Artificial Intelligence for Assisting People with Sensory and Cognitive Disabilities

Kostandina Veljanovska[1], Simona Gulevska[1] and Blagoj Ristevski[1]

[1] *University St. Kliment Ohridski – Bitola, Faculty of Information and Communication Technologies, Republic of North Macedonia*

kostandina.veljanovska@uklo.edu.mk; gulevska.simona@uklo.edu.mk; blagoj.ristevski@uklo.edu.mk

Abstract:

This paper explores and compares various applications of artificial intelligence (AI) to support people with sensory and cognitive disabilities. A detailed analysis and classification of existing intelligent solutions were conducted according to their primary functions. Furthermore, a convolutional neural network (CNN) for image classification was implemented and evaluated. The results indicate that appropriately selected AI technologies can provide functional support, and significantly improve the quality of everyday life for individuals with diverse disabilities. The study highlights both the potential and practical value of integrating AI-based systems in assistive technologies, emphasizing their role in enhancing accessibility, autonomy, and inclusion.

Keywords:

Artificial intelligence, assistive technologies, sensory disabilities, cognitive disabilities, convolutional neural networks

1. Introduction

In recent decades, artificial intelligence has opened new opportunities for the integration of people with disabilities into all spheres of life. A particularly important role in this transformation is played by intelligent systems, which enable automation, adaptation, and personalization of services and content. The term *intelligent systems* (IS) *[1]* refers to software and hardware solutions based on algorithms from artificial intelligence, machine learning (ML), and natural language processing (NLP), designed to improve user interaction and increase functional independence.

This paper focuses on intelligent systems aimed at supporting individuals with sensory and cognitive disabilities. Sensory disabilities include impairments of vision, hearing or sensory integration while cognitive disabilities involve conditions that affect learning, memory, communication, and decision-making. These difficulties often create significant barriers to accessing educational resources (e.g., the unavailability of braille textbooks or challenges in following digital lectures), communication tools (such as speech or visual applications without alternative modes of expression), and daily activities (including mobility, device management, or performing tasks at home and in public spaces).

Therefore, there is a growing need for inclusive intelligent systems that apply adaptive interfaces, multimodal interactions, and personalized algorithms to meet the diverse cognitive and sensory needs of these individuals. Such challenges can be significantly mitigated through the application of appropriate technological solutions that leverage the advantages of modern artificial intelligence.

2. Intelligent Systems Supporting People with Disabilities

Intelligent systems play an increasingly important role in facilitating the daily lives of people with disabilities by providing innovative technological solutions based on artificial intelligence. These systems offer practical tools that enhance communication, navigation, access to information, and the performance of everyday tasks. A classification of such systems according to their primary functionality is presented below. There are scientific studies that underline the importance of collaboration among several disciplines in order to create IS that promotes equity and empowerment, such as: AI specialists, technologists, policymakers, healthcare providers, and also, end-users [2]. Some of the scientists take into account services that can improve life quality, minimize environmental impacts, improve health service, improve security and accessibility for People with Disabilities, and therefore they propose a Model for Assistive Smart Cities [3], or implement identity recognizer, facial expression recognizer, gesture recognizer, vocalization recognizer as part of IS [4]. There are IS where AI holds transformative potential in advancing equitable and inclusive healthcare, promoting autonomy, accessibility, and well-being for people with disabilities and at the same time addressing ethical challenges, overcoming limitations, and fostering user-centered design [5, 6, 7], at the same time having in mind disability – friendly AI implementation according to the "rules to ensure that artificial intelligence systems are used safely and ethically, with a particular focus on protecting human rights" [8].

2.1. Classification of IS by Function

IS supporting people with disabilities can be classified into several categories according to their primary function. In order to research their way of work and perform comparative analysis of IS we use the most important implementations of those systems already applied across North and South America, Europe, South Africa, Asia, and Australia [9, 10]. They can be classified into the few categories:

Communication support IS are a category that consists of applications and devices designed to facilitate communication for individuals with various cognitive and motor impairments. These systems provide alternative methods of expression and interaction by employing advanced technologies such as speech processing, brain–computer interfaces (BCI), augmented reality (AR), and text-to-speech (TTS) synthesis. Some of the systems in this category include [11]:

- *Expressia* – facilitates communication for non-speaking individuals
- *TelepatiX* – generates speech through blinking control
- *Voiceitt* – recognizes non-standard speech
- *Cognixion ONE* – enables communication using brain signals and augmented reality
- *Prometheus* – translates brain signals and facial expressions into digital commands

Visual support IS are a category of IS designed for individuals with partial or complete visual impairment. These solutions rely on advanced sensory and cognitive technologies to enable improved orientation, object recognition, text reading, and safe navigation in the environment [12, 13, 14]. By combining artificial intelligence, computer vision, spatial audio, and mobile or wearable devices these systems provide practical support for daily living. Some of the systems in this category include:

- *Seeing AI* – describes the environment and text through the camera
- *Be My Eyes* – video assistance from volunteers
- *Aira* – navigation and spatial description via remote operators
- *Microsoft Soundscape* – audio navigation with 3D spatial sound
- *OrCam MyEye* – reads text and recognizes objects through a camera

Hearing support IS are a category of IS that provide inclusive communication through automatic translation, speech transcription, visual notifications, and digital sign language interpreters. Modern solutions leverage technologies such as automatic speech recognition (ASR), animated avatar generation, natural sign language synthesis, and mobile applications for real-time interaction. Some of these systems are based on machine learning algorithms and computer graphics, enabling highly accurate and context-aware translations that are essential for daily communication, and educational activities. Some of the systems in this category include:

- *ProDeaf* – translates speech into sign language (LIBRAS)
- *HandTalk* – automatically converts text and speech into sign language
- *Ava* – provides real-time speech transcription for deaf individuals

Motor support and accessibility IS are a category of intelligent solutions that provide innovative systems that enable interaction through alternative methods such as eye-tracking, voice control, head movement tracking, or direct brain–computer communication [15]. To overcome the challenges faced by individuals with motor impairments in using traditional devices and physically managing their environment these systems integrate computer vision, voice interaction, BCI technology, and personalization algorithms. Some of the systems in this category include:

- *eSSENTIAL Accessibility* – navigation through head movement or voice;
- *Tobii Dynavox* – device control via eye-tracking
- *Polly (Parrots Inc.)* – AI assistant for wheelchair users

Spatial and infrastructure support – this category of intelligent systems is of particular importance for individuals with disabilities, as the physical environment can directly facilitate or hinder their mobility, participation in public life and access to resources. These solutions support spatial orientation and accessibility awareness, representing a vital component in the creation of inclusive societies:

- *Wheelmap* – identifies wheelchair-accessible public spaces

3. Comparative analysis

3.1. Communication support systems

The intelligent solutions included in this category (Expressia, TelepatiX, Voiceitt, Cognixion ONE, and Prometheus) exhibit significant variation in their technological approaches, the complexity of their implementation, and the degree to which they adapt to the individualized needs of users. In Table 1 comparative analysis of these type of intelligent systems is shown.

Table 1:
Comparative analysis of communication support systems

System	Expressia	TelepatiX	Voiceitt	Cognixion ONE	Prometheus
Interaction method	Symbols, Icons, Touch	Eye blink	Speech (non-standard)	Brain signals + AR	EEG (electroencephalogram) + facial expressions
Technologies	TTS, Cloud Sync, AI adaptation	Eye tracking, Blink detection, TTS, AI	Custom ASR, ML, Internet of Things integration	Eye tracking, TTS, AR, EEG, Edge AI	EEG, AI, Eye tracking
Required level of Motor Control	Low	Very low	Average	None	None
Personalization	High	Average	High	High	High
Works offline	Yes	Partial	Yes	Yes	Yes
Platform	Android/iOS/ Web applications/ CRM solutions	Android/iOS	Android/iOS	Independent device	Cloud-based (Azure, Amazon)/ Local servers/ Internet of Things devices

All of the analyzed solutions incorporate some form of intelligent data processing, and provide alternative channels of communication, yet they differ in terms of technical sophistication and adaptability. Expressia and Voiceitt emerge as the most accessible options for broader user groups, whereas Cognixion ONE and Prometheus integrate advanced technologies that require more specialized configurations. TelepatiX represents a balanced compromise, particularly suitable for users with highly limited motor abilities.

When personalization and offline functionality are considered, the systems demonstrate an orientation toward individual needs, though the extent of adaptation, and the required initial setup vary. Consequently, the selection of a particular solution often depends on the type and degree of disability, as well as the availability of technical support, and the user's budget.

3.2. Visual support systems

The intelligent solutions in this category (Seeing AI, Be My Eyes, Aira, Microsoft Soundscape, and OrCam MyEye) take different approaches to supporting independence and spatial orientation for people with visual impairments. They combine mobile apps, wearable devices, and 3D audio technologies to provide information about the surrounding environment through sound or human guidance (shown in Table 2).

Table 2:
Comparative analysis of visual support systems

System	Seeing AI	Be My Eyes	Aira	Microsoft Soundscape	OrCam MyEye
Interaction method	Camera + Touch	Video call+ Speech	Video call + Smart glasses	3D audio navigation	Gestures + Pointing
Technologies	Computer vision, Optical character recognition, TTS	WebRTC, GPT-4, Push notifications	Cloud streaming, GPS, Smart glasses	Head Related Transfer functions (HRFT), Bluetooth	Edge AI, Optical character recognition, Computer vision, Gesture recognition
Required level of Motor Control	Low	Low	Low	Low	None
Personalization	Average	Low	Average	Low	Average
Works offline	Partial	No	No	Partial	Yes
Platform	iOS	Android/iOS	Android/iOS	iOS	Independent device

OrCam MyEye operates entirely offline, ensuring both privacy and uninterrupted availability, while Seeing AI offers partial offline functionality. In contrast, Be My Eyes and Aira rely on live human assistance, which can be highly effective but requires a stable internet connection. Microsoft Soundscape stands apart by adopting a unique navigation approach through 3D audio cues rather than visual or video-based interaction.

Systems that employ on-device intelligence (such as OrCam MyEye) provide the greatest autonomy, while human-assisted platforms (Be My Eyes and Aira) may be more advantageous in situations where contextual interpretation and accuracy are paramount.

3.3. Hearing support systems

The intelligent solutions in this category (ProDeaf, HandTalk, and Ava) use different technological approaches to enhance access to communication and information. They illustrate three practical solutions in this area: automatic transcription, sign language translation, and the use of visual interfaces.

All analyzed systems require only minimal motor control, which makes them accessible to a wide range of users. As shown in Table 3, ProDeaf and HandTalk primarily support visual communication via sign language, whereas Ava provides a textual channel through real-time speech transcription. HandTalk stands out for its level of personalization thanks to customizable avatars and web integrations. The main distinction lies in the communication medium: visual (ProDeaf and HandTalk) versus textual (Ava).

Table 3:
Comparative analysis of hearing support systems

System	ProDeaf	HandTalk	Ava
Interaction method	Text/Speech → Sign language	Text/Speech → Sign language	Speech → Text
Technologies	Text-to-Sign,3D avatar, TTS, STT	Text-to-Sign, Speech-to-Sign, 3D avatar, Cloud SDK	ASR, Speaker diarization, Cloud synchronization, Bluetooth
Required level of Motor Control	Low	Low	Low
Personalization	Average	High	Average
Works offline	Partial	Partial	Partial
Platform	Android/iOS	Android/iOS/Web	Android/iOS

All systems offer partial offline functionality, giving users a degree of flexibility. However, certain advanced features (such as cloud synchronization and group interactions) require an internet connection.

3.4. Motor support and accessibility systems

The intelligent solutions in this category (eSSENTIAL Accessibility, Tobii Dynavox, and Polly (Parrots Inc.)), are designed for users with limited or entirely absent hand function but they differ in their technological approaches, level of autonomy, and target user group. The comparative analysis is shown in Table 4.

Table 4:
Comparative analysis of motor support and accessibility systems

System	eSSENTIAL Accessibility	Tobii Dynavox	Polly (Parrots Inc.)
Interaction method	Voice, Head movement	Eye movement	Voice, Head movement
Technologies	Virtual Keyboard, Voice input, Head-tracking	IR eye-tracking, Augmentative and Alternative communication, TTS, ML	AI assistant, Computer vision, Head-tracking
Required level of Motor Control	Low	Very low	Low
Personalization	Average	High	High
Works offline	Partial	Partial	Partial
Platform	Android/iOS/Web	Independent device	Android/iOS

Tobii Dynavox employs highly precise eye-tracking technology, enabling direct device control without physical touch, while Polly combines mobility, artificial intelligence, and environmental interaction via IoT devices. eSSENTIAL Accessibility is the easiest to implement and serves as a software-based alternative accessible through existing devices.

Regarding personalization, Polly and Tobii Dynavox offer the greatest flexibility through adaptive algorithms and individual profiles, whereas eSSENTIAL Accessibility provides basic configuration options. Offline functionality is limited across all systems due to reliance on cloud services or remote communication.

4. Fundamental IS Architecture for Supporting People with Disabilities

Based on the previously analyzed intelligent solutions, it can be observed that assistive systems for individuals with diverse disabilities have fundamental architecture where several interrelated components manage data acquisition, processing, and feedback delivery, enabling the system to adapt to the specific needs of each user [16]. Regardless of their particular application, these systems typically feature a layered architecture, consisting of: an input layer, a processing layer, an adaptation and personalization layer, and an output layer. The in-depth analysis of the architecture shows that the input layer is responsible for receiving sensory, speech, visual, or textual data. Depending on the system, this may include cameras, microphones, eye-tracking devices, EEG sensors, or touch interfaces. The processing layer encompasses the intelligent logic of the system. It is typically based on artificial intelligence (AI), machine learning (ML), deep learning (DL), or natural language processing (NLP) algorithms. This layer performs tasks such as speech recognition, object classification, intention prediction, translation, or response generation. The adaptation and personalization layer enables the system to adjust based on the user profile, usage history, and contextual information. Techniques commonly applied in this layer include reinforcement learning, content-based filtering, or rules defined by caregivers or therapists. Finally, the output layer generates visual, auditory, or mechanical feedback (e.g., speech, text, or signals to IoT devices). It may include screens, speakers, haptic vibrations, or synchronization with other devices.

4.1. Key Technologies and Frameworks Utilized in the Reviewed Systems

The analysis of intelligent systems developed to support individuals with cognitive and sensory disabilities highlights a number of widely adopted technologies and software frameworks. Presented below are some of the most prominent technologies along with their respective applications in the design of assistive intelligent systems:

- *TensorFlow* and *PyTorch* – these frameworks are extensively utilized for implementing machine learning and deep learning algorithms
- *OpenCV* – a computer vision library applied in applications such as facial recognition, eye-tracking, and gesture analysis
- *Speech Recognition (ASR) libraries* – used for converting spoken language into text (such as Google Speech-to-Text, Mozilla DeepSpeech and Whisper (OpenAI))
- *Text-to-Speech technologies* – applied to generate synthesized speech (such as Google Text-to-Speech, Amazon Polly and Microsoft Azure TTS)
- *Natural Language Processing (NLP) tools* – libraries utilized for natural language understanding (such as Hugging Face Transformers and spaCy)
- *Internet of Things (IoT) and edge computing platforms* – facilitate the management of connected devices (e.g., in smart home environments) and enable real-time data processing close to the user (as exemplified in systems like Polly)
- *Brain–Computer Interface (BCI) SDKs* – applied for the analysis of EEG signals to support interaction based on brain activity (such as with Emotiv or OpenBCI)

5. Practical Implementation of ANN for Assisting People with Visual Disabilities

We would like to highlight the importance of protecting and ensuring the safety of people with visual disabilities (PVD). Devices typically used in IS for helping PVD include sensors like infrared, ultrasonic, and imagery to gather environmental data, which is then processed by machine learning techniques [17]. During the research an artificial neural network (ANN) model was created to perform image classification aiming to support assistive solutions for individuals with visual difficulties.

5.1. Model: Image Classification and Class Interpretation

The model performs image classification using a moderately complex Convolutional Neural Network (CNN) built with Keras and TensorFlow. It consists of three convolutional blocks with batch normalization, pooling, and dropout, followed by fully connected layers mapping features to output classes. The model was trained on the CIFAR-10 dataset for 30 epochs with validation on the test set. The evaluation included both accuracy and a confusion matrix to evaluate class-specific performance.

Accurate class interpretation is essential for assistive systems for individuals with visual impairments, as it provides the basis for generating accessible outputs, such as synthesized speech or haptic feedback.

The model was implemented in Python using the Jupyter environment via Google Colab. This setup facilitates straightforward access to the required libraries, GPU resources for accelerated computation and easy sharing of results.

5.1.1. Results of the CNN model for image classification

The CNN model, designed for image classification, was trained on the CIFAR-10 dataset and it was evaluated twice. In the first evaluation, the model was tested over 25 epochs, achieving a test accuracy of 83.17%. In the second evaluation, the model was tested over 30 epochs achieving a test accuracy of 84.17%.

The model correctly predicted the class of the examined image as dog in both cases, which matches the true label.

This demonstrates the model's capability to accurately interpret visual input, confirming its suitability as a foundational component in assistive systems for individuals with visual impairments, where reliable image classification can support the generation of meaningful output information.

Figure 1 shows the confusion matrix for the model for different objects and as seen, the confusion between cat and dog is observable. This is somehow understandable, taking into account the similarity of appearance of these two animals.

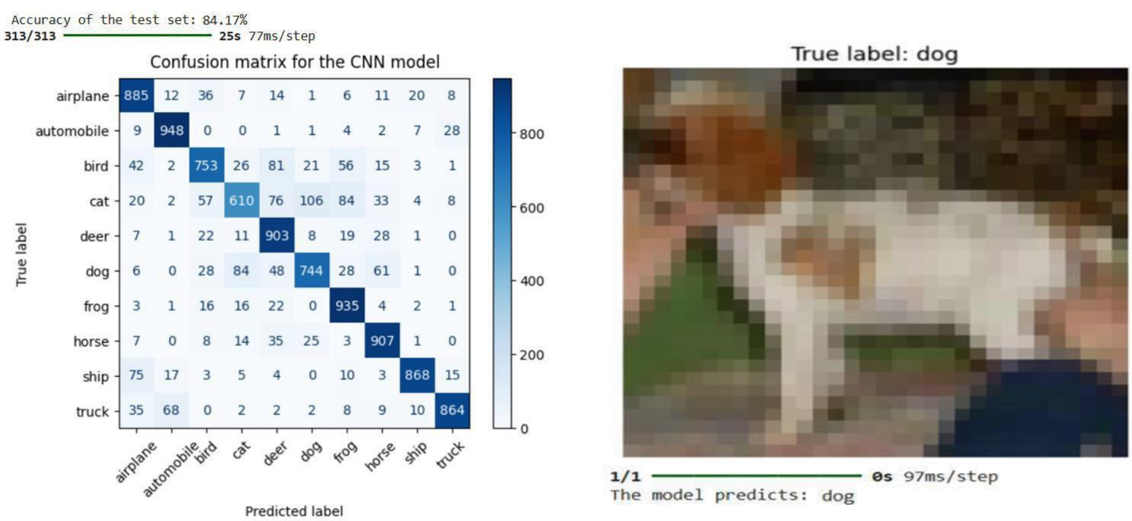

Figure 1: Result for the CNN model (after 30 epochs of the training process)

6. Conclusion

This paper presents a comprehensive investigation, analysis, and practical demonstration related to artificial intelligence aimed at supporting individuals with sensory and cognitive disabilities. By reviewing existing solutions, the main types of intelligent systems, including applications, devices, and platforms, were identified and categorized according to their functionality.

As a result of the comprehensive analysis, a generalized architecture for intelligent systems was proposed, informed by the current solutions. This architecture consists of multiple interconnected layers: an input layer, a processing layer, an adaptation and personalization layer, and an output layer, each serving a distinct role in facilitating interaction and responsiveness.

The practical component involved the implementation of a CNN model addressing visual impairments. Trained on the CIFAR-10 dataset, the model demonstrated robust image classification capabilities with an accuracy of 84.17%, providing a reliable foundation for integration into assistive technologies.

As a conclusion it is evident that the theoretical insights and practical results underscore the pivotal role of intelligent systems in fostering inclusive technological environments. The findings indicate that even relatively straightforward approaches, such as the CNN-based image classifier can deliver tangible support in everyday activities. Looking forward, the adoption of more advanced models, larger datasets, and greater personalization holds considerable promise for enhancing the quality of life for diverse users, contributing to a society where technology serves the needs of all individuals equally.

References:

[1] GeeksforGeeks, Intelligent systems in AI, 2025. URL: https://www.geeksforgeeks.org/artificial-intelligence/intelligent-systems-in-ai/

[2] Setiawan, Ariyono, Intelligent Systems for Disabilities Technology for a More Inclusive Life (November 15, 2024). Available at SSRN: https://ssrn.com/abstract=5062273 or http://dx.doi.org/10.2139/ssrn.5062273

[3] Telles, M. et al., An intelligent model to assist people with disabilities in smart cities, Journal of Ambient Intelligence and Smart Environments, Vol. 13, Issue 4, 2021

[4] Slapnicar, G., et al., Intelligent Supportive System for People with Profound Intellectual and Multiple Disabilities, Advanced intelligent Systems, Wiley, 2025

[5] Olawade, D. et al., The role of artificial intelligence in enhancing healthcare for people with disabilities, Social Science & Medicine, Volume 364, 2025

[6] Almufareh, M.F., et al., Intellectual Disability and Technology: An Artificial Intelligence Perspective and Framework. *JDR.* 2023. Vol. 2(4):58-70. DOI: 10.57197/JDR-2023-0055

[7] Harvard Business Review, Designing Generative AI to Work for People with Disabilities, 2023. URL: https://hbr.org/2023/08/designing-generative-ai-to-work-for-people-with-disabilities

[8] Noori, K., A Disability – inclusive Artificial Intelligence Act, European AI & Society Fund, 2024

[9] Key2Enable, 7 free apps to help people with disabilities, 2024. URL: https://key2enable.com/tecnologia-assistiva/7-aplicativos-gratuitos-para-ajudar-as-pessoas-com-deficiencia

[10] AIMResearch, 5 AI-Powered Assistive Technologies for People with Disabilities in the US, 2024. URL: https://aimresearch.co/market-industry/5-ai-powered-assistive-technologies-for-people-with-disabilities-in-us

[11] Expressia, 2025. URL: https://expressia.life/

[12] Aira, Aira Explorer App, 2023. URL: https://aira.io/aira-explorer-app/

[13] Microsoft Research, Soundscape, 2025. URL: https://www.microsoft.com/en-us/research/product/soundscape/

[14] OrCam Technologies Ltd., OrCam MyEye 3 Pro, 2024. URL: https://www.orcam.com/en-us/orcam-myeye-3-pro?srsltid=AfmBOopJpoAGlYBwbTDZjqVUg9k7Pd-

J292C4MS7v1sMuBCGyyby7FWH

[15] Tobii Dynavox, TD I-Series, 2025. URL: https://www.tobiidynavox.com/pages/td-i-series/

[16] Medium, Seven Layers of AI Model Architecture: A Deep Dive into the Foundation of AI Systems, 2025. URL: https://medium.com/design-bootcamp/seven-layers-of-ai-model-architecture-a-deep-dive-into-the-foundation-of-ai-systems-a8d9511c448a

[17] Dhumal, A. (2024). A Review of CNN-Based Object Recognition and Tracking Systems for Visually Impaired People. International Journal for Research in Applied Science and Engineering Technology. 12. 155-159. 10.22214/ijraset.2024.64986