

北京邮电大学学报



Journal of Beijing University of Posts and Telecommunications ISSN 1007-5321,CN 11-3570/TN

《北京邮电大学学报》网络首发论文

题目: 基于语义重要性的语义编码算法研究 作者: 林上豪,刘芳芳,郭彩丽,仝顽杰

DOI: 10.13190/j.jbupt.2023-090

收稿日期: 2023-04-25 网络首发日期: 2024-04-02

引用格式: 林上豪,刘芳芳,郭彩丽,仝顽杰.基于语义重要性的语义编码算法研究

[J/OL]. 北京邮电大学学报. https://doi.org/10.13190/j.jbupt.2023-090





网络首发: 在编辑部工作流程中,稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定,且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件,可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定;学术研究成果具有创新性、科学性和先进性,符合编辑部对刊文的录用要求,不存在学术不端行为及其他侵权行为;稿件内容应基本符合国家有关书刊编辑、出版的技术标准,正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性,录用定稿一经发布,不得修改论文题目、作者、机构名称和学术内容,只可基于编辑规范进行少量文字的修改。

出版确认:纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约,在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版,以单篇或整期出版形式,在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z),所以签约期刊的网络版上网络首发论文视为正式出版。

网络首发时间:2024-04-02 13:20:45

网络首发地址: https://link.cnki.net/urlid/11.3570.TN.20240329.1424.001

北京邮电大学学报 Journal of Beijing University of Posts and Telecommunications

DOI: 10.13190/j.jbupt.2023-090

基于语义重要性的语义编码算法研究

林上豪1, 刘芳芳1,2, 郭彩丽1,2, 全硕杰1

(1. 北京邮电大学 先进信息网络北京实验室, 北京 100876; 2. 北京邮电大学 网络体系构建与融合北京市重点实验室, 北京 100876)

摘要: 为了满足智能通信实体之间信息高效传输的需求,语义通信研究面临语义信息的表示与度量、编码以及传输等基本问题。针对通信资源紧张、任务时延敏感的万物智联传输场景,提出了基于语义重要性的语义编码算法,并设计、搭建了语义通信实验平台,实现了语义信息传输。首先,结合无线通信的基本流程建立了面向智能任务的语义通信系统模型; 其次,设计了语义重要性提取方法,通过聚合图像通道特征并自适应校准其特征响应以获得特征图对于任务结果的重要性权重; 然后,设计了语义压缩方法,在实际传输时根据重要性权重进行特征图筛选以实现语义数据压缩; 最后,在搭建的语义通信平台上以工业互联网场景中表面缺陷分类任务为例对所提算法的性能进行实测验证。实验结果表明,所提算法具有传输数据量小、抗噪声能力强和任务处理时延短等优势。

关键词:语义通信;语义编码;图像分类;实验平台

中图分类号: TN929.5

文献标志码: A

Research on Semantic Coding Algorithm Based on Semantic Importance

LIN Shanghao¹, LIU Fangfang^{1,2}, GUO Caili^{1,2}, TONG Wanjie¹

(1. Beijing Laboratory of Advanced Information Networks, Beijing University of Posts and Telecommunications, Beijing 100876, China;

2. Beijing Key Laboratory of Network System Construction and Integration, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: In order to meet the demand of efficient information transmission between intelligent communication entities, semantic communication research is faced with the basic problems of semantic information representation, encoding and transmission. In view of the communication resource shortage and task delay sensitive transmission scenario of everything intelligent link, a semantic coding algorithm based on semantic importance was proposed, and a semantic communication experiment platform was designed and built to realize semantic information transmission. Firstly, the intelligent task-oriented semantic communication system model was established based on the basic process of wireless communication. Secondly, a semantic importance extraction method was designed to obtain the importance weights of the feature maps to the task results by aggregating the channel features and adaptively calibrating the feature responses. Then, a semantic compression method was designed to filter the feature maps according to the importance weight in the actual transmission to achieve semantic data compression. Finally, the performance of the proposed algorithm was verified on the semantic communication platform by taking the task of surface defect classification in the industrial Internet scene as an example. Experimental result shows that the proposed algorithm has the advantages of small data transmission, strong anti-noise ability and short task processing delay.

Key words: semantic communication; semantic coding; image classification; experimental platform

为了对大规模物联网(IoT, internet of things) 等新兴应用场景的普及提供保障,第6代移动通

收稿日期: 2023-04-25

基金项目: 中央高校基本科研业务费专项项目(2021XD-A01-1)

作者简介: 林上豪(1998—), 男,硕士生。

通信作者: 刘芳芳(1984—), 女,硕士生导师,邮箱: fliu@bupt.edu.cn。

信系统 (6G, the 6th generation of mobile communicatians system)在全球范围的研究已经启动^[1]。在未来 6G 移动通信系统中,用户的智能需求将被进一步挖掘和实现,并以此为基准进行技术规划与演进布局^[2]。传统无线通信技术主要以准确传输数据或信号波形为目标,其系统性能指标逐渐逼近香农极限。在未来的万物互联时代,如何进一步提高通信效率,满足各种通信实体之间复杂多样的智能化信息传输需求,是无线通信领域面临的一大挑战^[3]。

为了应对上述挑战,以"达意"为目标的语义 通信成为近年来研究的热点。语义通信技术的根 本目的在于保证通信双方之间语义的准确传递, 并不力求原始数据或信号保真[4],因此大幅减小 在信息转化过程中产生的时延与失真。目前语义 通信主要面向智能任务相关的通信场景,其将任 务分别在 IoT 设备与云计算终端上进行协同处理, 一方面显著减少 IoT 设备传输的数据量,节约了 通信资源,另一方面将大量模型参数存储于云计 算终端中,有效减轻 IoT 设备的存储压力。语义 通信研究面临的基本问题包括语义信息的表示与 度量,语义信息编码以及语义信息传输。

第一,为了表示与度量语义信息,鲁晨光[5]以 香农信息论为基础引入广义信息论,并结合模糊 集合等理论统一度量语义信息。近年来,考虑到 信息语义理解的智能化需求, Xie 等[6]证明了语 义信息可以用神经网络提取的特征来表示。第二, 在语义信息编码方面, Basu 等[7]建立了无损的语 义数据压缩理论,表明语义层的数据压缩能让传 输的数据量显著减少。随着人工智能与无线通信 技术紧密融合, 语义编码利用信源数据的语义冗 余,提取与任务高度相关的语义特征[8]。针对不 需要恢复信源的场景, 语义编码旨在让机器理解 语义并完成智能任务,因此具有更大的压缩空间 [9]。在通信资源日益紧缺的大数据时代, 充分压 缩语义信息是语义通信研究的关键目标。第三, 在语义信息传输方面,基于深度学习技术提出了 面向文本、图像、语音等不同信源类型的语义通 信系统[6,10-11],通过传输提取的语义特征以完成 相应的智能任务,为语义通信的应用研究提供了 可供参考的网络和系统模型。但这些工作仅涉及 理想情况下的理论研究以及仿真性能,语义通信

系统突出的性能需要实测平台的验证和支撑。

基于以上提出的问题,针对未来通信资源紧张、任务时延敏感的无线传输场景,提出了基于语义重要性的语义编码算法,并且设计与搭建了基于通用软件无线电外设(USRP, universal software radio peripheral)的语义通信实验平台。首先通过聚合图像特征信息并自适应地校准特征响应的重要性权重,从而获取特征图之间的相关性,然后根据权重筛选重要的特征图进行传输以实现语义压缩;进一步地,以工业互联网场景中热轧带钢表面缺陷分类任务为例,在基于 USRP的 5G 链路级通信实验平台[12]上实现了语义信息传输,证明了该算法能够在信道条件稳定的情况下保证智能任务的性能,并且占用更少的带宽和时延,因此显著提高无线通信系统的整体容量。

1 系统模型

面向智能任务的语义通信系统模型如图 1 所示。首先对信源数据进行特征提取;然后基于特征的重要性实现数据压缩;经信道传输后,通过语义解压还原语义信息形式,并输入神经网络计算任务结果。

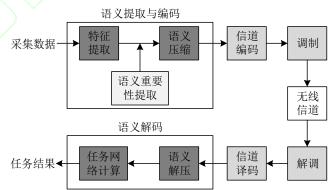


图 1 面向智能任务的语义通信系统模型

在面向图像任务时, IoT 设备采集数据后输入特征提取网络进行初步特征提取。将输入的数据记为 *X* , 发送端的特征图为

$$\boldsymbol{U}_{t} = S_{\boldsymbol{\alpha}}(\boldsymbol{X}) \tag{1}$$

其中: $S_{\alpha}(\cdot)$ 为特征提取网络, α 为模型参数。

为了保留重要语义信息,资源受限的 IoT 设备根据特征图对于任务结果的重要性实现语义压缩后再进行传输。语义压缩的过程可表示为

$$\boldsymbol{I}_{t} = C_{\sigma}(\boldsymbol{U}_{t}) \tag{2}$$

其中: $C_{\sigma}(\cdot)$ 为压缩函数, σ 为压缩比例。此过程一方面使实际传输的数据量进一步减少,大大降低端到端任务时延;另一方面显著降低了任务网络的计算复杂度。

语义压缩后进行独立的信道编码与调制,输出的数据记为Y_i,接着经无线信道传输后接收端接收到的数据可表示为

$$Y_{r} = hY_{t} + n \tag{3}$$

其中:h为信道系数,n为加性高斯白噪声。

完成解调与信道译码后,考虑到任务网络输入维度的限制,需要将语义信息进行维度还原。 具体方式是先匹配任务网络的输入顺序,再对缺失的维度进行填充,语义解压的过程可表示为

$$\boldsymbol{U}_{r} = D_{\sigma}(\boldsymbol{I}_{r}) \tag{4}$$

其中 $D_{\sigma}(\cdot)$ 为解压缩函数。

任务网络取决于智能任务的类别,如在面向 图像分类任务时采用分类器网络。任务网络计算 的结果为

$$Z = N_a(U_r) \tag{5}$$

其中 $N_{\alpha}(\cdot)$ 为任务网络。

2 基于语义重要性的语义编码算法

2.1 总体网络结构

语义通信的总体网络结构如图 2 所示,旨在保证智能任务性能的前提下实现语义信息的高效传输。首先,对输入图像进行尺寸调整等预处理操作,卷积神经网络主要负责对图像的局部特征进行提取,随着层数的叠加,深层的神经元具有更大的感受野,可以提取更加精细的特征。语义重要性提取网络由一层全局池化层和两层全连接层构成,为进一步降低实际传输的数据量,可以基于特征图的语义重要性进行压缩编码。

接收到特征图后先进行语义解码,对特征数据进行维度还原,然后输入任务网络进行计算。分类任务网络主要包括全连接层和 softmax 层,全连接层将特征图非线性映射至一维特征向量,其中包括全部特征信息;softmax 层可以限制输出值的和为 1,因此转化成分类为各个类别的概率,最终将概率最高的类别作为分类结果。

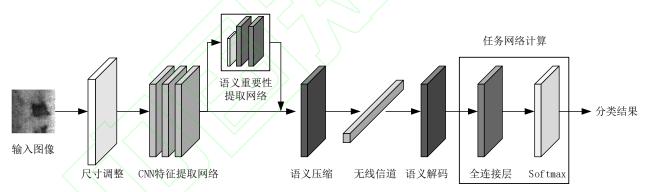


图 2 语义通信的总体网络结构

2.2 语义重要性提取

IoT 设备通过卷积神经网络对图像进行特征提取以后,可以通过筛选需要传输的特征图进一步减少实际传输的数据量,筛选的依据是特征图对于任务结果的语义重要性。下面介绍基于特征图权重的语义重要性提取方法。

在 输 入 图 像 $X \in \mathbb{R}^{H \cdot W \cdot C}$ 映 射 到 特 征 图 $U_{\cdot} \in \mathbb{R}^{H \times W \times C}$ 的卷积运算中, \mathbb{R} 表示矩阵向量空间,

 $V = [v_1, v_2, ..., v_C]$ 表示卷积神经网络学习的滤波器卷积核参数,因此可以得到输出为

$$u_c = v_c * X = \sum_{s=1}^{C'} v_c^s * x^s$$
 (6)

其中: *为卷积操作, \mathbf{v}_c^s 为对应通道中的卷积核参数。由于输出是所有通道的总和,通道之间的相关性被嵌入到 \mathbf{v}_c 中,同时引入了滤波器获取的局部空间相关性,导致特征图通道间的依赖关系难以确定。

通过聚合图像特征信息并自适应地校准滤波器响应,从而学习通道之间的关联,确定每个通道的重要程度,以便网络能够识别出其后续转换所使用的特征。下面分别介绍嵌入全局信息与自适应校准2个步骤。

为了得出通道之间的相关性,需要考虑每个通道的输出特性,因此将全局空间信息聚合到通道的权重中,其通过对特征通道进行全局平均池化来实现。对卷积神经网络提取的所有特征图进行压缩以后将得到通

道维度的全局特征。在空间维度 $H \times W$ 上对 U_t 进行压缩可以得到统计上的 $z \in R^c$,其中 z 的第 c 项为

$$z_{c} = \frac{1}{HW} \sum_{i=1}^{H} \sum_{j=1}^{W} u_{c}(i, j)$$
 (7)

为了从聚合的数据中获得通道之间的相关性,对 其特征响应进行自适应校准。选择函数时需要满足两 个条件:一是能对非线性的通道关系进行建模,二是必 须能够同时突出多个通道的特性。因此使用 sigmoid 激 活函数对门限机制进行设计,即

$$s = \frac{1}{1 + e^{-T_2 \delta(T_1 z)}} \tag{8}$$

其中: δ 为线性整流函数(ReLU, rectified linear unit),

 $T_{1} \in \mathbf{R}^{\frac{C}{r}}$, $T_{2} \in \mathbf{R}^{\frac{C}{r}}$ 。为了限制模型的复杂度且便于泛化,在非线性层中使用 2 个全连接层来参数化门限机制,即一个降维比为 r 的降维层与一个 ReLU 层,以及一个可以转换到 U,通道维度的升维层。通过带有

门限机制的激活函数 s 重缩放U, 得到最后的权重值为

$$\tilde{\boldsymbol{x}}_c = \boldsymbol{s}_c \boldsymbol{u}_c \tag{9}$$

其中 $\tilde{X} = [\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_C]$ 为权重矩阵,体现了每个特征图对任务结果的重要性。

语义重要性提取使模型更加关注与任务结果相关 的通道特征,最后得到特征图的重要性权重,为下一步 进行语义压缩提供了依据。

2.3 语义压缩

在实际传输时需要根据当前的通信资源等条件 对提取的特征图进行筛选,接着传输部分最重要的 特征图,实现语义压缩。

一般情况下,压缩重要性权重越小的特征图对 任务网络计算结果产生的影响越小。因此在给定压 缩比例的情况下,可以确定允许传输的特征图。

依据语义重要性提取得到的结果实现特征图压 缩,即

$$\boldsymbol{M}_{k} = \begin{cases} \boldsymbol{M}_{k}, & \omega_{k} \geqslant \omega_{n} \\ \boldsymbol{0}, & \omega_{k} < \omega_{n} \end{cases}$$
 (10)

其中: M_k 为最后一层卷积层的第 k 个特征图, ω_k 为该

特征图的重要性权重, ω_n 为压缩阈值。实际传输的特征图数量为

$$n = N(1 - \sigma) \tag{11}$$

其中: N 为提取的特征图总数, $\sigma \in [0,1)$ 为压缩因子,反映了 IoT 设备语义压缩的程度,因此实际传输的特征图数量所占比例为 $1-\sigma$ 。

式(10)的压缩阈值为

$$\omega_n = z(n) \tag{12}$$

其中: $z = \text{sort}([\omega_1, \omega_2, ..., \omega_N])$ 为特征图权重的重要性排序,且排列顺序为从大到小, ω_n 为第 n 项。

因此,语义压缩后的n个特征图索引为

$$\mathbf{s}_{t} = [s_{1}, s_{2}, ..., s_{n}] = \{k | \omega_{k} \ge \omega_{n}\}$$
 (13)

IoT 设备根据式(13)的索引将特征图经无线信道传输,接着云计算终端对接收的特征图数据进行维度还原,让未收到数据的索引位置特征图置零。最终由任务网络计算得到分类结果,并传回该结果至 IoT 设备。

3 实验平台设计与搭建

为验证所提语义编码算法的性能,搭建了基于USRP的语义通信平台,语义通信平台的整体架构如图3所示,分为语义层和技术层。语义层主要是通过神经网络算法实现对图像语义的提取与解析;技术层主要是通过5G端到端通信链路实现物理信号的发射和接收。为了实现语义层语义信息与技术层5G通信链路物理信号之间的映射,语义层与技术层的层接口通过用户数据报协议(UDP, user datagram protocol)端口完成数据类型匹配与软件数据交互。

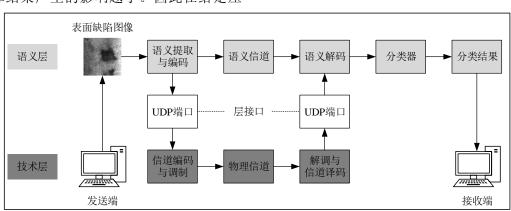


图 3 语义通信平台的整体架构

3.1 语义层

语义层的实现分为智能任务的选取、知识库的同步以及提取与编码语义信息,该层基于 PyTorch 深度学习框架实现。

在智能互联场景中,语义通信的目的是理解接收信息的内容,以高效地完成智能任务。缺陷检测与分类属于工业场景中的常见任务,目前工厂的检测方式多为目视检测,造成检测精度差、人工效率低等问题,因此以表面缺陷分类任务为例验证和分析所提算法的性能,将神经网络提取的与该分类任务相关的特征作为语义信息。

语义编解码基于海量数据训练的知识库。为了保证通信双方对语义信息有相同的理解,发送端和接收端需要实现语义知识库的同步。通过热轧带钢表面缺陷数据集^[13]以及卷积神经网络模型构建语义知识库,其中包含 6 类热轧带钢表面缺陷图像,分别为轧制氧化皮、开裂、内含物、斑块、点蚀表面以及划痕。在本地训练好神经网络的初始参数,并在离线的情况下加载到发送端与接收端。

提取与编码语义信息的关键在于部署特征提取网络并对实际传输的特征进行压缩编码。针对网络的部署,由于该数据集类间差异小、类内差异大,特征提取较为困难,因此特征提取网络采用深度残差网络

(ResNet,residual network)^[14]。考虑到该数据集需要进行 六分类任务,调整网络结构最后一层输出神经元数量 为 6。最后,将 ResNet 全连接层以前的神经网络部署 在发送端作为特征提取网络,将全连接层及以后的神 经网络部署在接收端作为任务网络。为了对提取的特 征进行压缩编码,平台实现了基于语义重要性的语义 编码算法。因此在实际传输时可以根据所需压缩比例 选择重要性权重较大的特征。

3.2 技术层

为了实现物理信号的传输,在技术层搭建了 5G 无线通信链路。语义通信平台技术层的结构设计如图 4 所示,分为硬件结构和软件结构。

考虑到快速开发的需求,选择支持 LabVIEW 通信系统设计套件开发的 USRP-2943R。该设备带有灵活的上下射频转换器以及带宽范围,支持发射和接收频率高达 6 GHz 的无线信号^[15]。TDJ-5158BKR 天线的频率范围为 5.15~5.85 GHz,满足平台工作频段的要求。

软件结构主要是通过 LabVIEW 通信系统设计套件控制硬件设备进行 5G 通信链路物理信号的发射和接收。信道编码采用译码复杂度低、更适用于低时延应用的低密度奇偶校验码,调制方式采用 64 阶正交振幅调制。

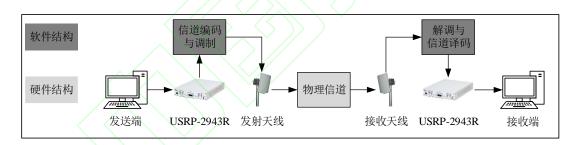


图 4 语义通信平台技术层的结构设计

3.3 语义层与技术层的接口

为了完成语义信息的传输,通过设计层接口 将语义层的语义信息映射为技术层 5G 通信链路 中的物理信号。

一方面,PyTorch 深度学习框架以张量的形式处理神经网络层间的浮点型数据类型,需要对其进行量化才能满足技术层二进制数据的传输需求。在 ResNet 网络中的特征提取网络后增加二值量化层,将特征均值作为判别门限分别量化为 0 与 1,然后进行联合训练,用于实现语义层语义信息到技术层物理信号的转换。另一方面,为了实现 PyTorch 深度学习框架与 LabVIEW 通信系

统设计套件之间的数据传输,使用通信双方的UDP端口作为二进制数据交互的接口。发送端PyTorch深度学习框架将特征量化后产生的比特流先转换成字节流,然后写入本地UDP端口;接着LabVIEW通信系统设计套件读取该端口的数据并执行后续发送流程。接收端将恢复的载荷数据写入UDP端口,程序监听相应端口并接收端口中的UDP数据报,最终将其中的字节流转换为比特流,输入PyTorch深度学习框架后续网络中计算任务结果。

4 实验设计与结果分析

4.1 实验设置

将基于所提算法的语义通信方法命名为语义重要性提取方式,将人工智能 IoT 中面向智能任务的语义通信方法^[16]命名为语义关系提取方式,并将随机传输特征图的语义通信方法作为对比方案。

所用数据集包含 6 类热轧带钢表面缺陷图像,每类 300 张,图片分辨率为 200×200。由于该数据集规模较小,通过旋转、缩放以及移位等方法进行数据增强,将数据集扩充到原来的 10 倍,其中 2000 张用于训练,1000 张用于测试。数据预处理时,首先将图像扩展到 256×256,然后随机裁剪为 224×224,作为特征提取网络的输入。为了加快模型收敛的速度,使用在 ImageNet 数据集上预训练好的参数初始化特征提取网络参数,再进行后续训练。采用交叉熵作为损失函数,模型训练的参数设置如表 1 所示。训练以及测试环境为 Ubuntu 16.04 + CUDA 10.2,深度学习框架为 PyTorch1.7.1。

	/
模型参数	数值
批尺寸	32
迭代次数	50
学习率	0.01
动量	0.9
优化器	随机梯度下降

表 1 模型训练的参数设置

4.2 带宽需求分析

不同编码位数条件下的分类准确率如图 5 所示。实验时将发送端信噪比固定为 30 dB。在语义重要性提取与语义关系提取方式中,通过改变压缩因子来实现不同的编码位数;在随机压缩方式中,通过随机选择对应数量的特征图来实现不同的编码位数。实验结果表明,在带宽条件有限时,语义重要性提取方式和语义关系提取方式可以进一步改善任务性能,是由于这两种方式在实际传输时较多地保留了与任务相关的重要语义信息,并且语义重要性提取方式在带宽资源紧张的场景下具有更好的性能。

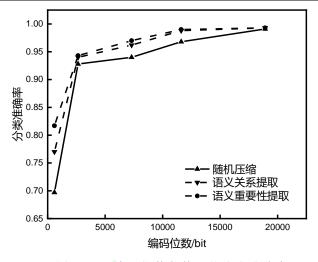


图 5 不同编码位数条件下的分类准确率

4.3 抗噪声性能分析

不同信噪比条件下的分类准确率如图 6 所示。实验时固定压缩因子为 0.8,保证上述通信方式具有相同的编码位数。实验结果表明,在相同信噪比条件下,语义重要性提取方式和语义关系提取方式具有更好的抗噪声性能。在信道条件恶劣时,由于代表重要语义信息的比特误差会导致语义失真,因此语义关系提取方式的性能优于语义重要性提取方式。在信道条件较好的情况下,语义重要性提取方式在保证任务性能的前提下能够实现更大程度的压缩,从而减轻通信传输负担。

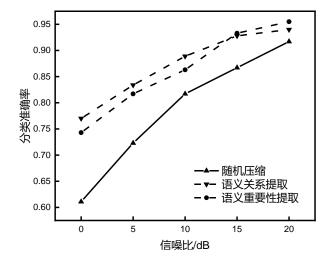


图 6 不同信噪比条件下的分类准确率

4.4 任务时延分析

不同通信方式完成分类任务的时延如表 2 所示。总时延包含传输时延与处理时延。传输时延即发收天线间信号传输所用的时间;处理时延包括网络模型计算、接口数据交互与通信链路基带

处理三部分。将传输全部特征图的方式命名为语义通信方式,语义关系提取与语义重要性提取方式的压缩因子固定为 0.8,对所有图像进行传输并计算平均任务时延。实验结果表明,由于语义关系提取和语义重要性提取 2 方式对提取的特征图进行了压缩,减少了无线传输以及网络模型处理的数据,因此传输时延与处理时延均有所下降。此外,语义重要性提取方式的模型复杂度更低,更适用于时延敏感场景。

表 2 任务时延

传生	諭时 延	处理时延	总时延
语义通信方式	9.6	26.9	36.5
语义关系提取	6.8	19.5	26.3
语义重要性提取	6.8	14.4	21.2

5 结束语

针对海量数据交互时代通信资源紧缺的挑战,提出了基于语义重要性的语义编码算法,并以工业场景中的具体任务为例搭建了语义通信实验验证平台。所提算法在特征提取网络的基础上,通过语义重要性提取与语义压缩方法对与任务结果相关的重要语义信息进行编码,进一步提高了语义传输的有效性。实验结果表明,所提算法能够有效缓解 IoT 设备带宽资源受限的负担,显著降低端到端任务时延,对信道条件稳定的大规模 IoT 场景下语义通信的应用具有重要参考价值。

参考文献:

- [1] 石光明, 肖泳, 李莹玉等. 面向万物智联的语义通信 网络[J]. 物联网学报, 2021, 5(2): 26-36. SHI G M, XIAO Y, LI Y Y, et al. Semantic communication networking for the intelligence of everything[J]. Chinese Journal on Internet of Things, 2021, 5(2): 26-36.
- [2] 牛凯, 戴金晟, 张平等. 面向 6G 的语义通信[J]. 移动通信, 2021, 45(4): 85-90.

 NIU K, DAI J S, ZHANG P, et al. 6G-oriented semantic communications[J]. Mobile Communications, 2021, 45(4): 85-90.
- [3] XIAO Y, SHI G M, LI Y Y, et al. Toward self-learning edge intelligence in 6G[J]. IEEE Communications Magazine, 2020, 58(12): 34-40.
- [4] 石光明, 李莹玉, 谢雪梅. 语义通讯——智能时代的产物[J]. 模式识别与人工智能, 2018, 31(1): 91-99. SHI G M, LI Y Y, XIE X M. Semantic communications: outcome of the intelligence era[J]. Pattern Recognition

- and Artificial Intelligence, 2018, 31(1): 91-99.
- [5] LU C G. The third kind of Bayes' theorem links membership functions to likelihood functions and sampling distributions[C]//International Conference on Cognitive Systems and Signal Processing. Berlin: Springer, 2019: 268-280.
- [6] XIE H Q, QIN Z J, LI G Y, et al. Deep learning enabled semantic communication systems[J]. IEEE Transactions on Signal Processing, 2021, 69: 2663-2675.
- [7] BASU P, BAO J, DEAN M, et al. Preserving quality of information by using semantic relationships[C]//IEEE International Conference on Pervasive Computing and Communications Workshops. Piscataway, NJ: IEEE Press, 2012: 58-63.
- [8] STRINATI E C, BARBAROSSA S. 6G networks: beyond Shannon towards semantic and goal-oriented communications[J]. Computer Networks, 2021, 190: 107930.
- [9] 刘传宏, 郭彩丽, 杨洋等. 面向智能任务的语义通信: 理论、技术和挑战[J]. 通信学报, 2022, 43(6): 41-57. LIU C H, GUO C L, YANG Y, et al. Intelligent taskoriented semantic communications: theory, technology and challenges[J]. Journal on Communications, 2022, 43(6): 41-57.
- [10] LEE C H, LIN J W, CHEN P H, et al. Deep learning-constructed joint transmission-recognition for Internet of Things[J]. IEEE Access, 2019, 7: 76547-76561.
- [11] WENG Z Z, QIN Z J. Semantic communication systems for speech transmission[J]. IEEE Journal on Selected Areas in Communications, 2021, 39(8): 2434-2444.
- [12]黄丁才. 基于 USRP 的极化全双工通信实验平台设计与实现[D]. 北京: 北京邮电大学, 2020: 35-42. HUANG D C. Design and implementation of polarized full-duplex communication experimental platform based on USRP[D]. Beijing: Beijing University of Posts and Telecommunications, 2020: 35-42.
- [13]HE Y, SONG K C, MENG Q G, et al. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(4): 1493-1504
- [14]HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2016: 770-778.
- [15] VOINA A, CODAU C, PASTRAV A, et al. Implementation of a SDR-based redundant access network using NI USRP-RIO[C]//RoEduNet Conference: Networking in Education and Research. Piscataway, NJ: IEEE Press, 2017: 1-6.
- [16]刘传宏,郭彩丽,杨洋等.人工智能物联网中面向智能任务的语义通信方法[J].通信学报,2021,42(11):97-108.
 - LIU C H, GUO C L, YANG Y, et al. Intelligent task-oriented semantic communication method in artificial intelligence of things[J]. Journal on Communications, 2021, 42(11): 97-108.