

引用格式: 衡佳鸣, 王宁浩, 董凤林, 等. 基于深度学习的超声成像技术研究现状[J]. 声学技术, 2023, 42(2): 174-183. [HENG Jiaming, WANG Ninghao, DONG Fenglin, et al. Research status of deep learning based ultrasonic imaging[J]. Technical Acoustics, 2023, 42(2): 174-183.] DOI: 10.16300/j.cnki.1000-3630.2023.02.008

基于深度学习的超声成像技术研究现状

衡佳鸣^{1,2}, 王宁浩^{1,2}, 董凤林³, 简小华^{1,2}

(1. 中国科学技术大学生物医学工程学院(苏州), 生命科学与医学部, 江苏苏州 215000;

2. 中国科学院苏州生物医学工程技术研究所, 江苏苏州 215163; 3. 苏州大学附属第一医院, 江苏苏州 215000)

摘要: 超声成像技术以其无损、无辐射、实时性好、成像深度深和检查费用低等优点, 是目前临床医学最常应用的影像技术之一。随着医疗技术的日益发展, 兼顾高帧率和高质量的超声成像已成为临床上的迫切需求。深度学习作为能快速提取信号特征的技术, 近年来已经在超声成像领域展开多种应用研究并产生了较好的效果, 具有很大的应用前景。文章总结了基于深度学习的超声成像技术现状, 重点介绍了超声成像技术中常用深度学习架构, 列举了深度学习技术在超声成像中的应用, 最后总结了目前深度学习在训练以及实现临床应用中可能会遇到的挑战。

关键词: 深度学习; 超声成像; 超声波束合成; 神经网络

中图分类号: TB553

文献标志码: A

文章编号: 1000-3630(2023)-02-0174-10

Research status of deep learning based ultrasonic imaging

HENG Jiaming^{1,2}, WANG Ninghao^{1,2}, DONG Fenglin³, JIAN Xiaohua^{1,2}

(1. School of Biomedical Engineering (Suzhou), Division of Life Sciences and Medicine, University of Science and Technology of China, Suzhou 215000, Jiangsu, China;

2. Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou 215163, Jiangsu, China;

3. The first affiliated hospital of Soochow University, Suzhou 215000, Jiangsu, China)

Abstract: Ultrasonic imaging is one of the most often used clinical imaging and diagnostic methods because of its moderate, real-time, well penetrative and cost-efficient performances. In recent years, the high frame-rate and high quality ultrasonic imaging method is urgently demanded. Meanwhile, as a popular tool to acquire signal features in a short period, deep learning has made impressive improvements and shown great potential in ultrasonic imaging. In this paper, the deep learning method and its applications related to ultrasonic imaging are reviewed, and challenges in deep learning training and clinical practices are also summarized.

Key words: deep learning; ultrasonic imaging; ultrasonic beamforming; neural network

0 引言

超声成像具有经济、便捷、安全(非电离性、无辐射性)等优点, 是目前临床应用最广泛和普及的影像学手段之一。随着超声成像技术和系统工艺的不断进步, 超声成像变得更加便携化、智能化和专业化, 超声成像技术在临床诊疗上的应用场景也更加广泛。然而, 由于个体差异、组织环境的复杂、声速的不均匀分布以及声波在组织中的散射现象, 超声在人体组织环境内的传播十分复杂^[1], 导致在一定帧率要求的情况下, 超声图像的质量较

差, 超声图像辨识困难, 提高了对诊疗医师的要求。由于诊疗医师的临床经验和知识储备不同, 超声成像诊疗结果往往有很大的差异。为了降低误诊率, 提高诊疗时效性, 获取实时高质量的超声图像已成为临床的迫切需求。

在传统超声成像技术里, 电子相控阵聚焦成像已经成为超声成像标准成像方法之一^[2]。其采用延迟激励阵元的方法使各阵元按设计的延时对指定位置依次先后发射声波, 然后接收各阵元产生的后向散射回波数据进行滤波、增益、波束合成、包络检波等操作, 最终获取成像结果。在线性扫描模式下, 超声成像传输时间为^[3]

$$t_a = (2 \times d / c + t_d) \times n \quad (1)$$

式中: t_a 为传输时间; c 为介质中的声速; d 为扫描深度; t_d 为延时等待时间; n 为扫描线数。由于若干阵元组成一组工作并形成一条扫描线, 阵元数目越多, 阵元宽度越大, 聚焦效果越好, 图像的横向分辨率越高, 但获得的信息量越大, 后期处理也越

收稿日期: 2021-10-11; 修回日期: 2021-11-15

基金项目: 国家自然科学基金(21927803); 江苏省科技项目(BE2022118, BZ2021037); 中科院项目(Y20196120A122062ZY)。

作者简介: 衡佳鸣(1996—), 女, 吉林吉林人, 硕士研究生, 研究方向为超声成像。

通信作者: 简小华, E-mail: jianxh@sibet.ac.cn

费时。因此，在扫描深度一定时，成像算法需要在帧率和分辨率之间进行权衡。

为了提高图像帧率和图像质量，许多学者对超声成像算法进行了优化改进，提出了多线采集(Multi-line Acquisition, MLA)^[4]和多线传输(Multi-line Transmission, MLT)^[5]等技术来提高成像帧率。此外，为了保证阵元数目的同时提高帧率，还提出了利用非聚焦波束如平面波、发散波来实现超快超声成像(Ultrafast Ultrasound Imaging)^[6]。然而这种非聚焦成像会导致声束宽度扩大，影响成像的分辨率，故需要采用额外的处理技术对回波进行处理，如相干合成技术(Coherence-based Beamforming)^[7]等。由于这种合成方法产生的数据量较大，会带来很大的通信负担与计算量、增加波束形成的时间，故而无法将非聚焦成像的帧率优势全部发挥出来。另一方面，为了实现高质量超声成像，一些自适应波束形成算法也被提出，如最小方差法(Minimum Variance, MV)^[8-9]以及延时-相乘-叠加(Delay Multiply and Sum, DMAS)^[10]方法等。这些方法提供了更高的图像可信度、更窄的主瓣和更小的旁瓣，但相应地也带来了更大的计算量^[3]，过大的计算量导致这些算法很难得到临床应用和普及。

深度学习是一种从数据中直接处理和自动学习特征的技术，其指数级增长的表达能力可有效降低计算维度，被视为提高医学成像质量以及速度的有效方法之一。近年来深度学习技术在正电子发射型计算机断层显像(Positron Emission Computed Tomography, PET)^[11]、核磁共振(Magnetic Resonance Imaging, MRI)^[12]、计算机断层扫描(X-ray Computed Tomograph, CT)^[13-14]和光声成像(Photoacoustic Tomography, PAT)^[15]等领域得到了快速发展。截至2020年12月，已有两家公司研发的基于深度学习的CT成像算法获得了美国食品药品监督管理局(Food and Drug Administration, FDA)的批准^[16]。而在超声成像领域，深度学习技术也得到了应用并产生较好的效果。例如：2017年Gasse等^[17]提出了基于深度学习的超声成像算法，利用3个平面波合成实现了相当于31个传统平面波合成的效果。2020年，Bell等为了促进基于深度学习的超声成像技术发展^[18]，开展了名为“深度学习超声波束形成挑战”的挑战，围绕基于深度学习的平面波成像和电子聚焦成像技术，思考实现图像帧率与质量的平衡以及实现更快、更简单动态聚焦成像的方法。通过其成像结果发现，基于深度学习的超声成像技术在保持散斑信噪比的同时有效改善了成像质量^[19]。相较于传统超声成像算法，基于深度学习的超声成像算法可以

从大量训练集数据中提取肉眼难以发现的抽象特征和学习图像间的映射关系，从而实现传统成像算法进行优化。此外，基于深度学习的超声成像算法还可以实现从原始射频数据到图像特征的直接分割。这种方法以原始射频数据为输入，提供了更多分割所需要的特征信息并且可以泛化到不同数据集中。

目前，在已有的一些基于深度学习超声成像技术的综述中，有的总结了深度学习技术在具体某超声成像领域的应用，如平面波超声成像技术领域^[20]；有的根据超声成像技术的流程讨论相应的深度学习解决方案^[21]；有的涉及整个医学分析领域^[22]。与这些文章相比，本文从深度学习技术的原理及应用出发，总结了超声成像技术中常用的深度学习架构，列举介绍了深度学习技术在超声成像中的应用和其面临的挑战，并进行了总结和展望。

1 超声成像下的深度学习技术

深度学习技术作为人工智能以及机器学习的子集，是对传统神经网络的进一步优化。该技术规避了传统算法对人工定义和特征提取的依赖，通过多层神经元网络形成的深度神经网络(Deep Neural Network, DNN)实现了类似通用函数近似器的效果。最初的深度神经网络采用一种全连接架构，以五层全连接深度神经网络为例，其架构如图1所示。

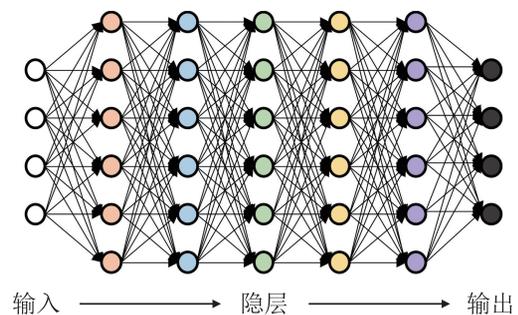


图1 五层全连接深度神经网络架构

Fig.1 Architecture of five-layer fully-connected deep neural network

首先，通过合理确定神经网络的层数以及每层的宽度近似构建目标函数。然后，将大量训练集输入到神经网络中，使神经网络能够自主学习并不断调整更新自身的参数。最终，找到输入到输出的最优映射关系，利用此时的参数搭建函数近似器来进行实际应用。训练一旦结束，所形成的神经网络常能以非常快的速度实现输入到输出的转换^[2]。但是需要注意，这种泛化能力强、具有高鲁棒性的神经网络是以牺牲可解释性为代价来

获得的^[23]。

除了全连接深度神经网络外，深度神经网络还有许多常用的神经网络结构与模型。目前基于深度学习的超声成像技术主要采用自编码网络(Autoencoder Network, AE network)、卷积神经网络(Convolutional Neural Network, CNN)、生成对抗网络(Generative Adversarial Network, GAN)以及它们的变体。接下来将详细介绍超声成像中常用的神经网络结构模型。

1.1 自编码网络

Rumelhart 团队在 1986 年的《Nature》中首次提出了自编码网络的概念^[24]，而后 Bourlard 等^[25]对其进行了进一步补充阐释。自编码网络构架图如图 2 所示。

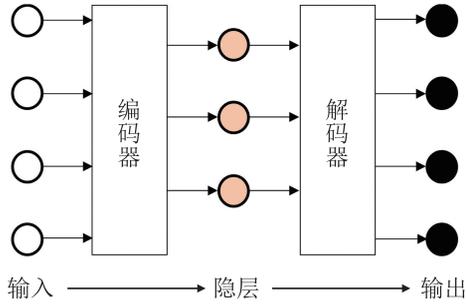


图2 自编码网络架构
Fig.2 The architecture of AE network

传统的自编码网络是一种自监督模型，训练集中的输入作为最终输出的对比标签。通过将输入样本作为网络的期望输出，对输入数据进行降维和升维操作来学习输入样本的抽象特征^[26]和输入到输出的映射关系。自编码网络主要包含编码器和解码器两部分。

编码器定义为

$$\mathbf{h} = f(\mathbf{x}) = s_c(\mathbf{W}\mathbf{x} + \mathbf{b}_c) \quad (2)$$

式中： \mathbf{h} 为编码器输出， s_c 为非线性激活函数， \mathbf{W} 为权重矩阵， \mathbf{x} 为编码器输入， \mathbf{b}_c 为编码器输入偏差。

解码器定义为

$$\mathbf{y} = g(\mathbf{h}) = s_d(\mathbf{W}^T\mathbf{h} + \mathbf{b}_d) \quad (3)$$

式中： \mathbf{y} 为解码器输出， s_d 为非线性激活函数， \mathbf{W}^T 为权重矩阵 \mathbf{W} 的转置矩阵， \mathbf{h} 为解码器输入(编码器输出)， \mathbf{b}_d 为解码器输入偏差。

编码器和解码器是对称结构，输入通过编码器网络进行压缩降维，然后通过解码器部分实现图像恢复，最后通过损失函数不断进行参数优化，实现最小失真^[27]。自编码网络的损失函数为

$$L_{AE}(\theta) = \arg \min_{\theta} \left[\sum_{\mathbf{x} \in D_n} L(\mathbf{x}, \mathbf{y}) \right] \quad (4)$$

式中： $L_{AE}(\theta)$ 为自编码网络的损失函数，用来优化参数 θ ，参数 θ 包括权重矩阵 \mathbf{W} 以及编码器解码器的偏差 \mathbf{b}_c 和 \mathbf{b}_d ， D_n 为训练集输入。

超声成像技术中常用的自编码网络变体技术包括堆积降噪自编码网络(Stacked Denoising Autoencoders, SDA)^[28-29]和卷积自编码网络(Convolutional Autoencoder, CAE)^[30]等。其中，堆积降噪自编码网络相当于将多个降噪自编码网络(Denoising Autoencoder, DAE)^[31]叠加在一起。这种降噪自编码网络以一定的概率分布将输入的部分节点置0，再投入到训练中，增加了神经网络的鲁棒性。卷积自编码网络将编码器与解码器中使用的全连接层替换为卷积块，这种方式会考虑输入信号之间的联系，有效提高神经网络的性能。此外，对输入数据进行编码解码这一自编码网络的中心思想还用在了全连接卷积神经网络中，相关内容将在 1.2 节进行介绍。

1.2 卷积神经网络

Lecun 等在 1998 年首次提出了卷积神经网络^[32]。这类网络源于人的视觉认知方式，具有分层学习特征的能力^[33]。卷积神经网络由多个模块连在一起，每个模块中包含卷积层和池化层，一些卷积神经网络还会加入激励层和归一化层等。此外，为了进行特征整合还会在尾端加入全连接层。卷积神经网络架构图如图 3 所示。

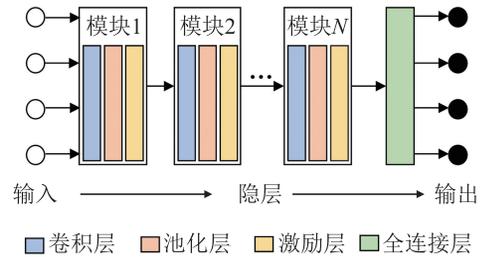


图3 卷积神经网络架构
Fig.3 The architecture of CNN network

卷积层定义：

$$\mathbf{h}_k^{(i)} = \mathbf{W}_k * \mathbf{x}_k^{(i)} + \mathbf{b}_k^{(i)} \quad (5)$$

式中：假设将输入信号分为 n 块，特征空间(核数)分为 m 个， $i \in [1, n]$ ， $k \in [1, m]$ 。 $\mathbf{h}_k^{(i)}$ 为卷积输出值， \mathbf{W}_k 为特征矩阵， $\mathbf{x}_k^{(i)}$ 为输入子块， $\mathbf{b}_k^{(i)}$ 为偏差。卷积层将输入信号分成小块依次与特征矩阵做卷积，实现特征图的提取。由于卷积运算代替了传统神经网络的乘法运算，对于同一特征卷积运算权值相同，降低了运算的复杂度^[34]。

池化层定义(以最大池化为例)：

$$\mathbf{g}_k^{(j)} = \max(\mathbf{H}_k^{(j)}) \quad (6)$$

式中：假设将卷积后输出分为 p 块， $j \in [1, p]$ ； $\mathbf{g}_k^{(j)}$

为池化后输出； $H_k^{(l)}$ 为卷积输出子块。池化层以一定规则(如最大池化)对特征图进行下采样，实现维度降低同时保持平移不变性，有效防止过拟合。

由于卷积神经网络常为有监督模型，神经网络的输出会同训练集提供的参考标签作比较来实现参数优化更新。卷积神经网络的损失函数为

$$L_{CON}(\theta) = \arg \min_{\theta} f(\hat{y}_k, y_k) \quad (7)$$

式中： $L_{CON}(\theta)$ 为卷积神经网络的损失函数，用来更新参数 θ ，参数 θ 包括特征矩阵 W_k 以及偏差 b_k 。 \hat{y}_k 为卷积神经网络输出， y_k 为标签输出。

为了让卷积神经网络提取出来的尺度特征能够映射回原信号，文献[35]中提出了一种全卷积神经网络(Fully Convolutional Networks, FCN)。全连接神经网络前半部分同卷积神经网络一致，不同在于全连接神经网络在全连接层前又引入了新模块。这些新模块中包含上采样层，使降维特征能够恢复到原图像大小。全连接卷积神经网络架构如图4所示。

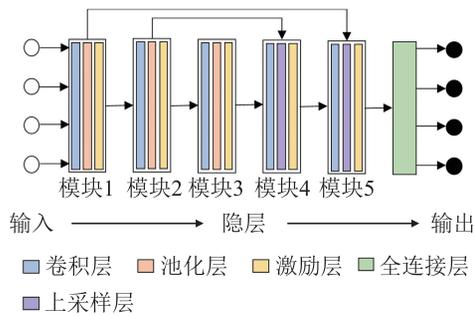


图4 全卷积神经网络架构
Fig.4 The architecture of FCN network

Ronneberger 等^[36]进一步对全连接神经网络进行了优化，形成了U-Net网络架构，这种网络架构引入了编码解码思想，在全连接神经网络所引入的新模块中加入了激励层来提取上采样层特征。U-Net网络架构图如图5所示。

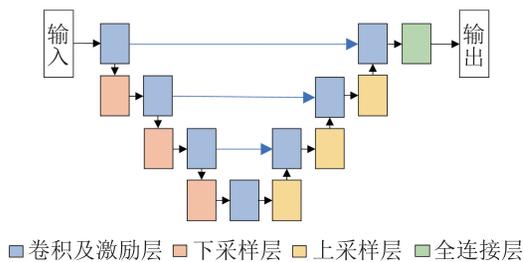


图5 U-Net网络架构图
Fig.5 The architecture of U-Net network

1.3 生成对抗网络

生成对抗网络是由文献[37]中首次提出，该网

络采用一种对抗博弈的思想，由生成模型和判别模型两部分组成，通过两部分模型的不断优化更新，获得最优估计结果。生成对抗网络的网络架构如图6所示。

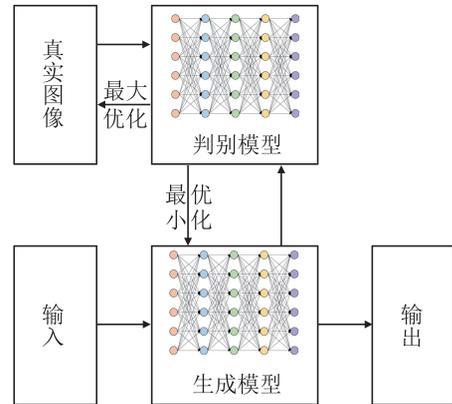


图6 生成对抗网络模型
Fig.6 The architecture of GAN network

首先将训练数据输入到生成模型中，然后将其输出结果送入判别模型行中进行优化估计，当生成模型优化到一定程度时，判别模型也会进行优化更新，如此反复，优化过程表示为

$$\min_G \max_D V(D, G) = F_{x \sim p_{data}(x)} [\lg D(x)] + F_{z \sim p_z(z)} \{1 - D[G(z)]\}$$

式中： $V(D, G)$ 为损失函数， D 为判别模型， G 为损失模型， $p_{data}(x)$ 为真实概率分布， $p_z(z)$ 为判别模型模拟产生的概率分布。

2 深度学习下超声成像技术

目前深度学习主要通过三种方式改善超声成像技术：优化传统超声成像算法、提取图像映射关系、噪声抑制。此外，基于深度学习的超声成像算法还可以实现由射频信号到超声图像，甚至到图像分割结果的直接转换。

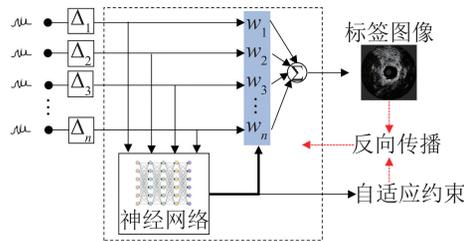
2.1 基于深度学习的超声成像应用

2.1.1 优化传统成像算法

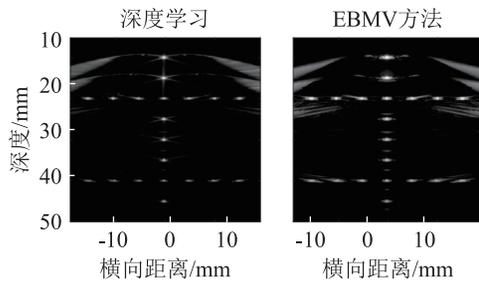
由于深度神经网络具有通用函数近似器的功能，深度学习技术可以在传统成像算法的基础上进行优化成像，设计深度神经网络构架优化或者替代传统成像过程^[2,38-42]。

深度学习技术可以通过学习给出超声成像算法的自适应最优参数。Luijten 等^[2]提出了自适应波束形成深度学习算法(Adaptive Beamforming by Deep Learning, ABLE)，通过全连接深度神经网络学习

MV波束形成方法,利用神经网络自适应获得最优变迹加权参数。该方法效果优于传统波束形成方法,甚至产生了比训练标签获取所用的基于特征的最小方差波束合成算法(Eigen-Based Minimum Variance Beamformer, EBMV)^[8]更高的横向分辨率。同时在计算速度上比EBMV算法提高了约400倍,系统流程及测试结果如图7所示。



(a) 基于自适应波束形成深度学习算法的波束合成流程



(b) 深度学习和EBMV方法波束合成结果

图7 自适应波束形成深度学习算法流程及结果^[2]

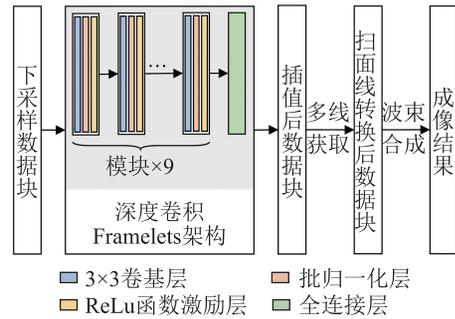
Fig.7 Flow of ABLE algorithm and its processing results^[2]

深度学习技术还可以实现传统成像算法的近似,并以更快的速度进行算法运算。Yoon等^[42]发现并将射频数据、汉克尔矩阵(Hankel Matrix)和深度学习^[43-44]结合在一起,采用一种带有编解码思想的卷积神经网络——深度卷积小帧架构近似基于湮没滤波器的低秩汉克尔矩阵(Annihilating Filter-based Low rank Hankel Matrix, ALOHA)实现射频数据的插补,然后通过多线采集和波束合成实现成像。同ALOHA结果相比,该方法有效提高了图像质量,重建时间也减少了几个数量级。以下采样射频数据插补为例,具体处理流程及测试结果如图8所示。其中,图8(b)中上图为B超成像结果,下图为不同算法下生成的下采样射频数据图^[42]。

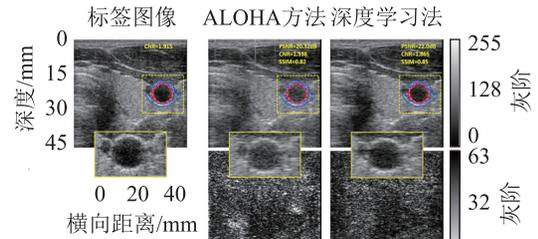
2.1.2 提取图像映射关系

深度学习技术可以通过获取低质量图像到高质量图像的非线性映射关系来进行神经网络的训练。所形成的神经网络可以有效提高低质量图片的分辨率及对比度^[6,17,45-49]。

深度学习技术可以学习高质量非聚焦波束的复合效果,并将其应用在更少的非聚焦波束合成中。Gasse等^[17]和Lu等^[6]通过卷积神经网络,学习3个非



(a) 基于深度卷积小帧架构的波束合成流程



(b) 颈动脉B超的下采样射频数据波束形成结果

图8 深度卷积小帧架构算法流程及结果^[42]

Fig.8 Flow of deep convolutional framelet architecture based algorithm and its processing results^[42]

聚焦波束复合成像图像与31个对应波束复合成像图像的映射关系,优化平面波与发散波的空间合成运算,最终用深度学习下3个波束复合成像实现了传统算法^[50]下31个对应波束复合成像效果。以平面波合成为例,测试结果如图9所示。图9中从左到右依次为传统下3平面波合成,深度学习下3平面波合成及传统下31平面波合成方法^[17]。

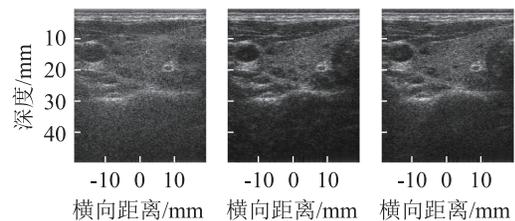


图9 基于卷积神经网络的平面波成像算法结果^[17]

Fig.9 Processing results of CNN based plane wave imaging^[17]

深度学习技术还可以通过学习高质量图像到低质量图像的映射关系,获得最优变迹权值。Vedula等^[45]和Senouf等^[46]针对多线传输和多线采集技术图像质量优化展开研究,通过U-Net架构获取MLT/MLA图像与单线传输(Single-line Transmission, SLT)/单线采集(Single-line Acquisition, SLA)图像的映射,获取MLT/MLA的自适应变迹权值,最终形成更高质量的超声成像图像。以MLT优化算法为例,MLT算法采用多个聚焦在不同方向的窄波束同时发射,考虑到波束之间的相互作用该方法会产生串扰伪影,但同时也会带来较高的能量。基于深度学习的MLT算法以SLT对应生成的MLT仿真数据

为输入，以对应的无伪影 SLT 图像为输出进行网络训练，算法流程和测试结果如图 10 所示，其中测试结果为同时发射 6 个不同方向的聚焦波束后回波信号经过 Tucky 窗和深度学习算法后的成像结果^[45]。

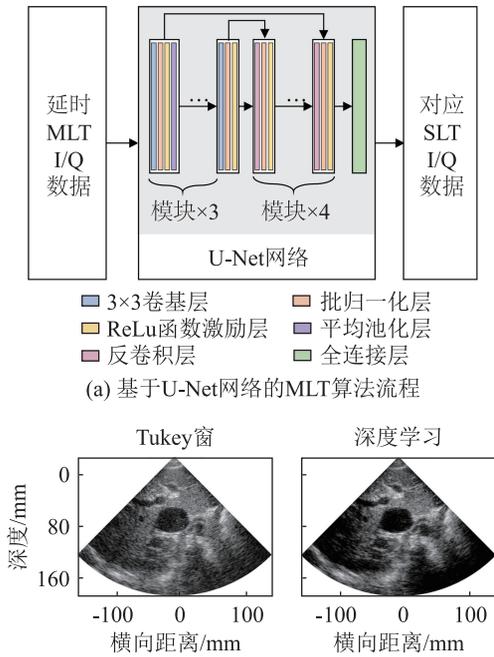


图 10 基于 U-Net 网络的 MLT 算法流程及结果^[45]

Fig.10 Flow of U-net based MLT algorithm and its processing results^[45]

2.1.3 噪声抑制

深度学习技术也是实现超声图像噪声抑制，减少图像退化的有效手段。传统超声图像噪声抑制方法常在后处理滤波器中进行斑纹抑制，然而由于波束形成过程中丢失了信道信息和相位信息，成像信息无法在斑纹抑制时得到完全利用。而将深度学习技术同传统噪声抑制方法结合起来，则可以通过原始射频数据在波束合成阶段实现噪声抑制，使成像信息得到充分利用。此外，深度学习技术还可以获取含噪声数据与无噪声数据之间的联系，并利用神经网络表达^[39,51-52]。

深度学习技术可以与传统噪声抑制方法相结合实现噪声抑制。由于孔径域模型图像重建算法^[53-54] (Aperture Domain Model Image Reconstruction, ADMIRE)将波束合成问题看作是一种非线性回归问题，Luchies 等^[51]将该方法的思想引用到深度学习技术中，通过抑制超声波束主瓣区域以外的信号再形成超声图像来提高图像质量，实现离轴散射和混响的抑制。为实现该任务，该方法将射频信号进行短期傅里叶变换，将变换后的虚部和实部作为输

入送入前馈全连接神经网络，然后通过短期傅里叶反变换将输出恢复为已经抑制离轴散射的射频信号。该方法有效提高了超声图像对比度，并成功抑制了囊肿内的噪声，算法流程及测试结果如图 11 所示。

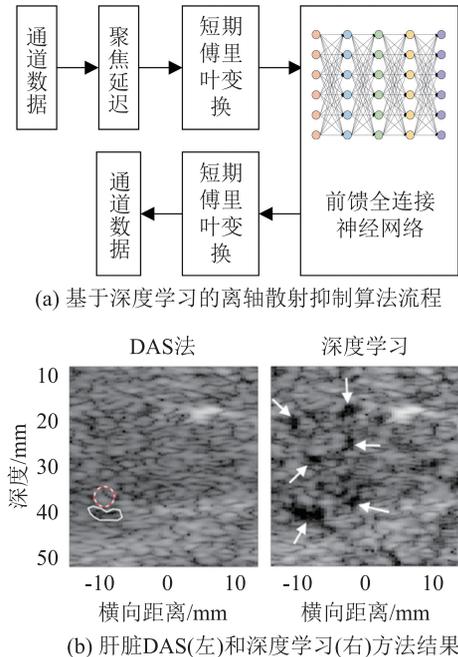


图 11 基于深度学习的离轴散射抑制算法流程及结果^[51]

Fig.11 Flow of deep learning based off-axis scattering suppression algorithm and its processing results^[51]

深度学习技术还可以通过学习含噪声图像与不含噪声图像之间的关系，通过神经网络进行噪声抑制。Hyun 等^[52]以射频信号为输入，采用深度卷积神经网络，将无散斑图像作为标签图像，通过对经过神经网络后成像的超声图像与无散斑图像之间的差距设置损失函数，进行参数优化，减少 B 超图像的散斑噪声。该算法最终实现了同优化的贝叶斯非局部均值 (Optimized Bayesian Nonlocal Means, OBNLM) 算法^[55]相当的分辨率，测试结果如图 12 所示。图 12 中，从左到右、从上到下依次为 DAS、空间组合 (Spatial Compounding, SC)、OBNLM 及深度学习方法^[52]。

2.2 实现射频数据直接成像

基于深度学习的超声成像技术常以射频(子)数据作为输入，通过神经网络产生输出。常见输出形式如经过插值修复后的射频数据^[42,51,56]、自适应优化参数及变迹权重^[2,40,57]、最终成像图像和图像分割结果等^[1,38,58-60]。其中，由射频数据到超声图像的直接生成技术，跳过了传统波束形成的方法，作为一种信号端到图像端的成像技术，体现了深度学习强大的生成能力以及特征提取能力。

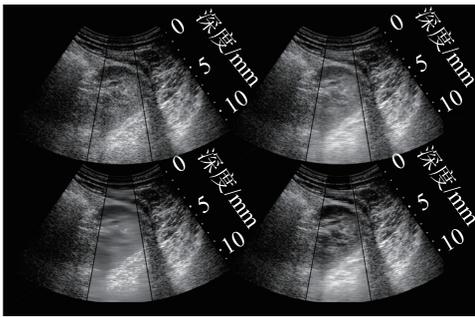
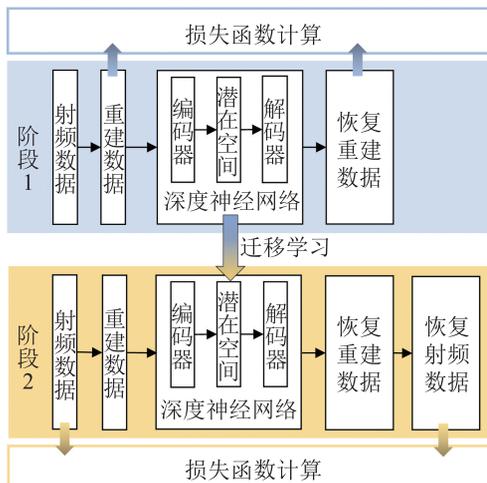


图 12 基于卷积神经网络的散斑噪声减少结果^[52]

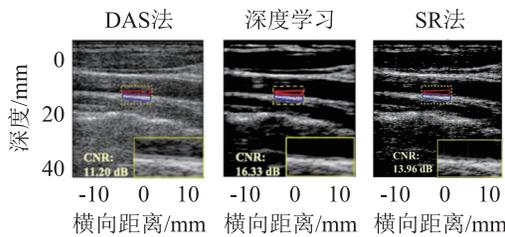
Fig.12 Processing results of CNN based speckle reduction algorithm^[52]

Luo 等^[1]将图像重建问题看作一种逆问题，采用自监督深度神经网络，利用类似堆积自编码网络实现从射频数据到图像的端到端直接恢复。该算法通过前层网络进行初始化，然后采用迁移学习，将前层网络参数迁移至后层网络，开始后层网络训练。在后层网络中损失函数引入了稀疏正则化参量 (Sparse Regularization, SR) 提高重建图像的质量和稳定性。与传统 SR 方法^[61]相比，该方法在单平面波条件下实现了与其相当的图像质量，并显著缩短了重构时间。系统流程图及测试结果如图 13 所示。

深度学习超声成像技术还可以从射频数据直接



(a) 基于自监督深度神经网络超声图像重建算法流程



(b) 单平面波图像重建结果^[61]

图 13 基于自监督深度神经网络的超声图像重建流程及结果^[1]
Fig.13 Flow of self-supervised deep neural network based ultrasonic image reconstruction algorithm and its processing results^[1]

实现图像分割。Nair 等^[59]采用 U-Net 网络架构实现了从平面波射频数据到超声图像分割结果的直接转换。该算法通过神经网络学习射频数据的多尺度特征，由全连接层对这些多尺度特征进行整合计算，获得图像各像素点对应囊肿区域与组织区域的像素置信值，形成相应的分割图像。系统测试结果如图 14 所示。图 14 中，从左到右、从上到下依次为射频数据，波束合成结果，实际分割图像和深度学习方法下图像分割结果^[59]。

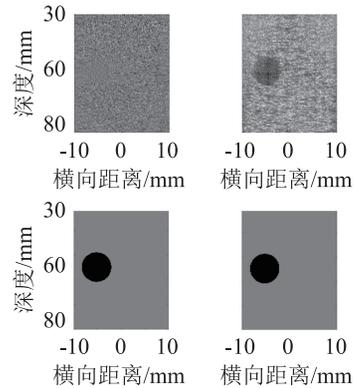


图 14 基于 U-Net 网络超声成像分割算法结果^[59]

Fig.14 Processing results of U-Net based ultrasonic image segmentation algorithm^[59]

3 基于深度学习的超声成像技术面临的挑战

超声成像领域应用深度学习技术能够有效的提高成像速度和成像质量。当然，目前基于深度学习的超声成像技术还面临着许多挑战，主要集中在以下方面：

(1) 缺乏公开的训练集

由于基于深度学习的超声成像技术研究正在起步阶段，部分成像方式缺乏公开可用的训练集进行测试，并以相同的参考数据评估新的深度学习方法性能。目前仅针对平面波及聚焦成像技术提供了可以作为标准的训练集^[18]。而其他相关成像技术如发散波成像等的标准训练集还有待构建。

(2) 训练标签获取困难

由于人体环境复杂，很难获取到可以作为标准的参考超声图像。目前，对于训练标签的选取问题解决方案主要分为 3 种。① 通过 Field II 软件仿真获取训练图像和训练标签，并将训练模型放在真实的体内超声图像与仿体上进行验证^[51-52]，这种方法虽然可以验证训练标签的准确度，但很难证明该方法是否可以推广到实际应用。② 以成像质量好的超声成像算法得到的超声图像作为训练标签，对深

深度学习算法进行训练，如提取图像映射关系的算法以及以图像为输出的算法^[2-3]等。这种方法的成像质量往往会受到其作为标准的成像算法质量的限制，很难发挥出深度学习技术的全部效果。③以射频数据作为标签^[42]，该算法以获取的射频数据作为参考，对其进行下采样获得下采样射频数据作为输入，进行以射频数据插值为目的的卷积神经网络训练。此外，文献[1]提出了一种新的解决方法，通过自监督深度神经网络，将输入本身作为输出，并在损失函数中加入正则化项提高成像质量。

(3) 深度学习如何解释

深度学习技术作为黑盒技术，通过训练集不断对神经网络进行训练，实现神经网络参数最优化，训练好的神经网络可以根据输入自适应输出。由于神经网络采用的是神经元提取特征的思想，故最终训练好的网络很难通过数学方式进行解释。而算法可以进行解释是能临床应用的前提。为了提高医生和患者对含有深度学习技术的系统的信任度，解释深度学习框架很有必要。目前，基于深度学习的超声成像算法^[6,46]常通过计算结构相似性指标值(Structural Similarity Index Measure, SSIM)，判断所设计算法与参考对照算法的相似度，并以此作为算法的解释。此外，Yoon等^[42]提出的算法以ALOHA方法为参考，设计了新的深度学习框架并对该框架进行了数学理论验证。Luchies等^[51]通过三个方面解释了前馈深度神经网络的训练过程。

4 结论

本文针对基于深度学习的超声成像技术的现状展开研究，总结了超声成像领域常应用的深度学习技术及其应用现状。基于深度学习技术的超声成像技术通过训练自适应学习改变自身参数，形成特定的信号处理器，在提高成像分辨率和对比度的同时加快了计算速度，并能有效抑制杂波。相比传统超声成像技术，深度学习技术一方面可以有效地提高超声成像的速度，有望使计算量大的高质量成像算法实现实时成像。另一方面，深度学习技术还可以通过学习高质量超声图像的图像特征，在高帧率下产生更高质量的超声图像。深度学习技术还可以实现从原始射频数据到图像分割结果的直接转换。

此外，深度学习技术不仅是提高超声成像质量及速度的有效方法，而且也将为超声领域带来新的发展。随着超声成像技术的日益发展，如3D、实时3D、平面波成像、弹性成像等各种功能和新型成像技术导致超声成像产生的图像数据量越来越

大，通过深度学习技术包括卷积、递归和自动编码器或生成对抗网络等方法，可以有效地改善相关超声成像算法，将这些数据进行有效地降维压缩，使数据得到充分利用，从而提高成像质量。另一方面，对于超声图像处理，深度学习技术不仅可以从射频数据入手进行图像成像及处理操作，减少图像成像时带来的数据损失，而且其在后期图像处理的组织器官分割、组织表征、特征匹配、辅助诊断方面都有很大的应用空间，并已有很多成功的研究应用，将最终改善超声诊断和治疗的效果。

最后，虽然目前基于深度学习的超声成像技术还面临不少挑战，但随着深度学习技术和超声成像技术的不断发展和深入研究，基于深度学习的超声成像技术有望在临床上实现应用和推广普及，并拓展超声成像技术的临床应用领域。

参 考 文 献

- [1] ZHANG J K, HE Q, XIAO Y, et al. Ultrasound image reconstruction from plane wave radio-frequency data by self-supervised deep neural network[J]. *Medical Image Analysis*, 2021, **70**: 102018.
- [2] LUIJTEN B, COHEN R, DE BRUIJN F J, et al. Adaptive ultrasound beamforming using deep learning[J]. *IEEE Transactions on Medical Imaging*, 2020, **39**(12): 3967-3978.
- [3] SIMSON W, PASCHALI M, NAVAB N, et al. Deep learning beamforming for sub-sampled ultrasound data[C]//2018 IEEE International Ultrasonics Symposium (IUS). Kobe, Japan. IEEE, 2018: 1-4.
- [4] SHATTUCK D P, WEINSHENKER M D, SMITH S W, et al. Exploscan: a parallel processing technique for high speed ultrasound imaging with linear phased arrays[J]. *The Journal of the Acoustical Society of America*, 1984, **75**(4): 1273-1282.
- [5] MALLART R, FINK M. Improved imaging rate through simultaneous transmission of several ultrasound beams[C]//SPIE Proceedings, New Developments in Ultrasonic Transducers and Transducer Systems. San Diego, CA, USA. SPIE, 1992, **1733**:120-130.
- [6] LU J F, MILLIOZ F, GARCIA D, et al. Reconstruction for diverging-wave imaging using deep convolutional neural networks[J]. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2020, **67**(12): 2481-2492.
- [7] MONTALDO G, TANTER M, BERCOFF J, et al. Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography[J]. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2009, **56**(3): 489-506.
- [8] DEYLAMI A M, JENSEN J A, ASL B M. An improved minimum variance beamforming applied to plane-wave imaging in medical ultrasound[C]//2016 IEEE International Ultrasonics Symposium (IUS). Tours, France. IEEE, 2016: 1-4.
- [9] KIM K, PARK S, KIM J, et al. A fast minimum variance beamforming method using principal component analysis[J]. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2014, **61**(6): 930-945.
- [10] MATRONE G, SAVOIA A S, CALIANO G, et al. The delay multiply and sum beamforming algorithm in ultrasound B-

- mode medical imaging[J]. *IEEE Transactions on Medical Imaging*, 2015, **34**(4): 940-949.
- [11] READER A J, CORDA G, MEHRANIAN A, et al. Deep learning for PET image reconstruction[J]. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 2021, **5**(1): 1-25.
- [12] CHANDRA S S, BRAN LORENZANA M, LIU X W, et al. Deep learning in magnetic resonance image reconstruction[J]. *Journal of Medical Imaging and Radiation Oncology*, 2021, **65**(5): 564-577.
- [13] SINGH R, WU W W, WANG G, et al. Artificial intelligence in image reconstruction: the change is here[J]. *Physica Medica*, 2020, **79**: 113-125.
- [14] WANG G, YE J C, DE MAN B. Deep learning for tomographic image reconstruction[J]. *Nature Machine Intelligence*, 2020, **2**(12): 737-748.
- [15] DAVOUDI N, DEÁN-BEN X L, RAZANSKY D. Deep learning optoacoustic tomography with sparse data[J]. *Nature Machine Intelligence*, 2019, **1**(10): 453-460.
- [16] ARNDT C, GÜTTLER F, HEINRICH A, et al. Deep learning CT image reconstruction in clinical practice[J]. *RoFo: Fortschritte Auf Dem Gebiete Der Rontgenstrahlen Und Der Nuklearmedizin*, 2021, **193**(3): 252-261.
- [17] GASSE M, MILLIOZ F, ROUX E, et al. High-quality plane wave compounding using convolutional neural networks[J]. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2017, **64**(10): 1637-1639.
- [18] BELL M A L, HUANG J Q, HYUN D, et al. Challenge on ultrasound beamforming with deep learning (CUBDL)[C]//2020 IEEE International Ultrasonics Symposium (IUS). Las Vegas, NV, USA. IEEE, 2020: 1-5.
- [19] HYUN D, WIACEK A, GOUDARZI S, et al. Deep learning for ultrasound image formation: CUBDL evaluation framework and open datasets[J]. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2021, **68**(12): 3466-3483.
- [20] 张经科, 何琼, 罗建文. 平面波超声成像中的波束合成方法研究进展[J]. *应用声学*, 2021, **40**(1): 22-32.
ZHANG Jingke, HE Qiong, LUO Jianwen. Research progress of beamforming methods in plane-wave ultrasound imaging [J]. *Journal of Applied Acoustics*, 2021, **40**(1): 22-32.
- [21] VAN SLOUN R J G, COHEN R, ELDAR Y C. Deep learning in ultrasound imaging[J]. *Proceedings of the IEEE*, 2020, **108**(1): 11-29.
- [22] 施俊, 汪琳琳, 王珊珊, 等. 深度学习在医学影像中的应用综述 [J]. *中国图象图形学报*, 2020, **25**(10): 1953-1981.
SHI Jun, WANG Linlin, WANG Shanshan, et al. Applications of deep learning in medical imaging: a survey[J]. *Journal of Image and Graphics*, 2020, **25**(10): 1953-1981.
- [23] AKKUS Z, CAI J, BOONROD A, et al. A survey of deep-learning applications in ultrasound: artificial intelligence - powered ultrasound for improving clinical workflow[J]. *Journal of the American College of Radiology*, 2019, **16**(9): 1318-1328.
- [24] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J]. *Nature*, 1986, **323**(6088): 533-536.
- [25] BOURLARD H, KAMP Y. Auto-association by multilayer perceptrons and singular value decomposition[J]. *Biological Cybernetics*, 1988, **59**(4): 291-294.
- [26] 袁非牛, 章琳, 史劲亭, 等. 自编码神经网络理论及应用综述 [J]. *计算机学报*, 2019, **42**(1): 203-230.
YUAN Feiniu, ZHANG Lin, SHI Jinting, et al. Theories and applications of auto-encoder neural networks: a literature survey[J]. *Chinese Journal of Computers*, 2019, **42**(1): 203-230.
- [27] RIFAI S, VINCENT P, MULLER X, et al. Contractive auto-encoders: explicit invariance during feature extraction[C]// *Proceedings of the 28th International Conference on International Conference on Machine Learning*. Bellevue, Washington, USA. Omnipress, 2011: 833-940.
- [28] VINCENT P, LAROCHELLE H, LAJOIE I, et al. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion[J]. *Journal of Machine Learning Research*, 2010, **11**: 3371-3408.
- [29] MOUSAVI A, PATEL A B, BARANIUK R G. A deep learning approach to structured signal recovery[C]//2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton). Monticello, IL, USA. IEEE, 2016: 1336-1343.
- [30] MJ, MEIER U, CIREŞAN D, et al. Stacked Convolutional Auto-Encoders for Hierarchical Feature Extraction[C]// *International Conference on Artificial Neural Networks*. Berlin, Heidelberg, Germany: Springer, 2011: 52-59.
- [31] VINCENT P, LAROCHELLE H, BENGIO Y, et al. Extracting and composing robust features with denoising autoencoders[C]//*Proceedings of the 25th International Conference on Machine Learning*. Helsinki, Finland: ACM, 2008: 1096-1103.
- [32] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, **86**(11): 2278-2324.
- [33] 李彦冬, 郝宗波, 雷航. 卷积神经网络研究综述 [J]. *计算机应用*, 2016, **36**(9): 2508-2515, 2565.
LI Yandong, HAO Zongbo, LEI Hang. Survey of convolutional neural network[J]. *Journal of Computer Applications*, 2016, **36**(9): 2508-2515, 2565.
- [34] LIU W B, WANG Z D, LIU X H, et al. A survey of deep neural network architectures and their applications[J]. *Neurocomputing*, 2017, **234**: 11-26.
- [35] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]// *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston, MA, USA. IEEE, 2015: 3431-3440.
- [36] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]//*International Conference on Medical Image Computing and Computer-Assisted Intervention*. Munich, Germany. Springer, 2015: 234-241.
- [37] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. *Communications of the ACM*, 2020, **63**(11): 139-144.
- [38] ZHUANG R X, CHEN J Y. Deep learning based minimum variance beamforming for ultrasound imaging[C]//*International Workshop on Preterm, Perinatal and Paediatric Image Analysis, International Workshop on Smart Ultrasound Imaging*. Shenzhen, China. Springer, 2019: 83-91.
- [39] MOR E, BAR-HILLEL A. A unified deep network for beamforming and speckle reduction in plane wave imaging: a simulation study[J]. *Ultrasonics*, 2020, **103**: 106069.
- [40] WIACEK A, GONZALEZ E, BELL M A L. CohereNet: a deep learning architecture for ultrasound spatial correlation estimation and coherence-based beamforming[J]. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2020, **67**(12): 2574-2583.
- [41] KHAN S, HUH J, YE J C. Deep learning-based universal beamformer for ultrasound imaging[C]//*International Confer-*

- ence on Medical Image Computing and Computer-Assisted Intervention. Shenzhen, China. Springer, 2019: 619-627.
- [42] YOON Y H, KHAN S, HUH J, et al. Efficient B-mode ultrasound image reconstruction from sub-sampled RF data using deep learning[J]. *IEEE Transactions on Medical Imaging*, 2019, **38**(2): 325-336.
- [43] JIN K H, HAN Y S, YE J C. Compressive dynamic aperture B-mode ultrasound imaging using annihilating filter-based low-rank interpolation[C]//2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). Prague, Czech Republic. IEEE, 2016: 1009-1012.
- [44] YE J C, HAN Y, CHA E. Deep convolutional framelets: a general deep learning framework for inverse problems[J]. *SIAM Journal on Imaging Sciences*, 2018, **11**(2): 991-1048.
- [45] VEDULA S, SENOUF O, ZURAKHOV G, et al. High quality ultrasonic multi-line transmission through deep learning [C]//International Workshop on Machine Learning for Medical Image Reconstruction. Granada, Spain. Springer, 2018: 147-155.
- [46] SENOUF O, VEDULA S, ZURAKHOV G, et al. High frame-rate cardiac ultrasound imaging with deep Learning[C]// International Conference on Medical Image Computing and Computer-assisted intervention. Granada, Spain. Springer, 2018: 126-134.
- [47] ZHOU Z X, WANG Y Y, YU J H, et al. High spatial-temporal resolution reconstruction of plane-wave ultrasound images with a multichannel multiscale convolutional neural network [J]. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2018, **65**(11): 1983-1996.
- [48] ZHANG X, LI J, HE Q, et al. High-quality reconstruction of plane-wave imaging using generative adversarial network[C]//2018 IEEE International Ultrasonics Symposium (IUS). Kobe, Japan. IEEE, 2018: 1-4.
- [49] GHANI M U, CAN MERAL F, VIGNON F, et al. High frame-rate ultrasound imaging using deep learning beamforming[C]//2019 IEEE International Ultrasonics Symposium (IUS). Glasgow, UK. IEEE, 2019: 295-298.
- [50] LU J Y. Experimental study of high frame rate imaging with limited diffraction beams[J]. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 1998, **45**(1): 84-97.
- [51] LUCHIES A C, BYRAM B C. Deep neural networks for ultrasound beamforming[J]. *IEEE Transactions on Medical Imaging*, 2018, **37**(9): 2010-2021.
- [52] HYUN D, BRICKSON L L, LOOBY K T, et al. Beamforming and speckle reduction using neural networks[J]. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2019, **66**(5): 898-910.
- [53] BYRAM B, JAKOVLJEVIC M. Ultrasonic multipath and beamforming clutter reduction: a chirp model approach[J]. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2014, **61**(3): 428-440.
- [54] BYRAM B, DEI K, TIERNEY J, et al. A model and regularization scheme for ultrasonic beamforming clutter reduction [J]. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2015, **62**(11): 1913-1927.
- [55] COUPÉ P, HELLIER P, KERVRANN C, et al. Nonlocal means-based speckle filtering for ultrasound images[J]. *IEEE Transactions on Image Processing*, 2009, **18**(10): 2221-2229.
- [56] GOUDARZI S, ASIF A, RIVAZ H. Ultrasound beamforming using MobileNetV2[C]//2020 IEEE International Ultrasonics Symposium (IUS). Las Vegas, NV, USA. IEEE, 2020: 1-4.
- [57] CHEN Y R, LIU J, LUO X B, et al. ApodNet: learning for high frame rate synthetic transmit aperture ultrasound imaging [J]. *IEEE Transactions on Medical Imaging*, 2021, **40**(11): 3190-3204.
- [58] WANG Y N, KEMPSKI K, KANG J U, et al. A conditional adversarial network for single plane wave beamforming[C]//2020 IEEE International Ultrasonics Symposium (IUS). Las Vegas, NV, USA. IEEE, 2020: 1-4.
- [59] NAIR A A, TRAN T D, REITER A, et al. A deep learning based alternative to beamforming ultrasound images[C]//2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Calgary, AB, Canada. IEEE, 2018: 3359-3363.
- [60] DU B, WANG J Y, ZHENG H T, et al. A novel transcranial ultrasound imaging method with diverging wave transmission and deep learning approach[J]. *Computer Methods and Programs in Biomedicine*, 2020, **186**: 105308.
- [61] WANG C Z, PENG X, LIANG D, et al. An easily-achieved time-domain beamformer for ultrafast ultrasound imaging based on compressive sensing[C]//Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Milan, Italy. IEEE, 2015: 7490-7493.