SUPPORT VECTOR MACHINE HARDWARE ACCELERATOR FOR TONGUE COLOUR DIAGNOSIS

AMANDA THIAH SU LIN

A project report submitted in partial fulfilment of the
requirements for the award of the degree of
Master of Engineering (Computer and Microelectronic Systems)

School of Electrical Engineering
Faculty of Engineering
Universiti Teknologi Malaysia

JANUARY 2020

DEDICATION

This project report is dedicated to my parents, who taught me never to give up and always strive to be the best. It is also dedicated to my siblings, who provided me with moral support throughout the entire length of the project.

ACKNOWLEDGEMENT

I would like to take this opportunity to thank everyone who have helped me throughout the entire duration of this project. First of all, I would like to express my sincere gratitude to my project supervisor, Dr. Mohd Shahrizal for his guidance, wisdom and brilliant ideas which he shared during the course of this project. I would also like to thank him for spending their valuable time coming all the way out to Penang to have discussion with me regarding the project and provided me with invaluable advice. Also, I would like to express my gratitude to Dr. Nur Diyana Kamaruddin for providing guidance that had significantly helped my work in the early stages of my project. Their insight have been the main reason I am able to complete my project.

Secondly, I would like to thank my panel of examiners Dr. Usman Ullah Sheikh and Dr Ismahani Ismail for providing constructive feedback during the viva session.

Last but no least, I would like to thank my parents, family members and friends for their encouragement and moral support thorughout the duration of this project.

ABSTRACT

Tongue body features such as colour is used in Traditional Chinese Medicine (TMC) practices to diagnose a patient's state of health. However, the diagnosis of one patient's health condition varies between practitioners. This gives rise to the need of a standard method such as using Support Vector Machine to identify the tongue body colour. SVM is a supervised machine learning algorithm that aims to explore the characteristics and correlations between the gathered data to deduce the most efficient way in classifying the data into groups. Typically, SVM classifiers are implemented in software where the classification performance is very dependent on the architecture of the general-purpose CPU. Since classification of tongue images is a recurring event, the design of a hardware accelerator is explored in this project. The purpose of design a hardware accelerator is to boost the classifier performance, execution time and latency so that it meets real-time constraints. Architectural optimization methods, such as loop unrolling, memory array partitioning, pipelining and adder tree implementation of the SVM classification algorithm are performed in Xilinx's Vivado HLS and later synthesized to target for FPGA implementation. To further optimize the resource utilization, 18-bits IEEE-754 floating-point representations for the floating point units are used. The SVM hardware is able to demonstrate 140x speed up with similar classification accuracy when compared to the software implementation in MATLAB.

ABSTRAK

Ciri-ciri tubuh lidah seperti warna digunakan dalam amalan Perubatan Tradisional Cina (TMC) untuk mendiagnosis keadaan kesihatan pesakit. Bagaimanapun, diagnosis keadaan kesihatan seorang pesakit berbeza antara pengamal. Ini menimbulkan keperluan kaedah standard seperti menggunakan Mesin Vektor Sokongan untuk mengenal pasti warna badan lidah. SVM adalah algoritma pembelajaran mesin yang diselia yang bertujuan untuk meneroka ciri-ciri dan korelasi antara data yang dikumpul untuk menyimpulkan cara yang paling berkesan dalam mengklasifikasikan data ke dalam kumpulan. Biasanya, pengelas SVM dilaksanakan dalam perisian di mana prestasi klasifikasi sangat bergantung pada seni bina CPU tujuan umum. Oleh kerana pengkelasan imej lidah adalah peristiwa yang berulang, reka bentuk pepatah perkakasan diterokai dalam projek ini. Tujuan merancang pemecut perkakasan adalah untuk meningkatkan prestasi pengelas, masa pelaksanaan dan kependaman supaya ia memenuhi kekangan masa nyata. Kaedah pengoptimuman seni bina, seperti gelung pembongkaran, pembahagian array memori, pipelining dan penambah pohon, algoritma klasifikasi SVM dilakukan di Xilinx's Vivado HLS dan kemudian disintesis untuk menargetkan pelaksanaan FPGA. Untuk mengoptimumkan penggunaan sumber, 18-bit perwakilan titik terapung IEEE-754 untuk unit terapung digunakan. Perkakasan SVM mampu menunjukkan kelajuan 140x dengan ketepatan klasifikasi yang sama apabila dibandingkan dengan pelaksanaan perisian di MATLAB.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CSD	-	Canonic Signed Digit
CPU	-	Central Processing Unit
CSE	-	Common Subexpression Elimination
FPGA	-	Field Programmable Gate Array
FSM	-	Finite State Machine
GPP	-	General Purpose Processor
GPU	-	Graphics Processing Unit
HDL	-	Hardware Description Language
HLS	-	High Level Simulation
HPC	-	High Performance Computing
LUT	-	Look Up Table
RBF	-	Gaussian Radial Basis Function
SMO	-	Sequential Minimal Optimization
SVM	-	Support Vector Machine
TIAS	-	Tongue Image Analysis System
TMC	-	Traditional Chinese Medicine

LIST OF APPENDICES

APPENDIX	TITLE	PAGE

CHAPTER 1

INTRODUCTION

1.1 Background of the Problem

The diagnosis of Traditional Chinese Medicine (TCM) is based on the information obtained in four diagnostic processes, namely, looking, listening, smelling, asking and touching. The most common task is to measure the pulse and examine the tongue. Usually, oriental medicine practitioners' uses features extracted from tongue tests, for example the colour of the tongue, to diagnose a patient's health [1]. Therefore, accurate recognition of tongue colour is crucial for predicting the patient's condition, as it provides useful information about blood coagulation, water imbalance and psychological problems [2].

There are three main groups of tongue colours: light red, red and dark red. The tongue body colour range is very narrow, as it compromises of different types of the identical colour. Colours are hard to distinguish with the naked eye, and there are many overlapping and similar pixels [2]. Traditionally, doctors diagnose diseases and recognize patterns by examining, describing, and experiencing them. As a result, the results are easily influenced by the doctor's expertise and the surrounding environment. Without objective evaluation criteria, the accuracy of pattern recognition and the repeat ability of verification are not clear [3]. Thus, we need a classifier to help effectively determine the colour clusters on the tongue, so as to make a correct and accurate diagnosis. In order to help daily clinical practice, the implementation of machine learning algorithm was introduced to develop an automatic digital tongue diagnosis system. Support vector machine (SVM) is a supervised machine learning technology. Due to the efficiency and accuracy of this algorithm, it has been widely used in many embedded detection systems, especially in image classification systems [4].

Precision and efficiency, however, is accompanied by a costly and time-consuming computation cost, this increases the demand for hardware acceleration.

Although the software implementation of SVM has high precision, it cannot effectively meet the constraints of real-time embedded system. In such, dedicated hardware must be implemented to meet resource utilization and computing time limitations. The tongue colour diagnosis method of support vector machine (SVM) is studied to characterize and identify the features of modules that affect the system performance, especially in term of computational time. Subsequently, a hardware accelerator will be designed using hardware description language (HDL) to improve performance. The existing and hybrid implementation of the systems are then compared in terms of performance, accuracy and required resources.

1.2 Problem Statement

While software implementation of SVM gives high accuracy, they cannot efficiently meet real time embedded systems' constraints, resource utilization and computation time [4]. This usually happens when the classification problem is huge and the number of support vectors for the solution increases, thus resulting it to be computationally demanding [5].

Many solutions have been proposed to overcome this aspect in software implementation of SVM. One of the solutions is using parallel processing. Indeed, systems with parallel processing processors can be very useful in accelerating such computations, but then the computation then becomes very high in cost [6].

Another problem faced when implementing SVM in software is that CPU implementation will be restricted by the instruction set and the architecture of the system. The CPU can sometimes lack the necessary functional logics to efficiently complete the task [7]. Lastly, running in software tools also mean that additional memory has to be allocated to store the programme code.

With all the problems stated, in order to save resource utilization while maintaining the accuracy of the SVM classifier, CPU implementations can be omitted. Platforms like GPU and FPGA acceleration can be studied to find a method of implementation that fits the requirements of a given application.

1.3 Objectives of Study

This section discusses the objectives set for this project. The objectives are listed below.

i. To study and identify the compute intensive segments of the baseline Support Vector Machine (SVM) implementation during the classification phase.

ii. To explore hardware optimization techniques and design hardware accelerator for the compute intensive segment of the classification algorithm using Hardware Description Language (HDL).

iii. To perform comparative analysis between the baseline and the hardware implementation of the SVM classification phase in terms of performance and accuracy achieved.

1.4 Scope of Study

This section discusses the scope of study set throughout the project. The scope of study are:

i. Implementation of the existing baseline binary linear SVM training and classification phase for tongue body colour in MATLAB. Red and pink tongue body colour are grouped as the same class, while deep red the other.

ii. Identify the compute intensive segment of the linear SVM classification algorithm and explore different hardware optimization techniques using Vivado HLS.

iii. Compare the results obtained from different optimization techniques in terms of maximum achievable frequency, latency in clock cycles, total execution time and hardware cost.

iv. Develop the hardware accelerator for SVM classification phase and analyse the performance in terms of execution time, total latency in terms of clock cycles and accuracy between the hardware and software implementation.

REFERENCES

1. Qi, Z., Tu, L.-P., Chen, J.-B., Hu, X.-J., Xu, J.-T. and Zhang, Z.-F. The Classification of Tongue Colors with Standardized Acquisition and ICC Profile Correction in Traditional Chinese Medicine. *BioMed Research International*, 2016. 2016: 1â9. doi:10.1155/2016/3510807.

2. Kawanabe, T., Kamarudin, N. D., Ooi, C. Y., Kobayashi, F., Mi, X., Sekine, M., Wakasugi, A., Odaguchi, H. and Hanawa, T. Quantification of tongue colour using machine learning in Kampo medicine. *European Journal of Integrative Medicine*, 2016. 8(6): 932â941. doi:10.1016/j.eujim.2016.04.002.

3. Zhang, J., Xu, J., Hu, X., Chen, Q., Tu, L., Huang, J. and Cui, J. Diagnostic Method of Diabetes Based on Support Vector Machine and Tongue Images. *BioMed Research International*, 2017. 2017: 1â9. doi:10.1155/2017/7961494.

4. Afifi, S., Gholamhosseini, H. and Sinha, R. Hardware Implementations of SVM on FPGA: AState-of-the-Art Review of Current Practice. *International Journal of Innovative Science, Engineering Technology (IJISET)*, 2015. 2: 733–752.

5. Kyrkou, C., Theocharides, T., Bouganis, C.-S. and Polycarpou, M. Boosting the Hardware-Efficiency of Cascade Support Vector Machines for Embedded Classification Applications. *International Journal of Parallel Programming*, 2017. 46(6): 1220â1246. doi:10.1007/s10766-017-0514-1.

6. Mandal, B., Sarma, M. P. and Sarma, K. K. Design of a systolic array based multiplierless support vector machine classifier. *2014 International Conference on Signal Processing and Integrated Networks (SPIN)*, 2014. doi:10.1109/spin.2014.6776917.

7. Wang, J.-F., Peng, J.-S., Wang, J.-C., Lin, P.-C. and Kuan, T.-W. Hardware/software co-design for fast-trainable speaker identification system based on SMO. *2011 IEEE International Conference on Systems, Man, and Cybernetics*, 2011. doi:10.1109/icsmc.2011.6083903.

8. Zhang, D., Zhang, H. and Zhang, B. Statistical Analysis of Tongue Color and Its Applications in Diagnosis. *Tongue Image Analysis*, 2017: 225â250. doi:10.1007/978-981-10-2167-1_12.

9. Gao, Z., Po, L., Jiang, W., Zhao, X. and Dong, H. A Novel Computerized Method Based on Support Vector Machine for Tongue Diagnosis. *2007 Third International IEEE Conference on Signal-Image Technologies and Internet-Based System*, 2007. doi:10.1109/sitis.2007.115.

10. Chiu, C.-C. A novel approach based on computerized image analysis for traditional Chinese medical diagnosis of the tongue. *Computer Methods and Programs in Biomedicine*, 2000. 61(2): 77â89. doi:10.1016/s0169-2607(99)00031-0.

11. OJI, T., Namiki, T., MITANI, K., Ueda, K., Nakaguchi, T., KAINUMA, M., Shibahara, N., MITUMA, T., ODAGUCHI, H., WATANABE, K., FUJII, Y., KITA, T., Kogure, T., Ogawa-Ochiai, K., TAHARA, E., OGIHARA, K., Yakubo, S., MINAMIZAWA, K., Muramatsu, S.-I. and HANAWA, T. Japanese Literature Survey of Tongue Findings for the Purpose of Creating a Unified Multicenter Description of Clinical Tongue Diagnoses. *Kampo Medicine*, 2014. 65: 224–230. doi:10.3937/kampomed.65.224.

12. HANAWA, T., ODAGUCHI, H., WAKASUGI, A., ITO, G., OIKAWA, T., SUZUKI, K., HAYASAKI, T., SAITO, E. and GONO, Y. Standardization of Kampo Medical Findings in a Specialized Kampo Medical Service Facility . *Kampo Medicine*, 2013. 64(6): 344–351. doi:10.3937/kampomed.64.344.

13. Lo, L.-C., Chen, Y.-F., Chen, W.-J., Cheng, T.-L. and Y. Chiang, J. The Study on the Agreement between Automatic Tongue Diagnosis System and Traditional Chinese Medicine Practitioners. *Evidence-based complementary and alternative medicine : eCAM*, 2012. 2012: 505063. doi:10.1155/2012/505063.

14. Li, Y.-Z., zheng Li, G., yi Gao, J., Zhang, Z.-F., chun Fan, Q., Xu, J.-T., Bai, G.-E., xian Chen, K., zhi Shi, H., Sun, S., Liu, Y., Shao, F.-F., Mi, T., Jia, X., Zhao, S., chang Chen, J., lian Liu, J., Guo, Y. R. and Tu, L. P. Syndrome Differentiation Analysis on Mars500 Data of Traditional Chinese Medicine. *TheScientificWorldJournal*. 2015.

15. Holzinger, A. and Jurisica, I. *Knowledge Discovery and Data Mining in Biomedical Informatics: The Future Is in Integrative, Interactive Machine Learning Solutions*. 2014, 1–18. doi:10.1007/978-3-662-43968-5_1.

16. Liu, C., Qiao, F., Yang, X. and Yang, H. Hardware acceleration with pipelined adder for Support Vector Machine classifier. 2014. ISBN 978-1-4799-3724-0. 13–16. doi:10.1109/DICTAP.2014.6821648.

17. Boser, B., Guyon, I. and Vapnik, V. A Training Algorithm for Optimal Margin Classifier. *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory*, 1996. 5. doi:10.1145/130385.130401.

18. Xu, H., Caramanis, C. and Mannor, S. Robustness and Regularization of Support Vector Machines. *Journal of Machine Learning Research*, 2008. 10.

19. Chien, Y. Pattern classification and scene analysis. *IEEE Transactions on Automatic Control*, 1974. 19(4): 462â463. doi:10.1109/tac.1974.1100577.

20. Cristianini, Nello and Shawe-Taylor, J. An introduction to support vector machines and other kernel-based learning methods. Repr. *Introduction to Support Vector Machines and other Kernel-Based Learning Methods*, 2001. 22. doi:10.1017/CBO9780511801389.

21. Burges, C. J. C. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 1998. 2(2): 121–167. ISSN 1573-756X. doi:10.1023/A:1009715923555. URL `https://doi.org/10.1023/A:1009715923555`.

22. Vapnik, V. N. *The nature of statistical learning theory*. Springer. 2010.

23. Kamarudin, N. D., Ooi, C. Y., Kawanabe, T., Odaguchi, H. and Kobayashi, F. A Fast SVM-Based Tongueâs Colour Classification Aided by k-Means Clustering Identifiers and Colour Attributes as Computer-Assisted Tool for Tongue Diagnosis. *Journal of Healthcare Engineering*, 2017. 2017: 1â13. doi:10.1155/2017/7460168.

24. Hastie, T., Friedman, J. and Tisbshirani, R. *The Elements of statistical learning: data mining, inference, and prediction*. Springer. 2017.

25. Chapelle, O., Haffner, P. and Vapnik, V. Support vector machines for histogram-based image classification. *IEEE Transactions on Neural Networks*, 1999. 10(5): 1055â1064. doi:10.1109/72.788646.

26. Song, X., Wang, H. and Wang, L. FPGA Implementation of a Support Vector Machine Based Classification System and Its Potential Application in Smart Grid. 2014. ISBN 978-1-4799-3188-0. 397–402. doi:10.1109/ITNG.2014.45.

27. Hsu, C.-W. and Lin, C.-J. A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks*, 2002. 13(2): 415â425. doi:10.1109/72.991427.

28. Nayak, J., Naik, B. and Behera, H. S. A Comprehensive Survey on Support Vector Machine in Data Mining Tasks: Applications Challenges. *International Journal of Database Theory and Application*, 2015. 8(1): 169â186. doi: 10.14257/ijdta.2015.8.1.18.

29. Saegusa, T., Maruyama, T. and Yamaguchi, Y. How fast is an FPGA in image processing? *2008 International Conference on Field Programmable Logic and Applications*, 2008. doi:10.1109/fpl.2008.4629911.

30. Eklund, A., Dufort, P., Forsberg, D. and Laconte, S. M. Medical image processing on the GPU â Past, present and future. *Medical Image Analysis*, 2013. 17(8): 1073â1094. doi:10.1016/j.media.2013.05.008.

31. Asano, S., Maruyama, T. and Yamaguchi, Y. Performance comparison of FPGA, GPU and CPU in image processing. *2009 International Conference on Field Programmable Logic and Applications*, 2009. doi:10.1109/fpl.2009.5272532.

32. Fowers, J., Brown, G., Cooke, P. and Stitt, G. A performance and energy comparison of FPGAs, GPUs, and multicores for sliding-window applications. *Proceedings of the ACM/SIGDA international symposium on Field Programmable Gate Arrays - FPGA 12*, 2012. doi:10.1145/2145694.2145704.

33. Wielgosz, M., Jamro, E., Å»urek, D. and Wiatr, K. FPGA Implementation of the Selected Parts of the Fast Image Segmentation. *Studies in Computational Intelligence*, 2012. 390: 203–216. doi:10.1007/978-3-642-24809-2_12.

34. Pietron, M., Wielgosz, M., Zurek, D., Jamro, E. and Wiatr, K. Comparison of GPU and FPGA Implementation of SVM Algorithm for Fast Image Segmentation. *Architecture of Computing Systems â ARCS 2013 Lecture Notes in Computer Science*, 2013: 292â302. doi:10.1007/978-3-642-36424-2_25.

35. Fan, R., Chen, P. and Lin, C.-J. Working Set Selection Using Second Order Information for Training SVM. *Journal of Machine Learning Research*, 2005. 6: 1889–1918.

36. He, B., Fang, W., Luo, Q., Govindaraju, N. and Wang, T. Mars: A MapReduce framework on graphics processors. 2008. 260–269. doi:10.1145/1454115.1454152.

37. Wen, Z., Shi, J., Li, Q., He, B. and Chen, J. ThunderSVM: A fast SVM library on GPUs and CPUs. *Journal of Machine Learning Research*, 2018. 19: 1–5.

38. Zeng, Z.-Q., Yu, H.-B., Xu, H.-R., Xie, Y.-Q. and Gao, J. Fast training Support Vector Machines using parallel sequential minimal optimization. *2008 3rd International Conference on Intelligent System and Knowledge Engineering*, 2008. doi:10.1109/iske.2008.4731075.

39. Kuan, T.-W., Wang, J.-F., Wang, J.-C. and Gu, G.-H. VLSI design of sequential minimal optimization algorithm for SVM learning. *2009 IEEE International Symposium on Circuits and Systems*, 2009. doi:10.1109/iscas.2009.5118311.

40. Cao, K.-K., Shen, H.-B. and Chen, H.-F. A parallel and scalable digital architecture for training support vector machines. *Journal of Zhejiang University SCIENCE C*, 2010. 11(8): 620â628. doi:10.1631/jzus.c0910500.

41. Peng, C.-H., Chen, B.-W., Kuan, T.-W., Lin, P.-C., Wang, J.-F. and Shih, N.-S. REC-STA: Reconfigurable and Efficient Chip Design With SMO-Based Training Accelerator. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 2014. 22(8): 1791â1802. doi:10.1109/tvlsi.2013.2278706.

42. Andraka, R. A survey of CORDIC algorithms for FPGA based computers. *Proceedings of the 1998 ACM/SIGDA sixth international symposium on Field programmable gate arrays - FPGA 98*, 1998. doi:10.1145/275107.275139.

43. Filho, J. G., Raffo, M., Strum, M. and Chau, W. J. A general-purpose dynamically reconfigurable SVM. *2010 VI Southern Programmable Logic Conference (SPL)*, 2010. doi:10.1109/spl.2010.5483031.

44. Bustio-MartÃnez, L., Cumplido, R., HernÃ¡ndez-Palancar, J. and Feregrino, C. On the Design of a Hardware-Software Architecture for Acceleration of SVMâs Training Phase. 2010, vol. 6256. 281–290. doi:10.1007/978-3-642-15992-3_30.

45. Ning, M., Shaojun, W., Yeyong, P. and Yu, P. Implementation of LS-SVM with HLS on Zynq. *2014 International Conference on Field-Programmable Technology (FPT)*, 2014. doi:10.1109/fpt.2014.7082816.

46. Wei, L., Chen, Z., Li, J. and Xu, W. Sparse and robust least squares support vector machine: A linear programming formulation. *2007 IEEE International Conference on Grey Systems and Intelligent Services*, 2007. doi:10.1109/gsis.2007.4443449.

47. Patil, R. A., Gupta, G., Sahula, V. and Mandal, A. Power Aware Hardware Prototyping of Multiclass SVM Classifier Through Reconfiguration. *2012 25th International Conference on VLSI Design*, 2012. doi:10.1109/vlsid.2012.47.

48. Hussain, H., Benkrid, K. and Seker, H. Novel dynamic partial reconfiguration implementations of the support vector machine classifier on FPGA. *Turkish Journal of Electrical Engineering and Computer Sciences*, 2014. 24. doi:10.3906/elk-1402-18.

49. Hussain, H. M., Benkrid, K. and Seker, H. Reconfiguration-based implementation of SVM classifier on FPGA for Classifying Microarray data. *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013. doi:10.1109/embc.2013.6610186.

50. Mandal, B., Sarma, M., Sarma, K. and Mastorakis, N. Implementation of systolic array based SVM classifier using multiplierless kernel. *Proceedings of the 16th International Conference on Automatic Control, Modelling Simulation (ACMOS '14)*, 2014: 288–294.

51. Vranjkovic, V. and Struharik, R. New architecture for SVM classifier and its application to telecommunication problems. *2011 19thTelecommunications Forum (TELFOR) Proceedings of Papers*, 2011. doi:10.1109/telfor.2011.6143852.

52. Koide, T., Hoang, A.-T., Okamoto, T., Shigemi, S., Mishima, T., Tamaki, T., Raytchev, B., Kaneda, K., Kominami, Y., Miyaki, R., Matsuo, T., Yoshida, S. and Tanaka, S. FPGA implementation of type identifier for colorectal endoscopie images with NBI magnification. *IEEE Asia-Pacific Conference*

on Circuits and Systems, Proceedings, APCCAS, 2015. 2015: 651–654. doi: 10.1109/APCCAS.2014.7032865.

53. Lesser, B., MÃŒcke, M. and Gansterer, W. Effects of Reduced Precision on Floating-Point SVM Classification Accuracy. *Procedia CS*, 2011. 4: 508–517. doi:10.1016/j.procs.2011.04.053.

54. Liu, C., Qiao, F., Yang, X. and Yang, H. Hardware acceleration with pipelined adder for Support Vector Machine classifier. *2014 Fourth International Conference on Digital Information and Communication Technology and its Applications (DICTAP)*, 2014. doi:10.1109/dictap.2014.6821648.

55. Afifi, S., Gholamhosseini, H. and Sinha, R. Hardware Acceleration of SVM-Based Classifier for Melanoma Images. *Image and Video Technology â PSIVT 2015 Workshops Lecture Notes in Computer Science*, 2016: 235â245. doi: 10.1007/978-3-319-30285-0_19.