

# 面向铁路入侵检测的语义通信技术



## Semantic Communication Technology for Railway Intrusion Detection

郭疆远/GUO Jiangyuan, 陈为/CHEN Wei, 艾渤/AI Bo

(北京交通大学, 中国 北京 100044)  
(Beijing Jiaotong University, Beijing 100044, China)

DOI: 10.12142/ZTETJ.202504003

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20250715.1637.002.html>

网络出版日期: 2025-07-16

收稿日期: 2025-05-25

**摘要:** 针对铁路入侵检测中海量视频数据传输效率低、检测精度低的问题, 提出了一种基于视频Transformer的自适应语义通信框架。该框架通过传输与任务高度相关的语义特征并集成信道自适应模块, 实现面向特定任务的高效且鲁棒的语义信息传输。在构建的铁路真实场景视频数据集上进行实验, 将该框架与传统的视频分离编码及视频联合编码方法进行对比。结果表明, 所提出的自适应语义通信框架在不同的高斯白噪声及多种衰落信道条件下均能取得更高的入侵检测精度, 并在低信噪比和复杂信道环境中展现出优越的鲁棒性和性能增益, 为提升铁路智能监控系统的视频分析与传输效能提供了新的技术途径。

**关键词:** 入侵检测; 语义通信; 深度联合信源信道编码; 信道自适应

**Abstract:** To address the challenges of low efficiency in transmitting massive video data and suboptimal detection accuracy under complex channel conditions for railway intrusion detection, an adaptive semantic communication framework based on Video Transformer is proposed. This framework achieves efficient and robust task-oriented semantic information transmission by conveying highly task-relevant semantic features and integrating a channel-adaptive module. Experiments conducted on a constructed real-world railway scenario video dataset show that, compared with traditional separate video coding and joint video coding methods, the proposed framework attains higher intrusion detection accuracy under various additive white Gaussian noise and diverse fading channel conditions. Furthermore, it exhibits superior robustness and performance gains in low signal-to-noise ratio and complex channel environments, offering a novel technical approach for enhancing the video analysis and transmission efficacy of intelligent railway monitoring systems.

**Keywords:** intrusion detection; semantic communication; deep joint source-channel coding; channel adaptation

**引用格式:** 郭疆远, 陈为, 艾渤. 面向铁路入侵检测的语义通信技术 [J]. 中兴通讯技术, 2025, 31(4): 13-18. DOI: 10.12142/ZTETJ.202504003

**Citation:** GUO J Y, CHEN W, AI B. Semantic communication technology for railway intrusion detection [J]. ZTE technology journal, 2025, 31(4): 13-18. DOI: 10.12142/ZTETJ.202504003

近年来, 人工智能、物联网和新一代无线通信技术的迅猛发展, 推动了铁路系统的数字化与智能化的深度变革。在“智慧铁路”“数智交通”战略引导下, 铁路行业正加快迈向高度自动化、感知化与智能决策的新阶段<sup>[1]</sup>。其中, 视频监控作为铁路感知体系的关键组成部分, 广泛部署于站台、咽喉、线路、道口等重点区域, 承担着入侵检测、施工监管、车流监控等重要任务。在这些智能任务中, 海量视频数据的实时、高效传输与分析处理已成为制约系统性能提升的核心瓶颈。一方面, 现有通信系统大多仍以传统

比特级传输为设计理念, 将视频看作像素流处理, 忽视了其承载的“任务语义信息”; 另一方面, 铁路场景中的通信链路环境复杂, 受地形遮挡、高速移动、异构终端功耗等因素的影响, 通信信道常处于低质量状态, 这进一步影响了数据传输的稳定性与下游智能任务的精度。

在铁路智能感知任务中, 入侵检测作为保障行车安全的核心应用之一, 正日益受到重视。该任务旨在及时识别并报警非法进入铁路防护区的人员、动物或障碍物, 避免异物入侵导致的行车事故。入侵检测系统通常依赖于部署在站台边缘、区间线路、道口等重点区域的高清摄像头, 通过对视频流的持续分析, 实现全天候、全场景的智能监控。然而, 铁路入侵检测面临诸多挑战: 一方面, 入侵事件通常具有突发

基金项目: 国家自然科学基金项目 (U2468201, 6221001)

性和稀疏性，要求系统具备更强的视频内容识别与理解能力；另一方面，系统需要在资源受限的边缘设备上实现低时延、高精度的实时检测，并通过通信链路将关键警报信息快速回传至控制中心。因此，实现高效准确的入侵检测，在保障检测精度的同时提升系统效率，是支撑铁路智能安防控制体系的关键技术支柱。

### 1 铁路入侵检测面临的挑战

传统的铁路入侵检测方法主要涵盖图像差分法和光流法，它们凭借像素域或光流域中的先验知识来检测入侵<sup>[2]</sup>，但容易造成次优的检测性能。基于深度学习的入侵检测方法因其巨大潜力备受关注，可基于单帧图像或连续视频帧实现目标检测。基于图像的入侵检测<sup>[3]</sup>借助卷积神经网络等神经网络处理特定关键帧来实现入侵识别。而基于视频的入侵检测能够剖析带有额外时间维度的连续视频帧，可提升入侵检测精度，但它依赖复杂的网络来处理视频中的时间关系，这使得在计算和存储资源受限的铁路监控设备上部署与运行该网络面临诸多挑战。

为了应对上述挑战，一种可行的解决方案是将计算负担卸载到云端服务器。在此方案中，监控摄像头捕获的原始视频通过无线信道传输至云端服务器，由其执行后续视频处理和入侵检测。然而，该方式需将完整视频逐帧传输至云端，这不仅带来巨大的带宽压力，而且由于大量图像内容（如空轨道、背景区域）与检测任务无关，存在显著冗余，导致通信效率极其低下。同时，传统视频通信系统普遍采用“视频压缩—信道传输—视频重建—智能识别”的串行架构，其视频压缩算法（如H.264、H.265）主要围绕像素误差进行优化，关注图像客观质量而非任务相关性，导致压缩过程中保留大量无用信息。在低信噪比条件下，基于该架构的视频重建质量大幅下降，进一步降低了后续识别算法（如目标检测、行为分析）的性能，容易引发漏检、误报等问题。此外，视频重建范式对接收端算力要求较高，接收端需顺序完成图像重建与智能推理，造成系统处理链路长、响应延迟高，且任一环节失效均可能影响整体识别效果。由此可见，仅依靠提升压缩效率或增强链路可靠性，难以从根本上解决传统视频通信系统带来的性能瓶颈问题。

作为一种融合通信技术与人工智能的新型范式，语义通信正逐渐受到学术界和产业界的广泛关注<sup>[4-5]</sup>。该范式主张在通信过程中

“只传递对任务有意义的信息”，本质上是从“面向信息的压缩”向“面向任务的压缩”的转变。通过在发送端提取任务相关的语义特征，语义通信可以有效减少传输数据量，降低带宽压力，同时保障下游智能任务的精度<sup>[6]</sup>。目前，语义通信技术已经在图像分类、目标识别、语音理解等任务中展示出优秀的性能<sup>[7-8]</sup>。然而，在铁路这样的强场景约束、高安全要求、高移动特性环境中，语义通信的设计仍面临一系列挑战，例如：语义的有效提取、端侧计算能力的匹配、信道质量波动下的鲁棒性保障等问题，亟需结合场景需求进行定制化设计。

本文将新兴的语义通信技术引入铁路场景并聚焦于铁路入侵检测这一典型智能任务，提出了一种面向铁路入侵检测的语义通信传输框架。该框架结合深度联合信源信道编码技术<sup>[9]</sup>，以视频Transformer作为骨干网络，在保障传输效率的同时提升任务识别精度。通过集成一个基于注意力机制的信道自适应模块，该框架可以有效地适应不同信道类型与信道质量，以提升铁路复杂无线环境下的传输鲁棒性。通过构建铁路咽喉区实景视频数据集，本文对所提出框架的检测准确率与传输性能进行了系统评估，验证了该方法在典型铁路通信场景中的有效性。

### 2 基于语义通信的传输框架设计

#### 2.1 系统模型

本文所提出的传输框架主要包含3个核心组件：

- 1) 发送端语义编码器：部署在前端摄像头设备或轻量化边缘节点中，负责提取原始视频帧中与任务相关的语义特征，并编码为嵌入向量；
- 2) 信道自适应传输机制：通过注意力机制将嵌入向量映射为适配无线信道特点的传输符号，增强抗干扰能力；
- 3) 接收端语义解码器：部署于车站云端服务器或调度中心，负责从接收到的传输符号中还原出语义表示，并完成入侵检测任务。

具体传输流程如图1所示，主要分为3步：

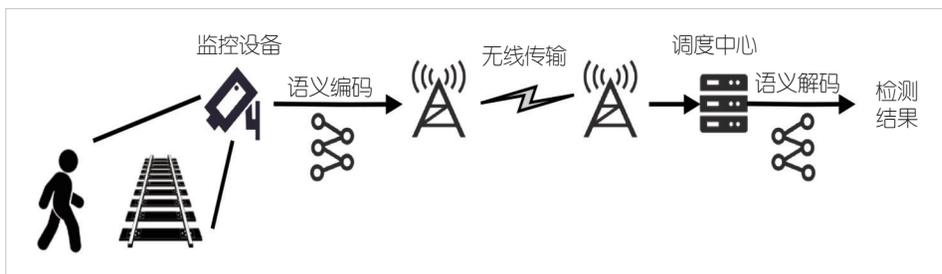


图1 面向铁路入侵检测的语义通信传输框架

1) 首先, 发送端捕获的监控视频可以表示为  $\mathbf{x} \in \mathbb{R}^{3 \times t \times h \times w}$ , 其中  $t$  表示时间步长, 即视频帧的数量,  $h$  表示视频的高度,  $w$  表示视频的宽度。发送端部署的语义编码器包含  $N_1$  个视频 Transformer 基本处理块。与视觉 Transformer 的处理类似, 为了降低计算量,  $\mathbf{x}$  首先沿着时间和空间维度被划分为  $l$  个不重叠的尺寸为  $1 \times p \times p$  的补丁块, 然后被输入至后续神经网络中处理。发送端利用语义编码器, 结合通过信道估计等方法获取的先验信道信息(信噪比、信道类型等)实现动态编码, 从原始视频  $\mathbf{x}$  中提取任务相关的深层视频语义表示  $\mathbf{y} \in \mathbb{R}^{d \times t \times l}$ , 其中  $d$  为视频语义表示的通道数。在发送端的处理过程可以表示为:

$$\mathbf{y} = E(\mathbf{x}, \boldsymbol{\mu}) \quad (1),$$

其中,  $E(\cdot)$  为语义编码函数,  $\boldsymbol{\mu}$  为先验信道信息。随后, 通过将提取到的深层视频语义表示  $\mathbf{y}$  的每两个实数值组合为一个复数符号, 以得到信道输入信号  $\mathbf{s} \in \mathbb{C}^m$  ( $m$  为信道传输符号数量)。假设一个符号的发送功率为  $P$ , 在进行实际传输之前, 需要对  $\mathbf{s}$  进行功率归一化操作以满足发送端的平均功率约束, 即:

$$\frac{1}{m} \mathbf{s} \mathbf{s}^* = P \quad (2).$$

2) 发送端将归一化后的信道输入信号  $\mathbf{s}$  以模拟信号形式直接通过无线噪声信道发送至接收端。

本文考虑两种无线信道下的通信过程, 即加性高斯白噪声信道与衰落信道。加性高斯白噪声信道的传输过程可以表示为:

$$\hat{\mathbf{s}} = \mathbf{s} + \mathbf{n} \quad (3),$$

其中,  $\hat{\mathbf{s}}$  为信道输出信号,  $\mathbf{n} \in \mathbb{C}^m$  为噪声向量, 包含  $m$  个遵从复高斯分布  $\mathcal{CN}(0, \sigma^2 \mathbf{I})$  的独立同分布元素,  $\sigma^2$  表示噪声功率,  $\mathbf{I}$  为单位矩阵。衰落信道的传输过程可以表示为:

$$\hat{\mathbf{s}} = \mathbf{h} \mathbf{s} + \mathbf{n} \quad (4),$$

其中,  $\mathbf{h} \in \mathbb{C}^l$  表示信道增益。

3) 接收端首先接收信道输出信号  $\hat{\mathbf{s}}$ , 并将每个复数符号分解为两个实数值以转换为重建的语义表示  $\hat{\mathbf{y}} \in \mathbb{R}^{d \times t \times l}$ 。随后, 接收端通过部署的语义解码器对  $\hat{\mathbf{y}}$  进行语义解码, 最终通过一个由线性映射层构成的分类头得到识别为入侵/非入侵的概率。与语义编码器类似, 语义解码器包含  $N_2$  个视频 Transformer 基本处理块, 并结合与发送端共享的先验信道信息实现动态解码。在接收端的处理过程可以表示为:

$$\mathbf{p} = D(\hat{\mathbf{y}}, \boldsymbol{\mu}) \quad (5),$$

其中,  $D(\cdot)$  为语义解码函数,  $\boldsymbol{\mu}$  为先验信道信息,  $\mathbf{p} \in \mathbb{R}^2$  为入侵检测概率向量, 表示对应视频被检测为入侵或非入侵的可能性。具有最高可能性的类别将作为入侵检测结果  $c$ :

$$c = \underset{i}{\operatorname{argmax}} p_i \quad (6),$$

其中,  $p_i$  表示  $\mathbf{p}$  的第  $i$  个元素。

### 2.2 基于视频 Transformer 的语义编解码

在铁路智能监控系统中, 入侵检测的核心挑战在于如何从复杂动态的视频流中快速、准确地识别稀疏且突发的入侵目标(如人员、车辆或障碍物)。传统方法依赖卷积神经网络或光流, 但受限于局部感受野与计算冗余, 难以兼顾时空维度实现联合建模。为此, 本文提出一种基于视频 Transformer 的语义编解码网络, 通过多层次的自注意力机制, 实现对铁路监控视频的高效语义特征提取与任务导向的压缩传输。

给定第  $i$  层输入特征  $\mathbf{y}_i$ , 输出特征  $\mathbf{y}_{i+1}$  由  $\mathbf{y}_i$  经一个完整的视频 Transformer 块处理获得, 如图 2 所示。每个视频 Transformer 块包含 3 个核心组件: 空间自注意力模块、时间自注意力模块以及前馈网络。其设计目标是通过多粒度特征

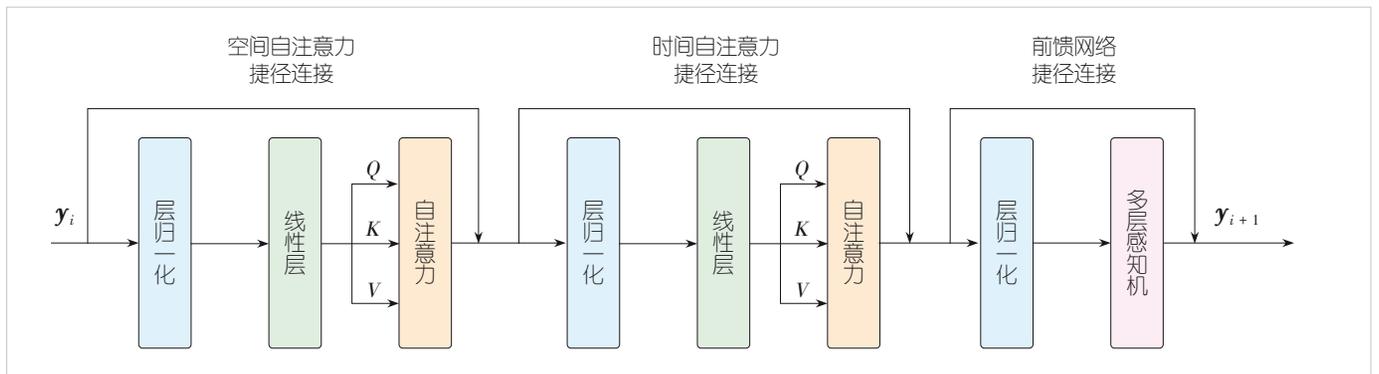


图2 视频Transformer块的网络结构

交互，精准捕捉铁路场景中入侵目标的静态形态与动态行为。其中，空间自注意力模块在空间维度实现注意力操作，聚焦单帧内的局部与全局空间关系，提取入侵目标的形态、位置等静态特征。空间自注意力执行完毕后，引入时间自注意力以建模连续帧间的时序依赖关系，分析入侵目标的运动轨迹与行为模式。最后，利用前馈网络融合时空特征，提升模型的非线性建模能力。考虑到边缘设备存在的算力限制，在本文所提出的传输架构中，发送端部署  $N_1 = 2$  个视频 Transformer 基本处理块，而接收端部署  $N_2 = 10$  个视频 Transformer 基本处理块。

### 2.3 基于注意力机制的信道自适应

铁路环境下的无线信道本质上是复杂且动态变化的，其特性受到多种因素的综合影响，表现为信噪比的时变性以及衰落特性的多样性。这种信道的不确定性和时变性给视频语义信息的可靠传输带来了严峻挑战，严重影响了下游入侵检测任务的精度和稳定性。

为了解决这一核心难题，本文为所提出的语义通信框架集成了一套信道自适应传输机制。该机制的核心目标在于赋予通信系统感知和适应信道变化的能力，通过在发送端和接收端共享并利用先验信道信息，动态优化语义信息的编码和解码，从而在多变信道条件下最大程度地保障通信的鲁棒性和语义信息的完整性与准确性。此处的先验信道信息不仅包括了量化的信道质量指标，如瞬时或平均信噪比，还涵盖了对信道类型的定性判断，例如当前链路符合的衰落模型。

所提出的信道自适应方法包含一个基于注意力机制的即插即用模块，该模块以在收发双方共享的信道先验信息为输入，实时生成一组缩放向量或仿射变换参数。这些参数用于动态地调整语义编码器与解码器内部的中间层特征，使整个通信系统能够动态地适应当前信道的特定状况。

具体而言，该信道自适应模块所利用的先验信道信息及其处理方式会根据当前无线信道的主要特性进行区分：

1) 针对加性高斯白噪声信道：在该信道条件下，其相对简单且主要受噪声功率影响，信道自适应模块所利用的关键先验信息是瞬时或平均信噪比。信噪比直接反映了信号强度与噪声强度之间的对比关系，是衡量此类信道质量的核心指标。

2) 针对更复杂的衰落信道：对于由多径效应主导，导致信号包络快速且剧烈起伏的衰落信道，仅仅依赖单一的信噪比信息，往往难

以全面而准确地刻画信道的动态特性和对信号的实际影响。信噪比主要反映的是信号平均功率与噪声平均功率的对比，但无法充分体现衰落信道中的信号时变特性。因此，为了更有效地应对这类复杂衰落信道带来的挑战，在继续利用实时信噪比作为重要的信道质量参考的同时，额外引入了“信道类型标签”这一关键的上下文信息。这个标签并非简单的数值，而是对当前信道所属的具体衰落模型（如瑞利、莱斯等）或其所处的典型物理场景特征（如城市峡谷、开阔地、隧道内等）的定性描述或分类。引入信道类型标签，系统能够获得关于信道本质特征的更深层次理解，为后续的自适应调整提供更为精准的依据。

信道自适应模块的结构如图3所示。给定第  $i - 1$  层视频 Transformer 块输出特征  $y_i \in \mathbb{R}^{d_i \times t_i \times l_i}$ ， $d_i$  表示特征通道数， $t_i \times l_i$  表示特征分辨率。首先利用多层感知机处理信噪比及信道类型等先验信道信息，生成缩放因子。其中，信道标签  $r$  为一个整数值，代表对应的信道类型（5个信道类型则对应1—5的信道标签）。信噪比（SNR）为一个浮点数值，表示对应的信道质量。二者首先被合并在一起，组成先验信道信息  $\mu$ ，表示为：

$$\mu = \text{concat}(\text{SNR}, r) \in \mathbb{R}^2 \tag{7}$$

其中， $\text{concat}(\cdot)$  表示合并操作。 $\mu$  经多层感知机处理得到缩放因子  $\alpha \in \mathbb{R}^{d_i}$  与  $\beta \in \mathbb{R}^{d_i}$ ，表示为：

$$[\alpha, \beta] = \text{MLP}(\mu) \in \mathbb{R}^{2d_i} \tag{8}$$

其中， $\text{MLP}(\cdot)$  表示由两个线性映射层以及中间的非线性激活构成的多层感知机。

随后， $\alpha$ 、 $\beta$  被扩展为  $\mathbb{R}^{d_i} \rightarrow \mathbb{R}^{d_i \times 1 \times 1}$  并与中间层特征  $y_i$  进行通道维度的点积与求和，生成调整后的特征  $\tilde{y}_i \in \mathbb{R}^{d_i \times t_i \times l_i}$ 。基于注意力机制的调整过程可以表示为：

$$\tilde{y}_i = \alpha \times y_i + \beta \tag{9}$$

其中， $\times$  和  $+$  分别表示通道维度的点积与求和操作。

本文中，每个视频 Transformer 基本处理块后均嵌入一个

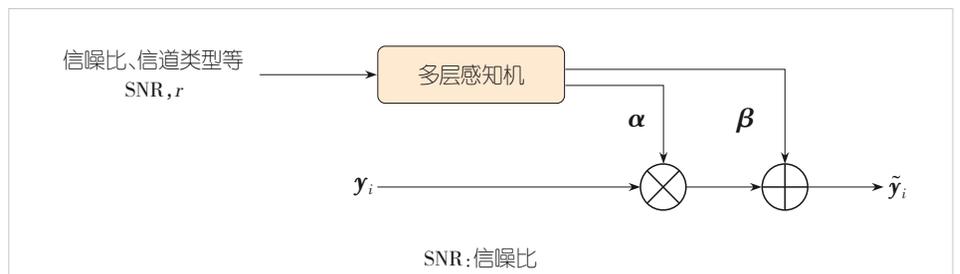


图3 信道自适应模块的结构

信道自适应模块以调整中间特征，从而使得语义编码和语义解码过程能够适配当前信道状态。每一层的处理包含一个视频 Transformer 块以及一个信道自适应模块。

### 3 实验结果与分析

#### 3.1 实验设置

本文构建了一个真实的视频监控数据集，以全面评估所提方法性能。数据采集自中国境内 65 个实际铁路咽喉区（通过云台摄像机在 2018 年 12 月—2019 年 4 月采集的监控视频）。这些区域因轨道分支和交叉复杂，安全风险较高。数据集包含 9 197 个样本，每个样本由 8 个连续帧构成，并统一调整分辨率至  $224 \times 224$  像素，用于训练与测试。样本以 5 : 1 的比例划分为训练集和测试集，确保入侵与非入侵样本在各数据集中数量均衡。视频 Transformer 模块中，时空自注意力机制均设置 8 个注意力头，每个头通道数为 64，语义特征的通道数为  $d = 256$ 。

为验证所提出方法的有效性，本文与多种基准方法对比，包括理想化的无差错传输、传统的 H.265 视频编码结合信道容量传输方案（使用 FFmpeg 进行 H.265 编码），以及一种先进的基于深度联合信源信道编码的视频传输方法<sup>[10]</sup>。同时，我们还对比了移除信道自适应模块的基础传输框架，以凸显信道自适应模块的性能增益。

本文主要考虑了高斯白噪声信道以及 3GPP 中定义的 3 种衰落信道下的视频语义传输。所有模型均在 0 ~ 20 dB 的混合信噪比条件下进行训练和测试，且在性能评估时各方法使用相同带宽资源（传输符号数量一致）。训练中，采用分类任务常见的交叉熵损失函数优化模型参数。

#### 3.2 实验分析

图 4 展示了不同传输方法在高斯白噪声信道下的入侵检测准确率对比。实验结果表明，本文提出的语义通信框架显著优于依赖视频重建的传统通信方案。具体而言，提出的自适应传输方法在整个 0 ~ 20 dB 的信噪比区间内均表现最佳，即使在 0 dB 的恶劣条件下，准确率仍高达 84.8%，并随信噪比改善迅速逼近约 88% 的理想传输上限。相比之下，传统的 H.265 编码结合信道容量的方案（视频分离编码方案）和先进的视频联合编码方法在低信噪比时性能严重恶化（0 dB 时准确率分别仅为约 66% 和 69%）。即使在高信噪比下，其准确率上限也与本文方法存在明显差距。这种性能差异源于设计理念的根本不同：传统基于重构的方法侧重于像素级保真，不仅对信道噪声敏感，还传输了大量与入侵检测任务无

关的冗余信息，导致任务精度受损。而本文的语义通信框架通过端到端优化，仅提取和传输与任务高度相关的语义特征，极大地提升了通信效率和噪声鲁棒性。此外，信道自适应机制可以使系统更有效地利用信道信息，以克服信道迅速恶化带来的影响。

图 5 展示了在信道描述语言（CDL）系列衰落信道及不同信噪比条件下，多种训练策略及是否利用信道类型先验信息对入侵检测性能的影响。首先，混合信道类型训练能有效提升模型的泛化能力。在不输入类型标签的情况下，经混合类型场景训练的模型准确率普遍高于仅在单一特定信道类型（CDL-A、C、D）下训练的模型，这表明多样化训练数据有助于学习更鲁棒的普适特征。其次，当模型能获取并利用当前信道的类型标签时，性能会得到显著提升。输入类型标签

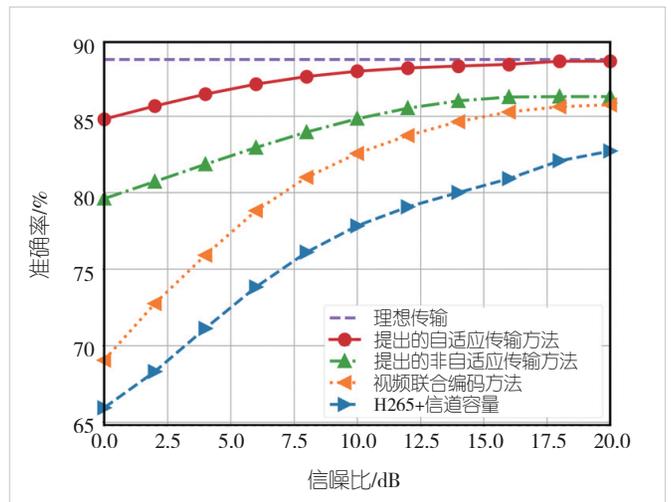


图 4 高斯白噪声信道下不同传输方法的性能比较

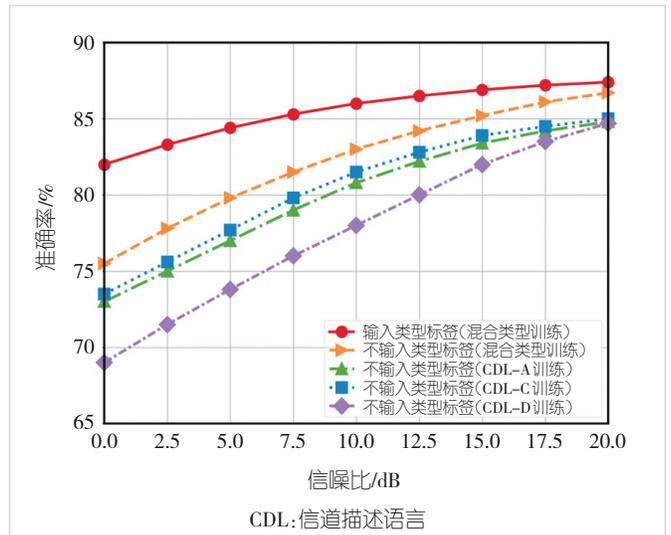


图 5 衰落信道下不同训练方法的性能比较

且经混合类型训练的模型,相较于所有不使用类型标签的模型(包括同样混合类型训练但不输入标签的模型),在各信噪比下均取得了更高的检测精度。这说明类型标签作为精确的信道先验,能引导模型进行针对性优化,有效克服特定衰落带来的影响。此外,在特定类型(CDL-A、C、D)下训练的模型间性能差异,反映了不同衰落场景本身的学习难易度与信道特性不同,进一步说明类型标签能够辅助模型训练,从而带来性能增益。因此,赋予模型场景感知能力并结合覆盖多种信道条件的混合训练,是确保在复杂衰落信道中实现高效、鲁棒入侵检测的关键。

#### 4 结束语

本文面向铁路场景下对系统安全至关重要的入侵检测任务,提出了一种基于视频Transformer的自适应语义通信框架。该框架通过直接提取并传输与入侵检测任务高度相关的语义特征,并结合信道自适应机制(考虑信噪比及信道类型先验),有效克服了传统视频传输方法在高带宽需求、信道噪声敏感以及任务信息冗余等方面的局限性。实验结果表明,无论在高斯白噪声信道还是复杂的CDL衰落信道下,本文方法均显著优于传统的视频分离编码方案及先进的视频联合编码方法。在各种信噪比条件下,本文方法均能以更低的传输开销实现更高的入侵检测精度,展现出强大的鲁棒性和传输效率。这项工作为铁路智能化系统的发展提供了新的技术思路,展现了语义通信在提升铁路场景智能任务性能方面的巨大潜力。

#### 参考文献

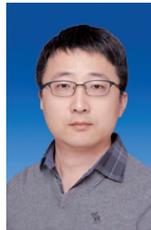
- [1] 钟章队, 官科, 陈为, 等. 铁路新一代移动通信的挑战与思考 [J]. 中兴通讯技术, 2021, 27(4): 44-50. DOI: 10.12142/ZTETJ.202104009
- [2] CAO Z W, QIN Y, JIA L M, et al. Railway intrusion detection based on machine vision: a survey, challenges, and perspectives [J]. IEEE transactions on intelligent transportation systems, 2024, 25(7): 6427-6448. DOI: 10.1109/tits.2024.3412170
- [3] GONG X, CHEN X, ZHONG Z, et al. Enhanced few-shot learning for intrusion detection in railway video surveillance [J]. IEEE transactions on intelligent transportation systems, 2021, 23(8): 11301-11313
- [4] 吕守晔, 戴金晟, 张平. 信源信道联合的新范式: 语义通信 [J]. 中兴通讯技术, 2022, 27(2): 2-8. DOI: 10.12142/ZTETJ.202302002
- [5] 施雨轩, 吴泳澎, 张文军. 基于信息论的语义通信: 理论与挑战 [J]. 中

- 兴通讯技术, 2023, 29(2): 13-18. DOI: 10.12142/ZTETJ.202302004
- [6] GUO J Y, CHEN W, SUN Y X, et al. VideoQA-SC: adaptive semantic communication for video question answering [J]. IEEE journal on selected areas in communications, 2025, 43(7): 2462-2477. DOI: 10.1109/jsac.2025.3559160
- [7] GUO L, CHEN W, SUN Y X, et al. Digital-SC: digital semantic communication with adaptive network split and learned non-linear quantization [EB/OL]. (2024-12-02) [2025-05-20]. <https://ieeexplore.ieee.org/document/10772628>
- [8] CHENG J M, CHEN W, LI J, et al. Efficient spatio-temporal predictive learning for massive MIMO CSI prediction [J]. ZTE Communications, 2025, 23(1): 3-10. DOI: 10.12142/ZTECOM.202501002
- [9] 许佳龙, 陈为, 艾渤. 基于深度联合信源信道编码的CSI反馈技术 [J]. 中兴通讯技术, 2022, 27(2): 29-33. DOI: 10.12142/ZTETJ.202302007
- [10] WANG S, DAI J, LIANG Z, et al. Wireless deep video semantic transmission [J]. IEEE journal on selected areas in communications, 2022, 41(1): 214-229. DOI: 10.1109/JSAC.2022.3221977

#### 作者简介



**郭疆远**, 北京交通大学在读博士研究生; 主要研究方向为深度联合信源信道编码、语义通信; 发表论文3篇。



**陈为**, 北京交通大学教授; 长期从事无线通信、信号处理、人工智能技术研究, 并围绕高维信息感知和处理、海量机器类通信、语义通信、智慧交通开展研究工作; 作为负责人承担国家优秀青年科学基金等多项省部级及以上项目; 获IEEE通信学会杰出讲师、2022年度信息通信领域十大科技进展、首届詹天佑铁道科学技术奖秦驰道专项奖(青年奖)等荣誉。



**艾渤**, 北京交通大学教授、电子信息工程学院院长, 国家6G技术研发总体专家组专家, 国家自然科学基金委创新群体带头人, 国家杰出青年科学基金、优秀青年科学基金、牛顿高级学者基金、中国科协“求是杰出青年奖”获得者, 中共中央组织部“万人计划”领军人才, 北京市优秀教师, IEEE Fellow, IEEE VTS杰出讲师; 入选斯坦福大学2020-2022年全球前2%顶尖科学家终身影响力榜单, 获省部级科技奖励9项。