Testing Different Models for Indoor Localization

MAC

Arif Zamil, Analeidi Barrera, and Susan Fox

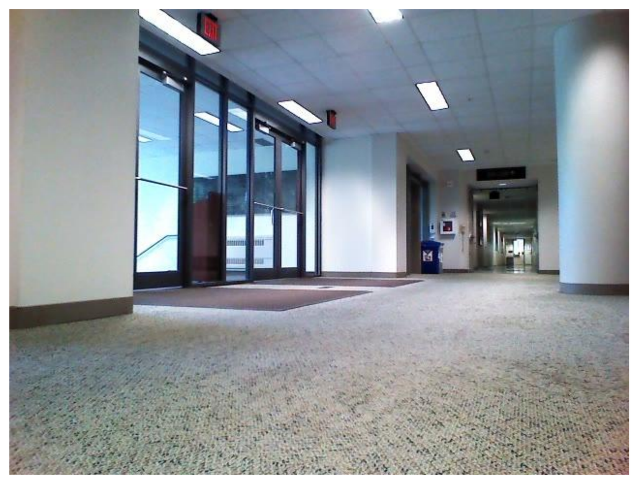
Macalester College – St. Paul, MN

Overview

- **Task:** given an image, predict its cell and heading
- **From 2019:** Two CNNs
  - Input: image + cell   Output: heading
  - Input: image + heading   Output: cell
  - Both with >90% accuracy
- **Summer 2020:** networks where input is image only
  - CNNs: image → cell, image → heading
  - CNN-LSTM
  - CNN-Regression

Dataset

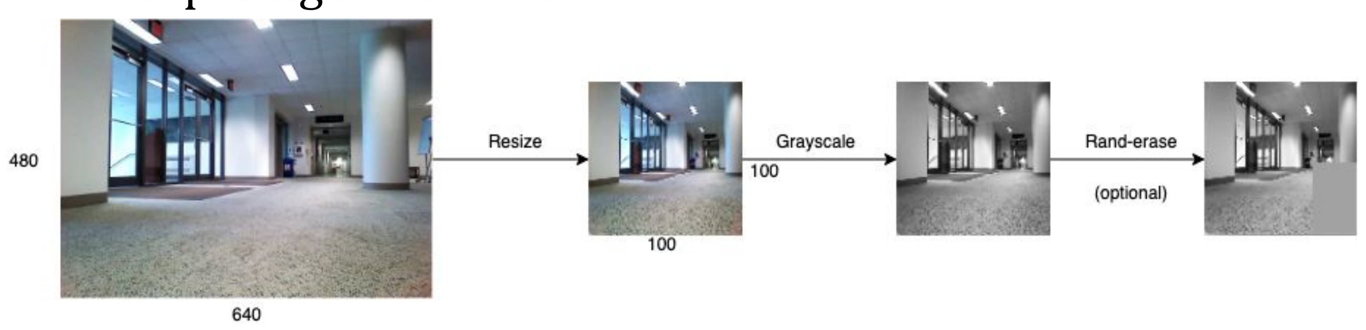
- 95,000+ images, 2nd floor of Olin-Rice, taken by our robot
- For each image we know:
  - cell number (cell = 2m by 2m section)
  - (x, y) coordinates (in meters)
  - heading (N, NE, E, SE, S, SW, W, NW)



cell=140, x=30.95, y=56.15, head=NE

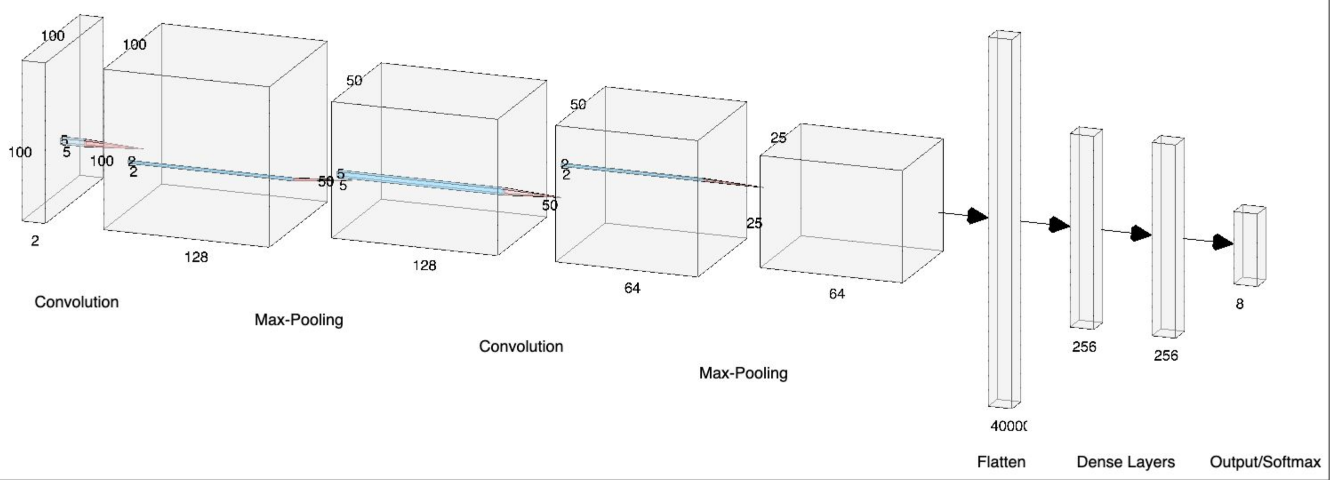
cell=263, x=47.00, y=78.00, head=S

- Preprocessing of images:
  - Selecting/creating duplicate images so each cell had 500 images
  - Preparing for network:



What is a CNN?

- CNN = Convolutional Neural Network
- Machine learning, datasets of images or similar data
- Learns to detect useful features in data
- Learns to associate features with specific outputs
- Layers perform different tasks
  - **Convolutional:** contains multiple "convolutional filters"; each filter learns to detect a feature
  - **Max-pooling:** reduces data size
  - **Flatten:** converts data to 1-dimensional
  - **Dense:** correlates patterns across data
  - **Softmax:** produces output patterns as probabilities



Localization with CNNs

- Although the previous CNN models had high accuracy in predicting cells and headings, we needed to simplify these models to work with an LSTM
  - New models take only image as input
  - CNN has reduced numbers of filters
  - Dropout was unified across network
- Performance of new CNNs
  - <Add here> new training data, how much training versus testing
  - <add> each network, report number of epochs and final accuracy

Data Arrangement

- Sample data:

1	1	2	2	2	3	4	4	5	6

- CNN
  - Data must be randomized
  - CNNs can "forget" data not seen for a while

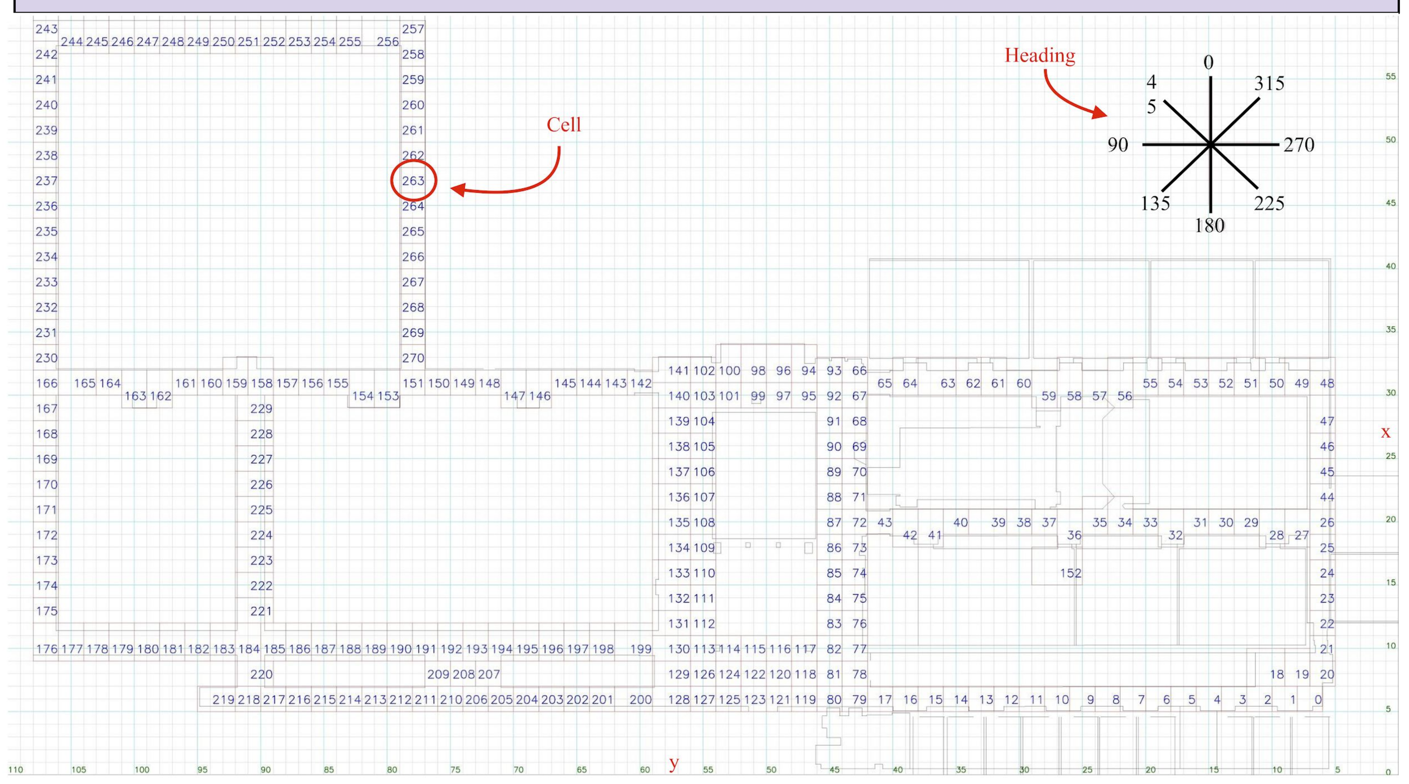
3	2	4	1	2	6	4	2	1	5

- LSTM
  - LSTM needs time/space ordered sequences of images
  - All sequences the same length
  - Our dataset was not configured this way by default!
  - We reorganized the data for use with LSTM

1	1	2	2	2
2	2	3	4	4
3	4	4	5	6

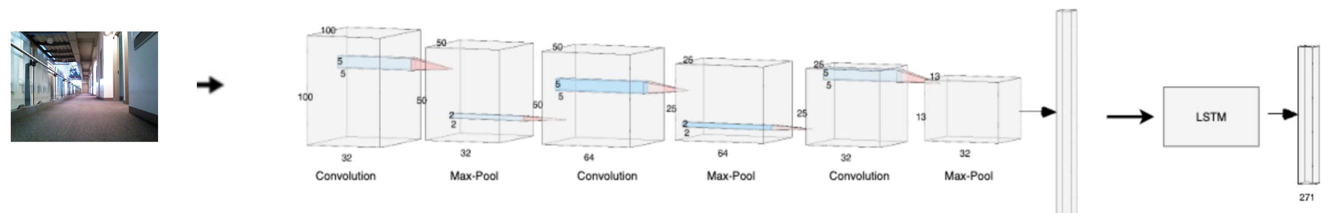
- CNN-Regression
  - Randomized data like CNN
  - (x, y) coordinate data had to be cleaned!

Map

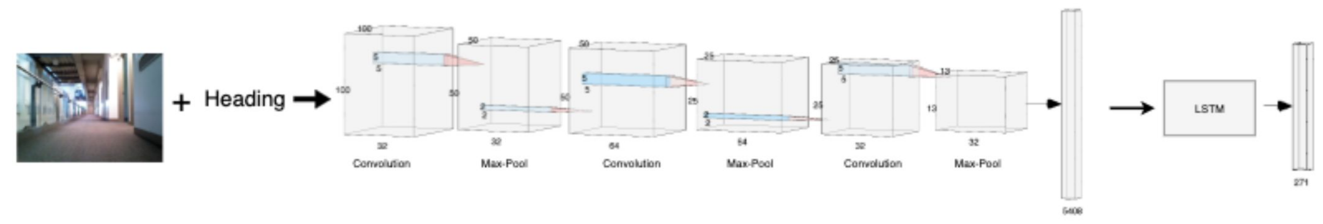


Localization with LSTMs

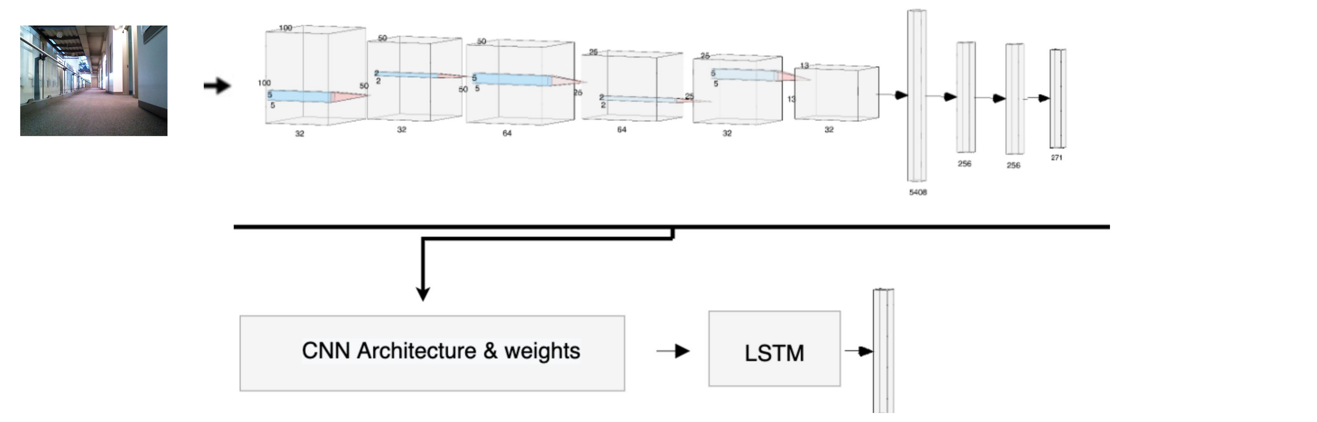
- LSTM = Long Short-Term Memory
- A recurrent neural network architecture, it remembers previous images to help predict the current step
- CNN-LSTM starts with CNN, then has LSTM layers
- **Results:**
  - CNN-LSTM succeeded predicting headings
  - CNN-LSTM did poorly predicting cells
  - No real improvement over CNN alone
- **Experimental architectures:**
  - CNN+LSTM, which consumed too much memory



  - CNN + LSTM with both images and headings as input



  - Trained CNN model & weights + LSTM (Transfer Learning)



Localization with Regression

- Regression Neural Networks predict an output variable as a function of the inputs: they can produce real-valued outputs
- Input features (independent variables) may be any type
- Output must be numeric (integer or floating point)
- <why use>
- First experiment: modify old CNN with regression layer at output
  - Network was too small, could not learn
  - <add details? dataset, number epochs, …>
- Second experiment: use a large existing regression network, GoogleNet and apply transfer learning
  - Unable to get this network working
  - Concerned about training/running time of network

Conclusions & Future Direction

- Basic CNNs work well for this task
  - We were able to simplify CNN and still get good performance
- Training CNN-LSTM is expensive and difficult
  - Requires complete revision of dataset
  - Benefits of time-sensitive learning were not obvious
  - More testing to be done
- Training Regression CNN requires more network complexity
  - More need to examine this in the future
- Testing on actual robot needs to happen
- Augmenting of dataset would be helpful