UNIVERSITI TEKNOLOGI MARA

FORENSIC LANGUAGE OF PROPERTY THEFT GENRE BASED ON MATHEMATICAL FORMULAE AND MACHINE LEARNING ALGORITHMS

HANA' BINTI ABDRAZAK

PhD

May 2020

AUTHOR'S DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Teknologi MARA. It is original and is the results of my own work, unless otherwise indicated or acknowledged as referenced work. This thesis has not been submitted to any other academic institution or non-academic institution for any degree or qualification.

I, hereby, acknowledge that I have been supplied with the Academic Rules and Regulations for Post Graduate, Universiti Teknologi MARA, regulating the conduct of my study and research.

Name of Student	:	Hana' Binti Abd Razak
Student I.D. No.	:	2011324131
Programme	:	Doctor of Philosophy (Electrical Engineering) – EE990
Faculty	:	Electrical Engineering
Thesis Title	:	Forensic Language of Property Theft Genre Based On Mathematical Formulae and Machine Learning Algorithms
Signature of Student	:	……..
Date	:	May 2020

ABSTRACT

One of the indicators that resemble rapid development of a country is the crime rate. Hence in this thesis, the forensic language focusing on property theft genre is investigated and explored. Forensic language can be categorised based on three properties namely forensic postures, frequency of movement and execution time. Thus, forensic language refers to the combination of any four forensic gait postures that will be performed sequentially or repeatedly in any order in front of residential unit. Firstly, a set of mathematical formulae is developed according to the consented forensic postures to verify the anthropometric traits of the postures. Four traits was discovered namely spine, pelvis, knee and ankle. Next, data collection of normal and anomalous behaviour is collected using Kinect sensor. Here, a camera is mounted on a specific location and height as CCTV at the residential unit gate in detecting anomalous behaviour based on the subjects' behaviour. Normal behaviours are varied to activities associated at the gate and housebreaking crime is selected as anomalous behaviour at the gate. Next, the forty joints with 1400 joint features acquired from X- and Y-plane via Kinect are further extracted using Orthogonal Least Squares (OLS) yielded twenty four significant joints and 840 gait features. Further, Principal Component Analysis (PCA) is used as feature optimisation in selecting the most significant features from the 840 joint features and resulted in 114 significant joint features. In addition Artificial Neural Network (ANN) and Support Vector Machine (SVM) are used as classifiers. Results showed the specificity and sensitivity of SVM classifier is approximately within 50% for all tested kernels. As for ANN, four different setting of learning function are used. Contrary to SVM classifier, ANN classifier has achieved high specificity and sensitivity up to 98.21 and 97.88% in classifying normal and anomalous behaviour, respectively. At this stage a set of anthropometric traits of forensic postures is computationally generated by OLS, PCA and ANN viz. head, shoulder, spine, elbow, hip, hand, wrist, foot, ankle and knee. Convolution Neural Network (CNN) using deep learning algorithm is chosen in identifying frequency of movement and execution time of housebreaking crime. Here, 19116 colour pixel-based images that comprised of both normal and anomalous behaviours acted as inputs to four samplings of CNN. The highest sensitivity achieved by Up-Down sampling is 98.21% with specificity of 97.37%. Few samplings are relatively good in offline test, however only Up Sampling produced the best accuracy in real-time test with recognition rate of 95%. Towards achieving better detection in real-time environment, colour pixel-based images were trained on five pre-trained CNNs using transfer learning algorithm. Remodelled ResNet-50 was demonstrated the highest performance with both specificity and sensitivity of 98.44% during classification. Remodelled Inception -v3 is the best network for both offline and real-time mode tests with recognition rate up to 99% in determining the frequency and duration of normal and anomalous movement at the gate. Thus, the proposed forensic language of property theft was proven its authenticity and acquires high ability in detecting anomalous behaviour at the gate of residential unit. The proposed method of anomalous behaviour detection in this thesis could be further applied for detection of potential crime at various locations.

ACKNOWLEDGEMENT

Firstly, Thanks to Allah S.W.T, the Most Sovereign, the Ever Wise, Who gave me courage, strength, knowledge and wisdom in order to attain our goal and making all things possible.

First and foremost, I wish to express my gratitude to my one and only supervisor, Prof. Dr. Nooritawati Md Tahir for her encouragement, support, inspiring discussions, and guidance throughout my study period.

I deeply appreciate the help from the staff members in Faculty of Electrical Engineering, Universiti Teknologi MARA who have assisted and managed schedule of my lengthy study. Special thanks go to Dr. Rohilah Sahak, Dr. Ahmad Ihsan Mohd Yassin, Dr. Azlee Zabidi, and Dr. Ali Abd Almisreb for their helpful suggestions and interesting discussions regarding machine learning algorithms.

Thanks to Dean of Electrical Engineering for providing ample research facilities, Premier Laboratory of Human Motion Gait Analysis, Institute of Research Management and Innovation Unit (IRMI), Universiti Teknologi MARA (UiTM) Shah Alam, Selangor.

Financial support of MyBrain15, MyPhD awarded by Ministry of Higher Education (MOHE) is greatly acknowledged. Equally appreciated for the grant awarded by Research Management Centre (RMC), Universiti Teknologi MARA (UiTM) Grant No: 600-IRMI/MyRA 5/3/BESTARI (041/2017).

I reserve my sincere thanks for my family members. I am deeply indebted to them for their dedication and time through my extraordinary journey of study. Their support and love has strengthened me spiritually throughout this study.

TABLE OF CONTENTS