International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Introduction to Writing for Children with A K-Nearest Neighbor Approach

Rini Wongso*[1], Violitta Yesmaya[2], Suryanto Wijaya[3], Arya Dwi Kurniawan[4], Mochamad Harsya Joe Andaru[5]

Submitted: 23/12/2023 **Revised**: 29/01/2024 **Accepted**: 07/02/2024

Abstract: This paper presents the design, development, and evaluation of an innovative mobile application targeted at promoting writing skills among young learners. Learning to write is a basic skill for children at an early age, but acquiring this competency is often a challenging process. The primary aim of this research was to design an application that is child-friendly, interactive, and effective in teaching children how to write. The application was built with a specific focus on aiding children in the journey of learning writing skills starting from individual characters and gradually progressing to words. So that this learning can make it easier for children to understand every letter and word that they can learn to write. The application utilizes the robustness of K-Nearest Neighbor (K-NN) algorithm for the recognition of children's handwriting. The K-NN algorithm was employed as the core engine to recognize and assess the child's handwriting and provide immediate feedback. Based on the results of preliminary testing shows promising results, with improved writing skills and high engagement levels among a group of test students.

Keywords: *K-NN, K-Nearest Neighbor, Handwriting Recognition*

1. Introduction

As we move into an increasingly digital era, technology plays a significant role in shaping educational landscapes globally. Despite various advancements in the field, a noticeable gap persists in the realm of digital learning aids for writing, particularly in the context of Indonesian young learners. According to the survey, Indonesian literacy rate reach 96% in 2020 [1]. However, the 4% illiteracy rate still covers around 11 million people. While the number is getting lower yearly, still raises a concern, which underlines the urgency of innovatively addressing this issue. Existing traditional pedagogies often fail to engage and stimulate the digital-native Gen Z, characterized by their immersion in technology from a very early age. Therefore, a technology-based solution is needed to improve their illiteracy, both in reading or writing.

Moreover, according to Statista [2], smartphone penetration rate in Indonesia reached 76.26% of total population in 2021, and it is predicted to grow higher each year. The forecast can be seen in the following Fig. 1. The data represents the enormous potential in and reach of digital learning platforms.

By developing an application specifically designed for Indonesian young learners to learn writing, we leverage the widespread accessibility of smartphones. It can provide a more engaging and effective learning environment, ultimately aiding in bridging the literacy gap, preparing students for future academic pursuits to increase literacy rates in the country.

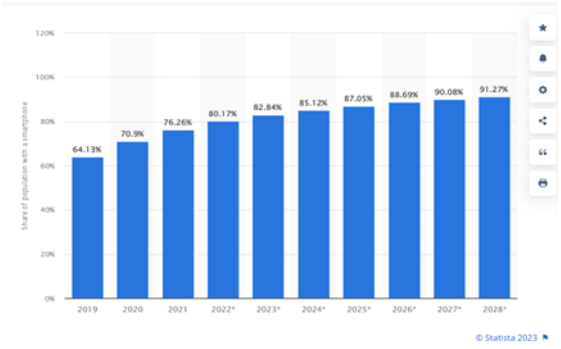

Fig.1. Smartphone Penetration in Indonesia

Among the smartphones available in the market, we can conclude that Android plays a significant role in the Indonesian market as described by GlobalStats (2023) in the following Fig. 2.

[1] *Computer Science Department, School of Computer Science, Bina Nusantara University*
ORCID ID : 0009-0001-4581-5380
[2] *Computer Science Department, School of Computer Science, Bina Nusantara University*
ORCID ID : 0000-0002-8000-0347
[3,4,5] *Computer Science Department, School of Computer Science, Bina Nusantara University*
**Corresponding Author E-mail: rwongso@binus.ac.id*

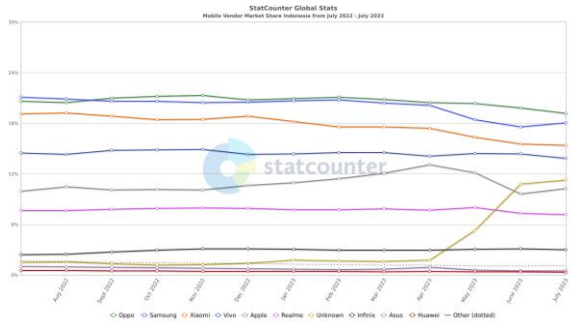

Fig.2. Mobile Vendor Market Share Indonesia

In this era of digital transformation, people begin to rely on technology more, hence handwriting is increasingly overshadowed by the convenience and speed of typing. However, learn to write using hands is in fact important as it has many benefits, especially for children [3] [4], where handwriting can enhance children's fine motor skills and improve their motor-perceptual skills, hence leading to the faster development of children's literacy and numeracy. According to [5], Dr. Sugiyanto, S.IP., an expert staff of Elementary Education Directorate of the Ministry of Education and Culture of the Republic of Indonesia, emphasizes the importance of learning handwriting in early years, where it is beneficial for brain development, it is also affecting positively on mental growth, hence supporting both motor and social skill. Moreover, some studies have also shown that the act of writing by hand can lead to better understanding and retention of material in both primary and secondary school's studies, as act of writing by hands require more cognitive engagement and activities in different areas of the brain [6]. In therapy, handwriting is also often used as a tool to help the individuals improve their motor skill and visual-motor coordination.

There are actually many Android applications available to assist young learners on how to write [7], such as: iTrace, hip hop hen: abc letter tracing, Dexteria Jr., Trace it, try it, and many more. However, although these applications can assist the children to learn tracing the alphabets and numbers, the application is only available in English and do not support handwriting recognition. Therefore, this study aims to create an application to introduce children to write.

2. Material and Methods

A survey is conducted to 100 primary 1 students, 50 teachers and parents in Indonesia area to obtain more details on how teachers introduce alphabets to students, how the students learn, and how the parents assist their children in learning. Based on the result, we conclude the following: (1) students are introduced to alphabets since kindergarten, however, they will learn to write since primary school; (2) students are first introduced to the shapes of each alphabet character, how to pronounce it, and then write them with assistance up until they can do it themselves; (3) the optimal age for children to learn on writing is: age 4 to 5, while writing intensively is more optimal for children ages > 5; (4) the order on how to complete an alphabet is important, both for lower and upper letter case character, and overall to achieve this, there is a need of tools that can encourage students to train to write.

Therefore, this research proposed to create an application to assist students in learning to write alphabets, both in lower and upper letter case. With three main features: writing lower letter case, writing upper letter case, and quiz to overall evaluate user performance. K-NN Algorithm will be implemented in the system to assist recognizing the alphabet character written by the user, as this simple algorithm is often used to assist this problem [8]. The data set used to train the model is obtained from the surveys where we collect children's handwriting.

The proposed method is as follows:

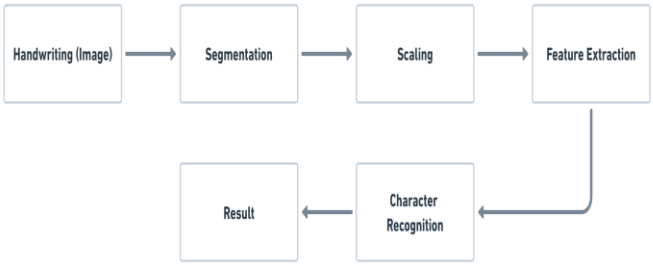

Fig.3. Proposed Methods

First, the digital handwriting (image) will be processed through Segmentation. Segmentation will divide the input handwriting into separate characters as described in Fig. 4 below. The purpose is to let the system recognize the character more accurately.

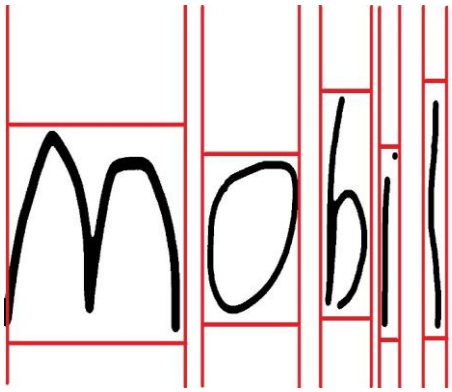

Fig.4. Segmentation Result

Segmentation results are then scaled to 25 x 25 pixels using OpenCV library, as the training data set, to ensure that there are no significant differences.

Writing recognition is done by using the K-Nearest Neighbor method to classify characters in writing. Here is how the K-Nearest Neighbor method works in this application (see Fig. 5).

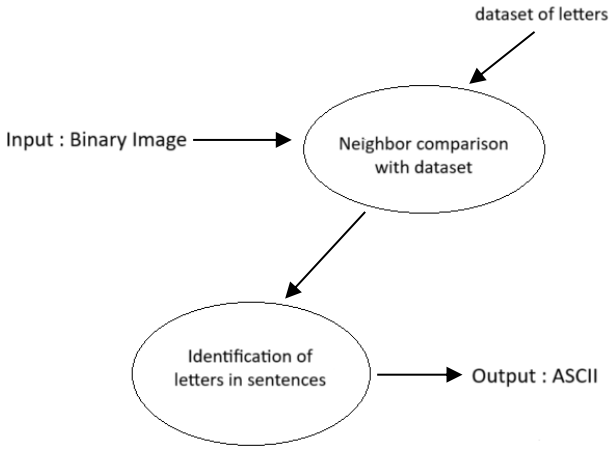

Fig.5. Diagram of how K-Nearest Neighbor works in the application.

Then, features are extracted from the characters, where the data is represented using binary value, 0 to represent black pixel, and 1 to represent white pixel.

Scaling is the next stage after the segmentation process. By using the grayscale method, which is a technique used to change an existing image to grey by taking the RGB (red, green, and blue) value for each pixel and making a unit value that will be the output. This technique is used to make searching based on content easier. The use of colour as a feature in content is commonly used in searching. This grayscale process is also used to get the strongest colour intensity, white [9].

Scaling will be described in Fig. 6. these are several stages in scaling:

(a) Initial image before processing

(b) Grayscaling on images

(c) Remove noise with gaussian blur

(d) Thresholding

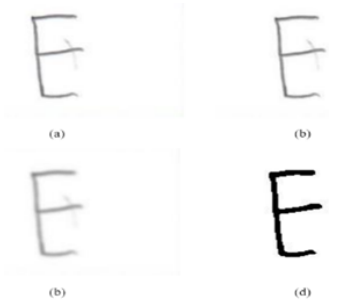

Fig.6. Stages in Scaling image preprocessing

The dataset to be processed requires a large data classification that includes all the required information. So, data mining is needed to be able to classify data [10]. Classification technique is needed to assistance from data mining which can determine the value of categorical variables by building a model based on one or more numerical or categorical variables. Classification is a data mining task to predict the value of a categorical variable by building a model based on one or more numeric or categorical variables.

Figure 7 below describes samples of features from each character.

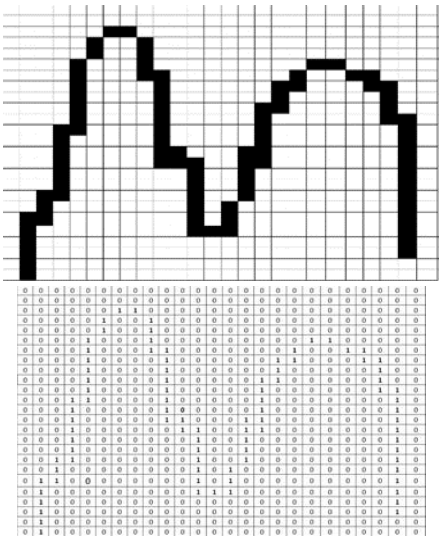

Fig.7. Character pattern from segmented from the handwriting (left) and binary pixel extracted from pattern (right)

The K-NN algorithms will then be used to recognize the character. Then the result will be obtained.

K-NN Algorithms

K-NN (K-Nearest Neighbor) Algorithm is a classification method towards new objects according to the K number of neighbors nearest to it [11] [12]. This algorithm is a supervised learning algorithm where results from query instances are classified according to most categories (classes) in K-N (see Fig. 8).

According to Gorunescu [13], some inputs such as training data set, metrics (distance) to compute object similarity between objects, and K value that describes number of categories of objects [14], are needed to implement K-NN. The steps are as follows: (1) calculate distance of all training data and new object, (2) Identify closest objects based on K values, by ordering training objects considering the distance calculated in the first step, and (3) gives a label based on the nearest neighbor that often appears from the value of K.

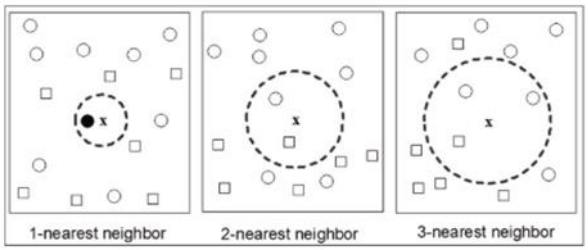

Fig.8. Nearest Neighbor Example with K = 1, 2, 3

3. Result and Discussion

The writing that has been collected from respondents is processed into a dataset that can be used in writing recognition, so the next stage is to test the dataset that has been created. Testing is carried out by dividing the existing data into test data and dataset. The test data is recognized using the dataset as a comparison to find out whether the writing can be recognized or not. Results from testing the dataset using K-Nearest Neighbor with 30 test data and 25 datasets.

From the test results in table 1, the average accuracy in recognizing uppercase letters reached 81.96% and lowercase letters was 79.38% and the total for both was 80.67%.

Table 1. Trial test data

Letter	Succeed	Not Successful	Accuracy (%)	Letter	Succeed	Not Successful	Accuracy (%)
A	26	4	86	a	21	9	70
B	24	6	80	b	27	3	90
C	25	5	83	c	28	2	93
D	27	3	90	d	28	2	93
E	25	5	83	e	26	4	86
F	26	4	86	f	26	4	86
G	22	8	73	g	6	24	20
H	25	5	83	h	27	3	90
I	24	6	80	i	28	2	93
J	23	7	76	j	25	5	83
K	25	5	83	k	26	4	86
L	26	4	86	l	25	5	83
M	23	7	76	m	26	4	86
N	24	6	80	n	26	4	86
O	25	5	83	o	25	5	83
P	24	6	80	p	24	6	80
Q	25	5	83	q	23	7	76
R	25	5	83	r	24	6	80
S	25	5	83	s	27	3	90
T	26	4	86	t	20	10	66
U	25	5	83	u	13	17	43
V	24	6	80	v	26	4	86
W	25	5	83	w	26	4	86
X	26	4	86	x	22	8	73
Y	22	8	73	y	23	7	76
Z	25	5	83	z	24	6	80

There are several experiments that have been carried out to determine the level of accuracy of the words being written.

In this experiment, several comparisons were made, namely with writing that did not consider the distance between letters, writing that was not scaled, and writing that had undergone spacing and scaling. This experiment can be seen in Table 2 for more details.

Table 2. Word Experiment Results

Case	No scaling		No spaces		Spacing and scaling	
	Readable	Unreadable	Readable	Unreadable	Readable	Unreadable
Apel	34	21	40	15	44	11
Pisang	32	23	38	17	41	14
Bola	26	29	32	23	38	17
Kucing	31	24	36	19	40	15
Sepatu	27	28	33	22	35	20

From the experimental results there are differences between experiments using the three conditions that have been determined. The following Table 3 is a comparative table of experimental results on 275 writing samples.

Table 3. Results of Experiments with Scaling and Calculation

Condition	**Not successful**	**Succeed**	**Present**
No calculations were made	125	150	54,54%
No scaling performed	96	179	65,09%
Performed scaling and calculations	77	198	72%

The figures below is an example of the results of the implementation of experiments that have been carried out on the system that has been created. Figure 9 below describes the features offered by the system, consisting of “Writing Lower Letter Case”, “Writing Upper Letter Case”, and “Quiz”. The application is served with Indonesian Language as the target is Indonesian young learners.

Fig. 9. Menu page (right)

In writing either lower or upper letter case, user may choose the alphabet (‘a’ - ‘z’ or ‘A’ - ‘Z’) to focus on (see Fig. 10). The user then may learn to write the character and obtain feedback of whether they write it correctly or incorrectly (see Fig. 11). There will be animation that shows the order to write the character, ensuring the children follows the correct way.

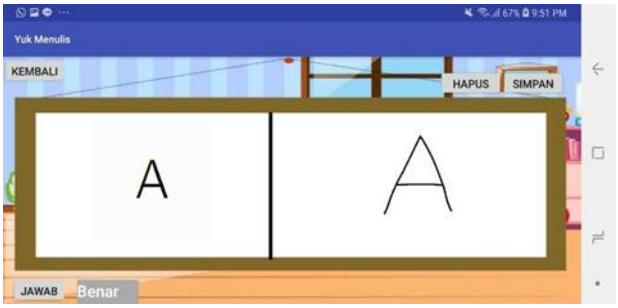

Fig.10. Alphabets available

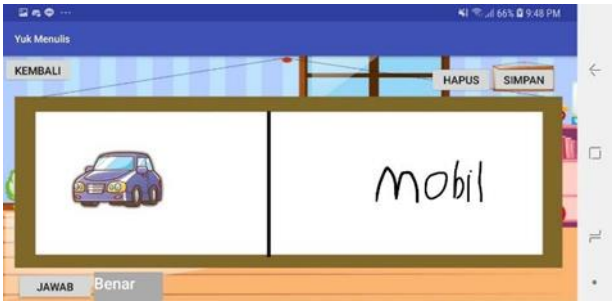

Fig.11. Learn Writing

After the user has successfully learned to write all the alphabet, then the user can choose the Quiz menu. This quiz will be evaluating whether the user can correctly write the answer according to the image given, as described in Figure 12.

Fig.12. Quiz

4. Conclusions

The test is done to students around 5-6 years old in at a kindergarten school in Indonesia with a total of 40 students. We observed that they are very interested in using this system to learn writing. They understand how to use the application with a little assistance. In the observation we conclude that students are familiar with the alphabet, however not all students can write well, as only 1 student can answer the quiz correctly. We also gathered some feedback from teachers and parents.

This research explains the use of the K-Nearest Neighbor method in recognizing handwriting. From the test results, it can be concluded that the K-Nearest Neighbor method is capable of recognizing writing with results of letter recognition accuracy of 80.67% and children's writing of 72%.

From the results of testing children's writing, the level of accuracy can be increased by carrying out additional calculations on the spaces in the word segmentation section into letters, as well as adjusting the data size to the dataset used. Apart from that, by adding a better amount of training data. The application is able to carry out the desired tasks by recognizing letters and writing.

References

[1] "knoema," 2023. [Online]. Available: https://knoema.com/atlas/Indonesia/topics/Education/Literacy/Adult-literacy-rate.

[2] "statista," 2023. [Online]. Available: https://www.statista.com/statistics/321485/smartphone-user-penetration-in-indonesia/.

[3] R. W. Wiley and B. Rapp, "The Effects of Handwriting Experience on Literacy Learning," Psychological Science, vol. 32, no. 7, pp. 1086-1103, 2021.

[4] "The effects of handwriting experience on functional brain development in pre-literate children," Trends in Neuroscience and Education, vol. 1, no. 1, pp. 32-42, 2012.

[5] T. Theresia, "APP Sinarmas," Majalah CIA, 24 May 2023. [Online]. Available: https://asiapulppaper.com/-/handwriting-key-competence-for-generation-z-#:~:text=A%20study%20published%20in%20the,of%20children's%20literacy%20and%20numeracy.. [Accessed 4 June 2023].

[6] OT Dude, "OTDUDE," OT Dude Occupational Therapy, APC, 22 January 2023. [Online]. Available: https://www.otdude.com/patients/handwriting-vs-typing-in-child-development-and-role-of-school-based-occupational-therapy/. [Accessed 6 June 2023].

[7] "IDN Times," 2 Aug 2019. [Online]. Available: https://www.idntimes.com/tech/trend/abraham-herdyanto/aplikasi-belajar-menulis-tangan-untuk-anak?page=all. [Accessed 4 June 2023].

[8] U. R. Babu, Y. Venkateswarlu and A. K. Chintha, "Handwritten Digit Recognition Using K-Nearest Neighbour Classifier," in 2014 World Congress on Computing and Communication Technologies, Trichirappalli, 2014.

[9] Debasish Biswas, Amitava Nag, Soumadip Ghosh, Arindrajit Pal, Sushanta Biswas, Snehasish Banerjee, Anjan Pal, "NOVEL GRAY SCALE CONVERSION

TECHNIQUES BASED ON PIXEL DEPTH," Journal of Global Research in Computer Science, vol. 2, no. June, 2011.

[10] Aman Kataria, M. D. Singh, "A Review of Data Classification Using K-Nearest Neighbour," International Journal of Emerging Technology and Advanced Engineering, vol. 3, no. 6, June 2023.

[11] U Ravi Babu, Aneel Kumar Chintha, Dr. Y Venkateswarlu, "Handwritten Digit Recognition Using Structural,," I.J. Information Engineering and Electronic Business, vol. 1, no. MECS, pp. 62-68, 2014.

[12] N. Hidayati and A. Hermawan, "K-Nearest Neighbor (K-NN) algorithm with Euclidean and Manhattan in classification of student graduation," Journal of Engineering and Applied Technology, vol. 2, no. 2, pp. 86-91, 2021.

[13] F. Gorunescu, Data Mining: Concepts, models and techniques, vol. 12, Springer Science & Business Media, 2011.

[14] Mohammad Imrul Jubair, Prianka Banik, "A Simplified Method for Handwritten Character Recognition from Document Image," International Journal of Computer Applications, vol. 51, no. 14, August 2012.