CNN based Indoor Localization using RSS Time-Series

Yang Yang  Meng Wei

Contents

- Introduction of CNN
- System Architecture
- Four Approaches
- Data Processing
- Achievements
- Problems
Basic Concepts of CNN

The Basic Structure of CNN

Why CNN?
Temporal Dependency
Processing data has grid-like topology. Such as time-series data and image data.

Convolution

The size of filter is $3 \times 3$ and the possible combination of $3 \times 3$ in $6 \times 6$ is $4 \times 4$. In general, if we have a $n \times n$ and a $f \times f$ filter, the dimension of the output will be $(n - f + 1) \times (n - f + 1)$.

Disadvantage

1. The image will shrink after every convolutional operator;
2. Pixels on the corners or on the edges will be thrown away;
Padding

To avoid the disadvantages of convolution, padding the image.

Valid convolution: no padding
Same convolution: pad so that output size is the same as the input size

\[ p = \frac{f - 1}{2} \]

(#the filter is always odd)

Strided Convolution

Stride=2;

In general, the output dimension is \( A \times (B \times C) = (A \times B) \times C \)
Convolutions or RGB Images

$6 \times 6 \times 3$ is the height $\times$ width $\times$ channel
Different parameters get different features detector

In general, the output dimension can be got by:

$$n \times n \times n_c \times f \times f = (n - f + 1) \times (n - f + 1) \times n_c'$$

$n_c' = $ the number of the filter

System Architecture Using Time-Series

- RSS readings are integer values from 0 to 2.
- RSS readings are integer values from 0 to 4.
- Trained DNNs for regression, and output two values represent the longitude and latitude.
Four Approaches

Floor Prediction Accuracy

<table>
<thead>
<tr>
<th>Input</th>
<th>Convolutional Layers</th>
<th>Hidden Layers Units</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single RSS Vector</td>
<td>-</td>
<td>128, 128, 128</td>
<td>99.9%</td>
</tr>
<tr>
<td>Averaged RSS Time-Series</td>
<td>-</td>
<td>256, 256, 256</td>
<td>97.9%</td>
</tr>
<tr>
<td>Concatenated RSS Time-Series</td>
<td></td>
<td>8, 8, 8</td>
<td>100%</td>
</tr>
<tr>
<td>RSS Time-Series Image</td>
<td>1 layer with 3 out channels and 2 × 2 kernel</td>
<td>8, 8</td>
<td>100%</td>
</tr>
</tbody>
</table>

The Fourth Approach

\[
\begin{bmatrix}
    RSS_{1,1} & RSS_{1,2} & \cdots & RSS_{1,T} \\
    RSS_{2,1} & RSS_{2,2} & \cdots & RSS_{2,T} \\
    \vdots & \vdots & \ddots & \vdots \\
    RSS_{N,1} & RSS_{N,2} & \cdots & RSS_{N,T}
\end{bmatrix}
\]

T: RSS vectors, equals 10 in this model
N: RSS values from the N WLAN access points, equals 520 in UJIIndoorLoc dataset

Data processing

drawing a grid of $D \times D$ (m2) cells on the area

\[
\begin{bmatrix}
    A_{11} & \cdots & A_{1T} \\
    \vdots & \ddots & \vdots \\
    A_{D1} & \cdots & A_{DT}
\end{bmatrix}
\]

assigning each record from the 19937 records to a cell based on its longitude and latitude
sorted them based on their timestamp and put them in groups of $T$ records
Each cell contains a $T \times N$ matrix

\[
\begin{bmatrix}
    RSS_{1,1} & RSS_{1,2} & \cdots & RSS_{1,T} \\
    RSS_{2,1} & RSS_{2,2} & \cdots & RSS_{2,T} \\
    \vdots & \vdots & \ddots & \vdots \\
    RSS_{N,1} & RSS_{N,2} & \cdots & RSS_{N,T}
\end{bmatrix}
\]

#Each group is less than or equal to $S$ seconds
• After training, the best parameter to get the largest data set is $D = 3\ m,$ $T = 10,$ and $S = 60\ seconds$

• gives 5484 records of 10 consecutive readings from the same cell

• Each group of $T$ records can then be used as an RSS time-series sample

---

**Achievements**

<table>
<thead>
<tr>
<th>Traditional Methods</th>
<th>Input</th>
<th>Building Prediction Accuracy</th>
<th>Floor Prediction Accuracy</th>
<th>Localization Mean Error (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single RSS Vector</td>
<td>100%</td>
<td>91.42%</td>
<td>10.25</td>
</tr>
<tr>
<td></td>
<td>Averaged RSS Time-Series</td>
<td>99.91%</td>
<td>96.90%</td>
<td>4.39</td>
</tr>
<tr>
<td></td>
<td>Concatenated RSS Time-Series</td>
<td>100%</td>
<td>99.66%</td>
<td>3.35</td>
</tr>
<tr>
<td><strong>RSS Time-Series Image</strong></td>
<td><strong>100%</strong></td>
<td><strong>100%</strong></td>
<td><strong>100%</strong></td>
<td><strong>2.77</strong></td>
</tr>
</tbody>
</table>

Improved the localization performance successfully.
Time-series of RSS readings is more robust to estimate location.
Time-series data affects the model accuracy.
Problems

- The size of RSS image is 520*10. The dataset is too small to train the model.
- The measurement method is not convenient.
- How to choose the value of T is not given in detail, and the source code.

END