

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2022年9月29日 (29.09.2022)

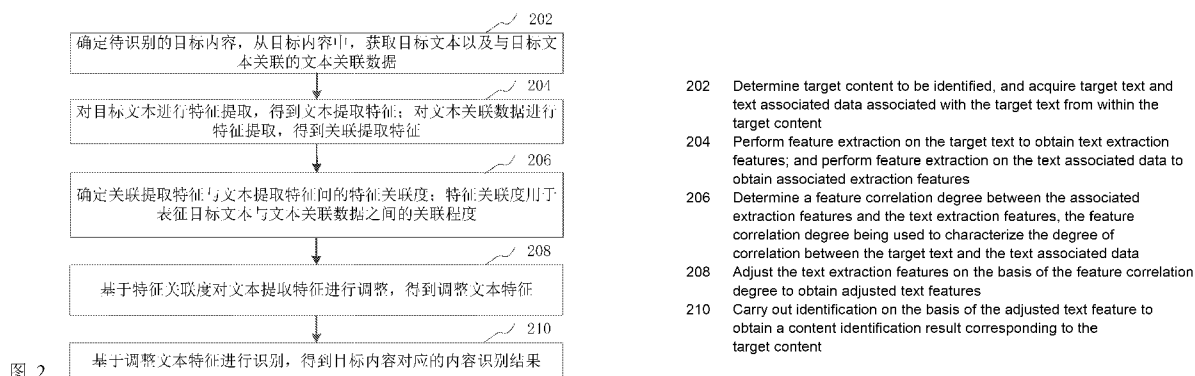


(10) 国际公布号
WO 2022/199504 A1

- (51) 国际专利分类号:
G06K 9/00 (2022.01)
- (21) 国际申请号: PCT/CN2022/081896
- (22) 国际申请日: 2022年3月21日 (21.03.2022)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202110325997.8 2021年3月26日 (26.03.2021) CN
- (71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。
- (72) 发明人: 徐启东 (XU, Qidong); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。 陈小帅 (CHEN, Xiaoshuai); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。
- (74) 代理人: 华进联合专利商标代理有限公司 (ADVANCE CHINA IP LAW OFFICE); 中国广东省广州市天河区珠江东路6号4501房(部位: 自编01-03和08-12单元)(仅限办公用途), Guangdong 510623 (CN)。
- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK,

(54) Title: CONTENT IDENTIFICATION METHOD AND APPARATUS, COMPUTER DEVICE AND STORAGE MEDIUM

(54) 发明名称: 内容识别方法、装置、计算机设备和存储介质



(57) Abstract: The present application relates to a content identification method and apparatus, a computer device and a storage medium. The method comprises: acquiring target text and text associated data associated with the target text from within target content to be identified; performing feature extraction on the target text to obtain text extraction features; performing feature extraction on the text associated data to obtain associated extraction features; determining a feature correlation degree between the associated extraction features and the text extraction features; adjusting the text extraction features on the basis of the feature correlation degree to obtain adjusted text features; and carrying out identification on the basis of the adjusted text features to obtain a content identification result corresponding to the target content.

(57) 摘要: 本申请涉及一种内容识别方法、装置、计算机设备和存储介质。所述方法包括: 从待识别的目标内容中, 获取目标文本以及与目标文本关联的文本关联数据; 对所述目标文本进行特征提取, 得到文本提取特征; 对所述文本关联数据进行特征提取, 得到关联提取特征; 确定所述关联提取特征与所述文本提取特征间的特征关联度; 基于所述特征关联度对所述文本提取特征进行调整, 得到调整文本特征; 基于所述调整文本特征进行识别, 得到所述目标内容对应的内容识别结果。

SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, WS, ZA, ZM, ZW。

- (84)** 指定国 (除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告 (条约第21条(3))。

内容识别方法、装置、计算机设备和存储介质

本申请要求于 2021 年 03 月 26 日提交中国专利局，申请号为 202110325997.8，申请名称为“内容识别方法、装置、计算机设备和存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及计算机技术领域，特别是涉及一种内容识别方法、装置、计算机设备和存储介质。

背景技术

随着自然语言处理技术以及人工智能技术的发展，在很多情况下都需要对内容进行识别，例如对视频进行识别。在对内容进行识别时，可以基于人工智能的模型对内容进行识别，并且可以从内容中获取所需要的信息。例如，可以对文本进行识别以从文本中获取所需要的内容实体。

目前，对内容进行识别的方法，存在不能准确识别内容的信息的情况，导致内容识别的准确度较低。

发明内容

根据本申请提供的各种实施例，提供了一种内容识别方法、装置、计算机设备、存储介质和计算机程序产品。

一种内容识别方法，由计算机设备执行，所述方法包括：确定待识别的目标内容，从所述目标内容中，获取所述目标文本以及与所述目标文本关联的文本关联数据；对所述目标文本进行特征提取，得到文本提取特征；对所述文本关联数据进行特征提取，得到关联提取特征；确定所述关联提取特征与所述文本提取特征间的特征关联度；所述特征关联度用于表征所述目标文本与所述文本关联数据之间的关联程度；基于所述特征关联度对所述文本提取特征进行调整，得到调整文本特征；及，基于所述调整文本特征进行识别，得到所述目标内容对应的内容识别结果。

一种内容识别装置，所述装置包括：目标内容确定模块，用于确定待识别的目标内容，从所述目标内容中，获取目标文本以及与所述目标文本关联的文本关联数据；特征提取模块，用于对所述目标文本进行特征提取，得到文本提取特征；对所述文本关联数据进行特征提取，得到关联提取特征；特征关联度得到模块，用于确定所述关联提取特征与所述文本提取特征间的特征关联度；所述特征关联度用于表征所述目标文本与所述文本关联数据之间的关联程度；调整文本特征得到模块，用于基于所述特征关联度对所述文本提取特征进行调整，得到调整文本特征；及，内容识别结果得到模块，用于基于所述调整文本特征进行识别，得到所述目标内容对应的内容识别结果。

一种计算机设备，包括存储器和一个或多个处理器，所述存储器存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述一个或多个处理器执行上述内容识别方法中的步骤。

一个或多个非易失性可读存储介质，其上存储有计算机可读指令，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器实现上述内容识别方法中的步骤。

一种计算机程序产品，包括计算机可读指令，所述计算机可读指令被处理器执行时实现上述图像处理方法的步骤。

本申请的一个或多个实施例的细节在下面的附图和描述中提出。本申请的其他特征、目的和优点将从说明书、附图以及权利要求书变得明显。

附图说明

- 图 1 为一些实施例中内容识别方法的应用环境图；
- 图 2 为一些实施例中内容识别方法的流程示意图；
- 图 3 为一些实施例中应用内容识别方法进行视频识别的原理图；
- 图 4 为一些实施例中内容识别模型的框架图；
- 图 5 为一些实施例中内容识别模型的框架图；
- 图 6 为一些实施例中利用实体识别网络进行实体识别的原理图；
- 图 7 为一些实施例中内容识别网络的框架图；
- 图 8 为一些实施例中利用实体识别模型进行实体识别的原理图；
- 图 9 为一些实施例中内容识别装置的结构框图；
- 图 10 为一些实施例中计算机设备的内部结构图；
- 图 11 为一些实施例中计算机设备的内部结构图。

具体实施方式

为了更清楚地说明本申请实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

本申请提供的内容识别方法，可以应用于如图 1 所示的应用环境中。该应用环境包括终端 102 以及服务器 104。其中，终端 102 以及服务器 104 通过网络进行通信。

具体地，服务器 104 可以响应于内容识别请求，获取待识别的目标内容，待识别的目标内容可以是内容识别请求中携带的，或者根据内容识别请求中携带的内容标识所获取的内容，服务器 104 可以从目标内容中，获取目标文本，并从目标内容中获取与目标文本关联的文本关联数据，对目标文本进行特征提取，得到文本提取特征，对文本关联数据进行特征提取，得到关联提取特征，确定关联提取特征与文本提取特征间的特征关联度，特征关联度用于表征目标文本与文本关联数据之间的关联程度，基于特征关联度对文本提取特征进行调整，得到调整文本特征，基于调整文本特征进行识别，得到目标内容对应的内容识别结果，服务器 104 可以将内容识别结果与目标内容关联存储，例如可以将内容识别结果作为目标内容的标签。其中，内容识别请求可以是服务器 104 触发的，也可以是其他设备例如终端发送至服务器 104 的。

其中，终端 102 上可以安装有客户端，例如可以安装有视频客户端、浏览器客户端、即时通信客户端或者教育客户端等中的至少一个。终端 102 可以通过客户端响应于用户触发的内容搜索操作，向服务器 104 发送内容搜索请求，内容搜索请求中可以携带搜索信息，服务器 104 可以将搜索信息与内容识别结果进行匹配，当搜索信息与内容识别结果匹配时，将该内容识别结果对应的内容发送至终端 102，终端 102 可以在客户端中展示服务器 104 返回的内容。

其中，终端 102 可以但不限于是笔记本电脑、智能手机、智能电视、台式电脑、平板电脑和便携式可穿戴设备，服务器 104 可以用独立的服务器或者是多个服务器组成的服务器集群或云服务器来实现。可以理解，上述应用场景仅是一种示例，并不构成对本申请实施例提供的内容识别方法的限定，本申请实施例提供的方法还可以应用在其他应用场景中，例如本

申请提供的内容识别方法可以由终端 102 或服务器 104 执行的，还可以是由终端 102 与服务器 104 协同执行的，终端 102 可以将识别出的内容识别结果上传至服务器 104，服务器 104 可以将目标内容与内容识别结果关联存储。

在一些实施例中，如图 2 所示，提供了一种内容识别方法，以该方法应用于图 1 中的服务器 104 为例进行说明，包括以下步骤：

步骤 202，确定待识别的目标内容，从目标内容中，获取目标文本以及与目标文本关联的文本关联数据。

其中，内容可以是视频、音频或文本中的任意一种，内容包括文本数据，还可以包括图像数据或音频数据中的至少一种，音频数据例如可以是语音数据。当内容为视频时，内容中的文本数据可以包括视频中的字幕、弹幕、评论或标题中的至少一种，内容中的图像数据可以是视频中的视频帧，内容中的音频数据可以是视频中的配音或音乐等音频数据。当内容为音频数据时，内容中的文本数据可以是音频数据对应的文本数据，例如当内容为歌曲时，内容中的文本数据可以是歌曲对应的歌词，内容中的音频数据可以是音频帧。音频帧是对音频进行分帧得到的，分帧指的是将音频分为多个小段，每个小段为一帧。

目标内容指的是待进行识别的内容，可以是待进行身份识别的内容或待进行场景识别的内容中的至少一种，身份识别指的是对目标内容中出现的人物的身份进行识别，例如可以通过识别目标内容中出现的人物信息确定人物的身份，人物信息可以包括人物的名称或人物的人脸中的至少一种，场景识别指的是对目标内容所属的场景进行识别，例如可以通过识别目标内容中出现的地点确定场景。目标文本指的是目标内容中的文本数据，可以包括目标内容中任意时刻的文本数据，例如，当目标文本为视频时，目标文本可以包括视频中任意时刻或时间段展示的字幕、弹幕、评论或标题中的至少一种。当目标内容为歌曲时，目标文本数据可以是歌曲对应的歌词。

文本关联数据指的是目标内容中与目标文本具有关联关系的数据，例如可以包括目标内容中与目标文本具有关联关系的目标图像数据或目标音频数据中的至少一种。目标图像数据为目标内容中与目标文本具有关联关系的图像数据，目标音频数据为目标内容中与目标文本具有关联关系的音频数据。目标图像数据可以包括一张或多张图像，多张指的是至少两张，目标音频数据可以包括一段或多段音频帧，多段指的是至少两段。

关联关系可以包括时间上的关联关系。例如文本关联数据可以包括目标内容中目标文本出现的时间内所出现的数据，或者包括目标内容中与目标文本出现的时间之间的时间间隔小于时间间隔阈值的时间内所出现的数据。例如当目标内容为目标视频，目标文本为视频的字幕文本时，则文本关联数据可以为与该字幕匹配的视频帧以及语音，例如目标文本与对应的文本关联数据可以为对同一视频场景进行描述的数据。例如目标文本为目标视频中的目标时间出现的字幕时，文本关联数据可以包括目标视频中目标时间出现的数据，例如可以包括目标视频中目标时间出现的视频帧、弹幕或音频帧中的至少一种，或者包括目标视频中与目标时间之间的时间间隔小于时间间隔阈值的时间所出现的数据。时间间隔阈值可以是预先设置的，也可以根据需要设置。其中，目标视频可以是任意的视频，可以是直接拍摄得到的视频，也可以是从拍摄得到的视频中截取得到的视频片段，目标视频可以是任意类型的视频，包括但不限于是广告类的视频、电视剧类的视频或新闻类视频中的至少一种，目标视频还可以是待推送至用户的视频。目标视频中目标时间出现的视频帧可以包括一帧或多帧，目标视频中目标时间出现的音频帧可以包括一帧或多帧，多帧指的是至少两帧。

关联关系还可以包括语义上的关联关系。例如文本关联数据可以包括目标内容中与目标文本的语义匹配的数据，与目标文本的语义匹配的数据可以包括与目标文本的语义一致的数据，或者包括语义与目标文本的语义的差异小于语义差异阈值的数据。语义差异阈值可以是预先设置的，也可以根据需要设置。

具体地，服务器可以获取待进行实体识别的内容，例如待进行实体识别的视频，将待进行实体识别的内容作为待识别的目标内容，利用本申请提供的内容识别方法对待进行实体识别的内容进行识别，得到识别出的实体词，基于识别出的实体词构建知识图谱，或者可以将识别出的实体词作为目标内容对应的标签。当需要对目标内容进行推送时，可以根据目标内容对应的标签，确定目标内容所匹配的用户，将目标内容推送至该匹配用户的终端。

其中，实体(Entity)是指具有特定意义的事物，例如可以包括地名、机构名或者专有名词等中的至少一种。目标文本中可以包括一个或者多个实体，实体词为表示实体的词语。例如假设目标文本为“猴子喜欢吃香蕉”，则目标文本中包括的实体为“猴子”以及“香蕉”，“猴子”为一个实体词，“香蕉”为一个实体词。知识图谱(Knowledge Graph)是一种基于图的数据结构，包括节点(point)和边(Edge)，每个节点表示一个实体，每条边为实体与实体之间的关系。

实体识别还可以称为实体词识别或命名实体识别(Named Entity Recognition, NER)。实体词识别是自然语言处理(Natural Language Processing, NLP)领域研究的一个重要方向，进行实体词识别的方法有很多，例如包括基于词典和规则的方法，包括隐马尔科夫模型(Hidden Markov Model, HMM)、最大熵马尔科夫模型(Maximum Entropy Markov Model, MEMM)、条件随机场(Conditional Random Fields, CRF)等机器学习方法，包括循环神经网络(RNN, Recurrent Neural Networks)和长短期记忆网络(LSTM, Long Short-Term Memory)等深度学习模型，以及包括LSTM与CRF结合的识别方法。其中，自然语言处理是计算机科学领域与人工智能领域中的一个重要方向。它研究能实现人与计算机之间用自然语言进行有效通信的各种理论和方法。

在一些实施例中，第一终端可以向服务器发送内容推送请求，内容推送请求中可以携带待推送的内容对应的内容标识，内容标识用于唯一识别内容，待推送的内容例如可以是待推送的视频，服务器可以响应于内容推送请求，获取内容推送请求中携带的内容标识对应的待推送的内容，作为待识别的目标内容。例如，第一终端可以展示内容推送界面，内容推送界面中可以展示推送内容获取区域以及内容推送触发控件，推送内容获取区域用于接收待推送的内容对应的内容信息，内容信息中包括一个或者多个内容标识，多个指的是至少两个，内容推送触发控件用于触发第一终端向服务器发送内容推送请求，当第一终端获取到对内容推送触发控件的触发操作时，获取推送内容获取区域中接收的内容信息，向服务器发送携带该内容信息的内容推送请求。服务器可以获取内容信息中包括的各个内容标识分别对应的内容，分别作为各个待识别的目标内容。服务器可以利用本申请提供的内容识别方法对各个待识别的内容进行识别，根据识别的结果确定与各个目标内容分别匹配的用户，将目标内容推送至所匹配的用户终端。例如，可以将识别的结果与用户的用户画像进行匹配，当匹配成功时，将目标内容推送至该用户的终端。

举例说明，当内容为视频时，内容识别方法还可以称为视频识别方法，内容推送请求还可以称为视频推送请求，内容推送界面例如可以为图3中的视频推送界面300，推送内容获取区域例如可以是图3中的区域302，内容推送触发控件例如可以是图3中的“确定”控件

304, 当终端获取到对“确定”控件 304 的点击操作时, 可以向服务器发送视频推送请求, 服务器根据视频识别方法对视频 A 和视频 B 进行识别, 确定视频 A 所匹配的用户 1, 以及视频 B 所匹配的用户 B, 将视频 A 推送至用户 1 的终端, 将视频 B 推送至用户 2 的终端。

步骤 204, 对目标文本进行特征提取, 得到文本提取特征; 对文本关联数据进行特征提取, 得到关联提取特征。

其中, 文本提取特征是对目标文本进行特征提取所得到的特征, 文本提取特征可以为目标文本对应的目标分词的目标词向量进行进一步的特征提取得到的特征。目标分词是对目标文本进行切分得到的, 切分的粒度可以根据需要设置, 例如可以是字、词语或句子为单位进行切分, 得到切分后的文本块, 将每个文本块作为一个分词。当以字为单位进行切分时, 一个字对应一个文本块, 即一个字为一个分词。例如, 当目标文本为“我好渴”时, 当以字为单位进行分词时, 得到的各个分词分别为“我”、“好”以及“渴”。目标词向量为目标分词的向量表示形式。目标文本切分得到的目标分词可以有一个或多个, 多个指的是至少两个。

关联提取特征是对文本关联数据进行特征提取所得到的特征。当文本关联数据为目标图像数据时, 关联提取特征可以是对目标图像数据进行图像特征提取所得到的目标图像特征。当文本关联数据为目标音频数据时, 关联提取特征可以是对目标音频数据进行音频特征提取所得到的目标音频特征。目标图像特征是对目标图像数据进行图像特征提取所提取的图像特征。目标音频特征是对目标音频特征进行音频特征提取所提取的音频特征。文本提取特征与关联提取特征可以是同维度的, 例如可以是同维度的向量或同纬度的矩阵。

具体地, 服务器可以将目标文本输入到已训练的内容识别模型中的文本特征提取网络, 利用文本特征提取网络对目标文本进行特征提取, 得到文本提取特征, 将文本关联数据输入到已训练的内容识别模型中的关联特征提取网络中, 利用关联提取网络对文本关联数据进行特征提取, 得到关联提取特征。已训练的内容识别模型用于对内容进行识别, 得到内容识别结果, 例如用于识别视频的字幕中包括的实体词或视频的场景中的至少一种。已训练的内容识别模型中的关联特征提取网络可以有多个, 例如关联特征提取网络可以包括图像特征提取网络或音频特征提取网络中的至少一种, 图像特征提取网络用于提取图像的特征, 音频特征提取网络用于提取音频的特征。当文本关联数据为目标图像数据时, 可以将文本关联数据输入到图像特征提取网络中, 将图像特征提取网络提取的图像特征作为关联提取特征。当文本关联数据为目标音频数据时, 可以将文本关联数据输入到音频特征提取网络中, 将音频特征提取网络提取的音频特征作为关联提取特征。

其中, 文本特征提取网络、图像特征提取网络以及音频特征提取网络可以是基于人工智能的神经网络, 例如可以是卷积神经网络 (Convolutional Neural Networks, CNN), 当然也可以是其他类型的神经网络。文本特征提取网络例如可以是 Transformer (转换器) 网络或基于 Transformer 的双向编码器 (Bidirectional Encoder Representations from Transformers, BERT) 网络。图像特征提取网络例如可以是残差卷积网络 (ResNet, Residual Network)。音频特征提取网络例如可以是 VGG (Visual Geometry Group) 卷积网络。VGG 代表了牛津大学的视觉几何组 (Visual Geometry Group, VGG)。例如, 服务器可以对目标图像进行尺度变换, 得到尺度变换后的图像, 将尺度变换后的图像数据输入到残差卷积网络中进行图像特征提取, 将残差卷积网络中的 feature map (特征图) 提取层输出的特征进行池化 (pooling), 例如池化为固定尺寸的 $n \times n$ 大小, 将池化后的特征作为关联提取特征。 n 为

大于等于 1 的正数。

在一些实施例中，对目标文本进行特征提取，得到文本提取特征的步骤包括：对目标文本进行切分，得到目标分词，对目标分词进行向量转化，得到目标分词对应的目标词向量，将目标词向量作为文本提取特征。

在一些实施例中，服务器可以将目标文本输入到基于 attention(注意力)的 transformer 模型中，transformer 模型作为文本特征的编码器，可以对目标文本进行编码，得到目标文本中每个字的嵌入式(embedding)表示形式的编码特征，可以将每个字对应的编码特征作为文本提取特征。

在一些实施例中，服务器可以对目标音频数据进行频谱计算，得到目标音频数据对应的频谱图，对目标音频数据对应的频谱图进行特征提取，将提取到的特征作为关联提取特征。例如，服务器可以对目标音频数据对应的频谱图进行声谱计算，得到目标音频数据对应的声谱信息，对目标音频数据的声谱信息进行特征提取得到关联提取特征。例如，服务器可以利用 hann(海宁窗)时窗对目标音频数据进行傅里叶变换得到目标音频数据对应的频谱图，通过 mel(梅尔)滤波器对频谱图进行计算，得到目标音频数据对应的声谱信息，利用 VGG 卷积网络对声谱信息进行特征提取，将特征提取得到的音频特征作为关联提取特征。

步骤 206，确定关联提取特征与文本提取特征间的特征关联度；特征关联度用于表征目标文本与文本关联数据之间的关联程度。

其中，特征关联度是对关联提取特征与文本提取特征进行关联计算所得到的结果，关联度越大，则表示关联关系越强。特征关联度可以包括图像关联度或音频关联度中的至少一个，图像关联度指的是对目标图像特征与文本提取特征进行关联计算所得到的结果，音频关联度指的是对目标音频特征与文本提取特征进行关联计算所得到的结果。关联计算可以为乘积运算或加和运算中的至少一种。

关联提取特征可以包括多个有序的关联特征值，文本提取特征可以包括多个有序的文本特征值，文本特征值指的是文本提取特征中包括的特征值，关联特征值指的是关联提取特征中包括的特征值。关联提取特征与文本提取特征可以是同维度的，例如可以是同维度的向量或矩阵，也就是说，关联提取特征中所包括的关联特征值的数量，与文本提取特征中所包括的文本特征值的数量可以是相同的。例如，假设文本提取特征为向量 $A=[a_1, a_2, a_3]$ ，关联提取特征为向量 $B=[b_1, b_2, b_3]$ ，其中，向量 A 包括 3 个元素，分别为 a_1, a_2 和 a_3 ，向量 A 中的每个元素为一个文本特征值，同样的，向量 B 中包括 3 个元素，分别为 b_1, b_2 和 b_3 ，向量 B 中的每个元素为一个关联特征值。

具体地，关联计算可以为乘积运算或加和运算中的至少一种。当关联计算为乘积运算时，可以将关联提取特征中的关联特征值与文本提取特征中对应位置的文本特征值进行乘积运算，得到乘积运算值，对各个乘积运算值进行统计运算，例如对各个乘积运算值进行加和运算或均值运算，得到统计运算结果，基于统计运算结果得到特征关联度，例如可以将统计运算结果作为特征关联度，也可以对统计运算结果进行归一化处理，将归一化处理的结果作为特征关联度。当关联计算为加和运算时，可以将关联提取特征中的关联特征值与文本提取特征中对应位置的文本特征值进行加和运算，得到加和运算值，对各个加和运算值进行统计运算，例如可以对各个加和运算值进行加和运算或均值运算，得到统计运算结果。

在一些实施例中，目标文本切分得到的目标分词有多个，服务器可以获取根据各个目标分词分别得到的文本提取特征，将各个文本提取特征组成矩阵，将组成的矩阵作为文本提取

特征矩阵，文本提取特征矩阵中的每一列为一个文本提取特征。服务器可以对关联提取特征与文本提取特征矩阵进行乘积运算，得到总乘积运算结果，基于总乘积运算结果确定各个文本提取特征分别对应的特征关联度。其中，对关联提取特征与文本提取特征矩阵进行乘积运算，得到总乘积运算结果的步骤可以包括：将文本提取特征矩阵中的文本提取特征分别与关联提取特征进行乘积运算，得到各个文本提取特征分别对应的子乘积运算结果，将各个子乘积运算结果作为总乘积运算结果。其中，将文本提取特征矩阵中的文本提取特征分别与关联提取特征进行乘积运算，得到各个文本提取特征分别对应的子乘积运算结果的步骤可以包括：将文本提取特征中的文本特征值与关联提取特征中对应位置的关联特征值进行乘积运算，得到文本提取特征对应的子乘积运算结果。基于总乘积运算结果确定各个文本提取特征分别对应的特征关联度的步骤可以包括：对总乘积运算结果中的各个子乘积运算结果进行归一化处理，得到各个归一化后的子乘积运算结果，将归一化后的子乘积运算结果作为文本提取特征对应的特征关联度。

在一些实施例中，当文本关联数据为目标图像数据，并且目标图像数据有多个时，服务器可以将各个目标图像数据对应的目标图像特征组成矩阵，将组成的矩阵作为图像特征矩阵，图像特征矩阵中的每一列为一个目标图像特征。服务器可以对目标图像特征矩阵的转置矩阵与文本提取特征矩阵进行矩阵相乘运算，得到第一乘积矩阵，对第一乘积矩阵中的各个矩阵值进行归一化处理，得到归一化后的第一乘积矩阵，基于归一化后的第一乘积矩阵确定各个文本提取特征分别对应的图像关联度，归一化后的第一乘积矩阵中包括各个文本提取特征分别对应的图像关联度。

例如，假设目标文本为“我好渴”，按照字为单元对目标文本进行切分，得到3个目标分词，分别为“我”、“好”以及“渴”，一个分词为一个文字，目标分词对应的目标词向量的维度为2，“我”对应的目标词向量为 $A=(a1, a2)^T$ ，“好”对应的目标词向量为 $B=(b1, b2)^T$ ，“渴”对应的目标词向量为 $C=(c1, c2)^T$ ，将每一个目标词向量作为一个文本提取特征，则文本提取特征矩阵 $feature_{text}$ 可以表示为公式（1）。假设有3个目标图像数据，这3个目标图像数据可以相同也可以不同，例如为3张图像，各个目标图像数据分别对应的目标图像特征为 $R=(r1, r2)^T$ ， $M=(m1, m2)^T$ ， $N=(n1, n2)^T$ ， R 、 M 以及 N 可以相同也可以不同，则目标图像特征矩阵 $feature_{image}$ 可以表示为公式（2）。则第一乘积矩阵 $L1$ 可以表示为公式（3）。

$$feature_{text} = [A, B, C] = \begin{bmatrix} a1 & b1 & c1 \\ a2 & b2 & c2 \end{bmatrix} \quad (1) \quad feature_{image} = [R, M, N] = \begin{bmatrix} r1 & m1 & n1 \\ r2 & m2 & n2 \end{bmatrix} \quad (2)$$

$$L1 = [feature_{image}]^T [feature_{text}] \quad (3)$$

在一些实施例中，对第一乘积矩阵中的各个矩阵值进行归一化处理，得到归一化后的第一乘积矩阵的步骤包括：确定缩放因子，将第一乘积矩阵中的各个矩阵值分别除以缩放因子，得到各个矩阵值对应的缩放值，对各个缩放值进行归一化处理，将各个缩放值组成的矩阵作为归一化后的第一乘积矩阵。其中，缩放因子可以是预先设置的，也可以是根据需要设置的，例如缩放因子可以根据文本提取特征的维度确定，例如缩放因子可以与文本提取特征的维度成正相关关系，例如可以对文本提取特征的维度进行开方计算，得到缩放因子，例如可以对文本提取特征的维度进行开平方处理，将开平方处理后的结果与第一数值的比值，作为缩放因子。第一数值可以是预设设置的。归一化处理所采用的方法，可以是任意的能够将输入数据转化为0到1之间的数的函数，例如可以采用函数 softmax 进行归一化处理。例如，可以利用公式（4）计算得到归一化后的第一乘积矩阵 $L2$ 。其中， d 为文本提取特征的维度， m 为

第一数值。

$$L2 = \text{soft max} \left(\frac{\begin{bmatrix} \text{feature}_{image} \end{bmatrix}^T \begin{bmatrix} \text{feature}_{text} \end{bmatrix}}{\sqrt{d}/m} \right) \quad (4)$$

同样的，当文本关联数据为目标音频数据，且目标音频数据有多个时，服务器可以将各个目标音频数据对应的目标音频特征组成目标音频特征矩阵，目标音频特征矩阵中的每一列为一个目标音频特征，服务器可以对目标音频特征矩阵的转置矩阵与文本提取特征矩阵进行矩阵相乘运算，得到第二乘积矩阵，对第二乘积矩阵中的各个矩阵值进行归一化处理，得到归一化后的第二乘积矩阵，基于归一化后的第二乘积矩阵确定各个文本提取特征分别对应的音频关联度，归一化后的第二乘积矩阵中包括各个文本提取特征分别对应的音频关联度。

步骤 208，基于特征关联度对文本提取特征进行调整，得到调整文本特征。

其中，调整文本特征是基于特征关联度对文本提取特征进行调整所得到的特征，调整文本特征可以包括第一调整文本特征或第二调整文本特征中的至少一种。第一调整文本特征指的是基于图像关注强度对文本提取特征进行调整所得到的特征。第二调整文本特征指的是基于音频关注强度对文本提取特征进行调整所得到的特征。

具体地，服务器可以基于特征关联度得到文本提取特征对应的特征关注强度，特征关联度与特征关注强度成正相关关系，基于特征关注强度对文本提取特征进行调整，得到调整文本特征。其中，特征关联度与特征关注强度成正相关关系。当文本提取特征对应的目标分词不同时，文本提取特征对应的特征关注强度可以不同。例如可以将特征关联度作为特征关注强度，或者对特征关联度进行线性运算或非线性运算，将运算的结果作为特征关注强度。线性运算包括加和运算或乘积运算中的至少一种，非线性运算包括指数运算或对数运算中的至少一种。正相关关系指的是：在其他条件不变的情况下，两个变量变动方向相同，一个变量由大到小变化时，另一个变量也由大到小变化。可以理解的是，这里的正相关关系是指变化的方向是一致的，但并不是要求当一个变量有一点变化，另一个变量就必须也变化。例如，可以设置当变量 a 为 10 至 20 时，变量 b 为 100，当变量 a 为 20 至 30 时，变量 b 为 120。这样，a 与 b 的变化方向都是当 a 变大时，b 也变大。但在 a 为 10 至 20 的范围内时，b 可以是没有变化的。

特征关注强度可以包括图像关注强度或音频关注强度中的至少一个。图像关注强度基于图像关联度得到，图像关注强度与图像关联度成正相关关系，音频关注强度基于音频关联度得到，音频关注强度与音频关联度成正相关关系。特征关注强度用于反映对特征进行关注的强度，特征关注强度越大，说明在进行内容识别时，需要更加关注该特征。

在一些实施例中，服务器可以对关联提取特征与文本提取特征进行相似度计算，得到特征相似度，将特征相似度作为特征关联度，基于特征关联度得到文本提取特征对应的特征关注强度。例如可以按照余弦相似度计算公式，对关联提取特征与文本提取特征进行相似度计算，将计算得到的余弦相似度作为特征相似度。

在一些实施例中，服务器可以利用特征关注强度对文本提取特征中的各个文本特征值进行调整，得到调整文本特征，例如可以对文本特征值与特征关注强度进行线性运算，得到线性运算后的文本特征值，基于各个线性运算后的文本特征值得到调整文本特征。其中，线性运算可以包括加和运算或乘积运算中的至少一种。例如，服务器可以将特征关注强度分别与文本提取特征中的各个特征值进行乘积运算，得到各个特征值乘积，并按照文本提取特征中

特征值的位置对特征值乘积进行排序，得到特征值序列，将特征值序列作为调整文本特征。文本特征值在文本提取特征中的位置与该文本特征值计算得到的特征值乘积在特征值序列中的位置相同。例如，假设文本提取特征为向量 $[a_1, a_2, a_3]$ ，则 a_1 、 a_2 和 a_3 为文本提取特征中的特征值，当特征关注强度为 c 时，特征值序列为向量 $[a_1*c, a_2*c, a_3*c]$ ， a_1*c 、 a_2*c 和 a_3*c 为特征值乘积， a_1*c 在特征值序列 $[a_1*c, a_2*c, a_3*c]$ 中位置与 a_1 在文本提取特征 $[a_1, a_2, a_3]$ 中的位置相同。

在一些实施例中，服务器可以利用归一化后的第一乘积矩阵对文本提取特征矩阵进行调整，得到第一调整文本特征矩阵，归一化后的第一乘积矩阵中包括各个文本提取特征分别对应的图像关联度，第一调整文本特征矩阵中可以包括各个文本提取特征分别对应的第一调整文本特征。例如，服务器可以对归一化后的第一乘积矩阵与文本提取特征矩阵的转置矩阵进行矩阵相乘运算，将相乘后得到的矩阵的转置矩阵作为第一调整文本特征矩阵。例如，可以采用公式(5)计算得到第一调整文本特征矩阵 $feature_{fusion1}$ ，其中 $feature_{fusion1}$ 表示第一调整文本特征矩阵， $[feature_{fusion1}]^T$ 表示 $feature_{fusion1}$ 的转置矩阵。同样的，服务器可以利用归一化后的第二乘积矩阵与文本提取特征的转置矩阵进行矩阵相乘运算，得到第二调整文本特征矩阵，归一化后的第二乘积矩阵中包括各个文本提取特征分别对应的音频关联度，第二调整文本特征矩阵中可以包括各个文本提取特征分别对应的第二调整文本特征。例如，可以采用公式(6)计算得到第二调整文本特征矩阵 $feature_{fusion2}$ ，其中， $feature_{audio}$ 为目标音频特征矩阵。 $[feature_{audio}]^T$ 表示目标音频特征矩阵对应的转置矩阵。

$$[feature_{fusion1}]^T = \text{soft max} \left(\frac{[feature_{image}]^T [feature_{text}]}{\sqrt{d/m}} \right) [feature_{text}]^T \quad (5)$$

$$[feature_{fusion2}]^T = \text{soft max} \left(\frac{[feature_{audio}]^T [feature_{text}]}{\sqrt{d/m}} \right) [feature_{text}]^T \quad (6)$$

步骤210，基于调整文本特征进行识别，得到目标内容对应的内容识别结果。

其中，内容识别结果是基于调整文本特征进行识别所得到的结果。内容识别结果可以根据识别时所采用的内容识别网络确定，内容识别网络不同，所得到的内容识别结果可以相同，也可以不同。内容识别网络可以包括场景识别网络或实体识别网络中的至少一种。场景识别网络用于识别场景，实体识别网络用于识别实体。当内容识别网络为场景识别网络时，内容识别模型还可以称为场景识别模型，当内容识别网络为实体识别网络时，内容识别模型还可以称为实体识别模型或实体词识别模型。

具体地，服务器可以将调整文本特征输入到已训练的内容识别模型的内容识别网络中，利用内容识别模型对调整文本特征进行识别，得到目标内容对应的内容识别结果。例如，当文本提取特征为目标文本中的目标分词对应的特征时，可以将各个目标分词分别对应的调整文本特征按照目标分词在目标文本中的顺序进行排序，将排序得到的序列作为特征序列，服务器可以基于特征序列进行识别，得到目标内容对应的内容识别结果，例如可以将特征序列输入到内容识别模型中的内容识别网络中，得到内容识别结果，例如当内容识别网络为实体识别网络时，可以识别出目标内容中包括的实体词。

如图4所示，展示了一个内容识别模型400，内容识别模型400包括文本特征提取网络、关联特征提取网络、关注强度计算模块、特征调整模块以及内容识别网络。其中，关注强度计算模块用于对关联提取特征与文本提取特征进行关联计算，得到特征关注强度。特征调整

模块用于基于特征关注强度对文本提取特征进行调整，得到调整文本特征，将调整文本特征输入到内容识别网络中，得到目标内容对应的内容识别结果。内容识别模型 400 中各个网络以及模块可以通过联合训练得到的。服务器从目标中获取目标文本以及文本关联数据，将目标文本输入到文本特征提取网络中，得到文本提取特征，将关联文本数据输入到关联特征提取网络中，得到关联提取特征，将文本提取特征以及关联提取特征输入到关注强度计算模块，得到特征关注强度，将特征关注强度以及文本提取特征输入到特征调整模块中，得到调整文本特征，将调整文本特征输入到内容识别网络中，得到内容识别结果。

在一些实施例中，服务器也可以对调整文本特征以及文本提取特征进行融合，得到融合文本特征，例如可以对调整文本特征与文本提取特征进行统计运算，例如进行加权计算或均值计算，得到融合文本特征，例如服务器可以确定调整文本特征对应的调整特征权重，基于调整特征权重对调整文本特征以及文本提取特征进行融合，得到融合文本特征，服务器可以基于融合文本特征进行识别，得到目标内容对应的内容识别结果。

在一些实施例中，调整文本特征包括第一调整文本特征以及第二调整文本特征，服务器可以基于第一调整文本特征、第二调整文本特征以及文本提取特征进行融合，例如加权计算或均值计算，得到融合文本特征。例如，调整特征权重可以包括第一调整文本特征对应的第一特征权重以及第二调整文本特征对应的第二特征权重，服务器可以基于第一特征权重对第一调整文本特征以及文本提取特征进行融合，得到第一融合特征，基于第二特征权重对第二调整文本特征以及文本提取特征进行融合，得到第二融合特征，将第一融合特征与第二融合特征进行统计运算，将统计运算的结果作为融合文本特征，例如将第一融合特征与第二融合特征中对应位置的特征值进行加和计算，得到各个加和值，按照特征值在第一融合特征或第二融合特征中的位置，对各个加和值进行排序，将排序后得到的序列作为融合文本特征。

上述内容识别方法中，确定待识别的目标内容，从目标内容中，获取目标文本以及与目标文本关联的文本关联数据，对目标文本进行特征提取，得到文本提取特征，对文本关联数据进行特征提取，得到关联提取特征，确定关联提取特征与文本提取特征间的特征关联度，特征关联度用于表征目标文本与文本关联数据之间的关联程度，基于特征关联度对文本提取特征进行调整，得到调整文本特征，基于调整文本特征进行识别，得到目标内容对应的内容识别结果。由于特征关联度用于表征目标文本与文本关联数据之间的关联程度，从而基于特征关联度对文本提取特征进行调整，实现了自适应的根据文本关联数据与目标文本之间的关联程度进行文本特征的调整，从而在基于调整文本特征进行识别时，使得识别结果受到文本关联数据的影响，目标文本与文本关联数据之间的关联程度越大，文本关联数据对识别结果的影响程度越大，使得在内容识别时更加关注关联程度较大的信息，提高了内容识别的准确度。

在一些实施例中，基于调整文本特征进行识别，得到目标内容对应的内容识别结果包括：对调整文本特征以及文本提取特征进行融合，得到融合文本特征；基于融合文本特征进行识别，得到目标内容对应的内容识别结果。

其中，融合文本特征是将调整文本特征与文本提取特征进行融合所得到的特征。融合文本特征、调整文本特征以及文本提取特征的维度可以相同，例如可以是相同维度的向量或矩阵。

具体地，服务器可以对调整文本特征与文本提取特征进行统计运算，例如均值运算或加和运算，将统计运算的结果作为融合文本特征。例如，服务器可以对文本提取特征进行编码，

得到文本提取特征对应的编码后的特征，作为第一编码特征，可以对调整文本特征进行编码，得到调整文本特征对应的编码后的特征，作为第二编码特征，对第一编码特征以及第二编码特征进行统计运算，例如均值运算或加和运算，将运算的结果作为融合文本特征。

在一些实施例中，服务器可以将融合文本特征输入到已训练的内容识别模型的内容识别网络中，利用内容识别网络对融合文本特征进行识别，得到目标内容对应的内容识别结果。

本实施例中，对调整文本特征以及文本提取特征进行融合，得到融合文本特征，基于融合文本特征进行识别，得到目标内容对应的内容识别结果，可以提高内容识别的准确度。

在一些实施例中，对调整文本特征以及文本提取特征进行融合，得到融合文本特征包括：对文本提取特征进行编码，得到第一编码特征，对调整文本特征进行编码，得到第二编码特征；将第一编码特征与第二编码特征进行融合，得到融合编码特征；基于融合编码特征得到调整文本特征对应的调整特征权重；基于调整特征权重对调整文本特征以及文本提取特征进行融合，得到融合文本特征。

其中，第一编码特征是对文本提取特征进行编码所得到的特征。第二编码特征是对调整文本特征进行编码所得到的特征。融合编码特征是将第一编码特征与第二编码特征进行融合所得到的特征。调整特征权重基于融合编码特征得到。

具体地，内容识别模型中还可以包括第一编码器、第二编码器以及特征融合模块，特征融合模块用于对第一编码特征与第二编码特征进行融合，得到融合编码特征。服务器可以将文本提取特征输入到已训练的内容识别模型中的第一编码器进行编码，得到第一编码特征，将调整文本特征输入到已训练的内容识别模型中的第二编码器进行编码，得到第二编码特征，将第一编码特征与第二编码特征进行融合，例如可以将第一编码特征以及第二编码特征输入到特征融合模块中，得到融合编码特征。其中，第一编码器以及第二编码器可以是基于人工智能的神经网络，内容识别模型中的各个网络以及模块可以是联合训练得到的，例如第一编码器与第二编码器是联合训练得到的。

在一些实施例中，服务器可以对第一编码特征与第二编码特征进行统计运算，得到编码统计特征，例如，可以将第一编码特征与第二编码特征进行相加运算，将相加运算的结果作为融合编码特征，或者对第一编码特征与第二编码特征进行均值运算，将计算得到的均值作为融合编码特征。服务器可以基于编码统计特征确定融合编码特征，例如可以将编码统计特征作为融合编码特征。

在一些实施例中，服务器可以对融合编码特征进行归一化处理，将归一化所得到的结果作为调整文本特征对应的调整特征权重。例如，已训练的内容识别模型中可以包括激活层，激活层可以将数据进行转化为 0 与 1 之间的数据，对融合编码特征进行归一化处理，将归一化所得到的结果作为调整特征权重的步骤可以包括：将融合编码特征输入到内容识别模型的激活层中进行激活处理，将激活处理的结果作为调整文本特征对应的调整特征权重。

在一些实施例中，服务器可以将调整特征权重与调整文本特征进行乘积计算，得到计算后的调整文本特征，将计算后的调整文本特征与文本提取特征进行统计运算，例如进行加和运算或均值运算，将统计运算的结果作为融合文本特征。

在一些实施例中，服务器可以确定文本提取特征对应的文本特征权重，例如可以获取预设权重，将预设权重作为文本特征权重，预设权重可以是根据需要预先设置的权重。文本特征权重也可以是根据调整特征权重确定的，例如调整特征权重可以与文本特征权重成负相关关系，调整特征权重与文本特征权重的和可以为预设数值，预设数值可以根据需要预先设置

的，例如可以是 1，例如可以将预设数值减去文本特征权重所得到的结果作为文本特征权重，例如，当调整特征权重为 0.3 时，文本特征权重可以为 0.7。其中，预设数值大于文本特征权重，并且预设数值大于调整特征权重。负相关关系指的是：在其他条件不变的情况下，两个变量变动方向相反，一个变量由大到小变化时，另一个变量由小到大变化。可以理解的是，这里的负相关关系是指变化的方向是相反的，但并不是要求当一个变量有一点变化，另一个变量就必须也变化。

在一些实施例中，第一编码器可以包括第一文本编码器或第二文本编码器中的至少一个，第二编码器可以包括图像编码器或音频编码器中的至少一个。第一编码特征可以包括第一文本特征或第二文本特征中的至少一个，第一文本特征是利用第一文本编码器对文本提取特征进行编码所得到的特征，第二文本特征是利用第二文本编码器对文本提取特征进行编码所得到的特征。第二编码特征可以包括图像编码特征或音频编码特征中的至少一个，图像编码特征是利用图像编码器对第一调整文本特征进行编码所得到的特征，音频编码特征是利用音频编码特征对第二调整文本特征进行编码所得到的特征。融合编码特征可以包括文本图像编码特征或文本音频编码特征中的至少一种。文本图像编码特征是将第一文本编码特征与图像编码特征进行融合所得到的特征。文本音频编码特征是将第二文本编码特征与音频编码特征进行融合所得到的特征。例如，当调整文本特征为第一调整文本特征时，服务器可以将文本提取特征输入到第一文本编码器进行编码，得到第一文本特征，将第一调整文本特征输入到图像编码器进行编码，得到图像编码特征，将第一文本特征与图像编码特征进行融合，得到文本图像编码特征。当调整文本特征为第二调整文本特征时，服务器可以将文本提取特征输入到第二文本编码器进行编码，得到第二文本特征，将第二调整文本特征输入到音频编码器进行编码，得到音频编码特征，将第二文本特征与音频编码特征进行融合，得到文本音频编码特征，可以将文本图像编码特征以及文本音频编码特征作为融合编码特征。其中，第一文本编码器与第二文本编码器可以为同一编码器，也可以为不同的编码器，图像编码器与音频编码器可以为同一编码器，也可以为不同的编码器。

本实施例中，对文本提取特征进行编码，得到第一编码特征，对调整文本特征进行编码，得到第二编码特征，将第一编码特征与第二编码特征进行融合，得到融合编码特征，基于融合编码特征得到调整文本特征对应的调整特征权重，基于调整特征权重对调整文本特征以及文本提取特征进行融合，得到融合文本特征，从而融合文本特征既能反映文本提取特征又能反映调整文本特征，提高了融合文本特征的表达能力，当基于调整文本特征进行识别时，可以提高识别的准确度。

在一些实施例中，第一编码特征是通过已训练的内容识别模型中的第一编码器编码得到的，第二编码特征是通过内容识别模型中的第二编码器编码得到的，基于融合编码特征得到调整文本特征对应的调整特征权重包括：将融合编码特征输入到内容识别模型中的目标激活层进行激活处理，得到目标激活值，将目标激活值作为调整文本特征对应的调整特征权重，激活层为第一编码器与第二编码器的共享激活层。

其中，激活层用于将数据转换为 0 与 1 之间的数据，可以通过激活函数实现，激活函数包括但不限于是 Sigmoid 函数、tanh 函数或 Relu 函数中的至少一种。目标激活层是已训练的内容识别模型中的激活层，是第一编码器与第二编码器共享的激活层，即目标激活层集可以接收第一编码器的输出数据，又可以接收第二编码器的输出的数据。目标激活值是利用目标激活层对融合编码特征进行激活处理所得到的结果，目标激活值与融合编码特征的维度可

以相同，例如可以是同维度的向量或矩阵。如图 5 所示，展示了一个内容识别模块 500，内容识别模块 500 中包括关联特征提取网络、关注强度计算模块、文本特征提取网络、特征调整模块、第一编码器、第二编码器、特征融合模块、目标激活层、融合文本特征生成模块以及内容识别网络，特征融合模块用于将第一编码特征与第二编码特征进行融合得到融合编码特征，融合文本特征生成用于基于调整特征权重对调整文本特征与文本提取特征进行融合，得到融合文本特征。

具体地，目标激活层可以包括第一文本编码器与图像编码器共享的第一激活层、第二文本编码器与音频编码器共享的第二激活层中的至少一个。目标激活值可以包括对文本图像编码特征进行激活得到的第一激活值、或对文本音频编码特征进行激活得到的第二激活值中的至少一个，当融合编码特征为文本图像编码特征时，服务器可以将文本图像编码特征输入到第一激活层进行激活，得到第一激活值，将第一激活值作为第一调整文本特征对应的第一特征权重；当融合编码特征为文本音频编码特征时，服务器可以将文本音频编码特征输入到第二激活层进行激活，得到第二激活值，将第二激活值作为第二调整文本特征对应的第二特征权重，将第一特征权重以及第二特征权重作为调整特征权重。

在一些实施例中，当调整文本特征为第一调整文本特征，并且第一调整文本特征有多个时，服务器可以将第一调整文本特征矩阵与文本提取特征矩阵进行矩阵融合，例如可以将文本提取特征矩阵输入到第一文本编码器进行编码，得到第一矩阵编码特征，第一调整文本特征矩阵输入到图像编码器进行编码，得到第二矩阵编码特征，对第一矩阵编码特征与第二矩阵编码特征进行统计运算，得到第一矩阵特征统计结果，对第一矩阵特征统计结果进行归一化处理，例如可以将第一矩阵特征统计结果输入至第一激活层进行激活，得到归一化后的第一矩阵特征统计结果，归一化后的第一矩阵特征统计结果中可以包括各个第一调整文本特征分别对应的第一特征权重。例如，可以采用公式 (7) 计算得到归一化后的第一矩阵特征统计结果 $gate_1$ 。其中， $gate_1$ 表示归一化后的第一矩阵特征统计结果， $sigmoid$ 为激活函数， W_1^T 为第一文本编码器的模型参数， W_2^T 为图像编码器的模型参数。

$$gate_1 = sigmoid(W_1^T feature_{text} + W_2^T feature_{fusion1}) \quad (7)$$

在一些实施例中，当调整文本特征为第二调整文本特征，并且第二调整文本特征有多个时，服务器可以将第二调整文本特征矩阵与文本提取特征矩阵进行矩阵融合，例如可以将文本提取特征矩阵输入到第二文本编码器进行编码，得到第三矩阵编码特征，将第二调整文本特征矩阵输入到音频编码器进行编码，得到第四矩阵编码特征，对第三矩阵编码特征与第四矩阵编码特征进行统计运算，得到第二矩阵特征统计结果，对第二矩阵特征统计结果进行归一化处理，例如可以将第一矩阵特征统计结果输入至第二激活层进行激活，得到归一化后的第二矩阵特征统计结果，归一化后的第二矩阵特征统计结果中可以包括各个第二调整文本特征分别对应的第二特征权重。例如，可以采用公式 (8) 计算得到归一化后的第二矩阵特征统计结果 $gate_2$ 。其中， $gate_2$ 表示归一化后的第二矩阵特征统计结果， W_3^T 为第二文本编码器的模型参数， W_4^T 为音频编码器的模型参数。

$$gate_2 = sigmoid(W_3^T feature_{text} + W_4^T feature_{fusion2}) \quad (8)$$

本实施例中，将融合编码特征输入到内容识别模型中的目标激活层进行激活处理，得到目标激活值，将目标激活值作为调整文本特征对应的调整特征权重，使得调整特征权重为归一化后的值，提高了调整特征权重的合理性。

在一些实施例中，基于调整特征权重对调整文本特征以及文本提取特征进行融合，得到融合文本特征包括：基于调整特征权重得到文本提取特征对应的文本特征权重；将调整特征权重与调整文本特征进行乘积计算，得到计算后的调整文本特征；将文本特征权重与文本提取特征进行乘积计算，得到计算后的文本提取特征；将计算后的调整文本特征与计算后的文本提取特征进行相加，得到融合文本特征。

其中，文本特征权重可以根据调整特征权重确定，文本特征权重可以与调整特征权重成负相关关系，例如可以将预设数值减去文本特征权重所得到的结果作为文本特征权重。预设数值大于文本特征权重，并且预设数值大于调整特征权重。

具体地，服务器可以将调整特征权重与调整文本特征相乘后的结果，作为计算后的调整文本特征，可以将文本特征权重与文本提取特征相乘后的结果，作为计算后的文本提取特征，可以将计算后的调整文本特征与计算后的文本提取特征进行相加所得到的结果，作为融合文本特征。

在一些实施例中，调整特征权重包括第一特征权重以及第二特征权重，文本特征权重可以包括基于第一特征权重得到的第一文本权重以及基于第二特征权重得到的第二文本权重。第一文本权重与第一特征权重成负相关关系。第二文本权重与第二特征权重成负相关关系。服务器可以利用第一特征权重、第二特征权重、第一文本权重以及第二文本权重，对第一调整文本特征、第二调整文本特征以及文本提取特征进行加权计算，将加权计算的结果作为融合文本特征。例如，服务器可以利用第一特征权重以及第一文本权重对第一调整文本特征以及文本提取特征进行加权计算，得到第一加权值，利用第二特征权重以及第二文本权重对第二调整文本特征以及文本提取特征进行加权计算，得到第二加权值，将第一加权值与第二加权值相加的结果作为融合文本特征。具体地，服务器可以将第一文本权重与文本提取特征进行乘积计算，得到第一乘积值，将第一特征权重与第一调整文本特征进行乘积计算，得到第二乘积值，将第二文本权重与文本提取特征进行乘积计算，得到第三乘积值，将第二特征权重与第二调整文本特征进行乘积计算，得到第四乘积值，将第一乘积值、第二乘积值、第三乘积值以及第四乘积值进行相加，将相加后的结果作为融合文本特征。

在一些实施例中，服务器可以利用归一化后的第一矩阵特征统计结果以及归一化后的第二矩阵特征统计结果，对第一调整文本特征矩阵以及第二调整文本特征矩阵进行加权计算，得到融合文本特征矩阵，融合文本特征矩阵中可以包括各个文本提取特征分别对应的融合文本特征。例如可以利用公式（9）计算得到融合文本特征矩阵 output。其中，output 指的是融合文本特征矩阵。

$$output = feature_{fusion1} \cdot gate_1 + (1 - gate_1) feature_{text} + feature_{fusion2} \cdot gate_2 + (1 - gate_2) feature_{text} \quad (9)$$

本实施例中，将调整特征权重与调整文本特征进行乘积计算，得到计算后的调整文本特征，将文本特征权重与文本提取特征进行乘积计算，得到计算后的文本提取特征，将计算后的调整文本特征与计算后的文本提取特征进行相加，得到融合文本特征，由于文本特征权重是基于调整特征权重得到的，故提高了文本特征权重的准确度，从而提高了融合文本特征的准确度。

在一些实施例中，目标内容为目标视频；从目标内容中，获取目标文本以及与目标文本关联的文本关联数据包括：获取目标视频中目标时间对应的文本，得到目标文本；获取目标视频中目标时间对应的视频相关数据，将视频相关数据作为与目标文本关联的文本关联数据，视频相关数据包括视频帧或者音频帧的至少一种。

其中，视频帧是组成视频的最小单元，视频有多个图像组成，视频中的一张图像称为一帧，也可以称为视频帧。目标视频可以是任意的视频，可以是直接拍摄得到的视频，也可以是从拍摄得到的视频中截取得到的视频片段，目标视频可以是任意类型的视频，包括但不限于是广告类的视频、电视剧类的视频或新闻类视频中的至少一种，目标视频还可以是待推送至用户的视频。目标时间可以是目标视频的起始时间点至终止时间点中的任意的时间点或时间段。视频相关数据指的是目标视频中目标时间展示或播放的任意的数据，可以包括目标视频中目标时间展示的视频帧或目标时间播放的音频帧中的至少一种，目标时间展示的视频帧可以包括一帧或多帧，目标时间播放的音频帧可以包括一帧或多帧。

具体地，服务器可以获取目标视频中目标时间展示的文本，作为目标文本，例如目标时间展示的字幕、弹幕或评论中的至少一种，作为目标文本。服务器可以获取目标视频中目标时间展示的视频帧或目标时间播放的音频帧中的至少一个，作为视频相关数据。

本实施例中，获取目标视频中目标时间对应的文本，得到目标文本，获取目标视频中目标时间对应的视频相关数据，将视频相关数据作为与目标文本关联的文本关联数据，由于视频相关数据包括视频帧或者音频帧的至少一种，从而获取了文本数据以及除文本数据之外的图像数据或音频数据，从而可以在文本数据的基础上结合图像数据或音频数据对视频进行识别，从而有利于提高识别的准确度。

在一些实施例中，调整文本特征包括根据视频帧调整得到的第一调整文本特征；基于调整文本特征进行识别，得到目标内容对应的内容识别结果包括：将第一调整文本特征以及文本提取特征进行融合，得到融合文本特征；基于融合文本特征进行识别，得到目标内容对应的内容识别结果。

具体地，服务器可以从文本关联数据中获取视频帧，对获取的视频帧进行特征提取，得到目标图像特征，基于目标图像特征得到第一调整文本特征，从文本关联数据中获取音频帧，对获取的音频帧进行特征提取，得到目标音频特征，基于目标音频特征得到第二调整文本特征。

在一些实施例中，服务器可以将第一调整文本特征以及文本提取特征进行加权计算，将加权计算的结果作为融合文本特征。例如，服务器可以将第一文本权重与文本提取特征进行乘积计算，得到第一乘积值，将第一特征权重与第一调整文本特征进行乘积计算，得到第二乘积值，将第二文本权重与文本提取特征进行乘积计算，得到第三乘积值，将第一乘积值、第二乘积值以及第三乘积值进行相加，将相加后的结果作为融合文本特征。

本实施例中，将第一调整文本特征以及文本提取特征进行融合，得到融合文本特征，从而使得融合文本特征是基于第一调整文本特征以及文本提取特征这两种特征得到的，从而提高了融合文本特征的特征丰富程度，从而当基于融合文本特征进行识别时，可以提高识别的准确度。

在一些实施例中，调整文本特征还包括根据音频帧调整得到的第二调整文本特征，将第一调整文本特征以及文本提取特征进行融合，得到融合文本特征包括：将第一调整文本特征、第二调整文本特征以及文本提取特征进行融合，得到融合文本特征。

具体地，服务器可以从文本关联数据中获取音频帧，对获取的音频帧进行特征提取，得到目标音频特征，基于目标音频特征得到第二调整文本特征。

在一些实施例中，服务器可以将第一调整文本特征、第二调整文本特征以及文本提取特征进行加权计算，将加权计算的结果作为融合文本特征。例如，服务器可以将第一文本权重

与文本提取特征进行乘积计算，得到第一乘积值，将第一特征权重与第一调整文本特征进行乘积计算，得到第二乘积值，将第二文本权重与文本提取特征进行乘积计算，得到第三乘积值，将第二特征权重与第二调整文本特征进行乘积计算，得到第四乘积值，将第一乘积值、第二乘积值、第三乘积值以及第四乘积值进行相加，将相加后的结果作为融合文本特征。

本实施例中，将第一调整文本特征、第二调整文本特征以及文本提取特征进行融合，得到融合文本特征，从而使得融合文本特征是基于第一调整文本特征、第二调整文本特征以及文本提取特征这三种特征得到的，从而提高了融合文本特征的特征丰富程度，从而当基于融合文本特征进行识别时，可以提高识别的准确度。

在一些实施例中，基于特征关联度对文本提取特征进行调整，得到调整文本特征包括：基于特征关联度得到文本提取特征对应的特征关注强度，特征关联度与特征关注强度成正相关关系；基于特征关注强度对文本提取特征进行调整，得到调整文本特征。

本实施例中，由于特征关联度与特征关注强度成正相关关系，从而基于特征关注强度对文本提取特征进行调整，实现了自适应的根据文本关联数据与目标文本之间的关联程度进行文本特征的调整，从而在基于调整文本特征进行识别时，使得识别结果受到文本关联数据的影响，目标文本与文本关联数据之间的关联程度越大，文本关联数据对识别结果的影响程度越大，使得在内容识别时更加关注关联程度较大的信息，提高了内容识别的准确度。

在一些实施例中，基于特征关注强度对文本提取特征进行调整，得到调整文本特征包括：将特征关注强度与文本提取特征的各个特征值相乘，得到特征值乘积；按照文本提取特征中特征值的位置对特征值乘积进行排列，将排列得到的特征值序列作为调整文本特征。

其中，特征值乘积指的是文本特征值与特征关注强度相乘所得到的结果。特征值序列是按照文本特征值在文本提取特征中的位置对文本特征值计算得到的特征值乘积进行排列得到的。即文本特征值在文本提取特征中的位置与该文本特征值计算得到的特征值乘积在特征值序列中的位置相同。

本实施例中，将特征关注强度与文本提取特征的各个特征值相乘，得到特征值乘积，从而特征值乘积可以反映文本关联数据对文本特征值的关注程度，按照文本提取特征中特征值的位置对特征值乘积进行排列，将排列得到的特征值序列作为调整文本特征，从而调整文本特征可以反映文本关联数据对文本提取特征的关注程度。

在一些实施例中，文本提取特征为目标文本中的分词对应的特征；各个调整文本特征按照分词在目标文本的顺序形成特征序列；基于调整文本特征进行识别，得到目标内容对应的内容识别结果包括：基于特征序列得到各个分词相对于命名实体的位置关系；基于各个位置关系从目标文本中获取目标命名实体，将目标命名实体作为目标内容对应的内容识别结果。

其中，特征序列是按照目标分词在目标文本中的顺序，对目标分词对应的调整文本特征进行排序所得到的序列，目标分词指的是目标文本中的分词。命名实体 (named entity) 指的是以名称为标识的实体，可以包括人名、地名或机构名中的至少一种，命名实体例如可以是“张三”、“A 地区”或“B 机构”。

相对于命名实体的位置关系可以包括命名实体位置或非命名实体位置的至少一种。命名实体位置指的是命名实体所在的位置，可以包括命名实体的起始位置、命名实体的结束位置或命名实体的中间位置的至少一个。命名实体的中间位置可以包括命名实体的起始位置与结束位置之间的各个位置。非命名实体位置指的是命名实体之外的分词所处的位置。

具体地，服务器可以基于特征序列，确定各个目标分词相对于命名实体的位置关系，得

到各个目标分词分别对应的位置关系，从各个位置关系中，获取位置关系属于命名实体位置的位置关系对应的目标分词，作为实体分词，基于各个实体分词得到目标命名实体。

在一些实施例中，已训练的内容识别模型可以包括实体识别网络，服务器可以将特征序列输入到实体识别网络中，利用实体识别网络对特征序列中的各个调整文本特征进行位置识别，例如实体识别网络可以基于调整文本特征确定该调整文本特征对应的目标分词处于命名实体位置的概率，得到命名实体概率，将命名实体概率大于命名实体概率阈值的目标分词的位置关系确定为命名实体位置。命名实体概率阈值可以根据需要设置。实体识别网络还可以基于调整文本特征，确定该调整文本特征对应的目标分词处于命名实体的起始位置的概率，得到起始概率，将起始概率大于起始概率阈值的目标分词的位置关系确定为命名实体的起始位置。起始概率阈值可以根据需要设置。实体识别网络还可以基于调整文本特征，确定该调整文本特征对应的目标分词处于命名实体的结束位置的概率，得到结束概率，将结束概率大于结束概率阈值的目标分词的位置关系确定为命名实体的结束位置。结束概率阈值可以根据需要设置。

本实施例中，基于特征序列得到各个分词相对于命名实体的位置关系，基于各个位置关系从目标文本中获取目标命名实体，将目标命名实体作为目标内容对应的内容识别结果，从而可以基于调整文本特征形成的特征序列进行内容识别，提高了内容识别的准确度。

在一些实施例中，基于各个位置关系从目标文本中获取目标命名实体包括：获取位置关系为命名实体的起始位置的分词，作为命名实体起始词；将命名实体起始词对应的后向分词中，位置关系为处于命名实体的内部的分词作为命名实体组成词；将命名实体起始词与命名实体组成词进行组合，得到目标命名实体。

其中，命名实体起始词指的是处于命名实体的起始位置的分词，命名实体起始词对应的后向分词指的是目标文本中排序在命名实体起始词之后的分词。命名实体组成词指的是目标文本中处于命名实体的内部的分词，命名实体的内部包括命名实体的结束位置以及命名实体的中间位置，命名实体词的结束位置与命名实体的中间位置可以为同一位置。例如，当分词为单个字时，假设目标文本为“张三喜欢花”，由于则命名实体为“张三”，为两个字，由于“张”处于命名实体的起始位置，故命名实体起始词为“张”，命名实体起始词对应的后向分词包括“三”、“喜”、“欢”以及“花”，由于“三”处于命名实体的内部，故命名实体组成词为“三”。目标命名实体是目标文本中包括的实体，是将命名实体起始词与对应的命名实体组成词组合得到的。目标文本中可以包括一个或多个目标命名实体，多个指的是至少两个。例如假设目标文本为“张三喜欢李四”，则目标文本包括2个目标命名实体，分别为“张三”以及“李四”。

具体地，服务器可以基于各个目标分词对应的位置关系，从目标文本中获取位置关系为命名实体的起始位置的分词，作为命名实体起始词，按照从前到后的排列顺序，依次从命名实体起始词的后向分词中获取一个后向分词，作为当前后向分词，当当前后向分词的位置关系为命名实体的内部时，将当前后向分词作为与命名实体起始词对应的命名实体组成词，当当前后向分词的位置关系为命名实体的外部时，停止从命名实体起始词的后向分词中获取后向分词，按照命名实体起始词以及命名实体组成词在目标文本中的位置，从前到后对命名实体起始词以及命名实体组成词进行排序，得到目标命名实体。例如，由于“张”的位置在“三”之前，因此排序得到的是“张三”，即“张三”为目标命名实体。

本实施例中，获取位置关系为命名实体的起始位置的分词，作为命名实体起始词，将命

名实体起始词对应的后向分词中，位置关系为处于命名实体的内部的分词作为命名实体组成词，将命名实体起始词与命名实体组成词进行组合，得到目标命名实体，从而可以基于调整文本特征形成的特征序列进行实体识别，提高了实体识别的准确度。

在一些实施例中，基于特征序列得到各个分词相对于命名实体的位置关系包括：基于特征序列得到各个分词相对于命名实体的位置关系以及分词对应的实体类型；将命名实体起始词对应的后向分词中，位置关系为处于命名实体的内部的分词作为命名实体组成词包括：将命名实体起始词对应的后向分词中，位置关系为处于命名实体的内部，且实体类型与命名实体起始词的类型相同的分词作为命名实体组成词。

其中，实体类型指的是命名实体的类型，包括人名、机构名或地名中的至少一种类型。命名实体起始词以及命名实体组成词可以分别对应有实体类型。

具体地，服务器可以对特征序列中的各个特征进行实体类型的识别，确定特征序列中的各个特征分别对应的实体类型，按照从前到后的排列顺序，依次从命名实体起始词的各个后向分词中获取一个后向分词，作为当前后向分词，当当前后向分词的位置关系为命名实体的内部，并且实体类型与命名实体起始词的实体类型相同时，将当前后向分词作为与命名实体起始词对应的命名实体组成词，当当前后向分词的位置关系为命名实体的外部或实体类型与命名实体起始词的实体类型不同时，停止从命名实体起始词的各个后向分词中获取后向分词。

在一些实施例中，文本提取特征为目标文本中的目标分词对应的特征；各个目标分词对应的融合文本特征按照目标分词在目标文本的顺序形成融合特征序列；基于调整文本特征进行识别，得到目标内容对应的内容识别结果包括：基于融合特征序列得到各个分词相对于命名实体的位置关系；基于各个位置关系从目标文本中获取目标命名实体，将目标命名实体作为目标内容对应的内容识别结果。

在一些实施例中，可以将融合特征序列输入到实体识别网络中，实体识别网络对融合特征序列中的各个融合文本特征进行实体词的识别。实体识别网络例如可以是图 6 中的 CRF 网络，图 6 中，目标文本为“张小花爱笑”，融合特征序列为[h1, h2, h3, h4, h5]，h1 为分词“张”对应的融合文本特征，h2 为分词“小”对应的融合文本特征，h3 为分词“花”对应的融合文本特征，h4 为分词“爱”对应的融合文本特征，h5 为分词“笑”对应的融合文本特征。将融合特征序列输入到 CRF 网络中进行实体识别，CRF 网络可以基于融合特征序列中的各个特征对目标文本中的分词进行打分，得到各个分词对应的分数，可以利用 softmax 对分词的分数进行归一化处理，得到分词对应的概率分布。利用 CRF 网络识别“张小花爱笑”中的人名所处的位置，CRF 网络可以采用“BIO”的标注方法对“张小花爱笑”中的各个目标分词进行标注，得到各个融合文本特征对应的标注，其中，B 为 begin 的缩写，表示实体词开始，I 为 inside 的缩写，表示实体词内部，O 为 outside 的缩写，表示实体词外部，如图所示“张小花爱笑”的标注为“B-PER, I-PER, I-PER, O, O”，其中，“PER”为表示实体词类型为人名。从“B-PER, I-PER, I-PER, O, O”可以确定“张小花爱笑”中的“张小花”为目标命名实体。

本实施例中，基于特征序列得到各个分词相对于命名实体的位置关系以及分词对应的实体类型，将命名实体起始词对应的后向分词中，位置关系为处于命名实体的内部，且实体类型与命名实体起始词的类型相同的分词作为命名实体组成词，提高了实体识别的准确度。

在一些实施例中，确定关联提取特征与文本提取特征间的特征关联度包括：将关联提取

特征中的关联特征值与文本提取特征中对应位置的文本特征值进行乘积运算，得到乘积运算值；对乘积运算值进行统计，得到关联提取特征与文本提取特征之间的特征关联度。

具体地，关联提取特征可以与文本提取特征为同维度的向量或矩阵，服务器可以从关联提取特征中获取目标排序处的关联特征值，作为第一目标特征值，从文本提取特征中获取目标位置处的文本特征值，作为第二目标特征值，则第二目标特征值与第二目标特征值具有位置对应关系，服务器可以对第一目标特征值与第二目标特征值的乘积运算，得到目标位置处的文本特征值与关联特征值计算得到的乘积运算值。目标位置可以是关联提取特征或文本提取特征中的任一位置，例如当关联提取特征为向量时，目标位置可以是任一排序位置，例如第一位。

在一些实施例中，服务器可以对各个乘积运算值进行统计，得到乘积统计值，对乘积统计值进行归一化处理，将归一化处理的结果作为特征关联度。服务器可以将特征关联度作为文本提取特征对应的特征关注强度。

本实施例中，将关联提取特征中的关联特征值与文本提取特征中对应位置的文本特征值进行乘积运算，得到乘积运算值，对各个乘积运算值进行统计运算，得到关联提取特征与文本提取特征之间的特征关联度，从而特征关联度可以准确的反映文本关联数据与目标文本之间的关联关系，从而当基于特征关联度对文本提取特征进行调整时，可以提高调整的准确度。

在一些实施例中，提供了一种内容识别方法，包括以下步骤：

步骤 A、确定待识别的目标视频，从目标视频中，获取目标文本以及与目标文本关联的目标图像数据以及目标音频数据。

步骤 B、对目标文本进行特征提取，得到文本提取特征，对目标图像数据进行特征提取，得到目标图像特征，对目标音频数据进行特征提取，得到目标音频特征。

具体地，如图 7 所示，展示了一个已训练的实体识别网络 700，服务器可以利用已训练的实体识别模型中的文本特征提取网络对目标文本进行特征提取，得到文本提取特征，同样的可以利用图像特征提取网络对目标图像数据进行特征提取，得到目标图像特征，利用音频特征提取网络对目标音频数据进行特征提取，得到目标音频特征。

步骤 C、对目标图像特征与文本提取特征进行关联计算，得到图像关联度，将图像关联度作为图像关注强度，对目标音频特征与文本提取特征进行关联计算，得到音频关联度，将音频关联度作为音频关注强度。

具体地，如图 7 所示，可以利用图像关注强度计算模块，对目标图像特征与文本提取特征进行关联计算，得到图像关注强度，利用音频关注强度计算模块，对目标音频特征与文本提取特征进行关联计算得到音频关注强度。图像关注强度计算模块中包括乘积运算单元以及归一化处理单元，图像关注强度计算模块可以通过乘积运算单元对目标图像特征与文本提取特征进行乘积运算，将运算的结果输入到归一化运算单元进行归一化处理，得到图像关注强度。音频关注强度计算模块计算得到音频关注强度的过程可以参考图像关注强度计算模块。

步骤 D、基于图像关注强度对文本提取特征进行调整，得到第一调整文本特征，基于音频关注强度对文本提取特征进行调整，得到第二调整文本特征。

具体地，如图 7 所示，可以将图像关注强度以及文本提取特征输入第一特征调整模块，第一特征调整模块可以将图像关注强度与文本提取特征各个特征值相乘，按照文本提取特征中特征值的位置对相乘得到的各个值进行排列得到第一调整文本特征。同样的，可以利用第二特征调整模块得到第二调整文本特征。

步骤 E、确定第一调整文本特征对应的第一特征权重，确定第二调整文本特征对应的第二特征权重。

具体地，如图 7 所示，服务器可以将第一调整文本特征输入到图像编码器进行编码，得到图像编码特征，将文本提取特征输入到第一文本编码器进行编码，得到第一文本特征，将第一文本特征以及图像编码特征输入到第一特征融合模块中，得到文本图像编码特征。服务器可以将第二调整文本特征输入到音频编码器进行编码，得到音频编码特征，将文本提取特征输入到第二文本编码器进行编码，得到第二文本特征，将第二文本特征以及音频编码特征输入到第二特征融合模块中，得到文本音频编码特征，将文本音频编码特征输入到第一激活层进行激活，得到第一调整文本特征对应的第一特征权重，将文本音频编码特征输入到第二激活层进行激活，得到第二调整文本特征对应的第二特征权重。

步骤 F、基于第一特征权重对第一调整文本特征以及文本提取特征进行融合，得到第一融合特征，基于第二特征权重对第二调整文本特征以及文本提取特征进行融合，得到第二融合特征，将第一融合特征与第二融合特征进行统计运算，将统计运算的结果作为融合文本特征。

具体地，如图 7 所示，服务器可以将第一特征权重、第一调整文本特征以及文本提取特征输入到第一融合文本特征生成模块中，得到第一融合特征，将第二特征权重、第二调整文本特征以及文本提取特征输入到第二融合文本特征生成模块中，得到第二融合特征。

步骤 G、基于融合文本特征进行命名实体识别，得到目标内容对应的目标命名实体，将目标命名实体作为目标内容对应的内容识别结果。

例如，如图 8 所示，目标视频为“张小花”的视频，目标文本为“张小花”的视频中的字幕“张小花爱笑”，目标图像数据为“张小花”的视频中与字幕“张小花爱笑”在时间上关联的图像，即包括“张小花”的图像，目标音频数据为“张小花”的视频中与字幕“张小花爱笑”在时间上关联的音频，即包括“张小花”的音频，将字幕“张小花爱笑”、包括“张小花”的图像以及包括“张小花”的音频输入到实体识别模型中，可以确定实体词“张小花”。

上述的内容识别方法在进行实体识别时，除了利用视频中的文本信息，例如视频中的标题、字幕或描述信息外，还利用了视频的音频特征以及图像特征，并且将多种模态特征进行融合，能更加精准有效的提取视频信息，增强实体词识别的识别效果，例如提高了实体词识别的准确度和效率。可以提升测试数据集上的准确率和召回率。其中，一种模态可以是一种数据类型，例如文本、音频以及图像分别一种模态，多模态包括至少两种模态。模态特征例如可以是文本特征、音频特征或图像特征中的任意一种。多模态特征包括至少两种模态特征。本申请提供的实体词识别模型（即实体识别模型），可以对视频信息进行有效的提取。

本申请还提供一种应用场景，该应用场景应用上述的内容识别方法，可以对视频中的文本进行实体识别。具体地，该内容识别方法在该应用场景的应用如下：

接收针对目标视频的视频标签生成请求，响应于视频标签生成请求，使用本申请提供的内容识别方法，对目标视频进行实体词识别，得到识别出的实体词，将识别出的各个实体词作为目标视频对应的视频标签。

本申请提供的内容识别方法，应用于视频识别，可以节省获取视频信息的时间，提高理解视频信息的效率。

本申请还提供一种应用场景，该应用场景应用上述的内容识别方法，可以对视频中的文

本进行实体识别。具体地，该内容识别方法在该应用场景的应用如下：

接收针对目标用户对应的视频推荐请求，获取候选视频，利用本申请提供的内容识别方法对候选视频进行实体词识别，将识别出的实体词作为候选视频对应的视频标签，获取目标用户对应的用户信息，当确定视频标签与用户信息匹配时，例如视频标签与用户的用户画像匹配时，向目标用户对应的终端推送该候选视频。

本申请提供的内容识别方法，应用于视频推荐中，可以为视频推荐算法提供优质特征，优化视频推荐效果。

应该理解的是，虽然上述各实施例的流程图中的各个步骤按照箭头的指示依次显示，但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明，这些步骤的执行并没有严格的顺序限制，这些步骤可以以其它的顺序执行。而且，上述各实施例中的至少一部分步骤可以包括多个步骤或者多个阶段，这些步骤或者阶段并不必然是在同一时刻执行完成，而是可以在不同的时刻执行，这些步骤或者阶段的执行顺序也不必然是依次进行，而是可以与其它步骤或者其它步骤中的步骤或者阶段的至少一部分轮流或者交替地执行。

在一些实施例中，如图9所示，提供了一种内容识别装置，该装置可以采用软件模块或硬件模块，或者是二者的结合成为计算机设备的一部分，该装置具体包括：目标内容确定模块902、特征提取模块904、特征关联度得到模块906、调整文本特征得到模块908和内容识别结果得到模块910，其中：目标内容确定模块902，用于确定待识别的目标内容，从目标内容中，获取目标文本以及与目标文本关联的文本关联数据。特征提取模块904，用于对目标文本进行特征提取，得到文本提取特征；对文本关联数据进行特征提取，得到关联提取特征。特征关联度得到模块906，用于确定关联提取特征与文本提取特征间的特征关联度；特征关联度用于表征目标文本与文本关联数据之间的关联程度。调整文本特征得到模块908，用于基于特征关联度对文本提取特征进行调整，得到调整文本特征。内容识别结果得到模块910，用于基于调整文本特征进行识别，得到目标内容对应的内容识别结果。

上述内容识别装置中，确定待识别的目标内容，从目标内容中，获取目标文本以及与目标文本关联的文本关联数据，对目标文本进行特征提取，得到文本提取特征，对文本关联数据进行特征提取，得到关联提取特征，确定关联提取特征与文本提取特征间的特征关联度，基于特征关联度对文本提取特征进行调整，得到调整文本特征，基于调整文本特征进行识别，得到目标内容对应的内容识别结果。由于特征关联度反映了目标文本与文本关联数据之间的关联程度的大小，特征关联度越大，则目标文本与文本关联数据之间的关联程度越大，特征关联度越小，则目标文本与文本关联数据之间的关联程度越小，从而在基于调整文本特征进行识别时，目标文本与文本关联数据之间的关联程度越大，文本关联数据对识别结果的影响程度越大，目标文本与文本关联数据之间的关联程度越小，文本关联数据对识别结果的影响程度越小，从而可以自适应的根据文本关联数据与目标文本之间的关系，调整用于识别时的特征，提高识别时所用的特征的准确度，提高内容识别的准确度。

在一些实施例中，内容识别结果得到模块910包括：第一融合文本特征得到单元，用于对调整文本特征以及文本提取特征进行融合，得到融合文本特征。第一内容识别结果得到单元，用于基于融合文本特征进行识别，得到目标内容对应的内容识别结果。

本实施例中，对调整文本特征以及文本提取特征进行融合，得到融合文本特征，基于融合文本特征进行识别，得到目标内容对应的内容识别结果，可以提高内容识别的准确度。

在一些实施例中，第一融合文本特征得到单元，还用于对文本提取特征进行编码，得到

第一编码特征，对调整文本特征进行编码，得到第二编码特征；将第一编码特征与第二编码特征进行融合，得到融合编码特征；基于融合编码特征得到调整文本特征对应的调整特征权重；基于调整特征权重对调整文本特征以及文本提取特征进行融合，得到融合文本特征。

本实施例中，对文本提取特征进行编码，得到第一编码特征，对调整文本特征进行编码，得到第二编码特征，将第一编码特征与第二编码特征进行融合，得到融合编码特征，基于融合编码特征得到调整文本特征对应的调整特征权重，基于调整特征权重对调整文本特征以及文本提取特征进行融合，得到融合文本特征，从而融合文本特征既能反映文本提取特征又能反映调整文本特征，提高了融合文本特征的表达能力，当基于调整文本特征进行识别时，可以提高识别的准确度。

在一些实施例中，第一编码特征是通过已训练的内容识别模型中的第一编码器编码得到的，第二编码特征是通过内容识别模型中的第二编码器编码得到的，第一融合文本特征得到单元，还用于将融合编码特征输入到内容识别模型中的目标激活层进行激活处理，得到激活值，将激活值作为调整文本特征对应的调整特征权重，激活层为第一编码器与第二编码器的共享激活层。

本实施例中，将融合编码特征输入到内容识别模型中的目标激活层进行激活处理，得到目标激活值，将目标激活值作为调整文本特征对应的调整特征权重，使得调整特征权重为归一化后的值，提高了调整特征权重的合理性。

在一些实施例中，第一融合文本特征得到单元，还用于基于调整特征权重得到文本提取特征对应的文本特征权重；将调整特征权重与调整文本特征进行乘积计算，得到计算后的调整文本特征；将文本特征权重与文本提取特征进行乘积计算，得到计算后的文本提取特征；将计算后的调整文本特征与计算后的文本提取特征进行相加，得到融合文本特征。

本实施例中，将调整特征权重与调整文本特征进行乘积计算，得到计算后的调整文本特征，将文本特征权重与文本提取特征进行乘积计算，得到计算后的文本提取特征，将计算后的调整文本特征与计算后的文本提取特征进行相加，得到融合文本特征，由于文本特征权重是基于调整特征权重得到的，故提高了文本特征权重的准确度，从而提高了融合文本特征的准确度。

在一些实施例中，目标内容为目标视频；目标内容确定模块 902 包括：目标文本得到单元，用于获取目标视频中目标时间对应的文本，得到目标文本。文本关联数据得到单元，用于获取目标视频中目标时间对应的视频相关数据，将视频相关数据作为与目标文本关联的文本关联数据，视频相关数据包括视频帧或者音频帧的至少一种。

本实施例中，获取目标视频中目标时间对应的文本，得到目标文本，获取目标视频中目标时间对应的视频相关数据，将视频相关数据作为与目标文本关联的文本关联数据，由于视频相关数据包括视频帧或者音频帧的至少一种，从而获取了文本数据以及除文本数据之外的图像数据或音频数据，从而可以在文本数据的基础上结合图像数据或音频数据对视频进行识别，从而有利于提高识别的准确度。

在一些实施例中，调整文本特征包括根据视频帧调整得到的第一调整文本特征；内容识别结果得到模块 910 包括：第二融合文本特征得到单元，用于将第一调整文本特征以及文本提取特征进行融合，得到融合文本特征。第二内容识别结果得到单元，用于基于融合文本特征进行识别，得到目标内容对应的内容识别结果。

本实施例中，将第一调整文本特征、第二调整文本特征以及文本提取特征进行融合，得

到融合文本特征，从而使得融合文本特征是基于第一调整文本特征、第二调整文本特征以及文本提取特征这三种特征得到的，从而提高了融合文本特征的特征丰富程度，从而当基于融合文本特征进行识别时，可以提高识别的准确度。

在一些实施例中，调整文本特征还包括根据音频帧调整得到的第二调整文本特征，第二融合文本特征得到单元还用于：将第一调整文本特征、第二调整文本特征以及文本提取特征进行融合，得到融合文本特征。

本实施例中，将第一调整文本特征、第二调整文本特征以及文本提取特征进行融合，得到融合文本特征，从而使得融合文本特征是基于第一调整文本特征、第二调整文本特征以及文本提取特征这三种特征得到的，从而提高了融合文本特征的特征丰富程度，从而当基于融合文本特征进行识别时，可以提高识别的准确度。

在一些实施例中，基于特征关联度对文本提取特征进行调整，得到调整文本特征包括：基于特征关联度得到文本提取特征对应的特征关注强度，特征关联度与特征关注强度成正相关关系；基于特征关注强度对文本提取特征进行调整，得到调整文本特征。

本实施例中，由于特征关联度与特征关注强度成正相关关系，故目标文本与文本关联数据之间的关联程度越大，特征关注强度越大，对文本提取特征调整的程度越大，目标文本与文本关联数据之间的关联程度越小，特征关注强度越小，对文本提取特征调整的程度越小，从而基于特征关注强度对文本提取特征进行调整，实现了自适应的根据文本关联数据与目标文本之间的关联程度进行文本特征的调整，从而在基于调整文本特征进行识别时，使得识别结果受到文本关联数据的影响，目标文本与文本关联数据之间的关联程度越大，文本关联数据对识别结果的影响程度越大，使得在内容识别时更加关注关联程度较大的信息，提高了内容识别的准确度。

在一些实施例中，调整文本特征得到模块 908 包括：特征值乘积得到单元，用于将特征关注强度与文本提取特征的特征值相乘，得到特征值乘积。调整文本特征得到单元，用于按照文本提取特征中特征值的位置对特征值乘积进行排列，将排列得到的特征值序列作为调整文本特征。

本实施例中，将特征关注强度与文本提取特征的特征值相乘，得到特征值乘积，从而特征值乘积可以反映文本关联数据对文本特征值的关注程度，按照文本提取特征中特征值的排序对特征值乘积进行排序，将排序得到的特征值序列作为调整文本特征，从而调整文本特征可以反映文本关联数据对文本提取特征的关注程度。

在一些实施例中，文本提取特征为目标文本中的分词对应的特征；各个调整文本特征按照分词在目标文本的顺序形成特征序列；内容识别结果得到模块 910 包括：位置关系得到单元，用于基于特征序列得到各个分词相对于命名实体的位置关系。第三内容识别结果得到单元，用于基于各个位置关系从目标文本中获取目标命名实体，将目标命名实体作为目标内容对应的内容识别结果。

本实施例中，基于特征序列得到各个分词相对于命名实体的位置关系，基于各个位置关系从目标文本中获取目标命名实体，将目标命名实体作为目标内容对应的内容识别结果，从而可以基于调整文本特征形成的特征序列进行内容识别，提高了内容识别的准确度。

在一些实施例中，第三内容识别结果得到单元，还用于获取位置关系为命名实体的起始位置的分词，作为命名实体起始词；将命名实体起始词对应的后向分词中，位置关系为处于命名实体的内部的分词作为命名实体组成词；将命名实体起始词与命名实体组成词进行组合，

得到目标命名实体。

本实施例中，获取位置关系为命名实体的起始位置的分词，作为命名实体起始词，将命名实体起始词对应的后向分词中，位置关系为处于命名实体的内部的分词作为命名实体组成词，将命名实体起始词与命名实体组成词进行组合，得到目标命名实体，从而可以基于调整文本特征形成的特征序列进行实体识别，提高了实体识别的准确度。

在一些实施例中，位置关系得到单元，还用于基于特征序列得到各个分词相对于命名实体的位置关系以及分词对应的实体类型；第三内容识别结果得到单元，还用于将命名实体起始词对应的后向分词中，位置关系为处于命名实体的内部，且实体类型与命名实体起始词的类型相同的分词作为命名实体组成词。

本实施例中，基于特征序列得到各个分词相对于命名实体的位置关系以及分词对应的实体类型，将命名实体起始词对应的后向分词中，位置关系为处于命名实体的内部，且实体类型与命名实体起始词的类型相同的分词作为命名实体组成词，提高了实体识别的准确度。

在一些实施例中，特征关联度得到模块 906 包括：乘积运算值得到单元，用于将关联提取特征中的关联特征值与文本提取特征中对应位置的文本特征值进行乘积运算，得到乘积运算值。特征关注强度得到单元，用于对乘积运算值进行统计，得到关联提取特征与文本提取特征之间的特征关联度。

本实施例中，将关联提取特征中的关联特征值与文本提取特征中对应位置的文本特征值进行乘积运算，得到乘积运算值，对各个乘积运算值进行统计运算，得到关联提取特征与文本提取特征之间的特征关联度，从而特征关联度可以准确的反映文本关联数据与目标文本之间的关联程度，从而当基于特征关联度对文本提取特征进行调整时，可以提高调整的准确度。

关于内容识别装置的具体限定可以参见上文中对于内容识别方法的限定。上述内容识别装置中的各个模块可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中，也可以以软件形式存储于计算机设备中的存储器中，以便于处理器调用执行以上各个模块对应的操作。

在一些实施例中，提供了一种计算机设备，该计算机设备可以是服务器，其内部结构图可以如图 10 所示。该计算机设备包括通过系统总线连接的处理器、存储器和网络接口。其中，该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统、计算机可读指令和数据库。该内存储器为非易失性存储介质中的操作系统和计算机可读指令的运行提供环境。该计算机设备的数据库用于存储内容识别数据。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机可读指令被处理器执行时以实现一种内容识别方法。

在一些实施例中，提供了一种计算机设备，该计算机设备可以是终端，其内部结构图可以如图 11 所示。该计算机设备包括通过系统总线连接的处理器、存储器、通信接口、显示屏和输入装置。其中，该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统和计算机可读指令。该内存储器为非易失性存储介质中的操作系统和计算机可读指令的运行提供环境。该计算机设备的通信接口用于与外部的终端进行有线或无线方式的通信，无线方式可通过 WIFI、运营商网络、NFC（近场通信）或其他技术实现。该计算机可读指令被处理器执行时以实现一种内容识别方法。该计算机设备的显示屏可以是液晶显示屏或者电子墨水显示屏，该计算机设备的输入装置可以是显示屏上覆盖的触摸层，也可以是计算机设备外壳上设置的按键、轨

迹球或触控板，还可以是外接的键盘、触控板或鼠标等。

本领域技术人员可以理解，图 10 和图 11 中示出的结构，仅仅是与本申请方案相关的部分结构的框图，并不构成对本申请方案所应用于其上的计算机设备的限定，具体的计算机设备可以包括比图中所示更多或更少的部件，或者组合某些部件，或者具有不同的部件布置。

在一些实施例中，还提供了一种计算机设备，包括存储器和一个或多个处理器，存储器中存储有计算机可读指令，该计算机可读指令被处理器执行时，使得一个或多个处理器执行计算机可读指令时实现上述各方法实施例中的步骤。

在一些实施例中，提供了一个或多个非易失性可读存储介质，存储有计算机可读指令，该计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器实现上述各方法实施例中的步骤。

一种计算机程序产品，包括计算机可读指令，所述计算机可读指令被处理器执行时实现上述图像处理方法的步骤。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机可读指令来指令相关的硬件来完成，所述的计算机可读指令可存储于一非易失性计算机可读存储介质中，该计算机可读指令在执行时，可包括如上述各方法的实施例的流程。其中，本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用，均可包括非易失性和易失性存储器中的至少一种。非易失性存储器可包括只读存储器（Read-Only Memory, ROM）、磁带、软盘、闪存或光存储器等。易失性存储器可包括随机存取存储器（Random Access Memory, RAM）或外部高速缓冲存储器。作为说明而非局限，RAM 可以是多种形式，比如静态随机存取存储器（Static Random Access Memory, SRAM）或动态随机存取存储器（Dynamic Random Access Memory, DRAM）等。

以上实施例的各技术特征可以进行任意的组合，为使描述简洁，未对上述实施例中的各个技术特征所有可能的组合都进行描述，然而，只要这些技术特征的组合不存在矛盾，都应当认为是本说明书记载的范围。

以上所述实施例仅表达了本申请的几种实施方式，其描述较为具体和详细，但并不能因此而理解为对发明专利范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本申请构思的前提下，还可以做出若干变形和改进，这些都属于本申请的保护范围。因此，本申请专利的保护范围应以所附权利要求为准。

权利要求书

1、一种内容识别方法，其特征在于，由计算机设备执行，所述方法包括：

确定待识别的目标内容，从所述目标内容中，获取目标文本以及与所述目标文本关联的文本关联数据；

对所述目标文本进行特征提取，得到文本提取特征；对所述文本关联数据进行特征提取，得到关联提取特征；

确定所述关联提取特征与所述文本提取特征间的特征关联度；所述特征关联度用于表征所述目标文本与所述文本关联数据之间的关联程度；基于所述特征关联度对所述文本提取特征进行调整，得到调整文本特征；及，

基于所述调整文本特征进行识别，得到所述目标内容对应的内容识别结果。

2、根据权利要求1所述的方法，其特征在于，所述基于所述调整文本特征进行识别，得到所述目标内容对应的内容识别结果包括：

对所述调整文本特征以及所述文本提取特征进行融合，得到融合文本特征；及，

基于所述融合文本特征进行识别，得到所述目标内容对应的内容识别结果。

3、根据权利要求2所述的方法，其特征在于，所述对所述调整文本特征以及所述文本提取特征进行融合，得到融合文本特征包括：

对所述文本提取特征进行编码，得到第一编码特征，对所述调整文本特征进行编码，得到第二编码特征；

将所述第一编码特征与所述第二编码特征进行融合，得到融合编码特征；

基于所述融合编码特征得到所述调整文本特征对应的调整特征权重；及，

基于所述调整特征权重对所述调整文本特征以及所述文本提取特征进行融合，得到融合文本特征。

4、根据权利要求3所述的方法，其特征在于，所述第一编码特征是通过已训练的内容识别模型中的第一编码器编码得到的，所述第二编码特征是通过所述内容识别模型中的第二编码器编码得到的，所述基于所述融合编码特征得到所述调整文本特征对应的调整特征权重包括：

将所述融合编码特征输入到所述内容识别模型中的目标激活层进行激活处理，得到目标激活值，将所述目标激活值作为所述调整文本特征对应的调整特征权重，所述激活层为所述第一编码器与所述第二编码器的共享激活层。

5、根据权利要求3所述的方法，其特征在于，所述基于所述调整特征权重对所述调整文本特征以及所述文本提取特征进行融合，得到融合文本特征包括：

基于所述调整特征权重得到所述文本提取特征对应的文本特征权重；

将所述调整特征权重与所述调整文本特征进行乘积计算，得到计算后的调整文本特征；

将所述文本特征权重与所述文本提取特征进行乘积计算，得到计算后的文本提取特征；

及，

将所述计算后的调整文本特征与所述计算后的文本提取特征进行相加，得到融合文本特征。

6、根据权利要求1所述的方法，其特征在于，所述目标内容为目标视频；所述从所述目标内容中，获取目标文本以及与所述目标文本关联的文本关联数据包括：

获取所述目标视频中目标时间对应的文本，得到目标文本；及，

获取所述目标视频中所述目标时间对应的视频相关数据，将所述视频相关数据作为与所述目标文本关联的文本关联数据，所述视频相关数据包括视频帧或者音频帧的至少一种。

7、根据权利要求6所述的方法，其特征在于，所述调整文本特征包括根据所述视频帧调整得到的第一调整文本特征；

所述基于所述调整文本特征进行识别，得到所述目标内容对应的内容识别结果包括：

将所述第一调整文本特征以及所述文本提取特征进行融合，得到融合文本特征；及，

基于所述融合文本特征进行识别，得到所述目标内容对应的内容识别结果。

8、根据权利要求7所述的方法，其特征在于，所述调整文本特征还包括根据所述音频帧调整得到的第二调整文本特征，所述将所述第一调整文本特征以及所述文本提取特征进行融合，得到融合文本特征包括：

将所述第一调整文本特征、所述第二调整文本特征以及所述文本提取特征进行融合，得到融合文本特征。

9、根据权利要求1所述的方法，其特征在于，所述基于所述特征关联度对所述文本提取特征进行调整，得到调整文本特征包括：

基于所述特征关联度得到所述文本提取特征对应的特征关注强度，所述特征关联度与所述特征关注强度成正相关关系；

基于所述特征关注强度对所述文本提取特征进行调整，得到调整文本特征。

10、根据权利要求9所述的方法，其特征在于，所述基于所述特征关注强度对所述文本提取特征进行调整，得到调整文本特征包括：

将所述特征关注强度与所述文本提取特征的各个特征值相乘，得到特征值乘积；及，

按照所述文本提取特征中特征值的位置对所述特征值乘积进行排列，将排列得到的特征值序列作为所述调整文本特征。

11、根据权利要求1所述的方法，其特征在于，所述文本提取特征为所述目标文本中的分词对应的特征；各个调整文本特征按照分词在所述目标文本的顺序形成特征序列；所述基于所述调整文本特征进行识别，得到所述目标内容对应的内容识别结果包括：

基于所述特征序列得到各个所述分词相对于命名实体的位置关系；及，

基于各个所述位置关系从所述目标文本中获取目标命名实体，将所述目标命名实体作为所述目标内容对应的内容识别结果。

12、根据权利要求11所述的方法，其特征在于，所述基于各个所述位置关系从所述目标文本中获取目标命名实体包括：

获取位置关系为命名实体的起始位置的分词，作为命名实体起始词；

将所述命名实体起始词对应的后向分词中，位置关系为处于命名实体的内部的分词作为命名实体组成词；及，

将所述命名实体起始词与所述命名实体组成词进行组合，得到目标命名实体。

13、根据权利要求12所述的方法，其特征在于，所述基于所述特征序列得到各个所述分词相对于命名实体的位置关系包括：

基于所述特征序列得到各个所述分词相对于命名实体的位置关系以及所述分词对应的实体类型；及，

所述将所述命名实体起始词对应的后向分词中，位置关系为处于命名实体的内部的分词作为命名实体组成词包括：

将所述命名实体起始词对应的后向分词中，位置关系为处于命名实体的内部，且实体类型与所述命名实体起始词的类型相同的分词作为命名实体组成词。

14、根据权利要求 1 所述的方法，其特征在于，所述确定所述关联提取特征与所述文本提取特征间的特征关联度包括：

将所述关联提取特征中的关联特征值与所述文本提取特征中对应位置的文本特征值进行乘积运算，得到乘积运算值；及，

对所述乘积运算值进行统计，得到所述关联提取特征与所述文本提取特征之间的特征关联度。

15、一种内容识别装置，其特征在于，所述装置包括：

目标内容确定模块，用于确定待识别的目标内容，从所述目标内容中，获取目标文本以及与所述目标文本关联的文本关联数据；

特征提取模块，用于对所述目标文本进行特征提取，得到文本提取特征；对所述文本关联数据进行特征提取，得到关联提取特征；

特征关联度得到模块，用于确定所述关联提取特征与所述文本提取特征间的特征关联度；所述特征关联度用于表征所述目标文本与所述文本关联数据之间的关联程度；

调整文本特征得到模块，用于基于所述特征关联度对所述文本提取特征进行调整，得到调整文本特征；及，

内容识别结果得到模块，用于基于所述调整文本特征进行识别，得到所述目标内容对应的内容识别结果。

16、一种计算机设备，包括存储器和一个或多个处理器，所述存储器存储有计算机可读指令，其特征在于，所述计算机可读指令被所述处理器执行时，使得所述一个或多个处理器执行权利要求 1 至 14 中任一项所述的方法的步骤。

17、一个或多个非易失性可读存储介质，存储有计算机可读指令，其特征在于，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器实现权利要求 1 至 14 中任一项所述的方法的步骤。

18、一种计算机程序产品，包括计算机可读指令，所述计算机可读指令被处理器执行时实现权利要求 1 至 14 中任一项所述的方法的步骤。

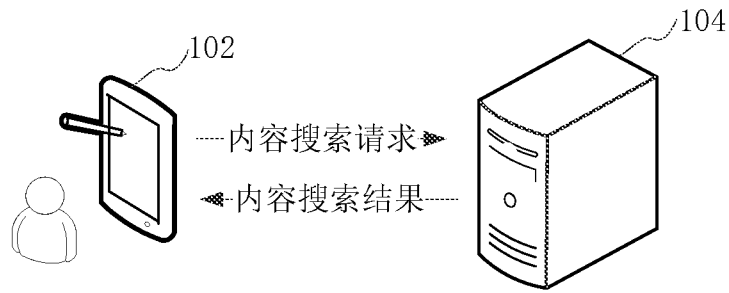


图 1

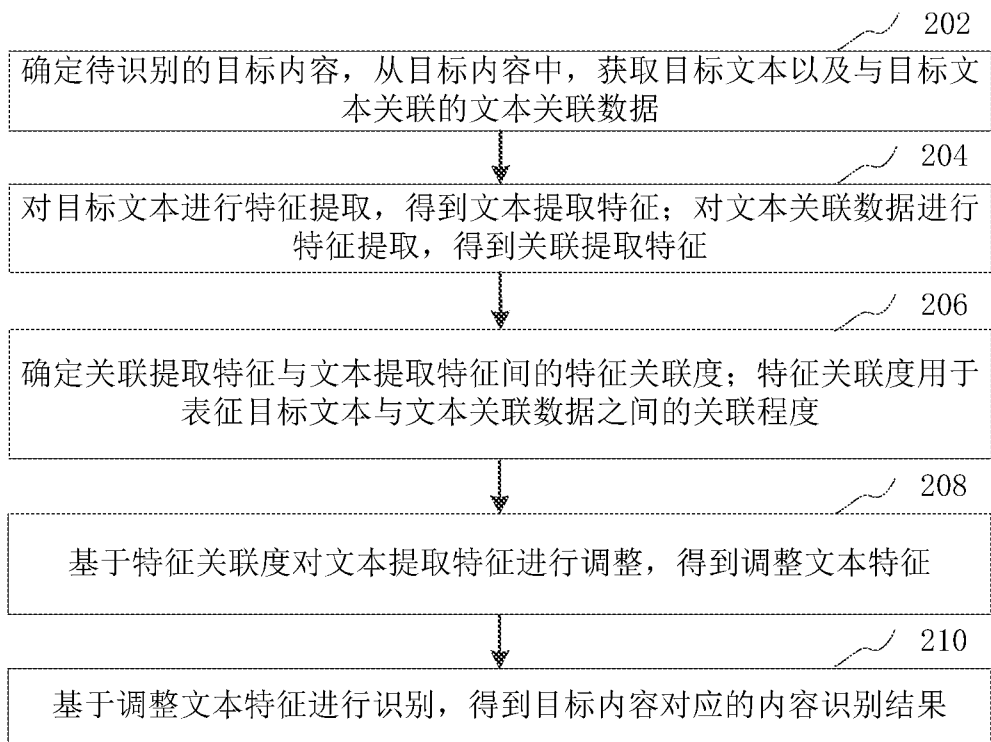


图 2

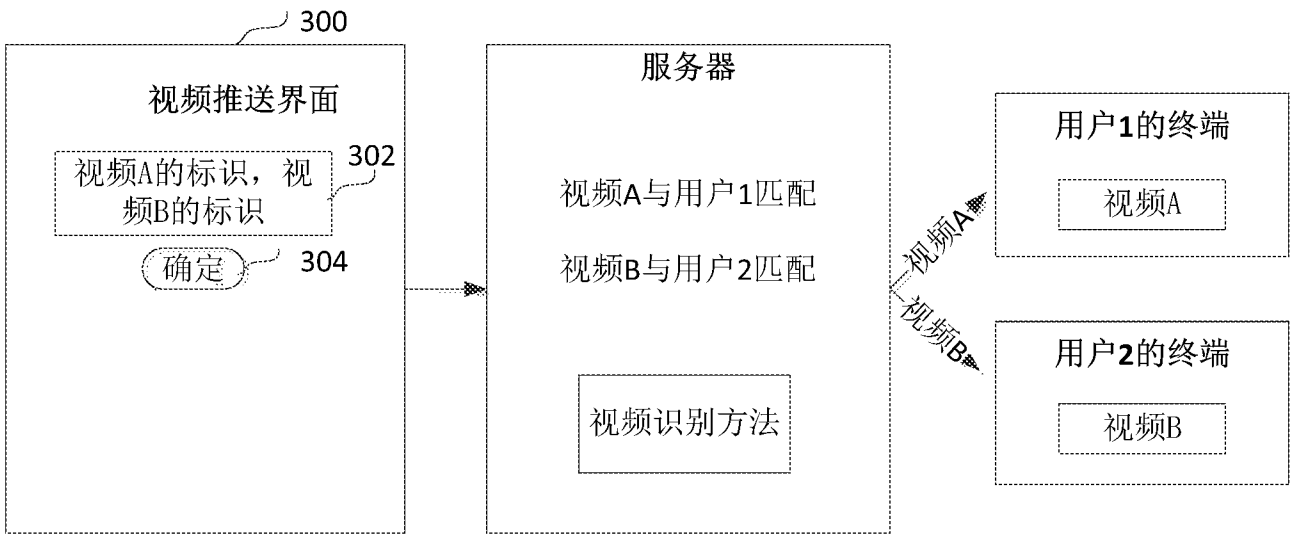


图 3

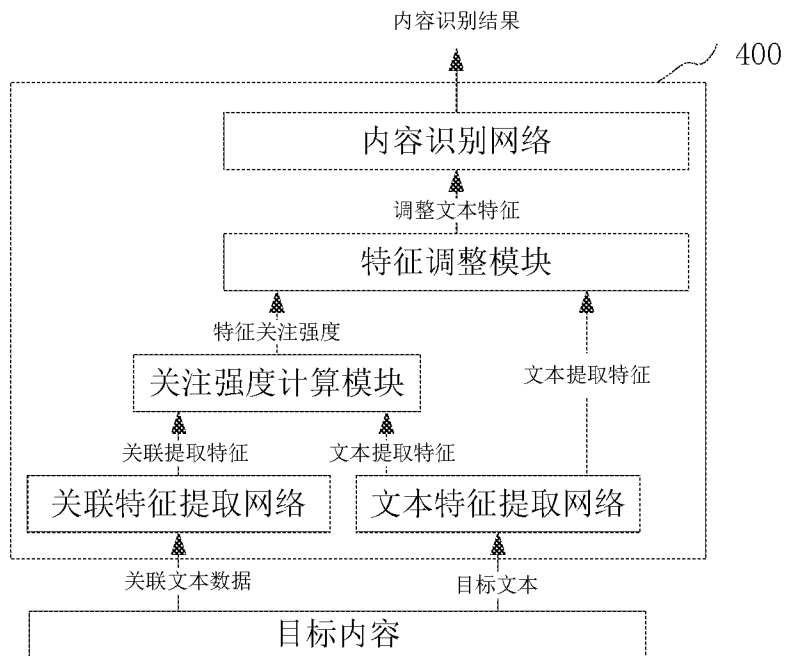


图 4

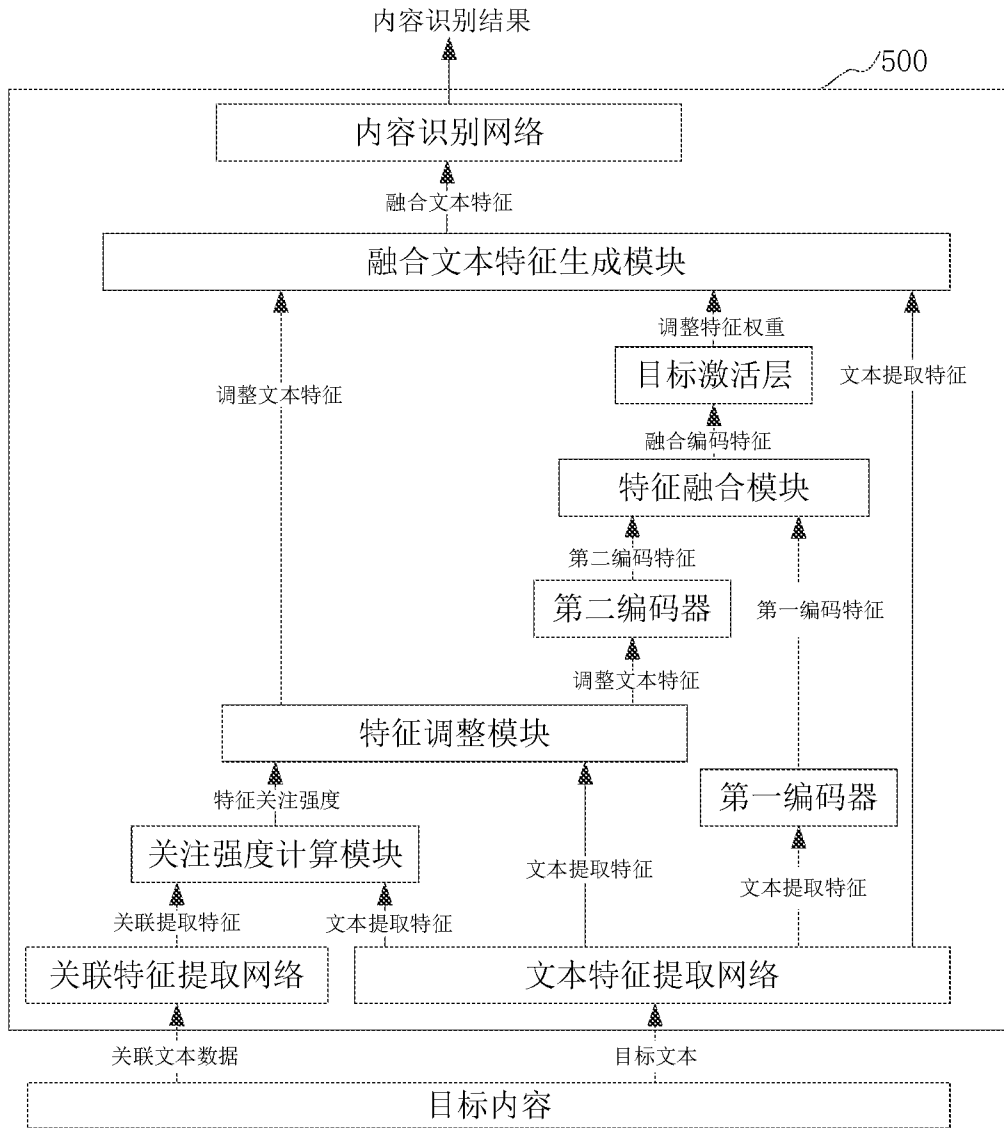


图 5

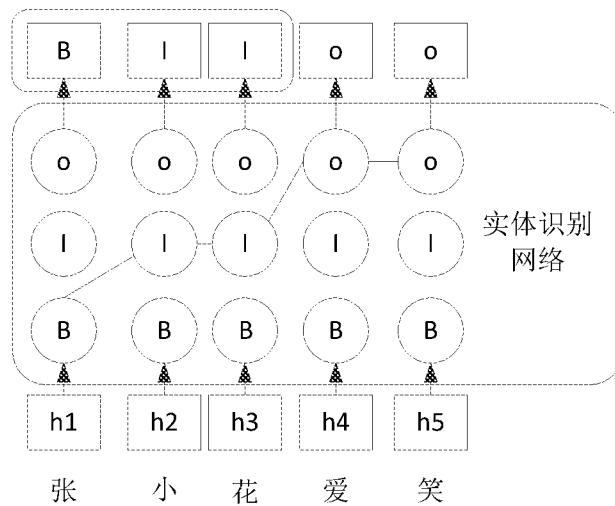


图 6

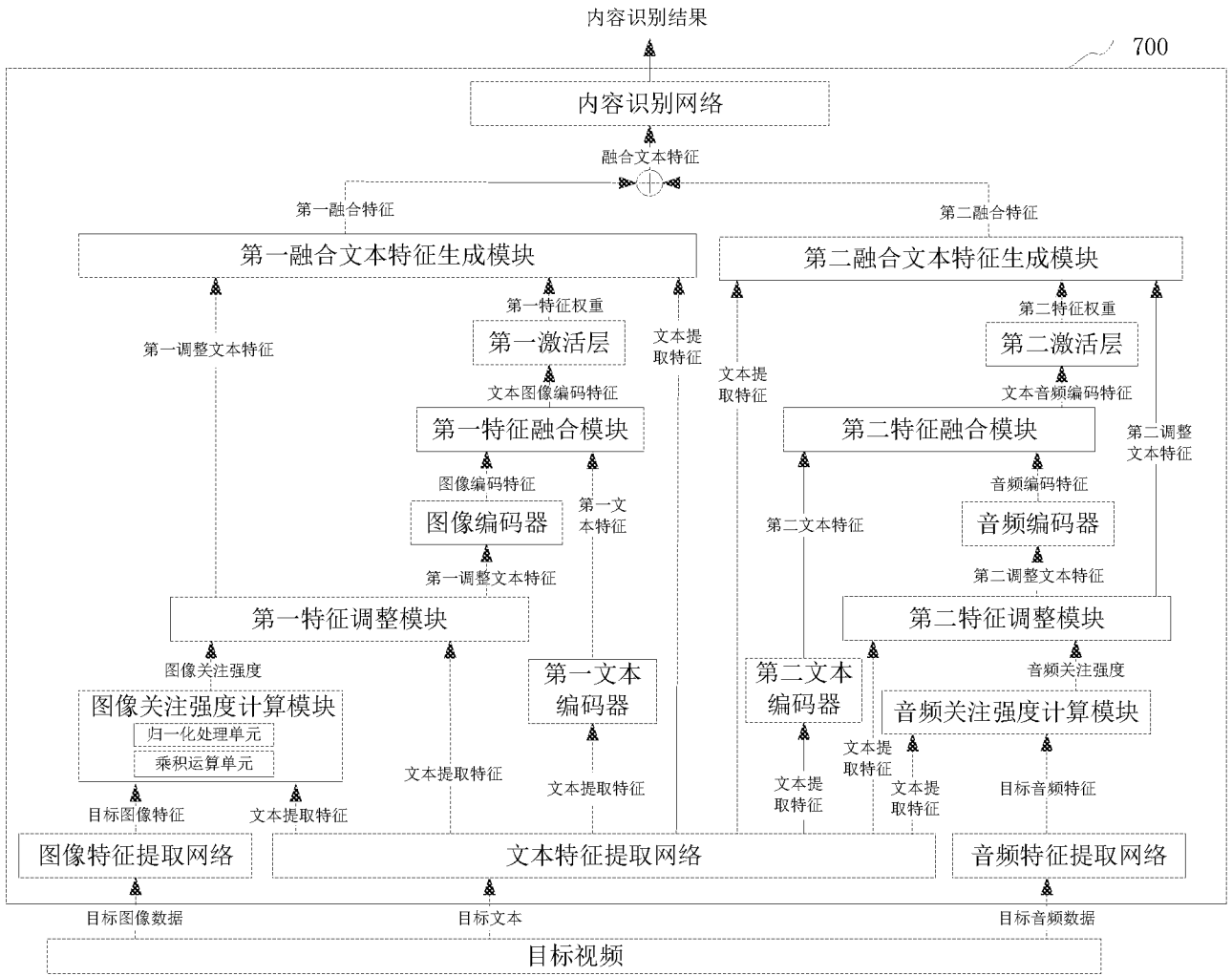


图 7

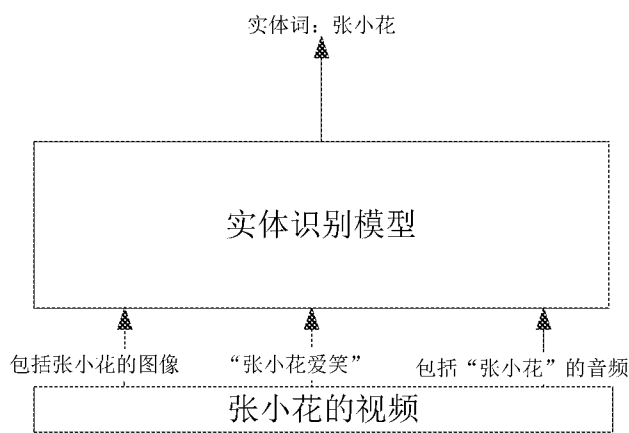


图 8

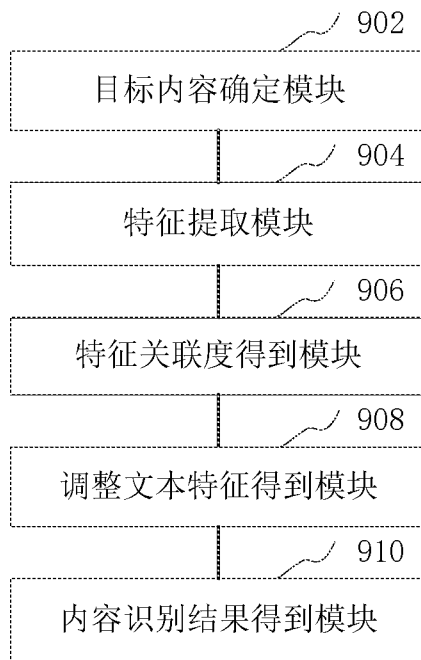


图 9

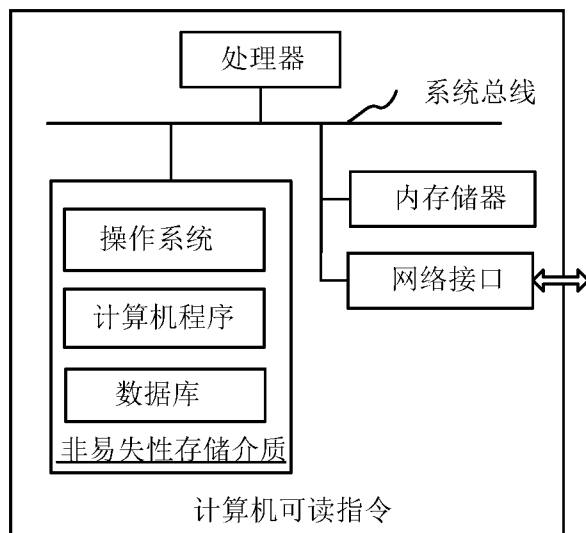


图 10

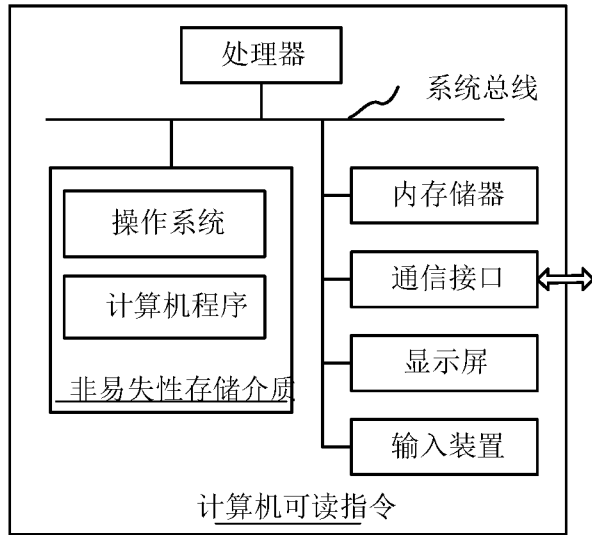


图 11

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/081896

A. CLASSIFICATION OF SUBJECT MATTER		
G06K 9/00(2022.01)i		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols)		
G06K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
CNABS; CNTXT; CNTXTC; VEN; USTXT; WOTXT; ENTXT; EPTXT; CNKI: 文本, 字幕, 标题, 视频, 音频, 帧, 图像, 识别, 特征, 提取, 关联, 相关, 语义, 时间, 调整, 更新, 修正, 融合, 权重, text, caption, title, video, audio, frame, image, recognize, feature, extract, associate, correlative, semantic, time, adjust, update, revise, fuse, weight		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 113723166 A (TENCENT TECHNOLOGY (BEIJING) CO., LTD.) 30 November 2021 (2021-11-30) claims 1-15, description, paragraphs 1-208, and figures 1-11	1-18
PX	CN 113010740 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 22 June 2021 (2021-06-22) description, paragraphs 2-43 and 71-198, and figures 1-12	1-18
X	CN 112418034 A (YUEN MENG HUMANITIES INTELLIGENT INTERNATIONAL LTD.) 26 February 2021 (2021-02-26) description, paragraphs 7-64 and 83-217, and figures 1-5	1-18
Y	CN 110991427 A (BEIJING BAIDU NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 10 April 2020 (2020-04-10) claims 1-20, description, paragraphs 32-91, and figures 1-4	1-18
Y	CN 111680541 A (HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY) 18 September 2020 (2020-09-18) description, paragraphs 77-116, and figures 1-3	1-18
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
19 May 2022		27 May 2022
Name and mailing address of the ISA/CN		Authorized officer
China National Intellectual Property Administration (ISA/CN) No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088, China		
Facsimile No. (86-10)62019451		Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/081896

C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 112364810 A (SHENZHEN HUANTAI TECHNOLOGY CO., LTD. et al.) 12 February 2021 (2021-02-12) claims 1-10, description, paragraphs 31-91, and figures 1-6	1-18
Y	CN 109885723 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 14 June 2019 (2019-06-14) claims 1-11, description, paragraphs 117-230, and figures 1-4	1-18
A	US 2015324350 A1 (INTERNATIONAL BUSINESS MACHINES CORP.) 12 November 2015 (2015-11-12) entire document	1-18

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No. PCT/CN2022/081896

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	113723166	A	30 November 2021	None	
CN	113010740	A	22 June 2021	None	
CN	112418034	A	26 February 2021	None	
CN	110991427	A	10 April 2020	None	
CN	111680541	A	18 September 2020	None	
CN	112364810	A	12 February 2021	None	
CN	109885723	A	14 June 2019	None	
US	2015324350	A1	12 November 2015	US 10642935 B2	05 May 2020

<p>A. 主题的分类</p> <p>G06K 9/00 (2022.01) i</p> <p>按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类</p>																										
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>G06K</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))</p> <p>CNABS;CNTXT;CNTXTC;VEN;USTXT;WOTXT;ENTXT;EPTXT;CNKI:文本, 字幕, 标题, 视频, 音频, 帧, 图像, 识别, 特征, 提取, 关联, 相关, 语义, 时间, 调整, 更新, 修正, 融合, 权重, text, caption, title, video, audio, frame, image, recognize, feature, extract, associate, correlative, semantic, time, adjust, update, revise, fuse, weight</p>																										
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 113723166 A (腾讯科技北京有限公司) 2021年11月30日 (2021 - 11 - 30) 权利要求1-15, 说明书第1-208段, 图1-11</td> <td>1-18</td> </tr> <tr> <td>PX</td> <td>CN 113010740 A (腾讯科技深圳有限公司) 2021年6月22日 (2021 - 06 - 22) 说明书第2-43、71-198段, 图1-12</td> <td>1-18</td> </tr> <tr> <td>X</td> <td>CN 112418034 A (元梦人文智能国际有限公司) 2021年2月26日 (2021 - 02 - 26) 说明书第7-64、83-217段, 图1-5</td> <td>1-18</td> </tr> <tr> <td>Y</td> <td>CN 110991427 A (北京百度网讯科技有限公司) 2020年4月10日 (2020 - 04 - 10) 权利要求1-20, 说明书第32-91段, 图1-4</td> <td>1-18</td> </tr> <tr> <td>Y</td> <td>CN 111680541 A (华中科技大学) 2020年9月18日 (2020 - 09 - 18) 说明书第77-116段, 图1-3</td> <td>1-18</td> </tr> <tr> <td>Y</td> <td>CN 112364810 A (深圳市欢太科技有限公司 等) 2021年2月12日 (2021 - 02 - 12) 权利要求1-10, 说明书第31-91段, 图1-6</td> <td>1-18</td> </tr> <tr> <td>Y</td> <td>CN 109885723 A (腾讯科技深圳有限公司) 2019年6月14日 (2019 - 06 - 14) 权利要求1-11, 说明书第117-230段, 图1-4</td> <td>1-18</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 113723166 A (腾讯科技北京有限公司) 2021年11月30日 (2021 - 11 - 30) 权利要求1-15, 说明书第1-208段, 图1-11	1-18	PX	CN 113010740 A (腾讯科技深圳有限公司) 2021年6月22日 (2021 - 06 - 22) 说明书第2-43、71-198段, 图1-12	1-18	X	CN 112418034 A (元梦人文智能国际有限公司) 2021年2月26日 (2021 - 02 - 26) 说明书第7-64、83-217段, 图1-5	1-18	Y	CN 110991427 A (北京百度网讯科技有限公司) 2020年4月10日 (2020 - 04 - 10) 权利要求1-20, 说明书第32-91段, 图1-4	1-18	Y	CN 111680541 A (华中科技大学) 2020年9月18日 (2020 - 09 - 18) 说明书第77-116段, 图1-3	1-18	Y	CN 112364810 A (深圳市欢太科技有限公司 等) 2021年2月12日 (2021 - 02 - 12) 权利要求1-10, 说明书第31-91段, 图1-6	1-18	Y	CN 109885723 A (腾讯科技深圳有限公司) 2019年6月14日 (2019 - 06 - 14) 权利要求1-11, 说明书第117-230段, 图1-4	1-18
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																								
PX	CN 113723166 A (腾讯科技北京有限公司) 2021年11月30日 (2021 - 11 - 30) 权利要求1-15, 说明书第1-208段, 图1-11	1-18																								
PX	CN 113010740 A (腾讯科技深圳有限公司) 2021年6月22日 (2021 - 06 - 22) 说明书第2-43、71-198段, 图1-12	1-18																								
X	CN 112418034 A (元梦人文智能国际有限公司) 2021年2月26日 (2021 - 02 - 26) 说明书第7-64、83-217段, 图1-5	1-18																								
Y	CN 110991427 A (北京百度网讯科技有限公司) 2020年4月10日 (2020 - 04 - 10) 权利要求1-20, 说明书第32-91段, 图1-4	1-18																								
Y	CN 111680541 A (华中科技大学) 2020年9月18日 (2020 - 09 - 18) 说明书第77-116段, 图1-3	1-18																								
Y	CN 112364810 A (深圳市欢太科技有限公司 等) 2021年2月12日 (2021 - 02 - 12) 权利要求1-10, 说明书第31-91段, 图1-6	1-18																								
Y	CN 109885723 A (腾讯科技深圳有限公司) 2019年6月14日 (2019 - 06 - 14) 权利要求1-11, 说明书第117-230段, 图1-4	1-18																								
<p><input checked="" type="checkbox"/> 其余文件在C栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p> <table border="0"> <tr> <td> <p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“P” 公布日先于国际申请日但迟于所要求的优先权日的文件</p> </td> <td> <p>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件</p> <p>“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性</p> <p>“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性</p> <p>“&” 同族专利的文件</p> </td> </tr> </table>			<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“P” 公布日先于国际申请日但迟于所要求的优先权日的文件</p>	<p>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件</p> <p>“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性</p> <p>“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性</p> <p>“&” 同族专利的文件</p>																						
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“P” 公布日先于国际申请日但迟于所要求的优先权日的文件</p>	<p>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件</p> <p>“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性</p> <p>“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性</p> <p>“&” 同族专利的文件</p>																									
<p>国际检索实际完成的日期</p> <p>2022年5月19日</p>		<p>国际检索报告邮寄日期</p> <p>2022年5月27日</p>																								
<p>ISA/CN的名称和邮寄地址</p> <p>中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088</p> <p>传真号 (86-10)62019451</p>		<p>授权官员</p> <p>周丹丹</p> <p>电话号码 86-(20)-28958352</p>																								

C. 相关文件		
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	US 2015324350 A1 (INTERNATIONAL BUSINESS MACHINES) 2015年11月12日 (2015 - 11 - 12) 全文	1-18

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2022/081896

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	113723166	A	2021年11月30日	无	
CN	113010740	A	2021年6月22日	无	
CN	112418034	A	2021年2月26日	无	
CN	110991427	A	2020年4月10日	无	
CN	111680541	A	2020年9月18日	无	
CN	112364810	A	2021年2月12日	无	
CN	109885723	A	2019年6月14日	无	
US	2015324350	A1	2015年11月12日	US 10642935 B2	2020年5月5日