Express Yourself Simply: Mobile AI for Bridging Communication Gaps

Anna Kramar
University College Cork
Cork, Ireland
122121663@umail.ucc.ie

Kinza Salim
University College Cork
Cork, Ireland
124106946@umail.ucc.ie

Hoang D. Nguyen
University College Cork
Cork, Ireland
hn@cs.ucc.ie

ABSTRACT

Doodle Me is an AI mobile application designed to bridge communication gaps and barriers. This tool is beneficial for individuals with articulation and phonological speech disorders. By using advanced machine learning algorithms, particularly a Convolutional Neural Network (CNN), the application interprets user-created drawings in real-time. This app provides a universal mode of expression that transcends traditional language barriers. The CNN model is trained on the QuickDraw dataset which achieves an accuracy of 92.6%.

The application is accessible and practical, functioning seamlessly across all smartphones. Doodle Me enhances communication by recognizing and categorizing doodles, and provides visual and auditory feedback to users. This tool provides a holistic approach by aiding at rehabilitation of numerous medical conditions. It deals with multilinguistics of different age groups and is also eco-friendly.

The application serves as an assistive tool in the healthcare system. It can support a wide range of medical conditions, including post-ischemic stroke aphasia, facial paralysis, stuttering/stammering, Autism Spectrum Disorder (ASD), post-operative laryngectomy or vocal cord paralysis, Alzheimer's disease, and rehabilitation for post-cochlear implant patients (where children are taught about speech recognition by giving visual cues) and elderly with presbycusis (hearing loss at high frequencies).

Doodle Me addresses communication barriers and promotes biosustainability by reducing the need for unnecessary stationary items. It is a practical application aimed at making a positive societal impact by facilitating seamless integration.

CCS CONCEPTS

• **Computing methodologies** → **Neural networks**; • **Human-centered computing** → *Natural language interfaces*; *Ubiquitous and mobile computing systems and tools.*

KEYWORDS

Convolutional Neural Network, Mobile AI, Doodle Recognition, Human-Computer Interaction, Mobile Application, Machine Learning

MOBILEHCI Adjunct '24, September 30-October 3, 2024, Melbourne, VIC, Australia

ACM ISBN 979-8-4007-0506-9/24/09
https://doi.org/10.1145/3640471.3680456

ACM Reference Format:
Anna Kramar, Kinza Salim, and Hoang D. Nguyen. 2024. Express Yourself Simply: Mobile AI for Bridging Communication Gaps. In *26th International Conference on Mobile Human-Computer Interaction (MOBILEHCI Adjunct '24), September 30-October 3, 2024, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3640471.3680456

1 INTRODUCTION

Communication is a fundamental human right for every individual worldwide [13]. However, some people face significant barriers due to speech-related disabilities or disorders. These conditions affect a wide range of individuals, including those with congenital or developmental disorders, brain injuries, stroke survivors, and older adults with degenerative diseases. Such conditions can severely hinder an individual's ability to communicate and express themselves, often leading to social isolation and hence a decreased quality of life [3]. However, with the right support and intervention, these communication barriers can be reduced.

Globally, over 1 billion people live with some form of disability, with many experiencing communication impairments. For instance, Broca's aphasia affects many stroke or brain tumor survivors, leading to speech disorders [10]. Similarly, individuals with stuttering, autism, facial paralysis, Alzheimer's disease, and those who have undergone post-laryngectomy or cochlear implant surgeries struggle with effective communication [4, 5]. Traditional aids often fail to meet the diverse needs of these individuals due to high costs, complexity, or limited functionality.

A solution to address these communication gaps is to design an AI application that is easy to use. Communication interventions often include assistive technology, personalized training, and a supportive environment, which enables users to express themselves more easily. These interventions can reduce social isolation and improve quality of life.

Doodle Me is designed to provide a solution that leverages AI techniques to interpret user-created doodles in real-time, converting them into recognizable outputs. The application uses a CNN for this purpose, providing a mode of expression that can be understood across different languages. The application is developed with a focus on usability, ensuring it is accessible for users of various ages and technical abilities. A key aspect of this project is the development of a cross-platform mobile application that works on both Android and iOS devices, thereby reaching a broad audience. The implemented machine learning algorithm uses a dataset from Google Quick, Draw! [8] to train the CNN model. Feedback from initial users shows its design allows for quick learning, making it a practical tool for everyday communication needs.

Inspired by real-world communication challenges, Doodle Me targets people with speech disabilities/disorders and non-native

language speakers. Doodles, simple user-made drawings that can convey ideas or emotions, emerged as a potential solution because of their universal understanding. By using doodles as a tool for augmentative and alternative communication, Doodle Me provides a practical and inclusive communication aid.

The objective of this project is to overcome cultural and linguistic barriers that hinder effective communication. By combining an easy-to-use interface with AI technology, Doodle Me facilitates communication, fosters inclusivity, and enhances quality of life. This paper discusses the development and implementation of Doodle Me, highlighting its potential as an assistive communication tool for individuals facing communication challenges worldwide.

2 RELATED WORK

The rapid advancement of mobile technologies has significantly improved assistive communication applications. Several existing solutions have been helpful for individuals with disabilities. Proloquo2Go, for instance, supports non-speaking individuals through symbols and text-to-speech functionalities [12]. However, it requires navigating numerous buttons to construct messages, which can be time-consuming and may limit its effectiveness for quick and simple communication. Similarly, applications like AutoDraw and QuickDraw use AI for visual recognition [1, 11]. AutoDraw enhances sketches, while QuickDraw identifies doodles in a game-like setting. Although these tools are primarily designed for entertainment, they demonstrate AI's potential in visual communication but do not focus on accessibility for those with communication disabilities. Research by Vaezipour et al. highlights a gap in previous mobile applications, noting a lack of interactive and engaging elements for users [17]. Doodle Me aims to address this gap with its user-friendly design and doodle recognition capabilities, aiming to provide a more practical and accessible communication tool.

Technically, the efficacy of different machine learning models in doodle recognition has been extensively explored. Studies by Guo, WoMa, and Xu compared k-nearest neighbors (KNN) and CNN for classifying doodles, finding that CNNs, with an accuracy of 60%, outperformed KNNs, which achieved only 35% [9]. Fortuna-Cervantes' research on CNN models with transfer learning for Art Media Classification reached an accuracy of 86%, highlighting the potential for further improvement [7].

The use of AI in assistive communication has also been explored in various studies. Floriano, Brito, and Rocha's work on Artificial Neural Networks for severe neuromotor disorders emphasized the importance of a user-centered approach [6]. Their findings on the efficacy of facial recognition technology in Augmentative and Alternative Communication (AAC) resonate with our aim to tailor the app experience to user needs. This approach ensures that the technology is accessible and effective for users with different types of communication impairments. Sigafoos' research on AAC for individuals with intellectual and developmental disabilities underscores the pivotal role of computer-based AAC devices in enhancing communication for those with severe impairments [16].

By integrating AI-based doodle recognition with a user-friendly interface, Doodle Me aims to address the limitations of previous applications and improve user interaction. This application simplifies the user experience, making it accessible for people with diverse communication needs and provides a practical solution that previous applications have not deliver.

3 ANALYSIS OF DOODLE ME

This project falls within the realm of Human-Computer Interaction (HCI), focusing on mobile solutions to enhance communication. Doodle Me aims to bridge communication gaps for individuals with speech disabilities/disorders and non-native language speakers by leveraging mobile technology and advanced AI.

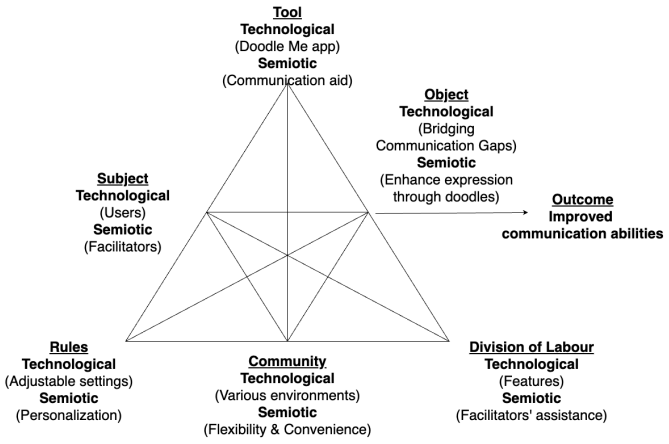


Figure 1: Theoretical Framework based on Activity Theory for Doodle Me

3.1 Theoretical Framework

Communication is essential for everyone, especially for individuals with speech-related disabilities and non-native language speakers. To create a tool that meets their needs, it's crucial to ground the development process in a theoretical framework. Grounded in Activity Theory, we analyzed Doodle Me as an activity system. In this framework, mobile devices act as interactive tools that enable users to communicate by interpreting doodles and converting them into understandable words and images. The various use cases of Doodle Me span everyday interactions at home, educational settings, healthcare environments, and social contexts. By supporting activities such as learning, therapy, daily communication, and social integration, Doodle Me follows the principles of Activity Theory, which focuses on the interaction between users, tools, and community to achieve outcomes. The following components have been developed based on this theoretical approach, as illustrated in Figure 1.

- *Subject.* The primary users of the Doodle Me app include individuals with speech-related disabilities and non-native language speakers. Facilitators, such as caregivers and speech therapists, assist users with the app.
- *Object.* The main objective is to provide a communication aid that helps users express themselves through doodles, addressing gaps caused by speech impairments or language barriers.
- *Tool.* The Doodle Me app, integrating AI-driven doodle recognition with an easy-to-use interface, serves as the essential tool. It operates on both Android and iOS platforms.

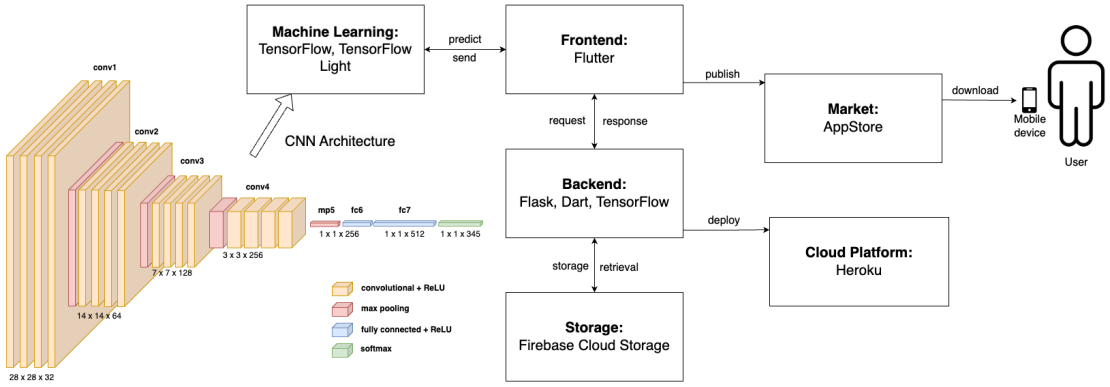


Figure 2: System Architecture of Doodle Me

- *Rules.* Users can personalize their experience through adjustable language settings and app themes.
- *Community.* The app is designed for various environments, including homes, schools, hospitals, and rehabilitation centers.
- *Division of Labour.* The app supports effective communication through features like real-time doodle recognition, word pronunciation, alternative predictions, and a history feature for recalling past communications.
- *Outcome.* The expected outcome is an improved ability for users to communicate their thoughts and needs, reducing social isolation and improving the overall quality of life for individuals with communication impairments.

3.2 Expert Feedback

Expert opinions have played a pivotal role in the development of the Doodle Me app. We have had the privilege of collaborating with Dr. Salim, a distinguished medical doctor with clinical experience in Otolaryngology. She brought invaluable insights to our project framework. During the initial phases, she meticulously identified various gaps in existing communication tools and proposed that additional features like incorporating a wider range of doodles that could significantly benefit users across the globe. Speech disorders may have variable presentations in patients. They range from mild form of disability like slurred speech(Bell's palsy), repetition of sounds(stuttering), husky voice(vocal cord paresis) to a total inability to speak(post-ischemic stroke) or a complete deafness(congenital sensorineural hearing loss). These liguistic conditions can be dealt by using Doodle Me. The application acts as a tool by providing an easy and independent communicative ability where users can interact any community member without any speech therapist or facilitator. The users are facilitated with a key feature to listen to a word's pronunciation repetitively and can also save their search in history for remembrance. The verbal and cultural gaps can be reduced by using this application as it helps the users to communicate in a simple language despite of any region or ethnicity.

4 DESIGN AND DEVELOPMENT OF DOODLE ME

Based on the development and implementation of Doodle Me, we propose this app as a technological solution to bridge communication gaps. It aims to enhance ease of communication for individuals with linguistic disabilities and non-native language speakers. It is an assistive AI tool that leverages doodle recognition by smartphones.

The design process began with user research and feedback analysis to understand potential users' needs and challenges. We used the MoSCoW (Must have, Should have, Could have, Won't have) method to prioritize features and ensure critical functionalities were developed first. This helped us create a Minimum Viable Product (MVP) that could be tested and iterated upon. The UX design started with prototypes developed in Miro, allowing us to visualize and refine the user interface and interactions before implementation [14].

4.1 System Architecture and Implementation Details

Creating Doodle Me involved innovation and user-focused design. Our aim was to develop a cross-platform mobile app with an AI component to recognize doodles. Figure 2 demonstrates the system architecture of Doodle Me.

We started by building the mobile app using Flutter, which allowed us to create a single codebase that works on both Android and iOS devices, giving users a consistent experience across different platforms.

The core functionality of the app is the AI component. We implemented a CNN because of its ability to recognize patterns in images, which is essential for identifying doodles. Using TensorFlow, we built and trained our neural network models with the Google Quick, Draw! dataset, which contains over 50 million drawings across 345 categories. To ensure our model was accurate, we split the dataset into training, validation, and testing sets. We repeatedly adjusted the model, tweaking parameters and layers based on the validation set results. As shown in Figure 4, the training and validation loss decreased consistently, and the accuracy improved with each epoch. Finally, we achieved an accuracy of 92.6% on the test dataset.

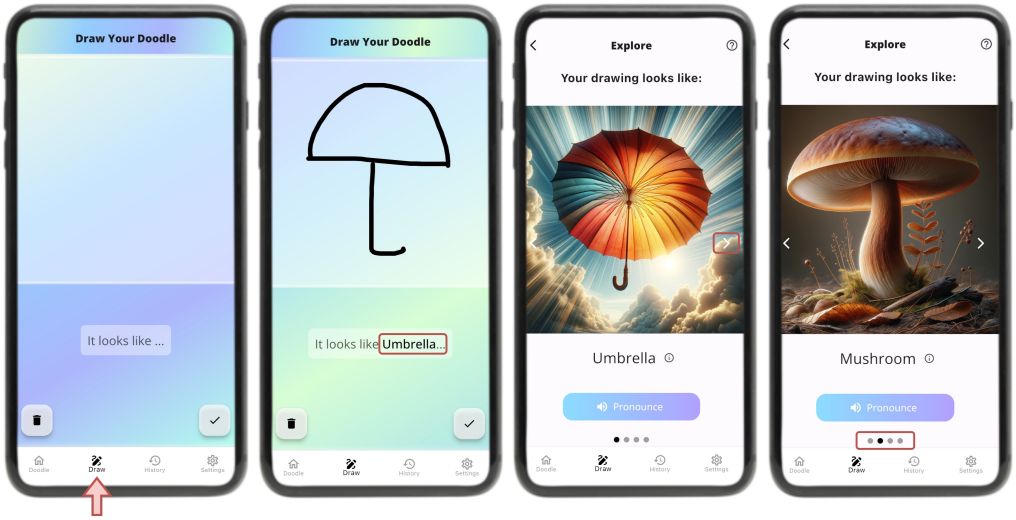


Figure 3: Screenshots of the Doodle Me app illustrating the doodle recognition process, from left to right: the drawing board where users create their doodles, real-time prediction displayed as the user draws, the predicted word and image with a button to listen to the pronunciation, and alternative predictions provided when the initial guess does not match the user's intended doodle

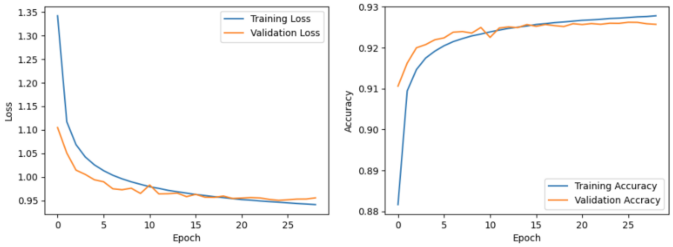


Figure 4: Training and validation loss and accuracy

To make the model work on mobile devices, we converted it to TensorFlow Lite, a lighter version of TensorFlow designed for mobile and embedded devices. This conversion was essential for making real-time predictions in the app.

For consistency in the input data for our neural network models, we developed a backend server using Flask and deployed it on Heroku. The Flask backend preprocesses raw doodle data received from the mobile app, converting raw vector data into 28x28 pixel grayscale images.

We used Firebase Cloud Storage for data management to store and retrieve visual representations of each class. AI-generated images corresponding to the 345 categories are stored in Firebase. Once a doodle is classified, these images are retrieved from Firebase, providing users with visual confirmation of the AI's interpretation.

4.2 Key Features and Functionalities of Doodle Me

Building on the design and functionality of Doodle Me, we have developed key features that enhance user experience and communication capabilities. Each feature was designed with a focus on user needs and usability, aligned with the principles of Activity Theory. These features include:

Doodle Recognition. When users open the Doodle Me app, they are presented with a drawing screen that includes a drawing board, as illustrated in Figure 3. Here, users can freely create their doodles. For example, a user might draw a simple umbrella. As they draw, the app continuously analyzes the doodle and predicts the most likely word that corresponds to the drawing, displaying this prediction in real-time. In this case, it would display "Umbrella". After completing their doodle, users can press the "OK" button. This action reveals an image of the predicted word alongside the word itself. For instance, the image of an umbrella would appear along with the word "umbrella". Additionally, users have the option to listen to the pronunciation of the word, aiding in verbal communication. The predicted image also includes detailed information, such as the accuracy percent of the prediction (see Figure 4). If the visual and verbal feedback aligns with what the user intended to convey, they can confidently use it to communicate their message. In cases where the initial prediction does not match the user's intended doodle, the app offers a solution. This feature acts as the primary tool within the Activity Theory framework, which enables the interaction between the user and the app. For example, if the user intended to draw a mushroom but the app predicted an umbrella, the user can scroll to the right to view up to five alternative predictions, such as "mushroom", ranked by their likelihood. This ensures that even if the first guess is incorrect, users have multiple options to find the correct match for their drawing.

Doodle Library. In addition to the drawing features, Doodle Me includes a library of objects. This library acts as a reference guide, showcasing the variety of objects that the AI can recognize (see Figure 5). Users can browse through the library to see examples of images that the system can identify. Users can easily search for

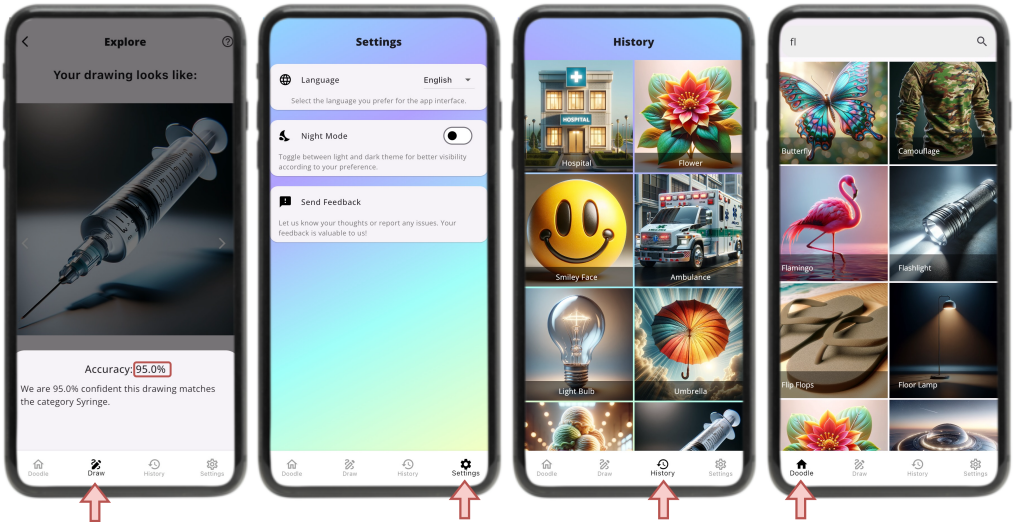


Figure 5: Screenshots of the Doodle Me app illustrating additional features, from left to right: accuracy percentage of the predicted doodle, settings for personalization, history of past predicted doodles, and the library of objects

any object within the library. This feature allows them to quickly find and view specific examples of objects the app can identify. The library contains 345 different objects, providing users with a view of what the app is capable of recognizing and helping them create doodles that are more likely to be accurately identified by the AI.

Personalization. For a personalized experience, Doodle Me interface allows users to set their preferences, including language settings and app themes. The app supports a variety of languages, making it accessible to a global audience (see Figure 5). Users can choose their preferred language to suit their linguistic needs. Additionally, users can also select between dark and light themes to customize the app's appearance to their liking. This personalization feature lets each user tailor the app to their specific needs and preferences, which aligns with the Activity Theory concept of rules.

History. Doodle Me includes a history feature where users can view their past predicted doodles. This feature is especially helpful for users who need to recall or show objects multiple times without redrawing them (see Figure 4). By accessing their history, users can easily find and use previous predictions. This is particularly beneficial for patients with Alzheimer's disease (memory loss/dementia), as it helps them retain and recall memories through revisiting and communicating their past doodles. It will also benefit patients with Facial paralysis (Bell's palsy) who have diminished speaking skills due to weakness in facial muscles. Their speech is difficult to understand and recognize that leads to a communication barrier. This application will provide help and ease the interactive ability.

5 EVALUATION AND DISCUSSION

The performance of Doodle Me was evaluated using quantitative and qualitative methods. We analyzed collected data from various analyses and user feedback surveys to provide insights into the app's capabilities.

The AI model achieved an accuracy rate of 92.6% when tested on the QuickDraw! dataset. Table 1 illustrates a comparative analysis of different models. This model's accuracy signifies an improvement over other models trained on the same dataset, where the highest recorded accuracy was 86% for another CNN implementation. We attribute this gain in performance to several factors, including the larger and more complex structure of our CNN and more extensive hyperparameter optimization.

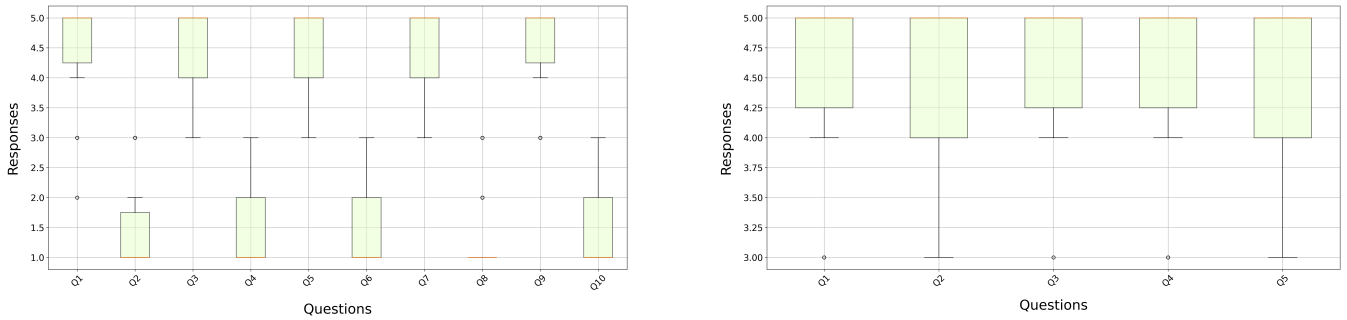
Model	Doodle Me CNN	CNN [9]	KNN [9]	CNN [7]
Accuracy	92.6%	60%	35%	86%

Table 1: Performance Comparison of Different Models

Further performance metrics include the F1 Score (Macro), which stood at 0.76, which indicates balanced precision and recall across all categories. Additionally, the Area Under the Curve (AUC) from the Receiver Operating Characteristic (ROC) curve was calculated to be 0.99, which underscores the model's capability to distinguish between different classes accurately.

A user feedback survey was conducted with 26 participants to gather insights into the app's usability and user satisfaction. The participants included individuals of different ages, non-native language speakers, and people who work with individuals with speech-related disabilities. The System Usability Scale (SUS) score achieved was 91 (see Figure 6a), indicating a positive response to the app's design and functionality. The Mean Opinion Score (MOS) was reported to be 4.65 (see Figure 6b), demonstrating generally favorable feedback from users.

The study results provide empirical evidence on the effectiveness of the Doodle Me application. The accuracy and performance metrics of the AI model with positive user feedback, indicate that the app meets its objectives of interpreting and analyzing user-created doodles and highlights the app's potential as a valuable tool for both general public use and for individuals with speech disorders. This study not only validates the concept but also sets a foundation for future enhancements and wider adoption of the application in various settings which includes healthcare and education.



(a) System Usability Scale Survey Results

(b) Mean Opinion Score Survey Results

Figure 6: Survey Results

6 CONCLUSION

This paper presents the development and implementation of an AI mobile app named Doodle Me. The application uses machine learning algorithms, specifically a CNN, to transform user-created doodles into recognizable communication outputs.

Doodle Me has demonstrated potential in enhancing communication and has been tested with various users, including the general public and individuals with speech and linguistic disabilities. For instance, postoperative cochlear implant patients in rehabilitation centers can benefit from this application [2, 15]. It enables users to express themselves easily, reducing emotions like fear and anxiety. Doodle Me fosters social inclusion and interaction. User feedback and engagement rates indicate the app's practicality and relevance in real-world scenarios, with potential to improve users' quality of life.

The implications extend beyond individual user benefits to broader societal impacts. This application shows the potential for similar tools to be developed in other areas where communication barriers exists, such as education, healthcare and rehabilitative centres. In the short term, Doodle Me may assist individuals in overcoming daily communication challenges. Over the long term, the app could contribute to improving communication skills and building confidence.

However, this project is not without limitations. The app has not yet been published on Android due to a lack of resources, which restricts its reach to a broader audience. Additionally, the number of doodle categories is limited to 345 due to the constraints of the dataset used. While the app has shown high accuracy in recognizing doodles, there is always room for improvement in the AI's ability to interpret more complex or abstract drawings. These limitations highlight the potential for future work to ensure that Doodle Me can serve a diverse and broad user base.

ACKNOWLEDGMENTS

This research work has emanated from research conducted with financial support from Science Foundation Ireland under Grant 12/RC/2289-P2 and 18/CRT/6223.

REFERENCES

[1] AutoDraw. 2019. *AutoDraw.* https://www.autodraw.com/

[2] Laila Banu, Tania Saad, Md Hossain, and Md Hossain. 2024. Speech Language Disorder in Children: An Overview. *Journal of Comilla Medical College Teachers' Association* 27 (02 2024), 75–82. https://doi.org/10.3329/jcomcta.v27i2.71568

[3] Lena Brandt, Sihong Liu, Christine Heim, and Andreas Heinz. 2022. The effects of social isolation stress and discrimination on mental health. *Translational Psychiatry* 12, 1 (2022), 398. https://doi.org/10.1038/s41398-022-02178-4

[4] Salvatore Cocuzza, Marco Bonfiglio, Calogero Grillo, Luigi Maiolino, Mariano Malaguarnera, Francesco Martines, and Agostino Serra. 2013. Post laryngectomy speech rehabilitation outcome in elderly patients. *European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery* 270 (03 2013). https://doi.org/10.1007/s00405-013-2430-3

[5] Irish Collective and Irwin Gill. 2023. Rehabilitation for Children and Young People in Ireland Following Acquired Brain Injury: current services and potential future directions. (11 2023).

[6] Carolina Floriano, Leonardo Brito, and Adson Rocha. 2024. Alternative Communication System for People with Severe Neuromotor Disorders using Artificial Neural Networks. 1–7. https://doi.org/10.21528/CBIC2023-026

[7] J. Fortuna-Cervantes, Carlos Soubervielle-Montalvo, Cesar Puente-Montejano, Oscar Perez-Cham, and Rafael Peña. 2024. Evaluation of CNN Models with Transfer Learning in Art Media Classification in Terms of Accuracy and Class Relationship. *Computación y Sistemas* 28 (03 2024). https://doi.org/10.13053/cys-28-1-4895

[8] Google Inc. 2017. Quick, Draw! Dataset. https://creativecommons.org/licenses/by/4.0/. Available under the Creative Commons Attribution 4.0 International License. Accessed: 2023-07-25.

[9] Kristine Guo, James WoMa, and Eric Xu. 2018. Quick, Draw! Doodle Recognition. Stanford University.

[10] Dragos Jianu, Tihomir Ilić, Silviana Jianu, Any Axelerad, Claudiu Bîrdac, Flavius Dan, Anca Gogu, and Georgiana Munteanu. 2021. *A Comprehensive Overview of Broca's Aphasia after Ischemic Stroke.* https://doi.org/10.5772/intechopen.101560

[11] Google Labs. 2019. *Quick, Draw!* https://quickdraw.withgoogle.com/

[12] Nicole Lambright. 2023. *Proloquo2Go (P2G) in the Classroom: Providing Speech, Communication, Behavioral Support, Academic, and Social Support for Students.* 140–155. https://doi.org/10.4018/978-1-6684-6424-3.ch007

[13] Sharynne McLeod. 2018. Communication rights: Fundamental human rights for all. *International journal of speech-language pathology* 20, 1 (2018), 3–11. https://doi.org/10.1080/17549507.2018.1428687

[14] Miro. [n. d.]. Miro: Online Collaborative Whiteboard Platform. https://miro.com/. Accessed: 2023.

[15] Mina Moradi, Masoud Fallahi-Khoshknab, Asghar Dalvandi, Mohammad Farhadi, Sadat Maddah, and Eesa Mohammadi. 2021. Rehabilitation of children with cochlear implant in Iran: A scoping review. *Medical Journal of The Islamic Republic of Iran* 35 (04 2021). https://doi.org/10.47176/mjiri.35.73

[16] Jeff Sigafoos, Larah van der Meer, Ralf Schlosser, Giulio Lancioni, Mark O'Reilly, and Vanessa Green. 2016. *Augmentative and Alternative Communication (AAC) in Intellectual and Developmental Disabilities.* 255–285. https://doi.org/10.1016/B978-0-12-802075-3.00010-3

[17] Atiyeh Vaezipour, Jessica Campbell, D. Theodoros, and Trevor Russell. 2020. Mobile Apps for Speech-Language Therapy in Adults With Communication Disorders: Review of Content and Quality. *JMIR mHealth and uHealth* 8 (10 2020), e18858. https://doi.org/10.2196/18858