



(12)发明专利申请

(10)申请公布号 CN 110909680 A

(43)申请公布日 2020.03.24

(21)申请号 201911156553.5

G06N 3/04(2006.01)

(22)申请日 2019.11.22

G06N 3/08(2006.01)

(71)申请人 咪咕动漫有限公司

地址 361021 福建省厦门市集美区集美大道1133号

申请人 咪咕文化科技有限公司

中国移动通信集团有限公司

(72)发明人 李鹏 裴峥 柯聪梁 陈艺勇

肖梦 李晋瑞 张伟 王堃

曾思达

(74)专利代理机构 北京路浩知识产权代理有限公司 11002

代理人 苗晓静

(51)Int.Cl.

G06K 9/00(2006.01)

权利要求书2页 说明书11页 附图2页

(54)发明名称

人脸图像的表情识别方法、装置、电子设备及存储介质

(57)摘要

本发明实施例提供一种人脸图像的表情识别方法、装置、电子设备及存储介质。方法包括：将待识别人脸图像输入表情识别模型，得到所述待识别人脸图像的表情识别结果；其中，所述表情识别模型为将人脸图像样本的五官距离信息和五官特征信息输入第一卷积神经网络训练得到；所述五官距离信息为根据人脸图像中五官的位置计算得到的距离信息；所述五官特征信息为根据人脸图像的五官特征提取的特征信息。本实施例通过人脸图像样本的五官距离信息训练表情识别模型，能够实现对人脸图像中五官位置的准确识别；通过人脸图像样本的五官特征信息训练表情识别模型，能够基于准确的五官位置，对人脸图像中丰富多样的人脸表情进行准确识别。

将待识别人脸图像输入表情识别模型，得到所述待识别人脸图像的表情识别结果 S101

1. 一种人脸图像的表情识别方法,其特征在于,包括:  
将待识别人脸图像输入表情识别模型,得到所述待识别人脸图像的表情识别结果;  
其中,所述表情识别模型为将人脸图像样本的五官距离信息和五官特征信息输入第一卷积神经网络训练得到;  
所述五官距离信息为根据人脸图像中五官的位置计算得到的距离信息;  
所述五官特征信息为根据人脸图像的五官特征提取的特征信息。
2. 根据权利要求1所述的人脸图像的表情识别方法,其特征在于,所述表情识别模型的训练过程包括:  
建立坐标系,根据人脸图像样本中五官各部分在所述坐标系中的位置,计算五官各部分之间的距离,得到所述五官距离信息;  
将五官标注后的人脸图像样本输入五官特征模型,得到所述五官特征信息;  
将所述五官距离信息和所述五官特征信息输入所述第一卷积神经网络进行训练,得到所述表情识别模型;  
其中,所述五官特征模型为将人脸图像样本的五官训练集输入第二卷积神经网络训练得到。
3. 根据权利要求2所述的人脸图像的表情识别方法,其特征在于,所述建立坐标系,根据人脸图像样本中五官各部分在所述坐标系中的位置,计算五官各部分之间的距离,得到所述五官距离信息,具体包括:  
根据中心法则确定人脸图像样本中五官的中轴,并根据所述中轴得到对应的曲面网格;  
在所述曲面网格中建立坐标系并确定坐标原点,根据所述坐标原点确定五官各部分在所述坐标系中的坐标;  
根据五官各部分在所述坐标系中的坐标,计算五官各部分之间的非欧氏距离,得到所述五官距离信息;  
其中,所述坐标原点根据人脸图像样本中鼻子的位置确定,或,根据人脸图像样本中眼睛和嘴巴的位置的平均值确定。
4. 根据权利要求2所述的人脸图像的表情识别方法,其特征在于,所述将五官标注后的人脸图像样本输入五官特征模型,得到所述五官特征信息,具体包括:  
对人脸图像样本进行五官标注,并根据预设尺寸对五官标注后的人脸图像样本进行裁剪,得到五官训练集;  
将所述五官训练集输入目标卷积神经网络进行训练,得到五官特征;  
将所述五官特征输入多通道卷积神经网络,得到所述五官特征信息;  
其中,所述五官特征信息包括以下任意组合:形态学特征、纹理特征和色彩特征;  
所述第二卷积神经网络包括所述目标卷积神经网络和所述多通道卷积神经网络。
5. 根据权利要求2所述的人脸图像的表情识别方法,其特征在于,所述将所述五官距离信息和所述五官特征信息输入所述第一卷积神经网络进行训练,得到所述表情识别模型,具体包括:  
获取不同作者的人脸图像样本,将同一人物不同作者的人脸图像样本作为数据集对,每个数据集对包括第一人脸图像样本和第二人脸图像样本;

构建生成式对抗网络,所述生成式对抗网络包括生成器和判别器;

在所述生成器中通过第三卷积神经网络对当前的第一人脸图像样本的五官距离信息和五官特征信息进行特征解码,生成伪目标人脸图像;

在所述判别器中通过第四卷积神经网络对所述伪目标人脸图像和当前的第二人脸图像样本进行判别,得到判别结果;

当所述判别结果不符合预设条件时,继续通过所述生成器和所述判别器进行训练;

当所述判别结果符合所述预设条件时,将当前的第三卷积神经网络作为所述表情识别模型。

6. 根据权利要求1-5任一项所述的人脸图像的表情识别方法,其特征在于,所述五官特征包括以下任意组合:眼睛轮廓特征、眼珠特征、眼睛变形后的结构特征、嘴巴轮廓特征、牙齿特征、舌头特征、耳朵特征、鼻子特征和发型头饰特征。

7. 根据权利要求1-5任一项所述的人脸图像的表情识别方法,其特征在于,所述表情识别结果包括以下任意组合:尺度不对称、尺度夸张、五官变形、五官比例失衡、眼睛瞳孔变形、眼睛轮廓变形和下三角变形。

8. 一种人脸图像的表情识别装置,其特征在于,包括:

表情识别模块,用于将待识别人脸图像输入表情识别模型,得到所述待识别人脸图像的表情识别结果;

其中,所述表情识别模型为将人脸图像样本的五官距离信息和五官特征信息输入第一卷积神经网络训练得到;

所述五官距离信息为根据人脸图像中五官的位置计算得到的距离信息;

所述五官特征信息为根据人脸图像的五官特征提取的特征信息。

9. 一种电子设备,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,其特征在于,所述处理器执行所述程序时实现如权利要求1-7任一项所述人脸图像的表情识别方法的步骤。

10. 一种非暂态计算机可读存储介质,其上存储有计算机程序,其特征在于,该计算机程序被处理器执行时实现如权利要求1-7任一项所述人脸图像的表情识别方法的步骤。

## 人脸图像的表情识别方法、装置、电子设备及存储介质

### 技术领域

[0001] 本发明涉及图像处理技术领域,尤其涉及一种人脸图像的表情识别方法、装置、电子设备及存储介质。

### 背景技术

[0002] 随着计算机技术、人工智能技术及其相关学科的迅猛发展,整个社会的自动化程度不断提高,人们对于人机交互的需求日益强烈,希望计算机和机器人能够像人类那样具有理解和表达情感的能力,使计算机和机器人能够更好地为人类服务。而人脸图像的表情识别是情感理解的基础,是计算机理解人类情感的前提,也是人们探索和理解人工智能的有效途径。

[0003] 现有技术对于人脸图像的表情识别仅基于五官的位置信息,对于丰富多样的表情,无法进行准确识别。

### 发明内容

[0004] 针对现有技术存在的问题,本发明实施例提供一种人脸图像的表情识别方法、装置、电子设备及存储介质。

[0005] 第一方面,本发明实施例提供一种人脸图像的表情识别方法,包括:

[0006] 将待识别人脸图像输入表情识别模型,得到所述待识别人脸图像的表情识别结果;

[0007] 其中,所述表情识别模型为将人脸图像样本的五官距离信息和五官特征信息输入第一卷积神经网络训练得到;

[0008] 所述五官距离信息为根据人脸图像中五官的位置计算得到的距离信息;

[0009] 所述五官特征信息为根据人脸图像的五官特征提取的特征信息。

[0010] 可选地,所述表情识别模型的训练过程包括:

[0011] 建立坐标系,根据人脸图像样本中五官各部分在所述坐标系中的位置,计算五官各部分之间的距离,得到所述五官距离信息;

[0012] 将五官标注后的人脸图像样本输入五官特征模型,得到所述五官特征信息;

[0013] 将所述五官距离信息和所述五官特征信息输入所述第一卷积神经网络进行训练,得到所述表情识别模型;

[0014] 其中,所述五官特征模型为将人脸图像样本的五官训练集输入第二卷积神经网络训练得到。

[0015] 可选地,所述建立坐标系,根据人脸图像样本中五官各部分在所述坐标系中的位置,计算五官各部分之间的距离,得到所述五官距离信息,具体包括:

[0016] 根据中心法则确定人脸图像样本中五官的中轴,并根据所述中轴得到对应的曲面网格;

[0017] 在所述曲面网格中建立坐标系并确定坐标原点,根据所述坐标原点确定五官各部

分在所述坐标系中的坐标；

[0018] 根据五官各部分在所述坐标系中的坐标，计算五官各部分之间的非欧氏距离，得到所述五官距离信息；

[0019] 其中，所述坐标原点根据人脸图像样本中鼻子的位置确定，或，根据人脸图像样本中眼睛和嘴巴的位置的平均值确定。

[0020] 可选地，所述将五官标注后的人脸图像样本输入五官特征模型，得到所述五官特征信息，具体包括：

[0021] 对人脸图像样本进行五官标注，并根据预设尺寸对五官标注后的人脸图像样本进行裁剪，得到五官训练集；

[0022] 将所述五官训练集输入目标卷积神经网络进行训练，得到五官特征；

[0023] 将所述五官特征输入多通道卷积神经网络，得到所述五官特征信息；

[0024] 其中，所述五官特征信息包括以下任意组合：形态学特征、纹理特征和色彩特征；

[0025] 所述第二卷积神经网络包括所述目标卷积神经网络和所述多通道卷积神经网络。

[0026] 可选地，所述将所述五官距离信息和所述五官特征信息输入所述第一卷积神经网络进行训练，得到所述表情识别模型，具体包括：

[0027] 获取不同作者的人脸图像样本，将同一人物不同作者的人脸图像样本作为数据集对，每个数据集对包括第一人脸图像样本和第二人脸图像样本；

[0028] 构建生成式对抗网络，所述生成式对抗网络包括生成器和判别器；

[0029] 在所述生成器中通过第三卷积神经网络对当前的第一人脸图像样本的五官距离信息和五官特征信息进行特征解码，生成伪目标人脸图像；

[0030] 在所述判别器中通过第四卷积神经网络对所述伪目标人脸图像和当前的第二人脸图像样本进行判别，得到判别结果；

[0031] 当所述判别结果不符合预设条件时，继续通过所述生成器和所述判别器进行训练；

[0032] 当所述判别结果符合所述预设条件时，将当前的第三卷积神经网络作为所述表情识别模型。

[0033] 可选地，所述五官特征包括以下任意组合：眼睛轮廓特征、眼珠特征、眼睛变形后的结构特征、嘴巴轮廓特征、牙齿特征、舌头特征、耳朵特征、鼻子特征和发型头饰特征。

[0034] 可选地，所述表情识别结果包括以下任意组合：尺度不对称、尺度夸张、五官变形、五官比例失衡、眼睛瞳孔变形、眼睛轮廓变形和下三角变形。

[0035] 第二方面，本发明实施例提供一种人脸图像的表情识别装置，包括：

[0036] 表情识别模块，用于将待识别人脸图像输入表情识别模型，得到所述待识别人脸图像的表情识别结果；

[0037] 其中，所述表情识别模型为将人脸图像样本的五官距离信息和五官特征信息输入第一卷积神经网络训练得到；

[0038] 所述五官距离信息为根据人脸图像中五官的位置计算得到的距离信息；

[0039] 所述五官特征信息为根据人脸图像的五官特征提取的特征信息。

[0040] 可选地，所述表情识别模型的训练过程包括：

[0041] 建立坐标系，根据人脸图像样本中五官各部分在所述坐标系中的位置，计算五官

各部分之间的距离,得到所述五官距离信息;

[0042] 将五官标注后的人脸图像样本输入五官特征模型,得到所述五官特征信息;

[0043] 将所述五官距离信息和所述五官特征信息输入所述第一卷积神经网络进行训练,得到所述表情识别模型;

[0044] 其中,所述五官特征模型为将人脸图像样本的五官训练集输入第二卷积神经网络训练得到。

[0045] 可选地,所述建立坐标系,根据人脸图像样本中五官各部分在所述坐标系中的位置,计算五官各部分之间的距离,得到所述五官距离信息,具体包括:

[0046] 根据中心法则确定人脸图像样本中五官的中轴,并根据所述中轴得到对应的曲面网格;

[0047] 在所述曲面网格中建立坐标系并确定坐标原点,根据所述坐标原点确定五官各部分在所述坐标系中的坐标;

[0048] 根据五官各部分在所述坐标系中的坐标,计算五官各部分之间的非欧氏距离,得到所述五官距离信息;

[0049] 其中,所述坐标原点根据人脸图像样本中鼻子的位置确定,或,根据人脸图像样本中眼睛和嘴巴的位置的平均值确定。

[0050] 可选地,所述将五官标注后的人脸图像样本输入五官特征模型,得到所述五官特征信息,具体包括:

[0051] 对人脸图像样本进行五官标注,并根据预设尺寸对五官标注后的人脸图像样本进行裁剪,得到五官训练集;

[0052] 将所述五官训练集输入目标卷积神经网络进行训练,得到五官特征;

[0053] 将所述五官特征输入多通道卷积神经网络,得到所述五官特征信息;

[0054] 其中,所述五官特征信息包括以下任意组合:形态学特征、纹理特征和色彩特征;

[0055] 所述第二卷积神经网络包括所述目标卷积神经网络和所述多通道卷积神经网络。

[0056] 可选地,所述将所述五官距离信息和所述五官特征信息输入所述第一卷积神经网络进行训练,得到所述表情识别模型,具体包括:

[0057] 获取不同作者的人脸图像样本,将同一人物不同作者的人脸图像样本作为数据集对,每个数据集对包括第一人脸图像样本和第二人脸图像样本;

[0058] 构建生成式对抗网络,所述生成式对抗网络包括生成器和判别器;

[0059] 在所述生成器中通过第三卷积神经网络对当前的第一人脸图像样本的五官距离信息和五官特征信息进行特征解码,生成伪目标人脸图像;

[0060] 在所述判别器中通过第四卷积神经网络对所述伪目标人脸图像和当前的第二人脸图像样本进行判别,得到判别结果;

[0061] 当所述判别结果不符合预设条件时,继续通过所述生成器和所述判别器进行训练;

[0062] 当所述判别结果符合所述预设条件时,将当前的第三卷积神经网络作为所述表情识别模型。

[0063] 可选地,所述五官特征包括以下任意组合:眼睛轮廓特征、眼珠特征、眼睛变形后的结构特征、嘴巴轮廓特征、牙齿特征、舌头特征、耳朵特征、鼻子特征和发型头饰特征。

[0064] 可选地,所述表情识别结果包括以下任意组合:尺度不对称、尺度夸张、五官变形、五官比例失衡、眼睛瞳孔变形、眼睛轮廓变形和下三角变形。

[0065] 第三方面,本发明实施例还提出一种电子设备,包括:

[0066] 至少一个处理器;以及

[0067] 与所述处理器通信连接的至少一个存储器,其中:

[0068] 所述存储器存储有可被所述处理器执行的程序指令,所述处理器调用所述程序指令能够执行上述方法。

[0069] 第四方面,本发明实施例还提出一种非暂态计算机可读存储介质,所述非暂态计算机可读存储介质存储计算机程序,所述计算机程序使所述计算机执行上述方法。

[0070] 本发明实施例提供的人脸图像的表情识别方法、装置、电子设备及存储介质,通过人脸图像样本的五官距离信息训练表情识别模型,能够实现对人脸图像中五官位置的准确识别;通过人脸图像样本的五官特征信息训练表情识别模型,能够基于准确的五官位置,对人脸图像中丰富多样的人脸表情进行准确识别。

## 附图说明

[0071] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作一简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施例,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。

[0072] 图1本发明实施例提供的人脸图像的表情识别方法的流程示意图;

[0073] 图2本发明实施例提供的人物五官特征谱的训练过程示意图;

[0074] 图3本发明实施例提供的动漫脸谱图的表情识别的流程示意图;

[0075] 图4本发明实施例提供的人脸图像的表情识别装置的结构示意图;

[0076] 图5为本发明实施例提供的电子设备的结构示意图。

## 具体实施方式

[0077] 为使本发明实施例的目的、技术方案和优点更加清楚,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0078] 图1示出了本实施例提供的一种人脸图像的表情识别方法的流程示意图,如图1所示,包括:

[0079] S101、将待识别人脸图像输入表情识别模型,得到所述待识别人脸图像的表情识别结果。

[0080] 其中,所述待识别人脸图像为将进行表情识别的人脸图像。

[0081] 所述表情识别模型为将人脸图像样本的五官距离信息和五官特征信息输入第一卷积神经网络训练得到。

[0082] 所述表情识别结果包括以下任意组合:尺度不对称、尺度夸张、五官变形、五官比例失衡、眼睛瞳孔变形、眼睛轮廓变形和下三角变形。

- [0083] 所述五官距离信息为根据人脸图像中五官的位置计算得到的距离信息。
- [0084] 所述五官特征信息为根据人脸图像的五官特征提取的特征信息。
- [0085] 所述第一卷积神经网络为用于训练表情识别模型的卷积神经网络。
- [0086] 所述五官特征包括以下任意组合:眼睛轮廓特征、眼珠特征、眼睛变形后的结构特征、嘴巴轮廓特征、牙齿特征、舌头特征、耳朵特征、鼻子特征和发型头饰特征。
- [0087] 具体地,根据人脸图像样本的五官距离信息和五官特征信息对第一卷积神经网络进行训练,训练完成后生成表情识别模型;在对待识别人脸图像进行表情识别时,只需要将待识别人脸图像输入该表情识别模型,即可输出表情识别结果,例如人脸的尺度不对称和眼睛瞳孔变形,进一步可根据表情识别结果确定人脸图像中人物的情感。
- [0088] 本实施例提供的人脸图像的表情识别方法、装置、电子设备及存储介质,通过人脸图像样本的五官距离信息训练表情识别模型,能够实现对人脸图像中五官位置的准确识别;通过人脸图像样本的五官特征信息训练表情识别模型,能够基于准确的五官位置,对人脸图像中丰富多样的人脸表情进行准确识别。
- [0089] 进一步地,在上述方法实施例的基础上,所述表情识别模型的训练过程包括:
- [0090] 建立坐标系,根据人脸图像样本中五官各部分在所述坐标系中的位置,计算五官各部分之间的距离,得到所述五官距离信息;
- [0091] 将五官标注后的人脸图像样本输入五官特征模型,得到所述五官特征信息;
- [0092] 将所述五官距离信息和所述五官特征信息输入所述第一卷积神经网络进行训练,得到所述表情识别模型;
- [0093] 其中,所述五官特征模型为将人脸图像样本的五官训练集输入第二卷积神经网络训练得到。
- [0094] 所述第二卷积神经网络为用于训练五官特征模型的卷积神经网络。
- [0095] 通过建立坐标系,方便计算五官各部分之间的距离,即方便确定五官距离信息;通过对人脸图像样本的五官进行标注后输入五官特征模型,方便确定五官特征信息。
- [0096] 进一步地,在上述方法实施例的基础上,所述建立坐标系,根据人脸图像样本中五官各部分在所述坐标系中的位置,计算五官各部分之间的距离,得到所述五官距离信息,具体包括:
- [0097] 根据中心法则确定人脸图像样本中五官的中轴,并根据所述中轴得到对应的曲面网格;
- [0098] 在所述曲面网格中建立坐标系并确定坐标原点,根据所述坐标原点确定五官各部分在所述坐标系中的坐标;
- [0099] 根据五官各部分在所述坐标系中的坐标,计算五官各部分之间的非欧氏距离,得到所述五官距离信息。
- [0100] 其中,所述坐标原点根据人脸图像样本中鼻子的位置确定,或,根据人脸图像样本中眼睛和嘴巴的位置的平均值确定。
- [0101] 所述曲面网格为根据人脸图像样本中五官的中轴建立的曲面的网格,用于在该曲面网格上建立坐标系,方便后续计算五官距离信息。
- [0102] 通过建立曲面网格中的坐标系,并在该坐标系中计算五官各部分在所述坐标系中的坐标,能够准确确定人脸图像样本的五官距离信息。



[0103] 进一步地,在上述方法实施例的基础上,所述将五官标注后的人脸图像样本输入五官特征模型,得到所述五官特征信息,具体包括:

[0104] 对人脸图像样本进行五官标注,并根据预设尺寸对五官标注后的人脸图像样本进行裁剪,得到五官训练集;

[0105] 将所述五官训练集输入目标卷积神经网络进行训练,得到五官特征;

[0106] 将所述五官特征输入多通道卷积神经网络,得到所述五官特征信息。

[0107] 其中,所述五官特征信息包括以下任意组合:形态学特征、纹理特征和色彩特征。

[0108] 所述第二卷积神经网络包括所述目标卷积神经网络和所述多通道卷积神经网络。

[0109] 具体来说,在表情识别模型的训练过程中,可以首先使用网络爬虫方法,在动漫网站与论坛中获取动漫人脸表情图片集,在获取图像后进行筛选,剔除分辨率低图像,保留较清晰图像;然后用图像处理软件进行五官标注,并进行分割裁剪,将图像批量缩放到高度宽度都为256的大小,作为五官训练集;进一步地,分别对五官训练集构建卷积神经网络进行训练,得到拓展后的五官特征,并从五官特征中提取眼睛(轮廓、眼珠)结构、眼睛变形结构、嘴巴(整体轮廓、牙齿、舌头)结构、鼻子结构、耳朵结构、发型结构和头饰结构,得到人物五官特征谱;最后通过多通道卷积网络单独提取形态学特征、纹理特征和色彩特征,能够快速实现表情识别模型的训练。

[0110] 具体地,如图2所示,通过各卷积神经网络提取训练样本中各五官部分的特征后得到各五官部分的语义特征,最终生成人物五官特征谱。

[0111] 通过提取人脸图像样本的形态学特征、纹理特征和色彩特征,能够从不同特征的角度描述人脸图像中五官的信息,尤其对于丰富多样的表情,能够准确完整地进行描述。

[0112] 进一步地,在上述方法实施例的基础上,所述将所述五官距离信息和所述五官特征信息输入所述第一卷积神经网络进行训练,得到所述表情识别模型,具体包括:

[0113] 获取不同作者的人脸图像样本,将同一人物不同作者的人脸图像样本作为数据集对,每个数据集对包括第一人臉图像样本和第二人脸图像样本;

[0114] 构建生成式对抗网络,所述生成式对抗网络包括生成器和判别器;

[0115] 在所述生成器中通过第三卷积神经网络对当前的第一人臉图像样本的五官距离信息和五官特征信息进行特征解码,生成伪目标人脸图像;

[0116] 在所述判别器中通过第四卷积神经网络对所述伪目标人脸图像和当前的第二人脸图像样本进行判别,得到判别结果;

[0117] 当所述判别结果不符合预设条件时,继续通过所述生成器和所述判别器进行训练;

[0118] 当所述判别结果符合所述预设条件时,将当前的第三卷积神经网络作为所述表情识别模型。

[0119] 其中,所述第三卷积神经网络为用于训练生成式对抗网络中生成器的卷积神经网络。

[0120] 所述第四卷积神经网络为用于训练生成式对抗网络中判别器的卷积神经网络。

[0121] 具体地,获取人脸图像样本的五官距离信息和五官特征信息后,需要将人脸图像样本进行分类,将同一人物不同作者的人脸图像样本作为数据集对,分别用于输入生成器和判别器中。

[0122] 或者,获取采集不同作者的人脸图像样本后,先对人脸图像样本进行分类,将同一人物不同作者的人脸图像样本作为数据集对,然后再获取各数据集对中人脸图像样本的五官距离信息和五官特征信息。

[0123] 举例来说,使用夸张表情编码器进行特征提取分析,并结合第三卷积神经网络提取的特征进行解码过程,作为生成式对抗网络的生成器;将分类用的第四卷积神经网络作为生成式对抗网络的判别器;然后使用数据集对对生成式对抗网络进行训练,直到满足预设条件后停止训练,得到表情识别模型。

[0124] 其中,所述预设条件为预先设置的判别结果的判断条件。

[0125] 具体地,在构建生成式对抗网络的过程中,可以设置训练数据每批大小为1个人脸图像样本。在训练生成式对抗网络时采用最小批量的随机梯度下降算法和Adam求解器更新学习网络的权重参数。

[0126] 生成器采用U-Net的结构,使输入输出图像可以共享底层信息。使用夸张表情编码器可以得到分离的各个五官信息和五官空间位置信息,使用多通道卷积神经网络对输入的图像提取形态学特征、纹理特征和色彩特征等信息;并将上述特征信息进行拼接叠加后,进行特征解码生成伪目标人脸图像。

[0127] 判别器采用第四卷积神经网络对生成器生成的伪目标人脸图像和真目标人脸图像(当前的第二人脸图像样本)一同输入。每层的卷积层按照卷积计算和ReLU激活函数的顺序方式进行对输入图像的处理。判别器中的所有层都不应用批量正规化处理。在判别器的最后一层是一个全连接层,用于匹配一维输出,接下来是Sigmoid函数。

[0128] 需要说明的是,对生成器的输入是数据集对的第一人脸图像样本,然后生成表情语义;而判别器使用生成器生成的表情语义。判别器的目标是提取人脸样本图像的表情后标注表情语义。

[0129] 上述两个卷积神经网络(第三卷积神经网络和第四卷积神经网络)以对抗的方式不断优化训练;在网络训练过程中,采用Dropout训练技巧,用于加速网络训练。

[0130] 具体来说,在训练第三卷积神经网络的过程中,具体可以包括以下步骤:

[0131] A1、将每批训练集数据大小设置为 $p$ 个,进行数据划分,其中 $p$ 为自然数;采用Adam随机梯度下降算法进行第三卷积神经网络的训练;

[0132] A2、构建具有5个卷积网络层的卷积神经网络,每个网络层带有 $k$  ( $k=16, 32, 64, 128, 256$ )个卷积核,按照卷积计算和ReLU激活函数的顺序处理输入信息。所有卷积计算中使用高和宽都是3的卷积核,1的步长。最后一层为全连接层,全连接层采用softmax激活函数,输出神经元个数为2表示对单独五官的二分类判断,如下表所示:

[0133]

名称	卷积核
卷积层1	{ $3 * 3.16$ }
卷积层2	{ $3 * 3.32$ }
池化层	$2 * 2$
卷积层3	{ $5 * 5.64$ }
卷积层4	{ $5 * 5.128$ }
卷积层5	{ $7 * 7.256$ }
全连接层	256

[0134] 另外,上述卷积神经网络有4个反卷积层,如下表所示:

[0135]	名称	卷积核
	反卷积层1	{3 * 3.128}
	反卷积层2	{3 * 3.64}
	反卷积层3	{5 * 5.32}
	反卷积层4	{7 * 7.3}

[0136] A3、设置网络层中所需要的卷积核数目和特征图数目,卷积核大小,学习率,初始权重等超参数;

[0137] A4、实现基于五官在脸部的空间位置关系变化的特征描述,针对几何形变后的五官空间关系,采用金字塔CNN网络、流形几何对五官关系进行空间上的相关性描述,根据中心法则定义五官的中轴后,画出曲面网格。以中轴线上鼻子高度位置点为坐标原点,若无鼻子则以双眼水平线与嘴巴水平线高度均值为坐标原点,建立坐标系,得到各个部分在曲面网格的非欧式距离;

[0138] A5、将分割出的五官距离信息和五官特征信息作为输入,训练出人脸图像中表情特征谱分析用的卷积神经网络,该卷积神经网络对于动漫人物的夸张表情的识别尤其有效。去掉卷积神经网络最后一层全连接,得到动漫夸张表情特征脸谱的编码器,生成目标图。

[0139] 以动漫脸谱图为例,如图3所示,将原目标动漫脸谱图输入夸张表情特征谱编码器,得到人物五官特征谱;将目标动漫脸谱图输入卷积神经网络后得到纹理色彩特征;将人物五官特征谱和纹理色彩特征输入解码器,即生成目标图。

[0140] 本实施例通过提取五官特征信息,进行语义分割后能提取到关于五官的细节部分的特征信息;基于五官在脸部的空间位置关系变化进行特征描述,结合几何形变后的五官空间关系能生成更为自然的五官位置分布;配合五官距离信息和五官特征信息能够基于准确的五官位置,对人脸图像中丰富多样的人脸表情进行准确识别。

[0141] 通过本实施例提供的人脸图像的表情识别方法,能够将用户拍摄的视频,直接与标准动作视频进行对比,不用借助其他仪器,可以得到比较准确的表情识别结果,尤其对于夸张动漫人物的表情进行分析,得到准确的表情识别结果。

[0142] 夸张动漫分析是指从给定的静态图像或动态视频序列中分离出特定的表情状态,从而确定被识别对象的表情含义,实现计算机对动漫表情的理解与识别,从根本上改变人与计算机的关系,从而达到更好的人机交互。动漫表情识别在心理学、智能机器人、智能监控、虚拟现实及合成动画等领域有很大的潜在应用价值。

[0143] 图4示出了本实施例提供的一种人脸图像的表情识别装置的结构示意图,如图4所示,包括:表情识别模块401,其中:

[0144] 所述表情识别模块401用于将待识别人脸图像输入表情识别模型,得到所述待识别人脸图像的表情识别结果;

[0145] 其中,所述表情识别模型为将人脸图像样本的五官距离信息和五官特征信息输入第一卷积神经网络训练得到;

[0146] 所述五官距离信息为根据人脸图像中五官的位置计算得到的距离信息;

[0147] 所述五官特征信息为根据人脸图像的五官特征提取的特征信息。

[0148] 本实施例提供的人脸图像的表情识别方法、装置、电子设备及存储介质,通过人脸图像样本的五官距离信息训练表情识别模型,能够实现对人脸图像中五官位置的准确识别;通过人脸图像样本的五官特征信息训练表情识别模型,能够基于准确的五官位置,对人脸图像中丰富多样的人脸表情进行准确识别。

[0149] 进一步地,在上述装置实施例的基础上,所述表情识别模型的训练过程包括:

[0150] 建立坐标系,根据人脸图像样本中五官各部分在所述坐标系中的位置,计算五官各部分之间的距离,得到所述五官距离信息;

[0151] 将五官标注后的人脸图像样本输入五官特征模型,得到所述五官特征信息;

[0152] 将所述五官距离信息和所述五官特征信息输入所述第一卷积神经网络进行训练,得到所述表情识别模型;

[0153] 其中,所述五官特征模型为将人脸图像样本的五官训练集输入第二卷积神经网络训练得到。

[0154] 进一步地,在上述装置实施例的基础上,所述建立坐标系,根据人脸图像样本中五官各部分在所述坐标系中的位置,计算五官各部分之间的距离,得到所述五官距离信息,具体包括:

[0155] 根据中心法则确定人脸图像样本中五官的中轴,并根据所述中轴得到对应的曲面网格;

[0156] 在所述曲面网格中建立坐标系并确定坐标原点,根据所述坐标原点确定五官各部分在所述坐标系中的坐标;

[0157] 根据五官各部分在所述坐标系中的坐标,计算五官各部分之间的非欧氏距离,得到所述五官距离信息;

[0158] 其中,所述坐标原点根据人脸图像样本中鼻子的位置确定,或,根据人脸图像样本中眼睛和嘴巴的位置的平均值确定。

[0159] 进一步地,在上述装置实施例的基础上,所述将五官标注后的人脸图像样本输入五官特征模型,得到所述五官特征信息,具体包括:

[0160] 对人脸图像样本进行五官标注,并根据预设尺寸对五官标注后的人脸图像样本进行裁剪,得到五官训练集;

[0161] 将所述五官训练集输入目标卷积神经网络进行训练,得到五官特征;

[0162] 将所述五官特征输入多通道卷积神经网络,得到所述五官特征信息;

[0163] 其中,所述五官特征信息包括以下任意组合:形态学特征、纹理特征和色彩特征;

[0164] 所述第二卷积神经网络包括所述目标卷积神经网络和所述多通道卷积神经网络。

[0165] 进一步地,在上述装置实施例的基础上,所述将所述五官距离信息和所述五官特征信息输入所述第一卷积神经网络进行训练,得到所述表情识别模型,具体包括:

[0166] 获取不同作者的人脸图像样本,将同一人物不同作者的人脸图像样本作为数据集对,每个数据集对包括第一人脸图像样本和第二人脸图像样本;

[0167] 构建生成式对抗网络,所述生成式对抗网络包括生成器和判别器;

[0168] 在所述生成器中通过第三卷积神经网络对当前的第一人脸图像样本的五官距离信息和五官特征信息进行特征解码,生成伪目标人脸图像;

[0169] 在所述判别器中通过第四卷积神经网络对所述伪目标人脸图像和当前的第二人

脸图像样本进行判别,得到判别结果;

[0170] 当所述判别结果不符合预设条件时,继续通过所述生成器和所述判别器进行训练;

[0171] 当所述判别结果符合所述预设条件时,将当前的第三卷积神经网络作为所述表情识别模型。

[0172] 进一步地,在上述装置实施例的基础上,所述五官特征包括以下任意组合:眼睛轮廓特征、眼珠特征、眼睛变形后的结构特征、嘴巴轮廓特征、牙齿特征、舌头特征、耳朵特征、鼻子特征和发型头饰特征。

[0173] 进一步地,在上述装置实施例的基础上,所述表情识别结果包括以下任意组合:尺度不对称、尺度夸张、五官变形、五官比例失衡、眼睛瞳孔变形、眼睛轮廓变形和下三角变形。

[0174] 本实施例所述的人脸图像的表情识别装置可以用于执行上述方法实施例,其原理和技术效果类似,此处不再赘述。

[0175] 图5示出了一种电子设备的实体结构示意图,如图5所示,该电子设备可以包括:处理器(processor) 501、通信接口(Communications Interface) 502、存储器(memory) 503和通信总线504,其中,处理器501,通信接口502,存储器503通过通信总线504完成相互间的通信。处理器501可以调用存储器503中的逻辑指令,以执行如下方法:

[0176] 将待识别人脸图像输入表情识别模型,得到所述待识别人脸图像的表情识别结果;

[0177] 其中,所述表情识别模型为将人脸图像样本的五官距离信息和五官特征信息输入第一卷积神经网络训练得到;

[0178] 所述五官距离信息为根据人脸图像中五官的位置计算得到的距离信息;

[0179] 所述五官特征信息为根据人脸图像的五官特征提取的特征信息。

[0180] 需要说明的是,本实施例中的电子设备在具体实现时可以为服务器,也可以为PC机,还可以为其他设备,只要其结构中包括如图5所示的处理器501、通信接口502、存储器503和通信总线504,其中处理器501,通信接口502,存储器503通过通信总线504完成相互间的通信,且处理器501可以调用存储器503中的逻辑指令以执行上述方法即可。本实施例不对电子设备的具体实现形式进行限定。

[0181] 此外,上述的存储器503中的逻辑指令可以通过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。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM, Read-Only Memory)、随机存取存储器(RAM, Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0182] 进一步地,本发明实施例公开一种计算机程序产品,所述计算机程序产品包括存储在非暂态计算机可读存储介质上的计算机程序,所述计算机程序包括程序指令,当所述程序指令被计算机执行时,计算机能够执行上述各方法实施例所提供的方法,例如包括:

[0183] 将待识别人脸图像输入表情识别模型,得到所述待识别人脸图像的表情识别结果;

[0184] 其中,所述表情识别模型为将人脸图像样本的五官距离信息和五官特征信息输入第一卷积神经网络训练得到;

[0185] 所述五官距离信息为根据人脸图像中五官的位置计算得到的距离信息;

[0186] 所述五官特征信息为根据人脸图像的五官特征提取的特征信息。

[0187] 另一方面,本发明实施例还提供一种非暂态计算机可读存储介质,其上存储有计算机程序,该计算机程序被处理器执行时实现以执行上述各实施例提供的传输方法,例如包括:

[0188] 将待识别人脸图像输入表情识别模型,得到所述待识别人脸图像的表情识别结果;

[0189] 其中,所述表情识别模型为将人脸图像样本的五官距离信息和五官特征信息输入第一卷积神经网络训练得到;

[0190] 所述五官距离信息为根据人脸图像中五官的位置计算得到的距离信息;

[0191] 所述五官特征信息为根据人脸图像的五官特征提取的特征信息。

[0192] 以上所描述的装置实施例仅仅是示意性的,其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元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。本领域普通技术人员在不付出创造性的劳动的情况下,即可以理解并实施。

[0193] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到各实施方式可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件。基于这样的理解,上述技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品可以存储在计算机可读存储介质中,如ROM/RAM、磁碟、光盘等,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备)执行各个实施例或者实施例的某些部分所述的方法。

[0194] 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,而非对其限制;尽管参照前述实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的精神和范围。

将待识别人脸图像输入表情识别模型，得到所述待识别人脸图像的表情识别结果 S101

图1

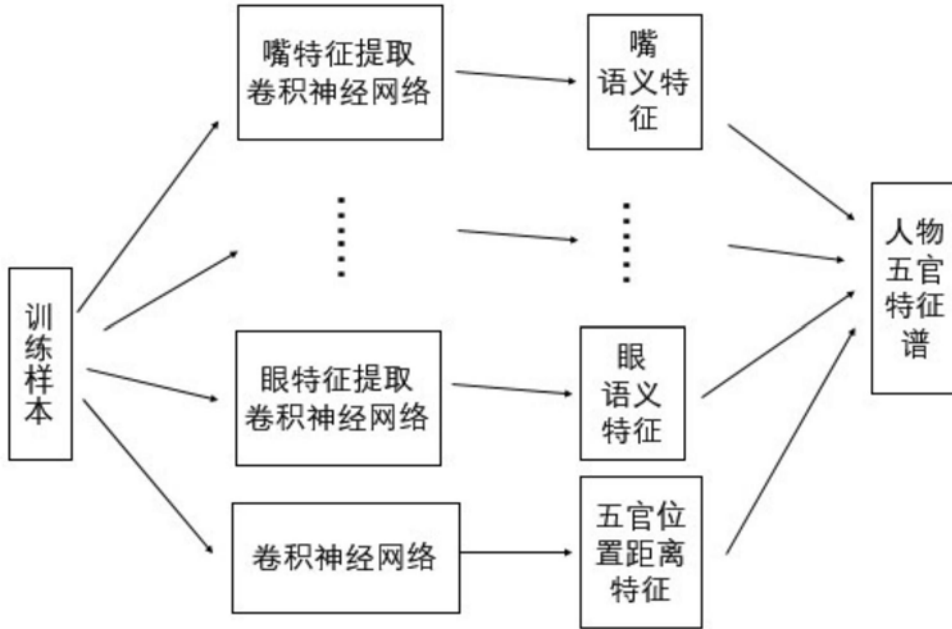


图2

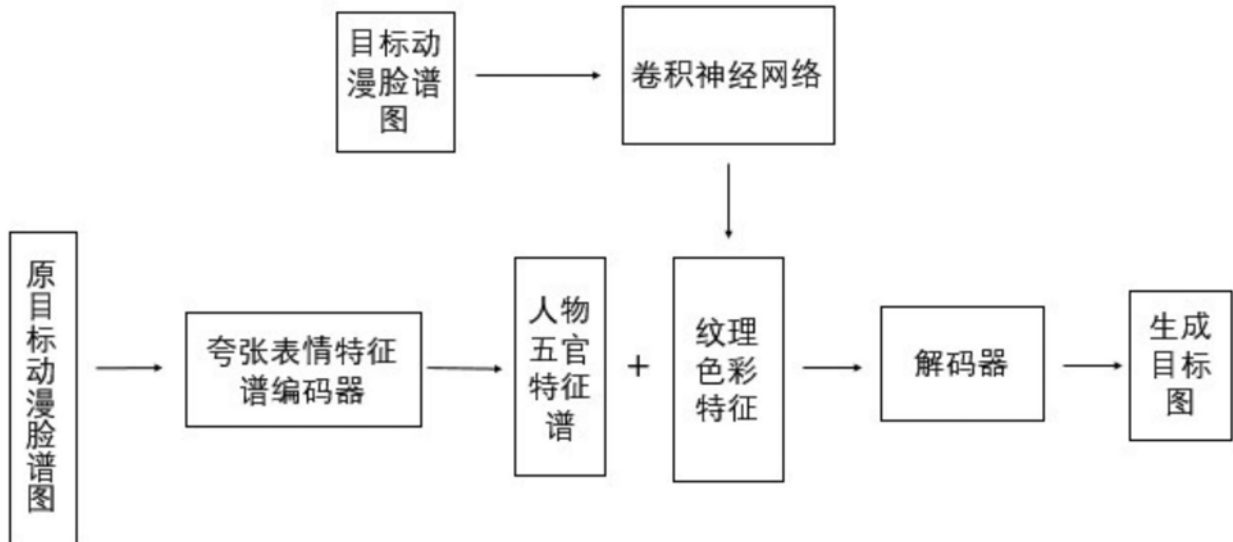


图3

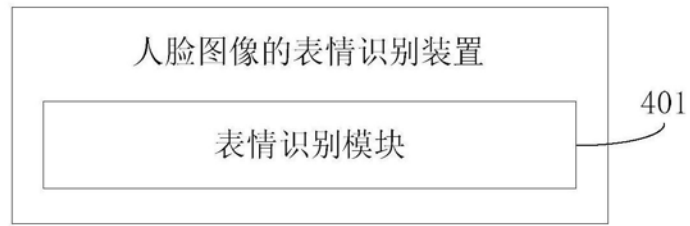


图4

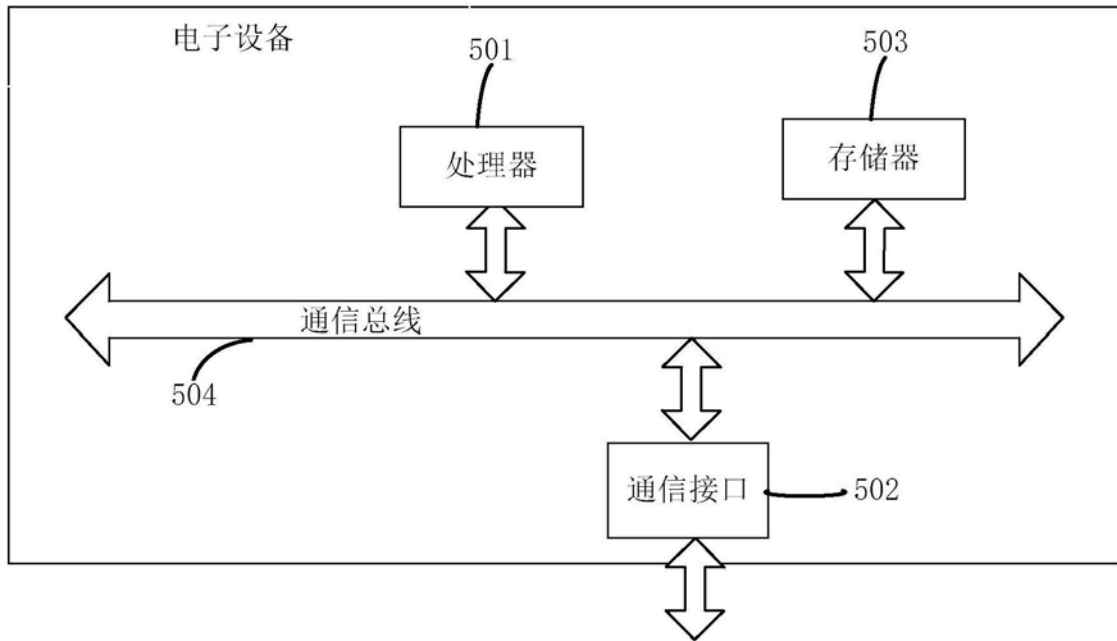


图5