Eye Strokes: An Eye-gaze Drawing System for Mandarin Characters

DEREK C.W. TONG, University of Glasgow, Singapore
XUET YING TAN, Tan Tock Seng Hospital, Singapore
ALEX Q. CHEN*, Singapore Institute of Technology, Singapore

Studies showed that 61% of the elderly in Singapore are English illiterate, and it is essential to find alternatives for non-English literate patients who have literacy in the next most common language, Mandarin. Many eye typing solutions use Mandarin hanyu pinyin, a phonetic system similar to eye typing in the English language. However, most Mandarin-speaking elderly in Singapore are not familiar with Mandarin hanyu pinyin despite being able to read and write Mandarin characters, which makes eye typing redundant. We propose Eye Strokes, a technique to capture eye gaze, as an input modality to identify Mandarin characters for motor neurone disease patients to communicate. This method uses the eye gaze as strokes in a Mandarin character to predict and identify the Mandarin word the user intends to communicate. The proof-of-ideation evaluation discussed in the paper shows that our technique is feasible with promising character prediction accuracy for further investigations, although some limitations exist.

CCS Concepts: • **Human-centered computing** → **Accessibility technologies**; • **Computing methodologies** → *Computer vision*; *Artificial intelligence*.

Additional Key Words and Phrases: Mandarin character detection, Chinese character detection, Eye strokes, Assistive technology

ACM Reference Format:
Derek C.W. Tong, Xuet Ying Tan, and Alex Q. Chen. 2024. Eye Strokes: An Eye-gaze Drawing System for Mandarin Characters. *Proc. ACM Comput. Graph. Interact. Tech.* 7, 2, Article etra24-fp1043 (June 2024), 15 pages. https://doi.org/10.1145/3654702

1 INTRODUCTION

Motor Neuron Disease (MND) is a medical condition that renders patients bed-bound and tracheostomized, meaning they cannot produce verbal speech. To make things worse, most of the patients suffer from paralysis or eventually will be. This is due to the disease affecting both the upper and lower motor neurons, which causes rapid loss of muscle control and eventual paralysis. [11] Patients who can't move or communicate struggle to express their thoughts, which makes it very difficult to provide the appropriate care they require. Thankfully, one of their other senses, eyesight, remains unaffected, which allows them to use it to communicate. However, this method of communication is rather binary and can be further improved through breakthroughs in technology, such as Eye Tracking.

*Correspondence author

Authors' Contact Information: Derek C.W. Tong, derektongcw@gmail.com, University of Glasgow, Singapore; Xuet Ying Tan, xuet_ying_tan@ttsh.com.sg, Tan Tock Seng Hospital, Singapore; Alex Q. Chen, Alex.Q.Chen@singaporetech.edu.sg, Singapore Institute of Technology, Singapore.

ACM 2577-6193/2024/6-ARTetra24-fp1043
https://doi.org/10.1145/3654702

Existing studies on eye typing systems focus primarily on English-literate users who use a visual keyboard as an input modality. [9] Users of these studies gaze and dwell at the desired letters on the visual keyboard to make their selection to form the words and sentences. However, it has been reported that only 61% of the elderly in Singapore are English illiterate. [12] Hence, it is necessary to find alternatives for the group of non-English literate patients with MND who have literacy in the next most common language in Singapore. In a study done at the Singapore General Hospital (SGH) over 10 years between 2004 and 2014, 68.5% were Chinese. The Chinese language is a complex language that has a long history and geographical variations that resulted in many dialects and subdialects, of which Mandarin is most widely spoken and used. Chinese characters are logographs, and most were developed as a form of image of an object or an idea. As such, two similar-looking characters can have very different meanings and pronunciations. Conversely, two similar-sounding words can have very different-looking characters. With standardisation by China, Simplified Chinese Characters are widely used however there may be variations across dialects. For this paper, we will use the term “Mandarin characters” to refer to Simplified Chinese Characters that are commonly used in Mandarin. [2] Many eye-typing investigations were done with Mandarin hanyu pinyin, a phonetic spelling system using Roman characters. [5] However, most Chinese elderly in Singapore are not familiar with Mandarin hanyu pinyin despite being able to read and write Mandarin characters. Thus, most existing solutions that use Mandarin hanyu pinyin to communicate are unusable.

We proposed Eye-Strokes, a system that detects and recognises Mandarin characters from the user’s eye gaze to draw out the Mandarin Chinese characters. The project aims to improve the well-being of elderly patients who suffer from Motor Neuron Disease (MND) and only have a basic level of Mandarin communication. It builds upon a pilot experiment to evaluate the concepts of such a system without a Graphical User Interface (GUI). This project leverages the lessons learnt from the pilot experiment to translate it to the current version proposed in this paper. It aims to enhance the system’s capabilities and accuracy and serve as proof of ideation for future development.

This project is done in collaboration with Tan Tock Seng Hospital’s Speech Therapy Department (TTSH-STD) to deliver a thoughtful solution for the target audience. They provide valuable insights such as the day-to-day communications between MND patients and the nurses, which are done by blinking their eyes when the nurses correctly point to the character, word, or item they wish to communicate on a core and fringe words printout.

2 LITERATURE REVIEW

Eye-tracking technology has advanced significantly and is widely adopted in many systems. Studies to improve eye-typing entry [14] and using eye-typing with alternative Mandarin keyboards [5] show the demand and extent of research in this space.

2.1 Eye Typing

Studies exploring eye typing techniques are widely covered. [7, 14] Existing technology uses dwelling for the selection of specific keys on standard keyboards. Attempts to use eye typing with Mandarin pinyin keyboards to input Mandarin characters/text exist. [5] This study uses a specially designed software keyboard based on a study of Mandarin Pinyin. However, it does not fit the needs of our targeted group of elderly MND users, as they are unfamiliar with the Mandarin hanyu pinyin system or English letters to type with their eye gaze.

BlinkWrite2 is another interesting study that used the switch scanning method with blink as an input modality. It increases efficiency for selection keys or predicted text and may be used to select common phrases. This method is existingly used in pictorial or word charts. The drawbacks, however, can be that if patients were keen to express something beyond the choices given, they would be limited by the options offered, as listing all existing Mandarin characters is impossible. This study suggested that auditory feedback was helpful and was considered in the design of Eye Strokes. [1]

2.2 Drawing Objects with Eye Gaze

EyeDraw is designed for children with severe mobility impairments, utilizing an LC Technologies Eyegaze eye tracker that reports gaze points 60 times per second. [3] The system was validated for nondisabled children and empowers those with limited mobility to draw pictures using eye movements and dwell time. User observation studies with disabled children are ongoing. To enhance the drawing experience, EyeDraw addresses saccades by averaging six consecutive gaze points, presenting them on screen as the eye cursor—a coloured square with minimal delay. The program lets users control line drawing by determining starting and ending points instead of pixel-by-pixel rendering.

The author's HCI research, including CPM-GOMS analysis, examined users' eye movement decision times and durations [3]. Emphasizing auditory feedback's significance for prompt decision-making, a 30 ms lead for eye movement or drawing initiation was suggested. Additionally, the proposal of grid dots with 1 cm spacing as visual guides enhances drawing precision, providing a better anchor for users compared to a blank white field.

2.3 Existing Handwriting and Drawing Character Recognition Methods

The study draws inspiration from two sources, particularly Lopresti et al.'s exploration of Symbolic Indirect Correlation (SIC) and Style Constrained Classification (SCC) for recognizing handwritten Arabic and Mandarin words. [8] SIC involves reassembling segments of an unknown query matching those of labelled reference words, relying on the order of feature vectors. SCC, on the other hand, predicts distortions in characters based on style, adapting with a long enough field.

Both Arabic and Mandarin characters present unique challenges in recognition. Arabic, with a small alphabet and word-position-dependent allographs, contrasts with the vast classes of Mandarin characters. Both scripts, with traditional roots dating back thousands of years, share commonalities and are incorporated into other languages. Jayech et al. delve into Arabic handwriting, highlighting the morphological complexity and variable writing styles, proposing a Dynamic Hierarchical Bayesian Network (DHBN) to address these challenges. [4] The DHBN employs segmentation, Zernike, and HU moments for feature extraction, showcasing a promising approach for tackling Arabic handwriting intricacies. HU movements are like unique fingerprints for shapes in pictures, which were used to help computers recognize Arabic handwriting. These insights provide a robust foundation for the current study's exploration of handwriting recognition methods.

2.4 Handwriting Chinese Character Recognition with Convolutional Neural Network

Another common approach involves training an Artificial Intelligence system for Mandarin character recognition from handwriting using Convolutional Neural Networks (CNN) on an offline dataset. [16] The CNN algorithm utilizes software functionality and cross-entropy loss to construct the model. The dataset, sourced from CASIA online and offline Chinese handwriting databases, undergoes necessary pre-processing, including cropping, scaling, and padding.

Three distinct models shared the same neural network structure. The first model employs the 200-batch SoftMax activation approach with stochastic gradient descent optimization. The SoftMax cross-entropy loss function is used to increase the distance between classes. [16] The second model utilizes a 180-batch size, combining SoftMax with Euclidean distance to minimize the distance between samples of the same class. With a 200-batch size, the third model incorporates SoftMax plus variance to minimize variance within the same class. Post-testing revealed that the variance SoftMax cross-entropy outperformed the other models.

The report discusses a CNN model where character classification and similarity ranking supervisory signals complement each other. [16] This combination effectively increases inter-class variations and reduces intra-class variations, resulting in superior classification performance compared to only SoftMax cross-entropy-based character classification. Notably, the variance similarity ranking function outperforms the Euclidean similarity ranking function.

3 METHODOLOGY

This section explains an overview of the technologies and techniques used to implement the Eye Strokes system. The application uses the Unity Engine and the Tobii Eye Tracker to assist users in drawing Mandarin characters on the screen using their eye gaze. The system's input is processed by an ML model to predict the intended Mandarin Character.

3.1 Requirement Analysis

The primary users of the application are MND patients in the hospital. These patients only have a basic level of Mandarin comprehension and are unable to vocalise or physically communicate their thoughts. The patients will be trained on how to navigate the application and have their eye profiles calibrated by the nurses. The application will utilise ML algorithms to process the eye-gaze drawings to identify the Mandarin characters for communication with healthcare practitioners and caregivers.

User stories were designed to provide concise descriptions of the desired functionalities by project. A total of 10 user stories covers the key functionalities.

#1 – Eye Profile Calibration
As a user, I would like to be able to sync my eye profile with the eye tracker so that I can use the software application with accurate eye tracking.

#2 – Tracking of Eye-Gaze
As a user, I would like my eye-gaze to be tracked accurately when I move my eyes around the screen so that I can interact with the application.

#3 – Tracking of Eye-Dwelling
As a user, I would like my eye-gaze to be tracked accurately when I keep my eyes on a certain area on the screen so that I can interact with the application.

#4 – Eye-Gaze Indicator
As a user, I would like to see a visual indicator on the screen to be able to know where I am gazing.

#5 – Drawing of intended Mandarin Character
As a user, I would like to be able to draw strokes on the application with my eye gaze to get closer to conveying my intended Mandarin Character to the nurses.

#6 – Fix mistakes or bad strokes
As a user, I would like to be able to go back and redo the previous stroke or restart from scratch when I make a mistake so that I can get closer to getting a better prediction for my intended Mandarin Character for the nurses.

#7 – Prediction of intended Mandarin Character
As a user, I would like to be able to receive predictions of my intended Mandarin Character to be able to narrow down and choose my actual intended Mandarin Character for the nurses.

#8 – Confirm Intended Mandarin Character
As a user, I would like to be able to confirm the Mandarin Character to let the nurse know that it is my intended Mandarin Character.

#9 – Identify Intended Mandarin Character
As a nurse, I would like to be able to tell what Mandarin Character the patient has chosen after they are done drawing and confirming their Mandarin Character.

#10 – Knowing the current state of the application
As a user, I would like to be able to tell what is happening at any point of time in the application so that I won't be confused or lost while using the application.

3.2 Design

The design considerations for the study include the needs of Motor Neuron Disease (MND) patients, the capability to recognise Mandarin characters, and the physical set-up required. More detailed discussions of the design of the system are presented in the following subsections. These discussions provide a high-level overview of the system's architecture, system flow and how the project is physically and digitally set up.

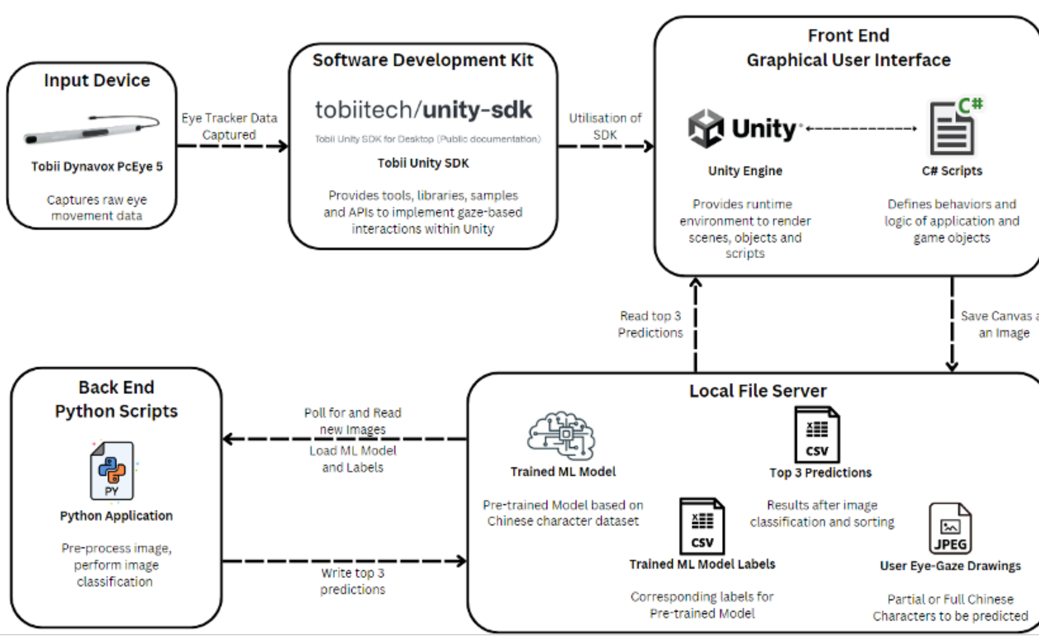

Fig. 1. Eye-Strokes Drawing System for Mandarin Characters Architecture

3.2.1 System Architecture. The system uses a modular architecture to enable easy changes, maintainability, and expandability as shown in Figure 1. Starting from the top left, the eye tracker collects the user's eye movement data. This data is relayed to the front end, which is the Unity engine, through the integration of the Tobii Unity Software Development Kit (SDK) within the Unity project. Both the front end and back end are always running concurrently, both actively reading and writing the files in the project folder in the local file server. When the user completes a stroke, the canvas is captured as an image, which is stored in the project folder on the local file server. At the same time this happens, the Python application, which is constantly polling for changes in the output images subdirectory, reads the new image, pre-processes it and uses the ML model and labels loaded to classify the image. The results are sorted, and the top 3 predictions are written into a CSV file, which will then be read in the update cycle of the Unity runtime and displayed to the user.

3.2.2 System's Mandarin Characters Detection Flowchart. The flowchart in Figure 2 shows the system's flow for detecting Mandarin characters. After the "Start System" activity, the patient will interact with the system by triggering the Start Button. This event will start a countdown to prepare the user for drawing their Mandarin character stroke. The countdown will give them time to orientate their eye gaze to the canvas and the spot where they wish to begin their stroke. Once the user has started drawing, a new countdown will begin as visual and audible feedback, to inform the duration for completing the drawing. After the countdown concludes, the canvas is immediately exported as an image, pre-processed, classified, sorted and lastly, displayed on the GUI for visual feedback. The patient is then free to do a total of 4 actions, add a new stroke, undo the previous stroke, clear all strokes, or select one of the top 3 predictions provided by the system. This cycle will repeat until the user has chosen a prediction, which will then be observed by the nurse, completing the flow of the system.

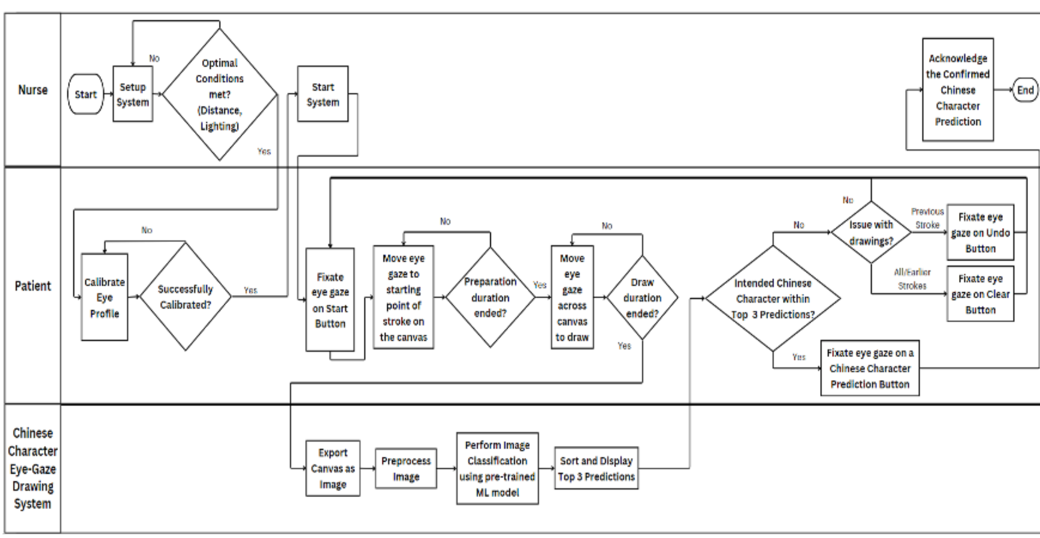

Fig. 2. Eye-Strokes Drawing System for Mandarin Characters Detection Flowchart

4 EXPERIMENT SETUP AND METHODS

Eye Strokes consist of two key components, the Graphical User Interface (GUI) component and the Machine Learning (ML) component. Unity Game Engine was used to create the GUI component inside a Unity application where game objects, C# scripts and the Tobii Unity SDK work together to provide an interactive front-end experience for the user. Python is used in the ML component to perform live image classification from the inputs of the Unity application processes, which acts as the backend of the system. The ML model is trained separately from the system's runtime using Python both locally and on Google Colab, a cloud based Jupyter Notebook environment. The following subsections detail how the system and experiment were designed and set up to capture meaningful insights into the feasibility of the system and the effectiveness of the GUI.

4.1 Project Physical Setup

The physical setup of the project is also essential in ensuring the Tobii eye tracker can perform as accurately as possible. As shown in Figure 3, the user is recommended to be within 50 to 95cm from the eye tracker for optimal eye gaze readings and must be parallel with the user's eyes. [10, 15] In the case of MND patients, they would be lying on their backs on the hospital bed. The room should also ideally be well-lit but without any bright light sources shining directly onto the screen or into the user's eyes.

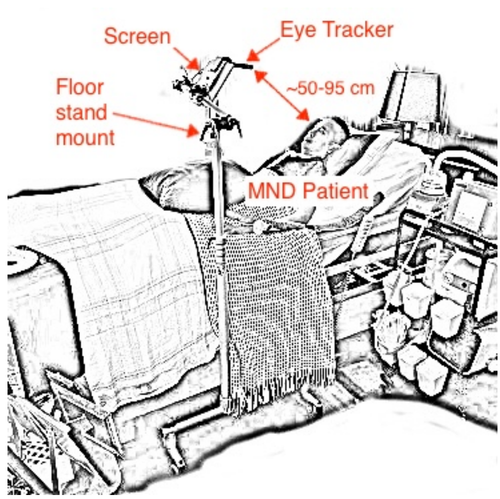

Fig. 3. Eye-Strokes Drawing System for Mandarin Characters Project Physical Setup

4.2 Tobii Eye Tracker Calibration

To use Eye Strokes, the user must first go through an eye tracker calibration process, which creates their eye profile before they can properly use the system. This process allows the eye tracker to record the unique eye characteristics and eye movement of the user. The calibration uses Tobii Dynavox's proprietary Software Bundle, as shown in Figure 4, where the user must follow the instruction of moving their eye gaze to follow the dot on the screen. The calibration process is only required once if the physical setup does not change. Any changes to the physical setup will require the eye profile to be recalibrated.

4.3 Graphical User Interface (GUI) Component

The design concepts are to keep the GUI simple, within reach, intuitive and easy on the eyes. These choices were based on the discussion in the literature review on the EyeDraw application in section 2.2.

4.3.1 Main Scene. The Main scene serves as the central hub for the system's functionalities, containing various objects that users can interact with. Figure 5 shows what the Main scene looks like after the user has achieved a satisfactory prediction from their drawing. The Main scene allows users to engage with the buttons and canvas elements immediately. All objects involving texts in the Main scene will be primarily in Mandarin, with their English equivalent for evaluation purposes.

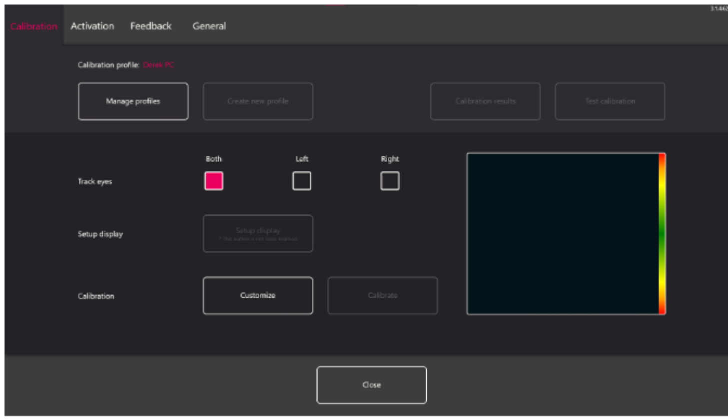

Fig. 4. Tobii Dynavox Software Bundle – Calibration Settings Page

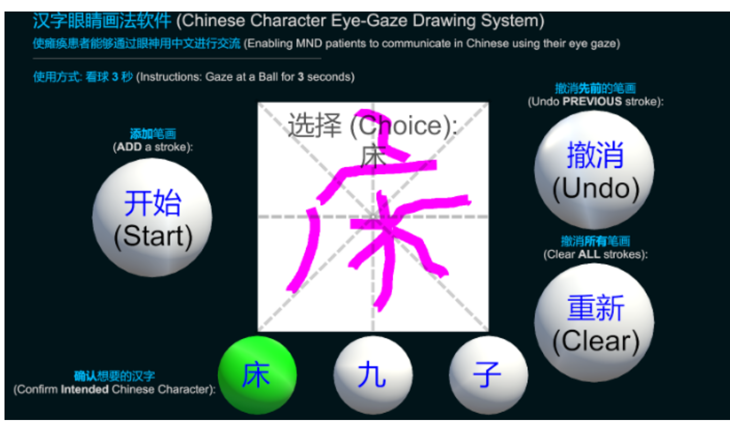

Fig. 5. Main Scene Screenshot (Completed state)

4.3.2 Buttons Implementation. Interactive buttons are designed around the Canvas to provide users with a way to inform the system what action they want to perform. Users can trigger the drawing process with the Start button, correct mistakes with the Undo or Clear buttons, and accept the correct prediction with the respective Prediction Buttons.

To design these buttons to be intuitive, they will gradually be filled with colour to indicate that the user's gaze has been fixated on that button. When the gaze dwell duration of 3 seconds is met, the button will be filled completely to indicate the selection of the button. Figure 6 shows the different states of gaze dwelling on the Start button. The text in the buttons is also dynamically changed based on the state of the system to inform the user of the system and the buttons' states. Their sizes and locations have been optimised and made consistent based on their functions, where the canvas interaction buttons are larger and situated on the left and right of the canvas due to their importance in the system flow, while the prediction buttons are slightly smaller and situated below the canvas as they are the final steps in the system flow.

The design of the buttons closely embodies Nielsen's 7 usability heuristics [14], which are Visibility of system status, Match between system and real world, User control and freedom, Consistency and standards, Error prevention, Recognition rather than recall and Flexibility and efficiency of use.

Fig. 6. Start Button Game Object (None, Partial, Complete dwell durations)

4.3.3 Canvas for Drawing the Characters. The placement of the canvas in the Main scene is intentionally designed for the user's gaze to easily return to the origin to perform the drawing of strokes for their Mandarin character. The left image in Figure 7 is when the canvas is empty, and the Unity application is in an idle state. The right image is the completed state where the Mandarin character is drawn, successfully predicted, and confirmed by the user.

The diagonal, vertical and horizontal dotted guidelines are included to enhance the user experience by giving guidance to the user and providing an anchor and reference points for drawing the Mandarin characters. This method is widely used in Mandarin calligraphy papers or Mandarin character practice papers and was also proposed by the EyeDraw application in subsection 2.2 to act as a visual guide to enhance drawing precision, providing a better anchor for users compared to a blank white field [3].

Visual cues were explicitly integrated into the system to complement the intuitiveness of the user interface. Guiding texts are added to the upper section of the canvas through various states of the Unity application, as shown in Figure 8. Additionally, there are other visual cues implemented to enhance the feedback to the user, such as every button's text will update and return to normal after triggering, and for the preparation phase, the canvas will blink every second to complement the countdown guide text and countdown audio.

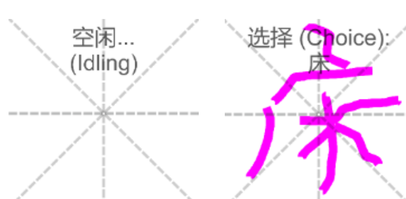

Fig. 7. Canvas Game Object (Idle, Complete states)

4.3.4 Audio Cues. Audio cues are essential in a system where the user's eye gaze is predominantly focused on being the input rather than receiving information. Auditory feedback was recommended in the literature review on the EyeDraw application in subsection 2.3. This was made apparent during the findings of early tests with healthy participants who mentioned that they did not want to let their gaze wander or settle on any location for too long as they were afraid they would trigger something unintentionally. Thus, having audio cues on top of the visual cues allows the user to be more focused on triggering the buttons and preparing to draw or are currently drawing, all whilst listening for the audio cues. Audio cues that were added were countdown sound effects between phases in the drawing sequence, system states and the button dwelling countdowns.

4.4 Machine Learning Component

The focus of the Machine Learning (ML) component was to complement the GUI component with a competent and robust prediction model. The following subsections detail the considerations and implementation of the ML component.

4.4.1 CASIA Offline Handwriting Database Dataset. The dataset that was used to train the main ML model was built by the National Laboratory of Pattern Recognition (NLPR), Institute of Automation of Chinese Academy of Sciences (CASIA), known as the CASIA Offline Chinese Handwriting Database (CASIA-HWDB). The handwritten samples were produced by 1,020 writers using Anoto pen on paper, such that both online and offline data were obtained. The samples include both isolated characters and handwritten texts from continuous scripts. [6]

The dataset consists of 7,185 classes of Mandarin characters with about 600 images for each class and a total dataset size of about 13 Gigabytes was downloaded from Kaggle. As the project is a pilot study to evaluate its feasibility, 26 characters out of 243 words were picked from the Core and Fringe Mandarin Words document provided by TTSH-STD, which is approximately 10% of the list. This list is curated by the team at TTSH-STD to train healthcare practitioners on what patients commonly use to convey to the nurses. To select the Mandarin characters fairly, the first two unique Mandarin characters from each category or type were chosen. Since numerals are commonly used, the equivalent to Roman numerals 1 to 10 were also chosen. The evaluation was restricted to these characters for this pilot study. In future iterations, a larger pool of trained characters will be incorporated.

4.4.2 Convolutional Neural Networks (CNN) Model Training. The training of the ML model was done outside the runtime of the system. Conceptually, a CNN Model was a popular technique used for image recognition and classification tasks due to its higher accuracy and was researched to be effective for the study in subsection 2.4. [16]

It consists of three layers: a convolutional layer for feature extraction, a pooling layer for dimensionality reduction, and a fully connected layer for making predictions. In this model, the Rectified Linear Unit (ReLU) activation function is utilised, known for its reliability and faster convergence compared to other activation functions. The final layer uses the SoftMax function, which was also recommended by the study in subsection 2.4, for multi-class classification. The model is implemented using the TensorFlow Keras library, with Categorical Cross Entropy as the loss function to handle multiple classes and Adam as the optimizer, well-suited for datasets with empty spaces to increase accuracy.

5 RESULTS AND ANALYSIS

In this section, the trained model and techniques used in Eye Strokes were evaluated and analysed. The usability of the Eye Stroke system is also examined to expose the strengths and limitations of the system. The results and discussions are presented in subsequent sections.

5.1 Convolutional Neural Networks Trained Model

To effectively train and evaluate the CNN model, the dataset is split into two subsets: a training set and a validation/test set. The 80:20 split ratio is commonly employed, where 80% of the data is used for training the model, while 20% is reserved for evaluating its performance. [13] This split allows the model to learn patterns and features from a substantial amount of training data while enabling assessment of its generalisation to unseen data using the validation/test set. The 80:20 split helps prevent overfitting and ensures reliable model performance. The split was achieved using the Python Scikit-learn library, which utilises a random seed to split the dataset into the training and testing set based on the given parameters. The images were sized at 64 x 64 pixels in all datasets used for the CNN model. This standardisation ensures consistency and simplifies the model's architecture whilst retaining sufficient features for an accurate ML model. The ideal epochs and batch sizes to train the dataset were found to be 30 epochs and a batch size of 64 after programmatically evaluating the dataset. These parameters give a balance of computational power and fall-off in accuracy. As seen in Figure 8, a test accuracy of approximately 95% was achieved at 30 epochs using a batch size of 64. The higher validation accuracy seen in Figure 8 can be attributed to regularization, data splitting, model complexity, and randomness. In our case, the validation set was done on more straightforward examples than the training set, which resulted in a higher validation accuracy.

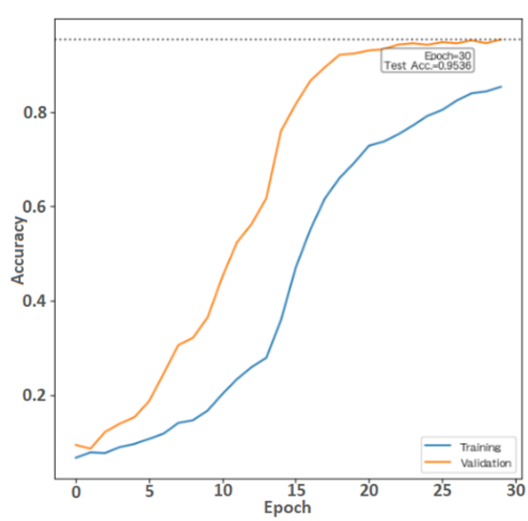

Fig. 8. Sample Training Accuracy graph of CNN Model Training on dataset

5.2 Usability Testing

The usability and accuracy of recognising the Mandarin characters drawn using eye gaze are the two Key factors to determine the feasibility of our approach. This study is a pilot investigation to demonstrate the feasibility of our approach, so only ten participants were recruited. The evaluation was conducted when the key functionalities of the GUI component were complete, and a lightweight ML model based on the Augmented Mandarin MNIST dataset was trained. These participants were undergraduate students who voluntarily partook in this test. All participants were healthy subjects between the ages of 23 and 28. As our target audience is actual MND patients in the hospital, conducting trials on patients is heavily regulated, and a common practice before clinical trials (evaluation on the targeted audience) is to demonstrate its feasibility, which is the stage we are at for the project. Thus, healthy subjects were selected for this study. The explicit and implicit feedback from the users as they interacted with the interface and the collected data allowed for the identification of usability issues, improvement of interface design, and enhancement of user satisfaction prior to introduction to the patient population. At this stage, only a small pool of healthy subjects was selected to aid in setting a strong foundation for the system. Future usability tests should include MND patients and a larger pool of subjects for a more comprehensive usability and robustness test in realistic use cases. This study focuses on the ability to draw and accurately identify relatively simple individual characters with the system without significant obstacles.

Each participant is assigned to three tasks. These tasks required the participant to draw character(s) and use the undoing and clearing the canvas features. If the tasked Mandarin character appears in the Prediction Buttons, the participants are required to choose one of the buttons that best resembles the character in the task to complete it. The participant's and application performance are captured and evaluated. A sample of the tasks that the participants were asked to conduct:

(1) User is to draw '一' (a single horizontal stroke)
   (a) Evaluate intuitiveness of Start and Prediction Buttons, Canvas, Visual and Audio Cues
(2) User is to draw '十' (a horizontal and vertical stroke, with allowance to undo the last stroke or clear the entire canvas)
   (a) Evaluate intuitiveness and usefulness of Undo and Clear Buttons
(3) User is to draw '床' (bed in Mandarin)
   (a) Evaluate the usefulness of Undo and Clear Buttons
   (b) Fatigue levels when drawing more complex Mandarin Characters

Besides the purposeful intentions of selecting Mandarin characters for the test, these characters are also selected from the Final Mandarin character pool as they are relatively basic Mandarin characters, therefore should keep the tasks simple and quick to complete. Individual characters are tested in this usability test; however, this does not mean that the system and its usage are only limited to single character recognition. If patients wish to express multiple Mandarin characters, they can continue to do so by adding on subsequent characters after each selection.

The survey includes questions to capture cognitive demands, mental effort, and frustration experienced by users. Questions related to software-specific factors such as interface intuitiveness and navigation challenges were also included. Participants were allowed to provide optional feedback and suggestions for improvements to the system at the end of the survey. Table 1 shows the truncated quantitative survey results[1] obtained from the 10 participants.

From the results obtained in the intermediate Usability Test, we can see that out of the 10 participants, 70% feel that the GUI is intuitive, 90% are satisfied with User Experience and when tasked in Task 3 to draw a realistic Mandarin character, '床', that would likely be used by the actual target audience, only 30% felt that the task was difficult. Other results obtained in the survey were used to validate and improve the GUI elements of the final iteration of the system. In particular, the dwell durations were extended by 1 second from 3 seconds, which was the original dwell duration chosen, and the button sizes were increased by 20%.

The bar chart in Figure 9 illustrates the results from questions 1 to 3, each task's difficulty, GUI intuitiveness and User Experience. The questions and the options for each question are available in the supplementary materials, section D.2 Usability Survey Questions. The bar chart in Figure 9 aligns with the surveyed responses from Table 1, suggesting that the usability of the GUI design is reliable and intuitive.

The questions 1.1, 1.2, and 1.3 regarding each task's difficulty were 3-point scale single-choice questions with the options of Easy, Neutral or Hard. Similarly, questions 2 and 3 regarding the GUI intuitiveness and overall user experience were also 3-point scale single-choice questions to clearly categorise whether they were pleased, neutral or displeased with the respective aspect of the system.

Additionally, task-related items relevant to software usage measurements, such as completion time for each task and performance and the number of times each of the buttons is triggered, were integrated programmatically so that the measurements are logged automatedly during the tests.

[1]The full survey questions and options is available in the supplementary materials, section D.2 Usability Survey Questions.

Table 1. Intermediate Usability Test Survey Results (Truncated)[1]

Question No.	Question (Simplified)	Participants									
		1	2	3	4	5	6	7	8	9	10
1.1	Task 1 Difficulty	Easy	Easy	Easy	Easy	Easy	Easy	Easy	Easy	Easy	Easy
1.2	Task 2 Difficulty	Easy	Neutral	Neutral	Easy	Easy	Neutral	Easy	Easy	Neutral	Neutral
1.3	Task 3 Difficulty	Neutral	Neutral	Easy	Hard	Neutral	Hard	Neutral	Neutral	Hard	Neutral
2	GUI Intuitiveness	Neutral	Intuitive	Intuitive	Neutral	Neutral	Intuitive	Intuitive	Intuitive	Intuitive	Intuitive
3	User Experience Satisfaction	Satisfied	Satisfied	Satisfied	Neutral	Satisfied	Satisfied	Satisfied	Satisfied	Satisfied	Satisfied
4	Unintentional Button Triggers	Never	Never	Never	Rarely	Never	Never	Rarely	Never	Never	Never
5	Failed Button Triggers	Rarely	Never	Never	Rarely	Never	Never	Never	Never	Never	Never
5.1	Change Dwell Duration	0s	+1s	0s	+1s	0s	+1s	0s	0s	0s	+1s
5.2	Change Button Size	0%	0%	0%	0%	0%	+25%	0%	0%	+10%	0%
6	Stroke Drawing Difficulty	Easy	Neutral	Easy	Easy	Easy	Neutral	Easy	Easy	Neutral	Neutral
6.1	Change Drawing Duration	0s	0s	0s	0s	0s	0s	0s	0s	0s	0s
7	Un-intentionally Leave Canvas	Never	Never	Never	Rarely	Never	Never	Rarely	Never	Never	Never
7.1	Change Canvas Size	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

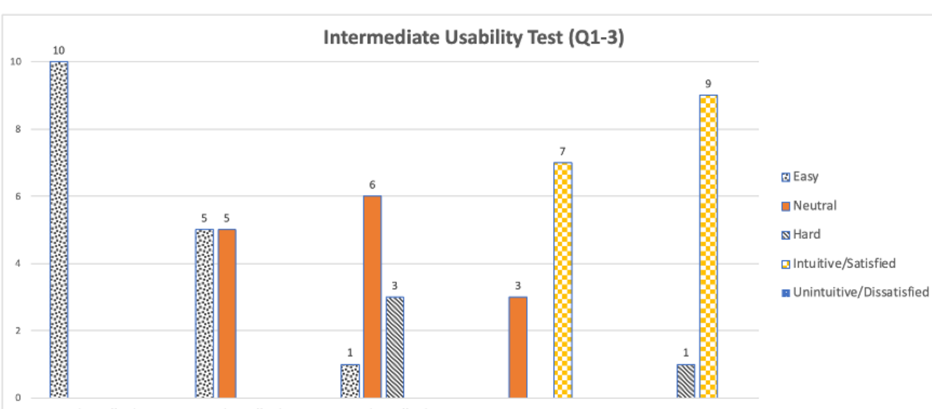

Fig. 9. Intermediate Usability Test Bar Chart (Questions 1-3)

Figure 10 shows the average time taken with the longest and shortest completion times for each task and Figure 11 shows the average number of triggers for each button across the tasks for the 10 participants. These results provide insights into the system's efficiency and usability. All participants rated "Easy" for Task 1, and for Task 2, an equal number of participants said it was "Easy" and "Neutral" (see Figure 11). Tasks 1 and 2 had a relatively short average time taken per task of 16.8s and 42.0s respectively, and low amounts of button triggers. Task 3 was designed to be more realistic and complex than the first two tasks, that is likely to be accounted for by the target audience. An average time of 184.4s, approximately three minutes, was recorded for task 3. The standard deviation for the time taken in task 1 is 0.13, while it is 10.40 in task 3. The

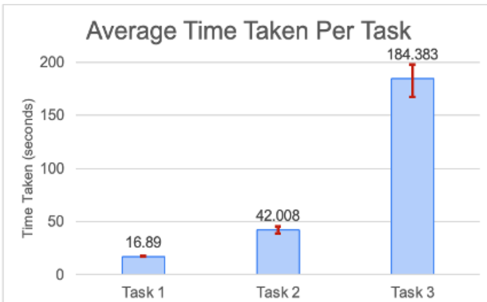

Fig. 10. Intermediate usability test for the average time taken per task with their longest and shortest completion times

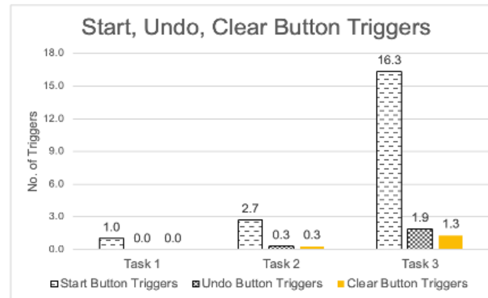

Fig. 11. Intermediate usability test for Start, Undo, Clear Button Triggers

longer time required to complete task 3 is expected since the task is more complex. From the individual participants' progression logs, we observed that each participant had their own method of completing the task, whether it was more backtracking with the undo and clear buttons or, instead, adding more strokes. The variation of methods resulted in a higher standard deviation. The results of task 3, shown in Figure 9, suggest that when users use Eye Strokes in a realistic setting, some may find it hard to use but still manageable.

These findings tell us that for healthy subjects, the system proves manageable, user-friendly, and effective. The majority find the GUI intuitive and express satisfaction with the user experience, while only a few find it challenging. Overall, all participants complete tasks relatively quickly.

6 CONCLUSION

Our proposed solution (Eye Strokes) aims to improve the well-being of Mandarin-speaking elderly patients with Motor Neuron Disease (MND) who are unable to type in English or use Mandarin hanyu pinyin. Eye Strokes employs eye-tracking technology to help patients communicate more effectively with healthcare practitioners and caregivers through text prediction from drawing the strokes of Mandarin characters.

Eye Strokes consists of two components: the GUI and the ML components. The system usability experiment tells us that the GUI is satisfactory and that the controlled scenarios to predict simple Mandarin Characters with Eye Strokes are manageable by healthy subjects. Which can be seen from out of the 10 participants who participated, 70% feel that the GUI is intuitive, 90% are satisfied with the User Experience and when tasked to draw a realistic Mandarin character, '床', that would likely be used by the actual target audience, only 30% felt that the task was difficult and their average time to complete the scenario was approximately 3 minutes.

More can be done to improve Eye Strokes' usability and capability in the future. (1) The current pilot study only evaluates a small pool of healthy subjects. To measure its true capability, future work should focus on MND patients. (2) More research is required to understand the needs of MND patients using such a system. So that related features can be included to improve the usability of Eye Strokes. (3) Expanding the pool of words/things trained will help understand the limits and capability of the techniques used in Eye Strokes.

7 ACKNOWLEDGMENTS

Ethical clearance for this study was granted by the Singapore Institute of Technology's Institutional Review Board (IRB) application approval reference number RECAS-0124.

REFERENCES

[1] Ashtiani, B., and MacKenzie, I. S. Blinkwrite2: an improved text entry method using eye blinks. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications - ETRA '10* (2010), ETRA '10, ACM Press.

[2] Deng, X., Hao, Y., Xiao, B., Tan, E.-K., and Lo, Y.-L. Risk factors for respiratory failure of motor neuron disease in a multiracial asian population. *Journal of Clinical Neuroscience 39* (May 2017), 137–141.

[3] Hornof, A., Cavender, A., and Hoselton, R. Eyedraw: a system for drawing pictures with eye movements. In *Proceedings of the 6th international ACM SIGACCESS conference on Computers and accessibility* (Sep 2003), ASSETS04, ACM.

[4] Jayech, K., Mahjoub, M. A., and Amara, N. E. B. Arabic handwritten word recognition based on dynamic bayesian network. *Int. Arab J. Inf. Technol. 13.6B* (2016), 1024–1031.

[5] Liang, Z., Fu, Q., and Chi, Z. Eye typing of chinese characters. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (Mar. 2012), ETRA '12, ACM.

[6] Liu, C.-L., Yin, F., Wang, D.-H., and Wang, Q.-F. Casia online and offline chinese handwriting databases. In *2011 International Conference on Document Analysis and Recognition* (Sept. 2011), IEEE.

[7] Liu, Y., Zhang, C., Lee, C., Lee, B.-S., and Chen, A. Q. Gazetry: Swipe text typing using gaze. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction* (Dec 2015), OzCHI '15, ACM.

[8] Lopresti, D., Nagy, G., Seth, S., and Zhang, X. *Multi-character Field Recognition for Arabic and Chinese Handwriting.* Springer Berlin Heidelberg, 2008, p. 218–230.

[9] Majaranta, P., and Räihä, K.-J. Twenty years of eye typing: systems and design issues. In *Proceedings of the symposium on Eye tracking research & applications - ETRA '02* (2002), ETRA '02, ACM Press.

[10] Nielsen, J. Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI conference on Human factors in computing systems celebrating interdependence - CHI '94* (1994), CHI '94, ACM Press.

[11] Nih.gov. Motor neuron diseases fact sheet - national institute of neurological disorders and stroke. https://www.ninds.nih.gov/health-information/disorders/motor-neuron-diseases/ (Last accessed Jul. 02, 2023).

[12] Suppiah, S., Tan, Y. W., Cheng, G. H.-L., Tang, W. E., and Malhotra, R. Mediators of the association of limited english health literacy with medication non-adherence among singaporean elderly. *Proceedings of Singapore Healthcare 29*, 1 (Jan. 2020), 25–32.

[13] TechMediaToday. What is the 80/20 rule in machine learning? https://www.techmediatoday.com/what-is-the-80-20-rule-in-machine-learning/ (Last accessed Jul. 02, 2023), 2023.

[14] Urbina, M. H., and Huckauf, A. Alternatives to single character entry and dwell time selection on eye typing. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications - ETRA '10* (2010), ETRA '10, ACM Press.

[15] US, T. D. How far away should my eyes be from the pceye 5 when using the device or calibrating? https://us.tobiidynavox.com/blogs/support-articles/how-far-away-should-my-eyes-be-from-the-pceye-5-when-using-the-device-or-calibrating (Last accessed Jul. 02, 2023).

[16] Zou, J., Zhang, J., and Wang, L. Handwritten chinese character recognition by convolutional neural network and similarity ranking. https://www.researchgate.net/publication/335565054_Handwritten_Chinese_Character_Recognition_by_Convolutional_Neural_Network_and_Similarity_Ranking (Last accessed Jul. 02, 2023), 2019.

Received 17 January 2024; revised 06 April 2024; accepted 08 April 2024