Body Mass Index (BMI) of Normal and Overweight/Obese Individuals based on Speech Signals

C. C. Lim[1], C. Y. Fook[1], V. Vikneswaran[1], M. A. H. Abdul Rahim[1] and M. Hariharan[2]
[1]*Biomedical Electronic Engineering Programme, School of Mechatronics, Pauh Putra Campus, Universiti Malaysia Perlis, Malaysia*
[2]*Department of Biomedical Engineering, SRM University, Chennai, India.*
cclim@unimap.edu.my

***Abstract*— Conventional method for measuring Body Mass Index (BMI) for individuals are using calibrated weight scale and measuring tape. However, there are certain cases in which the conventional method of measuring the BMI is not accessible. Thus, this experiment was proposed to overcome the problem using speech approach. In order to develop an effective BMI measuring system, speech signals of 30 subjects were recorded using a microphone. The dimension of the speech signal was reduced by extracting the relevant features using LPCC, MFCC and WPT based energy and entropy features. Lastly, both k-Nearest Neighbour (kNN) and Probabilistic Neural Network (PNN) were used to measure the BMI of an individual. The kNN classifier (97.50%) gives promising accuracy compared to PNN classifier (96.33%).**

***Index Terms*—Estimate BMI Using Speech; MFCC; LPCC; WPT; kNN; PNN.**

I. Introduction

The Body mass index (BMI) is an effort to measure the tissue mass that included muscle, fat, and bone in an individual. It can be used to categorise that person as underweight, normal weight, overweight, or obese based on BMI value. The BMI is very important for a patient that had been diagnosed with obesity-related diseases because they must constantly measure their body mass index on the daily basis of their weight because of the body weight changes over time. Based on American Association on Health and Disability reports, there are almost 20% of overweight adults, and approximately 30% of obese adults also have a disability, while roughly 36% of people with disabilities are also obese [1]. Nowadays, it also is widely used as a risk factor for the incidence of various health issues. In addition, it is also widely used as public health guidelines.

Recently, the speech signal of an individual utterance had used to measure BMI [2]. The conventional method of measure BMI has its drawbacks. First, overestimates in very muscular individuals. Some athletes have a pack of muscles than a normal person. Second, underestimates individuals who have lost muscle mass such as elderly like folks or children. Third, exaggerated in individuals with the extreme of body height between short and tall individuals. Lastly, exaggerated in the presence of oedema.

Based on the limitation of the conventional method, this experiment is proposed to investigate the relationship between speech signals and BMI categories. Besides that, the individuals will be predicted as normal, obese or overweight using speech signal in this experiment. The Mel Frequency Cepstral Coefficients (MFCCs), Linear Predictive Cepstral Coefficients (LPCCs) and Wavelet Packet Transform based features (WPT) techniques were used as feature extraction. The comparisons of three feature extraction techniques that used for classifying the BMI categorise were discussed. Finally, two different classifiers namely k-nearest neighbour (kNN) classifier and Probabilistic Neural Network (PNN) were used to identify the BMI categories of an individual [3].

II. Literature Review

Body mass index (BMI) is normally used to estimate your body fat. This helps to determine if the weight is within the normal range, underweight or overweight. Based on previous research works, people who are obese have a higher chance to have diseases such as diabetes, cancer, stroke or even cardiovascular disease [4]. The BMI frequently utilised by the healthcare personnel to diagnose any possibilities of disease development to the patients. Besides, the obesity can lead to hypertension and hyperlipidemia [1].

The mathematical method of calculating and estimating the BMI is using the formula which is the divided the weight in kilogram to the square of height in the meter which will present in the meter which will present in a metric unit of kg/m2 [5]. Lambert Adolphe Jacques Quetelet proposed that BMI is a measurement the relationship between body weight and height [6] and have been used commonly for classifying underweight, normal, overweight, and obese of an individual. However, the BMI varies by several factors. The BMI for individuals in Europe and Asia are different. It is because the body size, weight and height of the European people and the Asian people are different. Besides that, the BMI also varies regarding the age, sex, demographic and region [7].

According to Lee et al. [2], a correlation of body size of the speaker which includes the weight and height of the speaker are used to determine the BMI of a person [8], [9]. By using Goldwave v5.58 audio software, the speaker signals are recorded and saved in a wav file with the sampling frequency of 44.1kHz [2], [8]. The subjects were requested to seat comfortably, and each session will allow an hour rest before next session. Besides that, the subjects are also required to talk naturally. During the experiment, the subjects need to pronounce each vowel for 3 seconds while having 1 seconds silence in between. After that, the subject needed to read a given sentence repeatedly for two times before finishing the experiment process [1, 7].

A significant correlation can be observed by using speech to the body size of a speaker. However, certain criteria affected the body size. Consideration of people ages and sex need to be taken into account [8, 9]. In the previous study, they used formant frequencies and the average of the fundamental frequency (F0) to correlate between the speech and the speaker body size parameters which are the weight or height [11]. The study of vocal tract length (VTL) is used to underline the correlations between the weights of the speaker [11].

As for early research, a different task is done by using 15 males and 15 females. 30 people with the normal capability of hearing to be as the judges and their speech were recorded. The judges were asked to estimate the height of the speaker from the audio recording played to them [11].

In this experiment, speech processing is focused on predicting the BMI categories of normal, overweight and obese people accurately. Thus, certain criteria and aspects need to be considered before conduct the experiment as the speech signal is highly sensitive to noise.

III. METHODOLOGY

The overall BMI classification flowchart is shown in Figure 1. It generally consists of 4 main parts such as data collection, pre-processing, feature extraction and classification of BMI.

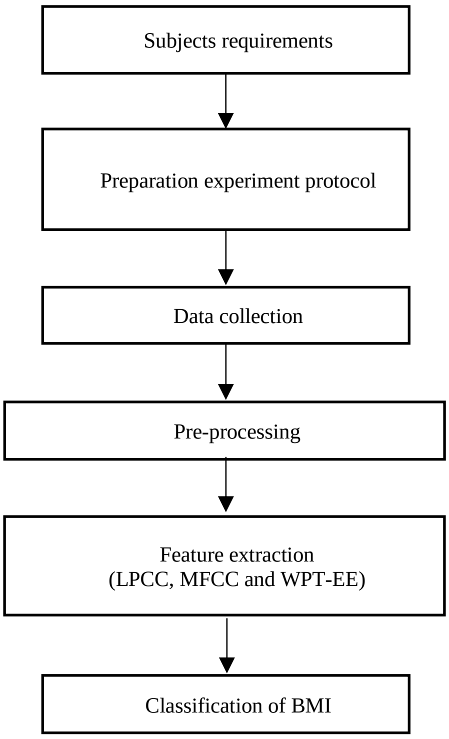


Figure 1: Flowchart of the BMI classification system

A. Database Development

This experiment consists of 30 subjects which are divided based on three different BMI categorise and gender (15 males and 15 females). The age range of the subjects is between 21 and 27 years old. The overall database information for BMI classification in this experiment was summarised in Table 1.

All the subject participated are healthy or without disease that affects their vocal cords and voices. The subject that smoker or heavy smoker will also be considered. The heavy smoker has a higher tendency towards disease that related with vocal [7].

Table 1
Database Information for BMI Classification

Classes	Normal	Overweight	Obese
Number of subjects	10	10	10
Gender	5 males 5 females	5 males 5 female	5 males 5 female
Age range (years old)	21 to 27	21 to 27	21 to 27

The preparation before data collection, the subject has to calm down. After that, the /ah/ sound will be recorded for 5 seconds and repeat for ten times for each subject. For each trial, the position of /ah/ sound is mixed in between the sequence with the digits from ‘0’ to ‘9’. The speech signals were sampled at 44.1 kHz and then down sampled to 16 kHz for speech signal analysis. The pre-emphasis filter had used to remove the noises and flatten the signal.

B. Feature Extraction

In this stage, three different feature extraction techniques were used in this experiment. They are Linear Predictive Cepstral Coefficients (LPCCs), Mel Frequency Cepstrum Coefficients (MFCCs), and Wavelet Packet Transform based Energy and Entropy (WPT-EE) features [13]–[16]. This short-term cepstral and nonlinear feature vectors can be used to represent the signal characteristic of the BMI categories, and it will use as input to the kNN classifier and PNN classifier. The LPCCs and MFCCs extraction process had shown in Figure 2 and Figure 3 respectively.

For wavelet packet transform based energy and entropy features, it had been extracted from the wavelet packet coefficients by using Equation (1) and Equation (2).

$$Log\ Energy = \sum_{i=1}^{m}\sum_{j=1}^{n}\log({W_{i,j}}^2) \quad (1)$$

$$Entropy = -\sum_{i=1}^{m}\sum_{j=1}^{n}\mathrm{p}(W_{i,j})\log_{10}p(W_{i,j}) \quad (2)$$

where: m = Denotes scale parameter
n = Frequency parameter
W = Wavelet packet coefficients

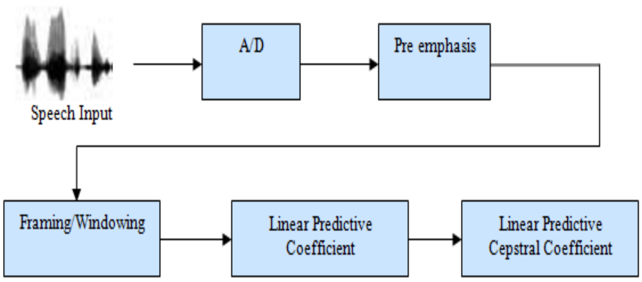


Figure 2: Block diagram of Linear Predictive parameter

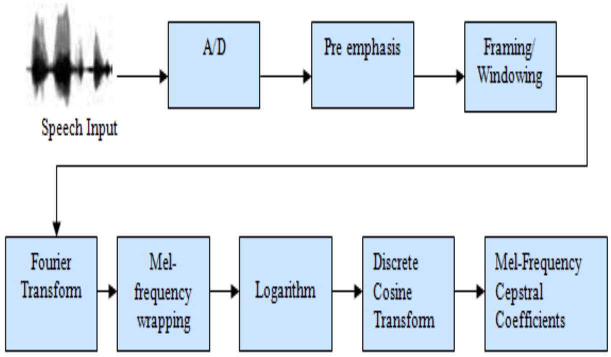


Figure 3: Block diagram of Mel-Frequency Cepstral Coefficient

C. Classification

The kNN classification system is a simple, supervised algorithm that employs lazy learning [17], [18]. The PNN classification is the most popular classifier for artificial neural networks [19], [20]. Therefore, both of these classifiers had been used to classify the BMI in this experiment. The accuracy of the kNN classifier and PNN classifier are calculated using Equation (3).

$$Accuracy = \frac{TC}{TT} x100\% \quad (3)$$

where: TC = Total number of correct classified
TT = Total number of testing samples for each categories

IV. Results and Discussions

This section provides the experimental results for classification of BMI with different feature vector sets and two classifiers. There is 270 speech sample files are used for training, and 30 speech sample files are used for testing in this experiment results.

Based on Figure 4, the hybrid features gives the highest accuracy of 97.28% with k-value equal to 5. Meanwhile, the results proved that hybrid feature vector helps to improve the accuracy than single feature vector alone.

Based on Figure 5, the hybrid features gives the highest accuracy of 97.00% with smoothing factor of (0.13, 0.14, 0.16, 0.18 and 0.19). Besides that, a smoothing factor of 0.13 was better for MFCC, LPCC, WPT-EE and hybrid features vectors. The optimum accuracy can obtain by PNN classifier was 97.00%.

Table 2
Confusion Matrix of Classification

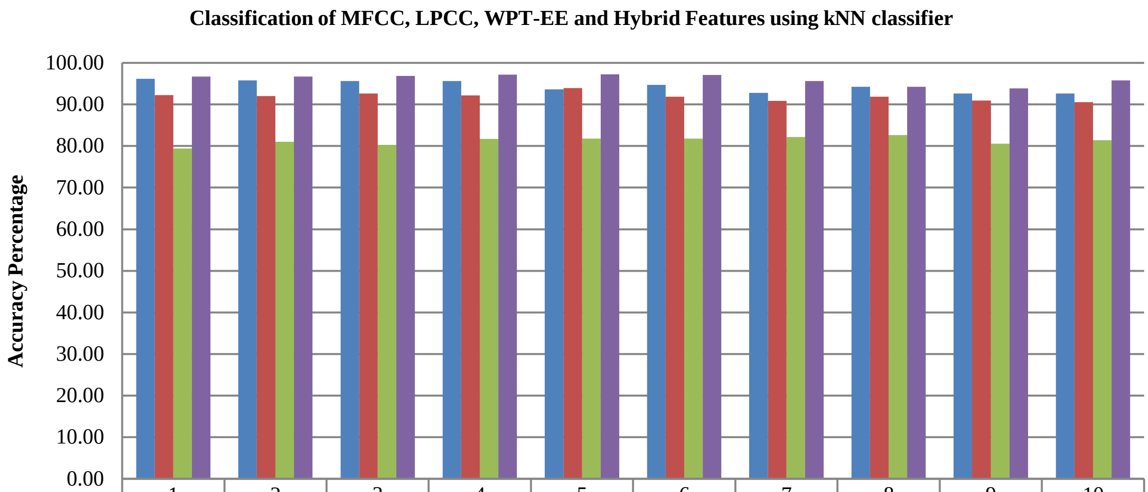
Output Class		Target Class: Normal	Overweight	Obese
	Normal	90	1	2
	Overweight	8	99	5
	Obese	2	0	93

In Table 2, the confusion matrix shows normal, overweight and obese that has 100 samples for each class. For normal, 90 samples had correctly classified, eight samples were wrongly classified as overweight, and two samples were wrongly classified as obese. For overweight, 99 samples had correctly classified, and 1 sample was wrongly classified as normal. Last, 93 samples had correctly classified as obese, two samples were wrongly classified as normal, and five samples were wrongly classified as obese.

V. Conclusion

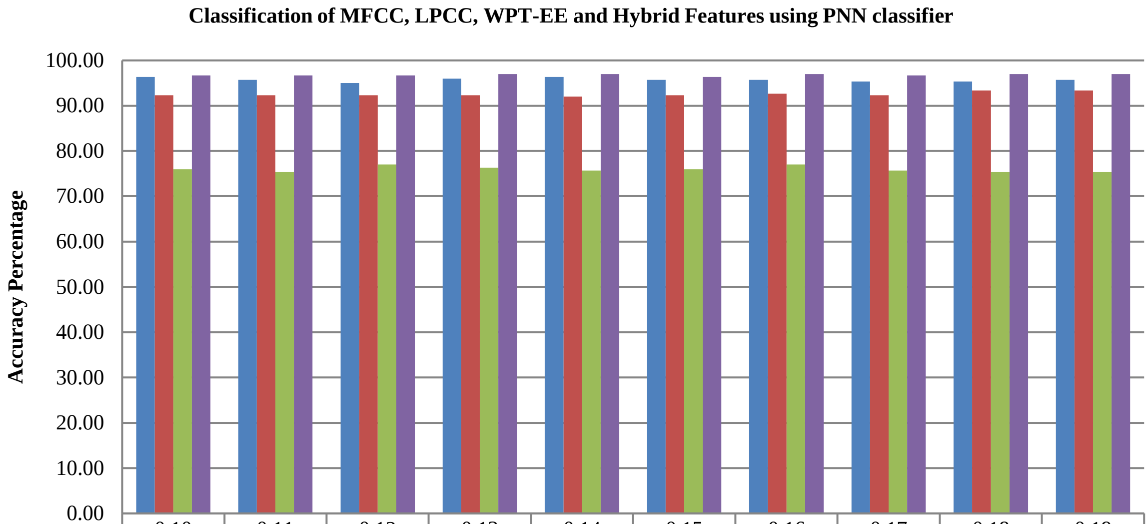
This study is conducted to propose an effective way to classify the Body Mass Index (BMI) status of an individual by using speech features. The previous research works for this area of study has proposed and suggested a few methods and approach to their study. The conventional method of measuring the BMI is already easy. However, there are certain cases that the method is insufficient and has limitations such as it can show where the body fat locate, muscular man, tends to fall into the overweight category and weight change rapidly over time, therefore, it is hard to work in the remote area. Hence, an automated classification system that will classify the different status of BMI has been developed using speech features to overcome the drawbacks of the current method thereby achieved the three main objectives of this study.

Researchers in the study have proposed several types of classifiers. Most of the researchers have tested their best features with one or several classifiers. As for this research work, k-Nearest Neighbour (kNN) and Probabilistic Neural Network (PNN) were used to test and validate the efficiency of the features extraction method used in this report. A kNN shows a better accuracy which is 97.28% when compared to PNN with 97.00% of accuracy. Difficulties faced during the study were finding suitable subjects for this research, selecting suitable features and classification method to classify the data.



	1	2	3	4	5	6	7	8	9	10
MFCC	96.17	95.78	95.61	95.61	93.67	94.72	92.78	94.22	92.61	92.61
LPCC	92.22	92.00	92.67	92.17	93.94	91.83	90.89	91.83	90.94	90.56
WPT-EE	79.39	81.00	80.28	81.72	81.83	81.78	82.17	82.67	80.56	81.44
Hybrid	96.72	96.72	96.89	97.17	97.28	97.11	95.67	94.28	93.83	95.78

Figure 4: Classification of MFCC, LPCC, WPT-EE and Hybrid Features using kNN classifier



	0.10	0.11	0.12	0.13	0.14	0.15	0.16	0.17	0.18	0.19
MFCC	96.33	95.67	95.00	96.00	96.33	95.67	95.67	95.33	95.33	95.67
LPCC	92.33	92.33	92.33	92.33	92.00	92.33	92.67	92.33	93.33	93.33
WPT-EE	76.00	75.33	77.00	76.33	75.67	76.00	77.00	75.67	75.33	75.33
Hybrid	96.67	96.67	96.67	97.00	97.00	96.33	97.00	96.67	97.00	97.00

Figure 5: Classification of MFCC, LPCC, WPT-EE and Hybrid Features using PNN classifier

REFERENCES

[1] E. Anuurad, K. Shiwaku, A. Nogi, K. Kitajima, B. Enkhmaa, K. Shimono, Y. Yamane, and K. Shiwaku, “The New BMI Criteria for Asians by the Regional Office for the Western Pacific Region of WHO are Suitable for Screening of Overweight to Prevent Metabolic Syndrome in Elder Japanese Workers,” *J Occup Heal. J. Occup. Heal.*, vol. 45, pp. 335–343, 2003.

[2] B. J. Lee, B. Ku, J.-S. Jang, and J. Y. Kim, “A novel method for classifying body mass index on the basis of speech signals for future clinical applications: a pilot study.,” *Evid. Based. Complement. Alternat. Med.*, vol. 2013, p. 150265, Mar. 2013.

[3] M. Hariharan, C. Y. Fook, R. Sindhu, A. H. Adom, and S. Yaacob, “Objective evaluation of speech dysfluencies using wavelet packet transform with sample entropy,” *Digit. Signal Process.*, vol. 23, no. 3, pp. 952–959, 2013.

[4] L. L. Yan, M. L. Daviglus, K. Liu, A. Pirzada, D. B. Garside, L. Schiffer, A. R. Dyer, and P. Greenland, “BMI and Health-Related Quality of Life in Adults 65 Years and Older,” *Obes. Res.*, vol. 12, no. 1, pp. 69–76, Jan. 2004.

[5] K. Schousboe, G. Willemsen, K. O. Kyvik, J. Mortensen, D. I. Boomsma, B. K. Cornes, C. J. Davis, C. Fagnani, J. Hjelmborg, J. Kaprio, M. de Lange, M. Luciano, N. G. Martin, N. Pedersen, K. H. Pietiläinen, A. Rissanen, S. Saarni, T. I. A. Sørensen, G. C. M. van Baal, and J. R. Harris, “Sex Differences in Heritability of BMI: A Comparative Study of Results from Twin Studies in Eight Countries,” *Twin Res.*, vol. 6, no. 5, pp. 409–421, Oct. 2003.

[6] D. Gallagher, M. Visser, D. Sepúlveda, R. N. Pierson, T. Harris, and S. B. Heymsfield, “How useful is body mass index for comparison of body fatness across age, sex, and ethnic groups?,” *Am. J. Epidemiol.*, vol. 143, no. 3, pp. 228–39, Feb. 1996.

[7] K. Pisanski, P. J. Fraccaro, C. C. Tigue, J. J. M. O’Connor, S. Röder, P. W. Andrews, B. Fink, L. M. DeBruine, B. C. Jones, and D. R. Feinberg, “Vocal indicators of body size in men and women: a meta-analysis,” *Anim. Behav.*, vol. 95, pp. 89–99, Sep. 2014.

[8] I. Mporas and T. Ganchev, “Estimation of unknown speaker’s height from speech,” *Int. J. Speech Technol.*, vol. 12, no. 4, pp. 149–160, Dec. 2009.

[9] C. Berkai, M. Hariharan, S. Yaacob, and M. I. Omar, “Estimation of BMI status via speech signals using short-term cepstral features,” in *2015 IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, 2015, pp. 195–199.

[10] K. H. Kim, B. Ku, N. Kang, Y.-S. Kim, J.-S. Jang, and J. Y. Kim, “Study of a Vocal Feature Selection Method and Vocal Properties for Discriminating Four Constitution Types,” *Evidence-Based Complement. Altern. Med.*, vol. 2012, pp. 1–10, Mar. 2012.

[11] J. González, “Research in acoustics of human speech sounds: Correlates and perception of speaker body size,” *Transw. Res. Netw. India Recent Res. Devel. Appl. Phys*, vol. 37661, no. 9, pp. 81–7895, 2006.

[12] B. J. Lee, K. H. Kim, B. Ku, J.-S. Jang, and J. Y. Kim, “Prediction of body mass index status from voice signals based on machine learning for automated medical applications,” *Artif. Intell. Med.*, vol. 58, no. 1, pp. 51–61, May 2013.

[13] C. A. Ooi, M. Hariharan, S. Yaacob, and S. C. Lim, “Classification of speech dysfluencies with MFCC and LPCC features,” *Expert Syst. Appl.*, vol. 39, no. 2, pp. 2157–2165, Feb. 2012.

[14] C. Y. Fook, M. Hariharan, S. Yaacob, and A. Adom, “Comparison of LPCC and MFCC for Isolated Malay Speech Recognition,” vol. 2012, no. i, pp. 884–888, 2012.

[15] C. Y. Fook, M. Hariharan, S. Yaacob, and A. H. Adom, “Malay speech recognition in normal and noise condition,” in *Proceedings - 2012 IEEE 8th International Colloquium on Signal Processing and Its Applications, CSPA 2012*, 2012, pp. 409–412.

[16] C. Y. Fook, M. Hariharan, S. Yaacob, and A. Adom, “A Review : Malay Speech Recognition and Audio Visual Speech Recognition,” in *International Conference on Biomedical Engineering*, 2012, no. February, pp. 27–28.

[17] K. Teknomo, “K-Nearest Neighbor (KNN) Tutorial,” 2006. [Online]. Available: http://people.revoledu.com/kardi/tutorial/KNN/HowTo_KNN.html.

[18] M. Kalamani, S. Valarmathy, and S. Anitha, “Automatic Speech Recognition using ELM and KNN Classifiers,” no. 2014, pp. 3145–3152, 2015.

[19] D. F. Specht, “Probabilistic neural networks,” *Neural Networks*, vol. 3, no. 1, pp. 109–118, 1990.

[20] W. Ford, K. Xiang, W. Land, R. Congdon, Y. L. Li, and O. Sadik, “A multi-class probabilistic neural network for pathogen classification,” *Procedia Comput. Sci.*, vol. 20, pp. 348–353, 2013.