

基于深度学习的无线图像跨层优化传输方法综述

叶兴涛¹, 宋佳铭¹, 王敏^{2,3*}

¹赣南师范大学数学与计算机科学学院, 江西 赣州

²赣南师范大学智能制造与未来能源学院, 江西 赣州

³江西省教育厅数据科学与人工智能重点实验室, 江西 赣州

收稿日期: 2026年3月18日; 录用日期: 2026年4月17日; 发布日期: 2026年4月24日

摘要

随着5G/6G技术的快速发展, 高清图像、虚拟现实等新兴应用对无线图像传输的速率、实时性与可靠性提出了更高要求。然而, 无线信道的时变特性使传统分离式的信源与信道编码方法在恶劣环境下性能急剧下降。为此, 文章综述了基于深度学习的无线图像跨层优化传输方法的研究进展, 系统回顾了从传统分离传输到端到端跨层联合优化的演变过程, 重点介绍了自适应机制在无线图像跨层优化传输方法中的应用。同时, 也指出了当前研究在更加复杂信道、实际部署以及其他传输场景等方面面临的挑战。

关键词

无线图像传输, 跨层优化, 深度学习, 自适应机制

Review of Deep Learning-Driven Cross-Layer Optimization Methods Wireless Image Transmission

Xingtao Ye¹, Jiaming Song¹, Min Wang^{2,3*}

¹School of Mathematics and Computer Science, Gannan Normal University, Ganzhou Jiangxi

²School of Intelligent Manufacturing and Future Energy, Gannan Normal University, Ganzhou Jiangxi

³Key Laboratory of Jiangxi Province for Data Science and Artificial Intelligence, Ganzhou Jiangxi

Received: March 18, 2026; accepted: April 17, 2026; published: April 24, 2026

*通讯作者。

文章引用: 叶兴涛, 宋佳铭, 王敏. 基于深度学习的无线图像跨层优化传输方法综述[J]. 计算机科学与应用, 2026, 16(4): 274-286. DOI: 10.12677/csa.2026.164129

Abstract

With the rapid development of 5G/6G technologies, emerging applications such as high-definition images and virtual reality have imposed higher requirements for the rate, real-time performance, and reliability of wireless image transmission. However, the time-varying characteristics of wireless channels cause the performance of traditional separated source and channel coding methods to drop sharply in harsh environments. Therefore, this article reviews the research progress of wireless image cross-layer optimization transmission methods based on deep learning, systematically examines the evolution process from traditional separated transmission to end-to-end joint optimization, and focuses on introducing the application of adaptive mechanisms in wireless image cross-layer optimization transmission methods. At the same time, it also points out the challenges faced by current research in more complex channels, actual deployment, and other transmission scenarios.

Keywords

Wireless Image Transmission, Cross-Layer Optimization, Deep Learning, Adaptive Transmission

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着 5G 移动通信技术的规模化商用以及 6G 技术的前瞻性布局,无线通信网络正以前所未有的深度和广度渗透到社会生产的各个角落。高清视频[1]、虚拟现实[2]、自动驾驶[3]等新兴应用对无线图像视频传输提出了更为严苛的要求,在传输速率、可靠性、实时性和用户体验方面提出了极致挑战。然而,无线信道受如路径损耗、多径衰落、噪声干扰等因素导致其状态具有时变性[4],从而成为高质量无线图像传输的关键。因此,探索高效、鲁棒的无线图像传输理论与技术,已成为通信与信息处理领域亟待解决的关键问题。

传统的无线图像数字传输系统,比如 JPEG [5]、H.264/AVC [6]和 HEVC 标准[7]等,遵循香农分离定理[8],采用“信源-信道”分离架构,信源编码致力于最大限度地去除冗余信息从而压缩图像数据,信道编码则通过添加受控冗余来抵抗信道噪声和衰落从而保证数据传输的可靠性。这种模块化设计思想极大地简化了系统设计的复杂度,并在过去数十年间取得了巨大成功。然而,香农分离定理的理想条件,如无限码长、无差错信源编码、信道统计特性精确已知等,在实际系统中难以满足。另外,这种分离设计在高压缩率或恶劣信道条件下,系统性能会急剧恶化,呈现“悬崖效应”。

为克服分离式架构的固有缺陷,信源信道联合编码(Joint source-channel coding, JSCC)思想应运而生。JSCC 通过打破信源编码与信道编码之间的壁垒,进行跨层联合优化设计,直接以最大化接收端图像质量为目标。传统的 JSCC 方案,如基于不等差错保护的编码[9]、联合量化调制[10]等,在一定程度上提升了系统性能,但通常依赖于复杂的数学模型和手工设计的特征,难以捕捉图像内容的复杂语义结构和信道的动态统计特性,导致其性能提升有限且泛化能力不足。另外一种代表性 JSCC 实现方案是伪模拟传输

SoftCast [11], 通过变换对输入图像进行处理, 并将系数直接调制为密集的星座图进行传输, 无须常规的量化、熵编码和信道编码。

近年来, 以深度学习为代表的人工智能技术取得了突破性进展, 为解决上述挑战学者们提出了新的语义通信范式。在该范式下, 编码器被设计为一个深度神经网络, 直接将原始图像像素映射为适合在物理信道上传输的符号; 解码器则作为另一个神经网络, 完成从接收符号到重建图像的逆映射。这种基于深度学习的信源信道联合编码(Deep learning-based JSCC, DeepJSCC)方案, 由 Bourtsoulatze 等人[12]开创性地提出并验证了其可行性, 实现了从信源到信道的真正意义上的跨层优化。但是, DeepJSCC 方案需要在特定信噪比(Signal-to-Noise Ratio, SNR)下进行训练。当实际信道 SNR 与训练 SNR 不匹配时, 性能会急剧下降。

于是, DeepJSCC 的研究重心开始转向如何赋予模型自适应的能力。早期的研究致力于让单个模型能适应一定范围内波动的信噪比[13] [14]。随后, 带宽自适应的研究[15] [16]也取得了进展, 如 Deep-JSCC-1 [16]通过训练编码器生成一个渐进式的符号流, 解码端则根据实时带宽截取前端最重要部分进行解码, 巧妙地实现了带宽自适应。与此同时, 注意力机制等方法的引入, 使得网络能够减少参数量、扩展感受野[17], 增强了对信道状态的感知与利用。随后, 能够同时应对信噪比与带宽联合变化的方案成为新的趋势。研究者们通过设计精巧的调制模块[18] [19]、动态调整网络层结构[20], 推动单一模型覆盖更广泛的信道区间, 向实用化迈出了坚实一步。随着研究的不断推进, DeepJSCC 研究的视野也从理想的加性高斯白噪声(Additive white gaussian noise, AWGN)信道, 转向更具现实意义的衰落信道[21]-[23]与 MIMO 系统[24]-[26]。

目前, DeepJSCC 的相关研究已经涌现了大量创新成果, 如模型架构的革新、动态信道的鲁棒性研究以及多样化场景与任务的拓展。但现有综述文献却表现出明显的滞后性, 难以全面、准确地反映该领域的当前进展和发展趋势。比如中国科学技术大学穆天杰等人[27]综述了基于深度学习的联合信源信道优化研究, 但聚焦于 2020 年之前的相关工作。因此, 本文系统综述基于深度学习的无线图像跨层优化传输方法的最新进展, 首先回顾从传统分离设计到深度学习驱动联合优化的演进脉络, 继而分析深度联合编码的核心机制与自适应策略, 最后探讨该领域面临的开放性挑战与发展趋势。通过梳理这一交叉领域的技术轨迹, 我们期望为后续研究提供参考, 推动高效可靠无线图像传输系统的应用。

2. 基于深度学习的分离信源信道传输方法

受深度学习在计算机视觉领域成功应用的启发, 有学者利用深度学习研究图像视频压缩传输。围绕“变换 - 量化 - 熵编码”技术架构, 这类方法用深度学习网络模型替换或优化其中的系统部件, 研发端对端的基于深度学习图像压缩算法, 并取得了比传统数字传输方法更优异的性能。

(一) 传统无线图像数字传输

经过几十年的发展, 视频编码的技术已经相当成熟, 众多图像视频编码标准被相继提出, 如 JPEG [5]、H.264/AVC [6], 以及高效视频编码 HEVC 标准[7]等。这些传统的数字传输系统采用分离信源信道编码模型, 以率失真函数为优化目标, 以“预测 - 变换 - 量化 - 熵编码”为技术框架, 实现去除冗余信息并达到信源压缩目的。图 1 给出了传统无线图像传输框架, 发送端首先对图像进行变换、量化和熵编码, 生成压缩后的比特流; 然后系统独立地进行信道编码以对抗传输过程中的噪声与干扰。接收端则按相反顺序执行信道解码、熵解码、反量化与逆变换, 最终重建出图像。为了提高自适应能力, 标准图像视频数字传输系统在信源和信道都有相应的自适应策略, 比如可伸缩视频编码(Scalable video coding, SVC) [28]和自适应编码调制(Adaptive modulation coding, AMC) [29]。由于门限效应和悬崖效应的存在, 数字视频通信系统的自适应能力较差。

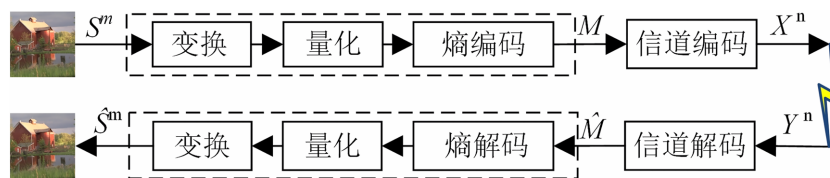


Figure 1. Traditional wireless image digital transmission framework
图 1. 传统无线图像数字传输框架

(二) 基于深度学习的图像压缩传输

2017年, 纽约大学的 Johannes Ballé 等人[30]开创性地提出了端到端优化的深度学习图像压缩框架。随后, 基于深度学习的图像压缩方法迅速发展, 涌现出多种技术路径。从深度学习模型结构角度可以划分为基于卷积神经网络(Convolutional neural network, CNN)方法、基于循环神经网络(Recurrent neural network, RNN)方法, 以及基于生成对抗网络(Generative adversarial network, GAN)方法。图 2 给出了基于深度学习的图像压缩传输框架, 与传统图像数字传输框架相比, 区别在于信源编码直接被深度编解码器替换。

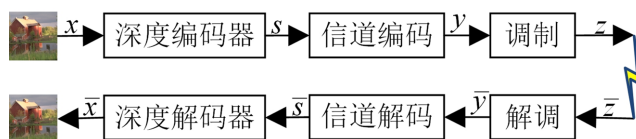


Figure 2. Framework for image compression and transmission based on deep learning
图 2. 基于深度学习的图像压缩传输框架

Ballé 等人[31]针对图像压缩任务优化了 CNN 中的激活函数, 提出了广义除性归一化(Generalized divisive normalization, GDN)方法, 该函数具有良好的去相关能力, 有助于提升压缩效率。进一步地, 通过端到端训练共同优化率失真性能, 研究人员相继提出了基于超先验的熵编码模型[31]以及结合超先验与自回归模型的混合方案[32], 这些方法在压缩性能上超越了传统图像编码方法 BPG。

RNN 因其能够有效捕捉数据中的长短期依赖关系, 在图像压缩中表现出显著潜力。Google 的 Toderici 团队[33]于 2015 年提出了基于长短期记忆(Long short-term memory, LSTM)网络的变压缩率图像压缩框架。随后, 该团队[34]在 2016 年进一步引入门控循环单元和残差结构, 提出一种支持全分辨率的有损压缩方法, 并结合熵编码技术, 在相同压缩率下实现了优于经典 JPEG 标准的性能。

GAN 凭借其强大的生成能力, 也为图像压缩提供了新的思路。2019 年, Agustsson 等人[35]提出一种基于 GAN 的图像压缩算法, 该算法依据语义标签对图像部分区域进行选择性的编码, 在解码阶段结合解码重建与生成器补充, 显著提升了低码率下的压缩效率, 性能优于 BPG 算法。

除了图像的深度图像压缩应用, 基于深度学习的视频编码也得到了发展。比如 Lu 等人[36]提出了一种端到端的深度视频编码方案, 对输入视频的每一帧进行压缩, 并使用光流估计模块来获取当前帧与前一帧之间的运动信息, 并使用两个自编码器压缩预测残差和运动信息取得了比 H.264 标准更低的压缩率。Felix 等人[37]提出了一种全连接网络用于帧内预测, 使用多模型的训练策略以提高压缩率。针对视频传输, Tung 等人[38]提出了端到端信源信道跨层优化系统 DeepWiVe。

在端到端优化网络的启发下, 研究者们提出了更多的图像压缩方案。这些工作主要围绕变换模块, 量化模块, 熵编码等方面展开了研究。尽管端到端的图像压缩模型能够自动调整相应参数, 但相对于传统的图像压缩算法来说压缩耗费时间, 训练耗费时间上有待提升。但是, 上述基于深度学习的图像视频压缩传输工作只考虑了信源如何压缩, 并没有考虑到无线信道时变的问题。

(三) 对比与讨论

首先,从模型架构的演进来看,不同网络结构在压缩任务中呈现出较为鲜明的差异化特性。基于 CNN 的方法,在率失真性能上确立了新的标杆。RNN 方法则凭借其记忆机制,在处理可变码率压缩和长距离依赖建模上展现出独特潜力,但受限于串行计算特性,其推理速度往往难以满足实时性要求。GAN 方法另辟蹊径,利用生成能力弥补低码率下的信息丢失,虽在主观视觉质量上表现优异,但在像素级保真度(如 PSNR 指标)上仍存在波动风险。

其次,先前研究在“信源压缩”与“信道传输”之间存在明显的割裂。上述工作大多遵循经典的分离定理假设,将研究重心置于信源端的冗余去除,默认传输信道是理想的或静态的。然而,在实际无线通信场景中,信道状态信息具有高度的时变性和不可预测性。当信道质量下降时,固定码率的压缩码流极易因误码累积导致重建质量急剧恶化。这一局限性也进一步印证了向联合信源信道编码(JSCC)及信道自适应技术转型的必要性。

3. 基于深度学习的图像信源信道联合优化传输方法

为了克服传统图像数字传输局限,深度学习技术为无线图像传输跨层优化带来了新变革。当前 DeepJSCC 是当前实现语义通信领域主流且成功的方法之一,从根本上挑战了香农信息论中经典的信源信道分离定理。文献[39]给出了 DeepJSCC 性能与 SSCC 性能的比较图,如图 3 所示,分离式传输架构存在典型的阈值效应:系统仅在特定 SNR_{test} 下最优,一旦实际信噪比跌破该阈值,便触发“悬崖效应”导致性能暴跌。相反,DeepJSCC 表现出优异的平滑过渡能力。针对特定 SNR_{test} 训练的模型呈现出“优雅退化”特征:当测试 SNR 低于训练值时,性能平缓下降;当测试 SNR 升高时,性能则逐步改善,从而避免了传统方案的性能突变问题。然而,图 3 也揭示了一个关键局限:尽管具备平滑退化特性,DeepJSCC 仅在训练与测试信噪比匹配时达到最佳性能。若要实现所有信噪比下的最优凸包性能,理论上需要针对每个信噪比值分别训练并存储神经网络,这种“一网一噪”的模式在实际部署中面临巨大的存储与计算开销瓶颈。

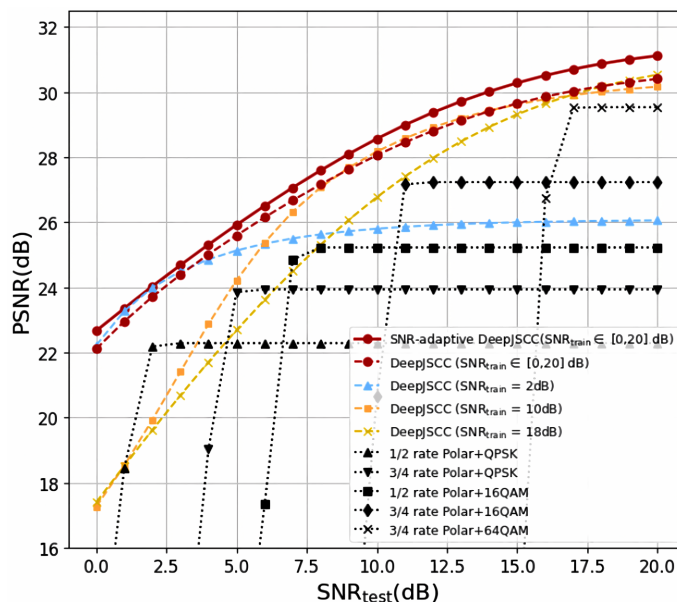


Figure 3. Comparison of the performance of DeepJSCC and SSCC on the CIFAR-10 dataset transmitted over additive white Gaussian noise channels with different signal-to-noise ratio values

图 3. 在不同信噪比值的加性高斯白噪声信道上传输的 CIFAR-10 数据集上,对 DeepJSCC 性能与 SSCC 性能的比较

针对上述“单一模型难以兼顾多场景”的挑战，当前 DeepJSCC 的研究主要沿两个方向演进：一是受深度学习启发，考虑模型归纳偏置和计算特性的差异导致重建图像在像素保真度、结构相似性、语义连贯性上产生系统性差异，研究不同网络模型结构以精确表达无线图像语义。二是考虑传统固定信道条件的 DeepJSCC 模型在时变信道中性能急剧下降，将研究重心放在信道自适应机制上。因此，如何使 DeepJSCC 系统具备对信道状态的感知与自适应能力，成为连接其理论优势与实际部署的关键桥梁。本节将介绍基于神经网络以及信道自适应的深度联合信源信道编码方案，阐述对应的设计思路。

图 4 给出了基于深度学习的信道自适应的无线图像传输框架，发送端实时获取信道的信噪比与带宽信息，并驱动深度联合编码网络动态调整其编码策略与符号映射方式，从而在压缩效率与抗干扰能力之间取得实时最优权衡。接收端同步获取信道状态，引导深度解码网络调整其重建策略，显式利用信道信息以增强对噪声的鲁棒性，最终输出高质量的重建图像。此后，学者们的研究重点逐渐转向提升系统的自适应能力。表 1 列出了近年来针对无线图像跨层优化传输的 DeepJSCC 方法。从自适应角度大致分为无自适应、SNR 自适应、带宽自适应，以及 SNR + 带宽自适应方法。

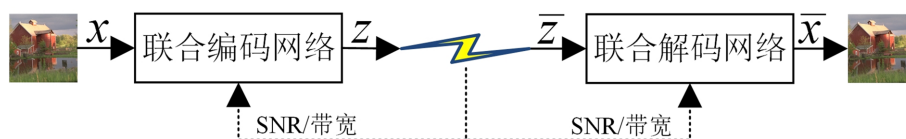


Figure 4. Wireless image transmission framework based on deep learning for channel adaptation

图 4. 基于深度学习的信道自适应的无线图像传输框架

Table 1. List of DeepJSCC image transmission methods

表 1. DeepJSCC 图像传输方法列表

序号	模型	年份	自适应类型	信道类型	网络结构
1	Deep-JSCC [12]	2019	无	AWGN 信道	CNN
2	Deep-JSCC- f [40]	2020	无	AWGN 信道	CNN
3	OFDM-Guided DeepJSCC [41]	2022	无	衰落信道	CNN
4	DeepJSCC-MIMO [42]	2022	无	MIMO 信道	ResNet
5	SNR-Adaptive Model [13]	2021	SNR 自适应	AWGN 信道	Autoencoder
6	ADJSCC [14]	2022	SNR 自适应	AWGN 信道	Attention
7	CBJSCC [17]	2023	SNR 自适应	AWGN 信道	CNN + Attention
8	CFA-JSCC [21]	2025	SNR 自适应	衰落信道	CNN
9	SwinSIT [22]	2025	SNR 自适应	衰落信道	Swin Transformer
10	MIMO-DJSCC [43]	2024	SNR 自适应	MIMO 信道	CNN + Attention
11	Deep-JSCC- l [16]	2021	带宽自适应	AWGN 信道	RNN
12	entropy-awareDeepJSCC [15]	2023	带宽自适应	AWGN 信道	CNN
13	DeepJSCC- $l++$ [44]	2023	SNR + 带宽自适应	AWGN 信道	Swin Transformer
14	ADWITT [18]	2024	SNR + 带宽自适应	AWGN 信道	Swin Transformer

续表

15	SwinJSCC [19]	2024	SNR + 带宽自适应	AWGN 信道	Swin Transformer
16	DD-JSCC [20]	2025	SNR + 带宽自适应	AWGN 信道	CNN + Attention
17	MambaJSCC [46]	2024	SNR + 带宽自适应	AWGN 信道	Mamba
18	DeepJSCC-ViT-f [47]	2024	SNR + 带宽自适应	AWGN 信道	ViT
19	CA-JSCC [23]	2022	SNR + 带宽自适应	衰落信道	Autoencoder
20	JFPNet [24]	2025	SNR + 带宽自适应	MIMO 信道	CNN + Attention
21	Weixuan Chen [25]	2025	SNR + 带宽自适应	MIMO 信道	CNN
22	Zhang G [26]	2023	SNR + 带宽自适应	MIMO 信道	CNN

(一) 基于深度学习的 DeepJSCC

2019年, D. B. Kurka 等人[12]提出了一个基于自编码器网络的信源信道联合优化的无线图像传输方法 Deep-JSCC, 将传统的信源信道编解码由编码网络和解码网络来完成。结果显示 Deep-JSCC 在低信噪比 AWGN 信道下比 JPEG 或者 JPEG2000 表现出更好的性能。随后, 涌现出一系列在反馈信道、OFDM 和 MIMO 的拓展 Deep-JSCC 方法。比如, 文献[40]则考虑信道输出反馈, 并提出了无线图像传输方法 Deep-JSCC-f。文献[41]将 deep-JSCC 拓展到具有非线性信号截断的多径衰落信道, 集成了一个用于应对多径衰落信道的 OFDM 数据通路。文献[42]将 MIMO 通信的经典理论融入端到端的 DeepJSCC 框架中, 设计并比较两种针对 MIMO 瑞利衰落信道的 DeepJSCC 方案, 一种旨在最大化分集增益, 另一种旨在最大化复用增益。然而, DeepJSCC 系列方法在实际部署过程中存在关键瓶颈, 其性能过度依赖于针对特定 SNR 训练的专用模型, 难以适应真实无线环境中动态变化的信道条件。

(二) SNR 自适应的 DeepJSCC

针对 AWGN 信道不同 SNR, M. Ding 等人[13]提出一种首个 SNR 自适应的 DeepJSCC 方案, 使单一模型能够适应不同 SNR 条件, 实现多用户场景下的图像传输, 各用户信道 SNR 不同, 但共享同一解码器。实验结果表明, 该方法对不同的信噪比具有较好的适应性, 并且对译码器信噪比估计中的噪声具有较强的鲁棒性。Xu 等人[14]提出了基于注意力机制的信源信道联合优化方法 ADJSCC, 训练一个网络模型能够自适应信道状态。具体而言, ADJSCC 在编解码网络模型中使用了特征学习模块和注意力特征模块。注意力特征模块以特征学习模块的输出和 SNR 作为输入, 学习产生尺度因子, 然后根据尺度因子校正特征学习模块的输出特征。

针对衰落信道, 文献[22]提出了一种基于 Swin Transformer 的端到端优化的深度学习联合源信道编码方案 SwinSIT, 通过引入 SNR 感知模块和基于 CNN 的信道估计与补偿模块, 在衰落信道中实现了高效的图像语义传输。2023年, Hongjie Yuan 等人[17]提出了一种不需要显式获取 SNR 信息, 神经网络自动适应不同信道条件的框架 CBJSCC, 结合卷积注意力和自注意力的网络结构, 扩大感受野, 增强长距离依赖建模能力。2025年, Hongjie Yuan 等人[21]提出了一种双阶段信道自适应的 DeepJSCC 框架 CFA-JSCC, 采用粗细两个阶段对信道变化进行适应, 粗阶段根据平均 SNR 调整编码策略, 细阶段则根据瞬时 SNR 动态细化编码策略, 并利用有限集的信道质量指示符(Channel Quality Indicator, CQI)和基于强化学习的 CQI 选择策略来学习信道 SNR 的映射, 通过粗细粒度调整编码策略, 提高系统在时变信道环境下的性能。

针对 MIMO 信道, Jiang 等人[43]提出了一种基于注意力模块的 MIMO-DJSCC 方案, 通过在串行和

并行结构中引入注意力模块,使系统能够根据 SNR 调整信息输出量,同时采用奇异值分解(Singular Value Decomposition, SVD)基线的预编码方案,将 MIMO 信道分解为多个子信道。

(三) 带宽自适应的 DeepJSCC

针对不同信道带宽,2021年,D. Gunduz 团队[16]针对上述 DeepJSCC 带宽固定的问题,提出一种带宽自适应的 DeepJSCC,能够自适应地不同带宽下工作,而无需为每个带宽训练一个单独的模型。核心思想是训练一个编码器产生一个无限长的信道符号流,解码器则根据当前可用带宽,只取前 L 个符号进行解码。训练时,随机截取不同长度的符号序列作为输入,迫使网络学会组织信息流,将最重要的信息放在最前面。该方法成功实现了单一模型支持连续带宽自适应。针对不同的码率,2023年,Chen 等人[15]提出了提出一种支持多码率的自适应 DeepJSCC 方法,通过引入策略网络,实现根据信道 SNR 和图像内容动态调整传输速率,仅使用单一模型即可覆盖多个码率点,提升带宽利用效率并简化系统部署。研究结论表明,策略网络成功学习到合理的速率分配策略,单一自适应模型性能与多个固定码率模型相当。

(四) SNR + 带宽自适应的 DeepJSCC

2023年,D. Gunduz 团队[44]提出了一种基于 Swin Transformer 的 DeepJSCC 方案,通过计算不同带宽比对应的重建损失,根据它们各自的重建质量动态地为不同带宽比的损失分配不同的权重,实现了仅使用一个模型适应多个目标带宽比以及不同的 SNR。2024年,Li 等人[18]引入了一种先进的信道自适应模块,该模块借鉴了深度学习方法及变分信息瓶颈(Variational Information Bottleneck, VIB)的自适应调制原理设计了速率调制模块和 SNR 调制模块。这一创新增强了原始的 WITT 模型[45],从而开发出了自适应无线图像传输架构 ADWITT。与 ADJSCC 与 WITT 相比,ADWITT 在推理速度以及计算复杂度上有了较大提升。如图 5 所示, Yang 等人[19]改进了以上两个模块,根据信道状态信息和目标传输率缩放潜在表示,实现了单一模型同时适应不同 SNR 和信道带宽比(Channel Bandwidth Ratio, CBR)。DeepJSCC 部分模型在参数大小与性能之间的比较图如图 6 所示, SwinJSCC 在减少参数大小的同时,保持了较高的性能。实验结果表明,与最先进的工程 BPG + 5G LDPC 编码传输系统相比, SwinJSCC 具有更快的端到端编码速度,并且性能更优或相近。2025年, Avi Deb Raha 等人[20]提出一种动态深度联合信源信道编码架构 DD-JSCC,核心思想是根据发射机和接收机的性能、功率限制、压缩比和当前信道条件实时动态调整其层结构。这种适应性是通过分层激活机制与通过顺序随机训练的隐式正则化相结合实现的,有效降低了组合复杂性,防止过拟合,并确保了不同配置下的一致性特征表示,实现编码器和解码器层结构的实时动态调整,以适应不同 SNR、压缩比(Compression Ratio, CR)、设备计算能力和功耗限制。

此外,研究者还考虑从信道状态信息(Channel State Information, CSI)及其反馈中实现对信道的自适应。例如, Wu 等人[46]提出的 CSI 嵌入方法,在编码器和解码器中都部署了一个共享的 CSI 编码模块,对每个 VSSM-CA 块进行编码和注入 CSI,提高了单个模型对变化的信道条件的适应性。2024年, Wu 团队[47]在之前的工作上提出了一种利用接收方反馈并从反馈信号中获取信道状态信息的方法,实现对不同的信道条件进行匹配。针对衰落信道, Gunduz 团队[23]提出了一种基于学习的信道(多径衰落信道)自适应联合源和信道编码 CA-JSCC 方案,核心思想是使用一个端到端的自动编码器架构,具有双注意力机制,并采用正交频分复用(Orthogonal Frequency Division Multiplexing, OFDM)传输,通过利用估计的 CSI 来适应信道增益和噪声功率的变化。

针对 MIMO 信道,文献[26]首先提出了适用于 MIMO 信道的 JSCC 系统,核心思想是使用自适应 CSI 反馈预编码方案,通过开发性能评估器预测每幅图像的重建质量,对预测质量高的图像自适应减少 CSI 反馈开销。实验结果表明,该方案在大幅降低反馈开销的同时,显著提升了图像传输性能。最近, Guo 等人[24]提出了一个基于深度学习的频分双工系统中的联合 CSI 反馈和多用户预编码方法,旨在以端到端

的方式最大化下行链路总速率性能。具体来说,使用深度联合源信道编码技术压缩 CSI 矩阵的特征向量。这种压缩方法增强了反馈 CSI 信息在反馈信道退化时的鲁棒性。设计了一个联合多用户预编码模块和一个功率分配模块,根据反馈 CSI 信息调整用户的多用户预编码方向和预编码功率。Chen 等人[25]提出了一种新颖的语义通信系统,通过结合熵感知和信道自适应机制,有效提升了多用户多输入多输出(MU-MIMO)衰落信道上传输的率失真性能和鲁棒性。

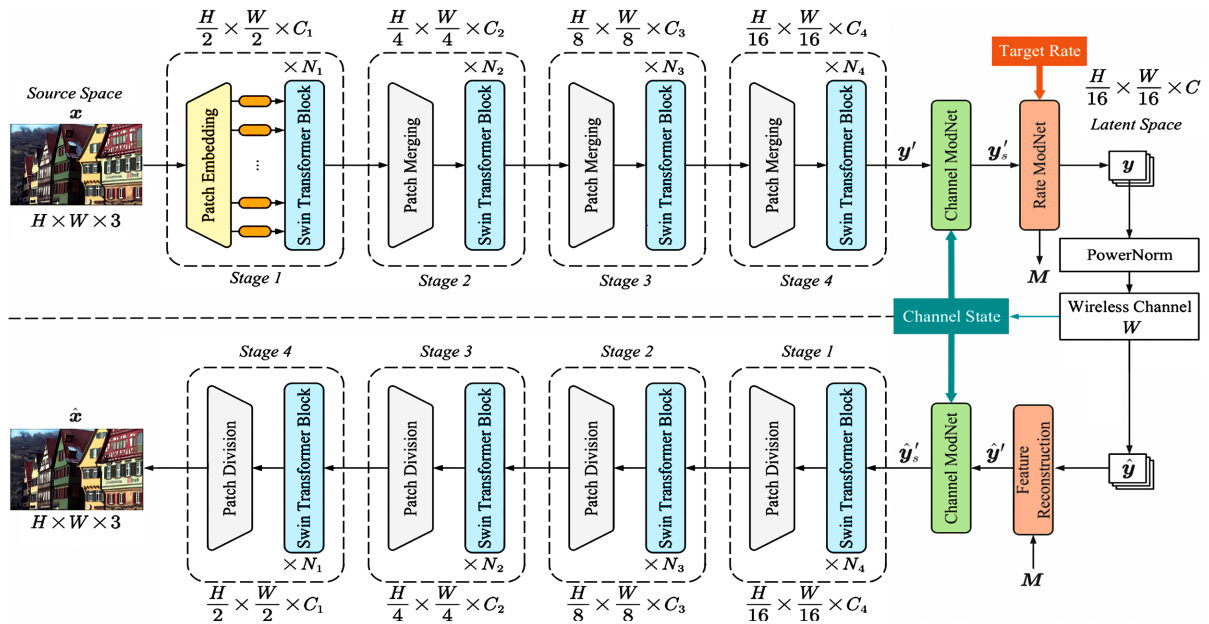


Figure 5. System architecture of SwinJSCC
图 5. SwinJSCC 系统架构

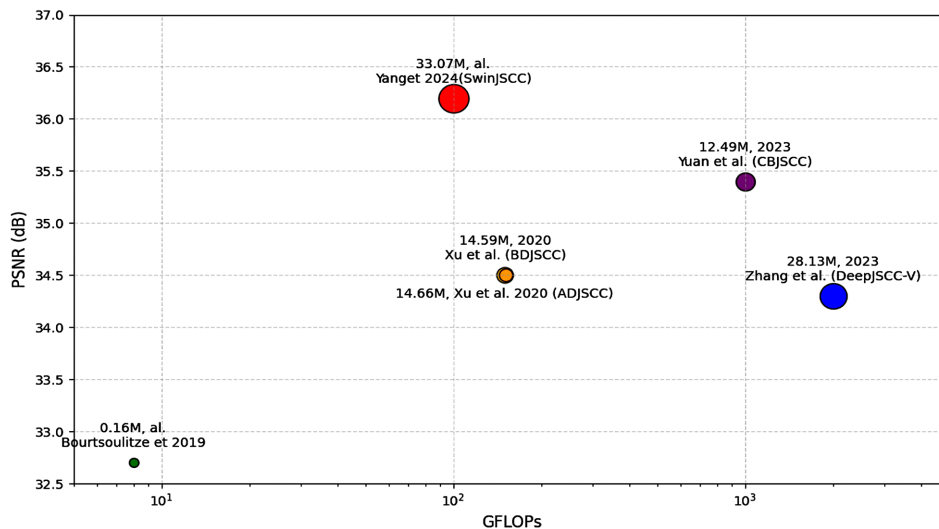


Figure 6. Reconstructed image PSNR performance versus processing complexity of different methods for Kodak dataset over AWGN channel where SNR = 7 dB and channel bandwidth ratio (CBR) = 1/6. Complexity measures include the floating point operations and the size of model parameters
图 6. 不同方法在 AWGN 信道下对柯达数据集的重建图像 PSNR 性能与处理复杂度的关系,其中信噪比(SNR) = 7 分贝,信道带宽比(CBR) = 1/6。复杂度的衡量指标包括浮点运算次数和模型参数的大小

(五) 对比与讨论

首先,从技术结构脉络来看,DeepJSCC的发展清晰地呈现出从“突破分离定理桎梏”向“多维自适应融合”跨越的趋势。早期的DeepJSCC研究(第一阶段)主要致力于验证端到端优化的可行性,成功打破了传统分离架构的性能瓶颈,特别是在低SNR区域展现出显著优势。然而,其“一网一噪”的静态训练模式导致了存储开销与环境的。随后的研究分别在SNR和带宽两个维度上寻求突破,通过引入注意力机制、特征调制等策略,实现了单一维度上的动态适配。最新的研究(第四阶段)则致力于解决SNR与带宽的联合自适应问题,并引入CSI反馈机制,标志着DeepJSCC从“特定场景优化”迈向了“全场景通用”的新阶段。

其次,在自适应机制的实现上,特征调制与注意力机制已成为主流技术范式之一。对比SNR自适应与带宽自适应的方法可以发现,两者在底层逻辑上具有高度一致性:即通过引入外部状态信息(如SNR、带宽比)作为控制信号,动态调整网络的特征表达。在SNR自适应方面,从早期简单的噪声注入或特征缩放,再到CFA-JSCC的粗细双阶段调节,对信道变化的响应精度不断提升;这些技术路径的共同点在于,它们赋予了神经网络根据物理信道约束“自主决策”的能力,实现了语义信息与信道容量的高度匹配。

再次,网络架构正经历从传统CNN向Transformer的结构转变。早期的DeepJSCC主要依赖CNN架构,虽然计算效率较高,但在捕捉长距离语义依赖和处理高分辨率图像时存在局限。近年来,以SwinJSCC、SwinSIT为代表的方案引入了Swin Transformer,利用其全局感受野和自注意力机制,不仅显著提升了图像重构质量,更适配于处理变长序列和多尺度特征融合任务。

最后,面向实际部署的考量正逐渐成为研究的核心驱动力之一。虽然自适应技术解决了模型泛化问题,但引入复杂的注意力模块或大型Transformer架构显著增加了推理时延和计算复杂度。近期的研究(如DD-JSCC、ADWITT)开始关注“性能-复杂度”的平衡,通过动态深度调整或轻量化模块设计,试图在保证自适应能力的同时降低对终端算力的需求。

4. 未来发展趋势

现有的DeepJSCC方法平衡了计算复杂度与系统性能,但仍然面临着实际无线时变信道的挑战。因此,基于深度学习的信源信道联合编码仍需解决诸多问题。

首先,大多数DeepJSCC方案是基于理想化信道假设下(如AWGN信道)进行模型训练,在面对真实环境中普遍存在的时变信道衰落、多普勒频移、以及“语义级扰动”(如对抗样本、图像模糊、局部遮挡等)时,仍表现出一定的鲁棒性缺陷。此外,在多用户通信场景,我们甚至没有分离的理论最优性,DeepJSCC方案在面对多址和广播信道模型仍存在挑战。

其次,现有性能优异的DeepJSCC模型大都采用Transformer和扩散模型等技术,在基站之间的通信具有优势。但对于计算资源和能源受限的系统,比如物联网系统或者信道状态随时间快速变化的移动端之间的通信,这些大模型明显不适应。因此,未来应当研究复杂度低、参数量少的DeepJSCC模型,以适应适合物联网系统场景。另一方面,考虑现有云端计算资源优势,云边端协同的DeepJSCC方法是未来重要的研究方向,以便适配现有的通信系统。

最后,现有DeepJSCC方法主要针对图像和视频应用,未来可以拓展到虚拟现实、虚拟增强和点云数据传输。当涉及到传输高速率的多媒体内容时,需要依赖有线连接或者高质量的稳定无线链路。因此,研究大量数据稳定传输场景的DeepJSCC方法模型显得尤为重要。

5. 总结

随着无线通信场景的不断变化,传统的无线传输方案难以满足越来越复杂的传输需求。本文系统地

展示了基于深度学习的无线图像跨层优化传输技术, 从 SNR 自适应、带宽自适应、SNR + 带宽自适应几个方面梳理了 DeepJSCC 的研究进展, 讨论了 DeepJSCC 的未来发展趋势。从本文分析来看, 深度学习方法在基于 AWGN 信道的自适应研究中已有较为显著的成果, 这使得研究者们开始探索基于其他信道模型的 DeepJSCC 方案。然而, 现有研究仍面临若干挑战, 包括拓展到其他复杂信道、研究轻量级的 DeepJSCC 方案, 以及拓展到虚拟现实、虚拟增强和点云数据传输等其他通信任务。

参考文献

- [1] Wang, J., Al-Khalidi, A., Ahearn, S. and Wasige, E. (2021) 1080P HD Video Transmission Using RTD Transmitter. 2021 14th UK-Europe-China Workshop on Millimetre-Waves and Terahertz Technologies (UCMMT), Lancaster, 13-15 September 2021, 1-3. <https://doi.org/10.1109/ucmmt53364.2021.9569915>
- [2] Chaccour, C., Soorki, M.N., Saad, W., Bennis, M. and Popovski, P. (2022) Can Terahertz Provide High-Rate Reliable Low-Latency Communications for Wireless VR? *IEEE Internet of Things Journal*, **9**, 9712-9729. <https://doi.org/10.1109/jiot.2022.3142674>
- [3] Zeng, T., Semiari, O., Saad, W. and Bennis, M. (2019) Joint Communication and Control for Wireless Autonomous Vehicular Platoon Systems. *IEEE Transactions on Communications*, **67**, 7907-7922. <https://doi.org/10.1109/tcomm.2019.2931583>
- [4] Wang, M., Wu, J., Shi, S.F., Luo, C. and Wu, F. (2012) Fast Decoding and Hardware Design for Binary-Input Compressive Sensing. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, **2**, 591-603. <https://doi.org/10.1109/jetcas.2012.2220291>
- [5] Wallace, G.K. (1992) The JPEG Still Picture Compression Standard. *IEEE Transactions on Consumer Electronics*, **38**, 18-36. <https://doi.org/10.1109/30.125072>
- [6] Wiegand, T., Sullivan, G.J., Bjontegaard, G. and Luthra, A. (2003) Overview of the H.264/AVC Video Coding Standard. *IEEE Transactions on Circuits and Systems for Video Technology*, **13**, 560-576. <https://doi.org/10.1109/tcsvt.2003.815165>
- [7] Sullivan, G.J., Ohm, J., Han, W. and Wiegand, T. (2012) Overview of the High Efficiency Video Coding (HEVC) Standard. *IEEE Transactions on Circuits and Systems for Video Technology*, **22**, 1649-1668. <https://doi.org/10.1109/tcsvt.2012.2221191>
- [8] Shannon, C.E. (2001) A Mathematical Theory of Communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, **5**, 3-55. <https://doi.org/10.1145/584091.584093>
- [9] Ninkovic, V., Vukobratovic, D., Hager, C., et al. (2021) Autoencoder-Based Unequal Error Protection Codes. *IEEE Communications Letters*, **25**, 3575-3579. <https://doi.org/10.1109/lcomm.2021.3108845>
- [10] 朱勇刚, 陈亮, 朱义勇. 一种基于鲁棒矢量量化的信源-信道联合编码算法[J]. 解放军理工大学学报(自然科学版), 2005, 6(2): 118-122.
- [11] Cui, H., Xiong, R.Q., Luo, C., Song, Z.H. and Wu, F. (2015) Denoising and Resource Allocation in Uncoded Video Transmission. *IEEE Journal of Selected Topics in Signal Processing*, **9**, 102-112. <https://doi.org/10.1109/jstsp.2014.2338279>
- [12] Bourtsoulatzé, E., Burth Kurka, D. and Gunduz, D. (2019) Deep Joint Source-Channel Coding for Wireless Image Transmission. *IEEE Transactions on Cognitive Communications and Networking*, **5**, 567-579. <https://doi.org/10.1109/tccn.2019.2919300>
- [13] Ding, M., Li, J., Ma, M. and Fan, X. (2021) SNR-Adaptive Deep Joint Source-Channel Coding for Wireless Image Transmission. 2021 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, 06-11 June 2021, 1555-1559. <https://doi.org/10.1109/icassp39728.2021.9414037>
- [14] Xu, J., Ai, B., Chen, W., Yang, A., Sun, P. and Rodrigues, M. (2022) Wireless Image Transmission Using Deep Source Channel Coding with Attention Modules. *IEEE Transactions on Circuits and Systems for Video Technology*, **32**, 2315-2328. <https://doi.org/10.1109/tcsvt.2021.3082521>
- [15] Chen, W., Chen, Y., Yang, Q., Huang, C., Wang, Q. and Zhang, Z. (2023) Deep Joint Source-Channel Coding for Wireless Image Transmission with Entropy-Aware Adaptive Rate Control. 2023 *IEEE Global Communications Conference*, Kuala Lumpur, 4-8 December 2023, 2239-2244. <https://doi.org/10.1109/globecom54140.2023.10437482>
- [16] Kurka, D.B. and Gunduz, D. (2021) Bandwidth-Agile Image Transmission with Deep Joint Source-Channel Coding. *IEEE Transactions on Wireless Communications*, **20**, 8081-8095. <https://doi.org/10.1109/twc.2021.3090048>
- [17] Yuan, H., Xu, W., Wang, Y. and Wang, X. (2024) Channel-Blind Joint Source-Channel Coding for Wireless Image

- Transmission. *Sensors*, **24**, Article 4005. <https://doi.org/10.3390/s24124005>
- [18] Li, H., Xia, G., Yu, C., Zhang, Y. and Li, H. (2024) Adaptive Wireless Image Transmission Transformer Architecture for Image Transmission and Reconstruction. *Sensors*, **24**, Article 6772. <https://doi.org/10.3390/s24216772>
- [19] Yang, K., Wang, S., Dai, J., Qin, X., Niu, K. and Zhang, P. (2025) SwinJSCC: Taming Swin Transformer for Deep Joint Source-Channel Coding. *IEEE Transactions on Cognitive Communications and Networking*, **11**, 90-104. <https://doi.org/10.1109/tccn.2024.3424842>
- [20] Raha, A.D., Adhikary, A., Gain, M., Park, Y., Saad, W. and Hong, C.S. (2025). DD-JSCC: Dynamic Deep Joint Source-Channel Coding for Semantic Communications. *IEEE International Conference on Communications*, Montreal, 8-12 June 2025, 3754-3759. <https://doi.org/10.1109/icc52391.2025.11162089>
- [21] Li, H., Zhang, G., Zhou, K., *et al.* (2024) Coarse-to-Fine: A Dual-Phase Channel-Adaptive Method for Wireless Image Transmission. <https://arxiv.org/abs/2412.08211>
- [22] Salim, M.M., Abdalzaher, M.S., Muqaibel, A.H., *et al.* (2025) SNR-Aware Semantic Image Transmission with Deep Learning-Based Channel Estimation in Fading Channels. <https://arxiv.org/abs/2504.20557>
- [23] Wu, H., Shao, Y., Mikolajczyk, K. and Gunduz, D. (2022) Channel-Adaptive Wireless Image Transmission with OFDM. *IEEE Wireless Communications Letters*, **11**, 2400-2404. <https://doi.org/10.1109/lwc.2022.3204837>
- [24] Guo, Y., Chen, W., Xu, J., Li, L. and Ai, B. (2025) Deep Joint CSI Feedback and Multiuser Precoding for MIMO OFDM Systems. *IEEE Transactions on Vehicular Technology*, **74**, 1730-1735. <https://doi.org/10.1109/tvt.2024.3452409>
- [25] Chen, W., Yang, Q., Chen, Y., *et al.* (2025) Semantic Communication with Entropy-and-Channel-Adaptive Rate Control over Multi-User MIMO Fading Channels. <https://arxiv.org/abs/2501.15414>
- [26] Zhang, G., Hu, Q., Cai, Y., *et al.* (2023) Adaptive CSI Feedback for Deep Learning-Enabled Image Transmission. *IEEE International Conference on Communications*, **74**, 1730-1735.
- [27] 穆天杰, 陈晓辉, 汪逸云, 等. 基于深度学习的信源信道联合编码方法综述[J]. 电信科学, 2020, 36(10): 56-66.
- [28] Secker, A. and Taubman, D. (2003) Lifting-Based Invertible Motion Adaptive Transform (LIMAT) Framework for Highly Scalable Video Compression. *IEEE Transactions on Image Processing*, **12**, 1530-1542. <https://doi.org/10.1109/tip.2003.819433>
- [29] Song, Y., Zhu, X., Fang, Y. and Zhang, H. (2010) Threshold Optimization for Rate Adaptation Algorithms in IEEE 802.11 WLANs. *IEEE Transactions on Wireless Communications*, **9**, 318-327. <https://doi.org/10.1109/twc.2010.01.090459>
- [30] Ballé, J., Laparra, V. and Simoncelli, E.P. (2025) End-to-End Optimized Image Compression. <https://arxiv.org/abs/1611.01704>
- [31] Ballé, J., Minnen, D., Singh, S., *et al.* (2018) Variational Image Compression with a Scale Hyperprior. arXiv:1802.01436. <https://arxiv.org/abs/1802.01436>
- [32] Minnen, D., Balle, J. and Toderici, G. (2018) Joint Autoregressive and Hierarchical Priors for Learned Image Compression. <https://arxiv.org/abs/1809.02736>
- [33] Toderici, G., O'malley, S.M., Hwang, S.J., *et al.* (2016) Variable Rate Image Compression with Recurrent Neural Networks. <https://arxiv.org/abs/1809.02736>
- [34] Toderici, G., Vincent, D., Johnston, N., *et al.* (2017) Full Resolution Image Compression with Recurrent Neural Networks. <https://arxiv.org/abs/1511.06085>
- [35] Agustsson, E., Tschannen, M., Mentzer, F., Timofte, R. and Van Gool, L. (2019) Generative Adversarial Networks for Extreme Learned Image Compression. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 221-231. <https://doi.org/10.1109/iccv.2019.00031>
- [36] Lu, G., Ouyang, W., Xu, D., *et al.* (2018) DVC: An End-to-End Deep Video Compression Framework. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **45**, 4569-4585.
- [37] Felix, A., Cammerer, S., Dorner, S., Hoydis, J. and Ten Brink, S. (2018) OFDM-Autoencoder for End-to-End Learning of Communications Systems. 2018 *IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Kalamata, 25-28 June 2018, 1-5. <https://doi.org/10.1109/spawc.2018.8445920>
- [38] Tung, T. and Gunduz, D. (2022) Deepwive: Deep-Learning-Aided Wireless Video Transmission. *IEEE Journal on Selected Areas in Communications*, **40**, 2570-2583. <https://doi.org/10.1109/jsac.2022.3191354>
- [39] Gündüz, D., Wigger, M.A., Tung, T., Zhang, P. and Xiao, Y. (2025) Joint Source-Channel Coding: Fundamentals and Recent Progress in Practical Designs. *Proceedings of the IEEE*, **113**, 888-919. <https://doi.org/10.1109/jproc.2024.3477331>
- [40] Kurka, D.B. and Gunduz, D. (2020) DeepJSCC-f: Deep Joint Source-Channel Coding of Images with Feedback. *IEEE Journal on Selected Areas in Information Theory*, **1**, 178-193. <https://doi.org/10.1109/jsait.2020.2987203>

-
- [41] Yang, M., Bian, C. and Kim, H. (2021) Deep Joint Source Channel Coding for Wireless Image Transmission with OFDM. 2021 *IEEE International Conference on Communications*, Montreal, 14-23 June 2021, 1-6. <https://doi.org/10.1109/icc42927.2021.9500996>
- [42] Wu, H., Shao, Y., Bian, C., Mikolajczyk, K. and Gündüz, D. (2024) Deep Joint Source-Channel Coding for Adaptive Image Transmission over MIMO Channels. *IEEE Transactions on Wireless Communications*, **23**, 15002-15017. <https://doi.org/10.1109/twc.2024.3422794>
- [43] Jiang, W., Chen, W. and Ai, B. (2024) Deep Joint Source Channel Coding with Attention Modules over MIMO Channels. 2024 *IEEE 99th Vehicular Technology Conference (VTC2024-Spring)*, Singapore, 24-27 June 2024, 1-5. <https://doi.org/10.1109/vtc2024-spring62846.2024.10683289>
- [44] Bian, C., Shao, Y. and Gündüz, D. (2023). Deepjscc-1++: Robust and Bandwidth-Adaptive Wireless Image Transmission. 2023 *IEEE Global Communications Conference*, Kuala Lumpur, 4-8 December 2023, 3148-3154. <https://doi.org/10.1109/globecom54140.2023.10436878>
- [45] Yang, K., Wang, S., Dai, J., Tan, K., Niu, K. and Zhang, P. (2023) WITT: A Wireless Image Transmission Transformer for Semantic Communications. 2023 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, 4-10 June 2023, 1-5. <https://doi.org/10.1109/icassp49357.2023.10094735>
- [46] Wu, T., Chen, Z., Tao, M., Xu, X., Zhang, W. and Zhang, P. (2024) Mambajsc: Deep Joint Source-Channel Coding with Visual State Space Model. 2024 *IEEE Global Communications Conference*, Cape Town, 8-12 December 2024, 1677-1682. <https://doi.org/10.1109/globecom52923.2024.10901192>
- [47] Wu, H., Shao, Y., Ozfatura, E., Mikolajczyk, K. and Gündüz, D. (2024) Transformer-Aided Wireless Image Transmission with Channel Feedback. *IEEE Transactions on Wireless Communications*, **23**, 11904-11919. <https://doi.org/10.1109/twc.2024.3386052>