

# 基于 GPREGAN 框架的人体康复运动对抗样本的研究

郑康洁<sup>1</sup>, 靳珊<sup>2</sup>, 张程伟<sup>2\*</sup>

<sup>1</sup>大连海事大学航海学院, 辽宁 大连 116026;

<sup>2</sup>大连海事大学信息科学技术学院, 辽宁 大连 116026

**摘要** 近年来机器学习的出现给人体康复运动领域的建模带来希望, 基于深度学习的分类识别已经实现了很高的识别率。深度模型特性会使得传感器在对抗噪声的攻击下, 识别率受到影响。因此基于生成式对抗网络 (Wasserstein Generative Adversarial Networks, WGAN), 提出了生成人体康复运动 GAN (Generative Physical Rehabilitation Exercise GAN, GPREGAN) 框架, 它可以将攻击性数据伪装成正常的的数据。这些对抗数据与原始数据高度相似, 检测算法无法区分。实验中将生成的对抗数据输入到基于卷积神经网络 (Convolutional Neural Networks, CNN) 和长短期记忆 (Long Short-Term Memory, LSTM) 网络的深度模型中, 检测率从 99% 降至 0, 成功攻击了网络。为了评估生成的对抗样本的有效性, 使用样本均方差进行评估。实验证明, GPREGAN 框架具有生成类似于人体康复运动领域时序数据的能力, 可以增加该领域中样本的多样性。

**关键词** 传感器; 人体康复运动; 深度模型; 对抗样本; 均方差

中图分类号 TP212.3

文献标志码 A

doi: 10.3788/LOP202158.2428004

## Research on Adversarial Examples in Human Physical Rehabilitation Exercises Based on GPREGAN Framework

Zheng Kangjie<sup>1</sup>, Jin Shan<sup>2</sup>, Zhang ChengWei<sup>2\*</sup>

<sup>1</sup>Navigation College, Dalian Maritime University, Dalian, Liaoning 116026, China;

<sup>2</sup>School of Information Science and Technology, Dalian Maritime University, Dalian, Liaoning 116026, China

**Abstract** The advent of machine learning in recent years has given a hope for modeling in the field of human physical rehabilitation exercises, and the classification recognition based on deep learning has achieved a high recognition rate. The characteristics of the depth model can make the sensor suffer from noise attacks in the recognition rate. Thus here based on the Wasserstein generative adversarial network (WGAN), the generative physical rehabilitation exercise GAN (GPREGAN) framework is proposed, which is improved to disguise aggressive data as normal data. This adversarial data is so highly similar to the original data that the detection algorithms cannot distinguish between them. The generated adversarial data is fed into a deep recognition model based on convolutional neural network (CNN) and long short-term memory (LSTM) network in the experiments, and the detection rate is reduced from 99% to 0 by successfully attacking the network. To evaluate the effectiveness of the generated adversarial samples, the paper uses the sample mean square error for evaluation. It is demonstrated that the GPREGAN framework has the ability to generate time-series data analogous to that in the field of human physical rehabilitation exercises and to increase the diversity of samples in this field.

**Key words** sensors; human rehabilitation exercise; depth model; adversarial examples; mean square error

**OCIS codes** 280.4788; 200.4260; 120.0280

收稿日期: 2021-01-05; 修回日期: 2021-02-22; 录用日期: 2021-03-02

基金项目: 国家自然科学基金(61906027, 61906135)、中国博士后科学基金资助项目(2019M661080)、中央高校基本科研基金(3132020211)

通信作者: \*cheny@dmlu.edu.cn

# 1 引言

随着人口老龄化的加剧,高血压、高血脂等疾病的发病率逐年上升。这些疾病患者会出现一定程度的运动障碍,生活质量严重受影响。研究表明,大量重复性和高密度性的人体康复运动可以使得患者受损的肌肉骨骼得到最大限度的恢复,从而改善患者的生活质量。由于缺乏专业康复人员对患者进行监督以及患者恐惧再伤害等心理因素,康复治疗不能成功实施。因此人体康复运动的工作重点主要集中在工具的开发,例如机器人辅助设备<sup>[1]</sup>、外骨骼、触觉设备和虚拟游戏环境等<sup>[2]</sup>。然而,运动数据复杂的随机性和非线性特征给人体运动建模带来了挑战。

近年来机器学习的出现给人体康复运动领域的建模带来了希望。其中,为特定问题提供适当的数据集对于增强现有算法的性能以及开发和评估新算法至关重要。前人的工作主要集中在运动识别和分类的研究上,传统的识别方法是基于骨骼关节的时间序列<sup>[3-9]</sup>。随着技术的发展,人体康复运动领域的运动数据采集方式主要集中在光学运动捕捉传感器上。Kinect 作为一款廉价的三维数据采集传感器,可以获取场景的影像信息与深度数据,可快速获取人体三维点数据,常被应用于人体的三维重建<sup>[10-12]</sup>。研究者发现,机器学习同样可以很好地解决分类识别问题,其中深度学习最为成功,可预测人体未来的运行模式<sup>[13-14]</sup>并进行分类<sup>[15]</sup>。但是,深度模型易受对抗噪声(对抗样本)的干扰<sup>[16-18]</sup>。对于分类任务,一点微小的干扰就可能影响分类的结果。

上述大部分工作都是基于前向传播神经网络,它们均假设数据间相互独立。但是传感器信息是有时序关系的,因此本文使用的基础模型除了卷积神经网络(CNN)之外,还使用了长短期记忆网络(LSTM)。受人体表面光滑程度等环境因素的影响,Kinect 传感器捕获的三维数据精度偏低,获取的数据噪声较大,不利于人体的三维重建。因此,本文基于生成式对抗网络(Wasserstein Generative Adversarial Networks, WGAN),提出了生成人体康复运动 GAN(Generative Physical Rehabilitation Exercise GAN, GPREGAN)框架。首先通过引入噪声验证深度模型对识别率的影响,发现识别率高达 99%。基于该实验,本文在原样本上添加了微小的干扰,通过 GPREGAN 框架产生对抗样本,并对深度模型进行实验,发现识别率从 99%降至 0,成功

攻击了深度模型。为了验证对抗样本的有效性,本文使用样本均方差进行评估,成功为人体康复运动领域提供了开源的对抗样本数据,增加了样本的多样性。

## 2 相关工作

### 2.1 人体康复运动识别

人体运动识别是指计算机对人体行为的描述和理解,主要研究内容有目标检测、特征提取和行为识别等。随着传感器技术的发展,基于骨骼视频序列和深度图像的人体运动识别方法被提出。文献[19]已经证明了人体行为模式可以用若干个人体关节来表示。Kinect 传感器则通过处理深度数据来得到人体各个关节的位置坐标。基于骨骼的人体运动识别方法大致分为两种:基于关节和基于身体部位。

对于关节的研究主要包括人体骨骼关节和关节组合的特征提取。文献[20]从骨骼关节的子集中提取空域和时域的局部特征向量。文献[21]提出一种基于稀疏编码和时间金字塔的匹配框架用于动作识别。在机器学习出现以后,文献[22]提出基于 CNN 和 LSTM 的模型进行分类识别。在关于身体部位的研究中,文献[23]将人体分为五个部分,使用各部分运动参数来表示人体动作并进行分类。文献[24]使用 3D CNN 来处理三维空间的人体动作,提取了时间和空间上的维度,从而捕捉到多个连续帧之间的动作信息。

基于骨架的行为识别技术有两个关键点:一方面是设计有强判别性的稳定特征,另一方面是利用时域相关性对行为动作的动态变化进行建模。

### 2.2 深度学习

深度学习是机器学习的一种,它的概念来源于人工神经网络。主要分为三类:卷积神经网络、自编码神经网络和深度置信网络。

CNN 是一种前馈神经网络,它的人工神经元可以响应部分覆盖范围内的周边单元。CNN 主要用来识别位移、缩放及其他形式扭曲不变性的二维图形。由于 CNN 的特征检测是通过训练数据进行的,因此在训练数据中进行隐式学习可避免显式特征的抽取。同一特征映射面上的神经元权值相同,所以网络可以并行学习,CNN 凭借局部权重共享的特殊结构在语音识别和图像处理领域占领优势。

循环神经网络也是深度学习中的一种网络,以序列数据为输入,在序列的演进方向进行递归且所

有节点按链式连接。它在语音识别、语言建模、机器翻译等自然语言处理领域有应用,也被用于各类时间序列的预报。在循环神经网络中,LSTM 是一种时间循环神经网络,可解决一般循环神经网络存在的梯度消失问题。相比普通的循环神经网络,LSTM 在更长的序列中有更好的表现。

近期研究发现,神经网络很容易受到欺骗,细微扰动构造出的对抗样本能够愚弄深度模型,导致模型的识别率大幅下降,甚至使模型错误地将毫无意义的噪音以极高的置信度识别为特定类型。因此文献[25]基于神经网络的特性提出了对抗样本的概念。而产生对抗攻击的根本原因是:深度学习在高维空间的线性特性已经足以产生这种攻击行为。探索深度学习中的对抗因素成为研究热点。

### 3 GPREGAN 框架的搭建

本节中我们将介绍一种能够生成人体康复运动领域内的对抗样本的框架。在保持 Kinect 传感器数据原始时域特征的情况下,我们引入了一种高信噪比噪声来干扰模型。我们在 WGAN 的基础上提

出了 GPREGAN 框架。GPREGAN 框架由干扰器、生成器、鉴别器组成。生成器的作用是生成一种干扰来扰乱神经网络的判断。为了模拟 Kinect 传感器识别不同噪声,我们在生成器中设计了一种基于 CNN 的网络结构。干扰器的作用是为了模拟 Kinect 传感器数据的加噪方式,将生成器生成的干扰与原始样本进行融合以形成对抗样本。而鉴别器的作用则是为了判别生成样本与原始样本的真伪。

GPREGAN 框架如图 1 所示,其中  $X$  为原样本, $G(X)$  为生成器生成的噪声。该框架中的原始数据会交互给生成器  $G$ , $G$  通过神经网络进行特征提取,再通过反卷积产生具有对抗能力的微小干扰,该干扰使得被攻击的目标网络识别出我们想要的结果;将生成的干扰与原始数据在干扰器  $O$  中进行结合以生成对抗样本,将生成的样本与原始样本传入鉴别器  $D$  中进行分类识别。 $G$ 、 $O$ 、 $D$  在博弈过程中不断地进行训练和自我优化,最终达到均衡,使得  $G$  在干扰极小的情况下, $O$  产生的对抗样本与原始样本的相似度极高,并且能够成功诱骗目标网络,使得识别出错。

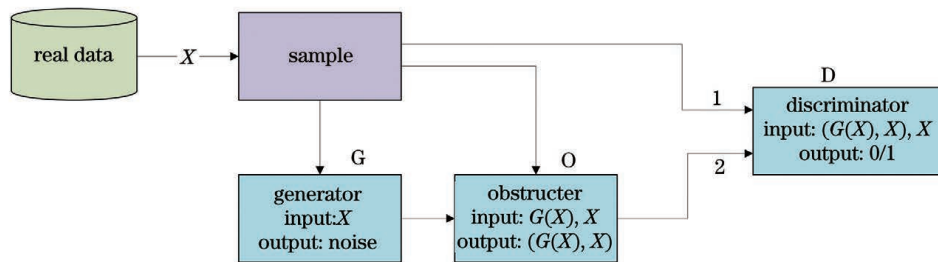


图 1 GPREGAN 框架图

Fig. 1 GPREGAN framework

#### 3.1 生成器 G

生成器  $G$  的功能是生成多传感器采集的时间序列数据。由于每一个传感器采集到的时间序列数据不同,且存在关联性,因此我们设计了一个三层的一维卷积(One-Dimensional Convolution, Conv1D)堆叠结构以读取原样本的特征。在读取到的原样本特征的基础上加入噪声,得到干扰的特征,然后再通

过三层一维反卷积(One-Dimensional Deconvolution, DeConv1D)堆叠结构来制造干扰误差。每层卷积层使用带泄露整流函数(Leaky Rectified Linear Unit, Leaky ReLU)作为激活函数,而最后一层使用双曲正切(tanh)函数。在每层之间我们还引入了群组归一化(Group Normalization, GN)<sup>[26]</sup>以最大化干扰样本的丰富性。生成器  $G$  的网络结构如图 2 所示。

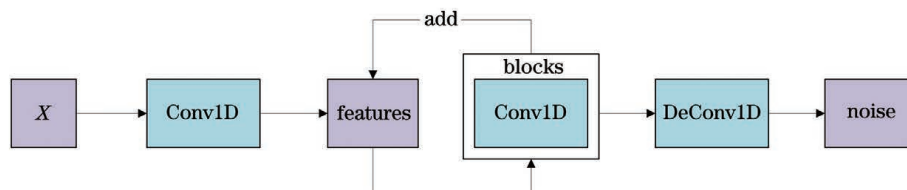


图 2 生成器的网络结构

Fig. 2 Network structure of generator

### 3.2 干扰器 O

由于 Kinect 传感器获取的深度信息噪声数据较大,因此我们需要模拟出实际情况的影响,即对样本添加噪声。其中,评价噪声对传感器信号影响强弱的参数是信噪比(SNR,  $R_{\text{SNR}}$ ),较高的 SNR 代表噪声对信号的影响较小,较小的 SNR 代表噪声对信号的影响较大。SNR 的计算公式为

$$R_{\text{SNR}} = 10 \lg \left( \frac{p_{\text{signal}}}{p_{\text{noise}}} \right), \quad (1)$$

式中:  $p_{\text{signal}}$  表示信号的功率;  $p_{\text{noise}}$  表示噪声的功率。连续信号  $f(t)$  在  $t$  时刻的瞬时功率( $P$ )可以记为  $f(t)$  的模的平方,信号的能量( $E$ )是瞬时功率的积分值,而一组离散信号  $[x_1, x_2, \dots, x_i, \dots, x_n]$  的平均功率可表示为

$$\bar{P} = \frac{1}{n} \sum_{i=1}^n x_i^2, \quad (2)$$

式中:  $\bar{P}$  为平均功率;  $n$  为信号的序列长度;  $x_i$  为第  $i$  个时域信号;  $i$  为信号编号。

因此在一组序列长度为  $n$  的时域信号中,一个功率为  $p$  的信号  $x_i$  的加噪方式  $x'_i$  可以表示为

$$x'_i = x_i + n \times \sqrt{p}. \quad (3)$$

干扰器 O 将生成器生成的噪声  $G(X)$  与原样本  $X$  结合以得到一个受干扰的对抗样本。对于生成器  $G$  生成的噪声,如果噪声太大,则结合后的对抗样本失去了原有特性。为了得到足够小的噪声,用 SNR 来衡量噪声对原样本的影响。我们设定传感器信号误差的接受范围为  $\eta$ , 则当生成噪声对应的原样本的 SNR 小于传感器信号误差接受范围  $\eta$  对应的原样本的 SNR 时,在保持生成噪声特征不变的情况下,通过减小噪声的功率可提高生成噪声对应的原样本的 SNR。

假设生成器  $G$  获得一个噪声信号并求得功率为  $p_{\text{noise}}$ ; 然后设定传感器信号误差的接受范围为  $\eta$ , 则可以求出原样本的功率为  $p_\eta$ 。分别计算噪声对应的原样本的信噪比  $R_{\text{SNR}}^{\text{noise}}$  和在传感器误差范围  $\eta$  内的原样本的信噪比  $R_{\text{SNR}}^{(\eta)}$ , 最后进行比较可以得出  $R_{\text{SNR}}^{\text{noise}} \leq R_{\text{SNR}}^{(\eta)}$ 。当  $p_{\text{noise}} \leq p_\eta$  时,保持噪声信号不变; 当  $p_{\text{noise}} > p_\eta$  时,我们将噪声信号功率按比例缩小到  $p_\eta$ , 从而保证信噪比在  $R_{\text{SNR}}^{(\eta)}$  内。然后将该生成噪声在干扰器 O 中与原样本进行结合,最后生成对抗样本。

### 3.3 鉴别器 D

在鉴别器 D 的设计中,不采用任何池化层,而是用带有步长的卷积来代替池化层。该网络使用批

归一化方法,先采样一小批数据,然后在网络各输出层对该批数据进行归一化处理,以此加快 D 网络的训练速度,防止过拟合。该网络同样使用 Leaky ReLU 作为激活函数。目的是输入一个样本  $X$ , 利用网络层不断尝试区分正常样本和对抗样本。在鉴别器 D 和生成器  $G$  的对抗训练过程中,  $G$  将生成较多的与正常数据相似的对抗样本。

### 3.4 训练 GPREGAN 框架

在训练的过程中,生成器  $G$  将原始样本  $X$  作为输入,生成了噪声  $G(X)$ ; 干扰器 O 将生成的噪声  $G(X)$  与原始样本  $X$  进行结合,形成了对抗样本 ( $G(X) + X$ ), 将这一结果输送到鉴别器 D 中以区分对抗样本和原始样本,当 D 无法区别对抗样本与原始样本时,训练终止。为了实现这一过程,我们首先要对网络进行攻击,将目标模型定义为  $f'$ , 则  $f'$  以  $G(X) + X$  作为输入,输出了损失函数  $L_{\text{tar}}$ 。该损失函数代表着预测值与攻击目标之间的距离,根据 GAN 的损失函数定义,该  $L_{\text{tar}}$  的表达式为

$$L_{\text{tar}} = E_X [\log D(X)] + E_X \{ \log \{ 1 - D[X + G(X)] \} \}, \quad (4)$$

式中:  $E_x(\cdot)$  为期望;  $D(X)$  为判别器输出的结果。

则在目标攻击中,目标模型  $f'$  的损失函数定义为

$$L_{\text{tar}}^{f'} = E_X l_{f'} [X + G(X), t'], \quad (5)$$

式中:  $t'$  是目标类型;  $l_{f'}$  是被攻击模型  $f'$  的交叉损失。  $L_{\text{tar}}^{f'}$  的目的是让对抗信号被目标模型  $f'$  识别为目标类型  $t'$ , 也就是让对抗信号被错误识别为原始样本中的一部分。为了进一步限制干扰的大小,我们在 L2 范数上添加了铰链损失(hinge loss):

$$L_{\text{hinge}} = E_X \max(0, \|G(X)\|_2 - c), \quad (6)$$

式中:  $c$  用来稳定 GAN 网络的训练,从而得到对抗样本。因此全部损失表示为

$$L = L_{\text{tar}}^{f'} + \lambda L_{\text{tar}} + \gamma L_{\text{hinge}}, \quad (7)$$

式中:  $\lambda$  代表  $L_{\text{tar}}$  的权重;  $\gamma$  代表  $L_{\text{hinge}}$  的权重。  $L_{\text{tar}}$  用来鼓励生成的对抗样本与原始样本尽可能相似,而  $L_{\text{tar}}^{f'}$  则用来生成对抗样本,优化攻击效果。

### 3.5 对抗样本的评估

评估和比较生成式对抗网络产生的数据极具挑战性,对于图像来说,可以通过人眼来客观评定图像生成的质量,但是这种方法无法直接应用到数据拟合上。在本实验中,我们的目的是尽可能生成较多的对抗样本又不缺失原始样本该有的特征属性。因此我们用样本均方差指标来进行评估。样本均方差是样本中各数据与样本平均数的差的平方和的平均数,是衡量一个样本波动大小的量,样本均方差越

大,样本数据的波动就越大。我们利用样本均方差来获得原始样本与生成的对抗样本之间的波动。相似性越大,则均方差越小,因此我们希望均方差足够小,这样对抗样本就具有原始属性。该模型的样本均方差计算公式为

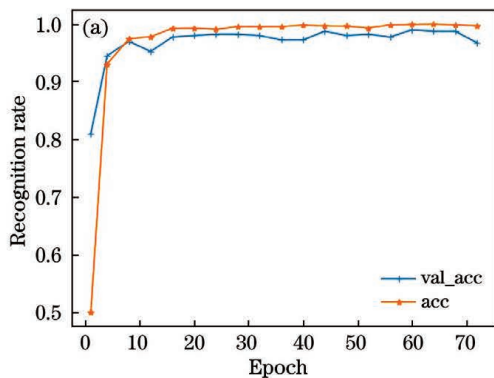
$$\bar{X} = \frac{x_1 + \dots + x_i + \dots + x_n}{n}, \quad (8)$$

$$S^2 = \frac{\sum (x_i - \bar{X})^2}{n - 1}, i = 1, \dots, n. \quad (9)$$

## 4 实验分析

### 4.1 实验数据

在本实验中我们使用的是爱达荷大学-身体康复运动数据(UI-PRMD)<sup>[27]</sup>,它是患者进行人体康复运动的公开数据集。该数据使用 Kinect 传感器对受试者的不同物理运动治疗进行捕捉。数据包括了人体全身关节的位置和角度。



在数据集的预处理过程中,我们对正负样本进行平铺,以方便分类。由于数据维度不同,我们使用插值方式将样本维度修正为固定值。最后我们将平铺后的数据进行打乱。我们还对数据进行了归一化的操作,因为归一化训练样本的输入值将有助于加快学习阶段。

### 4.2 深度模型下的识别结果

我们使用 80% 的数据作为训练集,20% 的数据作为测试集。CNN 模型的网络结构设计是将四个 Conv1D 层叠在一起,卷积核大小为 128×60。为了使模型能够提取更好的特征,最后一层使用全连接层,激活函数为 softmax。训练时,损失函数为交叉熵损失函数,优化器为 Adam,学习率为 0.002,结果如图 3(a)所示。基于 LSTM 模型,将三个 LSTM 层堆叠在一起,卷积核大小、激活函数、损失函数以及优化器设置与 CNN 模型相同,学习率设置为 0.001,结果如图 3(b)所示。

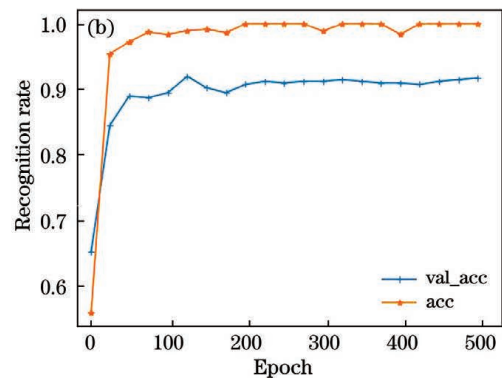


图 3 深度模型的识别率。(a)CNN 识别率;(b)LSTM 识别率

Fig. 3 Recognition rate of depth model. (a) CNN recognition rate; (b) LSTM recognition rate

从实验结果可以看出,经过训练后,CNN 在训练集上的准确率(acc)达到 100%,测试集的准确率(val\_acc)达到 0.9925,LSTM 在训练集上的准确率达到 100%,测试集的准确率为 0.9100。

### 4.3 噪声对深度模型的影响

针对 CNN 和 LSTM 模型,分别引入不同信噪

比的噪声以验证噪声对深度模型的影响,深度模型的网络结构设置和实验设置与 4.2 小节一致。实验中将测试集与训练集同时加上相同噪声,以测试不同 SNR 噪声对深度模型的影响,实验结果如表 1 所示。

表 1 不同 SNR 对深度模型的影响  
Table 1 Influence of SNR on depth model

SNR	-8 dB	-6 dB	-4 dB	-2 dB	0	2 dB	4 dB	6 dB	8 dB
CNN	0.7625	0.8175	0.9100	0.9300	0.9600	0.9725	0.9775	0.9850	0.9875
LSTM	0.2600	0.3625	0.4475	0.5375	0.7275	0.7525	0.8100	0.8700	0.8925

从表 1 可以看出,两种深度模型在低 SNR 时的识别率都受到了影响,但是在高 SNR 时,可以保持稳定的识别率,因此我们将在下面实验中引入对抗

网络生成的新样本以测试两种深度模型的识别率。

### 4.4 GPREGAN 框架对深度模型的影响

本实验中通过 GPREGAN 框架来生成对抗样

本,并攻击已训练好的深度模型 CNN 和 LSTM,目的是使生成的对抗样本在鉴别器中被分类为原始样本。根据实验 4.3 的结果,我们设定噪声的 SNR 在 6 dB 以上,并希望在范围内找到较多的对抗样本,因此我们定义 SNR 为 6 dB,接受域为 0.5。然

后测试对抗样本对训练好的 CNN 和 LSTM 模型的影响。在训练过程中,干扰器生成的噪声足够小,但是目标网络 CNN 和 LSTM 仍错误识别引入的对抗样本。实验结果如图 4 所示,其中 d\_loss 是鉴别器的损失, gan\_loss 是生成器的损失。

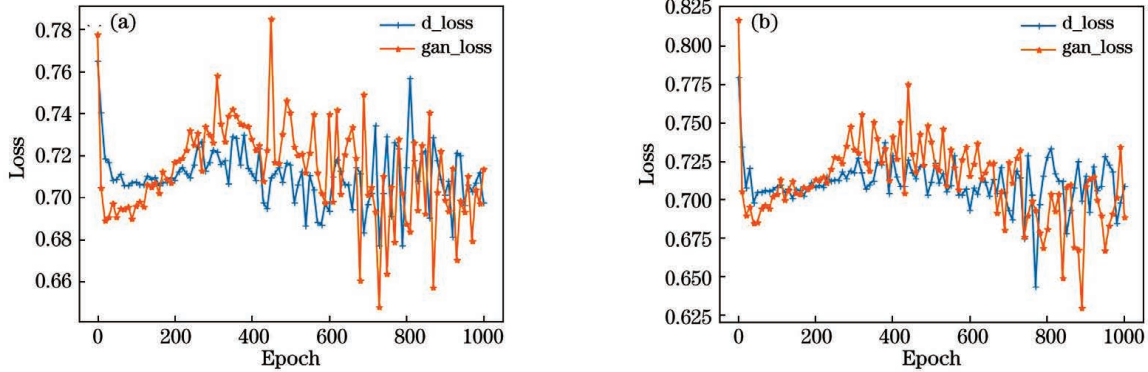


图 4 深度模型对对抗样本的识别率。(a)CNN 的识别率;(b)LSTM 的识别率

Fig. 4 Recognition rate of depth model for adversarial examples. (a) CNN recognition rate; (b) LSTM recognition rate

可以看出,GAN 在训练时鉴别器与生成器的损失(loss)很稳定,这样我们可以稳定地训练

GPREGAN 框架中的 D 和 G。引入了生成的对抗样本与原样本的距离比较,结果如图 5 所示。

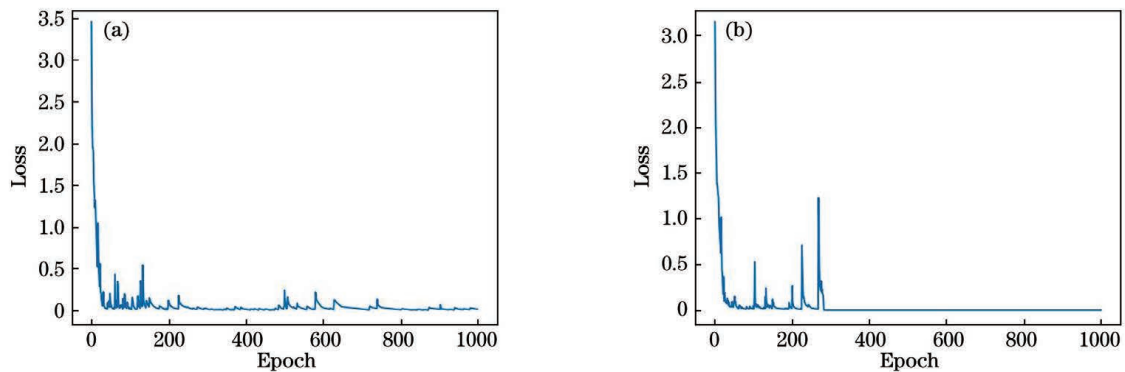


图 5 对抗样本与原样本的距离。(a)CNN;(b)LSTM

Fig. 5 Distance between adversarial example and original sample. (a) CNN; (b) LSTM

除了具有稳定的 GAN 训练以及较短的距离,如果进一步可以保证攻击率,那么我们的模型就

是有效的。幸运的是,图 6 所示结果是令人欣慰的。

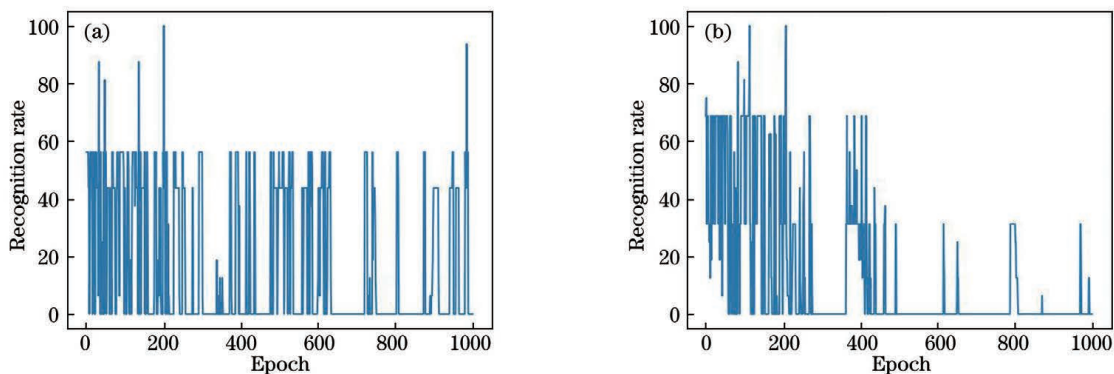


图 6 对抗样本的识别率。(a)CNN;(b)LSTM

Fig. 6 Recognition rate for adversarial examples. (a) CNN; (b) LSTM

可以看出,对于对抗样本,即使是在距离很短的情况下,CNN 与 LSTM 模型仍然无法识别。这进一步说明了我们的攻击模型的有效性。

#### 4.5 对抗样本的评估

本实验中分析了 GPREGAN 框架生成的对抗样本与原样本的样本均方差。对比了 8 组原样本与加入干扰后生成的对抗样本的均方差,结果如表 2 所示。

表 2 原样本与对抗样本的均方差

Table 2 Mean square error between original sample and adversarial example

Group No.	1	2	3	4	5	6	7	8
CNN	0.0163	0.0154	0.0147	0.0139	0.0132	0.0127	0.0122	0.0119
LSTM	0.0008	0.0168	0.0108	0.0127	0.0049	0.0134	0.0047	0.0175

我们可以看出,GPREGAN 框架生成的对抗样本与原始样本的均方差足够小,这表明生成的对抗样本与原始样本的相似性极高,再次验证了 GPREGAN 框架的有效性。

## 5 结 论

基于深度学习的人体康复运动识别算法已被研究了很多年,随着人工智能逐渐成为热门研究方向,GAN 是一个成功的生成模型并且被应用在图像领域。提出了一种 GPREGAN 框架,该框架可以将一组足够小又具有高信噪比的噪声与原样本进行融合以生成对抗样本,成功降低了基于 CNN 和 LSTM 的 Kinect 传感器的分类识别率。利用原样本与 GPREGAN 生成的对抗样本之间的均方差,可以监测它们之间的相似性。结果表明,所提 GPREGAN 框架可以稳定生成对抗样本,增加人体康复运动领域样本的多样性。

### 参 考 文 献

- [1] Maciejasz P, Eschweiler J, Gerlach-Hahn K, et al. A survey on robotic devices for upper limb rehabilitation[J]. *Journal of Neuroengineering and Rehabilitation*, 2014, 11: 3.
- [2] Gauthier L V, Kane C, Borstad A, et al. Video Game Rehabilitation for Outpatient Stroke (VIGoROUS): protocol for a multi-center comparative effectiveness trial of in-home gamified constraint-induced movement therapy for rehabilitation of chronic upper extremity hemiparesis[J]. *BMC Neurology*, 2017, 17(1): 109.
- [3] Vemulapalli R, Arrate F, Chellappa R. Human action recognition by representing 3D skeletons as points in a lie group[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 588-595.
- [4] Wang J, Liu Z C, Wu Y, et al. Mining actionlet ensemble for action recognition with depth cameras [C]//2012 IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence, RI, USA. New York: IEEE Press, 2012: 1290-1297.
- [5] Lv F, Nevatia R. Recognition and segmentation of 3-D human action using HMM and multi-class AdaBoost [M]//Leonardis A, Bischof H, Pinz A. *Computer vision-ECCV 2006. Lecture notes in computer science*. Cham: Springer, 2006, 3954: 359-372.
- [6] Wu D, Shao L. Leveraging hierarchical parametric networks for skeletal joints based action segmentation and recognition[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 724-731.
- [7] Xia L, Chen C C, Aggarwal J K. View invariant human action recognition using histograms of 3D joints[C]//2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, June 16-21, 2012, Providence, RI, USA. New York: IEEE Press, 2012: 20-27.
- [8] Ji S W, Xu W, Yang M, et al. 3D convolutional neural networks for human action recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(1): 221-231.
- [9] Du Y, Wang W, Wang L. Hierarchical recurrent neural network for skeleton based action recognition [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 1110-1118.
- [10] Ma Z F, Li J, Cao L X. Fall behavior detection and analysis using a Kinect sensor[J]. *Laser & Optoelectronics Progress*, 2020, 57(21): 210402. 马宗方, 李静, 曹陇鑫. 基于 Kinect 传感器的跌倒行为的检测与分析[J]. *激光与光电子学进展*, 2020,

- 57(21): 210402.
- [11] Bao Z Q, Lü C G. Real-time gesture recognition based on Kinect [J]. *Laser & Optoelectronics Progress*, 2018, 55(3): 031008.  
鲍志强, 吕辰刚. 基于 Kinect 的实时手势识别 [J]. *激光与光电子学进展*, 2018, 55(3): 031008.
- [12] Guo L P, Chen X N, Liu B, et al. 3D-object reconstruction based on fusion of depth images by Kinect sensor [J]. *Journal of Applied Optics*, 2014, 35(5): 811-816.  
郭连朋, 陈向宁, 刘彬, 等. 基于 Kinect 传感器多深度图像融合的物体三维重建 [J]. *应用光学*, 2014, 35(5): 811-816.
- [13] Badawi A A, Al-Kabbany A, Shaban H. Multimodal human activity recognition from wearable inertial sensors using machine learning [C]//2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), December 3-6, 2018, Sarawak, Malaysia. New York: IEEE Press, 2018: 402-407.
- [14] Chen M Y, He X J, Yang J, et al. 3-D convolutional recurrent neural networks with attention model for speech emotion recognition [J]. *IEEE Signal Processing Letters*, 2018, 25(10): 1440-1444.
- [15] Wu K L, Wei X M. Analysis of psychological and sleep status and exercise rehabilitation of front-line clinical staff in the fight against COVID-19 in China [J]. *Medical Science Monitor Basic Research*, 2020, 26: e924085.
- [16] Szegedy C, Zaremba W, Sutskever I, et al. Intriguing properties of neural networks [EB/OL]. (2013-12-21) [2021-01-01]. <https://arxiv.org/abs/1312.6199>.
- [17] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets [C]//Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13, 2014, Montreal, Quebec, Canada. [S.l.: s.n.], 2014: 2672-2680.
- [18] Hu W W, Tan Y. Generating adversarial malware examples for black-box attacks based on GAN [EB/OL]. (2017-02-20) [2021-01-01]. <https://arxiv.org/abs/1702.05983>.
- [19] Johansson G. Visual perception of biological motion and a model for its analysis [J]. *Perception & Psychophysics*, 1973, 14(2): 201-211.
- [20] Luvizon D C, Tabia H, Picard D. Learning features combination for human action recognition from skeleton sequences [J]. *Pattern Recognition Letters*, 2017, 99: 13-20.
- [21] Luo J J, Wang W, Qi H R. Group sparsity and geometry constrained dictionary learning for action recognition from depth maps [C]//2013 IEEE International Conference on Computer Vision, December 1-8, 2013, Sydney, NSW, Australia. New York: IEEE Press, 2013: 1809-1816.
- [22] Mahasseni B, Todorovic S. Regularizing long short term memory with 3D human-skeleton sequences for action recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 3054-3062.
- [23] Yacoob Y, Black M J. Parameterized modeling and recognition of activities [J]. *Computer Vision and Image Understanding*, 1999, 73(2): 232-247.
- [24] Zhu W T, Lan C L, Xing J L, et al. Co-occurrence feature learning for skeleton based action recognition using regularized deep LSTM networks [C]//Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA. Menlo Park: AAAI Press, 2016: 3697-3704.
- [25] Goodfellow I J, Shlens J, Szegedy C. Explaining and harnessing adversarial examples [EB/OL]. (2014-12-20) [2021-01-01]. <https://arxiv.org/abs/1412.6572>.
- [26] Wu Y X, He K M. Group normalization [J]. *International Journal of Computer Vision*, 2020, 128(3): 742-755.
- [27] Vakanski A, Jun H P, Paul D, et al. A data set of human body movements for physical rehabilitation exercises [J]. *Data*, 2018, 3(1): 2.