Check for updates

RESEARCH ARTICLE

Motion and Geometric Feature Analysis for Real-time Automatic Micro-expression Recognition Systems [version 1; peer review: 1 approved with reservations, 1 not approved]

Adamu Muhammad Buhari ¹, Chee-Pun Ooi¹, Vishnu Monn Baskaran², Wooi-Haw Tan¹

¹Faculty of Engineering, Multimedia University, Cyberjaya, Selangor, 63100, Malaysia

²School of Information Technology, Monash University Malaysia, Subang Jaya, 47500, Selangor, Malaysia

V1 First published: 11 Oct 2021, 10:1029
<https://doi.org/10.12688/f1000research.72970.1>

Latest published: 11 Oct 2021, 10:1029
<https://doi.org/10.12688/f1000research.72970.1>

Abstract

The trend of real-time micro-expression recognition systems has increased with recent advancements in human-computer interaction (HCI) in security and healthcare. Several studies in this field contributed towards recognition accuracy, while few studies look into addressing the computation costs. In this paper, two approaches for micro-expression feature extraction are analyzed for real-time automatic micro-expression recognition. Firstly, motion-based approach, which calculates motion of subtle changes from an image sequence and present as features. Then, secondly, a low computational geometric-based feature extraction technique, a very popular method for facial expression recognition in real-time. These approaches were integrated in a developed system together with a facial landmark detection algorithm and a classifier for real-time analysis. Moreover, the recognition performance were evaluated using SMIC, CASME, CAS(ME)2 and SAMM datasets. The results suggest that the optimized Bi-WOOF (leveraging on motion-based features) yields the highest accuracy of 68.5%, while the full-face graph (leveraging on geometric-based features) yields 75.53% on the SAMM dataset. On the other hand, the optimized Bi-WOOF processes sample at 0.36 seconds and full-face graph processes sample at 0.10 seconds with a 640x480 image size. All experiments were performed on an Intel i5-3470 machine.

Keywords

micro-expression recognition, facial feature extraction, real-time classification, geometric-based features, facial graph analysis, emotion classification

Open Peer Review

Approval Status ? X

	1	2
version 1
11 Oct 2021	view	view

1. **Madhumita Takalkar** , University of Technology Sydney, Ultimo, Australia
2. **Harshadkumar B. Prajapati** , Dharmsinh Desai University, Nadiad, India

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the **Research Synergy**
Foundation gateway.

Corresponding author: Adamu Muhammad Buhari (adam_m.buhari@yahoo.com)

Author roles: **Muhammad Buhari A:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Ooi CP:** Conceptualization, Formal Analysis, Methodology, Resources, Supervision, Writing – Review & Editing; **Baskaran VM:** Conceptualization, Formal Analysis, Supervision, Validation, Writing – Review & Editing; **Tan WH:** Conceptualization, Methodology, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2021 Muhammad Buhari A *et al.* This is an open access article distributed under the terms of the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Muhammad Buhari A, Ooi CP, Baskaran VM and Tan WH. **Motion and Geometric Feature Analysis for Real-time Automatic Micro-expression Recognition Systems [version 1; peer review: 1 approved with reservations, 1 not approved]** F1000Research 2021, 10:1029 <https://doi.org/10.12688/f1000research.72970.1>

First published: 11 Oct 2021, 10:1029 <https://doi.org/10.12688/f1000research.72970.1>

Introduction

A micro-expression is a brief, spontaneous facial expression that occurs on a human face in response to the emotions they are experiencing. The micro-expression contains a significant amount of information, and has attracted the interest of computer vision researchers because of its potential uses in security, interrogation, and healthcare.¹⁻³ However, due to the speed of facial muscle movement, it is difficult to extract this information, and the features must be more detailed. A typical period of micro-expression is 200 milliseconds or less.⁴ For real-time micro-expression analysis and emotion recognition, the process involves pre-processing, feature extraction, and recognition. This paper examines two popular feature extraction approaches: motion-based features and geometric-based features. These approaches are reported to have reliable details from uncontrolled image data, which makes it feasible for real-time analysis.

A motion-based feature is constructed based on non-rigid motion changes of subtle expressions where motion changes are extracted for spotting purposes. Facial motion analysis was first presented in⁵ using optical flow to spot micro-expressions. Since then, several studies in this field have explored this approach for facial landmark detection and micro-expression recognition. In,⁶ authors proposed an optical flow features from Apex-frame network (OFF-ApexNet), which combines optical flow guided context with the convolutional neural network (CNN) to compute features. Then, authors in⁷ presented a novel algorithm that combines a deep multi-task convolutional network for detecting facial landmarks and a fused deep convolutional network for micro-expression features. In another study,⁸ authors suggested the Riesz pyramid and a multi-scale steerable Hilbert transform. While, Merghani and Yap⁹ proposed a new region-based method with an adaptive mask. However, the motion-based features reported recognition accuracy from the current studies peaks at 74.06% over the CASMEII using leave-one-subject-out cross validation (LOSOCV).⁶

On the other hand, geometric facial analysis deals with the locations and shapes of facial components. As highlighted in Liu *et al.*,¹⁰ performance of landmark detection algorithms is limited, where only a few studies utilized landmarks in early facial graph representations. However, with recent advancements in face analysis study, improved facial landmark detection algorithms are presented in several studies.¹¹⁻¹⁴ Towards facial landmark graph features, Lei *et al.*,¹⁵ presented a method that only employed 28 brow and lip landmarks, which contributed significantly to micro-expressions. While, other studies¹⁶⁻¹⁹ presented graph-based methods using AU to define landmarks of interest. The reported recognition accuracies from these methods proved that micro-expression features can be extracted using facial graph approaches. However, the general problem with graph-based micro-expression recognition is the lack of large-scale in-the-wild datasets. To date, the recognition accuracy peaks at 87.33% over the SAMM dataset with LOSOCV as reported in Buhari *et al.*¹⁸

Methods

An automatic micro-expression recognition system is implemented for real-time facial analysis by integrating face landmark detection, feature extraction, and then classification. In the developed system, a trained model is generated using publicly available spontaneous micro-expression datasets. For micro-expression feature analysis, two different methods were implemented. Firstly, a Bi-Weighted Oriented Optical Flow (Bi-WOOF) feature descriptor by Liong *et al.*,²⁰ is used. This method is a motion-based approach that uses optical flow to compute features, and it requires an apex-frame spotting before the feature computation. The BiWOOF is considered in this study as the results of its performance improvement over the textural feature extraction methods such as local binary patterns on three orthogonal planes (LBP-TOP) as reported in Liong *et al.*²⁰ However, the computational cost poses challenges for real-time recognition as it requires apex-frame spotting. The second feature descriptor considered is a full-face graph by Buhari *et al.*¹⁸ This geometric-based feature computation method requires only facial landmarks to compute features. The full-face graph is considered in this study because the computational time is significantly low in comparison with motion-based. However, it is reported that the earlier geometric-based methods could not detect hidden changes in facial components due to its subtleness and briefness.

Motion-based framework

Figure 1 illustrates the implemented real-time micro-expression recognition system developed using Bi-WOOF. This feature extractor requires at least two frames (i.e., neutral frame and apex-frame) to compute. Firstly, the system captures images of faces using dlib-19.4.²¹ Next, apex-frame spotting is applied using an automatic apex frame spotting method by Liong *et al.*,²² to identify the frame with the highest facial expression within the captured image sequences (i.e., processing sample). As reported by the authors in Liong *et al.*,²² the performance of their method improved over the annotated apex frames provided in micro-expression databases. Here, image sequences from spontaneous micro-expression datasets were utilized. Upon identifying the onset and apex frames, optical flow vectors are computed to define the face motion patterns: (i) magnitude: pixel movement intensity; (ii) orientation: flow motion direction; and (iii) optical strain: modest deformation intensity. Then, using the computed optical flow vectors (i.e., the magnitude, orientation, and the optical strain), Bi-WOOF features are formed. Step-by-step details of this method can be found in

Figure 1. Framework of the designed real-time micro-expression recognition system.

Liong *et al.*²⁰ Figure 1 shows a framework of the real-time micro-expression recognition system using apex-frame spotting and BiWOOF feature extraction method.

Geometric-based framework

Figure 2 illustrates the implemented real-time micro-expression recognition system developed using full-face graph. At first, facial landmark detection is applied to detect coordinates of facial components using dlib-19.4. Then, segments of lines are generated using the detected coordinates of facial components by connecting each landmark point (denoted as p) with subsequent landmark points (denoted as q), for $p \in \{1, 2, \dots, N\}$ and $q \in \{1, 2, \dots, p\}$, where $N = 68$. This concept is described as a full-facial graph using landmark points, and segments are generated as follows:

$$\mathfrak{F}(k) = \{p, q\} \rightarrow k \in \left\{1, 2, \dots, \sum p \times q\right\}.$$

The indexes (i.e., p, q) of every landmark with subsequent landmark points are determined and stored as segments in $\mathfrak{F}$ for feature computations. After the graph is generated, features are computed by calculating the Euclidean distance and gradient of every segment, an idea presented in Buhari *et al.*¹⁸ The total number of features computed using this technique is $N \times (N - 1) \rightarrow K$, which translates to 4,556 features at $N = 68$.

To further analyse the potential performance improvement of the geometric-based features, Eulerian motion magnification (EMM) is applied to the images to amplify the micro-expressions prior to the landmark detection process. Eulerian-inspired approaches^{23,24} do not require explicit motion vectors but emulate motion magnification by magnifying property changes, such as amplitude (denoted as A-EMM) or phase (denoted as P-EMM). According to Le *et al.*²⁴ A-EMM outperformed P-EMM in terms of recognition rates over a broad range of magnification factors. Thus, this paper considered the A-EMM to the images before the feature computations. Details of the methods for A-EMM are detailed in Le *et al.*²⁴ Figure 3 illustrates the principle of the integrated a magnification sub-process to the implemented single-frame sample with a geometric-based features system.

Figure 2. Frame-based micro-expression recognition system.

Figure 3. Frame-based micro-expression recognition system with A-EMM.

Experiment settings

The experiments were performed using four spontaneous datasets: (i) spontaneous micro-expression dataset (SMIC) dataset,²⁵ (ii) Chinese Academy of Sciences Micro-expression (CASMEII) dataset,²⁶ (iii) spontaneous macro-expressions and micro-expressions (CAS (ME)²) dataset,²⁷ and (iv) spontaneous actions and micro-movements (SAMM) dataset.²⁸ These are spontaneous micro-expression datasets which were used in this study with full details of these datasets in Li *et al.*,²⁵ Yan *et al.*,²⁶ Qu *et al.*,²⁷ and Davison *et al.*²⁸ Details to acquire these datasets used in this study are available at www.oulu.fi/cmvs/node/41319 for SMIC,²⁵ fu.psych.ac.cn/CASME/casme2-en.php for CASMEII,²⁶ [fu.psych.ac.cn/CASME/cas \(me\) 2-en.php](http://fu.psych.ac.cn/CASME/cas(me)2-en.php) for CAS (ME)²,²⁷ and personalpages.manchester.ac.uk/staff/adrian.davison/SAMM.html for SAMM.²⁸ Moreover, to evaluate the performance using a larger dataset, this paper merged the four datasets to form a COMBINED dataset. The COMBINED dataset is created from the raw images of all the four datasets. The steps for generating the COMBINED datasets includes; face detection, face cropping, colour-space conversion to grayscale, and image re-scaling to 140×170 . From these steps, colour-space conversion is applied to adopt the SAMM dataset samples as provided in grayscale format, where the sample re-scaling to 140×170 adopts the SMIC dataset cropped image sizes (i.e., the smallest cropped image size considered to provide reliable features description, and achieve high speed performance for real-time micro-expression recognition). The image re-scaling utilises a down-sampling technique by Buhari *et al.*,²⁹ in order to re-produce a high quality down-scaled samples. In addition, the COMBINED dataset adopted the SMIC dataset sample labelling by re-grouping the seven classes of emotions (i.e., happiness, sadness, anger, surprise, fear, contempt, and disgust) to three classes (i.e., positive, negative, and surprise). Here, positive $\in \{\text{happiness}\}$, negative $\in \{\text{sadness, anger, fear, contempt, disgust}\}$, and surprise $\in \{\text{surprise}\}$. Figure 4 illustrates the COMBINED dataset formation from the four spontaneous datasets. Note that the participant images utilised in Figure 4 are the publishable images with the consent of participants, as stated in the documentation from each study. While, Table 1 summarises the selected spontaneous micro-expression datasets used in this study. In this table, the COMBINED dataset is denoted as δ .

Recognition accuracy

Table 2 records the recognition accuracy of the baseline Bi-WOOF,²⁰ (denoted as BBW), optimized Bi-WOOF, (denoted as OBW), full-face graph, (denoted as FFG) and full-face graph with A-EMM, (denoted as FFG+M). The baseline Bi-WOOF is referred to the original method by Liang *et al.*,²⁰ which was implemented using MATLAB, then the optimized Bi-WOOF refers to the implemented C++ version that accelerates the computation performance for real-time analysis. All the four experimental setups utilized the Support Vector Machines (SVM) classifier with a Radial basis function (RBF) kernel. The SVM hyper-parameter selection is based on the recommendation in Bergstra and Bengio.³⁰ This is described as an optimised hyper-parameter selection technique for best classification performance in comparison

Figure 4. COMBINED dataset formation, sample images were taken from SMIC, CASMEII, CAS (ME)², and SAMM datasets as labelled.

Table 1. Summary of spontaneous micro-expression datasets for analysis.

Datasets	Frame rate	Subjects	Samples	Classes
SMIC	100	20	164	3
CASMEII	200	35	247	5
CAS (ME) ²	30	22	341	3
SAMM	200	32	159	7
δ	—	94	911	3

Table 2. Recognition accuracy (%).

	SMIC	CASMEII	CAS (ME) ²	SAMM	δ
BBW	62.20	62.52	59.11	65.22	66.01
OBW	65.29	63.32	57.89	68.50	69.15
FFG	74.62	74.41	75.11	74.33	77.05
FFG+M	75.01	74.55	76.21	75.53	77.85

Table 3. Accuracy comparison table: Optimized Bi-WOOF with other motion-based features.

Method	Feature	Classifier	LOSOCV accuracy (%)			
			SMIC	CASMEII	CAS (ME) ²	SAMM
Liong <i>et al.</i> ²⁰	Bi-WOOF	SVM	62.20	58.85	59.26	-
Li <i>et al.</i> ³¹	MESR+LBP	LSVM	57.93	55.87	-	-
	MESR+HIGO	LSVM	65.24	57.09	-	-
	MESR+HOG	LSVM	57.93	57.49	-	-
Liong and Wong ³²	Bi-WOOF + Phase	SVM	68.29	62.55	-	-
Liong <i>et al.</i> ⁶	OFF-ApexNet	Softmax	67.68	74.06	-	68.18
Li <i>et al.</i> ⁷	deep-NN + Revised HOOF	SVM	-	58.03	-	-
Merghani and Yap ⁹	ROI + Adaptive Mask	SVM	-	68.20	-	56.10
Duque <i>et al.</i> ⁸	Riesz Pyramid + MORF	SVM	54.88	45.93	-	-
Re-implemented ²⁰	Optimized Bi-WOOF	SVM	65.29	63.32	57.89	68.50

to the sequential tuning in a context of a model with many hyper-parameters. In addition, all measured accuracies are based on LOSOCV. Similarly, the COMBINED dataset is denoted as δ in Table 2.

Results

Table 2 presents the recognition accuracies of the baseline Bi-WOOF (denoted as BBW), the optimized Bi-WOOF (denoted as OBW), full-face graph (denoted as FFG) and the full-face graph with A-EMM (denoted as FFG+M). Here, the BBW and OBW yield the highest recognition accuracies of 66.01% and 69.15%, respectively, over the COMBINED dataset. Similarly, FFG and FFG+M yield the highest recognition accuracy of 77.05% and 77.85%, respectively, over the COMBINED dataset. From these results, it is observed that the OBW improved the performance of the BBW by up to 3.28% over the SAMM dataset. On the other hand, the implemented full-face graph with A-EMM improved the performance by up to 1.20% over the SAMM dataset. Then, in comparison with optimized Bi-WOOF and full-graph with A-EMM improved the performance by 9.72%, 11.23%, 18.32%, 7.03% and 8.7% over SMIC, CASMEII, CAS (ME)², SAMM and COMBINED datasets, respectively. Moreover, Table 3 compares the accuracies of the optimized BiWOOF with other motion-based methods, where Table 4 compares the accuracies of the full-face graph with other geometric-based methods.

Table 4. Accuracy comparison table: Full-face graph with other geometric-base features.

Method	Feature	Classifier	LOSOCV accuracy (%)			
			SMIC	CASMEII	CAS (ME) ²	SAMM
Lei <i>et al.</i> ¹⁵	G-TCN	Softmax	-	73.98	-	75.00
Xie <i>et al.</i> ¹⁶	AU-GACN	Softmax	-	56.10	-	52.30
Buhari <i>et al.</i> ¹⁸	$\mathbb{R}$ +Full-face graph	SVM	66.90	73.45	72.83	80.23
	$\mathbb{R}$ +FACS-based graph	SVM	76.67	75.04	81.41	87.33
Zhou <i>et al.</i> ¹⁷	MER-auGCN	Softmax	-	70.80	-	66.20
Liu <i>et al.</i> ¹⁹	Sparse MDMO	SVM	70.51	66.95	-	-
Experiment I	Full-face Graph	SVM	74.62	74.41	75.11	74.33
Experiment II	Full-face Graph + A-EMM	SVM	75.01	74.55	76.21	75.53

Discussion

Table 3 lists the performance of benchmark motion-based methods,^{6-9,20,31,32} with the optimized Bi-WOOF. The best reported accuracy is 74.06% over the CASMEII dataset.⁶ Looking at Bi-WOOF+Phase,³² the highest performance reported is 68.29% over the SMIC dataset, which outperformed the baseline and the optimized Bi-WOOFs. However, the optimized Bi-WOOF outperformed the reported accuracies in other studies,^{7-9,31} as shown in Table 3.

On the other hand, Table 4 lists the benchmark geometric-based methods¹⁵⁻¹⁹ with the full-face graph and full-face graph + A-EMM, which are denoted as experiment I and experiment II, respectively. From these results, Buhari *et al.*,¹⁸ reported the highest accuracies of 76.67%, 75.04%, 81.41%, and 87.33% over the SMIC, CASMEII, CAS (ME)², and SAMM datasets, respectively. In Buhari *et al.*,¹⁸ the full-face graph utilized 68 landmarks from the raw images (denoted as $\mathbb{R}$). While, the full-face graph in experiment I and experiment II yield 74.62%, 74.41%, 75.11% and 75.01%, 74.55%, 76.21%, 75.53% over the SMIC, CASMEII, CAS (ME)² and SAMM datasets, respectively. From these results, the full-face graph in experiment II outperformed the full-face graph presented in¹⁸ with 8.11%, 1.1%, and 3.38% over the SMIC, CASMEII, and CAS (ME)² datasets, respectively. While, on the other hand, Buhari *et al.*,¹⁸ outperformed the results in experiment II with 4.7% over the SAMM dataset. While, in comparison with reported accuracies presented in Table 3, the full-face graph in experiment II achieved the highest performance.

Looking at the performance presented in Tables 3 and 4, the performance reported in Buhari *et al.*,¹⁸ registered the highest accuracy of 87.33% over SAMM datasets, respectively. While, the full-face with A-EMM outperformed the full-face graph performance presented in Buhari *et al.*¹⁸ From these results, the conclusion can be drawn that the geometric-based features are competing closely with the motion-based features. In terms of the computational time, the optimized Bi-WOOF running time is 0.36 seconds per sample (i.e., 2.7fps), while the full-face graph running time is 0.10 seconds per sample (i.e., 10fps), with a 640×480 image resolution, on an Intel i5-3470 machine. The reported running times include facial landmark detection and classification.

Conclusions

This paper analyzed the performance of motion-based features (i.e., Bi-WOOF) and geometric-based features (i.e., Full-face graph) for real-time micro-expression recognition systems. The results indicate that the optimized Bi-WOOF improved recognition accuracy of the baseline Bi-WOOF by up to 3.28% over the SAMM dataset. While, on the other hand, the full-face graph performance is improved by up to 1.20% with A-EMM over the SAMM dataset. Moreover, the full-face graph and full-face graph with A-EMM exhibit significant performance improvement over the baseline and the optimized Bi-WOOF by up to 18.32%. Though the full-face graph improved performance of recognition accuracy, the processing time could limit the readiness of the full-face graph features for real-time systems using high-speed cameras.

Data availability

Underlying data

The experiments were performed using four spontaneous datasets: (i) spontaneous micro-expression dataset (SMIC) dataset,²⁵ (ii) Chinese Academy of Sciences Micro-expression (CASMEII) dataset,²⁶ (iii) spontaneous macro-expressions and micro-expressions (CAS (ME)²) dataset,²⁷ and (iv) spontaneous actions and micro-movements (SAMM) dataset.²⁸ These are spontaneous micro-expression datasets which were used in this study with full details of these datasets in Li *et al.*,²⁵ Yan *et al.*,²⁶ Qu *et al.*,²⁷ and Davison *et al.*²⁸ Details to acquire these datasets used in this study are available at

www.oulu.fi/cmvs/node/41319 for SMIC,²⁵ fu.psych.ac.cn/CASME/casme2-en.php for CAS-MEII,²⁶ [fu.psych.ac.cn/CASME/cas \(me\) 2-en.php](http://fu.psych.ac.cn/CASME/cas%20en.php) for CAS (ME),²⁷ and personalpages.manchester.ac.uk/staff/adrian.davison/SAMM.html for SAMM.²⁸ Moreover, to evaluate the performance using a larger dataset, this paper merged the four datasets to form a COMBINED dataset. The COMBINED dataset is created from the raw images of all the four datasets with the source code available under extended data.

Extended data

Zenodo: Implementation of COMBINED micro-expression dataset and Setup files for real-time micro-expression recognition using motion and geometric features. <https://doi.org/10.5281/zenodo.5524141>.³³

The project contains the following extended data:

- Real-time micro-expression recognition using biwoof features (executable setup for micro-expression recognition using motion-based features).
- Real-time micro-expression recognition using full-face graph features (executable setup for micro-expression recognition using geometric-based features).
- Image re-scaler for COMBINED micro-expression dataset formation (Visual Studio 2010 source code written in C++).

Data are available under the terms of the Creative Commons Zero (CC0 v1.0 Universal).

Zenodo: Performance analysis of micro-expression recognition over different sample image sizes. <https://doi.org/10.5281/zenodo.5379773>.³⁴

This project contains the following extended data:

- Performance improvement over 140×170 sample size.
- Performance improvement over 240×340 sample size.
- Performance improvement over 560×680 sample size.
- Performance improvement over 1120×1360 sample size.

Data are available under the terms of the Creative Commons Zero (CC0 v1.0 Universal).

Author contributions

Conceptualization, Investigation, Methodology, Formal Analysis, Software, Visualization, Writing – Original Draft Preparation, Review & Editing; AM Buhari: Conceptualization, Resources, Formal Analysis, Methodology, Supervision, Writing – Review & Editing; CP Ooi: Conceptualization, Formal Analysis, Supervision, Validation, Writing – Review & Editing; VM Baskaran: Conceptualization, Methodology, Writing – Review & Editing; WH Tan.

Competing interests

No competing interests were declared.

Grant information

The authors declared that no grants were involved in supporting this work.

Acknowledgements

This work is supported in part by Research Management Centre (RMC), Multimedia University (MMU) Cyberjaya. In addition, I would like to thank Professor Raphaël C. W. Phan and Dr. KokSheik Wong for their guidance, assistance and support in all aspects of this study.

References

1. O'Sullivan M, Frank MG, Hurley CM, *et al.*: **Police lie detection accuracy: The effect of lie scenario.** *Law Hum Behav.* 2009; **330**(6): 530.
[PubMed Abstract](#) | [Publisher Full Text](#)
2. Frank MG, Maccario CJ, Govindaraju V: **Behavior and security. protecting airline passengers in the age of terrorism.** 2009.
3. Cohn JF, Kruez TS, Matthews I, *et al.*: **Detecting depression from facial actions and vocal prosody.** *2009 3rd Int Conf Affective Computing Intelligent Interaction Workshops.* IEEE; 2009; pages 1–7.
[Publisher Full Text](#)
4. Ekman P, Friesen WV: **Nonverbal leakage and clues to deception.** *Psychiatry.* 1969; **32**(1): 88–106.
[Publisher Full Text](#)
5. Shreve M, Godavarthy S, Manohar V, *et al.*: **Towards macro- and micro-expression spotting in video using strain patterns.** *2009 Workshop on Applications of Computer Vision (WACV).* IEEE; 2009; pages 1–6.
[Publisher Full Text](#)
6. Sze-Teng L, Gan YS, Yau W-C, *et al.*: **Off-apexnet on micro-expression recognition system.** *arXiv preprint arXiv.* 2018a: 1805.08699.
[Publisher Full Text](#)
7. Li Q, Zhan S, Xu L, *et al.*: **Facial micro-expression recognition based on the fusion of deep learning and enhanced optical flow.** *Multimedia Tools and Applications* 2019; **78** (20):0 29307–29322.
[Publisher Full Text](#)
8. Duque CA, Alata O, Emonet R, *et al.*: **Mean oriented riesz features for micro expression classification.** *Pattern Recognition Letters.* 2020; **135**: 382–389.
[Publisher Full Text](#)
9. Merghani W, Yap MH: **Adaptive mask for region-based facial micro-expression recognition.** *2020 15th IEEE Int Conf Automatic Face and Gesture Recognition (FG 2020)(FG).* IEEE Computer Society; 2020; pages 428–433.
[Publisher Full Text](#)
10. Yang L, Zhang X, Zhou J, *et al.*: **Graph-based facial affect analysis: A review of methods, applications and challenges.** *CoRR, abs/2103.2021a*: 15599.
[Reference Source](#)
11. Kazemi V, Sullivan J: **One millisecond face alignment with an ensemble of regression trees.** *Proc IEEE conf computer vision pattern recognition.* 2014: 1867–1874.
[Publisher Full Text](#)
12. Mohseni S, Zarei N, Ramazani S: **Facial expression recognition using anatomy based facial graph.** *2014 IEEE Int Conf Systems, Man, and Cybernetics (SMC).* IEEE; 2014; pages 3715–3719.
[Publisher Full Text](#)
13. Baltrušaitis T, Robinson P, Morency L-P: **Openface: an open source facial behavior analysis toolkit.** *2016 IEEE Winter Conf Applications Computer Vision (WACV).* IEEE; 2016; pages 1–10.
[Publisher Full Text](#)
14. Dapogny A, Bailly K, Dubuisson S: **Confidence-weighted local expression predictions for occlusion handling in expression recognition and action unit detection.** *Int J Computer Vision.* 2018; **126** (2): 255–271.
[Publisher Full Text](#)
15. Lei L, Li J, Chen T, *et al.*: **A novel graph-tcn with a graph structured representation for micro-expression recognition.** *Proc 28th ACM Int Conf Multimedia.* 2020: 2237–2245.
16. Xie H-X, Lo L, Shuai H-H, *et al.*: **Au-assisted graph attention convolutional network for micro-expression recognition.** *Proc 28th ACM Int Conf Multimedia.* 2020: 2871–2880.
[Publisher Full Text](#)
17. Zhou L, Mao Q, Dong M: **Objective class-based micro-expression recognition through simultaneous action unit detection and feature aggregation.** *arXiv preprint arXiv.* 2020: 2012.13148.
18. Buhari AM, Ooi C-P, Baskaran VM, *et al.*: **Facs-based graph features for real-time micro-expression recognition.** *J Imaging.* 2020; **6**(12): 130.
[Publisher Full Text](#)
19. Liu Y-J, Li B-J, Lai Y: **Sparse mdmo: Learning a discriminative feature for micro-expression recognition.** *IEEE Transactions on Affective Computing.* 2021b.
[Publisher Full Text](#)
20. Liong S-T, See J, Wong KS, *et al.*: **Less is more: Micro-expression recognition from video using apex frame.** *Signal Processing: Image Communication.* 2018b; **62**: 82–92.
[Publisher Full Text](#)
21. King DE: **Dlib-ml: A machine learning toolkit.** *J Machine Learning Res.* 2009; **100** (Jul): 1755–1758.
22. Liong S-T, See J, Wong KS, *et al.*: **Automatic apex frame spotting in micro-expression database.** *2015 3rd IAPR Asian conf pattern recognition (ACPR).* IEEE; 2015; pages 665–669.
[Publisher Full Text](#)
23. Wadhwa N, Rubinstein M, Durand F, *et al.*: **Riesz pyramids for fast phase-based video magnification** *US Patent 9,338,331.* May 10 2016.
24. Le Ngo AC, Yee-Hui O, Phan RC-W, *et al.*: **Eulerian emotion magnification for subtle expression recognition.** *2016 IEEE Int Conf Acoustics, Speech Signal Processing (ICASSP).* IEEE; 2016; pages 1243–1247.
[Publisher Full Text](#)
25. Li X, Pfister T, Huang X, *et al.*: **A spontaneous micro-expression database: Inducement, collection and baseline.** *2013 10th IEEE Int Conf Workshops Automatic Face Gesture Recognition (FG).* IEEE; 2013: 1–6.
[Publisher Full Text](#)
26. Yan W-J, Li X, Wang S-J, *et al.*: **Casme ii: An improved spontaneous micro-expression database and the baseline evaluation.** *PloS one.* 2014; **90**(1): e86041.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
27. Fangbing Q, Wang S-J, Yan W-J, *et al.*: **Cas (me)²: A database for spontaneous macro-expression and micro-expression spotting and recognition.** *IEEE Transactions on Affective Computing.* 2018; **90**(4): 424–436.
[Publisher Full Text](#)
28. Davison AK, Lansley C, Costen N, *et al.*: **Samm: A spontaneous micro-facial movement dataset.** *IEEE Transactions on Affective Computing.* 2018; **90**(1): 116–129.
[Publisher Full Text](#)
29. Buhari AM, Ling H-C, Baskaran VM, *et al.*: **Real-time high-resolution downsampling algorithm on many-core processor for spatially scalable video coding.** *J Electronic Imaging.* 2015; **24**(1): 013025.
[Publisher Full Text](#)
30. Bergstra J, Bengio Y: **Random search for hyper-parameter optimization.** *J Machine Lear Res.* 2012; **130**(2).
31. Li X, Hong X, Moilanen A, *et al.*: **Towards reading hidden emotions: A comparative study of spontaneous micro-expression spotting and recognition methods.** *IEEE transactions on affective computing.* 2017; **90**(4): 563–577.
[Publisher Full Text](#)
32. Liong S-T, Wong KS: **Micro-expression recognition using apex frame with phase information.** *2017 Asia-Pacific Signal Information Processing Association Annual Summit and Conference (APSIPA ASC).* IEEE; 2017; pages 534–537.
[Publisher Full Text](#)
33. Buhari AM, Ooi C-P, Baskaran VM, *et al.*: **Implementation of COMBINED micro-expression dataset and Setup files for real-time micro-expression recognition using motion and geometric features.** September 2021a.
[Publisher Full Text](#)
34. Buhari AM, Ooi C-P, Baskaran VM, *et al.*: **Performance analysis of micro-expression recognition over different sample image sizes.** September 2021b.
[Publisher Full Text](#)

Open Peer Review

Current Peer Review Status: ? ✖

Version 1

Reviewer Report 06 December 2021

<https://doi.org/10.5256/f1000research.76586.r101402>

© 2021 Prajapati H. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Harshadkumar B. Prajapati

Department of Information Technology, Dharmsinh Desai University, Nadiad, Gujarat, India

This manuscript performs experimental evaluation of two different methods: motion based and geometric based for facial expression classification. The manuscript contains some experiments and results, but the aims of the experiments appear to be random.

- The focus is on image sequences, which is addressed by motion based feature, but why geometric based feature? Does it deal with image sequences?
- The authors straightway use Bi-WOOF, a descriptor proposed by some other author. It is not justified why Bi-WOOF has been used; there are many other features. This choice should be scientifically justified.
- There are many descriptors that can specify motion in image sequences. The authors should discuss such descriptors.
- Current trend is applying Deep learning for ML problems; however, the authors use SVM (traditional ML model) for classification. Why? The readers must be convinced of the choice of SVM.
- The authors present results in Table 2 and Table 3, but it is not conveyed what are authors' contributions in those.
- In Table 3, at many places (-) is marked. Why? Have the authors have done these experiments? If YES, then why are many blank (-), if NO, then why are others' results included? Where are the results of the authors?
- Introduction section needs to be revised. An introduction section generally includes paragraphs on the following: (1) Motivation for the work, (2) Background/Technical information or terms, (3) Problem being addressed, (4) Existing major solutions and research gaps, and (5) Proposal of this paper and its significance.

- As Section (Methods) directly starts proposal using Bi-WOOF without any background concepts or theory, readers cannot follow/understand/appreciate the proposed work.
- The manuscript should discuss the proposal with suitable diagram(s), showing frame/video/image processing.
- The authors should show frame sequences of at least one sample from any one dataset.

Presentation

- The manuscript is very difficult to read for beginners; when any section starts, it must contain a brief paragraph discussing the content. E.g., in Section (Method), the authors directly discuss three different methods in Fig. 1, Fig. 2, and Fig. 3, but it is not clear how these are related to their proposal.

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

No

Are sufficient details of methods and analysis provided to allow replication by others?

No

If applicable, is the statistical analysis and its interpretation appropriate?

Not applicable

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

No

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Machine learning, image processing, distributed computing

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Reviewer Report 25 October 2021

<https://doi.org/10.5256/f1000research.76586.r96609>

© 2021 Takalkar M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Madhumita Takalkar

School of Computer Science, University of Technology Sydney, Ultimo, NSW, Australia

The paper talks about optimizing Bi-WOOF and using geometric based features to detect and recognise facial micro-expressions. Please find my comments as below.

1. The idea does not sound novel. There have been further developments in the detection and recognition of micro-expressions which includes deep learning models which exhibit higher accuracies than presented here in the paper. Why would the authors not compare with the latest methods? Following this comment, there are two papers the authors can refer: Madhumita *et al.* 2021¹ and Thuseethan *et al.* 2021².
2. Buhari *et al.* (2020)³ performs better and presents higher accuracy than the proposed methods. What is the message behind the proposed method? Is the computation time reduced with the proposed method if not the accuracy improvement?
3. SMIC database does not have the apex frame annotated. Did the author perform any methods to identify the apex frame? If so that description is missing.
4. The paper says Buhari *et al.* (2020)³ performs detection and recognition on the raw images. How are the authors comparing raw images with processed images? This may mean that processing the images is reducing the accuracy.
5. Considering the limited samples in the datasets, only accuracy may not be sufficient to prove the efficiency of the developed model. It would be advisable to include other performance evaluation metrics such as AUC, F1 score, and Confusion matrix.

References

1. Takalkar M, Thuseethan S, Rajasegarar S, Chaczko Z, et al.: LGAttNet: Automatic micro-expression detection using dual-stream local and global attentions. *Knowledge-Based Systems*. 2021; **212**. [Publisher Full Text](#)
2. Thuseethan S, Rajasegarar S, Yearwood J: Deep Continual Learning for Emerging Emotion Recognition. *IEEE Transactions on Multimedia*. 2021. [Publisher Full Text](#)
3. Buhari AM, Ooi CP, Baskaran VM, Phan RCW, et al.: FACS-Based Graph Features for Real-Time Micro-Expression Recognition. *J Imaging*. 2020; **6** (12). [PubMed Abstract](#) | [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Computer Vision, Deep Learning, Image Processing

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research