A Deep Learning Approach to Ultrasound Image Recovery

Dimitris Perdios, Adrien Besson, Marcel Arditi, and Jean-Philippe Thiran

Signal Processing Laboratory (LTS5)
École Polytechnique Fédérale de Lausanne (EPFL)

2017 IEEE International Ultrasonics Symposium (IUS)
Washington, D.C., USA
September 9, 2017

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Outline

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Introduction

Context

- Ultrasound (US) system design is pushing towards portability
- ADCs are incorporated in the probe head ➔ digital interface (e.g. wireless)
  - ⚠ **data transfer issues (esp. for ultrafast US imaging)**
  - ☺ **can “easily” add compression capability in the probe head**

Objective

- Recovering US signals from undersampled measurements
- In real time (if possible ☺) ➔ **fast compression and recovery**

Great candidate ➔ compressed sensing (CS)

- Provides a way to exactly recover a signal from undersampled measurements, under very specific assumptions (sparsity and RIP)[1]
- Main drawbacks:
  - Sparsity of US signals is very hard to obtain (esp. inside speckle regions)
  - Use of convex optimization algorithms (hundreds of iterations) ➔ **slow**

[1] http://statweb.stanford.edu/~markad/publications/ddek-chapter1-2011.pdf

Stacked Denoising Autoencoders (SDA)

- A DNN architecture successfully applied to **structured signal recovery**[2]
- Compression is considered as the first layer of the proposed architecture
- Recovery is performed by the hidden and output layers
- Two measurement cases are explored:
  1. **SDA–CNL**: Linear measurement case where the **compression is not learned**
  2. **SDA–CL**: Non-linear measurement case where the **compression is learned**

Imaging pipeline

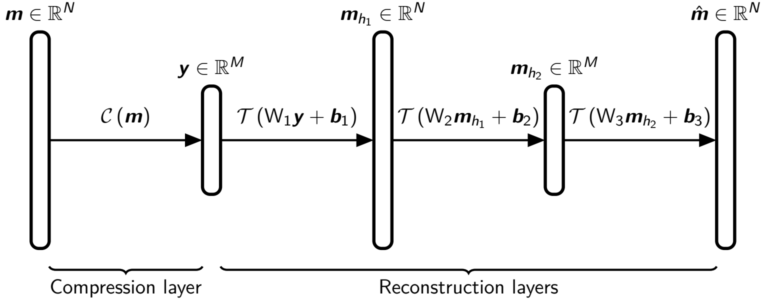
- Once trained, the first layer is used to compress each of the element-raw-data signals **independently**, the remaining layers are used for the recovery
- Both operations can be performed **in parallel** for all channels ➔ **fast** ☺
- The US image is then retrieved using any image reconstruction algorithm

[2] https://arxiv.org/abs/1508.04065

Deep learning for ultrasound image recovery

Proposed architectures

SDA with 4 fully-connected layers



Reconstruction layers

$\mathrm{W}_1 \in \mathbb{R}^{N\times M}$
$\boldsymbol{b}_1 \in \mathbb{R}^N$
$\mathrm{W}_2 \in \mathbb{R}^{M\times N}$
$\boldsymbol{b}_2 \in \mathbb{R}^M$
$\mathrm{W}_3 \in \mathbb{R}^{N\times M}$
$\boldsymbol{b}_3 \in \mathbb{R}^N$
$\mathcal{T}(\,.\,) = \tanh(\,.\,)$

Compression layer

- SDA–CNL: $\mathcal{C}(\boldsymbol{m}) = \Phi\boldsymbol{m}$, where $\Phi \in \mathbb{R}^{M\times N}$ is a random Gaussian matrix, **not learned** during training
- SDA–CL: $\mathcal{C}(\boldsymbol{m}) = \mathcal{T}(\mathrm{W}_{in}\boldsymbol{m} + \boldsymbol{b}_{in})$, where $\mathcal{T}(\,.\,) = \tanh(\,.\,)$, $\mathrm{W}_{in} \in \mathbb{R}^{M\times N}$ and $\boldsymbol{b}_{in} \in \mathbb{R}^M$, **learned** during training

Acquisition configuration

- Plane-wave imaging challenge in medical ultrasound (PICMUS)[3]

Parameter	**L11-4v**
Element number	128
Pitch	300 µm
Center frequency	**5.133 MHz**
Bandwidth	**67 %**
Element width	0.27 mm
Transmit frequency	5.208 MHz
Excitation	2.5 cycles
Sampling frequency	**20.832 MHz**

- Sampling frequency **extremely close to the Nyquist frequency** of US signals
- The sample number N is fixed to 1024 to fit typical DNN sizes

[3] https://www.creatis.insa-lyon.fr/EvaluationPlatform/picmus/index.html

Training set

- Simulated using the open-source k-Wave toolbox[4]
- $c_0 = 1540\,\mathrm{m\,s^{-1}}$, $Z_0 = 1.63 \times 10^6\,\mathrm{kg\,m^{-2}\,s^{-1}}$, $\alpha = 0.5\,\mathrm{dB\,MHz^{-1}\,cm^{-1}}$
- Simulation accounts for the element directivity
- Insonified medium is randomly generated from 3 main components:
  1. A fully diffusive background (echogenicity reference)
  2. 1 to 3 circular inclusions (random position) of variable radius and echogenicity:
     - Radius: drawn between 5 and 50 wavelengths
     - Echogenicity: anechoic (80 %) || −6 dB to 6 dB (15 %) || 10 dB to 20 dB (5 %)
  3. 0 to 5 point reflectors (random position)
- Transmit scheme: single PW insonification
- Each simulated acquisition is composed of 128 raw-data
- 20 000 simulated acquisitions ➔ **2.5 M element-raw-data signals**

[4] http://www.k-wave.org

Training set-up

- Implementation[5]: TensorFlow
- TGC is applied to raw-data
- Data normalized between -1 and 1 to fit the range of the non-linearity
- The training is performed on a NVIDIA GeForce GTX 1080 Ti
- Learning rate: 0.001
- Epoch number: 20 epochs
- Mini-batch learning with a batch size of 4096

Training set-up (cont.)

- Initialization:
  - Weights ➔ Xavier
  - Biases ➔ zero
- Optimizer: Adam
- Loss function: ℓ_2-loss
- Undersampling ratio M/N ranging from 0.05 to 0.5

[5] https://github.com/dperdios/us-rawdata-sda

Three approaches are compared

1. SDA–CNL: comp. Gaussian matrix, rec. 3 layers
2. SDA–CL: comp. learned, rec. 3 layers
3. A CS reconstruction based on a sparsity prior in a convolutional dictionary made of shifted pulses: comp. Gaussian matrix, rec. PDFB (1000 iterations)

Test set ➔ PICMUS datasets

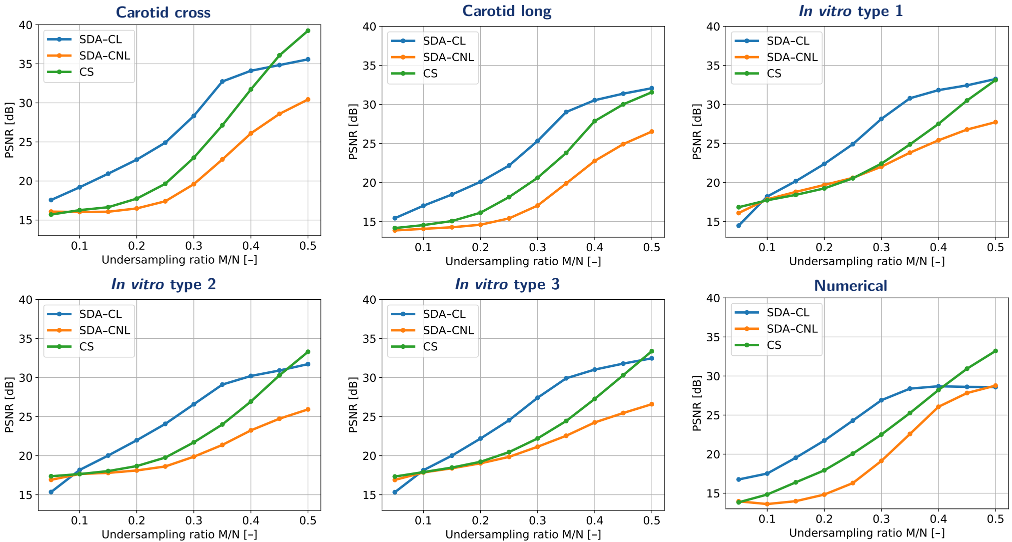
- 1 numerical image (PICMUS 2017)
- 3 *in vitro* images (PICMUS 2017)
- 2 *in vivo* images (PICMUS 2016)

Performance evaluation

- DAS (spline + elem. directivity) is performed on recovered signals ➔ RF image
- Envelope extraction ➔ normalization ➔ log-compression ➔ B-mode image
- PSNR on B-mode images (40 dB for *in vivo*, 60 dB for numerical and *in vitro*)

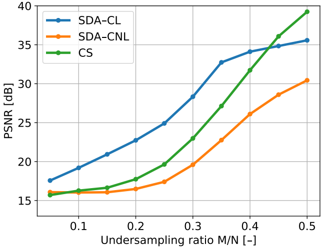
Results

Performance evaluation on the PICMUS dataset



Results

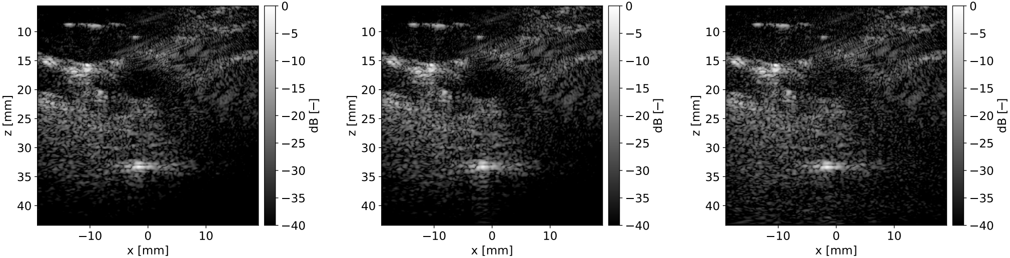
Visual assessment – Carotid cross



Reference (100 %)

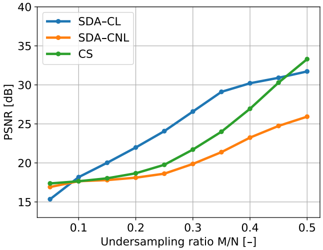
SDA–CL (30 %)

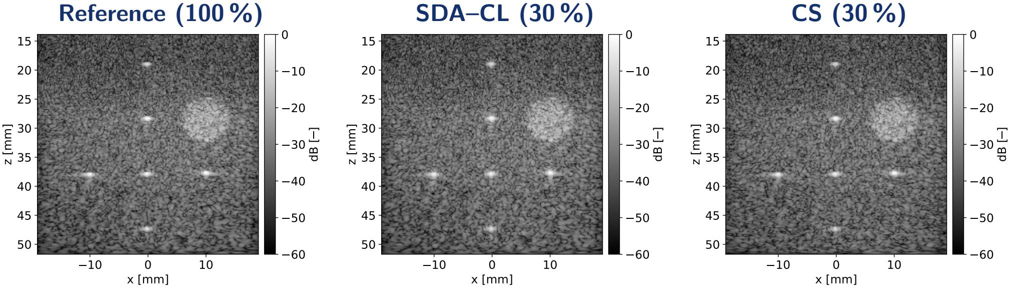
CS (30 %)



ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Results
Visual assessment – *In vitro* type 2





Conclusion

Compression capability

- Reference data extremely close to the Nyquist frequency
- Good recovery with 30 % of the data

Reconstruction complexity

- CS: $\geq 2 \times 1000 \times \mathcal{O}(MN)$
- SDA–CL: $3 \times \mathcal{O}(MN)$
- Almost **1000 times faster than CS**

Quality

- SDA–CL outperforms CS at low undersampling ratios
- Quite robust to variable image regions (speckle, anechoic, etc.)

Current drawbacks and future work

- Low generalizability: trained for 1024 time samples
- Seems to suffer from oscillating artifacts around hyperechoic regions
- Side information across the transducer elements is not exploited

THANK YOU FOR YOUR ATTENTION!

Dimitris Perdios

dimitris.perdios@epfl.ch

https://github.com/dperdios

Signal Processing Laboratory (LTS5)

https://lts5www.epfl.ch

École Polytechnique Fédérale de Lausanne