



面向通信的机器学习：通向智能传输与处理之路

王世雄, 李烨

摘要

在人工智能与大数据时代到来之前,无线通信研究主要遵循传统路线,包含问题分析、模型构建与校准、算法设计与调优,以及整体测试和经验验证等环节。然而,在处理大规模复杂问题和管理动态海量数据时,这种方法往往存在局限性,导致传统通信系统和技术效率低下、性能受限。因此,借助人工智能和机器学习的革命浪潮,无线通信领域开发出高度自适应、更高效且更智能的系统和算法。这一技术革新为信息传输与处理的智能化铺平了道路。本文将探讨机器学习在智能无线通信中的典型作用,及其特点、挑战与实际考虑因素。

关键词

机器学习, 智能传输, 智能处理

1 引言

自 19 世纪以来，无线电通信开启了人类社会信息传输的新纪元。在摩尔斯电码和电报机等早期无线电时代，传输技术严重依赖人工操作，限制了信息交换的效率和可靠性。为了实现信息传输与处理的自动化，第一代“智能传输与处理”概念应运而生。到了 20 世纪，模块化通信系统取得了重大发展。模块化通信系统由信源编码、信道编码、调制、发射波束赋形、无线传输、接收波束赋形、解调、信号检测、信道解码和信源解码等基本模块组成 [1-3]。从方法论上来看，模块化无线通信与信号处理遵循系统化研究路径，包括问题分析、模型开发与校准、算法设计与调优以及经验验证、反馈与改进，每一步都需要人类投入大量的脑力来完成。

进入 21 世纪后，无线通信系统需传输音频、视频、文本等各种形式的海量数据，并需保证低时延、高速率和高可靠性。此外，物联网和无人机中继网络等新型网络拓扑，以及通感一体化 (Integrated Sensing and Communications, ISAC) 和算通一体化 (Integrated Computing and Communications, ICAC) 等尖端技术的出现，增加了现代通信系统设计的复杂性，这主要体现在以下三个方面：

- 系统和单模块在建模时存在各种不确定性；
- 用户设备和基站产生各种形式的大数据；
- 网络实现中面临各种算法挑战。

传统设计方法高度依赖人类智力，难以处理大规模复杂问题和动态海量数据，导致信息传输处理效率低下、性能受限。在此背景下，无线通信与信号处理开始利用人工智能和机器学习进行革新 [4]。关于机器学习在通信领域的应用研究，请参见 [5-11]。这一技术和方法转变推动业界开发出高度自适应、高效、鲁棒且智能的系统和算法，催生了第二代“智能传输与处理”概念，以期大幅减少对人类智力的依赖，并提升通信系统的整体性能。

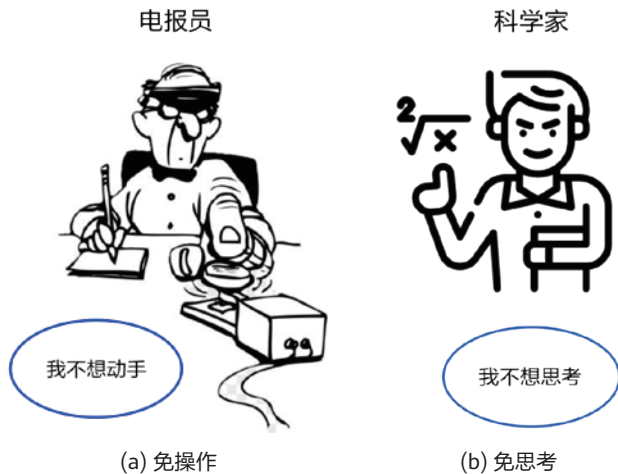


图 1 智能传输与处理的内涵 (图片来源: CLEANPNG.com 和 FLATICON.com)

图 1 展示了智能传输与处理的内涵。图 2 展示了基于机器学习的智能传输与处理系统的端到端架构。

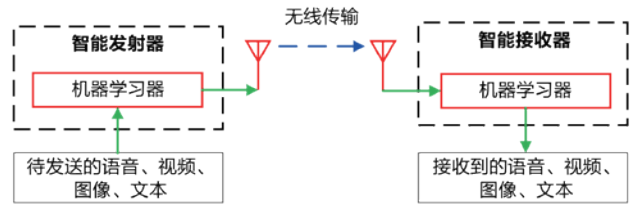


图 2 智能传输与处理系统的端到端架构。智能发射和接收模块作为端到端信息处理器 f ，均通过数据习得，并能实时自适应无线信道。智能化主要体现在不再需要人类研究大规模、动态和不确定的信息传递机制和处理方案。

为了展示机器学习在智能传输与处理中的重要作用，本文回顾了通信系统和方法中典型的机器学习应用，包括物理层通信 [6, 12]、语义通信 [13, 14]、通信资源分配 [15]、ICAC [16] 以及 ISAC [17]。受篇幅所限，本文并未穷尽列举该领域当前所有研究成果，仅阐明通往智能传输与处理的路径。

尽管机器学习有望革新无线通信理论和实践，基于机器学习的方法在带来机遇和优势的同时，也面临挑战和问题 [4]，如深度学习等“黑盒”学习方法缺乏可解释性导致的可靠性问题、训练数据有限及底层数据生成规律不平稳引发的泛化问题，以及在训练和存储大型机器学习模型（如深度学习）时资源不足的问题；参见图 3。除了上述主要挑战，还可能存在其他问题，如系统拓扑或硬件重配置（如添加或移除天线）引发的可扩展性问题（也可视为一种泛化问题），以及网络学习中的安全与隐私问题 [16]。在发展通信理论和系统时，切勿夸大机器学习的作用，机器学习（尤其是数据驱动深度学习）可以助力通信，但不是必须遵循的绝对规则；问题分析与机制建模历来都至关重要。相关技术综述，请参见 [18-20]。下例展示了数据驱动的机器学习方法的缺陷。

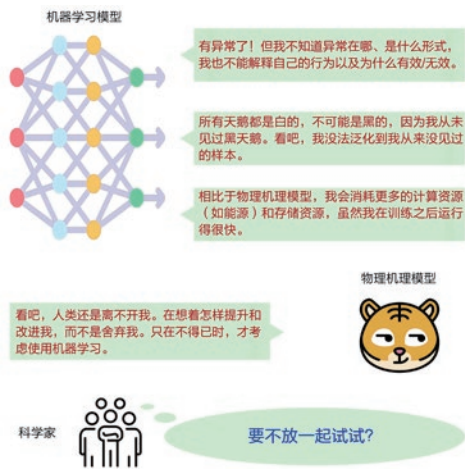


图 3 选择机器学习模型还是物理定律模型，这是个问题！我们需要选择最适合的方案！如果可能，将它们整合起来以提高整体系统性能。(图片来源: FLATICON.com)

示例（冰淇淋销售与鲨鱼袭击）：历史数据的回归分析表明，冰淇淋销售与鲨鱼袭击存在正相关性，但这显然不合理。然而，这种相关性背后的主要因素是气温——气温升高导致冰淇淋销量和海滩游客增加，而海滩游客增多则引发更多的鲨鱼袭击 [21]。可见，机制建模至关重要。

在探讨通信领域机器学习的应用前，我们将在第 2 节简要回顾其概念和方法，特别是可信赖机器学习。这些资料有助于读者理解无线通信中使用机器学习时的主要考虑因素，包括哲学和技术因素。

2 机器学习的概念与方法

机器学习的首要任务是从数据中发现隐含信息和模式，其优势在于能够自动解读数据，无需人类研究数据底层生成机制。这一特点从根本上使得机器智能能够参与到通信实践，尤其是信息传输与处理方面 [5, 7, 8, 11, 22]。

根据任务特点，机器学习可分为四类：有监督学习、无监督学习、半监督学习和强化学习。从数学角度看，所有机器学习任务的关键在于习得一个函数 f （亦称为“假设”），将观察数据映射到期望决策上。相关概念参见图 4，具体示例如下：



图 4 机器学习概念图。机器学习是一个寻找输入到决策的映射关系 f 的过程。上图是有监督学习，分类器 f 将图像识别为熊；下图是强化学习，动作生成器 f 结合当前场景为机器人提供路径选择建议。（图片均由 Microsoft Copilot 生成。）

- **有监督学习：**有监督学习挖掘了标记数据的隐含信息，可分为回归和分类两种任务类型。给定确定或随机变量对 (x, s) ，其中 x 表示特征向量， s 表示连续值期望响应。回归的目的是找到从 x 到 s 的函数关系 f ，使得预测标记 $f(x)$ 尽可能接近目标标记 s 。接近度由损失函数 $L[s, f(x)]$ 来衡量，如均方误差 $[s - f(x)]^2$ 。针对确定变量对 (x, s) ，可能存在函数 f ，使得每个 (x, s) 都有 $f(x) = s$ ，即 $L[s, f(x)] = 0$ 。而对于随机变量对 (x, s) ，一般无法保证完全相等。相反，需在 (x, s) 的联合分布 $\mathbb{P}_{x, s}$ 下计算损失，例如，损失的期望值

$\mathbb{E}_{(x, s) \sim \mathbb{P}_{x, s}} L[s, f(x)]$ 。当 s 为一维离散数值时，会出现分类问题， $L[s, f(x)]$ 由函数（如指示函数 $\mathbb{1}\{s \neq f(x)\}$ ）决定。在这种情况下， s 可表示不同的目标类，例如，对于二进制分类，有 $s \in \{-1, 1\}$ 。

- **无监督学习：**在无监督学习中，收集的数据没有标记 s ，只有特征 x 。因此，无监督学习主要用于从数据变量 x 的实现中发现隐含信息。聚类是一种典型的无监督学习任务。聚类与分类不同，分类是通过具有已知标记的训练数据集预测新数据点 x 的分类标记 s ，而聚类则根据相似数据点 x 的特征将其进行分组，无需预定义分类标记。简而言之，聚类就是寻找函数 f ，将数据 x 映射到一个合适的组。特征转换也是一种常见的无监督学习任务。它通过映射函数 f 将原始特征 x 转换到另一个特征空间，即 $y = f(x)$ 。例如，时域信号 x 及其傅里叶变换 y 就是一种特征转换。自编码器（Autoencoder）是一种人工神经网络，它的编解码操作也是一个很好的特征转换示例。另一个重要的无监督学习示例是分布估计，即估计出最能拟合（刻画）已收集数据的概率分布。在生成式任务中，分布估计尤为重要，例如根据已有样本生成新样本；具体比如给定一组猫的图片，利用拟合分布绘制出新的猫图像。
- **半监督学习：**半监督学习是监督学习的扩展，通过结合有标记数据 (x, s) 和无标记数据 x' 来提升模型性能。与完全依赖有标记数据的有监督学习不同，半监督学习利用无标记数据提高模型性能和泛化能力。当数据标记成本高耗时长，而无标记数据相对丰富且容易获取时，半监督学习就显得尤为有用。与单纯依赖标记数据的有监督学习相比，半监督学习能同时利用标记数据和无标记数据，习得的 f 能更准确地预测数据 x 的标记 s 。
- **强化学习：**强化学习关注的是动态、不确定环境中的决策问题。不同于依赖有标记数据的有监督学习或无标记数据的无监督学习，强化学习通过智能体与环境交互，接收奖励或惩罚反馈，从而习得各时间点的最优行为或策略。具体来说，智能体在当前状态 s 下执行动作 a ，通过自主学习决策获取最大化的累积奖励。因此，从数学上来看，需要习得从状态 s 到动作 a 的动作生成函数 f 。

关于这四类机器学习在无线通信中的具体应用，在 [8, 23] 中有具体介绍。

数据驱动学习与模型驱动学习：根据人类智力和领域知识的参与程度，机器学习可分为两类：数据驱动的机器学习和模型驱动的机器学习。数据驱动的机器学习完全依赖历史数据，不分析底层生成机制。而模型驱动的机器学习则在不同程度上结合了底层物理机制与数据生成模型，通过通信系统建模与大数据挖掘的协作，可以提升智能信息传输与处理的整体性能 [12, 23]。以信号检测为例，假设有 T 组导频数据 $\{(s_1, x_1), (s_2, x_2), \dots, (s_T, x_T)\}$ ，式中， x_i 表示接收信

号, s_i 表示传输符号 ($i = 1, 2, \dots, T$)。数据驱动的机器学习直接利用所有数据 (x 到 s) 训练出检测器 f 。而模型驱动的机器学习则首先考虑信号传输模型 $x = Hs + v$ (式中, H 表示信道矩阵, v 表示噪声), 然后基于该底层数据生成机制找到检测器 f 。详细的技术处理和讨论, 请参见 [19, 20, 24, 25]。

假设空间和深度学习: 为了找到最佳决策函数 f , 需要指定一个候选函数空间 \mathcal{H} , 即所谓的假设空间。有监督统计机器学习可以表述为:

$$\min_{f \in \mathcal{H}} \mathbb{E}_{(x,s) \sim \mathbb{P}_{x,s}} L(s, f(x))$$

式中, 未知的联合分布 $\mathbb{P}_{x,s}$ 可通过历史数据 (如经验分布) 估计。信号检测问题可以如上表述, 其中 f 表示检测器, x 表示天线接收的信号, s 表示传输的符号 (例如星座点) [20], 损失函数 L 可以是均方误差或误符号率。典型的假设空间 \mathcal{H} 有以下几种:

- 线性函数空间: \mathcal{H} 只包含输入 x 的线性变换。在信号检测中, \mathcal{H} 只包含线性检测器。
- 再生核希尔伯特空间: 基于特征映射函数 φ , \mathcal{H} 包含原始特征 x 的非线性提升特征 $\varphi(x)$ 的所有线性变换组合。实质上, \mathcal{H} 包含了输入 x 的一些特定类型的非线性函数。
- 神经网络函数空间: \mathcal{H} 由神经网络表征, 如多层感知机、递归神经网络、卷积神经网络 (Convolutional Neural Network, CNN)、径向基函数神经网络、自编码器和 Transformer 等。每个神经网络都定义了一个特定类型的函数空间 \mathcal{H} 。当使用的神经网络具有多个隐藏层时, \mathcal{H} 表示深度神经网络的函数空间。基于深度神经网络的机器学习又称为深度学习。

针对特定领域问题, 可以引入领域知识和专家设计, 调整或定制假设空间 [12, 18, 23]。因此, 模型驱动的机器学习通过利用已知问题特征和数据生成机制来设计一个专门且结构化的候选空间 \mathcal{H} 。

可解释性、可靠性和可持续性: 机器学习需要研究各种复杂问题, 包括模型的可解释性、可靠性和可持续性 [26, 27]。可解释性的目标是通过特征工程和物理建模等技术, 使学习模型透明、可解释、可纠错 [28]。基于底层物理数据生成机制的模型驱动机器学习就是这样的实现方式 [12, 18]。可靠性旨在构建鲁棒且准确的学习模型, 能够很好地泛化到新数据 (即训练未用到的数据), 以解决过拟合、泛化、知识迁移、小样本学习等问题 [20, 29-31]。可持续性旨在开发对环境和社会负面影响最小的学习模型, 解决能效、隐私安全、公平与偏见等问题 [16, 32]。对于智能信息传输与处理而言, 可解释性、可靠性和可持续性尤为重要, 在开发基于机器学习的无线通信方案时, 这些都是首要考虑因素。

集中式学习与分布式学习: 根据数据分布结构和计算架构, 可以采用不同方法训练机器学习模型, 主要有集中式

学习和分布式学习两种 [33, 34]。集中式学习将所有训练数据统一收集并存储在一个位置 (如数据中心或云服务器), 机器学习模型在此基础上进行训练。而分布式学习则是在多个设备或节点上分别训练机器学习模型, 每个设备或节点均存有部分数据。**联邦学习**是典型的分布式学习模式。联邦学习中, 多个客户端 (如智能手机、物联网设备或不同组织) 在不共享其本地数据的情况下共同训练模型, 各客户端利用其本地数据进行训练, 仅与中心节点共享模型更新 (梯度或权重), 由中心节点聚合这些更新, 形成新的全局模型。集中式和分布式学习方法适用于不同类型的现代通信网络, 有助于推动通信系统的进一步演进。

3 物理层通信

物理层通信旨在通过物理介质可靠地传输原始数据流, 如二进制比特。图 5 展示了传统的无线通信架构, 其功能模块由科学家根据基本数学和物理原理精心设计而来。这种模块化架构与图 2 所示的数据驱动的机器学习架构不同, 后者利用端到端操作取代了互连的功能模块。

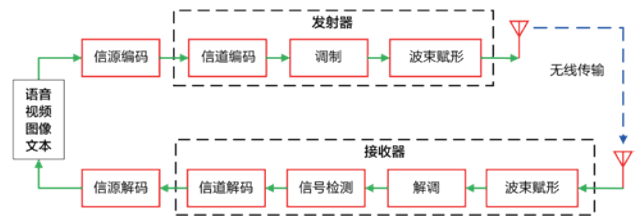


图 5 传统传输与处理系统的模块化架构。每个模块作为信息处理器 f , 由科学家根据底层物理机制和数学定律精心设计而来。

除了图 2 中高度集成化 (即高度智能化) 的架构外, 基于机器学习的传输与处理系统也可以实现部分智能化。例如, 在某些情况下, 仅由机器学习管理信道编码或解码模块, 即信道编解码方案是由机器而非信息科学家设计的。再如, 机器学习仅用于发射波束赋形器或接收波束赋形器, 如图 6 所示。

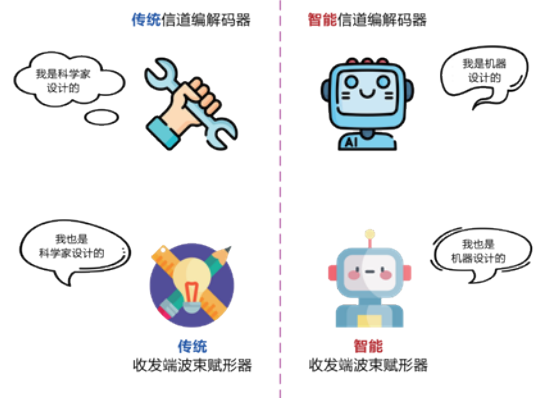


图 6 与图 2 所示的高度集成架构相比, 图 5 中的每个模块都可以通过机器学习增强 (图片来源: FLATICON.com)

技术上，机器学习在物理层通信中的应用包括整体端到端系统设计 [35, 36] (图 2) 和单模块设计 (图 6)。其中单模块设计包括：

- 编解码技术，如信源编码 [37]、信道编码 [38, 39] 以及联合信源信道编码 (Joint Source-Channel Coding, JSCC) [40, 41]
- 信号调制与检测 [25, 42]
- 发射和接收波束赋形 [20, 43-46]，如波束对齐和波束跟踪 [47-50]
- 信道估计与反馈 [51, 52]

文献 [6, 53, 54] 等已提供最新、最全的综述，本文不再赘述。

编解码技术在数字通信中至关重要，可确保数据的高效可靠传输。近年来，机器学习被广泛应用于增强信源编码、信道编码和 JSCC 等编解码技术。传统信源编码 (即数据压缩) 通过减少数据冗余实现高效传输和存储，而基于神经网络的自编码器机器学习技术已彻底改变了这一方式。自编码器通过将数据编码到低维空间并重构来学得高效的数据表征，从而在信息损失最小的情况下实现高压缩率 [55]。信道编码通过为数据添加冗余来检测并纠正由噪声信道引起的传输错误。伴随着机器学习模型 (尤其是深度学习) 模型的应用，一些新型纠错码开始涌现。例如，用神经解码器解码复杂方案，如低密度奇偶校验 (Low-Density Parity Check, LDPC) 码 [56] 和 Turbo 码 [57]，性能优于传统编码方式，尤其在噪声大的环境中表现更佳。JSCC 同时集成了信源编码与信道编码，提升了系统整体性能。变分自编码器 (Variational Autoencoder, VAE) [58]、CNN [59] 和生成式对抗网络 (Generative Adversarial Network, GAN) [60] 等机器学习模型可用于联合学习表征和纠错码，它们能够同时适应信源和信道特征，压缩和容错性能远胜过传统方式。总体而言，基于机器学习的编解码技术极大地促进了数字通信的发展，通过利用其预测与自适应功能，这些技术增强了数据压缩、纠错及传输效率，为构建更鲁棒、高效的通信系统奠定了基础。

在数字通信中，信号调制与检测是数据传输和解码的基本步骤。近年来，机器学习技术被应用于信道调制与检测，提高了传输效率和可靠性。调制通过改变载波信号的幅度、频率或相位等属性来编码信息。传统调制方案包括调幅 (Amplitude Modulation, AM)、调频 (Frequency Modulation, FM) 和相移键控 (Phase Shift Keying, PSK) 等技术。机器学习 (尤其是深度学习) 模型正应用于自适应调制方案设计，它们可以根据信道条件动态调整调制参数，实时优化性能。例如，神经网络可以学习复杂的调制模式，实现数据吞吐率最大化和误码率最小化 [61]。检测通过解调接收信号来恢复传输信息。传统方法使用预设算法估计传输数据，但通常是基于特定信道特征的假设。全连接深度神经网络 [42] 和迁移学习 [62] 等机器学习方法的出现，使得信号检测大大增强。这些模型能够从数据中学习，在复

杂多变的信道条件下准确检测信号，提高系统的抗噪声 / 干扰能力。总体而言，机器学习技术与信号调制检测的融合，是通信技术领域的重大进步，提高了数据传输效率、抗噪能力和系统整体性能。

在无线通信系统中，发射和接收波束赋形技术对增强信号质量和提高数据吞吐率至关重要。该技术通过天线阵列定向发射或接收信号，提高信号强度并减少干扰。近年来，机器学习显著提升了波束赋形性能，从而带来更广阔的应用场景。传统的发射波束赋形方法，如相控阵列系统，通过预设算法调整多天线信号的相位和幅度。而引入机器学习 (尤其是深度学习) 模型后，系统可从环境数据中自主学习、自优化。例如，强化学习能够根据通信环境的反馈实时调整波束赋形模式，从而在复杂多变的场景下提升性能 [63, 64]。传统的接收波束赋形方法，如最小方差无失真响应 (Minimum Variance Distortionless Response, MVDR) 和最大比合并 (Maximal Ratio Combining, MRC)，依赖于信号环境统计模型。而 CNN 等机器学习方法则通过直接从数据中习得最佳波束赋形权重，能在各种动态环境中实现准确、稳健地接收信号 [65, 66]。波束对准与跟踪是波束赋形的关键技术，在毫米波 (Millimeter Wave, mmWave) 和太赫兹通信等高频通信中尤为重要。它们能确保收发端始终保持最佳对准状态，从而实现信号强度和吞吐率的最大化 [49, 50, 67]。传统方法依赖穷举搜索或迭代算法，不仅耗时，而且对计算资源的需求很大。有监督学习、多臂老虎机、强化学习等机器学习方法可基于历史数据预测最优波束方向，提供更高效的解决方案，大幅缩小搜索空间。即使收发端移动或环境产生变化的情况下，波束也能始终保持对准。机器学习 (尤其是深度学习) 模型通过实时预测波束方向的变化来增强跟踪能力，特别是能捕捉时间依赖性的循环神经网络 (Recurrent Neural Network, RNN) 和长短期记忆 (Long Short-Term Memory, LSTM) 网络。在基于机器学习的波束对准与跟踪方面，[47-50, 67] 等文献提供了翔实的技术细节，此处不再赘述。总之，机器学习与波束赋形 (包括波束对准与跟踪) 技术的融合，通过预测和自适应能力提升信号质量、减少干扰并优化系统性能，对 5G 等下一代网络至关重要。

信道估计与反馈技术用于测量信道特性并向发射端提供必要反馈，确保准确表征信道特征并高效传输数据，在无线通信系统中至关重要。近年来，随着机器学习技术的引入，信道估计的准确性和效率显著提升 [68, 69]。信道估计通过预测信道状态优化信号收发。传统方法如最小均方差 (Minimum Mean Square Error, MMSE) 依赖统计模型，计算资源需求高。机器学习 (尤其是深度学习) 模型为提升信道估计精度及降低计算复杂度提供了新的思路。例如，CNN 可以基于接收信号数据自我学习并估计信道状态，为复杂场景提供了鲁棒性和自适应性更优的解决方案 [70]。LSTM 可以有效捕捉信道条件的时延依赖性，提高估计精度 [71]。反馈机制将信道状态信息 (Channel State Information, CSI) 从接收端传回发送端，便于实时调整发

射设置。传统反馈方法通常涉及 CSI 量化和编码，可能引入时延和误差。自编码器和 CNN 等机器学习技术通过压缩、重构 CSI，在尽量减少信息损失的同时提高反馈效率 [52]，从而更精确及时地调整传输策略。此外，机器学习模型可以同时处理信道估计和反馈，从而整体优化这两个过程 [69]，提升系统性能。

4 语义通信

与传统物理层通信不同，语义通信侧重于传输图像、文本、音频等数据中的语义信息，而非逐位传输原始数据。相比之下，语义通信能显著减少无线信道的传输负载，从而大幅提升信息传输速度和效率。在语义通信方面，[72-75] 等文献提供了最新研究，此处不再赘述。

语义通信的核心是从原始数据中提取语义信息，这可以通过精心设计的信源编解码策略或 JSCC 来实现。然而，原始数据中的语义信息通常仅限于特定任务（见图 7），目前尚未形成通用的数学分析、建模和计算框架。关于该方向的探索性研究，请参见 [76] 等文献。因此，对于特定通信任务，需要设计合适的语义编解码方案。在智能传输与处理系统中，语义通信由在高度集成的端到端收发器实现，参见图 2。

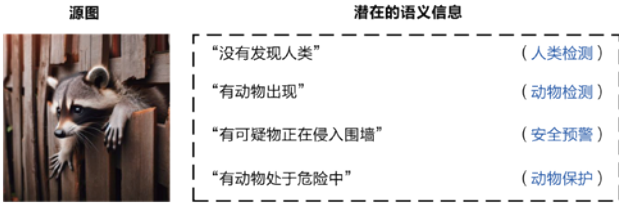


图 7 每个任务的原始数据包含不同的语义信息。无损传输高清图象既耗时又耗资源，而准确传输语义消息则相对简单且成本更低。（图片由 Microsoft Copilot 生成）

机器学习使系统可以更准确地理解、处理和传达语义，在语义通信中起着关键作用。深度学习模型（如 Transformer、CNN 和 RNN）已广泛应用于分析和预测数据的语义相关性，从而提高带宽利用率并通信效率。具体而言，自然语言处理（Natural Language Processing, NLP）算法可以从文本和音频中提取并解释语义内容，更高效地压缩与传输数据 [77, 78]；计算机视觉方法则可从图像和视频中提取并解释语义信息 [79]。

最近研究表明，基于机器学习的语义通信应用潜力巨大。例如，通过设计收发器神经网络直接传输文本语义 [13]，可显著减少通信资源需求并提升传输性能；开发高效视频会议系统 [80]，可提高传输效率。大多数语义通信研究致力于通过 JSCC 来节省资源，但这些成果需要改变现有基础设施，不利于实际应用。在此背景下，[14] 提出一个务实方法，仅修改现有基础设施中的部分模块来实现无线语义传输。为确保语义传输的可靠性和通信效率，有学者研究了语义域中的频谱效率及基于语义感知的资源分配问题 [81]。此外，语义

通信与物联网 (Internet of Things, IoT) [82]、边缘计算 [83] 等新兴技术协作，为智能化、上下文感知通信系统创造了新机遇。基于分布式机器学习模型，语义通信系统可以动态适应环境变化和用户需求，确保稳健高效的信息交换 [84]。

总之，语义通信依托先进的机器学习技术，为通信系统指明了发展方向，有望革新信息传输和解码方式，将会对未来的通信效率和效能产生深远影响。

5 通信资源分配

无线通信中的资源分配旨在高效管理和利用频谱、功率、计算、空间和时间等资源，以提高网络性能，实现更高的吞吐率、更低的时延、更广的覆盖范围和更高的可靠性 [15, 85, 86]。典型应用包括链路调度、消息路由、功率分配、信道选择、波束赋形、频谱接入与管理以及信道划分协议设计（即时分、频分等）等。从数学规划角度来看，资源分配可表示为优化问题。从运筹学角度，分配和调度是关键技术——分配解决静态资源分配问题，而调度则解决动态资源分配问题（静态和动态是相对时间而言的）。从计算和算法角度，主流解决方案框架包括：

- 连续优化、离散（如组合、整数）优化和混合优化
- 单目标优化和多目标优化
- 线性规划和非线性规划
- 凸优化与非凸优化
- 光滑优化与非光滑优化
- Min-Max 优化（如博弈论、最坏情况鲁棒性分析）
- 确定性规划和随机规划（即是否涉及随机变量；如涉及，考虑相应的概率分布）
- 单阶段优化（即静态规划）和多阶段优化（即动态规划）
- 启发式优化（如遗传算法、粒子群优化和模拟退火）
- 代理优化，也称黑盒优化（如贝叶斯优化）
- 基于机器学习的优化（如基于强化学习和深度学习的求解方法）

无线通信中的典型资源分配应用及解决方案框架如图 8 所示，更多信息请参考 [85, 87, 88]。

随着 ISAC 的出现 [89]，对资源分配的要求有所改变，即便是最佳的通信资源分配方案，也未必同样适用于感知业务（详见 [90]）。例如，由于通信和感知功能设计偏好存在差异甚至冲突，二者的最优波形也有所不同 [91, 92]。因此，应细化空口、计算、功率、时间、波束等资源的分配方案，满足通信和感知的性能要求。ICAC（如边缘计算 [93] 和网络控制* [94]）也面临同样的困境，需在计算与通信之间合理分配有限的资源（详见 [95, 96]）。

* 控制器本质上是信息处理器，因此属于特定用途的计算模块。

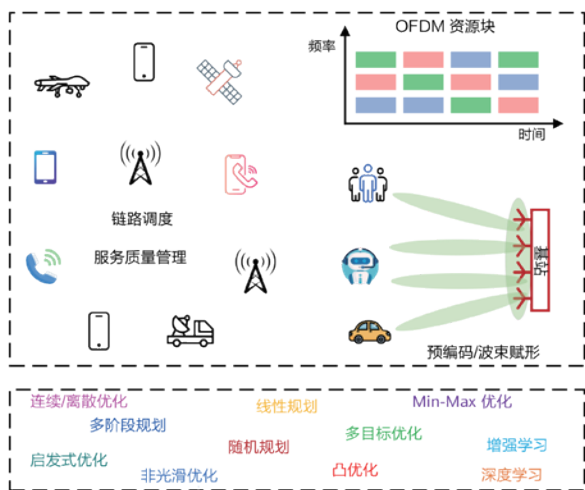


图8 无线通信中的典型资源分配应用及解决方案框架。OFDM：正交频分复用。（图片来源：FLATICON.com）

机器学习技术已成为解决无线通信中资源分配问题的有力工具。最近研究表明，机器学习在各类资源分配任务中具有巨大潜力。例如，深度强化学习已应用于优化异构网络中的频谱分配、功率控制和用户关联等问题，与传统方法相比，优化效果显著提升 [22, 97-102]。同样，监督学习算法已成功应用于解决混合整数非线性规划（Mixed Integer Nonlinear Programming, MINLP）等资源分配的复杂优化问题 [103]。此外，无监督学习技术还可用于解决资源分配问题，如图嵌入技术在链路调度中的应用 [104]。

传统资源分配方法在很大程度上依赖人类智力构建精确模型、开发特定方案，但在大规模动态复杂场景中可能表现欠佳甚至无法适应。机器学习，尤其是深度学习和强化学习等技术，可以对环境中的复杂交互进行建模，预测未来通信网络状态，并实时优化决策，从而提升无线网络性能和适应性 [15]。例如，现网信道条件会不断变化，但我们无法预知信道会如何随时间而变化，因此难以用数学手段处理这种不确定性。此外，数学规划模型计算复杂，很难高效解决问题。而机器学习则可以利用现实数据，发现数据中隐含知识和模式，并自动做出令人满意的资源分配决策。从技术角度来看，机器学习可以将资源的分配优化问题视为参数-决策的映射问题，从而帮助解决复杂的优化问题。这种映射依赖于深度神经网络强大的有监督学习函数拟合能力，标注数据-决策对由性能好的人工解法生成。此外，机器学习还可将资源分配问题中的效用函数作为训练阶段中的损失函数，无需依赖传统人工算法即可生成高质量的资源分配决策。除了这两种方案外，还有一种“算法展开”技术 [18]，即使用神经网络来展开现有高效迭代算法。具体来说，每层神经网络充当迭代算法的一个迭代步骤，通过多层级联模拟该算法的迭代过程。这种由算法指导的方法也称为“模型驱动的深度强化学习” [12]，根据领域知识定制深度神经网络架构，以提高网络泛化能力，并减少训练所需的数据量。强化学习是第

四种有前景的方案。强化学习通过探索未知且难以建模的环境（如复杂信道条件）并与其动态交互来获得智能资源分配方案。图9总结了机器学习在资源分配中的四个典型职能。机器学习方法的第一个优势是在运行阶段计算速度极快——尽管训练阶段的计算量可能较大（相比前三种方案）；第二个优势在于其不依赖建模即可应对动态、不确定甚至未知的环境（相比第四种方案，即强化学习方案）。

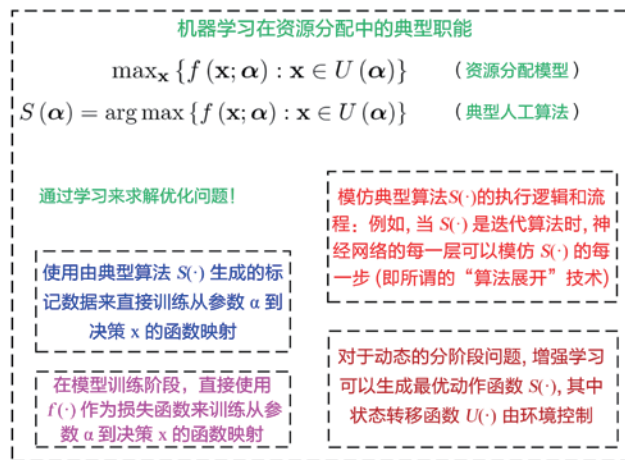


图9 机器学习在资源分配中的四个典型职能（以学习的方式求解优化问题）

6 不仅仅是数据传输：感知与计算

随着对低时延高速连接、高性能通信辅助感知（如用户定位和跟踪）及协同计算设备需求的增加，无线通信系统正在经历一场变革，这推动了 ISAC [89, 105] 和 ICAC [106, 107] 等新型系统范式的发展，使得传统上分离的功能得到了统一，从而提高资源利用率、降低硬件成本并增强系统能力。例如，通过波束管理和资源分配，利用环境和用户传感数据提升通信性能；跨网络节点的传感数据共享实现实时网络监控与态势感知，从而提高感知精度并扩大覆盖范围。再如，边缘本地数据处理可减少实时通信业务的时延，而高速通信则为大规模数据分析提供高效的分布式计算。如前所述，机器学习，尤其是深度学习，在现代通信系统中至关重要。它们利用算法从海量数据中学习，优化网络资源分配、信号处理和故障检测等性能。这些优势同样适用于新兴的 ISAC 和 ICAC 系统。在 ISAC 系统中，CNN 和 Transformer 等深度学习模型可以提升感知精度与鲁棒性 [108]，并实现语义信息传输 [109]。在 ICAC 系统中，联邦学习等机器学习算法能够保护用户数据隐私并优化计算任务，提高数据处理与通信效率 [110, 111]。总之，机器学习/深度学习与新兴的一体化范式相结合，将使通信系统变得更加智能、自适应且高效。

6.1 通感一体化

ISAC 是一种将感知与通信融合为一体的系统，感知与通信共享基础设施和频谱资源。该系统对自动驾驶、智慧城市和先进监控系统等应用至关重要。通过同步数据采集与通信，ISAC 提升系统效率和性能，降低硬件成本并缓解频谱拥塞。然而，ISAC 使通信波形设计、系统和硬件资源分配、干扰管理及整体网络操作更加复杂 [112, 17]。这些挑战推动了方法创新，有利于释放 ISAC 在现实应用中的潜力。因此，机器学习和深度学习技术为 ISAC 提供了先进的数据处理与决策能力，其重要性不言而喻。如欲系统了解机器学习在 ISAC 的作用，请参见 [17] 和 [112]。

6.2 算通一体化

ICAC 是一种将计算和通信融为一体的技术。为满足应用日益增长的计算需求，同时保持高效稳健的通信性能，在源端就近处理大规模数据任务，减少时延并提升边缘计算效率，实现连接设备智能化，ICAC 应运而生。该技术有助于实时处理和分析数据，对于工业自动化、虚拟现实以及物联网等应用至关重要。ICAC 典型应用包括边缘计算、联邦学习、普适计算、雾计算、物联网 / 车联网及自治系统等。机器学习和深度学习是 ICAC 的重要组成部分，能够实现动态资源分配、系统配置自适应和实时信息分析。这些技术确保计算与通信资源得到有效利用，从而提升性能和响应能力。关于机器学习在 ICAC 中的具体作用，请参阅 [16]、[113] 和 [114]。需要注意的是，群体智能和网络控制 [94] 也与 ICAC 密切相关，因为控制器本质上是将系统状态信号映射到系统控制输入信号的信息处理器，属于特定用途的计算模块。

7 讨论和结论

本文探讨了机器学习在物理层通信、语义通信、资源分配、ISAC 及 ICAC（如联邦学习、边缘计算）等关键无线应用中的变革潜力。从信号处理算法到整网管理，机器学习都有着广泛的影响。尽管如此，我们不能过分夸大机器学习的作用，在实际应用中机器学习仍会面临一些挑战。模型的可解释性、故障处理、大规模训练数据集需求以及训练和部署所需的计算资源（如电力、处理速度）等问题必须解决。此外，还需关注机器学习系统的可靠性和安全性，尤其是在涉及数据隐私（如联邦学习 [115]）、数据时效性（如小样本学习 [116, 117]）及实时决策（如自动驾驶）等场景中。为应对这些挑战，本文建议将传统物理定律模型与数据驱动的机器学习模型相结合，这样既能利用物理机制的可靠性和可解释性，又能发挥机器学习的自适应性和学习能力，从而提升整体系统性能。关于智能传输与处理的特征、挑战及未来考量，详见图 10。

在所有可预见的挑战中，以下三项是实现基于机器学习通信系统的最低要求，也是实际应用中的关键挑战：

- 如何解读机器学习模型的性能增益与故障，并在系统宕机时排查和修复故障，以提高整体可靠性？从这个角度看，图 6 中的范式比图 2 中的更可靠、更易管理。
- 如何利用有限数据获得更好的泛化能力？如何整合新的可用数据来提高泛化能力 [20, 31]，包括如何将已学得模型快速适配到新数据上（如当环境中的数据生成规则随时间变化时 [62, 117]）？在机器学习领域，数据时效性、样本效率和数据分布鲁棒性都与此问题密切相关。
- 如何构建基于领域知识的机器学习模型（而非多层感知等通用深度神经网络）？如何设计计算高效的训练算法（而非随机梯度下降法）以缩短响应时间、降低功耗 [12, 108, 118, 119]？如何缩小模型（尤其是深度神经网络）大小以节省存储空间 [120]？对于嵌入式设备和边缘设备，这三点考虑尤为重要。

简言之，机器学习与通信系统的融合是重大的技术进步，为建立更加智能、高效、可靠的通信网络提供了可能。

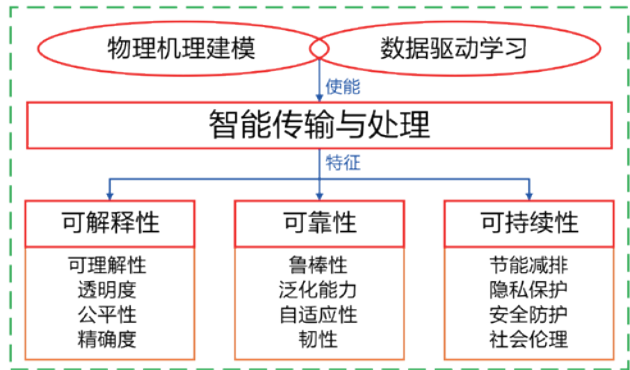


图 10 智能传输与处理的特点、挑战和注意事项（对照图 3）。尽管数据驱动的机器学习很强大，但机理建模（包括发现物理或数学定律）始终对提高可解释性、可靠性和可持续性至关重要。

参考文献

- [1] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, 2006.
- [3] G. L. Stüber and G. L. Steuber, *Principles of Mobile Communication*, 4th ed. Springer, 2017.
- [4] W. Tong and G. Y. Li, "Nine challenges in artificial intelligence and wireless communications for 6G," *IEEE Wireless Communications*, vol. 29, no. 4, pp. 140–145, 2022.
- [5] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, 2016.
- [6] Z. Qin, H. Ye, G. Y. Li, and B.-H. F. Juang, "Deep learning in physical layer communications," *IEEE Wireless Communications*, vol. 26, no. 2, pp. 93–99, 2019.
- [7] D. Gündüz, P. De Kerret, N. D. Sidiropoulos, D. Gesbert, C. R. Murthy, and M. Van der Schaar, "Machine learning in the air," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 10, pp. 2184–2199, 2019.
- [8] J. Wang, C. Jiang, H. Zhang, Y. Ren, K.-C. Chen, and L. Hanzo, "Thirty years of machine learning: The road to Pareto-optimal wireless networks," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1472–1514, 2020.
- [9] W. Yu, F. Sotiriou, and T. Jiang, "Role of deep learning in wireless communications," *IEEE BITS the Information Theory Magazine*, vol. 2, no. 2, pp. 56–72, 2022.
- [10] A. Alhammedi, I. Shayea, A. A. El-Saleh, M. H. Azmi, Z. H. Ismail, L. Kouhalvandi, and S. A. Saad, "Artificial intelligence in 6G wireless networks: Opportunities, applications, and challenges," *International Journal of Intelligent Systems*, vol. 2024, no. 1, p. 8845070, 2024.
- [11] A. Celik and A. M. Eltawil, "At the dawn of generative AI era: A tutorial-cum-survey on new frontiers in 6G wireless intelligence," *IEEE Open Journal of the Communications Society*, 2024.
- [12] H. He, S. Jin, C.-K. Wen, F. Gao, G. Y. Li, and Z. Xu, "Model-driven deep learning for physical layer communications," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 77–83, 2019.
- [13] H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, "Deep learning enabled semantic communication systems," *IEEE Transactions on Signal Processing*, vol. 69, pp. 2663–2675, 2021.
- [14] P. Jiang, C.-K. Wen, S. Jin, and G. Y. Li, "Wireless semantic transmission via revising modules in conventional communications," *IEEE Wireless Communications*, vol. 30, no. 3, pp. 28–34, 2023.
- [15] L. Liang, H. Ye, G. Yu, and G. Y. Li, "Deep-learning-based wireless resource allocation with application to vehicular networks," *Proceedings of the IEEE*, vol. 108, no. 2, pp. 341–356, 2019.
- [16] M. A. Ferrag, O. Friha, B. Kantarci, N. Tihanyi, L. Cordeiro, M. Debbah, D. Hamouda, M. Al-Hawawreh, and K.-K. R. Choo, "Edge learning for 6G-enabled Internet of things: A comprehensive survey of vulnerabilities, datasets, and defenses," *IEEE Communications Surveys & Tutorials*, 2023.
- [17] S. Lu, F. Liu, Y. Li, K. Zhang, H. Huang, J. Zou, X. Li, Y. Dong, F. Dong, J. Zhu, *et al.*, "Integrated sensing and communications: Recent advances and ten open challenges," *IEEE Internet of Things Journal*, 2024.
- [18] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Processing Magazine*, vol. 38, no. 2, pp. 18–44, 2021.
- [19] N. Shlezinger and T. Rottenberg, "Discriminative and generative learning for the linear estimation of random signals [lecture notes]," *IEEE Signal Processing Magazine*, vol. 40, no. 6, pp. 75–82, 2023.

- [20] S. Wang, W. Dai, and G. Y. Li, "Distributionally robust receive beamforming," *arXiv preprint arXiv:2401.12345*, 2024.
- [21] G. James, D. Witten, T. Hastie, R. Tibshirani, *et al.*, *An Introduction to Statistical Learning*, 2nd ed. Springer, 2021.
- [22] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3072–3108, 2019.
- [23] Y. C. Eldar, A. Goldsmith, D. Gündüz, and H. V. Poor, *Machine Learning and Wireless Communications*. Cambridge University Press, 2022.
- [24] N. Shlezinger, J. Whang, Y. C. Eldar, and A. G. Dimakis, "Model-based deep learning," *Proceedings of the IEEE*, vol. 111, no. 5, pp. 465–499, 2023.
- [25] H. He, C.-K. Wen, S. Jin, and G. Y. Li, "Model-driven deep learning for MIMO detection," *IEEE Transactions on Signal Processing*, vol. 68, pp. 1702–1715, 2020.
- [26] B. Thuraisingham, "Trustworthy machine learning," *IEEE Intelligent Systems*, vol. 37, no. 1, pp. 21–24, 2022.
- [27] K. R. Varshney, *Trustworthy Machine Learning*. Chappaqua, NY, USA: Independently Published, 2022.
- [28] C. Molnar, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, 2nd ed. Leanpub, 2020.
- [29] K. Kawaguchi, Z. Deng, K. Luh, and J. Huang, "Robustness implies generalization via data-dependent generalization bounds," in *International Conference on Machine Learning*. PMLR, 2022, pp. 10 866–10 894.
- [30] J. Wang, C. Lan, C. Liu, Y. Ouyang, T. Qin, W. Lu, Y. Chen, W. Zeng, and S. Y. Philip, "Generalizing to unseen domains: A survey on domain generalization," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 8, pp. 8052–8072, 2022.
- [31] S. Wang and H. Wang, "Distributional robustness bounds generalization errors," *arXiv preprint arXiv:2212.09962*, 2024.
- [32] A. Van Wynsberghe, "Sustainable AI: AI for sustainability and the sustainability of AI," *AI and Ethics*, vol. 1, no. 3, pp. 213–218, 2021.
- [33] S. AbdulRahman, H. Tout, H. Ould-Slimane, A. Mourad, C. Talhi, and M. Guizani, "A survey on federated learning: The journey from centralized to distributed on-site learning and beyond," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5476–5497, 2020.
- [34] M. Chen, D. Gündüz, K. Huang, W. Saad, M. Bennis, A. V. Feljan, and H. V. Poor, "Distributed learning in wireless networks: Recent progress and future challenges," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 12, pp. 3579–3605, 2021.
- [35] H. Ye, L. Liang, G. Y. Li, and B.-H. Juang, "Deep learning-based end-to-end wireless communication systems with conditional GANs as unknown channels," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3133–3143, 2020.
- [36] H. Ye, G. Y. Li, and B.-H. Juang, "Deep learning-based end-to-end wireless communication systems without pilots." *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 3, pp. 702–714, 2021.
- [37] S. Manouchehri, J. Haghghat, M. Eslami, and W. Hamouda, "A delay-efficient deep learning approach for lossless turbo source coding," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 6, pp. 6704–6709, 2022.
- [38] H. Ye, L. Liang, and G. Y. Li, "Circular convolutional autoencoder for channel coding," in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2019, pp. 1–5.
- [39] Y. Zhang, H. Wu, and M. Coates, "On the design of channel coding autoencoders with arbitrary rates for ISI channels," *IEEE Wireless Communications Letters*, vol. 11, no. 2, pp. 426–430, 2021.

- [40] M. Jankowski, D. Gündüz, and K. Mikolajczyk, "Deep joint source-channel coding for wireless image retrieval," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 5070–5074.
- [41] M. Yang, C. Bian, and H.-S. Kim, "Deep joint source-channel coding for wireless image transmission with OFDM," in *ICC 2021-IEEE International Conference on Communications*. IEEE, 2021, pp. 1–6.
- [42] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 114–117, 2017.
- [43] S. Mohammadzadeh, V. H. Nascimento, R. C. de Lamare, and N. Hajarolasvadi, "Robust beamforming based on complexvalued convolutional neural networks for sensor arrays," *IEEE Signal Processing Letters*, vol. 29, pp. 2108–2112, 2022.
- [44] A. M. Elbir, K. V. Mishra, M. R. B. Shankar, and B. Ottersten, "A family of deep learning architectures for channel estimation and hybrid beamforming in multi-carrier mm-wave massive MIMO," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 2, pp. 642–656, 2022.
- [45] D. d. S. Brilhante, J. C. Manjarres, R. Moreira, L. de Oliveira Veiga, J. F. de Rezende, F. Müller, A. Klautau, L. Leonel Mendes, and F. A. P. de Figueiredo, "A literature survey on AI-aided beamforming and beam management for 5G and 6G systems," *Sensors*, vol. 23, no. 9, p. 4359, 2023.
- [46] N. Shlezinger, M. Ma, O. Lavi, N. T. Nguyen, Y. C. Eldar, and M. Juntti, "Artificial intelligence-empowered hybrid multiple-input/multiple-output beamforming: Learning to optimize for high-throughput scalable MIMO," *IEEE Vehicular Technology Magazine*, 2024.
- [47] S. H. Lim, S. Kim, B. Shim, and J. W. Choi, "Deep learning-based beam tracking for millimeter-wave communications under mobility," *IEEE Transactions on Communications*, vol. 69, no. 11, pp. 7458–7469, 2021.
- [48] F. Sahrabi, T. Jiang, W. Cui, and W. Yu, "Active sensing for communications by learning," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 6, pp. 1780–1794, 2022.
- [49] Y. Wei, Z. Zhong, and V. Y. Tan, "Fast beam alignment via pure exploration in multi-armed bandits," *IEEE Transactions on Wireless Communications*, vol. 22, no. 5, pp. 3264–3279, 2022.
- [50] W. Yi, W. Zhiqing, and F. Zhiyong, "Beam training and tracking in mmWave communication: A survey," *China Communications*, 2024.
- [51] Q. Hu, F. Gao, H. Zhang, S. Jin, and G. Y. Li, "Deep learning for channel estimation: Interpretation, performance, and comparison," *IEEE Transactions on Wireless Communications*, vol. 20, no. 4, pp. 2398–2412, 2020.
- [52] J. Guo, C.-K. Wen, S. Jin, and G. Y. Li, "Convolutional neural network-based multiple-rate compressive sensing for massive MIMO CSI feedback: Design, simulation, and analysis," *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2827–2840, 2020.
- [53] B. Ozpoyraz, A. T. Dogukan, Y. Gevez, U. Altun, and E. Basar, "Deep learning-aided 6G wireless networks: A comprehensive survey of revolutionary PHY architectures," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 1749–1809, 2022.
- [54] N. Ye, S. Miao, J. Pan, Q. Ouyang, X. Li, and X. Hou, "Artificial intelligence for wireless physical-layer technologies (AI4PHY): A comprehensive survey," *IEEE Transactions on Cognitive Communications and Networking*, 2024.
- [55] Y. Yang, S. Mandt, L. Theis *et al.*, "An introduction to neural data compression," *Foundations and Trends® in Computer Graphics and Vision*, vol. 15, no. 2, pp. 113–200, 2023.
- [56] S. Han, J. Oh, K. Oh, and J. Ha, "Deep-learning for breaking the trapping sets in low-density parity-check codes," *IEEE Transactions on Communications*, vol. 70, no. 5, pp. 2909–2923, 2022.
- [57] Y. Jiang, H. Kim, H. Asnani, S. Kannan, S. Oh, and P. Viswanath, "Turbo autoencoder: Deep learning based channel codes for point-to-point communication channels," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

- [58] K. Choi, K. Tatwawadi, A. Grover, T. Weissman, and S. Ermon, "Neural joint source-channel coding," in *International Conference on Machine Learning*. PMLR, 2019, pp. 1182–1192.
- [59] E. Boursoulatze, D. B. Kurka, and D. Gündüz, "Deep joint source-channel coding for wireless image transmission," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 3, pp. 567–579, 2019.
- [60] E. Erdemir, T.-Y. Tung, P. L. Dragotti, and D. Gündüz, "Generative joint source-channel coding for semantic image transmission," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 8, pp. 2645–2657, 2023.
- [61] E. Bobrov, D. Kropotov, H. Lu, and D. ZaeV, "Massive MIMO adaptive modulation and coding using online deep learning algorithm," *IEEE Communications Letters*, vol. 26, no. 4, pp. 818–822, 2021.
- [62] N. Van Huynh and G. Y. Li, "Transfer learning for signal detection in wireless networks," *IEEE Wireless Communications Letters*, vol. 11, no. 11, pp. 2325–2329, 2022.
- [63] F. B. Mismar, B. L. Evans, and A. Alkhateeb, "Deep reinforcement learning for 5G networks: Joint beamforming, power control, and interference coordination," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1581–1592, 2019.
- [64] M. Chu, A. Liu, V. K. Lau, C. Jiang, and T. Yang, "Deep reinforcement learning based end-to-end multiuser channel prediction and beamforming," *IEEE Transactions on Wireless Communications*, vol. 21, no. 12, pp. 10 271–10 285, 2022.
- [65] H. Huang, Y. Peng, J. Yang, W. Xia, and G. Gui, "Fast beamforming design via deep learning," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 1, pp. 1065–1069, 2019.
- [66] P. Ramezanpour, M. J. Rezaei, and M. R. Mosavi, "Deep learning-based beamforming for rejecting interferences," *IET Signal Processing*, vol. 14, no. 7, pp. 467–473, 2020.
- [67] K. Chen, C. Qi, C.-X. Wang, and G. Y. Li, "Beam training and tracking for extremely large-scale MIMO communications," *IEEE Transactions on Wireless Communications*, 2023.
- [68] J. Guo, C.-K. Wen, S. Jin, and G. Y. Li, "Overview of deep learning-based CSI feedback in massive MIMO systems," *IEEE Transactions on Communications*, vol. 70, no. 12, pp. 8017–8045, 2022.
- [69] J. Guo, T. Chen, S. Jin, G. Y. Li, X. Wang, and X. Hou, "Deep learning for joint channel estimation and feedback in massive MIMO systems," *Digital Communications and Networks*, vol. 10, no. 1, pp. 83–93, 2024.
- [70] P. Jiang, C.-K. Wen, S. Jin, and G. Y. Li, "Dual CNN-based channel estimation for MIMO-OFDM systems," *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 5859–5872, 2021.
- [71] R. Shankar, "Bi-directional LSTM based channel estimation in 5G massive MIMO OFDM systems over TDL-C model with Rayleigh fading distribution," *International Journal of Communication Systems*, vol. 36, no. 16, p. e5585, 2023.
- [72] W. Yang, H. Du, Z. Q. Liew, W. Y. B. Lim, Z. Xiong, D. Niyato, X. Chi, X. Shen, and C. Miao, "Semantic communications for future internet: Fundamentals, applications, and challenges," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 1, pp. 213–250, 2022.
- [73] X. Luo, H.-H. Chen, and Q. Guo, "Semantic communications: Overview, open issues, and future research directions," *IEEE Wireless Communications*, vol. 29, no. 1, pp. 210–219, 2022.
- [74] Z. Lu, R. Li, K. Lu, X. Chen, E. Hossain, Z. Zhao, and H. Zhang, "Semantics-empowered communications: A tutorial-cum-survey," *IEEE Communications Surveys & Tutorials*, 2023.
- [75] C. Chaccour, W. Saad, M. Debbah, Z. Han, and H. V. Poor, "Less data, more knowledge: Building next generation semantic communication networks," *IEEE Communications Surveys & Tutorials*, 2024.

- [76] D. Gündüz, Z. Qin, I. E. Aguerri, H. S. Dhillon, Z. Yang, A. Yener, K. K. Wong, and C.-B. Chae, "Beyond transmitting bits: Context, semantics, and task-oriented communications," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 1, pp. 5–41, 2022.
- [77] K. Chowdhary and K. Chowdhary, "Natural language processing," *Fundamentals of Artificial Intelligence*, pp. 603–649, 2020.
- [78] D. Khurana, A. Koli, K. Khatter, and S. Singh, "Natural language processing: State of the art, current trends and challenges," *Multimedia Tools and Applications*, vol. 82, no. 3, pp. 3713–3744, 2023.
- [79] R. Szeliski, *Computer Vision: Algorithms and Applications*. Springer Nature, 2022.
- [80] P. Jiang, C.-K. Wen, S. Jin, and G. Y. Li, "Wireless semantic communications for video conferencing," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 1, pp. 230–244, 2022.
- [81] L. Yan, Z. Qin, R. Zhang, Y. Li, and G. Y. Li, "Resource allocation for text semantic communications," *IEEE Wireless Communications Letters*, vol. 11, no. 7, pp. 1394–1398, 2022.
- [82] H. Xie and Z. Qin, "A lite distributed semantic communication system for Internet of things," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 1, pp. 142–153, 2020.
- [83] W. Yang, Z. Q. Liew, W. Y. B. Lim, Z. Xiong, D. Niyato, X. Chi, X. Cao, and K. B. Letaief, "Semantic communication meets edge intelligence," *IEEE Wireless Communications*, vol. 29, no. 5, pp. 28–35, 2022.
- [84] H. Tong, Z. Yang, S. Wang, Y. Hu, O. Semiari, W. Saad, and C. Yin, "Federated learning for audio semantic communication," *Frontiers in Communications and Networks*, vol. 2, p. 734402, 2021.
- [85] Z. Han and K. R. Liu, *Resource Allocation for Wireless Networks: Basics, Techniques, and Applications*. Cambridge University Press, 2008.
- [86] Y. Teng, M. Liu, F. R. Yu, V. C. Leung, M. Song, and Y. Zhang, "Resource allocation for ultra-dense networks: A survey, some research issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2134–2168, 2018.
- [87] R. Zheng and C. Hua, *Sequential Learning and Decision-Making in Wireless Resource Management*. Springer, 2016.
- [88] E. Hossain, M. Rasti, and L. B. Le, *Radio Resource Management in Wireless Networks: An Engineering Approach*. Cambridge University Press, 2017.
- [89] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 6, pp. 1728–1767, 2022.
- [90] F. Dong, F. Liu, Y. Cui, W. Wang, K. Han, and Z. Wang, "Sensing as a service in 6G perceptive networks: A unified framework for ISAC resource allocation," *IEEE Transactions on Wireless Communications*, vol. 22, no. 5, pp. 3522–3536, 2022.
- [91] A. Liu, Z. Huang, M. Li, Y. Wan, W. Li, T. X. Han, C. Liu, R. Du, D. K. P. Tan, J. Lu *et al.*, "A survey on fundamental limits of integrated sensing and communication," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 994–1034, 2022.
- [92] S. Wang, W. Dai, H. Wang, and G. Y. Li, "Robust waveform design for integrated sensing and communication," *IEEE Transactions on Signal Processing*, 2024.
- [93] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2322–2358, 2017.
- [94] X. Ge, F. Yang, and Q.-L. Han, "Distributed networked control systems: A brief overview," *Information Sciences*, vol. 380, pp. 117–131, 2017.

- [95] Q. Luo, S. Hu, C. Li, G. Li, and W. Shi, "Resource scheduling in edge computing: A survey," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 4, pp. 2131–2165, 2021.
- [96] H. Djigal, J. Xu, L. Liu, and Y. Zhang, "Machine and deep learning for resource allocation in multi-access edge computing: A survey," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 4, pp. 2449–2494, 2022.
- [97] Z. Xu, Y. Wang, J. Tang, J. Wang, and M. C. Gursoy, "A deep reinforcement learning based framework for power-efficient resource allocation in cloud RANs," in *2017 IEEE International Conference on Communications (ICC)*. IEEE, 2017, pp. 1–6.
- [98] L. Liang, H. Ye, and G. Y. Li, "Spectrum sharing in vehicular networks based on multi-agent reinforcement learning," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 10, pp. 2282–2292, 2019.
- [99] H. Ye, G. Y. Li, and B.-H. F. Juang, "Deep reinforcement learning based resource allocation for V2V communications," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 3163–3173, 2019.
- [100] N. Zhao, Y.-C. Liang, D. Niyato, Y. Pei, M. Wu, and Y. Jiang, "Deep reinforcement learning for user association and resource allocation in heterogeneous cellular networks," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5141–5152, 2019.
- [101] X. Xiong, K. Zheng, L. Lei, and L. Hou, "Resource allocation based on deep reinforcement learning in IoT edge computing," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 6, pp. 1133–1146, 2020.
- [102] K. Xu, N. Van Huynh, and G. Y. Li, "Distributed-training-and-execution multi-agent reinforcement learning for power control in hetnet," *IEEE Transactions on Communications*, 2023.
- [103] M. Lee, G. Yu, and G. Y. Li, "Learning to branch: Accelerating resource allocation in wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 1, pp. 958–970, 2019.
- [104] —, "Graph embedding-based wireless link scheduling with few training samples," *IEEE Transactions on Wireless Communications*, vol. 20, no. 4, pp. 2282–2294, 2020.
- [105] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, "Joint radar and communication design: Applications, state-of-the-art, and the road ahead," *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3834–3862, 2020.
- [106] W. Xu, Z. Yang, D. W. K. Ng, M. Levorato, Y. C. Eldar, and M. Debbah, "Edge learning for 5G networks with distributed signal processing: Semantic communication, edge computing, and wireless sensing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 17, no. 1, pp. 9–39, 2023.
- [107] D. Wen, X. Li, Y. Zhou, Y. Shi, S. Wu, and C. Jiang, "Integrated sensing-communication-computation for edge artificial intelligence," *IEEE Internet of Things Magazine*, vol. 7, no. 4, pp. 14–20, 2024.
- [108] B. Zhang and G. Y. Li, "White-box 3D-OMP-transformer for ISAC," *arXiv preprint arXiv:2407.02251*, 2024.
- [109] B. Zhang, Z. Qin, and G. Y. Li, "Compression ratio learning and semantic communications for video imaging," *IEEE Journal of Selected Topics in Signal Processing*, 2024.
- [110] S. Zhou and G. Y. Li, "FedGiA: An efficient hybrid algorithm for federated learning," *IEEE Transactions on Signal Processing*, vol. 71, pp. 1493–1508, 2023.
- [111] —, "Federated learning via inexact ADMM," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 8, pp. 9699–9708, 2023.
- [112] U. Demirhan and A. Alkhateeb, "Integrated sensing and communication for 6G: Ten key machine learning roles," *IEEE Communications Magazine*, vol. 61, no. 5, pp. 113–119, 2023.
- [113] S. H. Alsamhi, A. V. Shvetsov, S. Kumar, J. Hassan, M. A. Alhartomi, S. V. Shvetsova, R. Sahal, and A. Hawbani, "Computing in the sky: A survey on intelligent ubiquitous computing for UAV-assisted 6G networks and Industry 4.0/5.0," *Drones*, vol. 6, no. 7, p. 177, 2022.

- [114] V. A. Nugroho and B. M. Lee, "A survey of federated learning for mmwave massive MIMO," *IEEE Internet of Things Journal*, 2024.
- [115] H. Ye, L. Liang, and G. Y. Li, "Decentralized federated learning with unreliable communications," *IEEE Journal of Selected Topics in Signal Processing*, vol. 16, no. 3, pp. 487–500, 2022.
- [116] O. Wang, S. Zhou, and G. Y. Li, "Few-shot learning for new environment adaptation," in *GLOBECOM 2023-2023 IEEE Global Communications Conference*. IEEE, 2023, pp. 351–356.
- [117] O. Wang, J. Gao, and G. Y. Li, "Learn to adapt to new environments from past experience and few pilot blocks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 9, no. 2, pp. 373–385, 2022.
- [118] Y. Liu, Z. Qin, and G. Y. Li, "Energy-efficient distributed spiking neural network for wireless edge intelligence," *IEEE Transactions on Wireless Communications*, 2024.
- [119] O. Wang, S. Zhou, and G. Y. Li, "BADMM: Batch ADMM for deep learning," *arXiv preprint arXiv:2407.01640*, 2024.
- [120] H. Cai, C. Gan, L. Zhu, and S. Han, "TinyTL: Reduce memory, not parameters for efficient on-device learning," *Advances in Neural Information Processing Systems*, vol. 33, pp. 11 285–11 297, 2020.