

Title: Deep learning in photoacoustic imaging

URL: [https://en.wikipedia.org/wiki/Deep\\_learning\\_in\\_photoacoustic\\_imaging](https://en.wikipedia.org/wiki/Deep_learning_in_photoacoustic_imaging)

PageID: 63664231

Categories: Category:Computing in medical imaging, Category:Deep learning, Category:Medical imaging

Source: Wikipedia (CC BY-SA 4.0).

-----

Photoacoustic imaging (PA) is based on the photoacoustic effect, in which optical absorption causes a rise in temperature, which causes a subsequent rise in pressure via thermo-elastic expansion. [ 1 ] This pressure rise propagates through the tissue and is sensed via ultrasonic transducers. Due to the proportionality between the optical absorption, the rise in temperature, and the rise in pressure, the ultrasound pressure wave signal can be used to quantify the original optical energy deposition within the tissue. [ 2 ]

Photoacoustic imaging has applications of deep learning in both photoacoustic computed tomography (PACT) and photoacoustic microscopy (PAM). PACT utilizes wide-field optical excitation and an array of unfocused ultrasound transducers. [ 1 ] Similar to other computed tomography methods, the sample is imaged at multiple view angles, which are then used to perform an inverse reconstruction algorithm based on the detection geometry (typically through universal backprojection, [ 3 ] modified delay-and-sum, [ 4 ] or time reversal [ 5 ] [ 6 ] ) to elicit the initial pressure distribution within the tissue. PAM on the other hand uses focused ultrasound detection combined with weakly focused optical excitation (acoustic resolution PAM or AR-PAM) or tightly focused optical excitation (optical resolution PAM or OR-PAM). [ 7 ] PAM typically captures images point-by-point via a mechanical raster scanning pattern. At each scanned point, the acoustic time-of-flight provides axial resolution while the acoustic focusing yields lateral resolution. [ 1 ]

#### Applications of deep learning in PACT

The first application of deep learning in PACT was by Reiter et al. [ 8 ] in which a deep neural network was trained to learn spatial impulse responses and locate photoacoustic point sources. The resulting mean axial and lateral point location errors on 2,412 of their randomly selected test images were 0.28 mm and 0.37 mm respectively. After this initial implementation, the applications of deep learning in PACT have branched out primarily into removing artifacts from acoustic reflections, [ 9 ] sparse sampling, [ 10 ] [ 11 ] [ 12 ] limited-view, [ 13 ] [ 14 ] [ 15 ] and limited-bandwidth. [ 16 ] [ 14 ] [ 17 ] [ 18 ] There has also been some recent work in PACT toward using deep learning for wavefront localization. [ 19 ] There have been networks based on fusion of information from two different reconstructions to improve the reconstruction using deep learning fusion based networks. [ 20 ]

#### Using deep learning to locate photoacoustic point sources

Traditional photoacoustic beamforming techniques modeled photoacoustic wave propagation by using detector array geometry and the time-of-flight to account for differences in the PA signal arrival time. However, this technique failed to account for reverberant acoustic signals caused by acoustic reflection, resulting in acoustic reflection artifacts that corrupt the true photoacoustic point source location information. In Reiter et al. , [ 8 ] a convolutional neural network (similar to a simple VGG-16 [ 21 ] style architecture) was used that took pre-beamformed photoacoustic data as input and outputted a classification result specifying the 2-D point source location.

#### Deep learning for PA wavefront localization

Johnstonbaugh et al. [ 19 ] was able to localize the source of photoacoustic wavefronts with a deep neural network. The network used was an encoder-decoder style convolutional neural network. The encoder-decoder network was made of residual convolution, upsampling, and high field-of-view convolution modules. A Nyquist convolution layer and differentiable spatial-to-numerical transform layer were also used within the architecture. Simulated PA wavefronts served as the input for

training the model. To create the wavefronts, the forward simulation of light propagation was done with the NIRFast toolbox and the light-diffusion approximation, while the forward simulation of sound propagation was done with the K-Wave toolbox. The simulated wavefronts were subjected to different scattering mediums and Gaussian noise. The output for the network was an artifact free heat map of the targets axial and lateral position. The network had a mean error rate of less than 30 microns when localizing target below 40 mm and had a mean error rate of 1.06 mm for localizing targets between 40 mm and 60 mm. [ 19 ] With a slight modification to the network, the model was able to accommodate multi target localization. [ 19 ] A validation experiment was performed in which pencil lead was submerged into an intralipid solution at a depth of 32 mm. The network was able to localize the lead's position when the solution had a reduced scattering coefficient of 0, 5, 10, and 15 cm<sup>-1</sup>. [ 19 ] The results of the network show improvements over standard delay-and-sum or frequency-domain beamforming algorithms and Johnstonbaugh proposes that this technology could be used for optical wavefront shaping, circulating melanoma cell detection, and real-time vascular surgeries. [ 19 ]

Removing acoustic reflection artifacts (in the presence of multiple sources and channel noise)

Building on the work of Reiter et al., [ 8 ] Allman et al. [ 9 ] utilized a full VGG-16 [ 21 ] architecture to locate point sources and remove reflection artifacts within raw photoacoustic channel data (in the presence of multiple sources and channel noise). This utilization of deep learning trained on simulated data produced in the MATLAB k-wave library, and then later reaffirmed their results on experimental data.

Ill-posed PACT reconstruction

In PACT, tomographic reconstruction is performed, in which the projections from multiple solid angles are combined to form an image. When reconstruction methods like filtered backprojection or time reversal, are ill-posed inverse problems [ 22 ] due to sampling under the Nyquist-Shannon's sampling requirement or with limited-bandwidth/view, the resulting reconstruction contains image artifacts. Traditionally these artifacts were removed with slow iterative methods like total variation minimization, but the advent of deep learning approaches has opened a new avenue that utilizes a priori knowledge from network training to remove artifacts. In the deep learning methods that seek to remove these sparse sampling, limited-bandwidth, and limited-view artifacts, the typical workflow involves first performing the ill-posed reconstruction technique to transform the pre-beamformed data into a 2-D representation of the initial pressure distribution that contains artifacts. Then, a convolutional neural network (CNN) is trained to remove the artifacts, in order to produce an artifact-free representation of the ground truth initial pressure distribution.

Using deep learning to remove sparse sampling artifacts

When the density of uniform tomographic view angles is under what is prescribed by the Nyquist-Shannon's sampling theorem, it is said that the imaging system is performing sparse sampling. Sparse sampling typically occurs as a way of keeping production costs low and improving image acquisition speed. [ 10 ] The typical network architectures used to remove these sparse sampling artifacts are U-net [ 10 ] [ 12 ] and Fully Dense (FD) U-net. [ 11 ] Both of these architectures contain a compression and decompression phase. The compression phase learns to compress the image to a latent representation that lacks the imaging artifacts and other details. [ 23 ] The decompression phase then combines with information passed by the residual connections in order to add back image details without adding in the details associated with the artifacts. [ 23 ] FD U-net modifies the original U-net architecture by including dense blocks that allow layers to utilize information learned by previous layers within the dense block. [ 11 ] Another technique was proposed using a simple CNN based architecture for removal of artifacts and improving the k-wave image reconstruction. [ 17 ]

Removing limited-view artifacts with deep learning

When a region of partial solid angles are not captured, generally due to geometric limitations, the image acquisition is said to have limited-view. [ 24 ] As illustrated by the experiments of Davoudi et al., [ 12 ] limited-view corruptions can be directly observed as missing information in the frequency domain of the reconstructed image. Limited-view, similar to sparse sampling, makes the initial reconstruction algorithm ill-posed. Prior to deep learning, the limited-view problem was addressed

with complex hardware such as acoustic deflectors [ 25 ] and full ring-shaped transducer arrays, [ 12 ] [ 26 ] as well as solutions like compressed sensing, [ 27 ] [ 28 ] [ 29 ] [ 30 ] [ 31 ] weighted factor, [ 32 ] and iterative filtered backprojection. [ 33 ] [ 34 ] The result of this ill-posed reconstruction is imaging artifacts that can be removed by CNNs. The deep learning algorithms used to remove limited-view artifacts include U-net [ 12 ] [ 15 ] [ 35 ] and FD U-net, [ 36 ] as well as generative adversarial networks (GANs) [ 14 ] and volumetric versions of U-net. [ 13 ] One GAN implementation of note improved upon U-net by using U-net as a generator and VGG as a discriminator, with the Wasserstein metric and gradient penalty to stabilize training (WGAN-GP). [ 14 ]

#### Pixel-wise interpolation and deep learning for faster reconstruction of limited-view signals

Guan et al. [ 36 ] was able to apply a FD U-net to remove artifacts from simulated limited-view reconstructed PA images. PA images reconstructed with the time-reversal process and PA data collected with either 16, 32, or 64 sensors served as the input to the network and the ground truth images served as the desired output. The network was able to remove artifacts created in the time-reversal process from synthetic, mouse brain, fundus, and lung vasculature phantoms. [ 36 ] This process was similar to the work done for clearing artifacts from sparse and limited view images done by Davoudi et al. [ 12 ] To improve the speed of reconstruction and to allow for the FD U-net to use more information from the sensor, Guan et al. proposed to use a pixel-wise interpolation as an input to the network instead of a reconstructed image. [ 36 ] Using a pixel-wise interpolation would remove the need to produce an initial image that may remove small details or make details unrecoverable by obscuring them with artifacts. To create the pixel-wise interpolation, the time-of-flight for each pixel was calculated using the wave propagation equation. Next, a reconstruction grid was created from pressure measurements calculated from the pixels' time-of-flight. Using the reconstruction grid as an input, the FD U-net was able to create artifact free reconstructed images. This pixel-wise interpolation method was faster and achieved better peak signal to noise ratios (PSNR) and structural similarity index measures (SSIM) than artifact free images created when the time-reversal images served as the input to the FD U-net. [ 36 ] This pixel-wise interpolation method was significantly faster and had comparable PSNR and SSIM than the images reconstructed from the computationally intensive iterative approach . [ 36 ] The pixel-wise method proposed in this study was only proven for in silico experiments with homogenous medium, but Guan posits that the pixel-wise method can be used for real time PAT rendering. [ 36 ]

#### Limited-bandwidth artifact removal with deep neural networks

The limited-bandwidth problem occurs as a result of the ultrasound transducer array's limited detection frequency bandwidth. This transducer array acts like a band-pass filter in the frequency domain, attenuating both high and low frequencies within the photoacoustic signal. [ 15 ] [ 16 ] This limited-bandwidth can cause artifacts and limit the axial resolution of the imaging system. [ 14 ] The primary deep neural network architectures used to remove limited-bandwidth artifacts have been WGAN-GP [ 14 ] and modified U-net. [ 15 ] [ 16 ] The typical method to remove artifacts and denoise limited-bandwidth reconstructions before deep learning was Wiener filtering, which helps to expand the PA signal's frequency spectrum. [ 14 ] The primary advantage of the deep learning method over Wiener filtering is that Wiener filtering requires a high initial signal-to-noise ratio (SNR), which is not always possible, while the deep learning model has no such restriction. [ 14 ]

#### Fusion of information for improving photoacoustic Images with deep neural networks

The complementary information is utilized using fusion based architectures for improving the photoacoustic image reconstruction. [ 20 ] Since different reconstructions promote different characteristics in the output and hence the image quality and characteristics vary if a different reconstruction technique is used. [ 20 ] A novel fusion based architecture was proposed to combine the output of two different reconstructions and give a better image quality as compared to any of those reconstructions. It includes weight sharing, and fusion of characteristics to achieve the desired improvement in the output image quality. [ 20 ]

#### Deep learning to improve penetration depth of PA images

High energy lasers allow for light to reach deep into tissue and they allow for deep structures to be visible in PA images. High energy lasers provide a greater penetration depth than low energy lasers. Around an 8 mm greater penetration depth for lasers with a wavelength between 690 and 900 nm. [ 35 ] The American National Standards Institute has set a maximal permissible exposure (MPE) for different biological tissues. Lasers with specifications above the MPE can cause mechanical or thermal damage to the tissue they are imaging. [ 35 ] Manwar et al. was able to increase the penetration of depth of low energy lasers that meet the MPE standard by applying a U-net architecture to the images created by a low energy laser. [ 35 ] The network was trained with images of an ex vivo sheep brain created by a low energy laser of 20 mJ as the input to the network and images of the same sheep brain created by a high energy laser of 100 mJ, 20 mJ above the MPE, as the desired output. A perceptually sensitive loss function was used to train the network to increase the low signal-to-noise ratio in PA images created by the low energy laser. The trained network was able to increase the peak-to-background ratio by 4.19 dB and penetration depth by 5.88% for photos created by the low energy laser of an in vivo sheep brain. [ 35 ] Manwar claims that this technology could be beneficial in neonatal brain imaging where transfontanelle imaging is possible to look for any lesions or injury.

#### Applications of deep learning in PAM

Photoacoustic microscopy differs from other forms of photoacoustic tomography in that it uses focused ultrasound detection to acquire images pixel-by-pixel. PAM images are acquired as time-resolved volumetric data that is typically mapped to a 2-D projection via a Hilbert transform and maximum amplitude projection (MAP). [ 1 ] The first application of deep learning to PAM, took the form of a motion-correction algorithm. [ 37 ] This procedure was posed to correct the PAM artifacts that occur when an in vivo model moves during scanning. This movement creates the appearance of vessel discontinuities.

#### Deep learning to remove motion artifacts in PAM

The two primary motion artifact types addressed by deep learning in PAM are displacements in the vertical and tilted directions. Chen et al. [ 37 ] used a simple three layer convolutional neural network, with each layer represented by a weight matrix and a bias vector, in order to remove the PAM motion artifacts. Two of the convolutional layers contain RELU activation functions, while the last has no activation function. [ 37 ] Using this architecture, kernel sizes of  $3 \times 3$ ,  $4 \times 4$ , and  $5 \times 5$  were tested, with the largest kernel size of  $5 \times 5$  yielding the best results. [ 37 ] After training, the performance of the motion correction model was tested and performed well on both simulation and in vivo data. [ 37 ]

#### Deep learning-assisted frequency-domain PAM

Frequency-domain PAM constitutes a powerful cost-efficient imaging method integrating intensity-modulated laser beams emitted by continuous wave sources for the excitation of single-frequency PA signals. [ 38 ] Nevertheless, this imaging approach generally provides smaller signal-to-noise ratios (SNR) which can be up to two orders of magnitude lower than the conventional time-domain systems. [ 39 ] To overcome the inherent SNR limitation of frequency-domain PAM, a U-Net neural network has been utilized to augment the generated images without the need for excessive averaging or the application of high optical power on the sample. In this context, the accessibility of PAM is improved as the system's cost is dramatically reduced while retaining sufficiently high image quality standards for demanding biological observations. [ 40 ]

See also

Photoacoustic imaging

Photoacoustic microscopy

Photoacoustic effect

References

External links

Photoacoustic imaging

Photoacoustic microscopy

Photoacoustic effect