

面向近垂直场景的RD图多任务运动信息提取方法

刘子源^① 王绍平^② 何屹廷^② 谷源涛^{*①}

^①(清华大学电子工程系 北京 100084)

^②(清华大学深圳国际研究生院 深圳 518055)

摘要: 脉冲多普勒雷达具备全天候工作能力, 可通过距离-多普勒(RD)图像同时获取目标的距离与速度信息。在近垂直飞行场景下, RD图的几何结构中蕴含飞行平台的高度、速度及俯仰角等关键运动参数, 但这些参数在RD域中呈现明显耦合, 传统信号处理方法在复杂地形和近垂直入射条件下难以实现有效解耦。近年来, 深度学习在运动信息感知领域展现出显著优势, 但多任务学习在该场景下仍难以同时满足实时性与高精度要求。为此, 该文提出一种新型的网络结构——RDMFNet, 通过共享编码器与并行解码器实现多表征信息融合, 并采用两阶段渐进式训练策略, 以提高参数估计的精度。实验表明, RDMFNet在高度、速度及俯仰角估计上的误差分别降低至14.447 m, 4.635 m/s和0.755°, 展现了其在高精度实时感知中的优势。

关键词: 运动参数估计; 近垂直场景; 距离-多普勒图; 多任务学习; 深度学习

中图分类号: TN957.52

文献标识码: A

文章编号: 2095-283X(2026)03-0909-10

DOI: 10.12000/JR25259

CSTR: 32380.14.JR25259

引用格式: 刘子源, 王绍平, 何屹廷, 等. 面向近垂直场景的RD图多任务运动信息提取方法[J]. 雷达学报(中英文), 2026, 15(3): 909–918. doi: 10.12000/JR25259.

Reference format: LIU Ziyuan, WANG Shaoping, HE Yiting, *et al.* A multitask motion information extraction method based on range-Doppler maps for near-vertical scenarios[J]. *Journal of Radars*, 2026, 15(3): 909–918. doi: 10.12000/JR25259.

A MultiTask Motion Information Extraction Method Based on Range-Doppler Maps for Near-vertical Scenarios

LIU Ziyuan^① WANG Shaoping^② HE Yiting^② GU Yuantao^{*①}

^①(Department of Electronic Engineering, Tsinghua University, Beijing 100084, China)

^②(Tsinghua Shenzhen International Graduate School, Shenzhen 518055, China)

Abstract: Pulse Doppler radar provides all-weather operational capability and enables simultaneous acquisition of target range and velocity through Range-Doppler (RD) maps. In near-vertical flight scenarios, the geometric structure of RD maps implicitly encodes key platform motion parameters, including altitude, velocity, and pitch angle. However, these parameters are strongly coupled in the RD domain, making effective decoupling difficult for traditional signal-processing-based inversion methods, particularly under complex terrain and near-vertical incidence conditions. Although recent advances in deep learning have shown strong potential for motion information sensing, multitask learning in this context still faces challenges in achieving both real-time performance and high estimation accuracy. To address these issues, this study proposes a novel network architecture, termed Range-Doppler Map Fusion Network (RDMFNet), that performs multirepresentation information fusion via shared encoders and parallel decoders, along with a two-stage progressive training

收稿日期: 2025-12-01; 改回日期: 2026-01-11; 网络出版: 2026-02-06

*通信作者: 谷源涛 gyt@tsinghua.edu.cn

*Corresponding Author: GU Yuantao, gyt@tsinghua.edu.cn

基金项目: 国家自然科学基金(U2230201)

Foundation Item: The National Natural Science Foundation of China (U2230201)

责任主编: 丁一鹏 Corresponding Editor: DING Yipeng

©The Author(s) 2026. This is an open access article under the CC-BY 4.0 License
(<https://creativecommons.org/licenses/by/4.0/>)

strategy to enhance parameter estimation accuracy. Experimental results show that RDMFNet achieves estimation errors of 14.447 m for altitude, 4.635 m/s for velocity, and 0.755° for pitch angle, demonstrating its effectiveness for high-precision, real-time perception.

Key words: Motion parameter estimation; Near-vertical scenario; Range-Doppler map; Multitask learning; Deep learning

1 引言

近垂直飞行平台(如无人机、直升机、航天器等)在执行精准着陆、灾害救援等典型任务时,其运动状态的高精度实时感知是控制系统闭环的核心前提^[1]。当前主流的运动参数估计系统通常采用惯性导航、激光/雷达测距等多源异构传感器的融合架构^[2]。其中,惯性测量单元(Inertial Measurement Unit, IMU)虽具有高频输出特性,但其位姿解算依赖加速度与角速度的积分运算,不可避免地存在积分误差累积效应,且易受机械振动与环境温度扰动影响^[3],依赖其他传感器的定期校准。光学测距设备(如激光测距仪)虽能获取沿光束轴向的相对距离,但无法直接解算垂向绝对高度^[4];而传统雷达高度计在复杂地形条件下会出现显著的测量偏差。由此可见,单一传感模态难以实现高度、速度与姿态的协同精确测量。

脉冲多普勒雷达具有全天候工作能力和多维信息感知特性,可通过距离-多普勒(Range-Doppler, RD)成像同时获取目标的距离与径向速度信息^[5]。理论研究显示,近垂直飞行条件下的RD图像几何特征蕴含着平台高度、速度及俯仰角等关键运动参数^[6]。然而,受限于RD图像的分辨率及地形敏感性,基于传统信号处理的参数反演方法存在两方面固有缺陷:其一,复杂地形会引起回波能量分布的畸变,导致高度估计精度恶化;其二,俯仰角与高度、速度参数在RD域呈现强耦合特性,难以建立解析解耦模型。

近年来,深度学习为解决上述非线性估计问题提供了新范式^[7]。研究表明,深度神经网络通过多层次特征提取可有效挖掘图像中隐含的特征^[8]。例如, Kwon等人^[9]使用ResNet从二维磁性体系实验观测获得的磁畴图像中反演磁哈密顿参量。尽管已有研究证实了数据驱动方法的可行性,但多任务联合学习^[10]仍面临严峻的理论挑战:当网络同时回归高度、速度与俯仰角时,不同任务梯度方向的竞争会导致模型陷入次优收敛状态。不确定性加权损失(Uncertainty Weighted Loss, UCLoss)函数^[11]通过为每个任务的损失项添加可学习的权重缓解了优化冲突,不过实验结果表明该方法虽然能使各个任务的估计性能变得更加平衡,但仍然无法达到每个任务均最优的效果。

针对以上挑战,为了系统性探索近垂直场景下基于RD图反演飞行器运动参数的可行性,本研究提出一种新型网络结构Range-Doppler Map Fusion Network (RDMFNet),其创新性体现在两个层面:(1)网络架构层面,设计了基于共享编码器与并行解码器的异构分支结构,通过特征层共享与任务层解耦实现多表征信息的有效融合;(2)训练策略层面,创造性地采用两阶段渐进式学习方法,首阶段通过预训练各编码器与解码器分支建立基础特征空间,次阶段冻结编码器进行精细化微调,从而实现多任务梯度的独立优化。实验表明,所提方法在高度、速度及俯仰角估计上的误差仅有14.447 m, 4.635 m/s和 0.755° ,相较于传统信号处理方法,在高度、速度估计误差上分别下降了33.26%和70.68%。同时RDMFNet与对比网络相比在网络参数量和推理时间上均有明显优势。

2 RD图分析与仿真建模

2.1 近垂直场景RD图回波形态分析

近垂直雷达场景如图1所示,其中 h , v , p 分别表示平台的高度、速度和俯仰角,本研究将 $p \in [70^\circ, 90^\circ]$ 定义为近垂直观测区间。基于雷达成像的几何原理,该区间内独特的RD图像形成机制导致了显著的参数耦合效应。

在理想平面地形条件下,雷达观测几何满足轴对称特性。此时,等距离线表现为以雷达星下点为中心的同心圆族,而等速度线符合双曲函数分布规律^[12]。特别地,当飞行平台实施严格垂直运动时,平台速度矢量与雷达视轴共线,导致等速度线为同心圆结构。在此特殊情况下,单个等距离环内仅存

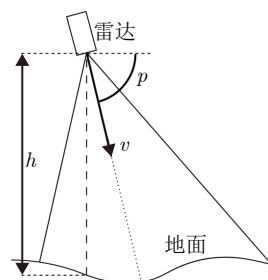


图1 近垂直雷达场景示意

Fig. 1 An illustration of a near-vertical radar scenario

在有限数量的多普勒分辨单元，反映在RD图像上表现为回波能量呈现高度聚集特性，使得高度与速度参数难以通过单一维度的谱峰特征进行解耦。

随着平台俯仰角偏离 90° ，等速度线的几何构型演化为椭圆，其离心率 e 与俯仰角 p 满足 $e = V_p \sin(p) / V_g$ 的函数关系(其中 V_p 为平台速度， V_g 为等速线的速度)。这种近垂直条件下的微小角度变化将引起多普勒谱宽的剧烈波动，导致单位等距离环内多普勒单元数的非线性增长。如图2所示，当俯仰角由 90° 降至 70° 时，图2(a)中用绿色标注的距离环内的有效多普勒单元数增加至图2(b)中的6个。相应地，RD图像中的回波能量分布呈现显著的弥散特点。

2.2 仿真RD图数据集生成

针对近垂直观测条件下高速运动平台雷达回波特性研究的样本稀缺性问题，本研究采用文献[13]提出的地杂波建模方法生成RD图数据集。仿真系统主要配置如表1所示，所生成的RD图包含251个距离单元，101个多普勒单元，距离分辨率与速度分辨率为

$$\Delta R = \frac{c}{2f_s} = 24 \text{ m} \quad (1)$$

$$\Delta V = \frac{c}{2(M-1)T_r f_c} = 37.5 \text{ m/s} \quad (2)$$

为实现模型的地形泛化能力，从ALOS PALSAR

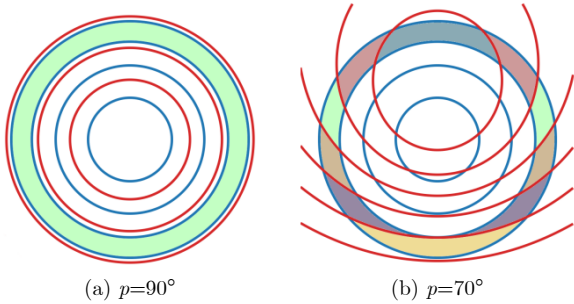


图2 近垂直场景等距离线(蓝色)与等速度线(红色)示意图

Fig. 2 A schematic illustration depicting isorange contours (in blue) and isovelocity contours (in red) in a near-vertical scenario

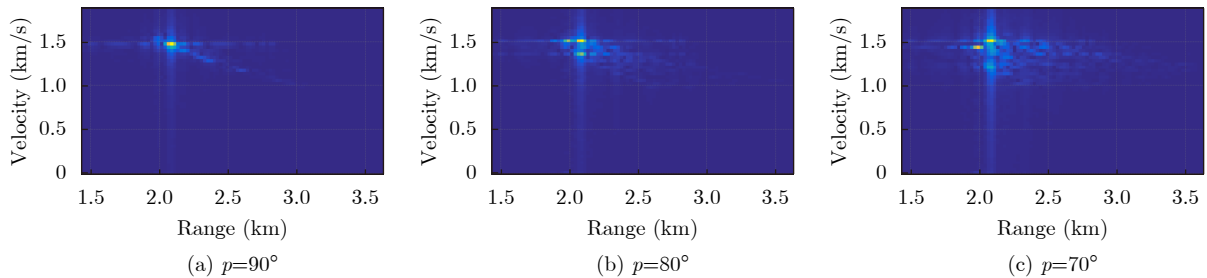


图3 相同高度和速度不同俯仰角下的仿真RD图

Fig. 3 Simulated RD maps of the same altitude and velocity with different pitch angles

高精度数字高程模型产品中选取200个具有地貌代表性的样本，其地表起伏度参数满足 $\Delta H \in [0, 500] \text{ m}$ 且空间分辨率为12.5 m。数据集按3:1:1比例将地形样本分配至训练集(120)、验证集(40)及测试集(40)。为确保训练数据能够充分覆盖近垂直场景下的物理状态空间，并避免模型对特定飞行轨迹的时间相关性产生过拟合，对于每个地形样本，我们在预设的动力学边界内通过均匀随机采样生成100组运动参数组合：

平台高度 $h \sim \mathcal{U}(1000, 2000) \text{ m}$

运动速度 $v \sim \mathcal{U}(1000, 1800) \text{ m/s}$

俯仰角 $p \sim \mathcal{U}(70, 90)^\circ$

最终构建包含20,000组RD图像的仿真数据集。

图3展示了典型仿真结果：图3(a)—图3(c)子图固定 $h = 2000 \text{ m}$ ， $v = 1500 \text{ m/s}$ ，当俯仰角从 90° 递减至 70° 时，回波形态逐渐发散，符合理论分析结果。

3 算法介绍

3.1 RDMFNet网络结构

图4为本文提出的RDMFNet网络结构示意图，其基础模块(用直角矩形表示)由卷积层(Conv)、批归一化层[14](Batch Normalization, BN)与参数化修正线性单元[15](PReLU)串联构成。各卷积层的参数详见表2(其中各卷积层与图4的颜色对应， C_{in} 、 C_{out} 分别表示输入/输出通道数， ks 为卷积核尺寸)。本节按数据流方向依次说明网络核心设计。

表1 仿真中设定的雷达参数

Tab. 1 Radar parameters set in simulaiton

参数	取值
载频 f_c	1 GHz
脉冲发射间隔 T_r	4E-5 s
采样率 f_s	6.25 MHz
发射信号时宽 T	5E-6 s
发射信号带宽 B	2.5 MHz
相参积累脉冲数 M	101

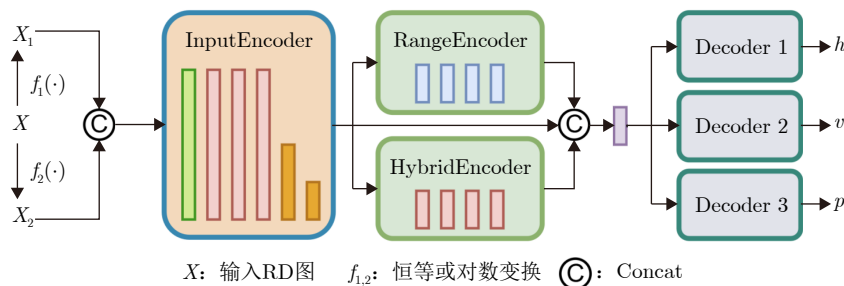


图4 RDMFNet网络结构

Fig. 4 Network architecture of RDMFNet

表2 RDMFNet各卷积层参数

Tab. 2 Parameters of the Convolutional Layers in RDMFNet

卷积层	C_{in}, C_{out}	ks (3, 3)	stride	pad
卷积1	(2, C)	(3, 3)	(1, 1)	(1, 1)
卷积2	(C, C)	(3, 3)	(1, 1)	(1, 1)
卷积3	(C, C)	(3, 3)	(2, 2)	(1, 1)
卷积4	(C, C)	(1, 3)	(1, 1)	(0, 1)
卷积5	($3C, C$)	(1, 1)	(1, 1)	(1, 1)

输入预处理模块：平台对地近垂直飞行场景中，地物相对平台的速度为正，首先对原始RD图进行速度维截断，仅保留正多普勒区域的特征图 $X \in \mathbb{R}^{50 \times 251}$ 以避免冗余计算，其中50, 251分别对应多普勒维与距离维。为了研究不同输入模式对参数估计性能的影响，本文对 X 分别施加两类预处理算子 f_1, f_2 ，并在通道维拼接得到网络输入：

$$X_{cat} = \text{Concat}(f_1(X), f_2(X)) \in \mathbb{R}^{50 \times 251 \times 2} \quad (3)$$

本文选择恒等映射(Raw)和对数变换(Log-Scaled)作为这两个算子。Raw模式直接使用经过脉冲压缩与相参积累处理后生成的RD图，不做额外变换，能够保持回波的原始幅度分布与距离维定位信息：

$$f_1(X) = X \quad (4)$$

Log-Scaled模式通过压缩动态范围，增强了低能量区域的频谱细节与回波弥散形态，有助于依赖多普勒谱宽与几何轮廓的速度 v 及俯仰角 p 的估计，定义为

$$f_2(X) = \log(X + \epsilon) \quad (5)$$

其中， $\epsilon = 10^{-6}$ 用于避免对接近零的噪声像素求对数导致数值不稳定。为了消除数值量级差异并加速网络收敛，在执行通道拼接之前，我们分别对经 f_1 和 f_2 变换后的图像独立执行了Min-Max归一化，使其数值分布映射至 $[0, 1]$ 。实验部分将系统分析Raw+Raw, Raw+Log-Scaled, Log-Scaled+Log-Scaled 3种组合模式对各参数估计的影响。

输入编码模块：InputEncoder的第1个卷积层

将特征图通道数扩展到 $C=32$ ，然后经过3个卷积块进行初级特征提取，最后经过两次步长为2的下采样将特征图空间维度压缩至原始尺寸的1/4 (13×63) 从而显著降低后续的计算复杂度。

特征编码模块：网络中间的编码器部分由RangeEncoder和HybridEncoder构成。考虑到高度信息主要依赖距离维的定位，且易受脉冲压缩旁瓣干扰，RangeEncoder采用4层膨胀卷积结构 ($ks = 1 \times 3$, dilation rate依次为1, 2, 4, 8)，通过逐步扩大的感受野捕获距离维度的多尺度特征，避免引入多普勒维度的噪声干扰。鉴于近垂直场景下速度与俯仰角表现为耦合的几何形态，HybridEncoder使用 3×3 标准卷积核实现距离-多普勒特征的融合，以弥补RangeEncoder对耦合特征提取能力的不足，同时通过跳跃连接来保留浅层空间特征。三者的输出经过通道拼接后使用 1×1 的卷积实现特征整合并将通道数调整回 C 。

输入编码和特征编码均为组合输入的共享编码器，原因在于(1)共享编码器显著减少了模型的总参数量与计算资源。(2)提升扩展能力：在需要处理RD图序列时，编码器的共享结构避免了为多张RD图单独分配独立编码器带来的资源消耗。(3)增强任务关联性：在多任务处理场景下，编码器的共享策略有助于多维度信息的综合分析，从而提高参数估计的精度。

解码输出模块：为实现 h, v, p 的独立估计，采用图5所示的并行解码架构。3个解码器具有相同的结构，特征图经过3个卷积层后，通过自适应平均池化将特征维度降到 C ，然后经过两个全连接层输出估计结果。

3.2 训练方法

在端到端训练中，3个参数的损失权重难以平衡，且各参数的收敛速度存在显著差异，本文在RDMFNet的基础上，提出了两阶段训练方法。第1阶段进行多参数联合训练，损失函数为各参数损失的加权，训练过程中记录各参数的验证集误差，

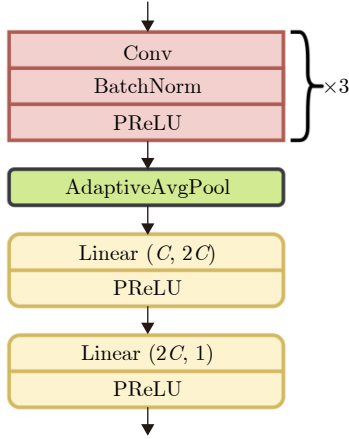


图5 解码器结构

Fig. 5 The architecture of the decoders

当某个参数 θ_i 的验证误差不再降低时,说明该参数已经收敛,此时停止训练,记此时的模型 M_{θ_i} 。第2阶段以 M_{θ_i} 作为初始权重,并冻结网络的编码器,此时由于各个decoder彼此隔离,实质上能够独立训练每个decoder,从而避免了损失竞争的问题。在具体实现中,在前向传播后计算出各参数的MSE损失,然后对这些损失分别执行反向传播,即可避免每个decoder都单独训练的开销。

3.3 损失函数

由于3个参数的取值范围差异巨大,训练网络前需要将各个参数归一化到[0, 1]:

$$\hat{y}_1 = \frac{h - h_{\min}}{h_{\max} - h_{\min}} \quad (6)$$

同理得到归一化后的速度和俯仰角 \hat{y}_2, \hat{y}_3 。设 y_1, y_2, y_3 表示网络输出,由于已事先对不同量纲的参数进行了归一化处理,本文采用3个任务的平均MSE作为损失函数:

$$L_{\text{MSE}} = \frac{1}{3} \sum_{i=1}^3 (y_i - \hat{y}_i)^2 \quad (7)$$

实验中发现使用MSE损失很难达到3个参数估计均最优的效果,因此本文还引入不确定性损失:

$$L_{\text{UC}} = \sum_{i=1}^3 \left(e^{-s_i} (y_i - \hat{y}_i)^2 + s_i \right) \quad (8)$$

其中, $s_1 \sim s_3$ 是可学习的参数,通过动态地调整各个任务损失的权重,达到平衡多个任务的效果。

4 实验结果与分析

4.1 实验参数设置

本研究所有实验均采用统一的训练配置。优化器选用Adam,其参数设置为 $\beta_1 = 0.9, \beta_2 = 0.999$,

批量大小(batch size)为16。学习率采用余弦退火策略进行调度,初始学习率设置为 2×10^{-4} ,最终学习率降至 1×10^{-6} ,总训练轮次(epoch)为100。所有实验均在NVIDIA A800 GPU上完成,深度学习框架采用PyTorch。

在两阶段训练策略中,第1阶段进行100个epoch的训练。如图6所示,图中蓝色实线绘制了网络的3个解码器(decoder)对应的输出 h, v, p 的验证集误差曲线。实验结果表明,3个参数的收敛速率存在显著差异。其中, p 在第33个epoch左右已基本收敛,而 h, v 需训练至接近80个epoch才能达到稳定状态,此时 p 已经出现明显的过拟合。因此,在第1阶段训练的100个epoch中,我们选择 p 性能最优的模型权重作为第2阶段的模型初始化权重。在第2阶段训练中,仅对3个decoder进行微调。由于 p 在第1阶段已实现良好收敛,为防止其过拟合,需设置较小的学习率;而 h 和 v 尚未完成学习,需采用较大的学习率。因此,我们调整学习率设置 h 和 v 继续采用余弦退火策略(初始学习率 1×10^{-4} ,最终学习率 1×10^{-6}),而 p 的学习率固定为 1×10^{-6} ,以避免其已收敛的特征表示发生显著变化。

图6红色虚线展示了第2阶段各个参数在验证集上的误差曲线,其纵坐标为归一化数值下的均方根误差(Root Mean Square Error, RMSE)。结果表明,引入第2阶段微调后,3个参数的估计误差均快速下降,且没有出现过拟合现象,这表明我们的策略在保证 h 和 v 收敛的同时,有效维持了 p 已学习到的特征,从而提升了模型的整体性能。

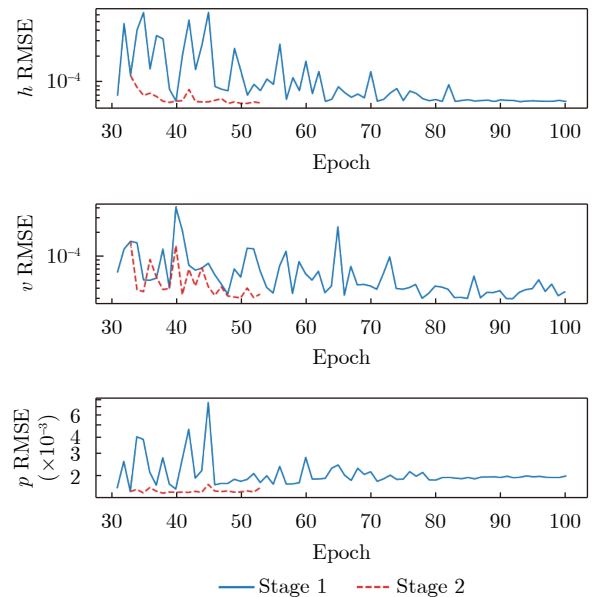


图6 两阶段训练过程中验证集上误差曲线

Fig. 6 RMSE curves on the validation set during the two-stage training process

4.2 对比模型介绍

为了验证所提出的RDMFNet的有效性, 本研究选择了以下算法进行对比实验。

传统算法: 为了构建一个强鲁棒性的对比基线, 本文采用峰值检测配合亚像素插值算法。首先, 对RD图进行与深度模型完全一致的预处理和裁剪操作, 以确保数据分布的一致性。在定位回波峰值所在的距离-多普勒单元后, 引入亚像素插值技术对峰值位置进行精细化修正, 以此作为 h 与 v 的最终估计值, 从而避免网格离散化带来的量化误差。

ResNet: 一种经典的深度神经网络, 通过跳跃连接(Skip Connection)有效缓解深层网络的梯度消失问题, 从而实现稳定训练^[16]。该模型广泛应用于计算机视觉任务, 本文采用ResNet-18版本作为对比基准。

TSCNet: 最初是一种用于语音增强的模型, 其核心模块TSCB利用了语谱图在时间维和频率维上具有不同物理意义的特点, 依次从两个维度提取特征^[17], 这一特性与RD图在距离维与多普勒维上所包含的差异化物理信息相契合, 因此本文将其引入用于运动参数估计。为了将TSCNet用于参数估计, 本文对其原结构进行了必要调整, 去除原用于复数域重建的Complex Decoder分支, 并在Mask Decoder之后新增预测器模块, 以实现高度、速度及俯仰角等参数的输出。原始TSCNet包含4个TSCB, 内部通道数为64。

TSCNet-mini: 鉴于原始TSCNet计算量较大, 推理过程中对时间与显存的需求均较高, 不利于在资源受限平台或具有实时性要求的近垂直场景感知任务中部署。为此, 本文进一步设计了轻量化版本TSCNet-mini, 将TSCB模块数减少至2个, 并将通道数缩减至32。该简化在保持模型对RD图双维度结构特征提取能力的同时, 显著降低了计算成本并提升了推理效率, 使网络更适合本研究中的多任务实时参数估计需求。

各模型的参数量(Params)、计算复杂度(MACs)及推理时间(Inference Time)总结如表3所示, 推理

表3 各模型的参数量, 计算复杂度和推理时间

Tab. 3 The parameter counts, MACs, and inference times for each model

模型	Params (M)	MACs (G)	Time (ms)
ResNet	11.172	0.52	0.881
TSCNet	1.576	6.65	21.984
TSCNet-mini	0.288	1.48	8.141
RDMFNet	0.204	0.51	0.815

时间基于测试100次尺寸为 50×251 的单张图像所得均值。实验结果表明, RDMFNet在参数量、计算复杂度和推理速度方面均表现最优, 体现了其在边缘计算或实时处理场景中的应用潜力。值得注意的是, 尽管ResNet具有最高的参数量(11.172M), 但其推理速度仅次于RDMFNet, 这主要得益于其较低的计算复杂度和硬件优化。而TSCNet的计算复杂度和推理时间均为最高, 限制了其在实时应用中的适用性。然而, TSCNet-mini通过削减网络规模, 在保持一定性能的同时大幅降低计算开销, 使其资源消耗得到明显改善。

4.3 不同方法在不同输入模式下的性能分析

为全面评估与直观了解所提RDMFNet模型的性能, 本文在不同输入模式下对高度 h 、速度 v 、俯仰角 p 进行了预测, 并在真实物理量纲下计算了不同方法的均方根误差(RMSE), 结果如表4所示, 并得出以下结论。

深度学习显著优于传统方法: 从实验结果来看, 所有深度学习对 h 和 v 的预测精度均优于传统方法, 表明基于深度学习的特征提取能力能够有效提升RD图的参数估计精度。

输入模式对预测精度的影响: 本文分别测试了Raw(原始输入)、Log-Scaled(对数输入)、Raw+Log-Scaled(原图-对数组合输入) 3种输入模式对不同模型的影响。与原图输入相比, 对数输入情况下, ResNet, TSCNet, TSCNet-mini模型对 h 的估

表4 各方法的均方根误差

Tab. 4 The RMSE of each method

算法	Input	h (m)	v (m/s)	p ($^{\circ}$)
传统方法	R	21.647	15.809	-
	R	15.928	9.205	0.991
ResNet	L	16.358	7.845	0.855
	R+L	15.886	6.530	0.851
TSCNet	R	<u>14.755</u>	5.852	0.777
	L	16.725	5.392	0.781
TSCNet	R+L	14.911	5.167	0.810
	R	15.432	5.859	0.834
TSCNet-mini	L	16.737	5.417	0.768
	R+L	15.099	5.484	0.788
TSCNet-mini	R	36.407	9.509	1.098
	R	36.407	9.509	1.098
RDMFNet (单阶段)	L	15.195	5.250	0.755
	R+L	14.820	6.341	0.780
RDMFNet (两阶段)	R	15.273	5.252	0.823
	L	14.447	4.635	0.755
RDMFNet (两阶段)	R+L	14.861	<u>4.690</u>	<u>0.758</u>
	R+L	14.861	<u>4.690</u>	<u>0.758</u>

注: 加粗数值表示最优性能, 下划线数值表示第二优性能。

计精度有所降低，但对 v 的估计精度均有所提升，ResNet和TSCNet-mini对 p 的估计性能也变得更好。这一现象的原因在于：高度信息依赖回波在距离维上的位置，但脉冲压缩过程会不可避免地产生旁瓣，使得回波位置存在一定模糊性。对图像取对数后，这种模糊性进一步加剧，导致高度估计精度下降；速度信息主要由多普勒维决定，而多普勒维信息不受脉冲压缩影响，对数处理能够增强回波细节，使得速度信息更容易提取，从而提高估计精度；俯仰角依赖回波区域的整体形态特征，对数变换后，回波形态更加突出，因此俯仰角估计精度提高。所以本文提出的组合输入模式(Raw+Log-Scaled)可在一定程度上弥补单一输入的不足，在ResNet18和TSCNet-mini上表现最佳，在RDMFNet单阶段训练模式下也表现出较好的综合性能。

两阶段训练策略的有效性：本文进一步探讨了兩阶段训练策略对RDMFNet预测精度的影响，实验结果表明，Raw或Log-Scaled作为输入时，两阶段训练均可有效提升 h, p, v 的估计精度；Raw输入时， h 估计的RMSE由36.407下降至15.273(降低58.0%)， v 和 p 的RMSE亦有明显下降，充分验证了微调训练对模型性能的优化作用；Raw+Log-Scaled组合输入时， h 预测RMSE略有上升(增加0.3%)，但 v 和 p 预测精度分别提升26.0%和2.8%，表明该输入模式在整体性能上的优势。

最佳方案选择：综上所述，以Log-Scaled作为输入，并采用两阶段训练的RDMFNet取得了最佳结果，但是其与其他模型的单阶段训练方案中会明显增大 h 的估计误差，Raw+Log-Scaled的输入方案也可以取得较好的结果，在ResNet和TSCNet-mini上得到了最好的表现。以上结论进一步验证了所提方法的有效性和可行性，为RD图参数估计任务提供了高效可靠的解决方案。

4.4 消融实验

4.4.1 网络结构消融实验

为验证所提网络结构的有效性，本文开展了消

融实验。实验采用Raw+Log-Scaled的输入模式，并选用UCLoss作为损失函数。RDMFNet在不同编码器组合下的实验结果如表5所示。

由实验结果可见，移除任何一个模块均会导致模型性能下降，表明各模块在特征提取与预测任务中的重要性。值得注意的是，当去除RangeEncoder后，模型性能下降最为显著。这表明，针对RD图两个维度物理意义不同的特点而设计的单维度特征提取器能够捕捉到距离维上的关键特征，在提升信息提取的精准度和预测精度方面具有重要贡献。HybridEncoder模块通过融合标准卷积，将距离-多普勒维度特征进行整合处理。移除HybridEncoder会导致特征间联系的削弱，影响信息融合的效果。该模块在实现全维度信息的综合分析，提升预测能力方面是不可或缺的。跳跃连接用于保留浅层空间特征并促进深层特征间的有效传递，去掉该模块会明显提高预测误差，这意味着跳跃连接在保持特征完整性和信息传递效率方面起到了关键作用。

特别是在单阶段训练模式下，完整模型在 h, p, v 预测方面均表现最佳。在两阶段微调后，尽管原始模型在 v 和 p 上的RMSE略高于最优结果，但在 h 的预测上仍显著优于缺失模块的情况。这进一步验证了完整网络结构的设计合理性与优越性能。

4.4.2 共享编码消融实验

本节旨在通过消融实验对共享编码器与独立编码器的性能差异进行对比和分析，以验证共享编码器在本任务中的有效性和优越性。共享编码即本文提出的不同输入组合先拼接再送入共享编码器，独立编码是指 $f_1(X)$ 和 $f_2(X)$ 分别送入各自的输入编码和特征编码模块，并且彼此不共享权重。在RDMFNet、单阶段、Raw+Log-Scaled组合下的结果比较如表6所示，采用共享编码策略的网络在多项指标上均优于独立编码策略。具体来说，使用共享编码器的模型不仅具有更小的参数量(0.204 M vs 0.315 M)与推理时间(0.815 ms vs 1.554 ms)，在3项关键参数的估计上也保持了更低的误差水平，高度、速度及俯仰角RMSE分别降低至14.820 m,

表5 移除各个模块后RDMFNet的估计误差

Tab. 5 The RMSE of RDMFNet after the removal of each module

Range Encoder	Hybrid Encoder	Skip Connection	一阶段			两阶段		
			h (m)	v (m/s)	p (°)	h (m)	v (m/s)	p (°)
-	√	√	23.177	<u>9.178</u>	0.790	15.551	4.950	0.792
√	-	√	<u>15.679</u>	9.887	0.787	<u>15.031</u>	4.460	0.769
√	√	-	16.002	23.073	0.759	15.148	5.088	0.740
√	√	√	14.820	6.341	0.780	14.861	<u>4.690</u>	<u>0.758</u>

注：加粗数值表示最优性能，下划线数值表示第二优性能。“-”表示去除该模块，“√”表示添加该模块。

表6 共享编码与独立编码下的误差结果比较

Tab. 6 Comparison of RMSE between shared encoding and independent encoding

编码策略	Params (M)	MACs (G)	Time (ms)	h (m)	v (m/s)	p (°)
共享编码	0.204	0.51	0.815	14.820	6.341	0.780
独立编码	0.315	1.17	1.554	15.123	6.440	0.792

注：加粗数值表示最优性能。

6.341 m/s和0.780°，验证了共享编码器在近垂直场景下的多任务学习环境中的有效性。

4.4.3 损失函数消融实验

本节进一步对损失函数进行了消融实验。为了验证UCLoss在本任务中的优越性，我们还引入了经典的多任务学习算法，动态损失加权^[18](Dynamic Loss Scaling, DLS)方法作为对比，即每个任务仍使用MSE损失，记为 L_{MSE}^i ，但是每项损失归一化后再反向传播梯度并更新网络参数，具体公式为

$$L_{DLS} = \frac{L_{MSE}^1}{|L_{MSE}^1|} + \frac{L_{MSE}^2}{|L_{MSE}^2|} + \frac{L_{MSE}^3}{|L_{MSE}^3|} \quad (9)$$

在应用DLS损失时，我们将高度、速度、俯仰角的归一化步骤从预处理阶段移至损失函数内进行。前述实验已验证，对于ResNet和TSCNet-mini而言，Raw+Log-Scaled作为输入模式最为合适；而对于RDMFNet，Raw+Log-Scaled和Log-Scaled两种输入均能取得良好效果。因此，我们针对这些输入设置，对比了UCLoss，DLSLoss和MSELoss的影响。

如表7所示，对于ResNet，TSCNet-mini以及单阶段训练的RDMFNet，UCLoss均表现出最佳的预测性能。在经过两阶段微调的RDMFNet中，MSELoss在RMSE指标上略占优势。尽管DLSLoss也展示了良好的性能，但在多数情况下，其效果不如UCLoss，这进一步验证了结合参数归一化与不确定性损失的组合策略的优越性。

综合而言，实验结果证实了UCLoss在多任务平衡训练中的有效性，其能够在绝大多数情况下提升模型的预测性能，显示出其在多任务学习中的应用潜力。

4.4.4 训练策略消融实验

为验证本文所提两阶段训练策略中阶段划分方式及epoch设置的合理性，本节对不同训练策略进行了消融实验。除单阶段训练和本文提出的“ p 收敛+微调”策略外，还进一步引入了“过拟合+微调”和“欠拟合+微调”两种对照方案，以全面评估不同初始化点对第2阶段训练效果的影响。

过拟合+微调：第1阶段训练100个epoch，使

模型在训练集上明显过拟合；第2阶段采用固定学习率 1×10^{-6} 对全部任务进行微调。

欠拟合+微调：第1阶段仅训练10个epoch，使高度、速度和俯仰角3个任务均未达到有效收敛；第2阶段将初始学习率设为 1×10^{-4} ，并采用余弦退火策略降至 1×10^{-6} 。

所有策略均在验证集上选择性能最优的权重进行测试。实验结果如表8所示，“过拟合+微调”由于第1阶段模型已偏离真实分布，第2阶段微调难

表7 使用不同损失函数时各方法的均方根误差

Tab. 7 The RMSE of each method when using different loss functions

算法	Input	Loss	h (m)	v (m/s)	p (°)
ResNet	R+L	UC	15.886	6.530	<u>0.851</u>
		DLS	17.365	6.943	0.986
		MSE	<u>16.277</u>	<u>6.693</u>	0.850
TSCNet	R+L	UC	15.099	5.484	0.788
		DLS	<u>15.107</u>	6.844	0.853
		MSE	15.234	<u>6.651</u>	<u>0.795</u>
RDMFNet (单阶段)	R+L	UC	14.820	6.341	0.780
		DLS	<u>19.741</u>	<u>7.900</u>	0.890
		MSE	21.471	8.329	<u>0.787</u>
	L	UC	15.195	5.250	0.755
		DLS	21.932	6.025	0.736
		MSE	<u>20.906</u>	<u>5.958</u>	<u>0.742</u>
RDMFNet (两阶段)	R+L	UC	<u>14.861</u>	<u>4.690</u>	<u>0.758</u>
		DLS	15.229	4.530	0.759
		MSE	14.767	4.861	0.755
	L	UC	14.447	<u>4.635</u>	<u>0.755</u>
		DLS	<u>14.441</u>	4.585	0.767
		MSE	14.432	4.525	0.742

注：加粗数值表示最优性能，下划线数值表示第二优性能。

表8 不同训练策略的均方根误差

Tab. 8 The RMSE of different training strategies

策略	epoch	h (m)	v (m/s)	p (°)
单阶段	100	14.820	6.341	<u>0.780</u>
p 收敛+微调	33+53	<u>14.861</u>	4.690	0.758
过拟合+微调	100+48	15.362	7.097	0.811
欠拟合+微调	10+93	14.897	<u>6.144</u>	0.790

注：加粗数值表示最优性能，下划线数值表示第二优性能。

以有效缓解过拟合效应，因此其3项参数的总体精度最差。“欠拟合+微调”由于第1阶段学习不足，第2阶段需进行更长时间的训练，尽管速度估计取得了第二好的结果，但整体性能仍不及本文提出的方案。综合来看，本文提出的“ p 收敛+微调”策略在总epoch次数和RMSE指标上均取得最优表现，验证了在第1阶段选择俯仰角任务收敛的最优点作为第2阶段初始化权重的合理性和有效性。

本文在研究中仍存在一些局限性值得探讨与进一步优化。本文以相对简单的场景检验了近垂直下RD图的雷达参数提取能力，尚未引入无人机、直升机等近垂直飞行场景的实际数据，在更大型和复杂地形中的泛化性仍需进一步验证。在未来工作

中，我们将在结合实测数据、优化复杂环境适应能力方面进行进一步研究。

4.4.5 噪声鲁棒性分析

为了进一步验证所提RDMFNet在实际应用场景中的适用性，本文进一步设置了-5, 0, 5, 10, 15 dB及无噪声6种信噪比(Signal-to-Noise Ratio, SNR)梯度，以评估噪声强度变化对参数估计精度的影响。在训练时，随机强度的噪声被添加到训练集中；在测试时，分别在这6种强度的噪声下进行测试。针对每种SNR设置，我们分别统计 h 、 v 、 p 的RMSE，并绘制了对应的SNR-RMSE曲线，如图7所示。

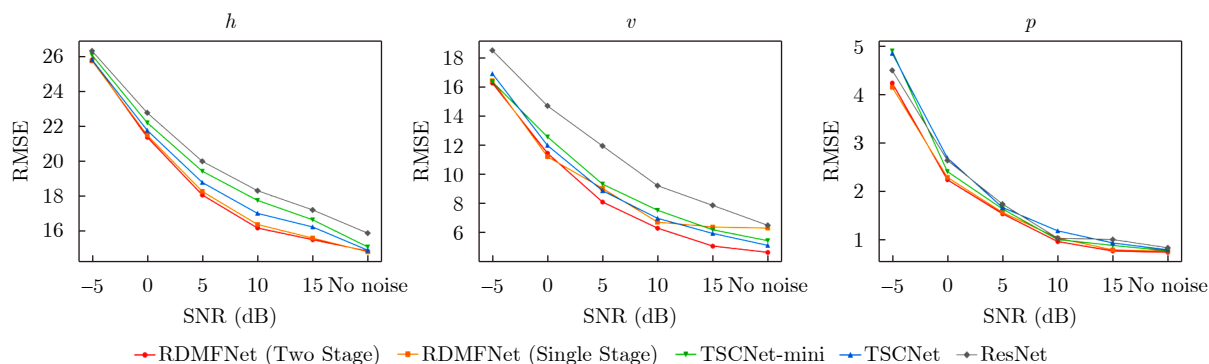


图7 不同SNR设置下的误差比较

Fig. 7 Comparison of RMSE under different SNR settings

从实验结果可以观察到：除了在对 v 的估计中，单阶段RDMFNet在高SNR条件下未表现出优势外，其余所有场景下，两阶段与单阶段RDMFNet均取得了优于其他对比方法的性能。在 h 和 v 的估计中，两阶段训练策略在多数SNR设置下进一步降低了预测误差，例如在15 dB信噪比下，两阶段模型的速度估计误差较单阶段模型降低约1.2 m/s，展现出更强的稳定性。上述结果表明，所提出的RDMFNet在不同噪声水平下均具有良好的鲁棒性。

5 结语

本文设计了基于共享编码器与独立解码器架构的RDMFNet，并提出了两阶段训练方法，第1阶段使编码器学习各个任务共享的特征，第2阶段对每个decoder进行独立微调。实验结果表明，所提方法高度、速度和俯仰角的估计误差分别降低至14.447 m, 4.635 m/s和0.755°，且在不同强度的噪声场景下也能对参数进行有效估计。与其他网络相比，所提网络不仅参数数量和运算量最小，估计精度也有显著优势。

本文在研究中仍存在一些局限性值得探讨与进

一步优化。本文以相对简单的场景检验了近垂直下RD图的雷达参数提取能力，尚未引入无人机、直升机等近垂直飞行场景的实际数据，在更大型和复杂地形中的泛化性仍需进一步验证。在未来工作中，我们将在结合实测数据、优化复杂环境适应能力方面进行进一步研究。

利益冲突 所有作者均声明不存在利益冲突

Conflict of Interests The authors declare that there is no conflict of interests

参考文献

- [1] MILLER S D, MWAFFO V, and COSTELLO III D H. Deep learning-based relative bearing estimation between naval surface vessels and uas in challenging maritime environments[C]. 2025 International Conference on Unmanned Aircraft Systems (ICUAS), Charlotte, USA, 2025: 742-748. doi: 10.1109/ICUAS65942.2025.11007882.
 - [2] 毛军, 付浩, 褚超群, 等. 惯性/视觉/激光雷达SLAM技术综述[J]. 导航定位与授时, 2022, 9(4): 17-30. doi: 10.19306/j.cnki.2095-8110.2022.04.003.
- MAO Jun, FU Hao, CHU Chaoqun, et al. A review of

- simultaneous localization and mapping based on inertial-visual-Lidar fusion[J]. *Navigation Positioning and Timing*, 2022, 9(4): 17–30. doi: [10.19306/j.cnki.2095-8110.2022.04.003](https://doi.org/10.19306/j.cnki.2095-8110.2022.04.003).
- [3] NARASIMHAPPA M, MAHINDRAKAR A D, GUIZILINI V C, *et al*. MEMS-based IMU drift minimization: Sage Husa adaptive robust Kalman filtering[J]. *IEEE Sensors Journal*, 2020, 20(1): 250–260. doi: [10.1109/JSEN.2019.2941273](https://doi.org/10.1109/JSEN.2019.2941273).
- [4] 李道京, 朱宇, 胡烜, 等. 衍射光学系统的激光应用和稀疏成像分析[J]. 雷达学报, 2020, 9(1): 195–203. doi: [10.12000/JR19081](https://doi.org/10.12000/JR19081).
LI Daojing, ZHU Yu, HU Xuan, *et al*. Laser application and sparse imaging analysis of diffractive optical system[J]. *Journal of Radars*, 2020, 9(1): 195–203. doi: [10.12000/JR19081](https://doi.org/10.12000/JR19081).
- [5] 王超, 王岩飞, 刘畅, 等. 基于参数估计的高分辨率SAR运动目标距离徙动校正方法[J]. 雷达学报, 2019, 8(1): 64–72. doi: [10.12000/JR18054](https://doi.org/10.12000/JR18054).
WANG Chao, WANG Yanfei, LIU Chang, *et al*. A new approach to range cell migration correction for ground moving targets in high-resolution SAR system based on parameter estimation[J]. *Journal of Radars*, 2019, 8(1): 64–72. doi: [10.12000/JR18054](https://doi.org/10.12000/JR18054).
- [6] 许京新. 基于深度学习的SAR图像舰船目标检测[D]. [硕士学位论文], 烟台大学, 2025. doi: [10.27437/d.cnki.gytd.2025.000610](https://doi.org/10.27437/d.cnki.gytd.2025.000610).
XU Jingxin. Deep learning-based ship target detection in SAR images[D]. [Master dissertation], Yantai University, 2025. doi: [10.27437/d.cnki.gytd.2025.000610](https://doi.org/10.27437/d.cnki.gytd.2025.000610).
- [7] DE HOOP M V, LASSAS M, and WONG C A. Deep learning architectures for nonlinear operator functions and nonlinear inverse problems[J]. *Mathematical Statistics and Learning*, 2022, 4(1/2): 1–86. doi: [10.4171/MSL/28](https://doi.org/10.4171/MSL/28).
- [8] DARA S and TUMMA P. Feature extraction by using deep learning: A survey[C]. 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2018: 1795–1801. doi: [10.1109/ICECA.2018.8474912](https://doi.org/10.1109/ICECA.2018.8474912).
- [9] KWON H Y, YOON H G, LEE C, *et al*. Magnetic Hamiltonian parameter estimation using deep learning techniques[J]. *Science Advances*, 2020, 6(39): eabb0872. doi: [10.1126/sciadv.abb0872](https://doi.org/10.1126/sciadv.abb0872).
- [10] KOLLIAS D. ABAW: Learning from synthetic data & multi-task learning challenges[C]. European Conference on Computer Vision, Tel Aviv, Israel, 2023: 157–172. doi: [10.1007/978-3-031-25075-0_12](https://doi.org/10.1007/978-3-031-25075-0_12).
- [11] CIPOLLA R, GAL Y, and KENDALL A. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 7482–7491. doi: [10.1109/CVPR.2018.00781](https://doi.org/10.1109/CVPR.2018.00781).
- [12] RICHARDS M A and MELVIN W L. Principles of Modern Radar: Basic Principles[M]. London: The Institution of Engineering and Technology, 2022: 360.
- [13] WANG Aiguo, ZHANG Wei, and CAO Jianshu. Terrain clutter modeling for airborne radar system using digital elevation model[C]. The 2012 International Workshop on Microwave and Millimeter Wave Circuits and System Technology, Chengdu, China, 2012: 1–4. doi: [10.1109/MMWCST.2012.6238182](https://doi.org/10.1109/MMWCST.2012.6238182).
- [14] IOFFE S and SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]. The 32nd International Conference on International Conference on Machine Learning, Lille, France, 2015: 448–456.
- [15] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *et al*. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification[C]. The IEEE International Conference on Computer Vision, Santiago, Chile, 2015: 1026–1034. doi: [10.1109/ICCV.2015.123](https://doi.org/10.1109/ICCV.2015.123).
- [16] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *et al*. Deep residual learning for image recognition[C]. The IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 770–778. doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- [17] ABDULATIF S, CAO Ruizhe, and YANG Bin. CMGAN: Conformer-based metric-GAN for monaural speech enhancement[J]. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2024, 32: 2477–2493. doi: [10.1109/TASLP.2024.3393718](https://doi.org/10.1109/TASLP.2024.3393718).
- [18] RUDER S. An overview of multi-task learning in deep neural networks[J]. *arXiv preprint arXiv: 1706.05098*, 2017. doi: [10.48550/arXiv.1706.05098](https://doi.org/10.48550/arXiv.1706.05098).

作者简介

刘子源, 博士生, 主要研究方向为SAR图像处理与雷达信号处理。

王绍平, 硕士, 主要研究方向为电子与通信。

何屹廷, 硕士生, 主要研究方向为雷达信号处理。

谷源涛, 教授, 主要研究方向为信号处理。

(责任编辑: 于青)