基于人物特征增强的拟人句要素抽取方法研究

李婧[1*]，王素格[1,2]，陈鑫[1]，王典[1]
[1]山西大学 计算机与信息技术学院，山西 太原 030006
[2]山西大学 计算机智能与中文信息处理教育部重点实验室，山西 太原 030006
{377143220}@qq.com, {wsg}@sxu.edu.cn
{1315614497}@qq.com，{547881490}@qq.com

摘要

在散文阅读理解的鉴赏类问题中，对拟人句赏析考查比较频繁。目前，已有的工作仅对拟人句中的本体要素进行识别并抽取，存在要素抽取不完整的问题，尤其是当句子中出现多个本体时，需要确定拟人词与各个本体的对应关系。为解决这些问题，本文提出了基于人物特征增强的拟人句要素抽取方法。该方法利用特定领域的特征，增强句子的向量表示，再利用条件随机场模型对拟人句中的本体和拟人词要素进行识别。在此基础上，利用自注意力机制对要素之间的关系进行检测，使用要素同步机制和关系同步机制进行信息交互，用于要素识别和关系检测的输入更新。在自建的拟人数据集上进行<本体，拟人词>抽取的比较实验，结果表明本文提出的模型性能优于其他比较模型。

关键词： 拟人句；向量表示；要素抽取；关系检测；同步机制

Research on Element Extraction of Personified Sentences Based on Enhanced Characters

Li Jing[1*], Wang Suge[1,2], Chen Xin[1], Wang Dian[1]
[1]School of Computer and Information Technology,Shanxi University,Taiyuan,Shanxi 030006,China
[2]Key Laboratory Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University,Taiyuan,Shanxi 030006,China
{377143220}@qq.com, {wsg}@sxu.edu.cn
{1315614497}@qq.com，{547881490}@qq.com

Abstract

In the appreciation questions of prose reading comprehension, the appreciation examination of anthropomorphic sentences is quite frequent. At present, the existing work only identification and extracts the ontological elements in the personification sentence, which has the problem of incomplete element extraction. Especially, when there are multiple ontologies in the sentence, the corresponding relationship between the personification word and each ontological element needs to be identified. In order to solve these problems, this paper proposes a method of identification and element extraction of personification sentences based on character feature enhancement. Using the characteristics of specific areas, the vector enhanced is represented for the sentence, the conditional random field model is used to identify the ontological and anthropomorphic elements in personification sentences. On this basis, the self-attention mechanism is used to detect the relationship between the elements, information exchange of between

国家自然科学基金(62076158)；山西省重点研发计划项目(201803D421024)

the element synchronization and relation synchronization mechanism is used to update input of element identification and relation detection. The experimental results show that the proposed method is better than other models for the extraction of <noumenon, personification word> on a self-built anthropomorphic dataset.

Keywords: Personification , Vector representation , Element extraction , Relation detection , Synchronization mechanism

1 引言

拟人作为最常见的修辞格之一，是将事物人格化，把原来不具有人动作和性格的事物比作和人一样的模样，在我们的日常交流和文学作品中有意识或无意识地使用，例如，童话故事里的动物、植物能讲话。拟人包括的三要素为本体、拟人词、拟体(赵琳玲, 2020)。本体：被描写和说明的事物，事物本身不是人，但是具有人的特点。拟人词：用来描绘人物特点的词语，如“夜空中的小星星在对你微笑。”，拟人词为“微笑”。拟体：与本体相对，就是人。由于拟人的修辞方式具有增强表达力，并生动刻画所描述对象的特点，常被用于散文的写作中，将物体、动物、植物、思想或抽象概念等物比拟为人，把事物人格化，使其具有人的动作、思想或情感。在近年的高考语文散文类鉴赏题中，多有涉及拟人句的考查。以2020年浙江省高考语文第10题为例。

原文：穿过小城，一片暮霭中，波塔波夫终于走到了房子跟前。小心翼翼地打开小门，可是小门还是咯吱地响了一声。花园仿佛抖动了一下。树枝上有雪花簌簌飘落，沙沙作响.....

问题：赏析文中画线部分的语言特点。

部分参考答案：语言具有诗化风格。如通过“花园仿佛抖动了一下”的拟人化描写，表现波塔波夫内心的情感波澜，情景交融，充满诗意。

根据上述的部分参考答案，如果能抽取拟人句中的本体和拟人词，不仅可以帮助解答鉴赏类问题，还可以进一步了解作者或主人公想表达的思想感情。

本文基于多任务学习，提出基于人物特征增强的拟人句要素抽取方法。该方法主要包含三个部分：表示增强、要素抽取及关系检测。具体地，在表示增强部分中，将人物特征词融入句子的表示中；要素抽取部分是利用条件随机场，确定标签之间的前后依赖关系；关系检测部分是使用自注意力机制，建模字间的关系。为了实现拟人句的要素抽取和关系检测部分间的信息交互，使用要素同步机制和关系同步机制。在创建的拟人数据集中进行<本体，拟人词>抽取的实验，结果表明本文提出的模型性能优于其他比较模型。

2 相关工作

对于要素抽取，研究者们利用多任务学习方法，通过在相关任务间共享表示信息，提升模型在原始任务上的泛化性能。由于CRF(Lafferty et al., 2001a)可以有效学习输出标签之间的前后依赖关系，近些年在自然语言处理领域中得到了广泛使用。(Huang et al., 2015)提出了一系列基于长短期记忆（LSTM）的序列标注模型，并首次将BiLSTM-CRF模型应用于NLP基准序列标记数据集，证明了此模型可以有效地利用过去和未来的输入特征，对于CRF层，使用句子级的标记信息，使方法具有较强的鲁棒性，而且对嵌入词的依赖性也小。但有关拟人句要素抽取的相关研究目前较少，赵琳玲(赵琳玲, 2020)通过对拟人修辞手法的分析，发现拟人句中包含显著的人物特征，因而，提出了基于人物特征的拟人句判别及要素抽取方法，但仅对拟人句中的本体进行了抽取，并没有对拟人词进行抽取且未判断二者存在的二元关系。

对于实体关系抽取，已有很多的研究工作。早期方法(Zelenko et al., 2003)、(Chan and Dan, 2013)将实体抽取和关系抽取视为两个独立的子任务，在抽取所有实体后，采用管道方法进行关系分类。为了在两个子任务之间建立桥梁，构建提取实体和关系的联合模型已经引起了研究者的广泛关注。Tagging方法通常使用标记策略构建实体和关系之间的连接。在这些方法中，NovelTagging(Zheng et al., 2017)首先将实体类型和关系角色作为标签的不同部分，将联合抽取任务建模为单个序列标注问题。但是，它不能处理重叠的情况。作为改进，(Takanobu et al., 2018)、(Dai et al., 2019)、(Yu et al., 2020)执行多轮标记过程。Seq2Seq方法尝试按顺序直接生成所有的三元组。CopyRE(Zeng et al., 2018)通过两个具有复制机制的对应实体生成关

系，但只能生成实体的最后一个字。因此，CopyMTL(Zeng et al., 2020)应用多任务学习框架提取多字实体。

由于拟人句中的本体和拟人词之间存在一定的隐式语义关系，若直接使用实体关系抽取方法不能将本体和拟人词进行准确地抽取。例如，“月亮那么明媚又充满哀伤。”一句中的本体是月亮，拟人词是哀伤。为了解决此问题，本文基于多任务学习，提出了一种基于人物特征增强的拟人句要素抽取方法模型。

3 拟人特征词库构建

由于目前没有开放的拟人句数据集，因而我们人工构建数据资源，通过收集筛选和标注处理，构建了4283条拟人句的数据集。数据来源于高中语文课文、查字典网、散文吧网站以及全国部分省市的高考语文真题，具有一定的代表性。

对于一个拟人句，拟人词是用来描绘人物特征的词语，将人物特征细分为人物的情感、动作、神态、性格、外貌和其他特征六类，通过对拟人数据进行人物特征统计，统计结果和人物特征示例如表1所示。

人物特征	人物特征词汇示例	条数（条）	占比（%）
情感	恋、陶醉、敬、失望、无奈、气愤、同情	892	20.83
动作	笑、哭、唱、说、对话、打哈欠、伸懒腰	3468	80.97
神态	炯炯有神、焦急、慌张、神采奕奕、眉开眼笑	235	5.49
性格	大度、宽容、旷达、洒脱、善良、调皮、乐观	486	11.35
外貌	眉清目秀、容光焕发、美如冠玉、出水芙蓉	400	9.34
其他特征	手指、长发、手臂、裙摆、啤酒肚、毅力、脉搏	544	12.70

表 1: 拟人数据统计结果和人物特征示例

从表1中可以看出，将人物特征归纳为六个方面，从不同的角度对人物的特点进行描述。同时，对拟人句进行分析发现，存在一个拟人句包含多种人物特征的情况，例如：“梨子穿上了金黄色的蓬蓬裙，兴致勃勃地去参与舞会。”，在此拟人句中，“穿”属于人物动作，“蓬蓬裙”属于其他特征，“兴致勃勃”属于人物情感，从多方面对梨子进行了人物特征描写。根据对拟人句的人物特征的统计结果，发现80.97%的拟人句中包含人物动作，其次是人物情感、其他特征、人物性格。因而，体现了人物特征在拟人句中的重要性。

在已构建的拟人数据集上，总结出较为常见的人物特征词汇1586个，利用哈工大的《同义词词林扩展版》和WordNet进行同义词查找，对特征词汇进一步扩充，使词库尽可能多的包含相关词汇，最终构建有2480个词汇或短语的人物特征词库即为DF，其中，人物特征词库包含表1中提到的六种人物特征，同时词汇带有褒、贬不同含义，覆盖面广较为全面，几乎涵盖了文学作品中常用到的人物特征，对于更准确的进行拟人句要素抽取，具有一定的辅助作用。

4 拟人句要素抽取方法

在拟人句中，本体和拟人词之间存在一定的隐式语义关系，这两个要素可以同时存在，但两者之间不一定存在二元关系。例如“黄昏时的村庄是那样的安逸，在晚风的抚摸下，与落日不舍地道别。”，在该句子中，存在两个本体为村庄和晚风，三个拟人词为安逸、抚摸、不舍地道别，若按照一般的要素抽取方法仅将本体和拟人词抽取，并未找到两个本体分别对应的拟人词，因此，为了解决这个问题，本文提出基于人物特征增强的拟人句要素抽取方法。在要素抽取时将其看作序列标注问题，采用BIO标注方法产生五种标记，其中B-T和I-T分别表示本体的首部和中部，B-P和I-P分别表示拟人词的首部和中部，O没有任何含义。同时，通过建模字间的关系最终推理出<本体，拟人词>，完成拟人句要素抽取。

<本体，拟人词>抽取任务的目标是，从给定句子S中获得本体与拟人词构成的集合$C = \{\langle a_i, o_i \rangle\}$，其中，$a_i$和$o_i$分别表示本体和拟人词，它们可以是一个词或短语。基于人物特征增强的拟人句要素抽取方法模型的总体框架如图1所示。

该模型框架中表示增强部分，将人物特征词作为特定领域的特征引入编码层，与Bert得到的上下文表示向量进行结合，得到句子的增强表示的特征。要素抽取部分和关系检测部分用于提取本体、拟人词以及判断二者存在的二元关系。此外，还使用了一个同步单元实现要素抽取

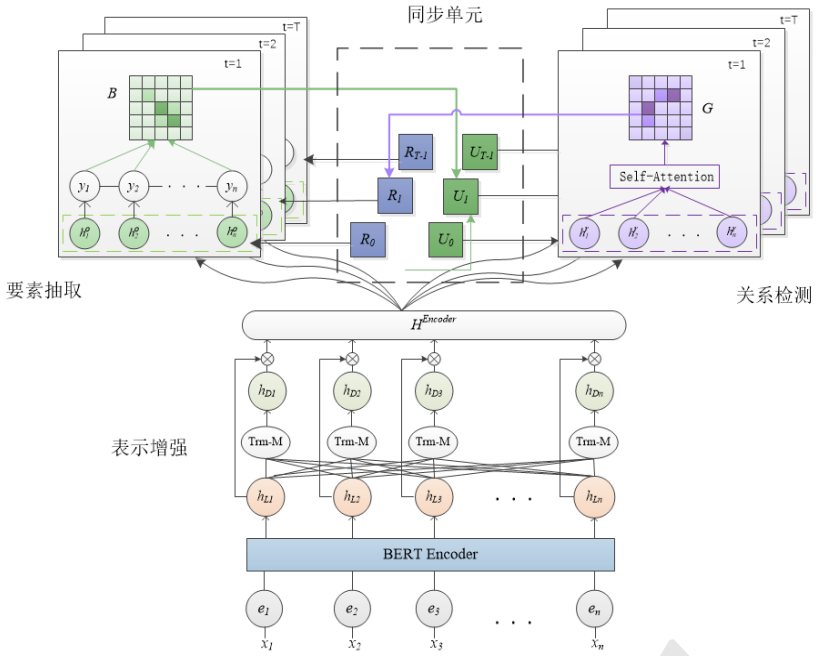

图 1: 基于人物特征增强的拟人句要素抽取方法模型

部分和关系检测部分之间的信息交互。整体模型需要多个递归过程，最后采用一个推理层捕获<本体，拟人词>。

4.1 表示增强部分

表示增强部分是将人物特征增强后的编码层。由于预训练模型的编码倾向于捕获一般文本表示，但缺乏领域知识。为了弥补相关领域信息的不足，在编码层中加入了人物特征进行增强。

对句子S的开头和结尾分别添加标记[CLS]和[SEP]，得到输入序列$X = \{x_1, x_2, ..., x_N\}$，每个句子有$N$个字。对于每个字$x_i$，得到初始嵌入$e_i$，则嵌入序列$E = \{e_1, e_2, ..., e_N\}$被送至由带有多个自注意头的堆叠变压器块组成的BERT中(Vaswani et al., 2017)。将最后一层的隐藏状态作为输入句子中每个标记的一般表示H^L，记作$H^L = \{h_1^L, h_2^L, ..., h_N^L\}$。

输入序列与已构建好的人物特征词库DF进行检索，找到所有可能构成人物特征的子序列。将$X[i:j]$定义为X的子序列，X以x_i开始，以x_j结束，再利用掩模矩阵M_D表示人物特征。其中第i行和第j列的元素m_{ij}表示子序列$X[i:j]$是否为人物特征的表达式。

$$m_{ij} = \begin{cases} 1 & X[i:j] \in DF \\ 0 & other \end{cases} \tag{1}$$

利用额外的Transformer编码器计算输入句子的人物特征的特定表示。该层包括两个子层，一个多头自注意力机制和一个前馈网络，每个子层后面都有一个残差连接和层规范化。融合了人物特征信息的特征掩蔽编码器的最终输出表示为H_D。最后，将H_L和H_D进行加权平均，得到人物特征增强表示$H^{Encoder}$。

$$H^{Encoder} = \gamma H^L + (1-\gamma) H^D \tag{2}$$

其中，γ为加权参数。在这项工作中，采用了$\gamma = 0.5$。

4.2 要素抽取部分

要素抽取部分作为模型的一部分，目的是提取拟人句中的本体和拟人词。将CRF(Lafferty et al., 2001b)耦合在编码层上，作为要素抽取部分。对于在第t个循环步骤的预测标签序列$Y^t = \{y_1^t, y_2^t, ..., y_N^t\}$，定义它的得分如下:

$$S(X,Y^t)=\sum_{i=1}^{N}Q_{y_{i-1}^t,y_i^t}+\sum_{i=1}^{N}P_{i,y_i^t}^t \tag{3}$$

$$P^t=H_t^oW_p+b_p \tag{4}$$

其中，H_t^o表示该部分第t个递归步骤的输入序列，由带有人物特征的表示序列$H^{Encoder}$和关系同步语义R_{t-1}计算得到。P为状态得分矩阵，Q为转移得分矩阵。W_p，b_p为模型参数。预测序列Y^t的概率可以计算如下:

$$P(Y^t|X)=\frac{exp(S(X,Y^t))}{\sum_{\widetilde{Y^t}\in Y_X^t}exp(S(X,\widetilde{Y^t}))} \tag{5}$$

其中，Y_X^t表示所有可能的标签序列。在解码过程中，采用Viterbi算法寻找得分最大的标签序列。

4.3 关系检测部分

由于本体和拟人词之间的二元关系结构可以是一对一也可以是一对多，甚至是多对多。因此，考虑到本体和拟人词之间关系的复杂性，采用自注意力作为关系检测部分，可以根据句子的上下文信息动态的建模字间关系，而不受时序限制。

在第t个递归步骤中，首先计算注意力矩阵G^t，它的元素$g_{i,j}^t$表示第i个字与第j个字的关联度如下:

$$g_{i,j}^t=\frac{exp(\gamma(h_{t,i}^r,h_{t,j}^r)}{\sum_{k=1}^{N}exp(\gamma(h_{t,i}^r,h_{t,k}^r)} \tag{6}$$

$$\gamma(h_{t,i}^r,h_{t,j}^r)=tanh(h_{t,i}^rW_r^1+h_{t,j}^rW_r^2)W_r^3 \tag{7}$$

其中，γ是一个分数函数，$h_{t,i}^r$为关系检测部分的第i个标记的输入表示，由$H^{Encoder}$和实体同步语义义U_{t-1}计算。W_r^1，W_r^2，W_r^3是模型参数。

在最后一步t中，通过最大化似然概率，进一步将监督信息引入到G^t的计算中，如下所示:

$$p(Z|X)=\prod_{i=1}^{N}\prod_{j=1}^{N}p(z_{i,j}|x_i,x_j) \tag{8}$$

其中，标准关系矩阵Z由元素$z_{i,j}$组成，关系概率$p(z_{i,j}|x_i,x_j)$可计算如下:

$$p(z_{i,j}|x_i,x_j)=\begin{cases}g_{i,j}^t & z_{i,j}=1\\ 1-g_{i,j}^t & z_{i,j}=0\end{cases} \tag{9}$$

其中，$z_{i,j}=1$表示第i个字与第j个字之间存在关系，反之亦然。有了这些监督信息，可以引导注意力更有效地捕捉字间的关联。

4.4 同步单元

由于要素抽取和关系检测两个部分是相互依赖的，为此，受到(Chen et al., 2020)的启发，使用了要素同步机制（ESM）和关系同步机制（RSM），通过高层信息的交互更新隐藏的表示序列H_t^o和H_t^r。

（1）要素同步机制

利用ESM可以捕获每个字对应的语义，并将这些语义集成到表示序列H_{t+1}^r中。根据要素抽取部分得到的预测标签序列Y^t及其概率，可计算第i个字在第t个递归步骤的每个要素语义$u_{i,j}$，最后得到关系检测部分的输入表示$h_{t+1,i}^r$:

$$u_{t,i}=\sum_{j=1}^{N}\varphi(B_{i,j}^t)h_j^s \tag{10}$$

$$\varphi(B_{i,j}^t) = \frac{B_{i,j}^t}{\sum_{k=1}^{N} B_{i,k}^t} \tag{11}$$

$$h_{t+1,i}^r = \sigma(u_{t,i}W_r^4 + h_i^s W_r^5) \tag{12}$$

其中，$B_{i,j}^t$为第i个字与第j个字属于同一要素时，第j个字的标签概率;否则$B_{i,j}^t$为0。$\varphi(\cdot)$是一个归一化函数。W_r^4和W_r^5是模型参数，σ是激活函数。使用零矩阵初始化要素语义序列$U_0 = \{u_{0,1}, u_{0,2}, ..., u_{0,N}\}$。

（2）关系同步机制

拟人句中两个要素之间的二元关系可以为要素抽取提供线索，因此对关系语义进行编码就显得尤为重要。因此，使用RSM捕获反映关系的语义，并更新隐藏的表示序列H_{t+1}^o。在第t个递归步骤，从关系检测部分计算出第i个关联度为$g_{i,j}^t$的字的关系语义$r_{i,j}$，最后得到要素抽取部分的输入表示$h_{t+1,i}^o$:

$$r_{t,i} = \sum_{j=1}^{N} \varphi(\phi(g_{i,j}^t))h_j^s \tag{13}$$

$$\phi(g_{i,j}^t) = \begin{cases} g_{i,j}^t & g_{i,j}^t \geq \beta \\ 0 & g_{i,j}^t < \beta \end{cases} \tag{14}$$

$$h_{t+1,i}^o = \sigma(r_{t,i}W_o^1 + h_i^s W_o^2) \tag{15}$$

其中，$\varphi(\cdot)$是与(11)式相同的归一化函数。为了避免噪声，用$\varphi(\cdot)$过滤低于给定阈值β的分数。W_o^1和W_o^2是模型参数。与ESM类似，初始关系语义序列$R_0 = \{r_{0,1}, r_{0,2}, ..., r_{0,N}\}$设为零矩阵。

4.5 联合学习

为了同步学习要素抽取部分和关系检测部分，将各自的损失函数进行融合。对于要素抽取部分，给定标准标签序列Y，最后一步最小化负对数似然损失函数如下:

$$L_E = log \sum_{\widetilde{Y} \in Y_X^T} exp(S(X, \widetilde{Y})) - S(X, Y) \tag{16}$$

对于关系检测部分，将标准注释转换为一个one-hot矩阵，其中0表示没有关系，1表示两个字间存在二元关系。最小化最后一步预测分布与标准分布之间的交叉熵损失:

$$L_R = -\sum_{i=1}^{N}\sum_{j=1}^{N} p(z_{i,j}|x_i, x_j) log[\hat{p}(z_{i,j}|x_i, x_j)] \tag{17}$$

将这两部分结合起来，构建整个模型的损失目标:

$$L(\theta) = L_E + L_R \tag{18}$$

4.6 推理层

由于本模型主要处理的是要素抽取和关系检测，因此引入推理层，根据两个部分的结果生成<本体，拟人词>。利用要素抽取部分的预测标签序列Y^t，得到本体集$A = \{a_1, a_2, ..., a_{l_A}\}$和拟人词集$O = \{o_1, o_2, ..., o_{l_O}\}$。其次，根据关系检测部分的权重矩阵$G^t$计算本体和拟人词之间的关系。例如，给定一个本体$a = \{x_{i_S^a}, ..., x_{i_E^a}\}$和一个拟人词$o = \{x_{i_S^o}, ..., x_{i_E^o}\}$，两者之间的关联度$\delta$可计算如下:

$$\delta = \frac{1}{2}(\frac{1}{|a|}\sum_{k=i_S^a}^{i_E^a}\sum_{l=i_S^o}^{i_E^o} g_{k,l} + \frac{1}{|o|}\sum_{l=i_S^o}^{i_E^o}\sum_{k=i_S^a}^{i_E^a} g_{l,k}) \tag{19}$$

其中，$|a|$和$|o|$分别表示本体和拟人词的长度。只有当δ大于给定的阈值$\hat{\delta}$时，才能提取$\langle a, o\rangle$。

5 实验

5.1 参数设置与评价指标

本文采用精确率P、召回率R和F1值作为评价指标。

$BERT_{base}$模型作为模型编码层的一部分，其中，嵌入和上下文表示的维数均为768。为了增强要素抽取部分和关系检测部分之间的信息交互，对递归步骤进行调参。训练时采用AdamW优化器。用于微调BERT的学习率设为2e-5，并将batchSize设为10，dropout为0.5，d_o、d_r均为250，同步关系机制中的阈值β为0.1。其他参数如递归步骤t、训练模型中的学习率lr、$H^{Encoder}$中的γ以及推理层的阈值$\hat{\delta}$，可通过对超参数的实验选定，其评价指标采用<本体，拟人词>的F1值。如图2所示，(a)、(b)、(c)、(d)分别为t、lr、γ、$\hat{\delta}$对F1值的影响。

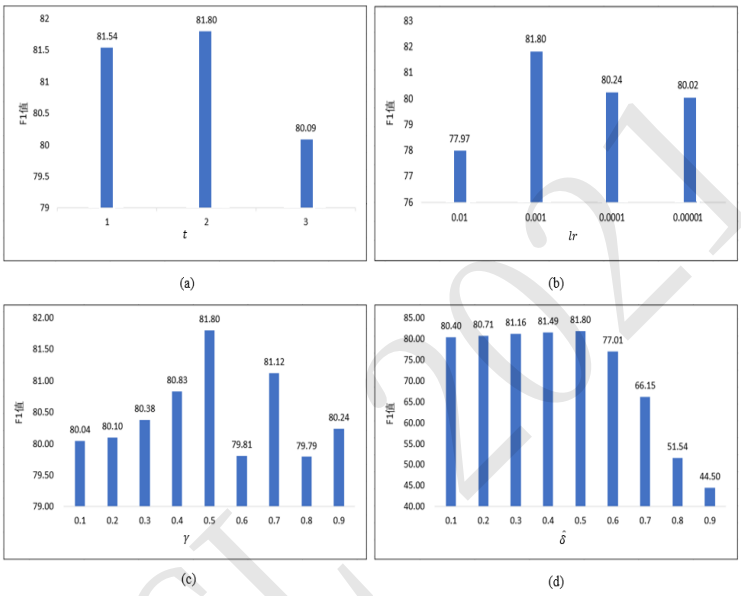

图 2: 部分超参数对实验性能的影响

通过图2对超参数的调试选择，可以发现递归步骤t为2，训练模型中的学习率lr为0.001，γ为0.5，推理层的阈值$\hat{\delta}$为0.5时实验结果达到最优。由于本次任务的目标是得到<本体，拟人词>，因此对推理层的阈值$\hat{\delta}$的设定尤为重要。从图2中可以看出，$\hat{\delta}$越小，可以取到的值就越多，涵盖多余不需要的结果；$\hat{\delta}$越大，则过滤掉许多有用的预测值，大大降低了模型的精度。而其他超参数不同的取值对模型的整体性能没有显著影响，这表明模型具有鲁棒性，对这些超参数的敏感度较小。

5.2 对比方法介绍

为了验证本文提出方法的有效性，将其与如下基线方法进行对比实验。

BERT+CH(赵琳玲, 2020)：该模型采用BiLSTM-CRF的方法抽取拟人句中的本体。编码层分为两部分，一是使用BERT得到上下文向量表示，二是微调BERT，segmentid用是否为人物特征（1或0）表示，拼接二者。但此模型没有对拟人词以及要素存在的二元关系进行进一步研究。

W+F(赵琳玲et al., 2021):该模型的Embedding层为每个词的向量和词性特征的拼接。此模型是对比喻句中的要素进行识别和抽取，现放入拟人数据。

SDRN(Chen et al., 2020)：该模型研究的是方面意见对抽取（AOPE）任务，目的是成对地提取方面和意见表达。

BERT+CH+SDRN(B+C+S):将上述BERT+CH和SDRN方法进行结合，在SDRN的编码层中微调BERT，segmentid用是否为人物特征（1或0）表示。

SDRN+SMHSA(Liu et al., 2020)(S+S)：该模型将SDRN模型中关系检测部分换为SMHSA模型中的多头自注意的方法。SMHSA的主要任务是联合实体和关系抽取，得到关系三元组。

5.3 实验结果与分析

利用第3节提出的模型，在已构建的拟人数据中进行对比实验，结果如表2所示。

模型	本体			拟人词			<本体，拟人词>		
	P	R	F1	P	R	F1	P	R	F1
BERT+CH	88.58	86.22	87.39	-	-	-	-	-	-
W+F	91.44	92.89	92.16	75.42	82.92	76.69	-	-	-
SDRN	90.57	92.78	91.66	81.42	85.25	83.29	80.23	80.63	80.43
B+C+S	88.54	87.71	88.12	81.87	85.33	83.65	78.58	70.27	74.36
S+S	91.39	92.85	91.97	82.09	**86.62**	84.10	71.21	61.38	65.77
Ours	**92.24**	**93.31**	**92.77**	**82.94**	85.87	**84.38**	**81.98**	**81.73**	**81.80**

表 2: 五种方法的对比实验结果

由表2实验结果可以看出:

(1)本文提出的模型在与其他模型进行比较，在<本体，拟人词>抽取任务的F1值达到了目前最优，验证了本文使用联合学习方法对<本体，拟人词>的抽取是有效的。

(2)由于本文的模型是对SDRN模型进行的改进，因此，本文所提出的方法与SDRN的结果比较。在<本体，拟人词>抽取的任务上，本文提出的模型比SDRN，在P值、R值、F1值分别提高了1.75，1.10，1.37个百分点，验证了在编码层中加入人物特征进行增强，弥补了预训练模型在编码时对相关领域信息获取不足的问题。

(3)由于之前的工作并没有对<本体，拟人词>抽取进行研究，而SDRN在很大程度上解决了判断两者间存在二元关系的问题，这说明自注意力机制有助于学习句子内部要素间相关联的依赖关系。BERT+CH+SDRN模型将BERT编码中的sigmentid进行修改，改变了上下文的语义。SDRN+SMHSA模型的要素抽取部分使用的是SDRN实体识别部分，而关系检测部分则采用到SMHSA模型中抽取实体关系任务的方法，导致实验结果不理想，其原因是在拟人句中本体和拟人词的关系不同于实体间的关系，因此，利用该方法存在关系无法判别的问题。而我们的模型使用了自注意力机制。

值得说明的是，本文使用联合学习模型的参数是在训练时仅考虑了<本体，拟人词>抽取的关系F1值达到最高，因此，仅仅抽取本体或拟人词的性能指标不是最佳。

5.4 消融实验

为了验证模型各个部分的性能，将模型中去掉部分信息进行消融实验。

-feature：表示将人物特征融合部分去掉后的模型。

-ESM：将模型中的要素同步机制（ESM）去掉，只保留全连接层更新关系隐藏表示。

-RSM：将模型中的关系同步机制（RSM）去掉，并采用全连接层更新拟人词隐藏表示。

-ESM+RSM：将模型中的要素同步机制（ESM）和关系同步机制（RSM）均去掉。

上述四种方法与本文的模型在拟人数据中的比较结果如表3所示。

由表3实验结果可以看出:

(1)-feature、-ESM和-RSM在<本体，拟人词>抽取任务的评价指标F1均有所下降。其中，-feature与本文模型的性能相比下降明显，说明具有人物特征增强的编码层对<本体，拟人词>抽取任务是有效的，在一定程度上弥补了一般编码层对相关领域信息的不足的问题。

(2)-ESM+RSM是所有方法中最差的，说明使用ESM或RSM，对模型的整体都是有帮助的，且二个同时使用的性能优于只使用一个。特别是ESM的贡献略大于RSM。另外在这种同步机制的作用下，我们的模型优于其他基线方法。

	<本体，拟人词>		
	P	R	F1
-feature	80.23	80.63	80.43
-ESM	80.97	79.76	80.36
-RSM	80.28	81.38	80.83
-ESM+RSM	78.39	81.75	80.04
Ours	**81.98**	**81.63**	**81.80**

表 3: <本体，拟人词>抽取消融实验对比结果

6 总结

针对拟人句的本体和拟人词抽取问题，本文提出了基于人物特征增强的拟人句要素抽取方法，首先通过表示增强部分将人物特征词作为特定领域的特征引入编码层，与BERT得到的上下文表示向量进行结合，得到能够增强表示的特征。其次，使用要素抽取部分和关系检测部分，同时提取本体、拟人词和二者存在的二元关系。此外，还用同步单元实现后两个部分之间的信息交互。经过多个递归过程后，最后采用推理层捕获<本体，拟人词>。并与其他模型进行对比实验，实验表明，人物特征增强和多任务学习的共同采用提高了本文所提出方法的性能。

参考文献

Y. S. Chan and R. Dan. 2013. Acl'11 exploiting syntactico-semantic structures for relation extraction. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*.

Shaowei Chen, Jie Liu, Yu Wang, Wenzheng Zhang, and Ziming Chi. 2020. Synchronous double-channel recurrent network for aspect-opinion pair extraction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*.

D Dai, X Xiao, Y. Lyu, S. Dou, and H. Wang. 2019. Joint extraction of entities and overlapping relations using position-attentive sequence labeling. In *Proceedings of the 33rd Association for the Advancement of Artificial Intelligence Conference on Artificial Intelligence*.

Z. Huang, X. Wei, and Y. Kai. 2015. Bidirectional lstm-crf models for sequence tagging. *Computer Science*.

J. Lafferty, A. Mccallum, and Fcn Pereira. 2001a. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the 18th International Conference on Machine Learning*.

J. Lafferty, A. Mccallum, and Fcn Pereira. 2001b. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the 18th International Conference on Machine Learning*.

J. Liu, S. Chen, B. Wang, J. Zhang, and T. Xu. 2020. Attention as relation: Learning supervised multi-head self-attention for relation extraction. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence and the 17th Pacific Rim International Conference on Artificial Intelligence*.

R. Takanobu, T. Zhang, J. Liu, and M. Huang. 2018. A hierarchical framework for relation extraction with reinforcement learning. In *Proceedings of the 32nd Association for the Advancement of Artificial Intelligence Conference on Artificial Intelligence*.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *CoRR*, abs/1706.03762.

B. Yu, Z. Zhang, X. Shu, Y. Wang, T. Liu, B. Wang, and S. Li. 2020. Joint extraction of entities and relations based on a novel decomposition strategy. In *Proceedings of the 24th European Conference on Artificial Intelligence*.

D. Zelenko, C. Aone, and A. Richardella. 2003. Kernel methods for relation extraction. *Journal of Machine Learning Research*, 3(3):1083–1106.

X. Zeng, D. Zeng, S. He, L. Kang, and J. Zhao. 2018. Extracting relational facts by an end-to-end neural model with copy mechanism. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*.

D. Zeng, H. Zhang, and Q. Liu. 2020. Copymtl: Copy mechanism for joint extraction of entities and relations with multi-task learning. In *Proceedings of the 34th Association for the Advancement of Artificial Intelligence Conference on Artificial Intelligence*.

S. Zheng, F Wang, H. Bao, Y. Hao, P. Zhou, and B. Xu. 2017. Joint extraction of entities and relations based on a novel tagging scheme. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*.

赵琳玲, 王素格, 陈鑫, 王典, and 张兆滨. 2021. 基于词性特征的明喻识别及要素抽取方法. 中文信息学报, 35(1):81.

赵琳玲. 2020. 面向高考鉴赏类问题的明喻与拟人识别方法研究. 硕士论文, 山西大学.