JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

BLIND ASSIST SYSTEM USING MACHINE LEARNING

Mr. FAIZ AMAN 1, Rakshitha C A2, Sindhu C S3, Upadhya Kshama B4, Vidya S U5

1. Guide, 2,3,4&5. Student Computer Science Department, Bahubali College of Engineering.

Abstract : Blindness is one of the most frequent and debilitating of the various disabilities. There are million visually impaired people in the globe, according to the World Health Organization (WHO). The proposed system is designed to aid visually impaired persons with real-time obstacle detection, avoidance, indoors and out navigation, and actual position tracking. The gadget proposed is a camera-visual detection hybrid that performs well in low light as part of the recommended technique, this method is utilized to detect and avoid impediments, as well as to aid visually impaired persons in identifying the environment around them. A simple and effective method for people with visual impairments to identify things in their environment and convert them into speech for improved comprehension and navigation. Along with these, the depth estimation, which calculates the safe distance between the object and the person, allowing them to be more self-sufficient and less reliant on others. This were able to achieve this model with the help of TensorFlow and pre-trained models. The approach suggested is dependable, inexpensive, practical, and practicable.
IndexTerms **-Yolo algorithm, For obstacle detection, For the purpose of navigation, Real time object detection .**

Eyesight is one of the essential human senses, and it plays a significant role in human perception about the surrounding environment. For visually impaired people to be able to provide, experience their vision, imagination mobility is necessary. The International Classification of Diseases 11 (2018) classifies vision impairment into two groups, distance and near presenting vision impairment. Globally, the leading causes of vision impairment are uncorrected refractive errors, cataract, age related macular degeneration, glaucoma, diabetic retinopathy, corneal opacity, trachoma, and eye injuries. It limits impaired ability to navigate, perform everyday tasks, and affect their quality of life and ability to interact with the surrounding world upon unaided. With the advancement in technologies, diverse solutions have been introduced such, as the Eye- ring project, the text recognition system, the hand gesture, and face recognition system, etc. However, these solutions have disadvantages such as heavyweight, expensive, less robustness, low acceptance, etc. visually impaired ability to navigate, perform everyday tasks, and affect their quality of life and ability to interact with the surrounding world upon unaided. With the advancement in technologies, diverse solutions have been introduced such, as the Eye- ring project, the text recognition system, the hand gesture, and face recognition system, etc. However, these solutions have disadvantages such as heavyweight, expensive, less robustness, low acceptance, etc. Hence, advanced techniques must evolve to help them. The early proposed system captures real-time images, then images are pre-processed, their background and foreground are separated and then the DNN module with a pre-trained YOLO model is applied resulting in feature extraction. The extracted features are matched with known object features to identify the objects. Once the object is successfully recognized, the object name is stated as voice output with the help of text-to-speech conversion. The key contributions of this project include: Robust and efficient object detection and recognition for visually impaired people to independently access familiar and unfamiliar environments and conversion and speech output.

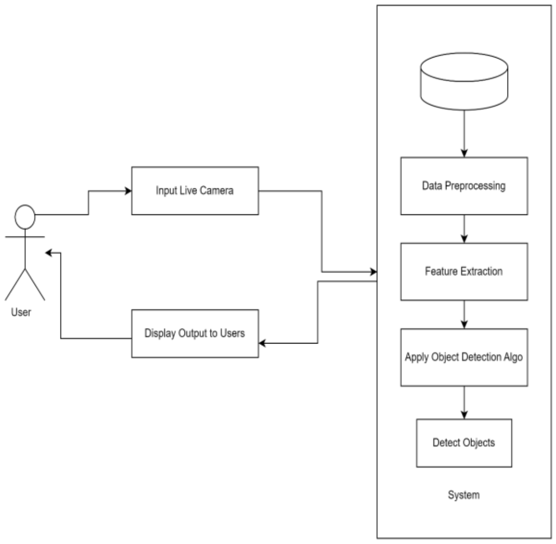


Figure: System architecture of Blind assist system usingmachine learning.

1. *Input Live Camera:* The system takes an input image in a specific format (typically RGB) on which object detection needs to be performed.

2. *Preprocessing:* The input image undergoes preprocessingsteps to prepare it for the detection process. This may include resizing the image to a fixed size and normalizing pixel values.

3. *Feature Extraction:* The backbone network extracts high- level features from the input image, capturing both low-leveland high-level visual information.

4. *Yolov3:* YOLOv3 has multiple detection layers at different scales to detect objects at various sizes. Each detection layer is responsible for detecting objects of a specific scale. These detection layers are usually implemented as convolutional layers followed by some additional layers for prediction.

5. *Output:* The system provides the final set of detected objects along with their class labels, bounding box coordinates, and confidence scores.

III. ISSUES OF EXISTING SYSTEM.

1. **Limited Environmental Adaptability**: Many systems struggle to adapt to diverse environments and situations, such as crowded streets or changing weather conditions.
2. **Object Recognition Accuracy**: While machine learning models have improved, they may still struggle with accurately identifying objects in real-time, especially in varied lighting conditions or with occluded objects.
3. **Processing Speed**: Some systems may face delays in processing and providing feedback due to computational limitations, which can be critical in fast-paced environments.
4. **Dependency on Data Quality**: The performance of machine learning models heavily relies on the quality and diversity of training data. Biases in the data or insufficient representation of certain scenarios can lead to inaccuracies.
5. **User Interface and Interaction**: The design of the user interface and interaction methods is crucial for user acceptance and effectiveness. Complex interfaces or confusing feedback canhinder usability for blind users.
6. **Privacy Concerns**: Some systems may raise privacy concerns if they collect and process sensitive information about the user ortheir surroundings.

OUR APPROACH :

1. **Multi-Sensor Integration**: We integrate various sensors such as cameras, LiDAR, and inertial sensors to gather comprehensive environmental data.

2. **Machine Learning Models**: Utilizing deep learning algorithms for robust object detection, scene recognition, and obstacle avoidance.

3. **Real-Time Processing**: Employing optimized algorithms and hardware to ensure fast and responsive feedback to the user.

4. **Adaptive Learning**: Implementing algorithms that continuously learn and adapt to user behavior and feedback, improving system performance over time.

5. **Data Augmentation**: Enhancing the diversity and quality of training data through techniques like augmentation and synthetic data generation.

6. **Privacy Preservation**: Implementing privacy-preserving techniques such as on-device processing and encryption to protect user data.

7. **Customizable Interfaces**: Providing customizable user interfaces and interaction methods to cater to individual preferences and needs.

8. **Feedback Mechanisms**: Incorporating clear and intuitive feedback mechanisms to convey environmental information tothe user effectively.

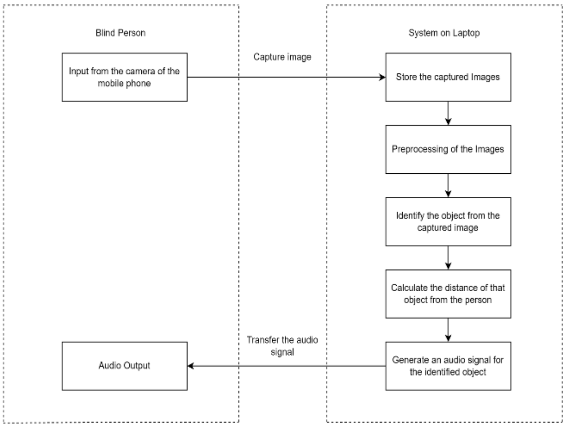
9. **Accessibility Features**: Ensuring the system is accessible tousers with varying degrees of visual impairment, including those with additional disabilities.

10. **Continuous Improvement**: Committing to ongoing research and development to address emerging challenges and improve system performance and usability over time.

ADVANTAGES OFPROPOSEDSYSTEM

- Our solution to the problem is to create a device which can recognize obstacle using camera using voice alert. This is small and compact also easy to carry.
- This blind assist can recognize the object in real time when the objects are visible the camera.
- This system will continuously record video of the surrounding and will convert it into frames. After analysing these frames, the system will alert the person about some obstacle or the surrounding.
- The main advantages are the portable, affordable and accessible system using image processing technologies is able to help visually impaired people. This system will help the visually impaired people to navigate their way through any obstacle and will give them a sense of visualization of world around them.

METHODOLOGY



Step1: Start by collecting different data.

To kickstart creating a system that helps blind people using machine learning, we first need to gather different types of information. This includes finding out about all sorts of things people might encounter, like chairs, tables, or even cars. We can get this information from different places, like the internet or by taking pictures ourselves. After that, we need to label these pictures with what's in them, like puttinga box around a chair and saying, "This is a chair." We can usetools to help us do this faster. Once we have lots of labeled pictures, we can make our computer program learn from them so it can recognize objects by itself. This helps blind people because the program can tell them what's around them, like ifthere's a chair or a person nearby. It's like teaching a smart robot to understand the world, and it can make life easier forpeople who can't see well.

Step2: Store the captured images.

In a blind assist system that uses machine learning forspotting objects, storing the pictures taken is like saving snapshots of what the camera sees. These pictures are super important because they're used to teach the computer how to recognize stuff like people, cars, or trees. Before saving them, each picture gets marked with labels saying what's in it, like "person" or "car". This helps the computer understand and learn. Once the pictures are stored, the computer looks at themover and over, learning from each one to get better at spottingobjects. Then, when the system is working, it can tell the userwhat's around them in real-time by processing the images fromthe camera and recognizing the objects. So, storing theseimages is like building a library of knowledge for the computerto learn from, helping it assist blind people by identifying things in their surroundings.

Step3: Preprocessing of the image.

Before the computer can understand what's in the pictures taken by the camera in a blind assist system, it needsto prepare them, kind of like getting them ready for school. This preparation is called preprocessing. During preprocessing, the computer does a few things to make the pictures easier to understand. First, it might resize them to allbe the same size, like making sure all the cookies are the samesize before baking. Then, it might adjust the colors to make them clearer, like changing the brightness or contrast on a TV. Next, it might crop the pictures to focus only on the importantparts, like cutting out a picture of a dog from a bigger photo of a park. Finally, it might flip or rotate the pictures to look at them from different angles, kind of like turning a book around to see it better. Once the pictures are preprocessed, thecomputer can start learning from them to recognize objects and help blind people understand what's around them. So, preprocessing is like getting the pictures ready for the computer to learn from, making it easier for it to spot objectsand help out.

Step4: Identify the object from the captured image to recognize the things like obstacles.

In a blind assist system using machine learning, identifying objects from captured images helps recognize things like obstacles to keep blind people safe. Imagine the system as a smart friend who looks at pictures taken by a camera and tells you what's in them. First, the camera takes pictures of the surroundings. Then, the computer looks at these pictures and figures out what's in them, like if there's a person, a chair, or a tree. It does this by learning from lots of examples, kind of like how you learn to recognize different animals by looking at pictures in a book. Once the computer knows what's in the pictures, it can warnthe user about any obstacles or things in their way, helping them navigate safely. So, the system works like a helpful buddy, using machine learning to spot objects in pictures andkeep blind people informed about their surroundings.

Step5: Calculate the distance of that object from the person.

In a blind assist system using machine learning, calculating the distance of an object from a person helpsdetermine how far away things are, kind of like knowing ifsomething is close or far in a game. Once the system identifiesobjects in pictures taken by the camera, it uses some clevermath and tricks to estimate how far those objects are from theperson. It's like guessing the distance between you and a friend standing across the room. The system might use things like the size of the object in the picture or how blurry it looksto make this guess. Then, based on these calculations, it cantell the person if something is nearby or farther away, helpingthem understand their surroundings better. So, calculating thedistance is like playing detective, using math and clues fromthe pictures to figure out how close or far away things are from the person.

Step6: Generate an audio signal for the identified object.

In a blind assist system using machine learning, generating an audio signal for the identified object means creating sound cues to tell the user what the system sees in the environment. Imagine it like having a friendly voicedescribe things around you. Once the system recognizes objects in the pictures from the camera, it assigns each object a unique sound, like a beep or a chime. For example, it mightuse a high-pitched sound for a person and a lower tone for a chair. Then, as the user moves around, the system plays thesesounds to let them know what's nearby. So, if the camera spots a person, it'll make the sound assigned to people. This way, even if the user can't see, they can still understand what's around them through sound, making it easier to navigate and stay safe. So, generating audio signalsis like giving the user a helpful soundtrack to their surroundings, helping them interpret the world with their earsinstead of their eyes.

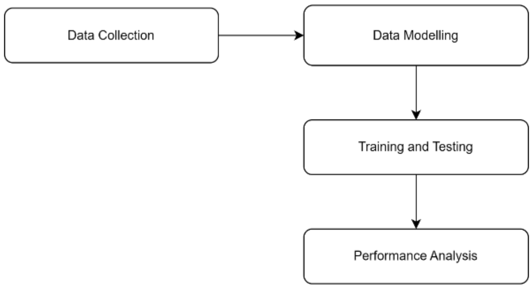
Step6: Generate an audio signal for the identified object.

In a blind assist system using machine learning, generating an audio signal for the identified object means creating sound cues to tell the user what the system sees in the environment. Imagine it like having a friendly voicedescribe things around you. Once the system recognizesobjects in the pictures from the camera, it assigns each objecta unique sound, like a beep or a chime. For example, it mightuse a high-pitched sound for a person and a lower tone for a chair. Then, as the user moves around, the system plays these sounds to let them know what's nearby. So, if the camera spots a person, it'll make the soundassigned to people. This way, even if the user can't see, they can still understand what's around them through sound, making it easier to navigate and stay safe. So, generating audio signals is like giving the user a helpful soundtrack to their surroundings, helping them interpret the world with their ears instead of their eyes.

Step7: Audio output.

In a blind assist system using machine learning, theaudio output is like having a friendly voice tell the user aboutthe objects detected in their surroundings. Think of it as a helpful narrator describing what's happening around you. Once the system identifies objects through the camera, it converts this information into spoken words or sounds. For example, it might say "person ahead" or make a specific sound for each object detected, like a beep for a chair or a whistle for a table. This way, the user can understand what's around them without needing to see it themselves. It's like having a guide by your side, helping you navigate safely through the world. So, the audio output provides clear and understandable information about the environment, making iteasier for blind individuals to move around independently.

Proposed System:



The proposed system was put into practice in which it serves as the foundation for YOLOv3 to find and categorize the objects. Data collecting, data modelling, training and testing the model, and performance analysis are the steps involved in the suggested system.

You Only Looked Once is also known as YOLO. Thistechnique is used to instantly find and identify various thingsin an image. In YOLO, object detection is accomplished as a regression issue that provides the likelihood of classes for theobserved photos.

How YOLO Algorithm Works: -

1. Yolo starts by requesting a picture.
2. Next, n X n grids are created from these input photos.
3. Each grid is used for image localization and categorization.

Then, YOLO finds the various bounding boxes andassociated class probabilities for each object in the image.

4. WORKTHROUGH TUTOR

1. **Understanding the Problem**: Begin by understanding the challenges faced by blind individuals in navigating their surroundings and the potential role of machine learning in addressing these challenges.
2. **Research and Background Study**: Conduct thorough researchon existing blind assist systems, machine learning algorithms, and relevant technologies such as computer vision and sensorfusion.
3. **Identify Goals and Requirements**: Define the goals and requirements of the blind assist system, including desired functionalities, target user demographics, and usability considerations.
4. **Selecting Tools and Technologies**: Choose appropriate tools and technologies for development, including programming languages (e.g., Python), machine learning frameworks (e.g., TensorFlow, PyTorch), and hardware components (e.g., cameras, sensors).
5. **Data Collection and Annotation**: Collect and annotatedatasets containing images, videos, and sensor data to train machine learning models for tasks such as object detection, scene recognition, and obstacle avoidance.
6. **Model Training**: Work with the tutor to train machine learning models using the collected datasets and appropriate algorithms (e.g., convolutional neural networks for imageprocessing).**Algorithm Implementation**: Implement algorithms for real-time processing of sensor data, object detection, and environmental analysis on the chosen hardware platform (e.g.,Raspberry Pi, NVIDIA Jetson).
7. **Integration and Testing**: Integrate the trained models and algorithms into a cohesive system and conduct thorough testing to

ensure accuracy, reliability, and responsiveness.

8. **User Interface Design**: Design intuitive and accessible userinterfaces for interacting with the blind assist system, considering the needs and preferences of visually impaired users.

9. **User Testing and Feedback**: Collaborate with visually impaired individuals to conduct user testing and gather feedback on the system's usability, effectiveness, and potentialimprovements.

10. **Documentation and Deployment**: Document the development process, including algorithms, data sources, and implementation details, and prepare the system fordeployment in real-world settings.

11. **Continued Learning and Improvement**: Encourage ongoing learning and improvement by staying updated on advancements in machine learning, assistive technologies, andfeedback from users and stakeholders.

5. *MODELLING:*

1. OBJECT DETECTION:

Description: Object detection models identify and locate objects within the user's surroundings, such as pedestrians, obstacles, and traffic signs.

Techniques: Convolutional Neural Networks (CNNs) are commonly used for object detection tasks due to their ability to learn spatial hierarchies of features.

Training Data: Large datasets containing annotated images are used to train the model to recognize different object classes.

2. SCENE RECOGNITION:

Description: Scene recognition models classify the overall sceneor environment the user is in, such as indoors, outdoors, or crowded spaces.

Techniques: Deep learning models, including CNNs and recurrent neural networks (RNNs), can be employed for scene recognition by learning patterns and features from input sensor data.

Training Data: Datasets containing labeled images or sensor data representing different environmental conditions are used to train the model.

3. OBSTACLE AVOIDANCE*:*

Description: Obstacle avoidance models predict potential collisions with obstacles in the user's path and plan safe navigation routes accordingly.

Techniques: Reinforcement learning algorithms or predictive models, such as decision trees or support vector machines (SVMs), can be used for obstacle avoidance by learning from pastexperiences and sensor inputs.

Training Data: Historical sensor data paired with collision outcomes or user feedback are used to train the model to predict safe navigation actions.

4. USER FEEDBACK GENERATION*:*

Description: Feedback generation models convert model outputsinto actionable feedback for the user, such as auditory alerts or haptic vibrations.

Techniques: Rule-based systems or generative models, such as sequence-to-sequence models or generative adversarial networks(GANs), can be used to generate feedback based on model predictions and user preferences.

Training Data: Human-labeled data or expert-designed rules are used to train the feedback generation model to produce relevant and interpretable feedback.

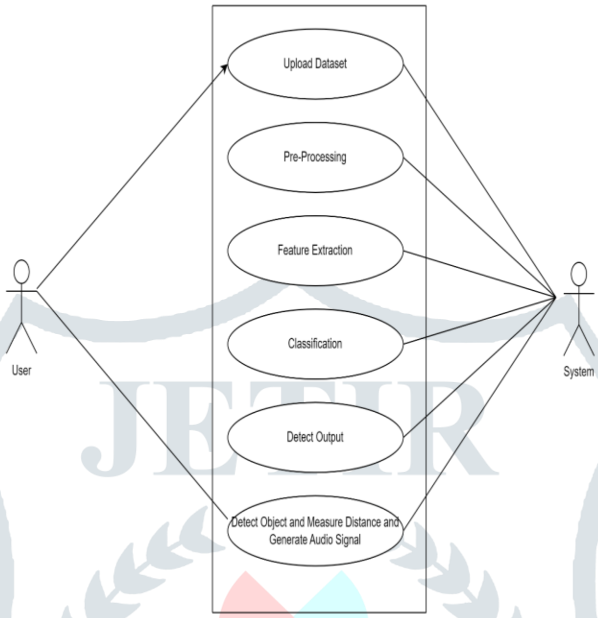
5. CONTINUOUS LEARNING AND ADAPTATION*:*

Description: Continuous learning models update and adapt over time based on user feedback and new data, improving system performance and adaptability.

Techniques: Online learning algorithms, transfer learning, or ensemble methods can be used to continuously update and refine the models based on incoming data and user interactions.

Training Data: Real-time user feedback and additional labeleddata collected during system operation are used to update and fine-tune the models

DATAFLOW DIAGRAM:



A. INPUT DATA*:*

Sensor Inputs: Data from sensors such as cameras, LiDAR, and inertial sensors capturing information about the user's surroundings.

User Inputs: Commands or gestures from the user interacting with the system.

B. PREPROCESSING*:*

Sensor Fusion: Integration of data from multiple sensors tocreate a comprehensive representation of the environment.

Data Filtering: Removal of noise and irrelevant informationfrom sensor inputs.

C. FEATURE EXTRACTION*:*

Computer Vision: Extraction of visual features from camerainputs, such as edges, colors, and object shapes.

LiDAR Processing: Extraction of spatial features and distancemeasurements from LiDAR data.

Inertial Data Analysis: Analysis of motion and orientation datafrom inertial sensors to understand the user's movement.

D. MACHINE LEARNING MODELS*:*

Object Detection: Identification and classification of objects inthe environment, such as pedestrians, vehicles, and obstacles.

Scene Recognition: Classification of the overall scene orenvironment, such as indoor, outdoor, or crowded.

Obstacle Avoidance: Prediction of potential collisions and planning of safe navigation paths. ***E.Decision Making***:

Path Planning: Determination of optimal routes for the user based on detected obstacles, scene information, and user preferences.

Feedback Generation: Generation of feedback for the user,including auditory alerts, haptic feedback, or voice instructions.

E. USER INTERFACE*:*

Audio Output: Conversion of feedback generated by the systeminto spoken instructions or auditory alerts.

Haptic Feedback: Provision of tactile feedback through vibration or other tactile interfaces.

Voice Interaction: Allowance for user interaction with the system through voice commands or speech recognition.

F. USER INTERACTION*:*

Input Handling: Processing of user inputs, such as gestures or voice commands, to control the system and provide feedback.

User Preferences: Customization of system settings and preferences based on individual user needs and preferences.

G. PRIVACY AND SECURITY:

Data Encryption: Encryption of sensitive user data to safeguardprivacy and prevent unauthorized access.

On-Device Processing: Execution of data processing and model inference on the user's device to minimize data transmission and enhance privacy.

H. CONTINUOUS LEARNING AND IMPROVEMENT*:*

Model Updating: Periodic updating of machine learning models using feedback from users and new training data to enhance performance and adaptability.

System Monitoring: Monitoring of system performance and user feedback to identify areas for improvement and prioritize future development efforts.

IV. RESULTS AND DISCUSSIONS

This technology enables blind people to navigate their surrounding by not taking help from others. The purpose of this project is to help the blind people to sense the visualizationof world around them without depending on others.

V. CONCLUSION AND FUTURE SCOPE

Overall, the system performed well during the testing with high success rates for human identification tasks. The system also complicated with relevant regulations and was found to be secure and user-friendly. However, there is room for improvement in the accuracy of identifying unknown individuals and tracking moving objects. It is recommended that the system undergo further testing and refinement to improve these areas. This paper presents a novel blind assistance system that harnesses the power of machine learning to provide real-time contextual information to visually impairedindividuals. Through efficient image acquisition, featureextraction and intelligent decision-making, the system successfully and navigation, object detection, obstacle avoidance and landmark recognition. The experimental evolution implements the systems effectiveness and benefits highlighting the potential for future enhancements and widespread adoption in assisting the visually impaired community.

REFERENCES

1. F. Catherine, Shiri Azenkot, Maya Cakmak, "Designing a Robot Guide for Blind People in Indoor Environments," ACM/IEEE International Conference onHuman-Robot Interaction Extended Abstracts, 2015.
2. H. E. Chen, Y. Y. Lin, C. H. Chen, I. F. Wang, "Blindnavi: a mobile navigation app specially designed for the visually

impaired," ACM Conference Extended Abstracts onHuman Factors in Computing Systems, 2015.

3. K. W. Chen, C. H. Wang, X. Wei, Q. Liang, C. S. Chen, M. H. Yang, and Y. P. Hung,"Vision-based positioning for Internet-of-Vehicles," IEEE Transactions on Intelligent Transportation Systems, vol. 18, no.2, pp. 364–376, 2016.

4. M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke,S. Roth, and B. Schiele, "The Cityscapes Dataset for Semantic Urban Scene Understanding," IEEE Conference on Computer Vision and Pattern Recognition,2016

5. J. Ducasse, M. Macé, M. Serrano, and C. Jouffrais, "Tangible Reels: Construction and Exploration of Tangible Maps by Visually Impaired Users," ACM CHI Conference on HumanFactors in Computing Systems, 2016.

6. J. Engel, T. Schops, and D. Cremers, "LSD-SLAM: Large- scale direct monocular SLAM," European Conference on Computer Vision, 2014

7. . S. Gilson, S. Gohil, F. Khan, V. Nagaonkar, "A Wireless Navigation System for the Visually Impaired," Capstone Spring, 2015.

8. J. Guerreiro, D. Ahmetovic, K. M. Kitani, and C. Asakawa, "Virtual Navigation for Blind People: Building Sequential Representations of the Real-World," International ACM SIGACCESS Conference on Computers and Accessibility, 2017.

9. Kendall, M. Grimes, and R. Cipolla, "PoseNet: a convolutional network for real-time 6-DOF camera relocalization," International Conference on Computer Vision, 2015.

10. Kendall, and R. Cipolla, "Geometric loss function for camera pose regression with deep learning," International Conference on Computer Vision, 2017