Deep Multitask Ultrasound Beamforming

Elay Dahan

M.Sc. Thesis Seminar

Supervision: Professor Israel Cohen

Outline

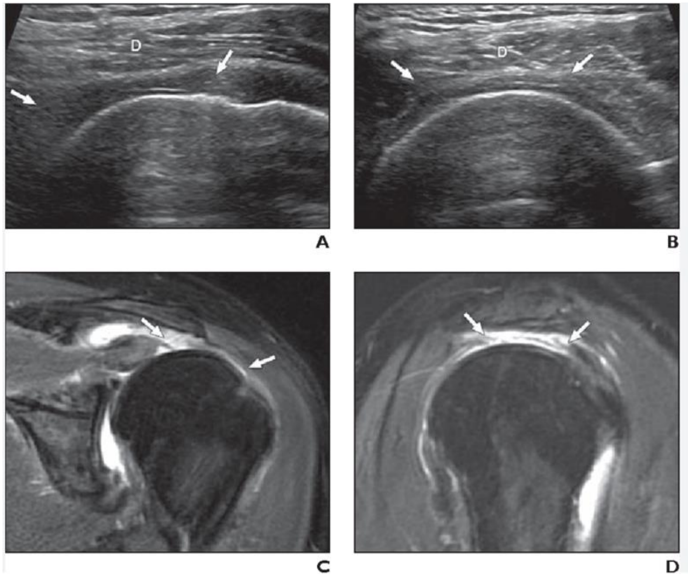
- Introduction
- Algorithms for ultrasound beamforming
- Main Contributions
- Conclusions

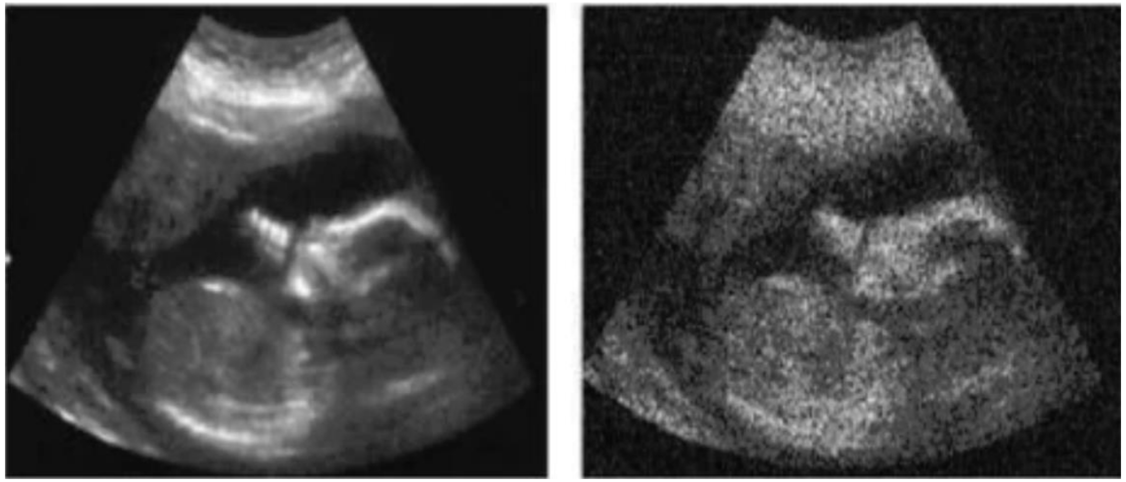
Ultrasound Imaging

- Noninvasive medical imaging device
- Utilizes ultrasonic waves to create and echogenicity map of the insonified area

Challenges

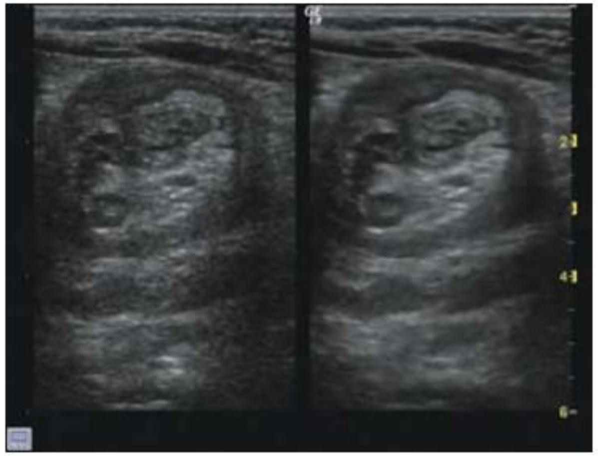
1. Diagnosis of ultrasound is challenging.
2. Ultrasound image are nosier and with lower resolution.



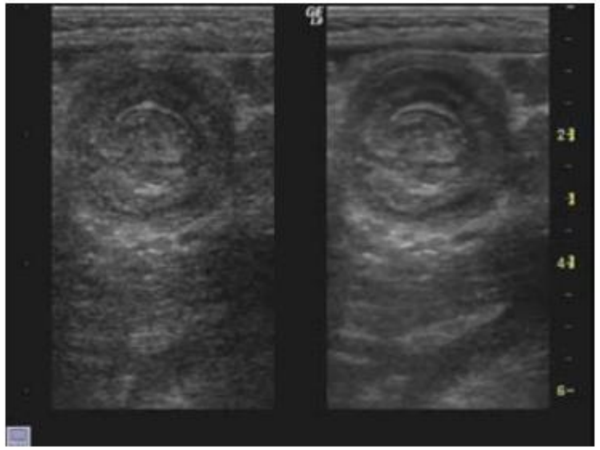


Post Processing of Ultrasound Image

- Reduce noise
- Must be lightweight – high frame rate is required.
- Will reduce framerate



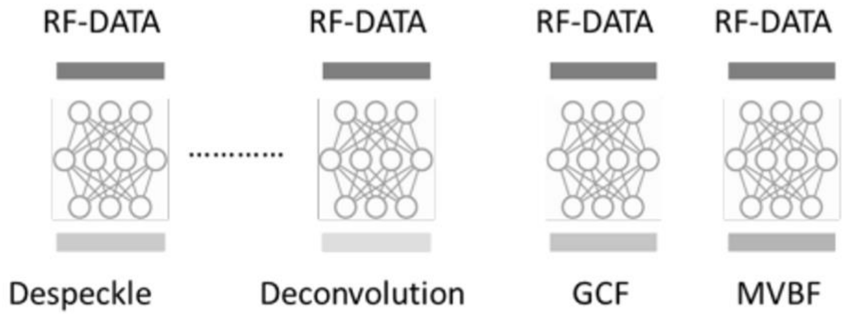
Noisy Denoised



Noisy Denoised

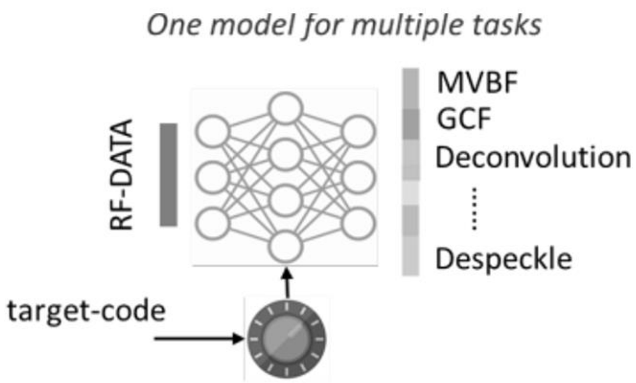
Limitations of Classical Ultrasound Reconstruction Algorithms

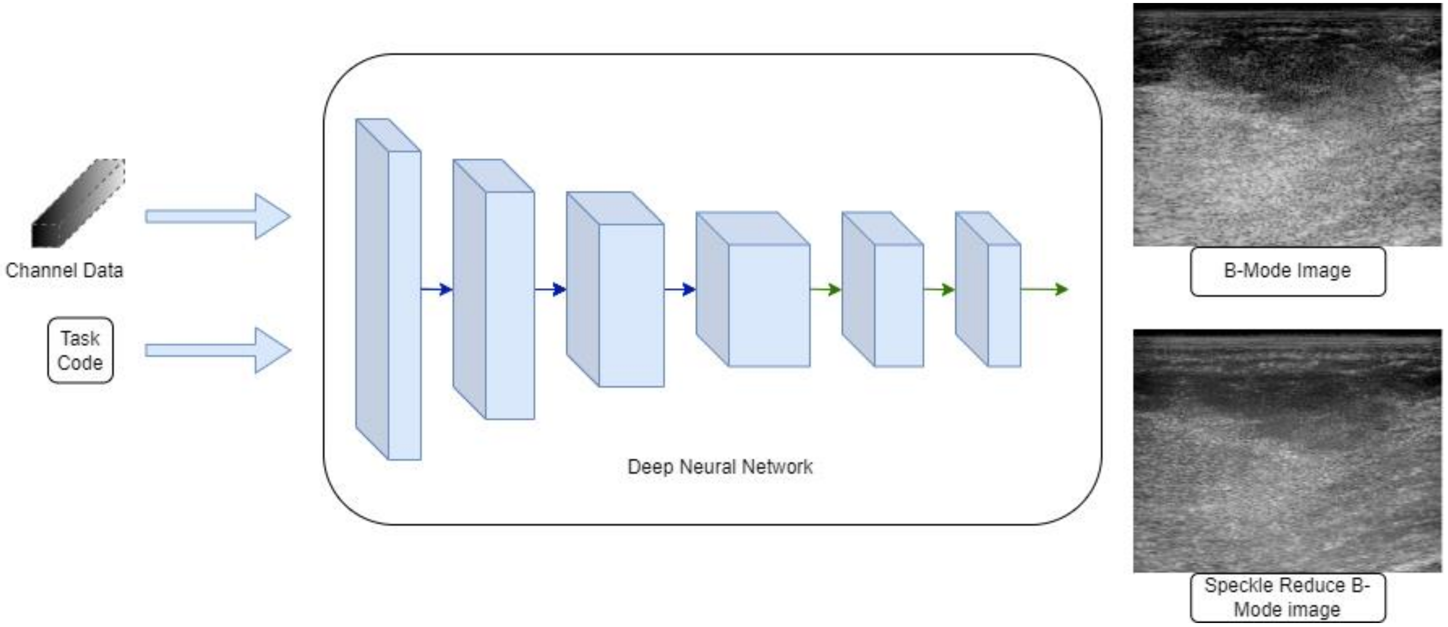
- On high end machines there are hundreds of different configuration/optimization to the ultrasound image.
- Ideally, a single model would encompass all desired optimizations.



Multitask Image Formation

- ✓ Will not affect frame rate
- ✓ Allow real time selection between image quality and frame rate tradeoff

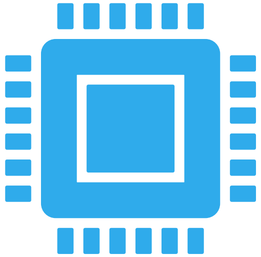




Main Research Contributions

Enhancing multitask beamformers robustness while maintaining efficiency.

Our approach outperformed classical algorithm on speckle reduction and reconstruction from subsample signal

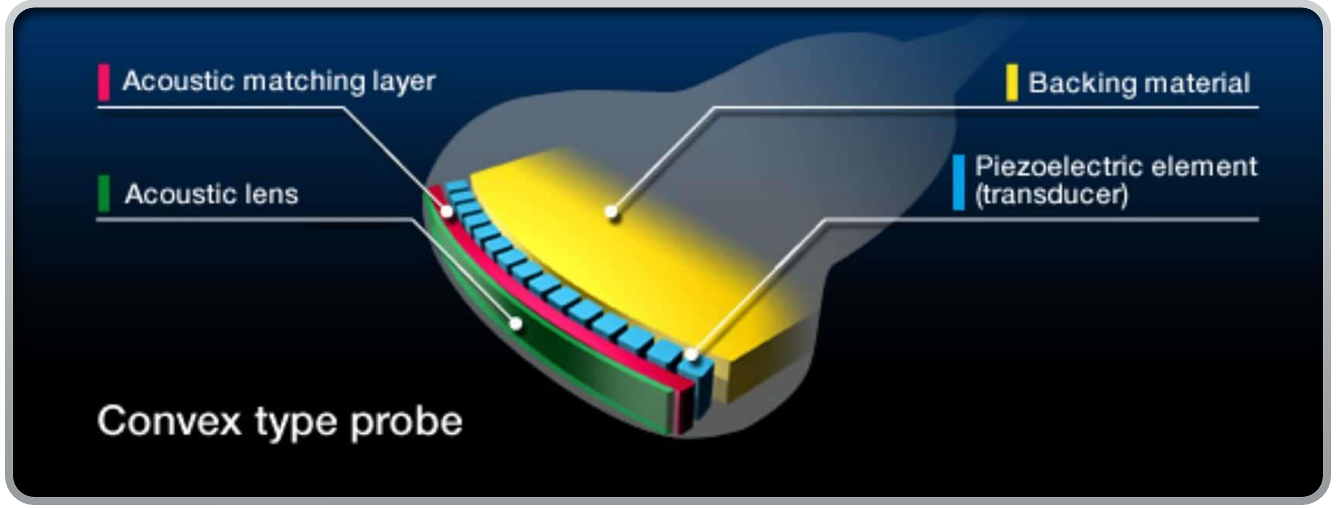


Enhancing computational and perceptual performance with a generalized linear weight transformation scheme.

We proposed an improved more general framework for multitask learning. This method produced results that are on par with specifically trained model.

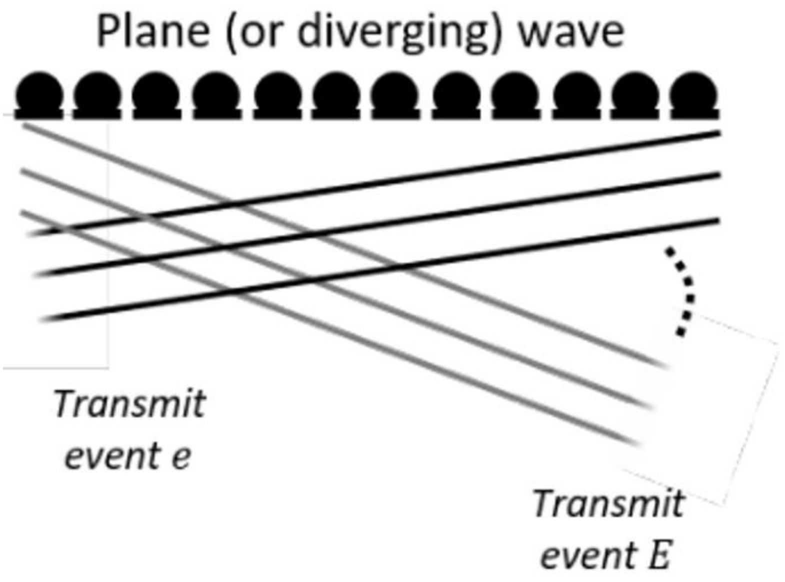
Ultrasound Probes

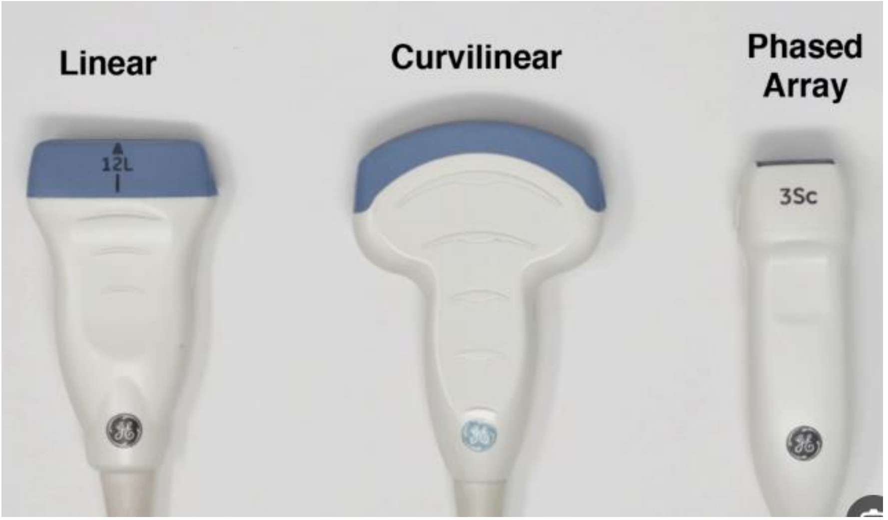
- Contains an array of piezo electric elements
- The geometry of the probe is dependent on the transmission scheme.
- Each transducer is suitable for specific application



Plane Wave Ultrasound

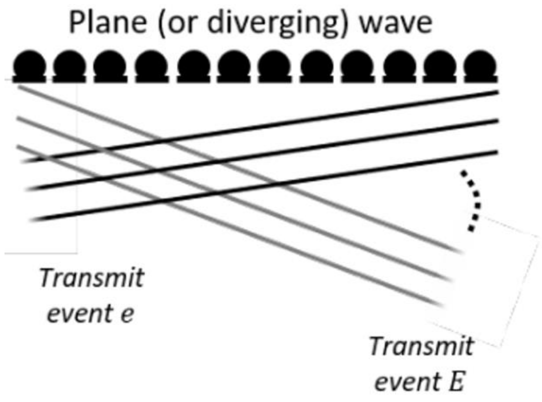
- In each transmit event all the elements are utilized.
- Each transmit event produces a plane wave at specific angle.
- The final image is aggregation of all the angles





Ultrasound Image Formation

- Let $x \in \mathrm{R}^{E \times C \times N}$ denote the received echo.
- The first step in the ultrasound image formation pipeline is beamforming.
- The goal is to aggregate the sample from E, C dimensions to generate unified signal.



Delay and Sum (and variations)

- The simplest beamforming algorithm, first we perform time-of-flight correction to the received echo, then summation is performed.

  Output signal is then given by:

  $$Y = \sum_E \sum_C X_{tof}$$

  $$Y_{weighted} = \sum_E \sum_C w_{ij} * X_{tof}$$

- Advantages: Fast, Linear time complexity
- Disadvantages: Low resolution & contrast.

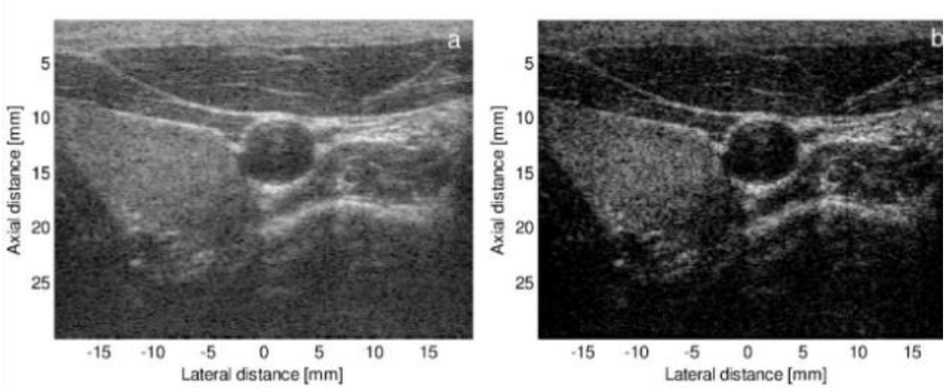
Adaptive beamforming - MVDR

MVDR – minimum variance distortion less response, the weights are calculated adaptively.

$$\mathrm{argmin}_{\boldsymbol{W}}\ \boldsymbol{W}^H \boldsymbol{R}_{x_e[r_x,r_y]} \boldsymbol{W}$$

$$\text{s.t.} \quad \boldsymbol{W}^H \boldsymbol{1} = 1$$

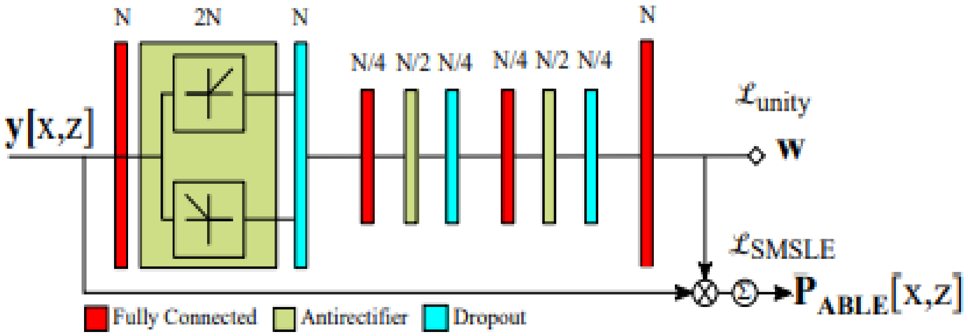
$$\boldsymbol{W}[r_x, r_y]_{\boldsymbol{mvdr}} = \frac{\boldsymbol{R}^{-1}_{x_e[r_x,r_y]}\,\boldsymbol{d}}{\boldsymbol{d}^H \boldsymbol{R}^{-1}_{x_e[r_x,r_y]}\,\boldsymbol{d}},$$

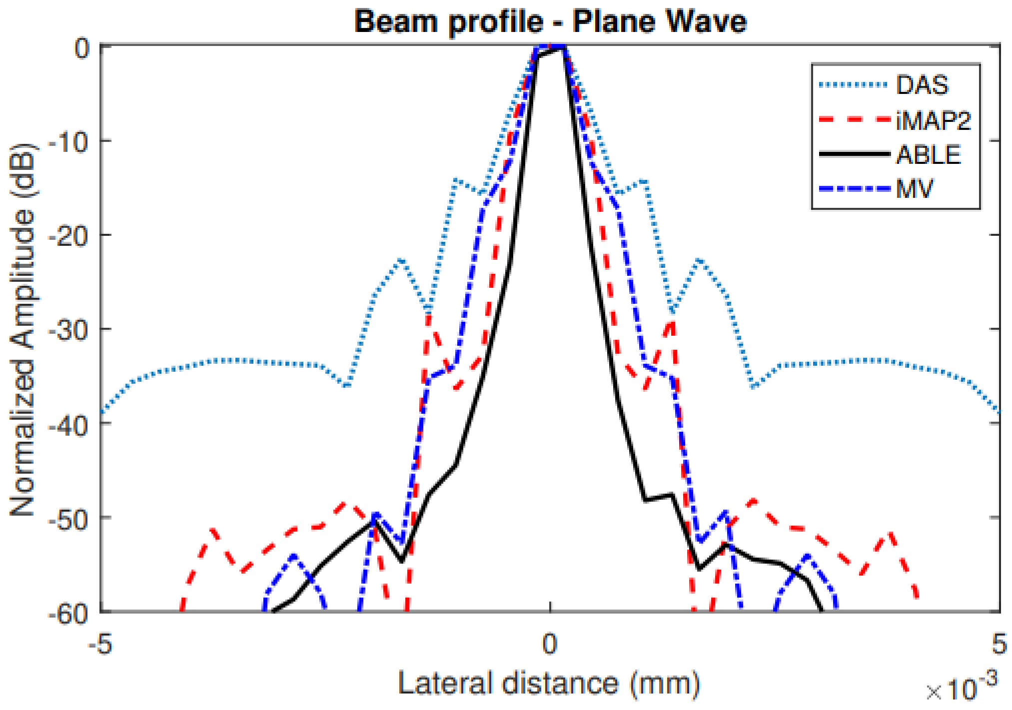


Where $\boldsymbol{d} = \boldsymbol{1}$, Since the signal is time-of-flight corrected

Adaptive beamforming – Deep Learning

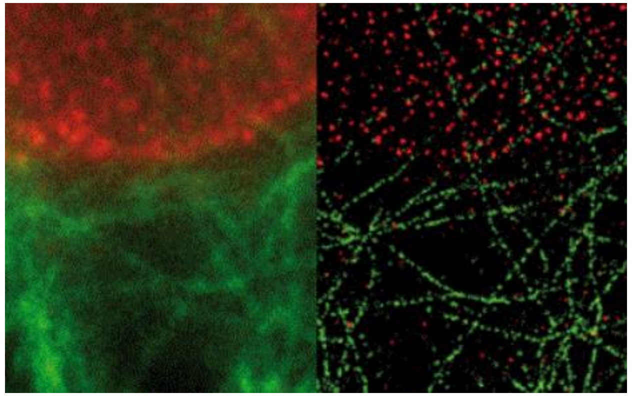
Main Idea: replicate the output of MV beamformer with lightweight neural network





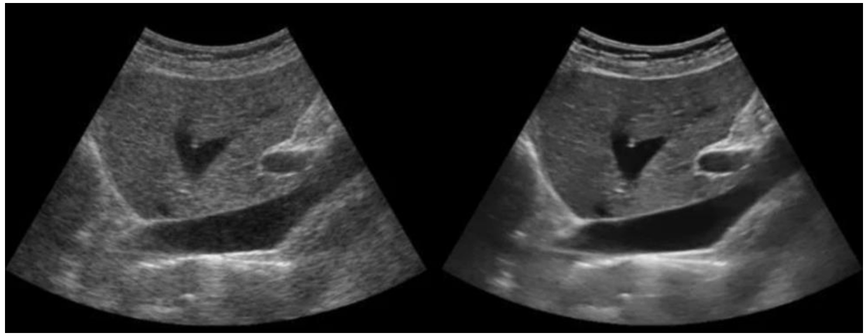
Other image targets

- Devices integrate various image enhancement and denoising algorithms for accurate diagnosis.
- Speckle reduction, and deconvolution are common example.
- Multiple beamformers allow real-time trade-offs between computation and perception.



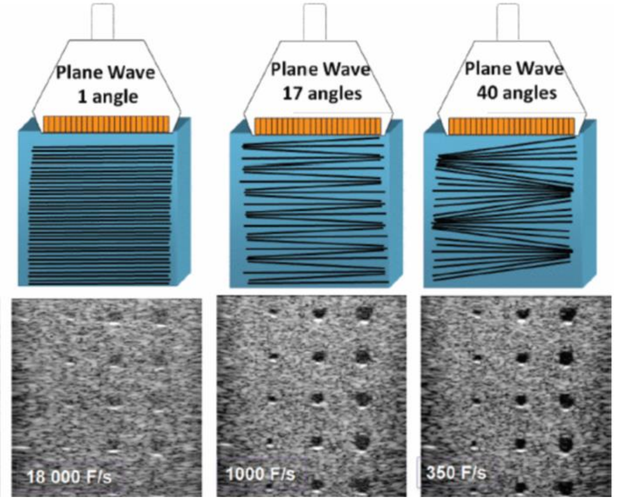
Speckle Noise

- **Speckle Noise**: Grainy interference in ultrasound images that obscures details. Reducing speckle enhances clarity for better diagnoses.
- **Filtering**: Median and Gaussian filters reduce noise while preserving features. Deep learning methods, like CNNs, adaptively learn noise patterns for improved reduction.



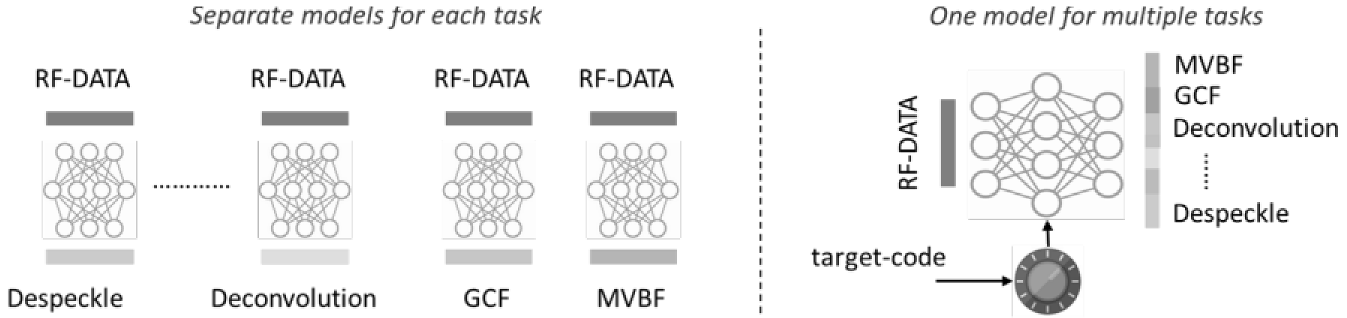
Reconstruction from subsampled signal

- In plane wave transmission, the final image is constructed from data acquired at multiple angles.
- Sequential data acquisition at different angles reduces frame rate.
- We benchmarked our approach on single-angle reconstruction.



Approach

- With classical algorithms the only solution is to implement all the algorithms on the device.
- When implementing the image formation algorithm (beamformer) with deep learning multitask learning approach's can be utilized.



Naive Solution

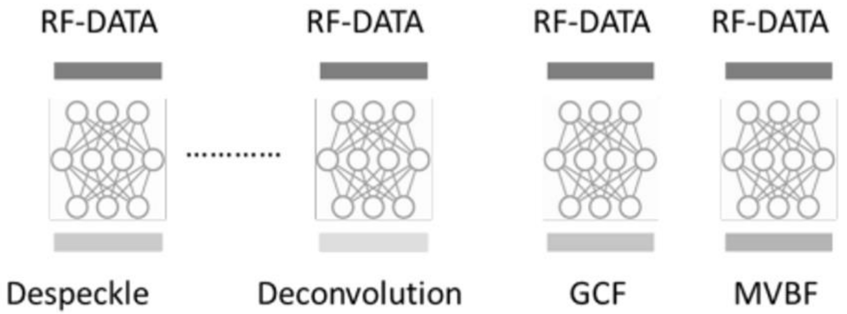
- Major constraint of ultrasound are computational performance

(The models must be lightweight)

- The naïve approach will be implementing a specific model for each task.

Advantages: optimal perceptual performance.

Disadvantages: not scalable, high memory and computational overhead, no knowledge sharing between tasks.



Shared Encoder Multi Decoder models

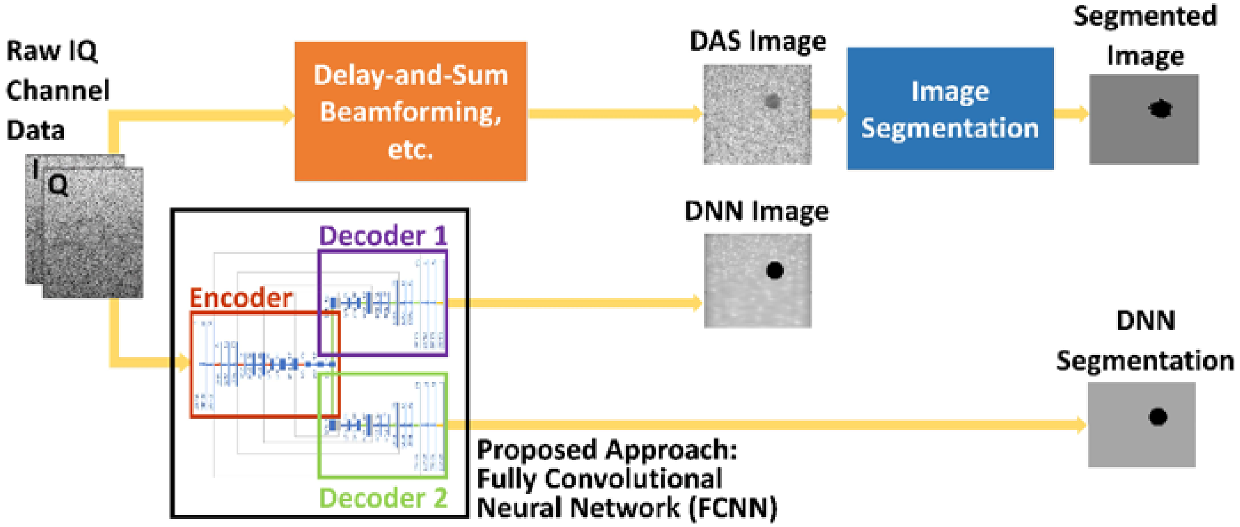
- **Advantages**: knowledge sharing,

simultaneous output

- **Disadvantages**: lack of scalability,

computational overhead grows linear

with the number of tasks.

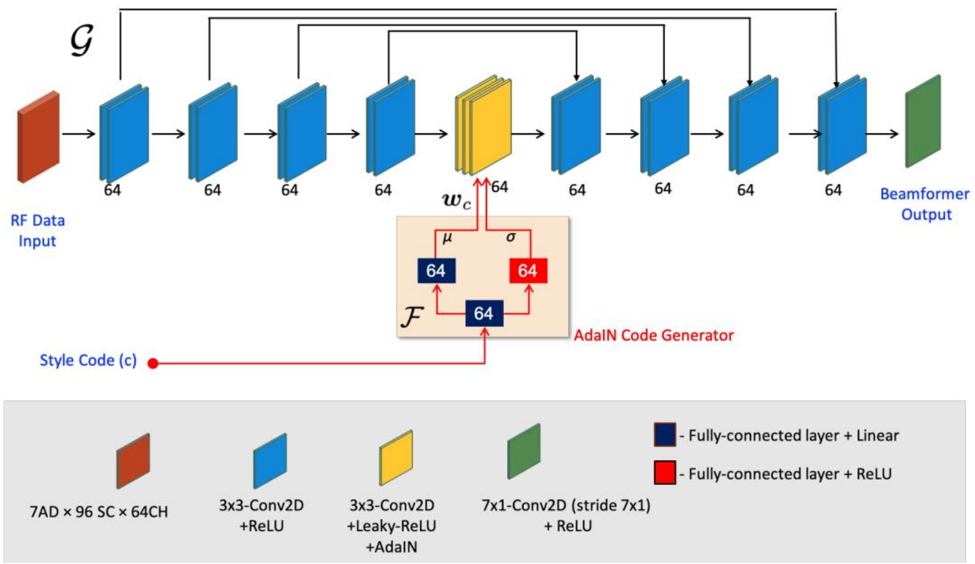


Instance normalization task shift

- **Advantages**: knowledge sharing,

Scalable, minimal computational overhead

- **Disadvantages:** task specific output, adaptation is only applied to the latent vector.



Gaps In Current Research

- Shared encoder-multi decoder approach[1] is robust since for each task a specific decoder is trained, however that effect scalability.
- Latent representation normalization-based methods are efficient in terms of computational overhead but are limited in robustness.

Weight Normalization For Multitask Learning

- Assumption: Trained model can be adapted with minor filter adaptation.
- w_i - learned convolutional filters for layer i, Then

  Adapt model to new tasks:

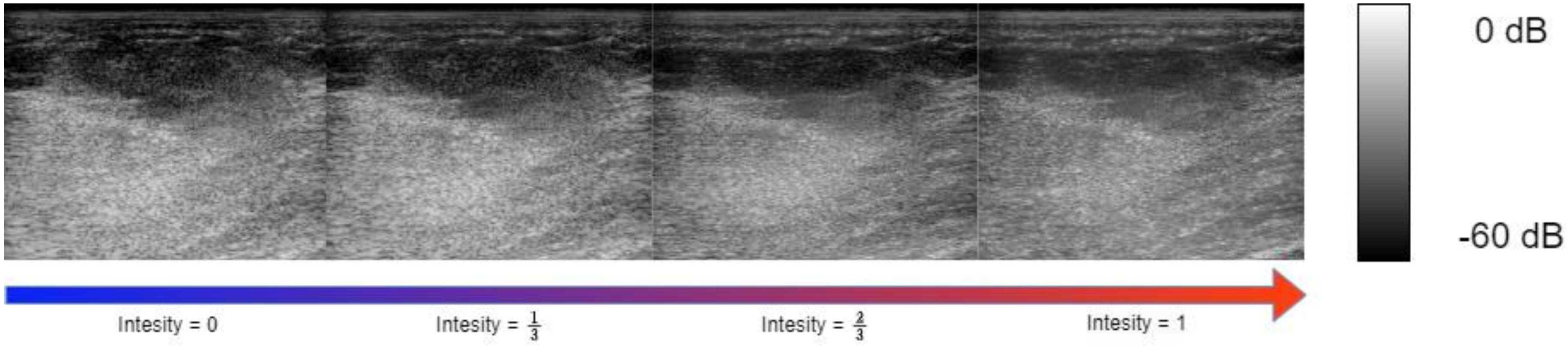
$$w_i' = \frac{w_i}{s_{i,k}} + b_{i,k},$$

Where $s_{i,k}, b_{i,k} \in \mathbb{R}^n$ are learned scale and bias vectors for layer i and task j respectively.
n is the number of convolution filter in layer i.

Controlling Task Intensity

- For post-processing like tasks sometimes its helpful to reduce the filter effect, thus we propose the following modification to the weight normalization scheme:

$$w_i' = \frac{w_i}{s_{i,k}^{\alpha}} + \alpha b_{i,k}.$$



Model Architecture

Table 1. Fully convolutional beamformer architecture.

Layer Type	Input Shape	Output Shape
ConvBlock	$2 \times Nc \times 256$	$64 \times Nc \times 256$
Max pooling	$64 \times Nc \times \times 256$	$64 \times \frac{Nc}{2} \times 256$
ConvBlock	$64 \times \frac{Nc}{2} \times 256$	$128 \times \frac{Nc}{2} \times 256$
max pooling	$128 \times \frac{Nc}{2} \times 256$	$128 \times \frac{Nc}{4} \times 256$
ConvBlock	$128 \times \frac{Nc}{4} \times 256$	$256 \times \frac{Nc}{4} \times 256$
max pooling	$256 \times \frac{Nc}{4} \times 256$	$256 \times \frac{Nc}{8} \times 256$
ConvBlock	$256 \times \frac{Nc}{8} \times 256$	$128 \times \frac{Nc}{8} \times 256$
ConvBlock	$128 \times \frac{Nc}{8} \times 256$	$64 \times \frac{Nc}{8} \times 256$
ConvBlock	$64 \times \frac{Nc}{8} \times 256$	$2 \times \frac{Nc}{8} \times 256$

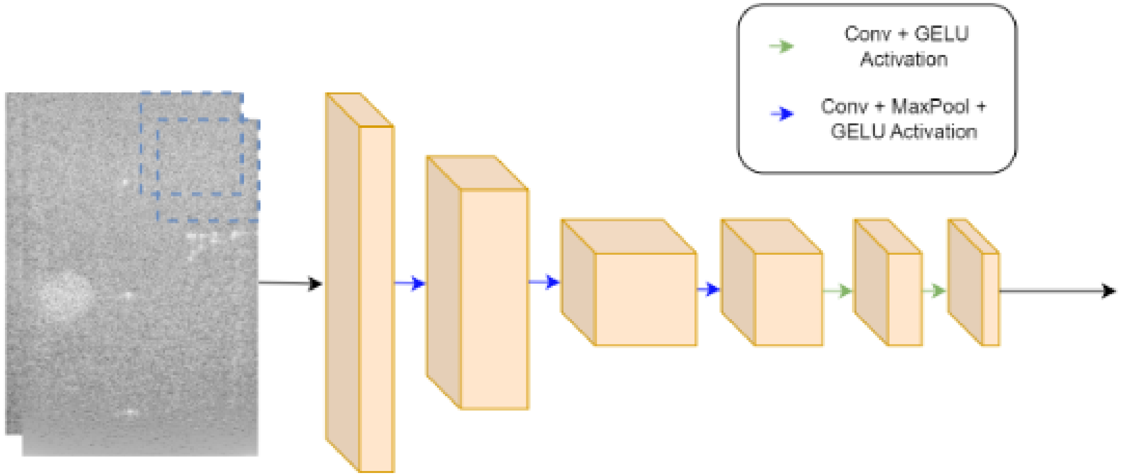


Table and Figure adapted from [5]

5. Dahan, Elay, and Israel Cohen. "Deep-Learning-Based Multitask Ultrasound Beamforming." *Information* 14.10 (2023): 582.

Challenge on ultrasound beamforming with deep learning (CUBDL)

- Evaluation was performed under the CUBDL framework.
- Dataset contains channel data of planewave ultrasound
- The data includes both in-vivo and simulated acquisitions.
- The task was to reconstruct multi-angle images from single-angle acquisitions.

Image Formation: Global Evaluation Metrics

$$\ell_1 = \frac{1}{N}\sum_{n=1}^{N} |x_n - y_n|,$$

$$\ell_2 = \sqrt{\frac{1}{N}\sum_{n=1}^{N} (x_n - y_n)^2},$$

$$\mathrm{PSNR} = 20\log_{10}\frac{\mathrm{DynamicRange}}{\frac{1}{N}\sum_{n=1}^{N}(x_n - y_n)^2},$$

$$\rho = \frac{\sum_n (x_n - \mu_x)(y_n - \mu_y)}{\sqrt{\sum_n (x_n - \mu_x)^2}\sqrt{\sum_n (y_n - \mu_y)^2}},$$

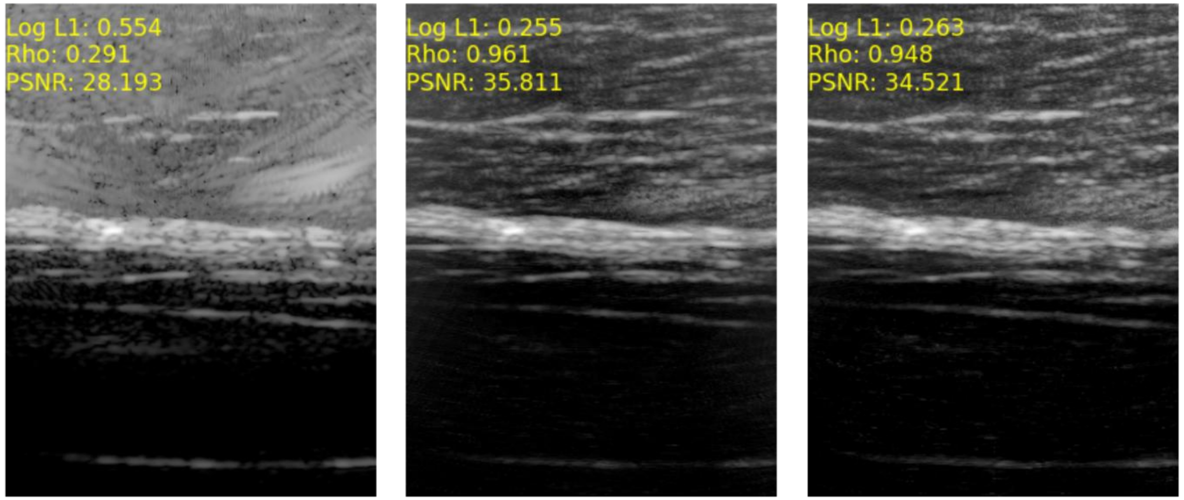


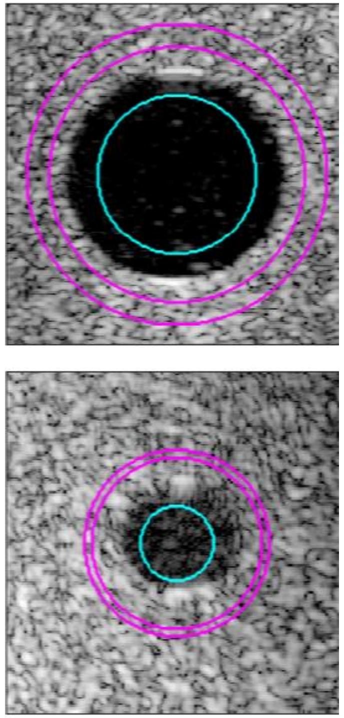
Image Formation: Local Evaluation Metrics

$$\text{Contrast} = 20\log_{10}\frac{\mu_1}{\mu_2},$$

$$\text{CNR} = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}},$$

$$\text{gCNR} = 1 - \sum_x \min\{f_1(x), f_2(x)\},$$

$$\text{SNR} = \frac{\mu_0}{\sigma_0},$$



Experimental Setup

- We first trained the proposed model for the base task in CUBDL of multi angle reconstruction from single angle acquisition.
- Then we adapt the trained model to 2 different tasks:

1. Reconstruction from subsampled signal – we adapt the model to reconstruct multi angle plane wave image from single angle acquisition that is subsample in the channel dimension.
2. Speckle noise reduction – The model is adapted to estimate the output of multiangle DAS reconstruction with speckle denoising postprocessing from single angle acquisition.

Training Procedure

- Loss function: A combination of $L1$ at logarithmic scale and SSIM.

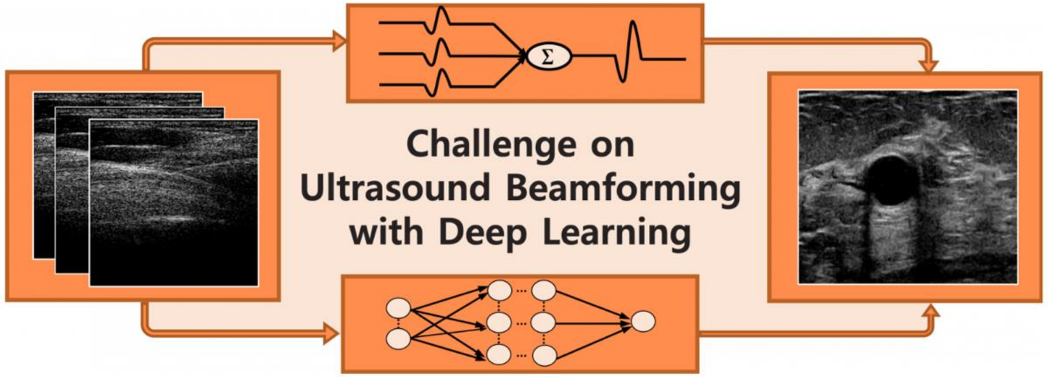
$$L = \frac{1}{N}\sum_{i}^{N} |\log_{10}(\hat{y}_i) - \log_{10}(y_i)| + \lambda * \mathrm{SSIM}(\hat{y}_i, y_i)$$

- We have used AdamW[1] optimizer for 50 epoch with learning rate of 0.003.

1. Loshchilov, Ilya, and Frank Hutter. "Decoupled weight decay regularization." *arXiv preprint arXiv:1711.05101* (2017).

Results: Image Reconstruction

- We first benchmark our model the base task of multiangle reconstruction from single angle acquisition. We compared our model to CUBDL winners.



Results: Image Reconstruction

CUBDL Test Set: Global Quality Metrics

Model	ℓ_1	$\log \ell_1$	ℓ_2	$\log \ell_2$	PSNR	ρ
Goudarzi [14]	0.03	0.42	0.05	0.56	29.10	0.91
Single Plane	0.03	0.39	0.042	0.53	30.36	0.93
Ours	**0.029**	**0.36**	**0.0408**	**0.49**	**30.4**	**0.93**
Ours—×2 Sub Sampled	0.035	0.41	0.0504	0.55	28.52	0.9
Single Plane—×2 Sub Sampled	0.0433	0.525	0.059	0.7	27.43	0.86

CUBDL Test Set: Local Quality Metrics

Model	Contrast dB	Speckle CNR	gCNR	log Speckle SNR	x Fwhm	z Fwhm
Goudarzi [14]	−13.77	1.345	0.814	1.827	**0.0003**	0.0004
Single Plane	−10.5	1.05	0.674	1.85	0.0004	0.0004
Ours	**−15.84**	**1.632**	**0.88**	**2.299**	0.0005	**0.0003**
Ours—×2 Sub Sampled	−5.07	1.083	0.62	1.94	0.0006	0.0003
Single Plane—×2 Sub Sampled	−4.95	0.542	0.394	1.842	0.0004	0.0004
Ground-Truth—Multi-Plane	−24.70	1.53	0.946	1.806	0.0004	0.0003

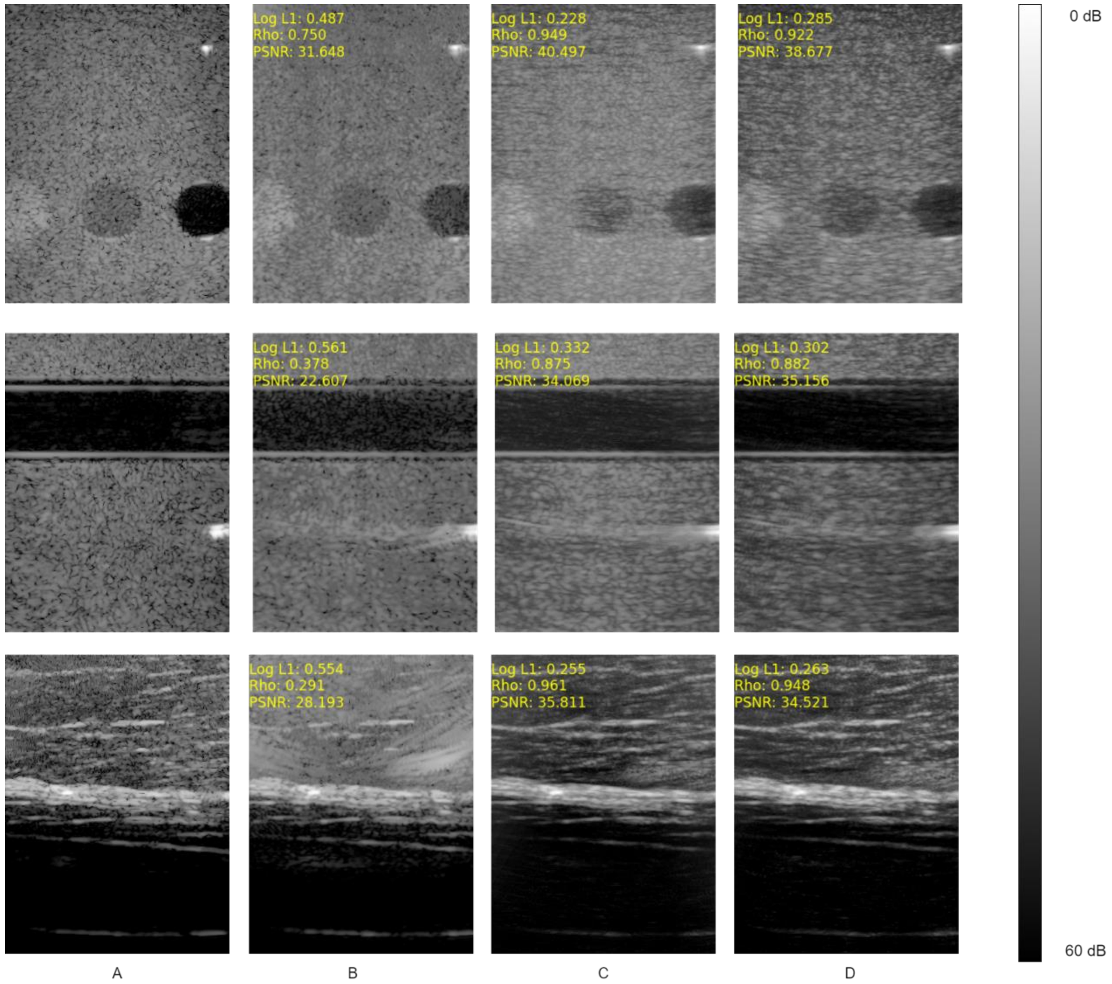


Figure 1. Comparison of 4 models on cubdl test set. Adapted from [5]

5. Dahan, Elay, and Israel Cohen. "Deep-Learning-Based Multitask Ultrasound Beamforming." *Information* 14.10 (2023): 582.

Results: Speckle Reduction

CUBDL Test Set: Global Quality Metrics - Speckle Reduction

Model	$\log \ell_1$	$\log \ell_2$	PSNR	ρ
Single Plane + speckle reduction	0.57	1.07	24.8	0.593
Ours-speckle reduction	**0.3**	**0.41**	**29.01**	**0.92**

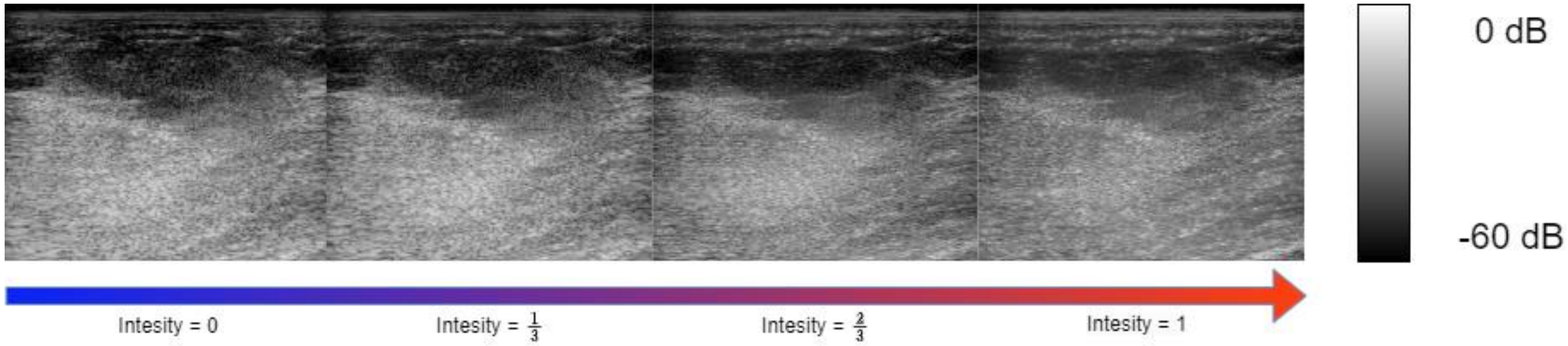
CUBDL Test Set: Local Quality Metrics - Speckle Reduction

Model	Contrast dB	Speckle CNR	gCNR	Speckle SNR
Single-Angle + speckle reduction	−10.5	1.05	0.674	1.85
Ours: speckle reduction	−10.07	**1.673**	**0.885**	**3.17**
Multi-Angle + speckle reduction (ground-truth)	−25.6	2.423	0.98	2.713

Speckle Reduction: Controlling Task Intensity

- With the following modified weight normalization formula, one can control the effect of the specific task

$$w_i' = \frac{w_i}{s_{i,k}^{\alpha}} + \alpha b_{i,k}$$



Improved Efficiency and Robustness: Linear Weight Adaptation Multitask Learning

Idea: With Apply more generic linear transformation to the learned weights.

Formulation: Let $W_i \in \mathbb{R}^{C_1 \times C_2 \times K_1 \times K_2}$ be the learned convolutional filters of i's layer.

C_1, C_2 - feature maps input and output dimension of the.

K_1, K_2 - Convolution kernel dimensions

Learn a transformation $T_{ij} \in \mathbb{R}^{K_1 K_2 \times K_1 K_2}$, such that the transformed filters for task j are:

$$W_{ij} = W_i \times T_{ij}$$

Improved Robustness: Visualization

- With weight normalization: $W' = \frac{W}{s} + b$, limited to affine transformation
- Affine transformation is a special case of $\mathrm{W}' = \mathrm{W} \times T$. For the 3x3 case:

$$\mathrm{T} = \begin{pmatrix} \frac{1}{s} & 0 & 0 \\ 0 & \frac{1}{s} & 0 \\ 0 & 0 & \frac{1}{s} \end{pmatrix} + \mathbf{b}$$

Experimental Setup – Speckle Reduction

- To evaluate to goodness of the proposed method, we have compared our method to:

I. Classical algorithms for speckle reduction: DAS + speckle denoising

II. Weight normalization for multitask learning.

III. Neural network trained specifically for the task of speckle denoising.

Results

CUBDL Test Set: Global Quality Metrics - Speckle Reduction

Model	$\ell_1 \downarrow$	$\log \ell_1 \downarrow$	$\ell_2 \downarrow$	$\log \ell_2 \downarrow$	PSNR $\uparrow$	$\rho \uparrow$
Base Model	**0.03**	**0.275**	**0.043**	**0.37**	**30.4**	**0.94**
Multitask model	**0.03**	0.3	0.048	0.41	29.5	0.93
Dahan et. al. [30]	0.032	0.3	0.048	0.41	29.01	0.92
Single Angle DAS + Speckle Denoising	0.043	0.57	0.1	1.07	24.78	0.593

CUBDL Test Set: Local Quality Metrics - Speckle Reduction

Model	Contrast $\downarrow$	CNR $\downarrow$	gCNR $\uparrow$	Speckle SNR $\uparrow$
Base Model	**-14.05**	**-1.95**	**0.89**	3.14
Multitask model	-13.32	-1.74	0.89	2.71
Single Angle DAS + Speckle Denoising	-10.5	-1.05	0.674	1.85
Dahan et. al. [30]	-10.07	-1.673	0.885	**3.17**
Ground Truth	-25.6	-2.423	0.98	2.71

Linear Weight Adaptation

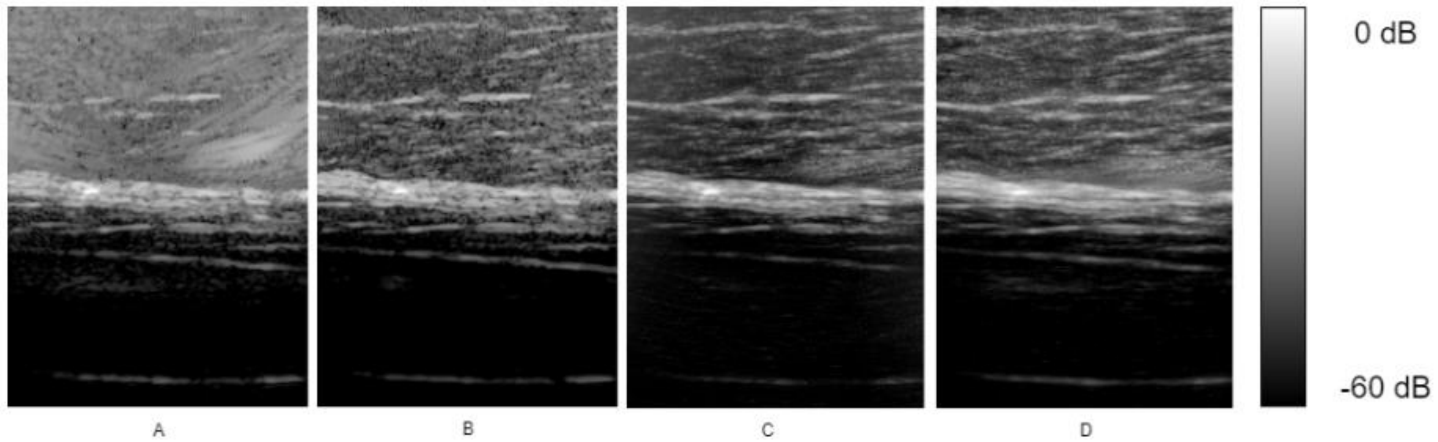


Fig. 2: In vivo inference of the different approach of speckle reduction. A) Single angle DAS with speckle reduction. B) 75-angle DAS with speckle reduction. C) Specifically trained neural network for speckle reduction. D) our multitask learning approach of speckle reduction. The results are very close to those of the specifically trained model while adapting the model to the target task with significantly fewer parameters, thus providing a more scalable approach.

Linear Weight Adaptation

Advantages	Disadvantages
Scalable	Increase inference computational complexity
results on par with task specific network training	Restricted to convolutional layers
Outperforms existing deep multitask learning methods both in robustness and efficiency.	

Parameter Efficiency Analysis

The ration between the number of learnable parameters and per-task extra parameters is given by:

$$R = \frac{K_x K_y}{C_1 C_2 + \frac{C_2}{K_x K_y}},$$

Where K_x, K_y be the horizontal and vertical dimension of 2D convolution kernel,
And C_1, C_2 are the input and output number of filters at the specific layer, respectively.

In modern applications, $C_1 C_2 \gg K_x K_y$ thus $R \ll 1$.
Our network had **250K** trainable parameters
while **our method required only 173 additional parameters** for speckle denoising.

Conclusions

We presented two methods for multitask learning applied to ultrasound beamforming.

We compared our approach to a model tailored to task of ultrasound beamforming.
Our model presented the best results at the challenge on ultrasound beamforming.

The proposed weight normalization scheme out-performed classical algorithms for image formation.

Linear weight transformation approach presented the best results. Both perceptually and computationally.

Future Research

- Future research should investigate the scalability of multitask learning.
- Robust deep learning-based beamforming algorithms from multi plane wave channel data.
- Apply modern deep learning architectures (like transformers) for ultrasound image formation and adapt the proposed methods for multitask learning to non-CNN architectures.

Thank You!

Publication Info

- Elay Dahan and Israel Cohen. Deep-learning-based multitask ultrasound beamforming. In-formation, 14(10):582, 2023.
- Elay Dahan and Israel Cohen. Lightweight Multitask Deep Learning for Ultrasound Image Formation. Under review for MDPI Applied Sciences Journal, Research topics Advances in Deep Learning and Intelligent Computing.