ASM Sc. J., 14, Special Issue 1, 2021 for ICSTSS2018, 48-54

Pattern-Matching for Real-Time Feedback on Pre-Writing Activity

Ahmad Saufi Mohd Amin[1], Norizan Mat Diah[1*], Sofianita Mutalib[1] and Haslizatul Fairuz Mohamed Hanum[1]

[1]*Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia*

Pre-writing activities have been administered to young school children to develop their hands' motor skills as one of the most vital skills in learning to handwrite. There are two main activities for the development of handwriting skill related to pre-writing, which are tracing and copying. For the most part of the development of handwriting skill, copying has been found to be more important than tracing as it requires more effort and more complex coordination skill. As per current practices, children perform pre-writing activities using simple writing instruments, such as pencils and papers or books. The training sessions include repetitive pre-writing activities of drafting straight and cursive shapes weighted from simple to complex difficulties. Children improve their skills through the training sessions based on feedback from instructors. Existing technologies enable pre-writing activities to be carried out using digital technology, which at the same time, provides feedback on the handwriting quality. The usage of online handwriting recognition algorithm was adopted in this study. Input data is entered online by users and will be classified based on the feature extracted upon the pre-processing using Freeman Chain Code (FCC) and pattern matching with regular expression. The algorithm provides feedback on the copying quality and keeps track of the participants' performances. Testing on 120 samples of copying activities, after which the accuracy was found to exceed 70%. Therefore, it could be suggested that real-time feedback via pattern-matching can assist children in digital learning or specifically, pre-writing activities.

Keywords: freeman chain code; pattern-matching; pre-writing

I. INTRODUCTION

Handwriting is a very important fine motor skill learned during the early school years and formal handwriting instruction may begin as early as the kindergarten years. The handwriting activity for children involves the use of hands and are an essential part of fine motor skills; young children engage in reading and writing skills to enhance their visual skills and fine motor skills (Steffani & Selvester, 2009). Handwriting is a form of alphabets composition performed using a pen, pencil, computerised stylus, or other instruments (Nordquist, 2017) controlled by holding it using hands. To attain good handwriting skill, one needs to acquire motor skills while holding an instrument and scribble on a piece of paper or a touch-input device. Evidently, most people have been made to sharpen their hands' motor skills since early childhood as handwriting is a basic tool in many school subjects (Kiefer *et al.,* 2015). The method of implementing pre-writing activities for pre-school children is a great approach to build essential foundation fine motor skills (Diah & Mat Zin, 2011; Husain *et. al.*, 2018, Kwok *et al.*, 2016).

Pre-writing skills are the motor skills for children to be able to develop their ability to hold and move a pencil smoothly and effectively and, therefore, produce legible handwriting. The skills are the fundamental skills children need to develop before they are able to write. According to Diego-Cottinelli and Barros (2010), children practice tracing and drawing with pencils, crayons or even with their fingers to acquire a basic pencil-control skill in the learning phase. Acquiring pre-writing skills is part of most states' pre-kindergarten standards and is considered an important early literacy skill.

*Corresponding author's e-mail: norizan@tmsk.uitm.edu.my

The instruction is distributed across the day school within developmentally appropriate activities (Grisham-Brown *et. al.*, 2006, Csikszentmihalyi, 2014). The pre-writing activities will include hand strengths, directional movement patterns, and effective hand positions, which will facilitate making lines, letters, and shapes. The stroke directions, which are top-to-bottom and left-to-right also constitute as factors in letter formation (Li & James, 2016). Normally, pre-school children should be developmentally ready to form the basic lines (vertical, horizontal, circular, and oblique) that constitute manuscript letters by the time they enter school at the age of five (Mangen & Balsvik, 2016). A pre-writing activity is defined as a whole set of tasks that should be accomplished to practice and assess motor skills for handwriting (Engel *et al.*, 2018).

These skills contribute to children's ability to hold and use a pencil as well as the ability to draw, write, copy, and colour (Freeman & Sanders, 1989; Levin & Bus, 2003). The phase of the pre-writing process emphasises the basic hand movements and muscle development necessary for future penmanship. Two types of pre-writing activities are predominantly used in early education, namely tracing and copying. Tracing involves the connection of dots until designs are formed; this usually does not require much effort. Instead, copying is the greatest contributor to handwriting skill in terms of the formation of letters. This activity necessitates good observation as well as memorising the letter samples, prior to sketching on a blank space, apart from maintaining a high degree of similarity between the original and copied versions (Cameron *et al.*, 2012). Pre-writing copying activity plays an important role in a different field - early education – in which it is used as a method to help young learners develop their hands' motor skills.

However, copying activity is yet to be implemented in technologies because most people are not aware of its importance and hence, they maintain the usage of conventional items, such as papers and writing instruments. As of now, technologies have been made available to a wide range of ages, including young children. The incorporation of handwriting technology that focuses on characters and numbers recognition (for example, computers should ideally be able to recognise the handwriting on documents), is an important technology in the management of digital businesses. Many researches have demonstrated the ability of technology to decipher and classify handwriting input using various techniques and algorithms. Furthermore, the usage of technology as a platform for practising handwriting has been proven to be more effective and appealing to the students (Yang *et. al.*, 2013; Diah & Mat Zin 2016).

Developing writing ability is not only important in building a child's self-esteem and confidence but is also considered an important element to success in school (Feder & Majnemer, 2007). Thus, to ensure the development of good handwriting skill, it is crucial for pre-writing copying to be embraced using modern technology to improve the modern education system

II. RELATED WORK

In this section, previous research on similar domain projects and techniques are discussed to help improve the understanding of the tasks that need to be done in the progress of our project.

A. Handwritten Character Recognition

Recognition is the action and process of identifying something from previous encounters or knowledge. In the case of our project, computers had to recognise patterns, forms, or characters so that they were able to use the input for calculations or classifications and in turn, provide solutions to real-world problems (John *et al.*, 2014). Most recognition systems - which focus on the recognition of numbers and characters - have been developed to aid businesses by scanning handwritten documents for later digital upkeep and processing using computers.

Handwriting recognition systems are used globally owing to the fact the diversity of languages from different countries results in a variety of types of fonts. There are also many different forms of handwriting within the same language, depending on the skills of the writers (Ghods *et al.*, 2016). Hence, handwriting-recognition becomes hard if the features being used are not appropriately selected. However, this does not stop the researchers from developing systems that can recognise and read the different characters in handwritten documents of different languages.

B. Image Pre-processing

Image pre-processing is vital so that the subsequent feature-extraction process can be performed in such a way to minimise the noise in the images. The former step usually involves the normalisation of the images into fixed alignments following segmentation of the images into suitable x*y-pixel grayscale ones with the preservation of their aspect ratios (Babu *et al.*, 2014). Then, the images are converted as binary images to ease the separation of the desired character from the background noise. Figure 1 shows an example of image pre-processing

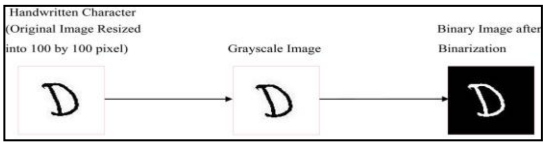


Figure 1. Image pre-processing

C. Feature Extraction

Feature extraction plays the most important part in the process of character recognition and there are many ways to do so. A larger number of extracted features assists to achieve higher accuracy. The purpose of feature extraction is to extract a set of features because their function is to maximise the recognition rate with the fewest number of elements. One of the commonly-used feature-extraction methods in character-recognition is a statistical feature. Statistical features represent the image of a character in terms of statistical distributions of its points in a two-dimensional plane.

The five types of statistical features, which are frequently used in character-recognition are zoning, projections, profiles, crossings, and distances:

i. Zoning

In zoning, the images are divided into x*y zones or areas. Feature vectors are formed when the features are extracted from each of the zones (Chherawala *et al.*, 2015).

ii. Project Histogram

This feature counts the number of pixels present in every row and column of the character, after which it presents them as a graph on the number of pixels (Suhasini *et al.*, 2017).

iii. Profiles

Profiling involves the counting of the number of pixels as distances between the boundaries of the box containing the image of the character and the edges of the character in all four perspectives: up, down, left, and right. It is used to describe the outer shapes of the characters.

iv. Crossings and Distances

Crossings count the number of pixels beyond the image of the character, starting from the background (which is the outer part of the character's image) to the foreground (which is usually the inner enclosed part of the character's image) along the horizontal and vertical lines. Subsequently, dressings are performed to calculate the distances of pixels between the boundaries of the box to the outer parts of the character, also along the vertical and horizontal lines.

v. Freeman Chain Code

The Freeman Chain Code (FCC) has been widely used in motion detection for handwriting ('Arif *et al.*, 2015). Basically, given a set of points representing an image, FCC locates the next point using the current point as the centre and assigns a direction to which the next point lies. Eight directions are usually used in the FCC (Althobaiti *et al.*, 2017), with numbers assigned to each.

D. Regular Expression-matching

Regular expression-matching is used to find the regularities or frequent-employed sequences of any character (Câmpeanu *et al.*, 2003). Regular expression provides many syntaxes that can be used in a variety of ways depending on one's creativity. Recently, regular-expression has not been frequently utilised in handwritten character-recognition, but with the correct feature input, such as a string of input, regular-expression can prove to be useful.

III. METHODOLOGY

For this project, data was collected from randomly-selected children of age 3 to 5 years using the drawing panel of the early version of the prototype. The input was the coordinates of the points drawn on the panel as shown in Figure 2.

Figure 2. Java drawing panel

Eight patterns of activities were provided in the prototype, each of which was assigned a number as shown in Figure 3.

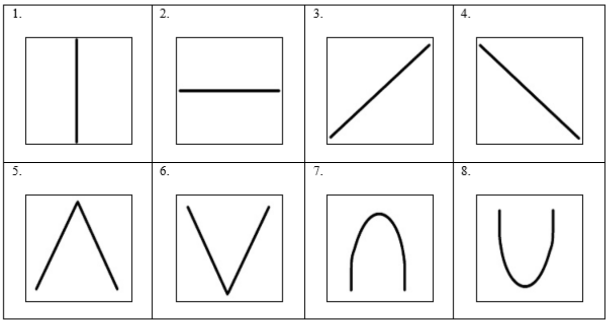


Figure 3. Eight patterns of activities

The input was drawn on the panel via the mouse-drag function. In addition, the input used for pre-processing was in the form of coordinates of the drawing. These coordinates were saved as an array list, which was easier to be managed regardless of the number of coordinates received as presented in Figure 4.

[205,92][205,94][205,96][205,98][205,100][205,102][205,104][205,106][205,107][205,110][205,112][204,113][204,120][204,124][204,126][204,128][203,133][203,136][203,139][203,141][203,143][202,146][202,149][202,152][201,155][201,157][201,159][201,161][201,163][201,165][201,168][201,169][201,172][201,173][201,177][201,181][201,183][201,187][201,190][201,192][201,195][201,196][201,202][201,205][201,2

Figure 4. Sample input

During pre-processing, the input was reduced to a resolution of 10 by 10 pixels only for the existing coordinates of the drawings. Subsequently, the input was significantly reduced and normalised using an equation that was created solely for pre-processing.

$$\text{Normalised value} = \frac{(x,y)\ of\ each\ point}{(distance\ x \cup distance\ y)/10} \qquad (1)$$

The equation works by finding the region in the drawing panel, which contains the drawing. First, the distance of x is calculated by finding the difference between maximum and minimum values of x while the same goes for the distance of y. Then, the longest distance becomes the length of each square side. The square box around the drawing was set as the region for the 10 by 10 segmentation. The drawing will always appear at the left side of the segmentation box. This ensures that the segmentation will always work only in the region in which the drawing appears. Figure 5 shows the sample pattern 1 that is being reduced to a resolution of 10 by 10 pixels.

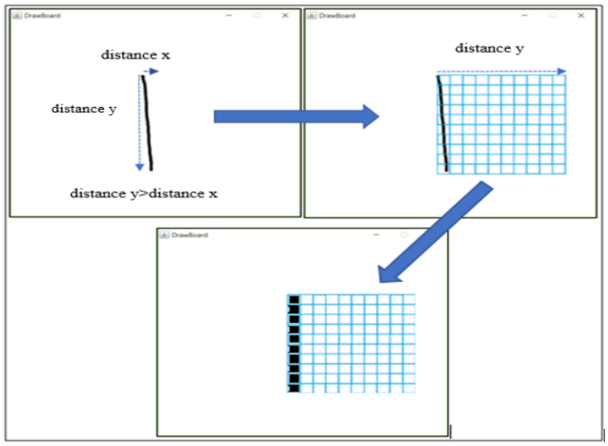


Figure 5. Segmentation of sample pattern 1

Upon calculating all the points using the given equation, the amount of data was significantly reduced as only the meaningful or distinct points were retained, refer Figure 6.

[10,4] [10,5][10,6][10,7][10,8][10,9][10,10][10,11][10,12][10,13]

Figure 6. Normalised and reduced data from the given sample

The features were extracted using the FCC and built into a string. The single numeric value was represented by a centre point and the next point was representing the position or direction from the centre. There are eight directions used in the feature extraction as shown in Figure 7.

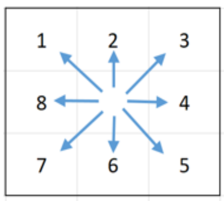


Figure 7. Eight directions of FCC

The distances between the points were calculated according to their directions in the indices of the array list from the current as of the centre to the next one. There were three steps in assigning directions to the coordinates. Prior to that, the pairs of x and y coordinates were separated and passed through the loop, which divided the nine squares of the FCC into horizontal divisions (x-values) and vertical divisions (y-values).

The horizontal division comprised three rows: above, middle, and below; while the vertical division comprised three columns: left, middle, and right. Following the identification of both horizontal and vertical divisions, a value of direction was assigned from the centre point. The order of the directions started from the first coordinate to the last one accordingly. Figure 8 shows the different ways in which the directions of the current point are possibly assigned, which then stored as FCC vectors.

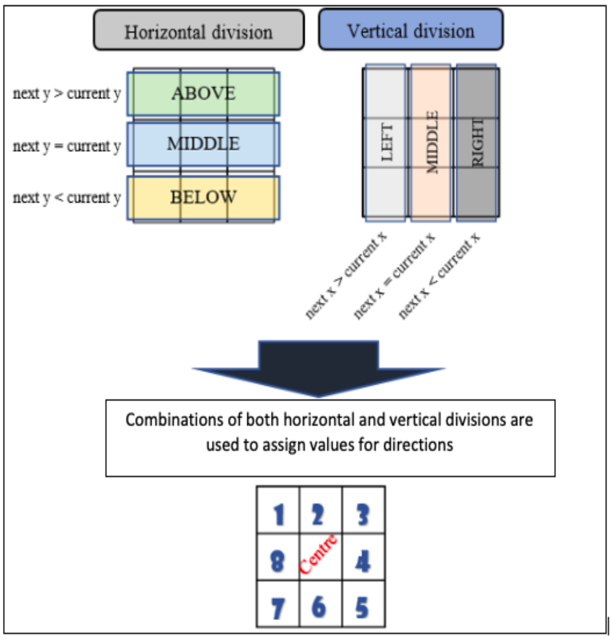
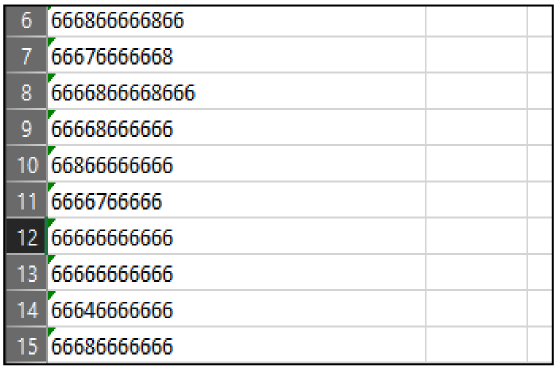


Figure 8. Point-direction assignment

All the FCCs of the given sample were appended into strings; the data would be identified based on the matching. The above-mentioned strings became the dataset of the prototype, which consists of 640 data points from eight patterns as shown in Figure 9.



6	666866666866		
7	66676666668		
8	6666866668666		
9	66668666666		
10	66866666666		
11	6666766666		
12	66666666666		
13	66666666666		
14	66646666666		
15	66686666666		

Figure 9. Selected data from the sample

A regular expression for each identified pattern was built using an online regular expression-checker tool. A total of 640 data points was gathered, or 80 data points for each pattern used in the building of regular expressions. Using a matching algorithm in the java library along with the constructed regular expressions, the prototype was able to distinguish the type of patterns drawn.

IV. RESULTS OF TESTING

The users (children) around ages of 3, 5, and 6 years were assigned as the testers. For each pattern, the data was collected from up to 5 trials with 120 sample data. Table 1 shows the recognition accuracy for results obtained through the prototype. The accuracy shows how accurate each pattern is written by each child. The accuracy of each pattern is calculated from the correct number of trials detected by the algorithm.

Table 1. Results from the prototype

Child	Pattern No.	Accuracy
	1	100%
	2	80%
	3	100%
3-year-old	4	100%
	5	60%
	6	60%
	7	80%
	8	80%
	1	100%
	2	80%
	3	100%
	4	100%
5-year-old	5	80%
	6	80%
	7	100%
	8	100%

	1	100%
	2	100%
	3	100%
6-year-old	4	100%
	5	100%
	6	80%
	7	100%
	8	100%

Meanwhile, Figure 10 depicts the average accuracy of each pattern, considering different age ranges.

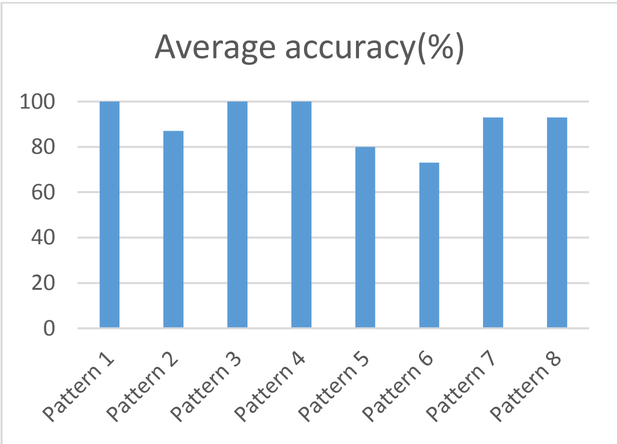


Figure 10. Average accuracy of each pattern

V. CONCLUSION

This research was conducted in the form of pre-writing copying activities so that the computer was able to automatically recognise the children's answers and provide results with reference to the chosen assignments. Eight assignments were stipulated in the prototype, all of which involved the drawing of a stroke. Real-time feedback on the performance of the eight different stroke patterns was made available with reference to the features extracted after pre-processing via the employment of Freeman Chain Code and regular expression-matching. The average accuracy of the application was up to 70%, which was relatively high. The accuracy of the results has shown that this algorithm and method can be used to develop writing-learning software for children. In addition, this study promotes the usage of digital technology for the enhancement of handwriting skills for children.

VI. ACKNOWLEDGEMENT

This work is supported by UiTM Internal Research Grants Scheme (LESTARI) (600-IRMI/DANA5/3/LESTARI (1210/2016)) from the Research Management Centre (RMC), Universiti Teknologi MARA, Shah Alam, Selangor. The appreciation also goes to Faculty of Computer and Mathematical Sciences (FSKM), Universiti Teknologi MARA, Shah Alam, Selangor for giving a moral support in the production of this paper.

VII. REFERENCES

Althobaiti, H, Shah, K & Lu, C 2017, 'Isolated handwritten Arabic character recognition using freeman chain code and tangent line', in Proceedings of the International Conference on Research in Adaptive and Convergent Systems (RACS '17), 20–23 September, Krakow Poland.

'Arif, M, Hassan, H, Nasien, D & Haron, H 2015, 'A review on feature extraction and feature selection for handwritten character recognition', International Journal of Advanced Computer Science and Applications, vol. 6, no. 2, pp. 204-2012.

Babu, RU, Venkateswarlu, Y & Chintha, A 2014, 'Handwritten digit recognition using k-nearest neighbour classifier', in World Congress on Computing and Communication Technologies, (WCCCT 2014), 27 February-1 March, Trichirappalli India.

Cameron, CE, Murrah, WM, Bell, LH, Worzalla, SL, Grissmer, D & Morrison, FJ 2012, 'Fine motor skills and executive function both contribute to kindergarten achievement', Child Development, vol. 83, no. 4, pp. 1229–1244

Câmpeanu, C, Salomaa, K & Yu, S 2003, 'A formal study of practical regular expressions', International Journal of Foundations of Computer Science, vol. 14, no. 06, pp 1007–1018

Chherawala, Y, Roy, PP & Cheriet, M 2015, 'Feature set evaluation for offline handwriting recognition systems: application to the recurrent neural network model', IEEE Transactions on Cybernetics, vol. 46, no. 10, pp. 2825–2836.

Csikszentmihalyi, M 2014, Applications of flow in human development and education, Springer, Netherlands.

Diego-Cottinelli, BD & Barros, B 2010, 'TRAZO: a tool to acquire handwriting skills using, tablet-pc devices', in The 9th International Conference on Interaction Design and Children, 9-11 June, Barcelona, Spain.

Diah, MN & Mat Zin, NA 2016, 'Tracing activity engine for learning writing the jawi character', Journal Asia-Pacific Journal of Information Technology and Multimedia, vol. 5, no. 1, pp. 61-69.

Diah, MN & Mat Zin, NA 2011, 'Digital training tool framework for jawi character formation', in Visual Informatics: Sustaining Research and Innovations, Lecture Notes in Computer Science, 9-11 November, Bangi Selangor.

Engel, C, Lillie, K , Zurawski, S & Travers, BG 2018, 'Curriculum-based handwriting programs: a systematic review with effect sizes', American Journal Occupational Therapy, vol. 72, no. 3, pp. 7203205010p1–7203205010p8.

Feder, KP & Majnemer A 2007, 'Handwriting development, competency, and intervention', Developmental Medicine and Child Neurology, vol. 49, no. 4, pp. 312-317.

Freeman, E & Sanders, T 1989, 'Kindergarten children"s emerging concepts of writing functions in the community', Early Childhood Research Quarterly, vol. 4 pp. 331-338.

Ghods, V, Sohrabi, M & Hosseini S 2016, 'Recognition of online Farsi handwriting based on freeman chain code using hidden Markov model', in 4th International Symposium on Computational and Business Intelligence, (ISCBI 2016), 5-7 September, Olten, Switzerland.

Grisham-Brown, J, Ridgley, R, Pretti-Frontczak, K, Litt, C, & Nielson, A 2006, 'Promoting positive learning outcomes for young children in inclusive classrooms: a preliminary study of children's progress toward pre-writing standards', Journal of Early and Intensive Behavior Intervention, vol. 3, no. 1, pp. 171-183.

Husain, NH, Diah, NM & Hanum, HFM 2018, 'The relevance of Freeman Chain Code for copying activities', in IEEE Conference on e-Learning, e-Management and e-Services, IC3e, 16–17 November, Kuching Serawak Malaysia.

John, J, Pramod, KV, Balakrishnan, K & Chaudhuri B 2014, 'A two stage approach for handwritten Malayalam character recognition', in Proceedings of International Conference on Frontiers in Handwriting Recognition, ICFHR (2014), 1-4 September, Heraklion, Greece.

Kiefer, M, Schuler, S, Mayer, C, Trumpp, NM Hille, K & Sachse, S 2015, 'Handwriting or typewriting? the influence of pen- or keyboard-based writing training on reading and writing Performance in preschool children', Advances in Cognitive Psychology, vol. 11, no. 4, pp.136–146.

Kwok, K, Ghrear, S, Li, V, Haddock, T, Coleman, P, Birch, S A 2016, Children can learn new facts equally well from interactive media versus face to face instruction, Frontiers in Psychology, vol. 7, pp. 1603.

Levin, I & Bus, AG 2003, 'How is emergent writing based on drawing? Analyses of children"s products and their sorting by children and mothers', Developmental Psychology, vol. 39, pp. 891-905.

Li, JX & James, KH 2016, 'Handwriting generates variable visual output to facilitate symbol learning', J. Exp. Psychol.: Gener., vol. 145, no. 3, pp. 298-313.

Mangen, A & Balsvik, AL 2016, 'Pen or keyboard in beginning writing instruction? Some perspectives from embodied cognition', Trends in Neuroscience and Education, vol. 5, no. 3, pp. 99-106.

Nordquist, R 2017, Handwriting glossary of grammatical and rhetorical terms, viewed 13 July 2018, <https://www.thoughtco.com/handwriting-definition-and-examples-1690831>.

Steffani, S & Selvester, PM 2009, 'The relationship of drawing', Writing, Literacy, and Math in Kindergarten Children, Reading Horizons, vol. 49, no. 2, pp. 125-142.

Suhasini, PS, Sri Rama Krishna, K & Murali Krishna, IV 2017, 'Content based image retrieval based on different global and local color histogram methods: a survey', Journal of The Institution of Engineers (India): Series B, vol. 98, no. 1, pp. 129–135.

Yang, CY, Chiu-Huang, TY & Wu, YT 2013, 'Effectiveness of learning Chinese character using tablet technology', in 7th International Conference on Universal Access in Human-Computer Interaction, (UAHCI 2013), 21-26 July, Las Vegas, USA.