COMPUTATIONAL CARDIO-PHYSIOLOGICAL MODEL OF EMOTION USING PHONOCARDIOGRAPHY

BY

ZELDI SURYADY

A thesis submitted in fulfilment of the requirement for the degree of Doctor of Philosophy in Computer Science

Kulliyyah of Information and Communication Technology
International Islamic University Malaysia

SEPTEMBER 2022

ABSTRACT

Several studies on physiological-based human emotion have suggested that emotion causes variations in various physiological parameters. As one of the physiological parameters, heart sound signals (also referred to as phonocardiography) may infer emotions and can possibly be used for emotion recognition. For this purpose, the use of Phonocardiography (PCG) signal is substantially cheaper, and the process of acquiring the signal for heart sound analysis is comfortable as compared to other physiological measures. Capturing heart-sound signals using PCG does not require touching the surface of the human body directly. Hence it offers a convenient and practical usage in various applications of emotion recognition.

Additionally, unlike the use of electrocardiography (ECG) that reflects only heartbeats through the electrically conductive system of the heart, the PCG can also reflect the muscle contraction sound of the heart. Nevertheless, the use of PCG in the emotion recognition domain is still scarce in the research literature. As such, this thesis explored usability and methods for modelling emotion recognition using PCG signals.

The thesis is developed with four major phases. (i) Since PCG data for emotion recognition are not widely available, the first phase performs the creation of the corpus for both PCG and EEG, hence, the performance for both modalities can be compared. (ii) The second phase investigates the most suitable method for building a computational model for PCG-based emotion recognition. Three cepstral-based features, namely, MFCC, LFCC, and GFCC, are considered in the experiment. The DNN, XGBoost, and Decision tree are selected as the classifiers. The initial experiments of this research indicate that the best model for recognizing emotion is achieved at 87% accuracy rate by using combination of MFCC feature extraction and DNN classifier, (iii) The third phase compares PCG-based emotion recognition using heart sound signal (PCG) with EEG modality. The experimental results implied that with techniques used in phase two, the PCG signal could achieve comparatively robust performance in recognizing emotion as compared to the EEG modality. (iv) In the fourth phase, a new computational approach is proposed and implemented by incorporating signal decomposition techniques such as Empirical Mode Decomposition (EMD). As the main issue with this approach is feature dimensionality, the PCA feature reduction technique is adopted in the proposed method. The proposed method demonstrated a robust and optimal performance of a PCG-based emotion recognition model, achieving overall accuracy rate at 98%.

Overall, this research has highlighted the potential use of PCG signals for emotion recognition as an alternative to other commonly discussed modalities such as EEG. Additionally, the thesis also empirically proved that with proper methods in pre-processing the signal and the right feature extraction process and the suitable classifier, the PCG signal could achieve optimal performance in recognizing emotion. As future works, the proposed approach can be used to build a wide range of practical application of emotion recognition such as Ambient Assisted Living (AAL), whereby the patient's mental state is required to be continuously monitored.

ملخص البحث

في السنوات الأخيرة، تم تطوير نظام التعرف على المشاعر البشرية الفسيولوجية باستخدام مخطط كهربائية الدماغ (EEG) باعتباره الطريقة الرئيسية. توفر الإشارة الفسيولوجية مثل بيانات EEG مزايا فريدة وموثوقة لاستنتاج المشاعر. لقد أشارت العديد من الدراسات إلى أن المشاعر تسبب اختلافات في مختلف المعايير الفسيولوجية لكونها احدى المعاملات الفسيولوجية، فإشارات صوت القلب (يشار إليها أيضًا باسم تخطيط صوت القلب) قد تستنتج مشاعرا عاطفية ويمكن استخدامها للتعرف على المشاعر. لأجل هذا يعد استخدام إشارة جهاز تخطيط أصوات القلب (PCG) أرخص عملية للحصول على إشارة لتحليل صوت القلب, أضف إلى ذالك,تكون إشارة جهاز تخطيط أصوات القلب (PCG) مريحة مقارنةً بالقياسات الفسيولوجية الأخرى, حيث أن التقاط إشارات صوت القلب باستخدام PCG لا يتطلب لمس سطح جسم الإنسان مباشرة. ومن ثم فإن إشارة تخطيط الصوت توفر استخدامًا مريحا ومناسبًا وعمليًا في تطبيق برامج مثل مساكن الرعاية الدائمة (AAL) Ambient Assisted Living، حيث يلزم مراقبة الحالة العقلية للمريض باستمرار. بالإضافة إلى ذلك، على عكس استخدام تخطيط القلب الكهربائي (ECG) الذي يعكس نبضات القلب فقط من خلال نظام التوصيل الكهربائي للقلب يمكن أن يعكس جهاز PCG أيضًا صوت تقلص عضلات القلب. ومع ذلك، فإن استخدام تخطيط أصوات القلب في مجال التعرف على المشاعر لا يزال نادرا في الأدبيات البحثية. على هذا النحو، استكشفت هذه الأطروحة قابلية الاستخدام وطرق نمذجة التعرف على المشاعر باستخدام إشارات PCG. تم تطوير الأطروحة في أربع مراحل رئيسية: (1) بسبب عدم توفر بيانات PCG للتعرف على المشاعر فى نطاق واسع، فقد قدمت هذه الأطروحة في المرحلة الأولى بإنشاء اجزاء أساسية جديدة لكل من PCG وEEG. لتتم مقارنة نتائج التعرف على المشاعر المستندة على الجهازين. (2) المرحلة الثانية هي التحقق من أنسب طريقة لبناء نموذج حسابي للتعرف على المشاعر القائمة على PCG. تم اعتبار ثلاث ميزات قائمة على تحليل سيبسترال وهي MFCC و LFCC و GFCC في التجربة. تم تحديد الشبكة العصبية العميقة DNN و XGBoost وشجرة القرار كمصنفات. اظهرت النتائج التجريبية في هذه المرحلة أن الجمع بين MFCC كاستخراج الميزات و DNN كمصنف للتعلم الآلي ساهم أكثر في تحسين الأداء عند مقارنته بالتقنيات الأخرى. (3) قامت المرحلة الثالثة بمقارنة التعرف على المشاعر المستندة إلىPCG باستخدام إشارة تخطيط أصوات القلب (PCG) مع طريقة EEG. أشارت النتائج التجريبية

إلى أنه من خلال الأساليب والتقنيات المناسبة، يمكن أن تحقق إشارة PCG أقوى أداء في التعرف على المشاعر مقارنة مع طريقة EEG. (4) في المرحلة الرابعة، تم اقتراح نهج حسابي جديد وتنفيذه من خلال دمج تقنيات تحلل الإشارة مثل تحليل الوضع التجريبي (EMD). نظرًا لأن المشكلة الرئيسية في هذا النهج هي أبعاد الميزة، فقد تم اعتماد تقنية تقليل ميزة PCA في الطريقة المقترحة. أظهرت الطريقة المقترحة أداءً قويًا ومثاليًا لنموذج التعرف على المشاعر المستند إلى PCG بتحسن قدره 98%. بشكل عام، سلط هذا البحث الضوء على الاستخدام المحتمل لإشارات جهاز تخطيط أصوات القلب (PCG) للتعرف على المشاعر كبديل للطرائق الأخرى التي نوقشت بشكل شائع مثل مخطط كهربائية الدماغ (EEG). بالإضافة إلى ذلك، أثبتت الأطروحة أيضًا بشكل تجريبي أنه من خلال الأساليب المناسبة في المعالجة المسبقة للإشارة وعملية استخراج الميزة الصحيحة والمصنف المناسب يمكن لإشارة PCG تحقيق الأداء الأمثل في التعرف على المشاعر.

APPROVAL PAGE

The thesis of Zeldi Suryady has been approved by the following:

Abdul Wahab Abdul Rahman
Supervisor

Norzaliza Binti Md.Nor
Co-Supervisor

Raini Binti Hassan
Co-Supervisor

Imad Fakhri Taha
Internal Examiner

Azlinah Mohamed
External Examiner

Ma'an Fahmi Rashid Al-Khatib
Chairman

DECLARATION

I hereby declare that this thesis is the result of my own investigations, except where otherwise stated. I also declare that it has not been previously or concurrently submitted as a whole for any other degrees at IIUM or other institutions.

Zeldi Suryady

Signature…………………………… Date 26 September 2020

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH

COMPUTATIONAL CARDIO-PHYSIOLOGICAL MODEL OF AFFECT USING PHONOCARDIOGRAPHY

Affirmed by Zeldi Suryady

……………………………… 26 September 2022

Signature Date

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH

COMPUTATIONAL CARDIO-PHYSIOLOGICAL MODEL OF AFFECT USING PHONOCARDIOGRAPHY

I declare that the copyright holder of this thesis is International Islamic University Malaysia.

By signing this form, I acknowledged that I have read and understand the IIUM Intellectual Property Right and Commercialization policy.

Affirmed by Zeldi Suryady

………………………………	26 September 2022
Signature	Date

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH

COMPUTATIONAL CARDIO-PHYSIOLOGICAL MODEL OF AFFECT USING PHONOCARDIOGRAPHY

I declare that the copyright holder of this thesis is Zeldi Suryady.

No part of this unpublished research may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the copyright holder except as provided below

1. Any material contained in or derived from this unpublished research may only be used by others in their writing with due acknowledgement.

2. IIUM or its library will have the right to make and transmit copies (print or electronic) for institutional and academic purpose.

3. The IIUM library will have the right to make, store in a retrieval system and supply copies of this unpublished research if requested by other universities and research libraries.

By signing this form, I acknowledged that I have read and understand the IIUM Intellectual Property Right and Commercialization policy.

Affirmed by Zeldi Suryady

……………………………… 26 September 2022
Signature Date

ACKNOWLEDGEMENTS

All praise to Allah the almighty, for giving me the blessing, the strength, the chance and endurance during PhD study and finally completed this Thesis. Shalawat and Salam are upon the Prophet Muhammad (S.A.W.), his family as well as his companions.

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Dr. Abdul Wahab Bin Abdul Rahman, whose sincerity and encouragement I will never forget. Prof. Dr. Abdul Wahab has been an inspiration as I hurdled through the path of this PhD degree. This thesis would not have been possible without his supervision, whose guidance from the initial step in research enabled me to develop an understanding of the subject. I am thankful for the extraordinary experiences he arranged for me and for providing opportunities for me to grow professionally. I would like to also express my gratitude to my co-supervisors Dr. Zaliza Md Noor and Dr. Raini Hassan.

I owe a special gratitude to my wife, Dr. Dini Handayani and our two lovely children, Raushan and Nuha. With them by my side, I have been constantly presented with an abundance of love, motivation and endurance to complete my PhD study in the best time possible. So this research work is dedicated for them.

Last but not least. I want to thank to my late parent, my beloved sisters for their prayers and support. This thesis is also dedicated to my nephews and nieces, I hope this PhD journey can become a little inspiration for them to study harder and achieve anything that they can dream of in life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND OF THE STUDY

The heart is one of the central organs for human beings, which has a superb mechanism. It is the central organ of the human circulatory system acting as a pump in relation to the blood vessels. Excepting the lungs, it is the only organ in the body through which the blood passes in every cycle. The activities of the heart may explain many aspects of human life. One mechanism to understand the heart activities is through the sounds produced during the heart's cardiac cycle.

In fact, the sound produced by the acoustic activity of the heart is one of the most important bio-physiological signals (bio-signal) in the human body. It can be used to measure or even analyse important bio variables of the heart (e.g., state of cardiovascular, vessels, atrial and ventricular heart, etc.). Heart auscultation – the process of listening and interpreting the acoustic wave of the heart – is known as a fundamental component of the physical assessment of the heart. However, with the advancement of Electrocardiographic or Echocardiogram (ECG), the use of heart auscultation is somewhat eclipsed in the research literature. In fact, the research explorations on heart sound bio-signals are considered lesser compared to other heart-related bio signals such as ECG.

As in other bio-physiological signal modalities, such as brainwave (Electroencephalography) and galvanic skin response (GSR), the heart sound signals (also referred as PCG – Phonocardiography) may as well infer emotions and can be used in the *Affective Computing* domain for recognizing human emotion (McCraty, 2016; Soosalu et al., 2019). For this purpose, the use of PCG signal is substantially cheaper, and the process of acquiring the signal for heart sound analysis is comfortable as compared to other physiological measures. Capturing heart-sound signals using PCG devices does not require touching the surface of the body directly. Hence it offers a convenient and practical usage in an application such as Ambient Assisted Living (AAL), whereby the patient's mental state is required to be continuously monitored. Additionally, unlike the use of electrocardiography (ECG) that reflects only heartbeats through the electrically conductive system of the heart (Chronotropy), the PCG can also reflect the muscle contraction sound of the heart (Inostropy) (Giordano & Knaflitz, 2019).

However, using the PCG modality in the Affective Computing domain to recognize human emotion is still scarce in the research literature. The current emotion recognition research using heart-related modality is merely based on the electrical signals of the heart captured using an invasive ECG device. It is known that the heart analysis based on ECG is generally cumbersome due to the many complexities of setting up electrodes over the body. Unlike ECG, the heart-sound modality offers a convenient and practical usage to be used in the affective computing domain. Therefore, this thesis investigates and develops a computational model for emotion recognition (affect recognition) using the PCG signals.

1.2 RESEARCH MOTIVATION

This research is motivated by and proceeds from a central interest in the importance of evaluating heart sound in recognizing individual affect. The aim is to contribute to the *Affective Computing* field by addressing heart sound as one of the modalities in identifying human emotion.

The ability of a machine or computer to recognize human emotions falls into a field of research called *Affective Computing*. Apparently, Affective computing is an emerging field in technologies, The Gartner Hype Cycle (Figure 1.1) tells us that affective computing is an area that is experiencing high levels of innovation. As can be seen from Gartner's hype cycle, affective computing is expected to reach its "*plateau of productivity*" (when mainstream adoption of technology starts to take off) in the next 5-10 years. Therefore, this practical fact triggers great motivation in contributing to this area of research.

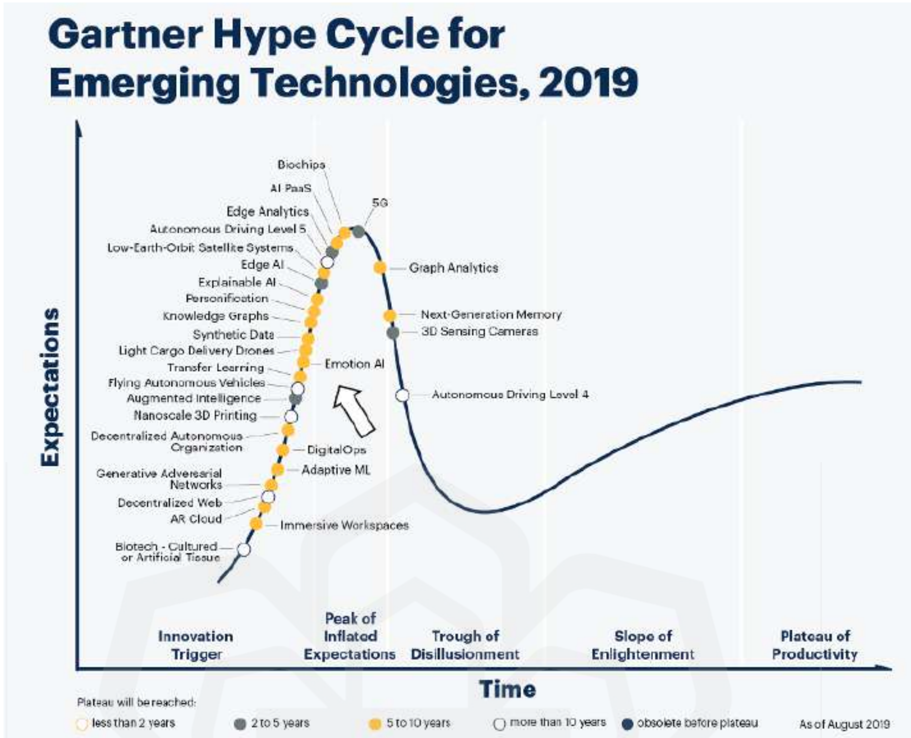

Figure 0.1 Gartner Hype Cycle (Gartner, 2019)

This research is also motivated by Islamic tradition on the heart. In the Quran – the primary source of Islamic doctrine, it has been implicitly shown that individual affect or emotion can be identified and understood from the state of the heart. The heart is referred in the Quran as the organ of emotions, feelings, and desires. The word *al-fuad* (فؤَاد) is often used in the Quran to describe the condition of the heart with a sense of emotion. Additionally, it is reported in a well-known Hadith (sayings of the Prophet Muhammad) that the state of the heart also indicates the general state of an individual:

"*There is a piece of flesh in the body if it becomes good, the whole body becomes good but if it gets spoilt the whole body gets spoilt and that is the heart*".

All the religious dogmas mentioned above are another source of motivation to proceed with this proposed research domain. In practice, if the heart sound is properly analysed, it can provide information and open various vertical applications in the affective computing domain and in other vertical domains.

However, technically, to use PCG signal (heart sound) as a modality to recognize emotion is relatively challenging due to the complex and highly non-stationary nature of the heart sound signals (Giordano & Knaflitz, 2019; Subasi, 2019). Thanks to the advances in signal processing and artificial intelligence, the technologies have provided various techniques to cope with this complexity. Various contributions have been made in the literature regarding the techniques to process and extract necessary information from the signals and classification method of non-stationary signals such as heart sound signals. Thus, the initial aim is to develop a PCG-based computational model for individual emotion recognition using the most efficient techniques. In this thesis, emotion recognition and classification are made by extracting features from the heart sound signals. Well established cepstral based feature engineering techniques, namely MFCC, LFCC and GFCC, are considered. Deep Neural Network (DNN), Extreme Gradient Boosting (XGBoost), and Decision Tree are the pattern classification techniques engaged for emotion classification. Furthermore, a new computational model based on the EMD technique is proposed to improve the computational performance in recognizing individual emotion using PCG signals.

1.3 PROBLEM STATEMENTS

In the context of automatic affect recognition, numerous works of literature have argued that the heart is more than an efficient pump that sustains life. It is also the source of emotion, courage, and wisdom (McCraty, 2016; Soosalu et al., 2019). As the heart sound is commonly used to measure the activities of the heart, it may be an essential bio-signal to be used and analyzed to understand the human affective state, such as emotion, mood or feeling. However, as shown in the literature review the effort in conducting experimental works, and analysing heart sound to recognize human emotions remain scarce.

Moreover, in the past years, there has been increasing evidence about the brain-heart interaction with significant potential implications for the treatment of cardiovascular diseases. Hence, It can be assumed that there is also a close relationship between the heart and the brain regarding human emotional states. However, as the effort to use heart sound for emotion recognition is still scarce, the effort to analyse heart sound and its relationship with the brain signal in recognizing human emotions remains uncommon.

In light of these facts, there are knowledge gaps that can be explored concerning heart sounds in the area of affective computing. As such, the following problem statements are the foundation and the rationale for the significance of this research:

- Heart sounds analysis for emotion recognition is generally less understood than other bio-physiological modalities, especially the brain-wave signal.
- There have been less attempts to infer emotion automatically from heart sounds to the best of my knowledge. In this scenario, most heart sounds analysis is dominated by clinical heart diagnosis and biometric system.
- Moreover, there have been no previous attempts to correlate and compare heart sounds and brain waves to recognize individual emotions.

1.4 ISSUES IN USING PCG SIGNAL FOR EMOTION RECOGNITION

As the use of PCG signal is essential in this research, below are several challenging tasks to be considered in realizing a computational system for emotion recognition using heart sound modality:

1. **Datasets:** As the research PCG-based emotion recognition is limited, there is a lack of publicly available PCG signal datasets of human emotions. Therefore, the research on PCG-based emotion recognition requires the development of new PCG datasets for emotion recognition. Furthermore, to validate its results, the development of PCG-based emotion recognition may require the comparison with other physiological-based emotion recognitions (e.g., EEG)
2. **Pre-processing**: The PCG signal collected from the participants during data collections may contains numerous noises like displacement of the digital stethoscope, participant’s movements etc. These kinds of noises must be removed using filters for efficient processing.

3. **Feature Extraction**: Patterns that are extracted from PCG signals are denoted by the features. The recognition stage or classification are mainly concerned with the retrieval of PCG characteristics. Feature extraction is defined as the process of converting a PCG signal into a sequence of feature vectors carrying characteristic information about the signal, which are used as the basis for various types of PCG analysis algorithms. The difficulty arises in finding the suitable feature extraction techniques for PCG signals. In the proposed work for this thesis, cepstral-based features namely MFCC, LFCC and GFCC are extracted from the PCG signals.
4. **Modelling**: One of the challenges for modelling emotion recognition is to find for a suitable classifier. In this study, features extracted from the PCG signal are modelled using Deep Neural Network (DNN), Extreme Gradient Boosting (XGBoost) and Decision Tree classifiers.

1.5 RESEARCH QUESTIONS

The research focuses primarily on analysing if and how the heart sound modality (phonocardiography) can be used to recognize human emotion from an individual. As PCG signals for emotion recognition are not common, this research will further compare its performance with the EGG modality. Overall, the study intends to answer the following questions:

RQ1: What is the effectiveness of using Phonocardiography Signal (PCG) for individual emotion recognition?

RQ2: What aspects of emotion recognition modelling constitute an effective emotion recognition system using PCG signals?

RQ3: How reliable is the Phonocardiography Signal (PCG) compared to the EEG modality in developing an emotion recognition model?

RQ4: What is the potential method to achieve optimal performance of a phonocardiography-based emotion recognition system?