MINISTERUL EDUCAȚIEI

UNIVERSITATEA TEHNICĂ
DIN CLUJ-NAPOCA

**FACULTATEA DE ELECTRONICĂ, TELECOMUNICAȚII
ȘI TEHNOLOGIA INFORMAȚIEI**

MASTER TELECOMUNICAȚII/ TEHNOLOGII MULTIMEDIA

ACP 1

Bracelets for people with visual impediments

Oscar GAL, I TM

Coordonator: Sl.dr.ing. Rajmond JANO

2021

str. Memorandumului nr. 28, 400114 Cluj-Napoca, România
tel. +40-264-401200, fax +40-264-592055, secretariat tel. +40-264-202209, fax +40-264-202280
www.utcluj.ro

Introduction

According to the World Health Organization statistics [1], there are approximately 285 million people who are blind, have severe low vision, or are near-blind. Out of these, 39 million are blind and 246 million have low vision problems. About 90% of these people live in developing countries. Romania is a country in which the cases of blindness are at some pretty high odds: 95.889 people who are blind from which 2.770 are children.

In this article, a solution that might come in handy for all the people with visual impairments, is presented. The solution combines the latest technologies from the Artificial Intelligence field, hardware and software, creating a product that could help people with vision impediments. Simply put, we started from the idea of having bracelets capable of "seeing" and perceiving the outside world and transmitting that information to the person in need.

The main idea of this project is to create 4 bracelets (one for each arm and one for each leg), which will have cameras and gyroscopes attached. The cameras with the help of artificial intelligence will be able to detect objects from the images. The gyroscope will also determine the positions of the arms and legs for a better understanding of the outside world. For instance, the bracelets from the legs will be able to determine if there is a pothole on the sidewalk and it will alert the person in need. Also, the bracelets from the hand will be able to determine on which object the person is going to put his or her hands on (Figure 1).

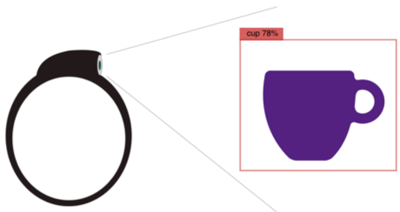

Figure 1. How the bracelets work

All the data retrieved from the bracelets will be stored in an application on the user's phone. This application will take care of all the computation for predicting what is in an image or not. We will also be able to compute the distance between the object and the hand of the user (an approximate distance) and we will use that to notify the user about the object that he or she is going to put the hand on, and not about all the objects that the camera is going to detect.

We are going to communicate the data to the user via a special pair of glasses which transmits the sound using vibrations behind the ear. The reason why we chose this method in order to let the user perceive the outside world, is that people with visual impediments base their lives on hearing more than people who doesn't have a visual problem, that is because they already lost one of their senses, we cannot use headphones or anything that would block the hearing.

This idea can have other multiple usages, for instance when a user with visual impediments goes to a shop and tries to buy clothes, these bracelets can help that person to better understand what color is that blouse and also what type.

The possibilities and usages of this idea can never stop. As the technology evolves, we could add more and more functionalities to the bracelets and also to the artificial intelligence in the app and the end game would be to have these bracelets that totally “replace” the capability of seeing.

Chapter 1: Artificial Intelligence in recognizing objects

When a person tries to figure out what object is in an image (identify and object), we are talking about object recognition. When we, humas, look at a photograph or we watch a movie, we can readily spot people, objects, landscapes, scenes and visual details. Object recognition is the output of deep learning and machine learning algorithms [2]. The main goal of our application will be to make the artificial intelligence app to gain a level of understanding of what an image contains. Example can be seen in Figure 2.

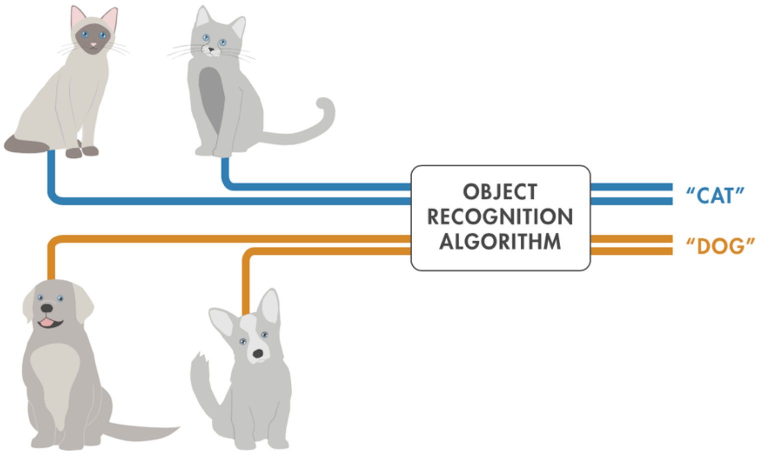

Figure 2 Using object recognition to identify different categories of objects. [3]

Object recognition is a key technology behind driverless cars, recognizing patterns and so on. This technology enables car for instance to recognize a stop sign or to distinguish a pedestrian from a lamppost. It is also useful in a variety of applications such as disease identification in bioimaging, industrial inspection, robotic vision and we will also use it to determine the objects for our application.

To create such an object recognition algorithm, there are a variety od approaches. Recently, techniques in machine learning and deep learning have become popular approaches to object recognition problems. Both techniques learn to identify objects in images, but they do have different execution.

The following section explains the differences between machine learning and deep learning for object recognition, and it shows how to implement both techniques.

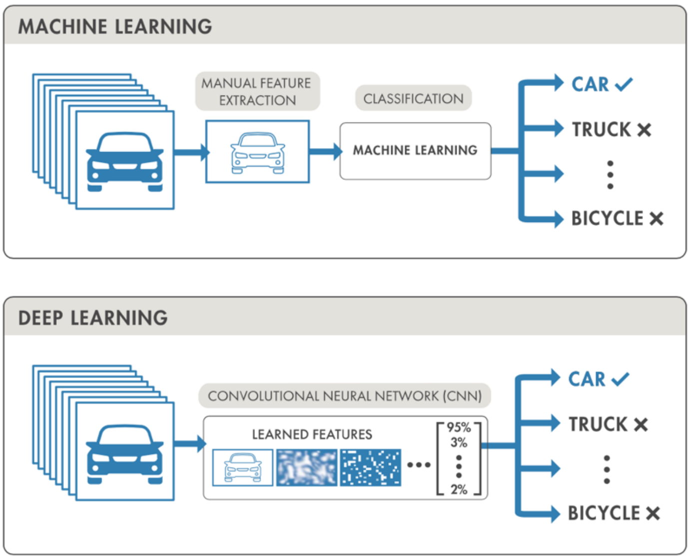

Figure 3 Machine learning and deep learning techniques for object recognition.

Deep learning techniques have become a popular method for doing object recognition. Deep learning models such as convolutional neural networks, or CNNs, are used to automatically learn an object's inherent features in order to identify that object. For example, a CNN can learn to identify differences between cats and dogs by analyzing thousands of training images and learning the features that make cats and dogs different.

There are two approaches to performing object recognition using deep learning:

1. Training a model from scratch: To train a deep network from scratch, we gather a very large labeled dataset and design a network architecture that will learn the features and build the model. The results can be impressive, but this approach requires a large amount of training data, and we need to set up the layers and weights in the CNN.
2. Using a pretrained deep learning model: Most deep learning applications use the transfer learning approach, a process that involves fine-tuning a pretrained model. We start with an existing network, such as AlexNet or GoogLeNet, and feed in new data containing previously unknown classes. This method is less time-consuming and can provide a faster outcome because the model has already been trained on thousands or millions of images.

Deep learning offers a high level of accuracy but requires a large amount of data to make accurate predictions.

Machine learning techniques are also popular for object recognition and offer different approaches than deep learning. Common examples of machine learning techniques are:

1. HOG feature extraction with an SVM machine learning model
2. Bag-of-words models with features such as SURF and MSER
3. The Viola-Jones algorithm, which can be used to recognize a variety of objects, including faces and upper bodies

To perform object recognition using a standard machine learning approach, we start with a collection of images (or video), and select the relevant features in each image. For example, a feature extraction algorithm might extract edge or corner features that can be used to differentiate between classes in our data.

These features are added to a machine learning model, which will separate these features into their distinct categories, and then use this information when analyzing and classifying new objects.

We can use a variety of machine learning algorithms and feature extraction methods, which offer many combinations to create an accurate object recognition model.

Using machine learning for object recognition offers the flexibility to choose the best combination of features and classifiers for learning. It can achieve accurate results with minimal data.

Chapter 2: The application

The application will have multiple components. The first component is the server on which all the computations and the image processing are done.

The images taken from the bracelets will be send to a server[4] which will determine what objects are in the image using the artificial intelligence and the image processing. After all the computation is done, the response will be sent back to the app and the app will alert the user about what is in the image.

The application will be implemented using React Native[5], because of its cross-platform feature, and we want to have as many user on the application as we can. Also, the application will have a desktop version on which the user can set up the bracelets and add them in their account for an easier management.

The application needs to have a UI/UX[6] that facilitates easy navigation, so people with visual impediments won't face difficulties while using the platform.

Down below, some screenshots from the application can be seen. (Figure 3, Figure 4)

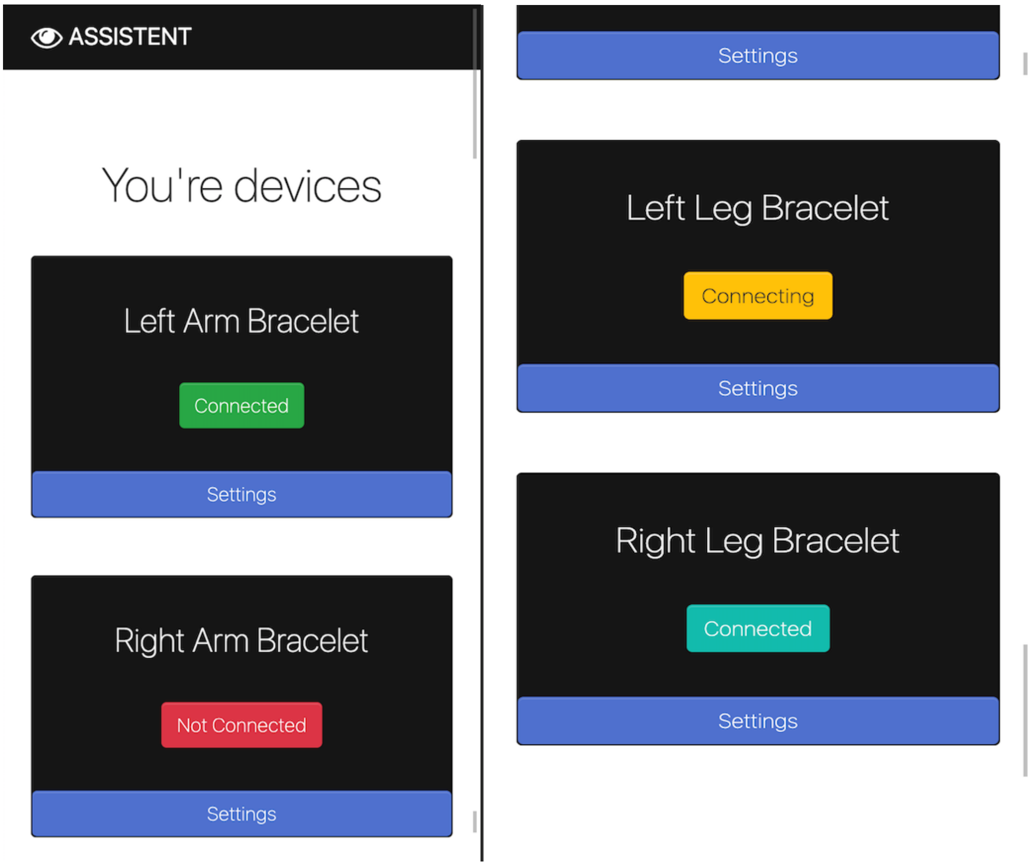

Figure 3-4 Devices connected to the application

In the figure is displayed the homepage after login. On that page the user can see all the bracelets connected in their account. In this page the user can see the status of each bracelet (Connected, Connecting and Not Connected). Also, we will add the capability to see the percentage of battery for each device connected. If the settings button is pressed, the user will be redirected to the settings page for each device. On the settings page we can edit the bracelets name, also the information stored on the app can be managed (due to the GDPR rules). On the same page the user can delete (unlink) a bracelet from their account and can also trigger a sound in case the bracelet was misplaced.

In Figure 5 we can see the log in page. All the authentication process will be one safe to use. We will have the creation of an account using email validation for more safety measures, because after all with these bracelets we can see what a person sees and also we can know the location of a person every moment. The login process will be secured with email verification, also for security reasons. We needed to add the feature for recover password in case someone misplaced their password.

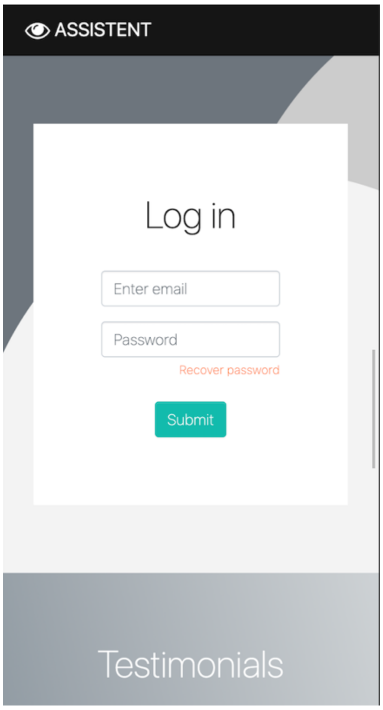

Figure 5 Log in

Bibliografie

[1] https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment
[2] https://www.edge-ai-vision.com/2020/04/object-recognition-3-things-you-need-to-know/
[3] https://www.mathworks.com/solutions/image-video-processing/object-recognition.html
[4] https://www.django-rest-framework.org
[5] https://reactnative.dev/docs/getting-started
[6] https://uxdesign.cc/how-visually-impaired-people-navigate-the-web-7f9eab9d9c37