# Liang Niu

## Personal Information

---

| | |
|---|---|
| **Status:** | **MS Student** |
| **Program:** | **Computer Science and Engineering** |
| **School:** | **Tandon School of Engineering, New York University** |
| **RA Period:** | **From 2017-09 to 2018-06** |

## Biography

---

**I'm a Ph.D. student at New York University. Before that, I was a research assistant in NYU Multimedia and Visual Computing Lab, advised by Professor Yi Fang. I am broadly interested in 3D Computer Vision, Pattern Recognition and Deep Learning.**

**Research Project:** A Wearable Assistive Technology for the Visually Impaired for for Smart Hand-To-Handle Manipulation

# 1 Description

The visually impaired are consistently faced with mobility restrictions due to the lack of truly accessible environments. Even in structured settings, people with low vision may still have trouble navigating efficiently and safely due to hallway and threshold ambiguity. Assistive technologies that are currently available do not provide door and door-handle object detections nor do they concretely help the visually impaired reaching towards the object. In this paper, we propose an AI-driven wearable assistive technology that integrates door handle detection, user's real-time hand position in relation to this targeted object, and audio feedback for "joy stick-like command" for acquisition of the target and subsequent hand-to-handle manipulation. When fully envisioned, this platform will help end users locate doors and door handles and reach them with feedback, enabling them to travel safely and efficiently when navigating through environments with thresholds. Compared to the usual computer vision models, the one proposed in this paper requires significantly fewer computational resources, which allows it to pair with a stereoscopic camera running on a small graphics processing unit (GPU). This permits us to take advantage of its convenient portability. We also introduce a dataset containing different types of door handles and door knobs with bounding-box annotations, which can be used for training and testing in future research.

# 2 Method

In this project, we develop an AI-driven wearable assistive technology that integrates door handle detection, user's real-time hand position in relation to this targeted object, and audio feedback for "joy stick-like command" for acquisition of the target and subsequent hand-to-handle manipulation. When fully envisioned, this platform will help end users locate doors and door handles and reach them with feedback, enabling them to travel safely and efficiently when navigating through environments with thresholds. As displayed in Figure.1, the pipeline

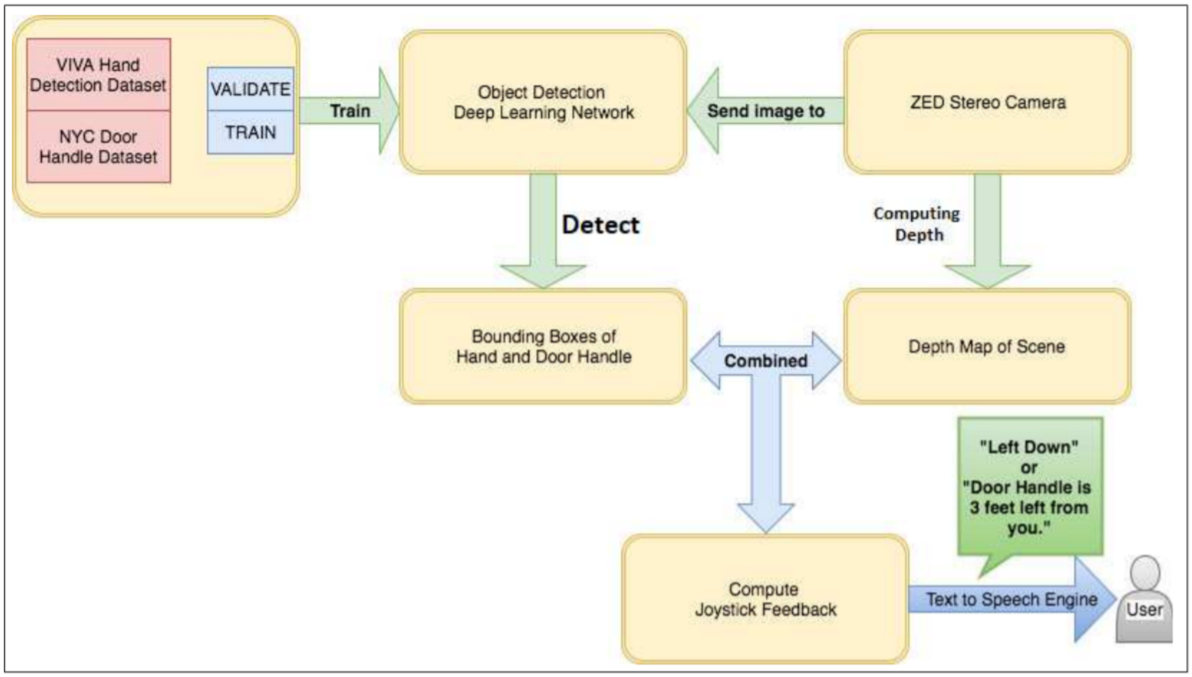



Figure 1: Pipeline of our proposed project.

of this proposed project has three major components. The first component is the deep learning neural network that can perform two-class object detection – door/door handles and human hands, and it will return the corner coordinates of the bounding box enclosing the object. The second component is the image flow and depth information extractor (stereo camera). The third component is a hard-coded program that integrates the information flow: combining the position of the detected objects along with the depth information extracted from the camera to obtain its 3D spatial location. It then transforms such information into descriptive sentences and outputs it with synthesized voice to the users, providing joy stick-like commands to the user, assisting in-hand control with feedback assistance.

## 3 Results

In this section, we compare the result of this research on the popular benchmark dataset for the Vision for Intelligent Vehicles Applications (VIVA) Hand Detection Challenge. These photos were taken from 7 possible viewpoints, including first person view, which is the most common viewpoint in our model (similar to Figure.2) since the camera is mounted in front of our user's chest. The object detection network will return bounding boxes with corresponding "confidence",

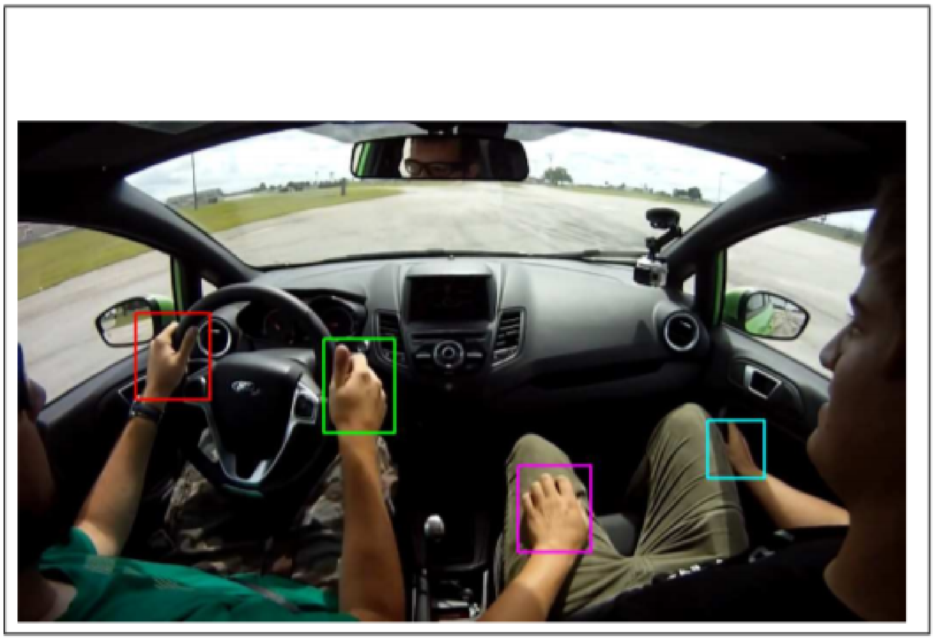
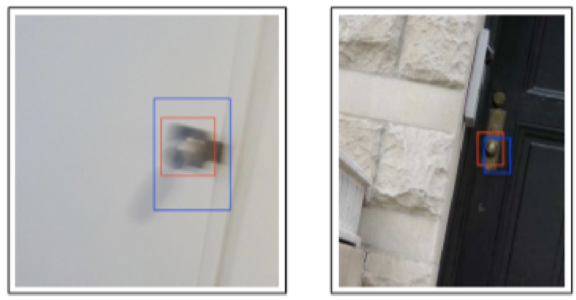


Figure 2: Example of the hand pictures with annotation..

i.e. the probability that it contains our target object, door knob. For example, if the model returns a box with a confidence value of 0.8, it is ”80% sure” that there is a door knob inside this box. Hence, we can set up a proper threshold such as 0.6 to filter out all the uncertain cases and reduce false positives. Some success and failed/inaccurate cases of the Model 1 predictions are shown below. The model successfully detected almost all door knobs appearing in the pictures (14 out of 15 testing samples) with any reasonable confidence threshold between 0 and 0.6, although lower thresholds tend to let our model return more false alarms. Moreover, among those successful cases, the model accurately and decisively highlights the localization of door knobs with high confidence. When a confidence threshold of 0.6 is set, the model prediction acts as if it is the ground truth data, where the bounding boxes look almost like human annotations – no false positive or false negative among the 14 selected testing samples (Figure.3). For experimentation purposes, different thresholds have been tested in order to observe its impact as well as to search for the optimal value. When the threshold drops to 0.2, only one picture starts with a false positive; when it drops even lower to 0.005, about one third (5 out of 15) of the tests resulted in failed/inaccurate predictions. Aside from some clean detections, there are some samples that give multiple bounding boxes with low threshold settings (0.005) and they are listed in Figure.4. In Figure.4a, despite the fact that both the shape and features of such a blurry door knob are barely recognizable by human eye, our model still detected it with confidence. It was also accompanied by another larger blue box, presumably because the blurriness obscured the boundaries. Nonetheless, the general position (the centroid) is not significantly affected. Secondly, in Figure.4b, the model gives two boxes, but this time the centroid coordinates are distinctively different. Although it correctly covers all parts of the knob, the blue box is severely skewed and it includes too much marginal space. The cause behind this skewness is somewhat unclear, but the dark color can be a possible factor.

(a) (b)

Figure 3: Successful cases.



(a) (b)

Figure 4: Inaccurate cases..