5 5G 通信技术应用于智慧城市建设应对策略

5.1 多方合作降低建设成本

政府可构建全方位政策支持体系，从资金扶持、资源整合、技术创新三个维度，系统性推进 5G 基础设施建设。在资金扶持方面，出台专项税收优惠政策，对参与 5G 建设的企业实施增值税即征即退、企业所得税 “三免三减半”，并设立 5G 产业发展专项基金，通过贴息贷款、财政补贴等方式，直接降低运营商建设成本。在资源整合层面，推广“政府引导 + 企业主导” 的共建共享模式，建立统一的 5G 基础设施资源管理平台，实现不同运营商间基站、传输线路等设施的动态共享。

5.2 加强网络安全防护

建立健全 5G 网络安全防护体系，是确保智慧城市稳定运行的关键。在技术应用层面，应采用端到端加密技术，对数据传输过程进行全链路保护，同时部署具备 AI 智能分析能力的入侵检测系统，通过机器学习算法实时识别异常流量和潜在攻击行为。防火墙需升级为下一代智能防火墙，实现基于应用层的访问控制与威胁阻断。

在安全管理机制方面，构建动态化的网络安全监测平台，结合大数据分析与态势感知技术，对网络安全事件进行 7×24 小时实时监测。建立多级预警响应机制，根据威胁等级自动触发相应处置流程，并通过应急演练不断优化事件响应速度，确保在遭受攻击时能快速隔离受影响区域，最大限度降低损失。

人才培养体系的完善同样不可或缺。一方面，联合高校与专业机构开设 5G 网络安全专项课程，培养掌握 5G 核心技术与安全防护技能的复合型人才；另一方面，定期组织行业安全攻防培训与竞赛，强化从业人员的实战能力。同时，建立完善的安全意识教育机制，通过定期开展网络安全科普与案例分析，提升全体人员的安全防范意识，从技术、管理、人员多维度筑牢 5G 网络安全防线，保障智慧城市建设稳步推进。

5.3 完善数据隐私保护机制

在数字化时代，数据隐私保护至关重要。我将从立法、技术、监管三方面深入扩写，补充具体案例、实施细节，增强说服力和实用性。

制定和完善相关的数据隐私保护法律法规迫在眉睫。政府部门应加快立法进程，针对智慧城市建设中数据全生命周期管理，制定专门的《智慧城市数据隐私保护条例》，细化数据收集环节的 " 最小必要原则 " 适用标准，明确规定未经用户明示同意，任何机构不得超范围收集生物特征、行踪轨迹等敏感数据。参考欧盟 GDPR 设立分级处罚机制，对违规企业处以年度全球营收 4% 的高额罚款，形成法律震慑。同时，建立跨部门的数据隐私执法协调机制，联合网信办、工信部等多部门开展专项治理行动。

在技术防护层面，需构建多层级的数据安全防护体系。传输过程中采用国密 SM4 算法对数据进行端到端加密，配合量子密钥分发技术，确保数据在 5G 网络中传输时的绝对安全；存储环节运用同态加密技术，实现数据在密文状态下的直接运算，避免明文泄露风险。通过联邦学习框架，在不泄露原始数据的前提下实现多方数据协同建模，例如医疗数据共享场景中，各医院可在加密状态下联合训练疾病预测模型。此外，引入差分隐私技术对公开数据进行匿名化处理，在发布统计信息时添加可控噪声，保障个体数据不可追溯性。

强化数据管理监管体系建设，要构建 " 事前 - 事中 - 事后 " 全流程监管机制。建立数据安全评估备案制度，要求所有接入智慧城市平台的企业提交数据安全风险评估报告，通过第三方机构认证后方可开展业务。在运行过程中，运用 AI 技术构建动态监测模型，实时分析数据流向，一旦发现异常访问行为（如短时间内高频次下载敏感数据）立即触发告警机制。定期开展数据安全审计工作，审计内容涵盖数据存储合规性、访问权限控制有效性等核心指标，审计结果向社会公开，并将违规记录纳入企业信用评价体系。同时，建立数据泄露应急响应标准流程，要求企业在发现数据泄露事件后 72 小时内完成上报，并启动损害评估和用户通知程序，最大限度降低数据泄露造成的负面影响。

6 结语

综上所述；5G 通信技术凭借其独特的优势，在智慧城市建设的各个领域展现出了巨大的应用潜力，为城市的智能化发展带来了革命性的变化。然而，在 5G 技术应用过程中，也面临着基础设施建设成本高、网络安全风险增加、数据隐私保护等诸多挑战。通过采取有效的应对策略，如多方合作降低建设成本、加强网络安全防护、完善数据隐私保护机制等，能够充分发挥 5G 技术的优势，推动智慧城市建设不断向前发展，构建更加智能、高效、安全、绿色的城市环境。

参考文献

[1] 刘洋.基于5G的物联网技术在智慧城市建设中的应用[J].智能建筑与智慧城市,2024,(10):47-49.

[2] 袁引.5G通信技术在智慧城市中的应用[J].电子技术,2024,53(04):396-397.

[3] 王琪.5G通信技术在智慧城市中的应用[J].数字技术与应用,2024,42(04):45-47.

 DOI: https://doi.org/10.12345/bdai.v6i6.34031

Research review on adversarial training—Three rights balance of robustness, generalization and computational cost

Hua Qi

Nanjing City Vocational College, Nanjing, Jinagsu, 211200, China

Abstract

Adversarial training enhances AI model robustness and is used in safety-critical areas like autonomous driving and medical diagnosis. However, the trade-off between robustness, generalization, and computational costs limits its implementation efficiency. This article explores these trade-offs using VC dimension theory, examines key parameters like generation methods and loss weights, and offers optimization insights, advancing adversarial defense theory and engineering.

Keywords

adversarial training; Robustness; Generalization; Calculate costs; Three dimensional trade-off

对抗训练研究综述——鲁棒性、泛化性与计算成本三维权衡

戚华

南京城市职业学院，中国·江苏 南京 211200

摘　要

对抗训练是提升AI模型抗攻击能力的关键手段，已被广泛应用于自动驾驶、医疗诊断等安全敏感场景中。但对抗训练中的鲁棒性提升、泛化性保持和计算成本控制的三维权衡问题制约了其工程落地效率。本文基于VC维理论揭示了鲁棒与泛化此消彼长的关系，深入探讨了生成方式、损失权重等关键参数，给出了模型优化依据，为对抗防御理论的深化与工程落地提供了参考。

关键词

对抗训练；鲁棒性；泛化性；计算成本；三维权衡

1 引言

1.1 研究背景与核心问题

随着深度学习在自动驾驶、医疗诊断等关键领域的不断应用，模型安全问题日益凸显，对抗样本就是一种典型威胁。2023 年美国某车企在自动驾驶测试中，将对抗贴纸贴在“STOP”标识上，导致识别置信度从 97% 降至 12%，导致了碰撞风险。当下，此类攻击越来越多地威胁到了自动驾驶、医疗诊断、面部识别等系统的可靠性上。因此，梳理对抗防御技术、明确对抗训练研究缺口，对推动机器学习安全发展至关重要。

【基金项目】南京市“十四五”规划“软件技术专业教学创新团队”项目；第五期江苏省职业教育教学改革研究课题（项目编号：ZYB560）。

【作者简介】戚华（1975–），女，高级工程师，硕士，从事软件工程、人工智能研究。

1.2 对抗训练研究阶段演进

对抗训练的发展经历了三个阶段。2015-2018 年为基础探索期，Goodfellow 首次引入 FGSM 对抗样本，Madry 提出 PGD 框架，使 CIFAR-10/ResNet-50 的 FGSM 攻击防御率从 58% 提升到 89%，奠定了理论基础；2019-2022 年为优化突破期，Zhang 等提出了梯度复用策略，降低了 40% 的训练耗时，但在未见过的攻击下鲁棒准确率却下降了 15-20%，泛化性问题仍然突出；2023-2025 年为跨场景适配期，周纯毅等提出了联邦对抗蒸馏，应用于肺癌诊断数据，实现了数据泄漏率 <0.1%、PGD 攻击准确率 89%，推动了医疗 AI 的安全发展，后来又拓展到了联邦学习和边缘设备方向。

2 对抗训练核心剖析

2.1 对抗样本攻击原理

对抗样本是在原始样本中添加微小扰动后得到的样本。其视觉或语义特征与原始样本没有显著差异，但能够使机器学习模型输出错误的预测结果。例如，在动物图像识别任务中，正常情况下模型预测图片为 Dog，但对抗样本攻击通过添加微小噪声，可以让图片看起来还是 Dog，但是预测结果

却是其他的动物。

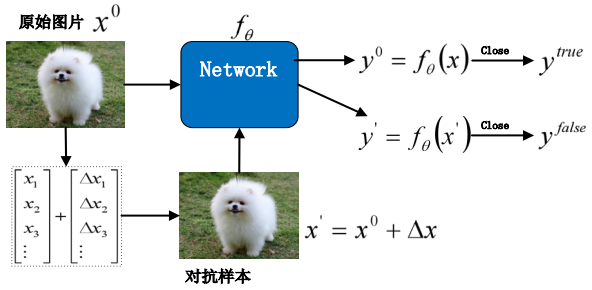

图 2–1 对抗样本攻击基本原理（图像识别场景）

2.2 对抗训练的理论基础：鲁棒性与泛化性的 trade–off

对抗训练的本质是通过“最小化原始样本损失 + 对抗样本损失”优化模型参数，其目标函数为：

$$\min_{\emptyset} \frac{1}{N}\sum_{i=1}^{N}\left[L(\theta, x_i, y_i) + \mu \cdot L(\theta, x_{adv,i}, y_i)\right]$$

其中：μ 为对抗损失权重，$x_{adv,i}$为 PGD 生成的对抗样本。

从 VC 维理论视角看，对抗训练的本质就是主动缩小模型的泛化边界。当模型在对抗样本分布上进行拟合时，会对原始样本分布的拟合能力产生约束，也就是模型需要牺牲部分对原始样本的精细拟合能力来换取对扰动样本的稳健识别能力，这一过程将必然导致鲁棒性与泛化性的平衡矛盾（trade-off）。

基于 FGSM 算法，可以通过指定目标分类使得网络针对任何输入图片均产生指定分类的对抗样本，即源 / 目标误分类，同时也可以不指定期待的分类，只需要使得生成的图片被网络识别为与正确分类不同的分类即可。如图 2-2 所示。

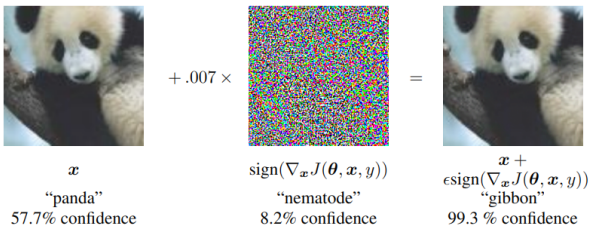

图 2–2 FGSM 生成演示

2018 年 Madry 等提出了 FGSM 的迭代优化版本：PGD 模型，该模型通过多步小扰动的方法提升攻击成功率。其迭代过程如下：

（1）初始化对抗样本$x_{adv}^0 = x + \delta$，其中 δ 为随机噪声（满足$\|\delta\|_\infty \le \epsilon$）；

（2）对 t=1,2,...,T（T 为迭代次数），计算梯度$\nabla_x J(\theta, x_{adv}^{t-1}, y)$，更新样本：

$$x_{adv}^t = \text{clip}(x_{adv}^{t-1} + \alpha \cdot \text{sign}(\nabla_x J), x, x-, x+\epsilon)$$

其中α 为步长（通常 α =ϵ/4），clip(·)函数确保扰动不超出ϵ范围；

（3）迭代 T 次后，输出x_{adv}^T作为最终对抗样本。

相比 FGSM，PGD 通过迭代优化能够让扰动更贴合模型的决策边界，进一步量化了基于 VC 维理论推导的鲁棒性与泛化性之间的 trade-off 约束关系。设原始样本泛化误差为，对对抗鲁棒误差为$R_{adv}(\theta)$，则存在常数C，使得$R_{adv}(\theta) \ge R(\theta) + C \cdot \epsilon$（其中，ϵ为对抗样本的扰动阈值）。

以 ImageNet 数据集与 ViT-B 模型为实验基准，当设置扰动阈值 ϵ=0.05 时，对抗训练的效果验证如下：

对抗鲁棒误差$R_{adv}(\theta)$：从优化前的 65% 降至 28%，鲁棒性显著提升；

原始样本泛化误差 R(θ)：从优化前的 5% 升至 8%，验证了鲁棒性提升需以少量泛化性为代价。

2.3 对抗训练效果的影响因素

消融实验是机器学习领域用于精准拆解关键变量影响的实验方法，其核心逻辑是通过固定多数条件、仅改变单一目标变量的控制变量设计，排除无关因素干扰，从而量化该变量对模型效果的独立作用。这一方法能有效地避免多变量混杂导致的结论模糊，是定位优化方向的关键手段。为了验证影响对抗训练效果的关键因素，我们以 CIFAR-10/ResNet-50 为实验基准，控制单一变量分析核心参数对对抗训练效果的影响，实验结果数据如下：

对抗样本生成方式	对抗损失权重	训练耗时（epoch）	原始样本准确率（%）	PGD 攻击准确率（%）	AutoAttack 准确率（%）
FGSM（1 步）	1.0	80	92.3	72.1	58.5
PGD（10 步）	1.0	80	90.1	89.2	74.3
PGD（10 步）	0.5	80	91.5	82.4	68.7
PGD（10 步）	1.5	80	88.7	90.5	75.1

分析数据得出以下结论：

（1）对抗样本生成方式可决定鲁棒泛化性：PGD（10 步）生成的对抗样本覆盖更广泛的扰动分布，因此模型对 AutoAttack 的鲁棒准确率显著高于 FGSM；

（2）对抗损失权重存在最优值：μ 过小（0.5）会导致鲁棒性不足，μ 过大（1.5）则会牺牲原始样本拟合精度。而当 μ=1.0 时，原始样本准确率与鲁棒准确率将达到最佳平衡；

（3）梯度复用策略可以有效地降低计算成本：采用梯度复用方法后，当 =1.0、PGD（10 步）生成对抗样本时，训练耗时从 80epoch 降至 48epoch，且各项性能指标损失均 <2%。

2.4 现存瓶颈与改进方向

2.4.1 计算成本高

在对抗训练中，PGD 方法能通过多步梯度迭代生成高强度对抗样本，但其大量重复计算、训练时间多出 2-3 倍的问题，严重制约了工程化落地进程。针对这一问题，2025 年，