MINISTRY OF EDUCATION AND RESEARCH OF THE REPUBLIC OF MOLDOVA

Technical University of Moldova

Faculty of Computers, Informatics, and Microelectronics

Department of Software Engineering and Automation

Approved for defense
Department head:
Ion FIODOROV, PHD, associate professor

„___" _____ 2025

SOLUTIONS TO IMPROVE COGNITIVE ASSISTANCE FOR PEOPLE WITH MENTAL DISABILITIES BASED ON NEURAL NETWORKS

Master's project

Student: _____ **Cerlat Pavel, IS-231M**

Coordinator: _____ **Cojuhari Irina, associate professor**

Consultant: _____ **Cojocaru Svetlana, university assistant**

Chișinău, 2025

ABSTRACT

The research involves the use of a neural network doodle and gesture recognition with emotional state detection to assist mentally disabled people. It involves the designing of the use of convolutional neural networks in an image-based input that will help in interpreting doodles being drawn by the user and recognizing hand gestures. This is further integrated with a recurrent neural network that includes long short-term memory units for the analysis of temporal data associated with gestures and emotional expressions. This emotional state detection uses facial expression and gesture pattern recognition to fetch the emotional state, which, after processing in real time, sends the intake to either a caregiver or parents via a mobile application showing the state of the individual's emotions. The proposed system architecture includes a pre-trained CNN model with custom-trained layers integrated for fine tuning on specific emotions relevant for the target population. Preliminary results taken from already-existing neural network results give a high accuracy in doodle recognition at 92% precision, while for gesture recognition, an accuracy rate of 98.44% was given. The emotional state detection feature can classify happiness, sadness, and distress. It is an integrated system whereby the capability of communication is assisted both by doodle and gesture recognition, adding emotional monitoring. Such a system would be very complete in order to improve life quality for mentally disabled people and their caregivers.The study concludes by discussing how this system has the potential for real-world deployment and further development to include a wider range of emotions and more sophisticated gestures. In the future, the work will focus on making this model much more robust with larger training datasets and real-time processing optimizations.

REZUMAT

Cercetarea implică utilizarea unei rețele neuronale pentru recunoașterea schițelor și a gesturilor, împreună cu detectarea stării emoționale, pentru a ajuta persoanele cu dizabilități mintale. Aceasta implică proiectarea și utilizarea rețelelor neuronale convolutionale într-o intrare bazată pe imagini. Sistemul folosește două modele CNN, unul specializat pentru recunoașterea schițelor și unul pentru recunoașterea gesturilor. Modelul pentru gesturi este antrenat pe un set de date de gesturi din limbajul semnelor, dar poate fi ușor antrenat și pentru alte tipuri de gesturi. Aceste modele sunt integrate cu o rețea neuronală recurentă care include unități de memorie pe termen lung pentru analiza datelor temporale asociate cu gesturi și expresii emoționale. Detectarea stării emoționale folosește expresii faciale și recunoașterea modelelor de gesturi pentru a obține starea emoțională, care, după procesarea în timp real, trimite informațiile fie unui îngrijitor, fie părinților printr-o aplicație mobilă care arată starea individului. Arhitectura de sistem propusă include un model CNN pre antrenat, cu straturi personalizate integrate pentru reglarea fină a emoțiilor specifice relevante pentru populația țintă, oferind precizie. Pentru recunoașterea gesturilor, s-a obținut o rată de acuratețe de 98,44%. Funcția de detectare a stării emoționale poate clasifica fericirea, tristețea și suferința. Este un sistem integrat prin care capacitatea de comunicare este asistată atât de recunoașterea schițelor, cât și de recunoașterea gesturilor, completate de monitorizarea emoțională. Un astfel de sistem ar fi foarte util pentru a îmbunătăți calitatea vieții persoanelor cu dizabilități mintale și a îngrijitorilor acestora. Studiul se încheie discutând despre modul în care acest sistem are potențialul de implementare în lumea reală și de dezvoltare ulterioară pentru a include o gamă mai largă de emoții și gesturi mai sofisticate. În viitor, se propune să se concentreze pe crearea unui model mult mai robust, cu seturi de date de antrenament mai mari și optimizări de procesare în timp real.

CONTENTS

List of Abbreviations:

WHO	World Health Organization
AI	Artificial Intelligence
CNN	Convolutional Neural Network
LSTM	Long Short-Term Memory
VGG	Very Deep Convolutional Network
ResNet	REsidual Neural Network
RNN	Recurrent Neural Network
ID	Intellectual Disabilities
QoL	Quality of Life
HCI	Human-Computer Interaction
AAC	Augmentative and Alternative Communication
SGD	Speech Generating Devices
VOCA	Voice Output Communication Aids
ASL	American Sign Language

INTRODUCTION

According to the World Health Organization (WHO), over a billion people worldwide live with some form of disability, including nearly 240 million children whose well-being is compromised [1]. As highlighted by UNICEF, factors such as the aging global population, the emergence of new diseases, and an increasing trend in chronic illnesses are contributing to a growing number of individuals living with disabilities. These individuals often require assistive solutions to help them lead independent lives with dignity.

Historically, assistive technologies have been developed to address various challenges faced by people with disabilities. Early examples include the walking cane for individuals with mobility issues, dating back to the 16th century, and the mass production of wheelchairs that began in 1933 (Woods & Watson) [1]. Sensory assistive devices, such as glasses for partially sighted individuals and ear trumpets for those with hearing impairments, laid the groundwork for modern assistive technologies.

In the 19th century, significant strides were made in the education of individuals with disabilities. Louis Braille's [2] invention of the Braille alphabet enabled blind individuals to learn independently. The advent of talking books and electronic text-to-speech synthesizers further improved access to information for the visually impaired.

In the modern era, artificial intelligence (AI) plays a critical role in the development of assistive technologies. AI has paved the way for innovations such as augmented and virtual reality, AI-based gesture recognition, and brain-computer interfaces. AI-supported systems enhance communication and assist individuals with cognitive, sensory, and physical impairments. For example, gesture-based text prediction applications that use AI are particularly useful for individuals with hearing impairments, while deep learning models have advanced the capabilities of EEG-based brain-computer interfaces. Despite the significant progress in assistive technologies, there is still much room for improvement, particularly in creating affordable, user-friendly, and accessible solutions [1].

This master's project proposes a novel solution that leverages the latest advancements in AI and machine learning to assist individuals with mental disabilities in communicating through drawings, gestures, and emotional expressions. The project will utilize convolutional neural networks (CNNs) and long short-term memory (LSTM) recurrent neural networks to develop an assistive system capable of recognizing and interpreting user inputs. The ultimate goal is to improve the quality of life for individuals with disabilities, enabling them to express themselves and communicate in more accessible ways, while also providing emotional support to their caregivers and families.

BIBLIOGRAPHY

[1] Katerina Zdravkova, Venera Krasniqi, Fisnik Dalipi,Mexhid Ferati, Cutting-edge communication and learning assistive technologies for disabled children: An artificial intelligence perspective, National Center for Biotechnology Information, 2022.

[2] Gaudenz Boesch, Very Deep Convolutional Networks (VGG) Essential Guide, viso.ai, 2021.

[3] Wenchao Gu,Lingxing Kong, Residual Neural Network, Image and Vision Computing, 2022.

[4] Nico Klingler, AlexNet: A Revolutionary Deep Learning Architecture, viso.ai, 2024.

[5] Nicole Agaronnik, Eric G. Campbell, Julie Ressalam, Lisa I. Iezzoni, Communicating with Patients with Disability: Perspectives of Practicing Physicians, National Center for Biotechnology Information, 2019, pp. 1-7.

[6] AUSFDO Editorial Team, Communication with people with disabilities, Australian Federation of Disability Organisations. 2024 pp. 2-3.

[7] Martine Smith, Beatrice Manduchi, Éilish Burke, Rachael Carroll, Philip McCallion, Mary McCarron Communication difficulties in adults with Intellectual Disability: Results from a national cross-sectional study,National Center for Biotechnology Information, 2019, pp-1.

[8]Juan Carlos García, Emiliano Díez, Dominika Z. Wojcik,Mónica Santamaría, Communication Support Needs in Adults with Intellectual Disabilities and Its Relation to Quality of Life, National Center for Biotechnology Information, 2020, pp. 6,8-10.

[9] Hilari K., Needle J.J., Harrison K.L. What Are the Important Factors in Health-Related Quality of Life for People With Aphasia? A Systematic Review. Arch. Phys. Med. Rehabil, 2022, pp-5.

[10] Lei Zhang, Hand-drawn sketch recognition with a double-channel convolutional neural network, EURASIP Journal on Advances in Signal Processing, 2021, pp. 2-6.

[11] Norah Meshari Alnaim, Hand Gesture Recognition using Deep Learning Neural Networks, Brunel University London, 2019.

[12] Phani Pajeet, What is the Convolutional Neural Network Architecture, 2023, pp. 1.

[13] Liufeng Fan, Zhan Zhang, Biao Zhu, Decheng Zuo, Xintong Yu and Yiwei Wang, Smart-Data-Glove-Based Gesture Recognition for Amphibious Communication, University of Science and Technology of China, 2023, pp- 3.

[14] Emily K. Unholz-Bowden, Shawn N. Girtler, Alefyah Shipchandler, Rebecca L. Kolb, Jennifer J. McComas Use of Augmentative and Alternative Communication by Individuals with Rett Syndrome: High-Tech and Low-Tech Modalities, National Center for Biotechnology Information. 2023, pp. 1-21.

[15] N. Sriram, M. Nithiyanandham, A hand gesture recognition based communication system for silent speakers, 2013 International Conference on Human Computer Interactions (ICHCI), 2014.

[16] What is the purpose of the game "Quick, Draw!" created by Google?, EITC/AI/GCML Google Cloud Machine Learning, 2023.

[17] Klimis Symeonidis, Hand Gesture Recognition Using Neural Networks, IEEE Xplore, 2021.
[21] Worked example of 2x2 max pooling with stride 2. Zhang, Aston; Lipton, Zachary; Li, Mu; Smola, Alexander J. (2024). "7.5. Pooling". Dive into deep learning. Cambridge New York Port Melbourne New Delhi Singapore: Cambridge University Press. ISBN 978-1-009-38943-3.
[22] Will Kenton, How to Perform a SWOT Analysis, investopedia.com/terms/s/swot.asp, 2024.
[23] MyDisabilityJobs, “Intellectual Disability Employment Statistics” , Accessed: Dec. 15, 2024. [Online]. Available: https://mydisabilityjobs.com/statistics/intellectual-disability-employment.
[24] Prankur Haldiya, “Google Play Store Statistics you must consider before building your app in 2021 & 2022”, Accessed: Dec. 16, 2024. [Online]. Available: https://ripenapps.com/blog/google-play-store-statistics/.
[25] Research and Markets, “Global Augmentative & Alternative Communication (AAC) Devices Market Analysis 2024-2029 by Device Type, Application, End-User, Region, and Country”, Accessed: Dec. 16, 2024. [Online]. Available: https://www.globenewswire.com/news-release/2024/01/15/2809369/0/en/Global-Augmentative-Alternative-Communication-AAC-Devices-Market-Analysis-2024-2029-by-Device-Type-Application-End-User-Region-and-Country.html.
[26] Tai Le, “Mobile App Development Costs 2023”, Accessed: Dec. 16, 2024. [Online] Available: https://kms-solutions.asia/blogs/factors-impacting-mobile-app-development-costs.