基于频域多目标优化的SAR图像对抗样本生成方法

刘洁怡,李明哲,杨曜铭,李 豪,周 宇,党可林
(西安电子科技大学电子工程学院,陕西西安 710071)

摘 要: 基于深度学习的合成孔径雷达(Synthetic Aperture Radar,SAR)目标识别方法在军事侦察、灾害监测等领域应用广泛,然而深度神经网络易受到对抗攻击的威胁,导致模型决策的可靠性下降.现有黑盒对抗攻击方法在SAR图像对抗样本生成过程中面临参数设计维度高、易被察觉等问题.针对以上问题,提出一种基于频域多目标优化的对抗攻击方法,通过二维离散傅里叶变换将SAR图像从空间域映射至频域,降低扰动设计维度,进而在频域中修改单一频率分量,以生成图像域纹理状扰动.同时,结合基于超体积度量的多目标进化算法平衡对抗样本的攻击性能与视觉隐蔽性.实验结果表明,以T62类别为例,运用本文方法后,在VGG16、AConvNet和YOLO系列模型架构上,对抗样本分别实现了90.39%、71.43%、44.28%以上的置信度错误分类.同时,生成的对抗样本与原始图像的相似度均高于99%,为SAR图像的安全性与鲁棒性测试提供了有效的技术支持.

关键词: SAR图像识别;对抗样本攻击;频域转换;黑盒攻击;多目标优化算法
基金项目: 国家自然科学基金(No.62106185)
中图分类号: TN974 **文献标识码:** A **文章编号:** 0372-2112(2025)06-1958-11
电子学报URL:http://www.ejournal.org.cn **DOI:**10.12263/DZXB.20250095

A Multi-Objective Optimization Method in the Frequency Domain for SAR Image Adversarial Sample Generation

LIU Jie-yi, LI Ming-zhe, YANG Yao-ming, LI Hao, ZHOU Yu, DANG Ke-lin
(*School of Electronic Engineering*, *Xidian University*, *Xi'an*, *Shaanxi* 710071, *China*)

Abstract: Deep learning-based synthetic aperture radar (SAR) target recognition methods are widely used in military reconnaissance and disaster monitoring. However, deep neural networks (DNNs) are vulnerable to adversarial attacks, which compromise the reliability of model decisions. Existing black-box adversarial attack methods for SAR images face challenges such as high-dimensional parameter design and perceptible perturbations. To address these issues, a frequency-domain multi-objective optimization-based adversarial attack method is proposed. By transforming SAR images from the spatial domain to the frequency domain via 2D Discrete Fourier Transform, the method reduces perturbation design complexity and modifies a single frequency component to generate texture-like perturbations in the spatial domain. A hypervolume metric-guided multi-objective evolutionary algorithm is integrated to balance attack performance and visual stealthiness. Experimental results demonstrate that, for the T62 category, the adversarial samples generated by our method achieve misclassification confidence rates of more than 90.39%, 71.43%, 44.28% on VGG16, AConvNet, and YOLO series models, respectively. Additionally, the similarity between adversarial and original images exceeds 99% across all cases, providing effective technical support for security and robustness evaluation of SAR imaging systems.

Key words: SAR image recognition; adversarial sample attack; frequency-domain transformation; black-box attack; multi-objective optimization algorithm
Foundation Item(s): National Natural Science Foundation of China (No.62106185)

1 引言

合成孔径雷达(Synthetic Aperture Radar,SAR)通过发射电磁波并接收目标的回波信号,能够生成高分辨率的地表图像,在云层覆盖或光照不足等光学传感器受限的条件下表现得尤为出色.这一特性使其在军事目标识别、灾害应急监测及国土安全等关键领域发挥

收稿日期:2025-02-06;修回日期:2025-05-29;责任编辑:赵克

着不可替代的作用[1,2]. 传统的SAR图像识别方法主要依赖于手工特征提取与经典算法,处理过程复杂且受限于特征的表达能力. 随着深度学习技术的发展,神经网络因其强大的特征提取和模式识别能力,在SAR图像的目标识别、分类和变化检测任务中得到了广泛应用[3~5]. 相比于传统技术,基于深度学习的SAR图像处理技术在自动特征学习和处理复杂模式上具有显著优势,极大地提升了识别精度和效率.

然而,基于深度神经网络的SAR识别模型易受对抗样本攻击威胁[6~8],攻击者通过添加人眼难以察觉的细微扰动即可误导模型输出,导致军事目标误判或关键信息泄露,严重威胁SAR系统的可靠性与应用安全性. 现有对抗样本生成方法主要分为白盒攻击和黑盒攻击. 在白盒攻击中,Li等人[9]将光学图像快速梯度符号法(Fast Gradient Sign Method,FGSM)和基本迭代法(Basic Iterative Method,BIM)两种经典攻击算法应用到SAR图像,仿真结果表明FGSM运行速度快但攻击成功率有限,BIM虽然运行速度慢但成功率更高. Du等人[10]不局限于直接使用光学图像对抗攻击方法,提出针对一种SAR图像生成对抗样本的算法Fast C&W,通过引入深度编码器网络加速传统C&W算法的优化过程,从而显著提高了对抗样本生成的效率和效果. 这些白盒攻击依赖模型梯度信息生成扰动,虽攻击效率高,但其强假设(需知晓模型参数与架构)在实际场景中难以满足. 因此,近几年黑盒攻击变得日益流行,如Square Attack[11]和simBA[12]. Du等人[13]提出了基于UNet和生成对抗网络(Generative Adversarial Networks,GAN)的Attack-UNet-GAN方法,利用UNet提取分离特征,并通过GAN训练生成高质量的对抗样本. 为了使攻击具有现实意义,Cui等人[14]提出SAR-AE-SFP-Attack方法. 该方法利用SAR成像的相干能量累积和电磁波散射特性,通过RaySAR模拟目标的反射假体,调整目标对象的散射特征参数来生成对抗样本. Zhang等人[15]提出了一种基于频域流场攻击的方法,通过离散余弦变换将SAR图像从空间域转换到频域,直接优化流场以生成对抗样本. Huang等人[16]提出了一种结合频域和空间域交替优化的对抗样本生成方法,通过迭代更新频域全局扰动和空域局部扰动,并利用梯度引导与裁剪约束,显著提升了对抗样本的迁移性.

尽管现有的黑盒攻击方法在图像领域取得了一定的成果,但在SAR图像上,攻击效果和隐蔽性仍面临显著挑战[17~19]. 与光学图像不同,SAR图像包含丰富的振幅信息与相位信息,这使得SAR图像的对抗攻击需融入更强的物理意义. 传统的空间域(原始图像的像素域)扰动优化方法无法充分利用SAR图像的频域特性,限制了其攻击效果. 现有的黑盒方法,尽管在一定条件下可取得成功,且算法的攻击成功率和效率较高,但在处理大尺寸图像时面临设计参数过大的瓶颈,也未充分考虑对抗样本的隐蔽性. 为克服上述设计参数过多、对抗样本与原始图像差异过大的问题,本文提出一种基于频域单一分量优化的对抗样本生成方法,通过深入探索SAR图像的成像原理,挖掘频域扰动的潜力,旨在生成更高效、更不易被察觉的对抗样本.

如图1所示,本文提出针对SAR图像识别的频域多目标黑盒对抗攻击(Frequency-domain Multi-objective black-box Adversarial Attack,FMAA)框架,该框架旨在通过优化频域信息的扰动,提升SAR对抗样本的攻击性能与隐蔽性. 与传统方法直接在空间域进行扰动不同,本文将优化过程转移至频域,利用频率分量的稀疏表示代替原始图像的高维像素数据. 这样可以使优化

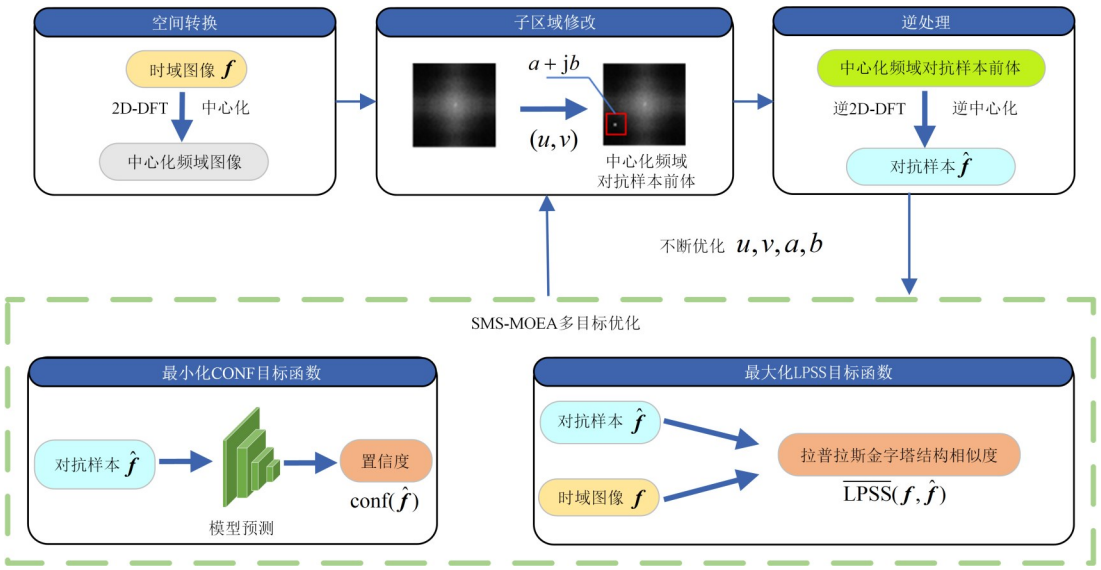


图1　算法流程图

过程集中在少量关键频率分量上，从而减少设计变量的数量，简化问题并降低计算开销. 此外，针对频域黑盒对抗攻击过程中的频域分量修改选择的权衡难题，还创新性地引入多目标优化框架，构建两个核心目标：一是最小化分类器对对抗样本的正确分类置信度，保证对抗样本的威胁性；二是最大化拉普拉斯金字塔结构下对抗样本与原始图像的相似度，确保样本在视觉效果上无法察觉变化. 最后，为权衡攻击性能与隐蔽性能，通过建立基于超体积度量的多目标优化框架，获得满足多目标优化需求的最优解集. 超体积度量能够有效评估多目标优化解集的质量，通过衡量解集的覆盖面积和目标空间的分布情况，确保在优化过程中不仅能提升对抗能力，还能维持良好的视觉隐蔽效果，使得本文在对抗性能和隐蔽性之间实现了最佳平衡.

2 本文方法

为了应对SAR图像的高分辨率、复数特性和现有黑盒攻击方法的局限性，本文提出了一种基于频域多目标优化的黑盒对抗攻击方法(FMAA)，该方法将图像从空间域映射到频域，利用跨域转换简化问题的复杂性，显著降低数据维度以突出关键特征，并优化解的搜索空间. 在此基础上，构建一个多目标优化框架，全局协调攻击性能与隐蔽效果之间的冲突目标，提升对抗样本的整体质量.

2.1 SAR图像频域映射

在高分辨率SAR图像中，直接在空间域添加扰动会导致设计变量数量庞大，优化问题变得复杂且计算开销巨大. 因此，将问题从空间域映射到频域可有效降低计算的复杂度，减少参数空间的维度，从而提升优化过程的效率. 并且，频域将图像的低频与高频信息清晰分离，使得修改频域分量既能更有效地保持图像质量，又能精准地控制扰动[20].

基于上述问题，将SAR图像从空间域映射到频域，便于聚焦关键频率分量进行优化，减少设计变量的数量，简化问题复杂性[21]. 具体处理方式为，对于一个M行N列的二维图像$\boldsymbol{f}$，$\boldsymbol{f}(x,y)$代表了图像$\boldsymbol{f}$在空间坐标(x,y)处的像素值. 先按行队列变量y做一次长度为N的一维离散傅里叶变换，再将计算结果按列向对变量x做一次长度为M的傅里叶变换，可以得到该图像的傅里叶变换结果$\boldsymbol{F}$：

$$\boldsymbol{F}(u,v)=\frac{1}{MN}\sum_{x=0}^{M-1}\left[\sum_{y=0}^{N-1}\boldsymbol{f}(x,y)\exp\left(\frac{-\mathrm{j}2\pi vy}{N}\right)\right]\exp\left(\frac{-\mathrm{j}2\pi ux}{M}\right) \tag{1}$$

其中，$\boldsymbol{F}$是$\boldsymbol{f}$经过2D-DFT的结果，为M行N列的复数矩阵，u和v是频域中的频率索引对. 为了更清晰地观察和分析频域中不同频率成分的影响，将频域中的低频成分移至图像的边缘，将高频成分移至中心. 这种转换方式便于观察频域扰动对不同频率区域的影响，并有助于更精确地分析攻击的效果. 频域中心化可以通过以下表达式实现：

$$\boldsymbol{F}_c(u,v)=(-1)^{u+v}\boldsymbol{F}(u,v) \tag{2}$$

其中，$\boldsymbol{F}_c$是频域矩阵$\boldsymbol{F}$的中心化结果，$(-1)^{u+v}$是中心化因子.

2.2 频域单一分量修改

在频域中，通过修改特定的频率成分，可以显著改变图像的纹理、边缘或其他视觉特征，从而干扰图像分类器的判断. 与传统的空间域扰动相比，频域扰动能够更加精确地作用于关键频率成分，以较少的扰动实现更大的攻击效果. 为了实现对图像的攻击，将中心化频域矩阵中频率索引对为(u_0,v_0)的频率成分修改成复数$a+\mathrm{j}b$：

$$\hat{\boldsymbol{F}}_c(u,v)=\begin{cases}a+\mathrm{j}b, & \text{如果 } u=u_0,v=v_0\\ \boldsymbol{F}_c(u,v), & \text{其他}\end{cases} \tag{3}$$

其中，$\hat{\boldsymbol{F}}_c(u,v)$为修改频率成分后的中心化频域复数矩阵. 然后，对$\hat{\boldsymbol{F}}_c(u,v)$逆中心化，再进行二维离散逆傅里叶变换，将频域变换回空间域，从而获得对抗样本$\hat{\boldsymbol{f}}$. 同样，做傅里叶逆变换时，先按行队列变量v做一次长度为N的一维离散傅里叶变换，再将计算结果按列向对变量u做一次长度为M的傅里叶变换：

$$\hat{\boldsymbol{F}}(u,v)=(-1)^{u+v}\hat{\boldsymbol{F}}_c(u,v) \tag{4}$$

$$\hat{\boldsymbol{f}}(x,y)=\sum_{u=0}^{M-1}\left[\sum_{v=0}^{N-1}\hat{\boldsymbol{F}}(u,v)\exp\left(\frac{\mathrm{j}2\pi vy}{N}\right)\right]\exp\left(\frac{\mathrm{j}2\pi ux}{M}\right) \tag{5}$$

其中，$\hat{\boldsymbol{f}}(x,y)$代表图像在空间坐标$(x,y)$处的像素值. 经过这一系列转换，实现对抗样本初步生成.

2.3 基于超体积度量的扰动优化方法

在频域单一分量修改策略中，数据通过2D-DFT从空间域转化到频域. 然而，仅仅进行频域转换并不足以确保对抗样本同时具备高效的攻击性能和良好的隐蔽效果. 为此，本文采用多目标优化方法，在提升攻击效果的同时最大化对抗样本与原始图像的视觉相似度，从而生成既高效又难以察觉的对抗样本.

2.3.1 目标函数构建

在对抗样本生成的研究中，置信度和拉普拉斯金字塔结构相似度作为两个关键的性能指标，分别衡量了对抗样本的攻击效果和相似情况. 置信度反映模型对其预测结果的可信程度，而拉普拉斯金字塔结构相似度则衡量图像在不同尺度上的重建质量. 通过结合这两个目标函数，可以实现对对抗样本的综合评估，从而推动对抗样本的精准生成.

置信度作为概率值,反映了模型对预测结果的可信程度. 这一指标在分类任务中尤为重要[22],高置信度意味着模型对其预测结果较为确信. 为了使得生成的对抗样本$\hat{\boldsymbol{f}}$有较好的攻击性,需要最小化对抗样本$\hat{\boldsymbol{f}}$的置信度:

$$\text{minmize}\quad \text{conf}\left(\hat{\boldsymbol{f}}\right) \tag{6}$$

其中,conf(·)表示从识别网络获取置信度. 将对抗样本输入预训练的识别网络,经过前向传播计算后,从网络的输出层获取模型对正确分类类别的预测置信度值.

拉普拉斯金字塔是一种多分辨率图像表示方法,具有良好的多分辨率特性,通过梯次向下采样获得一系列分辨率逐步降低的图像集合. 基于结构相似性的图像质量评价(Structural Similarity Index Measure, SSIM)是衡量图像重建质量的重要指标,其考虑了图像的结构信息,比传统的峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)更符合人眼视觉特性. 拉普拉斯金字塔结构相似度计算的具体步骤如下.

首先,通过高斯模糊和下采样,为两幅SAR图像$\boldsymbol{f}_1$和$\boldsymbol{f}_2$构建高斯金字塔. 对于图像$\boldsymbol{f}_i$,高斯金字塔的第$n+1$层W_{n+1}为

$$W_{n+1}=\text{pyrDown}\left(\text{GaussBlur}\left(W_n\right)\right) \tag{7}$$

其中,pyrDown表示下采样操作,GaussBlur表示高斯模糊操作. 从高斯金字塔的每一层生成下一层后,将下一层上采样回到当前层的分辨率,然后计算当前层与上采样后的图像的差值,得到当前层的拉普拉斯层. 对于第n层拉普拉斯LPSS_n,计算公式为

$$\text{LPSS}_n=W_n-\text{pyrUp}\left(\text{pyrDown}\left(\text{GaussBlur}\left(W_n\right)\right)\right) \tag{8}$$

其中,pyrUp表示上采样操作. 对于拉普拉斯金字塔的每一层,可以使用结构相似性指数(SSIM)来衡量两幅图像的相似度. SSIM的计算公式为

$$\text{SSIM}\left(\text{LPSS}_{1i},\text{LPSS}_{2i}\right)=\frac{\left(2\mu_{L1i}\mu_{L2i}+c_{1i}\right)\left(2\sigma_{L1iL2i}+c_{2i}\right)}{\left(\mu_{L1i}^2+\mu_{L2i}^2+c_{1i}\right)\left(\sigma_{L1i}^2+\sigma_{L2i}^2+c_{2i}\right)} \tag{9}$$

其中,LPSS_{1i}和LPSS_{2i}是$\boldsymbol{f}_1$和$\boldsymbol{f}_2$第i层拉普拉斯,μ_{L1i}和μ_{L2i}分别为其均值,σ_{L1i}和σ_{L2i}表示其方差,σ_{L1iL2i}是其协方差,c_{1i}和c_{2i}是常数,用于避免分母为0的情况. 对于每一层拉普拉斯金字塔,分别计算SSIM值,然后取所有层的平均值作为两张图像最终的相似度度量$\overline{\text{LPSS}}\left(\boldsymbol{f}_1,\boldsymbol{f}_2\right)$,如式(10)所示:

$$\overline{\text{LPSS}}\left(\boldsymbol{f}_1,\boldsymbol{f}_2\right)=\sum_{i=1}^{n}\text{SSIM}\left(\text{LPSS}_{1i},\text{LPSS}_{2i}\right)/n \tag{10}$$

为确保对抗攻击不易被察觉,在生成对抗样本过程中需严格控制其与原始图像的感知相似性,通过优化视觉差异指标实现有效伪装,故需要最大化原始图像$\boldsymbol{f}$和对抗样本$\hat{\boldsymbol{f}}$的拉普拉斯金字塔结构相似度$\overline{\text{LPSS}}\left(\boldsymbol{f}_1,\boldsymbol{f}_2\right)$,如式(11)所示:

$$\text{minimize}\quad -\overline{\text{LPSS}}\left(\boldsymbol{f}_1,\boldsymbol{f}_2\right) \tag{11}$$

2.3.2　基于超体积度量的多目标优化方法

在对抗样本生成的过程中,不仅需要生成能够成功攻击模型的样本,还要确保这些样本在视觉上足够隐蔽,难以被人眼察觉且避免被其他算法检测到. 通过多目标优化方法,可以在攻击性能与视觉相似度之间找到最佳平衡,确保生成的对抗样本具备高效的攻击效果,提高实际应用的可行性. 因此,本文构建最小化$\text{conf}\left(\hat{\boldsymbol{f}}\right)$和最大化$\overline{\text{LPSS}}\left(\boldsymbol{f}_1,\boldsymbol{f}_2\right)$作为目标函数,在频域矩阵范围内将$\boldsymbol{F}(u,v)$处的频域分量修改成$a+\text{j}b$,具体描述为

$$\begin{aligned}&\text{minimize}\quad \text{conf}\left(\hat{\boldsymbol{f}}\right)\\&\text{minimize}\quad -\overline{\text{LPSS}}\left(\boldsymbol{f}_1,\boldsymbol{f}_2\right)\\&\text{s.t.}\begin{cases}u\in[0,t]\\v\in[0,t]\\a\in\left[\min\boldsymbol{R}(u,v),\max\boldsymbol{R}(u,v)\right]\\b\in\left[\min\boldsymbol{I}(u,v),\max\boldsymbol{I}(u,v)\right]\end{cases}\end{aligned} \tag{12}$$

其中,t是图像矩阵行数和列数的尺寸大小,$\boldsymbol{R}(u,v)$是$\boldsymbol{F}(u,v)$的实部,$\boldsymbol{I}(u,v)$是$\boldsymbol{F}(u,v)$的虚部,$\min\boldsymbol{R}(u,v)$代表矩阵$\boldsymbol{F}(u,v)$的实部最大值,同理可得$\max\boldsymbol{R}(u,v)$、$\min\boldsymbol{I}(u,v)$、$\max\boldsymbol{I}(u,v)$,通过设定参数范围,能够有效限制扰动强度.

由于计算复杂度较高,且传统方法难以有效平衡多个优化目标,本文采用优化算法来获得最优的对抗样本生成策略. 针对FMAA框架,设计一种基于超体积度量的多目标对抗样本生成框架. 该方法简化了多目标优化过程,避免了复杂的非支配排序,从而降低了计算复杂度,并提高了对抗样本的生成效率. 基于超体积度量的多目标优化特别适合处理双目标优化任务,能够更好地平衡目标之间的权重,避免局部最优解. 具体算法流程如算法1所示.

在多目标优化问题中,Pareto前沿表示所有非支配解的集合. 每个解在至少一个目标上优于其他解,但不存在一个解在所有目标上均最优. 本文以最小化正确分类置信度和最大化拉普拉斯金字塔结构相似度为双目标,通过超体积度量驱动的多目标优化算法生成Pareto前沿. 该前沿包含一系列权衡解,某些解以增加被察觉的代价显著提升攻击性,而另一些解则优先保持高隐蔽性,但攻击效果较弱. 在实际应用中,可根据具体需求从Pareto前沿中选择不同策略的最优解.

算法1 本文对抗样本生成算法

输入：初始种群和每代种群数量num、最大迭代数$t_{\max}$、目标函数$G(k)=\left\{\text{conf}\left(\hat{\boldsymbol{f}}(\boldsymbol{f},k)\right),-\overline{\text{LPSS}}\left(\boldsymbol{f},\hat{\boldsymbol{f}}(\boldsymbol{f},k)\right)\right\}$.其中$\hat{\boldsymbol{f}}(\boldsymbol{f},k)$表示基于原始SAR图像$\boldsymbol{f}$和参数集$k$所获得的对抗样本,$k=\{u,v,a,b\}$

输出：基于频域图像的最优解集P^*

1. 步骤1：初始化种群
2. 随机生成大小为num的种群$P_t=\{k_1,k_2,\cdots,k_N\}$.其中,每个个体$k_i$是一个决策变量解,决定在$(u,v)$处的频域分量修改成$a+\text{j}b$
3. 从第1代到第$t_{\max}$代,重复以下操作:
4. 步骤2：适应度评估
5. 计算种群P_t中每个个体k_i的目标函数值$\text{conf}\left(\hat{\boldsymbol{f}}(\boldsymbol{f},k)\right)$和$\overline{\text{LPSS}}\left(\boldsymbol{f},\hat{\boldsymbol{f}}(\boldsymbol{f},k)\right)$
6. 步骤3：快速非支配排序
7. 将种群P_t划分为多个非支配前沿$G_1,G_2,\cdots,G_n$,并按照支配级别对个体进行排序
8. 步骤4：超体积贡献计算
9. 计算当前前沿中每个个体的超体积贡献$\Delta\text{HV}(k_i)=\text{HV}(S)-\text{HV}\left(S\backslash\{k_i\}\right)$
10. 步骤5：个体选择
11. 根据$\Delta\text{HV}(k_i)$排序,优先选择超体积贡献较大的个体
12. 步骤6：变异操作
13. 使用模拟二进制交叉$k'_i=k_i+\eta\cdot(k_j-k_i)$即多项式变异$k'=k+\delta\cdot(k_{\max}-k_{\min})$,生成子代种群$Q_t$
14. 步骤7：种群更新
15. 合并父代种群和子代种群$R_t=P_t\cup Q_t$,对R_t执行非支配排序,保留前num个体作为新种群P_{t+1}
16. 步骤8：循环结束,输出Pareto最优解

基于超体积贡献值的动态排序,通过移除对当前Pareto前沿超空间体积增长贡献度最低的个体,确保种群始终维持最优收敛态势.这一策略使其更注重解集对目标空间的整体覆盖,通过优先保留对整体目标超体积贡献大的个体,提升解集分布的全局性,避免可能由于快速非支配排序和拥挤度计算产生的局部最优解或解集中某些区域过度聚集的问题.本研究所采用的超体积度量的多目标优化框架,通过重构传统多目标进化算法的选择机制,在计算效率层面实现突破性改进,计算复杂度从NSGA-II的$O(mN^2)$降低到$O(N\log N)$,这种显著的复杂度改进在处理本文建立的优化问题时尤为明显,能够显著降低优化时间开销,提高方法的可扩展性[23].更重要的是,采用超体积度量的多目标优化在解的探索与开发之间实现了更好的平衡,使得解集不仅能够快速逼近Pareto最优前沿,还保持了解集中个体的多样性和代表性,从而在复杂多目标优化任务中展现出更优异的表现.通过引入基于超体积度量的多目标优化,本文实现了基于频域分量修改的对抗样本的进一步优化,对扰动达到精细化的寻优控制.

3 仿真实验

本文实验在PyTorch深度学习框架上进行,并在单个NVIDIA RTX 3060 GPU的环境下实施.为了验证所提模型在生成对抗样本方面的可行性,选择由美国国防高级研究计划局(DARPA)公开发布的移动和固定目标获取及识别(MSTAR)数据集.这个数据集包含了10种不同类型的军事目标,包括2S1(自行榴弹炮)、BRDM2(装甲侦察车)、BTR60(装甲人员运输车)、D7(推土机)、T62(坦克)、ZIL131(货车)、ZSU234(自行高射炮)、BTR70(装甲人员运输车)、BMP2(步兵战车)和T72(坦克)[24].在实验中,使用了俯仰角17°的全类别样本(共2 340张)和俯仰角15°的全类别样本(共585张)混合示例,其中80%的混合示例用作训练集,20%的混合示例用作测试集.训练前对所有图像进行归一化和中心裁剪.

在全面评估本文所提出方法的有效性与通用性时,本文使用MSTAR数据集,针对VGG16[25]、AConvNet[26]和YOLOv5[27]这类具有代表性的目标识别网络进行了系统的训练与验证.这3种网络涵盖了不同类型的网络架构,能够全面验证本文方法的有效性和通用性.经过严谨的训练过程与参数调优,所挑选模型训练准确度、测试准确度均在99.8%、98.9%以上,表明模型已具备显著的目标识别性能,可进行后续对抗样本生成方法的测试实验.

3.1 空间域图像变化过程

随机选取待攻击的原始图像$\boldsymbol{f}$,通过2D-DFT,获得频域图像$\boldsymbol{F}$、中心化频域图像$\boldsymbol{F}_c$.由于$\boldsymbol{F}$、$\boldsymbol{F}_c$均为复数图像,为了更好地可视化观察本文方法的图像变化过程,对$\boldsymbol{F}$、$\boldsymbol{F}_c$复数图像进行幅度谱计算,为增强可视化效果,同时避免log(0)的情况,对对应图像幅度谱进行以下处理:

$$\hat{\boldsymbol{H}}=\log(1+\boldsymbol{H}) \tag{13}$$

其中,$\boldsymbol{H}$为某图像的幅度谱,$\hat{\boldsymbol{H}}$为处理后的幅度谱.

本文方法图像从空间域转换到频域的变化过程如图2所示.左侧为原始图像,中间为原始图像经过2D-DFT得到的频域图像的幅度谱,幅度图展示了图像频率成分的强度分布,将图像的低频信息集中在图像的边缘区域,而高频信息则集中在图像的中心.低频部分通常代表了图像的整体结构和主要特征,反映了图像的宏观信息,如大范围的物体形状或背景色块等;高频部分则包含了细节特征,如图像的边缘、纹理和噪声等微观信息.右侧为中心化频域图像的幅度谱,可以观察到,通过中心化操作,将频域图像的边缘区域对应高频分量,包含SAR图像的整体结构信息,而中心区域为低

频分量,携带细节特征,提升可视化的直观性.

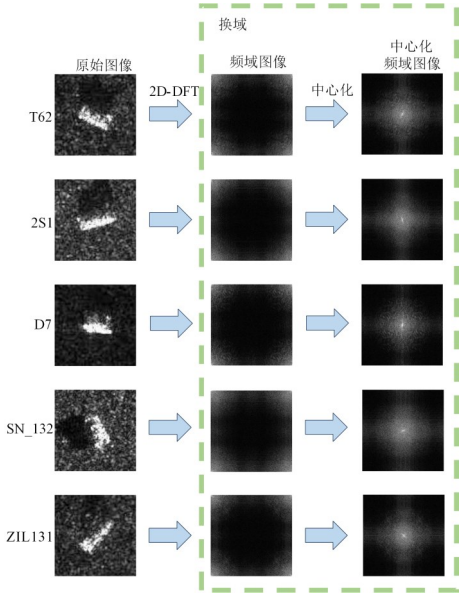


图2　原始图像转换成中心化频域图像示例图

3.2　扰动添加过程

在完成图像从空间域到频域的转换并中心化后,可对中心化频域图像$\boldsymbol{F}_c$进行随机频域分量修改,通过合理地调整幅度和相位信息,能够实现对图像特征的微小扰动,以生成初步的对抗样本. 本节对3.1节中作为示例的5张图像,分别对其对应的中心化频域图像开展随机频域分量修改工作,同时将所有频域图像进行式(13)的可视化处理.

如图3第一列所示,对于中心化频域图像,在$u=63$,$v=11$时将A处频域分量修改为$-108+\mathrm{j}8.8$,得到中心化频域对抗样本前体. 随后,进行逆中心化处理,获得频域对抗样本前体,最后通过逆2D-DFT将其变换回空间域,从而得到最终的对抗样本. 由于对频域单一分量进行修改,使得对抗样本相较原图像产生如第一行最后一列所示的扰动效果. 通过在频域中修改频域分量,大幅减少了所需优化参数的数量,使得在保持算法效率的前提下实现了大范围扰动注入. 其他4张修改示例的修改效果与位置A相类似,均能有效利用频域的特性,使得扰动的设计具有较高的灵活性.

3.3　基于超体积度量的多目标优化迭代过程

为了增强对抗样本的实际攻击效果,并在视觉特性上维持与原始图像的高相似度,需要对修改频域分量时的参数进行优化. 本节实验以224×224的原始图像为输入,针对SAR图像对抗样本生成问题进行多目标优化. 在优化过程中,通过不断演化Pareto前沿,逐步探索解空间中更优的解集,使得对抗扰动的设计在隐蔽性和攻击性之间达到最佳平衡.

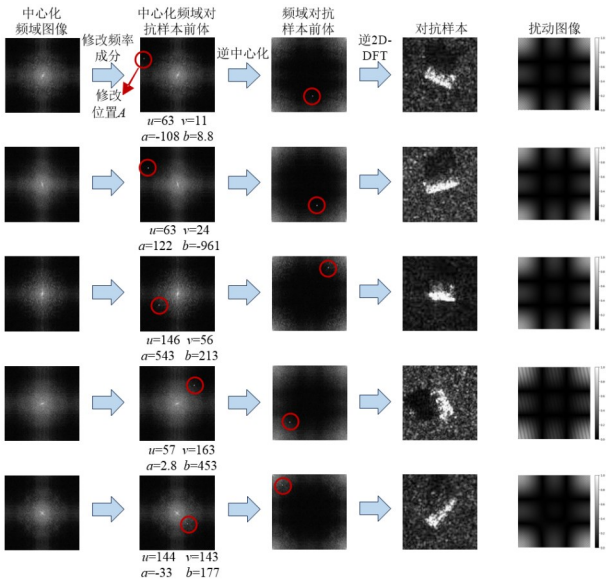


图3　扰动添加过程示例图

3.3.1　FMAA不同策略选择

本节选择VGG16作为示例识别网络对优化生成的Pareto前沿进行分析. 为全面评估不同策略的效果,本文从Pareto前沿中选取了3个具有代表性的解点进行仿真实验,包括2个极端点和1个拐点.

仿真实验如图4所示,VGG16将原始图像以100%置信度分类为2S1,策略1选择了以隐蔽性为优先目标的对抗攻击,生成的对抗样本与原始图像几乎完全一致,LPSS接近1,在视觉效果几乎无差异的情况下,将VGG16正确分类的置信度有效降低至8.39%,同时VGG16将其以41.15%的置信度误分类为ZSU_23_4. 然而,这种牺牲攻击性能的对抗样本在实际应用中可能存在失误,从而增加了实际应用的风险.

策略2则在隐蔽性和攻击性之间进行了折中选择,对抗样本在损失少量LPSS(99.95%)的情况下,成功将VGG16正确分类的置信度降低至0.69%,并使其以49.43%的置信度误分类为ZSU_23_4. 尽管这种折中策略在一定程度上平衡了攻击性能和隐蔽性,但在实际应用中可能面临如何在多样化场景下保持这种平衡的挑战.

策略3则针对攻击性能最大化,将正确分类置信度最大程度削弱至0.007%,成功实现置信度为67.25%的错误分类,将原始图像分类成ZSU_23_4,但在此过程中适度牺牲了一定的LPSS(99.86%),生成较高质量的对抗样本. 这种以牺牲隐蔽性为代价的攻击策略在实际应用中可能更容易被人眼或系统检测到,从而限制其在某些高安全要求场景中的应用.

实验结果表明,无论选择Pareto前沿上的哪种策略,本文方法均能够在保持对抗样本与原始SAR图像

较高相似度的前提下实现有效攻击. 这不仅验证了本文方法在频域扰动优化中的高效性和适应性,还表明其能够满足多样化的应用场景需求,为SAR图像的黑盒对抗性攻击提供了可靠的技术支撑.

3.3.2 FMAA对不同网络的对抗效果

在实验的第二部分,为验证本文方法在不同识别网络上的通用性和攻击有效性,本文随机选取了MSTAR数据集中4类目标图像作为实验对象,分别对VGG16、AConvNet、YOLOv5、YOLOv8、YOLOv11这5种主流网络进行了仿真实验,其中,YOLOv5、YOLOv8、YOLOv11的检测阈值设定为0,仅获取置信度最高的分类结果和对应置信度. 通过生成对抗样本并对其攻击效果进行分析,全面评估本文方法在不同网络架构和任务需求下的性能表现.

图5展示了本文方法生成的4组不同类别的对抗样本及其对应攻击效果,每组图像包含原始样本、Pareto前沿分布、对抗样本与所添加的扰动图像. 本文针对VGG16、AConvNet及YOLO系列5种网络架构的特性,对于Pareto前沿的最优解统一采用拐点解选择策略,确保生成的对抗样本既能高效攻击模型,又维持与原始图像的高相似度. 通过将拐点解生成的对抗样本输入目标网络,获取其输出类别及置信度数据,验证了所提方法在不同任务场景下的普适性与鲁棒性. 本文方法对不同网络的攻击效果如表1所示.

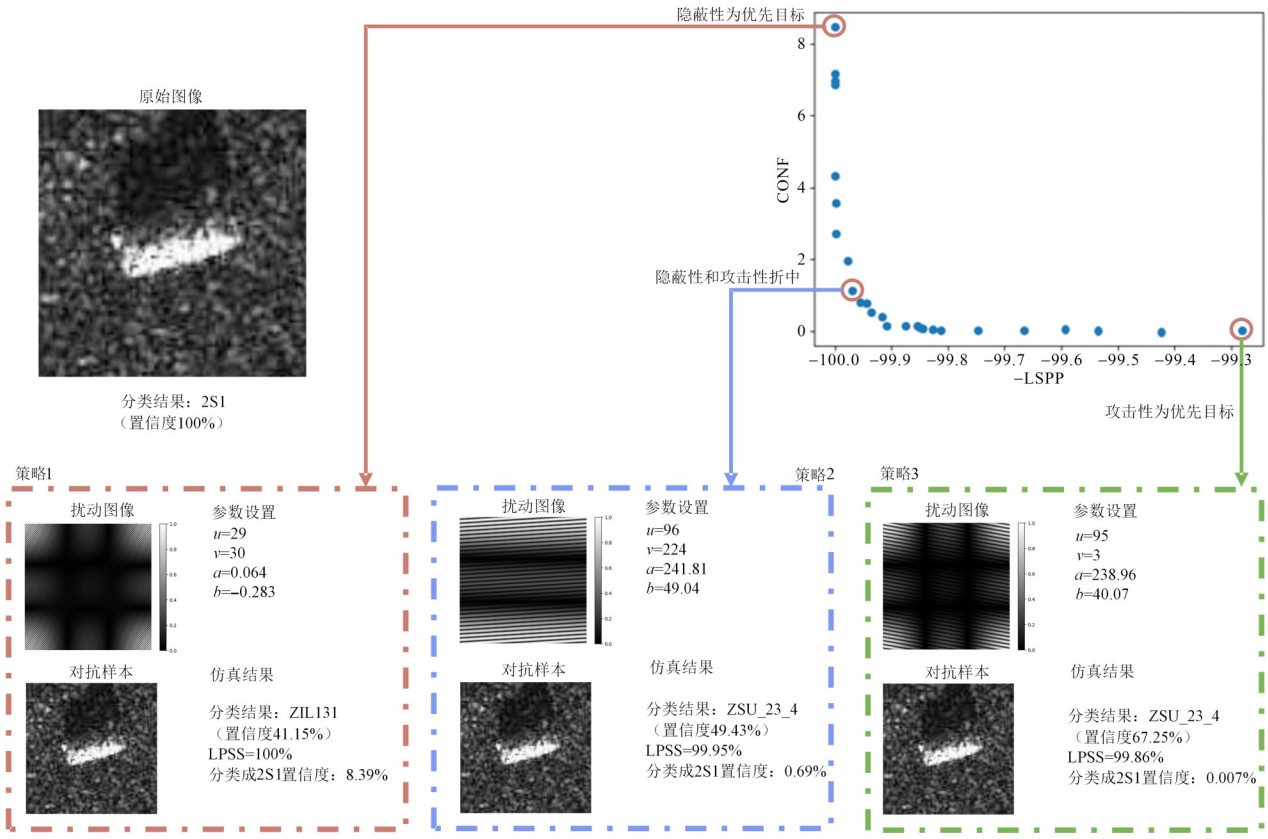
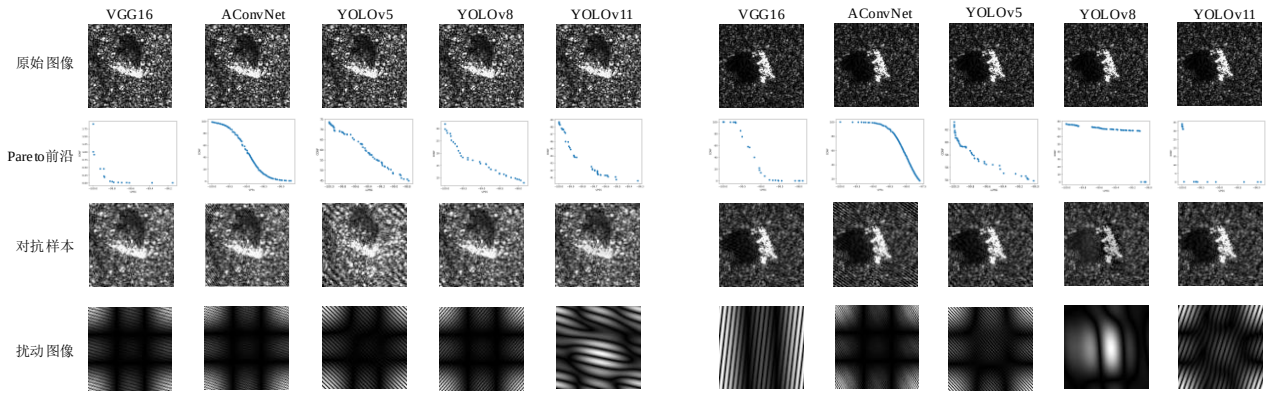


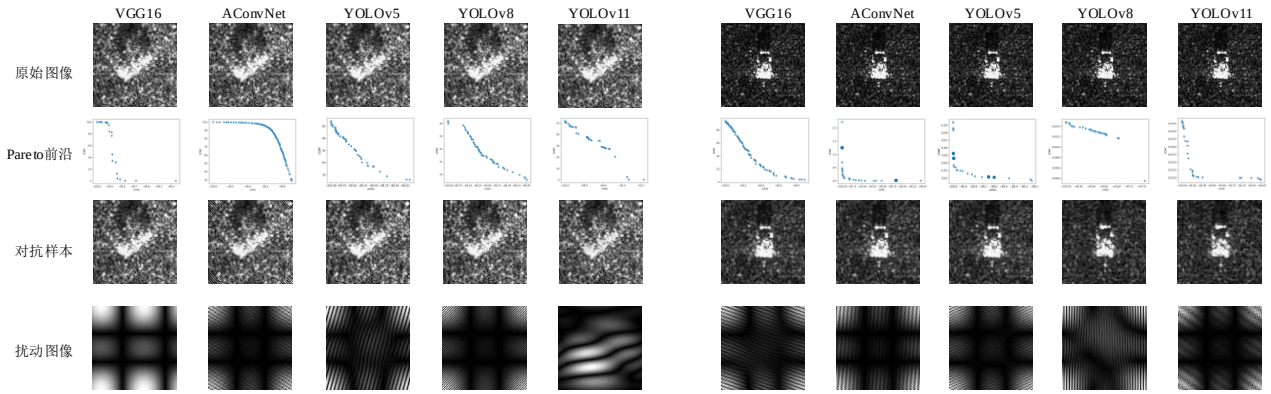
图4 示例攻击的Pareto前沿及3种策略选择、参数设置和仿真结果

实验结果表明,本文方法在不同模型架构的攻击性能存在差异. VGG16模型对于频域扰动表现出显著敏感性,对于BRDM_2、SN_132、ZIL131及T62这4类样本,本文生成的对抗样本均能成功误导VGG16,误分类置信度分别达到86.08%、88.27%、94.42%与90.39%,表明频域扰动可有效破坏其高置信度分类决策边界. 同时所生成的对抗样本LPSS均保持在99%以上,证明生成的对抗样本拥有较强的隐蔽性. AConvNet面对BRDM_2、SN_132、ZIL131及T62这4类样本同样均能被成功误导,但其误分类置信度为55.73%、68.96%、56.39%和71.43%,整体低于VGG16. 此外,本文方法针对AConvNet生成的对抗样本LPSS达到97.96%以上,比VGG16略低,说明需要较大的扰动幅度才能对AConvNet产生明显的影响. 实验表明频域扰动虽然依旧能够有效攻击AConvNet模型,但其优化扰动过程中面临更高的鲁棒性壁垒,相较于VGG16更难以突破.

相较于VGG16和AConvNet,YOLO系列网络对频域扰动效果展现出更强的鲁棒性,但修改频域分量添加扰动依旧可以显著削弱其判别能力. 在T62类别中,以YOLOv5、YOLOv8和YOLOv11为代表的YOLO系列



(*a*) BRDM_2对抗样本生成流程　　(*b*) SN_132对抗样本生成流程



(*c*) ZIL131对抗样本生成流程　　(*d*) T62对抗样本生成流程

图5　不同类别图像、识别网络通过本文方法的Pareto前沿、对抗样本及扰动图像

表1　本文方法对不同识别网络的攻击效果　　单位:%

		VGG16	AConvNet	YOLOv5	YOLOv8	YOLOv11
原始类别:BRDM_2	分类结果	ZIL131(86.08)	BTR_60(55.73)	2S1(47.84)	2S1(60.98)	BRDM_2(40.71)
	LPSS	99.63	98.78	98.90	98.88	99.31
原始类别:SN_132	分类结果	ZSU_23_4(88.27)	SN_9563(68.96)	SN_132(56.67)	SN_C71(57.52)	SN_9563(44.44)
	LPSS	99.18	98.03	99.63	99.05	99.46
原始类别:ZIL131	分类结果	BTR_60(94.42)	BTR_60(56.39)	ZIL131(20.17)	ZIL131(70.88)	BRDM_2(56.92)
	LPSS	99.86	97.96	99.26	98.55	99.47
原始类别:T62	分类结果	ZIL131(90.39)	ZSU_23_4(71.43)	2S1(57.75)	2S1(69.32)	2S1(44.28)
	LPSS	99.99	99.13	99.95	99.98	99.91

网络均被完全误导至2S1类别,置信度为57.75%、69.32%和44.28%,且LPSS值均高于99.91%,表明扰动在保持视觉隐蔽性的同时有效误导YOLO系列识别网络. 以BRDM_2类别为例,YOLOv5与YOLOv8分别以47.84%和60.98%的置信度误判为2S1,而YOLOv11虽保留正确分类能力,但置信度下降至40.71%. 在SN_132类别中,本文方法成功误导YOLOv8和YOLOv11,分别以57.52%的置信度误识别为SN_C71和以44.44%的置信度误识别为SN_9563,而未能成功误导YOLOv5,但正确分类置信度下降至56.67%. 对于ZIL131类别,YOLOv5和YOLOv8保留正确分类能力,但削减置信度至20.17%及70.88%,同时YOLOv11被成功误导,以56.92%的置信度误识别为BRDM_2. 上述实验结果表明,不同YOLO网络因架构对频域扰动的敏感性与鲁棒性存在差异,但均能有效降低YOLO网络架构对于正确类别的识别置信度. 此外,本文方法所生成的对抗样本与原始图像均保持大于98.5%的LPSS值,不易在视觉上被察觉.

综上,仿真实验结果表明,基于频域添加扰动和多目标优化的方法能够针对VGG16、AConvNet、YOLO系列等不同网络架构生成高效的对抗样本,无论是降低分类置信度还是误导模型决策,都表现出优异的通用性和鲁棒性. 同时,生成的对抗样本与原始图像并无明

显差别,生成的高质量对抗样本,在隐蔽性和攻击效果之间取得了优异的平衡.

3.3.3 消融实验

为验证本文所提频域多目标对抗攻击方法的有效性与优越性,本节设计了消融实验,对比现有不同的典型对抗攻击方法(包括白盒攻击FGSM、BIM和黑盒攻击Square Attack、simBA)在多种类别(BRDM_2、SN_132、ZIL 131和T62)SAR图像上的分类效果和LPSS. 其中,FGSM的扰动幅度设置为0.05;BIM的扰动幅度设置为0.03,迭代次数10;Square Attack的扰动幅度设置为0.5,迭代次数设置为10 000;simBA步长设置为0.01,迭代次数设置为10 000. FMAA框架均选取Pareto前沿的拐角点最优解,仿真数据记录于表2,加粗数据为本文方法结果.

表2 本文方法与经典算法效果对比 单位:%

攻击方法	指标	识别网络1	识别网络2
原始类别:BRDM_2			
		VGG16	AConvNet
FGSM	分类结果	2S1(100)	BRDM_2(99.79)
	LPSS	99.54	99.35
BIM	分类结果	ZIL131(100)	2S1(100)
	LPSS	99.63	99.45
FMAA	分类结果	**ZIL131(86.08)**	**BTR_60(55.73)**
	LPSS	**99.63**	**98.78**
Square Attack	分类结果	ZIL131(61.35)	BTR_60(59.84)
	LPSS	98.93	99.33
simBA	分类结果	BRDM_2(54.14)	BTR_60(52.62)
	LPSS	98.65	99.12
原始类别:SN_132			
		VGG16	AConvNet
FGSM	分类结果	SN_132(81.54)	SN_132(74.15)
	LPSS	99.42	99.41
BIM	分类结果	BRDM_2(99.99)	2S1(99.99)
	LPSS	99.55	99.63
FMAA	分类结果	**ZSU_23_4(88.27)**	**SN_9563(68.96)**
	LPSS	**99.18**	**98.03**
Square Attack	分类结果	SN_132(78.42)	SN_9563(64.15)
	LPSS	98.88	99.19
simBA	分类结果	SN_132(64.15)	SN_9563(53.64)
	LPSS	98.97	99.18
原始类别:ZIL131			
		VGG16	AConvNet
FGSM	分类结果	T62(99.97)	ZSU_23_4(99.98)
	LPSS	99.45	99.43
BIM	分类结果	D7(100)	BRDM_2(91.31)
	LPSS	99.65	99.95
FMAA	分类结果	**BTR_60(94.42)**	**BTR_60(56.39)**
	LPSS	**99.86**	**97.96**
Square Attack	分类结果	ZSU_23_4(73.65)	ZSU_23_4(53.15)
	LPSS	99.88	99.99
simBA	分类结果	ZSU_23_4(83.15)	ZIL131(64.15)
	LPSS	99.95	99.96
原始类别:T62			
		VGG16	AConvNet
FGSM	分类结果	ZIL131(100)	ZIL131(99.93)
	LPSS	98.75	98.92
BIM	分类结果	BRDM_2(100)	2S1(100)
	LPSS	99.55	99.41
FMAA	分类结果	**ZIL131(90.39)**	**ZSU_23_4(71.43)**
	LPSS	**99.99**	**99.13**
Square Attack	分类结果	T62(51.56)	ZSU_23_4(53.08)
	LPSS	99.49	99.71
simBA	分类结果	D7(55.04)	BTR_60(58.11)
	LPSS	99.11	99.86

以VGG16识别网络作为待攻击网络,本文FMAA框架所生成的对抗样本均能够在BRDM_2、SN_132、ZIL131和T62这4种类别的SAR图像上完成86.08%、88.27%、94.42%和90.39%的高置信度的误分类. 逼近FGSM、BIM等白盒攻击方法的高攻击性能,大幅高于Square Attack、simBA等黑盒攻击方法. 同时,本文方法所生成对抗样本和原始图像的LPSS与现有经典方法保持持平,这意味着本文方法能够在保持隐蔽性尽可能不变的同时完成攻击性能的提升.

针对AConvNet识别网络,其余方法针对4种类别图像所生成的对抗样本均存在未能成功误导分类的结果,例如FGSM针对原始类别为BRDM_2生成的对抗样本被AConvNet以99.79%的高置信度正确分类为BRDM_2,对于SN_132类别,FGSM所生成的对抗样本同样未能成功误导. 而本文方法所生成的对抗样本被AConvNet以55.73%的置信度错误分类为BTR_60,SN_132类别以68.96%的置信度成功误导为SN_9563. 虽然本文方法在AConvNet上的攻击效果不如VGG16,但依旧优于现有方法.

4 结论

本文提出的频域多目标对抗攻击(FMAA)框架,通过频域分量修改设计与多目标优化策略,有效解决了传统黑盒攻击在SAR图像中隐蔽性较差与攻击性不足之间平衡的问题. 基于频域转换扰动添加策略显著减少了扰动设计的变量. 利用基于超体积度量的多目标优化算法寻找频域分量修改参数解集,使得生成的对

抗样本能够在保证较强攻击能力的同时在视觉特性上维持与原始图像的高相似度. 实验验证了本文方法在多种主流网络架构上的通用性,及其生成的对抗样本与现有算法相比的高效生成能力,为SAR图像领域的安全性评估和对抗技术研究提供了新的思路. 未来研究可进一步探索该方法在其他复杂场景中的适用性与性能优化潜力.

参考文献

[1] XIA W J, LIU Z, LI Y. SAR-PeGA: A generation method of adversarial examples for SAR image target recognition network[J]. IEEE Transactions on Aerospace and Electronic Systems, 2022, 59(2): 1910-1920.

[2] 化青龙, 张云, 任航, 等. 基于最小熵准则与生成对抗网络的SAR三维转动舰船目标重聚焦方法[J]. 电子学报, 2024, 52(8): 2900-2912.
HUA Q L, ZHANG Y, REN H, et al. Refocusing for three-dimensional rotating ship targets in SAR images based on minimum entropy criteria and generative adversarial network[J]. Acta Electronica Sinica, 2024, 52(8): 2900-2912. (in Chinese)

[3] SHI J. SAR target recognition method of MSTAR data set based on multi-feature fusion[C]//2022 International Conference on Big Data, Information and Computer Network. Piscataway: IEEE, 2022: 626-632.

[4] WU Y, LIU J M, GONG M G, et al. Joint Semantic Segmentation using representations of LiDAR point clouds and camera images[J]. Information Fusion, 2024, 108: 102370.

[5] 化青龙, 魏晨曦, 张云, 等. 基于自监督复数域深度学习网络的SAR有源压制干扰抑制方法[J]. 电子学报, 2023, 51(4): 965-974.
HUA Q L, WEI C X, ZHANG Y, et al. Active jamming suppression for SAR images based on self-supervised complex-valued deep learning[J]. Acta Electronica Sinica, 2023, 51(4): 965-974. (in Chinese)

[6] XIE C H, ZHANG Z S, ZHOU Y Y, et al. Improving transferability of adversarial examples with input diversity[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2019: 2725-2734.

[7] MOOSAVI-DEZFOOLI S M, FAWZI A, FROSSARD P. DeepFool: A simple and accurate method to fool deep neural networks[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 2574-2582.

[8] CARLINI N, WAGNER D. Towards evaluating the robustness of neural networks[C]//2017 IEEE Symposium on Security and Privacy. Piscataway: IEEE, 2017: 39-57.

[9] LI H F, HUANG H K, CHEN L, et al. Adversarial examples for CNN-based SAR image classification: An experience study[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 14: 1333-1347.

[10] DU C, HUO C Y, ZHANG L, et al. Fast C&W: A fast adversarial attack algorithm to fool SAR target recognition with deep convolutional neural networks[J]. IEEE Geoscience and Remote Sensing Letters, 2021, 19: 4010005.

[11] ANDRIUSHCHENKO M, CROCE F, FLAMMARION N, et al. Square Attack: A Query-Efficient Black-Box Adversarial Attack via Random Search[M]//Computer Vision-ECCV 2020. Cham: Springer International Publishing, 2020: 484-501.

[12] GUO C, GARDNER J, YOU Y Y, et al. Simple black-box adversarial attacks[C]//International Conference on Machine Learning. Long Beach: PMLR, 2019: 2484-2493.

[13] DU C, ZHANG L. Adversarial attack for SAR target recognition based on UNet-generative adversarial network[J]. Remote Sensing, 2021, 13(21): 4358.

[14] CUI J H, DUAN J L, LUO B Y, et al. SAR-AE-SFP: SAR imagery adversarial example in real physics domain with target scattering feature parameters[EB/OL]. (2024-03-02)[2025-05-29]. https://doi.org/10.48550/arXiv.2403.01210.

[15] ZHANG L, JIANG T P, GAO S Y, et al. Generating adversarial examples on sar images by optimizing flow field directly in frequency domain[C]//IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium. Piscataway: IEEE, 2022: 2979-2982.

[16] HUANG X C, LU Z Z, PENG B. Frequency and spatial adversarial attack to SAR target recognition[C]//2024 7th International Conference on Electronics Technology. Piscataway: IEEE, 2024: 906-910.

[17] TESSARI G, FLORIS M, PASQUALI P. Phase and amplitude analyses of SAR data for landslide detection and monitoring in non-urban areas located in the North-Eastern Italian pre-Alps[J]. Environmental Earth Sciences, 2017, 76(2): 85.

[18] BEUME N, NAUJOKS B, EMMERICH M. SMS-EMOA: Multiobjective selection based on dominated hypervolume[J]. European Journal of Operational Research, 2007, 181(3): 1653-1669.

[19] LIN T W, MA Z Q, LI F, et al. Drafting and revision: Laplacian pyramid network for fast high-quality artistic style transfer[C]//2021 IEEE/CVF Conference on Com-

puter Vision and Pattern Recognition. Piscataway: IEEE, 2021: 5137-5146.

[20] TAO M L, SU J, HUANG Y, et al. Mitigation of radio frequency interference in synthetic aperture radar data: Current status and future trends[J]. Remote Sensing, 2019, 11(20): 2438.

[21] JIANG L M, DAI B, WU W, et al. Focal frequency loss for image reconstruction and synthesis[C]//2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2021: 13899-13909.

[22] 陈思伟, 周鹏. SAR图像对抗攻击的进展与展望[J]. 信息对抗技术, 2023, 2(Z1): 171-188.
CHEN S W, ZHOU P. Progress and prospect of SAR image adversarial attack[J]. Information Countermeasure Technology, 2023, 2(Z1): 171-188. (in Chinese)

[23] BEUME N, FONSECA C M, LOPEZ-IBANEZ M, et al. On the complexity of computing the hypervolume indicator[J]. IEEE Transactions on Evolutionary Computation, 2009, 13(5): 1075-1082.

[24] 章坚武, 能豪, 李杰, 等. 基于掩模提取的SAR图像对抗样本生成方法[J]. 电信科学, 2024, 40(3): 64-74.
ZHANG J W, NAI H, LI J, et al. Adversarial example generation method for SAR images based on mask extraction[J]. Telecommunications Science, 2024, 40(3): 64-74. (in Chinese)

[25] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2015-04-10)[2025-05-29]. https://doi.org/10.48550/arXiv.1409.1556.

[26] LIU Z, MAO H Z, WU C Y, et al. A ConvNet for the 2020s[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 11966-11976.

[27] ZHANG J, HUO Y B, YANG J L, et al. Automatic counting of retinal ganglion cells in the entire mouse retina based on improved YOLOv5[J]. Zoological Research, 2022, 43(5): 738-749.

作者简介

刘洁怡 女,1991年2月出生于陕西省西安市. 现为西安电子科技大学电子工程学院副教授. 主要研究方向为智能信号处理、多站雷达系统干扰对抗技术.
E-mail: jieyiliu@xidian.edu.cn

李 豪 男,1990年12月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副教授. 主要研究方向为计算智能、深度学习. 中国电子学会会员编号:E190022928M.
E-mail: haoli@xidian.edu.cn

李明哲 男,2004年5月出生于黑龙江省齐齐哈尔市. 现为西安电子科技大学电子工程学院本科生. 主要研究方向为深度学习、对抗样本生成.
E-mail: mingzheli@stu.xidian.edu.cn

周 宇 男,1983年8月出生于湖南省常德市. 现为西安电子科技大学电子工程学院教授. 主要研究方向为智能系统(集群)的自主感知和决策以及协同控制和优化.
E-mail: zhouyu@xidian.edu.cn

杨曜铭 男,2001年4月出生于海南省海口市. 现为西安电子科技大学硕士研究生. 主要研究方向为深度学习、对抗攻击.
Email: yangyaoming@stu.xidian.edu.cn

党可林 男,2000年8月出生于陕西省渭南市. 现为西安电子科技大学博士研究生. 主要研究方向为深度神经网络、对抗性攻击和遥感图像处理.
E-mail: 24021110902@stu.xidian.edu.cn