Handwriting Difficulty Screening Tool based on Dynamic Data from Drawing Process

Y. M. Ling and P. I. Khalid
Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Malaysia.
e-puspa@utm.my

***Abstract*—Children with handwriting difficulty are advised to join an intervention program to rectify the problem at an early stage. However, the available screening tools suffer from subjectivity judgement while lack of expertise reduces the chance for every student to be screened. Yet, digitalized screening tools that use dynamic data from writing activities are only applicable to those who know the language. These limitations had led this study to develop an objective handwriting difficulty screening tool based on dynamic data of drawings. Three attributes extracted from 120 sets of dynamic data from drawing process were found to be significant in differentiating below-average writers from average writers. The attributes were then used to train Support Vector Machine prediction model. To test the validity and reliability of the prediction model, additional sets of data were acquired from 36 pupils. The performance of the tool was compared with the results from the Handwriting Proficiency Screening Questionnaire (HPSQ) that employs teachers' observations on pupils' handwriting ability. With 78% reliability, 69% of the predictions made by the developed tool was in accordance with the teachers' observation. Most importantly, 53% of the average writers were screened as having handwriting problems. This denotes the objectivity of the developed tool in identifying below-average writers who failed to be recognized through teacher's observation.**

***Index Terms*—Dynamic Attributes; Drawing Process; Handwriting Difficulty; Support Vector Machine.**

I. INTRODUCTION

Handwriting skill, a complicated ability that engages with fine motor skill, visualization, and cognition, begins to develop at early childhood age [1]. Unfortunately, some children experience handwriting difficulty where they find a lack of muscle coordination, require extra time to complete a writing task, and experience muscle fatigue sooner than their peers [2]. Even though this problem has no direct relation with intelligence quotient [3], below-average writers often demotivate by the loads of written homework at schools. On top of this, the untidy handwriting often judged as laziness, impatience, and carelessness by teachers [4]. This would post a negative impact and may cause behavioral problems in their learning process [5].

Undeniably, early intervention may help to improve below-average writers' writing skill. In fact, there are several tests and assessments structured to screen handwriting difficulty among children, such as the Concise Assessment Scale of Children's Handwriting (BHK) [6], Development Test of Visual-Motor Integration (VMI) [7], Minnesota Handwriting Assessment Test (MHA) [8], Handwriting Proficiency Screening Questionnaire (HPSQ) [9], and Literacy and Numeracy Screening (LINUS) [10]. These tests assess written product quality such as size, form, and spacing, as well as time at various experimental settings. As these manual screening tools evaluate performance time and product legibility, involvement of expert is required [5]. Besides, most of the evaluations are based on a scoring system. The drawbacks are such that the evaluation provides no reflection on handwriting effort like strategies and pen pressure, lack of manpower to do the screening due to the limited number of occupational therapist available in the society, and the result is prone to subjective judgement especially when the assessment requires evaluator's observations. Obviously, the disadvantages of manual screening induced this field to venture into a digitalized evaluation mode to obtain an objective screening tool.

There are several digitalized handwriting difficulty evaluation tools under research and development, which focus on dynamic data analysis, such as pressure, time, coordinate and angles. Computerized Penmanship Evaluation Tool (ComPET) computes the mean width, height, pressure and tilting angle of every stroke within the entire paragraph [11]. Another instrument which is Chinese Handwriting Computer-Based Handwriting Assessment checks the accuracy, pausing time and writing speed for the written characters [12]. However, digitalized evaluation tools that use specific language for writing task have reduced the possibility to outreach people with different culture. The inconvenience caused in digitalized evaluation mode has further changes the writing task input to drawing task. Khalid *et al.* [13, 14] used drawing process to differentiate the below-average writers from the average writers, and had convincingly shown the feasibility of replacing writing tasks with drawing tasks for handwriting ability assessment [5]. Therefore, it is the focus of this research to develop an objective and automated handwriting screening tool based on dynamic data from drawing activities.

II. LITERATURE REVIEW

A. Grammar of Action

Grammar of Action is the applied rules in nature used in constructing geometry drawings or characters such as the selection of starting point, progression direction and the ending point [15]. The handwriting progression rule is largely influenced by the formal instruction that the children received in schools, whereby the writing principles are related to culture. When constructing a composite figure of more than one straight lines, vertical or oblique line would first be drawn and followed by a horizontal line [15]. According to the author, right-handed children prefer to pick up the topmost and/or the leftmost point as the starting point and then progress downward and/or rightward. For example, a horizontal line would be drawn from left to right. In contrast, left-handed children tend to draw it in the opposite direction. Research has shown that 85% of the progression rules in printing among children are obeyed [15]. However, the rules are not applicable in some drawings. The preference of progression is influenced by the structure of the drawing itself whereby children would have it completed with least effort and fewest total movement. This planning is also associated with hand motor planning, for examples the pencil positioning adjustment as well as finger flexion and extension [4]. Among below-average writers, limitation of hand motor capabilities cause them to less likely apply Grammar of Action in writing [5].

B. Writing and Drawing

The development of handwriting skill started with basic drawing. The relation between drawing and writing is significant, showing drawing is a fundamental to writing tasks [16]-[17]. The capability in copying basic lines and shapes such as Berry VMI shapes serves as an indicator to show that children are ready to write [7]. Moreover, research has shown that the dynamic attributes of drawings are capable of differentiating below-average writers from average writers [13]-[14]. This has ignited new page in analyzing handwriting difficulty as it improves the universality of the screening tool. Screening tool based on drawing process can be used worldwide, regardless of language and cultural background. Children who have just acquired drawing skills and yet to develop writing skills can be screened as well, thus allowing the screening test to be done at the earliest stage as possible to indicate the need for tailored advice and intervention.

C. Dynamic Attributes of Drawing

Khalid [5] utilized copying and tracing tasks, in which copying tasks involved four straight lines, each in two way directions, while tracing task involved four rotated semicircles, each different by 90°. From 85 extracted dynamic attributes, the four most significant attributes are the standard deviation of pen pressure when drawing the right oblique line in an upward direction (P1), the time ratio of drawing horizontal line in rightward direction to the leftward direction (P2), mean of pen pressure when tracing vertically flipped C (P3), and the use of progression rules when tracing the second, third and fourth semicircles (P4). However, due to its negative regression coefficient, the author had excluded the mean pressure of flipped C (P3) when training the classifiers.

Succeeding the research, another experiment was carried out with the setting of attributes focused on angle components, where threshold values were set to perform pattern recognition. Neo *et al.* [18] applied the copying tasks of straight lines in eight directions, circles in both clockwise and anticlockwise directions, and four semicircles without specific direction. The significant standard deviation angle was found in drawing left downward oblique, left upward oblique and right downward oblique. Gap angles in four rotated semicircles were also found to be significant.

D. Classification Techniques

Two commonly used techniques in machine learning are Support Vector Machine (SVM) and Artificial Neural Network (ANN) [19]. In the context of handwriting difficulty screening tool, Khalid [5] had used 120 samples and applied 10-fold cross validation procedure to train and test the Logistic Regression (LR) and ANN classifiers. With only 3 attributes (P1, P2, and P4), both classifiers had equally good ability (an average accuracy of 83%) to screen pupils who are at risk of handwriting difficulty.

Later, Hasseim [20] had used same attributes as Khalid [5] and compared the performance of Logistic Regression (LR), ANN, and SVM classifiers in screening below-average writers. SVM was claimed to have the best prediction accuracy. An improvement of accuracy was contributed by the parameter setting, specifically on the training epochs [20]. The suggested parameters are C set to 1 while γ tuned to 0.1. Furthermore, SVM has better stability as compared to ANN. With its shorter computation time as well as simpler architecture, it is a good choice as classifier for handwriting difficulty screening tools.

Based on the findings from Khalid [5] and Hasseim [20], an automated screening tool was developed using SVM classifier [21]. This tool used 72 samples to train and test the classifier. By using only two attributes from copying tasks (P1 and P2) and 4-fold cross validation procedure, the classification accuracy was improved to 89%. However, the author did not use a new set of data to test the validity of the tool. Meanwhile, it is believed that increasing the number of samples and predictors in the training phase could improve the accuracy of the tool. Therefore, this paper adopted the idea from [21] and increased the number of training samples as well as predictors in the predictive model to improve the classification accuracy.

III. METHOD

A. Participants

Data from one hundred and twenty Year-One children of 6 to 7 years old [5] were used in this study. The number of below-average writers, and average writers were sixty for each group. The scoring assessment of the Handwriting Proficiency Screening Questionnaire (HPSQ) by school teachers were used to categorize the participants into the two groups.

B. Instruments

The construction of the screening tool involved hardware and software intelligence. The hardware included a Wacom Intuous3 tablet and a stylus pen. Pen pressure, pen coordinates and time during the drawing process were recorded and stored in the format of ASCII file and transferred to laptop that was connected with the Wacom tablet.

From the aspect of software intelligence, Microsoft Visual Studio and R studio were used. Microsoft Visual Studio was used to develop attribute extraction algorithm in the C++ programming language while the graphical user interface (GUI) was developed from the C# Winforms application. The support vector machine extension was adopted from LIBSVM [22]. On the other hand, R studio was used in statistical analysis, especially in detecting significant attributes.

C. Experimental Setup

The design was separated into three phases: attribute identification, screening tool development, and system testing. In attribute identification phase, input data (x-y coordinates of drawing movement, pen pressure and time) from 120 Year-One students were adopted from [5]. These data were then injected into an attribute extraction algorithm and the output were tabulated. The four most significant attributes reported in [5] were extracted and the t-test and chi-square test were used to measure the significance of the extracted attributes. Only the significant attributes were selected to be used in the screening tool.

Next, in screening tool development phase, the output from the attribute extraction algorithm was fed into SVM classifier for system training. A GUI that linked the attributes extraction algorithm and SVM coding was created. Support Vector Machine Radial Basis Function (SVM RBF) kernel was employed. The two most important learning parameters for the SVM RBF kernel are C and γ. The tuning parameter γ is a control variable for the RBF amplitude and SVM generalization ability. The penalty parameter C, on the other hand, controls the influence of each individual support vector, which determines the trade-off between the complexity of decision rule and frequency of error [23].

To investigate the effect of the number of attributes on the classification accuracy, the significant attributes were grouped into groups of two, three and four attributes. The accuracy of the system based on each group of attributes was computed under the variation of parameters C and γ set at (1, 10 and 100) and (0.01, 0.1 and 1) respectively. To maximize the use of the 120 data samples, 10-fold cross validation procedure was applied. The best performed group of attributes was selected to be used as the model of prediction. The model was built using the same 120 data samples, but without cross validation procedure, which was then imparted into the screening system.

Lastly, dynamic data of drawing tasks from additional 36 students were acquired in system testing phase to check the system's validity and reliability. To test the system performance, the raw data went through attribute extraction algorithm; the extracted attributes were the significant attributes that were determined in attribute identification phase. The attributes were then fed into a support vector machine to do the prediction. To determine the validity of the developed tool, the system's prediction was compared with HPSQ result. The whole procedure was compiled in the GUI to ease user in injecting input as well as viewing prediction output.

In this last phase, HPSQ was distributed to the pupils' teachers. HPSQ was designed by Rosenblum [9] and applied in Khalid [5] and Chin [21] to test the validity of the findings. The teachers were required to observe the pupils writing behaviours in a classroom before completing the questionnaire. There were 10 questions with five-point Likert scale. The final scores ranging from 0 to 13 were categorized as average writer while scores greater than 14 were categorized as below-average writers. These results were then compared with the results from the developed screening tool to report the system's validity.

Thirty six participants were tested individually under similar environment settings. The tests were conducted in a private classroom, with a standard school chair and digitizing tablet on school desk. The drawing tasks were performed on A5 papers overlaid on the surface of Wacom Intous$_3$ using a wireless electronic pen. The data of x-y coordinates of drawing movement, pen pressure and time were sampled at 100Hz and stored in a computer for off-line processing.

Two types of drawing tasks were employed, namely copying and tracing. For copying task, the participants were required to draw the pattern appeared on the left column on the response frame in the right column, as shown in Figure 1 and Figure 2, in which the direction of drawings were verbally instructed. On the other hand, tracing task required the students to trace four rotated semicircles on the dotted line, as depicted in Figure 3, without direction assigned to lead the participant's progression. The sequence of the drawing activities is copying of horizontal rightward (HR), horizontal leftward (HL), right oblique upward (RU), and lastly is tracing of rotated semicircles.

As the exerted pen pressure varies among pupils, the raw pen pressure data should be normalized before extracting any related attribute. The value of pen pressure was normalized by using the following equation:

$$Normalized value = \frac{A - B}{C - B} \tag{1}$$

where: A = Original value
B = Minimum observed value except 0
C = Maximum observed value in a sample.

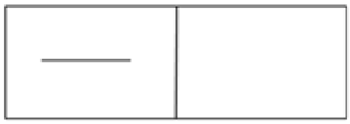


Figure 1: Copying task of horizontal line

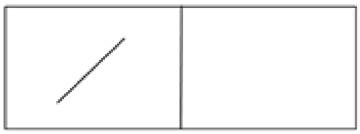


Figure 2: Copying task of RU

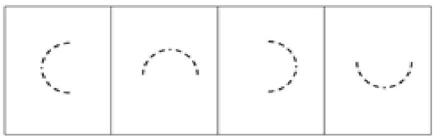


Figure 3: Tracing task of four rotated semicircles

To observe the consistency of the prediction (system's reliability), 23 out of the 36 pupils were asked to perform the drawing tasks four times repetitively. The system was considered reliable if it predicted 3 or 4 similar results for each pupil.

IV. RESULTS AND DISCUSSIONS

A. *Significance of the Attributes*

Table 1 highlights the four extracted attributes and Table 2 tabulates the significant outcome of the attributes. Statistical analysis showed that the extracted attributes had their p-values lesser than 0.05 which leads to a conclusion that these attributes can significantly distinguish below-average writers from average writers. This result is in accordance with the findings reported in Khalid [5].

Table 1
Attributes Extracted with Respective Drawings

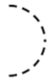
Attribute	Drawing	Figure
Progression rule categorical attribute	2nd, 3rd and 4th semicircles	
Mean pressure	3rd semicircle	
Time ratio	HR to HL	
Standard deviation pressure	RU	

Table 2
Significance of the extracted attributes

Test	Attributes	p-value	Result
Chi-square	Progression rule of the last three semicircles	2.055×10^{-6}	Significant
T	Mean pressure of 3rd semicircle	1.552×10^{-3}	Significant
T	Standard deviation pressure of RU	2.058×10^{-7}	Significant
T	Time ratio of HR to HL	3.835×10^{-6}	Significant

B. *Predictive Model of SVM*

Preceding results led to the combination of the attributes into 3 different groups. Table 3 lists the attributes involved in each group while Table 4 highlights the classification accuracy of SVM. The highest accuracy across all groups were observed at the parameter combination of (100, 0.1). Nevertheless, a combination of two attributes had the lowest accuracy while three and four attributes ranked similarly. It can be seen that the addition of progression rule categorical attributes is able to enhance the performance of the screening tool. However, further addition of attribute did not contribute to better prediction performance. This finding is parallel with Hasseim [20]. On top of this, it is also supporting the decision made in Khalid [5] to void negative regression coefficient in determining handwriting difficulty among students. Hence, the combination of three attributes with parameters C and γ set to (100, 0.1) was selected to be the predictive model for the screening tool.

Table 3
Attribute Groups and the Attributes Employed

Group	Number of Attributes	Attributes
A	2	Time ratio of HR to HL, Standard deviation (SD) pressure of RU
B	3	Time ratio of HR to HL, SD pressure of RU, Progression rule of three-rotated semicircles
C	4	Time ratio of HR to HL, SD pressure of RU, Progression rule of three-rotated semicircles, Mean pressure of 3rd semicircle

Table 4
10-folds Cross Validation Classification Accuracy (%)

Group	Number of Attributes	C \ γ	0.01	0.1	1
A	2	1	73.33	73.33	75.00
		10	73.33	75.00	74.17
		100	75.00	76.67	73.33
B	3	1	70.00	70.00	77.50
		10	70.00	78.33	77.50
		100	78.33	79.17	79.15
C	4	1	70.00	70.00	74.17
		10	70.00	75.00	76.67
		100	78.33	79.17	79.17

C. *Validity and Reliability of the Screening Tool*

The HPSQ scores classified 19 pupils as below-average writers and 17 pupils as average writers. The dynamic data of all students from the first trial were analysed for validity and the results are summarized in Table 5. Out of 36 sets of data, 69% of the tool's prediction agreed with the HPSQ results, or in other words, resembling teachers' observations. The tool had able to correctly screen 89% of the below-average writers. This indicates that these pupils have obvious symptoms of handwriting difficulty that can be easily noticed by the teachers. However, 53% of the average writers were screened as having handwriting problems. Literally, there are hidden symptoms among these pupils that failed to be recognized by the teachers. Most probably the symptoms were undercover by good results, cleverness, hardworking and other positive traits of the students. Overall, result discrepancy between the developed tool and the HPSQ is only 31%. The majority falls onto average writers.

Table 5
Validity of the Developed Screening Tool

Screening Tool \ HPSQ	Normal	At Risk
Similar to HPSQ	8	17
Dissimilar to HPSQ	9	2

The performance of the developed screening tool was better than the tool developed by Chin [21]. Summary of the comparison is tabulated in Table 6. Apparently, this tool can better screen those who do not have obvious symptoms, but yet may be at risk of handwriting difficulty.

Table 6
Performance Comparison of the Two Screening Tools

Performance		Developed Tool	Chin's Tool
Validity	Overall	69%	45%
	Average writer	47%	53%
	Below average writer	89%	37%

To test the system reliability, 23 participants (13 below-average writers and 10 average writers) repeated the tests four times. Consistent results across three or four trials were observed in most of the samples and the details is tabulated in Table 7. Compared to Chin's system which has reliability of 61%, this screening tool is able to achieve 78%, which is 17% better.

Table 7
Reliability of the Developed Screening Tool

Repeated test \ Handwriting ability	Normal	At Risk
Consistent results	7	11
Inconsistent results	3	2

V. Conclusion

Pertaining the obtained results, the extracted attributes are significant and capable of distinguishing below-average writer from the average writer at better consistency percentage. It provides an objective screening without relying on evaluators' observation. Also, the utilization of basic drawings makes it more universal as compared to other digitizing screening systems that used specific language of writing tasks. Hence, the feasibility of using this handwriting screening tool to assess students of various races and nationalities, as well as those who have yet to develop writing skills can be explored.

Acknowledgment

This research was supported by the Malaysian Ministry of Higher Education (Fundamental Research Grant Scheme, vote no. 4F870) and Universiti Teknologi Malaysia. The authors would like to express gratitude to the involvement of teachers and pupils of Sekolah Kebangsaan Taman Universiti (1), Sekolah Kebangsaan Taman Universiti (2), Sekolah Kebangsaan Taman Universiti (3), Sekolah Kebangsaan Taman Mutiara Rini (2), and Sekolah Kebangsaan Taman Desa Skudai.

References

[1] C. E. Cameron, L. L. Brock, W. M. Murrah, L. H. Bell, S. L. Worzalla, D. Grissmer and F. J. Morrison, "Fine motor skills and executive function both contribute to kindergarten achievement," *Child Development*, vol. 83, no. 4, pp. 1229-1244, 2012.

[2] M. J. Hartingsveldt, E. H. C. Cup, J. C. M. Hendriks, L. de Vries, I. J. M. de Groot, and M. W. G. Nijhuis-van der Sanden, "Predictive validity of kindergarten assessments on handwriting readiness," *Research in Developmental Disabilities*, vol. 36, pp. 114-124, 2015.

[3] E. Pagliarini, M. T. Guasti, C. Toneatto, E. Granocchio, F. Riva, D. Sarti, B. Molteni, and N. Stucchi, "Dyslexic children fail to comply with the rhythmic constraints of handwriting," *Human Movement Science*, vol. 42, pp. 161-182, 2015.

[4] K. P. Feder and A. Majnemer, "Handwriting development, competency and intervention," *Developmental Medicine and Child Neurology*, 49, pp. 312-317, 2007.

[5] P. I. Khalid, "Handwriting Ability Assessment Model Using Dynamic Characteristics of Drawing Process," *Ph.D thesis*, Universiti Teknologi Malaysia, Malaysia, 2012.

[6] M. J. M. Volman, B. M. Van Schendel, and M. J. Jongmans, "Handwriting difficulties in primary school children: A search for underlying mechanisms," *American Journal of Occupational Therapy*, vol. 60, no. 4, pp. 451-460, 2006.

[7] K. E. Beery and N. A. Beery, *The Beery-Buktenica Developmental Test of Visual-Motor Integration: Beery VMI Administration, Scoring, and Teaching Manual*. 5th. ed. Minneapolis, MN: NCS Pearson, 2004.

[8] K. L. Rostan, J. Hinojosa, and H. Kaplan, "Using the Minnesota Handwriting Assessmnet and Handwriting Checklist in screening first and second graders' handwriting legibility," *Journal of Occupational Therapy Schools, & Early Intervention*, vol. 1, no. 2, pp. 100-115, 2008.

[9] S. Rosenblum, "Development, reliability, and validity of the Handwriting Proficiency Screening Questionnaire (HPSQ)," *American Journal of Occupational Therapy*, vol. 62, no. 3, pp. 298-307, 2008.

[10] R. S. Kamini, "Raising the bar for all schools," *New Straits Times*, September 24th 2010 [Retrieved December 20, 2010, from http://www.nst.com.my].

[11] L. Hen, N. Josman and S. Rosenblum, "Tele-evaluation and intervention among adolescents with handwriting difficulties – computerized penmanship evaluation tool (ComPET) implementation," *Proceedings of the 7th International Conference on Disability, Virtual Reality, and Associate Technologies (ICDVRAT) with ArtAbilitation*, Maia, Portugal, 2008, pp. 143-149.

[12] T. F. Wu, G. S. Chen, and H. S. Lo, "The application of computer-based chinese handwriting assessment system to children with dysgraphia" in *ICCHP 2014, Part 2, LNCS*, vol. 8548, K. Miesenberger et al., Eds, 2014, Switzerland: Springer International Publishing, pp. 532-539.

[13] P. I. Khalid, J. Yunus, and R. Adnan, "Extraction of dynamic features from hand drawn data for the identification of children with handwriting difficulty," *Research in Developmental Disabilities*, vol. 31, pp. 256-262, 2010.

[14] P. I. Khalid, J. Yunus, R. Adnan, M. Harun, R. Sudirman, and N. H. Mahmood, "The use of graphic rules in grade one to help identify children at risk of handwriting difficulties," *Research in Developmental Disabilities*, vol. 31, no. 6, pp. 1685-1693, 2010.

[15] M. L. Simner, "The grammar of action, printing, and reversals in children's printing," *Developmental Psychology*, vol. 17, no. 6, pp. 866-871, 1981.

[16] S. Steffani and P. M. Selvester, "The relationship of drawing, writing, literacy and math in kindergarten children," *Reading Horizons*, vol. 49, no. 2, pp. 125-142, 2009.

[17] F. Bonoti, F. Vlachos and P. Metallidou, "Writing and drawing performance of school age children. Is there a relationship?" *School Psychology International*, vol. 26, no. 2, pp. 243-255, 2005.

[18] C. C. Neo, E. L. M. Su, P. I. Khalid and C. F. Yeong, "Computerized assessment for early screening of children with handwriting difficulty," *Applied Mechanics and Materials*, vol. 432, pp. 392-397, 2013.

[19] N. A. Kamaruddin, "Temporal Spectral Approach to Surface Electromyography Based Fatigue Classification of Biceps Brachii During Dynamic Contraction," *Master. thesis*, Universiti Teknologi Malaysia, Malaysia, 2016.

[20] A. A. Hasseim, "Classification Techniques for Handwriting Difficulties Among Children in Early Stage of Academic Life," *Master thesis*, Universiti Teknologi Malaysia, Malaysia, 2013.

[21] K. H. Chin, "The Screening Tool for Handwriting Difficulty," *Final Year Project Report*, Universiti Teknologi Malaysia, Malaysia, 2013.

[22] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 21 – 27:27, 2011.

[23] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273-297, 1995.