Place Recognition App for Visually Impaired Person

Problem Statement

Place recognition app is the problem of recognizing places for the visually impaired persons with the help of which blind people will see. In this research video is captured with a portable camera device on the client side, and is streamed to the server for real-time image recognition with existing object detection models (YOLO). The 3D location of the objects is estimated from the location and the size of the bounding boxes from the detection algorithm. Then, a 3D sound generation application based on Unity game engine renders the binaural sound with locations encoded. The sound is transmitted to the user with wireless earphones. Sound is play at an interval of few seconds, or when the recognized object differs from previous one, whichever earliest.There are various challenges which need to work upon such as real time responsesfacility, safety and security concerns.

Background Work

Despite of the challenges given in the problem statement self driving car is still an active area of research. Numerous approaches have been proposed over the years.One of them is the mobile app TapTapSee uses computer vision and crowdsourcing to describe a picture captured by blind users in about 10 seconds.The Blindsight offers a mobile app Text Detective featuring optical character recognition (OCR) technology to detect and read text from pictures captured from the camera. Facebook is developing image captioning technology to help blind users engaging in conversations with other users about pictures.introduced a system that uses spatial audio to facilitate discovery of points of interest in large, unfamiliar indoor environments (e.g. shopping mall). One of the recent research trained a CNN particularly for place recognition. Then integrate 3D sound into place recognition results produced by CNN helps visual impaired.

Methodology

1. **Data Collection:**The images in the dataset large 2.5 million) image place recognition data-set containing thousands of places and hundreds of examples of each place under a wide range of environmental conditions.
2. **Preprocessing:**Firstly, we automatically removed images which were pitch black. These images were usually captured at night time in areas where there was no illumination. Then we removed corrupted images produced when the cameras were not functioning correctly

3. **Training:** Training the deep convolution neural network for recognizing the places is done.CaffeNet architecture is used and adjusted to support our 15 categories (classes).Rectified Linear Units (ReLU) are used as substitute for saturating nonlinearities. This activation function adaptively learns the parameters of rectifiers and improves accuracy at negligible extra computational cost.
4. **Testing:**In this phase the author used the test set for prediction.

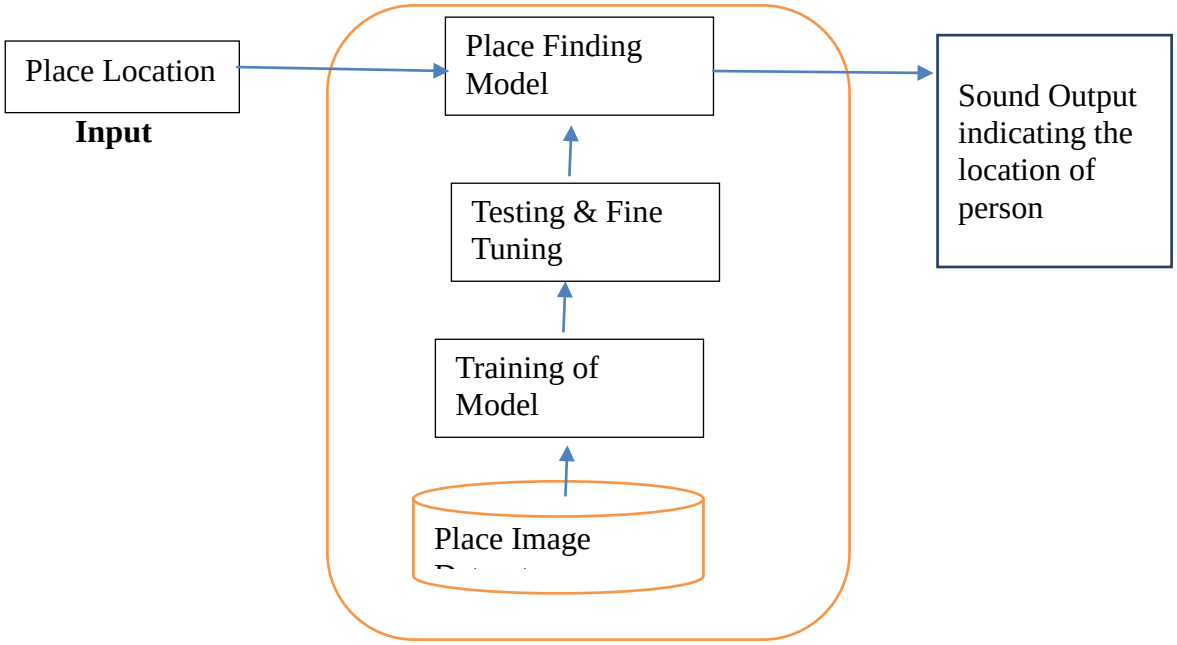

Figure 1: Architecture of Place finding app for visually impaired.

Figure 2:Data Flow Pipeline for visually impaired[Rui (Forest) Jiang Earth Science, Stanford,Qian Lin Applied Physics, Stanford Let Blind People See: Real-Time Visual Recognition with Results Converted to 3D Audio]

Experimental Design
Dataset: Place detection app use datasets such as public dataset of various places.It includes around 1 million place images collected to be used as dataset.

Evaluation Measures:Measures such as accuracy and subject(person) change robustness will be computed and measured as a parameter for accuracy.
Software and Hardware Requirements:Python based Computer Vision and Deep Learning libraries will be exploited for the development and experimentation of the project. Tools such as Anaconda Python, and libraries such as CaffeNet. Training will be conducted on GPU NVIDIA Tesla K40.