

文章编号: 1003-0077(2019)10-0031-07

## 会议场景下融合外部词典知识的领域个性化机器翻译方法

刘庆峰<sup>1</sup>, 刘晨璇<sup>1,2</sup>, 王亚楠<sup>2</sup>, 张为泰<sup>2</sup>, 刘俊华<sup>2</sup>

(1. 语音及语言信息处理国家工程实验室, 安徽 合肥 230027;

2. 认知智能国家重点实验室, 安徽 合肥 230088)

**摘要:** 会议场景下通过语音识别和机器翻译技术实现从演讲人语音到另外一种语言文字的翻译, 对于跨语言信息交流具有重要意义, 成为当前研究热点之一。该文针对由于会议行业属性带来的专业术语和行业用语的翻译问题, 提出了一种融合外部词典知识的领域个性化方法。具体而言, 首先采用联合占位符和拼接融合的编码策略, 通过引入外部词典知识, 在提升实体词、专业术语词翻译准确率的同时, 兼顾翻译结果的流畅性。其次提出基于分类的领域旁支参数个性化自适应策略, 在保持通用领域翻译效果的情况下实现会议相关领域翻译质量的提升。最后基于上述方案, 设计了一套领域个性化自动训练系统。实验结果表明, 在中英体育、商务和医学会议翻译任务上, 该系统在不影响通用翻译的情况下, 平均提升 9.22 个 BLEU, 获得较好翻译效果。

**关键词:** 机器翻译; 词典知识; 领域个性化

中图分类号: TP391

文献标识码: A

### Domain-Specific Machine Translation with External Dictionary Knowledge for Conference Scenarios

LIU Qingfeng<sup>1</sup>, LIU Chenxuan<sup>1,2</sup>, WANG Yanan<sup>2</sup>, ZHANG Weitai<sup>2</sup>, LIU Junhua<sup>2</sup>

(1. National Engineering Laboratory for Speech and Language Information Processing,

Hefei, Anhui 230027, China;

2. State Key Laboratory of Cognitive Intelligence, Hefei, Anhui 230088, China)

**Abstract:** Translation of a presenter's speech into other languages through speech recognition and machine translation in conference scenario is of great significance for cross-language communication. This paper proposes a domain-specific machine translation method based on external dictionary knowledge, so as to handle the translation of terminologies and professional expressions for a given conference. First, a constructed external dictionary is integrated through combining placeholder and concatenation fusion methods. The translation quality of domain related entity words and terminologies is significantly improved while maintaining coherence and fluency of the context. Second, classification-based domain adaptation can further improve the translation of the speech for a given conference while maintaining the overall quality of the general domain translation. Finally, an automatic domain adaptation training system is designed based on the above methods. The experimental results on Chinese to English translation task indicate that the proposed system achieves 9.22 BLEU improvement in average for sports, business and medical conference, without affecting the general translation quality.

**Keywords:** machine translation; dictionary knowledge; domain customization

## 0 引言

随着全球化经济的发展和我国改革开放力度的加大,越来越多的机构和企业正走向世界参与全球

经济的发展和管理工作。日益频繁的跨国交流,对包括口译、笔译及本地化等在内的语言服务提出了大量的、多样化的需求。在传统翻译行业中,译员培养和译员提供服务的时间成本与经济成本均较高,已逐渐进入瓶颈期。通过语音和翻译等技术实现机器自

动口语翻译,辅助跨语言国际沟通交流,成为当前人工智能领域最重要的研究热点之一<sup>[1-2]</sup>。

目前口语翻译的研究和应用有两个典型应用场景,一是不同语言使用者之间面对面的日常口语沟通交流,二是会议场景下讲者面向广大听众的演讲交流。会议场景下,由于会议具有领域专业性,其涉及的大量专业术语和行业相关的语言表达,给机器翻译带来较大挑战。

近年来,深度学习技术的进步推动神经机器翻译技术取得重大进展,同时,在大规模平行数据语料的支撑下,机器翻译效果不断提升。Sutskever 等<sup>[3]</sup>在 2014 年提出了一种编码器—解码器结构的翻译模型,相比传统基于统计的短语翻译方法,该方法生成的译文更加流畅。Bahdanau 等<sup>[4]</sup>在此基础上引入了注意力机制,使得神经机器翻译系统生成的译文全面超过统计机器翻译系统。2017 年, Vaswani 等<sup>[5]</sup>提出的完全基于注意力机制的神经机器翻译模型,能够完全不依赖递归神经网络或卷积神经网络,使得神经机器翻译整体效果又获得大幅提升。

针对会议场景下机器翻译面临的实体词或术语词翻译,以及领域用语翻译的专业性问题,目前也有一些研究工作。在术语词翻译优化上, Hokamp 等<sup>[6]</sup>提出使用 Grid Beam Search 方法进行限制解码,保证短语或词出现在解码译文中; Feng 等<sup>[7]</sup>通过对词典知识进行编码构建 Memory 表示,在训练和解码时引入相应知识表示来提升术语词的翻译效果; Pham 等<sup>[8]</sup>通过对训练数据进行额外标注,使用拼接融合的方法提升稀疏词的翻译效果。在领域自适应训练上, Luong 等<sup>[9]</sup>通过在领域数据上进行模型微调(fine tuning),能够在领域数据集上获得较好的性能提升; Britz 等<sup>[10]</sup>通过引入判别器网络加强翻译模型对领域信息的甄别,从而提升领域数据集的翻译效果。

但是,现有的方法在实际应用中仍存在一些问題。例如,基于占位符<sup>[11]</sup>的方法容易影响流利度,基于拼接融合的方法<sup>[8]</sup>不能保证词语翻译的准确度。因此,本文根据专业和低频词汇的类别不同,分别采用不同的处理策略进行定制优化,在提升翻译准确率的同时兼顾译文的流利度。但在领域个性化方面,现有的方法在提升特定领域效果的同时,其他领域效果一般会存在较明显下降。为此,我们通过基于分类和翻译模型旁支参数(Patch)的优化方法,对实际应用中的句子先进行分类,再通过不同的 Patch 参数进行解码,从而在特定领域获得较大幅度性能提升的同时,几乎不影响通用领域翻译效果。

## 1 基线系统介绍

目前主流的神经机器翻译方法采用编码器—解码器框架<sup>[3-4]</sup>。给定源语言句子  $X = (x_1, x_2, \dots, x_n)$ , 利用编码器将其映射为一组连续稠密的向量表示  $Z = (z_1, z_2, \dots, z_n)$ 。解码器在  $t$  时刻以自回归的形式将已生成词序列  $Y_{t-1} = (y_1, y_2, \dots, y_{t-1})$  和上下文向量  $c_t$  作为条件输入, 计算生成当前词  $y_t$ 。计算如式(1~3)所示。

$$z_i = h_{\text{encoder}}(z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_n) \quad (1)$$

$$d_t = h_{\text{decoder}}(d_{t-1}, y_{t-1}, c_t) \quad (2)$$

$$p(y_t | y_{<t}, X) = \text{softmax}(f(d_t, y_{t-1}, c_t)) \quad (3)$$

其中,  $z_i$  是  $x_i$  的向量表示,  $d_t$  和  $c_t$  分别为  $t$  时刻解码器隐状态和注意力函数给出的上下文向量,  $y_{<t}$  与  $y_{t-1}$  分别表示已生成词序列和前一时刻生成词序列。

本文采用基于注意力结构实现神经机器翻译中的编码、解码和对齐模型的 Transformer<sup>[5]</sup>模型作为基线系统。Transformer 的主要特点在于仅通过自注意力机制计算输入  $X$  序列和输出  $Y$  序列的表示, 通过自注意力机制(self-attention mechanism)将句子中的每个词和该句子中的所有词进行注意力计算, 目的是学习句子内部的依赖关系, 捕获句子的内部结构。

## 2 融合外部词典知识的领域个性化方法

针对会议场景下的领域专业性问题, 本文从专业词汇和语言表达领域个性化两个角度出发, 提出融合外部词典知识的领域个性化方案。在专业词汇个性化方面, 通过引入外部专业词汇双语词典知识提升翻译质量。在会议语言表达领域个性化方面, 则通过引入翻译模型旁支参数(Patch), 并基于会议场景相关数据, 设计一套领域个性化自动训练系统, 进行领域双语数据扩充和领域自适应训练, 在保持通用翻译效果相当的同时, 提升会议场景领域相关的翻译性能。

### 2.1 术语词个性化方法

神经机器翻译采用一套编解码神经网络模型结构实现源语言文字到目标语言文字的翻译, 在提升翻译效果的同时, 也使得词汇翻译的干预变得更加困难。因此, 针对术语词和低频词的翻译, 是目前神经机器翻译研究的一个重要任务。针对该问题, 本文根据专业和低频词汇的类别不同, 采用不同的处理策略进行定制优化。

针对人名、地名和数字等类别和用途明确的实

体词,本文采用占位符替换的翻译方法。该方法通过命名实体识别(name entity recognition, NER)模块获得待翻译的实体词信息,然后将其替换为相应类别的占位符进行翻译,最后将占位符替换为相应的目标译文。该方法对于关键实体词能够取得较好的翻译准确率,并且由于实体词在语句中的用法比较类似,对目标译文的流畅度影响较小。

针对上下文语义相关性较强的术语词翻译,本文借鉴文献[8]的方法,采用融合外部词典知识的拼接融合方法。该方法的初衷是考虑到该类术语词汇在句子中可能占据不同成分,直接采用占位符替换的方法进行翻译会破坏源语言不同词汇之间编码的相关性,同时在解码过程中也会影响上下文之间的连贯性和流畅性。因此本文采用了拼接融合的方法,即根据输入的源语言句子,从外部双语词典中查找相关术语对应的译文,并将其拼接到源语言句子中对应术语词之后的位置,如图 1 所示。由于输入源语言句子中既包括了术语词原始信息,也包括目标译文的信息,翻译模型编码过程中能够看到所有词汇之间的相关性信息,解码过程生成术语词目标译文时也能够兼顾上下文连贯性,从而提升了译文的质量。

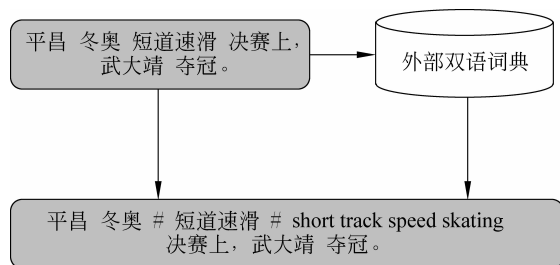


图 1 专业词汇拼接融合方法示意图

具体实现时,该方法将人地名表和术语词典等知识作为外部词典知识对部分训练数据源语言句子进行标注,构建基于术语词的拼接融合训练平行语料,使得翻译模型学习到类似的翻译模式。其次,对翻译模型训练准则也进行了改进,通过引入拷贝网络<sup>[12-13]</sup>和强化学习机制,提高了目标译文包含外部术语词典翻译方式的概率。

## 2.2 领域个性化方法

机器翻译性能与应用领域有很强的相关性,传统领域个性化方案通常基于领域平行数据对翻译模型进行自适应训练<sup>[9-10,14]</sup>。该类方法在提升领域场景翻译效果的同时,往往导致其他领域翻译效果的下降,影响了翻译模型的普适性。针对该问题,本文提出一种基于分类的领域 Patch 个性化定制优化方

法,如图 2 所示。待翻译文本首先经过预先训练好的分类(通用/会议领域)模型进行预测,当该文本被分到会议领域的概率大于设定阈值时,翻译模型选择使用领域 Patch 对其进行编码,否则使用通用编码器进行编码,然后依次进行解码预测,输出译文。

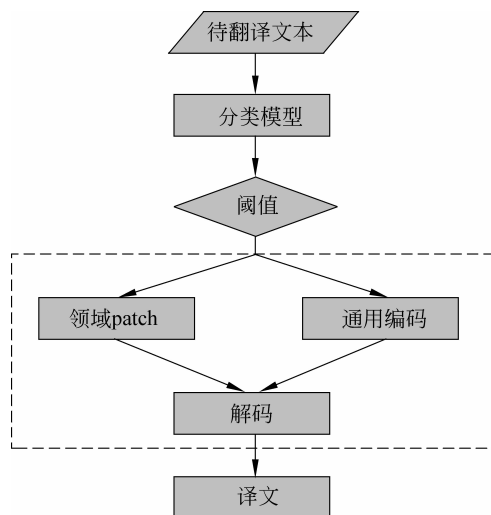


图 2 领域个性化定制方法流程

领域分类器模型方面,本文采用基于字符和词拼接融合的卷积神经网络(Convolutional Neural Networks, CNN)进行建模<sup>[15]</sup>,如图 3 所示。

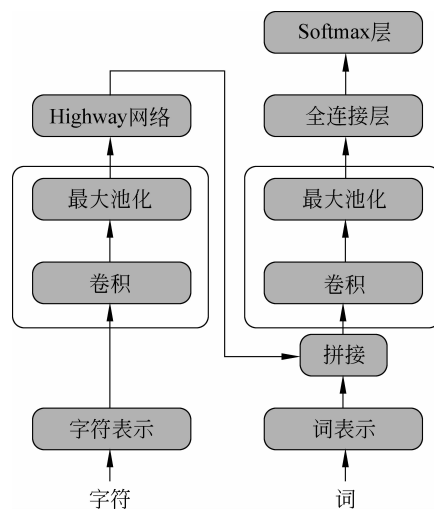


图 3 基于 CNN 的领域分类模型

采用字符的卷积编码与词向量表征拼接融合,能够增强模型对输入拼写错误和识别错误的容错度,进而提升分类模型的鲁棒性。在特征提取网络中,卷积层采用多个不同宽度的卷积核,从多个维度、不同粒度对语义特征进行建模,同时引入 Highway 网络<sup>[16]</sup>提高网络训练收敛速度。最后采用全连接层进行特征向量变换,使用 softmax 函数计算类别概率。

领域个性化模型训练方面,首先将训练好的通

用翻译模型的编码参数复制一份作为领域 Patch 参数的初始值,然后基于会议场景下领域相关的平行语料对领域 Patch 参数通过模型微调进行自适应训练,其他模型参数则和通用翻译模型共享。该方法在保持整体翻译模型大小增加有限的情况下,获得与通用翻译模型相当的效果,并有效提升会议领域场景下的翻译性能。

### 2.3 领域个性化自动训练系统

翻译模型的训练包含数据预处理、模型训练、模型评估和模型部署等流程,往往需要专业人员的参与。考虑到实际应用中具体会议涉及领域较多,人工成本较大,本文设计了一套自动训练系统,以解决会议场景下领域个性化模型训练和部署等问题。

如图 4 所示,本系统分为训练数据扩充、数据预处理、数据检索、模型训练评估和模型部署五个基本模块。训练数据扩充针对特定会议,用户可预先收集会议相关术语词典、单语数据、平行句对数据等,以达到领域数据增强的目的。在数据检索模块中,通过术语词搜索方法从大规模语料库检索平行句对,通过模糊匹配或语义相似度等算法可进一步抽取平行句对,以及采用回译技术<sup>[17-18]</sup>构建伪造句对,从而构建更丰富的训练数据集。在模型训练阶段,以通用模型为基线,使用领域数据对模型进行自适应训练,最终得到优化模型并完成部署。

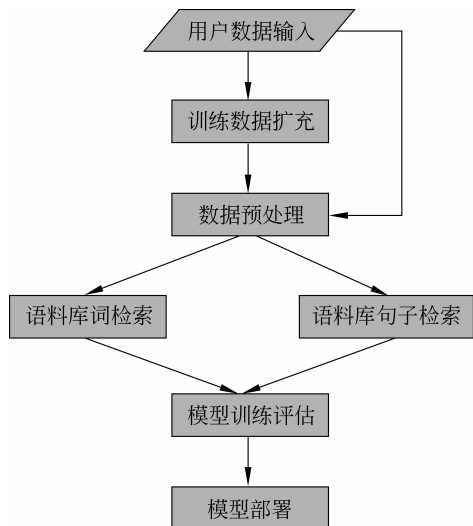


图 4 领域个性化自动训练系统架构图

## 3 实验

为验证本文所提方法的有效性,我们在中英方向的翻译任务上进行了实验,设定会议场景包含体育领域、商务领域和医学领域。我们分别从术语词

个性化定制和领域个性化定制两个方面验证方法的有效性,并在最后进一步验证了融合两个方法在翻译效果上的效果提升。

### 3.1 实验设置

基线系统采用 Transformer Big 模型结构<sup>[5]</sup>,并集成了拷贝网络<sup>[12-13]</sup>和强化学习策略,以便后续进行术语词个性化定制实验。模型中,编码器和解码器各 6 层,词向量维度为 1 024,多头注意力数量为 16,全连接层维度为 8 192,解码搜索束宽为 4。为解决词语的稀疏性问题,采用分词之后的字节对编码(Byte Pair Encoding, BPE)<sup>[19]</sup>子词作为建模单元,BPE 操作次数为 4 万次,中文词典大小为 5 万,英文词典大小为 4.1 万。训练语料共 3 000 万句对,其中 120 万句对来源于 Linguistic Data Consortium,2 300 万句对来源于 WMT18 国际翻译大赛,剩余 580 万句对来源于 CWMT18 机器翻译大赛等。

在术语词个性化定制实验中,我们从训练数据中抽取了约 200 万含有术语词的平行句对,构建拼接融合的训练数据,混入到原有训练语料中,共同用于模型训练。测试集合包括了体育领域会议场景下 3 725 句含有术语词的句对,包含了 8 583 对双语术语(含人名、地名等实体词);商务领域会议场景下 4 023 句含有术语词的句对,包含了 7 220 对双语术语;医学领域会议场景下 3 820 句含有术语词的句对,包含了 9 235 对双语术语。本文采用句子层面翻译客观评分 BLEU(Bilingual Evaluation Understudy)<sup>[20]</sup>和术语词典指定的专业术语词的翻译准确率作为评价指标。

在领域个性化实验中,我们使用自动化训练平台,针对体育领域搜集 28 万双语句对,抓取 200 万领域相关英语单语数据,并根据这些数据从原始训练集合中检索出约 500 万领域最相关的句对。针对商务领域搜集 25 万双语句对,抓取 300 万领域相关英语单语数据,并从原始训练集合中检索出约 500 万领域相关句对。针对医学领域搜集 25 万双语句对,抓取 800 万领域先关英语单语数据以及从原始训练集合中检索出 700 万领域相关句对。其中单语语料采用 Back Translation 算法进行回译得到伪平行语料,与其他领域数据一起共同构成了领域个性化训练集合。体育领域、商务领域和医学领域测试集与术语词个性化实验中相同,同时增加通用领域测试集合共 3 981 句。

### 3.2 术语词个性化实验

实验中,为了减少实体词识别错误对翻译性能带来的影响,我们采用了自己研发的基于 BiLSTM-

CRF<sup>[21]</sup>的NER工具,通过调参使得系统在体育、商务和医学三个测试集合上的平均正确率为94.21%,平均召回率为89.31%。基于该NER策略,术语词个性化实验如表1和表2所示。从表1可以看出,单独采用占位符替换方法,术语词翻译准确率较高,平均达到93.44%,而拼接融合方法由于需要靠模型自动学习指定术语译文的输出,导致部分术语词翻译准确率较低,平均为89.06%。术语词个性化方法在术语词准确率上与单独采用占位符替换方法相当,平均达到93.17%。从表2可以看出,拼接融合方法由于源句保留了术语词的原文和目标译文信息,BLEU分指标平均由基线的32.48提升至39.49,提升了7.01个BLEU,使得翻译结果更加流畅自然。占位符替换方法虽然术语词准确率最高,但由于句子整体流畅性差,存在译文直译等问题,其BLEU分平均为37.64,比拼接融合方法低了1.85个BLEU。术语词个性化方法在BLEU分指标上与拼接融合方法相当,达到了39.50,高于占位符替换方法1.86个BLEU。

综上,本文提出的术语词个性化方法,通过引入外部术语词典知识,融合了占位符替换和拼接融合方法的优点,能够在提升术语词翻译准确率的同时,兼顾翻译结果的整体流畅性。

表1 术语词个性化实验术语词准确率结果(%)

	体育	商务	医学	平均
基线系统	56.13	58.12	48.86	54.37
占位符替换 <sup>[11]</sup>	94.05	91.58	94.70	93.44
拼接融合 <sup>[8]</sup>	89.54	87.20	90.45	89.06
术语词个性化	93.57	91.40	94.55	93.17
相对基线系统提升	<b>37.44</b>	<b>33.28</b>	<b>45.69</b>	<b>38.80</b>

实验中也发现,由于NER存在一定的漏识别问题,影响了个别人名、地名等实体词的翻译。这主要是由于专业术语词中也存在个别词与词典指定翻译方式有些出入,因为这些术语词存在多种翻译方式,而翻译系统选择了出现频率较高的一种,整体上并不影响理解。

表2 术语词个性化实验BLEU分对比

	体育	商务	医学	平均
基线系统	33.57	35.32	28.54	32.48
占位符替换 <sup>[11]</sup>	39.10	39.85	33.97	37.64
拼接融合 <sup>[8]</sup>	40.82	41.20	36.45	39.49
术语词个性化	40.77	41.23	36.50	39.50
相对基线提升	<b>+7.20</b>	<b>+5.91</b>	<b>+7.96</b>	<b>+7.02</b>

3.3 领域个性化实验

在领域个性化实验中,我们分别将体育场景会议、商务场景会议和医学场景会议作为目标领域,使用自动化训练平台进行分类器训练、模型训练和预测等实验内容。以体育场景会议为例,首先从体育领域数据集和通用数据集中各随机挑选了15 000句中文作为分类器模型的训练数据集,从两个数据集剩余数据中随机挑选了4 000句中文作为开发集合,5 000句中文作为测试集合。商务场景会议和医学场景会议的相关实验配置相同。

表3给出了三个场景下通用与各领域分类效果。可以看出,在三个场景下,通用领域数据分类的F值均达到95%以上的较高水平,体育领域数据分类的F值为87.32%,商务领域数据分类的F值为87.14%,医学领域数据分类的F值为90.19%。在实际应用中,为了避免领域测试数据分类错误带来的通用性能损失,我们对分类阈值进行了调整,在牺牲领域数据召回率的情况下,使得其正确率处于较高水平。

表3 三个场景下领域分类效果

场景	测试集	正确率/%	召回率/%	F值/%
体育场景	通用集合	95.25	94.95	95.10
	体育集合	93.61	81.83	87.32
商务场景	通用集合	95.50	95.25	95.37
	商务集合	92.45	82.40	87.14
医学场景	通用集合	95.60	94.85	95.22
	医学集合	95.31	85.60	90.19

经过领域个性化模型自适应训练之后的实验结果,如表4所示。

表4 通用和各领域翻译效果

	通用	体育	商务	医学
基线方法	30.91	33.57	35.32	28.54
体育 FineTune	29.04	<b>37.75</b>	27.23	20.85
商务 FineTune	29.23	26.05	<b>38.88</b>	21.33
医学 FineTune	28.50	25.35	28.33	<b>33.15</b>
领域个性化	30.89	37.45	38.66	33.20
相对基线提升	<b>-0.02</b>	<b>+3.88</b>	<b>+3.34</b>	<b>+4.66</b>

可以看出,直接采用目前常用的FineTune方法<sup>[9]</sup>,使用领域内数据进行迭代训练和微调,虽然能够在各相应领域获得较明显的效果提升,但在非相

应领域会有较大的负面影响。例如,体育领域的 FineTune 模型虽然在体育集合上 BLEU 分由 33.57 提升至 37.75,提升了 4.18 个 BLEU,但是在通用集合、商务集合以及医学集合上都有不同幅度的下降,特别是在商务和医学集合上,由于领域差异性较大,下降幅度达到 7~8 个 BLEU。商务领域和医学领域 FineTune 模型结论一致。采用经过本文领域个性化训练之后,体育领域测试集合的翻译性能从基线系统的 33.57 提升到 37.45,提升了 3.88 个 BLEU;商务领域测试集合的翻译性能从 35.32 提升到 38.66,提升了 3.34 个 BLEU;医学领域测试集合的翻译性能从 28.54 提升到 33.20,提升了 4.66 个 BLEU。在通用领域测试集合上,个性化训练之后的性能下降了 0.02 个 BLEU,主要是因为一小部分通用数据错分成了体育领域数据,不过整体影响较小,可以忽略不计。该实验证明了本文所提方法在保持通用领域翻译性能基本相当时,能够大幅提升会议相关领域的翻译效果。

### 3.4 融合外部词典知识的领域个性化方法

最后,结合领域个性化自动训练系统,我们融合

术语词个性化定制方法和领域个性化定制方法,给出了最终基于融合外部词典知识的领域个性化方法在不同会议场景下的翻译性能提升结果。表 5 给出了体育场景会议、商务场景会议和医学场景会议下融合方法的实验结果。其中,体育领域测试集合的翻译性能提升至 42.89,相对于基线方法提升了 9.32 个 BLEU;商务领域测试集合的翻译性能提升至 42.98,提升了 7.66 个 BLEU;医学领域测试集合的翻译性能提升至 39.21,提升了 10.67 个 BLEU。

表 5 最终系统翻译效果对比

	通用领域	体育领域	商务领域	医学领域
基线方法	30.91	33.57	35.32	28.54
最终系统	30.89	42.89	42.98	39.21
提升幅度	-0.02	+9.32	+7.66	+10.67

以医学场景为例,我们在表 6 中给出不同系统和方法的翻译结果。从中可以看出,相比基线系统,融合了术语词个性化和领域个性化的最终系统翻译结果,在专业术语翻译准确率和行业用语表达上效果更好。

表 6 医学场景各系统译文举例对比

原文	将彩色血流成像与功率多普勒成像在灵敏度、速度混叠、光束角度依赖性、帧率和速度方向等方面进行比较。
参考答案	Compare the Color Flow Imaging versus Power Doppler Imaging in terms of sensitivity, velocity aliasing, beam angle dependency, frame rate and the direction of the velocity.
基线系统译文	Color flow and power imaging are compared insensitivity, velocity, beam angle dependence, frame rate and velocity.
最终系统译文	Color Flow Imaging was compared with Power Doppler Imaging insensitivity, velocity aliasing, beam angle dependency, frame rate and the direction of the velocity.

## 4 总结和展望

本文针对由于会议行业属性带来的机器翻译领域专业术语和行业用语的翻译问题,提出了一种融合词典知识的领域个性化方法,采用联合占位符和拼接融合的编码策略,通过引入外部词典知识,在提升实体词、专业术语词翻译准确率的同时,保持了译文的流畅性。同时,基于分类的领域旁支参数个性化自适应策略,在保持通用领域翻译效果的情况下,实现会议相关领域翻译质量的提升。基于上述方案,本文设计了一套领域个性化自动训练系统,在中英体育、商务和医学会议翻译任务上,基于会议领域相关数据,在不影响通用翻译的情况下平均提升

9.22 个 BLEU,获得较好应用成效。

不过会议场景下语音翻译除了行业属性带来了的术语词和行业用语翻译问题之外,还面临语音识别错误影响、演讲人口语化影响,以及翻译实时性等问题。未来我们将针对这些问题进行进一步研究,并探索不同的应用模式,提升会议场景下语音翻译的可用性,为不同语言使用者之间的信息沟通交流提供帮助。

## 参考文献

- [1] Takezawa T, Sumita E, Sugaya F, et al. Toward a broad-coverage bilingual corpus for speech translation of travel conversations in the real world[C]//Proceedings of LREC, 2002: 147-152.
- [2] Jan N, Cattoni R, Sebastian S, et al. The IWSLT

- 2018 evaluation campaign[C]//Proceedings of International Workshop on Spoken Language Translation, 2018: 2-6.
- [3] Sutskever I, Vinyals O, Le Q V. Sequence to sequence learning with neural networks[C]//Proceedings of Advances in Neural Information Processing Systems, 2014: 3104-3112.
- [4] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate[J]. arXiv preprint arXiv: 1409.0473, 2014.
- [5] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Proceedings of Advances in Neural Information Processing Systems. 2017: 5998-6008.
- [6] Hokamp C, Liu Q. Lexically constrained decoding for sequence generation using grid beam search[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, 2017,1: 1535-1546.
- [7] Feng Y, Zhang S, Zhang A, et al. Memory-augmented neural machine translation[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, 2017: 1390-1399.
- [8] Pham N Q, Niehues J, Waibel A. Towards one-shot learning for rare-word translation with external experts [C]//Proceedings of the 2nd Workshop on Neural Machine Translation and Generation, 2018: 100-109.
- [9] Luong M T, Manning C D. Stanford neural machine translation systems for spoken language domains[C]//Proceedings of the International Workshop on Spoken Language Translation, 2015: 76-79.
- [10] Britz D, Le Q, Pryzant R. Effective domain mixing for neural machine translation [C]//Proceedings of the 2nd Conference on Machine Translation, 2017: 118-126.
- [11] Crego J, Kim J, Klein G, et al. Systran's pure neural machine translation systems[J]. arXiv preprint arXiv: 1610.05540, 2016.
- [12] Gu J, Lu Z, Li H, et al. Incorporating copying mechanism in sequence-to-sequence learning[J]. arXiv preprint arXiv: 1603.06393, 2016.
- [13] Vinyals, Fortunato M, Jaitly N. Pointer networks [C]//Proceedings of Advances in Neural Information Processing Systems, 2015: 2692-2700.
- [14] Chu C, Wang R. A survey of domain adaptation for neural machine translation [C]//Proceedings of the 27th International Conference on Computational Linguistics 2018: 1304-1319.
- [15] Lecun Y, Bengio Y. Convolutional networks for images, speech, and time series[M]. The handbook of brain theory and neural networks. MIT Press, 1998.
- [16] Rupesh Kumar Srivastava, Klaus Greff, Jurgen Schmidhuber. Highway networks [C]//Proceedings of Presented at the Deep Learning Workshop, International Conference on Machine Learning, Lille, France, 2015.
- [17] Rico Sennrich, Barry Haddow, Alexandra Birch. Improving neural machine translation models with monolingual data [C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Berlin, Germany. Association for Computational Linguistics, 2016: 86-96.
- [18] Fadaee M, Monz C. Back-translation sampling by targeting difficult words in neural machine translation [C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Belgium, 2018: 435-446.
- [19] R Sennrich, B Haddow, A Birch. Neural machine translation of rare words with subword units [C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Berlin, Germany, 2016.
- [20] Papineni K, Roukos S, Ward T, et al. BLEU: A method for automatic evaluation of machine translation [C]// Proceedings of the 40th Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics, 2002: 311-318.
- [21] Guillaume Lample, Miguel Ballesteros, Sandeep Subramanian, et al. Neural architectures for named entity recognition [C]//Proceedings of NAACL, 2016.



刘庆峰(1973—),博士,教授,博士生导师,主要研究领域为信号处理,语音及语言信息处理。  
E-mail: nelslip@ustc.edu.cn



王亚楠(1993—),硕士,主要研究领域为自然语言处理、机器翻译。  
E-mail: ynwang8@iflytek.com



刘晨璇(1993—),通信作者,博士研究生,主要研究领域为自然语言处理、机器翻译。  
E-mail: cxliu3@iflytek.com