Using Convolutional Neural Networks For Indoor Robot Localization

Maddie AlQatami, Avik Bosshardt, Angel Sylvester, Susan Fox

Overview

- Initially, the robot partially relied on image feature matching and KD-Tree algorithms to identify its location. We fully replaced this system with two convolutional neural networks: one to identify robot heading, and one to identify location. These networks improved upon a network constructed in 2018 by taking additional input.
- A map of Olin-Rice is divided into cells. The turtlebot can travel to user-specified target cells, avoiding obstacles using lower-level potential field behaviors.
- In the event that the turtlebot gets lost, the localizer feeds image data into the networks, which will then tell it where it is.

Data Collection

- Continued mapping and dividing the Olin-Rice building into 271 approximately 2m x 2m cells.
- New data collection methods were developed to streamline the process.
  - Recording and Marking: We developed an android app which used a phone compass to automate the collection of heading data, and made it easy to manually mark the robot's location while it was driven by another person. A separate script recorded images from the turtlebot.
  - Data from the phone app is manually modified to include XY coordinates of each cell, then a script is run to pair the app measurements with the images.

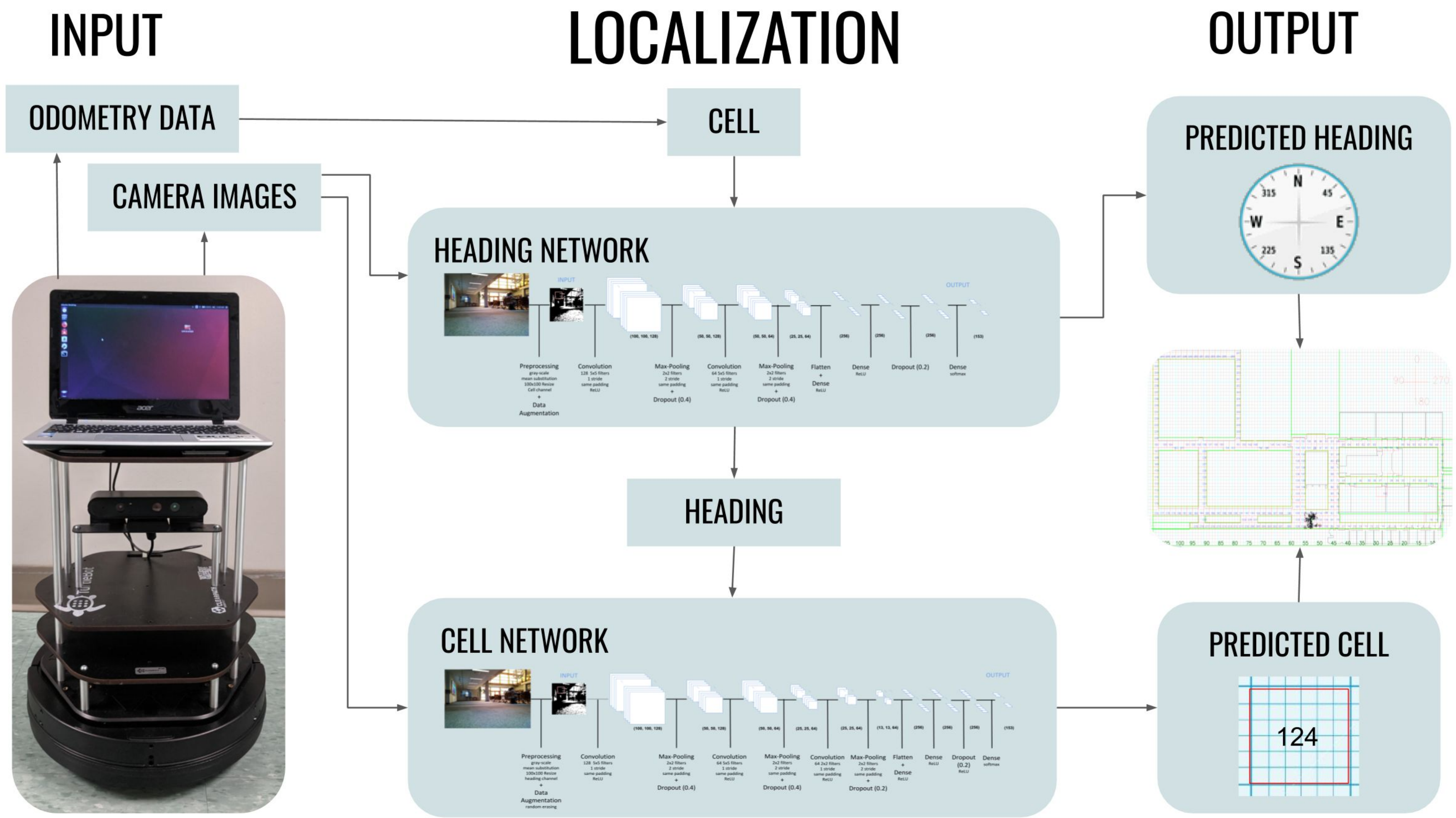

Figure 1. A diagram showing how the neural networks are used to localize the robot.

Convolutional Neural Network

- "A Convolutional Neural Network (CNN) is a deep learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other." (1)

References

1. /@_sumitsaha_. "A Comprehensive Guide to Convolutional Neural Networks-the ELI5 Way." Medium. December 17, 2018. Accessed July 17, 2019. https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a5
2. Zafari, F. (2019). 1A Survey of Indoor Localization Systems and Technologies. *2019*.

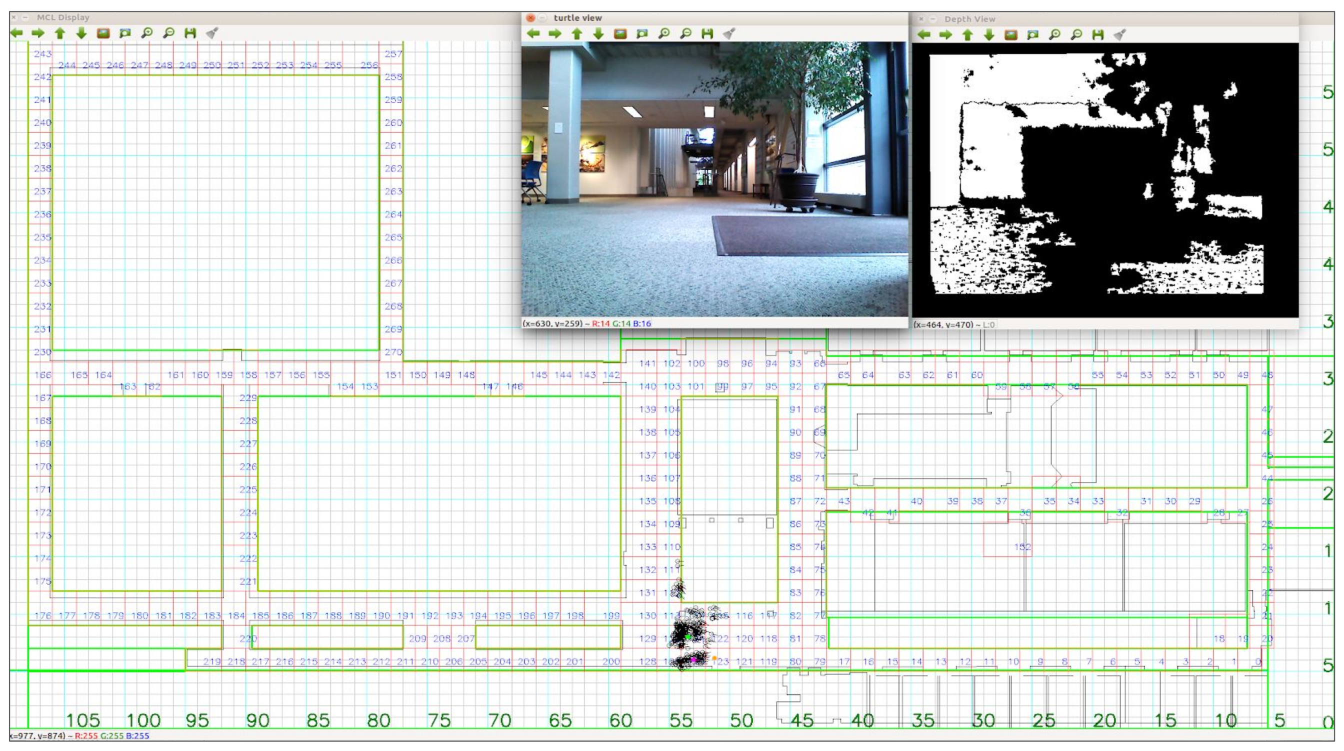

Figure 2. Screenshot of turtlebot using CNN localization to navigate Olin-Rice science building.

Localizing CNNs

- In order to improve localization using images, we built two CNNs: one to identify the robot's heading, and one to identify its location on the map.
- Preprocessing: Before training, images were preprocessed by converting them into gray-scale, resizing to 100 x 100, and subtracting the mean image. The network which predicted headings had the cell data embedded into an extra channel added to the image, and the network predicting cells had heading data embedded into an extra channel.
- Data Augmentation: Erased a random rectangle of pixels from each image to augment data before selecting 500 images per cell in such a way that each heading was relatively equally represented.
- Experiments
  - Training: Trained with different image sizes, preprocessing techniques, architectures, activation functions and hyperparameters to have better validation accuracy.
  - Retraining: Implemented and tested various published architectures including VGG16, PoseNet, ResNet50, InceptionV3, Xception, VGG19, DenseNet, and NasNet by retraining the top layers only, with limited success.

Applications

- Localization indoors or without the use of GPS signals
  - accessibility aids
  - asset management and tracking
  - disaster management and recovery (2)

Directions for Future Work

- Incorporate object classification to determine whether the robot is looking at something not useful, such as a blank wall, so that the robot could be instructed to turn.
- Explore Long-Short-Term Memory (LSTM) network architectures, which require sequential data as input.
- Improve dataset to increase network accuracy in situations where it may perform worse:
  - Remove blurred images
  - Collect data in different lighting conditions
  - Collect data with different artwork on the walls
- Transfer existing networks to other robots.

Acknowledgements

This research is funded by Clare Boothe Luce Foundation and Macalester College.