



(12) 发明专利申请

(10) 申请公布号 CN 116956067 A

(43) 申请公布日 2023. 10. 27

(21) 申请号 202310921185.9

G06F 17/16 (2006.01)

(22) 申请日 2023.07.25

(71) 申请人 京东科技信息技术有限公司

地址 100176 北京市大兴区北京经济技术
开发区科创十一街18号院2号楼6层
601

(72) 发明人 段超群 王培英 陈蒙

(74) 专利代理机构 北京律智知识产权代理有限
公司 11438

专利代理师 孙宝海

(51) Int. Cl.

G06F 18/22 (2023.01)

H04N 21/488 (2011.01)

H04N 21/4788 (2011.01)

G06F 18/214 (2023.01)

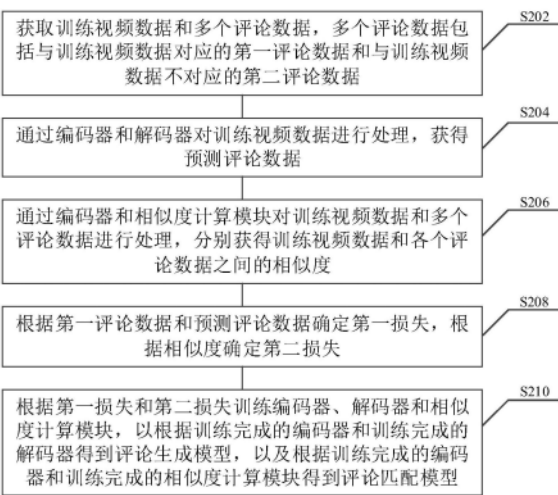
权利要求书3页 说明书15页 附图5页

(54) 发明名称

评论生成模型和评论匹配模型的联合训练
方法及相关设备

(57) 摘要

本公开提供了一种评论生成模型和评论匹配模型的联合训练方法及相关设备。该方法包括：获取训练视频数据和多个评论数据，多个评论数据包括与训练视频数据对应的第一评论数据和与训练视频数据不对应的第二评论数据；通过编码器和解码器对训练视频数据进行处理，获得预测评论数据；通过编码器和相似度计算模块对训练视频数据和多个评论数据进行处理，分别获得训练视频数据和各个评论数据之间的相似度；根据第一评论数据和预测评论数据确定第一损失，根据相似度确定第二损失；根据第一损失和第二损失训练编码器、解码器和相似度计算模块，根据训练完成的编码器和解码器得到评论生成模型，根据训练完成的编码器和相似度计算模块得到评论匹配模型。



1. 一种评论生成模型和评论匹配模型的联合训练方法,其特征在于,包括:

获取训练视频数据和多个评论数据,所述多个评论数据包括与所述训练视频数据对应的第一评论数据和与所述训练视频数据不对应的第二评论数据;

通过编码器和解码器对所述训练视频数据进行处理,获得预测评论数据;

通过所述编码器和相似度计算模块对所述训练视频数据和所述多个评论数据进行处理,分别获得所述训练视频数据和各个评论数据之间的相似度;

根据所述第一评论数据和所述预测评论数据确定第一损失,根据所述相似度确定第二损失;

根据所述第一损失和所述第二损失训练所述编码器、所述解码器和所述相似度计算模块,以根据训练完成的编码器和训练完成的解码器得到评论生成模型,以及根据所述训练完成的编码器和训练完成的相似度计算模块得到评论匹配模型。

2. 根据权利要求1所述的方法,其特征在于,所述训练视频数据包括训练视频图像和训练视频文本;

其中,通过编码器和解码器对所述训练视频数据进行处理,获得预测评论数据,包括:

将所述训练视频图像和所述训练视频文本进行拼接处理,获得第一输入序列;

分别将所述第一输入序列和所述第一评论数据进行矢量化,获得第一输入向量和第一评论向量;

将所述第一输入向量输入至所述编码器中进行处理,获得第一视频特征向量;将所述第一视频特征向量输入至所述解码器中进行处理,获得预测评论向量;

其中,根据所述第一评论数据和所述预测评论数据确定第一损失,包括:

根据所述第一评论向量和所述预测评论向量确定第一损失。

3. 根据权利要求1所述的方法,其特征在于,所述相似度包括第一相似度和第二相似度;

其中,通过所述编码器和相似度计算模块对所述训练视频数据和所述多个评论数据进行处理,分别获得所述训练视频数据和各个评论数据之间的相似度,包括:

通过所述编码器和相似度计算模块对所述训练视频数据和所述第一评论数据进行处理,获得所述训练视频数据和所述第一评论数据之间的第一相似度;

通过所述编码器和相似度计算模块对所述训练视频数据和所述第二评论数据进行处理,获得所述训练视频数据和所述第二评论数据之间的第二相似度。

4. 根据权利要求3所述的方法,其特征在于,所述训练视频数据包括训练视频图像和训练视频文本,所述相似度计算模块包括多层感知机模块;

其中,通过所述编码器和相似度计算模块对所述训练视频数据和所述第一评论数据进行处理,获得所述训练视频数据和所述第一评论数据之间的第一相似度,包括:

将所述训练视频图像、所述训练视频文本和第一评论数据进行拼接处理,获得第二输入序列;将所述第二输入序列进行矢量化,获得第二输入向量;

将所述第二输入向量输入至所述编码器中进行处理,获得第二视频特征向量;将所述第二视频特征向量输入至所述多层感知机模块中进行处理,获得所述训练视频数据和所述第一评论数据之间的第一相似度;

其中,通过所述编码器和相似度计算模块对所述训练视频数据和所述第二评论数据进

行处理,获得所述训练视频数据和所述第二评论数据之间的第二相似度,包括:

将所述训练视频图像、所述训练视频文本和第二评论数据进行拼接处理,获得第三输入序列;

将所述第三输入序列进行矢量化,获得第三输入向量;

将所述第三输入向量输入至所述编码器中进行处理,获得第三视频特征向量;将所述第三视频特征向量输入至所述多层感知机模块中进行处理,获得所述训练视频数据和所述第二评论数据之间的第二相似度。

5. 根据权利要求4所述的方法,其特征在于,根据所述相似度确定第二损失,包括:

根据所述第二相似度和所述第一相似度之间的差确定所述第二损失。

6. 根据权利要求1所述的方法,其特征在于,根据所述第一损失和所述第二损失训练所述编码器、所述解码器和所述相似度计算模块,包括:

若所述第一损失和所述第二损失的总损失大于或等于预设值,则调整所述编码器、所述解码器和所述相似度计算模块的模型参数,直至根据调整后的编码器、解码器和相似度计算模块得到的总损失小于所述预设值,获得训练完成的编码器、训练完成的解码器和训练完成的相似度计算模块。

7. 一种评论生成方法,其特征在于,包括:

获取待处理视频数据;

将所述待处理视频数据输入至评论生成模型中,获得所述待处理视频数据对应的目标评论;其中所述评论生成模型是根据权利要求1-6中任一项所述方法训练获得的。

8. 一种评论匹配方法,其特征在于,包括:

获取待处理视频数据和多个候选评论;

将所述待处理视频数据和所述多个候选评论输入至评论匹配模型中,获得所述待处理视频数据和各个候选评论之间的相似度;其中所述评论匹配模型是根据权利要求1-6中任一项所述方法训练获得的;

根据所述待处理视频数据和各个候选评论之间的相似度,从所述多个候选评论中确定目标评论。

9. 一种评论生成模型和评论匹配模型的联合训练装置,其特征在于,包括:

获取模块,用于获取训练视频数据和多个评论数据,所述多个评论数据包括与所述训练视频数据对应的第一评论数据和与所述训练视频数据不对应的第二评论数据;

处理模块,用于通过编码器和解码器对所述训练视频数据进行处理,获得预测评论数据;

所述处理模块还用于通过所述编码器和相似度计算模块对所述训练视频数据和所述多个评论数据进行处理,分别获得所述训练视频数据和各个评论数据之间的相似度;

确定模块,用于根据所述第一评论数据和所述预测评论数据确定第一损失,根据所述相似度确定第二损失;

训练模块,用于根据所述第一损失和所述第二损失训练所述编码器、所述解码器和所述相似度计算模块,以根据训练完成的编码器和训练完成的解码器得到评论生成模型,以及根据训练完成的编码器和训练完成的相似度计算模块得到评论匹配模型。

10. 一种评论生成装置,其特征在于,包括:

获取模块,用于获取待处理视频数据;

获得模块,用于将所述待处理视频数据输入至评论生成模型中,获得所述待处理视频数据对应的目标评论;其中所述评论生成模型是根据权利要求1-6中任一项所述方法训练获得的。

11.一种评论匹配装置,其特征在于,包括:

获取模块,用于获取待处理视频数据和多个候选评论;

获得模块,用于将所述待处理视频数据和所述多个候选评论输入至评论匹配模型中,获得所述待处理视频数据和各个候选评论之间的相似度;其中所述评论匹配模型是根据权利要求1-6中任一项所述方法训练获得的;

确定模块,用于根据所述待处理视频数据和各个候选评论之间的相似度,从所述多个候选评论中确定目标评论。

12.一种电子设备,其特征在于,包括:

至少一个处理器;

存储装置,用于存储至少一个程序,当所述至少一个程序被所述至少一个处理器执行时,使得所述至少一个处理器实现如权利要求1至8中任一项所述的方法。

13.一种计算机可读存储介质,其上存储有计算机可执行指令,其特征在于,所述可执行指令被处理器执行时实现如权利要求1至8中任一项所述的方法。

评论生成模型和评论匹配模型的联合训练方法及相关设备

技术领域

[0001] 本公开涉及计算机技术领域,具体而言,涉及一种评论生成模型和评论匹配模型的联合训练方法、评论生成模型和评论匹配模型的联合训练装置、评论生成方法、评论生成装置、评论匹配方法、评论匹配装置、电子设备及存储介质。

背景技术

[0002] 为了提高视频观看体验,在线视频应用提供了弹幕评论功能,该功能允许针对视频中的某个片段发表弹幕评论,并且将该弹幕评论在视频相应时段的画面中从左侧滑动到右侧,使得在观看视频的同时看到相关弹幕评论。

[0003] 相关技术中,可以使用生成模型自动生成弹幕评论,或者使用匹配模型从候选弹幕评论中匹配目标单品评论;但是,对于生成模型,仅仅使用目标弹幕评论作为监督信号来指导模型的学习,训练获得的生成模型对有效样本和无效样本的辨识能力较差;对于匹配模型,将目标弹幕评论作为正样例,随机抽取无效样例作为负样例,由于负样例是随机采样到的,因此训练获得的匹配模型对输入数据的表示不稳定,具有很大的随机性。

[0004] 需要说明的是,在上述背景技术部分公开的信息仅用于加强对本公开的背景的理解,因此可以包括不构成对本领域普通技术人员已知的现有技术的信息。

发明内容

[0005] 本公开实施例提供一种评论生成模型和评论匹配模型的联合训练装置、评论生成方法、评论生成装置、评论匹配方法、评论匹配装置、电子设备及存储介质,该方法提高了评论生成模型的样本辨识能力;增强了评论匹配模型对输入数据的表示能力,提升了评论匹配模型的稳定性。

[0006] 本公开的其他特性和优点将通过下面的详细描述变得显然,或部分地通过本公开的实践而习得。

[0007] 本公开实施例提供一种评论生成模型和评论匹配模型的联合训练方法,包括:获取训练视频数据和多个评论数据,所述多个评论数据包括与所述训练视频数据对应的第一评论数据和与所述训练视频数据不对应的第二评论数据;通过编码器和解码器对所述训练视频数据进行处理,获得预测评论数据;通过所述编码器和相似度计算模块对所述训练视频数据和所述多个评论数据进行处理,分别获得所述训练视频数据和各个评论数据之间的相似度;根据所述第一评论数据和所述预测评论数据确定第一损失,根据所述相似度确定第二损失;根据所述第一损失和所述第二损失训练所述编码器、所述解码器和所述相似度计算模块,以根据训练完成的编码器和训练完成的解码器得到评论生成模型,以及根据所述训练完成的编码器和训练完成的相似度计算模块得到评论匹配模型。

[0008] 在本公开一些示例性实施例中,所述训练视频数据包括训练视频图像和训练视频文本;其中,通过编码器和解码器对所述训练视频数据进行处理,获得预测评论数据,包括:将所述训练视频图像和所述训练视频文本进行拼接处理,获得第一输入序列;分别将所述

第一输入序列和所述第一评论数据进行矢量化,获得第一输入向量和第一评论向量;将所述第一输入向量输入至所述编码器中进行处理,获得第一视频特征向量;将所述第一视频特征向量输入至所述解码器中进行处理,获得预测评论向量;其中,根据所述第一评论数据和所述预测评论数据确定第一损失,包括:根据所述第一评论向量和所述预测评论向量确定第一损失。

[0009] 在本公开一些示例性实施例中,所述相似度包括第一相似度和第二相似度;其中,通过所述编码器和相似度计算模块对所述训练视频数据和所述多个评论数据进行处理,分别获得所述训练视频数据和各个评论数据之间的相似度,包括:通过所述编码器和相似度计算模块对所述训练视频数据和所述第一评论数据进行处理,获得所述训练视频数据和所述第一评论数据之间的第一相似度;通过所述编码器和相似度计算模块对所述训练视频数据和所述第二评论数据进行处理,获得所述训练视频数据和所述第二评论数据之间的第二相似度。

[0010] 在本公开一些示例性实施例中,所述训练视频数据包括训练视频图像和训练视频文本,所述相似度计算模块包括多层感知机模块;其中,通过所述编码器和相似度计算模块对所述训练视频数据和所述第一评论数据进行处理,获得所述训练视频数据和所述第一评论数据之间的第一相似度,包括:将所述训练视频图像、所述训练视频文本和所述第一评论数据进行拼接处理,获得第二输入序列;将所述第二输入序列进行矢量化,获得第二输入向量;将所述第二输入向量输入至所述编码器中进行处理,获得第二视频特征向量;将所述第二视频特征向量输入至所述多层感知机模块中进行处理,获得所述训练视频数据和所述第一评论数据之间的第一相似度;其中,通过所述编码器和相似度计算模块对所述训练视频数据和所述第二评论数据进行处理,获得所述训练视频数据和所述第二评论数据之间的第二相似度,包括:将所述训练视频图像、所述训练视频文本和第二评论数据进行拼接处理,获得第三输入序列;将所述第三输入序列进行矢量化,获得第三输入向量;将所述第三输入向量输入至所述编码器中进行处理,获得第三视频特征向量;将所述第三视频特征向量输入至所述多层感知机模块中进行处理,获得所述训练视频数据和所述第二评论数据之间的第二相似度。

[0011] 在本公开一些示例性实施例中,根据所述相似度确定第二损失,包括:根据所述第二相似度和所述第一相似度之间的差确定所述第二损失。

[0012] 在本公开一些示例性实施例中,根据所述第一损失和所述第二损失训练所述编码器、所述解码器和所述相似度计算模块,包括:若所述第一损失和所述第二损失的总损失大于或等于预设值,则调整所述编码器、所述解码器和所述相似度计算模块的模型参数,直至根据调整后的编码器、解码器和相似度计算模块得到的总损失小于所述预设值,获得训练完成的编码器、训练完成的解码器和训练完成的相似度计算模块。

[0013] 本公开实施例提供一种评论生成方法,包括:获取待处理视频数据;将所述待处理视频数据输入至评论生成模型中,获得所述待处理视频数据对应的目标评论;其中所述评论生成模型是根据上述任一种方法训练获得的。

[0014] 本公开实施例提供一种评论匹配方法,其特征在于,包括:获取待处理视频数据和多个候选评论;将所述待处理视频数据和所述多个候选评论输入至评论匹配模型中,获得所述待处理视频数据和各个候选评论之间的相似度;其中所述评论匹配模型是根据上述任

一种方法训练获得的；根据所述待处理视频数据和各个候选评论之间的相似度，从所述多个候选评论中确定目标评论。

[0015] 本公开实施例提供一种评论生成模型和评论匹配模型的联合训练装置，包括：获取模块，用于获取训练视频数据和多个评论数据，所述多个评论数据包括与所述训练视频数据对应的第一评论数据和与所述训练视频数据不对应的第二评论数据；处理模块，用于通过编码器和解码器对所述训练视频数据进行处理，获得预测评论数据；所述处理模块还用于通过所述编码器和相似度计算模块对所述训练视频数据和所述多个评论数据进行处理，分别获得所述训练视频数据和各个评论数据之间的相似度；确定模块，用于根据所述第一评论数据和所述预测评论数据确定第一损失，根据所述相似度确定第二损失；训练模块，用于根据所述第一损失和所述第二损失训练所述编码器、所述解码器和所述相似度计算模块，以根据训练完成的编码器和训练完成的解码器得到评论生成模型，以及根据训练完成的编码器和训练完成的相似度计算模块得到评论匹配模型。

[0016] 本公开实施例提供一种评论生成装置，包括：获取模块，用于获取待处理视频数据；获得模块，用于将所述待处理视频数据输入至评论生成模型中，获得所述待处理视频数据对应的目标评论；其中所述评论生成模型是根据上述任一种方法训练获得的。

[0017] 本公开实施例提供一种评论匹配装置，包括：获取模块，用于获取待处理视频数据和多个候选评论；获得模块，用于将所述待处理视频数据和所述多个候选评论输入至评论匹配模型中，获得所述待处理视频数据和各个候选评论之间的相似度；其中所述评论匹配模型是根据上述任一种方法训练获得的；确定模块，用于根据所述待处理视频数据和各个候选评论之间的相似度，从所述多个候选评论中确定目标评论。

[0018] 本公开实施例提供一种电子设备，包括：至少一个处理器；存储装置，用于存储至少一个程序，当所述至少一个程序被所述至少一个处理器执行时，使得所述至少一个处理器实现如上述任一种评论生成模型和评论匹配模型的联合训练方法或者评论生成方法或者评论匹配方法。

[0019] 本公开实施例提供一种计算机可读存储介质，其上存储有计算机程序，其特征在于，所述计算机程序被处理器执行时实现如上述任一种评论生成模型和评论匹配模型的联合训练方法或者评论生成方法或者评论匹配方法。

[0020] 本公开实施例提供的评论生成模型和评论匹配模型的联合训练方法，评论生成模型和评论匹配模型共享同一个编码器，将评论生成模型和评论匹配模型进行联合训练；针对评论生成模型，由于训练评论匹配模型引入了负样例（即与训练视频数据不对应的第二评论数据），因此评论生成模型在训练过程中也能够学习到负样例的信息，从而有助于生成模型区分目标弹幕评论和负样例，即提高了评论生成模型的样本辨识能力；针对评论匹配模型，由于与评论生成模型联合训练，评论生成模型为评论匹配模型提供了正确监督信号（即与训练数据对应的第一评论数据），从而增强了评论匹配模型对输入数据的表示能力，提升了评论匹配模型的稳定性。

[0021] 应当理解的是，以上的一般描述和后文的细节描述仅是示例性和解释性的，并不能限制本公开。

附图说明

[0022] 此处的附图被并入说明书中并构成本说明书的一部分,示出了符合本公开的实施例,并与说明书一起用于解释本公开的原理。显而易见地,下面描述中的附图仅仅是本公开的一些实施例,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。

[0023] 图1示出了可以应用本公开实施例的评论生成模型和评论匹配模型的联合训练方法或者评论生成方法或者评论匹配方法的示例性系统架构的示意图。

[0024] 图2是本公开一示例性实施例中的评论生成模型和评论匹配模型的联合训练方法的流程图。

[0025] 图3是本公开一示例性实施例中的评论生成模型和评论匹配模型进行联合训练的示意图。

[0026] 图4是本公开一示例性实施例中的评论生成方法的流程图。

[0027] 图5是本公开一示例性实施例中的评论匹配方法的流程图。

[0028] 图6是本公开一示例性实施例中的评论生成模型和评论匹配模型的联合训练装置的框图。

[0029] 图7是本公开一示例性实施例中的评论生成装置的框图。

[0030] 图8是本公开一示例性实施例中的评论匹配装置的框图。

[0031] 图9是根据一示例性实施方式示出的一种电子设备的结构示意图。

具体实施方式

[0032] 现在将参考附图更全面地描述示例实施例。然而,示例实施例能够以多种形式实施,且不应被理解为限于在此阐述的实施例;相反,提供这些实施例使得本公开将全面和完整,并将示例实施例的构思全面地传达给本领域的技术人员。在图中相同的附图标记表示相同或类似的部分,因而将省略对它们的重复描述。

[0033] 本公开所描述的特征、结构或特性可以以任何合适的方式结合在一个或更多实施方式中。在下面的描述中,提供许多具体细节从而给出对本公开的实施方式的充分理解。然而,本领域技术人员将意识到,可以实践本公开的技术方案而省略特定细节中的一个或更多,或者可以采用其它的方法、组元、装置、步骤等。在其它情况下,不详细示出或描述公知方法、装置、实现或者操作以避免模糊本公开的各方面。

[0034] 附图仅为本公开的示意性图解,图中相同的附图标记表示相同或类似的部分,因而将省略对它们的重复描述。附图中所示的一些方框图不一定必须与物理或逻辑上独立的实体相对应。可以采用软件形式来实现这些功能实体,或在至少一个硬件模块或集成电路中实现这些功能实体,或在不同网络和/或处理器装置和/或微控制器装置中实现这些功能实体。

[0035] 附图中所示的流程图仅是示例性说明,不是必须包括所有的内容和步骤,也不是必须按所描述的顺序执行。例如,有的步骤还可以分解,而有的步骤可以合并或部分合并,因此实际执行的顺序有可能根据实际情况改变。

[0036] 此外,在本公开的描述中,用语“一个”、“一”、“该”、“所述”和“至少一个”用以表示存在至少一个要素或组成部分;用语“包含”、“包括”和“具有”用以表示开放式的包括在内

的意思并且是指除了列出的要素或组成部分之外还可存在另外的要素或组成部分；用语“第一”、“第二”和“第三”等仅作为标记使用，不是对其对象的数量限制。

[0037] 图1示出了可以应用本公开实施例的评论生成模型和评论匹配模型的联合训练方法或者评论生成方法或者评论匹配方法的示例性系统架构的示意图。

[0038] 如图1所示，该系统架构可以包括服务器101、网络102、终端设备103、终端设备104和终端设备105。网络102用以在终端设备103、终端设备104或终端设备105和服务器101之间提供通信链路的介质。网络102可以包括各种连接类型，例如有线、无线通信链路或者光纤电缆等等。

[0039] 服务器101可以是提供各种服务的服务器，例如对用户利用终端设备103、终端设备104或终端设备105所进行操作的装置提供支持的后台管理服务器。后台管理服务器可以对接收到的请求等数据进行分析等处理，并将处理结果反馈给终端设备103、终端设备104或终端设备105。

[0040] 终端设备103、终端设备104和终端设备105可以是智能手机、平板电脑、笔记本电脑、台式计算机、智能音箱、可穿戴智能设备、虚拟现实设备、增强现实设备等，但并不局限于此。

[0041] 本公开实施例中，服务器101可以：获取训练视频数据和多个评论数据，多个评论数据包括与训练视频数据对应的第一评论数据和与训练视频数据不对应的第二评论数据；通过编码器和解码器对训练视频数据进行处理，获得预测评论数据；通过编码器和相似度计算模块对训练视频数据和多个评论数据进行处理，分别获得训练视频数据和各个评论数据之间的相似度；根据第一评论数据和预测评论数据确定第一损失，根据相似度确定第二损失；根据第一损失和第二损失训练编码器、解码器和相似度计算模块，以根据训练完成的编码器和训练完成的解码器得到评论生成模型，以及根据训练完成的编码器和训练完成的相似度计算模块得到评论匹配模型。

[0042] 本公开实施例中，服务器101可以：获取待处理视频数据；将待处理视频数据输入至评论生成模型中，获得待处理视频数据对应的目标评论；其中评论生成模型是根据上述方法训练获得的。

[0043] 本公开实施例中，服务器101可以：获取待处理视频数据和多个候选评论；将待处理视频数据和多个候选评论输入至评论匹配模型中，获得待处理视频数据和各个候选评论之间的相似度；其中评论匹配模型是根据上述方法训练获得的；根据待处理视频数据和各个候选评论之间的相似度，从多个候选评论中确定目标评论。

[0044] 应该理解，图1中的终端设备103、终端设备104、终端设备105、网络102和服务器101的数目仅仅是示意性的，服务器101可以是一个实体的服务器，还可以为多个服务器组成的服务器集群，还可以是云端服务器，根据实际需要，可以具有任意数目的终端设备、网络和服务。

[0045] 下面，将结合附图及实施例对本公开示例实施例中的评论生成模型和评论匹配模型的联合训练方法的各个步骤进行更详细的说明。

[0046] 图2是本公开一示例性实施例中的评论生成模型和评论匹配模型的联合训练方法的流程图。本公开实施例提供的方法可以由任意具备计算处理能力的电子设备执行，例如由图1所示服务器或终端设备执行，但本公开并不限于此。

[0047] 如图2所示,本公开实施例提供的评论生成模型和评论匹配模型的联合训练方法可以包括以下步骤。

[0048] 在步骤S202中,获取训练视频数据和多个评论数据,多个评论数据包括与训练视频数据对应的第一评论数据和与训练视频数据不对应的第二评论数据。

[0049] 本公开实施例中,训练视频数据可以是影视剧视频数据、综艺视频数据、短视频数据等;训练视频数据可以包括训练视频图像和训练视频文本,训练视频图像指的是影视剧视频、综艺视频、短视频的全部或片段,训练视频文本指的是从训练视频图像中提取到的文本信息,例如字幕等。

[0050] 本公开实施例中,评论数据可以是用户针对视频发表的展示在评论区的评论数据,也可以是用户针对视频中某个片段发表的展示在视频播放区域内的弹幕评论数据,评论数据可以是文本形式、也可以是图像形式,本公开对此不做限定。在下面的举例说明中,以评论数据是弹幕评论数据为例进行说明。

[0051] 本公开实施例中,第一评论数据指的是和训练视频数据对应的评论数据,例如第一评论数据是针对训练视频数据中的视频片段发出的评论数据;第二评论数据指的是和训练视频数据不对应的评论数据,例如第二评论数据和训练视频数据中的视频内容无关;第一评论数据和第二评论数据均可以有一个或多个,训练视频数据和第一评论数据可以看作是正例,训练视频数据和第二评论数据可以看作是负例。

[0052] 本公开实施例中,训练视频数据和每个第一评论数据组成一组正例,训练视频数据和每个第二评论数据组成一组负例,使用正例训练下述包括编码器和解码器的评论生成模型,使用正例和负例训练下述包括编码器和相似度计算模块的评论匹配模型。

[0053] 具体地,给定一批训练样本 $Batch = \{(F^{(j)}, C^{(j)}, y^{(j)+}, y^{(j)-})\}_{j=1}^{bs}$, 分别使用正例 $\{(F^{(j)}, C^{(j)}, y^{(j)+})\}_{j=1}^{bs}$ 和负例 $\{(F^{(j)}, C^{(j)}, y^{(j)+}, y^{(j)-})\}_{j=1}^{bs}$ 来训练下述编码器、解码器和相似度计算模块;其中 $F^{(j)}$ 表示第j个训练视频图像, $C^{(j)}$ 表示第j个训练视频文本, $y^{(j)+}$ 表示第j个与 $F^{(j)}$ 对应的评论数据, $y^{(j)-}$ 表示第j个与 $F^{(j)}$ 不对应的评论数据, bs表示该批训练样本的数量。

[0054] 在步骤S204中,通过编码器和解码器对训练视频数据进行处理,获得预测评论数据。

[0055] 本公开实施例中,参考图3,待训练的评论生成模型10可以是Transformer(转换器)模型,其可以包括编码器301和解码器302;在评论生成模型10的训练过程中,可以将训练视频数据输入至编码器301和解码器302中进行处理,输出预测评论数据。

[0056] 其中,编码器301可以是双向编码器,解码器302可以是自回归的且可以逐字生成序列。

[0057] 在示例性实施例中,通过编码器和解码器对训练视频数据进行处理,获得预测评论数据,包括:将训练视频图像和训练视频文本进行拼接处理,获得第一输入序列;分别将第一输入序列和第一评论数据进行矢量化,获得第一输入向量和第一评论向量;将第一输入向量输入至编码器中进行处理,获得第一视频特征向量;将第一视频特征向量输入至解码器中进行处理,获得预测评论向量。

[0058] 具体地,在输入至编码器之前,可以先通过公式(1)将训练视频数据中的训练视频

图像和训练视频文本进行拼接处理,得到第一输入序列:

$$[0059] \quad X = \langle BOS \rangle \circ F \circ \langle SEP \rangle \circ C \circ \langle EOS \rangle \quad (1)$$

[0060] 公式(1)中,X表示第一输入序列, $\circ$ 表示拼接, $\langle BOS \rangle$ 表示起始标记, $\langle SEP \rangle$ 表示分隔标记, $\langle EOS \rangle$ 表示结束标记;F表示训练视频图像, $F = \{f_1, f_2, \dots, f_{N_f}\}$,其中 f_i 表示训练视频图像中的第i个图像帧, N_f 表示训练视频图像中图像帧的数量;C表示训练视频文本, $C = \{c_1, c_2, \dots, c_{N_c}\}$,其中 c_i 表示训练视频文本中的第i个文本, N_c 表示训练视频文本中文本的数量。

[0061] 具体地,可以将第一输入序列进行矢量化,获得第一输入向量;将第一评论数据进行矢量化,获得第一评论向量;由于第一输入序列中的训练视频图像和训练视频文本属于两种不同模态的输入,可以使用不同的矢量化方法对其进行处理。同时,由于训练视频文本和第一评论数据的模态相同,第一评论数据的矢量化方法与训练视频文本的矢量化方法类似。

[0062] 对于训练视频图像 $F = \{f_1, f_2, \dots, f_{N_f}\}$,可以使用一个预训练好的ResNet(残差网络)模型将其中的每一帧图片转换成一个向量,即 $\{e_{f_1}, e_{f_2}, \dots, e_{f_{N_f}}\}$,例如可以使用一个预训练好的18层ResNet来编码视频图像,可以表示为:

$$[0063] \quad e_{f_i} = \text{ResNet}(f_i) \quad (2)$$

[0064] 对于训练视频文本 $C = \{c_1, c_2, \dots, c_{N_c}\}$,首先将这 N_c 条文本拼接成一个序列 $C = \{w_1, w_2, \dots, w_{l_c}\}$,其中 w_i 表示该序列中的第i个词, l_c 表示该序列的长度,然后,我们通过查找词向量矩阵M将该序列中每个词转换成向量表示,从而得到 $\{e_{w_1}, e_{w_2}, \dots, e_{w_{l_c}}\}$,其中 $M \in \mathbb{R}^{d \times |V|}$,d表示词向量的维度, $|V|$ 表示词表大小。

[0065] 由此,将第一输入序列X矢量化后获得第一输入向量x;类似地,使用相同的方法对第一评论数据(例如目标弹幕评论y)进行矢量化,获得第一评论向量。

[0066] 然后,将第一输入向量x输入至编码器(Encoder)中进行处理,编码器学习第一输入向量x各部分的上下文表示,获得第一视频特征向量:

$$[0067] \quad H_x = \text{Encoder}(x) \quad (3)$$

[0068] 在公式(3)和公式(4)中, H_x 表示第一视频特征向量。

[0069] 将第一视频特征向量输入至解码器(Decoder)中进行处理,解码器学习上下文表示,获得预测评论向量:

$$[0070] \quad H_y = \text{Decoder}(H_x, y) \quad (4)$$

[0071] 在公式(4)中, H_y 表示预测评论向量。

[0072] 进而可以通过下式计算第一评论数据(例如目标弹幕评论y)的生成概率 $P(y|F, C)$:

$$[0073] \quad p(y_i | y_{<i}, F, C) = \text{softmax} \left(\text{MLP}(H_{y_i}) \right) \quad (5)$$

$$[0074] \quad P(y|F, C) = \prod_i p(y_i | y_{<i}, F, C) \quad (6)$$

[0075] 其中, H_{y_i} 表示 H_y 的第 i 行元素, MLP 表示一个多层感知机模块, softmax 表示归一化指数函数, y_i 目标弹幕评论 y 中第 i 个词的上下文表示, $P(y|F, C)$ 表示生成整个目标弹幕评论 y 的概率, $p(y_i|y_{<i}, F, C)$ 表示弹幕评论 y 中第 i 个词的生成概率。

[0076] 由于目标弹幕评论 y 表示的是一个句子, 假设由 $k+1$ 个词组成, 即 $y = \{y_0, y_1, \dots, y_k\}$ 。在生成 y 的时候, 其实是从左到右依次生成 y 中的每一个词, 并且在生成第 i 个词的时候, 还会将 $y_{<i} = \{y_0, y_1, \dots, y_{i-1}\}$ 也作为输入信息, 因此使用 $p(y_i|y_{<i}, F, C)$ 表示生成弹幕评论 y 中第 i 个词时的概率。

[0077] 本公开实施例中, 通过预测评论向量计算获得第一评论数据的生成概率后, 可以根据该生成概率计算下述第一损失。

[0078] 在步骤 S206 中, 通过编码器和相似度计算模块对训练视频数据和多个评论数据进行处理, 分别获得训练视频数据和各个评论数据之间的相似度。

[0079] 本公开实施例中, 相似度计算模块可以包括多层感知机模块, 即可以通过多层感知机模块计算相似度。

[0080] 本公开实施例中, 参考图 3, 待训练的评论匹配模型 20 可以包括编码器 301 和相似度计算模块 303; 在评论匹配模型 20 的训练过程中, 可以将训练视频数据和评论数据输入至编码器 301 中进行编码, 然后输入至相似度计算模块 303, 输出训练视频数据和各个评论数据之间的相似度。

[0081] 本公开实施例中, 评论生成模型 10 和评论匹配模型 20 使用同一个编码器 301; 在训练评论生成模型 10 和评论匹配模型 20 时, 对于输入数据的先后顺序不做限定, 即对于执行步骤 S204 和步骤 S206 的先后顺序不做限定。

[0082] 在示例性实施例中, 通过编码器和相似度计算模块对训练视频数据和多个评论数据进行处理, 分别获得训练视频数据和各个评论数据之间的相似度, 包括: 通过编码器和相似度计算模块对训练视频数据和第一评论数据进行处理, 获得训练视频数据和第一评论数据之间的第一相似度; 通过编码器和相似度计算模块对训练视频数据和第二评论数据进行处理, 获得训练视频数据和第二评论数据之间的第二相似度。

[0083] 本公开实施例中, 第一评论数据和第二评论数据有多个时, 在计算第一相似度和第二相似度时, 是计算每个第一评论数据和训练视频数据的相似度、以及每个第二评论数据和训练视频的相似度。

[0084] 在示例性实施例中, 通过编码器和相似度计算模块对训练视频数据和第一评论数据进行处理, 获得训练视频数据和第一评论数据之间的第一相似度, 包括: 将训练视频图像、训练视频文本和第一评论数据进行拼接处理, 获得第二输入序列; 将第二输入序列进行矢量化, 获得第二输入向量; 将第二输入向量输入至编码器中进行处理, 获得第二视频特征向量; 将第二视频特征向量输入至多层感知机模块中进行处理, 获得训练视频数据和第一评论数据之间的第一相似度。

[0085] 在示例性实施例中, 通过编码器和相似度计算模块对训练视频数据和第二评论数据进行处理, 获得训练视频数据和第二评论数据之间的第二相似度, 包括: 将训练视频图像、训练视频文本和第二评论数据进行拼接处理, 获得第三输入序列; 将第三输入序列进行矢量化, 获得第三输入向量; 将第三输入向量输入至编码器中进行处理, 获得第三视频特征向量; 将第三视频特征向量输入至多层感知机模块中进行处理, 获得训练视频数据和第二

评论数据之间的第二相似度。

[0086] 对于评论匹配模型,可以将训练视频图像F、训练视频文本C和候选评论数据y(包括第一评论数据和第二评论数据)进行拼接,得到第二输入序列或第三输入序列:

$$[0087] \quad X = \langle BOS \rangle \circ F \circ \langle SEP \rangle \circ C \circ \langle SEP \rangle \circ y \circ \langle EOS \rangle \quad (7)$$

[0088] 在公式(7)中,在输入第一评论数据时,X表示第二输入序列;在输入第二评论数据时,X表示第三输入序列。

[0089] 然后使用上文的方法将第二输入序列和第三输入序列的各部分矢量化,得到第二输入向量或第三输入向量(使用x表示);将第二输入向量或第三输入向量输入到编码器中学习其上下文表示,获得第二视频特征向量或第三视频特征向量:

$$[0090] \quad H_x = \text{Encoder}(x) \quad (8)$$

[0091] 在公式(8)中, H_x 表示第二视频特征向量或第三视频特征向量。

[0092] 本公开实施例中,在第二视频特征向量或第三视频特征向量之后,可以利用池化操作(Pooling)将其转换成一个固定长度的向量作为最终的表示;例如使用平均池化操作(Average Pooling)和最大池化操作(Max Pooling)分别对 H_x 进行操作,并将他们的输出进行合并:

$$[0093] \quad V_{mean} = \text{mean}(H_{x_i}), V_{max} = \text{max}(H_{x_i}) \quad (9)$$

$$[0094] \quad V = [V_{mean} : V_{max}] \quad (10)$$

[0095] 其中, $[:]$ 表示向量拼接操作。

[0096] 在得到输出表示V之后,训练视频数据和评论数据(例如候选弹幕评论)之间的相似度可以按照如下方式计算:

$$[0097] \quad S(F, C, y) = \sigma(\text{MLP}_2(\text{ReLU}(\text{MLP}_1(V)))) \quad (11)$$

[0098] 其中, $S(F, C, y)$ 表示训练视频数据对应的训练视频图像F、训练视频文本C和评论数据y的相似度, σ 表示sigmoid函数, MLP_1 和 MLP_2 表示2个多层感知机模块,ReLU表示激活函数。

[0099] 在步骤S208中,根据第一评论数据和预测评论数据确定第一损失,根据相似度确定第二损失。

[0100] 本公开实施例中,可以根据第一评论数据和预测评论数据之间的差异确定第一损失,根据训练视频数据和各个评论数据之间的相似度确定第二损失。

[0101] 在示例性实施例中,可以先将第一评论数据进行矢量化,获得第一评论向量;再根据第一评论向量和预测评论向量确定第一损失。

[0102] 例如,使用交叉熵(Cross Entropy)作为评论生成模型的损失函数 $L_{gen}(\theta)$:

$$[0103] \quad L_{gen}(\theta) = -\sum_{j=1}^N \log P(y^{(j)} | F^{(j)}, C^{(j)}) \quad (12)$$

[0104] 其中,N表示训练集中样本个数, $y^{(j)}$ 、 $F^{(j)}$ 、 $C^{(j)}$ 表示第j个样本中的第一评论数据、训练视频图像和训练视频文本。

[0105] 在示例性实施例中,可以分别计算训练视频数据和第一评论数据(正例)之间的第一相似度、以及训练视频数据和第二评论数据(负例)之间的第二相似度;根据第一相似度和第二相似度确定第二损失。

[0106] 在示例性实施例中,根据第二相似度和第一相似度之间的差确定第二损失。

[0107] 例如,使用最大间隔损失函数(Max-margin Loss Function)作为评论匹配模型的损失函数 $L_{rank}(\theta)$:

$$[0108] \quad L_{rank}(\theta) = \frac{1}{N} \sum_{j=1}^N \max(0, M + S(F^{(j)}, C^{(j)}, y^{(j)-}) - S(F^{(j)}, C^{(j)}, y^{(j)+})) \quad (13)$$

[0109] 其中N表示训练集中样本个数, $(F^{(j)}, C^{(j)}, y^{(j)+})$ 表示第j个样本的正例, $(F^{(j)}, C^{(j)}, y^{(j)-})$ 表示第j个样本的负例, M表示正负例之间的间隔(Margin), 其值可以通过人工来指定。

[0110] 本公开实施例中,由于第一评论数据和训练视频数据是对应的,第二评论数据和训练视频数据是不对应的,所以根据第二相似度和第一相似度之间的差确定第二损失,在训练评论匹配模型时,可以使得评论匹配模型对正负例的识别效果更好。

[0111] 在步骤S210中,根据第一损失和第二损失训练编码器、解码器和相似度计算模块,以根据训练完成的编码器和训练完成的解码器得到评论生成模型,以及根据训练完成的编码器和训练完成的相似度计算模块得到评论匹配模型。

[0112] 本公开实施例中,可以使用第一损失调节编码器和解码器的参数,使用第二损失调节编码器和相似度计算模块的模型参数;也可以根据第一损失和第二损失得到总损失,根据总损失调节编码器、解码器和相似度计算模块的模型参数。

[0113] 在示例性实施例中,根据第一损失和第二损失训练编码器、解码器和相似度计算模块,包括:若第一损失和第二损失的总损失大于或等于预设值,则调整编码器、解码器和相似度计算模块的模型参数,直至根据调整后的编码器、解码器和相似度计算模块得到的总损失小于预设值,获得训练完成的编码器、训练完成的解码器和训练完成的相似度计算模块。

[0114] 具体地,若第一损失和第二损失的总损失大于或等于预设值,则调整编码器、解码器和相似度计算模块的模型参数,通过调整后的编码器和调整后的解码器对训练视频数据进行处理,获得预测评论数据;通过调整后的编码器和调整后的相似度计算模块对所述训练视频数据和所述多个评论数据进行处理,分别获得所述训练视频数据和各个评论数据之间的相似度;根据第一评论数据和所述预测评论数据确定第一损失,根据相似度确定第二损失;再次判断第一损失和第二损失的总损失和预设值的关系,若第一损失和第二损失的总损失仍大于或等于预设值,则继续调整编码器、解码器和相似度计算模块的模型参数;若第一损失和第二损失的总损失小于预设值,则完成训练,得到训练完成的编码器、训练完成的解码器和训练完成的相似度计算模块。

[0115] 本公开实施例提供的评论生成模型和评论匹配模型的联合训练方法,评论生成模型和评论匹配模型共享同一个编码器,将评论生成模型和评论匹配模型进行联合训练;针对评论生成模型,由于训练评论匹配模型引入了负样例(即与训练视频数据不对应的第二评论数据),因此评论生成模型在训练过程中也能够学习到负样例的信息,从而有助于生成模型区分目标弹幕评论和负样例,即提高了评论生成模型的样本辨识能力,从而提高了评论生成模型生成评论的准确性;针对评论匹配模型,由于与评论生成模型联合训练,评论生成模型为评论匹配模型提供了正确监督信号(即与训练数据对应的第一评论数据),从而增强了评论匹配模型对输入数据的表示能力,提升了评论匹配模型的稳定性。

[0116] 图4是本公开一示例性实施例中的评论生成方法的流程图。本公开实施例提供的方法可以由任意具备计算处理能力的电子设备执行,例如由图1所示服务器或终端设备执行,但本公开并不限于此。

[0117] 如图4所示,本公开实施例提供的评论生成方法可以包括以下步骤。

[0118] 在步骤S402中,获取待处理视频数据。

[0119] 本公开实施例中,待处理视频数据可以是影视剧视频数据、综艺视频数据、短视频数据等。

[0120] 在步骤S404中,将待处理视频数据输入至评论生成模型中,获得待处理视频数据对应的目标评论;其中评论生成模型是根据上述任一实施例提供的方法训练获得的。

[0121] 本公开实施例中,目标评论例如可以是目标弹幕评论;将待处理视频数据输入至评论生成模型中,评论生成模型可以自动生成与待处理视频数据对应的目标评论。

[0122] 本公开实施例提供的评论生成方法,将待处理视频数据输入至根据上述任一实施例得到的评论生成模型中,可以获得更准确的目标评论。

[0123] 图5是本公开一示例性实施例中的评论匹配方法的流程图。本公开实施例提供的方法可以由任意具备计算处理能力的电子设备执行,例如由图1所示服务器或终端设备执行,但本公开并不限于此。

[0124] 如图5所示,本公开实施例提供的评论匹配方法可以包括以下步骤。

[0125] 在步骤S502中,获取待处理视频数据和多个候选评论。

[0126] 本公开实施例中,待处理视频数据可以是影视剧视频数据、综艺视频数据、短视频数据等;候选评论中包括与待处理视频数据相关性较大的评论和与待处理视频数据相关性较小的评论;候选评论可以是候选弹幕评论。

[0127] 在步骤S504中,将待处理视频数据和多个候选评论输入至评论匹配模型中,获得待处理视频数据和各个候选评论之间的相似度;其中评论匹配模型是根据上述任一实施例提供的方法训练获得的。

[0128] 在步骤S506中,根据待处理视频数据和各个候选评论之间的相似度,从多个候选评论中确定目标评论。

[0129] 本公开实施例中,可以将相似度最大的(或相似度较大的)候选评论作为目标评论。

[0130] 本公开实施例提供的评论匹配方法,将待处理视频数据和多个候选评论输入至根据上述任一实施例得到的评论匹配模型中,获得的待处理视频数据和各个候选评论之间的相似度更加准确。

[0131] 还应理解,上述只是为了帮助本领域技术人员更好地理解本公开实施例,而非要限制本公开实施例的范围。本领域技术人员根据所给出的上述示例,显然可以进行各种等价的修改或变化,例如,上述方法中某些步骤可以是不必须的,或者可以新加入某些步骤等。或者上述任意两种或者任意多种实施例的组合。这样的修改、变化或者组合后的方案也落入本公开实施例的范围内。

[0132] 还应理解,上文对本公开实施例的描述着重于强调各个实施例之间的不同之处,未提到的相同或相似之处可以互相参考,为了简洁,这里不再赘述。

[0133] 还应理解,上述各过程的序号的大小并不意味着执行顺序的先后,各过程的执行

顺序应以其功能和内在逻辑确定,而不对本公开实施例的实施过程构成任何限定。

[0134] 还应理解,在本公开的各个实施例中,如果没有特殊说明以及逻辑冲突,不同的实施例之间的术语和/或描述具有一致性、且可以相互引用,不同的实施例中的技术特征根据其内在的逻辑关系可以组合形成新的实施例。

[0135] 上文详细介绍了本公开提供的评论生成模型和评论匹配模型的联合训练方法示例。可以理解的是,计算机设备为了实现上述功能,其包含了执行各个功能相应的硬件结构和/或软件模块。本领域技术人员应该很容易意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,本公开能够以硬件或硬件和计算机软件的结合形式来实现。某个功能究竟以硬件还是计算机软件驱动硬件的方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本公开的范围。

[0136] 下述为本公开装置实施例,可以用于执行本公开方法实施例。对于本公开装置实施例中未披露的细节,请参照本公开方法实施例。

[0137] 图6是本公开一示例性实施例中的评论生成模型和评论匹配模型的联合训练装置的框图,所述多模态数据表征模型包括图像编码器、文本编码器和所述多模态编码器。

[0138] 如图6所示,评论生成模型和评论匹配模型的联合训练装置600可以包括:获取模块602、处理模块604、确定模块606和训练模块608。

[0139] 其中,获取模块602用于获取训练视频数据和多个评论数据,所述多个评论数据包括与所述训练视频数据对应的第一评论数据和与所述训练视频数据不对应的第二评论数据;处理模块604用于通过编码器和解码器对所述训练视频数据进行处理,获得预测评论数据;所述处理模块604还用于通过所述编码器和相似度计算模块对所述训练视频数据和所述多个评论数据进行处理,分别获得所述训练视频数据和各个评论数据之间的相似度;确定模块606用于根据所述第一评论数据和所述预测评论数据确定第一损失,根据所述相似度确定第二损失;训练模块608用于根据所述第一损失和所述第二损失训练所述编码器、所述解码器和所述相似度计算模块,以根据训练完成的编码器和训练完成的解码器得到评论生成模型,以及根据训练完成的编码器和训练完成的相似度计算模块得到评论匹配模型。

[0140] 在本公开一些示例性实施例中,所述训练视频数据包括训练视频图像和训练视频文本;其中,处理模块604用于:将所述训练视频图像和所述训练视频文本进行拼接处理,获得第一输入序列;分别将所述第一输入序列和所述第一评论数据进行矢量化,获得第一输入向量和第一评论向量;将所述第一输入向量输入至所述编码器中进行处理,获得第一视频特征向量;将所述第一视频特征向量输入至所述解码器中进行处理,获得预测评论向量;其中,根据所述第一评论数据和所述预测评论数据确定第一损失,包括:根据所述第一评论向量和所述预测评论向量确定第一损失。

[0141] 在本公开一些示例性实施例中,所述相似度包括第一相似度和第二相似度;其中,处理模块604用于:通过所述编码器和相似度计算模块对所述训练视频数据和所述第一评论数据进行处理,获得所述训练视频数据和所述第一评论数据之间的第一相似度;通过所述编码器和相似度计算模块对所述训练视频数据和所述第二评论数据进行处理,获得所述训练视频数据和所述第二评论数据之间的第二相似度。

[0142] 在本公开一些示例性实施例中,所述训练视频数据包括训练视频图像和训练视频

文本,所述相似度计算模块包括多层感知机模块;其中,处理模块604用于:将所述训练视频图像、所述训练视频文本和第一评论数据进行拼接处理,获得第二输入序列;将所述第二输入序列进行矢量化,获得第二输入向量;将所述第二输入向量输入至所述编码器中进行处理,获得第二视频特征向量;将所述第二视频特征向量输入至所述多层感知机模块中进行处理,获得所述训练视频数据和所述第一评论数据之间的第一相似度;其中,处理模块604用于:将所述训练视频图像、所述训练视频文本和第二评论数据进行拼接处理,获得第三输入序列;将所述第三输入序列进行矢量化,获得第三输入向量;将所述第三输入向量输入至所述编码器中进行处理,获得第三视频特征向量;将所述第三视频特征向量输入至所述多层感知机模块中进行处理,获得所述训练视频数据和所述第二评论数据之间的第二相似度。

[0143] 在本公开一些示例性实施例中,确定模块606用于:根据所述第二相似度和所述第一相似度之间的差确定所述第二损失。

[0144] 在本公开一些示例性实施例中,训练模块608用于:若所述第一损失和所述第二损失的总损失大于或等于预设值,则调整所述编码器、所述解码器和所述相似度计算模块的模型参数,直至根据调整后的编码器、解码器和相似度计算模块得到的总损失小于所述预设值,获得训练完成的编码器、训练完成的解码器和训练完成的相似度计算模块。

[0145] 图7是本公开一示例性实施例中的评论生成装置的框图。

[0146] 如图7所示,评论生成装置700可以包括:获取模块702和获得模块704。

[0147] 其中,获取模块702用于获取待处理视频数据;获得模块704用于将所述待处理视频数据输入至评论生成模型中,获得所述待处理视频数据对应的目标评论;其中所述评论生成模型是根据上述任一种方法训练获得的。

[0148] 图8是本公开一示例性实施例中的评论匹配装置的框图。

[0149] 如图8所示,评论匹配装置800可以包括:获取模块802、获得模块804和确定模块806。

[0150] 其中,获取模块802用于获取待处理视频数据和多个候选评论;获得模块804用于将所述待处理视频数据和所述多个候选评论输入至评论匹配模型中,获得所述待处理视频数据和各个候选评论之间的相似度;其中所述评论匹配模型是根据上述任一种方法训练获得的;确定模块806用于根据所述待处理视频数据和各个候选评论之间的相似度,从所述多个候选评论中确定目标评论。

[0151] 需要注意的是,上述附图中所示的框图是功能实体,不一定必须与物理或逻辑上独立的实体相对应。可以采用软件形式来实现这些功能实体,或在一个或多个硬件模块或集成电路中实现这些功能实体,或在不同网络和/或处理器终端设备和/或微控制器终端设备中实现这些功能实体。

[0152] 关于上述实施例中的装置,其中各个模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述,此处将不做详细阐述说明。

[0153] 图9是根据一示例性实施例示出了适于用来实现本公开示例性实施例的电子设备的结构示意图。需要说明的是,图9示出的电子设备仅仅是一个示例,不应对本公开实施例的功能和使用范围带来任何限制。

[0154] 如图9所示,电子设备900包括中央处理单元(CPU)901,其可以根据存储在只读存

存储器 (ROM) 902 中的程序或者从存储部分 908 加载到随机访问存储器 (RAM) 903 中的程序而执行各种适当的动作和处理。在 RAM 903 中, 还存储有系统 900 操作所需的各种程序和数据。CPU 901、ROM 902 以及 RAM 903 通过总线 904 彼此相连。输入/输出 (I/O) 接口 905 也连接至总线 904。

[0155] 以下部件连接至 I/O 接口 905: 包括键盘、鼠标等的输入部分 906; 包括诸如阴极射线管 (CRT)、液晶显示器 (LCD) 等以及扬声器等的输出部分 907; 包括硬盘等的存储部分 908; 以及包括诸如 LAN 卡、调制解调器等网络接口卡的通信部分 909。通信部分 909 经由诸如因特网的网络执行通信处理。驱动器 910 也根据需要连接至 I/O 接口 905。可拆卸介质 911, 诸如磁盘、光盘、磁光盘、半导体存储器等等, 根据需要安装在驱动器 910 上, 以便于从其上读出的计算机程序根据需要被安装入存储部分 908。

[0156] 特别地, 根据本公开的实施例, 上文参考流程图描述的过程可以被实现为计算机软件程序。例如, 本公开的实施例包括一种计算机程序产品, 其包括承载在计算机可读介质上的计算机程序, 该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中, 该计算机程序可以通过通信部分 909 从网络上被下载和安装, 和/或从可拆卸介质 911 被安装。在该计算机程序被中央处理单元 (CPU) 901 执行时, 执行本公开的系统中限定的上述功能。

[0157] 需要说明的是, 本公开所示的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是一——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件, 或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于: 具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器 (RAM)、只读存储器 (ROM)、可擦式可编程只读存储器 (EPROM 或闪存)、光纤、便携式紧凑磁盘只读存储器 (CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中, 计算机可读存储介质可以是任何包含或存储程序的有形介质, 该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本公开中, 计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号, 其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式, 包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可读存储介质以外的任何计算机可读介质, 该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输, 包括但不限于: 无线、电线、光缆、RF 等等, 或者上述的任意合适的组合。

[0158] 附图中的流程图和框图, 图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上, 流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分, 上述模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意, 在有些作为替换的实现中, 方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如, 两个接连地表示的方框实际上可以基本并行地执行, 它们有时也可以按相反的顺序执行, 这依所涉及的功能而定。也要注意, 框图或流程图中的每个方框、以及框图或流程图中的方框的组合, 可以用执行规定的功能或操作的专用的基于硬件的系统来实现, 或者可以用专用硬件与计算机指令的组

合来实现。

[0159] 描述于本公开实施例中所涉及到的单元可以通过软件的方式实现,也可以通过硬件的方式来实现。所描述的单元也可以设置在处理器中,例如,可以描述为:一种处理器包括发送单元、获取单元、确定单元和第一处理单元。其中,这些单元的名称在某种情况下并不构成对该单元本身的限定,例如,发送单元还可以被描述为“向所连接的服务端发送图片获取请求的单元”。

[0160] 作为另一方面,本公开还提供了一种计算机可读存储介质,该计算机可读存储介质可以是上述实施例中描述的电子设备中所包含的;也可以是单独存在,而未装配入该电子设备中。上述计算机可读存储介质承载有一个或者多个程序,当上述一个或者多个程序被一个该电子设备执行时,使得该电子设备实现如上述实施例中所述的方法。例如,所述的电子设备可以实现如图2所示的各个步骤。

[0161] 根据本公开的一个方面,提供了一种计算机程序产品或计算机程序,该计算机程序产品或计算机程序包括计算机指令,该计算机指令存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令,处理器执行该计算机指令,使得该计算机设备执行上述实施例的各种可选实现方式中提供的方法。

[0162] 需要理解的是,在本公开附图中的任何元素数量均用于示例而非限制,以及任何命名都仅用于区分,而不具有任何限制含义。

[0163] 本领域技术人员在考虑说明书及实践这里公开的发明后,将容易想到本公开的其他实施方案。本公开旨在涵盖本公开的任何变型、用途或者适应性变化,这些变型、用途或者适应性变化遵循本公开的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的,本公开的真正范围和精神由下面的权利要求指出。

[0164] 应当理解的是,本公开并不局限于上面已经描述并在附图中示出的精确结构,并且可以在不脱离其范围进行各种修改和改变。本公开的范围仅由所附的权利要求来限制。

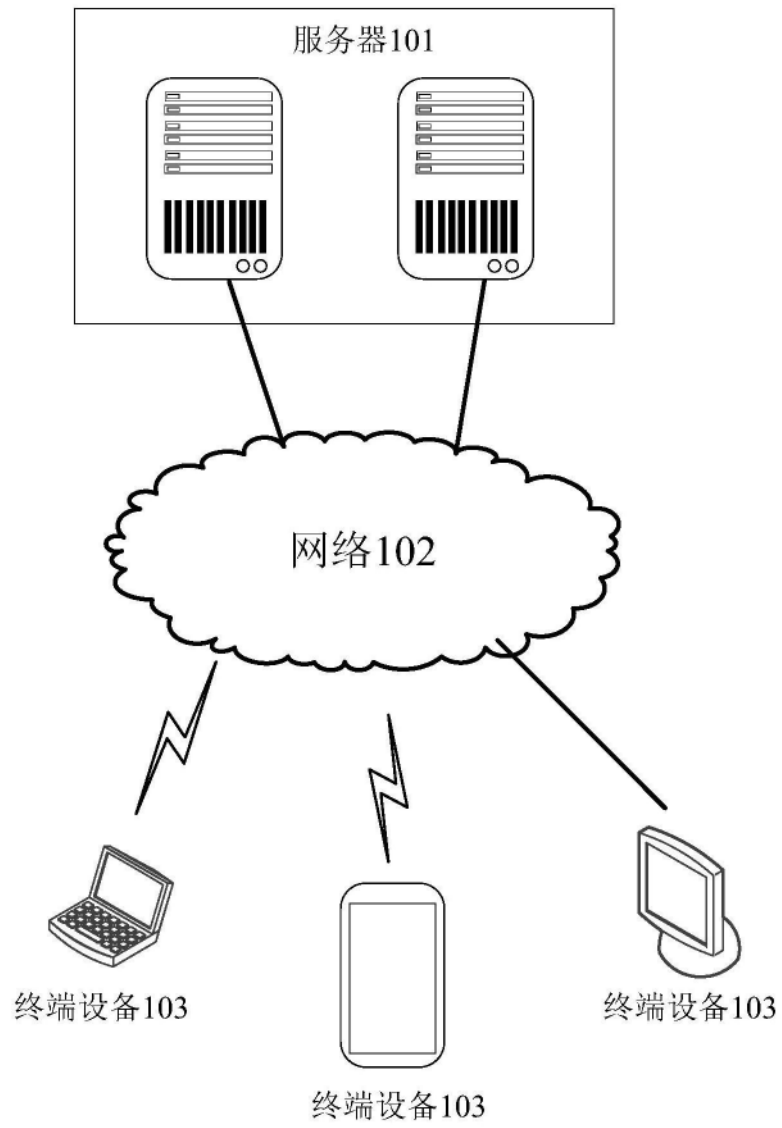


图1

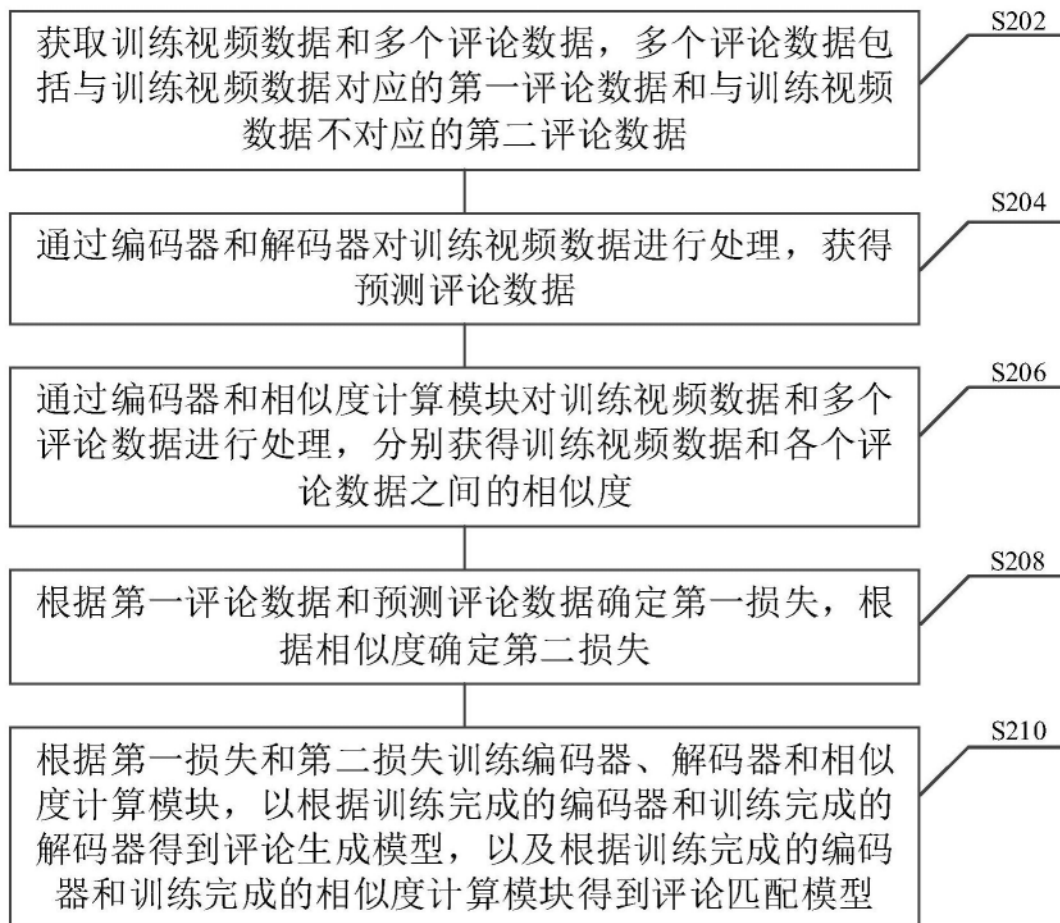


图2

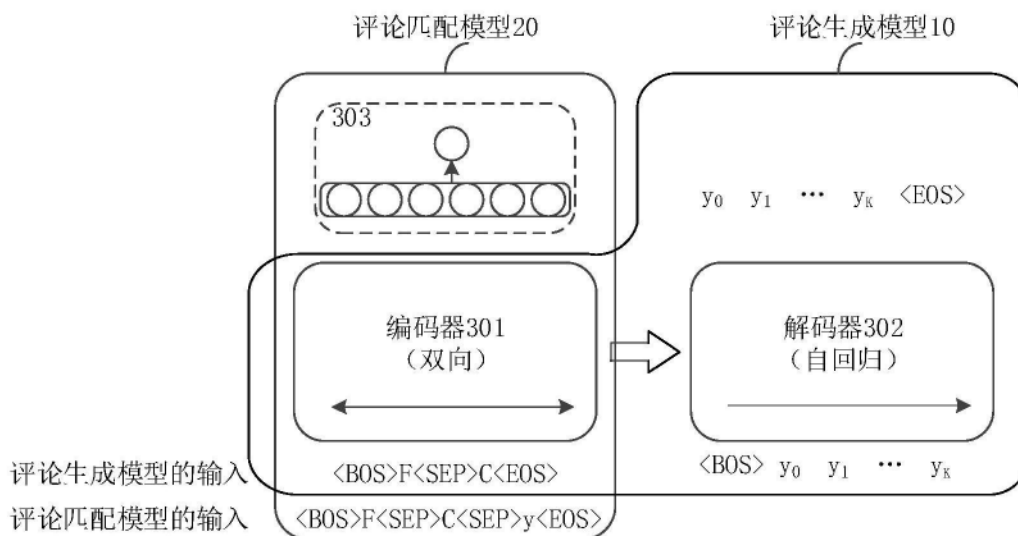


图3

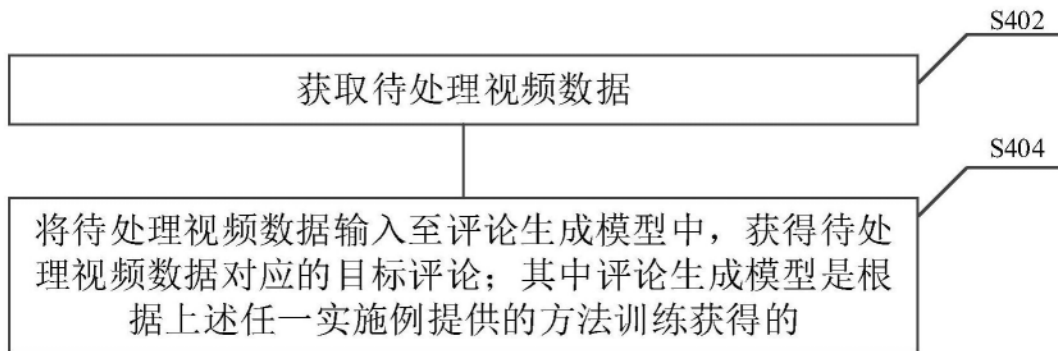
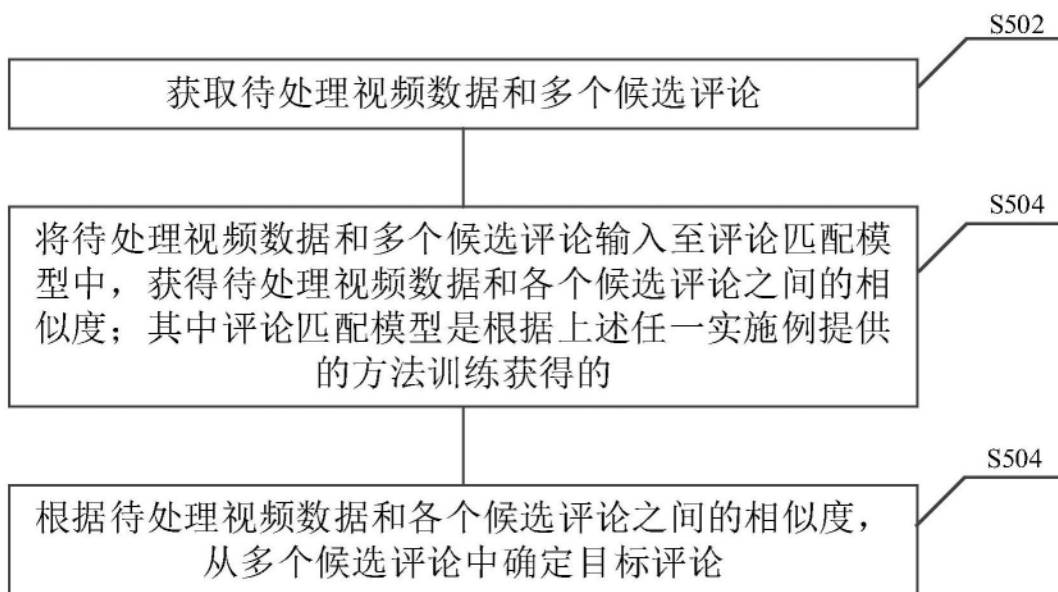




图6



图7



图8

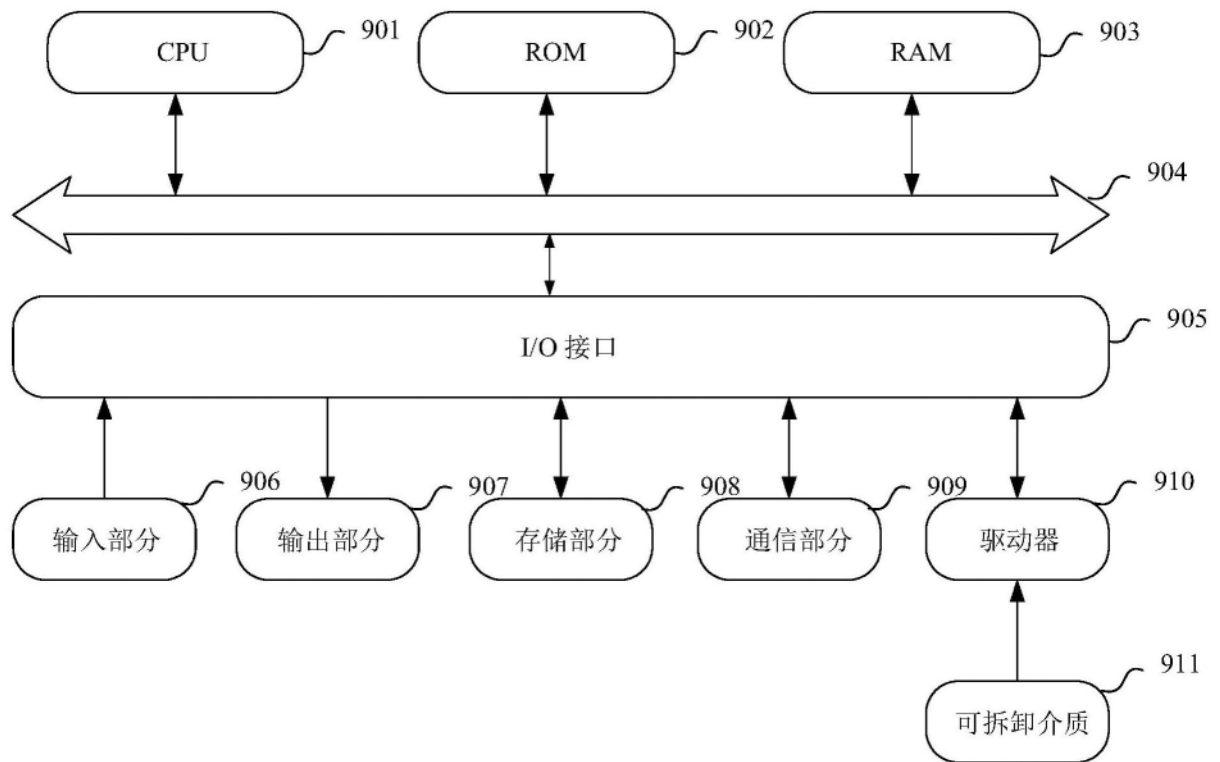
900

图9