文章编号：2095-4980(2022)08-0836-07

调制识别中目标对抗攻击

赵浩钧[a,b]，林　云[*a,b]，包志达[a,b]，史继博[a,b]，葛　斌[c]

(哈尔滨工程大学 a.信息与通信工程学院；b.先进船舶通信与信息技术工业和信息化部重点实验室；
c.数学科学学院，哈尔滨 黑龙江 150001)

摘　要：由于深度学习算法具有特征表达能力强、特征自动提取以及端到端学习等突出优势，因此被越来越多的研究者应用至通信信号识别领域。然而，对抗样本的发现使得深度学习模型极大程度地暴露在潜在的风险因素中，并对当前的调制识别任务造成严重的影响。本文从攻击者的角度出发，通过对当前传输的通信信号添加对抗样本，以验证和评估目标对抗样本对调制识别模型的攻击性能。实验表明，当前的目标攻击可以有效地降低模型识别的精确度，所提出的logit指标可以更细粒度地用于衡量攻击的目标性效果。

关键词：卷积神经网络；调制识别；对抗样本；无线电安全

中图分类号：TN911　　**文献标志码：**A　　**doi：**10.11805/TKYDA2020692

Targeted adversarial attack in modulation recognition

ZHAO Haojun[a,b]，LIN Yun[*a,b]，BAO Zhida[a,b]，SHI Jibo[a,b]，GE Bin[c]

(a.College of Information and Communication Engineering；b.Key Laboratory of Advanced Marine Communication and Information Technology，
c.College of Mathematical Sciences，Harbin Engineering University，Harbin Heilongjiang 150001，China)

Abstract：Since deep learning algorithms have outstanding advantages such as strong feature expression ability, automatic feature extraction, and end-to-end learning, more and more researchers have applied them to the field of communication signal recognition. However, the discovery of adversarial examples exposes deep learning models to potential risk factors to a great extent, which has a serious impact on current modulation recognition tasks. From the perspective of an attacker, adversarial examples are added to the currently transmitted communication signal to verify and evaluate the attack performance of the target countermeasure sample to the modulation recognition model. Experimental results show that the current targeted attack can effectively reduce the accuracy of model recognition, and the constructed logit indicator can be better applied to measure the targeted effect more fine-grained.

Keywords：Convolution Neural Network(CNN)；modulation recognition；adversarial examples；wireless security

随着各种通信系统的共存，无线电数据现在呈现出比以前更复杂和多样的特征，例如随机性、异构性和庞大性。由于其许多优点，深度学习越来越多地应用于通信领域，例如调制识别。为了满足用户之间的不同需求并充分利用信道容量，通信信号采用了不同的调制方式，基于调制识别，可以收集并分类信号特征如信号频谱、瞬时幅度和瞬时相位，以完成相关的推理和学习。同时，调制识别在缓解频谱资源短缺方面起着关键作用，并且是民用和军事应用中的重要技术手段[1-2]。在军事领域，调制识别为获取雷达电子战中的敌方情报以及选择最佳干扰和抑制方法提供了重要的技术手段。在民用领域中，它在日常无线电台的检测、无线电频谱资源的管理以及空中交通管制等信号的识别中起着重要作用。林云等[3]提出了一种新的数据转换算法，该算法基于深度学习模型来解决无线通信问题，进一步提升了通信信号调制的分类精确度。张智博等[4]探讨了深度学习在通信干扰模式中的识别能力，构建了一种基于功率谱谱图和双隐藏层神经网络的通信干扰模式识别方法。孙浩然[5]对深度神

收稿日期：2020-12-10；**修回日期：**2021-01-26

基金项目：国家自然科学基金面上资助项目(61771154)；中央高校基本科研业务费资助项目(3072020CF0813)

*通信作者：林　云　email:linyun@hrbeu.edu.cn

经网络进行了训练以用于无线通信中的信息管理。查雄等[6]提出了一种基于多端卷积神经网络的通信信号调制识别算法。涂涯等[7]使用基于生成对抗网络的半监督学习进行数字信号分类。吴灏等[8]基于卷积神经网络和稀疏滤波完成了调制识别。

尽管深度学习在解决无线电通信问题方面具有独特的优势，但深度神经网络(Deep Neural Networks，DNN)无法解释的黑盒属性[9]可能会带来许多安全风险。最近的研究发现，通过给当前数据集的输入样本添加微小扰动以形成对抗样本时，分类器模型会输出错误的类别，这些样本中的扰动是难以被检测到的，即使分类性能再优异的DNN模型也很容易被愚弄。对抗样本的存在给调制识别任务带来严峻的挑战。当前，Zhao等[10]将对抗样本添加至调制识别任务中，较为全面地验证了对抗攻击在信号领域中的有效性和可行性，然而，当前研究的攻击均为通用场景下的非目标攻击，即在该类攻击下，攻击者无法控制被扰动样本最终的输出类别标签是什么，唯一目标在于欺骗模型对输入样本进行错误的分类。考虑到越来越多的场景出于军事或者民用目的往往需要将模型输出诱导为期望的输出，即发起目标诱导攻击，然而当前学者们在目标诱导性场景中的研究进度缓慢，存在大量的短板和空缺，因此本文将DNN应用于调制信号分类任务，研究和讨论目标对抗攻击在信号域里引起的安全性问题，以丰富通信信号对抗样本的内涵，增进对对抗样本的认知，并为深度学习在通信任务中的实际部署和应用扫清障碍。进一步，本文提出了logit组合评价指标，更加细粒度地衡量对抗样本的目标攻击性能。

1　对抗样本算法介绍

对抗样本根据攻击的程度分为单步攻击和迭代攻击。单步攻击的对抗样本一次性地计算出梯度并进行生成，而迭代攻击的对抗样本则需要多步来完成更新。与单步攻击相比，迭代攻击可以产生更强的攻击效果，但是它需要不断地获取和更新被攻击模型的知识，从而花费更多的时间。为了确保不同攻击模型之间的可比性，本文选取了3类方法，来展开在调制识别任务中目标对抗攻击的研究，评估信号域的安全性问题。其中包括了两种迭代生成方法，投影梯度下降法(Projected Gradient Descent，PGD)[11]和基本迭代法(Basic Iterative Method，BIM)[12]，以及单步攻击方式的快速梯度标记法(Fast Gradient Sign Method，FGSM)[13]。

1.1　FGSM

FGSM由Goodfellow首次提出，可以快速生成对抗样本。以信号样本为例，希望对原始信号样本进行人眼无法识别的修改，但这会导致DNN模型的错误分类。假设原始信号样本为x，标识的类别结果为y，并且扰动为η，其中扰动要求足够小，即$\|\eta\|_\infty<\varepsilon$，其中$\varepsilon$为扰动的最大量级。记输入样本$x$的对抗样本为：

$$\tilde{x}=x+\eta \tag{1}$$

考虑权重向量$\boldsymbol{\omega}$和对抗样本$\tilde{x}$的内积：

$$\boldsymbol{\omega}^{\mathrm{T}}\tilde{x}=\boldsymbol{\omega}^{\mathrm{T}}x+\boldsymbol{\omega}^{\mathrm{T}}\eta \tag{2}$$

显然对抗扰动通过$\boldsymbol{\omega}^{\mathrm{T}}\eta$使得神经元的输出增大。如果权向量的维数为$n$并且权向量的均值为$m$，则$\boldsymbol{\omega}^{\mathrm{T}}\eta$的最大值为$\varepsilon mn$，此时$\eta=\varepsilon\mathrm{sign}(\omega)$。在高维空间中，即使是很小的扰动，也会对最终的输出值产生很大的影响。

1.2　PGD

PGD既是一种能有效抵御一阶对抗攻击的防御方法，从优化的角度研究神经网络的对抗鲁棒性，为之前对抗训练防御方法提供了统一的视角，又包含了生成对抗样本的方法，因此该方法也可以用来攻击目标模型。本文采用了PGD来展开对模型的对抗攻击。该方法的核心公式也称之为鞍点公式，具体如下：

$$\begin{aligned}&\min{}_{\theta}\rho(\theta)\\&\rho(\theta)=E_{(x,y)D}[\max{}_{\delta\in S}L(\theta,x+\delta,y)]\end{aligned} \tag{3}$$

式中：$E_{(x,y)D}[L]$是定义的总体风险；D是样本分布；S是允许的扰动。该公式由内部的最大化和外部的最小化两个问题构成。内部的最大化问题旨在找到实现最大损失的数据的扰动，这实际上是攻击问题，满足最大化条件的样本有很大概率是对抗样本。外部的最小化问题旨在找到模型的参数使得攻击的对抗损失最小，这实际上是训练鲁棒分类器的问题。其次，鞍点问题给出了理想的鲁棒模型所需要达到的确切目标，也就是衡量鲁棒性的标准。

1.3 BIM

BIM 的方法由 Kurakin 和 Goodfellow 等人提出，由于 FGSM 不需要迭代过程来计算对抗样本，因此相较于其他的方法快得多。BIM 是拓展的 FGSM，在 BIM 中对抗样本是在多次迭代中产生的，每次迭代的步长较小，并在每一步后截取中间的结果值，以确保它们位于原始输入的扰动 ε 邻域：

$$\begin{aligned} x_0 &= x \\ x_{n+1} &= \mathrm{Clip}_{x,\varepsilon}\left\{x_n + \varepsilon \mathrm{sign}\left(\nabla_x J(x_n, y)\right)\right\} \end{aligned} \tag{4}$$

式中 $\mathrm{Clip}_{x,\varepsilon}\{z\}$ 表示将 z 裁剪到 $[x-\varepsilon, x+\varepsilon]$ 的范围。该方法中，既可以通过设置总的扰动 ε 展开攻击，类似于 FGSM 的方法，也可以通过设置单步的迭代展开攻击。

2 目标攻击生成原理

在对抗攻击中，按照攻击是否有目标可分为目标攻击和非目标攻击。非目标攻击不会指定神经网络输出某个固定的类别，除了原始类别之外，预测的对抗样本类别可以是任意的，即信号攻击者可以产生导致任何错误分类的扰动。目标攻击通常发生在多类分类问题中，诱导深度神经网络输出到指定的类别。例如在调制信号中，攻击者欺骗分类器，将 QPSK、QAM64 等信号都无差别地输出为 BPSK，目标攻击通常会使目标对抗类的概率最大化。

令信号感知模型的输入信号 x，并输出最可能的 y 类。在通信信道中，x 表示复杂样本的单个通道。通过寻求最大化相同的损失函数，从而降低目标感知系统的准确性。此时，该优化问题可定义为：

$$\arg\max_{x'} J\left(f(\theta, x'), y\right) \tag{5}$$

分类器的参数固定，但可以通过操作输入 x，利用之前所述的对抗样本生成方法，产生通信信号对抗样本 x'，深度学习模型将输出错误类别，从而欺骗目标感知模型，实现非目标攻击。

目标攻击与非目标攻击使用了相同的损失函数，但也有几个重要区别：

$$\arg\min_{x'} J\left(f(\theta, x'), y_t\right) \tag{6}$$

式中原标签 y 替换为了己方需要指定目标错误分类的类别 y_t。为了提升目标类的可信度，需要最小化目标的损失函数，以使得目标分类器尽可能减少当前分类与目标类别之间的差距[14]。

将目标诱导与 FGSM 算法结合，可以得到：

$$x' = x + \varepsilon \cdot \mathrm{sign}\left(\nabla_x J_{F,t}(x)\right) \tag{7}$$

式中：F 可以看作目标函数，t 是目标类别。FGSM 每进行一次迭代时，梯度都会从原始的输入通信信号 x 开始更新，然后通过优化生成伪装信号 x'。进一步，在 BIM 中，由于它是对 FGSM 进行了迭代更新，因此通过每次迭代时进行小的修改而不是一次性更新，最终完成目标对抗样本的创建。

$$x'_i = x'_{i-1} - \mathrm{clip}\left\{\alpha \cdot \mathrm{sign}\left(\nabla_x J_{F,t}(x'_{i-1})\right)\right\} \tag{8}$$

式中 clip 是将每次迭代生成的值按照需要裁剪在一定范围内，每次的扰动程度为 α。通过进行 N 次迭代后，最终完成扰动总值 ε 的添加。

3 Logits 组合评价

通常情况下，研究者一般以 DNN 模型的 softmax 层的输出中最大的预测值作为置信度，以评估分类器的分类性能，而这主要利用了 softmax 的两个主要优势，一是归一化，即不同的预测概率值和为 1，每一个样本的预测值都是一个小于 1 的数值；二是扩大化，即 softmax 的指数运算有明显的放大效果，如果输入分别为 1 和 5，输出中 5 的占比能到 98% 以上。然而，当分析对抗样本的目标攻击性能时，更专注的是模型对不同样本预测的原始值，即能够从预测结果中观察到不同样本之间的真实差异，从而精准地衡量和评估扰动添加后模型对数据集的预测结果的实际影响。

因此，本文从分类器模型的倒数第二层即 logit 层入手，来对当前目标对抗样本的攻击性能展开评估。提出

了一套改进的通用指标，即logits组合评价，该指标由源类logits差值和目标类logits差值构成。通过两个不同角度的logits差值，可以直观地看出扰动添加前后模型对当前样本的分类效果。

源类logits差值表示从真实类输出中减去所有不正确类的最大输出，可以通过以下方式描述：

$$\begin{aligned}\Delta logits &= l_s - l_T \\ l_T &= \max(l_k \forall k \neq k)\end{aligned} \tag{9}$$

式中：l_s为原始类别的logit值；l_T为除了原始类别之外其他类别中具有最大预测的logit值。显然，在没有扰动添加的情况下，logits差值应当为正，即模型的预测标签即为样本的真正标签，而随着扰动的逐步添加，当样本被识别为其他类别后，显然l_T会大于l_s，此时logits差值就会为负值。从诱导目标的类别角度出发，又可以logits组合评价中的目标类logits差值：

$$\begin{aligned}\Delta logits &= l_t - l_S \\ l_S &= \max(l_i \forall i \neq i)\end{aligned} \tag{10}$$

式中：l_t为所选定的目标类型的logit值；l_S为除了目标类之外的最大的logit值。显然，从目标类的角度出发，可以更加清晰直观地衡量和判断出当前目标性攻击的性能和不同信号类型的抗扰动能力。

4　实验设置与性能分析

4.1　实验模型和数据集

为了确保调制数据与分类器模型相匹配，选择了参考文献[15]中使用的VT-CNN2模型。修改了网络参数、层数和初始权重，并使用2×128格式的输入样式阵重塑了110 000个输入信号。这些信号在长度、宽度和高度方面进行了调整，以确保模型能够更好地对信号特征进行提取。VT-CNN2模型的流程图如图1所示。

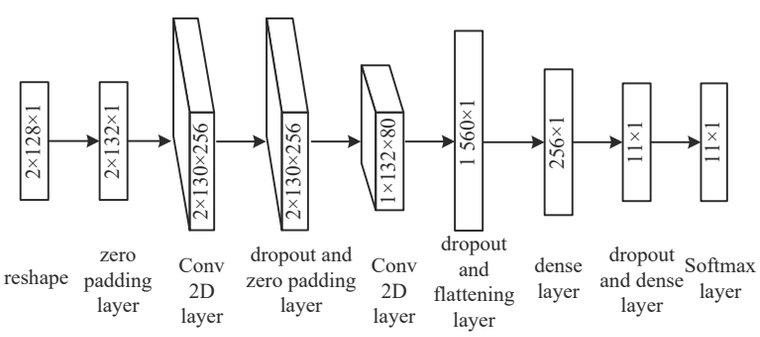


Fig. 1 Schematic diagram of VT-CNN2 structure

图1 VT-CNN2 结构示意图

选择了GNU Radio生成的RADIOML 2016.10A[16]作为调制识别的数据集。该数据集由11种不同信噪比的调制信号(8种数字信号和3种模拟信号)组成，其中包括的数字信号为BPSK、QPSK、8PSK、QAM16、QAM64、CPFSK、GFSK和PAM4，模拟信号为WBFM、AM-SSB和AM-DSB。数据集总共生成了220 000个在不同信噪比下的数据样本，信噪比从18 dB一直到−20 dB，步长为2。选择了80%的样本作为训练集，20%的样本作为测试集，每一个信号的向量长度为128，其中包括了同向分量和的正交分量。本文将随机生成几组不同的对抗样本攻击后的调制信号，并就目标攻击的效果展开分析和对比。

4.2　对抗样本的攻击效果分析

与计算机视觉领域相似，在调制识别中，给输入的信号样本添加细微的扰动，在使得分类器输出错误结果的同时，能够欺骗人类，使其无法察觉出区别来。令信号的输入样本为x，扰动为r_x，则扰动后深度学习模型输出的对抗样本可以认定为$x_{\text{ad}}=x+r_x$。本文从攻击者的角度出发，生成和添加这一类的扰动r_x，使得DNN模型对扰动后的信号进行错误分类，同时也能够朝着既定的目标生成目标对抗样本。

首先，本文在10 dB和−10 dB的信噪比下，对比和分析了VT-CNN2模型输出准确率随扰动的变化。如图2(a)所示，在没有攻击的情况下，VT-CNN2在10 dB时的预测精确度为76%。随着扰动的增加，模型的预测精确度大大降低，体现了分类器模型对扰动的高敏感性。当扰动系数为0.001时，在迭代方法BIM和PGD的攻击下，分类器模型预测精确度会显著降低50%左右，而随着扰动的进一步增加，模型的准确性不断降低，最终接近10%，达到了近乎随机化的效果。这说明，迭代攻击BIM和PGD相较于单步攻击方法FGSM，具有更加卓越的性能。

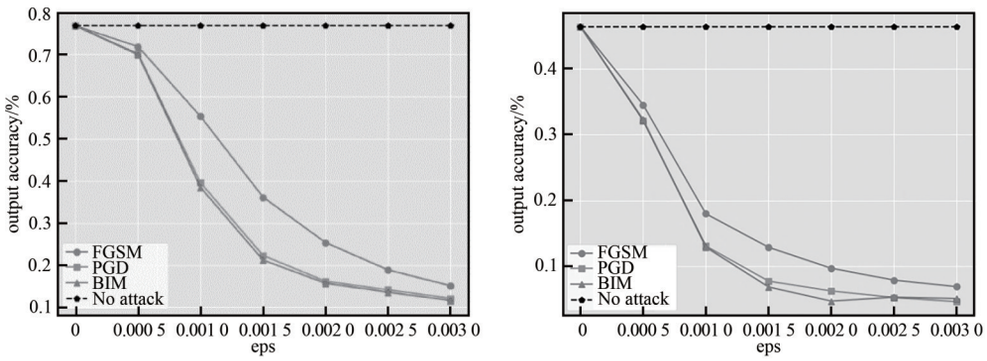


Fig.2 The prediction accuracy of the classifier model under different SNR changes with perturbation
图2 不同SNR下分类器模型预测精确度随扰动的变化

图2(b)给出了VT-CNN2在−10 dB扰动下的精确度。可以看出，在没有攻击的情况下，模型预测的精确度约为46%，低于10 dB时模型的预测精确度。在低SNR情况下，迭代算法的效果均优于FGSM，其中BIM的攻击效果强于其他两种迭代算法，在个别信噪比下有着更加突出的攻击性能。可以看出，不同的方法对噪声的敏感性不同，这就要求根据特定的情况选择合适的模型。

4.3 对抗样本的目标诱导分析

为了进一步分析当前对抗样本对调制识别模型的目标性性能，即当前的通信信号在添加扰动后是否可以诱导分类器模型归类于指定的目标类别，例如在军事场景下将一台坦克定向地错误分类为导弹或者火箭炮，本文采用了FGSM算法，进一步在VT-CNN2上发起了目标攻击，并且通过指定不同的目标类别来验证当前对抗样本的目标性能。

由图3(a)可见，在扰动为0时，模型对8PSK的预测值是最高的，印证了当前的样本为8PSK，而随着扰动强度逐步增大并达到0.001 5时，QAM16、QAM64信号的logit值开始超过了8PSK，并且随着扰动增加差距进一步扩大，此时样本已经被成功地诱导至8PSK类别。图3(b)中，扰动为0时可以看出当前信号为CPFSK信号，而随着扰动增加，CPFSK信号的预测logit值开始快速跌落，并且在扰动系数为0.001 4的时候开始被8PSK超过，并且在0.003的时候再次被QAM16反超，从而最终成功完成目标攻击。

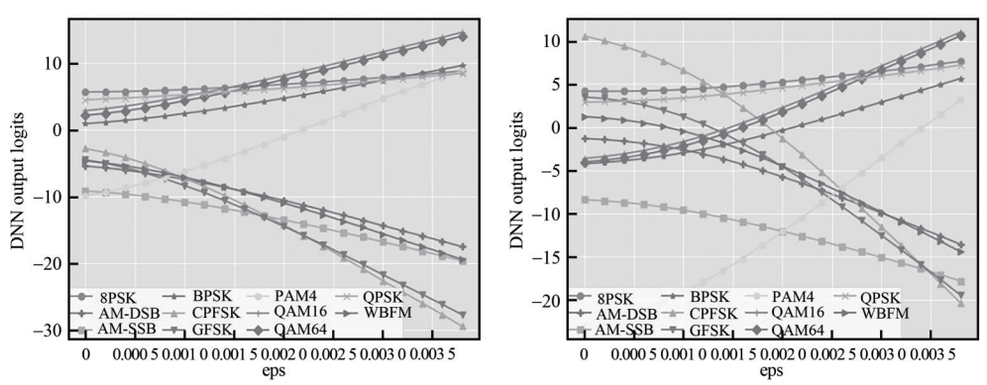
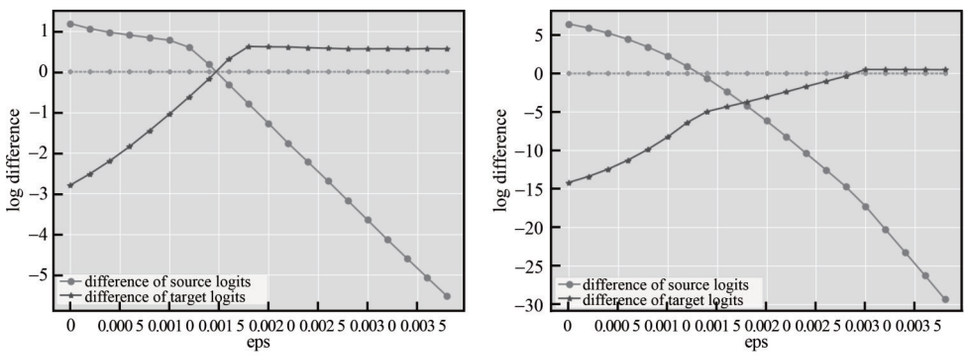


Fig.3 The predicted value output of the modulation recognition model in the logit layer
图3 调制识别模型在logit层的预测值输出

此外，不同信号对抗样本的攻击策略是不同的。对于8PSK而言，可以看出随着扰动增加，与8PSK最接近的信号类型，输出的logits均呈现出上升的趋势，也就是此时对抗样本的攻击侧重于增加其他类型信号的输出预测值，当其他信号的输出置信度高于当前信号的置信度时攻击就成功了。而对于CPFSK，对抗样本的攻击侧重于削弱对当前类型样本输出的预测值，因此其他信号的输出值高于当前信号的预测值时，攻击才会达成目标。

图4为对图3进一步计算得到的logits组合评价。如图4(a)所示，在0.001 5之前，源类logits差值为正，并且保持对8PSK的高度置信水平；而0.001 5之后，源类logits差值转为负，此时对抗样本开始生效，VT-CNN2的输出准确率开始降低。而目标类logits呈现先负后正，并也在0.001 5时与x轴交接，这说明了在该扰动值之后，对抗样本将原始类误导的类别即为目标类别。而图4(b)中，源类logits差值迅速下降，使得在0.001 4的时候logits

差值就与x轴交界并出现了误分类，即非目标攻击。只有到了0.003的扰动水平时，目标类logits差值才转负为正，表明此时才成功地实现了目标攻击。



(a) origin: 8PSK; target: QAM16　　(b) origin: CPFSK; target: QAM16

Fig.4 The logit difference between the source class and the target class of the modulation recognition model

图4 调制识别模型在logit层的预测差值

可见，有了logits差值，可以更加形象地将不同样本logit值给剥离出来，专注于对当前样本的预测变化趋势，并且可以根据logits差值曲线与x轴交点的远近，从而判断当前的样本的易扰动性以及是否容易被对抗样本诱导为其他类别。此外，根据logits组合评价，还可以进一步分析得到不同信号样本的抗扰动能力，即通过观察不同信号首次被成功攻击时对应的扰动系数，判断当前样本的易诱导能力和鲁棒性。

5　结论

本文分析了目标对抗性攻击在调制识别中引起的安全性问题，并结合所提出的logits组合评价指标评估了目标对抗样本的有效性和目标性。本文在不同的实验场景下对调制信号添加了细微的人眼无法察觉的扰动，然后将其输入至深度学习模型，以愚弄分类器并诱导其预测出错误的类别，过程中使用了logits组合评价来对目标攻击效果进行细粒度的评估。实验结果表明，对抗样本对基于深度学习的调制识别任务造成了严重的威胁和破坏，可以诱导模型将输入信号预测为所指定的类别，所提的logits指标可以帮助更细粒度地分析目标对抗攻击的能力，以及评估信号样本的易扰动性。

参考文献：

[1] O'SHEA T J,ROY T,CLANCY T C. Over-the-air deep learning based radio signal classification[J]. IEEE Journal of Selected Topics in Signal Processing, 2018,12(1):168-179.

[2] O'SHEA T, HOYDIS J. An introduction to deep learning for the physical layer[J]. IEEE Transactions on Cognitive Communications and Networking, 2017,3(4):563-575.

[3] LIN Y,TU Y,DOU Z,et al. The application of deep learning in communication signal modulation recognition[C]// 2017 IEEE/CIC International Conference on Communications in China(ICCC). Qingdao,China:IEEE, 2017:1-5.

[4] 张智博,樊雅玄,孟骁. 基于谱图和神经网络的通信干扰模式识别方法[J]. 太赫兹科学与电子信息学报, 2019,17(6):959-963. (ZHANG Zhibo, FAN Yaxuan, MEGN Xiao. Pattern recognition method of communication interference based on power spectrum density and neural network[J]. Journal of Terahertz Science and Electronic Information Technology, 2019, 17(6): 959-963.)

[5] SUN H,CHEN X,SHI Q,et al. Learning to optimize:training deep neural networks for wireless resource management[C]// 2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications(SPAWC). Sapporo,Japan:IEEE, 2017:1-6.

[6] 查雄,彭华,秦鑫,等. 基于多端卷积神经网络的调制识别方法[J]. 通信学报, 2019(11):30-37. (ZHA Xiong, PENG Hua, QIN Xin,et al. Modulation recognition method based on multi-inputs convolution neural network[J]. Journal of Communications, 2019 (11):30-37.)

[7] TU Y, LIN Y, WANG J, et al. Semi-supervised learning with generative adversarial networks on digital signal modulation classification[J]. Computers, Materials & Continua, 2018,55(2):243-254.

[8] 吴灏,周亮,李亚星,等. 基于卷积神经网络和稀疏滤波的调制识别方法[J]. 系统工程与电子技术, 2019(9):2114-2121. (WU Hao, ZHOU Liang, LI Yaxing, et al. Modulation classification based on convolutional neural network and sparse filtering[J].

Systems Engineering and Electronics, 2019(9):2114-2121.)

[9] SZEGEDY C,ZAREMBA W,SUTSKEVR I,et al. Intriguing properties of neural networks[C]// 2nd International Conference on Learning Representations. Banff,Canada:[s.n.], 2012:1-10.

[10] LIN Y,ZHAO H,TU Y,et al. Threats of adversarial attacks in DNN-based modulation recognition[C]// 2020 IEEE Conference on Computer Communications. Toronto,ON,Canada:IEEE, 2020:2469-2478.

[11] MADRY A, MAKELOV A, SCHMIDT L, et al. Towards deep learning models resistant to adversarial attacks[C]// The 6th International Conference on Learning Representations. Vancouver,Canada:[s.n.], 2016:1-23.

[12] KURAKIN A,GOODFELLOW I,BENGIO S. Adversarial examples in the physical world[C]// The 6th International Conference on Learning Representations. Vancouver,Canada:[s.n.], 2016:128-141.

[13] GOODFELLOW I J, SHLENS J, SZEGEDY C. Explaining and harnessing adversarial examples[C]// The 5th International Conference on Learning Representations. San Diego,CA,USA:2015:1-11.

[14] YUAN X, HE P, ZHU Q, et al. Adversarial examples: attacks and defenses for deep learning[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019,30(9):2805-2824.

[15] O'SHEA T, WEST N. Radio machine learning dataset generation with GNU radio[J]. Proceedings for the 6th GNU Radio Conference, 2016,1(1):1-6.

[16] DeepSig. Deepsig dataset:radioml 2016[EB/OL]. [2020-12-10]. https://www.deepsig.io/datasets.

作者简介：

赵浩钧(1996-)，男，硕士，主要研究领域为通信技术、机器学习和安全分析等.email:jeawray@163.com.

林　云(1980-)，男，博士，教授，主要研究领域为智能无线电技术、人工智能和机器学习、大数据分析与挖掘、软件和认知无线电、信息安全与对抗等.

包志达(1996-)，男，在读博士研究生，主要研究领域为对抗机器学习和人工智能安全等.

史继博(1996-)，男，在读博士研究生，主要研究领域为联邦学习和数据隐私安全等.

葛　斌(1979-)，男，博士，副教授，主要研究领域为非线性分析、偏微分方程和微分包含等.