Author Proof

Integrating Machine Learning with Augmented Reality for Accessible Assistive Technologies

Basel Barakat[1](✉), Lynne Hall[1], and Simeon Keates[2]

[1] School of Computer Science, University of Sunderland, Sir Tom Cowie Campus, St Peters Way, Sunderland SR6 0DD, UK
basel.barakat@sunderland.ac.uk
[2] University of Chichester, Chichester PO169 6PE, UK

Abstract. Augmented Reality (AR) is a technology which enhances physical environments by superimposing digital data on top of a real-world view. AR has multiple applications and use cases, bringing digital data into the physical world enabling experiences such as training staff on complicated machinery without the risks that come with such activities. Numerous other uses have been developed including for entertainment, with AR games and cultural experiences now emerging. Recently, AR has been used for developing assistive technologies, with applications across a range of disabilities. To achieve the high-quality interactions expected by users, there has been increasing integration of AR with Machine Learning (ML) algorithms. This integration offers additional functionality to increase the scope of AR applications. In this paper we present the potential of integrating AR with ML algorithms for developing assistive technologies, for the use case of locating objects in the home context.

AQ1

Keywords: Machine learning · Computer vision · Speech recognition · Emotion detection · Assistive technologies

1 Introduction

Technology is seen as a significant driver to enable, empower and include the 15% of the world's population that have a disability [1]. Recent technological advances, from voice recognition to brain computer interfaces offer a wealth of solutions and opportunities. Identifying and applying technologies can considerably improve the quality of life for disabled users. One of the technologies that is currently pushing the limits of our interaction with computers is Augmented Reality (AR).

AR technology uses 3-D images and animations that can be displayed in the physical environment. Most of the currently developed applications for AR are for entertainment [2] and training [3]. However, the specialised hardware and software developed for AR has massive potential for developing assistive technologies. For example, the available cameras and microphones, see Fig. 1, can provide data from the user surroundings. The software developed for imposing 3D images can benefit users by pointing them to a specific area or object.

M. Antona and C. Stephanidis (Eds.): HCII 2022, LNCS 13309, pp. 1–13, 2022.
https://doi.org/10.1007/978-3-031-05039-8_12

Fig. 1. An example of an AR kit, i.e., Microsoft Hololens2

The AR kits provide functionalities that can be extended, for example, using Machine Learning algorithms to improve understanding of contextual data, such as providing users with information about their surroundings that they might not have observed. These algorithms can be very useful as they can help the users with their unique requirements. The requirements this paper focuses on are provided by a visually impaired user in the home context aiming to locate objects. The goal is to is to show the potential of integrating computer vision and speech recognition algorithms with the AR to provide assistive technologies.

This paper investigates the exploratory research question: Can AR and ML be integrated to provide an assistive technology for the use case of locating objects in the home context? Sect. 2 briefly outlines recent assistive technologies for a range of sensory, physical and cognitive disabilities, impairments and neurodiversity. Section 3 focuses on visual impairment and the use of Augmented Reality as assistive technology. Section 4 considers empowering assistive technologies using Machine Learning algorithms. Section 5 presents a prototype system integrating computer vision and an AR headset. Section 6 outlines potential next steps and applications.

2 Assistive Technology

The World Health Organisation (WHO) defines assistive technologies as "*assistive products and related systems and services developed for people to maintain or improve functioning and thereby promote well-being*" [4]. The WHO estimated that over one billion people need some form of assistive technology with this figure expected to increase significantly in the coming decades. As the market for assistive technologies strengthens, grows and diversifies, commercialisation is increasing, with more assistive and inclusive technologies emerging.

Figure 1 highlights the wide range of impairments in the UK, many of which can be supported with assistive technologies. This growing requirement for assistive technology is emerging against the ongoing widening of participation in education, work, recreation and society, with a massive increase in the mainstreaming and inclusion of children, teenagers and adults with disabilities, impairments and neurodiversity. There is also an inevitable significant growth in the need for assistive technologies to enable independent living for an increasingly ageing population [4]. Such challenges, include, for example, visual impairment arising from age-related macular degeneration, cataracts or glaucoma; and cognitive impairment such as dementia and Alzheimer's. Supporting independent living is not only the most cost effective solution, but additionally improves quality of life allowing people to continue to live in their own homes.

Greater awareness of how assistive technologies can improve quality of life for those with disabilities, impairments and neurodiversity has also led to statutory regulation. The Convention on the Rights of Persons with Disabilities recognises access to assistive technology as a human right (Fig. 2).

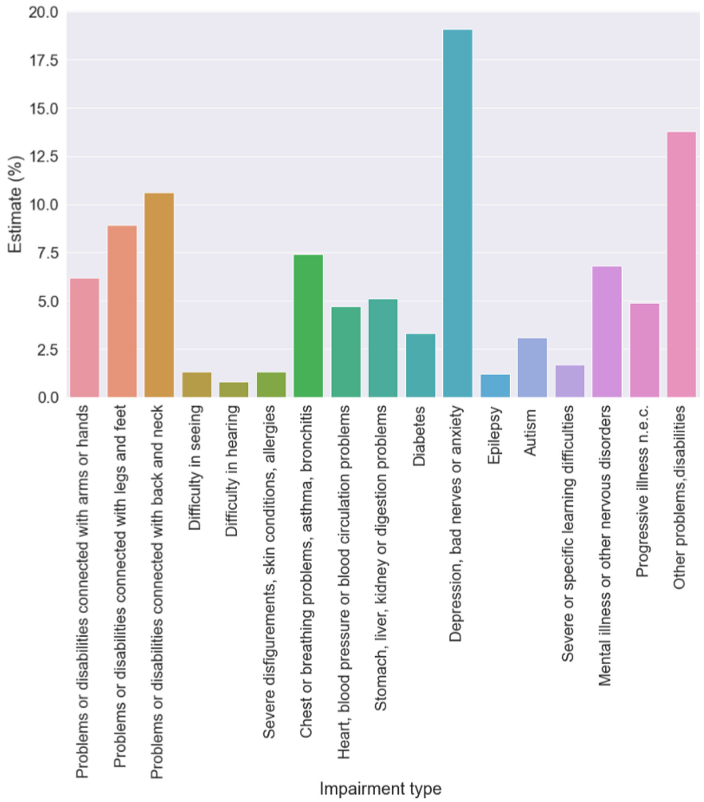

Fig. 2. Statistics on the wide range of disabilities in the UK, as shown by the Office for National Statistics - Annual Population Survey [5]

2.1 Visually Impaired

Assistive technologies for visual impairment include magnifying devices incorporated into wearables such as IrisVision, eSight or Oxsight smart glasses. These magnify, improve contrast, highlight edges and so on in real space. Additionally, some, such as eSight [6], can also be connected to digital experiences, such as games consoles and screens or Nu Eyes Pro glasses that incorporate voice and optical character recognition. There are also voice-only applications, such as Envision Glasses that enable the user to 'hear what they want to see' using AI to provide scene descriptions, facial recognition, to scan, remember and read text and make video calls [7].

In addition to wearables, there have been attempts to support the visually impaired with devices including for orientation and mobility training, such as augmented reality multi-sensory maps [8]; for navigation, including robot dogs [9]; and devices such as smart canes [10] for mobility support.

2.2 Hearing, Voice, Speech, or Language Disorders

Assistive technologies targeting hearing loss, range from hearing-related products being embedded in smartphones to implantable technologies [11]. Hearing aids will be increasingly enhanced with AI to optimise the hearing experience tailored to the user and context. Voice recognition continues to improve, with significant advances in processing language from those with disabilities impacting speech [12].

The assistive technologies currently used by health professionals can be classified as follows (Table 1): AQ2

Table 1. Assistive technologies for hearing, voice, speech, and language disorders [11]

Assistive listening devices (ALDs)	Help amplify the sounds you want to hear, especially where there's a lot of background noise. ALDs can be used with a hearing aid or cochlear implant to help a wearer hear certain sounds better
Augmentative and alternative communication (AAC) devices	Help people with communication disorders to express themselves. These devices can range from a simple picture board to a computer programme that synthesises speech from text
Alerting devices	Connect to a doorbell, telephone, or alarm that emits a loud sound or blinking light to let someone with hearing loss know that an event is taking place

2.3 Neurodiversity, Intellectual Disabilities and Cognitive Impairment

Assistive technologies have been developed to support neurodiverse and intellectually disabled populations, with mainstream as well as assistive technologies being effectively used. This includes aiming to provide neurodiverse and disabled users with the same experiences as their peers, using everyday devices, with EasyReading [13, 14], for example, providing an app that can make any webpage accessible across a wide range of disabilities. Text can be read-aloud, simplified, replaced by pictograms and so on, depending on the preference and needs of the user. Assistive technologies have been developed to support therapy for neurological conditions, such as Polipo [12] (Tam, Gelsomini & Garzotto, 2017) a multisensory interactive toy. Minecraft has been successfully used to support social skill development for children and teenagers with autism [15]. Similarly, virtual agents and social robots have been used to trigger pro-social behaviour [16], and to support social and emotional development.

3 Augmented Reality (AR)

Coughlin and Miele [17] categorise AR as an assistive technology into global applications that are used to augment the physical world to enable the user to navigate and interact more easily with it, and local applications, which augment physical objects that the user can touch, explore and use. Here, the focus is on the latter, creating a local application to assist a visually impaired user with locating objects in the home context.

AR has been used effectively to support users with visual impairments in a range of contexts and use cases, particularly in work, training [18] and educational contexts [19]. Through incorporating computer vision, the AR experience can enable visually impaired users to navigate and use spaces more effectively. For example, computer vision has been used to support navigation in indoor environments [20] with object recognition used to recognise doors and elevators enabling the visually impaired user to navigate an unfamiliar building [21] and to assist visually impaired users with sign reading using existing signage [22]. The Bright platform, an augmented reality assistive platform, provides contextual features enabled by vision processing, such as recognising contacts through facial recognition [23]. AR applications for visual impairment for users in the home include safety such as fall prevention applications [24].

4 Assistive Technologies Empowered by Machine Learning Algorithms

Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed [25]. Unlike the traditional approach where we ought to explicitly define the problem and write the rules to solve it, in machine learning we provide the computers with a dataset and let them propose a method for solving it. This approach has several advantages as we do not always know the optimal method for setting the rules.

Machine learning algorithms form rules by observing the datasets. They can extract rules that we might not know or understand. This is particularly beneficial for developing assistive technologies, as the range of problems and possible solutions are substantial. In this paper, we discuss two fields of ML that has have great potential to support visually impaired users: Speech recognition and Computer Vision.

4.1 Speech Recognition

Recently, Speech Recognition systems have become part of everyday life especially with the wide spread of smartphone assistants (e.g., Siri, Cortana, Google Now) and voice assistants such as Amazon Echo, and Kinect Xbox One [26, 27]. These systems uses accurate machine learning algorithms for ‘understanding’ the human voice and speech. Speech recognition has been an active research area for a long time [28, 29], with major contributions having been made using Hidden Markov Chains [28, 29]; machine learning algorithms [30] [31]; and Deep learning [31].

In essence the speech recognition algorithms take the speech signal as an input and extract information from it [30], as outlined in Table 2.

Table 2. Extracted information from speech signals

Extracted information	Details & related reference
Speech content	Recognise the content of words and phrases [30]
Speaker identity	Recognise the speaker by utilising the information embedded in his/her speech signal [32]
Emotions	Recognise the speaker emotion such as happy, scared, sad, etc.[33]
Language	Recognise the spoken langue can be challenging for languages that share phrases [34]
Accent	Recognition of a speaker’s regional accent, within a predetermined language [32]
Age and gender	Recognising the age and gender of the user [35]

Accruing this remarkable amount of information using speech has a great potential for developing assistive language. For instance, it can be used to help in developing assistive technologies for users who suffer from hearing loss and/langue difficulties. It can also be very helpful for the users that might have problems interacting with the peripherals.

4.2 Computer Vision

Computer vision is another very active research area in machine learning. Computer vision algorithms extract information from images and videos. The information we can extract can be very helpful in developing assistive technologies for the visually impaired. Figure 3 presents some of the capabilities of computer vision in developing assistive technologies.

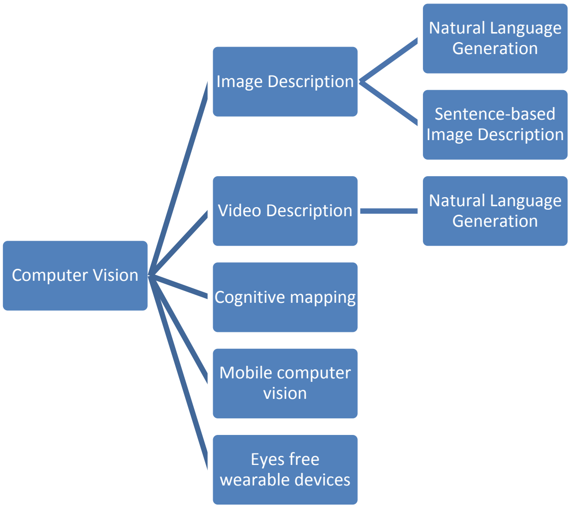

Fig. 3. Assistive technologies using Computer Vision

Recently major contributions in Computer vision have empowered the algorithms to detect multiple objects in an image in real time, as shown in Fig. 4. In particular, the algorithms that use deep learning have shown a huge potential for accurate and timely detection of objects. For instant You-Only-Look-Once (YOLO) algorithm has shown great performance and the potential scalability [32].

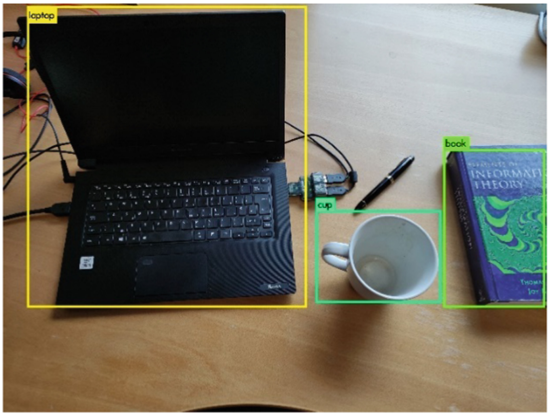

Fig. 4. Objects detection using Machine learning algorithm

An example of using object detection in assistive technologies had been developed in [33], where the presented framework (shown in Fig. 5) was developed to provide assistive technology to support a user in locating their items. The framework uses speech recognition, computer vision and optimises queue performance to locate items in real time. In this approach the user does need help from other people hence they can use it without worrying about invading their privacy. It also interacts with the user only using speech thus a visually impaired user can use it.

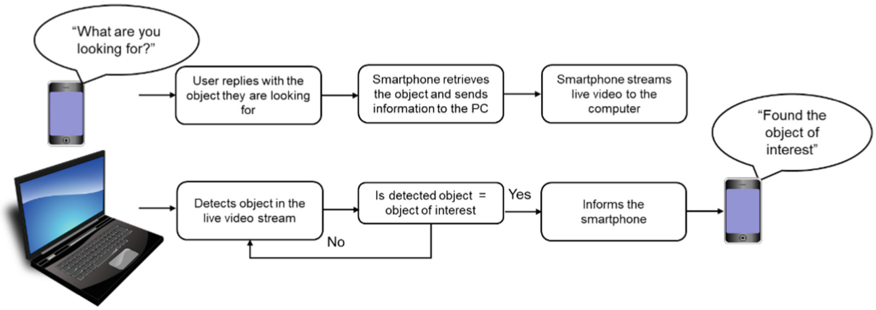

Fig. 5. Computer Vision assistive technology example [34]

5 Proposed Integration of AR with ML Algorithms

5.1 Overview

In this section, we present a potential assistive technology based on the framework outlined in Fig. 5, that integrates machine learning with augmented reality kits. The main idea is to utilise the AR capability of interacting with the physical word and integrate it with machine learning algorithms capabilities to make accurate and timely decisions. Generally, the proposed solution is for in-home assistance as shown in Fig. 6. However, it can be extended for outdoor applications.

Fig. 6. Potential solution for integrating computer vision and AR headsets

The user would interact with the AR through the microphone, then we can extract the content from the speech signal using a machine learning algorithm. For example, the user might ask to locate the keys. Then wearing the AR kit, they would walk around the home, in the meantime the AR would stream the live video to their personal computer where an object detection algorithm would analyze the frames and identify the object. If the item is in the sight of the AR, it would notify the user using an audio signal or even through imposing arrows to the location of the item, as shown in Fig. 7.

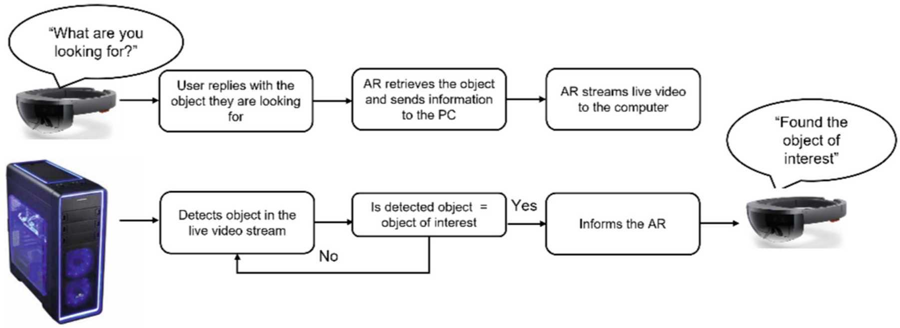

Fig. 7. Proposed framework for integrating AR with Machine learning

5.2 Practical Implantation Consideration and Challenges

Practical Implantation Consideration and Challenges. While implementing an integration between AR and ML, it is critical to the practical implications of these systems. One of the most challenging issues in complying ML algorithms is the required computational computer. As it has a massive effect on the processing duration and hence the usability of the assistive technologies. Hence, it is crucial to balance the ML algorithm accuracy with the processing duration.

ML algorithms can be complied on the AR kit or another device such as remote server. Having an ML algorithm compiled on the AR kit might sound like the more convenient method, however, it would generate several challenges such as compellability, RAM size and processor power. For example, if we consider the Microsoft HoloLens 2 it has 4GB RAM and a Qualcomm Snapdragon 850 processor. Which would make it very hard to develop onboard algorithms. On the other hand, if the algorithm was developed on an external server, then the computational power would not be as challenging. In this scenario it is critical to optimise the communication (between the AR kit and the server) performance.

In our implantation we took the second approach, as we had an external server running the ML algorithm. One of the problems we faced was the massive delay in transmitting the live video stream. To stream a video can be done by activating the developer mode on the kit and streaming the video using the 'Mixed Reality Capture' function on the device portal. This approach is quite easy to implant, and on the other hand, it has a significantly high streaming delay. To overcome this, we have used the Microsoft HoloLens app [35]. Further improvements can be made using the Microsoft Mixed reality Real-Time Communication for Web (Web-RTC) protocol [36].

6 Discussion

Integrating AR technologies with Machine Learning algorithms offers considerable potential for creating assistive technologies. The proof-of-concept in this paper demonstrates one such integration, that of enhancing AR with speech recognition and machine vision. This integration provides an improved assistive experience for a visually impaired user supporting them in locating objects.

Current work is focusing on testing the application in a lab with the goal of then moving to the home context of a visually impaired user to assess the application usability. This approach is also likely to have utility for users with cognitive disabilities or impairments related to age, enabling them to rapidly locate objects.

7 Conclusions

As technology advances the potential of using it to help people grows. Through integrating Augmented Reality with Machine Learning it becomes possible to create assistive technologies that can help users with a wide range of requirements. We presented a proof-of-concept example aimed at users with visual impairment integrating AR with machine learning algorithms for machine vision and speech recognition creating an interaction that helped users to locate their items.

References

1. World Report on Disability Summary. https://www.who.int/publications/i/item/WHO-NMH-VIP-11.01. Accessed 24 Feb 2022
2. Hung, S.-W., Chang, C.-W., Ma, Y.-C.: A new reality: exploring continuance intention to use mobile augmented reality for entertainment purposes. Technol. Soc. **67**, 101757 (2021). https://doi.org/10.1016/j.techsoc.2021.101757
3. Li, X., Yi, W., Chi, H.-L., Wang, X., Chan, A.P.C.: A critical review of virtual and augmented reality (VR/AR) applications in construction safety. Autom. Constr. **86**, 150–162 (2018). https://doi.org/10.1016/j.autcon.2017.11.003
4. "Assistive technology." https://www.who.int/news-room/fact-sheets/detail/assistive-technology. Accessed 24 Feb 2022
5. "Labour Force Survey - Office for National Statistics." https://www.ons.gov.uk/surveys/informationforhouseholdsandindividuals/householdandindividualsurveys/labourforcesurvey. Accessed 22 Feb 2022
6. "New Home gb - eSight." https://esighteyewear.com/gb. Accessed 24 Feb 2022
7. "Envision Glasses." https://www.letsenvision.com/envision-glasses. Accessed 24 Feb 2022
8. Laviole, J., Thevin, L., Albouys-Perrois, J., Brock, A.: Nectar. In: Proceedings of the Virtual Reality International Conference - Laval Virtual, April 2018, pp. 1–6 (2018). https://doi.org/10.1145/3234253.3234317

9. Chuang, T.-K., et al.: Deep trail-following robotic guide dog in pedestrian environments for people who are blind and visually impaired - learning from virtual and real worlds. In: 2018 IEEE International Conference on Robotics and Automation (ICRA), May 2018, pp. 5849–5855. https://doi.org/10.1109/ICRA.2018.8460994
10. Albouys-Perrois, J., Laviole, J., Briant, C., Brock, A.M.: Towards a multisensory augmented reality map for blind and low vision people. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, April 2018, vol. 2018, pp. 1–14 (2018). https://doi.org/10.1145/3173574.3174203
11. Assistive Devices for People with Hearing or Speech Disorders | NIDCD. https://www.nidcd.nih.gov/health/assistive-devices-people-hearing-voice-speech-or-language-disorders. Accessed 22 Feb 2022
12. Tam, V., Gelsomini, M., Garzotto, F.: Polipo. In: Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction, March 2017, pp. 11–20 (2017). https://doi.org/10.1145/3024969.3025006
13. Stone, B.G., Mills, K.A., Saggers, B.: Online multiplayer games for the social interactions of children with autism spectrum disorder: a resource for inclusive education. Int. J. Incl. Educ. **23**(2), 209–228 (2019). https://doi.org/10.1080/13603116.2018.1426051
14. EASY READING – Keeping the user at the digital original. https://www.easyreading.eu/. Accessed 24 Feb 2022
15. Pradhan, A., Mehta, K., Findlater, L.: Accessibility came by accident. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, April 2018, vol. 2018-April, pp. 1–13 (2018). https://doi.org/10.1145/3173574.3174033
16. Oliveira, R., Arriaga, P., Santos, F.P., Mascarenhas, S., Paiva, A.: Towards pro-social design: a scoping review of the use of robots and virtual agents to trigger prosocial behaviour. Comput. Hum. Behav. **114**, 106547 (2021). https://doi.org/10.1016/j.chb.2020.106547
17. Coughlan, J.M., Miele, J.: AR4VI: AR as an accessibility tool for people with visual impairments. In: 2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct), October 2017, pp. 288–292 (2017). https://doi.org/10.1109/ISMAR-Adjunct.2017.89
18. Cavus, N., Al-Dosakee, K., Abdi, A., Sadiq, S.: The utilization of augmented reality technology for sustainable skill development for people with special needs: a systematic literature review. Sustainability **13**(19), 10532 (2021). https://doi.org/10.3390/su131910532
19. Quintero, J., Baldiris, S., Rubira, R., Cerón, J., Velez, G.: Augmented reality in educational inclusion. A systematic review on the last decade. Front. Psychol. **10**, 1835 (2019). https://doi.org/10.3389/FPSYG.2019.01835/BIBTEX
20. Yoon, C., et al.: Leveraging augmented reality to create apps for people with visual disabilities. In: The 21st International ACM SIGACCESS Conference on Computers and Accessibility, October 2019, pp. 210–221 (2019). https://doi.org/10.1145/3308561.3353788
21. Tian, Y., Yang, X., Yi, C., Arditi, A.: Toward a computer vision-based wayfinding aid for blind persons to access unfamiliar indoor environments. Mach. Vis. Appl. **24**(3), 521–535 (2013). https://doi.org/10.1007/s00138-012-0431-7
22. Huang, J., Kinateder, M., Dunn, M.J., Jarosz, W., Yang, X.-D., Cooper, E.A.: An augmented reality sign-reading assistant for users with reduced vision. PLoS ONE **14**(1), e0210630 (2019). https://doi.org/10.1371/journal.pone.0210630
23. Bakshi, A.M., Simson, J., de Castro, C., Yu, C.C., Dias, A.: Bright: an augmented reality assistive platform for visual impairment. In: 2019 IEEE Games, Entertainment, Media Conference (GEM), June 2019, pp. 1–4 (2019). https://doi.org/10.1109/GEM.2019.8811556
24. lo Bianco, M., Pedell, S., Renda, G.: Augmented reality and home modifications. In: Proceedings of the 28th Australian Conference on Computer-Human Interaction - OzCHI 2016, November 2016, pp. 499–507 (2016). https://doi.org/10.1145/3010915.3010929

25. Géron, A.: Hands-on machine learning with Scikit-Learn, Keras and TensorFlow: concepts, tools, and techniques to build intelligent systems (2019). https://www.oreilly.com/library/view/hands-on-machine-learning/9781492032632/. Accessed 22 Feb 2022
26. Dabran, I., Avny, T., Singher, E., ben Danan, H.: Augmented reality speech recognition for the hearing impaired. In: 2017 IEEE International Conference on Microwaves, Antennas, Communications and Electronic Systems (COMCAS), November 2017, vol. 2017-November, pp. 1–4 (2017). https://doi.org/10.1109/COMCAS.2017.8244731
27. Zhang, Z., Geiger, J., Pohjalainen, J., Mousa, A.E.-D., Jin, W., Schuller, B.: Deep learning for environmentally robust speech recognition. ACM Trans. Intell. Syst. Technol. **9**(5), 1–28 (2018). https://doi.org/10.1145/3178115
28. Gales, M., Young, S.: The application of hidden markov models in speech recognition. Found. Trends® Sig. Process. **1**(3), 195–304 (2007). https://doi.org/10.1561/2000000004
29. Rabiner, L., Juang, B.: An introduction to hidden Markov models. IEEE ASSP Mag. **3**(1), 4–16 (1986). https://doi.org/10.1109/MASSP.1986.1165342
30. Abdel-Hamid, O., Mohamed, A.R., Jiang, H., Deng, L., Penn, G., Yu, D.: Convolutional neural networks for speech recognition. IEEE Trans. Audio Speech Lang. Process. **22**(10), 1533–1545 (2014). https://doi.org/10.1109/TASLP.2014.2339736
31. Nassif, A.B., Shahin, I., Attili, I., Azzeh, M., Shaalan, K.: Speech recognition using deep neural networks: a systematic review. IEEE Access **7**, 19143–19165 (2019). https://doi.org/10.1109/ACCESS.2019.2896880
32. Redmon, J., Farhadi, A.: YOLOv3: an incremental improvement, April 2018. http://arxiv.org/abs/1804.02767. Accessed 22 Feb 2022
33. Barakat, B., Steponenaite, A., Lall, G.S., Arshad, K., Wassell, I.J., Keates, S.: Assistive technology for the visually impaired: optimizing frame rate (freshness) to improve the performance of real-time objects detection application. In: Antona, M., Stephanidis, C. (eds.) HCII 2020. LNCS, vol. 12189, pp. 479–492. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-49108-6_34
34. Barakat, B., et al.: 6G opportunities arising from internet of things use cases: a review paper. Future Internet **13**(6), 159 (2021). https://doi.org/10.3390/fi13060159
35. Get Microsoft HoloLens - Microsoft Store en-GB. https://www.microsoft.com/en-gb/p/microsoft-hololens/9nblggh4qwnx?activetab=pivot:overviewtab. Accessed 25 Feb 2022
36. Microsoft. MixedReality-WebRTC Documentation. https://microsoft.github.io/MixedReality-WebRTC/. Accessed 25 Feb 2022