

基于CSI反馈的语义图像传输



Semantic Image Transmission Based on CSI Feedback

姜培文/JIANG Peiwen, 韩瑜/HAN Yu,
金石/JIN Shi, 李潇/LI Xiao

(东南大学移动通信国家重点实验室, 中国 南京 210096)
(National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China)

DOI: 10.12142/ZTETJ.202302006

网络出版地址: <https://kns.cnki.net/kcms/detail/34.1228.TN.20230411.1503.020.html>

网络出版日期: 2023-04-11

收稿日期: 2023-01-25

摘要: 语义通信通过构建传输背景知识库、感知传输内容, 大大节省传输带宽, 提升传输质量。该方面的研究逐步与无线通信物理层融合, 从而构建完整的语义传输系统, 但现存的端到端语义编码设计无法感知不断变化的无线信道。提出了基于信道状态信息反馈的图像语义分割编码(SS-CSI)方案, 根据不同子信道上的信噪比(SNR)对传输的关键特征信息进行分割和编码, 并根据实际分类任务需求保护相关语义特征。仿真测试表明, SS-CSI根据实际信道环境, 在低SNR下大幅提升了物体传输质量和分类性能, 在高信噪比上则进一步传输更多背景信息, 提升了整体的图像均方误差性能。

关键词: 语义分割; 无线通信; 图像传输

Abstract: Semantic communications greatly saves transmission bandwidth and improves transmission quality by constructing transmission background knowledge base and sensing transmission content. At present, this research is gradually integrated with the wireless communication physical layer to build a complete semantic transmission system. However, the existing end-to-end semantic coding design is unable to perceive the changing wireless channel. An image semantic segmentation coding based on CSI feedback (SS-CSI) is proposed. According to the signal-to-noise ratio (SNR) on different sub-channels, the key feature information transmitted is segmented and coded. And the relevant semantic features are protected according to the actual classification requirements. Simulation results show that SS-CSI greatly improves the object transmission quality and classification performance at low SNR, and further transmits more background information at high SNR, which improves the overall image mean square error performance.

Keywords: semantic segmentation; wireless communication; image transmission

智能通信系统在提升用户体验的同时, 也服务于大量的终端。由于新冠疫情大流行, 移动工作和在线会议成为人类生活中必不可少的一部分。据相关机构统计, 移动网络传输流量与大流行爆发前相比增加了60%以上。为了应对难以承受的负载, 一些服务提供商(如YouTube等), 只能选择在高峰时间降低视频质量。此外, 自动驾驶汽车依赖于数千个传感器, 同时又与其他车辆通信, 从而产生大量的数据。这些传输服务于特定的任务。因此, 在智能化进程中通信系统往往需要通过文本、语音、图像、视频等多模态数据的理解和传输, 来完成远程操控、辅助驾驶等特定任务^[1-3]。基于经典的香农编码范式^[4], 信源编码压缩传输内容并对内容进行解耦, 以生成独立分布的码字。信道编码通过对码字增加冗余来对抗无线信道带来的传输差错。因此, 传统的编解码模块通常遵循分而治之的设计原则。但是这种设计方式在全新的应用场景下仍然存在问题: 一方面, 诸如在线会议、自动驾驶等低延迟场景的码长受限, 最优编码方式难以获得; 另一方面, 特定任务下的传输特征具有较强的内

容相关性, 新的深度学习方法(如人脸重构^[5])无须使用额外码字即可完成错误修复。语义通信作为智能通信的关键技术之一, 通过构建传输背景知识库^[6]、感知传输内容, 极大地节省传输带宽, 提升传输质量, 成为极具潜力的研究方向。

传统的物理模块都是独立优化的, 不适用于基于语义的通信系统。例如, 调制和信号检测被设计为最小化误码率(BER)。信道估计是最小化估计信道与真实信道之间的均方误差。事实上, 对于语义通信, 不同传输比特具有不同的重要性, 并且这些传输比特也呈现出一定的语义相关性。文献[7]利用句子相似度形成一种新的调制方法, 这种方法的符号在星座图上的位置并不是等间隔分布的。在实际应用中, 峰均功率比也是一大挑战, 这是因为高峰均功率比对硬件设备要求很高。在文献[8]中, 峰均功率比被认为是一个额外的损失函数, 语义网络被训练后能同时实现最小化语义损失和峰均功率比。这意味着峰均功率比的降低应该与语义指标相匹配。在不同的信道环境下, 混合重传请求(HARQ)是保证传输成功的关键。传统的HARQ主要基于用于前向纠错

的信道编码和用于错误检测的循环冗余校验码。文献[9-10]引入基于语义的联合信源信道编码设计 HARQ, 大幅提高了传输性能。信道状态信息 (CSI) 反馈中的数据隐藏方法^[11]也是一种潜在的深度学习方法, 用于消除 CSI 反馈中的传输负载。

上述语义方案引入基于知识库的编解码网络、任务导向的语义损失函数等, 大幅提升了系统的性能。但这些方案需要借助端到端的训练, 才能实现发射端和接收端对知识库的共享, 以及发射端对任务需求的感知。因此, 语义通信方法通常训练在可导的高斯白噪声信道或者比特翻转信道上等, 无法充分利用实际的信道特征, 难以实现对语义码字灵活的保护。本文讨论了语义通信方法在无线通信系统中的应用, 提出了基于语义分割方法的图像传输系统, 并利用反馈的无线信道特征来指导编码网络对传输内容的保护。通过比较不同方法在分类准确率和均方误差上的性能, 所提系统可以灵活适应变化的信道, 在低信噪比下优先保护影响分类准确率的物体, 在高信噪比下进一步保护背景信息, 从而实现在不同的信噪比上都能取得比其他方法更好的分类准确率和均方误差性能。

1 系统模型

为了传输信源信息, 假设存在一张图片 p , 传统发射机首先对其进行信源编码 $S(\cdot)$ 和信道编码 $C(\cdot)$, 从而减少传输冗余, 对抗变化的无线信道。经过编码后的码字表示为:

$$s = C(S(p)). \quad (1)$$

对于传统的正交频分复用 (OFDM) 系统来说, 编码码字 s 会被调制到 K 个子载波上, 即 $s = [s_1, s_2, \dots, s_K]$ 。在通过无线信道后, 接收到的信号可以表示为:

$$y = h \cdot s + z, \quad (2)$$

其中, h 是信道的频率响应, \cdot 是哈达玛积, z 是均值为 0、方差为 σ^2 的加性高斯白噪声。对于频率选择性信道来说, h 可以被看作 K 个平坦的子信道, 每个子信道具有不同幅度的信道频率响应, 即 $h = [h_1, h_2, \dots, h_K]$ 。考虑到不同子信道的频率响应, 接收的码字可以表示为:

$$\hat{s} = \begin{bmatrix} \frac{y_1}{\hat{h}_1}, \dots, \frac{y_K}{\hat{h}_K} \end{bmatrix}. \quad (3)$$

因此, 整体的信噪比 (SNR) 可以表示为:

$$\text{SNR} = \frac{\sum_{k=1}^K \|h_k \cdot s_k\|^2}{K\sigma^2}. \quad (4)$$

在接收端恢复的图片为:

$$\hat{p} = S^{-1}(C^{-1}(\hat{s})), \quad (5)$$

其中, $S^{-1}(\cdot)$ 和 $C^{-1}(\cdot)$ 分别代表信源译码和信道解码。

传统 OFDM 系统的传输目标是使传输 BER 达到最小, 但在复杂的信道环境和海量的移动设备等情况下, 无线通信带宽是有限的。语义通信方法能够感知传输内容, 分割提取语义特征, 基于任务需求进行端到端训练, 从而大幅度降低传输带宽, 提升传输质量。现有的语义方法大多是在可导的高斯白噪声信道上训练获得的, 而实际 OFDM 系统下的信道是乘性的、不可导的, 因此端到端训练的语义通信方法无法充分利用信道特征, 传输性能受限。对此, 我们利用 CSI 反馈, 灵活调整编码方法。

2 网络结构和传输方案

本节中, 我们提出基于语义分割的图像传输方法。首先, 根据语义分割内容实现对信源信息的编码和解码, 以及对传输噪声的对抗; 接着, 考虑无线信道和 CSI 反馈对网络设计的影响, 并根据不同子载波上的 SNR, 实现对语义分割后的特征保护。

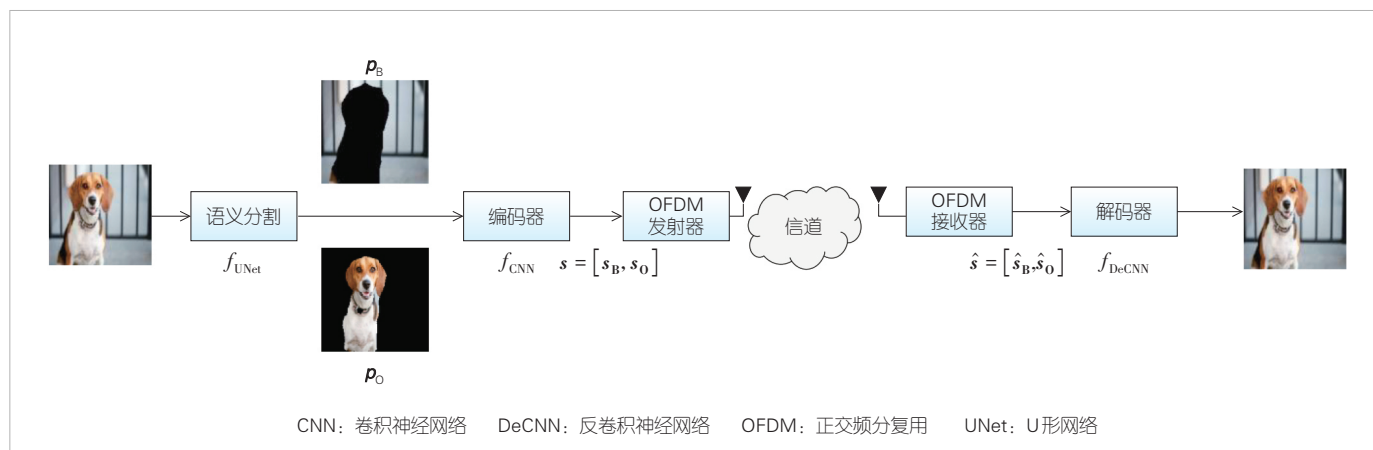
区别于传统的信源编码, 语义信源编码语义分割将源数据划分为不同的语义部分, 每个语义部分通常具有不同的严重程度。在特定的场景中, 分割依赖于一个训练好的网络^[12], 其蕴含着在这个训练场景下的语义知识库。比如: 一幅图像可以分为背景和物体, 其中背景通常较少受到用户关注且对物体分类识别性能影响不大。语义分割后, 信道编码可以保护信源的不同部分, 并在不同信道条件下保护最重要的语义部分。

针对图片传输的问题, 本文采用基于 UNet^[13] 的语义分割方法。如图 1 所示, 图片被分割为背景和物体两类, 分别表示为 p_B 和 p_O 。对于分割后的两部分待传图片信息, 分别采用卷积编码网络, 将其编码为码字 s , 即:

$$[p_B, p_O] = f_{\text{UNet}}(p), \quad (6)$$

$$s = [s_B, s_O] = f_{\text{CNN}}([p_B, p_O]), \quad (7)$$

其中, 语义分割网络表示为 $f_{\text{UNet}}(\cdot)$, p_B 和 p_O 分别为分割后的背景信息和物体图片信息, s_B 和 s_O 分别为背景信息和图片信息被卷积网络 $f_{\text{CNN}}(\cdot)$ 编码后得到的码字。卷积网络输入维度为 (160, 160, 3), 输出维度为 (1, 1, 16), 输出激活函数为 Tanh, 即 s_B 、 s_O 分别编码出 16 个 -1~1 的实数。该网络包含两个中间层, 中间层均由卷积和下采样组成, 其中第 1 个中



▲图1 基于语义分割的图片传输系统

间层输出维度为 (40,40,64)，第2个中间层输出维度为 (10,10,128)，激活函数为 Relu。在该码字通过上述 OFDM 传输系统后，码字的估计 $\hat{s} = [\hat{s}_B, \hat{s}_O]$ ，该码字的估计可以通过反卷积网络 $f_{DeCNN}(\cdot)$ 恢复成原图片，即：

$$\hat{p} = f_{DeCNN}([\hat{s}_B, \hat{s}_O]), \quad (8)$$

其中，反卷积网络输入为 (1,1,32)，其中间层尺寸与卷积编码网络一一对应。

在进行端到端训练时，由于实际信道影响不可导，假设从 s 到 \hat{s} 的变化符合高斯白噪声。图像训练通常采用均方误差 (MSE) 损失函数来恢复每个像素点的信息，但在训练时为了引导网络去保护对任务更重要的信息，会引入基于分类识别任务的损失函数。此损失函数依赖预训练好的视觉几何组 (VGG) 分类网络 $f_{VGG}(\cdot)$ ，并计算接收图像分类结果和发送图像分类结果的交叉熵 (CE) 损失 $L_{CE}(\cdot)$ 。训练的损失函数可以写为：

$$L = L_{MSE}(\hat{p}, p) + 0.1 L_{CE}(f_{VGG}(\hat{p}), p). \quad (9)$$

在获得 CSI 反馈后，计算出 K 个子信道上的传输 SNR 值，即可获得每个子信道的传输质量。因此，我们把语义分割和信道编码与 CSI 反馈相结合，以更加有效地保护传输特征。对于上述训练好的编码方式，可采用一层全连接网络将其分配在不同的子信道上，并且通过端到端训练得到合理的分配方式。具体方法如下：

$$s = [s_1, s_2, \dots, s_K] = \text{Tanh}(\mathbf{W}[s_B, s_O]), \quad (10)$$

其中， \mathbf{W} 为学习到的乘性可训练参数分配方式。由于每个信道的 SNR 值已知，可以按照从高到低的顺序依次将 $[s_1, s_2, \dots, s_K]$ 放置在不同的子信道上，即：

$$\mathbf{y} = [y_1, \dots, y_K] = [h_{\max}, \dots, h_{\min}] \cdot [s_1, s_2, \dots, s_K] + \mathbf{z}. \quad (11)$$

因此， s_1 总是传输在 SNR 最高的子载波上， s_K 则传输在 SNR 最差的子载波上。基于 2.1 节中已经训练好的模型参数，保持其他参数不变，只训练 \mathbf{W} ，则训练损失与 2.1 一致。以上网络训练均采用 Adam 优化器，学习率设置为 0.002，训练 1 000 轮。

由于采用不同子信道上的传输 SNR 来训练参数 \mathbf{W} ，并且考虑实际信道模型和信道估计方法对不同子载波上 SNR 的影响，待传输的语义特征可以更加有效地适应信道环境的变化。网络更倾向于将分割后的语义特征 $[s_B, s_O]$ 中的重要信息放置在较高 SNR 的子信道上，从而在低 SNR 下优先保护分类准确率。

3 仿真与测试

3.1 系统参数

传输一张图片会占用两个 OFDM 符号，每个 OFDM 符号占用 32 个子载波。其中，第 1 个 OFDM 符号为已知导频，用于信道估计并反馈各子载波上的信道响应；第 2 个 OFDM 符号携带图片信息。32 个子载波上的复数由 64 个实数组成，即发送信号 s 为 64 个实数。在语义分割下， s_B 、 s_O 各占用一半子载波数信道模型采用多径信道，在 [0,4,10] 采样点处有 3 条延时扩展路径，功率分布 [0,-5,-10] dB。测试数据采用视觉对象类别 (VOC) 中的动物分类数据集。我们使用物种识别准确率来衡量物体传输的质量，并采用 MSE 衡量整体图片像素信息传输的质量。

4 种比较方法为：只传输物体部分的图像语义分割编码方案 (SS-O)、基于信道状态信息反馈的图像语义分割编码 (SS-CSI)、传输全部内容的图像语义分割编码 (SC)，以及

传统的JPEG信源编码加上低密度奇偶校验码(LDPC)传输,其中低密度奇偶校验码码率为64/255。

3.2 测试结果

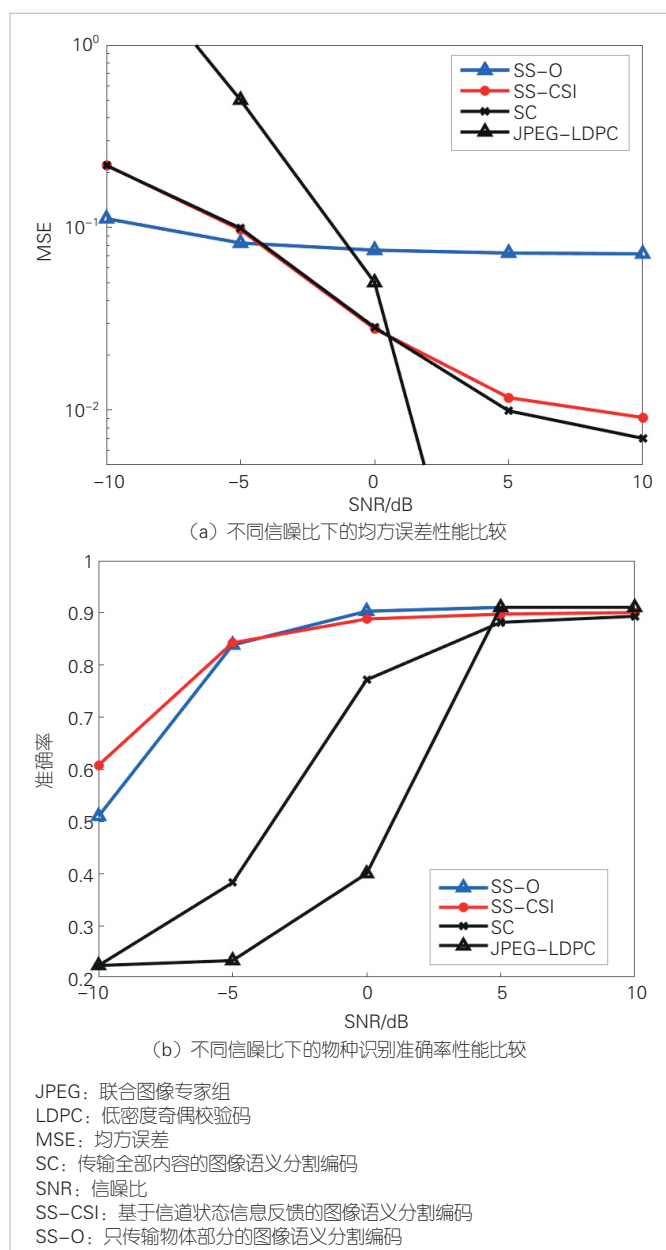
如图2所示,我们比较了不同方法下的图片传输分类准确率和MSE性能。分类准确率反映了图片中物体特征的还原度,MSE性能反映了图片的像素信息在传输中的损失程度。

在分类准确率性能方面,SS-O的所有32子载波均用于传输语义分割后的物体部分,因此它在较低SNR(例如-5~

5 dB)下也拥有很好的分类准确率。但是过低的SNR会使得SS-O的传输差错增大,例如:在-10~-5 dB下,SS-O的分类准确率反而不如SS-CSI。SS-CSI根据反馈的每个子载波上的SNR来分配不同的语义特征。实际上,由于在训练时考虑了分类准确率,SS-CSI通常把代表物体的语义特征 s_o 分配到SNR更高的子载波上,因此在SNR为-10~-5 dB时,部分较差信道带来的差错对SS-CSI影响更小,分类准确率更高。当SNR为0~10 dB时,由于大部分子信道条件都较好,而SS-O用于传输物体的子载波数大于SS-CSI,此时SS-O的分类准确率性能要略好于SS-CSI。SC中物体和背景占用的子载波数目相同且不用CSI反馈来辅助分配这些待传输的码字,因此在SNR较差的情况下,SC的分类准确率不如SS-O和SS-CSI。在SNR较高时,JPEG-LDPC的分类准确率达到分类网络的上限。但随着SNR的降低,传输差错超出LDPC编码的纠错能力,JPEG-LDPC分类准确率性能快速下降。

在MSE性能方面,SS-O完全舍弃了背景部分的传输方法,以便于在SNR很低时可以更好地保护物体部分的MSE性能。然而,在信道条件较好的情况下,SS-O也无法恢复背景部分,使得MSE几乎不会随着SNR的上升而降低。在SNR较低的情况下,SS-CSI和SC的MSE性能相当。CSI反馈会使得SS-CSI总倾向于将更好的子信道用于物体部分的传输,但在SNR较高的情况下,增加占比较大的背景部分会进一步降低MSE,导致SS-CSI在SNR较高的情况下略差于SC。传统的JPEG-LDPC方法在SNR较高时,使传输错误均被纠出,此时MSE几乎为0。同样地,当SNR降低导致传输错误超出其纠错能力后,MSE性能也会迅速降低。

在SNR=0 dB时,不同传输方法对图像的恢复结果如图3所示。其中,SS-O可将所有传输带宽用于恢复无图部分的像素信息。虽然物体部分恢复得最清晰,但背景却被完全舍弃。与SS-O不同,本文提出的方法SS-CSI在尽可能恢复物



▲图2 基于语义分割的图片传输系统性能比较



▲图3 信噪比为0 dB时3种方法的实例图

体信息的同时,利用SNR较低的信道,传输一部分背景信息。传统的语义传输方法SC不考虑信道条件,物体和背景均有不同程度的损失。从图3可以看出,SS-CSI能够根据CSI反馈的信息对不同传输部分进行灵活保护,即在MSE性能和分类准确率两种损失函数的指导下,网络学习到优先降低物体部分的MSE,在提高分类性能的同时对背景进行保护。通过这样的方法,网络既可以在高SNR下取得很好的MSE性能,又可以在低SNR下取得很好的分类性能。

4 结束语

本文中,我们介绍了语义通信系统在提升传输质量、节省传输带宽方面的优势;考虑到典型的OFDM无线通信系统以及无线信道的特性,阐述了语义信源信道编码在无线通信中的融合应用;设计了基于语义分割的无线图片传输方案,即通过对不同重要性的传输部分进行分割(如背景和物体信息),实现对重要语义特征的保护。考虑到CSI反馈在无线通信中的重要性,SS-CSI方法使用反馈的CSI信息指导网络对不同传输部分进行保护,并使用具体的任务指导网络训练。此外,我们比较了SS-CSI和CSI反馈对分类准确率和MSE的影响,并通过具体实例证明了分割和CSI反馈可以对接收图片的重要部分进行更好地保护。

参考文献

- [1] GUNDUZ D, QIN Z, AGUERRI I E, et al. Beyond transmitting bits: context, semantics, and task-oriented communications [EB/OL]. [2023-01-25]. <https://arxiv.org/abs/2207.09353>
- [2] LU K, ZHOU Q Y, LI R P, et al. Rethinking modern communication from semantic coding to semantic communication [J]. IEEE wireless communications, 2022, 30(1): 1-13. DOI: 10.1109/MWC.013.2100642
- [3] SHI G M, XIAO Y, LI Y Y, et al. From semantic communication to semantic-aware networking: model, architecture, and open problems [J]. IEEE communications magazine, 2021, 59(8): 44-50. DOI: 10.1109/MCOM.001.2001239
- [4] SHANNON C E, WEAVER W, WIENER N. The mathematical theory of communication [J]. Physics today, 1950, 3(9): 31-32. DOI: 10.1063/1.3067010
- [5] LI Y J, LIU S F, YANG J M, et al. Generative face completion [C]// Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017: 5892-5900. DOI: 10.1109/CVPR.2017.624
- [6] BAO J, BASU P, DEAN M K, et al. Towards a theory of semantic communication [C]// Proceedings of 2011 IEEE Network Science Workshop. IEEE, 2011: 110-117. DOI: 10.1109/NSW.2011.6004632
- [7] GUO S S, WANG Y H, ZHANG P. Signal shaping for semantic communication systems with A few message candidates [EB/OL]. [2023-01-25]. <https://arxiv.org/abs/2202.02072>
- [8] SHAO Y L, GUNDUZ D. Semantic communications with discrete-time analog transmission: a PAPR perspective [EB/OL]. [2023-01-25]. <https://arxiv.org/abs/2208.08342>
- [9] JIANG P W, WEN C K, JIN S, et al. Deep source-channel coding for sentence semantic transmission with HARQ [J]. IEEE transactions on communications, 2022, 70(8): 5225-5240. DOI: 10.1109/TCOMM.2022.3180997
- [10] ZHOU Q Y, LI R P, ZHAO Z F, et al. Adaptive bit rate control in semantic communication with incremental knowledge-based HARQ [EB/OL]. [2023-01-25]. <https://arxiv.org/abs/2203.06634>
- [11] GUO J J, WEN C K, JIN S. Eliminating CSI feedback overhead via deep learning-based data hiding [J]. IEEE journal on selected areas in communications, 2022, 40(8): 2267-2281. DOI: 10.1109/JSAC.2022.3180806
- [12] GARCIA-GARCIA A, ORTOS-ESCOLANO S, OPREA S, et al. A review on deep learning techniques applied to semantic segmentation [EB/OL]. [2023-01-25]. <https://arxiv.org/abs/1704.06857>
- [13] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation [M]// Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015: 234-241. DOI: 10.1007/978-3-319-24574-4_28

作者简介



姜培文, 东南大学移动通信国家重点实验室在读博士研究生;研究方向为智能通信物理层关键技术、语义通信等。



韩瑜, 东南大学移动通信国家重点实验室副教授;主要研究方向为大规模MIMO无线传输理论与关键技术、智能超表面辅助移动通信等。



金石, 东南大学移动通信国家重点实验室教授、博士生导师,教育部“长江学者奖励计划”特聘教授,国家自然科学基金杰出青年科学基金获得者,国家“万人计划”科技创新领军人才;主要研究方向为移动通信理论与关键技术、物联网理论与关键技术,以及人工智能在无线通信中的应用等;发表学术论文400余篇,获授权国际/国家发明专利50余项,出版专著2部、教材1部。



李潇, 东南大学移动通信国家重点实验室教授、博士生导师;主要研究方向为智能超表面辅助的移动通信、人工智能在无线通信中的应用等。