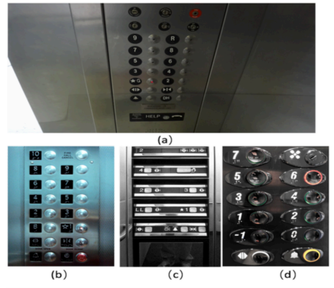
Recognizing Elevator Buttons and Labels for Blind Navigation

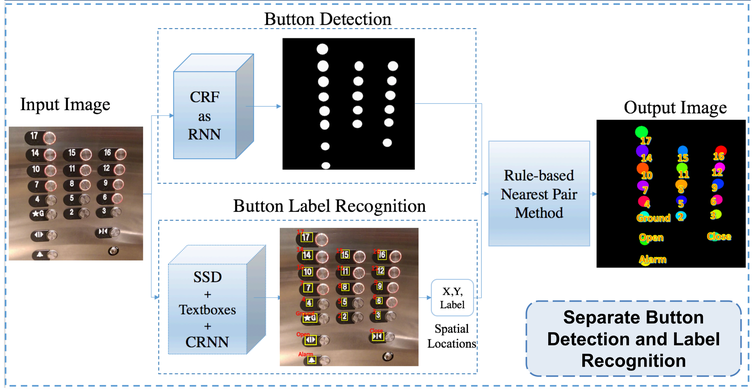
Jingya Liu and Yingli Tian Email: {jliu1, ytian}@ccny.cuny.edu

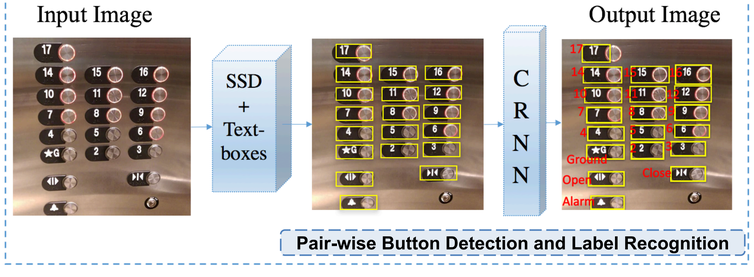
Motivation



For indoor navigation and way-finding, elevators are the most common tool to access multiple floors. Although there are braille labels on elevator button panels to assist blind users, it is still very challenging for them to locate the elevator buttons of the floors they want to go. Therefore, we develop a method to detect elevator buttons and recognize the associated text labels from images.

Frameworks





Methods

Separate Button Detection and Label Recognition

Step 1: Elevator Button Detection

This component estimates a semantic segmentation mask for elevator buttons at pixel-level using conditional random fields as recurrent neural networks (CRFasRNN). It takes an image as the input and outputs the pixel level segmentation mask which classifies the elevator buttons from background. The elevator button mask contains pixel-level segmentation regions of elevator buttons and reserves their location and shape information.

Step 2: Elevator Button Label Recognition

The text and spatial locations of button labels are recognized by descriptions a single-shot detector. A single-shot detector (SSD) based network is integrated with non-maximum suppression to generate a set of text candidates at region level including letters and numbers.

Step 3: Elevator Button Identification

By applying a rule-based method to search the nearest label around the button, each button is then associated with its corresponding label. For each detected button, we calculate the Euclidian distance between the centers of button and the recognized labels. The label with the closest distance is identified as the corresponding label of the button.

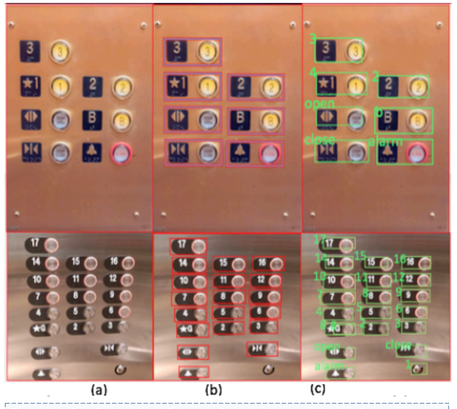
Pair-wise Button Detection and Label Recognition

To avoid the the accuracy eliminates due to the accumulated loss of features through each component step, we consider button and label as a whole area. SSD network is employed to detect and generate text candidate regions for pair-wise button and label for recognition. CRNN is then applied to recognize the semantic meaning, which refers to digital numbers, or symbols in these regions of each label panel in the candidate region.

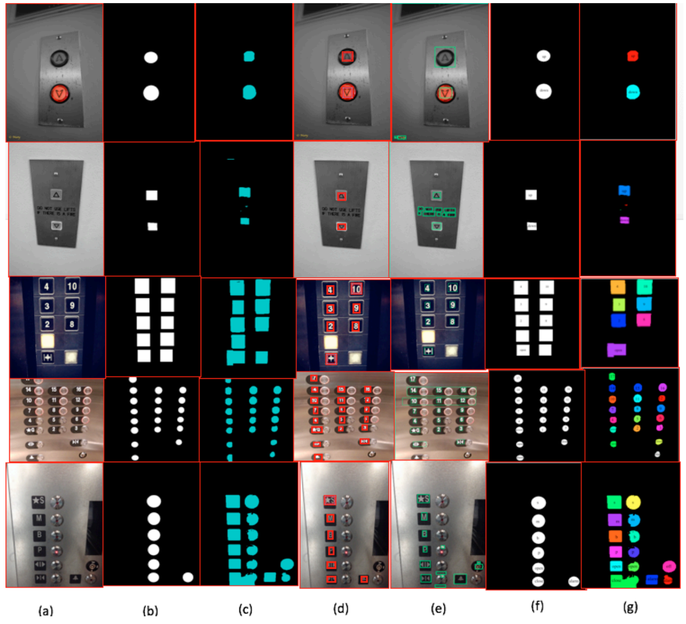
Experimental Results

Model	Step	Accuracy
Separate Button Detection and label Recognition	Pre-selection	92.32%
	Elevator Button Detection	74.2%
	Elevator Button Label Recognition	71.9%
	Elevator Button Identification	73%
Pair-wise Button Detection and label Recognition	-	79.2%

Results for Button and Label Detection and Recognition



Results for Pair-wise Button Detection and Label Recognition (a) Input images. (b) Ground-truths of button detection. (c) Results of button recognition.



Results for Separate Button Detection and Label Recognition

(a) Input images. (b) Ground-truths of button detection. (c) The results of button recognition. (d) Ground-truths of label. (e) The results of label recognition. (f) Ground-truths of button identification. (g) The results of button. identification.

Conclusion

We present a new method for assistive blind navigation by combining the object semantic segmentation with text recognition to detect and recognize elevator buttons and labels using a cascade network. The proposed method can be extended to segment objects with associated text descriptions for many applications. Our future work will focus on implementing the proposed method on a mobile device, developing a user-friendly interface, providing effective audio feedback about the location and label of the queried elevator button, and evaluating the developed system and interface by blind users.

This work was supported in part by NSF grants EFRI-1137172, IIP-1343402, and IIS-1400802.