Radar Image Reconstruction from Raw ADC Data using a Parametric Variational Autoencoder with Domain Adaptation

Michael Stephan\textsuperscript{12}, Thomas Stadelmayer\textsuperscript{12}, Avik Santra\textsuperscript{2}, Georg Fischer\textsuperscript{1}, Robert Weigel\textsuperscript{1}, Fabian Lurz\textsuperscript{1}
\textsuperscript{1}Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany
\textsuperscript{2}Infineon Technologies AG, Neubiberg, Germany

**Introduction**

**Goal**
Localization of human targets in indoor environment using a low-cost radar sensor.

**Challenges:**
- Indoor environment occlusions, ghost targets, multipath reflections, etc.
- Limited training data

**Contribution:**
Single-frame human position estimation from the raw ADC data with a variational autoencoder and domain adaptation to overcome the limited training data

**Architecture – Proposed Solution**

**Proposed Solution:**
- Localization via a variational autoencoder
- Parametric layer for FFT sampling learned through proposed CFEL
- Domain adaptation using synthetic data
- Cross-intenna information fusion only after the VAE

\[
\text{loss} = L_{FL} + \beta \cdot L_{DA} + \phi \cdot L_{KL} \\
\text{Focal Loss} \quad \text{Domain Adaptation} \quad \text{KL-Divergence}
\]

**Domain Adaptation**

- To overcome limited real world dataset we use domain adaptation
- Source domain: Synthetic data generated by Matlab
- Target domain: Real sensor recordings
- Pretrain especially angle estimation
- Using weight difference regularization as loss term
- The image shows the loss divergence depending on the influence of the difference regularization to the overall loss term

\[
L_{DA} = \sum_{s \in T} \left( \frac{|a_\text{g} - a_\text{gt}|^2}{|a_\text{g}|^2 + |a_\text{gt}|^2} \right)
\]

**Complex Frequency Extraction Layer**

- Parametric layer: Filter kernels are defined by a finite set of parameter
- Filter kernels are complex frequencies in slow and fast time
- Set of parameter are the frequencies in slow and fast time
- Therefore the CFEL acts as trainable 2D DFT that mimics the 2D FFT preprocessing
- The image shows (a) the raw time-domain data, (b) the range-Doppler image generated by a common $\text{MATLAB}$ and (c) the feature map of the proposed CFEL

**Dataset**

**Synthetic Dataset:**
- Reflections from point targets are simulated for all discretized range-angle bins using Matlab’s phased array toolbox

**Human Target Dataset:**
- Measurements with one real moving target in a typical conference room
- Two to four target measurements created via superposition of one target measurements
- Some augmentation by shifting the one target measurements in range via multiplication with different complex exponentials

**Results**

- Clear improvement over traditional single-frame localization
- Slightly better performance than other autoencoder-based approach in [1], with better performance on general point target angle estimation

**Approach**

<table>
<thead>
<tr>
<th>Description</th>
<th>P.T-Score</th>
<th>Model Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional DSS-CFAR with DIBSCAN</td>
<td>0.61</td>
<td>.</td>
</tr>
<tr>
<td>AE</td>
<td>0.77</td>
<td>1.23 MB</td>
</tr>
<tr>
<td>VAE with complex RDI as input</td>
<td>0.80</td>
<td>1.71 MB</td>
</tr>
</tbody>
</table>

**References**