Climb-o-Vision: A Computer Vision Driven Sensory Substitution Device for Rock Climbing

Mike Richardson

University of Bath, Department of Psychology, m.richardson@bath.ac.uk

Karin Petrini

University of Bath, Department of Psychology, CAMERA and REVEAL Research Centers, k.petrini@bath.ac.uk

Michael J. Proulx

University of Bath, Department of Psychology, CAMERA and REVEAL Research Centers, m.j.proulx@bath.ac.uk

The benefits of taking part in adventurous activities are many; particularly, for people with visual impairments. Sports such as rock climbing can improve feelings of skillfulness, autonomy, and confidence for people with low or no vision as they strive to overcome environmental and personal challenges. In this late-breaking work we present Climb-o-Vision, a novel sensory substitution software that utilizes YOLOv5 computer vision object-detection architecture, to aid navigation for rock climbers with visual impairments. Climb-o-Vision uses commercially available and cost-effective hardware to detect, track, and convert climbing hold spatial locations on to the surface of the tongue, via an electrotactile tongue interface. Preliminary testing of the device highlights the possibility of using sensory substitution as a sporting aid for people with visual impairments. Furthermore, it demonstrates the potential for adapting and improving current sensory substitution systems by employing computer vision techniques to filter useful task-specific information to users with visual impairments.

CCS CONCEPTS • Human-centered computing • Accessibility • Accessibility technologies

Additional Keywords and Phrases: Sensory substitution, Computer vision, Accessible sport, Rock climbing

ACM Reference Format:

First Author's Name, Initials, and Last Name, Second Author's Name, Initials, and Last Name, and Third Author's Name, Initials, and Last Name. 2018. The Title of the Paper: ACM Conference Proceedings Manuscript Submission Template: This is the subtitle of the paper, this document both explains and embodies the submission format for authors using Word. In Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY. ACM, New York, NY, USA, 10 pages. NOTE: This block will be automatically generated when manuscripts are processed after acceptance.

1 Introduction

Access to exercise for people with visual impairments is reduced compared to others [7, 27, 40]. As such, many people with visual impairments often suffer with comorbid health issues relating to insufficient exercise [48]. For many years, targeted applications of technology have been deployed to help increase participation in sport and exercise for people with disabilities [51]. For instance, Blind Football, Goalball (a ball game designed specifically for people with visual impairments), and Visually Impaired Tennis all utilize a ball that creates noise when in motion. For target sports, such as archery, often a tactile or auditory 'sight' can be used to align the participant's shot [4]. More recently, the field of human-computer interaction (HCI) has helped to develop novel prototype devices that further expand exercise possibilities to less 'typical' sports, such as hiking or kayaking [3, 23]. Exergames, video games that involve an exercise component to play or progress [35], are becoming increasingly popular, particularly since the development of interactive console hardware such as the Nintendo Wii and Xbox Kinect [14, 55]. HCI research has generated interesting prototypes from exergames and exergame concepts specifically created for people with visual impairments. Notably, a team developed VI-Tennis and VI-Bowling [28, 29] exergames built on Wii Sports, but with additional tactile and audio feedback to improve players' experiences.

However, exergames may lack many of the positive features of sports, such as social development and benefits associated with being in a dedicated exercise environment [26, 32, 47]. This is particularly the case for adventurous sports, where an individual tends to compete against the environment itself, or themselves, rather than another person or team. For example, in rock climbing, a major component of the sport involves using balance and stability to maintain traction on small holds off the ground, in a competition against gravity. Virtual climbing simulators struggle to successfully emulate this critical component as the feet tend to remain on the floor in virtual settings [17].

HCI research also offers technology to help exercisers with visual impairments increase their participation in more naturalistic exercise environments; for example, by attaching sound emitting devices to slalom gates to help kayakers navigate [19], or to climbing hold locations to help rock climbers orientate themselves [16]. Adventure sports such as these can increase feelings of confidence, resilience, and generate supportive relationships [18, 24, 46]; however, manipulating the environment with sound emitters may have some drawbacks. Each emitter must first be placed into the environment and subsequently moved or changed for each new variation of the activity or route. This immediately reduces accessibility as there must be at least one additional individual available to provide support, and drastically slows the rate at which a person could perform different routes, which may limit feelings of autonomy and skillfulness [8, 44], and impede positive associated mental states, such as flow and creativity [33, 54]. Furthermore, inside climbing gyms, routes are frequently 're-set' to create more variety for climbers but limits the feasibility of fixed navigation methods (like learning a tactile map).

Here, we present a new prototype software approach, called Climb-o-Vision (see Figure 1), that uses computer vision and sensory substitution to convey climbing hold locations in a climbing gym via a tongue interface (the Cthulhu Shield), without the need to adapt the environment or simulate the activity virtually, while also leaving the ears and auditory modality unoccluded (an important safety feature for climbing).

2 Related work

The concept of Climb-o-Vision was inspired by sensory substitution research, specifically, previous research with a different tongue interface, the BrainPort (Wicab, WI, USA). The BrainPort is an FDA and CE approved vision-into-tactile sensory substitution device (SSD), marketed as a vision aid for people with severe visual impairment [12]. It converts visual information from a camera mounted on the forehead to electrotactile stimulation on the tongue. The tongue display consists of a 20×20 grid of electrodes (minus 6 corner electrodes), with each electrode representing a pixel and encoding luminance as amplitude of signal, in essence, creating a 394-pixel greyscale display on the surface of the tongue.

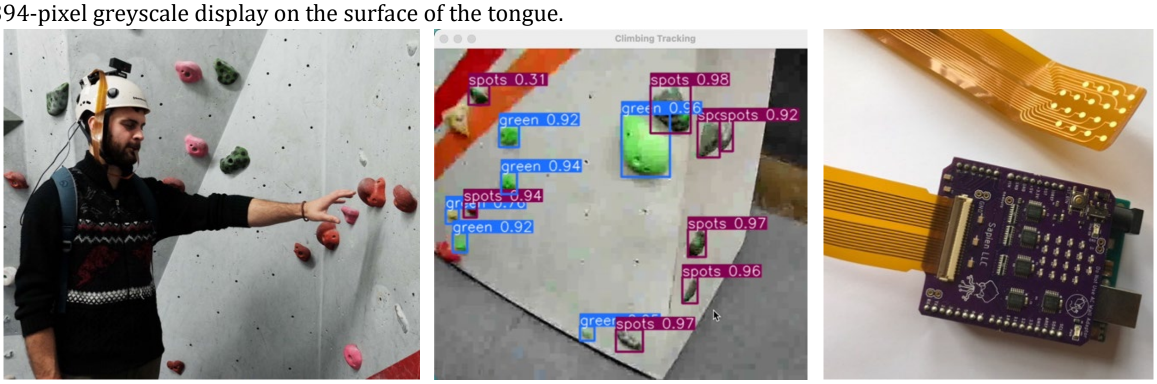

Figure 1: Climb-o-Vision in use (left). Climbing hold identification visual render (center). The Cthulhu Shield tongue interface output device (right).

Past research with the BrainPort has offered some compelling results for tasks, such as shape identification and navigation [12, 34, 49], however, this is often caveated by a user's performance being dependent on having enough time to actively explore the display with their tongue [36]. This may be potentially problematic when using the BrainPort for sporting purposes as environmental interaction needs to be rapidly adaptive in response to change. The BrainPort has previously been used for the purpose of rock climbing by adventure sport athlete Erik Weihenmayer [52], but this is yet to be adopted by others, or reflected in academic literature.

The tongue presents both unique possibilities and challenges as a receptor for sensory substitution. For instance, the tongue is incredibly sensitive to touch sensations, but is not uniform in its sensitivity [13, 45]. The anterior section possesses a better capacity for electrotactile sensation compared to the posterior section [2, 30]. Because of this, information presented to the posterior of the tongue's surface is generally less well perceived [37]. Additionally, visual information occurs in the vertical plane, perpendicular to gaze, while the surface of the tongue is aligned on the horizontal plane, meaning perception must be mentally rotated by 90° either forwards or backwards [42]. The BrainPort rotates information so that the top of the camera's field of view (FoV) is presented to the posterior section of the tongue and the bottom of the FoV to the anterior section. The perceptual sensitivity gradient on the tongue means that information at the bottom of the camera's FoV will be more accurately identified [37], and attentionally prioritized, compared to information at the top of the FoV. For climbing, this is particularly an issue, as navigation typically occurs while ascending.

One of the common issues of SSDs is that they can overload the user with too much information [5, 6, 11]. Vision is the highest bandwidth sense [43], which inevitably means perceptual loss when substituting information. Many SSDs, such as the BrainPort, currently deal with this issue by downscaling the information prior to substitution, for example, by reducing pixel density, removing color, or narrowing the camera's FoV. However, in the case of the tongue, this method may not be enough to counteract the sensitivity gradient. We argue that computer vision techniques can be applied to current sensory substitution systems to filter out task-relevant information and help prevent overstimulation, to create task-specific devices, and in this case, increase the available opportunities for accessible exercise.

3 Climb-o-vision prototype

In this section we put forward the prototype SSD, called Climb-o-Vision, that uses computer vision to identify climbing hold locations and then it converts the spatial information into tactile stimulations on the tongue. This prototype consists of commercially available and affordable devices (namely, the Cthulhu Shield for Arduino, an Arduino Uno, and a USB webcam), that connect to a laptop running the Climb-o-Vision program, stored in a backpack while climbing.

3.1 Hardware

The Cthulhu Shield (Sapien LLC, CO, USA) is an Arduino compatible tongue interface that utilizes an 18-electrode grid that, like the BrainPort, can write information to the tongue (but unlike the BrainPort can also read information from the tongue, a feature not used here). For the Climb-o-Vision prototype, we used the device as a display to stimulate the tongue with spatial information depending on climbing hold location. The Cthulhu Shield is driven by an Arduino Uno (Rev 3, Arduino, Italy), which is then connected to a computer via a serial port. The Cthulhu Shield was chosen over the BrainPort for a number of reasons, including, its accessible and customizable open-source programming, and cost effectiveness (it is approximately 1% of the price of the BrainPort).
The Climb-o-Vision software can run on any laptop with enough processing power, the version demonstrated here used a MacBook Pro (2020, Apple M1, 16GB RAM) running on macOS Big Sur (Version 11.4). Likewise, the program could use any external USB camera, however, the webcam used here was the NexiGo N60 (1080p, 30 fps, USB 2.0). The climbing hold detection model was trained on an Alienware Area-51 R2 PC, with an Intel® Core i7 (5820k @ 3.30GHz), 16 GB of RAM, running on Windows 10 (Microsoft, WA, USA), and utilizing a Nvidia 1080 Ti graphics card, allowing for CUDA (version 10.2.89) optimization of the model training.

3.2 Software

The software was developed in Python 3 and predominately utilizes PyTorch (version 1.9.0) and YOLOv5 (Ultralytics) object detection architecture [9]. YOLOv5, the fifth version of the seminal detection architecture 'You Only Look Once' [38], is pretrained using the Common Objects in Context (COCO) dataset and can then be adapted with custom categories [21]. The program uses the OpenCV library (version 4.5.2) to perform image processing and apply the model frame by frame. A copy of the program file can be found here: https://tinyurl.com/yck93jyd

We programmed the software to only provide lateral cues for hold location to bypass the issue created by the tongue's sensitivity gradient. To do this, we restricted each video frame's aspect ratio from 16:9 to 8:1 (800:100 pixels), creating a narrowed vertical field of view (FoV), meaning that the user must estimate verticality through active sensing by moving their head up and down. This method of substitution reduces the problematic posterior section of the tongue, by providing the same lateral information to both the anterior and posterior sections and allowing the user to solely attribute lateral cues in line with their head. A restricted FoV has been employed by other SSDs to encourage users to actively explore the environment and to minimize overstimulation [15, 25].

3.3 Model Training and Validation

To collect the training image dataset, video files were captured via the webcam mounted to a climbing helmet. Video clips were captured during the process of climbing on bouldering routes inside a climbing gym to maintain use-case validity. Each video file was broken down into individual frames and one frame in every sixty was recorded into the training image file (i.e., one frame for every two seconds of video), creating 565 images in total. For each of the training images, an accompanying annotation file was created using LabelImg [53], that marked the coordinates of each climbing hold. YOLOv5 then divides the dataset into two, creating batches of training images, and batches of validation images (see Table 1 for model performance, and Figure 2 for a sample validation batch).

In climbing gyms, different routes are typically designated by their colored holds. An issue with using the BrainPort for climbing, is that it only presents visual information in greyscale, so while it may be possible to correctly identify the hold location with practice, it would be far more difficult to correctly identify one route from another (see Figure 3C for an example of this). For the present version of Climb-o-Vision we focused on five categories of route color to maintain simplicity, however, this can easily be scaled up in future versions of the program. We chose to use three color categories: red, yellow, and green, respectively, and two multicolored categories: orange/green swirls, and black with green spots.

The smallest and fastest pretrained model was selected (YOLOv5s) to increase the speed of the training phase, and the model was trained using the settings: image size = 320 pixels; batch size = 16; workers = 1; epochs = 500. Table 1 shows the performance metrics at various epochs throughout the training process (100th, 200th, 300th, 400th, and 500th epoch, respectively); precision and recall metrics were high even by the 100th epoch, while mean average precision (mAP) progressively improved throughout, and by the 500th epoch, offered a good indication of accurate model performance. Precision, recall, and mAP are common markers of object-detection performance [38, 39].

Table 1: Model performance metrics

Epoch	Precision	Recall	mAP@.50:.95[a]
100	0.931	0.964	0.753
200	0.987	0.983	0.813
300	0.996	0.993	0.856
400	0.997	0.996	0.896
500	0.997	0.998	0.913

[a] Mean Average Precision at an Intersection over Union (IoU) thresholds of .50 to .95; a common performance indicator for the Common Objects in Context dataset. All metrics have a range of 0 to 1.

3.4 Simulated Testing Phase

To initially test the performance of the model after training and validation we ran the program on pre-recorded video which was not used in the training process, but was recorded at a similar time on the same climbing routes. This phase demonstrated that Climb-o-Vision could successfully apply the model to frames that were captured in the same conditions with the same holds, but that were not used in the model's training or validation, and therefore, were novel to the model (see Figure 3D). Inspection of the rendered video file demonstrated that the model could successfully be applied to new frames, but that a degree of overfitting existed, where some hold colors that had been excluded from the model were mis-identified as colors that were included in the model (e.g., predicting that pink holds were red). The model's inference speed results were: 0.2 ms pre-process; 98.5 ms inference; 0.2 ms non-maximum suppression (NMS) per image at shape (1, 3, 320, 320).

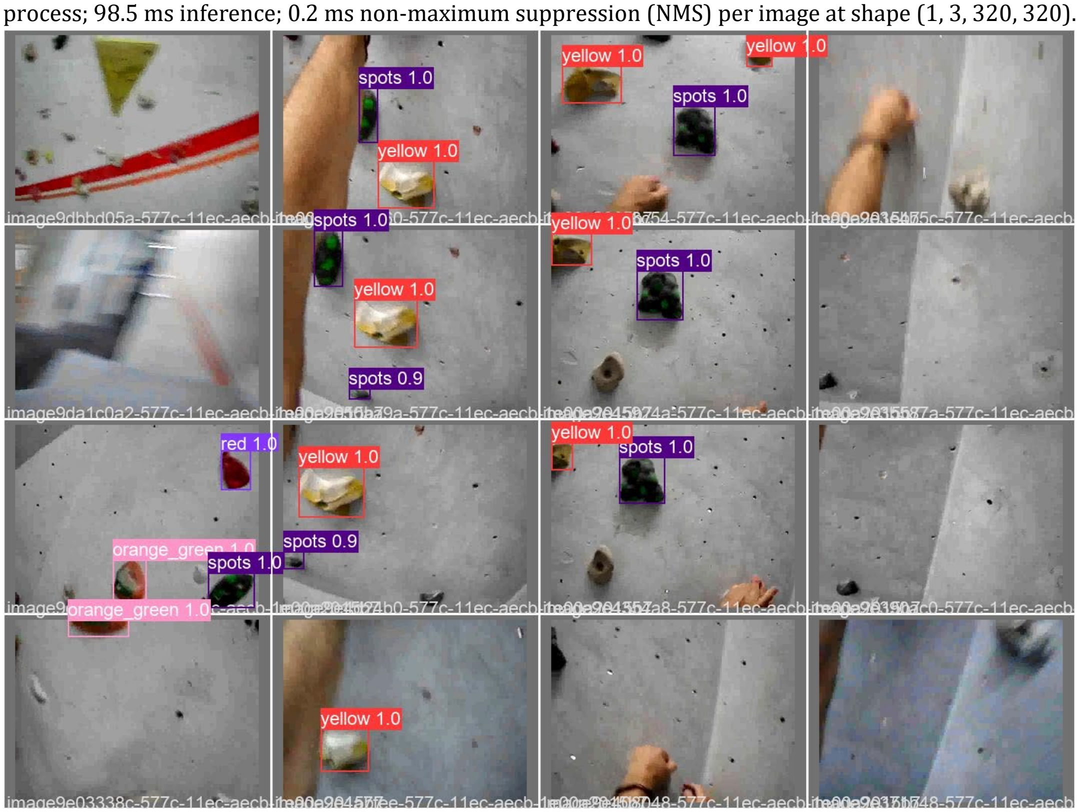

Figure 2: A sample batch of validation images and predictions created by YOLOv5 during the hold detection training process. Each batch contained 16 images, each with a resolution of 320 × 240 pixels.

3.5 Applicative Testing Phase

To test the model's actual performance, we took Climb-o-Vision back to the climbing gym after routes had been re-set. The re-set enabled us to test the program on new holds in different positions and in real-time. We found that Climb-o-Vision could still successfully detect climbing holds despite frames containing different holds in different locations (see Figure 4), and therefore, were more novel to the model than the pre-recorded video.

However, we observed that the overfitting seemed to occur more frequently than the pre-recorded video. Many excluded colors were mis-identified as included colors, and for some frames, the model mixed up trained colors (e.g., identifying red holds as yellow), perhaps due to slightly different lighting conditions [1]. This would be problematic for the user, as they could be directed to a hold color that is not part of their intended climbing route. Interestingly, the model identified holds with different colored spots quite robustly, perhaps suggesting that multi-toned holds are easier for the detection process.

4 Discussion

The Climb-o-Vision prototype demonstrates that computer vision techniques can be successfully applied to sensory substitution for the purpose of sport and exercise, creating task-specific applications that filter out useful information. The present prototype focused on the sport of rock climbing, as tongue displays have been utilized for this activity before in popular media [20, 52]. More importantly, rock climbers with blindness and visual impairments also often mention the positive impact of participating in an adventure sport [50], particularly on boosting confidence and overcoming fear associated with sight loss [50]. However, despite rock climbing taking the present focus, there is no reason why a similar method could not be applied to other sports. The field of sensory substitution can offer unique non-visual displays that translate important visual features from a given sport (e.g., other players in football/soccer, a ball in goalball, obstacles while trail running), which can be highlighted by custom trained object detection architecture. Computer vision techniques have been integrated into a vision-into-audio SSD previously for the purpose of alerting the user to approaching hazards [10], and likewise spatialized audio principles have been applied to computer vision devices for orientating the user to another person [31]. Climb-o-Vision represents the next logical steps of combining these two fields to develop a wider scope of visually assistive technology.

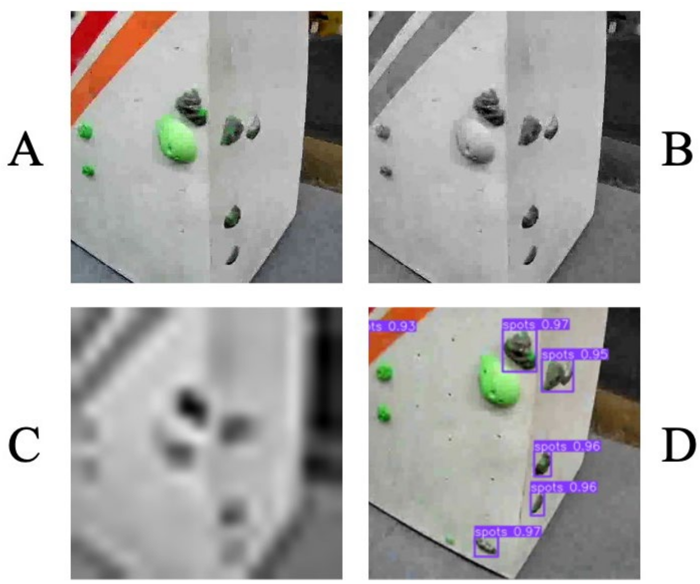

Figure 3: Demonstrating potential interpretations of climbing route navigation using restricted information. A = an unprocessed video frame of two separate climbing routes; B = a greyscale version of the frame; C = the greyscale frame rendered to 400 pixels, as the BrainPort would display to its user; D = the frame with object detection applied, highlighting only one climbing route, as would be displayed to the Climb-o-Vision user. The lighter shaded and darker shaded holds separate the routes.

4.1 Limitations

To reduce file size and increase processing time, the model was trained using low resolution images, however due to the size and shape of climbing holds, we expect that increasing the image size would increase the detection performance, and prevent the model overfitting (i.e., reduce the likelihood of misidentifying non-climbing holds as climbing holds). Furthermore, during applicative testing we found that the model regularly mis-identified holds with the wrong color. To combat this, we aim to train future versions of the model with a larger and more varied dataset, including images from different climbing walls in diverse lighting conditions. Increasing the number of categories to encapsulate more colors of climbing hold should also help to prevent the model from misidentifying the incorrect color (e.g., by differentiating pink holds from red, or grey holds from black). Our method of splitting the dataset into training and validation images could be further improved as the images were in close temporal proximity to one another thus reducing environmental variability. For future iterations of the device, we will collect images in three different sets: training images, validation images, and test images. These datasets would be best collected in isolation of each other and preferably contain distinct and isolated climbing routes. Furthermore, ensuring to include every color of climbing hold in the training dataset will reduce the likelihood of the model mistaking hold colors and improve overall performance.

4.2 Future Directions

The further next steps for Climb-o-Vision are twofold. Firstly, we plan to objectively evaluate climbing performance with the device, compared to the BrainPort, and self-motion, to give an indication of how it performs compared to traditional navigation methods for rock climbing (self-motion), but also compared to a non-computer vision tongue interface (BrainPort). To do this we will recruit rock climbers with blindness and visual impairments to test how Climb-o-Vision affects their current navigation strategies while rock climbing, and we will also recruit non-climbers to examine how it influences the learning process of climbing. Secondly, we aim to improve the hardware constraints of Climb-o-Vision. Currently, the software operates on a laptop stored in a rucksack while climbing (a method previously adopted by earlier iterations of the BrainPort when used while climbing), however, this adds additional weight to the climber. We hope to utilize modern powerful mobile processing, as other vision aids have done [15, 31], to reduce weight and increase the usability of the device. Specifically, the pretrained PyTorch YOLOv5 Climb-o-Vision model can be exported to a smaller model, such as TensorFlow Lite, to enable processing on a microcontroller or single board computer. Moving to a smaller model should also improve the inference speed, although the image processing latency speeds found in the simulated testing phase were adequate for perception via the tongue, which has a limited attentional capacity [41].

Figure 4: The model overfitting trained hold categories onto untrained hold colors.

We also hope that the field of sensory substitution begins to seriously consider the use of computer vision for improving generalized SSDs, and vice versa for computer vision assistive devices. SSDs can provide the ideal technology for intuitive non-visual displays [22], while visually assistive applications, such as SeeingAI (Microsoft, WA, USA), can provide a selective filter for incoming information; both fields can improve from mutually beneficial collaboration.

5 Conclusion

This late-breaking work demonstrates the concept that computer vision techniques can be successfully applied to the field of sensory substitution to improve sport and exercise opportunities. Computer vision can help to filter out background or irrelevant information that could potentially overload SSD users, particularly for specialized activities. While Climb-o-Vision has more refinements required before it is useable as an assistive technology, it marks the first steps towards unified and targeted approach of vision assistive technologies for increasing access to exercise.

ACKNOWLEDGMENTS

We would like to thank Alex Shaw and The Climbing Academy for their continuing help in the development process of Climb-o-Vision, and for their commitment to improving climbing accessibility for people with visual impairments. We would also like to acknowledge the team at Sapien LLC, as their development of an affordable open access tongue interface, provides the ideal non-visual display for developing assistive technology. MJP's and KP's research is partly funded by the UKRI Centre for the Analysis of Motion, Entertainment Research and Applications (CAMERA 2.0; EP/T022523/1). MR's research is funded by a SWDTP ESRC +3 PhD studentship.

REFERENCES

<bib id="bib1"><number>[1]</number>Acunzo, D., Zhu, Y., Xie, B. and Baratoff, G. 2007. Context-adaptive approach for vehicle detection under varying lighting conditions. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC* (2007), 654–660.</bib>

<bib id="bib2"><number>[2]</number>Allison, T.S., Moritz, J., Turk, P. and Stone-Roy, L.M. 2020. Lingual electrotactile discrimination ability is associated with the presence of specific connective tissue structures (papillae) on the tongue surface. *PLOS ONE*. 15, 8 (Aug. 2020), e0237142. DOI:https://doi.org/10.1371/journal.pone.0237142.</bib>

<bib id="bib3"><number>[3]</number>Anthierens, C., Groux, D. and Hugel, V. 2018. Sensory navigation guide for visually impaired sea kayakers. *Journal of Field Robotics*. 35, 5 (Aug. 2018), 732–747. DOI:https://doi.org/10.1002/rob.21775.</bib>

<bib id="bib4"><number>[4]</number>Archery - British Blind Sport: *https://britishblindsport.org.uk/membership/bbs-sports/archery/*. Accessed: 2021-12-07.</bib>
<bib id="bib5"><number>[5]</number>Brown, D.J. and Proulx, M.J. 2016. Audio-Vision Substitution for Blind Individuals: Addressing Human Information Processing Capacity Limitations. *IEEE Journal on Selected Topics in Signal Processing*. 10, 5 (Aug. 2016), 924–931. DOI:https://doi.org/10.1109/JSTSP.2016.2543678.</bib>
<bib id="bib6"><number>[6]</number>Brown, D.J., Simpson, A.J.R. and Proulx, M.J. 2014. Visual objects in the auditory system in sensory substitution: How much information dowe need? *Multisensory Research*. 27, 5–6 (Nov. 2014), 337–357. DOI:https://doi.org/10.1163/22134808-00002462.</bib>
<bib id="bib7"><number>[7]</number>Capella-McDonnall, M. 2007. The need for health promotion for adults who are visually impaired. *Journal of Visual Impairment and Blindness*. AFB Press.</bib>
<bib id="bib8"><number>[8]</number>Deci, E.L. and Ryan, R.M. 2000. The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*. 11, 4 (2000), 227–268. DOI:https://doi.org/10.1207/S15327965PLI1104_01.</bib>
<bib id="bib9"><number>[9]</number>GitHub - ultralytics/yolov5: YOLOv5 in PyTorch >ONNX >CoreML >TFLite: *https://github.com/ultralytics/yolov5*. Accessed: 2021-12-07.</bib>
<bib id="bib10"><number>[10]</number>Gomez, J.D., Bologna, G. and Pun, T. 2014. See ColOr: An extended sensory substitution device for the visually impaired. *Journal of Assistive Technologies*. 8, 2 (2014), 77–94. DOI:https://doi.org/10.1108/JAT-08-2013-0025.</bib>
<bib id="bib11"><number>[11]</number>Gori, M., Cappagli, G., Tonelli, A., Baud-Bovy, G. and Finocchietti, S. 2016. Devices for visually impaired people: High technological devices with low user acceptance and no adaptability for children. *Neuroscience and Biobehavioral Reviews*. 69, (Oct. 2016), 79–88. DOI:https://doi.org/10.1016/j.neubiorev.2016.06.043.</bib>
<bib id="bib12"><number>[12]</number>Grant, P., Maeng, M., Arango, T., Hogle, R., Szlyk, J. and Seiple, W. 2018. Performance of Real-world Functional Tasks Using an Updated Oral Electronic Vision Device in Persons Blinded by Trauma. *Optometry and Vision Science*. 95, 9 (Sep. 2018), 766–773. DOI:https://doi.org/10.1097/OPX.0000000000001273.</bib>
<bib id="bib13"><number>[13]</number>Haggard, P. and de Boer, L. 2014. Oral somatosensory awareness. *Neuroscience & Biobehavioral Reviews*. 47, (Nov. 2014), 469–484. DOI:https://doi.org/10.1016/J.NEUBIOREV.2014.09.015.</bib>
<bib id="bib14"><number>[14]</number>Hall, A.K., Chavarria, E., Maneeratana, V., Chaney, B.H. and Bernhardt, J.M. 2012. Health Benefits of Digital Videogames for Older Adults: A Systematic Review of the Literature. *Games for Health Journal*. 1, 6 (Dec. 2012), 402–410. DOI:https://doi.org/10.1089/g4h.2012.0046.</bib>
<bib id="bib15"><number>[15]</number>Hamilton-Fletcher, G., Alvarez, J., Obrist, M. and Ward, J. 2021. SoundSight: a mobile sensory substitution device that sonifies colour, distance, and temperature. *Journal on Multimodal User Interfaces*. (Jul. 2021), 1–17. DOI:https://doi.org/10.1007/s12193-021-00376-w.</bib>
<bib id="bib16"><number>[16]</number>Ilich, M. 2008. MetaHolds: A Rock Climbing Interface for the Visually Impaired.</bib>
<bib id="bib17"><number>[17]</number>Jenny, S.E. and Schary, D.P. 2015. Virtual and “real-life” wall/rock climbing: motor movement comparisons and video gaming pedagogical perceptions. *Sports Technology*. 8, 3–4 (Oct. 2015), 100–111. DOI:https://doi.org/10.1080/19346182.2015.1118110.</bib>
<bib id="bib18"><number>[18]</number>Jessup, G.M., Cornell, E. and Bundy, A.C. 2010. The treasure in leisure activities: Fostering resilience in young people who are blind. *Journal of Visual Impairment and Blindness*. 104, 7 (Jan. 2010), 419–430. DOI:https://doi.org/10.1177/0145482x1010400705.</bib>
<bib id="bib19"><number>[19]</number>Kirtland, D., Phife, R., Gardner, W. and Johnson, A. 2013. Guidance System, Kayaking for the Visually Impaired. *Biomedical Engineering*. (Jun. 2013).</bib>
<bib id="bib20"><number>[20]</number>Levy, B. 2008. The Blind Climber Who “Sees” With His Tongue. *Discover Magazine*.</bib>
<bib id="bib21"><number>[21]</number>Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P. and Zitnick, C.L. 2014. Microsoft COCO: Common objects in context. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (2014), 740–755.</bib>
<bib id="bib22"><number>[22]</number>Lloyd-Esenkaya, T., Lloyd-Esenkaya, V., O’Neill, E. and Proulx, M.J. 2020. Multisensory inclusive design with sensory substitution. *Cognitive Research: Principles and Implications*. Springer.</bib>
<bib id="bib23"><number>[23]</number>Long, S.K., Karpinsky, N.D., Döner, H. and Still, J.D. 2016. Using a mobile application to help visually impaired individuals explore the outdoors. *Advances in Intelligent Systems and Computing* (2016), 213–223.</bib>
<bib id="bib24"><number>[24]</number>Macpherson, H. 2009. The Intercorporeal Emergence of Landscape: Negotiating Sight, Blindness, and Ideas of Landscape in the British Countryside. *Environment and Planning A: Economy and Space*. 41, 5 (May 2009), 1042–1054. DOI:https://doi.org/10.1068/a40365.</bib>
<bib id="bib25"><number>[25]</number>Maidenbaum, S., Hanassy, S., Abboud, S., Buchs, G., Chebat, D.R., Levy-Tzedek, S. and Amedi, A. 2014. The “EyeCane”, a new electronic travel aid for the blind: Technology, behavior & swift learning. *Restorative Neurology and Neuroscience*. 32, 6 (Jan. 2014), 813–824. DOI:https://doi.org/10.3233/RNN-130351.</bib>
<bib id="bib26"><number>[26]</number>Martin, J.J. 2006. Psychosocial aspects of youth disability sport. *Adapted Physical Activity Quarterly*. 23, 1 (Jan. 2006), 65–77. DOI:https://doi.org/10.1123/apaq.23.1.65.</bib>
<bib id="bib27"><number>[27]</number>Matoso, G.P. and Portela, B.S. 2020. Level of physical activity and perceived barriers to its practice in adults with visual impairment. *Revista Brasileira de Atividade Física & Saúde*. 24, (Mar. 2020), 1–7. DOI:https://doi.org/10.12820/rbafs.24e0094.</bib>
<bib id="bib28"><number>[28]</number>Morelli, T., Foley, J., Columna, L., Lieberman, L. and Folmer, E. 2010. VI-Tennis: A vibrotactile/audio exergame for players who are visually impaired. *FDG 2010 - Proceedings of the 5th International Conference on the Foundations of Digital Games* (2010), 147–154.</bib>
<bib id="bib29"><number>[29]</number>Morelli, T., Foley, J. and Folmer, E. 2010. VI-bowling: A tactile spatial exergame for individuals with visual impairments. *ASSETS'10 - Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility* (2010), 179–186.</bib>
<bib id="bib30"><number>[30]</number>Moritz Jr., J., Turk, P., Williams, J.D. and Stone-Roy, L.M. 2017. Perceived Intensity and Discrimination Ability for Lingual Electrotactile Stimulation Depends on Location and Orientation of Electrodes. *Frontiers in Human Neuroscience*. 11, (Apr. 2017), 186. DOI:https://doi.org/10.3389/fnhum.2017.00186.</bib>

<bib id="bib31"><number>[31]</number>Morrison, C., Cutrell, E., Grayson, M., Roumen, G., Marques, R.F., Thieme, A., Taylor, A. and Sellen, A. 2021. PeopleLens. *Interactions*. 28, 3 (May 2021), 10–13. DOI:https://doi.org/10.1145/3460116.</bib>

<bib id="bib32"><number>[32]</number>Mutz, M. and Müller, J. 2016. Mental health benefits of outdoor adventures: Results from two pilot studies. *Journal of Adolescence*. 49, (Jun. 2016), 105–114. DOI:https://doi.org/10.1016/j.adolescence.2016.03.009.</bib>

<bib id="bib33"><number>[33]</number>Nakamura, J. and Csikszentmihalyi, M. 2014. The concept of flow. *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi*. Springer Netherlands. 239–263.</bib>

<bib id="bib34"><number>[34]</number>Nau, A.C., Pintar, C., Arnoldussen, A. and Fisher, C. 2015. Acquisition of visual perception in blind adults using the BrainPort artificial vision device. *American Journal of Occupational Therapy*. 69, 1 (Jan. 2015). DOI:https://doi.org/10.5014/ajot.2015.011809.</bib>

<bib id="bib35"><number>[35]</number>Oh, Y. and Yang, S. 2010. *Defining exergames & exergaming*.</bib>

<bib id="bib36"><number>[36]</number>Pamir, Z., Canoluk, M. U., Jung, J.-H. and Peli, E. 2019. The variable spatial resolution is compensated by tip of the tongue tracing in pattern recognition. *ARVO Annual Meeting* (2019).</bib>

<bib id="bib37"><number>[37]</number>Pamir, Z., Canoluk, M.U., Jung, J.H. and Peli, E. 2020. Poor resolution at the back of the tongue is the bottleneck for spatial pattern recognition. *Scientific Reports*. 10, 1 (Dec. 2020), 1–13. DOI:https://doi.org/10.1038/s41598-020-59102-3.</bib>

<bib id="bib38"><number>[38]</number>Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. 2016. *You Only Look Once: Unified, Real-Time Object Detection*.</bib>

<bib id="bib39"><number>[39]</number>Redmon, J. and Farhadi, A. 2018. YOLOv3: An Incremental Improvement. (Apr. 2018).</bib>

<bib id="bib40"><number>[40]</number>Richardson, M., Petrini, K. and Proulx, M.J. 2022. Access to exercise for people with visual impairments during the Coronavirus-19 pandemic. *British Journal of Visual Impairment*. (Jan. 2022), 026461962110673. DOI:https://doi.org/10.1177/02646196211067356.</bib>

<bib id="bib41"><number>[41]</number>Richardson, M., Petrini, K. and Proulx, M.J. In press. Orientation of Tactile Attention on the Surface of the Tongue. Journal of Experimental Psychology: Human Perception and Performance.</bib>

<bib id="bib42"><number>[42]</number>Richardson, M.L., Lloyd-Esenkaya, T., Petrini, K. and Proulx, M.J. 2020. Reading with the Tongue: Individual Differences Affect the Perception of Ambiguous Stimuli with the BrainPort. CHI 2020 (New York, NY, USA, Apr. 2020), 1–10.</bib>

<bib id="bib43"><number>[43]</number>Richardson, M.L., Thar, J., Borchers, J., Ward, J. and Hamilton-Fletcher, G. 2019. How much spatial information is lost in the sensory substitution process? Comparing visual, tactile, and auditory approaches. Perception. (2019). DOI:https://doi.org/10.1177/0301006619873194.</bib>

<bib id="bib44"><number>[44]</number>Ryan, R.M. and Deci, E.L. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist. 55, 1 (2000), 68–78. DOI:https://doi.org/10.1037/0003-066X.55.1.68.</bib>

<bib id="bib45"><number>[45]</number>Sakamoto, K., Nakata, H., Yumoto, M. and Kakigi, R. 2010. Somatosensory Processing of the Tongue in Humans. Frontiers in Physiology. 1, 1 (Nov. 2010), 136. DOI:https://doi.org/10.3389/fphys.2010.00136.</bib>

<bib id="bib46"><number>[46]</number>Schroeder, C. and Weihenmayer, E. 2002. Climbing Blind. About Campus: Enriching the Student Learning Experience. 7, 4 (Sep. 2002), 2–9. DOI:https://doi.org/10.1177/108648220200700402.</bib>

<bib id="bib47"><number>[47]</number>Seymour, H., Reid, G. and Bloom, G.A. 2009. Friendship in inclusive physical education. Adapted Physical Activity Quarterly. 26, 3 (Jul. 2009), 201–219. DOI:https://doi.org/10.1123/apaq.26.3.201.</bib>

<bib id="bib48"><number>[48]</number>Shields, N., Synnot, A.J. and Barr, M. 2012. Perceived barriers and facilitators to physical activity for children with disability: a systematic review. British journal of sports medicine.</bib>

<bib id="bib49"><number>[49]</number>Stronks, H.C., Mitchell, E.B., Nau, A.C. and Barnes, N. 2016. Visual task performance in the blind with the BrainPort V100 Vision Aid. Expert Review of Medical Devices. 13, 10 (Oct. 2016), 919–931. DOI:https://doi.org/10.1080/17434440.2016.1237287.</bib>

<bib id="bib50"><number>[50]</number>The Art of Climbing - RNIB - See differently: https://www.rnib.org.uk/the-art-of-climbing. Accessed: 2021-12-07.</bib>

<bib id="bib51"><number>[51]</number>Thomas, N. and Smith, A. 2008. Disability, sport and society: An introduction. Routledge Taylor & Francis Group.</bib>

<bib id="bib52"><number>[52]</number>Twilley, N. 2017. Seeing with Your Tongue. The New Yorker.</bib>

<bib id="bib53"><number>[53]</number>Tzutalin 2015. labelImg. Git code.</bib>

<bib id="bib54"><number>[54]</number>Wang, C.-C. 2012. Exploring the relationship among rock climbing engagement, sensation-seeking, and creativity: A multi-method approach to understanding the adventure seeking creativity model.</bib>

<bib id="bib55"><number>[55]</number>Willems, M.E. and Bond, T.S. 2009. Comparison of Physiological and Metabolic Responses to Playing Nintendo Wii Sports and Brisk Treadmill Walking. Journal of Human Kinetics. 22, (2009), 43–50. DOI:https://doi.org/10.2478/v10078-009-0022-5.