Computer Science and Application 计算机科学与应用, 2025, 15(4), 335-344
Published Online April 2025 in Hans. https://www.hanspub.org/journal/csa
https://doi.org/10.12677/csa.2025.154106

Hans汉斯

生成对抗网络与胶囊网络协同优化的窃电检测方法

刘毅宁*，白　玉

沈阳航空航天大学电子信息工程学院，辽宁 沈阳

收稿日期：2025年3月20日；录用日期：2025年4月20日；发布日期：2025年4月27日

摘　要

在现代智能电网系统中，窃电是导致经济损失的主要因素。然而，现有窃电检测方法面临数据量不足的问题，尤其在窃电样本稀缺的情况下，检测器往往会偏向大多数类，从而降低窃电样本的检测精度。为了解决这一问题，本文提出一种在小样本条件下基于深度学习的窃电检测方法(CapsGAN)。该方法融合胶囊网络与卷积神经网络。卷积神经网络用于提取用电数据中的局部浅层特征，识别基本的用电模式，而胶囊网络通过动态路由机制对这些特征进行整合，进一步构建更深层次的高级特征，最终实现对窃电行为的精准识别。同时为了进一步解决数据不平衡问题，本文还引入了一种基于生成对抗网络的三方结构模型，该模型结合了鉴别器、生成器和分类器，并引入了决策边界约束。生成器与分类器通过协同训练，共同生成窃电检测中的异常样本，逐步扩展少数类的决策边界，从而有效缓解窃电样本的类别不平衡问题。通过在中国国家电网公司提供的真实数据集上的实验，结果充分验证了该方法的有效性和准确性。

关键词

深度学习，窃电检测，生成对抗网络，胶囊网络，卷积神经网络

Collaborative Optimization of Generative Adversarial Networks and Capsule Networks for Electricity Theft Detection

Yining Liu*, Yu Bai

College of Electronic and Information Engineering, Shenyang Aerospace University, Shenyang Liaoning

*通讯作者。

文章引用：刘毅宁，白玉．生成对抗网络与胶囊网络协同优化的窃电检测方法[J]．计算机科学与应用，2025，15(4)：335-344. DOI: 10.12677/csa.2025.154106

Received: Mar. 20th, 2025; accepted: Apr. 20th, 2025; published: Apr. 27th, 2025

Abstract

In modern smart grid systems, electricity theft is a major cause of economic losses. However, existing electricity theft detection methods face the problem of insufficient data, especially when theft samples are scarce. In such cases, detectors tend to be biased toward the majority class, leading to a reduction in the detection accuracy of theft samples. To address this issue, this paper proposes a deep learning-based electricity theft detection method under small sample conditions (CapsGAN), which integrates Capsule Networks (CapsNet) and Convolutional Neural Networks (CNNs). In this approach, CNNs are used to extract local shallow features from electricity consumption data, recognizing basic consumption patterns, while the CapsNet integrates these features through dynamic routing, further constructing deeper-level high-level features to accurately identify theft behaviors. Additionally, to further address the data imbalance problem, this paper also Introduces a tripartite structure model based on Generative Adversarial Networks (GANs), which combines a discriminator, generator, and classifier, and introduces decision boundary constraints. The generator and classifier are collaboratively trained to jointly generate abnormal samples for theft detection, gradually expanding the decision boundary of the minority class, thus effectively alleviating the class imbalance issue in theft samples. Experiments on a real dataset provided by the State Grid Corporation of China validate the effectiveness and accuracy of this method.

Keywords

Deep Learning, Electricity Theft Detection, Generative Adversarial Networks, Capsule Networks, Convolutional Neural Networks

1. 引言

窃电检测任务本质上是一个二元分类问题，其目标是将用户划分为正常用户和窃电用户两类。传统的窃电检测方法依赖于人工排查或者安装计量装置防止窃电，这些方法既费时又费力且检测效果较差。近年来，随着智能电表[1]在电力系统中的不断普及，电网运行过程中积累了大量传感数据用于跟踪个体终端用户的能耗情况。这些海量数据的可用性为基于深度学习的方法提供了新的机遇。卷积神经网络(CNN)、长短期记忆网络(LSTM)等模型已被成功应用于电力盗窃检测，取得了显著的效果。例如，在[2]中，提出了一种将宽卷积神经网络和深卷积神经网络相结合的新框架。在[3]中，使用循环神经网络(RNN)来建立电力使用的时序特征，从而有效识别窃电用户。在[4]中，采用了长短期记忆网络(LSTM)来捕捉长期依赖关系，提升窃电检测的效果。[5]中介绍的工作构建了一个动态生成的残差图，并利用卷积神经网络提高电力盗窃检测的准确性。虽然卷积神经网络是最受欢迎和有效的深度学习模型之一，但其输出仅反映特定特征的变体，且在检测特征变化时存在局限。尤其当卷积神经网络使用最大池化(max-pooling)来增强判别能力，但这会导致许多局部特征的丢失。

考虑到这一点，本文探索了胶囊网络(CapsNet) [6]，并将其与卷积神经网络结合，以更好地提升窃电检测的分类效果。胶囊网络的胶囊结构与卷积神经网络的特性相辅相成，同时禁用了池化操作，以更好地保留局部特征。同时为了进一步解决窃电样本不足的问题，本文引入了一种三重生成对抗网络

(TripleGAN) [7]模型架构进行协同训练。

2. 数据分析及预处理

2.1. 数据分析

本文提出的方法基于消费者每日用电量的智能电表数据，该数据来自中国国家电网公司(SGCC) [8]，包含正常电力用户和存在窃电行为用户的用电记录。该数据集涵盖了 2014 年 1 月 1 日至 2016 年 10 月 31 日期间 42,372 名家庭用户的用电数据，其中 3615 名用户被标记为异常用户，38,757 名用户为正常用户，欺诈性用户占比为 8.53%。图 1 展示了数据集中两个用户在 2015 年一年时间中的典型用电模式，直观体现了正常用户与窃电用户在能耗特征上的差异。

通过对能耗曲线的初步分析，我们发现正常用户的能耗模式通常呈现出鲜明的周期性和规律性。这些特征与用户的日常生活习惯、工作周期以及外部环境因素(如季节变化对空调或取暖设备使用的影响)密切相关。以图 1(a)为例，正常用户在夏季(7 月、8 月、9 月)用电量显著高于其他月份，反映了空调设备高频使用的趋势，而其他月份的用电量相对平稳，波动较小，整体用电模式较为稳定。相比之下，窃电用户的能耗曲线则缺乏明显的周期性和规律性，通常呈现出混乱的波动模式。例如，在图 1(b)中，某些月份的用电量出现异常大幅下降，明显偏离正常用电模式。为了进一步分析，我们提取了正常用户和电力盗窃用户为期四周(2015 年 6 月)的用电数据。可以发现，在正常情况下(图 1(c))，用户的周用电量具有显著的周期性波动，每周通常在第 2 天或第 5 天达到峰值，随后逐步下降。然而，窃电用户的用电模式(图 1(d))仅在最初三天呈现周期性波动，从第四天开始，用电量显提升随后持续下降到较低水平。

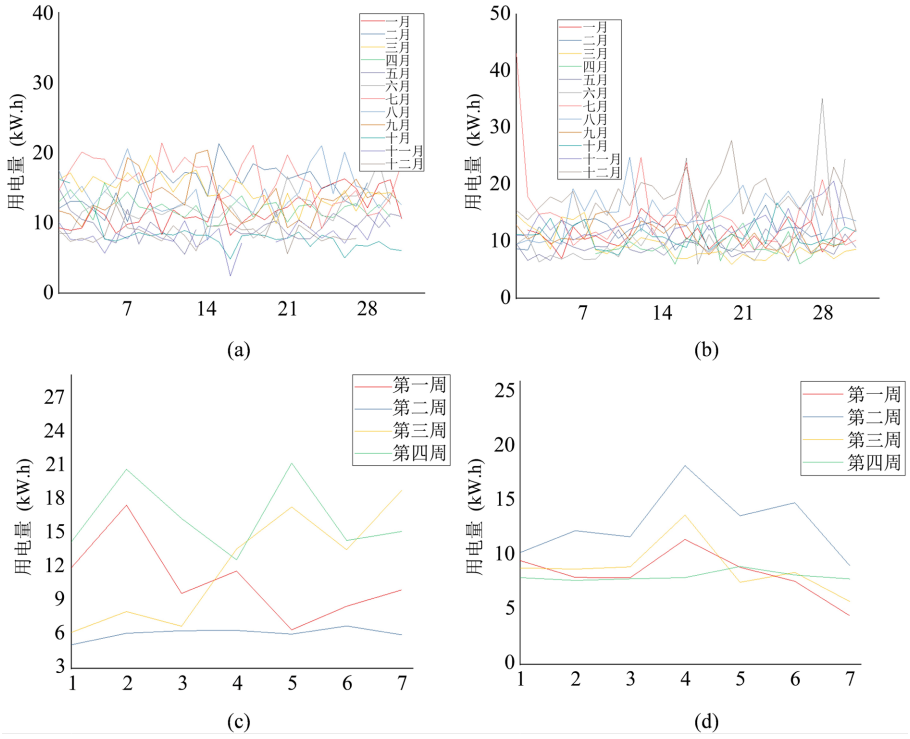

Figure 1. Power consumption description of users. (a) The monthly average power consumption of a normal power user in 2015; (b) Monthly average power consumption of a power stealing user in 2015; (c) Daily average power consumption of a normal user in June 2015; (d) Daily average power consumption of a power stealing user in June 2015

图 1. 用户的用电量说明。(a) 某正常用电用户 2015 年月平均用电量；(b) 某窃电用电用户 2015 年月平均用电量；(c) 某正常用电用户 2015 年 6 月日平均用电量；(d) 某窃电用电用户 2015 年 6 月日平均用电量

为了深入分析正常用户的周期性与窃电用户的非周期性特征，我们对用电数据进行了相关性分析。图 2 展示了两位用户在四周时间内用电量的皮尔逊相关系数(PCC)。图 2(a)呈现了正常用户的 PCC 值，而图 2(b)展示了窃电用户的 PCC 值。从图 2(a)可以看出，正常用户的用电数据通常表现出较强的正相关性，PCC 值大致在 0.7 左右，部分用户的 PCC 值甚至接近 0.9 (PCC 值接近 1 表明更强的正相关性[9])。与之相比，窃电用户的 PCC 值通常低于 0.4(如图 2(b))，甚至出现负相关现象，表明其用电数据之间存在反向关系。通过对正常用户与窃电用户用电数据的统计分析，可以发现，电力盗窃用户的用电模式通常不具有周期性或不规则性。因此，窃电检测的核心在于有效提取周期性特征，从而区分正常用电行为与异常用电行为。

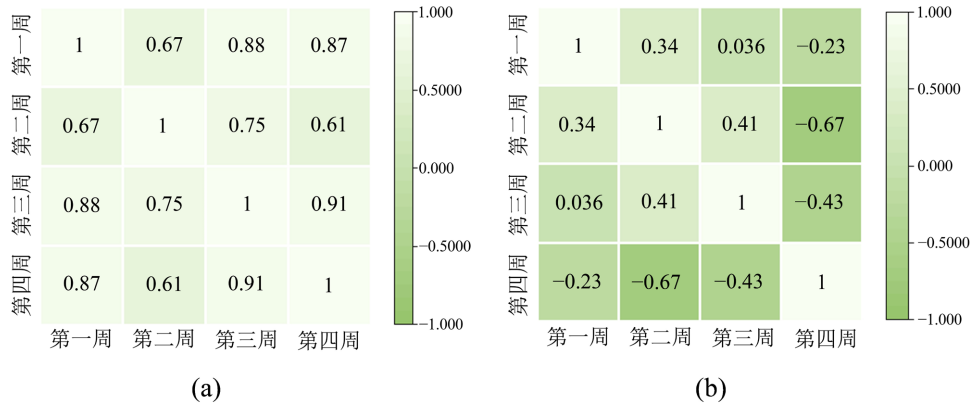

Figure 2. Pearson correlation coefficient (PCC) of weekly power consumption. (a) A normal power user; (b) A power stealing user
图 2. 每周用电量的皮尔逊相关系数(PCC)。(a) 某正常用电用户；(b) 某窃电用电用户

2.2. 数据预处理

电力数据丢失的原因多种多样，通常包括设备故障、数据传输问题以及人为操作失误。本文采用了一种基于二进制掩码的解决方案。具体而言，首先识别所有缺失数据的索引，并为每个缺失值创建一个二进制掩码，将缺失值标记为 1，非缺失值标记为 0。同时，在原始数据通道中，缺失数据位置被设为 0。最终，这两条通道(原始数据通道和二进制掩码通道)一起输入到二维卷积神经网络(CNN) [10]。

异常值的产生通常源于设备读取错误、数据录入问题或外部环境干扰等因素。在极端情况下，例如设备过载或数据记录不准确，也可能导致异常值的出现。根据文献[11]中提出的“三西格玛经验法则”，异常值可通过统计分析进行处理。首先计算每个样本的均值和标准差，然后依据“三西格玛经验法则”识别出超出三倍标准差范围的异常值，具体公式定义如下：

$$f\left(x_i\right)=\begin{cases} avg\left(x\right)+2\cdot std\left(x\right) & if\ x_i\ avg\left(x\right)+2\cdot std\left(x\right) \\ x_i & \text{otherwise} \end{cases} \tag{1}$$

其中，x 是由 x_i 逐日组成的向量，$avg(x)$是 x 的平均值，$std(x)$是 x 的标准差。通过这种方法可以有效地缓解异常值标准化操作改善了数据的数值分布，有助于提升模型的训练稳定性。通过这种调整，模型的收敛速度得以加快，整体算法效率也得到提升。同时，标准化有助于平衡不同特征之间的差异，避免某些特征因数值较大而对模型产生过度影响，从而提高了模型的泛化能力和抗噪性。在本文中我们选择最大最小归一化操作的缩放方法，根据下面的公式对数据进行标准化：

$$n\left(x\right)=\frac{x-\min\left(x\right)}{\max\left(x\right)-\min\left(x\right)} \tag{2}$$

其中 min(x)是整个数据集的最小值，max(x)整个数据集的最大值。

3. 方法

本文引入了一种结合三重生成对抗网络(Triple GAN)与胶囊网络(Capsule Network)的创新框架，以解决窃电检测中的数据不平衡和小样本学习问题。

3.1. 用于窃电检测的胶囊网络分类器

分类器采用了胶囊网络与卷积神经网络(CNN)相结合的混合架构，其总体设计如图 3 所示。由于一维的用电数据难以直接捕捉周期性特征，而用电行为往往受到周期性因素(如一周或一个月的模式)影响，因此将一维时间序列转换为二维格式有助于更好地挖掘数据的时序特征和邻域关系。计算公式如下：

$$N=\left[\frac{(L-P)}{S}+1\right] \tag{3}$$

每个用户的用电数据跨度为 L 天，相比其他转换方式，按周转换的效果最佳。因此，我们设置了参数 $P=7$ 和 $S=7$，将电力数据划分为多个子序列，并转化为二维结构，尺寸为 2 × 7 × 147。这种二维数据结构能更好地捕捉周期性用电模式，提升模型对异常时序特征的理解，减少噪声干扰，并为窃电检测提供更丰富的特征输入。

在特征提取阶段，输入数据依次通过大小为 1 × 1、3 × 3 和 5 × 5 的卷积核进行三次卷积操作，以捕捉时间序列数据中的多尺度局部模式并提取浅层特征。该过程为后续的胶囊网络(Capsule Network)生成高质量的初始特征表示，提升其对复杂时序模式和局部异常的建模与表达能力。每次卷积操作后，输出特征均经过批归一化(BN)和 ReLU 激活函数的非线性映射处理，最终输出的特征尺寸为 32 × 4 × 74 的特征图集 X。

为了使胶囊向量的模长更准确地反映特征的存在概率，需要应用一种非线性挤压函数(Squash Function)。该函数会保持向量方向不变，同时将向量模长压缩至(0, 1)区间内，使其更好地表示特征的存在性。其公式如下：

$$Squash(sj)=\frac{\left\|\mathbf{s}_j\right\|^2}{1+\left\|\mathbf{s}_j\right\|^2}\frac{\mathbf{s}_j}{\left\|\mathbf{s}_j\right\|} \tag{4}$$

经过压缩后的特征图集被整合为长度为 8 的向量，其中每 4 × 74 的区域包含 296 个胶囊单元。这些胶囊单元由 4 个特征图堆叠而成，总计 1184 个初级胶囊单元(4 × 4 × 74 个位置)，每个位置都代表一个初级胶囊的输出。

胶囊网络通过初级胶囊来表示电力用户的各种特征，每个初级胶囊对应一个浅层特征，并可对高级胶囊中的类别进行预测。首先，每个初级胶囊会基于其自身特征向量，生成相对于高级胶囊的预测向量。

$$U_{j|i}=w_{ij}u_i \tag{5}$$

其中 w_{ij} 为权重矩阵。随后，胶囊网络依据动态路由机制调整耦合系数 c_{ij}，确保高相关性的模式得到更高权重。最终，所有加权后的预测向量求和，得到高级胶囊的输入 v_j。

$$v_j=\sum_i c_{ij}\left(w_{ij}u_i\right) \tag{6}$$

从初级胶囊到高级胶囊的预测是通过全连接层实现的，每个预测向量与每个高级胶囊之间存在耦合系数 c_{ij}，对于每个初级胶囊，其与所有高级胶囊的耦合度之和为 1。

$$\sum_j c_{ij}=1 \tag{7}$$

它是一个概率分布，表示某一初级胶囊对特定类别的支持度，并满足归一化条件：

$$c_{ij} = \frac{exp\left(b_{ij}\right)}{\sum_k exp\left(b_{ik}\right)} \tag{8}$$

k 表示高级胶囊的类别数量。在初始化阶段，高级胶囊与预测向量之间的一致性被初始化为零向量，即 $b_{ij}=0$。在每次路由迭代过程中，b_{ij} 以累积的方式逐步更新，从而动态调整高级胶囊与预测向量之间的一致性关系。

$$b_{ij} \leftarrow b_{ij} + U_{j|i} \cdot \mathbf{v}_j \tag{9}$$

为了确保输出的稳定性并增强模型的判别能力，系统对高级胶囊的输入再次进行压缩。

$$\mathbf{v}_j = Squash\left(\mathbf{v}_j\right) \tag{10}$$

最后，模型的输出层生成两个高级胶囊($\mathbf{v}_0$ 和 $\mathbf{v}_1$)，每个胶囊分别表示一种用电行为，如正常用电和窃电行为。

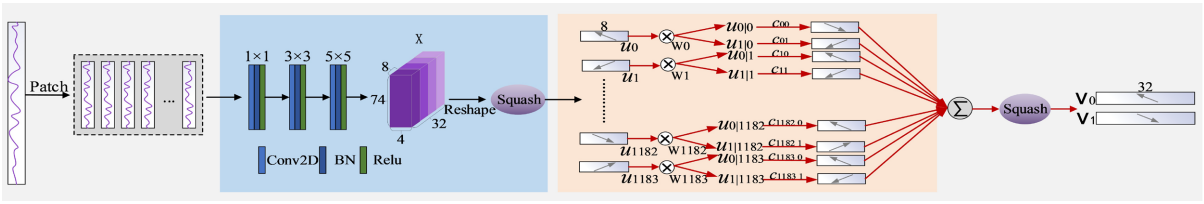

Figure 3. Classifier model
图 3. 分类器模型

3.2. 三重生成对抗网络模型架构

传统的三重生成对抗网络(TripleGAN)由生成器(G)、鉴别器(D)和分类器(C)三个模块组成如图 4 所示，生成器负责生成高质量样本，分类器用于提升分类性能，而鉴别器则用于区分真实样本与生成样本。然而，传统的 TripleGAN 在训练过程中仍存在一些局限性，例如生成器和分类器的优化可能无法完全同步，且对少数类样本的泛化能力有限。为了解决这些问题，本文在传统 TripleGAN 的基础上进行了改进，引入了协同优化机制和边界正则化策略。改进后的框架通过生成器、鉴别器和分类器的动态交互，结合边界正则化，进一步增强了模型对少数类样本的识别能力和泛化性能。具体训练流程分为以下三个阶段：

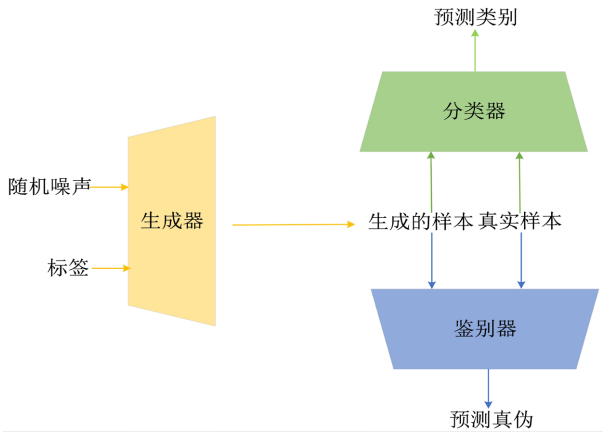

Figure 4. Traditional triple generation countermeasure network model architecture
图 4. 传统的三重生成对抗网络模型架构

1) 预训练阶段：生成器(G)和鉴别器(D)的优化遵循经典的 GAN 框架，其目标是通过对抗训练使生成样本的分布与真实样本分布尽可能接近。生成器通过生成逼真的样本来“欺骗”鉴别器，而鉴别器则学习区分真实样本和生成的伪造样本。在这一阶段，二者通过对抗训练不断提升各自的能力，因此通过优化其对抗损失函数来增强区分能力。对抗损失函数可以表示为以下形式：

$$V(D,G)=\mathbb{E}_{p(\mathbf{x},y)}\left[\log D(\mathbf{x},y)\right]+\mathbb{E}_{p_g(G(\mathbf{z},y),y)}\left[\log\left(1-D\left(G(\mathbf{z},y),y\right)\right)\right] \tag{11}$$

2) 分类器训练阶段：采用动态生成策略，在每次训练迭代中，仅根据当前所需的样本量，利用生成器动态生成等量的同类样本，确保每个批次中的正常样本与少数类样本的比例始终均衡，从而构建平衡的训练批次。按需生成的方式有效避免了传统方法中一次性生成大量样本所导致的内存浪费和过拟合风险，同时能够动态适应分类器的需求，显著提升其对少数类样本的识别能力。分类器的损失函数为：

$$V(C)=\mathbb{E}_{p(\mathbf{x},y)}\left[-\log pc(y\mid\mathbf{x})\right] \tag{12}$$

3) 协同训练阶段：生成器、鉴别器和分类器通过对抗与合作机制共同优化。生成器不仅需要生成真实可信的样本，还需特别关注生成有助于提升分类器在少数类识别上的表现的样本。鉴别器则通过优化对抗损失增强对生成样本与真实样本的区分能力，而分类器专注于基于这些样本进行准确分类，特别是少数类样本的识别。为了避免生成器和分类器在少数类样本生成与识别上的过度拟合，模型引入了边界正则化。该正则化技术控制生成样本与判别边界的距离，确保生成的样本不仅符合数据的真实分布，还能帮助分类器形成合理的决策边界，从而提升分类器的稳定性和识别能力。总的损失函数定义为：

$$\begin{aligned}&\min_{G,C}\max_{D}V(G,C,D)\\&=\mathbb{E}_{p(\mathbf{x},y)}\left[\log D(\mathbf{x},y)\right]+\mathbb{E}_{p_g(G(\mathbf{z},y),y)}\left[\log\left(1-D\left(G(\mathbf{z},y),y\right)\right)\right]+\mathbb{E}_{p(\mathbf{x},y)}\left[-\log pc(y\mid\mathbf{x})\right]+\\&\quad(1-\lambda)\mathbb{E}_{pg}\left(G(\mathbf{z},y),y\right)\left[-log\ pc\left(y\mid G(\mathbf{z},y)\right)\right]+\lambda R(G,C)\end{aligned} \tag{13}$$

在上式中通过最大化真实数据标签分布的概率 $log\ D(\mathbf{x},y)$ 鉴别器 D 学习识别真实样本；同时，生成器 G 通过最小化生成样本的识别误差 $log\left(1-D\left(G(\mathbf{z},y),y\right)\right)$ 来生成逼真的假样本，从而实现与鉴别器的对抗训练。分类器 C 则通过最小化真实样本标签的分类损失 $-log\ pc(y\mid\mathbf{x})$ 提升对真实数据的分类能力。参数 λ 是一个控制生成器和分类器贡献的权重系数。为了确保生成器、分类器和鉴别器能够收敛到唯一的最优解目标函数中引入了正则化项 $R(G,C)$用于约束生成器生成的样本分布，防止其过度集中在某些区域或引入噪声，同时鼓励生成样本分布靠近类别边界，进一步促进分类器的有效训练，其定义为：

$$R(G,C)=\begin{cases}\min\left(G_{mi}(x_g)-C_{ma}(x_g)\right) & 若x_g位于少数区域\\ 0 & 若x_g位于多数区域\end{cases} \tag{14}$$

当 x_g 在少数区域时，生成器会将样本分布调整至多数区域边界，使其少数得分 $C_{mi}(x_g)$接近多数得分 $C_{ma}(x_g)$，从而促进少数类区域扩展增加同类样本间的多样性。

4. 实验结果和分析

4.1. 实验建立

本实验使用 PyTorch 框架实现，数据集被随机分为训练集、验证集和测试集，分别占总数据量的 40%、20%和 40%，同时保持正常与窃电样本的比例不变。训练时使用批量大小为 64 的配置，并针对数据不平衡问题，每批次随机选择相同数量的正常和异常样本。优化器采用 Adam，以最小化二元交叉熵损失，训练共 200 个 epoch，第一阶段为前 50 个 epoch，学习率保持为 0.0001；第二阶段为接下来的 50 个 epoch，

学习率仍保持为 0.0001；第三阶段为最后 100 个 epoch，学习率从 0.0001 线性衰减至零。所有实验均在配置为 Intel(R) Core(TM) i5-1035G1 CPU、128GB 内存和 NVIDIA GeForce RTX 3090 Ti GPU 的服务器上进行，软件环境使用 Ubuntu 操作系统、PyTorch 1.10.0 深度学习框架和 Python 3.9 编译器。为了减少随机因素的影响，实验通过十次重复训练并报告结果的平均值和标准差。

4.2. 评估指标

为了评估该窃电检测模型的性能，使用以下三个指标包括 F1 分数(F1)，曲线下面积(AUC)和平均平均精度(MAP)。测量包括四个主要错误率，即假阳性(FP)、假阴性(FN)、真阳性(TP)和真阴性(TN)。具体介绍如下：

召回率是模型正确识别出的电力盗窃实例占所有真实电力盗窃案例的比例。

$$recall = \frac{TP}{TP + FN} \tag{15}$$

精度(Precision)是衡量模型预测准确性的指标，它表示在所有被模型判定为电力盗窃的实例中，真正属于电力盗窃案例的比例。

$$precision = \frac{TP}{TP + FP} \tag{16}$$

F1 分数用来衡量分类器模型精确率和召回率的调和平均数，从而综合衡量模型性能。

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \tag{17}$$

AUC 值，即曲线下面积，是衡量二分类模型性能的重要指标。它基于 ROC 曲线计算得出，取值范围在 0 到 1 之间，值越接近 1 表示分类器性能越好，与随机猜测相当时为 0.5，小于 0.5 则表示性能较差。

$$AUC = \frac{\sum_{i \in positiveClass} Rank_i - \frac{M(1+M)}{2}}{M \times N} \tag{18}$$

$Rank_i$ 表示样本 i 的秩值，M 为正常样本的个数，N 为异常样本的个数。

根据用户的概率预测值，对所有用户进行降序排序，计算前 100 名和前 200 名用户的平均精度(mAP)，分别记为 MAP100 和 MAP200。

$$MAP_{100} = \frac{1}{\sum_{i=1}^{100} y_i} \sum_{i=1}^{100} \left(\frac{y_i}{i} \sum_{j=1}^{i} y_j \right) \tag{19}$$

$$MAP_{200} = \frac{1}{\sum_{i=1}^{200} y_i} \sum_{i=1}^{200} \left(\frac{y_i}{i} \sum_{j=1}^{i} y_j \right) \tag{20}$$

4.3. 对比实验

在本文中，我们将提出的基于 CapsGAN 的方法与多种主流的机器学习(ML)和深度学习(DL)算法进行了全面的对比分析。首先，我们选取了几种具有代表性的机器学习和深度学习算法，如长短期记忆网络(LSTM)和随机森林(RF) [12]，以验证它们在窃电检测任务中的表现。接着，我们实现了几种经典的深度学习模型，包括宽深度卷积神经网络(Wide and Deep CNN)、混合阶表示学习(Hybrid-Order Representation) [13]和卷积自编码器(Convolutional Autoencoder) [14]。此外，我们还引入了最新的深度主动学习(Deep Active Learning, DAL) [15]和图卷积神经网络(Graph Convolutional Network，GCN)方法做进一步的对比。为了公平比较，所有方法均在预处理好的数据集上进行了实验，实验结果如表 1 所示。

Table 1. Performance comparison of different methods
表 1. 不同方法的性能比较

方法	F1	AUC	MAP100	MAP200
RF	0.575	0.801	0.715	0.698
LSTM	0.621	0.813	0.736	0.726
Wide and Deep CNN	0.679	0.827	0.759	0.737
Hybrid-Order Representation	0.733	0.831	0.783	0.755
Convolutional Autoencoder	0.836	0.896	0.825	0.792
Deep Active Learning	0.784	0.856	0.808	0.801
graph convolutional network	0.875	0.912	0.891	0.861
CapsGAN	**0.906**	**0.947**	**0.917**	**0.905**

表 1 总结了所有比较方法的性能。RF 在分类能力上表现较好，但并非专为窃电检测任务设计，因此存在一定局限性。LSTM 能够有效捕捉时间序列中的长时间依赖性，适用于窃电检测中的时序特征，但其计算复杂度较高，训练时间较长。Wide and Deep CNN 通过结合全局知识和局部模式，能较好地提取时间序列中的周期性特征，但由于卷积层和全连接层的限制，无法有效提取远距离的时序特征。Hybrid-Order Representation 通过结合多层次的信息提取策略增强模型表示能力，但其特征提取效果仍有提升空间，且在复杂模式建模上存在局限。Convolutional Autoencoder 擅长特征降维和数据压缩，在处理稀疏和高维数据时表现出色，但在时序特征的建模上不如 LSTM 等专门的时间序列模型。Deep Active Learning 通过智能选择样本进行训练，显著减少了标注数据的需求，但在复杂的窃电模式下，可能因过度依赖主动选择样本而影响准确性。Graph Convolutional Network 能够有效建模数据的结构信息，尤其适用于用户间关系等图形依赖的任务，但其计算资源要求较高，性能受限于图结构的准确构建。综合来看，CapsGAN 方法在 F1 评分、AUC (Area Under Curve)和 MAP 上均表现出色。CapsGAN 不仅能够与最优 CNN 模型相媲美，还在多个性能指标上超越了其他方法，特别适用于窃电检测的实际应用场景。

4.4. 消融实验

本文通过消融实验，逐步验证了 CapsGAN 框架中各个组成部分的有效性。实验结果如表 2 所示。首先，我们移除了生成器模块，将 CapsGAN 简化为仅包含分类器的网络，发现模型在 F1、AUC 和 MAP 等指标上出现大幅下降，特别是在 AUC 和 F1 分数方面。这一结果表明，生成器在解决了数据不平衡问题方面起到了关键作用，通过生成多样化的少数类样本，增强了分类器对窃电用户的识别能力。接着，我们移除了生成器(G)、分类器(C)和鉴别器(D)三者之间的协同训练机制。结果显示，模型在 F1、AUC、MAP100 和 MAP200 等评估指标上均出现显著下降，这证明了三者协同训练的策略在窃电检测任务中具有重要作用，能够有效提升生成样本的质量并增强分类器对正常样本与窃电样本的区分能力。

Table 2. Effect evaluation of CapsGAN variants
表 2. CapsGAN 各变体的效果评估

	生成器	分类器	协同训练	F1	AUC	MAP100	MAP200
(1)	×	√	×	0.763	0.832	0.875	0.846
(2)	√	√	×	0.817	0.894	0.893	0.875
(3)	√	√	√	0.906	0.947	0.917	0.905

5. 总结

本文提出了一种结合生成对抗网络(GAN)、胶囊网络(Capsule Network)和卷积神经网络(CNN)的窃电检测方法(CapsGAN)。该模型通过生成器与分类器的协同训练，共同生成少数类样本，并逐步扩展少数类的决策边界，从而显著提升窃电检测的分类性能。生成器通过生成逼真的窃电样本，帮助分类器更好地区分正常用户和窃电用户，特别是在样本稀缺和类别不平衡的情况下，CapsGAN 展示了其独特的优势。为了验证该方法的有效性，我们在公开的 SGCC 数据集上进行了实验，并与现有的主流窃电检测方法进行了对比。实验结果表明，CapsGAN 在 F1、AUC、MAP100 和 MAP200 等评估指标上均表现出色，尤其在少数类样本的识别能力上，CapsGAN 相较于现有的先进方法具有显著提升。这表明该模型不仅提高了窃电检测的准确性，也有效缓解了窃电检测中常见的数据不平衡问题。

参考文献

[1] Alahakoon, D. and Yu, X. (2016) Smart Electricity Meter Data Intelligence for Future Energy Systems: A Survey. *IEEE Transactions on Industrial Informatics*, **12**, 425-436. https://doi.org/10.1109/tii.2015.2414355

[2] Zheng, Z., Yang, Y., Niu, X., Dai, H. and Zhou, Y. (2018) Wide and Deep Convolutional Neural Networks for Electricity-Theft Detection to Secure Smart Grids. *IEEE Transactions on Industrial Informatics*, **14**, 1606-1615. https://doi.org/10.1109/tii.2017.2785963

[3] Eddin, M.E., Albaseer, A., Abdallah, M., Bayhan, S., Qaraqe, M.K., Al-Kuwari, S., et al. (2022) Fine-Tuned RNN-Based Detector for Electricity Theft Attacks in Smart Grid Generation Domain. *IEEE Open Journal of the Industrial Electronics Society*, **3**, 733-750. https://doi.org/10.1109/ojies.2022.3224784

[4] Gao, H., Kuenzel, S. and Zhang, X. (2022) A Hybrid ConvLSTM-Based Anomaly Detection Approach for Combating Energy Theft. *IEEE Transactions on Instrumentation and Measurement*, **71**, 1-10. https://doi.org/10.1109/tim.2022.3201569

[5] Ezhilarasan, D. and Bhavani, N.P.G. (2023) An Effective DNN Based ResNet Approach for Satellite Image Classification. 2023 4*th International Conference on Smart Electronics and Communication* (*ICOSEC*), Trichy, 20-22 September 2023, 1055-1062.

[6] Sabour, S., Frosst, N. and Hinton, G.E. (2017) Dynamic Routing Between Capsules. arXiv: 1710.09829.

[7] Choi, H., Jung, D., Kim, S. and Yoon, S. (2022) Imbalanced Data Classification via Cooperative Interaction between Classifier and Generator. *IEEE Transactions on Neural Networks and Learning Systems*, **33**, 3343-3356. https://doi.org/10.1109/tnnls.2021.3052243

[8] Nie, Y., Nguyen, N.H., Sinthong, P. and Kalagnanam, J. (2022) A Time Series Is Worth 64 Words: Long-Term Forecasting with Transformers. arXiv: 2211.14730.

[9] Kocaman, B. and Tümen, V. (2020) Detection of Electricity Theft Using Data Processing and LSTM Method in Distribution Systems. *Sādhanā*, **45**, Article No. 286. https://doi.org/10.1007/s12046-020-01512-0

[10] Finardi, P., Campiotti, I., Plensack, G., de Souza, R.D., Nogueira, R., Pinheiro, G. and Lotufo, R. (2020) Electricity Theft Detection with Self-Attention. arXiv: 2002.06219.

[11] Chandola, V., Banerjee, A. and Kumar, V. (2009) Anomaly Detection: A Survey. *ACM Computing Surveys*, **41**, 1-58. https://doi.org/10.1145/1541880.1541882

[12] Breiman, L. (2001) Random Forests. *Machine Learning*, **45**, 5-32. https://doi.org/10.1023/a:1010933404324

[13] Zhu, Y., Zhang, Y., Liu, L., Liu, Y., Li, G., Mao, M., et al. (2023) Hybrid-Order Representation Learning for Electricity Theft Detection. *IEEE Transactions on Industrial Informatics*, **19**, 1248-1259. https://doi.org/10.1109/tii.2022.3179243

[14] Cui, X., Liu, S., Lin, Z., Ma, J., Wen, F., Ding, Y., et al. (2022) Two-step Electricity Theft Detection Strategy Considering Economic Return Based on Convolutional Autoencoder and Improved Regression Algorithm. *IEEE Transactions on Power Systems*, **37**, 2346-2359. https://doi.org/10.1109/tpwrs.2021.3114307

[15] Zhu, L., Wen, W., Li, J., Zhang, C., Zhou, B. and Shuai, Z. (2024) Deep Active Learning-Enabled Cost-Effective Electricity Theft Detection in Smart Grids. *IEEE Transactions on Industrial Informatics*, **20**, 256-268. https://doi.org/10.1109/tii.2023.3249212