

# Research on News Recognition of “Clickbait” Fusing Image Information

Lin Yang, Jichao Ding, Sheng Zhu, Shuai Wang

School of Mechanical Electronic & Information Engineering, China University of Mining and Technology (Beijing), Beijing  
Email: d.jichao.dp@gmail.com

Received: May 13<sup>th</sup>, 2020; accepted: Jun. 4<sup>th</sup>, 2020; published: Jun. 11<sup>th</sup>, 2020

---

## Abstract

In recent years, online news has gradually replaced the traditional paper news, becoming the main way for people to get news daily. Internet news has thus become an industry. The main body of the industry is composed of news producers, users, and click-through rates. News producers convert users' clicks into click-through rates to obtain benefits, which leads to the emergence of the “Clickbait”. The “Clickbait” is a huge hazard to users, industries and even society. In the past, the identification method of “Clickbait” ignored the feature that the main body of online news was composed of two parts: image and text. It only detected the text information of the news and ignored the image information of the news. At present, there is a lot of news on the Internet that uses eye-catching or irrelevant images to attract users to click. This article designs an image information extraction model using deep learning related technology for the image information in the news. Using this model, the news Extract the information of the image, design the features of the extracted information, integrate the image features into the “Clickbait” recognition model, and finally verify the necessity and effectiveness of using the image information to identify the “Clickbait” news.

## Keywords

Image Caption, Clickbait, Machine Learning, Text Similarity

---

# 融合图片信息的“标题党”新闻识别研究

杨林, 丁继超, 朱胜, 王帅

中国矿业大学(北京)机电与信息工程学院, 北京  
Email: d.jichao.dp@gmail.com

收稿日期: 2020年5月13日; 录用日期: 2020年6月4日; 发布日期: 2020年6月11日

## 摘要

近年来,网络新闻逐渐取代传统的纸质新闻,成为人们日常获取新闻的主要方式。网络新闻因此而成为一个产业,产业的主体是由新闻制作者、用户、和点击率构成,新闻制作者通过用户的点击转换为点击率获取利益,由此导致了“标题党”的产生。“标题党”对用户、行业乃至社会都有巨大的危害。以往的“标题党”识别方法都忽略了网络新闻的主体是由图像和文本两个部分组成这一特点,只针对新闻的文本信息进行检测,而忽略了新闻的图片语义信息。目前网络上有大量利用引人眼球或者与文章毫不相关的图片吸引用户点击的新闻,本文针对新闻中的图像信息利用深度学习相关技术设计了图像语义描述与信息提取模型,使用这一模型,对新闻中的图片进行信息提取,对提取到的信息进行特征设计,将图片特征融合进“标题党”识别模型中,最后通过实验验证了使用图片信息识别“标题党”新闻的必要性和有效性。

## 关键词

图像描述, 标题党, 机器学习, 文本相似度

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

网络新闻在传播方式上具有流量化、去中心化、题文分离的特点,新闻标题既承担文章内容的概括作用,同时还作为文章链接负责将用户引导到新闻正文部分[1]。网络新闻内容多、数量大、更新速度快的特点,使用户逐渐养成快速阅读新闻标题并根据标题内容决定是否点击链接的“读题”式浏览习惯,也在素材趋于同质化的新闻行业中催生了“标题党”的产生[2][3]。“标题党”新闻常使用过度夸张、歪曲事实、制造虚假新闻等手段,加工出引人眼球、耸人听闻的标题,以迎合、吸引用户。“标题党”新闻的泛滥,不仅影响用户获取所需的新闻内容,而且对网络新闻行业的健康发展以及网络和谐环境的构建造成恶劣影响。“标题党”现象已受到国家相关部门和社会各界的关注,相关部门积极采取各种措施对“标题党”新闻进行打击和消除,并加强在“标题党”新闻识别方面的研究。

## 2. “标题党”新闻识别方法的介绍及分析

### 2.1. “标题党”新闻识别方法研究现状

目前,针对“标题党”新闻识别的研究主要以“标题党”写作中使用的语言特征为主要研究对象,常用的“标题党”新闻识别方法可分为基于文本相似度的识别方法和基于机器学习模型的识别方法两大类。

基于文本相似度的识别方法主要针对“标题党”新闻中“题文不符”的现象,利用文本相似度计算对“标题党”新闻进行判断。2011年,王志超提出基于主题句相似度的“标题党”新闻鉴别技术,利用正文主题句集合与标题的文本相似度对“标题党”新闻进行识别[4]。2015年,罗佳提出基于潜在语义的“标题党”新闻识别技术,利用正文中的词频构建向量空间模型,奇异值分解后得到正文的塌陷矩阵,生成标题的文档坐标,通过计算文档坐标和与各段落对应向量间的余弦相似度进行判断[5]。2018年,赵帅提出基于改进型VSM-How Net融合相似度算法的“标题党”新闻识别方法,使用同义词词组的向量替

代词向量对正文进行表达, 利用改进型 VSM 结合余弦相似度来度量文本相似度[6]。2019 年, 陈彤从语言学角度, 提取主题词后计算标题和正文相似度对“标题党”新闻进行识别[7]。

随着机器学习的兴起和运用, 相关学者提基于机器学习的“标题党”识别方法, 利用提取到的新闻标题及正文特征进行分类模型的训练, 使模型不断学习“标题党”新闻的各种特征, 并利用模型对“标题党”新闻进行判断。2016 年, Potthast 等人针对 twitter 中“Clickbait”现象, 提取了三大类共 215 个特征进行模型训练, 构建出“Clickbait”检测模型[8]。同年, Chakraborty 等人利用多种机器学习分类模型, 设计了一个浏览器扩展, 以达到检测“Clickbait”的目的[9]。

## 2.2. 现有“标题党”新闻识别方法及本文方法介绍

“标题党”新闻识别方法的不断成熟, 对“标题党”新闻起到了一定的控制作用。但随着网络技术与新闻行业的发展, 新闻图片作为提高视觉吸引力和视觉感受的媒介, 逐渐成为网络新闻中不可获取的一部分, 目前网络新闻的内容往往也是由新闻文本与新闻图片组合而成的, 新闻图片更是成为“标题党”新闻用于吸引用户的新手段[10]。因此, 仅仅利用新闻文本的语言特征对“标题党”新闻进行识别的方法可靠性逐渐降低, 不能够完全满足保障网络新闻健康传播的需求。因此, 研究结合新闻图片信息和文本信息的“标题党”新闻识别方法具有十分重要的意义。

本文提出一种综合考虑新闻图片与新闻文本特征的“标题党”新闻识别方法, 以提高对当前流行的“标题党”新闻的识别准确率。本文基于深度学习思想, 利用图像描述等技术, 充分提取新闻中的图片语义信息, 并对新闻图片中的人脸图片及文本图片进行模型构建。对于提取到的多类型图片信息, 本文提出一种图片和文本相关度评价指标, 并将该指标融入“标题党”新闻识别模型进行“标题党”新闻的识别与判断。

## 3. 新闻图片特征提取

对图片信息的信息提取是本文的主要研究内容, 通过对大量新闻的定量统计发现, 新闻中的配图主要可以分为两类: 一是人物、场景及物体信息等非文本图片, 二是评论截图、通知截图等文本图片。二者在信息提取过程中需要侧重点略有不同, 因此, 为了提高对新闻图片信息的有效提取, 针对不同类型的图片采用不同的提取方法。本文使用相关深度学习技术进行图片信息提取, 主要由四个模块组成, 分别是: 文本图片信息提取、非文本图片信息提取, 人物信息提取, 和文本图片相关度计算, 整体模型结构如图 1。

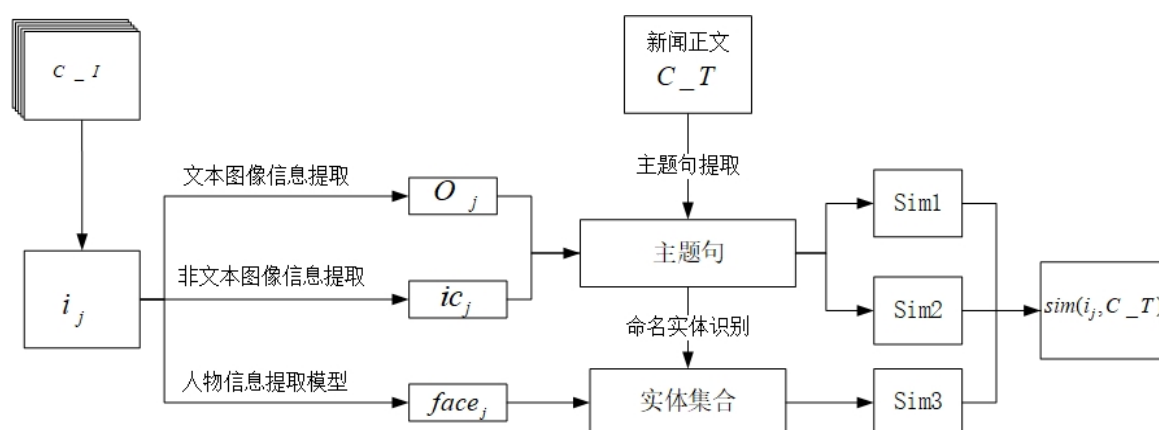


Figure 1. Image feature extraction flowchart  
图 1. 图像特征提取流程图

### 3.1. 图像语义描述与信息提取

本文使用融合场景信息的图像描述模型[11]获取新闻图片中的语义描述信息，获取新闻图像的图像描述，模型主要结构如图 2 所示。

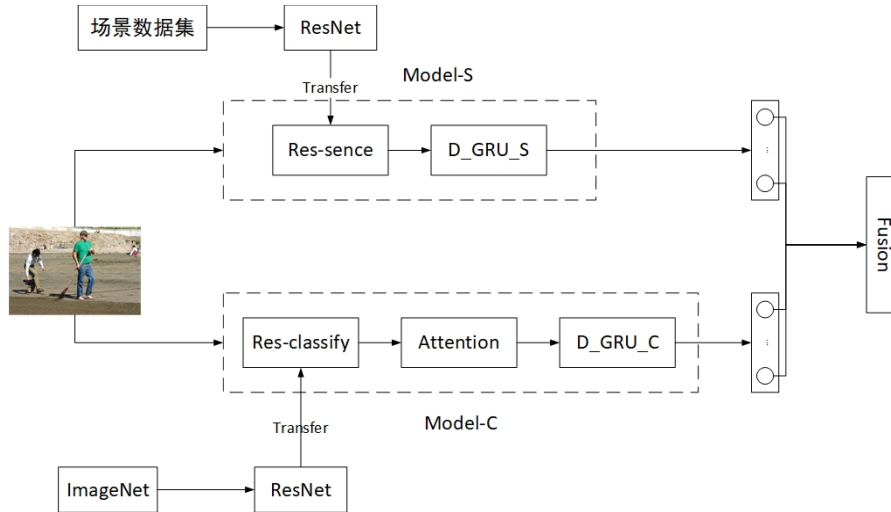


Figure 2. Scene image description structure diagram  
图 2. 场景图像描述结构图

Res-sence 是用于提取图片场景信息的模型，Res-classify 是用于提取图片中物体信息的模型，两者在结构上使用的都是 ResNet 101 网络结构。区别在于 Res-sence 是在场景数据集 places-365 和 AI challenge 的场景分类模型数据集上预训练的，而 Res-classify 是在 ImageNet 物体分类数据集上训练的。两个模型分别连接双层 GRU 解码器，得到侧重场景信息和侧重物体信息的图像语义表达，最后对概率进行加权平均计算，得到最终的输出。Attention 代表软注意力机制，考虑到本文研究方法中，新闻图像信息的提取需在最大程度上与新闻内容相匹配，因此，本文在使用 Res-classify 模型对非文本图像进行物体信息提取时，融入软注意力机制的思想[12]。Fusion 代表加权融合操作，在  $t$  时刻 Model-S 和 Model-C 的输出概率值为  $p'_S$  和  $p'_C$ ， $\alpha_S$  和  $\alpha_C$  表示权重，得到  $p'_{Fusion}$  表达式：

$$p'_{Fusion} = \frac{\alpha_S \cdot p'_S + \alpha_C \cdot p'_C}{\alpha_S + \alpha_C}$$

最终模型在 AIC-ICC 数据集(AI challenger 图像描述竞赛数据集)上训练得到本文的场景图像信息提取模型，AIC-ICC 数据集中数据分布表 1 所示。

Table 1. AIC-ICC data set distribution table  
表 1. AIC-ICC 数据集分布表

数据集	训练集	测试集	验证集
AIC-ICC	210000	30000	30000

图像描述常用的客观评价指标有 BLEU, CIDEr 等，但是考虑到客观评价指标主要是评价模型生成语句与参考语句的相似程度或者是评价生成语句与人类表达的差异性，这与本文的应用场景并不完全符合，本文选择人工评价的方式，具体的是设立四个评价等级，分别是非常好、好、一般、差，具体等级及对应等级描述如表 2 所示。

**Table 2.** Manual evaluation standard grade**表 2.** 人工评价标准等级

等级	描述
非常好	模型对于一幅图像生成的描述完全体现出图像中所有的细节信息
好	模型对生成的描述是正确的, 但部分词语使用不当
一般	模型对于一幅图像生成的描述场景和物体基本正确, 但是物体之间的属性关系有错误, 或者遗漏了一些图像中的信息
差	描述完全错误, 生成的描述和图像完全不一样, 在语义空间中没有任何相似的地方

**模型训练参数设置:**

因为对于循环神经网络来说时间步的大小决定了网络的记忆能力, 直接决定了在推断阶段网络能够生成的单词数, 因此在解码阶段, 将双层 GRU 模型的时间步设为 30。在模型编码、解码结构联合训练过程中, 考虑到 GPU 显存大小, 设置 Batchsize 为 8, 本文使用 Adam 算法训练网络的参数, 初始学习率为 0.001, Adam 的参数包括动量和衰减系数, 分别设为 0.9 和 0.99。

**实验结果:**

本文在 AIC-ICC 训练集中随机选取了 500 幅图片, 对模型生成的语句进行了人工评价, 具体实验数据如表 3 所示。

**Table 3.** Experimental data**表 3.** 实验数据

非常好	好	一般	差
36.0%	29.8%	19.0%	15.2%

对于图片中的人脸信息提取和文本图像信息提取, 本文使用开源的模型实现这两个模块的功能, 文本图像信息提取使用的是文献[13]提出的 CTPN 文本框检测模型和文献[14]提出的 CRNN 文本识别模型。在人脸信息提取上, 本文使用了文献[15]提出的 MTCNN 模型和百度的公众人物识别的 API。

**3.2. 图片与文本相关度计算**

因为新闻中的图片作为新闻的一部分, 图片的语义表达往往是和新闻的内容息息相关的。类似于文本新闻的“标题党”识别方法中使用正文文本与标题文本的相似度计算, 本文考虑使用新闻图片和正文文本的相关度作为模型识别的一个特征, 本文使用表 4 所使用的符号分别代表新闻中每个组成部分。

**Table 4.** The symbol of the main content in the news**表 4.** 新闻中主要内容符号示意

新闻组成	符号
超链接页面文本标题	$T\_T$
超链接页面标题图片	$T\_I = \{i_1, i_2, \dots, i_n\}$
正文页面文本内容	$C\_T$
正文页面图片集合	$C\_I = \{i_1, i_2, \dots, i_m\}$

其中  $T\_I \subseteq C\_I$ , 即  $m \geq n$ 。故本文在研究新闻中的图片与新闻主体的相关度的时候主要考虑的对象是正文图片和正文文本。

针对  $C\_I$  中一副图片  $i_j$ ，因为  $i_j$  可能同时包含图像描述、文本信息以及人脸信息。将  $i_j$  经过图像语义描述与信息提取模型进行处理，得到一个三元组  $\{\{O_j\}, \{ic_j\}, \{face_j\}\}$ ，其中  $\{O_j\}$  表示这幅图片经过 OCR 模型转换得到的文本信息； $\{ic_j\}$  表示这幅图片进过图像描述模型处理得到的图像描述； $\{face_j\}$  表示这幅图片经过人物信息提取模型处理得到的人物信息。因为考虑到  $\{O_j\}$  和  $\{ic_j\}$  是直接与新闻正文相关的文本信息，且本身具有一定的语义信息，我们可以将得到的  $\{O_j\}$  和  $\{ic_j\}$  分别与新闻的正文的主题句进行直接的文本相似度计算，可以得到这两个部分与新闻正文的相关度的度量值  $sim1$ ， $sim2$ 。

对于人物图像信息的使用，本文在对大量“标题党”新闻分析后发现，一些“标题党”新闻往往使用名人、政要的人脸图像吸引用户点击，让用户产生一种这篇新闻是与这个人物有关系的错觉，但是往往新闻的正文文本信息和使用的人物图像相关度很低。针对这个特点，本文考虑对新闻正文进行主题句提取后，对主题句进行命名实体识别操作，识别句子中的人名得到一个包含人名的集合  $CH\_Theme$ ，对  $ch_j$  和  $CH\_Theme$  进行比较，具体如下：

判断  $ch_j$  是否是  $CH\_Theme$  的子集，如果  $ch_j$  是  $CH\_Theme$  的子集，则认为这张图片和文章有关联，令  $sim3$  的值为 1；如果  $CH\_Theme$  中并未包含  $ch_j$  则  $sim3$  的值为 0；如果两个集合中有交集，则利用 Jaccard 相似度计算方法，得到  $sim3 = \frac{|ch_j \cap CH\_Theme|}{|ch_j \cup CH\_Theme|}$ ，由此可以得到关于一张图片中的人物信息的相关度的度量  $sim3$ ，且  $sim3$  的取值范围为  $[0, 1]$ 。

考虑到一副图片中并不一定全部包含这三类图像信息，故这三个值可能为空值，所以单独将这三个相似度值分别作为图片和文本的相似度度量是不合适的，会导致一幅图片因为某个方面的信息的缺失而被模型认为这张图片与文本的相关度为 0。因为  $sim1$  和  $sim2$  本身是通过文本相似度计算得到，故本身的取值范围是  $[0, 1]$ ，而我们从  $sim3$  的计算过程也可以发现， $sim3$  本身的取值范围也为  $[0, 1]$ ，我们考虑使用计算三个值的均值，将得到的均值作为一幅图片和文本的相关度的度量值。计算方法如下：

$$sim(i_j, C\_T) = (sim1 + sim2 + sim3) / N$$

其中， $N$  为一幅图片中存在的  $sim$  值的个数，又因为  $C\_I = \{i_1, i_2, \dots, i_m\}$ ，所以

$$sim(C\_I, C\_T) = \frac{1}{m} \sum_{j=1}^m sim(i_j, C\_T)$$

最终提取的新闻图片特征如表 5 所示。

**Table 5.** News image features

**表 5.** 新闻图片特征

符号	解释
$Num\_Image$	新闻中包含图片的数目
$CI\_CT\_Sim$	新闻中图片与正文文本的相关度

## 4. 实验过程及结果分析

### 4.1. 实验设计与评价标准

#### 4.1.1. 实验方案设计

本文设计了对比实验，验证本文提出的融合图像信息的“标题党”新闻识别方法的有效性。本文



使用文献[13]的模型作为 baseline，与本文的融合图像信息的“标题党”新闻识别模型进行对比实验。与原文不同，但是由于本文考实验必须使用、包含图像和文本两种模态的数据集，所以本文通过爬虫制作了包含文本数据和图像数据两种模态的新闻分类数据集，在这个数据集上对比 baseline 和在 baseline 中加入图片与文本相关度的模型效果，具体的实验模型设计如表 6 所示。

**Table 6.** Experimental model design  
**表 6.** 实验模型设计

模型	模型设计
baseline	文本特征 + 随机森林算法
baseline + sim(CI, CT)	图片特征 + 文本特征 + 随机森林算法

#### 4.1.2. 评价标准

**Table 7.** Confusion matrix  
**表 7.** 混淆矩阵

	Positive	Negative
正类	TP	FN
负类	FN	TN

通过表 7 中的混淆矩阵，可以得到精确度、召回率、 $F_1$ -值的计算公式：

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$\frac{1}{F_1} = \frac{1}{2} \left( \frac{1}{Precision} + \frac{1}{Recall} \right)$$

#### 4.2. 实验数据获取及预处理

本文使用 Scrapy 框架从一些新闻分发平台通过爬虫的方式构建本文的实验数据集，每条数据闻的正文、标题，文章类别，文章的作者，文章的发表时间，文章的评论数，文章的浏览量等数据。对爬取的新闻进行人工筛选和人工标注，最后获得实验数据 3356 条，其中非“标题党”新闻 1700 条，“标题党”新闻 1656 条。

数据预处理主要包括对图像信息的转换过程以及特征提取过程，具体如下：

- 1) 使用本文提出的图像信息提取模型对每篇新闻的图片进行信息提取；
- 2) 对文本进行分词处理，本文同样采用 Jieba 分词工具，Jieba 分词目前被广泛使用，在多个自然语言处理相关任务中均取得较好的表现；
- 3) 本文在文本信息的表示上使用 Word2Vec 工具对数据集中的原始文本信息和图片转换出来的文本信息进行词嵌入模型训练，使用 Skip-gram 模型的训练方式，词向量的维度设置为 128；
- 4) 使用余弦相似度计算每篇新闻的标题文本和正文文本的相似度，利用本文提出的文本图片相关度计算方法，计算获得每篇新闻的图片集合和正文的图片文本相关度。编写脚本，依据文献[16]的研究提取新闻的文本特征，并对全部特征进行特征筛选，最终得到模型的训练特征及每个特征的重要度评分如表 8 所示。

**Table 8.** Model training features and importance ranking  
**表 8.** 模型训练特征及重要度排序

符号	解释	贡献度
TT_CT_Sim	新闻正文文本和标题文本之间的相似度	1.9211
Have_S_Word	新闻标题中是否包含高频词汇, 高频词汇是指有具有吸引眼球效果的词汇	1.4416
CT_CI_Sim	图片文本相关度	1.1455
Have_Pron	新闻标题中是否含有“她”, “他”这类指代词	1.0762
Have_Symbol	新闻标题中是否具有情感倾向性的标点符号	1.0263
CT_Length	新闻正文长度	0.9376
Have_Num	新闻标题中是否包含数量词汇	0.8824
Is_Origin	新闻是否是原创新闻	0.7698
TT_Length	新闻标题的长度	0.6936
Have_Names	新闻标题中是否使用名人姓名	0.6182
News_Tag	新闻标签	0.5928
Nums_Image	新闻中图片的数量	0.5496
Nums_Review	新闻评论数	0.4096
Avg_Sen_Length	正文中平均句子的长度	0.2651

### 4.3. 模型训练

在模型选择和训练方式上, 为了保持单一变量, 我们在 **baseline** 上采用与文献中相同的采样方式和分类模型, 即采用有放回随机抽取的采样方式, 分类算法使用随机森林分类算法。本文的分类算法模型实现主要使用的是 Scikit-learn 机器学习库, Scikit-learn 是一个开源的机器学习 Python 库, 其中集成了包含分类、聚类、回归等众多机器学习算法。为了避免模型训练过拟合同时可以使模型学习到更多的有效信息, 本文使用  $K$  折交叉验证法训练模型避免过拟合,  $K$  值取 10。

### 4.4. 实验结果

**Table 9.** Experimental data of “Clickbait” news classification  
**表 9.** “标题党”新闻分类实验数据

Model	精确度	召回率	F <sub>1</sub> -score
baseline	0.862	0.842	0.852
Baseline + sim(CI, CT)	0.896	0.878	0.887

从表 9 中我们可以直观的看出, **baseline** 在同时含有文本信息和图像信息的多模态新闻数据集上精确度是 0.862, 召回率是 0.842, F<sub>1</sub>-score 是 0.852, 这个结果与文献[16]中使用随机森林的最好成绩 0.87 相比, 精确度有所下降, 导致这个现象的原因应该是数据集差异, 本文使用数据集一共是 3356 条, 而 **baseline** 使用的数据集为 808 条, 这对最后结果有所影响。而在 **baseline** 融合入图片特征信息后精确度是 0.896, 召回率是 0.878, F<sub>1</sub>-score 是 0.887, 相对于 **baseline**, 模型识别效果在精确度上有所提高, 但是提高的幅度并不是很大, 经过分析, 本文认为可能有两个方面的原因, 第一是因为数据集中部分图片“标题党”新闻中也具有“标题党”的文本特征, 这一部分的信息, 在忽略图片信息以后, 依然可以使用新闻的文本特征去识别, 而在加入图像信息之后, 模型的检测效果有所提升, 被提升出来的这一部分新闻, 是仅



仅包含图片“标题党”特征的新闻，这一类新闻在数据集中较少，因此导致了模型在评价指标上的提升效果有限。第二个导致模型在评价指标上提升较小的原因应该是本文使用的图像语义描述与信息提取模型在部分新闻图片与文字转换的结果上有所偏差，因为图像描述模型尽管已经在大部分图片文字转换任务中取得了较好的效果，但是依然有部分图片转换出来的效果较差，导致无法准确提取图片中的语义信息，直接影响到图片和文本相关度计算的准确性，使得图像信息并没有被充分利用。“标题党”新闻的识别效果有所提升，可以检测出一些原本只依靠文本信息无法检测出来的新闻，在一定程度上验证了本文提出的图片和文本相关度这一特征的有效性。

## 5. 结论

本文给出了对于网络新闻“标题党”现象检测的一种全新的思路，对于网络新闻由文本和图像两种模态数据组成的情况，本文提出使用图像语义描述与信息提取模型提取图片中的信息，并设计了一种融合场景先验信息和注意力机制的图像描述模型，提出了一种使用图像语义描述与信息提取模型提取的文本信息计算新闻图像和文本相关度的方法，并对方法的有效性进行了实验验证，实验结果表明，融合了图像信息的“标题党”新闻检测模型，综合考虑了新闻的图片特征和文本特征信息，可以更加准确、全面的检测“标题党”新闻。

## 参考文献

- [1] 张丹. 互联网时代“标题党”新闻的泛滥与有效遏制[J]. 人文与科技(第三辑), 2019(3): 207-215.
- [2] 王雅凤, 张文婷. “读题时代”新闻标题的研究[J]. 科技传播, 2016, 8(15): 6+11.
- [3] 王辰瑶, 金亮. 网络新闻“标题党”的现状与叙述策略——对 8 家网站新闻排行榜的定量分析[J]. 新闻记者, 2013(2): 65-71.
- [4] 王志超, 翁楠, 王宇. 基于主题句相似度的标题党新闻鉴别技术研究[J]. 现代图书情报技术, 2011(11): 48-53.
- [5] 罗佳. 基于潜在语义分析的“标题党”新闻识别技术研究[D]: [硕士学位论文]. 武汉: 湖北工业大学, 2015.
- [6] 赵帅. 基于改进型 VSM-How Net 融合相似度算法在“标题党”新闻识别中的研究[D]: [硕士学位论文]. 长春: 吉林大学, 2018.
- [7] 陈彤. 基于主题词对比的标题党新闻识别与分析[D]: [硕士学位论文]. 锦州: 渤海大学, 2019.
- [8] Potthast, M., Köpsel, S., Stein, B. and Hagen, M. (2016) Clickbait Detection. In: Ferro, N., et al., Eds., *Advances in Information Retrieval. ECIR 2016. Lecture Notes in Computer Science*, Springer, Cham, 810-817. [https://doi.org/10.1007/978-3-319-30671-1\\_72](https://doi.org/10.1007/978-3-319-30671-1_72)
- [9] Chakraborty, A., Paranjape, B., Kakarla, S. and Ganguly, N. (2016) Stop Clickbait: Detecting and Preventing Clickbaits in Online News Media. 2016 *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, San Francisco, CA, 18-21 August 2016. <https://doi.org/10.1109/ASONAM.2016.7752207>
- [10] 赵军朋. 会“说话”的新闻图片——网络新闻图片的价值及应用[J]. 新闻战线, 2016(12): 85-86.
- [11] 汤鹏杰, 谭云兰, 李金忠. 融合图像场景及物体先验知识的图像描述生成模型[J]. 中国图象图形学报, 2017, 22(9): 1251-1260.
- [12] Xu, K., Ba, J., Kiros, R., et al. (2015) Show, Attend and Tell: Neural Image Caption Generation with Visual Attention. *32nd International Conference on Machine Learning, PMLR 37*, 2048-2057.
- [13] Tian, Z., Huang, W. and He, T. (2016) Detecting Text in Natural Image with Connectionist Text Proposal Network. . In: Leibe, B., Matas, J., Sebe, N. and Welling, M., Eds., *Computer Vision-ECCV 2016. Lecture Notes in Computer Science*, Springer, Cham, 56-72. [https://doi.org/10.1007/978-3-319-46484-8\\_4](https://doi.org/10.1007/978-3-319-46484-8_4)
- [14] Shi, B., Bai, X. and Yao, C. (2017) An End-to-End Trainable Neural Network for Image-Based Sequence Recognition and Its Application to Scene Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 2298-2304. <https://doi.org/10.1109/TPAMI.2016.2646371>
- [15] Zhang, K., Zhang, Z., Li, Z. and Yu, Q. (2016) Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks. *IEEE Signal Processing Letters*, **23**, 1499-1503. <https://doi.org/10.1109/LSP.2016.2603342>
- [16] 陈春志. 基于多特征监督学习的“标题党”新闻识别技术研究[D]: [硕士学位论文]. 北京: 中国矿业大学(北京), 2018.