UNIVERSITI TEKNOLOGI MARA

SALIVARY BASED SERS ANALYSIA IN RECOGNITION OF NS1 FOR PCA-SVM CLASSIFICATION OF DENGUE FEVER

AFAF ROZAN MOHD RADZOL

Thesis submitted in fulfilment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Electrical Engineering

May 2018

ABSTRACT

NS1 is one of the seven non-structural proteins encoded in the RNA of flavivirus genome. It has been recognized as an early biomarker for flavivirus infection diseases. In the case of DF, NS1 is detectable in patients' blood serum from day 1 to 9 before the formation of IgM and IgG antibodies. ELISA is the most common technique used to detect NS1 from patients' blood serum. However, this technique suffers from limitations such as being invasive, delayed diagnosis, tedious sample preparation, in need of seasoned laboratory, technician and bulky expensive laboratory equipment. The presence of NS1 in blood can also be detected using immune-chromatic strip. While portable, this technique suffers from disadvantages such as non-type-specific, un-validated diagnostic accuracy, non-cost effective in addition to limitations mentioned before. Recently, the presence of NS1 in saliva via ELISA has been reported, but with low sensitivity (64.7%). SERS is a form of Raman spectroscopy which can provide fingerprint spectra, unique of each molecule at a higher signal intensity. It has been applied to detect a variety of diseases. Yet, its being used to identify NS1 molecule in saliva of DF patients remains unexplored. From SERS analysis, NS1 protein has been proven to be Raman active. It is found to produce a molecular fingerprint with a distinct characteristic peak at $1000cm^{-1}$. This peak is then chosen as the marker for automated classification algorithm for detecting NS1 in saliva. This study intends (i) to establish Raman fingerprint of NS1; (ii) to design algorithms for optimal classification of positive and negative DF subjects from salivary Raman spectra; (iii) to benchmark performance of optimised classifiers against two recommended serological diagnostic tests by WHO, NS1-ELISA and NS1-Rapid. Saliva samples from healthy and suspected dengue patients are collected and analysed using SERS technique to obtain the salivary Raman spectra. The spectra are pre-processed to remove unwanted features using signal processing algorithms customized and optimized for this study. Then, the clean spectra are analysed using PCA for feature extraction and dimension reduction. Finally, the extracted principal components are classified into dengue positive and negative using SVM algorithm. NS1 ELISA and NS1 Rapid serum tests result are used as benchmark against sensitivity, specificity and accuracy performance of the SVM algorithms in which three types of kernels i.e; Linear, RBF and MLP are optimized and compared. From the results, it is observed that RBF kernel gives the best performance. The highest performance of SVM-RBF classifier against NS1-ELISA benchmark is [83.22%, 88.27%, 78.13%] using 95 principal components proposed by CPV criterion, while the highest performance against NS1-Rapid benchmark is [81.90%, 80.32%, 83.49%] using 110 principal components proposed by Kaiser criterion. Both performances achieved is found better than detection of NS1 in saliva using ELISA technique by researchers for acute cases with performance of [NA, 64.7%, 95.8%]. The finding supports that SERS technique integrated with signal processing techniques is sensitive to detect the presence of NS1 in saliva.

ACKNOWLEDGEMENT

All praises and thank be to Allah, the most Merciful and most Gracious, Lord of the universe. To whom we belong and to Him we shall return. Peace be upon him, Prophet Muhammad S.A.W, his family, his companion and those who follow his path. Thanks to the almighty Allah for this opportunity, strength and ability to complete this work.

My sincere appreciation to my supervisor, Associate Prof. Dr. Lee Yoot Khuan for the patient guidance, encouragement and advice she has provided throughout my time as her student. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries at the best she can. I would also like to express my thanks to my co-supervisors, Associate Prof. Dr. Datin Wahidah Mansor for the ideas and suggestion to improve my work and thesis in many ways.

I also wish to thank Dr. Wong Peng Shyan from Hospital Pulau Pinang, Dr Ireene Looi from Hospital Seberang Jaya and Dr B. Venugopalan from Jabatan Kesihatan Negeri Selangor for the willingness, cooperation and collaboration.

My deepest gratitude to my parents, my husband, my children and family for their never-ending supports and understanding.

Last but not least, thanks to all who have directly and indirectly contributed to the completion of this thesis. May Allah repay your kindness.

TABLE OF CONTENT

CHAPTER ONE
INTRODUCTION

1.1 Research Background

Non-structural protein 1(NS1) is an antigen used as early biomarker for detection of Flavivirus infection. Dengue Fever (DF), Japanese Encephalitis (JE), Murray Valley Encephalitis (MVE), Tick borne Encephalitis (TBE), West Nile Encephalitis (WNE) and Yellow Fever (YF) are diseases with fatal consequences related with Flavivirus. NS1 is one of the non-structural protein encoded on the single strand of ribonucleic acid (RNA) virus genome [1]. Upon infected mosquito bite, dengue virus particle is carried into the host cell by endosome. The virus RNA strand is released from the endosome into the host cytoplasm where it is synthesized to create building blocks for new viruses. NS1 protein is synthesized inside the host cell and detectable in patient's blood serum from day 1 to 9 following the onset of disease where the amount is estimated to be in the range of 0.01 to 50ppm [2-4]. NS1 protein acts as antigen to induce the production of antibodies such as Immunoglobulin M (IgM), Immunoglobulin G (IgG) and Immunoglobulin A (IgA).

DF is an outbreak and a major national health concern in Malaysia. Since 2014, there are more than 100,000 reported dengue infection cases with more than 200 fatalities, yearly [5]-[10]. Haematological (platelet and haematocrit values) test is commonly used as warning indicators in screening DF. Beside haematological test, serology tests such as IgM-enzyme-linked immunosorbent assay (ELISA), IgG ELISA, IgM-IgG ratio, IgA, haemagglutination (HI) and NS1-ELISA detect the presence of antibodies and/or antigens in the blood serum of the suspected subjects. The word serology indicates it involves the study of blood serum with respect to its reaction and properties [11]. Antigen (NS1) detection is superior as an early detection method.

Surface Enhanced Raman Spectroscopy (SERS) is Raman Spectroscopic technique with the ability to enhance the weak signal of Raman scattering. By adsorbing the analyte molecules on noble surface such as gold, silver and copper, the intensity of the Raman scattering is observed to have increased by an average of 10^7 to 10^{10}enhancement [12]-[13]. This has empowered Raman spectroscopy to be competent to detect up to a single molecule [14]-[15]. Among others, the advantages of SERS