

## 基于深度学习的专利分类方法

马建红,王瑞杨,姚 爽,刘双耀

(河北工业大学 计算机科学与软件学院,天津 300401)

**摘 要:** 现有的效应概念图匹配方法多数存在匹配容错性差的问题。为此,从大数据的角度提出一种新的挖掘专利与效应对应关系的方法。利用长短期记忆网络(LSTM)与基于 attention 的双向 LSTM 相结合形成模型训练专利语料,通过 Softmax 分类模型进行分类,得到专利所属的效应。实验结果表明,该方法利用 Bi-LSTM-ATT 模型进行训练对判定专利所属效应具有一定的可用性,准确率可以达到 70% 以上。

**关键词:** 深度学习;attention 机制;Bi-LSTM-ATT 模型;专利分类;产品创新

**中文引用格式:**马建红,王瑞杨,姚 爽,等.基于深度学习的专利分类方法[J].计算机工程,2018,44(10):209-214.

**英文引用格式:**MA Jianhong,WANG Ruiyang,YAO Shuang,et al. Patent classification method based on depth learning[J]. Computer Engineering,2018,44(10):209-214.

## Patent Classification Method Based on Depth Learning

MA Jianhong,WANG Ruiyang,YAO Shuang,LIU Shuangyao

(School of Computer and Engineering,Hebei University of Technology,Tianjin 300401,China)

**[Abstract]** The existing concept graph matching methods have the problem of poor matching fault tolerance. Therefore, this paper proposes a new method to explore the relationship between patent and effect. Long Short-term Memory (LSTM) is used to combine with bidirectional LSTM which based on the attention mechanism to train the patent corpus. The Softmax classification model is used to classify the patent and the effect of the patent is obtained. Experimental results show that the method uses the Bi-LSTM-ATT model to train has certain validity in determining the effect of patents, and the accuracy can reach more than 70% .

**[Key words]** depth learning;attention mechanism;Bi-LSTM-ATT model;patent classification;product innovation

**DOI:**10.19678/j.issn.1000-3428.0048159

### 0 概述

产品创新即是功能的创新,所有功能都是基于效应而存在的。在创新理论中,效应是指完成某种功能的科学原理或现象。专利中往往蕴含着功能创新和效应的核心知识,因而可以从专利中挖掘相关效应知识。文献[1]提出类比设计是扩充传统概念生成方法的有力手段,它通过解决方案与类似问题的相似性关系从而扩大一系列新的想法。根据文献[1]的设计类比思想,在不同的行业领域里检索专利,借鉴专利中的原理和方法,就能打破思维惯性,为产品设计者提供类似的专利进行参考。这种设计类比不仅能够有效地进行专利规避和专利布局,还能提高专利等级和产品竞争力,使产品创新更加快捷。为达到以上效果,需要功能-效应-专利这样的检索方式来实现,即每个功能对应着不同的效应,每个效应对应着多个专利。在解决问题时,通过检索到

的效应就能找到应用了此效应的一组专利,进而帮助设计者确定正确的效应,将效应概念解转化为具体的技术方案。在以上检索方式中,研究挖掘专利与效应的对应关系或把专利按照效应分类是关键。

目前,学者们也提出了思路来寻找专利中隐含的效应知识。文献[2]提出的思路是将专利、效应分别从功能、行为、属性、结构方面的特征进行表达,并对这些特征进行归类用概念图的方式表示,再对专利概念图和效应概念图进行匹配得到专利的效应,这里存在概念图匹配容错性差的问题。文献[3]提出从国际专利分类表(IPC)中提取设计知识的思想,定义功能关联度、效应互换度、效应通用度与作用结构知识兼容度等定量参数,建立知识之间的关联关系。这种方法只能从中检索到抽象的知识,没有达到专利的层次,不能满足设计者要求的检索,设计者更想要的是具体形象的知识。在上述背景下,为解决现有方法中需要修订以及实用性差的问题,本文

**作者简介:**马建红(1965—),女,教授、博士,主研方向为计算机辅助创新设计、TRIZ 软件工程;王瑞杨,硕士研究生;姚 爽,助理研究员、硕士;刘双耀,硕士研究生。

**收稿日期:**2017-07-29 **修回日期:**2017-10-12 **E-mail:**m\_zh2002@126.com

提出一种基于深度学习的专利分类方法。该方法采用以效应作为分类标准,对专利进行文本分类以更具有实用性。

## 1 相关研究

近年来,在自然语言处理领域用深度学习模型解决文本分类问题受到人们的关注。文献[4]用 s-lstm 的记忆模块代替递归模型中的组合层,通过语义组合来理解一段文本。文献[5]基于句法树的长短时记忆神经网络(Tree-LSTM),将标准 LSTM 的时序结构改为语法树结构,在文本分类上得到非常好的提升。文献[6]在 LSTM 的基础上引入 attention 以此来捕获不同上下文信息对给定情感的重要性。文献[7]提出一种适用于前馈神经网络的简化的注意力模型,证明了 attention 机制能够在文本较长的情况下,有效解决信息丢失等长期依赖问题。文献[8]提出的层次化注意网络有 2 个层次的注意机制在单词和句子层次上应用,使它能够在构造文档表示时关注更重要的内容。充分说明了 attention 机制能够给文本中的关键部分分配更多的注意力,突出其重要性。

构建树结构的方式一般是用于短文本,不太适用于专利长文本。因此,在双向 LSTM 的基础上增加 attention 层,将基于 attention 的双向 LSTM 模型与单纯的 LSTM 进行组合,既考虑了上下文信息,又能计算注意力分配概率,有效防止信息的丢失,突出关键词的作用。本文在自然语言处理的基础上,把专利文本内容和功能部分分别进行预处理、文本表示,然后把专利文本内容作为基于 attention 的双向 LSTM 模型的输入,学习到专利文本特征。把文本表示后的功能部分作为 LSTM 模型的输入,训练得到功能特征。将文本特征与功能特征通过 pooling 层连接作为专利效应分类的总体特征,最后用这些特征去构建或调用 softmax 分类模型,达到预测分类的目的。最终的输出即是专利所属的效应类别。

## 2 专利效应分类模型

### 2.1 基本概念

输入部分涉及到的功能部分包括效应功能和专利功能。效应功能指专利所应用的效应的功能,即指系统或子系统在特定约束条件下输入、输出时,参数或状态变化的一种抽象描述。比如超声波效应的功能有产生热能、产生化学变化、破坏物体、控制物质运动、检测内部性质。专利功能指的是专利中某技术点的功能。比如,“通过自平衡电路控制舱室的径向摆动使其平衡”,专利功能为“控制摆动”;“通过引入在静电力作用下可绕轴旋转的微反射镜和“工”字形金属谐振器,实现对入射太赫兹波入射角

的动态调控,从而实现对吸波体吸收峰位置的动态调控。”提取的专利功能为“控制位置”。

### 2.2 专利效应分类流程

专利效应分类分为训练过程和测试过程,如图 1 所示。

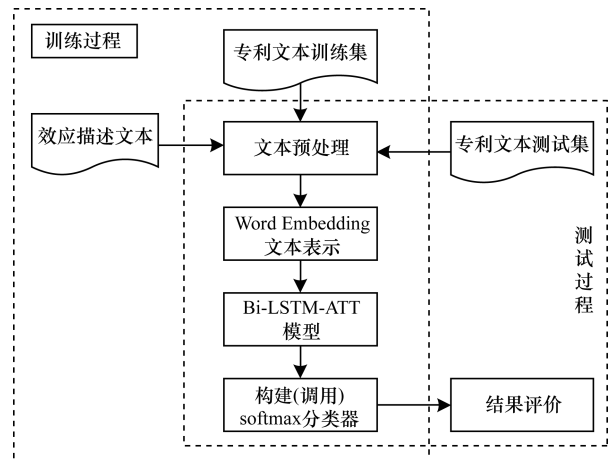


图 1 专利效应分类流程

从专利中能够分析得到效应的部分一般出现在专利的权利要求说明书和发明内容、具体实施方式部分,把这三部分合为一部分作为专利文本训练集的第一部分输入内容。从专利中分析的专利功能、效应功能从功能实现的角度反映专利的效应特征,由此,把它们作为专利文本训练集的第二部分输入内容。效应描述文本参与 word2vec 的训练,从而得到较好的向量词典,使文本表示更具代表性。

在进行文本预处理时,本文选择中科院分词系统 ICTCLAS 来对文本进行分词处理。由于专利中领域术语较多,调用分词时有些术语词往往是分开的。为了避免这种情况,本文添加一个领域术语词典作为用户字典,以减少词数目,降低运算复杂度。分词后所得到的词条集中包含着很多像“了”“总体来说”等没有实际意义的词,因此,在特征选择之前,构建了一个停用词表来去除文本的停用词。

文本预处理之后采用 Word Embedding 技术<sup>[9]</sup>进行文本表示,是因为 Word Embedding 技术可以避免传统词向量的维度过高或向量稀疏问题,还能提供含有语义信息的词向量。本文采用的是 2013 年 Google 开源推出的 word2vec 工具包。它将初始特征词映射到实数低维向量,用向量空间里的向量运算来代替对文本内容的处理,向量空间上的相似度就代表文本语义上的相似度。Word2vec 主要分为 CBOW 和 Skip-Gram 两种模式,CBOW 对小型数据比较合适,Skip-Gram 在大型语料中表现更好,本文采用的是 Skip-Gram 模型。

数据集内容主要由不定长的专利文本内容和功能两部分构成。这两部分分别经过预处理和文本表示得到专利文本向量表示和功能向量表示。专利文本向量

表示经由基于 attention 机制的双向 LSTM (即 Bi-LSTM-ATT 模型) 训练得到的向量作为专利分类的文本特征;而功能向量表示经标准 LSTM 学习得到的向量作为专利分类的功能特征。通过 pooling 层连接以上两部分特征作为专利效应分类的总体特征。之后用这些特征向量去构建或调用 softmax 分类模型,最终输出即是专利所属的效应。Bi-LSTM-ATT 模型相关内容将在下一节进行详细的介绍。

### 3 Bi-LSTM-ATT 模型的构建

#### 3.1 LSTM 基础

LSTM 由文献[10]提出,并在近期进行了改良和推广。LSTM 是在循环神经网络(RNN)的基础上发展起来的,它的结构与 RNN 的结构一样,只是对隐层单元内部结构进行了复杂化改造。RNN 在训练时,如果输入的是很长的序列,梯度向量的分量会呈指数增长或衰减,这个问题也可称之为梯度消失问题<sup>[11-12]</sup>,它会使 RNN 模型很难学习长序列中的长距离相关性。LSTM 引入了 memory cell<sup>[13]</sup>,memory cell 能够保存很长时间的的状态,具有记忆功能,由此可以解决 RNN 的梯度消失问题。LSTM 有很多变体,本文介绍的是文献[14]所使用的版本。图 2 所示为 LSTM 隐层神经元结构。

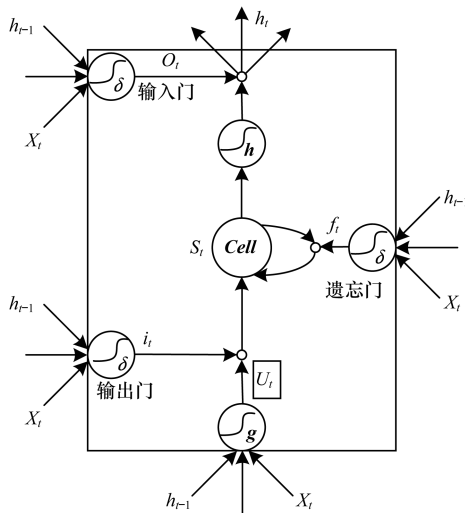


图 2 LSTM 隐层神经元结构

LSTM 隐层神经元由输入门、遗忘门、输出门这 3 个门来保护和控制细胞状态,这些门具有删除或者增加信息到细胞状态的能力。定义 LSTM 单元在时间步  $t$  时是一个  $d$  维的空间向量集合  $\mathbf{R}^d$ :  $i_t, f_t, o_t$  分别为时间为  $t$  时的输入门、遗忘门、输出门,它们的取值范围为  $[0, 1]$ 。  $u_t$  为输入值,  $s_t$  为记忆单元状态值 (memory cell),  $h_t$  为隐藏单元状态,即隐藏单元的输出值 (hidden state)。LSTM 的方程式为:

$$i_t = \delta(W^{(i)}x_t + U^{(i)}h_{t-1} + b^{(i)}) \quad (1)$$

$$f_t = \delta(W^{(f)}x_t + U^{(f)}h_{t-1} + b^{(f)}) \quad (2)$$

$$o_t = \delta(W^{(o)}x_t + U^{(o)}h_{t-1} + b^{(o)}) \quad (3)$$

$$u_t = \tanh(W^{(u)}x_t + U^{(u)}h_{t-1} + b^{(u)}) \quad (4)$$

$$s_t = i_t \odot u_t + f_t \odot s_{t-1} \quad (5)$$

$$h_t = o_t \odot \tanh(s_t) \quad (6)$$

其中,  $x_t$  表示时间步为  $t$  时的输入,  $\delta$  表示 logistic sigmoid 激活函数,  $\odot$  表示元素乘法。直观地说,遗忘门控制之前的 memory cell 遗忘信息的程度,输入门控制每个单元要更新的值,输出门决定要输出的内容。

#### 3.2 Bi-LSTM-ATT 分类模型

LSTM 属于一种 Encoder-Decoder 框架,当文本较长时,无论输入包含多少信息量,最终所有语义完全通过一个中间语义向量来表示,单词自身的信息已经消失。encoder 中丢失很多细节信息,decoder 的结果也就会明显变差,由此引入 attention 机制。单纯的 LSTM 模型在解决文本分类任务时,仅考虑了上文信息,没有用到下文信息,丢失部分语义信息。而双向 LSTM 前向传递层可以获取到输入序列的上文信息,后向传递层可以获取到输入序列的下文信息,从而丰富了语义信息。如图 3 所示为基于 attention 的双向 LSTM 神经网络,该 Bi-LSTM-ATT 分类模型分为编码过程、attention 层和 softmax 分类器层。

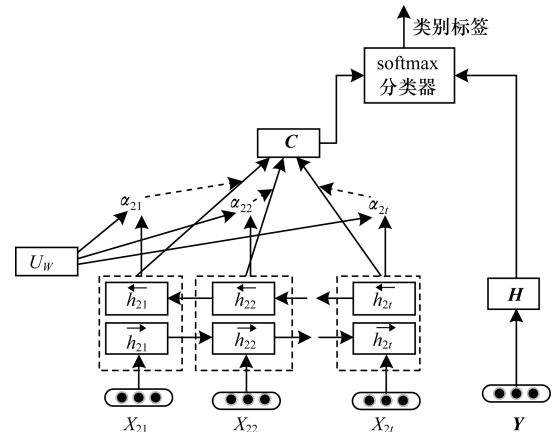


图 3 基于 attention 的双向 LSTM 神经网络

##### 3.2.1 编码过程

本文给出一个单词向量表示的文本序列,单词向量用  $x_{ij}$  表示,表示第  $j$  个词向量在第  $i$  个文本中,在这里把像  $x_{ij}$  这样的单词向量称为节点。为得到专利文本的特征向量,使用双向 LSTM 对句子进行注释,计算公式如式(7)所示。

$$\begin{aligned} \vec{h}_{it} &= \overrightarrow{LSTM}(x_{it}), t \in [1, T] \\ \overleftarrow{h}_{it} &= \overleftarrow{LSTM}(x_{it}), t \in [T, 1] \end{aligned} \quad (7)$$

连接  $\vec{h}_{it}$  和  $\overleftarrow{h}_{it}$  得到节点  $x_{it}$  的注释,如  $h_{it} = [\vec{h}_{it}, \overleftarrow{h}_{it}]$ 。  $h_{it}$  总结了与节点  $x_{it}$  邻近的一些节点,但注意力主要还是在节点  $x_{it}$  上。

### 3.2.2 Attention 机制

为突出关键词的作用,减少非关键词对于文本整体语义的影响,引入 Attention 机制抽取对文本重要的词向量,这些重要的词向量构成文本的最终特征向量。Attention 机制的相关计算公式<sup>[8]</sup>如式(8)所示。

$$\begin{aligned} u_{ii} &= \tanh(W_w h_{ii} + b_w) \\ \alpha_{ii} &= \frac{\exp(u_{ii}^T u_w)}{\sum_i \exp(u_{ii}^T u_w)} \\ C &= \sum_i \alpha_{ii} h_{ii} \end{aligned} \quad (8)$$

首先通过输入各节点的隐藏状态  $h_{ii}$ ,经由  $\tanh$  函数学习得到其隐式地表示  $u_{ii}$ 。然后把  $u_{ii}$  和词水平的上下文矢量  $u_w$  的相似度作为衡量节点的重要度。然后得到一个规范化的概率权重  $\alpha_{ii}$ ,表示词向量在文本句子中的注意力概率。接着计算基于权重单词注释的加权和作为第  $i$  个文本的文本向量  $C$ ,即文本的语义编码。

### 3.2.3 文本分类

文本向量  $C$  是文本的更高层次的表达,它可以作为专利效应分类的部分特征向量。图 3 中的  $Y$  表示经 word2vec 映射后的功能向量,它在经过标准的 LSTM 的训练学习得到向量  $H$  作为专利效应分类的另一部分依据。这 2 种特征向量在分类中的权重分别为 0.6 和 0.4,则文本属于某一类效应的概率计算公式如式(9)所示。

$$p = \text{softmax}(0.6 \times W_c C + 0.4 \times W_c H + b_c) \quad (9)$$

使用正确标签的负对数似然作为训练损失,计算公式如式(10)所示。

$$L = - \sum_d \log_a p_{dj} \quad (10)$$

其中, $j$  是文本  $d$  的标签, $R = \frac{A}{A+C}$  则表示文本  $d$  属于效应  $j$  的概率。

### 3.2.4 分类效果评价

评价分类器分类效果的指标有很多,本文采用常用的评价指标精确率、召回率和 F1 测量指标<sup>[15]</sup>。本文将从这 3 个指标来衡量算法的有效性。各评测指标计算如表 1 所示。

表 1 评测指标相关表

类别	正确分类	不正确分类
某效应类 $y$	$A$	$B$
其他类	$C$	$D$

在表 1 中, $A$  表示经分类器划分到效应类  $y$ ,且实际属于  $y$  类的专利文本数, $B$  表示经分类器划分到效应类  $y$ ,但实际不属于  $y$  类的专利文本数, $C$  表示经分类器没有划分到效应类  $y$ ,但实际属于  $y$  类的专利文本数, $D$  表示经分类器没有划分到  $y$  类,且实际也不属于  $y$  类的专利文本数。

召回率( $R$ )衡量的是实际属于某个类别的所有文本中被分类器正确划分到该类别中的比率。召回

率越高说明分类器在该类上漏掉的文本越少,式(11)为召回率数学公式。

$$R = \frac{A}{A+C} \quad (11)$$

精确率( $P$ )是衡量划分到某个类别中的文本中属于此类别的比率,精确率越高说明分类器分类越准确。式(12)为精确率的数学公式。

$$P = \frac{A}{A+B} \quad (12)$$

F1 值是召回率和精确率的加权调和平均。它平衡了精确率和召回率。式(13)为 F1 值的计算公式。

$$F1 = \frac{2PR}{P+R} \quad (13)$$

## 4 实验结果与分析

### 4.1 实验数据准备

本文选择了 100 个常用效应作为类标签,例如吸附、光的折射、X 射线、安培力、电磁感应等效应,并用 0~100 的数字来作为这些效应类标签。在使用类标签数据的时候,程序会将类标签首先转换成一定的格式。假设有 5 个专利文本,分别属于类标签  $[0, 0, 1, 2, 2]$ ,转换后的格式则为:  $[[1, 0, 0], [1, 0, 0], [0, 1, 0], [0, 0, 1], [0, 0, 1]]$ 。实验所用的专利主要为机械物理类专利。在实验训练阶段,专利语料主要包括以下部分:专利文本内容,指的是专利中的权利要求说明书、发明内容、具体实施方式三部分内容,它们均属于专利中原有的内容;功能部分分为效应功能(theoryFuction)和专利功能(functions),用于训练得到文本分类的功能特征;效应字段用来标识各专利样本所属的效应。在此阶段,借助的 100 条常用效应解释说明文本(包括效应名称、效应 id、效应说明)除了辅助分析专利所属效应外,还用于 Word2vec 的训练得到向量词典,效应说明即对效应的普遍解释,说明效应的本质。在测试阶段,所有样本使用的专利语料格式与训练阶段的相同,但在测试阶段没有用到效应解释说明文本。实验时使用的训练语料包括 100 条常用效应解释说明文本和 5 000 个专利文本。在测试时,使用了 500 个专利文本来作为测试语料。

### 4.2 实验对比及分析

本文采用 mini-batch 训练,经实验分析,针对专利文本内容而言,最终选择每个 mini-batch 样本大小为 32,隐层记忆单元数为 100,输出层节点数为 100。通过分析结果得到,网络可输入的最大长度(MAX\_SEQUENCE\_LENGTH)不变,随着迭代次数的增加,精确率也随之增大,但是迭代次数为 20 时,召回率和 F1 值相对较大,并且精确率也较大;当迭代次数不变时,MAX\_SEQUENCE\_LENGTH 为 800 时能够包含文本的所有具有代表性的词,效果也较好。针对功能字段而

言,在实验时,每个 mini-batch 样本大小为 32,网络可输入的最大长度为 20,隐层记忆单元数为 10,输出层节点数为 10,迭代次数取 10 次。将引入 Bi-LSTM-ATT 模型的方法分别与经典机器学习算法 SVM、逻辑回归(Logistic Regression, LR)、单纯的 LSTM、LSTM 与统计算法结合的方法进行对比,结果如表 2 所示。

表 2 文本分类结果比较

方法	准确率	召回率	F1 值
LR 方法	0.533 2	0.405 8	0.460 9
SVM 方法	0.598 1	0.441 0	0.507 7
LSTM 方法	0.729 9	0.647 1	0.686 0
LSTM + IG 方法	0.651 3	0.576 0	0.611 3
LSTM + MI 方法	0.645 8	0.558 3	0.598 9
LSTM + WRRL 方法	0.652 0	0.577 4	0.612 4
Bi-LSTM 方法	0.742 1	0.658 9	0.698 0
Bi-LSTM-ATT 方法	0.760 5	0.671 1	0.713 0

如表 2 所示,分别用 8 种方法实现专利效应的分类,由结果可知,基于深度学习的文本分类要优于基于机器学习的文本分类。其中 LSTM + IG (MI 或 WRRL) 是先利用统计的思想抽取专利文本的初始特征,再利用 LSTM 训练学习得到特征向量,最终借助 softmax 分类模型进行分类,这 3 种分类效果差别不大。SVM 方法采用 IG 来选取文本特征。这 4 种方法要比使用 LSTM 或 LSTM 变体的效果差,是因为引入了统计的思想去找初始特征,这些初始特征是一个个独立存在的词,词与词之间在原文中的位置关系、语义依存关系等将不复存在,即破坏了文本的时序关系和语义信息。从结果可以看到,Bi-LSTM 的精确率要比 LSTM 的精确率提高了约 2%,说明针对本任务 Bi-LSTM 比 LSTM 的效果要好,其原因是在 LSTM 中, $t$  时刻隐层编码的信息是到  $t$  为止之前的所有输入信息,也就是说  $t$  时刻的 LSTM 能够看到历史,但看不到未来。但是 Bi-LSTM 也能获取未来的信息,它不仅

考虑上文信息,还考虑了下文信息,对学习训练有积极作用。本文提出的基于 attention 的双向 LSTM (即 Bi-LSTM-ATT) 结果比 Bi-LSTM 提高了约 3%,比 LSTM 提高了约 2%,其主要原因是前者在合并输出向量时,attention 机制为每个输出向量赋予了不同的权值,使模型将注意力集中在对专利文本分类任务更重要的向量上,从而降低了无关向量的作用。这个模型能够更好地表征文本,突出了关键词的作用。此外,Bi-LSTM-ATT 模型既保留了文本的时序关系和语义关系,还考虑了文本上下文信息。综上考虑,说明基于 Bi-LSTM-ATT 模型的分方法对判定专利所属类别有效。

### 4.3 工程应用

与普通的文本相比,本文按效应分类的专利能够辅助设计者进行创新设计,为人们提供更多的思路去解决一些问题。举例来说,传统机翼的襟翼和副翼控制面利用铰链连接到机翼上,因为铰链处的过渡不光滑,流过机翼的气流过早地分离,造成气动效率低的问题,当控制面偏转角越大时,性能降低得越明显<sup>[16]</sup>。

1) 问题分析:如果流过机翼气流的附着时间能够延长,就能改善整体压力的分布,并且如果外侧机翼攻角增大,就能提高机翼的升力和俯仰力矩。这样能解决以上气动效率低、性能低的问题。

2) 功能确定:由于偏转控制面板及扭转机翼能够实现以上条件,由此确定功能为“改变物体尺寸”。

3) 检索专利:检索“改变物体尺寸”这一功能就会得到热膨胀、形状记忆合金、形变、压电效应、磁弹性、压磁效应,这些效应均可实现“改变物体尺寸”这一功能。通过建立的分类模型可以知道各专利所属的效应,进而得到可借鉴的专利。针对这个案例,逐一分析每一个效应下对应的专利,检索结果如表 3 所示。表 3 仅列出了属于热膨胀和形状记忆合金的部分专利及其概述。

表 3 效应与专利对应关系

效应名称	专利号	专利内容
热膨胀	CN01456045A	使用可流动且热膨胀系数大的材料作为板材成形中的凸模,当凸模温度增加时,其增大的体积将推动板材和凹模相贴,成形出工件
热膨胀	CN1149363C	日照外管,内管升温,不冻液受热膨胀,内管压力增加,热膨胀吸收装置被压缩,将热空气排除管外,日落内管压力减小,空气进入热膨胀吸收装置并恢复到原来状态
...	...	...
形状记忆合金	CN101132155A	通过对形状记忆合金丝施加电流调控形状记忆合金丝的温度,使其产生相应的收缩形变和复原,进而这种收缩运动通过旋转机构带动微电机转轴旋转
形状记忆合金	CN101581346A	在消能器中布置了形状记忆合金间隔板,使用过程中,钢板与形状记忆合金板将会发生内碰-摩擦-形变三级耗能,共同抑制结构或构件在平面内的扭转振动,提高了消能器的耗能能力
形状记忆合金	CN107076334A	凸轮从动件和凸轮从动件槽之间的作用将形状记忆合金弹簧作用在往复环上的线性动作转化成旋转阀的旋转动作,该动作是环境温度变化的结果。进而调节流体的流动
...	...	...

4) 解决方案确定:对以上效应进行取舍。结合这些专利,可以发现热膨胀是物质的体积发生变化,温度升高,体积变大,温度降低,体积缩小,此时的物质不一定能恢复原来的形状;形状记忆合金在某种条件下经过一定的塑性变形后,加热到一定温度,材料又恢复到原来的形状,形状记忆合金更具有可控性,适合窄小空间条件下使用。可以判断出形状记忆合金效应适合控制面的偏转和机翼的扭转,因而能够确定最终解决方案为使用嵌入形状记忆合金丝的柔性光滑结构连接襟翼、副翼和机翼,使用形状记忆合金扭力管驱动机翼扭转。当然也可以想像形状记忆合金是否能够作为机翼的材料使用,从而解决飞机升起时存在的问题。

通过案例可以看出,按效应分类的专利能够辅助设计者确定正确的效应和最终的方案。

## 5 结束语

本文基于自然语言处理,在双向 LSTM 层上引入了 attention 机制,即增加了注意力概率模型,依此来提取文本特征向量。由于功能字段较短,因此使用标准 LSTM 训练提取功能特征向量,从 2 个方面来表示专利所属效应。提取特征向量之后,再利用 softmax 回归来解决文本多分类问题。在已有的经人工分析的 100 条效应数据及选择的 5 500 个专利文本数据集上进行的实验结果表明,该方法在分析专利所应用的效应知识上具有一定的可用性,可减少人工分析专利所需的人力、时间,能够辅助设计者拓宽创新思路。本文只是从专利文本内容和功能字段 2 个方面进行考虑,专利文本内容过于宽泛。因此,下一步将专利中应用了效应的一小段文本或是几句话整理出来作为样本数据集,使文本信息更具代表性,更精确地表示文本,以达到提高准确率的目的。

### 参考文献

- [1] MURPHY J, FU K, OTTO K, et al. Function based design-by-analogy: a functional vector approach to analogical search [J]. *Journal of Mechanical Design*, 2014, 136(10).
- [2] 王朝霞. 专利知识获取及其支持概念创新设计的方法研究[D]. 杭州:浙江大学, 2009.
- [3] 冀瑜, 邱清盈, 冯培恩, 等. 国际专利分类表中设计知识的提取和利用[J]. *浙江大学学报(工学版)*, 2016, 50(3):412-418, 427.
- [4] ZHU X, SOBHANI P, GUO H. Long short-term memory over tree structures [EB/OL]. [2017-06-21]. <https://www.researchgate.net/>.
- [5] TAI K S, SOCHER R, MANNING C D. Improved semantic representations from tree-structured long short-term memory networks [J]. *Computer Science*, 2015, 5(1):36.
- [6] WANG Y, HUANG M, ZHU X, et al. Attention-based LSTM for aspect-level sentiment classification [C]// *Proceedings of Conference on Empirical Methods in Natural Language Processing*. Washington D. C., USA: IEEE Press, 2016:606-615.
- [7] RAFFEL C, ELLIS D P W. Feed-forward networks with attention can solve some long-term memory problems [EB/OL]. [2017-06-21]. <http://pdfs.semanticscholar.org>.
- [8] YANG Z, YANG D, DYER C, et al. Hierarchical attention networks for document classification [C]// *Proceedings of Conference of the North American Chapter of the Association for Computational Linguistics; Human Language Technologies*. Washington D. C., USA: IEEE Press, 2017: 1480-1489.
- [9] 孟欣, 左万利. 基于 word embedding 的短文本特征扩展与分类[J]. *小型微型计算机系统*, 2017, 38(8): 1712-1717.
- [10] GERAEV S. Longshort-term memory [J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [11] GERAEV A. Supervised sequence labelling with recurrent neural networks [M]. Berlin, Germany: Springer, 2012.
- [12] LIPTON Z C, BERKOWITZ J, ELKAN C. A critical review of recurrent neural networks for sequence learning [EB/OL]. [2017-06-21]. <http://zacklipton.com/media/papers/>.
- [13] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. *Nature*, 2015, 521(7553): 436-444.
- [14] JOZEFOWICZ R, ZAREMBA W, SUTSKEVER I. An empirical exploration of recurrent network architectures [C]// *Proceedings of the 32nd International Conference on Machine Learning*. Washington D. C., USA: IEEE Press, 2015: 2342-2350.
- [15] 张青, 吕钊. 基于主题扩展的领域问题分类方法[J]. *计算机工程*, 2016, 42(9): 202-207.
- [16] 李杰锋, 沈星, 杨学永. 形状记忆合金在变体机翼中的应用现状[J]. *材料导报*, 2014, 28(7): 104-108.

编辑 索书志