Machine learning for medical ultrasound: status, methods, and future opportunities

Laura J. Brattain,[1] Brian A. Telfer,[1] Manish Dhyani,[2,4] Joseph R. Grajo,[3]
Anthony E. Samir[4]

[1]MIT Lincoln Laboratory, 244 Wood St, Lexington, MA 02420, USA
[2]Department of Internal Medicine, Steward Carney Hospital, Boston, MA 02124, USA
[3]Department of Radiology, Division of Abdominal Imaging, University of Florida College of Medicine, Gainesville, FL, USA
[4]Division of Ultrasound, Department of Radiology, Center for Ultrasound Research & Translation, Massachusetts General Hospital, Boston, MA 02114, USA

Abstract

Ultrasound (US) imaging is the most commonly performed cross-sectional diagnostic imaging modality in the practice of medicine. It is low-cost, non-ionizing, portable, and capable of real-time image acquisition and display. US is a rapidly evolving technology with significant challenges and opportunities. Challenges include high inter- and intra-operator variability and limited image quality control. Tremendous opportunities have arisen in the last decade as a result of exponential growth in available computational power coupled with progressive miniaturization of US devices. As US devices become smaller, enhanced computational capability can contribute significantly to decreasing variability through advanced image processing. In this paper, we review leading machine learning (ML) approaches and research directions in US, with an emphasis on recent ML advances. We also present our outlook on future opportunities for ML techniques to further improve clinical workflow and US-based disease diagnosis and characterization.

Key words: Deep learning—Elastography—Machine learning—Medical ultrasound—Sonography

Ultrasound (US) is one of the core diagnostic imaging modalities, and is routinely used as the first line of medical imaging for evaluation of internal body structures, including solid organ parenchyma, blood vessels, the musculoskeletal system, and the fetus. US has become a ubiquitous diagnostic imaging tool owing to

several major advantages over other medical imaging methods such as computed tomography (CT) and magnetic resonance imaging (MRI). These key advantages include real-time imaging, no use of ionizing radiation, and better cost effectiveness than CT and MRI in many situations. In addition, US is portable, requires no shielding, and utilizes conventional electrical power sources and is therefore well suited to point-of-care applications, especially in under-resourced settings. As the field progresses, US, especially when combined with other technologies, has the potential to be an in-home biosensor, providing ambulatory, long duration, and non-intrusive monitoring with real-time biofeedback.

US also presents unique challenges, including operator dependence, noise, artifacts, limited field of view, difficulty in imaging structures behind bone and air, and variability across different manufacturers' US systems. Dependence on operator skill is particularly limiting. Many healthcare providers who are not imaging specialists do not use US at the point of care owing to a lack of skill in acquiring and interpreting images. For those that do, high inter- and intra-operator variability remains a significant challenge in clinical decision making. As a result of high inter-operator variability, US-derived tumor measurements are not accepted in most cancer drug trials, and US is therefore generally not used clinically for serial oncologic imaging. Automated US image analysis promises to play a crucial role in addressing some of these challenges.

Recent surveys of ML for medical imaging, such as [1–4], primarily focus on CT, MRI, and microscopy. In this review, we focus on the use of machine learning (ML) in US. The objective of this paper is to review how

Correspondence to: Laura J. Brattain; *email:* brattainl@ll.mit.edu

recent advances in ML have helped accelerate US image analysis adoption by modeling complicated multidimensional data relationships that answer diagnosis and disease severity classification questions. We have two goals: (1) to highlight contributions that utilize ML advances to solve current challenges in medical US, (2) to discuss future opportunities that will utilize ML techniques to further improve clinical workflow and US-based disease diagnosis and characterization. Our survey is non-exclusive, as we mainly focus on work within the past 5 years, where ML, particularly deep learning (DL), has started to have a major impact. We also emphasize solutions at the system level, which is an important aspect, due to the unique characteristics of the US image generation workflow. Figure 1 shows that US image processing involves more than simply a classification step, but additionally includes preprocessing and various types of analyses depending on several possible applications.

This article is divided into four sections: (i) an overview of basic principles of US, (ii) an overview of ML, (iii) ML for US, and (iv) summary and outlook.

Overview of US imaging

US imaging

Medical US images are formed by using an US probe to transmit mechanical wave pulses into tissue. Sound echoes are generated at boundaries where different tissues exhibit acoustic impedance differences. These echoes are recorded and displayed as an anatomic image, which may contain characteristic artifacts including signal dropout, attenuation, speckle, and shadows. Image quality is highly dependent on multiple factors, including force exerted on the US transducer, transducer location, and orientation.

Using various signal-to-image reconstruction approaches, several different types of images can be formed using US equipment. The most well-known and routinely used clinically is a B-mode image, which displays the acoustic impedance of a two-dimensional cross section of tissue. Other types of US imaging display blood flow (Doppler imaging and contrast-enhanced US), motion of tissue over time (M-mode), the anatomy of a three-dimensional region (3D US), and tissue stiffness (elastography).

US elastography

US elastography is a relatively new imaging technique of which there are two main types in current clinical use: (1) strain elastography, where image data are compared before and after application of external compression force to detect tissue deformation, and (2) shear wave elastography (SWE), which uses acoustic energy to move tissue, generating shear waves that extend laterally in tissue. These shear waves can be tracked to compute shear wave velocity, which is algebraically related to tissue stiffness measured as the tissue Young's modulus.

Tissue stiffness is a useful biomarker for pathologic processes, including fibrosis and inflammation, leading to several additional clinical applications for medical US. A diagnostic imaging gap recently addressed by US elastography is the evaluation of chronic liver disease [5–8].
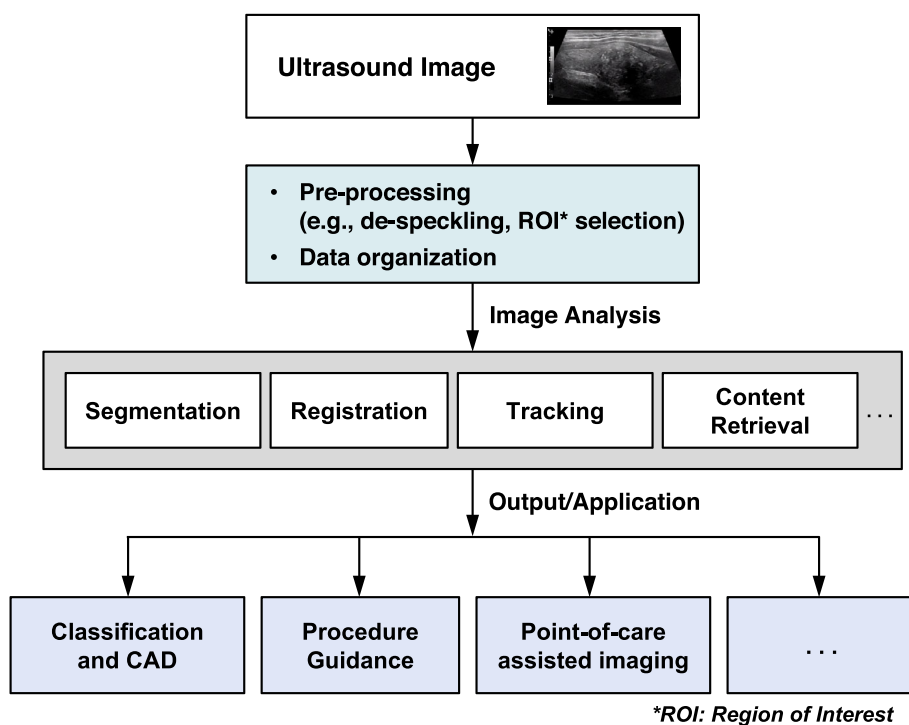

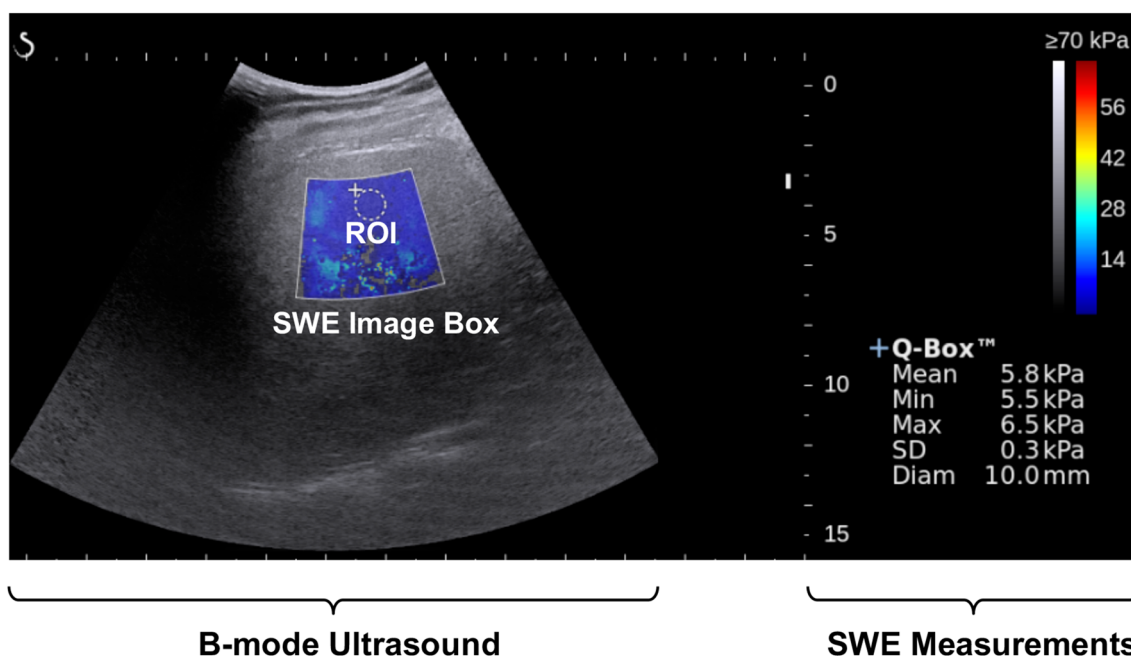Fig. 1. Overview of ultrasound processing system workflow.

*ROI: Region of Interest

Fig. 2. Example SWE color map overlaid on a B-mode ultrasound image.

US elastography liver stiffness measurements have been shown to be a promising liver fibrosis staging biomarker, and as a result highly relevant to chronic liver disease risk stratification. These technologies have the potential to replace liver biopsy as the diagnostic standard of care for key biologic variables in chronic liver disease. SWE has also been used to assess breast lesions [9–11], thyroid nodules [12–15], musculoskeletal conditions [16–20], and prostate cancer [21–23].

Figure 2 is an example of a SWE image (the colored pixels) overlaid on top of a B-mode US image, acquired for liver fibrosis staging. Tissue stiffness measurements are obtained by placing a region of interest (ROI) inside the SWE image box. Similar to B-mode US, elastography also suffers from inter- and intra-observer variability [8]. This represents an area of opportunity for ML-based automated image analysis improvement. We will discuss this in detail in Sect. "Additional applications of machine learning to US."

Contrast-enhanced US (CEUS)

Contrast-enhanced US utilizes gas-filled microbubbles for dynamic evaluation of microvasculature and macrovasculature. At present, US contrast agents are exclusively intravascular blood pool agents. Differentiation between benign and malignant focal liver lesions is an application of particular clinical interest [24–26]. The late phase of contrast enhancement allows for real-time characterization of washout, a critical feature in the differentiation of benign liver lesions (e.g., hemangioma, focal nodular hyperplasia, adenoma, regenerative nod-ule) and malignant liver lesions (e.g., hepatocellular carcinoma, cholangiocarcinoma, metastasis). The DE-GUM study—a multicenter German study that analyzed 1328 focal liver lesions—reported a 90.3% accuracy of CEUS for focal liver lesions with a 95.8% sensitivity, 83.1% specificity, 95.4% positive predictive value, and 95.9% negative predictive value for distinguishing benign and malignant liver lesions [27]. Other areas of clinical interest include evaluation of focal renal lesions [28], thyroid nodules [29], splenic lesions [30], and prostate cancer [31]. CEUS limitations include operator dependence, motion sensitivity, and the need for a good acoustic window. Advanced US image processing offers potential opportunities to augment CEUS by mitigating these limitations.

Overview of machine learning (ML)

ML is an interdisciplinary field that aims to construct algorithms that can learn from and make predictions on data [32], [33]. It is part of the broad field of artificial intelligence and overlaps with pattern recognition. Substantial progress has been made in applying ML to natural language processing (NLP), computer vision (e.g., image and text search, face recognition), video surveillance, financial data analysis, and many other domains. Recent progress in deep learning, a form of ML, has been dramatic, resulting in significant performance advances in international competitions and wide commercial adoption. The application of ML to diverse areas of computing is gaining popularity rapidly, not only because of more powerful hardware, but also because of the increasing availability of free and open source software,

which enable ML to be readily implemented. The purpose of this section is to introduce ML approaches and capabilities to US researchers and clinicians. Historical reviews of the field and its relationship with pattern recognition can be found elsewhere [34–36]. The following essential concepts are introduced at a level appropriate for understanding this review: supervised and unsupervised learning, learning based on hand-crafted features, deep learning, testing, and performance metrics.

Supervised vs. unsupervised learning

Most ML applications for US involve supervised learning, in which a classifier is trained on a database of US images labeled with desired classification outputs. For example, a classifier could be trained to output a value of 1 for input images of malignant tumors and a value of 0 for benign tumors. Once a classifier is trained, it can be used to classify previously unseen test images.

Unsupervised learning involves finding clusters or similarities in data, with no labels provided. This can be useful for applications such as content-based retrieval, or to determine features that can distinguish different classes of data.

A type of learning that falls between supervised and unsupervised learning is termed weakly supervised learning [37]. A significant challenge in building up large US image databases has been the time involved for expert annotation to support supervised learning. Annotation effort can be simplified by reducing the detail of information provided by the expert. For example, an image containing a tumor can be labeled as such, without having to annotate the precise location or boundaries. The ability to train a classifier with this type of less detailed information is termed weakly supervised learning. These and other types of learning, such as reinforcement learning, are described in detail elsewhere [38].

Learning based on handcrafted features

Traditionally, ML has involved computing handcrafted features that are believed to be able to distinguish between classes of data. These features are then used to train and test a classifier. For US, common types of features are morphologic, e.g., lesion area or perimeter, or textural [39], based on information in the frequency domain [40], or parameter fitting. Often a large number of candidate features are computed and then a feature selection algorithm is applied to select the best features or a dimensionality reduction algorithm [41] is applied to combine the features into a smaller composite set.

A classifier is then trained to form a feature mapping to compute desired outputs. It is important to constrain the classifier so that it does not overfit to the training data, because overfitting results in model errors that do not generalize beyond the training set to new data. The need to avoid overfitting is one of the main reasons feature selection or extraction algorithms are applied before training a classifier. Avoiding overfitting is a special concern for US research, which has thus far involved relatively small databases. Over the years, many supervised learning classification algorithms have been developed and many have been applied to US for handcrafted features. The most common approaches applied in the surveyed papers are random forests [42], support vector machines [43, 44], and multilayer feed-forward networks [45–47], also known as artificial neural networks.

Deep learning (DL)

The effort and domain expertise involved in handcrafting features has led researchers to seek algorithms that can learn features automatically from data. DL is a particularly powerful tool for extracting non-linear features from data. This is particularly promising in US, where predictable acoustic patterns are typically neither obvious nor easily hand-engineered. Figure 3 illustrates high-level differences between conventional ML and DL. The fast adoption of DL has been enabled by faster algorithms, more capable Graphics Processing Unit (GPU)-based computing, and large data sets.

DL extends multilayer feedforward networks from the two layers of weights used in the past to multiple layers. Figure 4 is an example of a generic supervised DL pipeline that includes both the learning phase and the deployment phase. In the learning phase, labeled samples (e.g., labeled US thyroid nodule images) are randomly divided into training/test sets or training/validation/test sets. The training data are used for finding the weights for each of the layers. During the process, features are discovered automatically and a model is learned. The validation set is used for optimizing the network parameters. The test data are used for estimating the performance of the learned network. This model estimation and selection technique is called cross-validation [48]. During the deployment phase, the machine applies the model learned to make a prediction on a new, unlabeled input (e.g., an unlabeled US thyroid nodule image that the machine has not seen before).

The multiple processing layers have been demonstrated to learn features of the data with multiple levels of hierarchy and abstraction [49]. For example, in imagery of humans, a low level of abstraction is edges; higher levels are body parts. A variety of deep learning structures have been explored. Among them, convolutional neural networks (CNNs) are one of the most popular choices for classifying images, due to unprecedented classification accuracy [50, 51] in applications such as object detection [52–54], face detection [55–57], and segmentation [58, 59]. In a typical CNN, convolu-
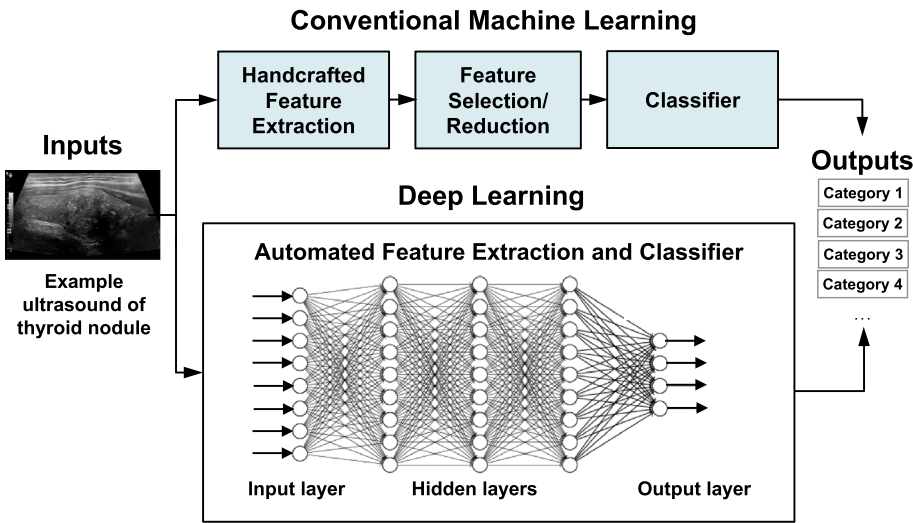
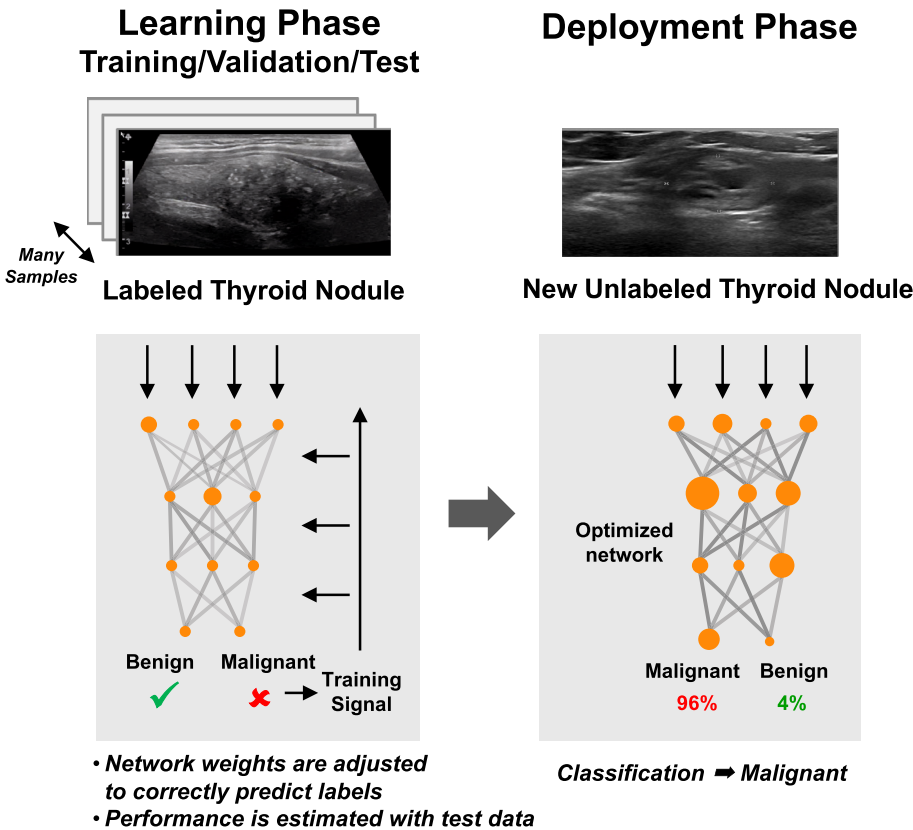**Fig. 3.** Conventional machine learning vs. deep learning.

Fig. 4. Supervised learning with deep neural networks.

tional filters are applied in each CNN layer to automatically extract features from the input image at multiple scales (e.g., edges, colors, and shapes), and a pooling process (termed 'max pooling') is often used between CNN layers in order to progressively reduce the feature map size. The last two layers are typically fully connected layers, from which classification labels are predicted (Fig. 5).

Testing and Performance Metrics

As mentioned in Sect. "Deep learning," classifier development and testing typically involve splitting the randomized labeled data into training/test sets or training/validation/test sets. A validation set is used to determine the best network structure and other classifier variations based on several training runs, and an independent test set held aside to evaluate performance until the classifier has been completed.
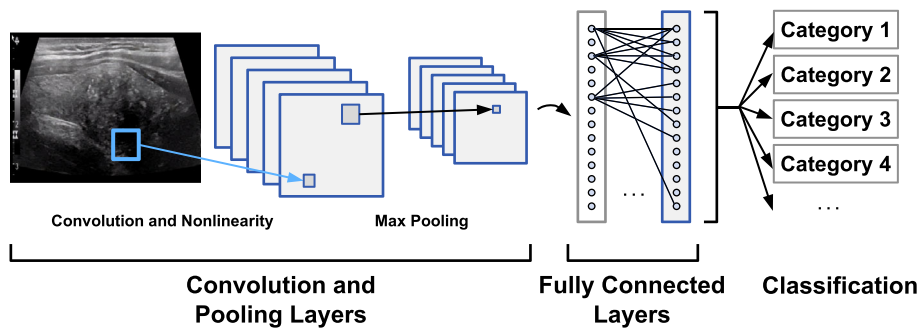
Fig. 5. Example convolutional neural network (CNN).

When a database is sufficiently large, it can be partitioned a priori into these distinct sets above. For smaller data sets such as those commonly seen in US research, k-fold cross-validation is often used. Cross-fold testing can be performed up to a maximum of N times for a database of N samples. In this case, termed leave-one-out testing, all of the samples in the database are used for training except for one sample data (e.g., one image), which is used for testing. Details of these techniques and other cross-validation techniques can be found in [60, 61].

Classifier performance is reported in a variety of ways. The most common for two classes is area under the receiver operating characteristic curve (AUROC), often simplified as "area under the curve" (AUC). An operating characteristic is formed by measuring true positive and false positive rates as the decision threshold applied to the classifier output is varied [62]. The AUC is then computed from the operating characteristic.

ML for US

Principal applications of ML to US include classification or computer-aided diagnosis, regression, and tissue segmentation. Other applications include image registration and content retrieval. Each of these applications is surveyed in the following subsections, with an aim to provide insights into progress and best approaches. In particular, advances in approaches using deep learning are highlighted, compared to approaches that use handcrafted features. Table 1 provides a summary of the applications in the papers surveyed.

Table 1. List of applications for papers surveyed

Organ or body location	Modality	Application	# Of papers	References
Breast	US	Classification (lesions)	10	[69, 71, 74, 93, 114–120]
Liver	US	Classification (lesions, cirrhosis, other focal and diffuse conditions)	4	[64, 121–123]
Lung	US	Classification (several diseases, B-lines)	2	[124, 125]
Muscle	US	Classification (atrophy, myositis)	2	[126, 127]
Intravascular	US	Classification (plaque)	2	[72, 128]
Spine	US	Classification (needle placement)	1	[129]
Ovary/uterus	US	Classification (masses)	2	[130, 131]
Kidney	US	Classification (renal disease, cysts)	1	[132]
Spleen	US	Classification (lesion)	1	[133]
Eye	US	Classification (cataracts)	1	[134]
Abdominal cavity	US	Classification (free fluid from trauma)	1	[135]
Thyroid	US	Classification (nodules)	1	[70]
Other	US	Classification (abdominal organs)	1	[73]
Heart	US	Classification (viewing planes, congenital heart disease) Content retrieval	8	[95–102]
Heart	US	Segmentation	1	[88]
Placenta	US	Segmentation	1	[91]
Lymph node	US	Segmentation	1	[90]
Brain	US	Segmentation	1	[92]
Prostate	US	Segmentation	2	[82, 83]
Carotid artery	US	Segmentation	1	[89]
Breast	US	Segmentation	1	[81]
Gastrocnemius muscle	US	Regression (orientation)	1	[76]
Fetal brain	US	Regression (gestational age)	1	[77]
Spine	US	Registration	1	[93]
Prostate	US (TRUS)	Registration	1	[94]
Liver	Elastography	Classification (fibrosis, chronic liver disease)	2	[106, 109]
Thyroid	Elastography	Classification (nodule)	1	[108]
Breast	Elastography	Classification (tumor)	3	[107, 110, 111]
Liver	CEUS	Classification (lesions)	1	[112]
Total			56	

Classification

Computer-aided disease diagnosis and classification in radiology have received extensive attention and have benefited greatly from the recent advances in ML. A variety of applications have been addressed in computer-aided diagnosis, but primarily for detecting or classifying lesions, mainly in the breast and liver. Most of the recent papers surveyed follow the classic approach of computing handcrafted features, applying a feature selection algorithm, and training a classifier on the reduced feature set. This basic approach has been investigated for over 20 years, e.g., [63, 64]. Specific feature and algorithm choices for each step vary. Preprocessing includes despeckling.

Features considered are primarily texture-based or morphological. The largest number of publications has been on classifying breast lesions. A review of breast image analysis [65] places US in the context of several imaging modalities. For classifying breast lesions, computerized methods have been developed to automatically extract features from the BI-RADS (Breast Imaging Reporting and Data System) lexicon, relating to shape, margin, orientation, echo pattern, and acoustic shadowing [66]. These features are standardized and readily understandable by radiologists. Typically, a large number of features is reduced in dimension by either selecting the most informative features or by linearly combining features with principal components analysis [41]. Commonly used classifiers include multilayer networks (neural networks) [67], support vector machines [43], and random forests [68], the details of which extend beyond the scope of this review.

Although these papers indicate the promise of US computer-aided diagnosis, the reported studies have several limitations. These classification studies typically rely on manual region-of-interest (ROI) selection of the portion of the image that includes candidate pathology; that subimage is then classified. Manual ROI selection assumes significant involvement by a radiologist in practice, or at least neglects the problem of ROI selection. The number of patients and images available for training and testing is typically small; in nearly all cases, the number of images was < 300. In addition, the US images were often collected at a single location by a single type of US device. Each paper reports results obtained on a different validation database, making results difficult or impossible to compare.

Two recent papers have compared the performance of commercial diagnosis systems vs. radiologists. In [69], performance of a system from ClearView Diagnostics (Piscataway, New Jersey, USA) for diagnosing breast lesions was compared to that of three certified radiologists. At the time of publication, the system was being reviewed for FDA clearance. The study was co-authored by ClearView Diagnostics employees and thus was not an independent evaluation. Ground truth for 1300 images was determined based on biopsy or one-year follow-up. Likelihood of malignancy and the preliminary BI-RADS assessment were assessed. The comparison focused on images; the reading radiologists did not have access to other information, such as patient history and previous imaging studies. Based on likelihood of malignancy, the computer system was determined to have outperformed the radiologists. Fusing the radiologist and computer assessments was also found to improve sensitivity and specificity over radiologist assessments alone.

In [70], performance of a system from Samsung (Seoul, South Korea) for assessing malignancy of thyroid nodules was compared to that from an experienced radiologist. One hundred two nodules with a definitive diagnosis from 89 patients were included in the study. The system's performance was lower than that of the radiologist's. It was speculated that improved segmentation would improve the performance.

The number of papers that have applied deep learning techniques to US disease classification has dramatically increased in the last 2–3 years [71, 72]. For deep learning, it has been unclear until very recently whether CNNs that have been trained on non-medical color images can be used as a starting point and partially retrained to classify US images that do not resemble optical images. However, recent work, such as [73], has shown that this method, referred to as "transfer learning," can be effective. This technique avoids overfitting on small data sets, which is often the case for US imagery. Fusing handcrafted features with those computed with deep learning has been shown to further improve performance [74]. Weakly supervised learning has also been successfully applied to US [75].

Regression

Regression involves estimating continuous values as opposed to discrete classes of data. Deep learning has been applied to regression, for example by [76] to estimate muscle fiber orientation from US imagery. Deep learning was found to improve over previous approaches using handcrafted features, specifically a well-established wavelet-based method. However, another regression application provides an example of how handcrafted features may still be the preferred approach. In this application, gestational age is estimated from 3D US images of the fetal brain [77]. A semi-automated approach based on deformable surfaces is used to compute standard biometric features, e.g., head circumference, as well as information on local structural changes in the brain.

Segmentation

Segmentation is the delineation of structural boundaries. Automated US segmentation is challenging; in that US data are often affected by speckle, shadow, and missing boundaries, as well as by tradeoffs between US frequency, depth, and resolution during image acquisition.

Many US segmentation approaches have been developed, including methods based on intensity thresholding, level sets, active contours [78], and other model-based methods. These techniques are reviewed in [79, 80]. Intensity-based approaches are sensitive to the noise and image quality. Active contour and level sets require initialization, which can affect the results. Most conventional approaches are not fully automated.

Segmentation methods based on ML typically involve two steps: first, a pixel-wise classification of the desired structure, followed by a clean-up or smoothing step since the pixel-wise classification is noisy. In recent papers, several classification approaches have been investigated involving handcrafted features [81–87], and various types of neural networks, including deep-learning [88–92]. Three papers [82, 91, 92] made use of 3D US.

Additional applications of machine learning to US

In addition to US segmentation, ML has also been applied to US registration, for example for imagery of vertebrae [93] and transrectal US [94].

One key advantage of US over some other modalities is that it is well suited for real-time guidance (e.g., needle guidance, intra-cardiac procedures, and robotic surgeries), but the real-time performance has not been fully realized due to the limitations in US image processing, including lack of robust content retrieval from US video clips. A number of very recent papers focus on using deep learning techniques for frame labeling or content interpretation [95–98]. One approach [99] was evaluated on a database of about 30,000 images, which is very large for US. Techniques that integrate spatiotemporal information have started to emerge, particularly in dealing with echocardiograms acquired from different views, to capture key information of the motion of heart [100–102]. We predict ML will play a major role in the near future in enabling US guided interventions.

US elastography and CEUS

Elastography, particularly SWE, is being increasingly used in conjunction with US as a quantitative measurement to characterize tissue lesions [103]. Key limitations of SWE, as summarized in [13], include variability in stiffness cutoff thresholds, lack of image quality control, and variability in ROI selection. It has been shown that SWE measurements depend greatly on the quality of the

acquired data [104, 105]. Using liver fibrosis staging as an example, Fig. 6A illustrates the existing clinical workflow and challenges. As such, the current clinical protocol requires multiple image acquisitions as a way to mitigate measurement variability. Figure 6B presents a potential solution to improving the clinical workflow. It includes algorithms to automatically check image quality and ML methods to quantify SWE and classify disease stages. In addition, algorithms can also assist with assessing additional useful biomarkers (e.g., subcutaneous fat content, steatosis, inflammation), which are currently not used because of the time-intensive manual interpretation required.

Among the surveyed papers from the past 5 years, the most common ML approach is to extract statistical features from the SWE images and then apply a classifier [106–110].

SWE images often contain irrelevant patterns (e.g., artifacts, noise, areas absence of SWE information), which can be difficult for both handcrafted feature extraction approaches and for typical DL methods such as CNN. Very recently, [111] reported using a two-layer DL network for automated feature extraction from SWE breast data. The work focuses on differentiating task-relevant (i.e., patterns of interest) vs. task-irrelevant patterns (i.e., distracting patterns).

CEUS is a non-invasive diagnostic tool for focal liver lesion evaluation. Typically, time intensity curves (TICs) are extracted from manually selected ROI in CEUS. Results are often subjected to operator variability, motion sensitivity, and speckle noise. Recently, DL has been applied to CEUS to improve the classification of benign and malignant focal liver lesions from automatically extracted TICs with respiratory compensation [112]. DL shows higher accuracy than conventional ML methods.

Discussion and outlook

While the use of medical US is becoming ubiquitous, advanced US image analysis techniques lag behind other modalities such as CT and MRI. As with CT and MRI, ML is a promising approach to improve US image analysis, disease classification, and computer-aided diagnosis.

Overall, application of ML to US is at an early stage, but is rapidly progressing, as evidenced by the large number of 2016 and 2017 surveyed papers. Most of the recent papers surveyed use databases of a few hundred images. Only a few papers use databases of at least one thousand images, which is three orders of magnitude smaller than large challenge databases of optical images. On the other hand, it is unrealistic to expect US databases will reach that size in the foreseeable future, pointing to the need for ML techniques that can train on smaller databases. In many cases, databases are generated from a single device type and a single collection site,
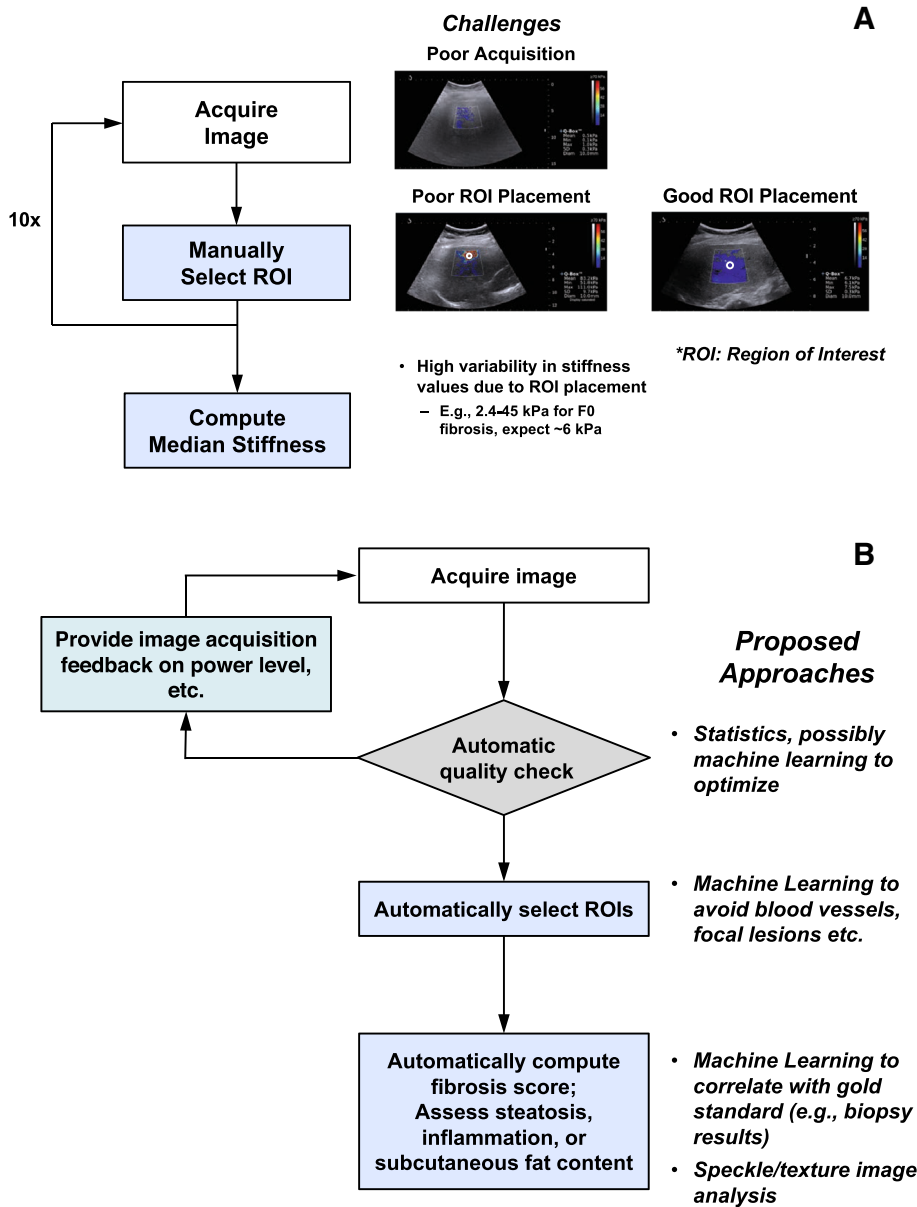
Fig. 6.   A Example pipeline of using SWE for liver fibrosis staging. **B** Proposed semi-automated SWE acquisition and analysis workflow.

limiting the generalizability of ML classification models derived from these databases. Large, publicly accessible challenge databases such as ImageNet that have significantly advanced conventional image classification performance are currently unavailable for US. Most of the present US ML research has concentrated on single functions within an overall system, such segmentation or classification.

Within the past few years, deep learning approaches have been shown to significantly improve performance when compared with classifiers operating on handcrafted features. Transfer learning, which involves retraining a portion of a network originally trained on other images, has been shown to be effective for classifying the relatively small databases that are currently available. These results address early skepticism that transfer learning

would not be useful for US because US images appear to be quite different than the optical color imagery on which the networks were originally trained. Deep learning approaches have also obviated the need for sophisticated preprocessing, such as despeckling. On the other hand, certain applications are based on sophisticated handcrafted features that are unlikely to be surpassed by deep learning with currently available databases. Moreover, surveyed papers combining classifiers with both deep learning and handcrafted features have shown improved results over either approaches, indicating that Deep Learning techniques alone are unlikely to achieve the potential of ML in US.

There are several challenges in applying ML to US and other medical imaging modalities: (1) because US is often used as a first-line imaging modality, there is often

an imbalance with an excess of normal "no-disease" images, and (2) obtaining consistently annotated data is a challenge as there is significant inter-operator and inter-observer variability among expert US physicians. The variability that this subjectivity adds to the annotations requires a larger database so the classifier can be trained to smooth over the variations. Transfer learning has been widely adopted to address the challenge of operating with relatively small databases. Weakly supervised learning was also successfully used in one surveyed paper; its use is likely to increase, although challenges have been found in unpublished work. In addition to these techniques, other approaches commonly used by the deep learning community to address small, annotated databases are unsupervised learning, database augmentation and active learning. Interestingly, these techniques have been rarely used in US, and are likely to be promising approaches. Active learning requires an interactive annotation tool that is somewhat more complex than a static tool, but once developed, has the value of focusing the expert's time on images most important to annotate. Another approach to annotating images would be to apply natural language processing tools to extract annotations from the patient reports. This is still an area of research that has its own challenges to address.

Another algorithmic challenge is the need for the results to be interpretable by radiologists, as opposed to a "black box" result that might suffice in domains other than clinical medicine. Although interpretability is not an intrinsic characteristic of deep learning, it is an active area of research. Within the past few years, new techniques for interpreting CNNs have emerged, and other classification techniques are being developed that are intrinsically interpretable [113].

One key strength of US is its ability to produce real-time video. ML applied to echocardiography and obstetrics has increasingly exploited the advantages of spatiotemporal data to improve results. Even in the case of detecting tumors and other pathologies, video clips provide more information than a single image frame. None of the surveyed papers about classifying pathologies exploited video data. This is an aspect that will likely advance in future work.

Returning to the system view in Fig. 1, advances across the workflow are needed. ML enables part of the system solution, but not all of it. For example, a unique challenge of US is the expertise required for image acquisition, which currently contributes to variable interpretations. Operating on freehand US is preferred. In the future, it will be important for ML systems to provide real-time feedback to the sonographer during image acquisition, and not only to interpret freehand US post hoc. Also, manual ROI selection and caliper placement for measurements are still common, which also result in significant variability. Image quality control, automatic ROI selection, and attention to computer–human integration are needed to replace manual ROI selection and caliper placement for measurements.

Based on recent rapid progress summarized in this review, we expect ML for US will continue to progress, and will be one of the most important trends in diagnostic US in the coming years. Broadly speaking, US will likely become one of the many inputs of a ML-based intelligent diagnostic assistant system, where multimodal and multiscale observations are learned over time and
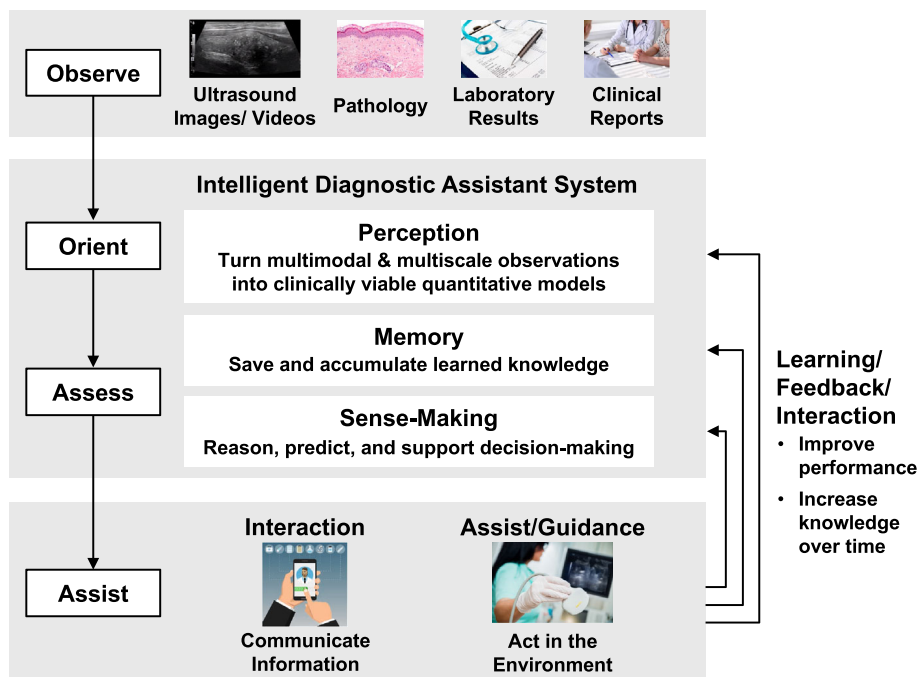

Fig. 7. Proposed framework for machine learning-based intelligent diagnostic assistant.

are turned into clinical viable quantitative models (Fig. 7); the aggregated machine intelligence will have the ability to observe data, orient the end user, assess new information, and assist with decision making. Such a system has the potential to greatly improve not only the clinical workflow but also the overall outcome of care.

Compliance with ethical standards

Conflict of interest All the authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Wang S, Summers RM (2012) Machine learning and radiology. Med Image Anal 16(5):933–951
2. Shen D, Wu G, Suk H-I (2017) Deep learning in medical image analysis. Annu Rev Biomed Eng 19:221–248
3. Litjens G et al. (2017) A survey on deep learning in medical image analysis. ArXiv Prepr. ArXiv170205747
4. Ravi D, et al. (2017) Deep Learning for health informatics. IEEE J Biomed Health Inform 21(1):4–21
5. Cassinotto C, et al. (2014) Non-invasive assessment of liver fibrosis with impulse elastography: comparison of Supersonic Shear Imaging with ARFI and FibroScan®. J. Hepatol. 61(3):550–557
6. Ferraioli G, Parekh P, Levitov AB, Filice C (2014) Shear wave elastography for evaluation of liver fibrosis. J. Ultrasound Med 33(2):197–203
7. Poynard T, et al. (2013) Liver fibrosis evaluation using real-time shear wave elastography: applicability and diagnostic performance using methods without a gold standard. J Hepatol 58(5):928–935
8. Samir AE, et al. (2014) Shear-wave elastography for the estimation of liver fibrosis in chronic liver disease: determining accuracy and ideal site for measurement. Radiology 274(3):888–896
9. Liu B, et al. (2016) Breast lesions: quantitative diagnosis using ultrasound shear wave elastography—a systematic review and meta-analysis. Ultrasound Med Biol 42(4):835–847
10. Wang M, et al. (2017) Differential diagnosis of breast category 3 and 4 nodules through BI-RADS classification in conjunction with shear wave elastography. Ultrasound Med Biol 43(3):601–606
11. Wang ZL, Li Y, Wan WB, Li N, Tang J (2017) Shear-wave elastography: could it be helpful for the diagnosis of non-mass-like breast lesions? Ultrasound Med Biol 43(1):83–90
12. Anvari A, Dhyani M, Stephen AE, Samir AE (2016) Reliability of shear-wave elastography estimates of the young modulus of tissue in follicular thyroid neoplasms. Am J Roentgenol 206(3):609–616
13. Dhyani M, Li C, Samir AE, Stephen AE (2017) Elastography: applications and limitations of a new technology. *Advanced thyroid and parathyroid ultrasound.* New York: Springer, pp 67–73
14. Ding J, Cheng HD, Huang J, Zhang Y, Liu J (2012) An improved quantitative measurement for thyroid cancer detection based on elastography. Eur J Radiol 81(4):800–805
15. Park AY, Son EJ, Han K, et al. (2015) Shear wave elastography of thyroid nodules for the prediction of malignancy in a large scale study. Eur J Radiol 84(3):407–412
16. Eby SF, et al. (2015) Shear wave elastography of passive skeletal muscle stiffness: influences of sex and age throughout adulthood. Clin Biomech 30(1):22–27
17. Pass B, Jafari M, Rowbotham E, et al. (2017) Do quantitative and qualitative shear wave elastography have a role in evaluating musculoskeletal soft tissue masses? Eur J Radiol 27(2):723–731
18. Taljanovic MS, et al. (2017) Shear-wave elastography: basic physics and musculoskeletal applications. RadioGraphics 37(3):855–870
19. Aubry S, Nueffer J-P, Tanter M, et al. (2014) Viscoelasticity in Achilles tendonopathy: quantitative assessment by using real-time shear-wave elastography. Radiology 274(3):821–829
20. Zhang ZJ, Ng GY, Lee WC, Fu SN (2014) Changes in morphological and elastic properties of patellar tendon in athletes with unilateral patellar tendinopathy and their relationships with pain and functional disability. PLoS ONE 9(10):e108337
21. Rouvière O, et al. (2017) Stiffness of benign and malignant prostate tissue measured by shear-wave elastography: a preliminary study. Eur Radiol 27(5):1858–1866
22. Sang L, Wang X, Xu D, Cai Y (2017) Accuracy of shear wave elastography for the diagnosis of prostate cancer: a meta-analysis. Sci Rep 7(1):1949
23. Woo S, Suh CH, Kim SY, Cho JY, Kim SH (2017) Shear-wave elastography for detection of prostate cancer: a systematic review and diagnostic meta-analysis. Am J Roentgenol 209:1–9
24. D'Onofrio M, Crosara S, De Robertis R, Canestrini S, Mucelli RP (2015) Contrast-enhanced ultrasound of focal liver lesions. Am J Roentgenol 205(1):W56–W66
25. Kim TK, Jang H-J (2014) Contrast-enhanced ultrasound in the diagnosis of nodules in liver cirrhosis. World J Gastroenterol 20(13):3590–3596
26. Strobel D, et al. (2008) Contrast-enhanced ultrasound for the characterization of focal liver lesions–diagnostic accuracy in clinical practice (DEGUM multicenter trial). Ultraschall Med Stuttg Ger 29(5):499–505
27. Westwood M, et al. (2013) Contrast-enhanced ultrasound using SonoVue® (sulphur hexafluoride microbubbles) compared with contrast-enhanced computed tomography and contrast-enhanced magnetic resonance imaging for the characterisation of focal liver lesions and detection of liver metastases: a systematic review and cost-effectiveness analysis. Health Technol Assess Winch Engl 17(16):1–243
28. Oh TH, Lee YH, Seo IY (2014) Diagnostic efficacy of contrast-enhanced ultrasound for small renal masses. Korean J Urol 55(9):587–592
29. Yuan Z, Quan J, Yunxiao Z, Jian C, Zhu H (2015) Contrast-enhanced ultrasound in the diagnosis of solitary thyroid nodules. J Cancer Res Ther 11(1):41–45
30. Li W, et al. (2014) Real-time contrast enhanced ultrasound imaging of focal splenic lesions. Eur J Radiol 83(4):646–653
31. Baur ADJ, et al. (2017) A direct comparison of contrast-enhanced ultrasound and dynamic contrast-enhanced magnetic resonance imaging for prostate cancer detection and prediction of aggressiveness. Eur Radiol . https://doi.org/10.1007/s00330-017-5192-2
32. Bishop CM (2006) *Pattern recognition and machine learning.* New York: Springer
33. Mitchell TM (1997) *Machine learning. WCB.* Boston: McGraw-Hill
34. De Mantaras RL, Armengol E (1998) Machine learning from examples: inductive and lazy methods. Data Knowl Eng 25(1–2):99–123
35. Dutton DM, Conroy GV (1997) A review of machine learning. Knowl Eng Rev 12(4):341–367
36. Kotsiantis SB, Zaharakis ID, Pintelas PE (2006) Machine learning: a review of classification and combining techniques. Artif Intell Rev 26(3):159–190
37. Torresani L (2014) Weakly supervised learning". *Computer vision.* New York: Springer, pp 883–885
38. Mohri M, Rostamizadeh A, Talwalkar A (2012) *Foundations of machine learning.* Cambridge: MIT press

39. Soh L-K, Tsatsoulis C (1999) Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. IEEE Trans Geosci Remote Sens 37(2):780–795

40. Materka A, Strzelecki M et al. (1998) Texture analysis methods–a review. Technical University of Lodz, Institute of Electronics, COST B11 Report, Brussels, pp 9–11

41. Wold S, Esbensen K, Geladi P (1987) Principal component analysis. Chemom Intell Lab Syst 2(1–3):37–52

42. Liaw A, Wiener M, et al. (2002) Classification and regression by randomForest. R News 2(3):18–22

43. Cortes C, Vapnik V (1995) Support vector machine. Mach Learn 20(3):273–297

44. Suykens JA, Vandewalle J (1999) Least squares support vector machine classifiers. Neural Process Lett 9(3):293–300

45. White H (1990) Connectionist nonparametric regression: multilayer feedforward networks can learn arbitrary mappings. Neural Netw 3(5):535–549

46. Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal approximators. Neural Netw 2(5):359–366

47. Leshno M, Lin VY, Pinkus A, Schocken S (1993) Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. Neural Netw 6(6):861–867

48. Geisser S (1993) *Predictive inference: an introduction*. New York: Chapman & Hall

49. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444

50. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp 1097–1105

51. Szegedy C et al.(2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1–9

52. Sermanet P, Kavukcuoglu K, Chintala S, LeCun Y (2013) Pedestrian detection with unsupervised multi-stage feature learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 3626–3633

53. Szegedy C, Toshev A, Erhan D (2013) Deep neural networks for object detection. In: Advances in neural information processing systems, pp 2553–2561

54. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition, ArXiv Prepr. ArXiv14091556

55. Lawrence S, Giles CL, Tsoi AC, Back AD (1997) Face recognition: a convolutional neural-network approach. IEEE Trans Neural Netw 8(1):98–113

56. Matsugu M, Mori K, Mitari Y, Kaneda Y (2003) Subject independent facial expression recognition with robust face detection using a convolutional neural network. Neural Netw 16(5):555–559

57. Farfade SS, Saberian MJ, Li LJ (2015) Multi-view face detection using deep convolutional neural networks. In: Proceedings of the 5th ACM on International Conference on Multimedia Retrieval, pp 643–650

58. Turaga SC, et al. (2010) Convolutional networks can learn to generate affinity graphs for image segmentation. Neural Comput 22(2):511–538

59. Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 3431–3440

60. Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. Ijcai 14:1137–1145

61. Arlot S, Celisse A (2010) A survey of cross-validation procedures for model selection. Stat Surv 4:40–79

62. Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143(1):29–36

63. Garra BS, Krasner BH, Horii SC, et al. (1993) Improving the distinction between benign and malignant breast lesions: the value of sonographic texture analysis. Ultrason Imaging 15(4):267–285

64. Maclin PS, Dempsey J (1992) Using an artificial neural network to diagnose hepatic masses. J Med Syst 16(5):215–225

65. Giger ML, Karssemeijer N, Schnabel JA (2013) Breast image analysis for risk assessment, detection, diagnosis, and treatment of cancer. Annu Rev Biomed Eng 15(1):327–357

66. Shan J, Alam SK, Garra B, Zhang Y, Ahmed T (2016) Computer-aided diagnosis for breast ultrasound using computerized BI-RADS features and machine learning methods. Ultrasound Med Biol 42(4):980–988

67. Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117

68. Carbonell JG, Michalski RS, Mitchell TM (1983) An overview of machine learning. *Machine learning*. New york: Springer, pp 3–23

69. Barinov L, Jairaj A, Paster L et al. (2016) Decision quality support in diagnostic breast ultrasound through artificial Intelligence. In: Signal Processing in Medicine and Biology Symposium, pp 1–4

70. Choi YJ, et al. (2017) A computer-aided diagnosis system using artificial intelligence for the diagnosis and characterization of thyroid nodules on ultrasound: initial clinical assessment. Thyroid 27(4):546–552

71. Hiramatsu Y, Muramatsu C, Kobayashi H, Hara , Fujita H (2017) Automated detection of masses on whole breast volume ultrasound scanner: false positive reduction using deep convolutional neural network. Med Imaging . https://doi.org/10.1117/12.2254581

72. Lekadir K, et al. (2017) A convolutional neural network for automatic characterization of plaque composition in carotid ultrasound. IEEE J Biomed Health Inform 21(1):48–55

73. Cheng PM, Malhi HS (2017) Transfer learning with convolutional neural networks for classification of abdominal ultrasound images. J Digit Imaging 30(2):234–243

74. Antropova N, Huynh BQ, Giger ML (2017) A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. Med Phys. https://doi.org/10.1002/mp.12453

75. Qi H, Collins S, Noble A (2017) Weakly supervised learning of placental ultrasound images with residual networks. In: Annual Conference on Medical Image Understanding and Analysis, pp 98–108

76. Cunningham R, Harding P, Loram I (2017) Deep residual networks for quantification of muscle fiber orientation and curvature from ultrasound images. In: Hernández MV, González-Castro V, González-Castro V (eds) *Medical image understanding and analysis*, vol. 723. Cham: Springer, pp 63–73

77. Namburete AI, Stebbing RV, Kemp B, et al. (2015) Learning-based prediction of gestational age from ultrasound images of the fetal brain. Med. Image Anal. 21(1):72–86

78. Cary TW, Reamer CB, Sultan LR, Mohler ER, Sehgal CM (2014) Brachial artery vasomotion and transducer pressure effect on measurements by active contour segmentation on ultrasound: brachial artery vasomotion and transducer pressure effect. Med Phys 41(2):022901

79. Noble JA, Boukerroui D (2006) Ultrasound image segmentation: a survey. IEEE Trans Med Imaging 25(8):987–1010

80. Noble JA (2010) Ultrasound image segmentation and tissue characterization. Proc Inst Mech Eng Part H 224(2):307–316

81. Torbati N, Ayatollahi A, Kermani A (2014) An efficient neural network based method for medical image segmentation. Comput Biol Med 44:76–87

82. Yang X, Rossi PJ, Jani AB, et al. (2016) 3D transrectal ultrasound (TRUS) prostate segmentation based on optimal feature learning framework. Med Imaging. https://doi.org/10.1117/12.2216396

83. Ghose S, et al. (2013) A supervised learning framework of statistical shape and probability priors for automatic prostate segmentation in ultrasound images. Med Image Anal 17(6):587–600

84. Sultan LR, Xiong H, Zafar HM, et al. (2015) Vascularity assessment of thyroid nodules by quantitative color doppler ultrasound. Ultrasound Med Biol 41(5):1287–1293

85. Chauhan A, Sultan LR, Furth EE, et al. (2016) Diagnostic accuracy of hepatorenal index in the detection and grading of hepatic steatosis: factors affecting the accuracy of HRI. J Clin Ultrasound 44(9):580–586

86. Noe MH, et al. (2017) High frequency ultrasound: a novel instrument to quantify granuloma burden in cutaneous sarcoidosis. Sarcoidosis Vasc Diffuse Lung Dis 34(2):136–141

87. Xiong H, Sultan LR, Cary TW et al. (2017) The diagnostic performance of leak-plugging automated segmentation vs. manual tracing of breast lesions on ultrasound images. Ultrasound http://journals.sagepub.com/doi/pdf/10.1177/1742271X17690425#articleCitationDownloadContainer. Accessed 17 Jan 2018

88. Carneiro G, Nascimento JC, Freitas A (2012) The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods. IEEE Trans Image Process 21(3):968–982

89. Menchón-Lara RM, Sancho-Gómez JL (2015) Fully automatic segmentation of ultrasound common carotid artery images based on machine learning. Neurocomputing 151(P1):161–167

90. Zhang Y, Ying MT, Yang L, Ahuja AT, Chen DZ (2016) Coarse-to-fine stacked fully convolutional nets for lymph node segmentation in ultrasound images. In: Bioinformatics and Biomedicine (BIBM), 2016 IEEE International Conference, pp 443–448

91. Looney P et al. (2017) Automatic 3D ultrasound segmentation of the first trimester placenta using deep learning. In: Biomedical Imaging (ISBI 2017), IEEE 14th International Symposium on, pp 279–282

92. Milletari F, et al. (2017) Hough-CNN: deep learning for segmentation of deep brain regions in MRI and ultrasound. Comput Vis Image Underst 164:92–102

93. Chen F, Wu D, Liao H (2016) Registration of CT and ultrasound images of the spine with neural network and orientation code mutual information. In: Zheng G, Liao H, Jannin P, Cattin P, Lee S-L (eds) *Medical imaging and augmented reality*, vol. 9805. Cham: Springer, pp 292–301

94. Yang X, Fei B (2012) 3D prostate segmentation of ultrasound images combining longitudinal image registration and machine learning. In: Proceedings of SPIE, vol 8316, p 83162O

95. Gao Y, Maraci MA, Noble JA (2016) Describing ultrasound video content using deep convolutional neural networks. In: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), pp 787–790

96. Baumgartner CF, Kamnitsas K, Matthew J et al.(2016) Real-time standard scan plane detection and localisation in fetal ultrasound using fully convolutional neural networks. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol 9901 LNCS, pp 203–211

97. Kumar A et al. (2017) Plane identification in fetal ultrasound images using saliency maps and convolutional neural networks. In: Proceedings of the IEEE International Symposium on Biomedical Imaging, pp 791–794

98. Chen H et al. (2015) Automatic fetal ultrasound standard plane detection using knowledge transferred recurrent neural networks. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol 9349, pp 507–514

99. Yaqub M, Kelly B, Papageorghiou AT, Noble JA (2015) Guided random forests for identification of key fetal anatomy and image categorization in ultrasound scans. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp 687–694

100. Gao X, Li W, Loomes M, Wang L (2017) A fused deep learning architecture for viewpoint classification of echocardiography. Inf Fusion 36:103–113

101. Sundaresan V, Bridge CP, Ioannou C, Noble JA (2017) Automated characterization of the fetal heart in ultrasound images using fully convolutional neural networks. In: Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on, pp 671–674

102. Khamis H, Zurakhov G, Azar V, et al. (2017) Automatic apical view classification of echocardiograms using a discriminative learning dictionary. Med Image Anal 36:15–21

103. Sigrist RM, Liau J, El Kaffas A, Chammas MC, Willmann JK (2017) Ultrasound elastography: review of techniques and clinical applications. Theranostics 7(5):1303

104. Rouze NC, Wang MH, Palmeri ML, Nightingale KR (2012) Parameters affecting the resolution and accuracy of 2-D quantitative shear wave images. IEEE Trans Ultrason Ferroelectr Freq Control 59:1729–1740

105. Pellot-Barakat C, Lefort M, Chami L, et al. (2015) Automatic assessment of shear wave elastography quality and measurement reliability in the liver. Ultrasound Med Biol 41(4):936–943

106. Wang J, Guo L, Shi X, et al. (2012) Real-time elastography with a novel quantitative technology for assessment of liver fibrosis in chronic hepatitis B. Eur J Radiol 81(1):e31–e36

107. Xiao Y, et al. (2014) Computer-aided diagnosis based on quantitative elastographic features with supersonic shear wave imaging. Ultrasound Med Biol 40(2):275–286

108. Bhatia KSS, Lam ACL, Pang SWA, Wang D, Ahuja AT (2016) Feasibility study of texture analysis using ultrasound shear wave elastography to predict malignancy in thyroid nodules. Ultrasound Med Biol 42(7):1671–1680

109. Gatos I, et al. (2017) A machine-learning algorithm toward color analysis for chronic liver disease classification, employing ultrasound shear wave elastography. Ultrasound Med Biol 43:1797–1810

110. Zhang Q, Xiao Y, Chen S, Wang C, Zheng H (2015) Quantification of elastic heterogeneity using contourlet-based texture analysis in shear-wave elastography for breast tumor classification. Ultrasound Med Biol 41(2):588–600

111. Zhang Q, et al. (2016) Deep learning based classification of breast tumors with shear-wave elastography. Ultrasonics 72:150–157

112. Wu K, Chen X, Ding M (2014) Deep learning based classification of focal liver lesions with contrast-enhanced ultrasound. Opt-Int J Light Electron Opt 125(15):4057–4063

113. Zeng J, Ustun B, Rudin C (2017) Interpretable classification models for recidivism prediction. J R Stat Soc Ser A 180(3):689–722

114. Shi J, Zhou S, Liu X, et al. (2016) Stacked deep polynomial network based representation learning for tumor classification with small ultrasound image dataset. Neurocomputing 194:87–94

115. Singh BK, Verma K, Thoke AS (2016) Fuzzy cluster based neural network classifier for classifying breast tumors in ultrasound images. Expert Syst Appl 66:114–123

116. Wu WJ, Lin SW, Moon WK (2015) An artificial immune system-based support vector machine approach for classifying ultrasound breast tumor images. J Digit Imaging 28(5):576–585

117. Shan J, Cheng HD, Wang Y (2012) Completely automated segmentation approach for breast ultrasound images using multiple-domain features. Ultrasound Med Biol 38(2):262–275

118. Nascimento CDL, Silva SDS, da Silva TA, et al. (2016) Breast tumor classification in ultrasound images using support vector machines and neural networks. Rev Bras Eng Biomed 32(3):283–292

119. Marcomini KD, Carneiro AAO, Schiabel H (2016) Application of artificial neural network models in segmentation and classification of nodules in breast ultrasound digital images. Int J Biomed Imaging 2016:13

120. Jamieson AR, Giger ML, Drukker K, et al. (2009) Exploring nonlinear feature space dimension reduction and data representation in breast CADx with Laplacian eigenmaps and t-SNE: nonlinear dimension reduction and representation in breast CADx. Med Phys 37(1):339–351

121. Hwang YN, Lee JH, Kim GY, Jiang YY, Kim SM (2015) Classification of focal liver lesions on ultrasound images by extracting hybrid textural features and using an artificial neural network. Biomed Mater Eng 26:S1599–S1611

122. Suganya R, Kirubakaran R, Rajaram S (2014) *Classification and retrieval of focal and diffuse liver from ultrasound images using machine learning techniques*. Cham: Springer, pp 253–261

123. Kalyan K, Jakhia B, Lele RD, Joshi M, Chowdhary A (2014) Artificial neural network application in the diagnosis of disease conditions with liver ultrasound images. Adv Bioinforma. https://doi.org/10.1155/2014/708279

124. Brattain LJ, Telfer BA, Liteplo AS, Noble VE (2013) Automated B-line scoring on thoracic sonography. J Ultrasound Med 32(12):2185–2190

125. Veeramani SK, Muthusamy E (2016) Detection of abnormalities in ultrasound lung image using multi-level RVM classification. J Matern Fetal Neonatal Med 29(11):1844–1852

126. Konig T, Steffen J, Rak M, et al. (2015) Ultrasound texture-based CAD system for detecting neuromuscular diseases. Int J Comput Assist Radiol Surg 10(9):1493–1503

127. Srivastava T, Darras BT, Wu JS, Rutkove SB (2012) Machine learning algorithms to classify spinal muscular atrophy subtypes. Neurology 79(4):358–364

128. Sheet D, et al. (2014) Joint learning of ultrasonic backscattering statistical physics and signal confidence primal for characterizing atherosclerotic plaques using intravascular ultrasound. Med Image Anal 18(1):103–117

129. Yu S, Tan KK, Sng BL, Li S, Sia AT (2015) Lumbar ultrasound image feature extraction and classification with support vector machine. Ultrasound Med Biol 41(10):2677–2689

130. Pathak H, Kulkarni V (2015) Identification of ovarian mass through ultrasound images using machine learning techniques. In: Research in Computational Intelligence and Communication Networks (ICRCICN), 2015 IEEE International Conference, pp. 137–140

131. Aramendía-Vidaurreta V, Cabeza R, Villanueva A, Navallas J, Alcázar JL (2016) Ultrasound image discrimination between be-nign and malignant adnexal masses based on a neural network approach. Ultrasound Med Biol 42(3):742–752

132. Subramanya MB, Kumar V, Mukherjee S, Saini M (2015) SVM-based CAC system for B-mode kidney ultrasound images. J Digit Imaging 28(4):448–458

133. Takagi K, Kondo S, Nakamura K, Takiguchi M (2014) Lesion type classification by applying machine-learning technique to contrast-enhanced ultrasound images. IEICE Trans Inf Syst E97D(11):2947–2954

134. Caxinha M, et al. (2015) Automatic cataract classification based on ultrasound technique using machine learning: a comparative study. Phys Procedia 70:1221–1224

135. Sjogren AR, Leo MM, Feldman J, Gwin JT (2016) Image segmentation and machine learning for detection of abdominal free fluid in focused assessment with sonography for trauma examinations: a pilot study. J Ultrasound Med 35(11):2501–2509