

# Deep Learning Enabled Semantic Communication Systems

Huiqiang Xie, Zhijin Qin, Member, IEEE, Geoffrey Ye Li, Fellow, IEEE, and Bing-Hwang Juang Life Fellow, IEEE

## 摘要

近年来,深度学习支持的端到端 (E2E) 通信系统被开发出来,用于合并传统通信系统中的所有物理层块,从而使联合收发器优化成为可能。在深度学习的推动下,自然语言处理 (NLP) 在分析和理解大量语言文本方面取得了巨大的成功。受到这两个领域研究成果的启发,我们旨在从语义层面为交际系统提供一个新的视角。特别地,我们提出了一个基于深度学习的语义传输系统 DeepSC,用于文本传输。在变压器的基础上,DeepSC 的目标是最大限度地提高系统容量,最大限度地减少语义错误,恢复句子的意义,而不是传统的比特或符号错误通信。在此基础上,利用迁移学习确保 DeepSC 适用于不同的通信环境,加快模型的训练过程。为了准确地验证语义通信的性能,我们还初始化了一个新的度量,称为句子相似度。仿真结果表明,与不考虑语义信息交换的传统通信系统相比,DeepSC 具有更强的抗信道变化能力,能够获得更好的性能,尤其是在低信噪比 (SNR) 情况下。

**关键词:** 深度学习; 语义通信

## 1 引言

在深度学习的推动下,自然语言处理 (NLP) 在分析和理解大量语言文本方面取得了巨大的成功。本篇论文旨在从语义层面为交际系统提供一个新的视角,提出了一个基于深度学习的语义传输系统 DeepSC,用于文本传输。在 Transformer 的基础上,DeepSC 的目标是最大限度地减少语义错误,恢复句子的意义,而不是传统的比特或符号错误通信。并且利用迁移学习确保 DeepSC 适用于不同的通信环境,加快模型的训练过程。与不考虑语义信息交换的传统通信系统相比,DeepSC 具有更强的抗信道变化能力,能够获得更好的性能,尤其是在低信噪比 (SNR) 情况下。

## 2 相关工作

本部分简要回顾了端到端物理层通信系统和自然语言处理中所采用的深层神经网络 (DNN) 技术的相关工作。

### 2.1 端到端物理层通信系统

DL 技术在处理计算机视觉和自然语言处理等各种智能任务方面显示出巨大的潜力。同时,由于硬件计算能力的提高,训练神经网络并在移动设备上运行它们是可能的。在通信领域,基于 DL 的 E2E 物理层通信系统已经进行了一些开创性的工作,它将传统通信系统中的模块融合在一起。通过采用 DL 语言的自动编码器结构和去除块结构,将 E2E 系统中的发射机和接收机作为 E2E 重构任务进行联合优化。已经证明,这样一个 E2E 系统的性能优于未编码的二进制相移键控 (BPSK) 和汉明编码的 BPSK 的误码率。此外,针对训练过程中缺失的信道梯度问题,也进行了一些初步的研究。提出了一种基于 DNN 的两阶段训练处理方法,其中收发信机通过随机信道模型进行训练,接收机在实际信道下进行微调。在未知信道模型下,强化学习已经被用来获取信道梯度,这种模型在实际信道下比差

分正交相位偏移调制 (DQPSK) 获得更好的性能。在中应用了条件生成对抗网络 (GAN) 来模拟信道输出的分布, 以便在 E2E 通信系统的训练期间, 梯度可以通过一个未知的信道传递给发射机 DNN。考虑到信源的类型, 文本和图像的联合信源信道编码旨在直接恢复接收端的信源信息, 而不是恢复数字比特。同时, 传统的度量方法, 如误码率, 不能很好地反映这类系统的性能。因此, 我们采用误码率和峰值信噪比率 (PSNR) 来衡量信源资料恢复的准确性。

## 2.2 自然语言处理中的语义表示

NLP 使机器能够理解人类语言, 其主要目标是理解语法和文本。语言可以根据上下文用联合概率模型来描述。因此, 语言模型提供上下文来区分具有相似语义的词和短语。尽管这种基于统计模型的自然语言处理 (NLP) 技术是为了描述一个句子中某个单词连续出现的概率, 但它很难处理超过 15 个单词的长句和句法。为了理解长句, 人们提出了单词模型来捕捉单词之间的关系, 使得相似单词在向量空间中的距离更近。即使这些密集的词向量能够捕获词之间的关系, 它们也无法描述语法信息。为了解决这些问题, 文本的潜在意义通过使用各种数字语言技术来表示, 这种技术能够提取长句中的语义信息及其句法。然而, 上面的单词表示方法是为特定任务设计的, 可能需要在任务发生变化时重新设计。

## 2.3 现代自然语言处理技术的比较

有三种类型的神经网络用于自然语言处理任务, 包括递归神经网络 (RNN), 卷积神经网络 (CNN) 和全连接神经网络 (FCNs)。通过引入 RNN, 语言模型可以有效地学习整个句子并获取句法信息。然而, 对于长句, 尤其是主谓距离超过 10 个词以上的长句, RNN 无法找到正确的主谓。例如, 对于句子, 在新邮局工作的人正走向商店, RNN 不能识别人与人之间的关系。此外, 由于神经网络的线性序列结构, 使其缺乏并行计算能力, 这意味着神经网络的时间消耗。CNN 生来就具有并行计算的能力。然而, 即使有线神经网络可以使用更深层的网络来提取长句中的语义信息, 其性能也不如有线神经网络, 因为有线神经网络的内核尺寸很小, 无法保证计算效率。通过结合注意机制, 基于模糊神经网络的语言模型, 如跑马灯, 使模型能够更加关注在各种自然语言处理任务中提高性能的有用语义信息。值得注意的是, 变压器具有 RNN 和 CNN 的优点。特别是采用了自我注意机制, 使模型能够理解不同长度的句子。

# 3 本文方法

## 3.1 本文方法概述

此部分对本文将要复现的工作进行概述, 所考虑的模型包括两个层次: 语义层、传输层如图 1 所示:

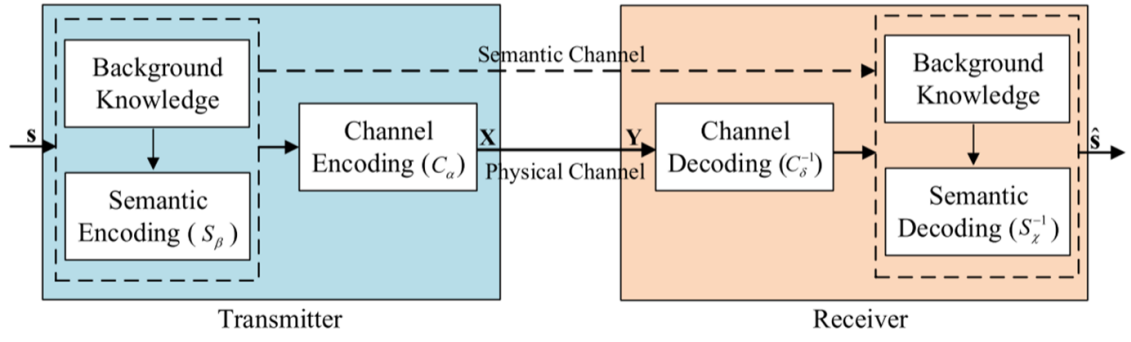


图 1: 模型示意图

语义层解决了编码和解码的语义信息处理，从而提取语义信息。传输级别保证语义信息可以在介质上正确交换。

### 3.2 性能指标

性能标准对系统设计很重要。在 E2E 通信系统中，发射机和接收机通常将误码率作为训练目标，忽视了通信的其他方面的目标。对于文本传输，BER 不能很好地反映性能。双语评价替代评分 (BLEU) 除了用于判断句子之间的相似性外，通常用来衡量机器翻译的结果，并将其作为本文的绩效指标之一。然而，BLUE 分数只能比较两个句子中单词之间的差异，而不能比较它们的语义信息。因此，我们初始化一个新的度量标准，命名为句子相似度，以描述两个句子的相似程度根据他们的语义信息。

## 4 复现细节

### 4.1 与已有开源代码对比

此部分为必填内容。如果没有参考任何相关源代码，请在此明确申明。如果复现过程中引用参考了任何其他入发布的代码，请列出所有引用代码并详细描述使用情况。同时应在此部分突出你自己的工作，包括创新增量、显著改进或者新功能等，应该有足够差异和优势来证明你的工作量与技术贡献。

---

**Initialization:** Initial the weights  $\mathbf{W}$  and bias  $\mathbf{b}$ .

- 1: **Input:** The background knowledge set  $\mathcal{K}$ .
  - 2: Create the index to words and words to index, and then embedding words.
  - 3: **while** Stop criterion is not met **do**
  - 4:   Train the mutual information estimated model.
  - 5:   Train the whole network.
  - 6: **end while**
  - 7: **Output:** The whole network  $S_\beta(\cdot), C_\alpha(\cdot), C_\delta^{-1}(\cdot), S_\chi^{-1}(\cdot)$ .
- 

图 2: 训练 DeepSC 网络

- 
- 1: **Input:** The knowledge set  $\mathcal{K}$ .
  - 2: **Transmitter:**
  - 3:     BatchSource( $\mathcal{K}$ )  $\rightarrow \mathbf{S}$ .
  - 4:      $S_\beta(\mathbf{S}) \rightarrow \mathbf{M}$ .
  - 5:      $C_\alpha(\mathbf{M}) \rightarrow \mathbf{X}$ .
  - 6:     Transmit  $\mathbf{X}$  over the channel.
  - 7: **Receiver:**
  - 8:     Receive  $\mathbf{Y}$ .
  - 9:     Compute loss  $\mathcal{L}_{\text{MI}}$  by (9).
  - 10:    Train  $T \rightarrow$  Gradient descent ( $T, \mathcal{L}_{\text{MI}}$ ).
  - 11: **Output:** The mutual information estimated model  $f_T(\cdot)$ .
- 

图 3: 训练互信息估计模型

- 
- 1: **Input:** The knowledge set  $\mathcal{K}$ .
  - 2: **Transmitter:**
  - 3:     BatchSource( $\mathcal{K}$ )  $\rightarrow \mathbf{S}$ .
  - 4:      $S_\beta(\mathbf{S}) \rightarrow \mathbf{M}$ .
  - 5:      $C_\alpha(\mathbf{M}) \rightarrow \mathbf{X}$ .
  - 6:     Transmit  $\mathbf{X}$  over the channel.
  - 7: **Receiver:**
  - 8:     Receive  $\mathbf{Y}$ .
  - 9:     Compute loss  $\mathcal{L}_{\text{MI}}$  by (9).
  - 10:    Train  $T \rightarrow$  Gradient descent ( $T, \mathcal{L}_{\text{MI}}$ ).
  - 11: **Output:** The mutual information estimated model  $f_T(\cdot)$ .
- 

图 4: 训练互信息估计模型

- 
- 1: **Input:** The knowledge set  $\mathcal{K}$ .
  - 2: **Transmitter:**
  - 3:     BatchSource( $\mathcal{K}$ )  $\rightarrow \mathbf{S}$ .
  - 4:      $S_{\beta}(\mathbf{S}) \rightarrow \mathbf{M}$ .
  - 5:      $C_{\alpha}(\mathbf{M}) \rightarrow \mathbf{X}$ .
  - 6:     Transmit  $\mathbf{X}$  over the channel.
  - 7: **Receiver:**
  - 8:     Receive  $\mathbf{Y}$ .
  - 9:      $C_{\delta}^{-1}(\mathbf{Y}) \rightarrow \hat{\mathbf{M}}$ .
  - 10:     $S_{\chi}^{-1}(\hat{\mathbf{M}}) \rightarrow \hat{\mathbf{S}}$ .
  - 11:    Compute loss function  $\mathcal{L}_{\text{total}}$  by (14).
  - 12:    Train  $\beta, \alpha, \delta, \chi \rightarrow$  Gradient descent  $(\beta, \alpha, \delta, \chi, \mathcal{L}_{\text{total}})$ .
  - 13: **Output:** The whole network  $S_{\beta}(\cdot), C_{\alpha}(\cdot), C_{\delta}^{-1}(\cdot), S_{\chi}^{-1}(\cdot)$ .
- 

图 5: 训练整个网络

---

**Initialization:** Load the pre-trained model  $S_{\beta}(\cdot), C_{\alpha}(\cdot), C_{\delta}^{-1}(\cdot), S_{\chi}^{-1}(\cdot)$ .

**Function:** Training for different background knowledge

- 1: **Input:** The different background knowledge set  $\mathcal{K}_1$ .
- 2: Freeze  $C_{\alpha}(\cdot)$  and  $C_{\delta}^{-1}(\cdot)$ .
- 3: Redesign and train part of  $S_{\beta}(\cdot)$  and  $S_{\chi}^{-1}(\cdot)$ .
- 4: **while** Stop criterion is not met **do**
- 5:     Train the mutual information estimated model.
- 6:     Train the whole network.
- 7: **end while**
- 8: **Output:** The adopted whole network.

**Function:** Training for different channel conditions

- 9: **Input:** The background knowledge set  $\mathcal{K}$  with the different channel parameters.
  - 10: Freeze  $S_{\beta}(\cdot)$  and  $S_{\chi}^{-1}(\cdot)$ .
  - 11: Redesign and re-train part of  $C_{\alpha}(\cdot)$  and  $C_{\delta}^{-1}(\cdot)$ .
  - 12: **while** Stop criterion is not met **do**
  - 13:     Train the mutual information estimated model.
  - 14:     Train the whole network.
  - 15: **end while**
  - 16: **Output:** The re-trained network.
- 

图 6: 基于迁移学习的动态环境训练

## 4.2 复现要点

实现原文章中的伪代码，验证实验结果，并得出相应结论。

## 5 实验结果分析

图 7 显示了训练后信噪比与相互信息之间的关系：互信息随着信噪比的增加而增加，使用互信息估计模型训练的收发信机的性能优于没有这种模型的收发信机。

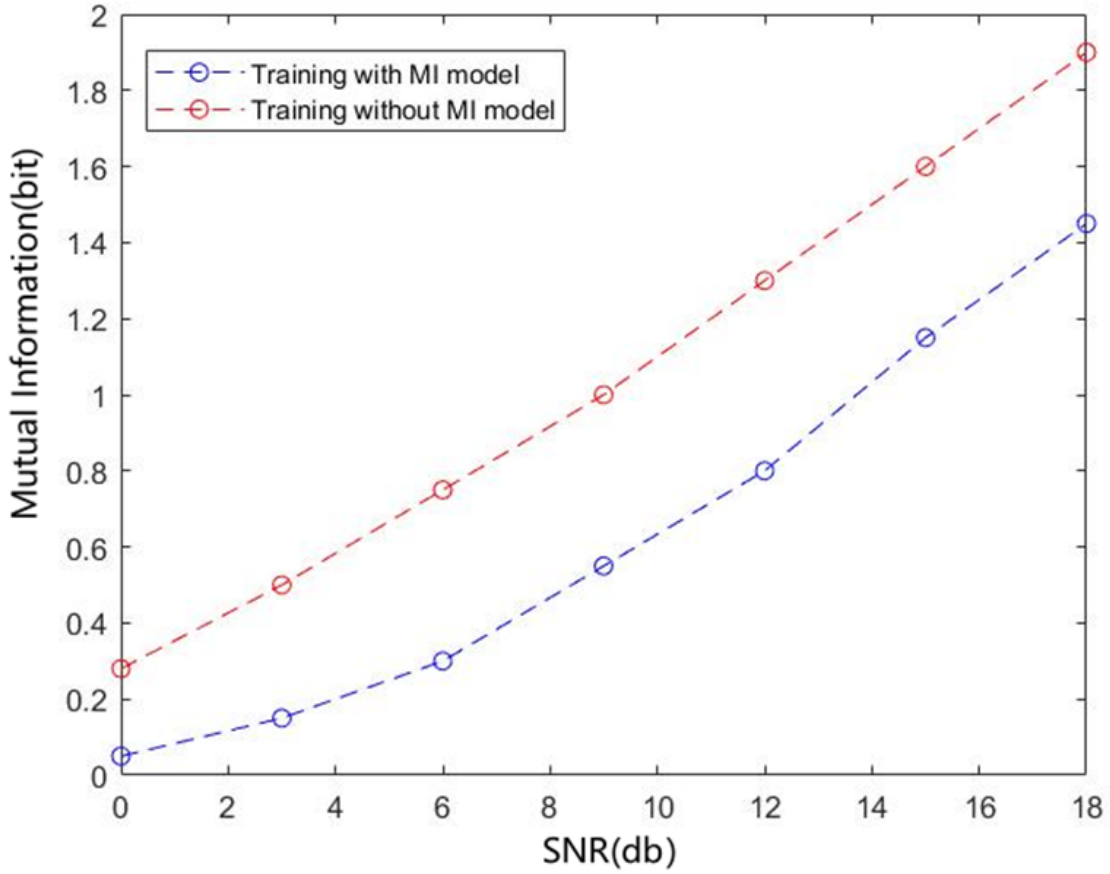


图 7: 训练后的信噪比与互信息

图 7 和 8 显示了不同信道的训练效率和性能，其中 DeepSC 收发信机在 AWGAN 信道下进行预训练，然后在相同的背景知识下分别在擦除信道进行再训练。模型具有相同的结构，并在每个场景中使用相同的参数进行重新训练。采用预训练模型可以加快消除信道和的训练过程。使用预训练模型的 DeepSC 的性能与没有预训练模型通道的性能相似，而在再训练过程中需要更少的 Epoch。

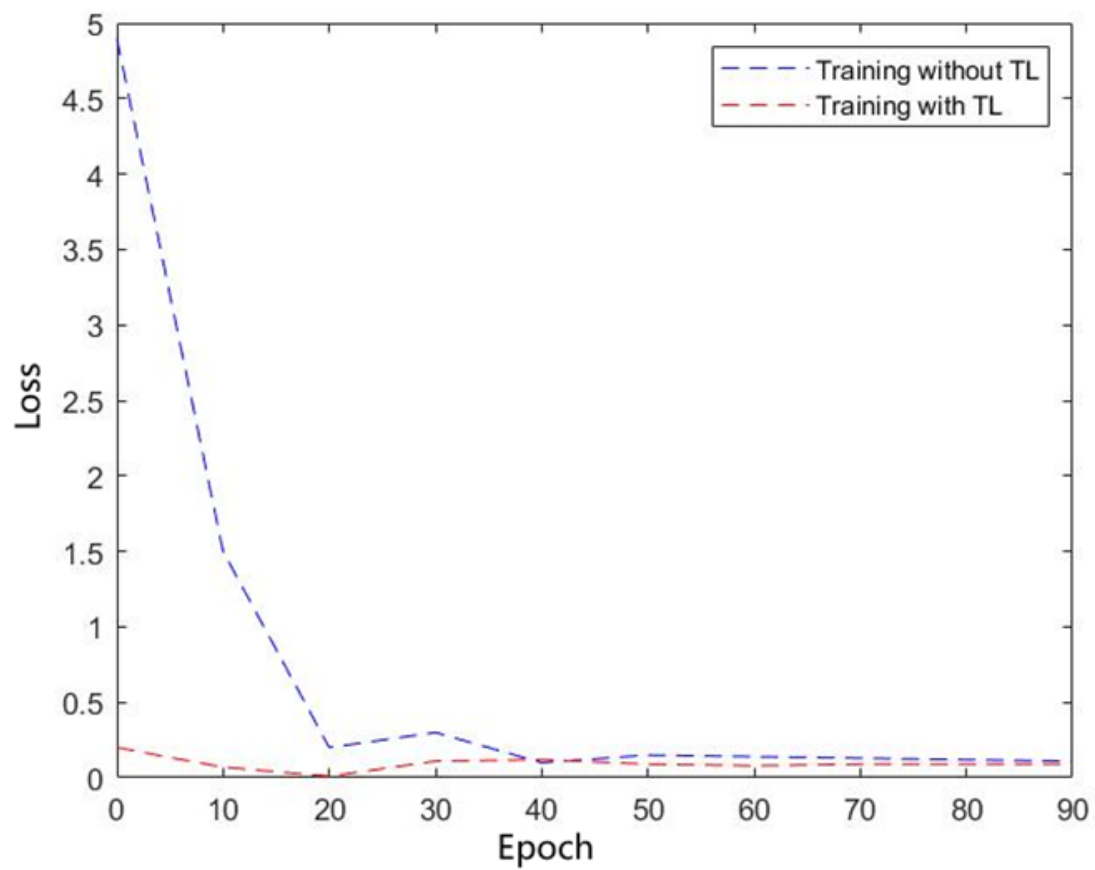


图 8: 不同信道的训练损失

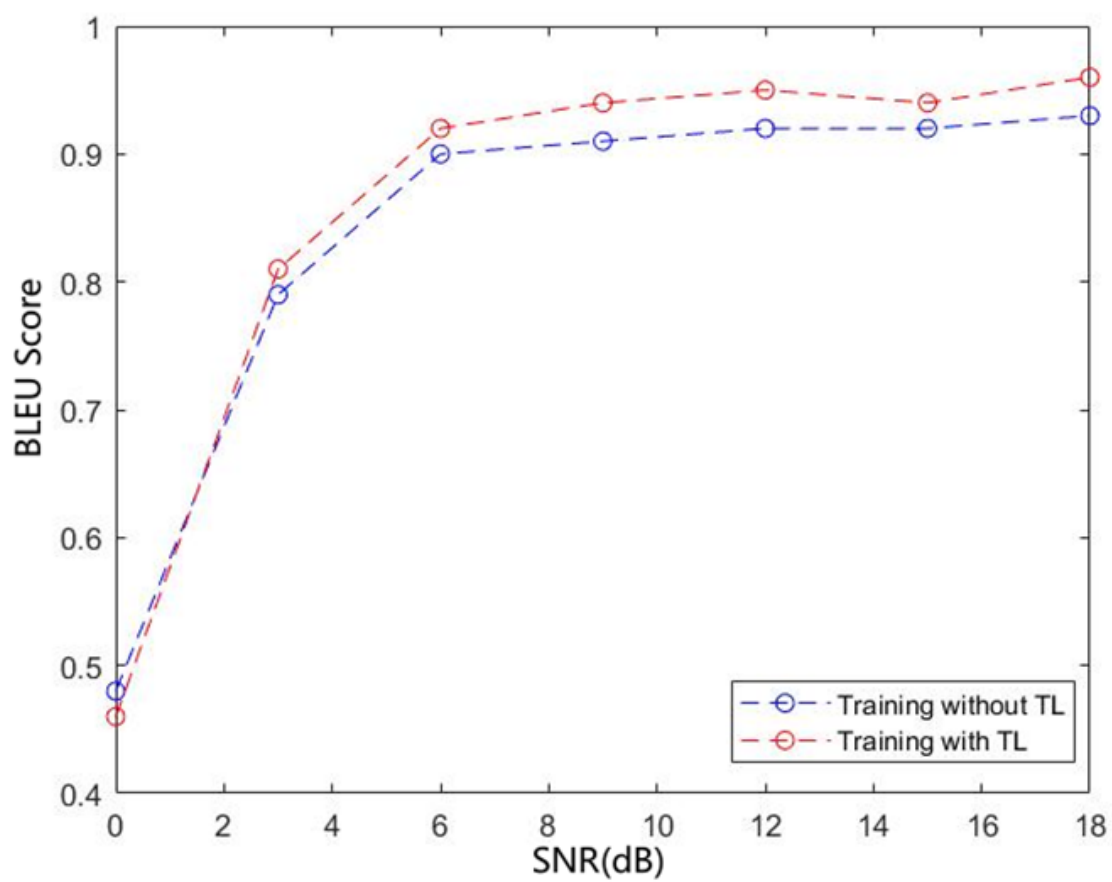


图 9: 不同的信道下的性能评分

## 6 总结与展望

文中提出了一个端到端的语义通信系统 DeepSC,它联合执行文本传输的语义信道编码。在 DeepSC 中,输入文本和输出符号的长度不受限制,互信息作为丢失函数的一部分来实现更高的数据速率。为了满足现实生活中不同的需求,并利用已有的背景知识,加快新网络的训练,采用了深度迁移学习方法。仿真结果表明,DeepSC 在低信噪比情况下性能优于各种基准测试。所提出的迁移学习辅助 DeepSC 已经证明了它具有快速收敛的能力,能够适应不同的通道和知识。因此,文章提出的 DeepSC 是一个很好的文本传输方案,特别是在低信噪比情况下。

复现总结:按照文中伪代码的思路实现模型,并验证其性能和文中结论相符。