CLASSIFICATION TECHNIQUES FOR HANDWRITING DIFFICULTIES AMONG CHILDREN IN EARLY STAGE OF ACADEMIC LIFE

ANITH ADIBAH BINTI HASSEIM

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Master of Engineering (Electrical)

Faculty of Electrical Engineering
Universiti Teknologi Malaysia

JUNE 2015

Specially dedicated to my beloved father and mother

Hasseim bin Shaaban and Robiah binti Romli

Also my beloved husband

Ismadi bin Ibrahim

My daughter Nur Auni Imthithal binti Ismadi

brothers, sisters and all my friends

for their inspiration, support and encouragement

throughout my adventure of educations

ACKNOWLEDGEMENT

First and foremost, I would like to express my grateful to Allah S.W.T because give me a good health and strength to complete my study. Without His permits, I would not be able to reach up until this level.

I wish to express my sincere appreciation to my project supervisor, Associate Professor Dr. Rubita binti Sudirman for her guidance and enthusiasm given throughout the progress of this study. Thanks for giving a worth knowledge that's help to improve myself.

My appreciation also goes to my husband and family who have been so tolerant and supported me all these years. Thanks for their encouragement, love and emotional supports that they had given to me.

I would also like to thank Dr. Puspa Inayat binti Khalid for her co-operations, guidance and helps in this study.

Nevertheless, my great appreciation dedicated to all my friends and those whom involve directly or indirectly with this study. There is no such meaningful word than thank you so much.

ABSTRACT

In today's era, all aspects of complex occupational task, plus the importance of early identification of developmental disorders in children, demand the essential need for screening children’s handwriting at elementary schools. Many underlying competence structures may interfere with handwriting performance. Children starting their academic programme should be tested for their handwriting abilities and readiness through regular routine screening. Screening a vast majority of 4 to 7+ years old necessitate the use of automated systems to collect data, keep tracks, and increase the speed of analysis and accuracy. Based on Handwriting Proficiency Screening Questionnaire (HSPQ) evaluated by their teachers, 120 pupils were individually tested on their use of graphic production rules. Then, the samples were divided into two group of writers; below average writers (test group) and above average writers (control group) based on the score of HSPQ. Each participant was required to copy four basic lines in two opposite directions and trace a sequence of rotated semi circles. This research examines the dynamic features such as ratio of time taken and standard deviation of pen pressure. In this study, three classification methods: Artificial Neural Network, Logistic Regression and Support Vector Machine (SVM) were chosen to classify children with handwriting problem. 10-fold cross-validation method is used for testing and training. At the end of this study, the results among these classifiers and features were compared. Based on the results, it can be concluded that the performance of SVM with Radial Basis Function kernel is the best among classifiers as it gives 100% of screening accuracy.

ABSTRAK

Dalam era hari ini, semua aspek dalam tugas pekerjaan yang kompleks, termasuk kepentingan pengenalan awal dalam gangguan perkembangan kanak-kanak, menuntut kepentingan membuat pemeriksaan awal tulisan tangan kanak-kanak awal persekolahan. Banyak struktur kecekapan asas boleh mengganggu prestasi tulisan tangan. Kanak-kanak yang memulakan program akademik mereka perlu diuji berdasarkan kebolehan tulisan tangan dan kesediaan melalui pemeriksaan rutin biasa. Pemeriksaan terhadap majoriti 4 hingga 7+ tahun, memerlukan sistem automatik untuk mengumpul data, menyimpan trek, dan meningkatkan kelajuan dan ketepatan analisis. Berdasarkan Kemahiran Soal Selidik Pemeriksaan Tulisan Tangan (HSPQ) yang dinilai oleh guru-guru mereka, 120 murid diuji secara individu ke atas penggunaan mereka terhadap kaedah pengeluaran grafik. Kemudian, sampel dibahagikan kepada dua kumpulan penulis; penulis di bawah purata (kumpulan ujian) dan di atas purata (kumpulan kawalan) berdasarkan skor HSPQ. Setiap peserta dikehendaki menyalin empat baris asas dalam dua arah yang bertentangan dan mengesan urutan separuh bulatan. Kajian ini telah mengkaji ciri-ciri dinamik seperti nisbah masa yang diambil dan sisihan piawai tekanan pen. Dalam projek ini juga, tiga kaedah klasifikasi: Rangkaian Neural Buatan, Regresi Logistik dan Mesin Vektor Sokongan (SVM) telah dipilih untuk mengelaskan kanak-kanak yang mempunyai masalah tulisan tangan. Kaedah 10 ganda pengesahan silang telah digunakan untuk ujian dan latihan. Di akhir projek ini, keputusan antara tiga kaedah pengelas dan ciri-ciri dinamik ini dibandingkan. Berdasarkan keputusan, dapatlah disimpulkan bahawa prestasi SVM dengan kernel Fungsi Asas Jejari adalah yang terbaik antara pengelas yang lain dengan mencapai 100% ketepatan seringan.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANN	-	Artificial Neural Network
HL	-	Horizontal Leftward
HR	-	Horizontal Rightward
HSPQ	-	Handwriting Proficiency Screening Questionnaire
LD	-	Left Oblique Downward
LR	-	Logistic Regression
LU	-	Left Oblique Upward
MSE	-	Mean Square Error
RBF	-	Radial Basis Function
RD	-	Right Oblique Downward
RU	-	Right Oblique Upward
SD	-	Standard Deviation
SVM	-	Support Vector Machine
VD	-	Vertical Downward
VU	-	Vertical Upward

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Overview

All aspects of complex occupational task as well as the importance of early identification of developmental problems in children, arises the importance to screen children's handwriting at elementary schools. On the development of digitalized modern world today such as keyboard with its fast typing task, can never be substitute handwriting. When writing or drawing, children perceptually extract relationships and perform according to their ability and skills. Many underlying competence structures may interfere with handwriting performance.

Penmanship must be a process associated with joy, esprit and self expression; instead of pressure, boring and complicated job for beginners. Academics argue handwriting is more than the transcription of ideas; it is directly related to how people generate and process those ideas. Handwriting is a language act, rather than just a motor act. It is about training the memory and hand to work together to correct and generate mental images and patterns of letters then interpret these into motor patterns of letters, automatically and without effort (Bloom, 2008). Reports emphasize that by ignoring poor handwriting may leads to fails to address a significant and continuing barrier to children's' achievement. The average or poor handwriting children have only 40% chance of achieving level 4, in national tests (Bloom, 2008); which indicates that bad handwriting can lead to unsuccessful in the examination by 40%. Therefore children need to develop the sub skills for handwriting at a young age, improving their basic hand eye and motor coordination, balance and visual perception.

Children starting their academic program should be tested for their handwriting abilities and readiness the same way as they are tested for optometry and dentistry through regular routine screening. Screening a vast majority of 4 to 7+ year old children not only to increase the speed of analyzing and processing data but also to collect data and keep track and increase the accuracy. Additionally, it significantly helps psychologists to monitor improvements and come up with modified new standard assessments based on the real large database. Moreover, this system will truly give a great benefit to the educational system and consequently the whole society's psychological and physical health.

Various softwares have been presented for handwriting recognition and movement analysis but, softwares directly related to child handwriting analysis with the prospective of screening children in general, and addressing difficulties are rare and the research is in its early stage. This research will hypothesize and examine the dynamic features that used by Khalid (Khalid, *et al*., 2010(a, b)). In contrast to Khalid approach (Khalid, 2012), different technique will be used to measure the influence of each feature in classifying pupil's handwriting performance. In this project, we propose using three classification methods which are Artificial Neural Network (ANN), Logistic Regression (LR) and Support Vector Machine (SVM).

1.2 Problem Statement

Writing is an important development skill for a child to master even though computers are widely used nowadays. Several study have been done to analyze pupils with handwriting difficulties, however most of the studies that involve in handwriting movement only give an attention to children with known physical or psychological problem. Nevertheless, not all these problems can be categorized as clear cut disease and condition. Hence, an effective solution should be identifying to indicate pupils who have difficulties in writing.

Several classifiers had been used in the literature reviews (Khalid, 2012; Guest *et. al*., 2003; Chindaro *et. al*., 2004) such as Hidden Marcov Model, Artificial Neural Network and Logistic Regression to select those who have handwriting problem. However, the maximum classification accuracy of these classifiers is just around 83% which is not highly enough to highlight the behavior differences between average and below average writers. Thus, it is the attention of this study to pilot the objective and select the best classification technique to increase the accuracy of prediction.

In contrast with similar method known by Khalid (2012), each different feature was tested individually and this study describe experiments carried out using Support Vector Machine (SVM) in addition to those classification methods used in previous researches. SVM is a supervised learning method that has proven it's efficiently over classic Neural Networks and its subset (Burges, 1998). The advantages of SVM are good generalization performance, able to handle high dimensional data and able to map the data into new high dimensional feature space for better classification using kernel functions.

1.3 Objectives of Study

The objectives of this study are:

i. To compare three classification methods to determine pupils who have difficulties in writing.
ii. To classify children either at risk of handwriting difficulties or not based on drawing task.

1.4 Scope of Study

The scope of the study is used as the guideline of the study. In order to achieve the objectives, the scope of the project has been confined as follows:

i. Data collection: normal healthy children between 7 to 12 years old in Skudai district performing copying and tracing tasks.
ii. Apply the different types of classifiers which are Artificial Neural Network, Logistic Regression and Support Vector Machine.
iii. Analyse the parameters based on the collected data.
iv. The standard deviation of pen pressure, ratio of time taken and the used of progression rules are the significant features that were used to identify children who are at risk of handwriting difficulties.

1.5 Significant of Study

This study investigates handwriting performance of a normal children aged between 7 and 12 years old. From the results, it may be helpful to the teachers because it can serve as a guide to deal with the problems and topic related to handwriting difficulties. The teachers also can monitor the students who have been indentified to have difficulties in writing and enable them to plan an action based on instructional programme that suited to the students' strengths and weaknesses.

Moreover, this study significantly helps psychologists to monitor improvements and come up with modified new standard assessments based on the real large database. On the other hand, this system will truly give a great benefit to the educational system and consequently the whole society’s psychological and physical health.

The student will be the most benefited by the results of this study. The finding will guide them in terms of their limitation and weaknesses. In addition, the students who have difficulties in handwriting can have more attention on the writing's physical process , thus limiting use of higher order cognitive planning, skills and generation of content. Therefore, the students will manage to write and complete their writing task efficiently.

1.6 Thesis Organization

This thesis consists of five chapters. Chapter 1 introduces the background of the research, problem statement, objective of the study, scope of the study and the overall thesis outline.

Chapter 2 focuses on the handwriting problems as well as the development of handwriting project. This chapter also give the explanation about dynamic features that used in this study. Furthermore, the classification methods which are artificial neural network, support vector machine and logistic regression have been discussed at the end of this chapter.

Chapter 3 reports the details of the dataset used in this study. It includes the process of developing the supplementary screening as well as extraction of dynamic features. Besides, this chapter also discusses the implementation of three classification methods used in this study.

Chapter 4 explains and discusses the results obtained and the analysis made. Comparisons between the classifiers' results are made in order to achieve the final objective of this study. Among these classifiers, further analysis has done with increasing the number of features for the best classifier only.

Finally, Chapter 5 gives the conclusion of this study. Last but not least, this chapter also gives some recommendation on future development of this study.

REFERENCE

Berninger V. (1999). *"Coordinating transcription and text generation in working memory during composing: Automatic and constructive processes."*Learning Disability Quarterly, 22, 99-112.

Berninger V., Cartwright A., Yates C., Swanson H. L. and Abbott R. (1994). *"Developmental skills related to writing and reading acquisition in the intermediate grades: Shared and unique variance."*Reading and Writing: An Interdisciplinary Journal, 6: 161-196.

Berninger V., Mizokawa, D., and Bragg, R. (1991). "*Theory-based diagnosis and remediation of writing*." Journal of School Psychology, 29: 57-59.

Bezine H., Alimi A. D. and Sherkat N. (2004). *"Generation and analysis of handwriting script with the beta-elliptic model,"* Proceedings of the Ninth International Workshop on Frontiers in Handwriting Recognition, 515-520.

Bian Z. Q. and Zhang X. G. (2000). "*Pattern recognition*," Tsinghua University Press, Beijing.

Bloom A. (2008). *"Not just a matter of neatness"*. Published in The TES [News]. Available: http://www.tes.co.uk

Bonny A. M. (1992). *"Understanding and assessing handwriting difficulties: Perspective from the literature"*. Australian Occupational Therapy Journal, 39: 7-15.

Bulacu M., Schomaker L. (2007). "*Automatic handwriting identification on medieval documents"*, 14th International Conference on Image Analysis and Processing (ICIAP 2007), 279-284.

Burges C. J. C. (1998). *"A tutorial on support vector machines for pattern recognition"*, Data Mining and Knowledge Discovery, 2: 121-167.

Chindaro S., Guest R. M., Fairhurst M. C., and Potter J. M. (2004).*"Assessing visuo-spatial neglect through feature selection and combination from geometric shape drawing performance and sequence analysis."*International Journal of Pattern Recognition & Artificial Intelligence, 18: 1253-1266.

Chowriappa A., Subrahmaniyan, N., Srimathveeravalli, G., Bisantz, A., Kesavadas, T. (2009). "*Modeling and defining expert handwriting behavior*," in Systems, Man and Cybernetics, IEEE International Conference on, 2918-2923.

Chu, S. (1997). *"Occupational therapy for children with handwriting difficulties: A framework for evaluation and treatment",* British Journal of Occupational Therapy, 60: 514-520.

Fairhurst M. C., Linnell T., Glinet S., Guest R. M. (2008). "*Developing a generic approach to online automated analysis of writing and drawing tests in clinical patient profiling*," Behavior Research Methods, 40:290-303.

Feder K. P. and Majnemer A. (2007). "*Handwriting development, competency, and intervention*", Developmental Medicine Child Neurol, 49: 312-7.

Freeman H.(1974). *"Computer processing of line-drawing images",* ACM Computing Surveys, 6: 57-97.

Gillespie H. M. (2003). *"Component handwriting skills among early elementary children with average and below average printing ability"*, Northwestern University.

Graham, S., and *Weintraub, N.* (*1996*). *"A review of handwriting research: Progress and prospects from 1980-1994"*. Educational Psychology Review , 8, 7-87.

Graham S., Berninger V. W., Weintraub N. and Schafer W. (1997).*"Development of handwriting speed and legibility in grades 1–9,"*Journal of Educational Research, 92: 42–52.

Graham, S., Berninger, V.W., and Weintraub, N (1998). *"The relationship between handwriting style and speed and legibility",* The Journal of Educational Research, 5: 290-296.

Guest R. M., Chindaro S., Fairhurst M. C., and Potter J. M. (2003).*"Automatic classification of hand drawn geometric shapes using constructional sequence analysis,"* Seventh International Conference on Document Analysis and Recognition (ICDAR'03), Edinburgh, Scotland, UK, 2: 990.

Guest R. M., Fairhust M. C., and Potter J. M. (2012).*"Diagnosis of visuo-spatial neglect using dynamic sequence features from a cancellation task",* Pattern Recognition Analysis & Applications, 5: 261–270.

Gunn S. (1998). "*Support Vector Machines for Classification and Regression*", ISIS Technical Report ISIS-1-98, Image Speech & Intelligent Systems Research Group, University of Southapton.

Hamstra-Bletz, L., and Blote, A. (1993). *"A longitudinal study on dysgraphic handwriting in primary school".* Journal of Learning Disability, 26, 689–699.

Ishida H., Takahashi T., Ide I. and Murase H. (2010). *"A Hilbert warping method for handwriting gesture recognition",* Pattern Recognition, 43: 2799-2806.

Jones, D. and Christensen, C. A. (1999). "*Relationship between automaticity in handwriting and student's ability to generate written text*." Journal of Educational Psychology, 91: 44-49.

Keerthi S. S. and Lin C. J. (2003).*"Asymptotic behaviors of support vector machines with Gaussian kernel",* Neural Computation, 15(7):1667-1689.

Khalid P. I. (2012).*"Handwriting ability assessment model using dynamic characteristics of drawing process."* PhD Thesis, Universiti Teknologi Malaysia.

Khalid P. I., Yunus J., Adnan R. (2010(a)). "*Extraction of dynamic features from hand drawn data for the identification of children with handwriting difficulty,*" Research in Developmental Disabilities, 31: 256-262.

Khalid P. I., Yunus J., Adnan R., Harun M., Sudirman R., Mahmood N. H. (2010(b)). "*The use of graphic rules in grade one to help identify children at risk of handwriting difficulties,*" Research in Developmental Disabilities, 31: 1685-1693.

Kim Y. S. (2010). *"Performance evaluation for classification method: A comparative simulation study,"* Expert System with Application, 37, 2292-2306.

Kimeldorf G. and Wahba G. (1971). *"Some results on Tchebycheffian spline functions"*, Journal on Mathematical Analysis Application, 33: 82-95.

Koziatek S. M. and Powell N. J. *"Pencil grips, legibility, and speed of fourth-graders' writing in cursive".* American Journal of Occupational Therapy, 57: 284–288.

Lee M. C. and To C. (2010). "*Comparison of support vector machine and back propagation neural network in evaluating the enterprise financial distress*". International Journal of Artificial Intelligence & Applications (IJAIA), 1; 3.

Lefteri H. T. and Robert E. U. (1997). "*Fuzzy and neural approaches in Engineering*", John Wiley and Sons.

Liu Y. Y., Yang M., Ramsay M., Li X. S., and Coid J. W. (2011). "*A comparison of logistic regression, classification and regression tree, and neural network* *models in predicting violent re-offending*". Journal of Qualitative Criminology, 27(4): 547-573.

Lorigo L. M. and Govindaraju V. (2006). *"Offline Arabic handwriting recognition: a survey",*Pattern Analysis and Machine Intelligence, IEEE Transactions on, 28: 712-724.

McHale K. and Cermak S. A. (1992).*"Fine motor activities in elementary school: Preliminary findings and provisional implications for children with fine motor problems",*American Journal of Occupational Therapy, 46: 898-903.

Pavlekovic M., Bensic M., and Zekic-Susac M. (2010). "*Modeling children's mathematical gift by neural networks and logistic regression*", Expert System with Application. 37(10): 7167-7173.

Reisman, J. E. (1993). "*Development and reliability of the research version of the minessotahandwriting test."*Physical and Occupational Therapy in Pediatrics, 13: 41-55.

Richardson A. (2008). "*COACH - Cumulative online algorithm for classification of handwriting deficiencies*," IAAI'08 20th National Conference on Innovative Applications of Artificial Intelligence, Chicago, IL, 3: 1725-1730.

Rmi C. (2002). "*Automatic analysis of the structuring of children's drawings and writing*," Pattern Recognition, 35: 1059-1069.

Rosenblum S., Parush S., Weiss P. L. (2003), *"The In Air phenomenon: temporal and spatial correlates of the handwriting process"*. Perceptual Motor Skills 96 (3 pt 1), 933-954.

Rosenblum S., Dvorkin A. Y., Weiss P. L. (2006), *"Automatic segmentation as a tool for examining the handwriting process of children with dysgraphic and proficient handwriting",* Human Movement Science, 25, 608-621.

Rosenblum S. (2008). "*Development, reliability and validity of the Handwriting Proficiency Screening Questionnaire (HPSQ)*". American Journal of Occupational Therapy, 62(3), 298-307.

Rosenblum S., Aloni T., and Josman N. (2010), "*Relationship between handwriting performance and organizational abilities among children with and without dysgraphia: A preliminary study."* Research in Developmental Disabilities. 31: 502-509.

Srihari S., Collins J., Srihari R., Srinivasan H., Shetty S., Griffler J. B. (2008). *"Automatic scoring of short handwritten essays in reading comprehension tests",* Artificial Intelligent, 172: 300-324.

Stoean, R. and Stoean, C. (2013). *"Modeling medical decision making by support vector machines, explaining by rules of evolutionary algorithms with feature selection."*Expert Systems with Applications, 40(7): 2677-2686.

Sudirman R., Mashadi N. T., Ariffin I. (2011). "*Aspects of a standardized automated system for screening children's handwriting,*" First International Conference on Informatics and Computational Intelligence, 49-54.

Vapnik V. N. (1995). *"The nature of statistical learning theory"*, Springer, New York.

Whyte E. (2007, August 9). *"Copying with reading, writing issues".* New Straits Times. p. C14.

Wilamowski B. M. (2009). *"Neural Networks or Fuzzy Systems"*, Workshop on Intelligent Systems, Budapest, Hungary, 1-12.

Yong Z., Tan T. and Wang Y. (2000). *"Biometric personal identification based on handwriting,"* Proceedings of the 15^{th} International Conference on Pattern Recognition, 2: 797-800.

Yu Z. (2000). *"Feed forward neural network and their applications in forecasting".* M. Sc. Thesis, University of Houston.

Yuan G. X., Ho C. H., Lin C. J. (2012). *"Recent advances of large-scale linear classification",* Proceedings of the IEEE 100(9): 2584-2603.