



PAPER • OPEN ACCESS

Classification of Body Mass Index Based Facial Images using Empirical Mode Decomposition

To cite this article: H Ali *et al* 2021 *J. Phys.: Conf. Ser.* **1878** 012010

View the [article online](#) for updates and enhancements.

You may also like

- [Design of Smart Medical Mirror using Raspberry Pi](#)
R Mariappan, P Naga Prasanna, SK Tehmeem Sulthana *et al.*
- [Improving brain-machine interface performance by decoding intended future movements](#)
Francis R Willett, Aaron J Suminski, Andrew H Fagg *et al.*
- [Urban sprawl and health: a review of the scientific literature](#)
Dario Genovese, Stefania Candiloro, Antonio D'Anna *et al.*

ECS The Electrochemical Society
Advancing solid state & electrochemical science & technology

247th ECS Meeting
Montréal, Canada
May 18-22, 2025
Palais des Congrès de Montréal

Showcase your science!

Abstracts due December 6th

Classification of Body Mass Index Based Facial Images using Empirical Mode Decomposition

H Ali, O S Yee, W K Wan Ahmad

Mechatronic Engineering Programme, Faculty of Electrical Engineering Technology, University Malaysia Perlis, 02600 Arau, Perlis, Malaysia.

Abstract. Human faces contain rich information. Recent studies found that facial features have relation with human weight or body mass index (BMI). Decoding “facial information” from the face in predicting the BMI could be linked to the various health marker. This paper proposed the classification of body mass index (BMI) based on appearance based features of facial images using empirical mode decomposition (EMD) as feature extraction technique. The facial images that describe the body mass index was extracted using EMD to obtain a set of significant features. In this framework, the facial image was decomposed using EMD to produce a small set of intrinsic mode functions (IMF) via sifting process. The IMF features which exhibit the unique pattern were used to classify the BMI. The obtained features were then fed into machine learning classifier such as k-nearest neighbour and support vector machines (SVM) to classify the three BMI classes namely normal, overweight and obese. The obtained results show that the IMF2 feature using SVM classifier achieved recognition rate of 99.12% which show promising result.

1. Introduction

Human faces contain rich information such as facial or emotion expression, age, gender, ethnicity and personal traits. It also contains a number of cues that could predict the health status for example, the facial attractiveness was found to predict future longevity [1]. Recent studies in human perception show that the facial features have relation with human weight body mass index (BMI). Decoding “facial information” from the face, reveal useful cues, thus it received great attention from psychologist, computer scientists and behavioural sciences. BMI is a measure of body fatness based on body height and weight of an individual [1]. Decoding information of body mass index (BMI) from the face could also be linked to the various health marker. Obesity can highly affect our body health and have significant relation associated with the risk of various diseases such as diabetes, cardiovascular, stroke and some common adults cancer [2]. Hence, by utilizing BMI based facial images it would be able to is very crucial for us to aware the level of healthiness through BMI for each individual in order to encourage the awareness of health.

According to World Health Organization (WHO), BMI is an index used to measure body fat based on body height and weight for individual. According to the values of BMI, people are divided into four categories such as underweight (< 18.5), normal (18.5 - 24.9), overweight (25.0 - 29.9) and obesity (> 30.0) [3]. Given an individual's height and weight, the body mass index (BMI) can be calculated as follow [4]:

$$BMI = \frac{Weight(kg)}{Height(m)^2} \quad (1)$$



Traditionally, to measure the weight and height of an individual, a person will have to use some scales or tools to get it measured accurately. The person will have to stand straight and upright so that the measurement can be taken accurately. However, it is impossible for a disable person to stand in straight. Therefore, unlike the traditional method, the BMI prediction of an individual using computer vision is more easily and reliable. In previous study, Wen and Guo [5] utilized Active Shape Model (ASM) which proposed by Cootes et al. [6] to detect the facial fiducial points. The detected facial fiducial points detected was used to compute the geometry and ratio of a face. Meanwhile, Wen [7] employed regression methods namely, least square estimation, Gaussian processes for regression and support vector regression in predicting the BMI value. Besides, Barr et al. [8] have utilized regression function to predict BMI value using facial image (photograph). They measured the distance and ratio between the extracted facial landmarks features in order to characterize the facial fatness before fed to regression functions. Similarly, Dantcheva et al. [9] have proposed regression technique based on the 50-layers ResNet-architecture on single-shot face image in predicting the weight, height and BMI using proposed dataset. Their results suggested that the facial images contain discriminatory information pertaining to the height, weight and BMI comparable to that body-images and videos.

The aforementioned of previous works have utilized the geometric based facial features of face images in estimating the body mass index. Although extracting the facial features based on geometric approach offer some advantages however, decoding “facial information” using geometric based facial features may contribute to low accuracy due to an adequate or insufficient information associated with BMI and also time consuming of manual labour. Therefore, this work proposes an appearance based features of facial images in predicting the BMI using Empirical Mode Decomposition (EMD) technique. The key benefit of using EMD is that the basic functions can be directly extracted from the signal itself, which offers a truly data-driven approach and also brings not only high decomposition efficiency but also sharp frequency [10]. In this work, the facial images were subjected to EMD technique to produce a set of intrinsic mode functions (IMFs) via sifting process. The obtained features were then fed to two different classifiers, which are k-Nearest Neighbour (KNN) and Support Vector Machines (SVM) independently for evaluation of classification performances. In this work, empirical mode decomposition technique is used as texture feature extraction to classify the three classes of BMI, which are normal, overweight and obese, based on facial images. The spatial relations between the neighbouring pixels can determine the texture features.

2. Material and methods

This section presents the methodology of the proposed method in classifying three different classes of BMI based on the face images. Figure 1 shows the block diagram of the proposed method. It consists of four (4) steps; BMI facial image database, image pre-processing, features extraction and classification. In the image pre-processing stage, it consists of two main steps which are face detection and cropping of face. In this framework, the facial frame firstly is being detected and cropped, then the cropped images will subject to grayscale transformation and scaling to uniformize all the images. Next, pre-processing steps such as histogram equalization was conducted before fed into features extraction stage for quality improvement of the quality of facial images into order to obtain a robust and high accuracy estimation system. The extracted features are used as input to KNN and SVM classifiers for BMI classification.

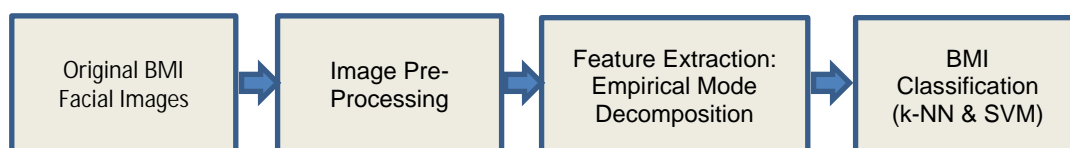


Figure 1. Block diagram of the proposed method.

2.1. BMI facial image database

Publicly available BMI face image database, namely MORPH-II [11] has been employed in this experiment. The database contains of 4,206 face images that consist of four classes of BMI such as underweight, normal, overweight and obese. The subjects consist of 1,768 females and 2,438 males in the dataset, together with the individual BMI information are provided. Due to the limited number of underweight subjects, only three classes of BMI have been considered. Totally 300 facial images (100 normal, 100 overweight and 100 obese) have been used in this experiment.

2.2. Pre-processing of BMI facial images

In this research, the image pre-processing is used to eliminate the unwanted noise or irrelevant parts of face region. The pre-processing such as face detection, face cropping, grayscale transformation, image rescaling and histogram equalization have been applied. For face detection, we employed the Viola Jones face detection algorithm [12]. The advantages of Viola Jones are they have high detection accuracy and fast processing speed, thus become effective in detecting the face area in the image. Figure 2 depicts the example of face detection for the three classes. The face automatically detected via bounding box. Then the face region were cropped and the unnecessary information were removed. The cropped image was then transform from RGB (Red, Green, Blue) into grayscale having scale from 0 to 255. Normalization of images usually performs various steps such as image resizing, translation and rotation. Each raw image has different size and resolution, hence the size of cropped face area will also different. Therefore, resizing of images is very important to have uniform size of face image. In this study, the original facial image was scaled into fixed size of dimension 152 x 152 pixels. Histogram equalization is one of the techniques used to increase the contrast of the images [13]. Histogram equalization enhances the contrast of images by transforming the values of image intensity. Figure 3 shows the pre-processed image after the cropped, grayscale, scaling and histogram equalization for normal, overweight and obese.

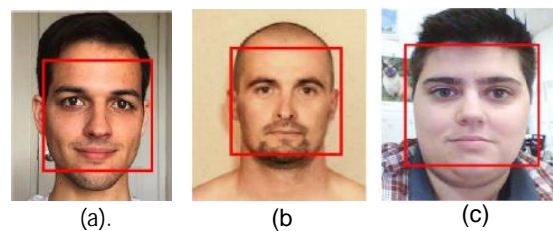


Figure 2. Face detection using Viola-Jones face detection: (a) normal, (b) overweight and (c) obese.

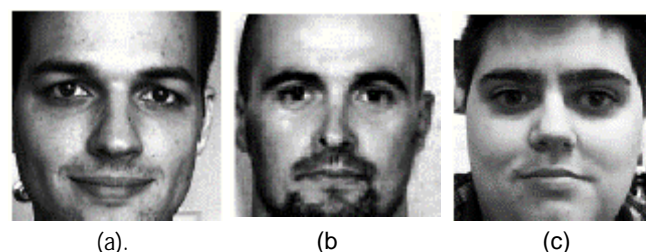


Figure 3. Example of the result after face cropped, grayscale, scaling and histogram equalization: (a) normal, (b) overweight and (c) obese

2.3. Feature extraction: Empirical mode decomposition (EMD)

Empirical Mode Decomposition (EMD) is a multiresolution decomposition techniques introduced by Huang et al. [10] which is adaptive and appear to be suitable for nonlinear, non-stationary data analysis. EMD decomposes signal into a small set of frequency oscillation namely, intrinsic mode functions (IMFs). The EMD is data driven and no prior knowledge is needed. The study of EMD was initially proposed for ocean waves, and then it was found successfully in face analysis [14, 15], biomedical

applications [16, 17] and etc. The EMD (one-dimensional) can be extended into a two-dimensional version for textural analysis [18]. The major advantage of EMD is that it able to capture information about local trends in the signal by measuring and quantizing oscillations. Such oscillation can be quantized by a local high frequency (local details) and correspondingly a local low frequency (local trend). The basic concepts of EMD lie on the utilization of the interpolation technique such as cubic spline, sifting process to extract the intrinsic mode function and numerical convergence criteria to stop the sifting process.

In EMD algorithm, there are two criteria must be achieved to extract the intrinsic mode functions:

- (1) the number of extrema and the number of zero crossings are either equal or differ at most by one; and
- (2) the mean of its upper and lower envelopes equal zero.

The above two conditions fulfill the physically necessary condition to define a meaningful instantaneous frequency. The whole procedure of EMD algorithm as follow:

For given the signal $s(t)$,

1. Locate the local maxima and local minima of $s(t)$. Let $d_0(t) = s(t)$.
2. Interpolate the local maxima and the local minimum to create upper envelope $e_u(t)$ and lower envelope $e_l(t)$, respectively using cubic B spline.
3. Compute the mean of envelope $m(t)$: $m(t) = (e_u(t) + e_l(t))/2$.
4. Compute the details by $d_1(t) = d_0(t) - m(t)$ (sifting process).
5. Iterate steps 1-4 on the residual signal until the details signal $d_k(t)$ can be considered as IMF: $c_1(t) = d_k(t)$
6. Iterate steps 1-5 on the residual $r_n(t) = s(t) - \sum_{i=1}^n c_i(t)$ in order to obtain all the IMFs $c_1(t), c_2(t), \dots, c_n(t)$ of the signal.

The process stops when the residual signal is either a constant, a monotones or a function with only one extrema. The power of EMD method is that once these IMF's have been found, they can easily go back and forth from them to the original data. In other words, the original data can be recovered by summing all the IMF's together accommodating for minor variation due to the interpolation present in the algorithm. The EMD can be expressed mathematically as: $x = \sum_{n=1}^N c_n(t) + r_n(t)$.

From the process we can see that, the lower order IMFs capture fast oscillation modes while the higher order IMFs capture low oscillation modes. In this work, we apply two-dimensional EMD to decompose the BMI facial images into a small set of intrinsic mode functions (IMFs). The IMFs is the features extracted at multiple scales or spatial frequencies via sifting process.

2.4. BMI classification

In this experiment, k-nearest neighbor and support vector machines classifiers have been used to classify the three classes of BMI. The following subsections discuss the aforementioned classifiers.

K-Nearest Neighbour (k-NN) classifier. k-NN is a basic classifier in machine learning that predicts the class of a new test data based on the class labels of its k neighbors in the feature space. Euclidean distance is used as metric measurement. KNN algorithm is a simple and basic machine learning algorithm that can be implemented easily and can perform well in solving classification and regression problems despite of its simplicity. k-NN is defined as a 'lazy learner' because training is not required while using k-NN because it does not learn anything from the training data [20]. Its instance-based learning characteristic leading to it only learns the training data stored when making the real time predictions. However, this algorithm can adapt to the changes of input data very quickly.

Support vector machines (SVM) classifier. SVM is a supervised machine learning techniques introduced by Vapnik [19] and gaining the popularity due to attractive features and promising empirical performance. The goal of SVM is to find the 'optimal' hyperplane that maximizes the margin (maximizes the distance between it and the nearest data point of each class). SVM can be categorized into linear separable and non-linear separable class. Figure 4 illustrates the SVM with linear separable case for binary class.

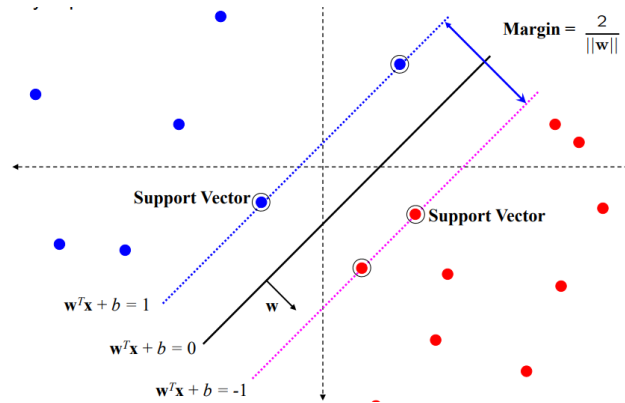


Figure 4. SVM with linear separable [19].

The equation of hyperplane given by:

$$(w \cdot x) + b = 0 \quad (1)$$

which implies,

$$y_i[(w \cdot x_i) + b] \geq 1, \quad i = 1, 2, \dots, k \quad (2)$$

Support vectors are the points that lie on the margin. Such a hyperplane with maximizes margin is called the optimal hyperplane. Margin is a distance between the hyperplane and the closest data points in which:

$$d_1 + d_2 = \frac{2}{\|w\|} \quad (3)$$

Hence, the hyperplane that *optimally separate* the data is the one that minimizes:

$$\Phi(w) = \frac{1}{2} \|w\|^2 \quad (4)$$

When the data is non-linear separable, slack variables ξ are introduced. The resulting problem is then to minimize:

$$\frac{1}{2} \|w\|^2 + C \left(\sum_i L(\xi_i) \right) \quad (5)$$

Where C is the adjustable penalty term and L is the loss function which in the case is a linear loss function. The optimization of Equation (1) using a Lagrangian method with linear loss function is to maximize:

$$L_D(w, b, \alpha) = \sum_i \alpha_i - \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (6)$$

Subject to:

$$0 < \alpha_i \leq C, \quad i = 1, 2, \dots, k \quad (7)$$

And

$$\sum_i \alpha_i y_i = 0 \quad (8)$$

Where, α_i is a set of Lagrangian multipliers. This optimization can be solved by using the standard quadratic programming technique. Once the problem is optimized, the parameters of optimal hyperplane are:

$$w = \sum_i^k \alpha_i y_i x_i \quad (9)$$

Note that, α_i is zero for every x_i except the ones that lie on the margin. The training data with non-zero α_i are called support vectors.

In the case of nonlinear separable problem, a kernel function is used to map the feature space into higher dimensional feature space so as to make the problem become linearly separable. A common choice of kernel functions includes linear, polynomial, radial basis function and sigmoid functions. In this experiment, radial basis function or radial kernel was adopted to map the IMF features into higher dimensional IMF features space. The results will be presented in the subsequence section.

3. Results and discussion

To evaluate the effectiveness of the proposed method, a publicly available BMI Facial Image database called MORPH-II database has been employed in this experiment. The preprocessed facial images were subjected to features extraction technique using Empirical Mode Decomposition. As the name implies, the EMD technique was used to extract features at multiple scales or spatial frequencies of nonlinear and non-stationary data. These features called intrinsic mode functions (IMFs) were extracted by sifting process. In this work, the EMD was used to decomposed facial images into a number of intrinsic mode functions (IMFs) via a sifting process. Then, the extracted IMFs were fed as input into k-NN and SVM classifiers in order to classify the BMI classes which are normal, overweight and obese.

3.1. Applied EMD on BMI Facial Images

The 2D decomposition via sifting procedure of an image provides a representation that is able to interpret. Every mode (IMF) contains information of a specific scale, which is conveniently separated. Spatial information is retained within the mode. Figure x shows the EMD decomposition of BMI facial images into a set of intrinsic mode functions which is IMF1, IMF2, IMF3 and residue.). The first component (IMF1) contains the finest spatial scale in the signal. Whereas, the residue now contains information about larger scales [18]. More specifically, the first IMF associated with the smallest time scale corresponds to the fastest time variation of the data. As the decomposition process proceeds, the time scale is increasing and hence, the mean frequency of the mode is decreasing. It can be seen from Figure 5 that the a set of IMF exhibits the pattern structure from the finest to the coarsest relative to original image. The IMF1 gives the most distinct details including the boundary lines along the face region such as eyes, nose and mouth. In this research, IMF1 and IMF2 were extracted for further analysis and classification since IMF1 and IMF2 have the characteristics of having greater extreme magnitude in which they contribute to the higher local information describing the behavior of BMI classes. For the sake of accuracy and efficiency of the proposed system, IMF3 and residue are discarded because they preserve minor facial structure which is insufficient for classification. To stop the sifting process, we used the standard deviation (SD) which is 0.5.

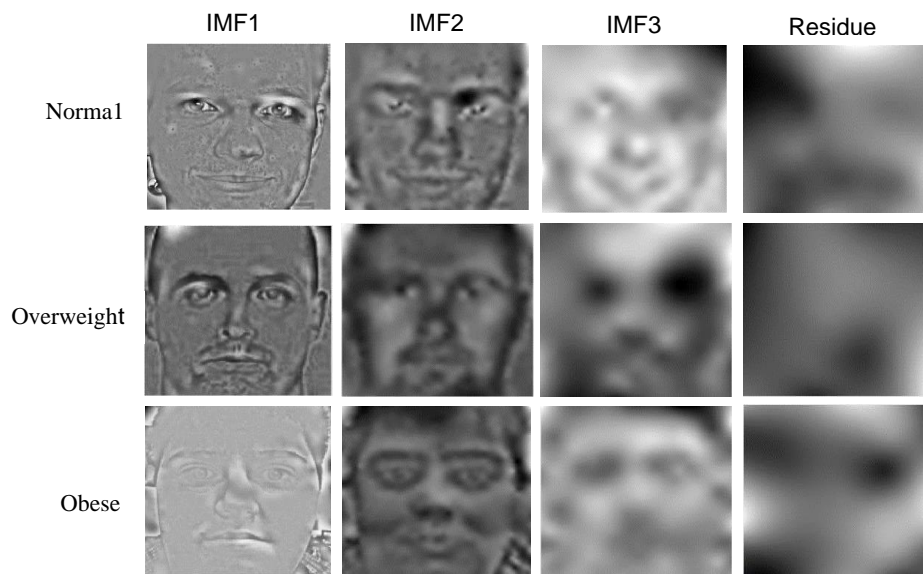


Figure 5. Performance of applied two-dimensional EMD on facial image for normal, overweight and obese.

3.2. Recognition results and discussion

In this experiment, 10-fold cross validation technique was employed. The whole database (300 images) is randomly divided into ten sets of having the same distribution of different BMI facial images. At each time of the process, nine sets (270 images) used for training and the remaining set (30 images) used for testing. The process was repeated for 10 times using different sets and the final average was calculated. For the performance comparison, 5-fold cross validation was used as the baseline reference. Figure 6 shows the recognition rates of IMF1 features using k-NN classifier. It can be seen from Figure 6 that the recognition rate of IMF1 features using $k = 2$ has achieved the highest accuracy which is 98%.

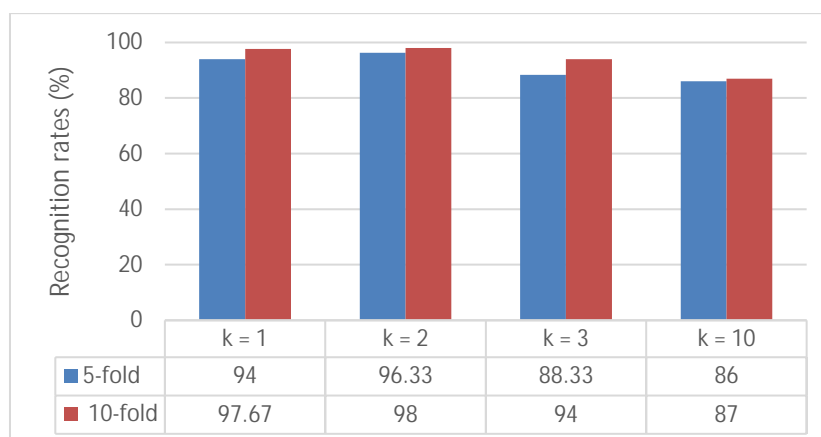


Figure 6. Recognition rates of IMF1 features + k-NN classifier

Table 1 shows the confusion matrix of IMF1 features for $k = 2$. As observed in Table 1 the normal BMI has achieved perfect recognition rate, thus there is no problem in classifying the normal BMI. In contrast, the overweight BMI contributes to the lowest recognition rate which is 96% or 4 out of 100 are misclassified. It is clear that the system mostly confused to classify between the overweight with obese. This may be due to interclass similarity that exist between overweight and obese.

Table 1. Confusion matrix of using IMF1 features with k-NN classifier.

Status	Normal	Overweight	Obese	Average
Normal	100	1	0	
Overweight	0	96	2	
Obese	0	3	98	
Total				98.0%

On the other hand, Figure 7 depicts the recognition rates of IMF2 features using similar classifier. Recognition rate using IMF2 features with $k = 1$ has the highest accuracy which is 99.33% using 10-fold cross validation. However, when $k = 2$, both fold give comparable results which approximate $\sim 97\%$. The recognition rates shows decreasing as we increase the value of k of nearest neighbour.

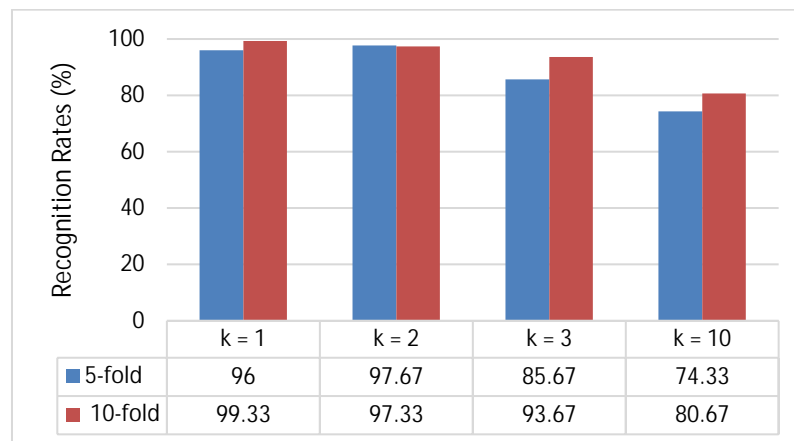
**Figure 7.** Recognition rates of IMF2 features + k-NN classifier

Table 2 illustrates the confusion matrix of IMF2 features using nearest neighbour of $k = 2$. It can be seen from Table 2 that both BMI overweight and obese achieved perfect recognition rates, whereas as BMI normal contributes 98% or 2 out of 100 are misclassified. This can be inferred that, the IMF2 features able to distinguish the behaviour of overweight and obese even there is exist of the interclass similarity.

Table 2. Confusion matrix of using IMF2 features with k-NN classifier.

Status	Normal	Overweight	Obese	Average
Normal	98	0	0	
Overweight	2	100	0	
Obese	0	0	100	
Total				99.33%

The experiments were further investigated using the SVM classifier. Table 3 shows the confusion matrix of IMF1 features using SVM classifier. As observed, SVM classifier was perfectly classify the BMI normal with optimal parameters ($C = 1000$ and kernel option = 13) while BMI overweight and obese sharing the same recognition rates which is 99% accuracy.

Table 4 shows the results of confusion matrix using component frequency IMF2 features with SVM classifier. It can be observed in Table 4 that the recognition rates of IMF2 features + SVM classifier slightly improved as compared with the results presented in Table 3. The recognition rate of

IMF2 features + SVM classifier has achieved 100% accuracy for both BMI overweight and obese with optimal parameter using $C = 100$ and kernel option = 12. In addition, recognition rate of BMI normal was nearly perfect with accuracy of 99%, thus considered as promising results. This can be inferred that the features component of IMF2 able to distinguish the three BMI classes. Furthermore, the IMF2 features may have the characteristics of having greater extreme magnitude in which they contribute to the higher local information describing the behaviour of BMI classes. Even though, the results of classification of BMI using EMD show promising, however the system is facing computational expensive due to high dimensional space of IMF features.

Table 3. Confusion matrix of using IMF1 features with SVM classifier.

Status	Normal	Overweight	Obese	Average
Normal	100	0	0	
Overweight	0	99	1	
Obese	0	1	99	
Total				99.33%

Table 4. Confusion matrix of using IMF2 features with SVM classifier.

Status	Normal	Overweight	Obese	Average
Normal	99	0	0	
Overweight	1	100	0	
Obese	0	0	100	
Total				99.67%

4. Conclusion

This paper has presented the classification of body mass index (BMI) based facial images using empirical mode decomposition technique for e-health awareness. In this experiment, publicly available MORPH-II database has been employed. The BMI facial images were firstly pre-processed to remove unwanted background as well as enhancing the facial images. The pre-processed image were then subjected to feature extraction using EMD. In this framework, EMD which is adaptive to nature used to decompose the BMI facial images to produce a small set of intrinsic mode functions (IMFs). Three IMFs (IMF1, IMF2, IMF3) and a residue were obtained from original facial images by sifting process whose exhibit the unique pattern of texture information. The extracted IMF1 and IMF2 features then were used as input to the k-NN and SVM classifier independently in order to classify the three BMI classes (normal, overweight and obese).

Based on the results obtained, it shows that the IMF2 features + SVM classifier has the highest recognition rate which is 99.67%. This can be inferred that the IMF2 has the ability to distinguish different class of BMI based on facial images with the assist of machine learning classifier specifically SVM. The used of EMD technique as feature extraction show a promising results in recognizing the BMI classes. Even though, the EMD has achieved good performance, however this technique has computational expensive as the number of database increases. Therefore, further study should be conducted in minimizing the computational cost so that real time health awareness system could be developed.

Acknowledgments

A portion of this work has used MORPH-II database as BMI facial images.

References

- [1] Gallagher D, Visser M, Sepulveda D, Pierson R, Harris T, Heymsfield S 1996 *J. Epidemiol.* **143** p 228–239
- [2] Renehan A G, Tyson M, Egger M, Heller R F, Zwahlen M. 2008 *Lancet* **371** p. 536–546
- [3] Henderson J J and Anglin J M 2003 *Evol. Hum. Behav.* **24** p. 351-356
- [4] Shackelford T K and Larsen R J 2005 *Evol. Hum. Behav.* **26** p. 213-226
- [5] Wen L and Guo G D 2013 *Image and Vision Computing* **31** p. 392–400
- [6] Cootes T F Taylor C J, Cooper D H and Graham J 1995 *Computer vision and image understanding* **61** p. 38-59
- [7] Wen L 2015 *Graduate Theses, Dissertations, and Problem Reports.* 6938.
- [8] Barr M L, Guo G, Colby S E and Olfert M D 2018 *Technologies* **6** 83 doi:10.3390/technologies6030083
- [9] Dantcheva A, Bremond F, Bilinski P 2018 *International Conference on Pattern Recognition (ICPR)*, Beijing, China
- [10] Huang N E , Shen Z , Long S R , Wu M C , Shih H H , Zheng Q , Yen N-C , Tung C C and Liu H H 1998 *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* **454** p. 903-995
- [11] Ricanek K, Tesafaye T 2006 *IEEE Conf. on AFGR* p. 341–345
- [12] Viola P and Jones M 2001 *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* **1** p. 511-518
- [13] Wu X, Liu X, Hiramatsu K, and Kashino K 2017 *IEEE International Conference on Image Processing (ICIP)*, 2017.
- [14] Bhagavatula R. & Savvides M 2007 *IEEE International Conference on Acoustics, Speech and Signal Processing* 2007.
- [15] Ali H, Hariharan M, Yaacob S, Adom A H 2015 *Expert Systems With Applications* **42** 1261
- [16] Gallix A, Gorriz J M, Ramirez J, Illan I A, Lang E W 2012 *Expert Systems with Applications*, **39** 13451
- [17] Panchori R B 2008 *Research Letter in Signal Processing* **2008**.
- [18] Nunes J, Bouaoune Y, Delechelle, E, Niang O & Bunel P 2003 *Image and Vision Computing* **21** 1019
- [19] Gunn S R 1998 *Technical Report*.
- [20] Bronshtein A 2019 *Medium*.