

基于深度无监督变分网络的杂波建模与雷达目标检测方法

刘要强^{1,2}, 陈文超^{1,2*}, 施力行^{1,2}, 田 隆³, 王鹏辉^{1,2}, 陈 渤^{1,2}

(1. 西安电子科技大学雷达信号处理全国重点实验室, 陕西西安 710071; 2. 西安电子科技大学电子工程学院, 陕西西安 710071;
3. 西安电子科技大学计算机科学与技术学院, 陕西西安 710071)

摘要: 现代雷达目标检测往往面临复杂多变的杂波环境, 传统的基于模型驱动的恒虚警率(Constant False Alarm Rate, CFAR)检测器容易出现模型失配的问题, 现有的基于数据驱动的有监督深度学习方法存在烦琐且昂贵的标签问题. 针对上述问题, 本文提出了一种基于深度无监督变分网络的杂波建模方法, 该方法利用面向雷达回波高维分布特征学习的变分自编码器(Variational Auto-Encoder, VAE), 针对雷达回波处理后的距离-多普勒谱, 实现对复杂杂波分布的重构建模. 首先, 在VAE的无监督推理-生成构架中引入卷积神经网络(Convolutional Neural Network, CNN)和循环神经网络(Recurrent Neural Network, RNN), 分别利用CNN网络的局部特征捕捉能力和RNN网络的时序相关性信息提取能力来实现对距离-多普勒谱的重构建模. 其次, 为了充分地捕获距离-多普勒谱中不同区域的杂波分布特征及二维时空信息, 本文提出了一种基于时空变分Transformer的杂波建模方法, 该方法将Transformer架构引入到所提的深度无监督杂波建模的变分网络中, 借助Transformer网络的自注意力机制来捕获R-D谱数据的全局相关性. 再次, 为了充分挖掘不同场景下R-D谱的杂波分布特征及保留原始数据的二维时空信息, 设计了开关机制和二维位置编码机制以匹配Transformer网络架构. 最后, 结合分布外(Out-Of-Distribution, OOD)检测策略, 本文提出了一种基于深度无监督变分网络的杂波建模与雷达目标检测方法, 重构似然表示无监督变分网络准确重构出输入样本的难易程度. 重构似然越大, 重构样本与输入样本越相似. 因此, 本文利用重构似然定义OOD分数, 作为划分目标与杂波的依据, 实现雷达目标检测任务. 通过仿真数据验证, 本文所提的无监督杂波建模方法能够实现对雷达距离-多普勒谱杂波分布的精细重构, 且相比传统CFAR方法, 当达到80%检测概率时, 本文提出的方法所需信杂噪比(Signal to Clutter plus Noise Ratio, SCNR)优化了5.6 dB.

关键词: 雷达目标检测; 无监督变分网络; 变分自编码器(VAE); 杂波建模; 距离-多普勒谱; 重构似然

基金项目: 国家自然科学基金(No.6220010437, No.U21B2006); 雷达信号处理全国重点实验室基金(No.JKW202X0X, No.KGJ202401)

中图分类号: TN957.51 **文献标识码:** A **文章编号:** 0372-2112(2025)08-2691-16

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20250370

Clutter Modeling and Radar Target Detection Method Based on Deep Unsupervised Variational Networks

LIU Yao-qiang^{1,2}, CHEN Wen-chao^{1,2*}, SHI Li-xing^{1,2}, TIAN Long³, WANG Peng-hui^{1,2}, CHEN Bo^{1,2}

(1. National Key Laboratory of Radar Signal Processing, Xidian University, Xi'an, Shaanxi 710071, China;

2. School of Electronic Engineering, Xidian University, Xi'an, Shaanxi 710071, China;

3. School of Computer Science and Technology, Xidian University, Xi'an, Shaanxi 710071, China)

Abstract: Modern radar target detection often faces complex and changeable clutter environments. Traditional model-driven constant false alarm rate (CFAR) detectors are prone to model mismatch problems, and existing data-driven supervised deep learning methods require cumbersome and expensive label problems. In response to the above problems, this paper proposes a clutter modeling method based on deep unsupervised variational networks. This method utilizes a variational autoencoder for learning the high-dimensional distribution features of radar echoes to achieve the reconstruction modeling

of complex clutter distributions for the range-doppler spectrum after radar echo processing. Firstly, convolutional neural network (CNN) and recurrent neural network (RNN) are introduced into the unsupervised inference-generation framework of the variational autoencoder. The reconstruction modeling of range-doppler spectra is achieved by respectively utilizing the local feature capture ability of CNN networks and the temporal correlation information extraction ability of RNN networks. To fully capture the clutter distribution characteristics and two-dimensional spatiotemporal information in different regions of the range-doppler spectrum, this paper proposes a clutter modeling method based on spatiotemporal variational Transformer. This method introduces the Transformer architecture into the proposed deep unsupervised clutter modeling variational network. Capture the global correlation of R-D spectral data by leveraging the self-attention mechanism of the Transformer network. In order to fully explore the clutter distribution characteristics of R-D spectra in different scenarios and retain the two-dimensional spatiotemporal information of the original data, a switching mechanism and a two-dimensional position encoding mechanism are designed to match the Transformer network architecture. Finally, combined with the out-of-distribution (OOD) detection strategy, this paper proposes a clutter modeling and radar target detection method based on deep unsupervised variational networks, and reconstructs the likelihood representation of the unsupervised variational network to accurately reconstruct the difficulty level of the input samples. The greater the reconstruction likelihood, the more similar the reconstructed sample is to the input sample. Therefore, the OOD score is defined by using the reconstructed likelihood as the basis for dividing the target from clutter to achieve the radar target detection task. Verified by simulation data, the unsupervised clutter modeling method proposed in this paper can achieve fine reconstruction of the clutter distribution in the radar range-Doppler spectrum. Moreover, compared with the traditional CFAR method, when the detection probability reaches 80%, the signal to clutter plus noise ratio (SCNR) required by the method proposed in this paper The SCNR is optimized by 5.6 dB.

Key words: radar target detection; unsupervised variational network; variational auto-encoder (VAE); clutter modeling; range-Doppler spectrum; reconstruction likelihood

Foundation Item(s): National Natural Science Foundation of China (No.6220010437, No.U21B2006); National Key Laboratory of Radar Signal Processing Fund (No.JKW202X0X, No.KGJ202401)

1 引言

雷达目标检测过程往往面临着复杂多变的非均匀杂波环境,杂波分布具有显著的非单一性^[1].在此背景下,针对目标检测任务,尤其是对弱小目标的检测,传统的动目标检测(Moving Target Detection, MTD)技术^[2]已经不能对信号能量进行有效的积累,导致雷达目标的检测性能严重下降.因此,探究适用于复杂背景下的雷达目标检测技术具有重要的实际意义.目前,雷达目标检测方法可以分为两类:传统的恒虚警率(Constant False Alarm Rate, CFAR)目标检测方法和基于深度学习的目标检测方法.

传统的CFAR检测过程是:先对目标邻近的参考单元内的杂波水平进行估计,并根据估计值设置门限;再将该门限值与检测单元的能量水平进行比较,从而判断检测单元是否为目标.其中,单元平均CFAR检测^[3](Cell Averaging-CFAR, CA-CFAR)是一种典型的非相干积累检测方法,其主要依靠回波的能量判断目标的有无.在均匀的瑞利包络杂波背景下,CA-CFAR能够达到最优的检测性能,但当目标处于杂波边缘或非均匀杂波环境中时,参考单元中的数据将不再满足独立同分布,该方法的检测性能会随着信杂噪比(Signal to Clutter plus Noise Ratio, SCNR)的降低而急剧下降^[4].

后续又相继出现了最大CFAR(Greatest Of CFAR, GO-CFAR)、最小CFAR(Smallest Of CFAR, SO-CFAR)、加权单元平均CFAR(Weighted Cell-Averaging CFAR, WCA-CFAR)等同属于均值类的CFAR检测方法^[5].上述的CFAR检测算法主要应用于具有高斯特性的杂波环境,当面临复杂非高斯特性的杂波环境时,CFAR检测方法的自适应阈值通常偏高,易造成目标的漏检.此外,针对参考单元杂波分布与假定的杂波分布失配的问题,有学者还提出了一种基于杂波统计特征的CFAR检测器^[6],将背景划分为噪声区和杂波区,并针对不同区域进行拟合优度检验,随后对分布进行指数化处理,并验证了其有效性.但这类CFAR检测器未考虑杂波抑制问题,当杂波抑制滤波器和杂波的多普勒谱不匹配时,会造成目标SCNR的损失,导致检测性能下降.

随着人工智能的发展和深度学习技术的不断提高,基于深度学习的目标检测算法在计算机视觉(Computer Vision, CV)^[7]、医学成像等领域迅速出现^[8].深度学习中的雷达目标检测任务可以看作是一个基于有监督学习的二元分类问题^[9].然而,这些有监督学习方法^[10]需要依赖大量的目标样本,实际应用中,非合作目标的信息难以获取,导致样本类别严重不平衡,监督学习方法很难充分发挥其优势.此外,对庞大的样本集进行标注,会消耗大量的人工成本.因此,无需样本标签

的无监督深度学习方法更适合解决在复杂环境下的雷达目标检测问题. 文献[11]将无监督深度学习方法应用于 SAR (Synthetic Aperture Radar) 图像, 以完成目标检测任务. 接着, 文献[12]在卷积神经网络 (Convolutional Neural Network, CNN) 基础上提出了一种无监督跨域的 SAR 目标检测方法, 旨在解决在 SAR 图像难以获取条件下的目标检测问题.

尽管基于无监督学习的检测方法取得了较好的进展, 目前仍有两个问题亟待解决. 一是现有的基于无监督方法训练的模型泛化性能差, 难以应对复杂多变的环境背景. 二是面向雷达 R-D 谱中的点目标检测任务, 杂波干扰对小目标的检测影响较大^[13], 如何精准地区分杂波与目标对模型方法的设计提出较大的挑战. 本文针对复杂环境下的雷达目标检测问题, 采用深度无监督的变分自编码器 (Variational Auto-Encoder, VAE) 架构^[14]提出一种基于时空变分 Transformer 的杂波 R-D 谱建模的方法, 结合本文提出的分布外 (Out Of Distribution, OOD) 检测策略^[15], 实现目标检测性能的提升. 本文的主要贡献总结如下:

(1) 为了应对实际检测环境中杂波分布复杂多变的问题, 本文提出了一种基于深度无监督变分网络的杂波建模方法, 实现对杂波复杂分布特性的精准建模.

(2) 为了充分地捕获 R-D 谱中不同区域的杂波分布特征及二维时空信息, 本文提出了一种基于时空变

分 Transformer 的杂波建模方法, 该方法将 Transformer 结构引入到所提的深度无监督杂波建模方法中, 并设计了开关 Transformer 机制和二维位置编码机制^[16], 以匹配 R-D 谱样本特点.

(3) 在无监督杂波分布建模基础上, 本文进一步开发了 OOD 检测策略指导的目标检测技术, 进而应用于雷达目标检测的任务, 利用仿真数据验证了所提方法框架相较于传统检测方法的高效性和优越性.

2 基于深度无监督变分网络的杂波建模方法

2.1 雷达 R-D 谱成像

雷达 R-D 谱成像过程^[17]是通过距离维进行脉冲压缩, 在多普勒维对脉冲串进行相参积累, 来完成 R-D 谱成像. R-D 谱成像的整个流程如图 1 所示, 雷达回波信号中最常见的调制信号是线性调频 (Linear Frequency Modulation, LFM) 信号, LFM 信号的数学表达式为

$$s_r(t) = \text{rect}\left(\frac{t}{T}\right) \exp(j2\pi f_c t + j\pi\mu t^2) \quad (1)$$

其中, f_c 为载波频率; μ 为调制斜率; T 为脉宽; $\text{rect}(\cdot)$ 为窗函数, 其表达式为

$$\text{rect}\left(\frac{t}{T}\right) = \begin{cases} 1, & \left|\frac{t}{T}\right| < 1 \\ 0, & \text{elsewise} \end{cases} \quad (2)$$

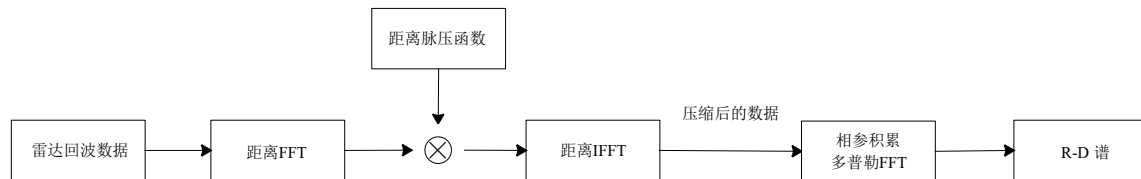


图 1 雷达 R-D 谱成像流程

若忽略噪声的影响, 雷达接收的目标反射回波信号可写为

$$s(t, t_m) = A_0 \times \text{rect}\left(\frac{t - 2R(t_m)/c}{T}\right) \times \exp\left[j2\pi f_c \left(t - \frac{2R(t_m)}{c}\right) + j\pi\mu \left(t - \frac{2R(t_m)}{c}\right)^2\right] \quad (3)$$

其中, A_0 为回波信号的幅度; $R(t_m)$ 为雷达位置到目标位置之间的斜距; c 为光速. 然后, 对回波进行下变频操作, 得到

$$s_d(t, t_m) = \sigma \times \text{rect}\left(\frac{t - 2R(t_m)/c}{T}\right) \exp\left[-j2\pi \frac{2R(t_m)}{\lambda}\right] \times \exp\left[j\pi\mu \left(t - \frac{2R(t_m)}{c}\right)^2\right] \quad (4)$$

其中, $\lambda = c/f$ 为中心频率对应的波长.

对距离维作脉冲压缩的匹配滤波器的时域脉冲响应为

$$h_r(t) = s_r^*(-t) = \text{rect}\left(\frac{t}{T}\right) \exp(j2\pi f_c t) \exp(-j\pi\mu t^2) \quad (5)$$

为了提高计算效率, 距离维的脉冲压缩通常在频域进行. 时域与频域之间的变换通过快速傅里叶变换 (Fast Fourier Transform, FFT) 及其逆变换 (Inverse Fast Fourier Transform, IFFT) 完成, 从而获得匹配滤波器的输出:

$$\begin{aligned} s_r(t, t_m) &= \text{IFFT}\left\{\text{FFT}\left[s_d(t, t_m)\right] \times \text{FFT}\left[h_r(t)\right]\right\} \\ &= A_1 \times \text{sinc}\left[\pi B \left(t - \frac{2R(t_m)}{c}\right)\right] \times \exp\left[-j2\pi f_c \frac{2R(t_m)}{c}\right] \end{aligned} \quad (6)$$

其中, A_1 为点目标信号经过距离压缩后的幅度; B 为 LFM 信号的带宽; sinc 函数为

$$\text{sinc}(v) = \frac{\sin(v)}{v} \quad (7)$$

在距离维度进行脉冲压缩处理后,在多普勒维进行相参脉冲的积累处理,相参积累处理的意义在于提高目标的信噪比,提高多普勒频率的准确度.相参积累的过

程如图2所示,每个行向量表示一个雷达脉冲重复周期内的距离单元向量,每连续 M 个行向量排列好后,形成一个二维数组.再对每个列向量做FFT处理,形成图2右侧的相参积累后的结果,得到“距离-距离勒”二维分布图,图中 M 、 N 分别表示 R-D 谱中多普勒维和距离维的维度.

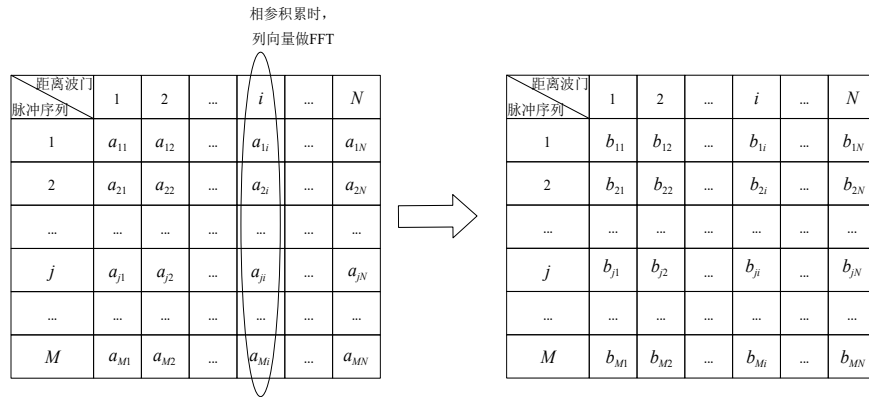


图2 距离-多普勒二维分布图

2.2 面向 R-D 谱建模的深度无监督变分网络

本文引入VAE模型实现对雷达回波R-D谱分布特性的无监督建模.VAE是一种经典的深度变分生成模型,2014年由Kingma提出^[14].VAE将变分贝叶斯推理和

自编码器相结合,其思想是利用低维空间中的隐变量来建模高维复杂数据的分布.相比于自编码器,VAE学习的是特征在低维空间中的分布参数而非唯一表示.VAE模型由推理网络和生成网络组成,其结构图如图3所示.

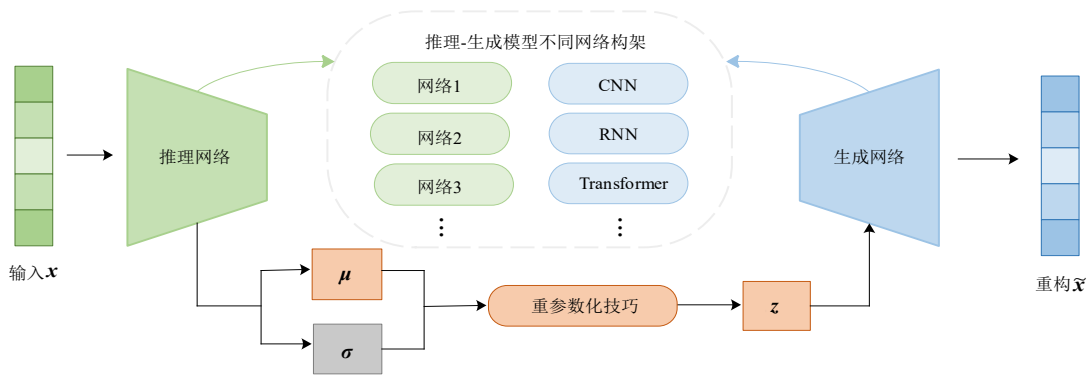


图3 VAE示意图

VAE的生成网络利用服从标准高斯分布的隐变量 z 重构出与输入数据 x 相似的数据.网络的生成过程可以表示为 $p_\theta(x, z) = p_\theta(x|z)p_\theta(z)$, 其中,若隐变量的先验分布 $p_\theta(z)$ 选定为标准高斯分布 $p_\theta(z) = N(z; 0, I)$, 则可求得条件概率 $p_\theta(x|z)$, 即给定隐变量 z 时生成数据 x 的分布,实现生成近似输入数据 x 的任务.

推理网络是将输入数据 x 的高维信息特征映射为隐空间的低维信息特征,即隐变量 z 的均值和方差.由于真实的后验分布 $p_\theta(z|x)$ 难以直接计算,需要引入变分分布 $q_\phi(z|x)$ 来近似推断真实的后验分布.利用变分推理网络得到隐变量 z 所服从高斯分布的均值与方差,然后通过采样操作得到蕴含输入数据 x 特征信息的隐

变量 z . VAE模型的优化函数即为最大化带正则项的对数似然函数,也被称为证据下界 (Evidence Lower Bound, ELBO):

$$\text{ELBO} = E_{q_\phi(z|x)} [\ln p_\theta(x|z)] - D_{\text{KL}} [q_\phi(z|x) \| p_\theta(z)] \quad (8)$$

其中,ELBO分为两项,第一项的对数似然期望是保证生成模型对输入数据的重构能力,第二项的KL (Kullback-Leibler) 散度是约束变分分布 $q_\phi(z|x)$ 接近其在生成模型中的先验分布 $p_\theta(z)$.

2.3 基于 CNN 和 RNN 结构骨干的 VAE 杂波重构模型

VAE模型的整个网络构架包括推理网络和生成网络,针对雷达R-D谱的特性,VAE的推理-生成构架可使

用不同的网络结构,神经网络不同的结构特点在雷达 R-D 谱建模任务中将有不同的表现. 本文在所提的无监督 VAE 变分网络构架下,首先引入了 CNN 和循环神经网络(Recurrent Neural Network, RNN)^[18].

2.3.1 CNN-VAE 重构网络

CNN 网络结构以卷积层为核心,利用局部感受野提取空间特征,适合处理图像数据. R-D 谱本身作为二维图像数据,适合由 CNN 网络直接处理. 输入 R-D 谱数据经过 CNN 卷积编码网络映射到隐变量 z 的均值 μ_z 和方差 σ_z , 利用重参数化技巧从分布中采样得到 z , 再经过反卷积解码网络重构出 R-D 谱数据 \tilde{x} . CNN 网络重构过程可以描述为

$$z \sim N(\mu_z, \text{diag}(\sigma_z)) \quad (9)$$

$$\begin{cases} \mu_z = \text{MLP}_{1\mu}(\text{MultiConv}(x)) \\ \sigma_z = \text{MLP}_{1\sigma}(\text{MultiConv}(x)) \end{cases} \quad (10)$$

$$\tilde{x} = \text{MultiDeconv}(\text{MLP}_2(z)) \quad (11)$$

其中, $\text{MultiConv}(\cdot)$ 和 $\text{MultiDeconv}(\cdot)$ 分别为多层卷积操作和多层反卷积操作; $\text{MLP}_{1\mu}(\cdot)$ 、 $\text{MLP}_{1\sigma}(\cdot)$ 为全连接网络,将高维 R-D 谱数据特征映射为低维隐变量的均值 μ_z 和方差 σ_z ; $\text{MLP}_2(\cdot)$ 作为全连接网络将采样得到的隐变量 z 映射回高维特征向量,以便通过多层反卷积网络重构出高维的 R-D 谱数据 \tilde{x} . 由于对隐变量 z 的采样操作是不可导的,训练过程梯度无法进行反向传播. 可以通过下面的重参数化技巧来实现对隐变量 z 的采样:

$$z_i = \mu_i + \sigma_i \odot \varepsilon_i, \quad \varepsilon_i \sim N(0, 1) \quad (12)$$

其中, ε_i 服从标准均匀分布; z_i 为对隐变量的第 i 次采样,这种方式采样的目的是确保梯度能够进行反向传播. 雷达 R-D 谱作为二维数据,具有明显的局部相关性, CNN 网络能够有效地捕捉 R-D 谱的结构特征. 但 CNN 网络无法建模长距离的依赖关系,捕捉全局特征的能力较弱.

2.3.2 RNN-VAE 重构网络

RNN 网络通过递归结构的隐藏状态建模序列数据,适合处理短时间序列数据. 长短期记忆网络(Long Short-Term Memory, LSTM)、门控循环单元(Gated Recurrent Unit, GRU)等变体的出现,解决了 RNN 网络在处理长时间序列时存在的梯度消失的问题. 本文采用 LSTM 网络来探究基于动态的网络架构在杂波 R-D 谱建模中的表现. 由于雷达 R-D 谱为二维数据,首先应处理成适合 LSTM 网络输入的时间序列形式,即将 R-D 谱进行展平操作:

$$\mathbf{x}_k = \text{Flatten}(\mathbf{x}) = [x_{1,1}, x_{1,2}, \dots, x_{M,1}, x_{M,2}, \dots, x_{M,N}] \quad (13)$$

其中, \mathbf{x}_k 为第 k 个 R-D 谱经过展平处理后的序列数据; $\text{Flatten}(\cdot)$ 为展平操作; M, N 分别为 R-D 谱的多普勒维

度和距离维度. 预处理后的序列数据 \mathbf{x}_k 经过 LSTM 推理网络,映射为低维的隐变量 z 的均值和方差,利用重参数化技巧从分布中采样得到 z ,再经过 LSTM 生成网络重构出序列数据 $\tilde{\mathbf{x}}_k$,最后恢复成原始维度 $\tilde{\mathbf{x}} = \text{Unflatten}(\tilde{\mathbf{x}}_k)$,即可得到重构后的 R-D 谱. RNN 网络重构过程可以描述为

$$z \sim N(\mu_z, \text{diag}(\sigma_z)) \quad (14)$$

$$\begin{cases} \mu_z = \text{LSTM}_{1\mu}(\text{Flatten}(\mathbf{x})) \\ \sigma_z = \text{LSTM}_{1\sigma}(\text{Flatten}(\mathbf{x})) \end{cases} \quad (15)$$

$$\tilde{\mathbf{x}} = \text{Unflatten}(\text{LSTM}_2(z)) \quad (16)$$

其中, $\text{LSTM}_{1\mu}(\cdot)$ 、 $\text{LSTM}_{1\sigma}(\cdot)$ 均为 RNN-VAE 网络中的推理网络; $\text{LSTM}_2(\cdot)$ 为 RNN-VAE 网络中的生成网络. RNN 网络能够很好地捕获 R-D 谱中的时间相关性信息,但 R-D 谱作为二维图像,将其展开为一维序列数据后序列长度过长, RNN 递归网络无法并行计算,导致训练速度较慢. 同时,输入序列数据过长,会造成 RNN 网络反传梯度过小,参数更新慢的问题.

面向不同的场景进行不同的结构骨干的选择,可更好地支撑所提框架在不同场景下回波建模的适应能力.

3 基于时空变分 Transformer 的杂波建模方法

在 VAE 的推理-生成构架下, CNN 和 RNN 模型由于自身网络骨干的限制,未能充分发挥无监督重构模型的优势, 本文将 Transformer 模型^[19]引入到深度无监督杂波建模方法中. Transformer 网络依靠自注意力机制进行特征提取,能够建模 R-D 谱中任意位置之间的关系,捕捉 R-D 谱数据中的全局相关性. 多头自注意力机制能够关注到 R-D 谱中各个位置的杂波特征信息,且 Transformer 网络支持数据并行处理,训练和推理速度快. 因此,本文在 VAE 的推理-生成框架下,提出时空变分 Transformer 模型,旨在针对雷达 R-D 谱,学习杂波特性和分布,实现对其精细化建模. 其中,在原始的 Transformer 模型^[20]基础上,所提时空变分 Transformer 模型引入了开关 Transformer 机制^[21]和二维位置编码机制,如图 4 所示.

3.1 开关 Transformer 机制

受 MoE^[22]的启发,本文在 Transformer 模型中加入开关机制,设计开关 transformer 模型. 如图 4 所示,将 Transformer 生成网络中的前馈神经网络(Feed-Forward Neural network, FFN)层替换为开关 FFN 层,用 K 个 FFN 模型来代替原单个 FFN 模型. 具有不同参数的 FNN 模型可以有效地捕获 R-D 谱中不同区域的杂波分布特征,帮助实现更加精准的杂波分布建模. 具体地,预处理后的 R-D 谱经过 Transformer 推理网络,在得到相应的隐变量 z 的均值 μ_z 和方差 σ_z 的同时,也会输出权重向

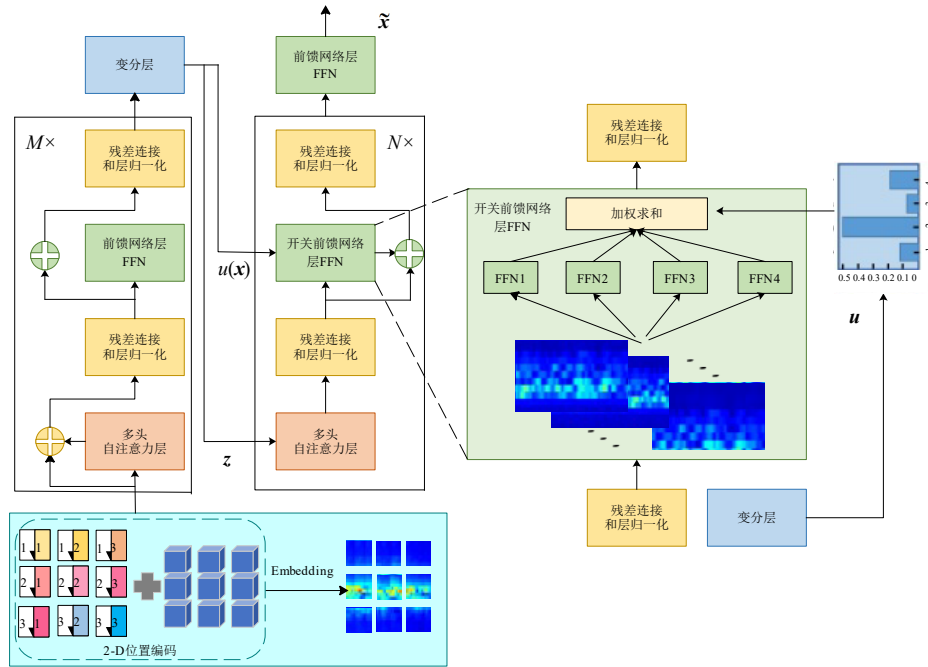


图4 基于时空变分Transformer的重构网络结构图

量 $\mathbf{u}(\mathbf{x}) \in \mathbb{R}^{K \times 1}$, 用于分配生成网络中 K 个 FFN 模型的输出所占的比重, 公式如下:

$$\tilde{\mathbf{h}}^l = \text{MSA}(\mathbf{h}^{l-1}) \quad (17)$$

$$\mathbf{h}^l = \sum_{i=1}^K \mathbf{u}_i(\mathbf{x}) \text{FFN}_i(\tilde{\mathbf{h}}^l), \mathbf{u}(\mathbf{x}) = \mathbf{f}_\theta(\mathbf{x}) \quad (18)$$

$$\text{FFN}_i(\tilde{\mathbf{h}}^l) = \text{ReLU}(\tilde{\mathbf{h}}^l \mathbf{W}_{1i} + \mathbf{b}_{1i}) \mathbf{W}_{2i} + \mathbf{b}_{2i} \quad (19)$$

其中, $\text{MSA}(\cdot)$ 为多头注意力机制操作; \mathbf{h}^{l-1} 为第 $l-1$ 层经过开关 FFN 层后加权求和输出的隐藏变量; $\tilde{\mathbf{h}}^l$ 、 \mathbf{h}^l 分别为在生成网络第 l 层输入开关 FFN 层之前的隐藏变量和经过开关 FFN 层后加权求和输出的隐藏变量; K 为开关 FFN 层中 FFN 模型的个数; $\mathbf{f}_\theta(\cdot)$ 为推理网络; θ 为推理网络的参数. $\text{ReLU}(\cdot)$ 为激活函数, \mathbf{W}_{1i} 、 \mathbf{W}_{2i} 、 \mathbf{b}_{1i} 、 \mathbf{b}_{2i} 为对应的第 i 个 FFN 模型的可学习参数. 我们将基于开关 Transformer 获取隐层表示的操作, 即式(17)~式(19), 定义为 ST(\cdot).

3.2 二维位置编码机制

3.2.1 预处理环节的切片操作

在预处理环节, 一个 R-D 谱样本的切片数量由切片大小 L_r 、 L_d 和重叠长度 O_r 、 O_d 共同决定. R-D 谱的切片过程如图 5 所示. 图中 R-D 谱样本维度为 23×128 , L_r 、 L_d 分别表示距离维和多普勒维的切片大小; O_r 、 O_d 表示切片分别沿距离维和多普勒维平移一次后, 得到的新切片与原切片在两个维度下的重叠长度. 则 R-D 谱样本的切片数量可以表示为

$$N_{\text{R-D}} = I \times J \quad (20)$$

$$\begin{cases} I = (M - L_d) / (L_d - O_d) + 1 \\ J = (N - L_r) / (L_r - O_r) + 1 \end{cases} \quad (21)$$

其中, $N_{\text{R-D}}$ 为一个 R-D 谱样本所得到切片数量; I 为遍历多普勒维需要的切片个数; J 为遍历距离维需要的切片个数; M 、 N 分别为 R-D 谱中多普勒维和距离维的维度; $S_d = L_d - O_d$ 和 $S_r = L_r - O_r$ 则分别为切片每次沿多普勒维和距离维移动的步长.

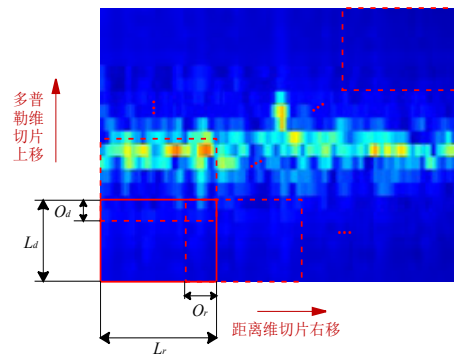


图5 R-D谱的切片过程

在检测实验中, 设置太大的切片尺寸会导致单个切片内杂波建模精度降低, 导致整体 R-D 谱建模精度不高; 太小的切片尺寸容易导致信息冗余, 并且容易破坏杂波的局部分布特征. 同时, 过度的重叠长度也会造成信息冗余, 并且使目标包含在多个切片内, 检测过程中目标信息也容易被重构, 导致检测性能下降. 通过实验验证, 本实验采用的切片大小为 $L_d = 7$ 、 $L_r = 16$; 重

长度大小为 $O_d=2, O_r=0$.

3.2.2 二维位置编码机制

值得注意的是,原始的 Transformer 网络使用的是 1-D(1-Dimension)位置编码,当直接用于雷达 R-D 谱处理时,会损失掉距离维和多普勒维的二维关联信息,因此,本文引入 2-D(2-Dimension)位置编码来引导模型更好地对杂波分布的时空特征进行学习,如图 6 所示.

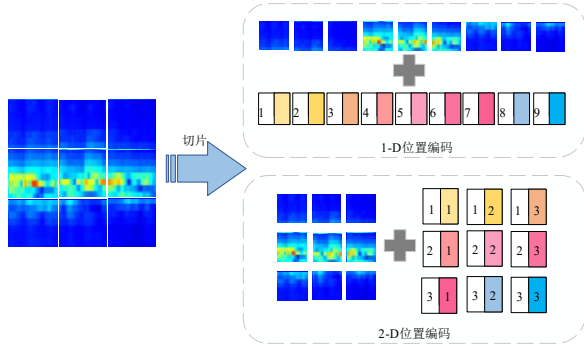


图 6 Transformer 网络输入 R-D 谱预处理

在 1-D 位置编码嵌入中,对应位置 n 的位置向量在偶数位和奇数位的值分别为

$$PE_{n,2i} = \sin\left(\frac{n}{10000^{2i/d}}\right), PE_{n,2i+1} = \cos\left(\frac{n}{10000^{2i/d}}\right) \quad (22)$$

其中, d 为输入向量嵌入的维度. 得到的 1-D 位置嵌入向量为 $\mathbf{PE}_{1D} = [\mathbf{PE}_1, \mathbf{PE}_2, \dots, \mathbf{PE}_{N_{R-D}}]$, N_{R-D} 表示输入 R-D 谱样本切片的数量. 对于二维的雷达 R-D 谱数据,数据预处理后得到切片的数量为 $N_{R-D} = I \times J$, J, I 分别为沿距离维和多普勒维的切片数量. 将距离维、多普勒维对应的位置嵌入向量分别沿列和行进行复制得到二维的位置嵌入向量:

$$\mathbf{PE}_{\text{Range}} = \begin{bmatrix} \mathbf{PE}_1 & \mathbf{PE}_2 & \dots & \mathbf{PE}_J \\ \mathbf{PE}_1 & \mathbf{PE}_2 & \dots & \mathbf{PE}_J \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{PE}_1 & \mathbf{PE}_2 & \dots & \mathbf{PE}_J \end{bmatrix} \quad (23)$$

$$\mathbf{PE}_{\text{Doppler}} = \begin{bmatrix} \mathbf{PE}_1 & \mathbf{PE}_1 & \dots & \mathbf{PE}_1 \\ \mathbf{PE}_2 & \mathbf{PE}_2 & \dots & \mathbf{PE}_2 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{PE}_I & \mathbf{PE}_I & \dots & \mathbf{PE}_I \end{bmatrix}$$

经过上式操作后, $\mathbf{PE}_{\text{Range}}$ 和 $\mathbf{PE}_{\text{Doppler}}$ 的矩阵维度相同,将距离维和多普勒维的位置嵌入连接起来,可以表示为

$$\mathbf{PE}_{2D} = \text{Concat}(\mathbf{PE}_{\text{Range}}, \mathbf{PE}_{\text{Doppler}})$$

$$= \begin{bmatrix} [\mathbf{PE}_1, \mathbf{PE}_1] & [\mathbf{PE}_2, \mathbf{PE}_1] & \dots & [\mathbf{PE}_J, \mathbf{PE}_1] \\ [\mathbf{PE}_1, \mathbf{PE}_2] & [\mathbf{PE}_2, \mathbf{PE}_2] & \dots & [\mathbf{PE}_J, \mathbf{PE}_2] \\ \vdots & \vdots & \vdots & \vdots \\ [\mathbf{PE}_1, \mathbf{PE}_I] & [\mathbf{PE}_2, \mathbf{PE}_I] & \dots & [\mathbf{PE}_J, \mathbf{PE}_I] \end{bmatrix} \quad (24)$$

其中, $\mathbf{PE}_{\text{Range}}$ 和 $\mathbf{PE}_{\text{Doppler}}$ 中每个表达式 \mathbf{PE}_n 与 1-D 位置

嵌入中的表达式相同. 采用 2-D 位置编码处理雷达 R-D 谱数据,可以保留杂波分布原有的二维时空信息特征,实现更加精准的建模.

3.3 基于时空变分 Transformer 的推理-生成网络

本文基于开关 Transformer 和二维位置编码,进一步提出时空变分 Transformer 模型,并将其简称为 SVT (Spatiotemporal Variational Transformer) 模型,本节介绍 SVT 模型的推理和生成过程. 首先,将雷达 R-D 谱转化为与 Transformer 网络兼容的数据形式,与 ViT (Visual Transformer, ViT) 模型^[19]类似,对 R-D 谱数据进行切片操作,并加入二维位置编码信息,得到预处理后的嵌入矩阵 \mathbf{x}' :

$$\mathbf{x}' = E(\text{cut}(\mathbf{x})) + \mathbf{PE}_{2D} \quad (25)$$

式中, $\text{cut}(\cdot)$ 为对输入 R-D 谱进行切片操作; $E(\cdot)$ 为嵌入函数; \mathbf{PE}_{2D} 为二维位置编码矩阵. 原始的 R-D 谱经过预处理后,输入到时空变分 Transformer 的推理网络映射到隐变量 \mathbf{z} 的均值 $\boldsymbol{\mu}_z$ 和方差 $\boldsymbol{\sigma}_z$,同时在推理网络中输出开关 FFN 层的权重向量 $\mathbf{u}(\mathbf{x})$ 传递到生成网络的开关 FFN 层中. 时空变分 Transformer 网络的推理过程可以描述为

$$\begin{cases} \boldsymbol{\mu}_z = \text{SVT}_{1\boldsymbol{\mu}}(\mathbf{x}') \\ \boldsymbol{\sigma}_z = \text{SVT}_{1\boldsymbol{\sigma}}(\mathbf{x}') \end{cases} \quad (26)$$

$$\mathbf{u}(\mathbf{x}) = \text{SVT}_{1\mathbf{u}}(\mathbf{x}') \quad (27)$$

其中, $\text{SVT}_{1\boldsymbol{\mu}}(\cdot)$ 、 $\text{SVT}_{1\boldsymbol{\sigma}}(\cdot)$ 、 $\text{SVT}_{1\mathbf{u}}(\cdot)$ 均为时空变分 Transformer 的编码网络,使用 vanilla Transformer 网络分别映射得到隐变量的均值 $\boldsymbol{\mu}_z$ 、方差 $\boldsymbol{\sigma}_z$ 和开关 FFN 层权重向量 $\mathbf{u}(\mathbf{x})$. 经过式 (12) 所示的重参数化技巧从分布中采样得到 \mathbf{z} ,再经过时空变分 Transformer 的生成网络重构出 R-D 谱样本 $\tilde{\mathbf{x}}$:

$$\mathbf{z} = \boldsymbol{\mu} + \boldsymbol{\sigma} \odot \boldsymbol{\varepsilon} \quad (28)$$

$$\tilde{\mathbf{x}} = \text{ST}_2(\mathbf{z}, \mathbf{u}(\mathbf{x})) \quad (29)$$

其中, $\boldsymbol{\varepsilon}$ 服从标准均匀分布; $\text{ST}_2(\mathbf{z}, \mathbf{u}(\mathbf{x}))$ 为时空变分 Transformer 的生成网络,使用引入开关机制的 Switching Transformer 结构,如图 4 所示. 利用采样得到的隐变量 \mathbf{z} 和开关 FFN 层权重向量 $\mathbf{u}(\mathbf{x})$ 经过时空变分 Transformer 生成网络,重构出与输入样本更加相似的 R-D 谱.

具体地,一个 R-D 谱样本经过时空变分 Transformer 推理-生成网络的尺寸参数变化过程如图 7 所示,为了更好地学习杂波分布特性,本文将 Transformer 模型作为重构网络的骨干网络, R-D 谱样本输入重构网络之前,需要进行切片操作,将 R-D 谱切分为图像块后再经过 rearrange 层, rearrange 层主要是将切片后的多维数组进行维度转换,之后经过线性层,其输出作为重构网络模型的输入. 经过变分 Transformer 模型后的输出与输入具有相同的维度,然后经过与预处理相对应的反向操作恢复到原始样本尺

寸,得到重构样本. Transformer 网络依靠自注意力机制能够捕获 R-D 谱数据的全局相关性,为了匹配 VAE 模型的推理-生成架构,在 Transformer 模型的基础上引入了开关 Transformer 机制,捕获全局相关性的同时,

充分学习 R-D 谱不同区域杂波分布的特征. 同时针对 R-D 谱的二维数据结构,引入二维位置编码机制,保留原始数据的二维时空信息,充分挖掘杂波分布的特征.

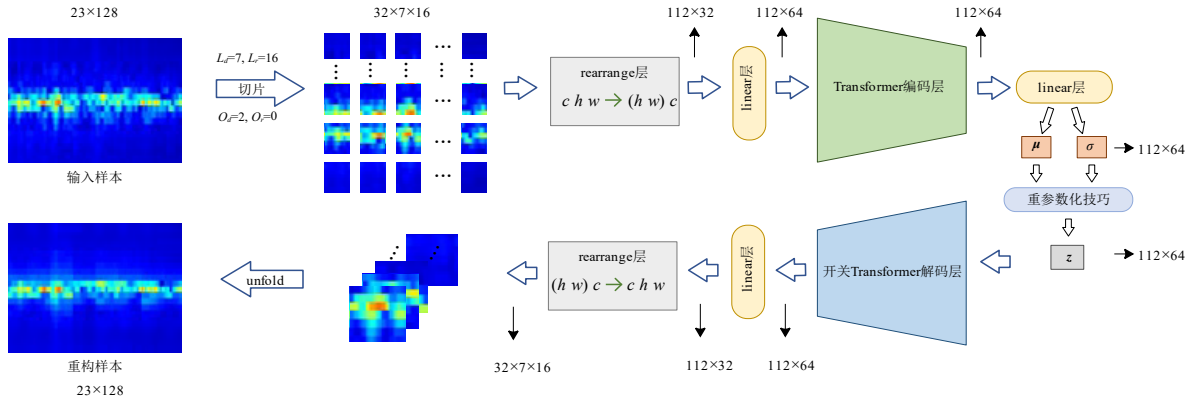


图7 R-D谱重构过程的尺寸参数变化图

需要说明的是,在考虑时空变分 Transformer 模型训练时的目标函数时,开关 Transformer 机制的引入会导致负载不平衡的问题,即在样本中最初几个 FNN 模型表现比较好,但在之后的样本中会被更加频繁地选择,并得到充分的优化. 这使得其余的 FNN 模型无法被充分地训练. 针对这个问题,需要在原有的目标函数上加入一个额外的损失项,让每个 FNN 模型都能够得到充分训练. 据此,结合 2.2 节 VAE 内容,用于杂波 R-D 谱建模的时空变分 Transformer 模型的目标函数可以进一步表示为

$$L(\theta, \phi) = -\text{ELBO} + \alpha \cdot K \cdot \sum_{i=1}^K G^2(i) \quad (30)$$

$$G(i) = \frac{1}{N_i} \sum_{j=1}^{N_i} g_j(i) \quad (31)$$

$$g_j(i) = \begin{cases} 1, & \text{if } \operatorname{argmax} \mathbf{u}(\mathbf{x}_j) = i \\ 0, & \text{others} \end{cases} \quad (32)$$

其中,式(30)中第一项 ELBO 即为 2.2 节式(8);第二项引入针对开关 Transformer 机制的损失项,确保 K 个 FNN 模型都能够得到充分训练;式(31)中 $G(i)$ 表示总的训练样本选取第 i 个 FNN 模型的平均概率; N_i 为训练样本总数; K 为开关 FNN 层中 FNN 模型的个数.

4 结合杂波建模与 OOD 策略的雷达目标检测方法

为了将本文所提的无监督变分重构模型应用于雷达目标检测任务,结合 OOD 检测策略^[23],本文提出了一种基于深度无监督变分网络的杂波建模与雷达目标检测方法.

4.1 训练阶段

具体地,首先对输入的 R-D 谱数据进行卷积、展平和切片等预处理,以匹配推理-生成模型中的不同网络架构. 其次,利用预处理后的杂波 R-D 谱训练集离线训练 2.3 节所提的基于 CNN、RNN 结构骨干和基于时空变分 Transformer 的三种重构模型,更新网络模型参数直到网络收敛,训练好的模型参数学习到了不同杂波的分布信息. 最后,将预处理后的杂波 R-D 图输入到训练好的模型中,模型会输出与输入数据相似杂波分布的重构数据,如图 8 所示.

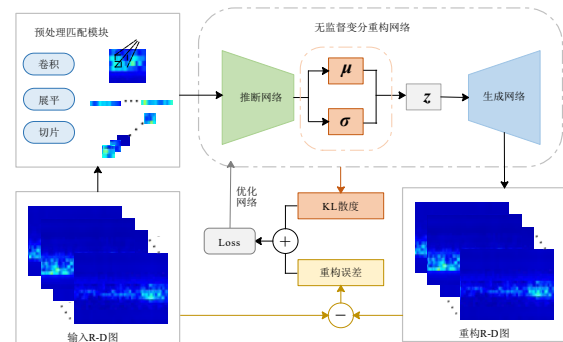


图8 模型训练过程

在模型训练过程中,式(30)目标函数的值会越来越小,输入 R-D 谱样本与重构样本之间也越来越相似. 在模型训练完成后,输入不同类型下的杂波分布,模型都能够重构出相似杂波分布的样本,说明模型通过参数的梯度更新学习了不同复杂环境下杂波分布的特征信息,并且模型具有了一定的泛化能力.

4.2 测试阶段

基于深度无监督变分网络的杂波重构模型训练完

成后,结合 OOD 检测策略,对待测样本进行检测. 根据 OOD 检测策略,利用训练好的模型进行测试,如果输入的样本是杂波数据,则模型能够很好地重构出杂波分布. 如果输入的样本是目标数据,则模型不能很好地重构出目标数据. 测试过程如图 9 所示. 根据上述现象,模型所输出的样本的重构显然代表了重构出输入样本的难易程度.

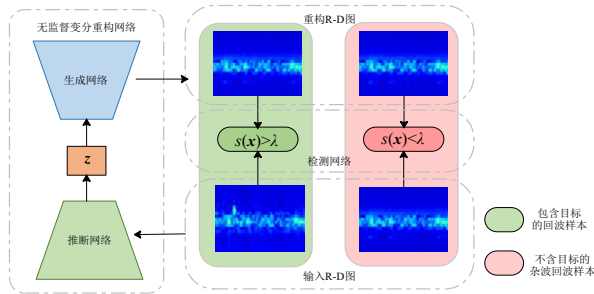


图9 模型测试过程

重构似然越大,重构样本与输入样本越相似;重构似然越小,说明对样本进行准确地重构就越困难. 因此,本文依据测试样本 x 的重构似然来定义最终得分,称作为 OOD 分数,记为 $s(x)$,并将其与阈值 λ_0 进行比较,以确定测试样本是否为目标,OOD 分数表示如下:

$$s(x) = 1 - M(\ln p(x|z)) \quad (33)$$

$$M(x) = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \quad p(x|z) = \frac{1}{\sqrt{2\pi} \sigma_x} \exp\left(-\frac{(x - \mu_x)^2}{2\sigma_x^2}\right) \quad (34)$$

其中, x 为输入的测试样本; μ_x, σ_x 为生成过程所服从的高斯分布的均值和方差; $p(x|z)$ 为重构似然; $M(\cdot)$ 为最值归一化函数. 对于含有目标的样本,模型不能很好地重构,其重构似然通常小于只含有杂波的样本. 因此,如果 $s(x)$ 分数大于阈值 λ_0 ,则将待检测的样本判定为 OOD 样本,即判定为目标;否则,将待检测样本判定为纯杂波.

5 实验结果与分析

为了验证本文提出的基于深度无监督变分网络的杂波建模与雷达目标检测方法的有效性,本文首先在 VAE 模型的推理-生成网络中分别引入了 CNN、RNN 网络,设计了 CNN-VAE、RNN-VAE 杂波重构模型. 其次,引入基于深度学习的自编码器检测方法,自编码器模型相比于所提的时空变分 Transformer 模型,去掉了变分模块. 模型由原来的概率映射变为确定性映射,且编码-解码网络均使用 vanilla Transformer 模型,增加自编码器网络检测实验有助于对比变分模块对检测性能的影响. 最后,本文提出了一种基于时空变分 Transformer 的杂波重构方法,结合 OOD 检测策略,对四种方法在雷达

目标检测中的性能表现进行分析.

5.1 实验数据

本文采用雷达仿真数据作为实验数据,仿真依据是针对某一地区警戒雷达实测数据,用不同的统计模型对各种类型的实测杂波数据进行参数拟合,利用实测数据计算出不同杂波模型对应的参数,最后利用得到的统计模型参数仿真出不同杂波类型场景的数据. 为了适应不同地形下不同的杂波分布特性,实验仿真了 4 种杂波类型,对应的 R-D 谱 3D 图如图 10 所示,图 10(a)、(b)、(c) 分别表示 3 种不同类型的地杂波,图 10(d) 表示一种气象杂波,本文采用零线性非记忆变换法来仿真雷达杂波数据^[24]. 其中,三种地杂波幅度谱由 K 分布和 Weibull 分布描述^[25],气象杂波幅度谱由对数正态分布描述^[26]. 实验仿真数据是均不包含目标的纯杂波背景,不同场景下的数据按照 4:1 的比例划分为训练数据和测试数据,其中训练数据共 5 050 帧,测试数据 1 220 帧. 对于测试数据,其中的 600 帧数据不做变动,仍作为杂波数据,其余的 620 帧数据在靠近杂波位置加入仿真目标,实验分别仿真了靠近地杂波的慢速目标和靠近气象杂波的快速目标,仿真目标的 SCNR 范围为 $-5 \sim 30$ dB, SCNR 计算公式表示为 $SCNR = 10 \lg(\alpha^2/P_{cn})$, α^2 表示目标功率, P_{cn} 表示杂波与噪声的平均功率. 训练数据中只存在杂波和噪声,杂噪比计算公式表示为 $CNR = 10 \lg(P_c/P_n)$, P_c 表示平均杂波功率, P_n 表示平均噪声功率. 仿真得到地杂波的平均杂噪比为 35 dB,气象杂波的平均杂噪比为 25 dB. 具体的仿真参数如表 1 所示.

5.2 实验设置

实验中关于 CNN-VAE、RNN-VAE、自编码器和时空变分 Transformer 四种杂波重构模型的超参数设置如下: CNN-VAE 模型的编码器的卷积层数和解码器中的反卷积层数设置为 6 层; RNN-VAE 模型的编码器和解码器采用 LSTM 结构,层数设置为 2 层, CNN-VAE 模型和 RNN-VAE 模型的隐变量维度大小均设置为 256. 时空变分 Transformer 模型的编码器和解码器层数设置为 2 层;隐变量维度大小设置为 256;数据预处理中 R-D 谱的切片大小为 $L_d=7, L_r=16$;重叠长度大小为 $O_d=2, O_r=0$.

开关 Transformer 模型中 FFN 网络个数 K 设置为 4;自编码器模型相比于时空变分 Transformer 模型,去掉了变分模块,且编码-解码网络不做任何改进,均使用 vanilla Transformer 模型,其余参数保持相同. 四种模型均训练了 500 次,训练批次大小为 64,模型使用 Adam 优化器进行训练,学习率设置为 0.000 5. 试验在 Intel i9-14 900 3.2 GHz 的 CPU 主机上进行,采用 NVIDIA RTX 4080 加速.

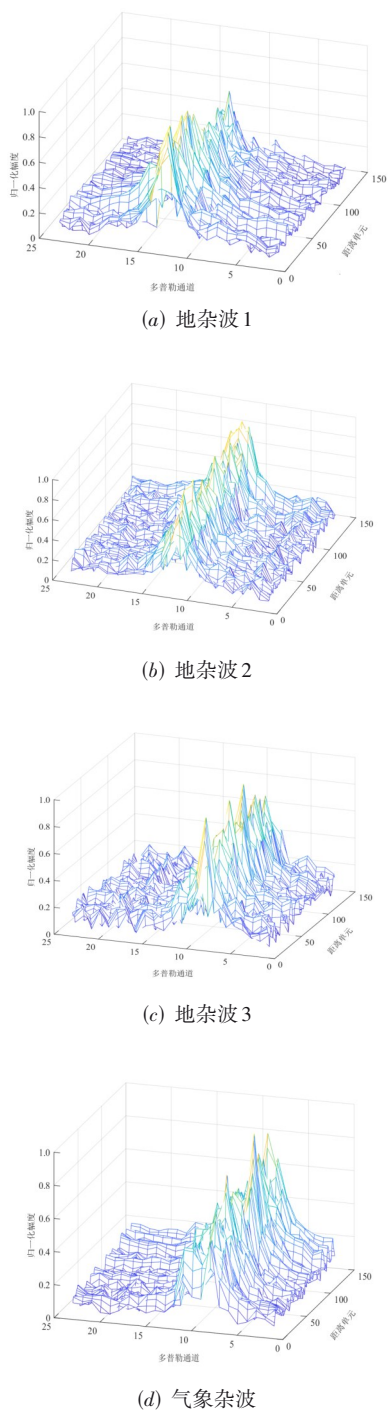


图10 训练样本中的四种杂波类型 R-D 谱

5.3 评价指标

本文采用检测率和虚警率^[27]作为雷达目标检测任务的评估指标,来探究 CNN-VAE、RNN-VAE、时空变分 Transformer 三种杂波重构模型的性能. 检测率和虚警率可表示为

$$P_d = \frac{TP}{TP + FN}, \quad P_{fa} = \frac{FP}{TP + FP} \quad (35)$$

其中,检测率 P_d 反映了被正确判定的正例占总的正例的比重; TP 为正确分类的目标样本数量; FN 为错误分类到杂波样本的目标样本数量; 虚警率 P_{fa} 反映了被判为正例的样本中, 负例所占的比例; FP 为错误分类到目标样本的杂波样本数量.

在仿真实验分析中, 实际测得的虚警率与预先设置的虚警率可能存在较大的差异, 本文将按照如下方法步骤来控制实际虚警率:

(1) 4.1 节介绍的无监督变分重构模型训练完成后, 首先, 将 5 050 帧训练数据输入网络模型得到对应的重构数据, 根据重构似然得到对应的 5 050 帧检测结果, 每一帧检测结果是大小为 23×128 的图像矩阵, 存放对应的 OOD 值, OOD 值越大, 说明该点越易被检测为虚警点.

(2) 由预先设置的虚警率来决定检测阈值 λ_0 . 步骤(1) 中得到的 5 050 帧的所有 OOD 值按照降序排列, 按照设置的虚警率找到对应索引下的 OOD 值, 将该 OOD 值作为目标测试过程中的检测阈值 λ_0 .

(3) 在测试过程中, 将 1 220 帧测试数据输入无监督变分重构网络得到对应的重构样本, 同时, 根据重构似然得到 1 220 帧检测结果. 对于检测结果中 OOD 值大于检测阈值 λ_0 的点, 被判定为目标点; 反之, 被判定为杂波点.

考虑到模型在训练过程中学习到了杂波的分布特性, 在测试过程中会重构生成类似的杂波分布, 本文采用上述方法策略来控制虚警率. 因此, 用设置的虚警率在训练数据上测试得到的检测阈值可以适用于测试数据, 通过实验验证, 该策略实际检测得到的虚警率接近预先设置的虚警率.

5.4 实验结果及分析

为了评估所提方法的有效性, 本文进一步选择传统的 CA-CFAR 检测方法并结合动目标显示 (Moving Target Indicator, MTI)^[28] 的 CA-CFAR 检测方法与本文所提的基于深度无监督变分网络的杂波建模与雷达目标检测方法进行比较. 首先, 本文设计了三个目标检测实验. 第一个实验设计在三种不同类型的地杂波附近检测低速目标, 测试所提方法相较于传统 CFAR 方法的性能改进情况以及不同网络模型在复杂背景下的建模能力. 第二个实验设计在气象杂波附近检测高速目标, 验证所提方法在多普勒中心偏移条件下仍然具备检测性能. 第三个实验探究了在更低的虚警率设置下, 各个检测模型在四种杂波环境下的目标检测能力. 其次, 针对时空变分 Transformer 模型, 设计了所提 2-D 位置编码与原始 Transformer 使用的 1-D 位置编码检测性能的对比实验. 最后, 针对本文所提的 CNN-VAE、RNN-VAE、时空变分 Transformer 重构模型, 设计实验探究了三种模型各自的训练效率.

5.4.1 检测性能分析

图 11 研究了在三种地杂波环境下, 对于不同 SCNR

表 1 雷达信号仿真参数

雷达信号仿真参数	信号调制类型	LFM			
	带宽 B	3 MHz			
	脉宽 τ	150 μ s			
	脉冲重复频率 f	500 Hz			
	距离分辨率 Δr	25 m			
	波束宽度 θ	2°			
杂波信号仿真参数	距离维度	128			
	多普勒维度	23			
	杂波类型	type1	type2	type3	type4
	幅度服从分布	Weibull分布	K 分布	K 分布	对数正态
	分布参数	$\lambda=0.9,$ $k=1.5$	$a=1.0,$ $v=0.9$	$a=1.5,$ $v=1.2$	$\mu_0=1.2,$ $\sigma_0=1.0$
	地杂波杂噪比	35 dB			
	气象杂波杂噪比	25 dB			
	SCNR	-5~30 dB			
	慢速目标速度	2.7 m/s			
	快速目标速度	13.6 m/s			
	网络训练参数	训练样本数	5 050 帧		
测试样本数		1 220 帧			
地杂波测试样本数		915 帧			
气象杂波测试样本数		305 帧			
训练轮次		500			
批处理大小		64			
学习率		0.000 5			

的目标,所提方法与传统 CFAR 检测器的 SCNR- P_d 曲线和 SCNR- P_{fa} 曲线. 其中,目标 SCNR 的设置为一5~30 dB,检测虚警率设置为 10^{-3} . 由于目标位于杂波附近,传统的 CFAR 检测方法在检测目标时受到杂波严重干扰,导致检测率下降. MTI 处理后的 CFAR 检测器在一定程度上降低了杂波干扰,但同时目标也有所抑制,检测率提升并不高. 相比之下,本文提出的基于深度无监督变分网络的建模方法通过对杂波进行精细化建模,结合 OOD 检测策略实现了更好的检测性能. 在 80% 的检测概率下,时空变分 Transformer 模型相比于 CFAR-MTI, SCNR 优化了 5.6 dB,这主要得益于时空变分 Transformer 模型对杂波分布的精细建模,从而对强杂波进行有效抑制,使得凸显出的目标能够更容易被检测. 此外,在所提的深度无监督的检测框架下,时空变分 Transformer 模型,相较于 CNN-VAE、RNN-VAE 模型,SCNR 均优化了 2 dB 以上,说明了在所提深度无监督的检测框架下,Transformer 模型相比于 CNN 和 RNN 模型更具适配性和优越性. 基于深度学习的自编码器模型的检测性能相对较低,主要原因是样本在重构过程中目标出现了信息泄露,杂波分布被模型重构出来的同时,部分目标信息也会被重构出来,导致自编码器方法的检测性能下降. 在虚警率方面,如图 11(b)所示,

所提方法中,RNN-VAE、时空变分 Transformer 模型的虚警率均低于传统 CFAR 检测器,CNN-VAE 模型因主干网络限制,泛化能力较差,对一些少量的异常样本不能很好地重构,导致在这些少量的样本上检测出大量的虚警点,使得虚警率偏高. 由于目标信息泄露问题,自编码器模型的虚警率也会偏高,随着 SCNR 增加,目标信息泄露的现象加重,导致检测的虚警率会逐渐升高.

图 12 研究了比地杂波更复杂的气象杂波环境下,对于不同 SCNR 的目标,所提方法与传统 CFAR 检测器的 SCNR- P_d 曲线和 SCNR- P_{fa} 曲线. 由图 12 可知,在更复杂的气象杂波环境下,时空变分 Transformer 模型再次拉开了与其他基线的距离,在 80% 的检测概率下,相较于 CNN-VAE、RNN-VAE 模型,时空变分 Transformer 模型的 SCNR 优化了约 2.5 dB. 在气象杂波环境中,自编码器检测方法同样出现了目标的信息泄露,导致检测性能降低. 同时,本文所提无监督重构方法的检测率均高于 CFAR 检测器,这说明基于无监督变分网络的建模方法具有很好的鲁棒性,能够应对各种杂波分布下的检测任务.

在实际场景中,通常要求检测方法在更低的虚警率下做到精确检测,本文将实验虚警率设置为 10^{-6} ,重新在四种杂波环境下进行性能检测. 图 13 展示了在四种仿真杂波环境下,对于不同 SCNR 的目标,各种检测方法的

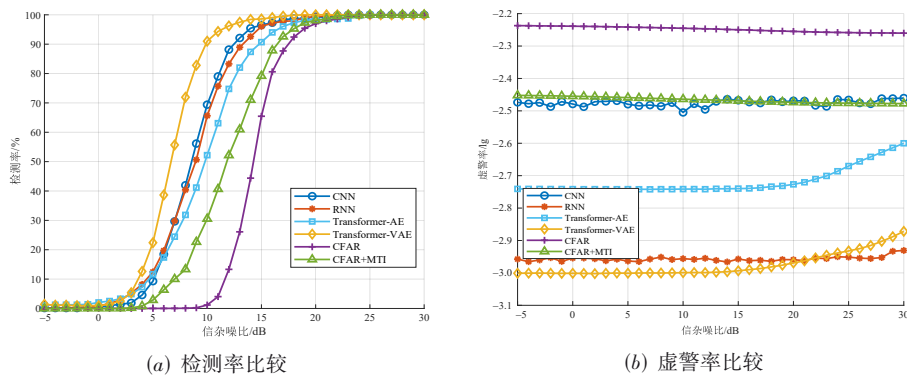


图 11 三种地杂波环境下各个模型在不同 SCNR 下的比较结果

SCNR- P_d 曲线和 SCNR- P_{fa} 曲线. 可以看出, 在低虚警率场景下, 传统 CFAR 方法已经对中低 SCNR 的目标失去了检测能力, 本文所提的时空变分 Transformer 模型在所

有基线方法中保持最优性能, 由图 13(b) 可以看出, 时空变分 Transformer 模型的实测虚警率也保持稳定在约 10^{-6} .

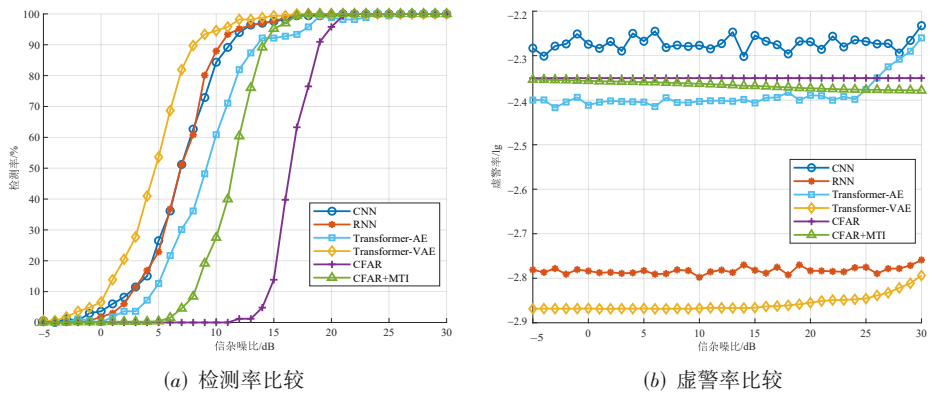


图 12 气象杂波环境下各个模型在不同 SCNR 下的比较结果

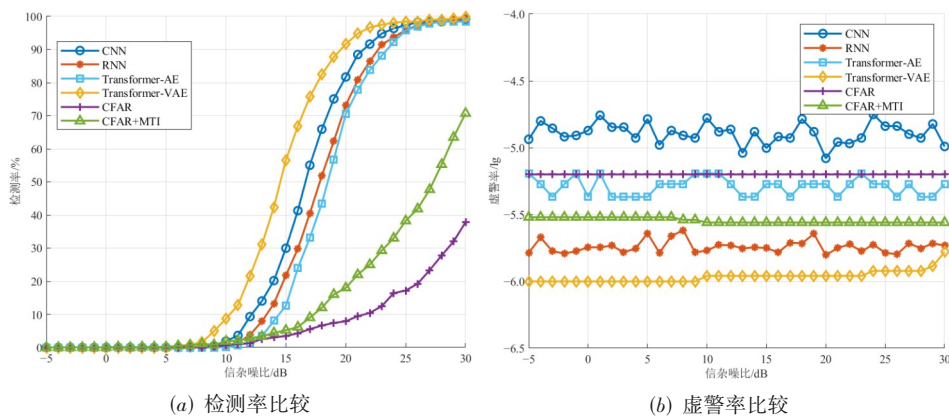


图 13 在四种杂波环境下各个模型在不同 SCNR 下的比较结果

针对时空变分 Transformer 模型, 本文设计了模型使用 2-D 位置编码与使用 1-D 位置编码时对目标检测性能的影响. 1-D 位置编码和 2-D 位置编码方式如图 6 所示, 与原始 Transformer 使用的 1-D 位置编码相比, 本文引入了 2-D 位置编码方式, 结合了多普勒维和距离维的二维时空信息, 指导模型更好地学习杂波的分布特

征. 图 14(a) 展示了在气象杂波环境下, 虚警率为 10^{-3} 时, 两种位置编码方式与传统 CFAR 的检测率随 SCNR 的变化曲线, 可以看出, 相比于 1-D 位置编码, 采用 2-D 位置编码方式的检测效果得到明显提升, 在 80% 的检测概率下, 2-D 位置编码方式相比于 1-D 位置编码, SCNR 优化了 2.2 dB. 说明二维位置编码方式考虑到样

本的时空信息,更适用于处理雷达 R-D 谱这种二维数据.图 14(b)结果表明使用 2-D 位置编码的 R-D 谱样本

重构似然普遍更高,对杂波建模的精度也就更高,尤其在少数的异常样本下,也能保持较高的重构水平.

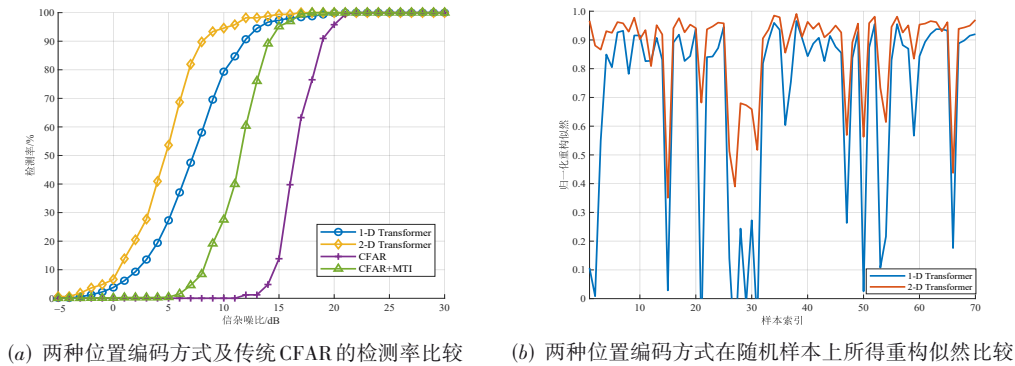


图 14 不同位置编码方式重构结果和检测结果比较

5.4.2 重构结果可视化

图 15(a)比较了测试数据中 100 个纯杂波 R-D 谱样本分别输入到 CNN-VAE、RNN-VAE、时空变分 Transformer 三种模型后得到的重构似然.图 15(a)可以看出在三种模型中,时空变分 Transformer 模型所得结果的重构似然基本保持在最高,说明该模型对杂波 R-D 谱的建模相较于 CNN-VAE、RNN-VAE 重构模型更加精细.图 15(b)针对某个带目标的 R-D 谱检测样本,其目标位于第 27 个距离单元上,比较三种模型在 128 个距

离单元上的 OOD 得分.可以看出,三种模型均在第 27 个距离单元上获得了较高的 OOD 得分,可认为在此距离单元上检测到了目标.但是在某些杂波区域也出现了较高的 OOD 分数,如第 21、126 个距离单元,其中,时空变分 Transformer 模型相比于 CNN-VAE 和 RNN-VAE 模型,在第 27 个距离单元以外的其他杂波区域上的 OOD 分数更低,使杂波与目标更容易区分,说明时空变分 Transformer 模型对杂波进行了更加精准的建模,避免了在检测过程中出现更多的虚警.

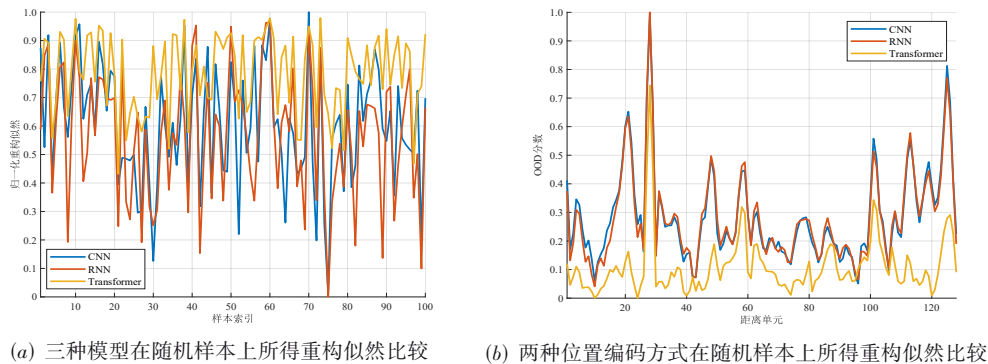


图 15 CNN、RNN 和 Transformer 重构结果可视化比较

图 16、图 17 研究了 CNN-VAE、RNN-VAE、自编码器模型和时空变分 Transformer 模型对带有目标的待测样本的重构效果.图 16 展示了在地杂波下三种模型的重构可视化结果,可以看出,时空变分 Transformer 模型对杂波的建模更加精细,这主要得益于时空变分 Transformer 模型的全局特征建模能力.自编码器检测模型的重构样本图中出现了目标的信息泄露,造成了部分目标信息的重构,导致目标的检测性能下降.CNN-VAE 模型中的卷积层通过局部感受野提取 R-D 谱的局部特征,并利用转置卷积操作来逐步恢复 R-D 谱的空间维度.因此,CNN-VAE 模型也能够获得较好的重构效果.

RNN-VAE 模型的重构效果不如 CNN-VAE 和时空变分 Transformer 模型,原因是 RNN 擅长处理序列数据间的依赖关系,二维 R-D 谱的特征信息难以通过 RNN 的时间步长机制很好地捕捉.图 17 展示了在气象杂波下三种模型的重构可视化结果,在杂波分布更为复杂的环境下,时空变分 Transformer 模型仍然能够保持很好的重构效果.

5.4.3 模型训练效率分析

为了探究 CNN-VAE、RNN-VAE、时空变分 Transformer 三种杂波重构模型的训练效率,本文在虚警率为 10^{-3} , SCNR 为 9 dB 的条件下,对三种模型不同训练次数下的检测性能设计了对比实验.图 18(a)展示了三种模

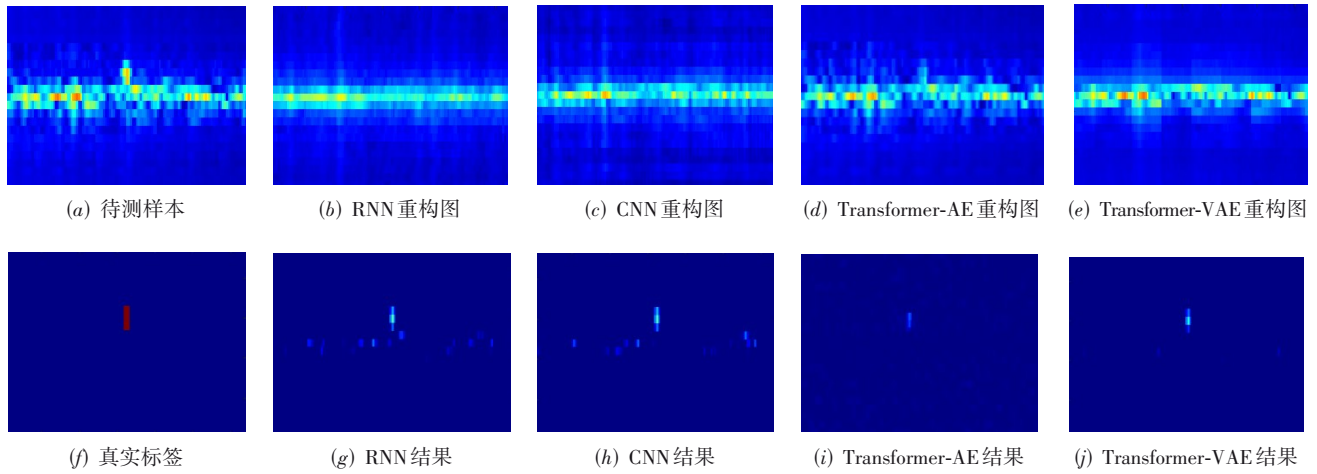


图 16 地杂波环境下不同模型 R-D 图建模可视化结果

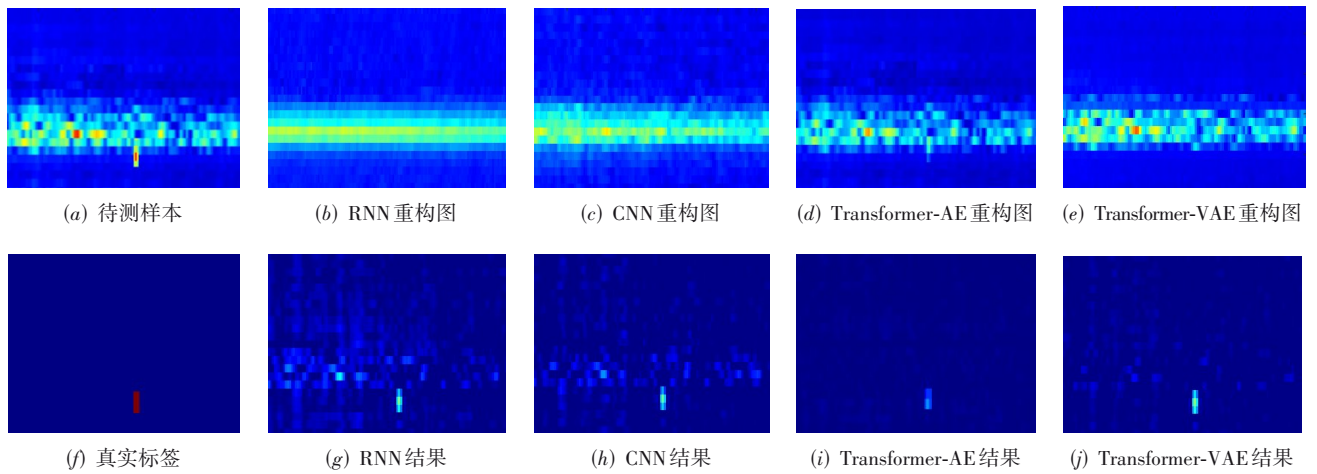


图 17 气象杂波环境下不同模型 R-D 图建模可视化结果

型的检测率随着训练次数的变化曲线,曲线的陡度反映了模型学习过程的收敛速度.

可以看出,相较于 CNN-VAE 和 RNN-VAE 模型,时空变分 Transformer 重构模型的收敛速度相对较慢,需要更多的训练步骤来对杂波分布进行精细地学习.但当模型都训练完成后,时空变分 Transformer 模型的检测

精度达到最高.图 18(b)展示了检测阈值随着训练次数的变化曲线.可以看出,训练过程中,模型的阈值越来越低,即杂波对应的 OOD 分数越来越低,模型对杂波分布的建模也越来越精细,模型均收敛后,时空变分 Transformer 模型检测阈值最低,说明其对杂波 R-D 谱建模更加精细.

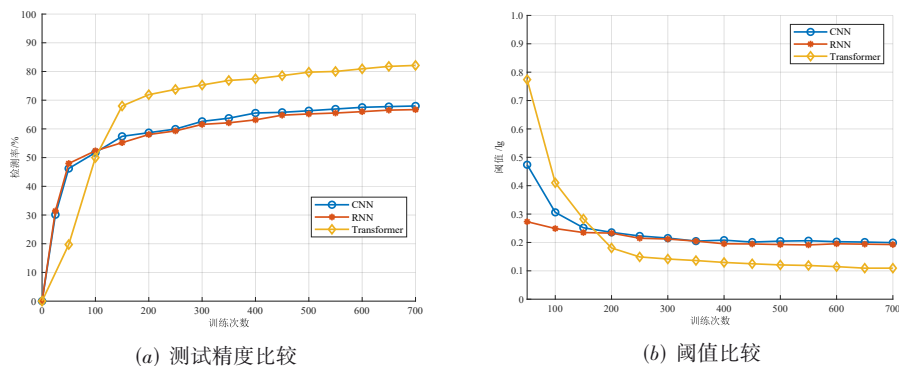


图 18 不同训练次数下三种模型比较结果

6 结论

本文针对复杂杂波背景下的雷达目标检测问题展开研究,首先,面向雷达 R-D 谱数据,提出了一种基于深度无监督变分网络的杂波建模方法.该方法利用面向高维分布特征学习的 VAE 模型对杂波分布进行建模,重构出与输入样本相似杂波分布的 R-D 谱.其次,为了更加精细化建模杂波分布特征,本文提出了一种基于时空变分 Transformer 的杂波建模方法,该方法将 Transformer 结构引入到所提的深度无监督杂波建模方法中,并设计了开关 Transformer 机制和二维位置编码机制以匹配网络架构.最后,结合 OOD 检测策略,将重构似然作为划分目标与杂波的依据,完成对雷达目标的检测.本文通过仿真实验表明:在复杂杂波分布背景下,所提的基于时空变分 Transformer 的杂波建模与目标检测方法在细粒度建模方面表现出色,能够有效地重构出不同杂波分布的场景.同时,融入 OOD 检测策略,实现了最先进的检测性能.所提的基于时空变分 Transformer 的目标检测方法主要面向的任务场景是对空,后续针对更多复杂场景,如海面场景,将进行更详细的分析与研究.

参考文献

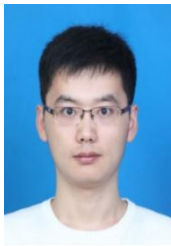
- [1] ZHANG C X, XU Y S, CHEN W C, et al. IfCMD: A novel method for radar target detection under complex clutter backgrounds[J]. *Remote Sensing*, 2024, 16(12): 2199.
- [2] NAVAS R E, CUPPENS F, BOULAHIA CUPPENS N, et al. MTD, where art thou? A systematic review of moving target defense techniques for IoT[J]. *IEEE Internet of Things Journal*, 2021, 8(10): 7818-7832.
- [3] WATTS S. Cell-averaging CFAR gain in spatially correlated K-distributed clutter[J]. *IEEE Proceedings - Radar, Sonar and Navigation*, 1996, 143(5): 321.
- [4] WEISS M. Analysis of some modified cell-averaging CFAR processors in multiple-target situations[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 1982, AES-18(1): 102-114.
- [5] 邹成晓, 张海霞, 程玉莹, 等. 雷达恒虚警技术处理方法综述[J]. *雷达与对抗*, 2021, 41(2): 29-35.
ZOU C X, ZHANG H X, CHENG Y K, et al. Review of radar constant false alarm technology processing methods[J]. *Radar and Countermeasures*, 2021, 41(2): 29-35. (in Chinese)
- [6] 郭辰锋. 复杂背景目标检测技术研究[D]. 哈尔滨: 哈尔滨工业大学, 2020.
GUO C F. Research on Target Detection Technology with Complex Background[D]. Harbin: Harbin Institute of Technology, 2020. (in Chinese)
- [7] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 770-778.
- [8] TIAN J M, WANG C X, CAO J W, et al. Fully convolutional network-based fast UAV detection in pulse Doppler radar[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5103112.
- [9] LE H, DOAN V S, LE D P, et al. Micro-Doppler-radar-based UAV detection using inception-residual neural network[C]//2020 International Conference on Advanced Technologies for Communications. Piscataway: IEEE, 2020: 177-181.
- [10] YANG Y P, YANG F, SUN L G, et al. Echoformer: Transformer architecture based on radar echo characteristics for UAV detection[J]. *IEEE Sensors Journal*, 2023, 23(8): 8639-8653.
- [11] SHI Y, DU L, GUO Y C. Unsupervised domain adaptation for SAR target detection[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 6372-6385.
- [12] SHI Y, DU L, GUO Y C, et al. Optical knowledge assisted unsupervised cross-domain SAR target detection[C]//IET International Radar Conference. London: IET, 2024: 1474-1480.
- [13] 张智, 易华挥, 郑锦. 聚焦小目标的航拍图像目标检测算法[J]. *电子学报*, 2023, 51(4): 944-955.
ZHANG Z, YI H H, ZHENG J. Focusing on small objects detector in aerial images[J]. *Acta Electronica Sinica*, 2023, 51(4): 944-955. (in Chinese)
- [14] KINGMA D P, WELLMING M. Auto-encoding variational Bayes[EB/OL]. (2022-12-10) [2025-03-10]. <https://arxiv.org/abs/1312.6114>.
- [15] ZHOU Y B. Rethinking reconstruction autoencoder-based out-of-distribution detection[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 7369-7377.
- [16] HE K M, CHEN X L, XIE S N, et al. Masked autoencoders are scalable vision learners[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 15979-15988.
- [17] JING H, CHENG Y Q, WU H, et al. Radar target detection with multi-task learning in heterogeneous environment[J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 4021405.

- [18] ZAREMBA W, SUTSKEVER I, VINYALS O. Recurrent neural network regularization[EB/OL]. (2015-02-19) [2025-03-10]. <https://arXiv.org/abs/1409.2329>.
- [19] VASWANI A. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30: 5998-6008.
- [20] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[EB/OL]. (2021-06-03) [2025-03-10]. <https://arXiv.org/abs/2010.11929>.
- [21] FEDUS W, ZOPH B, SHAZEER N. Switch transformers: Scaling to trillion parameter models with simple and efficient sparsity[EB/OL]. (2022-06-16) [2025-03-10]. <https://arXiv.org/abs/2101.03961>.
- [22] SHAZEER N, MIRHOSEINI A, MAZIARZ K, et al. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer[EB/OL]. (2017-01-23) [2025-03-10]. <https://arXiv.org/abs/1701.06538>.
- [23] LIU Z Z, ZHOU J J, WANG Y F, et al. Unsupervised out-of-distribution detection with diffusion inpainting[C]// Proceedings of the 40th International Conference on Machine Learning. Honolulu: Proceedings of Machine Learning Research, 2023
- [24] 杜勇, 李依林, 杨海粟. 基于 ZMNL 法的相关雷达杂波建模与仿真[J]. 火控雷达技术, 2012, 41(4): 11-14.
- DU Y, LI Y L, YANG H S. Modeling and simulation of ZMNL-based correlated radar clutter[J]. Fire Control Radar Technology, 2012, 41(4): 11-14. (in Chinese)
- [25] 汪斌. 基于概率统计模型的复杂杂波背景下雷达目标检测方法研究[D]. 西安: 西安电子科技大学, 2019.
- WANG B. Research on Radar Target Detection Method in Complex Clutter Background Based on Probabilistic Statistical Model[D]. Xi'an: Xidian University, 2019. (in Chinese)
- [26] 刘宁. 地杂波背景下雷达目标检测方法的研究[D]. 西安: 西安电子科技大学, 2018.
- LIU N. Research on Radar Target Detection Method Under Ground Clutter Background[D]. Xi'an: Xidian University, 2018. (in Chinese)
- [27] LIANG X L, CHEN B, CHEN W C, et al. Unsupervised radar target detection under complex clutter background based on mixture variational autoencoder[J]. Remote Sensing, 2022, 14(18): 4449.
- [28] BRENNAN L, MALLETT J, REED I. Adaptive arrays in airborne MTI radar[J]. IEEE Transactions on Antennas and Propagation, 1976, 24(5): 607-615.

作者简介



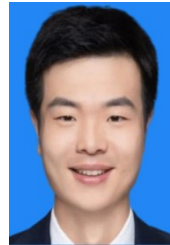
刘要强 男, 1999 年出生于河南省周口市。现为西安电子科技大学电子工程学院硕士生。主要研究方向为雷达目标检测。
E-mail: 23021212022@stu.xidian.edu.cn



陈文超 男, 1992 年出生于山东省济宁市。现为西安电子科技大学电子工程学院副教授。主要研究方向为统计机器学习、雷达目标识别与检测、大规模数据处理以及深度学习。中国电子学会会员编号: E190045537M。
E-mail: chenwenchao@xidian.edu.cn



施力行 男, 1998 年出生于浙江省湖州市。现为西安电子科技大学电子工程学院博士研究生。主要研究方向为雷达信号处理、雷达目标检测。
E-mail: shicyann@stu.xidian.edu.cn



田隆 男, 1992 年出生于陕西省西安市。现为西安电子科技大学计算机科学与技术学院副教授。主要研究方向为稀疏样本表示学习方法研究和智能信号处理方法研究。
E-mail: tianlong@xidian.edu.cn



王鹏辉 男, 1984 年出生于河南省洛阳市。现为西安电子科技大学电子工程学院教授。主要研究方向为雷达目标识别与跟踪、信号处理以及机器学习等。中国电子学会会员编号: E190017546S。
E-mail: wangpenghui@mail.xidian.edu.cn



陈渤 男, 1979 年出生于河南省郑州市。现为西安电子科技大学电子工程学院教授。主要研究方向为统计信号处理、雷达信号处理、机器学习、深度学习网络及在雷达目标检测识别方面的应用。中国电子学会会员编号: E190015313M。
E-mail: bchen@mail.xidian.edu.cn