

Using Tesseract OCR to assist KG pupils to learn writing

Ephrim Lawrence

Abstract

The paper explores the use of Optical Character Recognition (OCR) technology to assist kindergarten pupils in learning how to write English characters. OCR is a technology that allows computers to recognize printed or handwritten characters and convert them into machine-encoded text, allowing it to be analyzed by a computer. The software presented in this paper provides a platform for pupils to practice writing English alphabets on a computer device, and the OCR technology recognizes and evaluates the accuracy of the pupils' writing. The software also gives immediate feedback to the pupils, enabling them to identify and correct errors in their writing.

1 Introduction

Handwriting is a fundamental skill for literacy development, particularly in the early years of education. For young learners, developing legible and fluent handwriting skills is a critical component of academic success [Berninger et al., 2002, Dinehart and Manfra, 2013, Longcamp et al., 2005, Longcamp et al., 2006]. However, teaching young children to write can be challenging, particularly for pupils with different learning abilities. In recent years, there has been growing interest in the use of technology to support early literacy development, including the use of Optical Character Recognition (OCR) technology to assist pupils in learning how to write English characters. OCR technology has been used in a variety of applications, including docu-

ment scanning and digitization, automated data entry, and language translation. In the context of literacy education, OCR technology has been used to evaluate and provide feedback on handwriting skills.

For example, a study by [Alvaro et al., 2010] conducted with professional school teachers and parents revealed that the handwriting learning software (developed by the authors) for children, which uses a stylus on a touch screen monitor, is appropriate for children's comprehension level and provides a user-friendly, functional, reliable, and correct alternative way of teaching basic handwriting.

While previous studies have demonstrated the potential benefits of OCR technology and other software programs in supporting early literacy development, there is a need for further research to explore ef-

efficient OCR algorithms for supporting the development of handwriting skills in young learners. Therefore, this study aims to develop software that uses the Tesseract OCR engine [Smith, 2007] to assist kindergarten pupils in learning how to write English characters.

The paper is organized into three (4) main sections; Section 1 is the introduction; Section 2 discusses related works on the topic; Section 3 discusses how the systems work and its various components; Section 4 ends the paper by discussing some limitations in the proposed system as well as what can be improved in future works on the topic.

2 Related Works

Handwriting is a fundamental skill for literacy development, particularly in the early years of education. For young learners, developing legible and fluent handwriting skills is a critical component of academic success. However, teaching young children to write can be challenging, particularly for pupils with different learning abilities. In recent years, there has been growing interest in the use of technology to support early literacy development, including the use of Optical Character Recognition (OCR) technology to assist pupils in learning how to write English characters.

[Alvaro et al., 2010] conducted a study with professional school teachers and parents revealed that the handwriting learning software for children, which uses a stylus on a touch screen monitor, is appropriate for children’s comprehension level and provides a user-friendly, functional, reliable, and cor-

rect alternative way of teaching basic handwriting.

In [Tang et al., 2006] study, the authors presented a web-based Chinese handwriting education system that provides automatic feedback and analysis to assist students in learning how to write Chinese characters. The system uses handwriting recognition and machine learning techniques to evaluate students’ handwriting and provide immediate feedback on the accuracy and fluency of their writing. The study found that the system was effective in improving students’ handwriting skills and increasing their motivation and engagement in the learning process.

[Longcamp et al., 2006] compares the effects of handwriting and typing on letter recognition. The study found that handwriting leads to better recognition of letter orientation than typing, especially after some time has passed since learning. The study suggests that the motor activity involved in handwriting helps to form more stable letter representations in memory.

A study by [Zin et al., 2021] proposed an offline self-learning application to help primary-level students in developing countries learn written English and basic calculations. Handwritten characters or words written on tablets were saved as input images, and a proposed character segmentation method is used to recognize segmented characters with the help of a Convolutional Neural Network (CNN). Handwritten data was collected from primary-level students to build the dataset. The proposed system achieved 95.6% accuracy on 1000 randomly selected words and 98.7% for each character, making it a promising solution for improving access to quality education in de-

veloping countries.

These studies demonstrate the feasibility and effectiveness of using OCR to assist young people to learn how to write. However, there are still some challenges and limitations that need to be addressed. For example, OCR accuracy may vary depending on the quality of the images, the complexity of the characters, and the diversity of the handwriting styles [Asad et al., 2016, Faizullah et al., 2023, Ye and Doermann, 2013]. OCR feedback may also not be sufficient or appropriate for all learners, as different children may have different needs and preferences.

3 Proposed System

The proposed system is a web-based application that enables pupils to learn the writing of English alphabets. The system uses Tesseract [Smith, 2007] as the Optical Character Recognition (OCR) engine as well as the HTML Canvas API for the drawing functionality.

When the user opens the app, randomly generated English alphabets are presented to the user. After drawing the alphabet, the user then clicks on the *Check Answer* button to get feedback on whether what was drawn is correct or not. If it matches the generated alphabet, congratulations confetti is shown (this makes the application fun to use), else failure confetti is shown to the user. Figure 1 shows the state of the web app on the first start. Figure 2 shows the state of the app if the alphabet written by the user matches the generated alphabet whilst Figure 3 shows the state of the app if the written alphabet is different from the

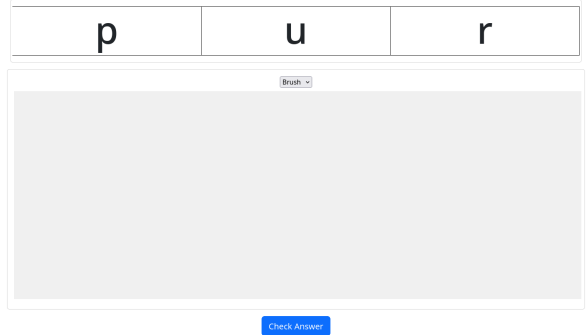


Figure 1: UI of the app when it is opened

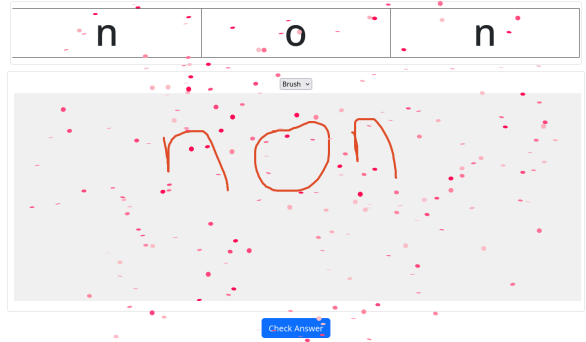


Figure 2: Success confetti when the written alphabets is correct

generated one.

Figure 4 shows the flowchart of the application.

3.1 Architecture

The proposed system is divided into three (3) main components; Front-end, Backend and the OCR component. The front end communicates with the backend component via a RESTful API. When the user clicks on the *Check Answer* button, the drawing is converted to an image and then encoded as a Base64 string. The Base64 string is sent via the RESTful API to the backend server. The server decodes the Base64 string back

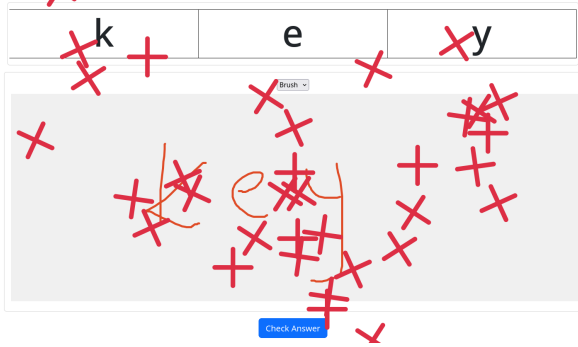


Figure 3: Failure confetti when the written alphabets is wrong

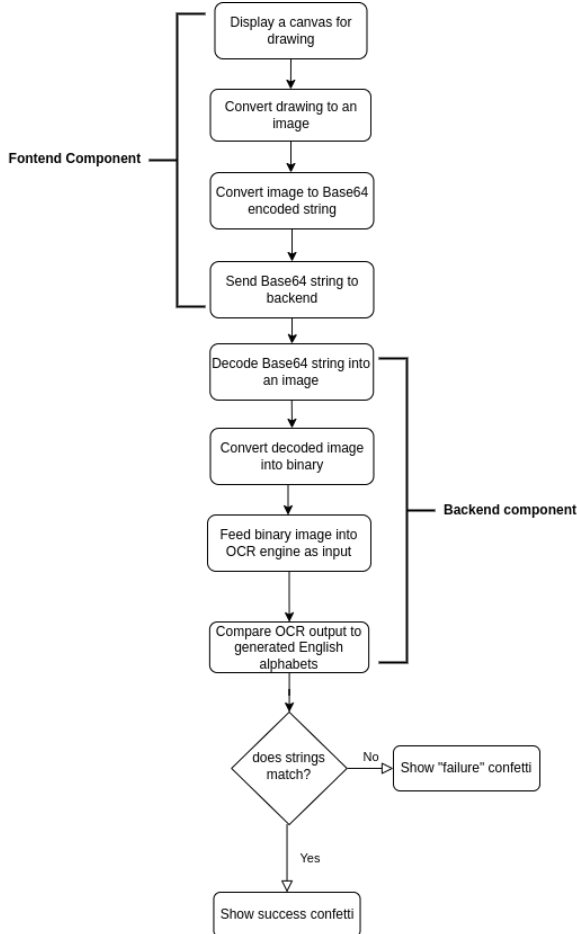


Figure 4: Flowchart of the web application

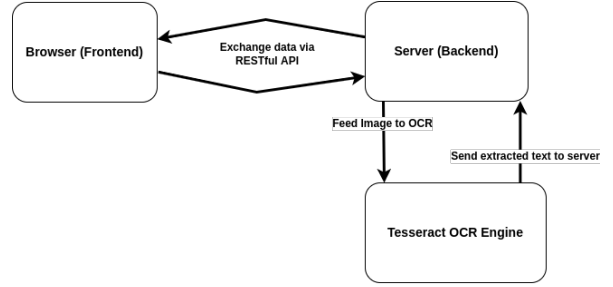


Figure 5: Architecture of the web application

to an actual image and converts the resulting image into a binary image. The binary image is then fed into the Tesseract OCR Engine. The resulting output from the OCR engine is submitted back to the frontend component. On the frontend, the OCR output is compared to the original alphabet, if both strings match success confetti is shown to the user otherwise a failure confetti is displayed to the user. Figure 5 shows the architecture of the proposed system.

3.2 Frontend Component

The frontend component was built with ReactJS [noa, 2023b]. ReactJS is a popular open-source JavaScript library used for building user interfaces for web applications; it allows developers to create reusable User Interface (UI) components and efficiently manage the state of the UI. The application uses KonvaJS [noa, 2023a] plugin for the drawing functionality. KonvaJS is a 2D graphics library for JavaScript that enables developers to create and manipulate graphical elements on an HTML5 canvas. It provides an easy-to-use API and supports various features such as drag and drop, an-

imations, and event handling.

3.3 Backend Component

The backend server was built using Python programming language and FastAPI web framework. FastAPI is a modern, fast, and type-safe web framework for building APIs with Python 3.9+ based on the open standards for APIs (OpenAPI); it provides a simple, intuitive, and high-performance interface for building APIs with minimal boilerplate code.

3.4 Image Processing

The image is sent as a Base64 encoded string to the server via RESTful API. The server decodes the string into an actual image. The colour image is converted into a binary image using Otsu's thresholding algorithm [Bangare et al., 2015]. The resulting binary image is then fed into Tesseract OCR to identify the characters in the image. The output is compared to the original alphabet to check for accuracy. Figure 6 shows graphical representation of the image processing pipeline used in this paper.

4 Limitation

It was discovered in the testing phase of the application that the accuracy of the Tesseract OCR engine is poor when using an image which has few characters. This is because Tesseract OCR relies heavily on contextual information to recognise characters correctly. In the case of single-character images, little or no contextual information is available, which can lead

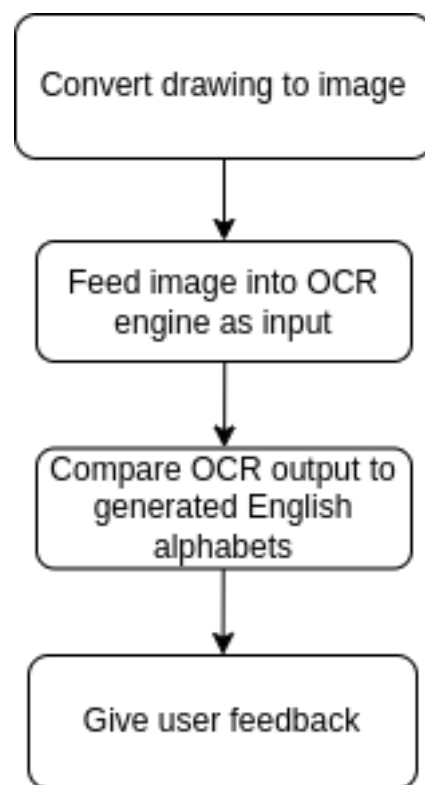


Figure 6: Image processing pipeline

to errors. Tesseract OCR uses machine learning algorithms to recognize characters, and the accuracy of the OCR can depend on the quality and quantity of the training data available. If the training data does not include enough examples of single-character images, the OCR accuracy may be lower. Unlike [Zin et al., 2021], this study did not collect any handwritten dataset to retrain the Tesseract OCR neural network. This, therefore, resulted in low accuracy when recognising handwritten text. However, there are methods to improve the accuracy of the Tesseract OCR engine as mentioned in [Hengaju and Bal, 2023, Sporici et al., 2020].

cess more interactive and fun. The Text-to-Speech functionality can spell out the alphabet to write and also give step-by-step instructions on how to write the alphabet.

5 Conclusion

The paper explored how to leverage the Tesseract OCR engine to assist KG pupils in learning the writing of the English alphabet. A web application that allows pupils to learn writing was presented in this paper. The application randomly generates a set of English alphabets. Users (KG pupils) then draw the alphabets and Tesseract OCR is used to check if the alphabets were written correctly.

6 Future Works

Future studies on leveraging OCR to assist young people in learning writing can implement or (using existing OCR) which extracts text images with few or even single characters. Text to Speech (in the native language of the learner) functionality can also be added to make the learning pro-

References

- [noa, 2023a] (2023a). Konva.js.
- [noa, 2023b] (2023b). React.
- [Alvaro et al., 2010] Alvaro, A. K. S., Dela Cruz, R. L. D., Fonseca, D. M. T., and Samonte, M. J. C. (2010). Basic handwriting instructor for kids using OCR as an Evaluator. In *2010 International Conference on Networking and Information Technology*, pages 265–268. ISSN: 2324-8203.
- [Asad et al., 2016] Asad, F., Ul-Hasan, A., Shafait, F., and Dengel, A. (2016). High Performance OCR for Camera-Captured Blurred Documents with LSTM Networks. In *2016 12th IAPR Workshop on Document Analysis Systems (DAS)*, pages 7–12.
- [Bangare et al., 2015] Bangare, S. L., Dubal, A., Bangare, P. S., and Patil, S. (2015). Reviewing Otsu’s Method For Image Thresholding. *International Journal of Applied Engineering Research*, 10(9):21777–21783.
- [Berninger et al., 2002] Berninger, V. W., Abbott, R. D., Abbott, S. P., Graham, S., and Richards, T. (2002). Writing and Reading: Connections Between Language by Hand and Language by Eye. *Journal of Learning Disabilities*, 35(1):39–56.
- [Dinehart and Manfra, 2013] Dinehart, L. and Manfra, L. (2013). Associations Between Low-Income Children’s Fine Motor Skills in Preschool and Academic Performance in Second Grade. *Early Education and Development*, 24(2):138–161. Publisher: Routledge _eprint: <https://doi.org/10.1080/10409289.2011.636729>.
- [Faizullah et al., 2023] Faizullah, S., Ayub, M. S., Hussain, S., and Khan, M. A. (2023). A Survey of OCR in Arabic Language: Applications, Techniques, and Challenges. *Applied Sciences*, 13(7):4584. Number: 7 Publisher: Multidisciplinary Digital Publishing Institute.
- [Hengaju and Bal, 2023] Hengaju, U. and Bal, D. B. K. (2023). Improving the Recognition Accuracy of Tesseract-OCR Engine on Nepali Text Images via Preprocessing. *Advancement in Image Processing and Pattern Recognition*, 3(2,3):40–52. Number: 2,3.
- [Longcamp et al., 2006] Longcamp, M., Boucard, C., Gilhodes, J.-C., and Velay, J.-L. (2006). Remembering the orientation of newly learned characters depends on the associated writing knowledge: A comparison between handwriting and typing. *Human Movement Science*, 25(4):646–656.
- [Longcamp et al., 2005] Longcamp, M., Zerbato-Poudou, M.-T., and Velay, J.-L. (2005). The influence of writing practice on letter recognition in preschool children: a comparison between handwriting and typing. *Acta Psychologica*, 119(1):67–79.
- [Smith, 2007] Smith, R. (2007). An Overview of the Tesseract OCR Engine. In *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007) Vol 2*, pages 629–633, Cu-

ritiba, Parana, Brazil. IEEE. ISSN: 1520-5363.

[Sporici et al., 2020] Sporici, D., Cuşnir, E., and Boiangiu, C.-A. (2020). Improving the Accuracy of Tesseract 4.0 OCR Engine Using Convolution-Based Preprocessing. *Symmetry*, 12(5):715. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute.

[Tang et al., 2006] Tang, K.-T., Li, K.-K., and Leung, H. (2006). A Web-Based Chinese Handwriting Education System with Automatic Feedback and Analysis. pages 176–188.

[Ye and Doermann, 2013] Ye, P. and Doermann, D. (2013). Document Image Quality Assessment: A Brief Survey. In *2013 12th International Conference on Document Analysis and Recognition*, pages 723–727. ISSN: 2379-2140.

[Zin et al., 2021] Zin, T. T., Thant, S., Pwint, M. Z., and Ogino, T. (2021). Handwritten Character Recognition on Android for Basic Education Using Convolutional Neural Network. *Electronics*, 10(8):904. Number: 8 Publisher: Multidisciplinary Digital Publishing Institute.