# Deep Learning for Image Processing and Reconstruction to Enhance LED-Based Photoacoustic Imaging

Kathyayini Sivasubramanian and Lei Xing

**Abstract** Photoacoustic imaging is a rapidly growing imaging technique which combines the best of optical and ultrasound imaging. For the clinical translation of photoacoustic imaging, a lot of steps are being taken and different parameters are being continuously improved. Improvement in image reconstruction, denoising and improvement of resolution are important especially for photoacoustic images obtained from low energy lasers like pulsed laser diodes and light emitting diodes. Machine learning and artificial intelligence can help in the process significantly. Particularly deep learning based models using convolutional neural networks can aid in the image improvement in a very short duration. In this chapter we will be discussing the basics of neural networks and how they can be used for improving photoacoustic imaging. We will also discuss few examples of deep learning networks put to use for image reconstruction, image denoising, and improving image resolution in photoacoustic imaging. We will also discuss further the possibilities with deep learning in the photoacoustic imaging arena.

## 1 Introduction

### 1.1 Photoacoustic Imaging

Photoacoustic (PA) imaging is a hybrid imaging technique which has gained importance in the last several years [1–4]. It is based on the photoacoustic effect discovered by Alexander Graham Bell in 1881 [5]. He observed that light energy absorbed by a material results in an acoustic signal. He demonstrated this with an apparatus called photophone which he designed. Almost after a century, it started gaining importance

K. Sivasubramanian (✉) · L. Xing
Department of Radiation Oncology, School of Medicine, Stanford University, Stanford, CA, USA
e-mail: kathya19@stanford.edu

L. Xing
e-mail: lei@stanford.edu

as we discovered that it can be used for the purpose of imaging. The major advantage of photoacoustic imaging is that it combines the best features of two different imaging modalities: the contrast of optical imaging and the resolution of ultrasound imaging [1, 6]. Photoacoustic imaging is based on the principle that when a pulsed laser light (pulse width in the range of nanoseconds) falls on a sample, if the sample absorbs the light at the particular wavelength then it undergoes a small increase in temperature in the order of milli kelvin (mK). Following the heating, there occurs thermoelastic expansion of the sample, which leads to the generation of the pressure waves. The pressure waves can be detected as photoacoustic waves by ultrasound transducers. The sound waves captured by the transducers are then reconstructed to form images known as photoacoustic images [7, 8]. Contrast agents are very important for photoacoustic imaging because when a sample is irradiated with a particular wavelength of light, only if the contrast agent absorbs light at that wavelength, detectable photoacoustic waves will be emitted from the sample [9–11]. The imaging wavelength is usually in the visible and the near-infra red (NIR) region of the optical spectrum, very recently the second NIR window is being explored for PA imaging [12]. Therefore, availability of contrast agents in these wavelength regions is very critical for PA imaging. Fortunately, there are some intrinsic contrast agents like blood (hemoglobin), melanin, lipids, etc. present in the human body which provides great contrast in the visible and near infrared spectrum [13–16]. However, the contrast from these are only sufficient and suitable for imaging certain body parts and for certain applications. Therefore, for imaging other organs and for different applications, the use of extrinsic contrast agents becomes inevitable. Some of the most commonly used extrinsic contrast agents are organic dyes, inorganic dyes, nanoparticles, nanomaterials, etc. Constant research is being done to develop highly efficient photoacoustic contrast agents [17–25]. Different forms of photoacoustic imaging are available like photoacoustic microscopy, tomography, endoscopy etc. [16, 26–32]. The applications of photoacoustic imaging ranges from cellular level imaging to systems imaging. Photoacoustic imaging can be used for obtaining both structural and functional data from the sample. Some of the most explored applications of photoacoustic imaging includes sentinel lymph node imaging, brain imaging, blood vasculature imaging, tumor imaging and monitoring, oxygen saturation monitoring etc. [33–44].

The PA wave equation is given by

$$(\nabla^2 - v_s^{-2}\partial^2/\partial t^2)p(\vec{r}, t) = -(\beta/C_P)\partial H(\vec{r}, t)/\partial t$$

Here $v_s$ refers to acoustic speed, $p(\vec{r}, t)$ refers to the acoustic pressure at location r and time t, $\beta$ refers to the thermal expansion coefficient, $C_P$ refers to the specific heat constant at constant pressure, and $H$ denotes the heating function which can be described as the thermal energy converted per unit volume and per unit time. The left-hand side of this equation describes the wave propagation, whereas the right-hand side represents the source term.

Traditionally, large and bulky Nd:YAG or dye based lasers are used as illumination source for photoacoustic imaging. They often need an optical table for housing them and are non-portable. Even the smallest misalignment will alter the results greatly

[45]. Very recently portable, mobile Nd:YAG laser with optical parametric oscillator for tuning different wavelengths has been commercially available from opotek [37]. The biggest advantage of using these lasers is the high laser energy, which in turn translates to higher penetration and high-resolution images. The catch with this laser is that it is very difficult to combine the light to the ultrasound transducer, which makes the clinical use of these lasers very limited. However, in recent times compact, lightweight lasers are starting to be used for imaging like the pulsed laser diode (PLD) and the light emitting diodes (LED) [46, 47]. The pulsed laser diode is very small often palm size, and very light weight which makes it easily portable to use. It can also be integrated with the ultrasound transducer much easily than the OPO laser. The frequency of these lasers is very high therefore, they can provide large number of frames in a short period of time. The problem with this laser is that the pulse energy is very low and will often need averaging over multiple frames to obtain a high-resolution image. LEDs are similar to the PLDs but with lesser energy. The frequency of the LEDs is also very high. Multiple LEDs are placed in an array to generate light for imaging. But, even with an array of LEDs the pulse energy of the system is very low [48, 49]. The system requires a lot of averaging to obtain an acceptable photoacoustic image. One major disadvantage with these systems is that they are usually single wavelength and cannot be tuned. Therefore, cannot be used for spectroscopic studies like the Nd:YAG laser. However, in the last few years, multiple commercial systems have been developed for real-time photoacoustic imaging using different types of lasers like the Nd:YAG, PLD, LEDs etc. Ongoing research is being done on how to improve the resolution of images from the low energy laser sources like PLD, LEDs etc.

## 1.2 Photoacoustic Image Acquisition and Reconstruction

As much as the light plays a crucial role in photoacoustic imaging, equally important are the ultrasound transducers. The signal form the sample can be acquired using ultrasound transducer [50, 51]. There are many ways in which an ultrasound transducer is used for photoacoustic imaging, a single element transducer can be used for signal acquisition or a raster scan be performed to obtain a 2D or 3-D image or it can be rotated around the sample to obtain a cross-sectional image. However, it can be very time consuming to scan a big area. In order to complete scanning in a very short time multiple transducers can be combined to make an array of transducers to obtain images [33, 52, 53]. When using commercial systems, linear array, concave array and convex array-based transducers are also available for data acquisition. These transducers are supported by the data acquisition cards (DAQ) for image acquisition. Once the data is acquired using the ultrasound transducers, it goes through the reconstruction process to form the final image. Different types of reconstruction methods, such as filtered back-projection, Fourier transform, alternative algorithm, time reversal, inversion of the linear Radon transform, and delay and sum beamforming, have been developed under different assumptions and approximation for ultrasound and

photoacoustic imaging [54–59]. The issue with these image reconstruction methods is that these methods assumes that the wave propagation is through a homogenous media, but in reality, that is often not the case. Another issue with using these image reconstruction methodologies is that these methods often generates artifacts like reflection artifact, etc. and cannot remove these artifacts. Post-processing of images is often used to remove some of the artifacts from the reconstructed images. However, the existing reconstruction and post-processing techniques are not sufficient to improve the quality of the images. There is a great need to improve the imaging resolution, reduce noise and remove artifacts of the photoacoustic images for clinical translation. Continuous research is required and being done on how to improve the image resolution from the perspective of reconstruction and post-processing.

Among the various image reconstruction methods, the delay-and-sum beamforming method is the most widely used algorithm for the reconstruction of both PA and US images. This algorithm works by summing the corresponding US signals while adjusting their time delays in accordance to the distance between the detectors and the sample. However, it has few drawbacks like low resolution, low contrast, and strong side lobes which results in artifact generation. Matrone et al. proposed a modification to the DAS algorithm leading to a novel beamforming algorithm, called the delay-multiply-and-sum (DMAS) beamformer, in order to help in overcoming the limitations of DAS in ultrasound imaging. The DMAS provides the high contrast and enhanced image quality, it also helps in obtaining narrow main lobes, and weaker side lobes in comparison to DAS. Owing to these advantages, several researchers extended the ultrasound DMAS algorithm to PA imaging also. Park et al. introduced a DMAS-based synthetic aperture focusing technique to PA microscopy. Alshaya et al. demonstrated the DMAS based PA imaging can be useful when using a linear array transducer also and additionally they introduced a subgroup of DMAS method to improve the signal to noise ratio (SNR) and the speed of image processing. To improve the quality of the image obtained from DMAS algorithm even more, Mozaffarzadeh et al. proposed using a double-stage DMAS operation, a minimum variance beamforming algorithm, or modified coherence factor [60–62]. In spite of all these advances, it has been difficult to use DMAS for image reconstruction clinically because of the heavy computation complexity involved in the incorporation of this algorithm to a clinical PA imaging system.

Another commonly used reconstruction technique for photoacoustic imaging is the back-projection (BP) method. This reconstruction technique and its derivatives like the filtered back projection (FBP) are one of the major reconstruction algorithms used for the photoacoustic computed tomography (PACT) specifically. This algorithm makes use of fact that the pressure propagating from an acoustic source reach the detectors at different time delays, which depends on a myriad of factors like the speed of sound, the distance between the source and the detectors, etc. The BP algorithm requires large number of signals collected from various view angles as its input. These signals can be collected by a single transducer or use an array of transducers rotating around the sample. Both the methods have their own pros and cons. This is a faster reconstruction technique, back-projection (BP) algorithms are

capable of producing good images for common geometries (planar, spherical, cylindrical) in simulations and is also applied widely for volumetric image reconstruction in PA imaging. Constant development of BP algorithms leads to improved image quality, which has improved the possibilities with PA imaging and the capabilities of PA imaging in the various biomedical applications. The formulas of back projection techniques are implemented either in the spatio-temporal domain or in the Fourier domain. The BP algorithms are constantly modified to improve the applications and the image quality, one of the modifications is based on a closed-form inversion formula. This modified algorithm was very successful in detection of the position and shape of absorbing objects in turbid media.

Although filtered back-projection (FBP) reconstruction techniques has proven its use in solving for time-dependent partial differential equations through Fourier spectral methods, there are still many critical problems that needs addressing to further improve the quality of FBP-reconstructed images [63, 64]. One of the shortcomings of the conventional back-projection algorithm is that they are not exact in experimental setting and may lead to the generation of substantial artifacts in the reconstructed image, such as the accentuation of fast variations in the image, which is accompanied by negative optical-absorption values that otherwise have no physical interpretation [59, 65]. The presence of the artifacts has not restricted the use of BP algorithms for structural PA imaging, they do affect the quantification capacity, the image fidelity, and the accurate use of the method for functional and molecular imaging applications.

Time reversal is another reconstruction method used in photoacoustic imaging. In the typical time reversal imaging reconstruction method, the recorded pressure time series are enforced in time reversed order as a Dirichlet boundary condition as the position of detectors on the measurement surface [66–69]. If the array of detectors is placed sparsely to collect the measurement rather instead of a continuous surface, the time reversed boundary condition will be discontinuous. This can cause severe blurring in the reconstructed images. To solve the problem, Treeby et al., improved time reversal image reconstruction technique with the usage of interpolated sensor data. In the course, the interaction can be avoided by interpolating the recorded data onto a continuous rather than discrete measurement surface within the space grid used for the reconstruction. The edges of the reconstructed image are considerably sharper, and the magnitude has also been improved. After that, they used the enforced time reversal boundary condition to trap artifacts in the final image, and by truncating the data, or introducing an adaptive threshold boundary condition, this artifact trapping can be mitigated to some extent.

## 1.3  Types of Artifact

Artifacts are one of the major problems in photoacoustic imaging. The presence of artifacts limits the application of photoacoustic imaging from a clinical perspective and hampers the clinical translation of the imaging modality greatly. Reflection artifact is one of the most commonly observed artifacts in photoacoustic imaging [67,

70, 71]. These reflections are not considered by traditional beamformers which use a time-of-flight measurement to create images. Therefore, reflections appear as signals that are mapped to incorrect locations in the beamformed image. The acoustic environment can also additionally introduce inconsistencies, like the speed of sound, density, or attenuation variations, which makes the propagation of acoustic wave very difficult to model. The reflection artifacts can become very confusing for clinicians during diagnosis and treatment monitoring using PA imaging. Until these are corrected the possibility of clinical translation is very slim.

In order to minimize the effect of artifacts in photoacoustic imaging different signal processing approaches have been implemented to enhance signal and image quality. These signal processing techniques use singular value decomposition and short-lag spatial coherence. But these techniques are not so efficient in the removal of intense acoustic reflection artifacts. A technique called photoacoustic-guided focused ultrasound (PAFUSion) was developed which differs from other traditional photoacoustic artifact reduction methodologies as it uses ultrasound to mimic wavefields produced by photoacoustic sources in order to identify reflection artifacts for removal [72, 73]. A slight modification of this approach was developed which uses plane waves instead of focused waves, but the implementation was very similar. Both of these methods make the assumption that acoustic reception pathways are identical, which may not always be true. When performing simultaneous ultrasound and photoacoustic imaging in real-time it is not always possible to have an exact overlay of the image because of the motion induced artifact caused by the moving organs inside the body especially organs like heart, abdominal cavity, blood vessels etc. Certain reconstruction methods have been proposed to overcome these types of artifacts, but the problem is they don't account for the inter patient variability and sometimes variability in the same patient when imaging different body parts (Fig. 1).

## 1.4  LED Based Photoacoustic Imaging

LED based photoacoustic system can play a very important role in the clinical translation of photoacoustic imaging. LEDs are less expensive compared to the traditional lasers for photoacoustic imaging, they are very compact, and capable of imaging in multiwavelength (e.g., 750, 810, 930, and 980 nm) [47–49]. The energy output from the LED arrays is much lesser than the energy from powerful class-IV lasers, therefore these can be used for clinical applications easily. But they have very low energy and usually produce low resolution images and are noisier. They also have higher laser pulse width, which limits the spatial resolution of the images. In order to obtain better images, signal averaging in the order of 1000s is required to obtain one image, which increases the image acquisition time. But, in spite of all the shortcomings LED based photoacoustic imaging has gained a lot of momentum with different types of applications that are possible with the system [47–49, 75–77]. The LED based photoacoustic imaging system from Cyberdyne INC (Tsukuba, Japan) can be operated at multiple wavelengths in the visible and the near infrared region. It has a
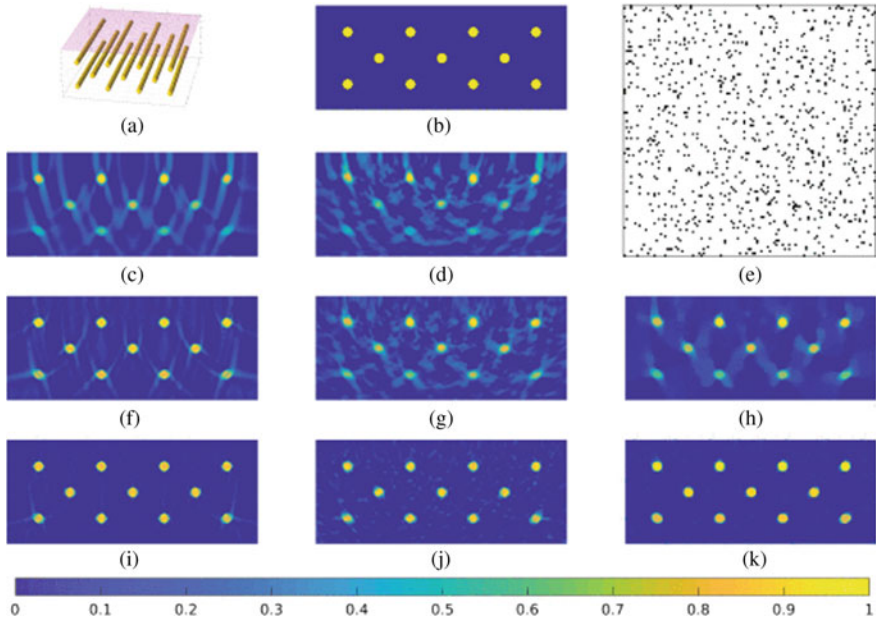
**Fig. 1** Images of errors of different image reconstruction methods using a simple numerical phantom consisting of tubes. **a**, **b** Images of the numerical phantoms. **e** Illustration of the sub-sampling pattern. **c**, **d** Slice view of full and sub-sampled data respectively. **f**–**k** Slice views through the reconstructions of the tube phantom by different methods and for full or sub-sampled data. **f**, **i** non negative least squares (NNLS) of full data at different iterations. **g**, **j** NNLS of sub-sampled data at different iterations. **h**, **k** total variation (TV) of sub-sampled data at different iterations. Reprinted with permission from Ref. [74]

linear array transducer for image acquisition and a 128-channel data acquisition card. The system comes with inbuilt image reconstruction algorithms based on delay and sum model, the image further undergoes post-processing through various filters [48, 77–79]. Figure 2 shows the schematic and the photograph of the LED-photoacoustic imaging system (PLED-PA). It has been demonstrated using this system that it can be used for applications like blood vessel imaging, diagnosis of inflammatory arthritis, detection of head and neck cancer, etc. It has also been used for functional imaging of blood oxygen saturation.

With the current reconstruction techniques and post-processing methodologies in photoacoustic imaging it is really difficult to generate artifact and noise free images in shorter time, with lesser averaging and minimal post-processing. This is especially more relevant to the low energy laser sources like the PLD, LED etc. [80, 81]. In order to improve the image reconstruction and reduce noise in the images in a shorter duration, artificial intelligence can be made use of, specifically deep learning using convolution neural networks could be very useful for this purpose. In the rest of the chapter we will focus on how to make use of deep learning for photoacoustic imaging.
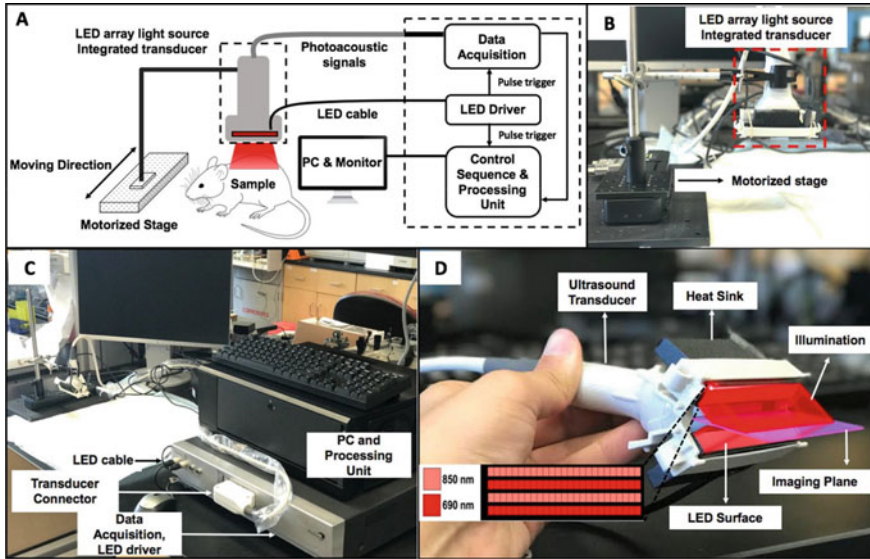
**Fig. 2** LED photoacoustic imaging system. **A** Schematic representation of the PA system using LED array light source. **B** Photograph of PLED-PA probe associated with motorized stage. **C** Whole imaging setup. **D** PLED-PA probe with imaging plane and illumination source are shown schematically. LED array design is also shown in the inset—there were alternating rows of LEDs with different wavelengths. Reprinted with permission from Ref. [47]

## 2 Machine Learning and Artificial Intelligence

In the year 1950, Alan Turing proposed a 'Learning Machine' that could learn and become artificially intelligent. Research in neurology had shown that synapses worked like a network firing electric impulses, based on this idea, the construction of an electronic brain was suggested. Marvin Minsky and Dean Edmonds build the first neural network machine in the year 1951, it was called the Stochastic Neural Analog Reinforcement Calculator (SNARC) [82]. Starting from the 80s the golden age of machine learning began, in that period many ground-breaking discoveries were made but due to non-availability of the infrastructure for higher computing power and speed, further developments were hindered. In the year of 1981, the government of Japan funded a project with the goal to develop machines which could carry on conversations, translate languages, interpret pictures and reason like human beings, some of which are not realized even today. In 1997 IBMs computer 'Deep Blue' bet the world chess champion, Garry Kasparov and in 2005 a robot from Stanford was able to drive autonomously for 131 miles. There are countless other examples of the success of deep learning approaches in numerous fields [83–88].

Artificial intelligence (AI) is a technology that aims to make machines which tries to mimic human brain and the field has grown exponentially in the last few years and continues to impact the world significantly. The applications of artificial intelligence

and machine learning are plenty, almost across all fields with a profound impact on improvement of human lives. Machine learning (ML) and deep learning (DL) has played a very significant role in the improvement of healthcare industry at different levels, right from diagnosis to patient monitoring. It is widely used in the areas of image processing, image analysis, diagnostics, treatment planning and follow-up, thus benefitting a large number of patients. It also helps the clinicians by reducing their workload and helps in making quicker decisions in many cases. Its impact on image processing and analysis especially are noteworthy.

Machine learning is a subset of AI which relies on pattern recognition and data analytics. In ML it is tested to identify if a computer can learn on its own from data without being programmed to perform different tasks. The iterative aspect of machine learning is critical because as models are exposed to new data, they are able to independently adapt. The system learns effectively from previous computations to make predictions and decisions.

## 2.1 Neural Networks

Artificial intelligence and machine learning are not complete without mentioning neural networks (NN). Neural networks are a set of algorithms, modeled inspired and based on human neural system, which are designed to recognize patterns [84, 89–91]. They interpret sensory data through a kind of machine perception, labeling or clustering raw input. Neural networks are capable of recognizing patterns from various input formats such as images, sound, text, etc. The information from the different types of inputs are translated into numerical values that can be understood by the machine. Some of the key areas in which neural networks help are in clustering and classification [88, 92–94]. When group of unlabeled data is presented to the neural networks, they are capable of grouping them according to the similarities between them. When the neural network is presented with a group of labelled data for training, they can classify data effectively. Neural networks are capable of extracting features which are then provided to the clustering and classification algorithms.

Deep learning is a subset of machine learning and is a rapidly growing field of research that targets to significantly enhance the performance of many pattern recognition and machine learning applications. Deep learning makes use of neural network designs for representing the nonlinear input to output map together with optimization procedures for adjusting the weights of the network during the training phase. In the last few years, deep learning-based algorithms were developed for achieving highly accurate reconstruction of tomography images.

## 2.2 *Convolution Neural Network*

The convolutional neural network (CNN), is a special neural network model that is designed to predominantly work with two-dimensional image data. Among neural networks, CNNs are primarily used for image recognition, images classification, Objects detections, recognition faces etc. [88, 93, 95, 96]. As the name suggests CNN derives its name from the convolutional layer and as it suggests this layer performs the "convolution" operation. In CNN, convolution is classified as a linear operation which involves the moving one or more convolution filters (with a set of assigned weights) across the input image [89, 97, 98]. Each of the weight from the convolution layer gets multiplied with the input data from the image on which it is scanned to yield a matrix that is smaller than the input image. The convolution filter is always chosen to be smaller than the input image data and element-wise multiplication followed by summation (dot product) is carried on between the convolution filter and the filter-sized patch of the input. The CNN always uses a filter smaller than the input image because, this filter can be used multiple times on the input data, and it can be moved across the entire input image at different times also. This can happen with data overlap or without overlap (top to bottom and left to right). Each of the convolution filter is designed in such a manner that they can detect a specific type of feature from the input image. As the filter is moved across the image it starts detecting the specific feature it is supposed to [89, 99, 100]. For more efficient and high-quality feature extraction, the filter can be passed through the input image multiple times. The result that is obtained after the filter performs the function of feature extraction is called the feature map, which is a two-dimensional array of the filtered input. Once the feature map is generated, it is passed through a nonlinearity like ReLU. Many numbers of convolution filters can be used on the same input image to extract and identify different types of feature maps. The more the feature maps for a given image, more accurate is the performance of the neural networks.

All CNN consists of at least three different types of layers, an input layer, an output layer and several hidden layers. Initial versions of CNNs were shallow (one input and one output layer with a hidden layer). Deep learning networks is classified as anything that has a minimum of three layers. In deep learning, each passing layer trains on features that are generated as the output from the previous layers and as the number of layers progresses, they are able to recognize more complex features which is known as feature hierarchy. This feature of deep-learning enables the networks to handle very large, high-dimensional data sets with a multitude of parameters. Some of the most commonly used layers in a CNN are discussed below. After the input layer, the very first layer of a CNN is the convolutional layer. The convolution layer performs the convolution operation on the input data. Once the convolution layer extracts the features to generate the feature map, it is passed through a Rectified linear units (ReLu), which are activation functions. Leaky ReLus allow a small, non-zero gradient when the unit is not active. Once the data passes through the ReLu, it goes through the pooling layers. The pooling layers combines the outputs from the neuron clusters at each layer into a single neuron in the next layer. Max Pooling is one of
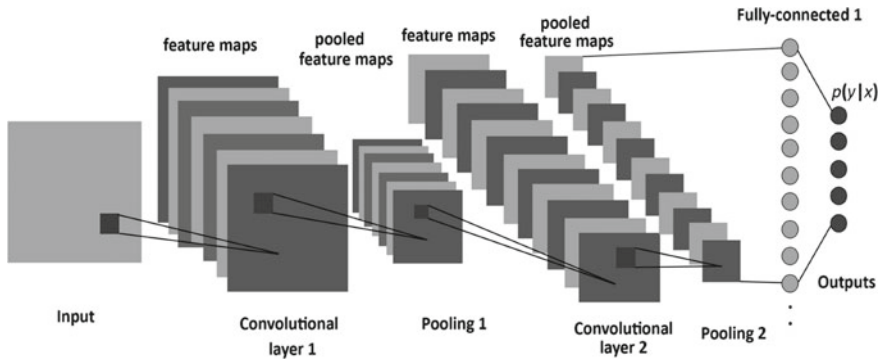
**Fig. 3** Representation of a typical CNN, consisting of convolutional, pooling, and fully-connected layers. Reprinted with permission from Ref. [89]

the examples of pooling layers, this layer chooses and uses the maximum value from each cluster of neurons in the previous layer and sends only the maximum values of the cluster to the next layer. Next, upsampling layers performs an upsampling using nearest neighbor, linear, bi-linear and tri-linear interpolation. Finally, fully connected layers connect every neuron of one layer to every other neuron in another layer. Model of a traditional CNN is shown in Fig. 3.

Feature extraction is a highly time-consuming task and it can be very tedious to perform especially by humans. Deep-learning networks can perform feature extraction with very minimal or without human intervention. This comes in very handy in the medical community, especially in the field like radiology, where there are always limited personnel to scan all the diagnostic images of a patient. This is specifically important because diagnosis from the images is very crucial for the treatment planning and monitoring of a patient.

## 2.3 Learning by Neural Networks

There are two different ways in which neural networks learns, which are supervised learning and unsupervised learning.

1. Supervised learning: Supervised learning is a machine learning method and is widely used, in this method a large dataset is required with corresponding labels. A supervised learning algorithm is trained using ground truth images which is a set of labelled data. Therefore, the algorithm attempts to reproduce the label and calculates a loss function that measures the error between the output from the machine and the label. The algorithm then considers the error value, which is then further factored to modify the internal adjustable parameters (weights), to further minimize this error and improve the efficiency of the model. The performance rate of any machine learning algorithm is based on its how accurately

it handles previously unseen data [101–105]. This can be evaluated with some data that the algorithm has not been exposed during the training process. This data is called test set. The algorithm is said to be more generalized if it is able to predict closer to the ground truth on unseen data. In contrast, if an algorithm can perform with accuracy on previously exposed data but perform very badly on the new test data, it shows that the algorithm only tries to memorize known solutions without any abstraction and does not generalize well. This problem is called overfitting. Overfitting is one of the most frequently encountered problem in machine learning and it can be avoided in many ways [106–108]. Some of them are to either train the algorithm with more data or data augmentation or by completely using a different neural network. Choosing a different neural network works best when the current network cannot handle the complexity of the data.

2. Unsupervised learning: Unsupervised learning is a method in which the algorithms train themselves automatically as they are trained on unlabeled data. In this method each node in every layer of the network tries to learn the features automatically by repeatedly trying to reconstruct the data from the input set, it tries to minimize the variation between the guesses of the network and the probability distribution of the input data itself [109–111]. Also, in this process, the neural networks learn to identify similarities and relationships between certain relevant features and optimal results. The networks try to find connections between feature signals and what it represents, whether it be a full reconstruction, or with labeled data [112, 113]. A deep-learning network can first be trained on labeled data can then be applied to unlabeled data as well. This way, it gives the network access to much more input than just the machine-learning nets. The key to the performance of any deep learning model is data, the greater the amount of data a network trains on, the network's probability of accuracy improves likewise. The output layer of any deep-learning network is either a softmax or logistic layer, the classifier assigns a probability to a specific outcome or label, this type of network is predictive in nature. Neural network follows a corrective feedback loop, that rewards the weights which support the correct guesses, and punishing weights that leads to error. The network tests extensively which combination of input is significant as it tries to reduce error.

Gradient descent is a very commonly used function for optimization. It further adjusts the weights according to the error values obtained. The slope of a neural networks depicts the relationship between the allotted weights and the error function. As a neural network continues to learn, it gradually starts adjusting many weights so that they can map signal to meaning more accurately [114–117]. The relationship between each weight of the network and the error is a derivative, every weight of a network is just one factor in whole deep network which involves multitude of transforms; the signal of each weight passes through activations and gets summed over several layers. The basic crux of a deep learning network is to constantly adjust and modify its weights in response to the error calculated in each iteration. This continues to happen until the error can't be reduced any more. The activation function layer of a network determines the possible output from a given node, based on the

input data [100, 118]. The activation function is set in the layer level and gets applied to all neurons present in that layer. Every output node produces binary output (0 or 1) as the two possible outcomes, as it determines whether an input variable either deserves a label or it does not. Neural networks working on labeled data only produces a binary output, as the input they receive is often continuous. That is, the signals that the network receives as input will be over a range of values and include any number of metrics, depending on the problem it is attempting to solve. The mechanism that is used for the conversion of continuous signals into binary output is known as logistic regression. It calculates the probability that a set of inputs match the label. For continuous inputs to be expressed as probabilities, they must output positive results, since there is no such thing as a negative probability.

(a) Training the network: The neural network starts by randomly initializing weights to the model and calculates the output from the first image. The obtained output image is compared with the ground truth with the help of a loss function. The loss is then back propagated to update and modify the weights of the network. This process is performed multiple times to optimize the performance of the networks.

(b) Testing of the network: After training, a network, testing will be done to evaluate the networks performance. In the testing data no labels are used. The network with the previously trained weights is evaluated on new data that was not encountered by the network previously. These weights determine the prediction of the network.

The Cost functions minimum is searched and an easy way to find a minimum is using gradient descent. Hence, the cost function needs to be differentiable. To perform the adjustment of the weights that are calculated by the gradient descent, the machine learning algorithm computes a gradient vector that, for each weight, which gives an indication on the error amount would increase or decrease if the weight were modified (increased or decreased) by a small amount. By updating the weights step-by-step the cost functions minimum is approached.

The learning rate parameter is introduced to improve the working efficiency of the algorithm. The learning rate is multiplied to the cost function, which thereby decides the step size for each iteration. If for a given algorithm the learning rate is chosen too low, then the algorithm takes a long time to converge to the minima, in contrary if the learning rate is chosen too large, then there is a possibility for the algorithm to overshoot the minima. In the state-of-the-art deep learning algorithms, the learning rate is made flexible which adapts continuously [119]. A lot of work has been done to optimize gradient descent algorithms in recent years [118]. One state-of-the-art algorithm called Adam [120] that is based on adaptive estimates of lower order moments, it is made with high computational efficiency and can deal with large datasets with ease [120].

## *2.4  Backpropagation*

Backpropagation is a methodology in which the contribution of each neuron towards error is calculated after the completion of processing of a batch of data. Using backpropagation, after the calculation of loss function and propagation of the error backwards, the weight of the neurons can be modified. The recently developed networks using back propagation are faster than earlier approaches, thus enabling the neural networks to be used for solving problems which were previously unsolvable. Backpropagation based algorithm is the most commonly used optimization approach in neural networks [121]. Using backpropagation, the networks weights are continuously adapted and thereby facilitating the network to learn the best parameters [121]. Back propagation-based algorithms are being used extensively in medical image processing.

## *2.5  Improving the Networks Performance*

As the applications of the neural networks keeps growing, it becomes very important to constantly improve the performance of the network to optimize their functions better and to improve their efficiency. Some of the ways in which the neural networks performance be improved are as follows.

1. Batch computing: Batch computing is used to improve the computational performance of a neural network. A group of data is consolidated and grouped to form a batch which helps in improving the computational performance as most of the libraries are optimized better for array computing [122].
2. Data Augmentation: Data augmentation is commonly used to increase the amount of data on which the algorithms are being trained on. When there is an increase in the amount of data on which the algorithms learn it leads to an increase in prediction accuracy of the algorithm. Therefore, data augmentation can improve an algorithms performance [123].
3. GPU Computing: GPU computing is a technique used to increase computational speed of processing by using a graphics processing unit (GPU), this unit traditionally handles only computations for computer graphics but it can also be used to compute tasks that are normally carried out using the central processing unit (CPU). The GPU is usually designed with more cores than a CPU and are capable of processing far more graphical data per second than the handling capacity of a CPU. Thus, if the data is transferred to the GPU instead of the CPU and processed there, it can lead to a significant speedup of the computing time.

## 2.6 Evaluation Indices

For the quantitative evaluation of a neural networks based on their performance on the test set, some of the most commonly used evaluation parameters includes signal-to-noise-ratio (SNR), peak-signal-to-noise-ratio (PSNR) and structural similarity index (SSIM). These are calculated for each of the test set data for comparison.

SNR

Signal to noise ratio (SNR) can be defined as the ratio of peak signal intensity from the sample to standard deviation of the background intensities represented in decibels. It is based on absolute signal strength and noise statistics of a given image. SNR can be mathematically represented as follows:

$$\text{SNR} = 20\log_{10}(\mu I/\sigma b)$$

where, $\mu I$ and $\sigma b$ represent the peak signal amplitude of the target area and the standard deviation of the background, respectively.

PSNR

The term peak signal-to-noise ratio (PSNR) can be defined as the ratio between the maximum possible value of a signal in a given image and the power of distorting noise which affects the image quality. Because a variety of signals have a very wide dynamic range, (ratio between the largest and smallest possible values of a changeable quantity) the PSNR is usually expressed in terms of the logarithmic decibel scale.

The mathematical representation of the PSNR is as follows:

$$PSNR = 20\log_{10}\left(\frac{MAX_f}{\sqrt{MSE}}\right)$$

The PSNR is a conventional measurement of the image quality in decibels (dB) based on the mean square differences between the estimated and reference images as:

where the MSE (Mean Squared Error) is given by:

$$MSE = \frac{1}{mn}\sum_{0}^{m-1}\sum_{0}^{n-1}\|f(i,j) - g(i,j)\|^2$$

Here, f represents the matrix data from the original image, g represents the matrix data from the degraded image, m represents the numbers of rows of pixels of the images and i represents the index of that row n represents the number of columns of pixels of the image and j represents the index of that column and MAXf is the maximum signal value that exists in the ground truth image.

SSIM

The Structural Similarity Index (SSIM) metric that is used to quantify the image quality degradation which can be caused due to the image processing tools like as data compression or by loss due to data transmission. It is a reference metric which requires two images from the same image capture namely the reference image and the processed image. The processed image is usually the compressed version.

SSIM measures the perceived quality of a digital image; a higher SSIM (in a scale of 1.0) indicates a better representation of an estimated image in terms of perception.

## 2.7 Training Data

Training data is very important for deep learning for photoacoustic imaging. Generating training data and ground truth images for training algorithms is very crucial as this data determines the efficiency of the model. Also, the number of training data available, the quality of the images and the variety of images in the training data pool helps the neural network model to learn more effectively and be able to handle any kind of images that it might come across in real-time scenario. For different imaging modalities the training data can be acquired in different ways. For the most commonly used clinical imaging techniques like MRI, CT etc. there are multiple open source libraries with thousands of data. We can choose the dataset which is most appropriate for our application and train the neural network with the dataset. However, in case of certain applications where relevant data set might not be available online or for imaging modalities that are not so commonly used for clinical imaging, training data needs to be custom generated. This can be done in two ways, the first is to use different imaging systems to acquire high quality images for the specific application, these types of images are more realistic, and it is easy to get a good ground truth image. But it can be very expensive to acquire enough number of training data to train model and it can also be very time consuming. The other method to generate training data is through simulation models. Simulation is a cost-effective way to generate images and the ideal case scenario can be obtained through simulation images. We can also add any type of artifact on the image to help the model to perform better for a specific application. Simulation images make good training data, but the shortcoming with this method is that it can sometimes be very far from reality that when the model comes across a real image, it may not be trained to work on the image.

## 2.8 Neural Networks for Medical Imaging

Now that what neural networks are and what they do is clear let's explore its applications especially in the field of medical imaging [93, 124–128]. Neural networks are

starting to have a huge impact on different aspects of medical imaging like segmentation, detection, classification etc. especially in the field of radiology. Classification is one of the most important tasks in radiology, it typically consists of predicting some target class like a lesion category or condition in the patient from an image or region of interest in a dataset [99, 104, 129, 130]. This task is used for a wide range of applications, right from determining the presence or absence of a disease to identifying the type of malignancy. Deep learning is very frequently used for the segmentation task which can be defined as the identification of pixels or voxels composing an organ or structure of interest [88, 131, 132]. For a machine learning algorithm, it can be considered as a pixel-level classification task, where the end goal is to determine whether a given pixel belongs to the background or to a target class (e.g., prostate, liver, lesions). For this, from image classification tasks, image masks can be used to perform various quantitative analyses such as virtual surgery planning, radiation therapy planning, or quantitative lesion follow-up. Detection is another common task for the deep learning, it can be used to identify focal lesions such as lung nodules, hepatic lesions, or colon polyps. This can be used as a screening technique before a radiologist can take a look at it [105, 133]. Detection is a subset of the classification task however, classification only aims to predict labels, detection tasks aim to predict the location of potential lesions, often in the form of points, regions, or bounding boxes of interest. All of the three tasks are extremely useful for diagnosis, treatment planning of a disease condition. The labeling of the images varies based on the task it performs. Classification of images requires image labeling. Detection of images requires marking the region of interest, such as a boxplot. Segmentation of images requires pixel-wise delineation of the desired object.

### 2.8.1 Deep Learning for Radiology

Among the various clinical imaging techniques, radiology is one place where deep learning is being explore more extensively. Radiology is one of the most important and widely used clinical imaging tool for diagnosis of many diseases and clinicians depend on it every day. Therefore, using deep learning in radiology can have more impact in the clinics than any other imaging technique [134–140]. In this section we will see how deep learning has come to play in the hospitals.

CheXNeXt is a convolutional neural network that was developed by a team of researchers at Stanford, it has the potential to concurrently detect up to 14 different pathologies, including pneumonia, pleural effusion, pulmonary masses, and nodules in frontal-view chest radiographs. The CheXNeXt CNN was trained and validated internally on dataset of ChestX-ray8 images [141]. A set of 420 images were used for training and kept for validation purpose including images of all the original pathology labels. 3 board-certified cardiothoracic specialist radiologists voted on the images which served as reference standard. The performance of the CheXNeXt's was compared with the performance of 9 radiologists using the area under the receiver operating characteristic curve (AUC) on the validation dataset. It was observed that

the performance of CheXNeXt was similar to the level of radiologists on 11 different pathologies but was not able to the achieve performance level of radiologist on 3 pathologies. The radiologists significantly higher performance on three different pathologies (cardiomegaly, emphysema, and hiatal hernia). CheXNeXt has performed significantly better than radiologists in detecting atelectasis. For the other 10 pathologies there was no statistical significance in differences between radiologists and the CheXNeXt. For the radiologists, the average time to interpret the validation set (420 images) was significantly longer than CheXNeXt. Radiologists took about 240 min but the CheXNeXt took only 1.5 min. One of the drawbacks in this study was that both the CNN and the radiologists were not given any patient history. Another limitation is that all the data acquired for this study was from a single institution only. So, the performance of the algorithm may be biased and limited by it. Figure 4 shows the performance of the algorithm for various disease models in comparison to a doctor. Figure 5 shows the predictions of the algorithm for disease conditions.

This is one example of how a CNN can aid the physicians in the field of radiology. Similarly, different types of algorithms are attempting to solve different types of problems in radiology.

### 2.8.2   Deep Learning for Ultrasound Imaging

Ultrasound imaging is a commonly used imaging technique in the clinics for patient diagnosis. There are many different types of artifacts present in ultrasound imaging, which needs efficient methods for artifact reduction or elimination. Deep learning is being explored for image classification, segmentation and artifact removal problems in ultrasound [142–145]. One such example of a classification problem of thyroid nodules is discussed below.

In the ultrasound images, thyroid nodules appear very heterogeneous in nature with unclear boundaries with various internal components, this makes it very difficult for physicians to discriminate between the benign thyroid nodules and malignant ones. A study was proposed for the diagnosis of thyroid nodules using a hybrid method. The model was developed using a combination of two different pre-trained convolutional neural networks. The two CNNs have different convolutional layers and fully connected layers. Initially, the two which are pretrained with the ImageNet database are trained individually. After individual training the two neural networks, the feature maps are learned by the trained convolutional filters, pooling and normalization operations of the two CNNs. After this the two obtained feature maps are fused and a softmax classifier is used to diagnose (classify) the thyroid nodules. This method was validated on 15,000 ultrasound images obtained from two different hospitals.

For CNN1 and CNN2, a single testing was performed on the training step. A multi-view was adapted to improve the performance of the network. For the input of the trained CNNs 256 views of the thyroid nodule images were cropped and was sampled randomly and used. The output was the average of the result of 256 views.
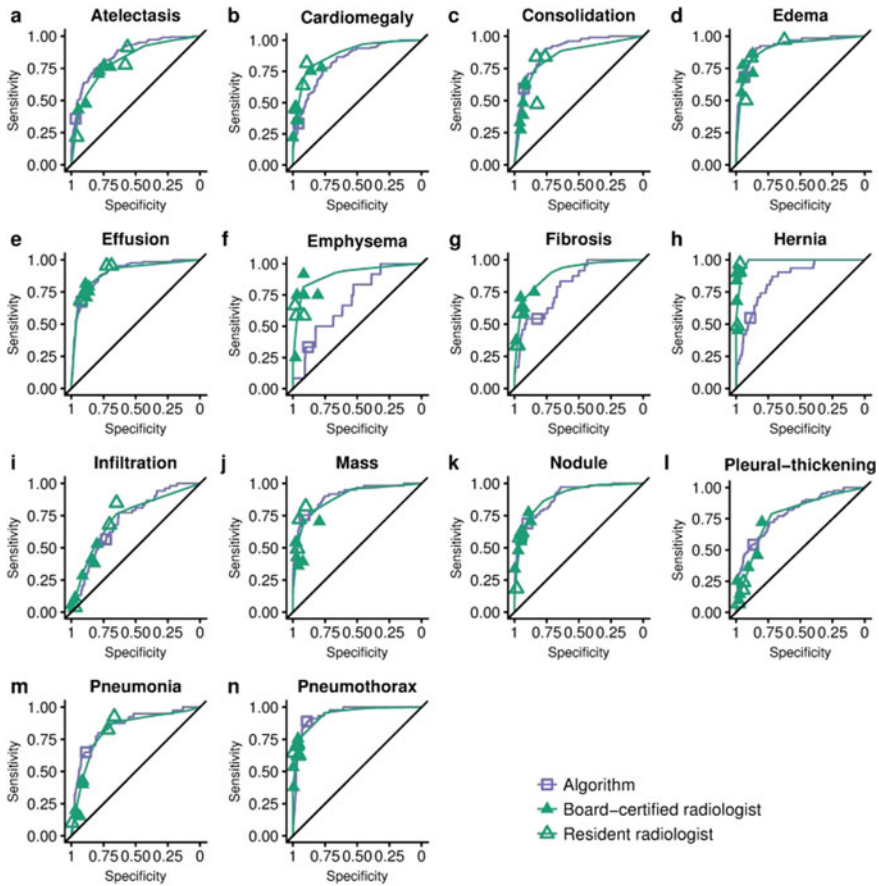
**Fig. 4** ROC curves of radiologists and algorithm for each pathology on the validation set. Each plot illustrates the ROC curve of the deep learning algorithm (purple) and practicing radiologists (green) on the validation set, Individual radiologist (specificity, sensitivity) points are also plotted. The ROC curve of the algorithm is generated by varying the discrimination threshold. Reprinted with permission from Ref. [141]

The two fused pretrained CNN used the fused feature maps that was generated by the two CNNs in multi-view testing as shown in Fig. 6. The softmax layer was trained for thyroid nodule classification. To compare the performance of the CNNs a well-established classification method called SVM was also implemented. The SVM with radial basis function (RBF) kernel was used for experiments [146].

The accuracy of the classification algorithm was tested and represented graphically in Fig. 7; this graph compares the classification accuracy of different methods used in this study. It can be noted from the results that CNN based methods outperform the various other methods significantly in the classification of thyroid nodules. Especially, the combination of CNN1 and CNN2 achieved a classification accuracy
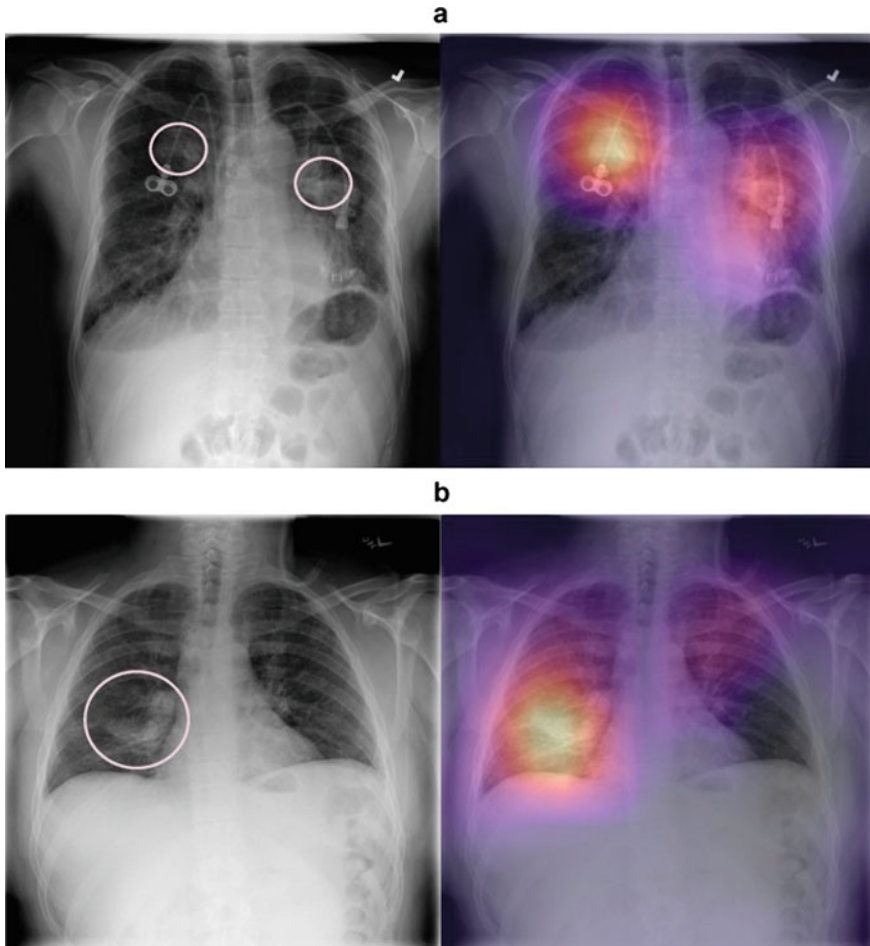
**Fig. 5** In the normal chest radiograph images (left), the pink arrows and circles highlight the locations of the abnormalities. **a** Frontal chest radiograph (left) demonstrates 2 upper-lobe pulmonary masses in a patient with both right- and left-sided central venous catheter. The algorithm correctly classified and localized both masses as indicated by the heat maps. **b** Frontal chest radiograph demonstrates airspace opacity in the right lower lobe consistent with pneumonia. The algorithm correctly classified and localized the abnormality. Reprinted with permission from Ref. [141]

as 83.02% ± 0.72%, sensitivity as 82.41% ± 1.35%, and specificity as 84.96% ± 1.85%. These demonstrate the potential clinical applications of this method.

Photoacoustic imaging is very similar to ultrasound imaging and the techniques from ultrasound can be easily adapted for photoacoustic with minimal modifications.
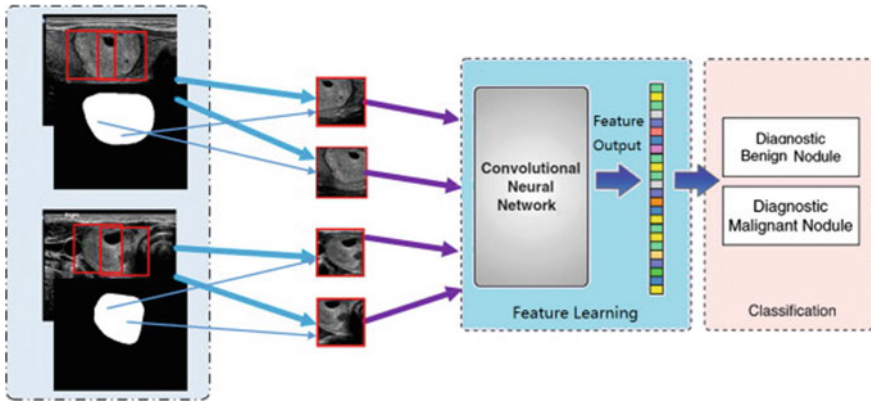
**Fig. 6** An overview testing of CNNs. This CNN based approach first extract multiple nodule patches to capture the wide range of nodule variability from 2D ultrasound images. The obtained patches are then fed into the networks simultaneously to compute discriminative features. Finally, a softmax is applied to label the input nodule. Reprinted with permission from Ref. [146]

**Fig. 7** Box plots of performance measures for classifying between benign and malignant thyroid nodules. In each box plot, the center red line is the median and the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually. Reprinted with permission from Ref. [146]
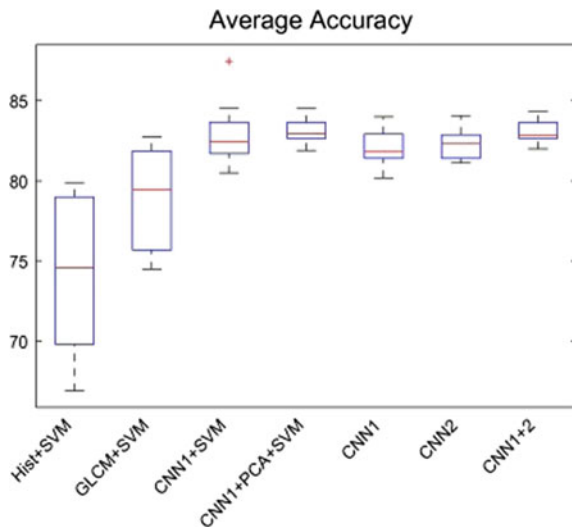


### 2.8.3 Deep Learning for Photoacoustic Imaging

Photoacoustic imaging is not being used in clinics yet. It aspires to become a clinical tool for diagnosis. Image quality and easy interpretability is very crucial for that to happen. Improvement in reconstruction and post processing of images is just one part of it. As evident from the other types of clinical imaging modalities, deep learning can be used for the improvement of photoacoustic imaging for artifact removal and reduction [74, 147–154]. One of the major limitations for using deep learning for

photoacoustic imaging is that there are not much clinically recorded data for training and validation of the neural networks. Therefore generating data through other means is very important for using deep learning for photoacoustic imaging. Using data from simulations is one possible solution for generating data for photoacoustic imaging.

## 3 Monte Carlo Simulation

With respect to photoacoustic imaging monte carlo simulations for light propagation can be used to generate training data. Using monte carlo simulations, the light absorbance by the sample can be calculated [155–158]. The absorbance is usually directly proportional to the photoacoustic signal intensity. Therefore, we can get an idea of how the photoacoustic image will look like. In monte carlo simulation a sample object (of desired shape and size) is simulated in medium like tissue or water, with the properties of the tissue specified. The absorption coefficient and the transmission coefficient of the sample and tissue are predetermined from literature and the number of layers are also mentioned [157–161]. Photon packet is launched from the light source and the movement of the photon is tracked as it propagates through the tissue. It loses weight as it passes through each layer where it either gets absorbed or transmitted. It loses weight as it moves across the tissue and some of the photons might hit the sample of interest and can get reflected, transmitted or absorbed [155, 162, 163]. To obtain a high-resolution image millions of photons are launched from the light source simultaneously and at the end of it, the light absorbance by the sample is calculated. This is reconstructed to form the absorbance map. These are equivalent to the photoacoustic images that are obtained from the imaging systems.

While building a MC simulation model, a large number of photons are modelled to propagate through the simulation medium (tissue). While passing through any medium photons undergo either reflection or refraction or absorption or scattering or a combination of these. The path that the photon takes is determined by the optical properties of the medium such as refractive index (n), absorption coefficient ($\mu_a$), scattering coefficient ($\mu_s$), and scattering anisotropy (g). Absorption coefficient ($\mu_a$) of a sample can be defined as the probability absorption by the photon in the medium per unit (infinitesimal) path length. This physical quantity is measured by Beer's law. Similarly, scattering coefficient ($\mu_s$) can be defined as the probability scattering of light in a medium per unit (infinitesimal) path length. Scattering anisotropy (g) is defined as the mean of the cosine of the scattering angle. In biological tissues, the typical values for the various optical parameters are as follows, $\mu_a = \sim 0.1$ cm$^{-1}$, $\mu_s = \sim 100$ cm$^{-1}$, g = 0.9, and n = 1.4. The flow chart for MC for an embedded sphere as object is shown in Fig. 8.

The images generated from the monte carlo simulation of light propagation through tissues can be used to train the neural networks. An example of absorbance maps generated from monte carlo simulation for a spherical object is shown in Fig. 9. For training networks on artifact detection and correction, monte carlo simulations can generate artifacts on the images as well. Many different types of artifacts can
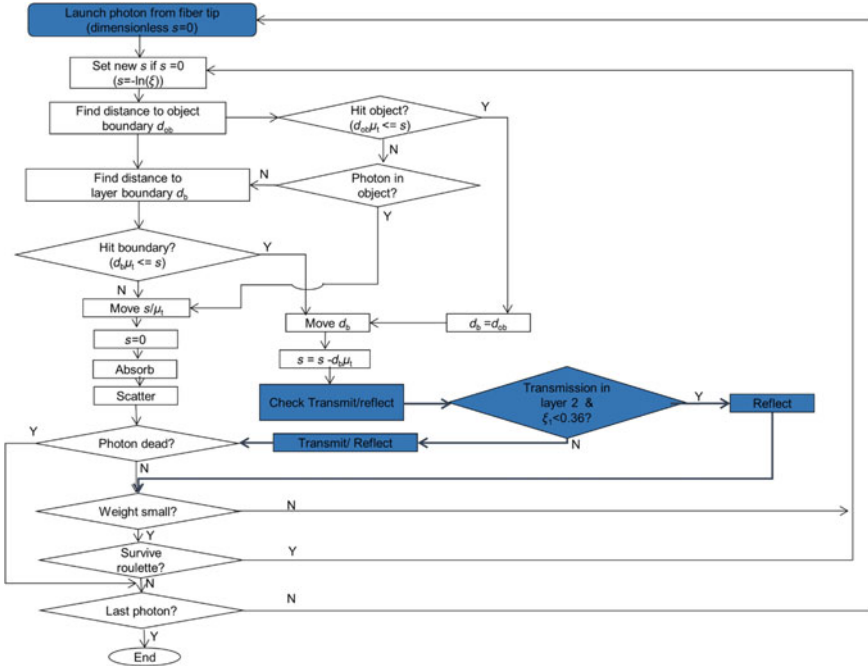
**Fig. 8** Flow chart of Monte Carlo with embedded sphere (MCES). Reprinted with permission from Ref. [164]
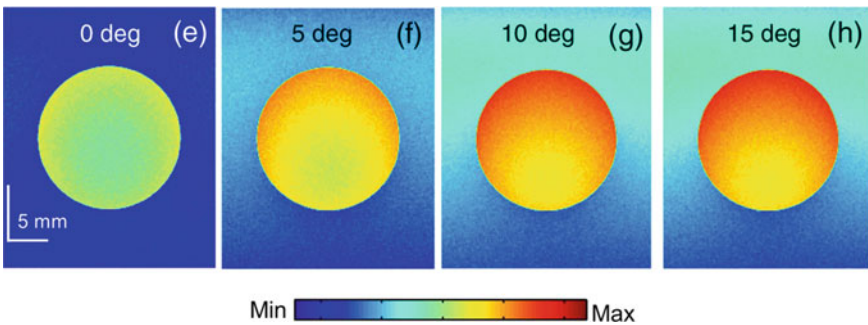


**Fig. 9 e–h** The absorbance maps of sphere at depth 0.5 cm for illumination angles 0°, 5°, 10°, and 15°. Reprinted with permission from Ref. [164]

be generated in the images to train the neural networks model appropriately. Images from the simulation can also be used to test a neural network. The major advantage of using monte carlo simulation for photoacoustic imaging is that a large amount of training data can be obtained very easily, and the data can be customized based on the problem.

# 4  Applications of Deep Learning in Photoacoustic Imaging

There are two ways in which deep learning can be applied for photoacoustic imaging. First, is during the image reconstruction process itself. When the traditional photoacoustic reconstruction techniques are being used, deep learning algorithms can be used on the raw data during reconstruction process to make the images better by reducing artifacts. Second, as a post-processing step after image reconstruction. In this case the traditionally reconstructed photoacoustic images are passed through deep learning algorithms to reduce the artifacts. Examples of both of these methods will be discussed in the following sections.

Photoacoustic signals collected at the boundary of a tissue surface and are most often band limited. In a recent work, in an attempt to improve the bandwidth of the photoacoustic signal detected from the sample, a deep neural network was proposed. Using the neural network would help in improving the quantitative accuracy of the reconstructed PA images. A least square-based deconvolution method which involves the Tikhonov regularization framework was used for comparison with the proposed network. The proposed deep learning method was evaluated with numerical and experimental data as well.

The network proposed contains five fully connected layers, out of the five, one layer is the input layer and one other layer is the output layer. The rest of the three layers are hidden layers. The architecture of this network very similar to that of the decoder network. Three different numerical phantoms (different from the training data) were used to evaluate the performance of the network: (a) a blood vessel network is frequently used as PA numerical phantom for imaging blood vasculature, (b) Derenzo phantom containing different sizes of circular distribution of pressure, and (c) PAT phantom to simulate sharp edges. The bandwidth enhancement using the proposed neural network can be evidently observed from the images as shown in Fig. 10. Here, frequency response of the signal calculated using the proposed neural networks was very similar to full bandwidth signal response. These results indicate that the proposed method using neural networks are capable of enhancing detected PA signal's bandwidth [165]. This further improves the contrast recovery and quality of reconstructed PA images without increasing any computational complexity significantly.

Another example of using deep learning is photoacoustic imaging for artifact reduction is discussed here. In one of the recent works, a novel technique with the help of a deep learning neural network which are trained layer-by-layer to reconstruct 3D photoacoustic images with high resolution was proposed. This network incorporates the physical model into the reconstruction procedure to iteratively reduce artefacts [74]. In this method a U-Net was used to post process data from direct reconstruction, the limitation of using neural networks for post processing is that the result from the neural networks are highly dependent on the quality of the initially reconstructed photoacoustic image. The U-Net is one of the commonly used deep neural network for image denoising, it is a state-of-the-art deep learning technique. The U-Net consists of equal number of contracting and expansive layers. In this network, the number of
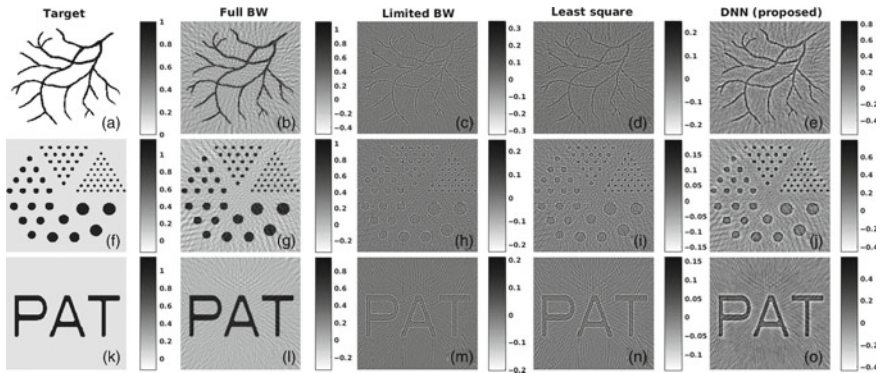
**Fig. 10** Numerical phantoms used for evaluation: **a** blood vessel network, **f** Derenzo phantom, and **k** PAT phantom. Reconstructed (backprojected) initial pressure images with 100 detectors using **b**, **g**, **l** full BW signal, **c**, **h**, **m** limited BW signal, **d**, **i**, **n** predicted signal from least square deconvolution method, and **e**, **j**, **o** predicted signal from the proposed DNN. The SNR of the data is at 40 dB. Reprinted with permission from Ref. [165]

feature channels is the same in the first and last layer, similarly, the number of feature channels in the second layer is two times the first layer and the same is true for the second-last layer. The resolution is being halved in of each contraction step and gets doubled in each of the expansion step. Every single layer in the neural network has a large number of feature channels, this aspect of it allows the propagation of context information to higher resolution layers in the network. Because of this the network assumes a symmetry and providing a U-shape architecture. For the down sampling, the contracting layers of the network consists of unpadded convolutions which are followed by rectified linear units and a pooling operation. Number of feature channels gets doubled in each of the down sampling step [166]. The expansive layers consist of up sampling feature map followed by an up-convolution where the number of feature channels gets halved, each of this is followed by a rectified linear unit. Owing to the high complexity of the photoacoustic forward operator, the training and computation of the gradient information was separated. This network used data from a set of segmented vessels from lung computed tomography scans for training and testing. The network was then applied to in-vivo photoacoustic data measurement.

Use of directly reconstructed images on the neural networks to remove artifacts is a valid approach in many applications, specifically if the goal is to achieve fast and real-time reconstructions. This approach only needs an initial direct reconstruction and one application of the trained network. In the case of a full-view data, this is a promising approach, but it has been demonstrated that even with limited-view images this technique performs very well. A comparison of DGD and U-Net for simulated data is shown in Fig. 11 (top row). The final image is cleaned up and many vessels are properly reconstructed although, some of the minor details are missing in the image and could not be recovered from the initially reconstructed data. The difference to the true target is also shown in Fig. 11 (bottom row). The
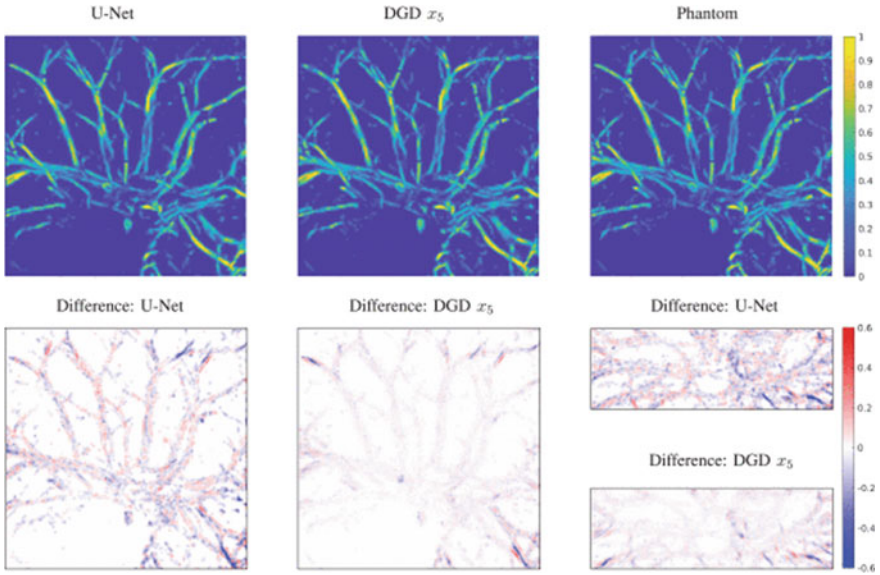
**Fig. 11** Comparison of reconstructions for a test image from the segmented CT data. Left: top and bottom shows the result by applying U-Net to the initialization x0 and the difference to the phantom, maximal value of difference is 0.6012. Middle: shows the result of the DGD after 5 iterations and the difference to the phantom, maximal value of difference is 0.4081. Right bottom: difference images as side projections for the results of DGD and U-Net. Reprinted with permission from Ref. [74]

differences are most pronounced in the outer parts of the domain as a consequence of the limited view geometry. In comparison the reconstruction by DGD has a much smaller overall error, but this is especially true in the center of the domain. The maximal error of the U-net reconstruction is 0.6012 (on the scale of [0, 1]) and of the DGD reconstruction 0.4081 as can be observed form Fig. 12. In conclusion we can say that the U-net architecture performs very well and is even capable of removing
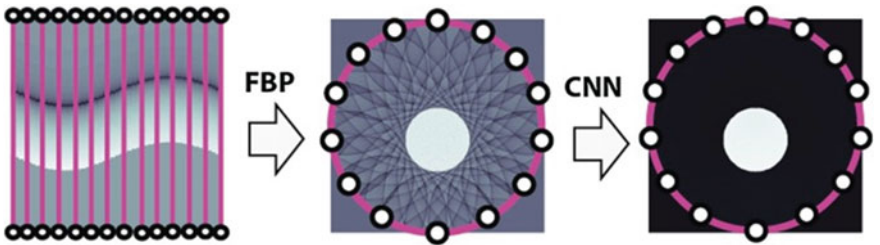


**Fig. 12** Illustration of the proposed network for PAT image reconstruction. In the first step, the FBP algorithm is applied to the sparse data. In a second step, a deep CNN is applied to the intermediate reconstruction which outputs an almost artefact-free image. Reprinted with permission from Ref. [150]

some limited-view artifacts but is ultimately limited by the information contained in the initial reconstruction.

In another work, deep learning approach was used for photoacoustic imaging from sparse data. In this approach, linear reconstruction algorithm was first applied to the sparsely sampled data and the results were further applied to a CNN with weights adjusted based on the training data set. Evaluation of the neural networks is non-iterative process and it takes similar numerical effort as a traditional FBP algorithm for photoacoustic imaging. This approach consists of two steps: In the first step, a linear image reconstruction algorithm was applied to the photoacoustic images, this method provides an approximate result of the original sample including under-sampling artifacts. In the next step, a deep CNN is applied for mapping the intermediate reconstruction to form an artifact-free end image.

The neural network is first trained using simulated ellipse shaped phantoms samples. 1000 pairs of images were generated and used for training. One part of the training data includes pressure data without any noise and the second part of the data random noise was introduced to the simulated pressure data. The neural network was evaluated on similar simulated images of ellipse samples which was not introduced to the network during training. The network performed well by eliminating all the artifacts from the test images. The network was further tested on Shepp-Logan type phantoms and as expected the network was not able to remove all the artifacts from the image as it was not trained on this data [167]. Hence, additional CNNs were trained on 1000 randomly generated ellipse phantoms and 1000 randomly generated Shepp–Logan type phantoms. The newly retrained network was once again tested on the Shepp-Logan type phantoms. It is evident from the images in Fig. 13 that when the neural network is trained with appropriate and correct training data, the performance of the neural networks improves significantly.

## 4.1 Deep Learning for LED Based Photoacoustic Imaging

As discussed earlier, the image quality of the LED based photoacoustic imaging system is not great. To improve the image resolution, improve artifact removal and reduce the averaging for these images would greatly help in the clinical use of this system. Deep learning can be applied to the photoacoustic images from LED systems to improve the overall system efficiency. One of the recent works uses deep neural networks-based performance improvement of the system for improving the quality of the images and also to reduce the average scanning time (averaging) of LED-based PA images. The proposed architecture of the neural networks consists of two important components; the first is a CNN which is used for the spatial feature extraction, and the second one is the recurrent neural networks (RNN) to leverage the temporal information from the PA images. RNN is a form of neural networks in which the output of each step is fed as input to the next step. It varies from the traditional neural networks in the sense that, in CNNs the input and output through different steps are independent of each other. The most unique and important feature of the
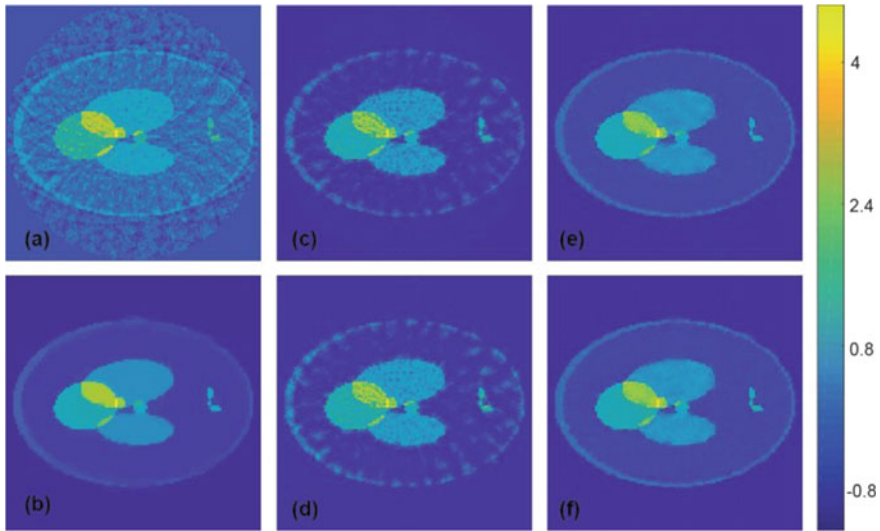
**Fig. 13** Reconstruction results for a Shepp–Logan type phantom from data with 2% Gaussian noise added. **a** FBP reconstruction; **b** reconstruction using TV minimization; **c** proposed CNN using wrong training data without noise added; **d** proposed CNN using wrong training data with noise added; **e** proposed CNN using appropriate training data without noise added; **f** proposed CNN using appropriate training data with noise added. Reprinted with permission from Ref. [150]

RNN is the hidden state, which helps the network to remember the information about a sequence. The neural networks are built based on the state-of-the-art algorithm of densenet-based architecture which uses a series of skip-connections to enhance the image content. For the RNN component, convolutional variant of short-long-term-memory was used to make use of the temporal dependencies in a given PA image sequence. Skip connections was introduced in the both the networks, CNN and RNN for effective feature propagation and elimination of vanishing gradient.

Figure 14a shows the densenet-based CNN architecture. The neural network accepts a low-quality PA image as input and as output generates high quality PA image. The number of feature maps are shown in Fig. 14. The architecture of the network consists of three dense blocks, where each dense block consists of two convolutional layers followed by a ReLU. One of the major advantages of using the dense convolutional layer is that it utilizes all the generated features from previous layers as inputs through skip connections. This enables the propagation of features more effectively through the network which leads to the elimination of the vanishing gradient problem. Finally, to obtain the output image, all the features from the dense blocks are concatenated, a single convolution with one feature map is performed at the end.

In order to train the network experimental study was done using the LED based photoacoustic system and evaluate the performance. The experiment including
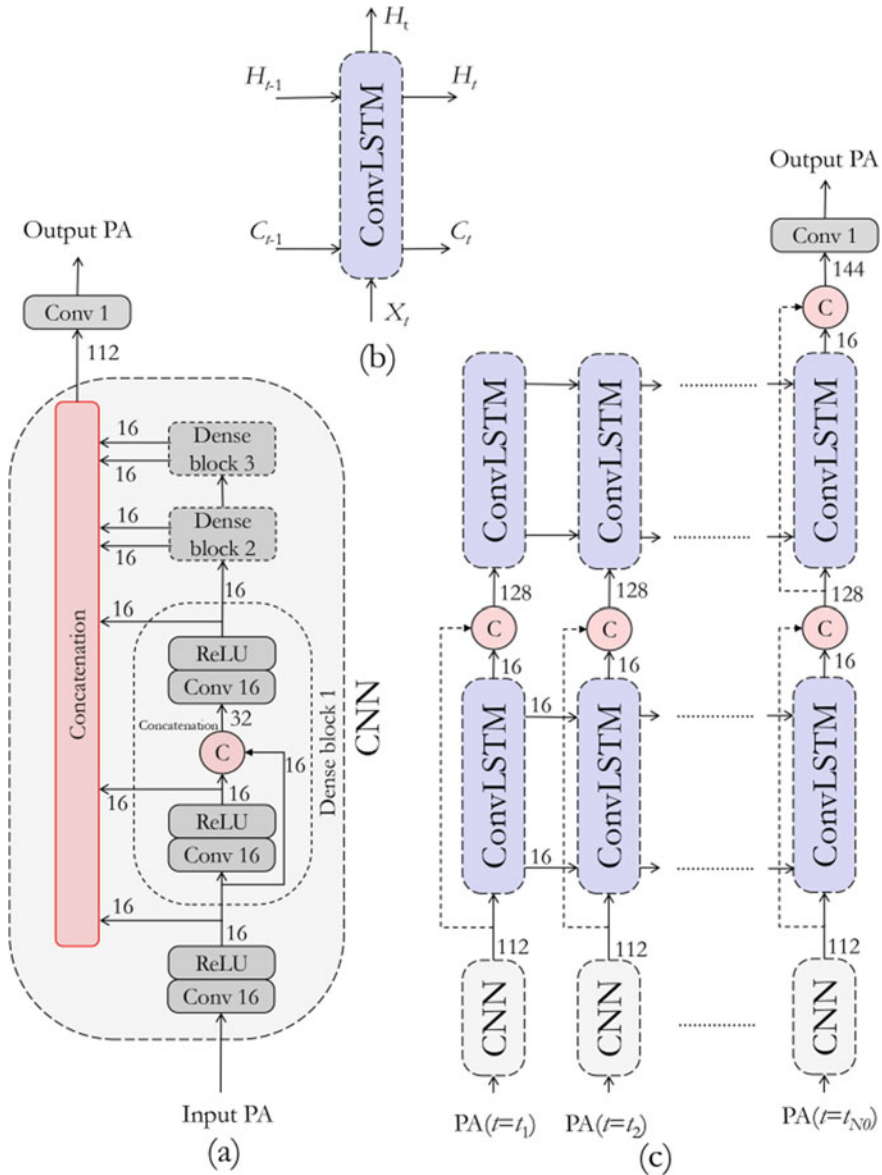
**Fig. 14** A schematic of the neural network. **a** The densenet-based CNN architecture to improve the quality of a single PA image. **b** A schematic of ConvLSTM cell. In addition to current input $X_t$, it exploits previous hidden and cell states to generate current states. **c** The architecture that integrates CNN and ConvLSTM together to extract the spatial features and the temporal dependencies, respectively. Reprinted with permission from Ref. [168]

acquiring images from phantoms and also in vivo human fingers. For the phantom experiments, PA signal was acquired for a time period of 11 s leading to the generation of 11,000 frames of pre-beamformed signals. To obtain a noise free image through averaging having a steady set up without any motion is critical. This is possible with phantoms whereas, maintaining a steady position for in-vivo imaging is very challenging therefore was only done for 5 s. After data acquisition, PA signals were averaged over certain number of frames, followed by beamforming using delay-and-sum technique, subsequently detecting the envelope to reconstruct the PA image.

Two different types of phantoms were used in this study, wire and magnetic nanoparticle phantoms because of their high optical absorption coefficients. For the wire phantom, a total of 62 sets of PA data from 62 different image planes was acquired, and each of the data set consists of a total of 11,000 frames. The phantom was built with fives cylindrical tubes that are placed at multiple depths. The tubes were varied in concentration and depth to perform a comprehensive evaluation of the performance of the neural networks. This helps in evaluating the sensitivity of the system at various depths. Tubes 1–3 were of same concentration but placed in decreasing depths. At the maximum depth along with tube 3, tubes 4 and 5 were placed with decreasing concentrations. For the phantom experiment suing nanoparticle tubes, a total of 10 sets of PA data was acquired from 10 different image planes.

For effective training of the neural networks, different qualities of input PA images was used. As stated previously, greater the averaging, better is the image quality and resolution. This aspect was made use of to obtain images of varying quality. The averaging works well with phantom data as the the three is not motion artifacts involved. The number of frames to be averaged (N) was chosen from a range starting with very low value and increased to the highest possible value (11,000).

Figure 15 depicts the photoacoustic images from the two different phantoms and in-vivo human finger here. The performance of the various networks can be clearly observed from the difference in the quality of images from the different neural networks for all the different samples and at various depths. This work is an example of how a neural network can be trained on very simple data that can be easily acquired to improve the image quality and reduce the scanning time for image acquisition.

## 5   Limitations of Deep Learning

Deep learning has been very successful in the recent times for a variety of applications. In spite of its success, there are many limitations associated with the application of the technique. Firstly, deep learning is not the best machine learning technique for all the different types of data analysis problems. For various issues in which the data is already well structured or if optimal features are well-defined, instead of deep learning a lot of other simple machine learning methods like logistic regression, support vector machines, and random forests can be applied to solve it. It will be much
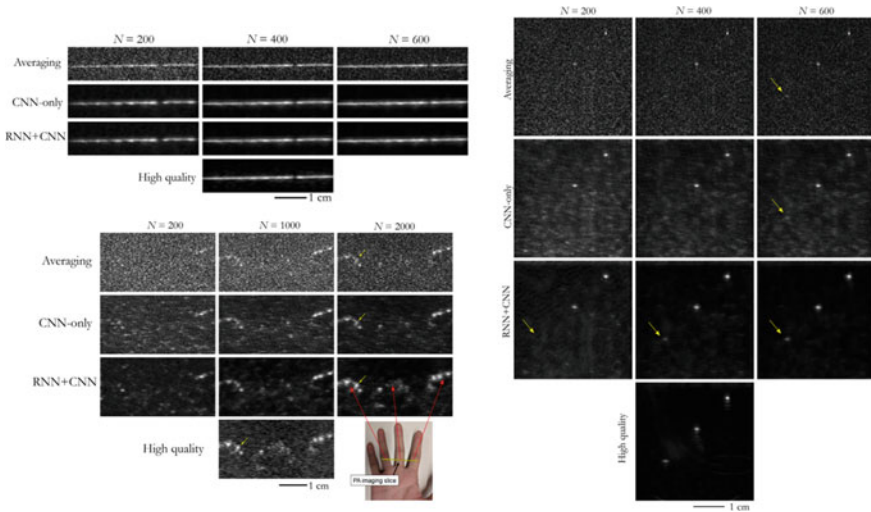
**Fig. 15** Qualitative comparison of our method with the simple averaging and CNN-only techniques for a wire phantom, in vivo example. The in vivo data consists of proper digital arteries of three fingers of a volunteer. Example effect of depth on the PA image quality on nanoparticles. Reprinted with permission from Ref. [168]

easier to apply and are also usually more effective with such datasets. CNNs have become very dominant in the field of computer vision, there are some limitations there as well. One of the most significant limitation is that deep learning is a technology that requires a large amount of data; for the network to learn the weights from scratch for a large network requires a huge number of labeled examples to achieve accurate classification. Deep learning scales very well with large datasets. Therefore, computing resources, time needed for training a deep learning model is very high. Also, obtaining so much of labelled training data is very difficult.

Transfer learning is receiving more research for moving to an effective way of reducing the data requirements. In recent transfer learning approaches, it reuses weights from networks trained on ImageNet (a labeled collection of low-resolution 2D color images). For most applications in radiology, higher-resolution volumetric images are required, for which pretrained networks are not yet available. As a result, creating a large labelled medical image library is really important step for further progress in applying deep learning, which is not easy due to cost, privacy etc. Also, with more future breakthroughs in deep learning, data requirements can be significantly reduced for training of deep learning systems.

# 6 Future Directions for Deep Learning

Deep learning models has shown expert-level or better performance at few tasks. Deep learning algorithms are capable of extracting or identifying more features than humans. Data availability and curation of data into repositories is becoming more organized now for better handling and usage of data. This will further help in developing better models for deep learning as there will be more availability of a variety of training data including different scenarios. In the recent past there have been approaches where they use data from one imaging modality to train a network for better performance on another imaging modality. This will help in boosting the performance of the neural networks as they train on better ground truth images. The importance of deep learning will keep increasing in the days to come in the hospitals.

For photoacoustic imaging, deep learning will have a more important role to play. Deep learning for photoacoustic is not much explored till now, so the potential of it has fully not been understood. Some of the major areas in which deep learning can be used for photoacoustic in general and LED based photoacoustic systems also includes better and faster reconstruction algorithms, reduction of artifacts in images, reduction in averaging to produce a high resolution image, decreasing the data acquisition time, possibility of reducing the laser power used for image acquisition and lesser exposure time. Further research in all the above-mentioned area will greatly improve the performance of photoacoustic imaging system and may make the clinical translation and utilization of photoacoustic for diagnosis and real-time monitoring more feasible in the near future.

# 7 Conclusion

In this chapter we discussed the limitations of the current image reconstruction and denoising techniques in photoacoustic imaging. The basic concepts of machine learning and artificial intelligence was established with a focus on deep learning. The applications of deep learning in various medical imaging techniques was discussed. Based on this, the use of deep learning in photoacoustic imaging was analysed especially for improvement in areas of image reconstruction, image denoising and image resolution. Although, deep learning has a lot of potential applications for improving photoacoustic imaging, it comes with certain limitations, especially in terms of training data. Upon overcoming the limitations, deep learning will definitely help in clinical translation and utilization for various clinical applications in the near future.

Next section of this book will focus on preclinical imaging applications and early clinical pilot studies using LED-based photoacoustics.

# References

1. P. Beard, Biomedical photoacoustic imaging. Interface Focus **1**, 602–631 (2011)
2. M. Erfanzadeh, Q. Zhu, Photoacoustic imaging with low-cost sources: a review. Photoacoustics **14**, 1–11 (2019)
3. P.K. Upputuri, M. Pramanik, Recent advances toward preclinical and clinical translation of photoacoustic tomography: a review. J. Biomed. Opt. **22**(4), 1–19 (2016)
4. V. Ntziachristos et al., Looking and listening to light: the evolution of whole-body photonic imaging. Nat. Biotechnol. **23**(3), 313–320 (2005)
5. S. Manohar, D. Razansky, Photoacoustics: a historical review. Adv. Opt. Photonics **8**(4), 586–617 (2016)
6. M. Xu, L.V. Wang, Photoacoustic imaging in biomedicine. Rev. Sci. Instrum. **77**(4), 041101 (2006)
7. J.L. Su et al., Advances in clinical and biomedical applications of photoacoustic imaging. Expert Opin. Med. Diagn. **4**(6), 497–510 (2010)
8. I. Steinberg et al., Photoacoustic clinical imaging. Photoacoustics **14**, 77–98 (2019)
9. J. Weber, P.C. Beard, S.E. Bohndiek, Contrast agents for molecular photoacoustic imaging. Nat. Methods **13**, 639 (2016)
10. Q. Fu et al., Photoacoustic imaging: contrast agents and their biomedical applications. Adv. Mater. **31**(6), 1805875 (2019)
11. S.W. Yoo et al., Biodegradable contrast agents for photoacoustic imaging. Appl. Sci. **8**(9), 1567 (2018)
12. P.K. Upputuri, M. Pramanik, Photoacoustic imaging in the second near-infrared window: a review. J. Biomed. Opt. **24**(4), 1–20 (2019)
13. G. Xu et al., Photoacoustic spectrum analysis for microstructure characterization in biological tissue: analytical model. Ultrasound Med. Biol. **41**(5), 1473–1480 (2015)
14. Y. Wang et al., Toward in vivo biopsy of melanoma based on photoacoustic and ultrasound dual imaging with an integrated detector. Biomed. Opt. Express **7**(2), 279–286 (2016)
15. P. Hai et al., High-throughput, label-free, single-cell photoacoustic microscopy of intratumoral metabolic heterogeneity. Nat. Biomed. Eng. **3**(5), 381–391 (2019)
16. H.D. Lee et al., Label-free photoacoustic microscopy for in-vivo tendon imaging using a fiber-based pulse laser. Sci. Rep. **8**(1), 4805 (2018)
17. M. Mathiyazhakan et al., In situ synthesis of gold nanostars within liposomes for controlled drug release and photoacoustic imaging. Sci. China Mater. **59**(11), 892–900 (2016)
18. W. Li, X. Chen, Gold nanoparticles for photoacoustic imaging. Nanomedicine **10**(2), 299–320 (2015)
19. D. Wu et al., Contrast agents for photoacoustic and thermoacoustic imaging: a review. Int. J. Mol. Sci. **15**(12), 23616–23639 (2014)
20. M. Pramanik et al., In vivo photoacoustic (PA) mapping of sentinel lymph nodes (SLNs) using carbon nanotubes (CNTs) as a contrast agent, in *Proc SPIE*, San Francisco (2009)
21. C.-W. Wei et al., In vivo photoacoustic imaging with multiple selective targeting using bioconjugated gold nanorods, in *Proc SPIE* (SPIE, Bellingham, 2008)
22. D. Das et al., On-chip generation of microbubbles in photoacoustic contrast agents for dual modal ultrasound/photoacoustic in vivo animal imaging. Sci. Rep. **8**(1), 6401 (2018)
23. K. Sivasubramanian et al., Near-infrared light-responsive liposomal contrast agent for photoacoustic imaging and drug release applications. J. Biomed. Opt. **22**(4), 041007 (2016)
24. G.P. Luke et al., Silica-coated gold nanoplates as stable photoacoustic contrast agents for sentinel lymph node imaging. Nanotechnology **24**(45), 455101 (2013)
25. W. Lu et al., Photoacoustic imaging of living mouse brain vasculature using hollow gold nanospheres. Biomaterials **31**(9), 2617–2626 (2010)
26. R. Ansari et al., All-optical forward-viewing photoacoustic probe for high-resolution 3D endoscopy. Light Sci. Appl. **7**(1), 75 (2018)
27. J.-M. Yang et al., Photoacoustic endoscopy. Opt. Lett. **34**(10), 1591–1593 (2009)

28. L.V. Wang, Multiscale photoacoustic microscopy and computed tomography. Nat. Photonics **3**(9), 503 (2009)
29. D. Cai et al., Dual-view photoacoustic microscopy for quantitative cell nuclear imaging. Opt. Lett. **43**(20), 4875–4878 (2018)
30. H. Zhang et al., Functional photoacoustic microscopy for high-resolution and noninvasive in vivo imaging. Nat. Biotechnol. **24**, 848 (2006)
31. S. Jeon et al., Review on practical photoacoustic microscopy. Photoacoustics **15**, 100141 (2019)
32. W. Liu, J. Yao, Photoacoustic microscopy: principles and biomedical applications. Biomed. Eng. Lett. **8**(2), 203–213 (2018)
33. L. Li et al., Single-impulse panoramic photoacoustic computed tomography of small-animal whole-body dynamics at high spatiotemporal resolution. Nat. Biomed. Eng. **1**, 0071 (2017)
34. A. Horiguchi et al., Pilot study of prostate cancer angiogenesis imaging using a photoacoustic imaging system. Urology **108**, 212–219 (2017)
35. P.K. Upputuri et al., A high-performance compact photoacoustic tomography system for in vivo small-animal brain imaging. J. Visualized Exp. **124**, e55811 (2017)
36. S. Wang et al., Recent advances in photoacoustic imaging for deep-tissue biomedical applications. Theranostics **6**(13), 2395 (2016)
37. J. Kim et al., Programmable real-time clinical photoacoustic and ultrasound imaging system. Sci. Rep. **6**, 35137 (2016)
38. R. Li et al., Assessing breast tumor margin by multispectral photoacoustic tomography. Biomed. Opt. Express **6**(4), 1273–1281 (2015)
39. A. Garcia-Uribe et al., Dual-modality photoacoustic and ultrasound imaging system for non-invasive sentinel lymph node detection in patients with breast cancer. Sci. Rep. **5**, 15748 (2015)
40. E.I. Galanzha, V.P. Zharov, Circulating tumor cell detection and capture by photoacoustic flow cytometry in vivo and ex vivo. Cancers **5**(4), 1691–1738 (2013)
41. M. Jeon, J. Kim, C. Kim, Photoacoustic cystography. J. Visualized Exp. **76**, e50340 (2013)
42. M. Pramanik, L.V. Wang, Thermoacoustic and photoacoustic sensing of temperature. J. Biomed. Opt. **14**(5), 054024 (2009)
43. P.K. Upputuri et al., Recent developments in vascular imaging techniques in tissue engineering and regenerative medicine. BioMed Res. Int. **2015** (2015)
44. K. Sivasubramanian et al., Hand-held, clinical dual mode ultrasound—photoacoustic imaging of rat urinary bladder and its applications. J. Biophotonics **11**, e201700317 (2018)
45. X. Wang et al., Noninvasive laser-induced photoacoustic tomography for structural and functional in vivo imaging of the brain. Nat. Biotechnol. **21**(7), 803–806 (2003)
46. K. Sivasubramanian, M. Pramanik, High frame rate photoacoustic imaging at 7000 frames per second using clinical ultrasound system. Biomed. Opt. Express **7**(2), 312–323 (2016)
47. A. Hariri et al., The characterization of an economic and portable LED-based photoacoustic imaging system to facilitate molecular imaging. Photoacoustics **9**, 10–20 (2018)
48. Y. Zhu et al., Light emitting diodes based photoacoustic imaging and potential clinical applications. Sci. Rep. **8** (2018)
49. T.J. Allen, P.C. Beard, High power visible light emitting diodes as pulsed excitation sources for biomedical photoacoustics. Biomed. Opt. Express **7**(4), 1260–1270 (2016)
50. S.K. Kalva, M. Pramanik, Experimental validation of tangential resolution improvement in photoacoustic tomography using a modified delay-and-sum reconstruction algorithm. J. Biomed. Opt. **21**(8), 086011 (2016)
51. R.A. Kruger et al., Photoacoustic ultrasound (PAUS)—reconstruction tomography. Med. Phys. **22**(10), 1605–1609 (1995)
52. Z. Deng, W. Li, C. Li, Slip-ring-based multi-transducer photoacoustic tomography system. Opt. Lett. **41**(12), 2859–2862 (2016)
53. S.K. Kalva, Z.Z. Hui, M. Pramanik, Calibrating reconstruction radius in a multi single-element ultrasound-transducer-based photoacoustic computed tomography system. J. Opt. Soc. Am. A **35**(5), 764–771 (2018)

54. R.A. Kruger et al., Photoacoustic ultrasound (PAUS)—reconstruction tomography. Med. Phys. **22**, 1605 (1995)
55. C. Huang et al., Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media. IEEE Trans. Med. Imaging **32**, 1097 (2013)
56. H. Jiang, Z. Yuan, X. Gu, Spatially varying optical and acoustic property reconstruction using finite-element-based photoacoustic tomography. J. Opt. Soc. Am. A **23**, 878 (2006)
57. M. Xu, Y. Xu, L.V. Wang, Time-domain reconstruction-algorithms and numerical simulations for thermoacoustic tomography in various geometries. IEEE Trans. Biomed. Eng. **50**, 1086 (2003)
58. P. Omidi et al., A novel dictionary-based image reconstruction for photoacoustic computed tomography. Appl. Sci. **8**(9), 1570 (2018)
59. J. Wang, Y. Wang, Photoacoustic imaging reconstruction using combined nonlocal patch and total-variation regularization for straight-line scanning. BioMed. Eng. OnLine **17**(1), 105 (2018)
60. M. Mozaffarzadeh et al., The double-stage delay-multiply-and-sum image reconstruction method improves imaging quality in a LED-based photoacoustic array scanner. Photoacoustics **12**, 22–29 (2018)
61. M. Mozaffarzadeh, A. Mahloojifar, M. Orooji, Medical photoacoustic beamforming using minimum variance-based delay multiply and sum, in *SPIE Digital Optical Technologies*, vol. 10335 (SPIE, Bellingham, 2017)
62. M. Mozaffarzadeh et al., Linear-array photoacoustic imaging using minimum variance-based delay multiply and sum adaptive beamforming algorithm. J. Biomed. Opt. **23**(2), 1–15 (2018)
63. H. Huang et al., An adaptive filtered back-projection for photoacoustic image reconstruction. Med. Phys. **42**(5), 2169–2178 (2015)
64. Z. Ren, G. Liu, Z. Huang, Filtered back-projection reconstruction of photo-acoustic imaging based on an modified wavelet threshold function, in *International Symposium on Optoelectronic Technology and Application 2016*, vol. 10155 (SPIE, Bellingham, 2016)
65. C. Zhang, Y. Zhang, Y. Wang, A photoacoustic image reconstruction method using total variation and nonconvex optimization. BioMed. Eng. OnLine **13**(1), 117 (2014)
66. E. Bossy et al., Time reversal of photoacoustic waves. Appl. Phys. Lett. **89**(18), 184108 (2006)
67. B.T. Cox, B.E. Treeby, Artifact trapping during time reversal photoacoustic imaging for acoustically heterogeneous media. IEEE Trans. Med. Imaging **29**(2), 387–396 (2010)
68. X. Minghua, X. Yuan, L.V. Wang, Time-domain reconstruction algorithms and numerical simulations for thermoacoustic tomography in various geometries. IEEE Trans. Biomed. Eng. **50**(9), 1086–1099 (2003)
69. Y. Xu, L.V. Wang, Time reversal and its application to tomography with diffracting sources. Phys. Rev. Lett. **92**, 033902 (2004)
70. M.A. Lediju Bell, J. Shubert, Photoacoustic-based visual servoing of a needle tip. Sci. Rep. **8**(1), 15519 (2018)
71. H.N.Y. Nguyen, A. Hussain, W. Steenbergen, Reflection artifact identification in photoacoustic imaging using multi-wavelength excitation. Biomed. Opt. Express **9**(10), 4613–4630 (2018)
72. M.K.A. Singh, W. Steenbergen, Photoacoustic-guided focused ultrasound (PAFUSion) for identifying reflection artifacts in photoacoustic imaging. Photoacoustics **3**(4), 123–131 (2015)
73. M.K.A. Singh et al., Reflection-artifact-free photoacoustic imaging using PAFUSion (photoacoustic-guided focused ultrasound), in *SPIE BiOS*, vol. 9708 (SPIE, Bellingham, 2016)
74. A. Hauptmann et al., Model-based learning for accelerated, limited-view 3-D photoacoustic tomography. IEEE Trans. Med. Imaging **37**(6), 1382–1393 (2018)
75. M. Tabei, T.D. Mast, R.C. Waag, A k-space method for coupled first-order acoustic propagation equations. J. Acoust. Soc. Am. **111**(1), 53 (2002)
76. Y. Adachi, T. Hoshimiya, Photoacoustic imaging with multiple-wavelength light-emitting diodes. Jpn. J. Appl. Phys. **52**(7S), 07HB06 (2013)

77. Y. Zhu et al., LED-based photoacoustic imaging for monitoring angiogenesis in fibrin scaffolds. Tissue Eng. Part C Methods **25**(9), 523–531 (2019)
78. J. Jo et al., Detecting joint inflammation by an LED-based photoacoustic imaging system: a feasibility study. J. Biomed. Opt. **23**(11), 1–4 (2018)
79. J. Leskinen et al., Photoacoustic tomography setup using LED illumination, in *European Conferences on Biomedical Optics*, vol. 11077 (SPIE, Bellingham, 2019)
80. P.K. Upputuri et al., Pulsed laser diode photoacoustic tomography (PLD-PAT) system for fast in vivo imaging of small animal brain, in *SPIE BiOS*, vol. 10064 (SPIE, Bellingham, 2017)
81. P.K. Upputuri, M. Pramanik, Performance characterization of low-cost, high-speed, portable pulsed laser diode photoacoustic tomography (PLD-PAT) system. Biomed. Opt. Express **6**(10), 4118–4129 (2015)
82. D. Crevier, *AI: The Tumultuous History of the Search for Artificial Intelligence* (Basic Books, Inc., New York, 1993), p. 386
83. O. Vinyals, M. Fortunato, N. Jaitly, *Pointer Networks*. arXiv e-prints (2015)
84. W. Samek, T. Wiegand, K.-R. Müller, *Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models*. arXiv e-prints (2017)
85. H. Chen et al., The rise of deep learning in drug discovery. Drug Discovery Today **23**(6), 1241–1250 (2018)
86. H. Chen et al., Low-dose CT via convolutional neural network. Biomed. Opt. Express **8**(2), 679–694 (2017)
87. J.X. Wang et al., *Learning to Reinforcement Learn*. arXiv e-prints (2016)
88. F. Isensee et al., *Brain Tumor Segmentation Using Large Receptive Field Deep Convolutional Neural Networks* (Springer, Berlin, 2017)
89. S. Albelwi, A. Mahmood, A framework for designing the architectures of deep convolutional neural networks. Entropy **19**(6), 242 (2017)
90. G.I. Parisi et al., Continual lifelong learning with neural networks: a review. Neural Networks **113**, 54–71 (2019)
91. J. Schmidhuber, Deep learning in neural networks: an overview. Neural Networks **61**, 85–117 (2015)
92. O.I. Abiodun et al., State-of-the-art in artificial neural network applications: a survey. Heliyon **4**(11), e00938 (2018)
93. G. Litjens et al., A survey on deep learning in medical image analysis. Med. Image Anal. **42**, 60–88 (2017)
94. J. Zhou et al., *Graph Neural Networks: A Review of Methods and Applications*. arXiv e-prints (2018)
95. X. Hu, W. Yi, L. Jiang, S. Wu, Y. Zhang, J. Du, T. Ma, T. Wang, X. Wu, Classification of metaphase chromosomes using deep convolutional neural network. J. Comput. Biol. **26**(5), 473–484 (2019)
96. A. Kensert, P.J. Harrison, O. Spjuth, Transfer learning with deep convolutional neural networks for classifying cellular morphological changes. SLAS DISCOVERY Advancing Life Sci R&D **24**(4), 466–475 (2019)
97. K.H. Jin et al., Deep convolutional neural network for inverse problems in imaging. IEEE Trans. Image Process. **26**(9) (2017)
98. Y. Ren, X. Cheng, Review of convolutional neural network optimization and training in image processing, in *10th International Symposium on Precision Engineering Measurements and Instrumentation* (*ISPEMI 2018*), vol. 11053 (SPIE, Bellingham, 2019)
99. W. Wang et al., Development of convolutional neural network and its application in image classification: a survey. Opt. Eng. **58**(4), 1–19 (2019)
100. X. Zhou, Understanding the convolutional neural networks with gradient descent and backpropagation. J. Phys. Conf. Ser. **1004**, 012028 (2018)
101. M.C. Belavagi, B. Muniyal, Performance evaluation of supervised machine learning algorithms for intrusion detection. Procedia Comput. Sci. **89**, 117–123 (2016)
102. A.N. Dalrymple et al., A supervised machine learning approach to characterize spinal network function. J. Neurophysiol. **121**(6), 2001–2012 (2019)

103. S. Klassen, J. Weed, D. Evans, Semi-supervised machine learning approaches for predicting the chronology of archaeological sites: a case study of temples from medieval Angkor, Cambodia. PLoS ONE **13**(11), e0205649 (2018)
104. M. Rucco et al., A methodology for part classification with supervised machine learning. Artif. Intell. Eng. Des. Anal. Manuf. **33**(1), 100–113 (2018)
105. H. Yao et al., MSML: a novel multilevel semi-supervised machine learning framework for intrusion detection system. IEEE Internet Things J. **6**(2), 1949–1959 (2019)
106. D.M. Hawkins, The problem of overfitting. J. Chem. Inf. Comput. Sci. **44**(1), 1–12 (2004)
107. S. Salman, X. Liu, *Overfitting Mechanism and Avoidance in Deep Neural Networks*. arXiv e-prints (2019)
108. R. Hingorani, C.L. Hansen, Can machine learning spin straw into gold? J. Nucl. Cardiol. **25**(5), 1610–1612 (2018)
109. M. Ceriotti, Unsupervised machine learning in atomistic simulations, between predictions and understanding. J. Chem. Phys. **150**(15), 150901 (2019)
110. A. Dik, K. Jebari, A. Ettouhami, An improved robust fuzzy algorithm for unsupervised learning. J. Intell. Syst. **29**(1) (2018)
111. M. Usama et al., Unsupervised machine learning for networking: techniques applications and research challenges. IEEE Access **7**, 65579–65615 (2019)
112. S. Becker, Unsupervised learning procedures for neural networks. Int. J. Neural Syst. **02**(01n02), 17–33 (1991)
113. S. Guan et al., Application of unsupervised learning to hyperspectral imaging of cardiac ablation lesions. J. Med. Imaging **5**(4), 1–12 (2018)
114. P. Baldi, Gradient descent learning algorithm overview: a general dynamical systems perspective. IEEE Trans. Neural Networks **6**(1), 182–195 (1995)
115. N. Cui, Applying gradient descent in convolutional neural networks. J. Phys. Conf. Ser. **1004**, 012027 (2018)
116. Q. Mercier, F. Poirion, J.-A. Désidéri, A stochastic multiple gradient descent algorithm. Eur. J. Oper. Res. **271**(3), 808–817 (2018)
117. D. Newton, F. Yousefian, R. Pasupathy, Stochastic gradient descent: recent trends, in *Recent Advances in Optimization and Modeling of Contemporary Problems* (INFORMS, Aliso Viejo, 2018), pp. 193–220
118. S. Ruder, *An Overview of Gradient Descent Optimization Algorithms*. arXiv e-prints (2016)
119. R.A. Jacobs, Increased rates of convergence through learning rate adaptation. Neural Networks **1**(4), 295–307 (1988)
120. D.P. Kingma, J. Ba, *Adam: A Method for Stochastic Optimization*. arXiv e-prints (2014)
121. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. Nature **521**, 436 (2015)
122. D. Masters, C. Luschi, *Revisiting Small Batch Training for Deep Neural Networks*. arXiv e-prints (2018)
123. L. Perez, J. Wang, *The Effectiveness of Data Augmentation in Image Classification using Deep Learning*. arXiv e-prints (2017)
124. S.M. Anwar et al., Medical image analysis using convolutional neural networks: a review. J. Med. Syst. **42**(11), 226 (2018)
125. J. Jiang, P. Trundle, J. Ren, Medical image analysis with artificial neural networks. Comput. Med. Imaging Graph. **34**(8), 617–631 (2010)
126. D. Shen, G. Wu, H.-I. Suk, Deep learning in medical image analysis. Annu. Rev. Biomed. Eng. **19**(1), 221–248 (2017)
127. Z. Shi et al., Survey on neural networks used for medical image processing. Int. J. Comput. Sci. **3**(1), 86–100 (2009)
128. I. Wolf et al., The medical imaging interaction toolkit. Med. Image Anal. **9**(6), 594–604 (2005)
129. T.J. Brinker et al., Skin cancer classification using convolutional neural networks: systematic review. J. Med. Internet Res. **20**(10), e11936 (2018)
130. P. Chang et al., Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas. Am. J. Neuroradiol. **39**(7), 1201 (2018)

131. Y. Guo et al., A review of semantic segmentation using deep neural networks. Int. J. Multimedia Inf. Retrieval **7**(2), 87–93 (2018)
132. G. Wang et al., Interactive medical image segmentation using deep learning with image-specific fine tuning. IEEE Trans. Med. Imaging **37**(7), 1562–1573 (2018)
133. D.A. Ragab et al., Breast cancer detection using deep convolutional neural networks and support vector machines. PeerJ **7**, e6201 (2019)
134. F. Pesapane, M. Codari, F. Sardanelli, Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine. Eur. Radiol. Exp. **2**(1), 35 (2018)
135. L. Saba et al., The present and future of deep learning in radiology. Eur. J. Radiol. **114**, 14–24 (2019)
136. G. Chartrand et al., Deep learning: a primer for radiologists. RadioGraphics **37**(7), 2113–2131 (2017)
137. A. Fourcade, R.H. Khonsari, Deep learning in medical image analysis: a third eye for doctors. J. Stomatology Oral Maxillofac. Surg. **120**(4), 279–288 (2019)
138. S. Soffer et al., Convolutional neural networks for radiologic images: a radiologist's guide. Radiology **290**(3), 590–606 (2019)
139. D. Ueda, A. Shimazaki, Y. Miki, Technical and clinical overview of deep learning in radiology. Jpn. J. Radiol. **37**(1), 15–33 (2019)
140. G. Zaharchuk et al., Deep learning in neuroradiology. Am. J. Neuroradiol. **39**(10), 1776 (2018)
141. P. Rajpurkar et al., Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLoS Med. **15**(11), e1002686 (2018)
142. L.J. Brattain et al., Machine learning for medical ultrasound: status, methods, and future opportunities. Abdom. Radiol. (New York) **43**(4), 786–799 (2018)
143. S. Liu et al., Deep learning in medical ultrasound analysis: a review. Engineering **5**(2), 261–275 (2019)
144. A.S. Becker et al., Classification of breast cancer in ultrasound imaging using a generic deep learning analysis software: a pilot study. Br. J. Radiol. **91**(1083), 20170576 (2018)
145. Q. Huang, F. Zhang, X. Li, Machine learning in ultrasound computer-aided diagnostic systems: a survey. Biomed. Res. Int. **2018**, 5137904 (2018)
146. J. Ma et al., A pre-trained convolutional neural network based method for thyroid nodule diagnosis. Ultrasonics **73**, 221–230 (2017)
147. D. Allman, A. Reiter, M.A.L. Bell, Photoacoustic source detection and reflection artifact removal enabled by deep learning. IEEE Trans. Med. Imaging **37**(6), 1464–1477 (2019)
148. D. Allman, A. Reiter, M.A.L. Bell, A machine learning method to identify and remove reflection artifacts in photoacoustic channel data, in *2017 IEEE International Ultrasonics Symposium (IUS)* (2017)
149. S. Antholzer et al., Photoacoustic image reconstruction via deep learning, in *Photonics West*, USA (2018)
150. S. Antholzer, M. Haltmeier, J. Schwab, Deep learning for photoacoustic tomography from sparse data. Inverse Prob. Sci. Eng. **27**(7), 987–1005 (2019)
151. C. Cai et al., End-to-end deep neural network for optical inversion in quantitative photoacoustic imaging. Opt. Lett. **43**(12), 2752–2755 (2018)
152. N. Davoudi, X.L. Deán-Ben, D. Razansky, Deep learning optoacoustic tomography with sparse data. Nat. Mach. Intell. **1**(10), 453–460 (2019)
153. J. Schwab et al., *Real-Time Photoacoustic Projection Imaging Using Deep Learning* (2018)
154. D. Waibel et al., Reconstruction of initial pressure from limited view photoacoustic images using deep learning, in *SPIE BiOS*, vol. 10494 (SPIE, Bellingham, 2018)
155. S.L. Jacques, Coupling 3D Monte Carlo light transport in optically heterogeneous tissues to photoacoustic signal generation. Photoacoustics **2**(4), 137–142 (2014)
156. V. Periyasamy, M. Pramanik, Monte Carlo simulation of light transport in tissue for optimizing light delivery in photoacoustic imaging of the sentinel lymph node. J. Biomed. Opt. **18**(10), 1–8 (2013)

157. V. Periyasamy, M. Pramanik, Monte Carlo simulation of light transport in turbid medium with embedded object—spherical, cylindrical, ellipsoidal, or cuboidal objects embedded within multilayered tissues. J. Biomed. Opt. **19**(4), 1–10 (2014)
158. V. Periyasamy, M. Pramanik, Advances in Monte Carlo simulation for light propagation in tissue. IEEE Rev. Biomed. Eng. **10**, 122–135 (2017)
159. Y. Liu, Z. Yuan, Monte Carlo simulation predicts deep-seated photoacoustic effect in heterogeneous tissues, in *Biomedical Optics 2016* (Optical Society of America, Fort Lauderdale, Florida, 2016)
160. G.S. Sangha, N.J. Hale, C.J. Goergen, Adjustable photoacoustic tomography probe improves light delivery and image quality. Photoacoustics **12**, 6–13 (2018)
161. A. Sharma et al., Photoacoustic imaging depth comparison at 532-, 800-, and 1064-nm wavelengths: Monte Carlo simulation and experimental validation. J. Biomed. Opt. **24**(12), 1–10 (2019)
162. G. Paltauf, P.R. Torke, R. Nuster, Modeling photoacoustic imaging with a scanning focused detector using Monte Carlo simulation of energy deposition. J. Biomed. Opt. **23**(12), 1–11 (2018)
163. L. Wang, S.L. Jacques, L. Zheng, MCML—Monte Carlo modeling of light transport in multi-layered tissues. Comput. Methods Programs Biomed. **47**(2), 131–146 (1995)
164. K. Sivasubramanian et al., Optimizing light delivery through fiber bundle in photoacoustic imaging with clinical ultrasound system: Monte Carlo simulation and experimental validation. J. Biomed. Opt. **22**(4), 041008 (2016)
165. S. Gutta et al., Deep neural network-based bandwidth enhancement of photoacoustic data. J. Biomed. Opt. **22**(11), 1–7 (2017)
166. O. Ronneberger, P. Fischer, T. Brox, *U-Net: Convolutional Networks for Biomedical Image Segmentation*. arXiv e-prints (2015)
167. L.A. Shepp, B.F. Logan, Reconstructing interior head tissue from X-ray transmissions. IEEE Trans. Nucl. Sci. **21**(1), 228–236 (1974)
168. E.M.A. Anas et al., Enabling fast and high quality LED photoacoustic imaging: a recurrent neural networks based approach. Biomed. Opt. Express **9**(8), 3852–3866 (2018)