivu et al., 2008] V. Ojansivu and J. Heikkilä, *Blur insensitive texture classification using local phase quantization*, International conference on image and signal processing. Springer, 2008, pp. 236–243.

[Xiaobai Li et al., 2013] Xiaobai Li; Pfister, T.; Xiaohua Huang; Guoying Zhao; Pietikainen, M., *A Spontaneous Micro-expression Database: Inducement, collection and baseline,Automatic Face and Gesture Recognition (FG)*, 2013 10th IEEE International Conference and Workshops on , vol., no., pp.1-6, 22-26 April 2013

[P. Ekman, 2003] P. Ekman, *Mett. micro expression training tool*, CD-ROM. Oakland, 2003.

[Cootes TF et al., 1995]Cootes TF, Taylor CJ, Cooper DH, Graham J (1995), *Active shape models-their training and application*, Comput Vis Image Underst 61: 38–59.

[Goshtasby A, 1988] Goshtasby A (1988), *Image registration by local approximation methods*, Image Vis Comput 6: 255–261.

[Z. Zhou et al., 2011] Z. Zhou, G. Zhao, and M. Pietik¨ainen, *Towards a practical lipreading system*, Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011, pp. 137–144.

[N. Dalal et al., 2005] N. Dalal and B. Triggs, *Histograms of oriented gradients for human detection*, in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 1. IEEE, 2005, pp.886–893.

[Sze-Teng Liong et al., 2017]. Sze-Teng Liong a, John See c, KokSheik Wong d,*, Raphael C.-W. PhanLess is more: Micro-expression recognition from video using apex frame, 14 December 2017, 0923-5965 2017 Elsevier, https://doi.org/10.1016/j.image.2017.11.006