DOI 10.15302/J-SSCAE-2021.02.013

深度学习与工业互联网安全：应用与挑战

杨晨，马瑞成，王雨石，翟岩龙，祝烈煌

（北京理工大学网络空间安全学院，北京 100081）

摘要：工业互联网安全是制造强国和网络强国建设的基石，深度学习因其具有表达能力强、适应性好、可移植性高等优点而可支持“智能自主式”工业互联网安全体系与方法构建，因此促进深度学习与工业互联网安全的融合创新具有鲜明价值。本文从产业宏观、安全技术、深度学习系统等角度全面分析了发展需求，从设备层、控制层、网络层、应用层、数据层的角度剖析了深度学习应用于工业互联网安全的发展现状；阐述了工业互联网深度学习应用在模型训练、模型预测方面的安全挑战，前瞻研判了未来研究的重点方向，如深度神经网络可解释性、样本收集和计算成本、样本集不均衡、模型结果可靠性、可用性与安全性平衡等。研究建议，在总体安全策略方面，深化促进两者的融合发展，建立动态的纵深防御体系；在技术攻关研究方面，采用应用驱动和前沿探索相结合的攻关方式，加快领域关键技术问题的攻关突破；在政策支持与引导方面，合理增加交叉领域的资源投入，建立“产学研”联合研发与应用的生态体系。

关键词：工业互联网安全；物联网安全；深度学习；数据安全

中图分类号：TP3　　**文献标识码：**A

Deep Learning and Industrial Internet Security: Application and Challenges

Yang Chen, Ma Ruicheng, Wang Yushi, Zhai Yanlong, Zhu Liehuang

(School of Cyberspace Science and Technology, Beijing Institute of Technology, Beijing 100081, China)

Abstract: Industrial Internet security is crucial for strengthening the manufacturing and network sectors of China. Deep learning, owing to its strong expression ability, good adaptability, and high portability, can support the establishment of an intelligent and autonomous industrial Internet security system and method. Therefore, it is of great value to promote the integrated innovation of deep learning and industrial Internet security. In this study, we analyze the development demand for industrial Internet security from the perspective of macro industrial environment, security technology, and deep learning system, and summarize the application status of deep learning to industrial Internet security in terms of device, control, network, application, and data layers. The security challenges faced by deep learning application to industrial Internet primarily lie in model training and prediction. Furthermore, we identify key research directions including interpretability of deep neural networks, cost control of sample collection and calculation, imbalance of sample sets, reliability of model results, and tradeoff between availability and security. Finally, some suggestions are proposed: a dynamic defense system in depth should be established in terms of overall security strategy; an application-driven and frontier exploration integrated method should be adopted to achieve breakthroughs regarding key technologies; and resources input should be raised for such interdisciplinary fields to establish an industry–university–research institute joint research ecosystem.

Keywords: industrial Internet security; Internet of Things security; deep learning; data security

收稿日期：2021-01-20；**修回日期：**2021-03-06

通讯作者：祝烈煌，北京理工大学网络空间安全学院教授，研究方向为网络与信息安全；E-mail: liehuangz@bit.edu.cn

资助项目：中国工程院咨询项目“新一代工业互联网安全技术发展战略研究”(2020-XZ-02)

本刊网址：www.engineering.org.cn/ch/journal/sscae

一、前言

工业互联网是新一代信息技术与制造业深度融合的新兴工业生态与应用模式，通过“人、机、物”的泛在可靠互联，连接生产全要素、全产业链、全价值链，推动制造业生产方式和企业形态变革。工业互联网安全是实现工业互联网高质量发展的前提条件。《加强工业互联网安全工作的指导意见》（2019 年）强调了工业互联网安全的重要价值，要求探索利用人工智能等新兴技术来提升安全防护水平。

深度学习具有较强的自动特征提取能力，为大数据时代的工业互联网安全（以应用场景复杂、数据规模庞大为特征）提供了更智能、更准确、更先进的分析工具 [1]：基于原始数据，使用一系列非线性处理层来学习不同抽象级别的数据表示；通过端到端的优化来定义、识别隐藏模式，提取高度非线性、极为复杂的特征；无需人工从领域知识中提取最佳特征，支持数据驱动的工业应用。也要注意到，深度学习的引入和应用，使得工业互联网系统更易面临恶意攻击或非法利用（见图 1），具有导致决策判断失准、造成工业制造损失的潜在风险 [2]。

深度学习应用于工业互联网，有关安全方面的研究开始出现，但依然缺乏较为完善的应用图景，且对深度学习系统自身安全问题关注较少 [3]。因此，本文针对这一空白领域展开前瞻研究，分析工业互联网安全需求，概括深度学习的具体应用，凝练新技术引入后面临的安全挑战，研判领域重点研究方向，以期为我国工业互联网安全发展提供策略参考。

二、工业互联网安全的需求分析

（一）工业互联网自身的安全需求

工业互联网安全是制造强国和网络强国建设的基石，关系到我国经济高质量发展。制造要素全面互联、接入开放的工业互联网网络，带来了规模和效率优势，也伴生了潜在安全问题：原本处于封闭状态的海量制造资源暴露于网络，面临更加开放的互联网环境，更容易被外部组织触达和发起恶意攻击；制造要素本身的计算资源有限、原生于封闭环境的防护能力普遍薄弱，易于被攻破和非法利用；鉴于工业系统普遍对可靠性、准确性、低时延等要求很高，即使网络化协同工业系统的单点被破坏，所造成的危害也可能很大。因此，工业互联网应用对安全保障提出了更高要求，需要利用诸如深度学习等先进技术来解决这些挑战。

从技术层面看，传统工业互联网的安全防护措

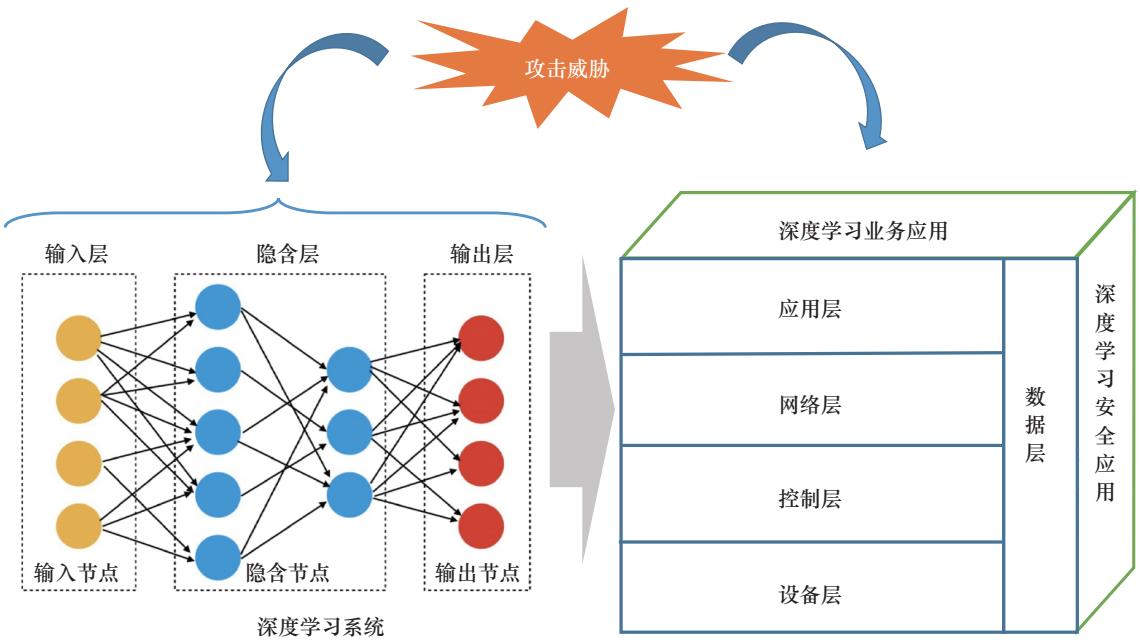

图 1　工业互联网安全面临深度学习引入的攻击威胁

施可以防御许多已知的安全威胁；但随着工业互联网应用领域的不断拓宽，接入设备数量与种类的不断增加，加之各类攻击方式的“推陈出新”，目前工业互联网攻击的数量、规模、速度、种类正在持续增加，现有的传统型安全防御工具和技术已难以全面应对这些新型攻击行为，亟需引入更加快速、高效、智能的安全防护新方法。深度学习的自学习能力强，在特征发现及自动分析方面具有优异性能，因此将之用于工业互联网设备、控制、网络、应用、数据等多个层次的安全防范，成为防护新型攻击形式的可行技术方向 [4]。

（二）工业互联网中深度学习系统的安全需求

深度学习技术能够广泛应用于工业互联网的五层体系架构、全生命周期各个阶段（见图 2），可显著减少人工操作、提高自动化水平与生产效率。例如，设备层采用有监督的深度学习，检测机器设备的使用情况与故障原因，与基于声纹的产品质量检测系统结合，实现质量检测自动化及智能化 [5]；应用层采用基于深度学习的图像识别技术进行视觉检测、分拣、定位等，提高流水线的效率和智能化水平；还有需求 / 销量预测、客户画像、供应链优化等可辅助企业进行决策的深度学习应用 [6]。

当前已有一些面向工业互联网安全的深度学习技术研究，如基于深度学习的入侵检测系统，可实现范围、速度、适应性等更优的恶意行为检测；基于深度学习的数据审计系统，可支撑从海量工业数据中提取关键信息，寻找威胁工业互联网安全的行为。随着这些深度学习应用的拓展和深入，深度学习系统自身存在的安全问题也引起了关注，如不防范这些安全问题，对可靠性、稳定性、可预测性等要求较高的工业互联网可能带来重大隐患。

三、工业互联网安全深度学习应用的发展现状

工业互联网安全可细分为设备、控制、网络、应用、数据等层次的安全 [7]，以下分别讨论各个层次的安全需求及深度学习应用。

（一）深度学习应用于设备层安全

工业互联网的设备安全包括设备身份鉴别与访问控制、固件安全保护等。深度学习对特征自动智能发现的能力、在二进制分析方面所具有的强大性能，为工业互联网中非加密设备的身份识别及固件代码分析提供了新思路。

工业互联网的开放性决定了大量非加密设备的接入导致相应设备易受身份欺骗攻击；攻击者会模仿合法设备的身份，在工业互联网中发送虚假信息

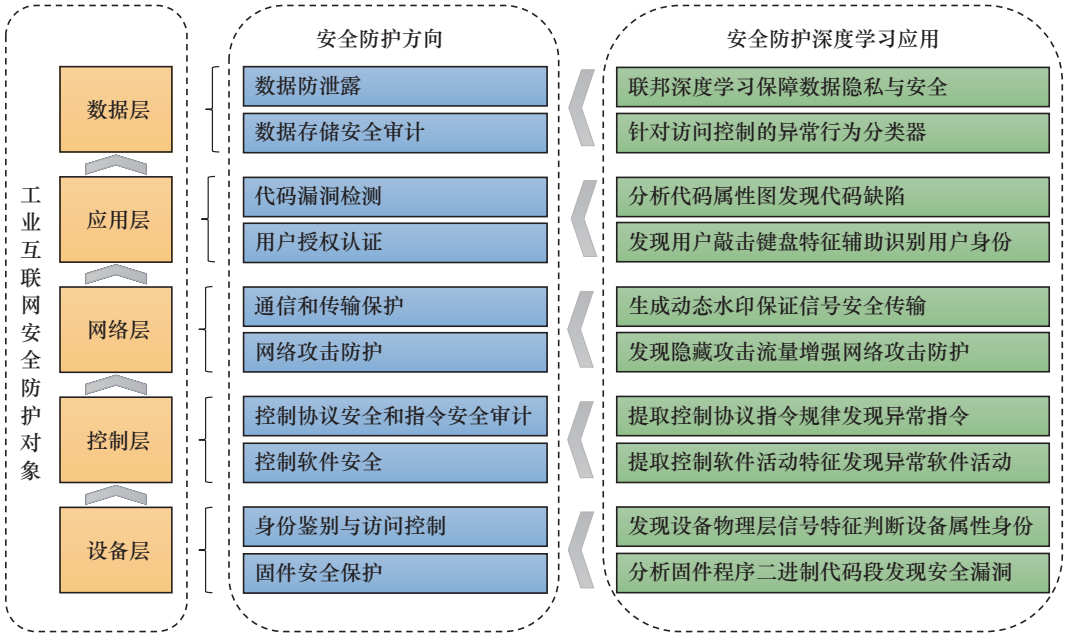

图 2　深度学习在工业互联网安全方面的应用分类

或进行其他恶意活动。这类攻击对关键工业设施而言非常危险，可能造成物理损坏。防止身份欺骗的传统方法是使用加密算法来验证设备身份，然而许多现有的工业互联网系统并未使用密码密钥操作。例如，全球航空领域广泛使用的 ADS-B 系统就未采用任何加密认证，对该系统进行密码安全改造将需要重大投资。接入工业互联网的设备在制造过程中会随机得到某些细微特征，这些特征会反映在设备产生的脉冲驱动信号中。一种可行的技术路径是利用自动编码器、卷积神经网络对接入工业互联网设备发出的物理层信号进行学习，在不知道设备发出信号具体特征的条件下建立对设备的辨识能力，进而判断设备的属性与身份，开展对所有已知设备的身份鉴别、对未知设备的情况报告 [8]，

工业互联网平台和固件众多，固件安全对工业互联网整体安全架构起着至关重要的作用。跨平台固件代码的二进制相似性分析是常用的设备固件漏洞安全检测方法，旨在检测来自不同平台的两段二进制函数是否相似。常规检测方法是近似图匹配算法，检测速度慢，如果仅存在几个指令不同的微小差异则会发生误判，在对速度和安全性要求很高的工业互联网领域难以应用。深度神经网络可以将二进制代码函数段的图嵌入表示为一个神经网络，通过对两个相似二进制代码函数的图嵌入接近程度进行比对，即可准确高效地开展二进制相似度分析；比传统检测速度提高 3~4 个数量级，能够克服传统方法的误判问题 [9]。因此深度学习技术可用于二进制代码段的相似性推断、漏洞检测，有效支持固件安全分析工作。

（二）深度学习应用于控制层安全

工业互联网的控制系统向上接入网络层、向下连接海量工业设备，其安全防护措施极为重要。工业互联网控制安全包括控制协议安全机制、指令安全审计、控制软件安全加固等。利用深度学习的自动特征发现能力，为控制协议指令攻击检测、控制软件检测提供了新思路。

工业互联网的控制系统分为过程控制子系统、监控与数据采集子系统、分布式控制子系统、现场总线控制子系统等。这些子系统都是利用控制协议进行控制指令下发，而针对控制协议的攻击较多通过在协议传递的控制指令中注入错误数据来实现。常规的指令攻击检测方式是分析攻击消息的异常规律，发现相似攻击行为；但在攻击方式不断更新的工业互联网环境下，这种检测方法并不能可靠地发现新的攻击形式。基于深度神经网络的特征发现能力，有望解决这一问题：深度神经网络从过程控制装置获取的传感器和执行器信号中学习正常控制协议下的通信规律，进行控制协议和指令的安全检测；既可以检测已知的指令攻击，还能识别新的攻击形式 [10]。

工业互联网的控制软件面临恶意软件注入等安全威胁，常见的恶意软件样本是精心制作的计算机程序片段，意图在不被发现的前提下对受感染工业互联网资产进行控制和监视。传统的恶意软件检测方式是人工发现恶意软件攻击特征，利用已知特征进行软件检测；但涉及多态性蠕虫或病毒检测时，这种方法不再可行。当前，诸多反病毒软件供应商对增强恶意软件检测能力的深度学习方法开展了深入研究，在实际测试中取得了很好的效果 [11]。因此，在工业互联网控制层中引入深度学习技术，发挥其对特征自动提取的固有优势，动态分析工业互联网控制软件活动的特征；持续分析软件活动情况、软件执行某些特定命令的活动情况，检测控制软件的行为，提高控制软件抵御恶意软件注入等安全威胁的能力 [12]。

（三）深度学习应用于网络层安全

工业互联网的网络层安全包括通信与传输保护、网络攻击防护等。利用深度学习的特征提取能力、自学能力、信息压缩能力，为工业互联网的通信数据加密、网络入侵检测提供新思路。

工业互联网包含数量众多的传感器、终端、控制、计算、存储等设备，设备之间需要实时、可靠、安全地传输来自周围环境、自身状态、控制指令等各种信息。尤其在资源受限的工业互联网终端节点，因其组成相对简单、计算和存储能力较弱，数据的安全传输是重大挑战。依靠加密算法的传统传输方式可靠性较高，但攻击检测的复杂度、延迟均比较高，不适合在通信低延迟、组成复杂的工业互联网环境进行大规模部署。因此，考虑基于工业互联网信号的深度学习框架，采用长短期记忆模块（LSTM）从工业互联网信号中提取随机特征（如谱平坦度、偏度、峰度、中心矩等），将之转换为水

印并加载在原始信号中，利用云计算或边缘计算节点验证水印信息以保证信号的可靠性，据此完成针对工业互联网的网络攻击行为检测 [13]。

工业互联网因其复杂性、敏感性而易受各种针对性的网络攻击，需要配置入侵检测系统来扫描网络流量活动、识别恶意或异常行为。传统的入侵检测系统通常采用（浅层）机器学习技术，无法有效解决具有实时性要求、来自环境的海量数据入侵分类检测问题。深度学习是十分理想的隐藏流量发现手段，可用于区分攻击流量和检测正常流量。例如，使用双向长短期记忆递归神经网络（BLSTM-RNN）方法，详细学习异常入侵所具有的网络流量特征，快速准确地识别针对工业互联网的网络攻击和网络欺诈等异常活动 [14]。

（四）深度学习应用于应用层安全

工业互联网的应用层安全包括用户授权认证、代码安全等。利用深度学习在“理解”自然语言、特征提取等方面的独特优势，为工业互联网的代码安全分析、用户授权认证提供新思路。

工业互联网的构成和功能复杂，涉及软件众多，对软件源代码的安全性提出了很高要求。传统的代码漏洞检测较多依赖分析人员对代码的人工分析、对安全问题的认识和经验积累，这一模式很难满足工业互联网的代码漏洞分析需求。一种可行的思路是借鉴自然语言处理方法，利用深度学习在“理解”自然语言方面的独特优势、LSTM 对自然语言上下文的“记忆”功能，对由源代码的抽象语法树、控制 / 数据流图、程序依赖图等构成的代码属性进行理解与分析，在源代码编程阶段及时发现并修正代码缺陷，主动完成代码漏洞分析检测 [15]。

工业互联网覆盖面广，对安全性和隐私性要求高，涉及大量用户授权认证过程。传统上基于密码和个人识别码的认证系统虽然有效，但不足以抵御多类恶意攻击行为。因此，利用深度学习在生物特征发现的优势发展形成的人脸识别等技术，已成功应用于应用层安全 [16]，起到配合传统认证系统、提高用户授权认证能力的作用。此外，为了有效提升用户授权认证的安全性、降低认证成本，有研究者提出了在键盘端利用深度学习技术提取用户每次敲击键盘的时间、键入时施加的压力以及移动设备、触摸面积、触摸位置等特征信息，辅助进行用户身份判断 [17]。这一方案为提升工业互联网用户授权认证能力提供了新途径。

（五）深度学习应用于数据层安全

工业互联网数据安全的主要工作之一是数据防泄露。在包含大量碎片化数据的工业互联网中，减少不必要的跨地域、跨组织的原始数据共享和流动，是提高数据安全性的重要方向，而这也是联邦深度学习技术的优势所在。在联邦深度学习系统中，自有数据不出本地，通过加密机制进行参数交换，在不违反数据隐私保护法规的情况下建立虚拟的共有模型。关于数据安全审计，敏感度低的工业互联网数据可以存储在成本价格相对低的工业大数据云平台，采用基于深度学习的数据安全审计机制来监管数据的访问等行为，防止数据被窃取、篡改、破坏，实现数据存储安全。

工业互联网存在传感器、边缘计算节点、云端、用户端等多点通信需求。传统的数据处理流程是数据产生于传感器端，初步采集的数据会先存储在相关软件中 [18]。数据入侵、非法访问较多隐藏在合理的授权之下，不容易被发现。使用无监督深度学习对异常行为进行分类，确保数据不受到窃取、篡改、破坏。感知数据会反馈至边缘计算节点，经清洗、预处理分析后，再上传至云端并经进一步加工处理供用户调用。敏感性、碎片化、海量数据流动十分不利于工业互联网环境下的数据安全保护，因此针对工业互联网数据的安全多方计算需求，可行的解决方案是引入联邦深度学习技术，在不直接共享敏感数据的前提下开展数据处理，最大限度地减少数据流动和不必要的数据传输，确保工业互联网的数据安全 [19]。

四、工业互联网安全深度学习应用面临的挑战

深度学习技术在赋予工业互联网安全新前景的同时，可能存在被攻击者利用的漏洞，可能受到高级可持续威胁攻击。例如，攻击者可以针对性地修改恶意文件来绕过基于深度学习的检测工具，加入一些不易察觉的噪音使得工厂语音控制系统被恶意调用，在交通指示牌或其他车辆上贴一些小标志使得基于深度学习的自动驾驶系统出现误判。在高价

值或高风险的工业生产过程中，如果深度学习系统被恶意攻击，可能会造成设备损坏，甚至威胁人员生命安全。针对深度学习的攻击分为 5 种（见图 3）：投毒攻击、模型逆向攻击、模型提取攻击、物理攻击、对抗性攻击，主要发生在模型训练阶段和模型预测阶段。

（一）模型训练阶段

投毒攻击指通过攻击训练数据集，使得模型无法正常工作。在工业互联网中，竞争对手可能通过篡改传感器的测量值来操纵训练数据。对于基于深度学习的故障检测器来说，微小的数据篡改可能会导致有针对性的错误分类或不良行为。后门攻击也是一种投毒攻击，在训练数据过程中添加特殊标志（后门触发器）来绕过模型的分类，如向标注为非恶意的文件中加入一段特殊的代码（文字），将之用于训练深度学习模型。训练好的模型能够正常识别恶意文件，但是当检测到带有这段特殊代码的恶意文件时，模型将会把它识别为非恶意的，从而绕过检测。这种攻击大多数时间不影响模型正常工作，极为隐蔽 [20]。

模型逆向攻击发生在训练完成阶段，可以通过模型的输出（黑盒攻击）、模型参数（白盒攻击）将训练数据集信息从模型中逆向提取出来；换言之，通过已经训练好的模型数据，还原出模型的训练数据成员。在工业互联网中数据是宝贵资源，特别是涉及到商业价值的敏感数据。例如，产品质量检测模型的训练需要的产品参数（如重量、尺寸型号等），入侵检测系统需要的工业生产系统中传感器数据，这些都涉及到产品隐私，具有一定的商业价值 [21]。

（二）模型预测阶段

不同于数字样本攻击，物理攻击属于实体样本攻击，通过在现实生活中改变目标物体的形态或者贴上特殊标记来欺骗深度学习模型。物理攻击不需要对模型的训练数据“做手脚”，只需通过一定次数的模型功能测试，就能发现模型的漏洞和缺陷，进而设计实现物理攻击。例如，自动驾驶汽车的视觉系统能够使用深度学习技术对道路上的行人、车辆、道路标志等进行分类，但在道路标志上粘贴精心设计的纸条后，视觉系统便无法正确识别该道路标志。对于基于深度学习的人脸识别系统而言，带有特殊标记的眼镜便能干扰其正常工作；即使在相对稳定的物理条件下，只需针对性调整姿势、距离、光线，也能使人脸识别系统发生识别错误。工业互联网中有很多人脸识别、产品质量检测的深度学习类应用，如果内部人员具有恶意，则这种物理攻击将比较隐蔽且威胁明显 [22]。

模型提取攻击指通过公开的应用程序接口（API）来模拟功能类似甚至相同的模型，具体参数很难被掌握，且攻击目的是复刻模型而不是还原数据成员 [23]。训练 1 个模型通常需要 20~30 d，较为复杂的模型甚至需要更长的时间，一些应用于工业互联网的模型具有一定的商业价值，如异常行为分类系统。这些模型具有一定的可移植性，如果将

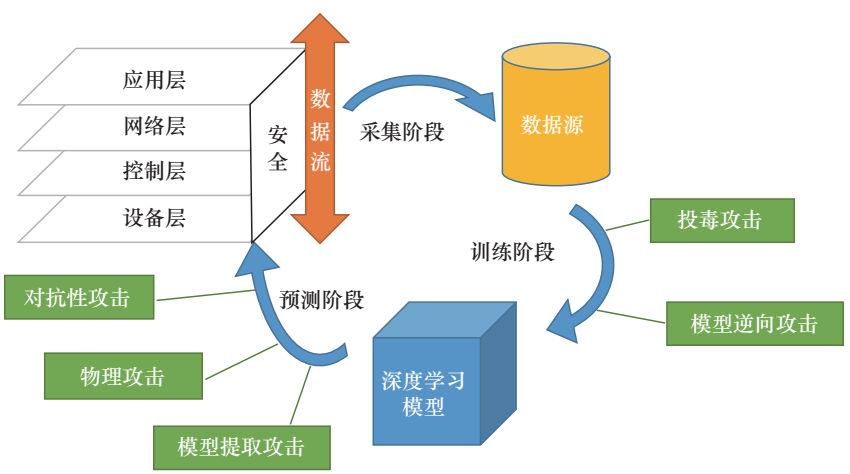

图 3　工业互联网中深度学习系统面临的安全挑战

之公开在网络上，即使只提供 API 接口，不法分子也能通过随机组合的输入来获取输入与输出关系，从而复刻功能相同的模型，使得公司知识产权利益受到损失。

对抗性攻击又称为躲避攻击，指在正常样本中加入了一些人眼难以察觉的干扰，从而造成模型预测错误；分为无特定目标攻击、特定目标攻击，前者只是干扰模型的正确判断，后者需要模型将特定的输入判断为指定的一种输出。目前，深度学习模型对于对抗性攻击是比较脆弱的，较为轻微的扰动就可以干扰到模型的正常工作。例如，工业互联网的产品检测模型易受无特定目标的对抗性攻击，攻击者只需在产品图片上加入肉眼不可见的噪声点，就可以使得模型失去判断能力，严重时甚至可以破坏整个工业生产流程。对于面向安全检测的深度学习系统，攻击者也可以通过添加一些特殊语句来绕过安全检测模型，从而干扰工业系统的安全运行 [24]。

五、工业互联网安全深度学习应用的未来研究方向

（一）深度神经网络的可解释性

人类思维很难理解深度神经网络的决策依据，这是因为深度神经网络通常被当成“黑盒”模型使用，每个神经元都是由上一层的线性组合再叠加 1 个非线性函数得到的，具有高度非线性的特征。对于工业互联网安全应用，除了模型输出的最终结果外，人们还应知道模型是基于哪些因素考量得出结论的；如果模型不可解释，则意味着模型本身是不可知、不安全的。因此，只有确保信息可靠性（如没有受到投毒攻击、对抗性攻击等）、明晰模型输入输出的因果关系，模型的预测结果才能令人信服，也才能交由深度学习来承担工业互联网安全体系中的核心任务。

（二）样本收集和计算成本

随着深度学习方法的发展，神经网络层数越来越深，所需的训练样例数目、算力要求（电力消耗）也在迅猛增加。即使深度学习模型相比于传统方法具有更好的效果，但提升效率带来的收益甚至可能无法弥补增加成本，这将直接制约深度学习技术在工业互联网安全中的推广应用。工业互联网安全的应用场景多样，需要针对性地收集数量可观的数据并加以手工标注，人力成本较高；深度神经网络规模庞大，为达到精度、实时性等要求，需要高性能计算系统的支持，带来较高能耗需求。因此，需要研究更高效、自动化的数据集构建方法，更低功耗的深度学习模型与计算系统。

（三）样本集不均衡

深度学习在消费互联网应用方面体现出了优势，但工业互联网的应用领域及场景千变万化，难以为深度学习模型提供足够多的样本量。因此，需要研究通过自动化工具增加样本量的方法，基于小样本的深度学习方法。面对碎片化、复杂多变的工业互联网安全应用及场景，构建具有均衡性、可全面反映数据真实分布的样本集，将之用于训练深度学习模型仍是挑战。目前已有过采样、欠样本等方法来缓解深度学习中样本不均衡的问题，但依然缺乏实际可用的系统性研究成果 [25]。

（四）模型结果的可靠性

工业互联网涉及领域众多，框架构成复杂，对系统的整体可靠性要求很高。例如，航空、航天类飞行器的零部件生产，要求设备达到可靠性不低于 99.999%；如果重要节点发生故障，会造成批次性的产品损坏或性能降级。在实际工业生产过程中，模型的稳定性要比表达能力更为重要，一旦某个生产环节出现问题，可能影响整条生产线的运转；很多深度学习模型的预测准确率不足 90%，几乎无法移植到对可靠性要求极高的工业互联网应用。因此，研究提高深度学习技术模型准确度、确定性、可靠性的方法显得尤为重要。

（五）平衡可用性与安全性

深度学习应用自身也存在安全性和隐私性的问题。深度学习模型的训练需要大量数据样本，在公开模型以实现商业价值的同时，保护模型与训练数据不被非法窃取和使用是值得关注的课题。对于工业生产而言，模型的安全性非常重要。有学者提出了通过差分隐私、同态加密方法保护模型隐私的办法，通过对抗性训练来侦测对抗性，进而提高模型的安全性；但这些方法在一定程度上降低了模型的

可用性，影响了模型的性能表现。因此，未来需要研究深度学习模型可用性与安全性的平衡措施。

六、对策建议

（一）完善工业互联网总体安全策略

建议在工业互联网总体安全策略中纳入深度学习技术方面的内容，构建可覆盖安全业务全生命周期、以主动智能响应为核心特征的工业互联网纵深安全防御体系。深度学习应体现在整个工业互联网安全架构中，据此连接工业互联网的各个层次，建立对安全事件“预警、监测、处置、防护”的动态防御体系，系统综合地维护工业互联网安全。

（二）攻克深度学习应用的重大问题

开展深度学习在工业互联网安全方面的应用，仍然存在一些亟待解决的关键技术问题，建议计算机、神经科学、自动化等学科领域的研究人员共同努力，协同开展应用突破，瞄准国际领先的发展目标来构建工业互联网安全生态。前瞻论证交叉领域创新性研究的重点方向，通过示范效应带动整个技术链的深化拓展。立足工业生产的实际场景和迫切需求，采取应用与问题联合驱动的模式，稳步攻关两者融合中存在的关键技术瓶颈。

（三）合理保障深度学习与工业互联网安全交叉融合领域的资源投入

深度学习在工业互联网安全领域的应用前景广阔、潜在价值显著，应合理增加在深度学习与工业互联网交叉融合方向的人、财、物投入。建议加强管理政策或行业性规划研究，鼓励科研人员自主联合，深化工业企业、高校、科研机构的三方合作关系，形成“政、产、学、研”合作体系，更好完善深度学习技术体系及其与工业互联网安全的融合应用；在实践中验证技术以凸显实效，与科学研究形成相互促进的新发展格局。

参考文献

[1] 李瑞琪, 韦莎, 程雨航, 等. 人工智能技术在智能制造中的典型应用场景与标准体系研究 [J]. 中国工程科学, 2018, 20(4): 112–117.
Li R Q, Wei S, Cheng Y H, et al. Research on typical application scenarios and standard system of artificial intelligence technology in intelligent manufacturing [J]. Strategic Study of CAE, 2018, 20(4): 112–117.

[2] Li J H. Cyber security meets artificial intelligence: A survey [J]. Frontiers of Information Technology & Electronic Engineering, 2018, 19(12): 1462–1474.

[3] Amanullah M A, Habeeb R A A, Nasaruddin F H, et al. Deep learning and big data technologies for IoT security [J]. Computer Communications, 2020, 151(1): 495–517.

[4] Ha T, Dang T K, Le H, et al. Security and privacy issues in deep learning: A brief review [J]. SN Computer Science, 2020, 1(5): 1–15.

[5] Tsai S Y, Chang J Y. Parametric study and design of deep learning on leveling system for smart manufacturing [C]. Hsinchu: 2018 IEEE International Conference on Smart Manufacturing, Industrial & Logistics Engineering (SMILE), 2018.

[6] Wang J J, Ma Y L, Zhang L B, et al. Deep learning for smart manufacturing: Methods and applications [J]. Journal of Manufacturing Systems, 2018, 48(C): 144–156.

[7] 余晓晖, 刘默, 蒋昕昊, 等. 工业互联网体系架构2.0 [J]. 计算机集成制造系统, 2019, 25(12): 2983–2996.
Yu X H, Liu M, Jiang X H, et al. Industrial Internet architecture 2.0 [J]. Computer Integrated Manufacturing Systems, 2019, 25(12): 2983–2996.

[8] Liu Y X, Wang J, Li J Q, et al. Zero-bias deep learning for accurate identification of Internet of things (IoT) devices [J]. IEEE Internet of Things Journal, 2020, 11(4): 2627–2634.

[9] Xu X J, Liu C, Feng Q, et al. Neural network-based graph embedding for cross-platform binary code similarity detection [C]. Dallas: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, 2017.

[10] Potluri S, Diedrich S. Deep learning based efficient anomaly detection for securing process control systems against injection attacks [C]. Vancouver: 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE), 2019.

[11] Kyadige A, Ethan M, Rudd B K, et al. Learning from context: A multi-view deep learning architecture for malware detection [C]. San Francisco: 2020 IEEE Security and Privacy Workshops (SPW), 2020.

[12] Kozik R. Distributing extreme learning machines with apache spark for NetFlow-based malware activity detection [J]. Pattern Recognition Letters, 2018, 101: 14–20.

[13] Ferdowsi A, Saad W. Deep Learning-based dynamic watermarking for secure signal authentication in the Internet of things [C]. Kansas City: 2018 IEEE International Conference on Communications (ICC), 2018.

[14] Roy B, Cheung H. A deep learning approach for intrusion detection in Internet of things using bi-directional long short-term memory recurrent neural network [C]. Sydney: 2018 28th International Telecommunication Networks and Applications Conference, 2018.

[15] Wang X M, Zhang T, Wu R P, et al. CPGVA: Code property graph based vulnerability analysis by deep learning [C]. Stockholm: 2018 10th International Conference on Advanced Infocomm Technology (ICAIT), 2018.

[16] Masi I, Wu Y, Hassner T, et al. Deep face recognition: A survey [C].

Parana: 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2018.

[17] Bernardi M, Cimitile M, Martinelli F, et al. Keystroke analysis for user identification using deep neural networks [C]. Budapest: 2019 International Joint Conference on Neural Networks (IJCNN), 2019.

[18] Yang C, Shen W M, Wang X B. The Internet of things in manufacturing: Key issues and potential applications [J]. IEEE Systems, Man, and Cybernetics Magazine, 2018, 4(1): 6–15.

[19] Yin B, Yin H, Wu Y L, et al. FDC: A secure federated deep learning mechanism for data collaborations in the Internet of things [J]. IEEE Internet of Things Journal, 2020, 7(7): 6348–6359.

[20] Saha A, Subramany A, Pirsiavash H. Hidden trigger backdoor attacks [EB/OL]. (2020-07-15)[2020-12-15]. https://www.csee.umbc.edu/~hpirsiav/papers/hidden_aaai20.pdf.

[21] Hidano S, Murakami T, Katsumata S, et al. Exposing private user behaviors of collaborative filtering via model inversion techniques [J]. Proceedings on Privacy Enhancing Technologies, 2020 (3): 264–283.

[22] Boloor A, He X, Gill C, et al. Simple physical adversarial examples against end-to-end autonomous driving models [C]. Las Vegas: 2019 IEEE International Conference on Embedded Software and Systems (ICESS), 2019.

[23] Shafique M, Naseer M, Theocharides T, et al. Robust machine learning systems: Challenges current trends perspectives and the road ahead [J]. Design & Test IEEE, 2020, 37(2): 30–57.

[24] Wan M, Han M, Li L, et al. Effects of and defenses against adversarial attacks on a traffic light classification CNN [C]. New York: Proceedings of the 2020 ACM Southeast Conference, 2020.

[25] Buda M, Maki A, Mazurowski M A. A systematic study of the class imbalance problem in convolutional neural networks [J]. Neural Networks, 2017, 106: 249–259.