



(12) 发明专利申请

(10) 申请公布号 CN 116340887 A

(43) 申请公布日 2023. 06. 27

(21) 申请号 202310610867.8

(22) 申请日 2023.05.29

(71) 申请人 山东省人工智能研究院

地址 250000 山东省济南市历下区科院路
19号

申请人 山东中联视听信息科技股份有限公司
天津理工大学

(72) 发明人 高赞 杨红旗 赵一博 马春杰
李传森 刘冬冬

(74) 专利代理机构 山东知圣律师事务所 37262
专利代理师 张露露

(51) Int. Cl.

G06F 18/25 (2023.01)

G06F 18/2431 (2023.01)

G06N 3/0464 (2023.01)

G06N 3/0455 (2023.01)

G06N 3/0499 (2023.01)

G06N 3/08 (2023.01)

G06V 40/30 (2022.01)

G06F 16/34 (2019.01)

G06F 16/35 (2019.01)

G06V 10/764 (2022.01)

G06V 10/80 (2022.01)

G06V 10/82 (2022.01)

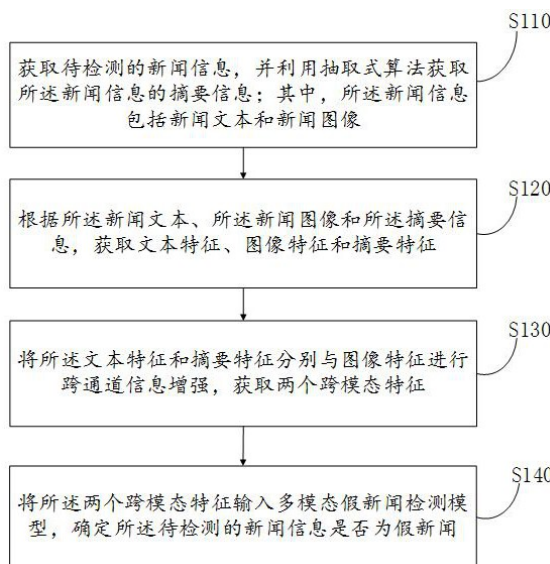
权利要求书2页 说明书12页 附图4页

(54) 发明名称

多模态假新闻检测方法及系统

(57) 摘要

本发明提供一种多模态假新闻检测方法及系统,属于电子信息技术领域,获取待检测的新闻信息,并利用抽取式算法获取新闻信息的摘要信息;根据新闻文本、新闻图像和摘要信息,获取文本特征、图像特征和摘要特征;将文本特征和摘要特征分别与图像特征进行跨通道信息增强,获取两个跨模态特征;将两个跨模态特征输入多模态假新闻检测模型,确定待检测的新闻信息是否为假新闻。本发明使得假新闻检测模型更加关注于新闻文本的有效信息,并通过将文本特征和摘要特征分别与图像特征进行跨通道信息增强,对于多模态之间的交互和关联捕捉效果较好;具有对多样性和复杂性较高的虚假新闻检测准确率高的技术效果。



1. 一种多模态假新闻检测方法,其特征在于,包括:

获取待检测的新闻信息,并利用抽取式算法获取所述新闻信息的摘要信息;其中,所述新闻信息包括新闻文本和新闻图像;

根据所述新闻文本、所述新闻图像和所述摘要信息,获取文本特征、图像特征和摘要特征;

将所述文本特征和摘要特征分别与图像特征进行跨通道信息增强,获取两个跨模态特征;

将所述两个跨模态特征输入多模态假新闻检测模型,确定所述待检测的新闻信息是否为假新闻。

2. 如权利要求1所述的多模态假新闻检测方法,其特征在于,将所述文本特征和摘要特征分别与图像特征进行跨通道信息增强,获取两个跨模态特征的方法包括,

将所述文本特征与图像特征作为第一特征向量,将摘要特征与图像特征作为第二特征向量,将所述第一特征向量和第二特征向量分别输入注意力网络;

将所述第一特征向量和第二特征向量分别映射为Q、K和V三个矩阵;利用第一特征向量的Q矩阵和第二特征向量的K矩阵获取相互注意力矩阵,并使用梯度信息生成掩码矩阵;

利用所述相互注意力矩阵、所述掩码矩阵分别和所述第一特征向量的V矩阵与第二特征向量的V矩阵,获取两个共同注意力向量;

将所获取的两个共同注意力向量和第一特征向量和第二特征向量相融合,获得两个跨模特征态特征。

3. 如权利要求2所述的多模态假新闻检测方法,其特征在于,

利用所述相互注意力矩阵、所述掩码矩阵和第二特征向量的V矩阵,获取共同注意力向量通过以下公式实现:

$$h_i = \text{softmax} \left(\frac{MLP^Q(input1).MLP^K(input2)^T}{\sqrt{d^h}} + mask \right) . MLP^V(input2)$$

其中, h_i 代表共同注意力向量, d^h 代表 h 特征的维度, $input1$ 和 $input2$ 代表输入的第一特征向量和第二特征向量, MLP 代表对输入的特征向量进行聚合并与初始特征向量融合后提取的两个模态的 q 、 k 、 v 值, $mask$ 代表使用梯度信息生成的掩码矩阵。

4. 如权利要求2所述的多模态假新闻检测方法,其特征在于,所述第一特征向量的获取方法包括,

利用多层感知机对文本特征中的每一个token计算和文本特征中所有token的相关性;

在文本特征中筛选和文本特征中所有token的相关性高于预设阈值的token;

将所筛选的文本特征中的token和图像特征进行模态聚合,获得第一特征向量。

5. 如权利要求1中所述的多模态假新闻检测方法,其特征在于,

将所述两个跨模态特征输入多模态假新闻检测模型之前,还包括将所述两个跨模态特征使用粗分类机制进行跨模态优化获得优化后的跨模态特征的步骤,包括,

多层感知机利用至少两个粗分类层对所述两个跨模态特征进行重复分类,获得优化后的跨模态特征;其中,

所述粗分类层为利用加权函数分别计算所述两个跨模态特征的权重,将跨模态特征和

对应的权重做内积运算后获得；

所述优化后的跨模态特征通过以下公式实现：
$$\omega_i = \frac{\exp(\alpha S_i + \beta)}{\sum_{i=1}^2 (\alpha e^{e_i} + \beta)} e_i ;$$

其中， ω_i 为第*i*个优化后的跨模态特征， e_i 为优化前的跨模态特征， α 和 β 均为超参数， S_i 为第*i*个跨模态特征的粗分类的得分。

6. 如权利要求1中所述的多模态假新闻检测方法，其特征在于，所述多模态假新闻检测模型利用损失函数进行约束训练后获得；

所述损失函数通过以下公式实现：
$$L = L_f + (L_{is} + L_{it}) / 2 ;$$

$$L_f = -[y \log \hat{y} + (1 - y) \log(1 - \hat{y})] ;$$

其中， L_f 表示跨模态特征优化后的分类损失， L_{is} 表示图像特征和摘要特征融合后粗分类的损失， L_{it} 表示图像特征和文本特征融合后粗分类的损失。

7. 如权利要求1中所述的多模态假新闻检测方法，其特征在于，根据所述新闻文本、所述新闻图像和所述新闻信息的摘要信息，获取文本特征、图像特征和摘要特征通过以下公式实现：

$$T = \{t_1, \dots, t_m\} = \text{BERT}(W)$$

$$I = \{i_1, \dots, i_m\} = \text{ResNet50}(R)$$

$$S = \{s_1, \dots, s_m\} = \text{BERT}(X)$$

其中， W 代表新闻文本， t 代表经过Bert模型提取后的文本特征， m 代表数据的序号， R 代表新闻图像数据， i 代表经过ResNet50模型提取后的图像特征， X 代表新闻信息的摘要信息， s 代表经过Bert模型提取后的摘要特征。

8. 一种多模态假新闻检测系统，其特征在于，包括：

摘要获取单元，用于获取待检测的新闻信息，并利用抽取式算法获取所述新闻信息的摘要信息；其中，所述新闻信息包括新闻文本和新闻图像；

特征获取单元，用于根据所述新闻文本、所述新闻图像和所述摘要信息，获取文本特征、图像特征和摘要特征；

特征增强单元，用于将所述文本特征和摘要特征分别与图像特征进行跨通道信息增强，获取两个跨模态特征；

检测单元，用于将所述两个跨模态特征输入多模态假新闻检测模型，确定所述待检测的新闻信息是否为假新闻。

9. 一种电子设备，其特征在于，所述电子设备包括：

至少一个处理器；以及，

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行如权利要求1至7中任一所述的多模态假新闻检测方法中的步骤。

10. 一种计算机可读存储介质，存储有计算机程序，其特征在于，所述计算机程序被处理器执行时实现如权利要求1至7中任一所述的多模态假新闻检测方法。

多模态假新闻检测方法及系统

技术领域

[0001] 本发明属于电子信息技术领域,具体涉及一种多模态假新闻检测方法及系统。

背景技术

[0002] 基于人工智能技术的发展,人们可以轻松地使用图像、视频和音频生成以假乱真的虚假新闻消息。随着互联网的普及和社交媒体的兴起,虚假新闻泛滥已经成为一个严重的问题。虚假新闻检测是指利用计算机视觉和自然语言处理技术从给定的信息来判断新闻的真假,通过利用多模态如文本、图像、视频等以提高虚假新闻检测的准确性和可靠性。因此,多模态虚假新闻检测已经成为研究和应用的热点之一。

[0003] 在现有技术中,多模态虚假新闻检测存在的弊端如下:

1) 现有的虚假新闻检测模型使用拼接或者点乘方式实现多模态特征的融合;虽然具有操作简单、不需要额外参数和可以保留模态的原始信息的优点;但是,对于多模态之间的交互和关联捕捉效果欠佳;2) 现有的虚假新闻检测模型在识别多样性和复杂性较高的虚假新闻时,有效性较差。

[0004] 因此,亟需一种虚假新闻识别效率高的多模态假新闻检测方法及系统。

发明内容

[0005] 本发明提供一种多模态假新闻检测方法、系统、电子设备及存储介质,用以克服现有技术中存在的至少一个技术问题。

[0006] 为实现上述目的,本发明提供一种多模态假新闻检测方法,方法包括:

获取待检测的新闻信息,并利用抽取式算法获取新闻信息的摘要信息;其中,新闻信息包括新闻文本和新闻图像;

根据新闻文本、新闻图像和摘要信息,获取文本特征、图像特征和摘要特征;

将文本特征和摘要特征分别与图像特征进行跨通道信息增强,获取两个跨模态特征;

将两个跨模态特征输入多模态假新闻检测模型,确定待检测的新闻信息是否为假新闻。

[0007] 进一步的,优选的方法为,将文本特征和摘要特征分别与图像特征进行跨通道信息增强,获取两个跨模态特征的方法包括,

将文本特征与图像特征作为第一特征向量,将摘要特征与图像特征作为第二特征向量,将第一特征向量和第二特征向量分别输入注意力网络;

将第一特征向量和第二特征向量分别映射为Q、K和V三个矩阵;利用第一特征向量的Q矩阵和第二特征向量的K矩阵获取相互注意力矩阵,并使用梯度信息生成掩码矩阵;

利用相互注意力矩阵、掩码矩阵分别和第一特征向量的V矩阵与第二特征向量的V矩阵,获取两个共同注意力向量;

将所获取的两个共同注意力向量和第一特征向量和第二特征向量相融合,获得两

个跨模特征态特征。

[0008] 进一步的,优选的方法为,利用相互注意力矩阵、掩码矩阵和第二特征向量的V矩阵,获取共同注意力向量通过以下公式实现:

$$h_i = \text{softmax} \left(\frac{MLP^Q(input1).MLP^K(input2)^T}{\sqrt{d^h}} + \text{mask} \right) .MLP^V(input2)$$

[0009] 其中, h_i 代表共同注意力向量, d^h 代表 h 特征的维度, $input1$ 和 $input2$ 代表输入的第一特征向量和第二特征向量, MLP 代表对输入的特征向量进行聚合并与初始特征向量融合后提取的两个模态的 q,k,v 值, $mask$ 代表使用梯度信息生成的掩码矩阵。

[0010] 进一步的,优选的方法为,第一特征向量的获取方法包括,

利用多层感知机对文本特征中的每一个token计算和文本特征中所有token的相关性;

在文本特征中筛选和文本特征中所有token的相关性高于预设阈值的token;

将所筛选的文本特征中的token和图像特征进行模态聚合,获得第一特征向量。

[0011] 进一步的,优选的方法为,将两个跨模态特征输入多模态假新闻检测模型之前,还包括将两个跨模态特征使用粗分类机制进行跨模态优化获得优化后的跨模态特征的步骤,包括,

多层感知机利用至少两个粗分类层对两个跨模态特征进行重复分类,获得优化后的跨模态特征;其中,

粗分类层为利用加权函数分别计算两个跨模态特征的权重,将跨模态特征和对应的权重做内积运算后获得;

优化后的跨模态特征通过以下公式实现: $\omega_i = \frac{\exp(\alpha S_i + \beta)}{\sum_{i=1}^2 (\alpha e_i + \beta)} e_i$; 其中, ω_i 为

第 i 个优化后的跨模态特征, e_i 为优化前的跨模态特征, α 和 β 均为超参数, S_i 为第 i 个跨模态特征的粗分类的得分。

[0012] 进一步的,优选的方法为,多模态假新闻检测模型利用损失函数进行约束训练后获得;

损失函数通过以下公式实现:

$$L = L_f + (L_{is} + L_{it}) / 2;$$

$$L_f = -[y \log \hat{y} + (1 - y) \log(1 - \hat{y})];$$

其中, L_f 表示跨模态特征优化后的分类损失, L_{is} 表示图像特征和摘要特征融合后粗分类的损失, L_{it} 表示图像特征和文本特征融合后粗分类的损失。

[0013] 进一步的,优选的方法为,根据新闻文本、新闻图像和新闻信息的摘要信息,获取文本特征、图像特征和摘要特征通过以下公式实现:

$$T = \{t_1, \dots, t_m\} = \text{BERT}(W)$$

$$I = \{i_1, \dots, i_m\} = \text{ResNet50}(R)$$

$$S = \{s_1, \dots, s_m\} = \text{BERT}(X)$$

其中, W 代表新闻文本, t 代表经过Bert模型提取后的文本特征, m 代表数据的序号, R 代表新闻图像数据, i 代表经过ResNet50模型提取后的图像特征, X 代表新闻信息的摘要信息, s 代表经过Bert模型提取后的摘要特征。

[0014] 为了解决上述问题,本发明还提供一种多模态假新闻检测系统,包括:

摘要获取单元,用于获取待检测的新闻信息,并利用抽取式算法获取新闻信息的摘要信息;其中,新闻信息包括新闻文本和新闻图像;

特征获取单元,用于根据新闻文本、新闻图像和摘要信息,获取文本特征、图像特征和摘要特征;

特征增强单元,用于将文本特征和摘要特征分别与图像特征进行跨通道信息增强,获取两个跨模态特征;

检测单元,用于将两个跨模态特征输入多模态假新闻检测模型,确定待检测的新闻信息是否为假新闻。

[0015] 为了解决上述问题,本发明还提供一种电子设备,电子设备包括:

至少一个处理器;以及,

与所述至少一个处理器通信连接的存储器;其中,

所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行如上述的多模态假新闻检测方法中的步骤。

[0016] 本发明还保护一种计算机可读存储介质,存储有计算机程序,计算机程序被处理器执行时实现如上述的多模态假新闻检测方法。

[0017] 本发明的一种多模态假新闻检测方法、系统、电子设备以及存储介质,获取待检测的新闻信息,并利用抽取式算法获取新闻信息的摘要信息;其中,新闻信息包括新闻文本和新闻图像;根据新闻文本、新闻图像和摘要信息,获取文本特征、图像特征和摘要特征;将文本特征和摘要特征分别与图像特征进行跨通道信息增强,获取两个跨模态特征;将两个跨模态特征输入多模态假新闻检测模型,确定待检测的新闻信息是否为假新闻。本发明具有有益效果如下:通过抽取待检测的新闻信息的摘要特征,使得模型更加关注于新闻文本的有效信息,减少了噪声对模型的影响;并通过将文本特征和摘要特征分别与图像特征进行跨通道信息增强,对于多模态之间的交互和关联捕捉效果较好;在识别多样性和复杂性较高的虚假新闻场景中,有效性较高;本发明的多模态假新闻检测模型在更加关注于文本的强信息的同时模型的鲁棒性得到了增强。

附图说明

[0018] 图1为根据本发明实施例的多模态假新闻检测方法的流程示意图;

图2为根据本发明实施例的多模态假新闻检测方法的原理示意图;

图3为根据本发明实施例的假新闻示例图;

图4为根据本发明实施例的多模态假新闻检测系统的逻辑结构框图;

图5根据本发明实施例的实现多模态假新闻检测方法的电子设备的内部结构示意图。

[0019] 本发明目的的实现、功能特点及优点将结合实施例,参照附图做进一步说明。

具体实施方式

[0020] 应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。

[0021] 本申请实施例可以基于人工智能技术和计算机视觉技术对相关的数据进行获取和处理。其中,人工智能(Artificial Intelligence, AI)是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能,感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。换句话说,人工智能是计算机科学的一个综合技术,它企图了解智能的实质,并生产出一种新的能以人类智能相似的方式做出反应的智能机器。人工智能也就是研究各种智能机器的设计原理与实现方法,使机器具有感知、推理与决策的功能。人工智能技术是一门综合学科,涉及领域广泛,既有硬件层面的技术也有软件层面的技术。人工智能基础技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理技术、操作/交互系统、机电一体化等技术。人工智能软件技术主要包括计算机视觉技术、语音处理技术、自然语言处理技术以及机器学习/深度学习等几大方向。

[0022] 具体的,作为示例,图1为本发明一实施例提供的多模态假新闻检测方法的流程示意图。参照图1所示,本发明提供一种多模态假新闻检测方法,该方法可以由一个装置执行,该装置可以由软件和/或硬件实现。多模态假新闻检测方法,包括步骤S110~ S140。

[0023] S110、获取待检测的新闻信息,并利用抽取式算法获取新闻信息的摘要信息;其中,新闻信息包括新闻文本和新闻图像。

[0024] 传统多模态虚假新闻检测往往利用新闻文本和新闻图像进行分类,但这些文本信息往往含有与新闻主题无关的噪声,本发明的创新在于利用抽取式算法为每条新闻数据生成一个摘要,减少噪声对模型的影响,同时保留原始的新闻文本信息,避免有效的上下文内容被去除。在具体的实施过程中,通过抽取式摘要算法生成对应新闻的摘要。示例如下:对于Weibo数据集由于其新闻内容较短,该数据集的摘要控制在50字以内;对于FakeNewsNet,由于其数据内容较长,该数据集摘要控制在200字以内。抽取式算法是从原文中摘取完整的句子以概括文章的主要内容,从而减少不必要的噪声对模型的影响。

[0025] 在一个具体的实施例中,新闻文本为“[鲸现场:品牌+电商开启汽车后服务市场蓝海]美国固铂轮胎与专业汽车后市场o2o平台----汽车超人(隶属金固股份)战略合作签署仪式在上海举行。Cooper美国固铂轮胎将借助汽车超人的平台优势,实现产品在线销售与线下渠道布局完美结合。汽车后服务或成为汽车领域另一大蓝海。”通过利用抽取式算法获取新闻信息的摘要信息后,摘要信息为“汽车后服务或成为汽车领域另一大蓝海”。

[0026] S120、根据新闻文本、新闻图像和摘要信息,获取文本特征、图像特征和摘要特征。

[0027] 具体地说,需要抽取两种文本特征和一个图像特征。对于文本特征而言,为了提取更加全面的层次信息,在本实施例中以Bert模型作为文本的编码器,该模型的hidden_states输出共有13层,其中,第1层为embedding层,另外12层为各层隐藏状态的输出,使用坐标为第3、6、9、12的隐藏状态,每个输出特征的形状是(batch_size, sequence_length, hidden_size),特征通道为768,为了降低模型的参数同时与图像的特征维度保持一致,每一层的输出后使用了一个卷积层将元素形状控制为(batch_size, 64),最后将得到的特征进行融合得到文本的最终表示 t_1 ;对于图像特征,本实施例以ResNet50模型作为图像的编码器,对模型进行改进提取Layer1、Layer2、Layer3、Layer4层的特征,并在每个Layer层的后面使用卷积改变特征形状并降低特征维度,减少模型参数的数量,使元素的形状控制为

(batch_size,64),最后将得到的特征进行融合。多尺度特征提取网络还可以包括新闻文本特征提取子网络、新闻图像特征提取子网络、新闻摘要特征提取子网络,并通过该新闻文本特征提取子网络提取新闻文本特征、通过新闻图像特征提取子网络提取新闻图像特征、通过新闻摘要特征提取子网络提取摘要特征。

[0028] 也就是说,在具体的实施过程中,根据新闻文本、新闻图像和新闻信息的摘要信息,获取文本特征、图像特征和摘要特征通过以下公式实现:

$$\begin{aligned} T &= \{t_1, \dots, t_m\} = \text{BERT}(W) \\ I &= \{i_1, \dots, i_m\} = \text{ResNet50}(R) \\ S &= \{s_1, \dots, s_m\} = \text{BERT}(X) \end{aligned}$$

其中, W 代表新闻文本, t 代表经过Bert模型提取后的文本特征, m 代表数据的序号, R 代表新闻图像数据, i 代表经过ResNet50模型提取后的图像特征, X 代表新闻信息的摘要信息, s 代表经过Bert模型提取后的摘要特征。

[0029] 总之,将新闻文本、新闻图像和摘要信息分别输入多尺度编码器,从多个层提取模态信息,通过对模态不同粒度的处理可以观察到不同的特征;也就是说通过多尺度编码器获取模态更全面的表征信息。针对现有的假新闻检测模型中语义信息丢失问题,本发明采用多尺度编码器的方法提取新闻文本和新闻图片的低维、中维、高维特征,使多模态假新闻检测模型获得更全面的表征信息。

[0030] S130、将文本特征和摘要特征分别与图像特征进行跨通道信息增强,获取两个跨模态特征。即,将通过多尺度编码器获取模态更全面的表征信息,引入注意力模块进行跨模态特征的交互,通过跨通道信息增强模块解决多模态特征信息交互融合的问题。

[0031] 具体地说,通过改进共同注意力设计了跨通道信息增强模块,在本实施例中跨通道信息增强模块由两个信息增强模块(MS blocks)并行排列而成。

[0032] 将文本特征和摘要特征分别与图像特征进行跨通道信息增强,获取两个跨模态特征的方法包括步骤S131~ S134。

[0033] S131、将文本特征与图像特征作为第一特征向量,将摘要特征与图像特征作为第二特征向量,将第一特征向量和第二特征向量分别输入注意力网络。

[0034] 第一特征向量的获取方法包括,S1311、利用多层感知机对文本特征中的每一个token计算和文本特征中所有token的相关性;S1312、在文本特征中筛选和文本特征中所有token的相关性高于预设阈值的token;S1313、将所筛选的文本特征中的token和图像特征进行模态聚合,获得第一特征向量。其中,多层感知机是一种前向结构的人工神经网络,包含输入层、输出层及多个隐藏层融合特征描述子。

[0035] 也就是说,在输入到共同注意力之前,使用多层感知机计算每个token的重要性,并根据重要性分数将模态聚合为一个表示,以达到使模型只关注有用的特征信息的目的。

[0036] S132、将第一特征向量和第二特征向量分别映射为Q、K和V三个矩阵;利用第一特征向量的Q矩阵和第二特征向量的K矩阵获取相互注意力矩阵,并使用梯度信息生成掩码矩阵。S133、利用相互注意力矩阵、掩码矩阵分别和第一特征向量的V矩阵与第二特征向量的V矩阵,获取两个共同注意力向量。

[0037] 也就是说,将两个输入的特征张量(即将文本特征与图像特征作为第一特征向量,将摘要特征与图像特征作为第二特征向量,第一特征向量和第二特征向量两个特征向量)

分别映射为query、key和value(即Q、K和V三个矩阵),用于共同注意力。具体地说,利用一个特征向量输入的query和另一个特征向量输入的key计算一个相互注意力矩阵,表示两个输入的特征向量之间的相关性;然后通过使用梯度信息生成两个特征向量的掩码矩阵,以实现只关注两个输入的特征向量的对模型输出有贡献的部分。进而利用相互注意力矩阵、掩码矩阵和两个特征向量输入的value计算两个共同注意力向量,以表示两个输入的特征向量对对方的关注;最后,将共同注意力向量与两个输入的特征张量相融合,完成通道级的多模态信息增强,得到最终的输出。

[0038] 在具体的实施过程中,

$$Q_i = QW_i^Q, K_i = KW_i^K, V_i = VW_i^V$$

W_i^Q, W_i^K, W_i^V 代表第*i*个头的投影矩阵,多头自注意函数的计算过程可以表示如下:

$$Co(Q, K, V) = hW^0$$

其中, $h = h1 \oplus h2 \oplus \dots \oplus h3$; h_i 代表经过加权函数的共同注意力向量, \oplus 代表拼接特征向量, $W^0 \in R^{md_h \times 1}$ 。

[0039] 利用相互注意力矩阵、掩码矩阵和第二特征向量的V矩阵,获取共同注意力向量通过以下公式实现:

$$h_i = \text{softmax} \left(\frac{MLP^Q(input1).MLP^K(input2)^T}{\sqrt{d^h}} + \text{mask} \right) \cdot MLP^V(input2);$$

其中, h_i 代表

共同注意力向量, d^h 代表*h*特征的维度, $input1$ 和 $input2$ 代表输入的第一特征向量和第二特征向量, MLP 代表对输入的特征向量进行聚合并与初始特征向量融合后提取的两个模态的 q, k, v 值, $mask$ 代表使用梯度信息生成的掩码矩阵。

[0040] 同样的,对于获取第一特征向量对应的共同注意力向量的方式也是参照上述公式,只是将上述第二特征向量的V矩阵替换为第一特征向量的V矩阵即可,具体的实现方式不再赘述。

[0041] S134、将所获取的两个共同注意力向量和第一特征向量和第二特征向量相融合,获得两个跨模特征态特征。

[0042] 综上,将通过多尺度编码器获取模态更全面的表征信息,通过引入共同注意力模块进行跨模态特征的交互解决多模态特征信息交互融合的问题。然后,需要将融合后的多模态特征信息进行分类。

[0043] S140、将两个跨模态特征输入多模态假新闻检测模型,确定待检测的新闻信息是否为假新闻。

[0044] 将新闻文本、摘要信息分别与新闻图像进行跨通道信息增强会得到两个跨模态特征,并将两个跨模态特征进行交互融合后,本发明创新性的采用粗分类方法对两个跨模态特征进行优化。具体地说,将多模态新闻分类任务分解为多个层次,前几层粗分类主要学习两个跨模态特征对最终结果的贡献度,将跨模态特征进行多层次优化从而提高分类的准确性和效率,最后一层细分类作为最终的分类结果。

[0045] 具体地说,将两个跨模态特征输入多模态假新闻检测模型之前,还包括将两个跨模态特征使用粗分类机制进行跨模态优化获得优化后的跨模态特征的步骤,包括,多层感

知机利用至少两个粗分类层对两个跨模态特征进行重复分类,获得优化后的跨模态特征;其中,粗分类层为利用加权函数分别计算两个跨模态特征的权重,将跨模态特征和对应的权重做内积运算后获得。即,将两个跨模态特征通过多层感知机进行分类后,使用sigmoid进行激活,得出该特征属于真新闻的概率值,同时从多视图的角度出发将两个特征向量看作共同决定分类结果的两个视图使用加权函数重新计算权重,将得到的权重与特征进行相乘,完成一层跨模态特征的优化,如此重复粗分类层完成多层次特征优化使得模型在最终分类中更准确的识别出新闻的真伪。

[0046] 优化后的跨模态特征通过以下公式实现:

$$\omega_i = \frac{\exp(\alpha S_i + \beta)}{\sum_{i=1}^2 (\alpha e^{e_i} + \beta)} e_i ; \text{其中, } \omega_i \text{ 为第 } i \text{ 个优化后的跨模态特征, } e_i \text{ 为优化前的}$$

跨模态特征, α 和 β 均为超参数, S_i 为第 i 个跨模态特征的粗分类的得分。

[0047] 也就是说,在粗分类模块,使用MLP(多层感知机)预先将融合特征进行聚合分类,得到的粗分类结果为属于真新闻的概率。即从多视图的角度出发,将两个跨模态特征看作是影响结果的两个视图,使用加权函数重新计算两者的权重,权重代表对于结果的贡献度,将权重与输入特征做内积运算,从而实现一层跨模态特征的优化;如此重复粗分类层完成多层次特征优化使得模型在最终分类中更准确的识别出新闻的真伪;在进行粗分类后再次经过跨通道信息增强模块进行融合,最终完成跨模态特征的优化。本发明通过设置粗分类机制使模型在最终分类前进行跨模态特征的优化,实现了增强模型的鲁棒性的技术效果。

[0048] 在具体的实施过程中,在粗分类层完成多层次特征优化后,在最后一层细分类作为最终的分类结果。也就是说,本发明将优化后的跨模态特征作为最后的分类特征,并使用交叉熵函数计算分类损失,同时在最后一次粗分类分别计算出两个跨模态特征的损失,联合三个损失约束网络进行模型训练,模型可以建模多模态信息之间和之内的关系,同时建模文本内容和图像内容各自的层次语义关系,分类器根据网络输出的特征向量,将其映射为一个概率分布,表示该新闻属于每个类别的可能性。

[0049] 多模态假新闻检测模型利用损失函数进行约束训练后获得;

损失函数通过以下公式实现:

$$L = L_f + (L_{is} + L_{it}) / 2;$$

$L_f = -[y \log \hat{y} + (1 - y) \log(1 - \hat{y})]$;其中, L_f 表示跨模态特征优化后的分类损失, L_{is} 表示图像特征和摘要特征融合后粗分类的损失, L_{it} 表示图像特征和文本特征融合后粗分类的损失。

[0050] 综上,本发明的多模态假新闻检测方法通过抽取式摘要算法生成新闻的摘要,解决了多模态信息噪声较多的问题;通过多尺度提取多模态的特征,解决了传统多模态虚假新闻检测提取特征的问题;利用跨模态融合模块解决了多模态特征信息交互的问题;使用粗分类机制优化了跨模态特征,最终实现了假新闻的有效检测。

[0051] 图2为根据本发明实施例的多模态假新闻检测方法的原理示意图;如图2所示,在具体的实施过程中,第一步:利用抽取式算法为每条新闻数据生成一个摘要,减少噪声对模型的影响,同时保留原始的新闻文本信息,避免有效的上下文内容被去除。第二步:将新闻

文本、新闻图像和摘要信息分别输入多尺度编码器,从多个层提取模态信息,通过对模态不同粒度的处理可以观察到不同的特征;采用多尺度编码器的方法提取新闻文本和新闻图片的低维、中维、高维特征,使多模态假新闻检测模型获得更全面的表征信息,即并通过文本编码器提取新闻文本特征、通过图像编码器提取新闻图像特征、通过文本编码器提取摘要特征。第三步:通过引入共同注意力模块进行跨模态特征的交互解决多模态特征信息交互融合的问题,跨通道信息增强模块由两个信息增强模块(MS blocks)并行排列而成;即将摘要与图像进行融合得到跨模态增强特征;将文本与图像进行融合得到另一个跨模态增强特征。第四步:将两者放到粗分类模块进行特征优化后再进行融合,这样使模型更加关注于文本的强信息同时增强了模型的鲁棒性。其中,跨通道信息增强模块同样由两个信息增强模块(MS blocks)并行排列而成。最后一步:将优化后的跨模态特征作为最后的分类特征,并使用交叉熵函数计算分类损失,同时在最后一次粗分类分别计算出两个跨模态特征的损失,联合三个损失约束网络进行模型训练,模型可以建模多模态信息之间和之内的关系,同时建模文本内容和图像内容各自的层次语义关系,分类器根据网络输出的特征向量,将其映射为一个概率分布,表示该新闻属于每个类别的可能性。

[0052] 图3为利用本发明的多模态假新闻检测方法识别的虚假新闻示例;其中,图3-A的待识别的原始新闻文本为战争中的越南兄妹;经过本发明的多模态假新闻检测方法进行检测后发现这是两个越南兄弟姐妹的真实照片,但这张照片是2015年尼泊尔地震期间拍摄的;因此,判定为假新闻。图3-B的待识别的原始新闻文本为飓风桑迪期间马路上出现鲨鱼;经过本发明的多模态假新闻检测方法进行检测后发现这是拼接后获取的鲨鱼图片,因此,判定为假新闻。图3-C的待识别的原始新闻文本为2015年3月20日出现日食,经过本发明的多模态假新闻检测方法进行检测后发现上述图片是人工智能生成的艺术品,因此,判定为假新闻。

[0053] 利用本发明的多模态假新闻检测方法在Weibo、FakeNewsNet等各个数据集上进行检测,结果表1所示:

表1:本发明的多模态假新闻检测模型在各个数据集中检测结果表

数据集	模型	准确率	假新闻			真新闻		
			查准率	召回率	F1分数	查准率	召回率	F1分数
Weibo	textual	0.643	0.662	0.578	0.617	0.609	0.685	0.647
	visual	0.608	0.61	0.605	0.607	0.607	0.611	0.609
	VQA	0.736	0.797	0.634	0.706	0.695	0.838	0.76
	Neural Talk	0.726	0.794	0.713	0.692	0.684	0.84	0.754
	SAFE	0.763	0.833	0.659	0.736	0.717	0.868	0.785
	att-RNN	0.772	0.797	0.713	0.692	0.684	0.84	0.754
	EANN-	0.795	0.827	0.697	0.756	0.752	0.863	0.804
	EANN	0.827	NA	NA	NA	NA	NA	NA
	MVAE-	0.743	NA	NA	NA	NA	NA	NA
	MVAE	0.824	0.854	0.769	0.809	0.802	0.875	0.837
	HMCAN	0.885	0.92	0.845	0.881	0.856	0.926	0.89
SpotFake	0.892	0.902	0.964	0.932	0.847	0.656	0.739	
our	0.913	0.78	0.81	0.79	0.95	0.94	0.95	
GossipCop	RoBERTa-MWSS	0.8	-	-	-	0.8	-	-
	SAFE	0.838	-	-	-	0.857	0.937	0.895
	Spotfake+	0.856	-	-	0.864	-	-	-
	LSTM-ATT	0.842	-	-	-	0.839	0.842	0.821
	our	0.864	0.89	0.96	0.92	0.72	0.48	0.58
PolitiFact	RoBERTa-MWSS	0.82	-	-	-	0.82	-	-
	SAFE	0.874	-	-	-	0.889	0.903	0.896
	Spotfake+	0.846	-	-	-	-	-	-
	LSTM-ATT	0.832	-	-	-	0.836	0.832	0.829
	our	0.92	0.92	0.97	0.95	0.92	0.79	0.85

[0054] 通过观察表1可见,本发明的多模态假新闻检测方法在各个数据集上的检测的平均准确率均达到了86%以上。尤其是在Weibo数据集上的真新闻的查准率为0.95远远超过其他模型的查准率。

[0055] 综上,本发明的多模态假新闻检测方法,针对新闻文本含有大量无用信息使用多层感知机首先减少噪声的干扰,并将提取的上下文有效信息与图像进行融合得到跨模态辅助特征;其次提取的新闻摘要极大程度上去除了新闻文本的噪声信息,将摘要与图像进行融合得到跨模态增强特征,并将两者放到粗分类模块进行特征优化后再进行融合,最终获得更加关注于文本的强信息以及强鲁棒性的多模态假新闻检测模型。

[0056] 与上述多模态假新闻检测方法相对应,本发明还提供一种多模态假新闻检测系统。图4示出了根据本发明实施例的多模态假新闻检测系统的功能模块。

[0057] 如图4所示,本发明提供的多模态假新闻检测系统400可以安装于电子设备中。根据实现的功能,所述多模态假新闻检测系统400可以包括摘要获取单元410、特征获取单元420、特征增强单元430和检测单元440。本发明所述单元也可以称之为模块,指的是一种能够被电子设备的处理器所执行,并且能够完成某一固定功能的一系列计算机程序段,其存储在电子设备的存储器中。

[0058] 在本实施例中,关于各模块/单元的功能如下:

摘要获取单元410,用于获取待检测的新闻信息,并利用抽取式算法获取新闻信息的摘要信息;其中,新闻信息包括新闻文本和新闻图像;

特征获取单元420,用于根据新闻文本、新闻图像和摘要信息,获取文本特征、图像特征和摘要特征;

特征增强单元430,用于将文本特征和摘要特征分别与图像特征进行跨通道信息增强,获取两个跨模态特征;

检测单元440,用于将两个跨模态特征输入多模态假新闻检测模型,确定待检测的新闻信息是否为假新闻。

[0059] 本发明所提供的上述多模态假新闻检测系统的更为具体的实现方式,均可以参照上述对多模态假新闻检测方法的实施例表述,在此不再一一列举。

[0060] 本发明所提供的上述多模态假新闻检测系统,通过获取待检测的新闻信息,并利用抽取式算法获取新闻信息的摘要信息;其中,新闻信息包括新闻文本和新闻图像;根据新闻文本、新闻图像和摘要信息,获取文本特征、图像特征和摘要特征;将文本特征和摘要特征分别与图像特征进行跨通道信息增强,获取两个跨模态特征;将两个跨模态特征输入多模态假新闻检测模型,确定待检测的新闻信息是否为假新闻。本发明具有有益效果如下:通过抽取待检测的新闻信息的摘要特征,使得模型更加关注于新闻文本的有效信息,减少了噪声对模型的影响;并通过将文本特征和摘要特征分别与图像特征进行跨通道信息增强,对于多模态之间的交互和关联捕捉效果较好;在识别多样性和复杂性较高的虚假新闻场景中,有效性较高;本发明的多模态假新闻检测模型在更加关注于文本的强信息的同时模型的鲁棒性得到了增强。

[0061] 如图5所示,本发明提供一种多模态假新闻检测方法的电子设备5。

[0062] 该电子设备5可以包括处理器50、存储器51和总线,还可以包括存储在存储器51中并可在所述处理器50上运行的计算机程序,如基于生成对抗网络的腹腔镜影像除烟程序52。

[0063] 其中,所述存储器51至少包括一种类型的可读存储介质,所述可读存储介质包括闪存、移动硬盘、多媒体卡、卡型存储器(例如:SD或DX存储器等)、磁性存储器、磁盘、光盘等。所述存储器51在一些实施例中可以是电子设备5的内部存储单元,例如该电子设备5的移动硬盘。所述存储器51在另一些实施例中也可以是电子设备5的外部存储设备,例如电子设备5上配备的插接式移动硬盘、智能存储卡(Smart Media Card, SMC)、安全数字(Secure Digital, SD)卡、闪存卡(Flash Card)等。进一步地,所述存储器51还可以既包括电子设备5的内部存储单元也包括外部存储设备。所述存储器51不仅可以用于存储安装于电子设备5的应用软件及各类数据,例如基于生成对抗网络的腹腔镜影像除烟程序的代码等,还可以用于暂时地存储已经输出或者将要输出的数据。

[0064] 所述处理器50在一些实施例中可以由集成电路组成,例如可以由单个封装的集成电路所组成,也可以是由多个相同功能或不同功能封装的集成电路所组成,包括一个或者多个中央处理器(Central Processing unit, CPU)、微处理器、数字处理芯片、图形处理器及各种控制芯片的组合等。所述处理器50是所述电子设备的控制核心(Control Unit),利用各种接口和线路连接整个电子设备的各个部件,通过运行或执行存储在所述存储器51内的程序或者模块(例如基于生成对抗网络的腹腔镜影像除烟程序等),以及调用存储在所述存储器51内的数据,以执行电子设备5的各种功能和处理数据。

[0065] 所述总线可以是外设部件互连标准(peripheral component interconnect,简称PCI)总线或扩展工业标准结构(extended industry standard architecture,简称EISA)总线等。该总线可以分为地址总线、数据总线、控制总线等。所述总线被设置为实现所述存储器51以及至少一个处理器50等之间的连接通信。

[0066] 图5仅示出了具有部件的电子设备,本领域技术人员可以理解的是,图5示出的结

构并不构成对所述电子设备5的限定,可以包括比图示更少或者更多的部件,或者组合某些部件,或者不同的部件布置。

[0067] 例如,尽管未示出,所述电子设备5还可以包括给各个部件供电的电源(比如电池),优选地,电源可以通过电源管理装置与所述至少一个处理器50逻辑相连,从而通过电源管理装置实现充电管理、放电管理、以及功耗管理等功能。电源还可以包括一个或一个以上的直流或交流电源、再充电装置、电源故障检测电路、电源转换器或者逆变器、电源状态指示器等任意组件。所述电子设备5还可以包括多种传感器、蓝牙模块、Wi-Fi模块等,在此不再赘述。

[0068] 进一步地,所述电子设备5还可以包括网络接口,可选地,所述网络接口可以包括有线接口和/或无线接口(如WI-FI接口、蓝牙接口等),通常用于在该电子设备5与其他电子设备之间建立通信连接。

[0069] 可选地,该电子设备5还可以包括用户接口,用户接口可以是显示器(Display)、输入单元(比如键盘(Keyboard)),可选地,用户接口还可以是标准的有线接口、无线接口。可选地,在一些实施例中,显示器可以是LED显示器、液晶显示器、触控式液晶显示器以及OLED(Organic Light-Emitting Diode,有机发光二极管)触摸器等。其中,显示器也可以适当的称为显示屏或显示单元,用于显示在电子设备5中处理的信息以及用于显示可视化的用户界面。

[0070] 应该了解,所述实施例仅为说明之用,在专利申请范围上并不受此结构的限制。

[0071] 所述电子设备5中的所述存储器51存储的基于生成对抗网络的腹腔镜影像除烟程序52是多个指令的组合,在所述处理器50中运行时,可以实现:S110、获取待检测的新闻信息,并利用抽取式算法获取新闻信息的摘要信息;其中,新闻信息包括新闻文本和新闻图像。S120、根据新闻文本、新闻图像和摘要信息,获取文本特征、图像特征和摘要特征;S130、将文本特征和摘要特征分别与图像特征进行跨通道信息增强,获取两个跨模态特征;S140、将两个跨模态特征输入多模态假新闻检测模型,确定待检测的新闻信息是否为假新闻。

[0072] 具体地,所述处理器50对上述指令的具体实现方法可参考图1对应实施例中相关步骤的描述,在此不赘述。需要强调的是,为进一步保证上述基于生成对抗网络的腹腔镜影像除烟程序的私密和安全性,上述基于生成对抗网络的腹腔镜影像除烟程序存储于本服务器集群所处区块链的节点中。

[0073] 进一步地,所述电子设备5集成的模块/单元如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。所述计算机可读取介质可以包括:能够携带所述计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器(ROM,Read-Only Memory)。

[0074] 本发明实施例还提供一种计算机可读存储介质,所述存储介质可以是非易失性的,也可以是易失性的,所述存储介质存储有计算机程序,所述计算机程序被处理器执行时实现:S110、获取待检测的新闻信息,并利用抽取式算法获取新闻信息的摘要信息;其中,新闻信息包括新闻文本和新闻图像。S120、根据新闻文本、新闻图像和摘要信息,获取文本特征、图像特征和摘要特征;S130、将文本特征和摘要特征分别与图像特征进行跨通道信息增强,获取两个跨模态特征;S140、将两个跨模态特征输入多模态假新闻检测模型,确定待检测的新闻信息是否为假新闻。

[0075] 具体地,所述计算机程序被处理器执行时具体实现方法可参考实施例多模态假新闻检测方法中相关步骤的描述,在此不赘述。

[0076] 在本发明所提供的几个实施例中,应该理解到,所揭露的设备,装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如,所述模块的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式。

[0077] 所述作为分离部件说明的模块可以是或者也可以不是物理上分开的,作为模块显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。

[0078] 另外,在本发明各个实施例中的各功能模块可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用硬件加软件功能模块的形式实现。

[0079] 对于本领域技术人员而言,显然本发明不限于上述示范性实施例的细节,而且在不背离本发明的精神或基本特征的情况下,能够以其他的具体形式实现本发明。

[0080] 因此,无论从哪一点来看,均应将实施例看作是示范性的,而且是非限制性的,本发明的范围由所附权利要求而不是上述说明限定,因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本发明内。不应将权利要求中的任何附关联图标记视为限制所涉及的权利要求。

[0081] 本发明所指区块链是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链(Blockchain),本质上是一个去中心化的数据库,是一串使用密码学方法相关联产生的数据块,每一个数据块中包含了一批网络交易的信息,用于验证其信息的有效性(防伪)和生成下一个区块。区块链可以包括区块链底层平台、平台产品服务层以及应用服务层等。

[0082] 此外,显然“包括”一词不排除其他单元或步骤,单数不排除复数。系统权利要求中陈述的多个单元或装置也可以由一个单元或装置通过软件或者硬件来实现。第二等词语用来表示名称,而并不表示任何特定的顺序。

[0083] 最后应说明的是,以上实施例仅用以说明本发明的技术方案而非限制,尽管参照较佳实施例对本发明进行了详细说明,本领域的普通技术人员应当理解,可以对本发明的技术方案进行修改或等同替换,而不脱离本发明技术方案的精神和范围。

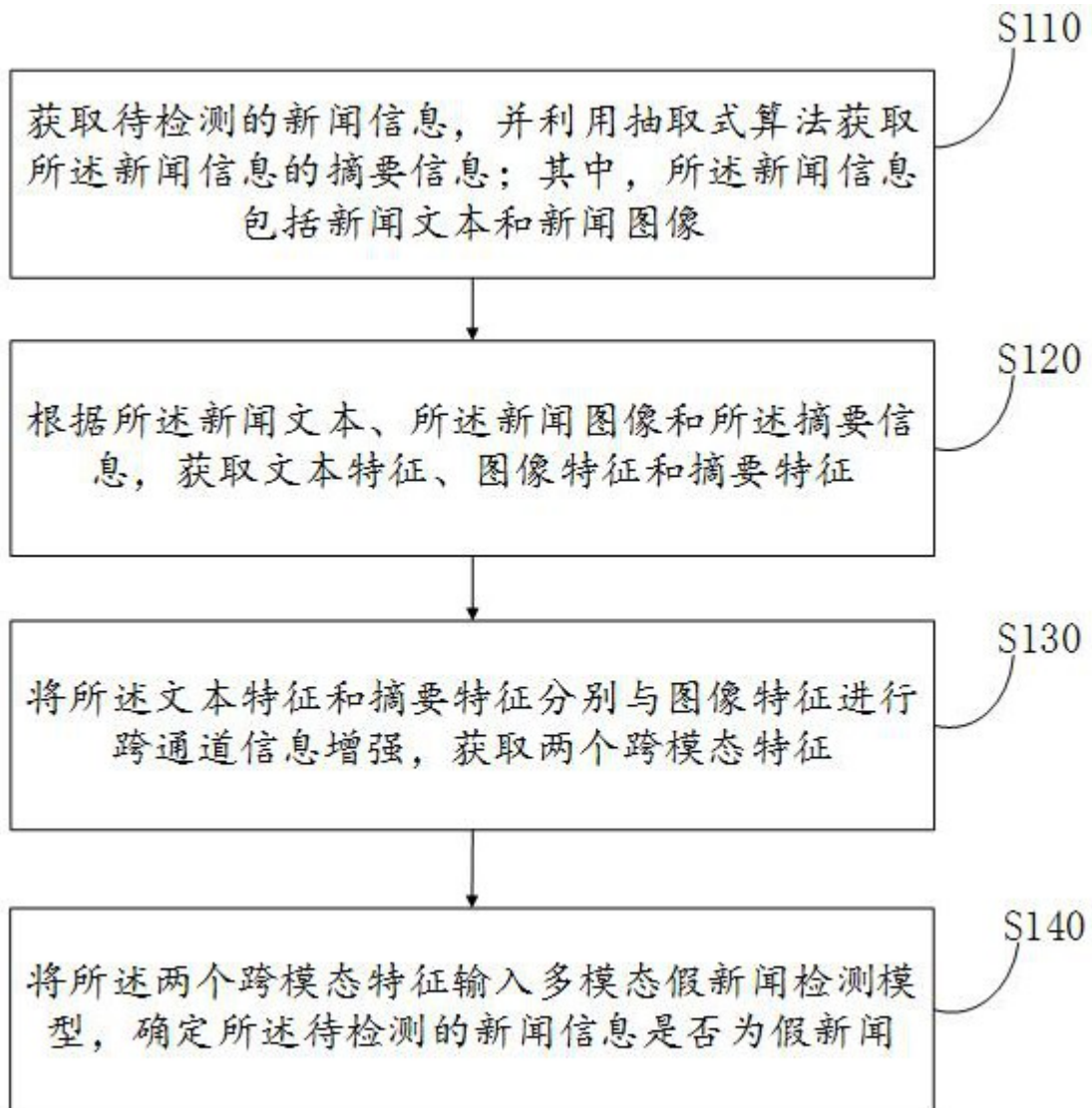


图 1

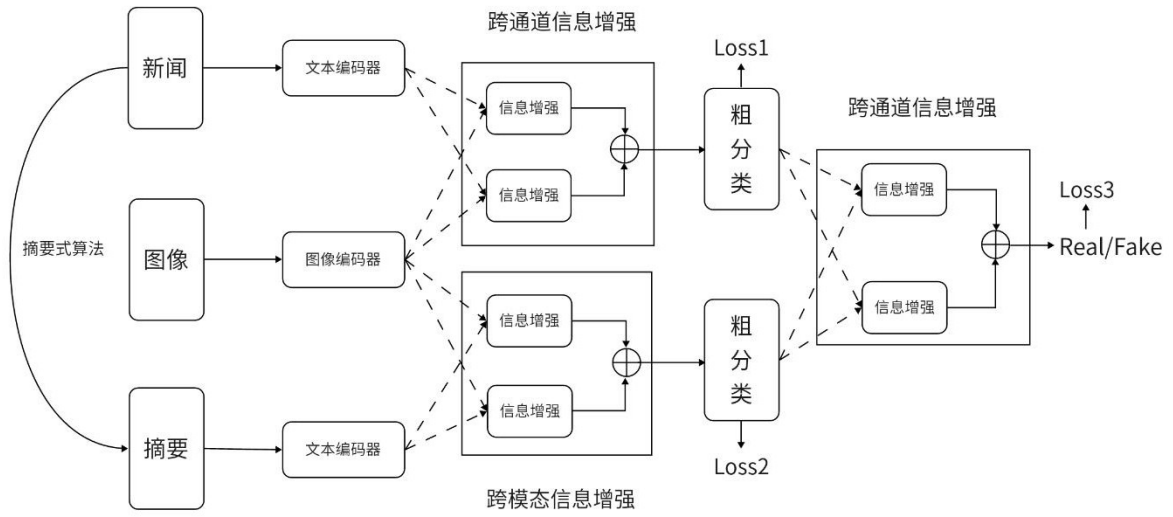


图 2

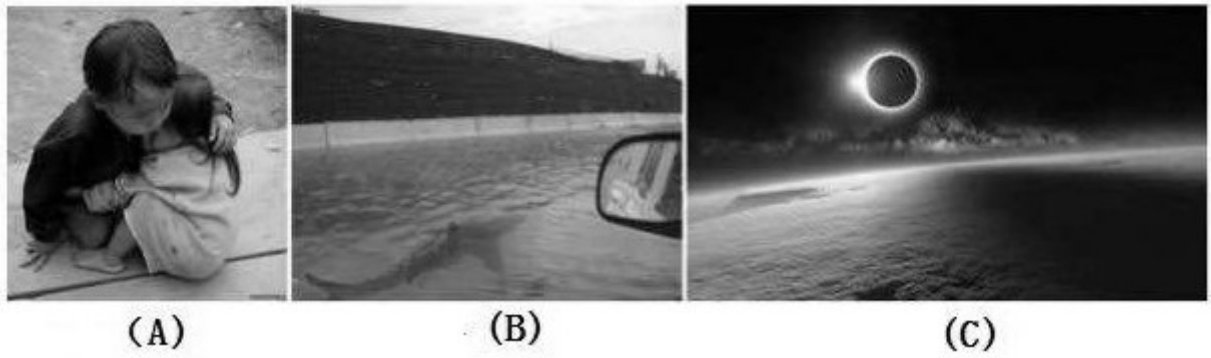


图 3



图 4

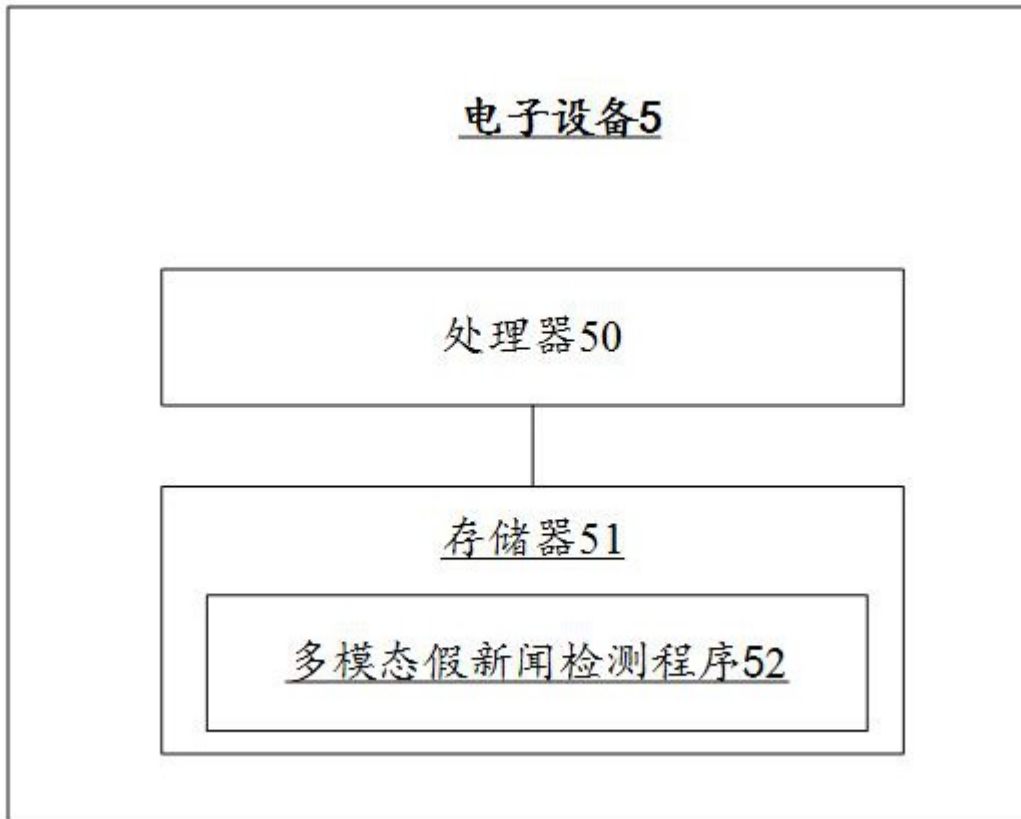


图 5