Demo Abstract: Lightweight, Deep RNNs for Radar Classification

Dhrubojyoti Roy*
Sangeeta Srivastava*
roy.174@osu.edu
srivastava.206@osu.edu
The Ohio State University

Pranshu Jain
anz178419@cse.iitd.ac.in
Indian Institute of Technology Delhi

Aditya Kusupati
kusupati@cs.uw.edu
University of Washington
Microsoft Research India

Manik Varma
manik@microsoft.com
Microsoft Research India
Indian Institute of Technology Delhi

Anish Arora
arora.9@osu.edu
The Ohio State University
The Samraksh Company

ABSTRACT

We demonstrate Multi-Scale, Cascaded RNN (MSC-RNN)[1], an energy-efficient recurrent neural network architecture for real-time micropower radar classification. Its two-tier architecture is jointly trained to reject clutter and discriminate displacing sources at different time scales, with a lighter lower tier running continuously and a heavier upper tier invoked infrequently on an on-demand basis. It offers for single microcontroller devices a better trade-off in accuracy and efficiency, as well as in clutter suppression and detectability, over competitive shallow and deep alternatives.

ACM Reference Format:
Dhrubojyoti Roy, Sangeeta Srivastava, Pranshu Jain, Aditya Kusupati, Manik Varma, and Anish Arora. 2019. Demo Abstract: Lightweight, Deep RNNs for Radar Classification. In *BuildSys '19: ACM Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, November 13–14, 2019, New York, NY*. ACM, New York, NY, USA, 3 pages. https://doi.org/xx.xxxx/xxxxxxx.xxxxxxx

(a) Hardware **(b) Example setting for indoor demo**

Figure 1: We use an integrated micropower Pulsed Doppler Radar (PDR)-ARM Cortex M3 mote in our demo. An example deployment in an indoor setting is illustrated

1 INTRODUCTION

Efficient data-driven discrimination of targets and activities is apropos in diverse embedded contexts of smart cities and other built environments. Many applications have resulted for activity recognition, noise complaint discrimination, active transportation monitoring and building occupancy estimation [2, 6, 7]. With substantial growth in the computing capability of Internet of Things (IoT) devices in the past decade, there is an increased motivation for embedding sophisticated sensing applications in situ on these devices. Simultaneously, there is increased momentum in migrating computation from the cloud to the edge for reasons of privacy, cost, and latency. Among the growing ecosystem of diverse edge sensors that satisfy the above desiderata is the Micropower Pulsed Doppler Radar. Pictured in Figure 1(a) is a sub-20 mW version of the device that can be easily deployed in isolation or integrated with existing smart city infrastructure.

A power efficient mote-scale radar classifier system needs to continuously suppress background clutter and discriminate legitimate sources (humans, vehicles, animals, etc.) that displace through the scene, which are a relatively rare occurrence. However, existing shallow and deep solutions such as SVMs [6] or RNNs [3, 4] are either inefficient or compromising in sensing quality in terms of clutter suppression, source discrimination, or both. We demonstrate a Multi-Scale, Cascaded Recurrent Neural Network (MSC-RNN) architecture that jointly achieves high sensing accuracy and efficiency in radar classification. It uses RNNs at two different scales to address the two aforementioned components of radar sensing. It is significantly more efficient than heavier RNN alternatives and SVM solutions with feature handcrafting, and offers better sensing quality when compared with vanilla EMI-RNN solutions. Thus, MSC-RNN inherits the best of both worlds – the efficiency of EMI-RNNs for common case clutter rejection, and the accuracy of windowed RNNs for source separation.

In this demo, we would demonstrate MSC-RNN in action on a single microcontroller mote, and highlight its sensing quality and runtime power consumption.

*Both authors contributed equally to this research.
[1]MSC-RNN is a conditionally accepted BuildSys 2019 paper by the same authors, available at https://tinyurl.com/yxe3hxe9

BuildSys '19, November 13–14, 2019, New York, NY

ACM ISBN 978-1-xxxx-xxxx-x/xx/xx…$15.00
https://doi.org/xx.xxxx/xxxxxxx.xxxxxxx

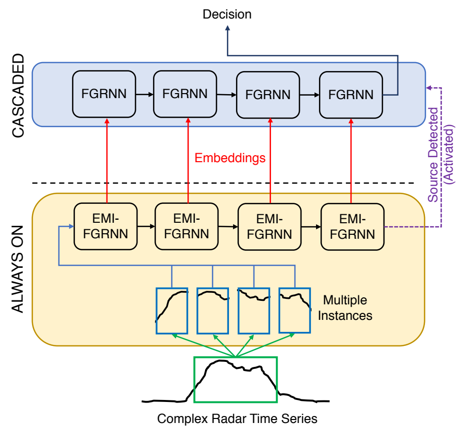

Figure 2: MSC-RNN architecture – the lower EMI-FastGRNN runs continuously, while the higher FastGRNN is invoked only for legitimate displacements

2 SYSTEM OVERVIEW

2.1 Components

MSC-RNN handles the two sub-problems of clutter rejection and source discrimination at different time scales of featurization (see Figure 2). Since clutter discrimination from displacements is easier, an early stopping, multi-instance (EMI-) RNN [3] is used. And since source classification is a harder problem that requires longer time scales, an upper-tier RNN [5] is invoked infrequently and on demand, only when the lower layer detects a legitimate displacement. Both tiers are jointly trained, and the upper-tier RNN, when active, operates on the instance-level embeddings computed by the lower tier. The two components of the architecture are outlined below:

2.1.1 FastGRNN. Recurrent Neural Networks (RNNs) have been state-of-the-art for analyzing sequences and time series. Traditional RNNs, even though theoretically all-powerful, fail to reach the best performance due to unstable training owing to the exploding and vanishing gradient problem (EVGP). Gated RNNs like LSTM circumvent the EVGP issue at the cost of significant compute overheads. FastRNN [5] provably stabilizes RNN training by helping to avoid EVGP by using only two additional scalars over the traditional RNN. FastGRNN is built over FastRNN and it extends the scalars of FastRNNs to vector gates while maximizing the information reuse.

2.1.2 EMI-RNN. Time-series signals when annotated are never precise and are noisy while also being coarsely labeled due to various factors like human errors and smaller time frames of activities themselves. EMI-RNN [3] tackles the problem of signal localization using a Multi-Instance Learning approach (MIL). EMI-RNN alternates between training the classifier and re-labeling the data based on the learnt classier until convergence. In the end, EMI-RNN produces precise signal signatures which are much smaller than the raw input, thus reducing the computation and memory overhead over traditional sequential techniques. EMI-RNN also ensures early detection of clutter or displacement sources, thereby removing the need for going through the entire signal before making a decision.

2.2 Implementation Details

The radar-mote device used in our demonstration has an ARM Cortex-M3 microcontroller with 96 KB of RAM and 4 MB of flash storage. It runs eMote [8], a low-jitter near real-time operating system with a small footprint. We take several measures to efficiently implement the multi-scale RNN to run at a low duty cycle on the device. These include low-rank representation of hidden states, integer quantization, and piecewise-linear approximations of non-linear functions. For example, $tanh(x)$ can be approximated as: $quantTanh(x) = \max(\min(x, 1), -1)$, and $sigmoid(x)$ as: $quantSigm(x) = \max(\min(\frac{x+1}{2}, 1), 0)$. All matrix and vector operations are implemented using the CMSIS-DSP library [1].

2.3 Performance Estimates

In our experiments using micropower radar datasets, MSC-RNN has a test accuracy of 97.2% with a 100% clutter recall, 92% human recall and 96% nonhuman recall over 2-second windows. It outperforms competitive algorithms at mote-scale and offers better resilience to source type imbalance as is common in radar data. Our demonstrated solution on the device consumes ~35 KB of working memory and, assuming sources are present <3% of the time, ~25 mW of power including the radar sensor.

3 DEMONSTRATION PLAN

The demo involves subjects parading in front of the radar occasionally, emulating a real-world deployment scenario that is dominated by environmental clutter. An example setup in an indoor amphitheater is illustrated in Figure 1(b). The demo aims to emphasize two aspects of the radar classifier:

(1) *Clutter Rejection*: The absence of detections when there is nothing in the scene,
(2) *Source Discrimination*: The correct classification of displacing sources.

The classification outputs from the radar would be displayed on a laptop in real time. An accompanying poster would explain the key design concepts of the classifier architecture and outline the training process. After showcasing scripted displacements of sources, we would invite the attendees to participate as sources as well. The displacement source types would be selected depending on the final demo location, for logistical reasons.

REFERENCES

[1] [n. d.]. CMSIS-DSP Software Library. http://www.keil.com/pack/doc/CMSIS/DSP/html/index.html.
[2] Juan P. Bello, Claudio Silva, et al. 2019. SONYC: A System for Monitoring, Analyzing, and Mitigating Urban Noise Pollution. *CACM* 62, 2 (Jan. 2019), 68–77.
[3] Don Dennis, Chirag Pabbaraju, et al. 2018. Multiple Instance Learning for Efficient Sequential Data Classification on Resource-constrained Devices. In *NIPS*. Curran Associates, Inc, 10975–10986.
[4] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.
[5] Aditya Kusupati, Manish Singh, et al. 2018. FastGRNN: A Fast, Accurate, Stable and Tiny Kilobyte Sized Gated Recurrent Neural Network. In *NIPS*. Curran Associates, Inc., 9030–9041.
[6] Dhrubojyoti Roy, Christopher Morse, et al. 2017. Cross-Environmentally Robust Intruder Discrimination in Radar Motes. In *MASS*. IEEE, 426–434.
[7] Oliver Shih and Anthony Rowe. 2015. Occupancy estimation using ultrasonic chirps. In *Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems*. ACM, 149–158.
[8] The Samraksh Company. [n. d.]. .NOW with eMote. https://goo.gl/C4CCv4.

SPECIAL REQUIREMENTS

If selected, we would request to be located in a relatively open space, such as a ballroom or a wide hallway. Our training has not incorporated obstacle rich indoor environments, and as a result, testing in crammed indoor areas with high multipath in the radar returns may affect the sensing quality. We could also be located outdoors, or in an area away from the presentation venue, in which case we would arrange to live stream the demo. If we are assigned an indoor area, care would be needed to avoid human movements within a 10 meter radius of the radar, including behind walls.