(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
2024 年 9 月 6 日 (06.09.2024)

WIPO | PCT

(10) 国际公布号
WO 2024/179575 A1

(51) 国际专利分类号:
G06V 10/764 (2022.01)

(21) 国际申请号: PCT/CN2024/079587

(22) 国际申请日: 2024 年 3 月 1 日 (01.03.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202310237098.1 2023年3月2日 (02.03.2023) 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: 尹邦杰 **(YIN, Bangjie)**; 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)。 姚太平 **(YAO, Taiping)**; 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)。 张克越 **(ZHANG, Keyue)**; 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)。 李博 **(LI, Bo)**; 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)。 丁守鸿 **(DING, Shouhong)**; 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: 广州三环专利商标代理有限公司 **(SCIHEAD IP LAW FIRM)**; 中国广东省广州市越秀区先烈中路80号汇华商贸大厦1508室, Guangdong 510070 (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,

(54) **Title:** DATA PROCESSING METHOD, AND DEVICE AND COMPUTER-READABLE STORAGE MEDIUM

(54) 发明名称: 一种数据处理方法、设备以及计算机可读存储介质

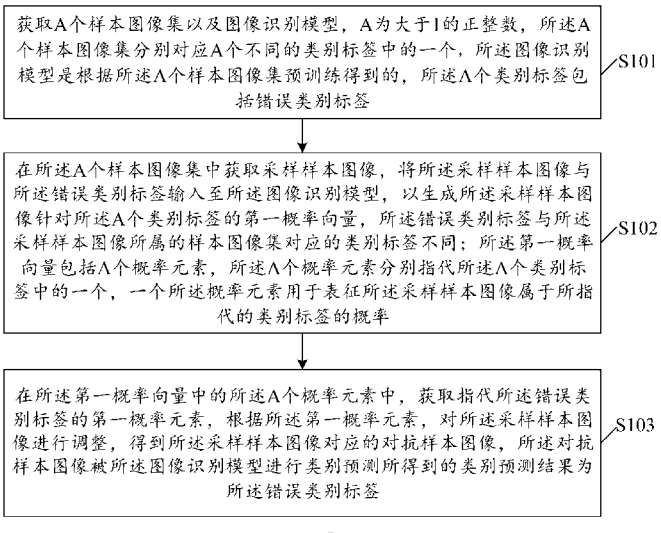


图 3

(57) **Abstract:** Provided are a data processing method, and a device and a computer-readable storage medium. The method comprises: inputting a sampling sample image in A sample image sets and an error category label among A category labels into an image recognition model, so as to generate a first probability vector of the sampling sample image in respect of the A category labels, wherein the error category label is different from a category label corresponding to a sample image set to which the sampling sample image belongs; and adjusting the sampling sample image according to a first probability element, which indicates the error category label, among A probability elements in the first probability vector, so as to obtain an adversarial sample image corresponding to the sampling sample image, wherein a category prediction result, which is obtained by means of the image recognition model performing category prediction on the adversarial sample image, is the error category label.

WO 2024/179575 A1

[见续页]

CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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))。

(57) 摘要: 提供了一种数据处理方法、设备以及计算机可读存储介质，方法包括：将A个样本图像集中的采样样本图像与A个类别标签中的错误类别标签输入至图像识别模型，以生成采样样本图像针对A个类别标签的第一概率向量，错误类别标签与采样样本图像所属的样本图像集对应的类别标签不同；根据第一概率向量中的A个概率元素中指代错误类别标签的第一概率元素对采样样本图像进行调整，得到采样样本图像对应的对抗样本图像，对抗样本图像被图像识别模型进行类别预测所得到的类别预测结果为错误类别标签。

说 明 书

一种数据处理方法、设备以及计算机可读存储介质

本申请要求于2023年3月2日提交中国专利局、申请号为202310237098.1、申请名称为“一种数据处理方法、设备以及计算机可读存储介质”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

本申请涉及互联网技术领域，尤其涉及一种数据处理方法、设备以及计算机可读存储介质。

背景技术

在图像识别场景、图像分类场景、图像推荐场景、视频属性识别场景等业务场景下，正确确定图像或视频帧的类别标签（即正确识别图像或视频帧的类别）是至关重要的。

主流的图像识别方法主要是基于机器学习的方法，即通过训练样本集训练一个图像识别初始模型，得到图像识别模型，明显地，训练数据可以决定图像识别模型的类别标签识别准确度以及识别泛化性。在现有技术中，针对训练数据的增强方式，都是直接将随机生成的噪声叠加至训练样本图像。明显地，随机生成的噪声与训练样本图像是相互独立的数据，故叠加有随机噪声的训练样本图像中的噪声特征，与图像自身的特征是相互独立的，此时，图像识别模型很容易区分噪声特征以及图像自身的特征。因此，采用现有技术所扩增的训练样本只能扩大训练数据的数量，而不能提高训练数据的多样性，故通过现有技术所扩增的训练样本，再次对图像识别模型进行训练后，并不能提高训练好的图像识别模型的类别标签识别准确度以及识别泛化性。

发明内容

本申请实施例提供一种数据处理方法、设备以及计算机可读存储介质，可以提高类别标签识别准确度以及识别泛化性。

本申请实施例一方面提供了一种数据处理方法，包括：

获取A个样本图像集以及图像识别模型，A为大于1的正整数，所述A个样本图像集分别对应A个不同的类别标签中的一个，所述图像识别模型是根据所述A个样本图像集预训练得到的，所述A个类别标签包括错误类别标签；

在所述A个样本图像集中获取采样样本图像，将所述采样样本图像与所述错误类别标签输入至所述图像识别模型，以生成所述采样样本图像针对所述A个类别标签的第一概率向量，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同；所述第一概率向量包括A个概率元素，所述A个概率元素分别指代所述A个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率；

在所述第一概率向量中的所述A个概率元素中，获取指代所述错误类别标签的第一概率元素，根据所述第一概率元素，对所述采样样本图像进行调整，得到所述采样样本图像对应的对抗样本图像，所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签。

本申请实施例一方面提供了一种数据处理方法，包括：

获取目标图像，将所述目标图像输入至图像识别优化模型；所述图像识别优化模型是根据混合样本训练集继续对图像识别模型进行优化训练所得到的；所述混合样本训练集包括A个样本图像集、A个类别标签、正确类别标签以及对抗样本图像；所述A个样本图像集分别对应A个不同的类别标签中的一个，所述A个样本图像集包括采集样本图像；所述正确类别标签属于所述A个类别标签；A为大于1的正整数；所述图像识别模型是根据所述A个样本图像集预训练得到的；所述对抗样本图像是通过第一概率向量中的A个概率元素中指代错误类别标签的第一概率元素对采样样本图像进行调整得到，所述第一概率向量是由所述图像识别模型针对所述采样样本图像所生成的，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同，所述正确类别标签为所述采样样本图像所属的样本图像集对应的类别标签；所述第一概率向量包括A个概率元素，所述A个概率元素分别指代所述A个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率；所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签；

通过所述图像识别优化模型，生成所述目标图像针对所述A个类别标签的目标概率向量；

将所述目标概率向量中的最大概率元素对应的类别标签，确定为所述目标图像的类别预测结果。

本申请实施例一方面提供了一种数据处理装置，包括：

数据获取模块，用于获取A个样本图像集以及图像识别模型，A为大于1的正整数，所述A个样本图像集分别对应A个不同的类别标签中的一个，所述图像识别模型是根据所述A个样本图像集预训练得到的，所述A个类别标签包括错误类别标签；

第一输入模块，用于在所述A个样本图像集中获取采样样本图像，将所述采样样本图像与所述错误类别标签输入至所述图像识别模型，以生成所述采样样本图像针对所述A个类别标签的第一概率向量，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同；所述第一概率向量包括A个概率元素，所述A个概率元素分别指代所述A个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率；

第一调整模块，用于在所述第一概率向量中的所述A个概率元素中，获取指代所述错误类别标签的第一概率元素，根据所述第一概率元素，对所述采样样本图像进行调整，得到所述采样样本图像对应的对抗样本图像，所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签。

其中，第一调整模块，具体用于生成第一概率元素对应的负概率元素，对负概率元素以及第一概率向量中的最大概率元素进行求和处理，得到采样样本图像针对错误类别标签的标签损失值，根据标签损失值，对采样样本图像进行调整，得到采样样本图像对应的对抗样本图像。

其中，第一调整模块用于执行根据标签损失值，对采样样本图像进行调整，得到采样样本图像对应的对抗样本图像时，具体用于根据标签损失值，生成针对采样样本图像的初始梯度值，获取初始梯度值的数值符号，根据数值符号生成初始梯度值对应的符号单位数值，对初始学习率以及符号单位数值进行乘积处理，得到第一待裁剪数值，通过由攻击强度所生成的第一裁剪区间，对第一待裁剪数值进行裁剪处理，得到用于调整采样样本图像的梯度值，对采样样本图像以及梯度值进行求差处理，得到第二待裁剪数值，通过由像素等级所生成的第二裁剪区间，对第二待裁剪数值进行裁剪处理，得到采样样本图像对应的对抗样本图像。

其中，第一调整模块，包括：

第一调整单元，用于根据第一概率元素，对采样样本图像进行调整，得到采样样本图像对应的初始对抗样本图像；

第一生成单元，用于将所述错误类别标签和所述初始对抗样本图像输入至所述图像识别模型，通过图像识别模型，生成初始对抗样本图像针对A个类别标签的第二概率向量；

第一获取单元，用于在所述第二概率向量所包含的A个概率元素中，获取针对所述错误类别标签的第二概率元素；

第一调整单元，还用于若所述第二概率元素不满足边界约束条件，或针对所述对抗样本图像进行类别预测所得到的类别预测结果不为所述错误类别标签，则根据所述第二概率元素继续对所述初始对抗样本图像进行调整；

第一获取单元，还用于若所述第二概率元素满足边界约束条件，且针对所述对抗样本图像进行类别预测所得到的类别预测结果为所述错误类别标签，则将所述初始对抗样本图像确定为所述采样样本图像对应的对抗样本图像；在所述类别预测结果为所述错误类别标签的情况下，所述第二概率元素为所述第二概率向量所包含的A个概率元素中的最大数值。

其中，第一获取单元，还用于将所述第二概率元素以及概率约束值进行对比；所述概率约束值为A的倒数；

第一获取单元，还用于若所述第二概率元素小于所述概率约束值，则确定所述第二概率元素不满足边界约束条件；

第一获取单元，还用于若所述第二概率元素等于或大于所述概率约束值，则确定所述第二概率元素满足边界约束条件。

其中，数据处理装置，还包括：

第二输入模块，用于将所述A个样本图像集、所述A个类别标签、正确类别标签以及所述对抗样本图像均输入至所述图像识别模型；所述正确类别标签是指所述采样样本图像所属的样本图像集对应的类别标签；

第二输入模块，还用于通过图像识别模型，生成A个样本图像集分别对应的第一预测类别，以及对抗样本图像对应的第二预测类别；

第二调整模块，用于根据所述第一预测类别和所述A个类别标签生成第一损失值，根据所述第二预测

类别以及所述正确类别标签生成第二损失值，根据所述第一损失值和所述第二损失值对所述图像识别模型中的参数进行调整，得到所述图像识别优化模型；所述图像识别优化模型针对所述对抗样本图像进行类别预测所得到的类别预测结果为所述正确类别标签。

其中，所述 A 个样本图像集包括样本图像集 Cb；b 为正整数且 b 小于或等于 A；第一输入模块，包括：

第一确定单元，用于确定采样比例，根据采样比例，对样本图像集 Cb 中的样本图像进行采样处理，得到采样样本图像；

第二获取单元，用于在 A 个类别标签中，获取与所述样本图像集 Cb 的类别标签不同的类别标签，将获取到的类别标签确定为错误类别标签。

其中，数据获取模块，包括：

第三获取单元，用于获取 A 个样本图像集，将 A 个类别标签和所述 A 个样本图像集分别输入至所述图像识别初始模型；

第二生成单元，用于通过图像识别初始模型，生成所述 A 个样本图像集分别对应的预测初始类别；

第二确定单元，用于根据所述预测初始类别以及所述 A 个类别标签，确定所述 A 个样本图像集分别对应的类别损失值；

第三确定单元，用于根据 A 个样本图像集分别对应的类别损失值，确定图像识别初始模型对应的总损失值；

第二调整单元，用于根据所述总损失值，对所述图像识别初始模型中的参数进行调整，得到图像识别模型；所述图像识别模型针对样本图像集 Cb 进行类别预测所得到的类别预测结果为所述样本图像集 Cb 的类别标签；所述 A 个样本图像集包括样本图像集 Cb；b 为正整数且 b 小于或等于 A。

其中，第三获取单元，包括：

图像获取子单元，用于获取 A 个原始图像集，将 A 个原始图像集均输入至对象检测模型；所述 A 个原始图像集分别对应 A 个不同的类别信息中的一个；A 个原始图像集包括原始图像集 Db；原始图像集 Db 包括原始图像 Ef；f 为正整数，且 f 小于或等于原始图像集 Db 中原始图像的总数量；

第二确定子单元，用于通过对象检测模型，确定原始图像 Ef 中关键区域的区域坐标，根据区域坐标，生成原始图像 Ef 对应的待标注样本图像；

第三确定子单元，用于将原始图像集 Db 中每个原始图像对应的待标注样本图像，确定为待标注样本图像集；

第四确定子单元，用于根据原始图像集 Db 对应的类别信息，生成类别标签 Hb，将标注有类别标签 Hb 的待标注样本图像集，确定为样本图像集 Cb。

其中，第二确定子单元，包括：

第三处理子单元，用于根据区域坐标，生成包括关键区域的初始检测框，对初始检测框进行扩张处理，得到包括目标对象的检测框；关键区域属于目标对象；

第四处理子单元，用于在检测框中，获取包括目标对象的待缩放图像，获取第一图像尺寸，根据第一图像尺寸，对待缩放图像进行缩放处理，得到待裁剪图像；

第五处理子单元，用于获取第二图像尺寸，根据第二图像尺寸，对待裁剪图像进行裁剪处理，得到原始图像 Ef 对应的待标注样本图像；第二图像尺寸小于第一图像尺寸。

本申请实施例一方面提供了一种数据处理装置，包括：

图像获取模块，用于获取目标图像，将所述目标图像输入至图像识别优化模型；所述图像识别优化模型是根据混合样本训练集继续对图像识别模型进行优化训练所得到的；所述混合样本训练集包括 A 个样本图像集、A 个类别标签、正确类别标签以及对抗样本图像；所述 A 个样本图像集分别对应 A 个不同的类别标签中的一个，所述 A 个样本图像集包括采集样本图像；所述正确类别标签属于所述 A 个类别标签；A 为大于 1 的正整数；所述图像识别模型是根据所述 A 个样本图像集预训练得到的；所述对抗样本图像是通过第一概率向量中的 A 个概率元素中指代错误类别标签的第一概率元素对采样样本图像进行调整得到，所述第一概率向量是由所述图像识别模型针对所述采样样本图像所生成的，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同，所述正确类别标签为所述采样样本图像所属的样本图像集对应的类别标签；所述第一概率向量包括 A 个概率元素，所述 A 个概率元素分别指代所述 A 个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率；所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签；

向量生成模块，用于通过图像识别优化模型，生成目标图像针对 A 个类别标签的目标概率向量；

类别确定模块，用于将目标概率向量中的最大概率元素对应的类别标签，确定为目标图像的类别预测结果。

本申请一方面提供了一种计算机设备，包括：处理器、存储器、网络接口；

上述处理器与上述存储器、上述网络接口相连，其中，上述网络接口用于提供数据通信功能，上述存储器用于存储计算机程序，上述处理器用于调用上述计算机程序，以使得计算机设备执行本申请实施例中的方法。

本申请实施例一方面提供了一种计算机可读存储介质，上述计算机可读存储介质中存储有计算机程序，上述计算机程序适于由处理器加载并执行本申请实施例中的方法。

本申请实施例一方面提供了一种计算机程序产品，该计算机程序产品包括计算机程序，该计算机程序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程序，处理器执行该计算机程序，使得该计算机设备执行本申请实施例中的方法。

在本申请实施例中，计算机设备可以获取 A 个样本图像集以及图像识别模型；其中，所述 A 个样本图像集分别对应 A 个不同的类别标签中的一个，图像识别模型是根据 A 个样本图像集预训练得到的，所以图像识别模型可以准确地确定 A 个样本图像集中的样本图像的类别标签；进一步，计算机设备在 A 个样本图像集中获取采样样本图像，将错误类别标签和采样样本图像输入至图像识别模型，其中，所述错误类别标签属于所述 A 个类别标签，且所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同；通过图像识别模型可以生成采样样本图像针对 A 个类别标签的第一概率向量，所述第一概率向量包括 A 个概率元素，所述 A 个概率元素分别指代所述 A 个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率；由于图像识别模型可以准确地确定 A 个样本图像集中的样本图像的类别标签，故在第一概率向量中，指代正确类别标签的概率元素最大，或者说，指代正确类别标签的概率元素远远大于指代错误类别标签的第一概率元素，所述正确类别标签是指所述采样样本图像所属的样本图像集对应的类别标签；进一步，根据第一概率元素，计算机设备可以对采样样本图像进行调整，得到采样样本图像对应的对抗样本图像，其中，所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签，即通过对采样样本图像进行调整，使得图像识别模型可以错误地确定对抗样本图像的类别标签。上述可知，本申请实施例提出一种对抗样本图像的生成方法，通过该方法，可以生成图像识别模型错误识别的对抗样本图像，该对抗样本图像可以提高用于优化训练图像识别模型的训练集的样本多样性，通过具有样本多样性的训练集，可以提高图像识别模型的优化训练准确度，进而可以提高优化训练好的图像识别模型的类别标签识别准确度以及识别泛化性。

附图说明

为了更清楚地说明本申请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。

图 1 是本申请实施例提供的一种系统架构示意图；

图 2a 是本申请实施例提供的一种数据处理的场景示意图一；

图 2b 是本申请实施例提供的一种数据处理的场景示意图二；

图 3 是本申请实施例提供的一种数据处理方法的流程示意图一；

图 4 是本申请实施例提供的一种通过干净样本数据所训练的分类边界面示意图；

图 5 是本申请实施例提供的一种通过第一混合样本数据所训练的分类边界面示意图；

图 6 是本申请实施例提供的一种通过第二混合样本数据所训练的分类边界面示意图；

图 7 是本申请实施例提供的一种数据处理方法的流程示意图二；

图 8 是本申请实施例提供的一种数据处理方法的流程示意图三；

图 9 是本申请实施例提供的一种数据处理装置的结构示意图一；

图 10 是本申请实施例提供的一种数据处理装置的结构示意图二；

图 11 是本申请实施例提供的一种计算机设备的结构示意图。

具体实施方式

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，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，都属于本申请保护的范围。

为了便于理解，首先对部分名词进行以下简单解释：

人工智能(Artificial Intelligence, AI)，是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能，感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。换句话说，人工智能是计算机科学的一个综合技术，它企图了解智能的实质，并生产出一种新的能以人类智能相似的方式做出反应的智能机器。人工智能也就是研究各种智能机器的设计原理与实现方法，使机器具有感知、推理与决策的功能。

人工智能技术是一门综合学科，涉及领域广泛，既有硬件层面的技术也有软件层面的技术。人工智能基础技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理技术、操作/交互系统、机电一体化等技术。人工智能软件技术主要包括计算机视觉技术、语音处理技术、自然语言处理技术以及机器学习/深度学习、自动驾驶、智慧交通等几大方向。

计算机视觉技术(Computer Vision, CV)是一门研究如何使机器“看”的科学，更进一步的说，就是指用摄影机和电脑代替人眼对目标进行识别和测量等机器视觉，并进一步做图形处理，使电脑处理成为更适合人眼观察或传送给仪器检测的图像。作为一个科学学科，计算机视觉研究相关的理论和技术，试图建立能够从图像或者多维数据中获取信息的人工智能系统。计算机视觉技术通常包括图像处理、图像识别、图像语义理解、图像检索、光学字符识别（Optical Character Recognition, OCR)、视频处理、视频语义理解、视频内容/行为识别、三维物体重建、3D 技术、虚拟现实、增强现实、同步定位与地图构建、自动驾驶、智慧交通等技术。在本申请实施例中，计算机视觉技术可以用于识别图像中的类别标签（例如人、猫、狗，或，真人、纸片人、模具人等)。

机器学习(Machine Learning, ML)是一门多领域交叉学科，涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。专门研究计算机怎样模拟或实现人类的学习行为，以获取新的知识或技能，重新组织已有的知识结构使之不断改善自身的性能。机器学习是人工智能的核心，是使计算机具有智能的根本途径，其应用遍及人工智能的各个领域。机器学习和深度学习通常包括人工神经网络、置信网络、强化学习、迁移学习、归纳学习、式教学习等技术。在本申请实施例中，图像识别模型以及对象检测模型均是基于机器学习的 AI 模型，图像识别模型可用于对图像进行识别处理，对象检测模型可用于检测图像中目标对象的关键区域。

请参见图 1，图 1 是本申请实施例提供的一种系统架构示意图。如图 1 所示，该系统可以包括业务服务器 100 以及终端设备集群，终端设备集群可以包括：终端设备 200a、终端设备 200b、终端设备 200c、…、终端设备 200n，可以理解的是，上述系统可以包括一个或者多个终端设备，本申请不对终端设备的数量进行限制。

其中，终端设备集群之间可以存在通信连接，例如终端设备 200a 与终端设备 200b 之间存在通信连接，终端设备 200a 与终端设备 200c 之间存在通信连接。同时，终端设备集群中的任一终端设备可以与业务服务器 100 存在通信连接，例如终端设备 200a 与业务服务器 100 之间存在通信连接，其中，上述通信连接不限定连接方式，可以通过有线通信方式进行直接或间接地连接，也可以通过无线通信方式进行直接或间接地连接，还可以通过其它方式，本申请在此不做限制。

应当理解，如图 1 所示的终端设备集群中的每个终端设备均可以安装有应用客户端，当该应用客户端运行于各终端设备中时，可以分别与上述图 1 所示的业务服务器 100 进行数据交互，即上述的通信连接。其中，该应用客户端可以为视频应用、直播应用、社交应用、即时通信应用、游戏应用、音乐应用、购物应用、小说应用、浏览器等具有识别目标图像功能的应用客户端。

其中，该应用客户端可以为独立的客户端，也可以为集成在某客户端（例如，社交客户端、教育客户端以及多媒体客户端等）中的嵌入式子客户端，在此不做限定。以购物应用为例，业务服务器 100 可以为包括购物应用对应的后台服务器、数据处理服务器等多个服务器的集合，因此，每个终端设备均可以通过该购物应用对应的应用客户端，与业务服务器 100 进行数据传输，如每个终端设备均可以通过购物应用的应用客户端，将目标图像上传至业务服务器 100，进而业务服务器 100 可以确定目标图像的类别标签。

可以理解的是，在本申请的具体实施方式中，涉及到用户信息（例如目标图像）等相关的数据，当本申请中的实施例运用到具体产品或技术中时，需要获得用户许可或者同意，且相关数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准。

为便于后续理解和说明，本申请实施例可以在图 1 所示的终端设备集群中选择一个终端设备作为目标终端设备，例如以终端设备 200a 作为目标终端设备。当获取到图像识别模型，并接收到针对图像识别模型的模型优化指令时，终端设备 200a 可以通过应用客户端，发送针对图像识别模型的模型优化请求至业务服务器 100。其中，图像识别模型是指用于识别图像并确定图像的类别信息的神经网络模型。

进一步，业务服务器 100 接收到终端设备 200a 发送的模型优化请求后，业务服务器 100 可以获取 A 个样本图像集以及图像识别模型，其中，A 个样本图像集即多个样本图像集，A 个样本图像集分别对应不同的类别标签，因此，类别标签的数量也可以为 A 个，即 A 个样本图像集分别对应 A 个不同的类别标签中的一个。可以理解的是，一个类别标签用于指示一个类别信息，故 A 个样本图像集分别对应的类别信息互不相同，本申请实施例不对类别信息进行限制，可以根据实际应用场景进行设定。可选的，类别信息可以是对象信息，此时，A 个样本图像集分别对应的对象信息互不相同，例如，A 个样本图像集中的第一样本图像集的对象信息是猫（即第一样本图像集中的每个样本图像均与猫相关），A 个样本图像集中的第二个样本图像集的对象信息是狗（即第二样本图像集中的每个样本图像均与狗相关）。可选的，类别信息可以是对象载体信息，此时，A 个样本图像集分别对应的对象信息可以相同，也可以不相同，但 A 个样本图像集分别对应的对象载体信息互不相同，例如在关键特征识别场景中，A 个样本图像集中的第三样本图像集的对象载体信息是翻拍屏幕，A 个样本图像集中的第四样本图像集的对象载体信息是关键特征模具，A 个样本图像集中的第五样本图像集的对象载体信息是纸片。

本申请实施例不对业务服务器 100 获取 A 个样本图像集以及图像识别模型的方式进行限定，可以根据实际应用场景进行设定，例如终端设备 200a 发送的模型优化请求中携带 A 个样本图像集或图像识别模型，或者同时携带 A 个样本图像集和图像识别模型，例如从数据库中获取 A 个样本图像集和图像识别模型，例如从区块链网络中获取 A 个样本图像集和图像识别模型。此外，图像识别模型是根据 A 个样本图像集预训练得到的，此处暂不对图像识别模型的生成过程展开描述，请参见下文图 3 所对应的实施例中步骤 S101 的的描述。

为了便于叙述以及理解，本申请实施例以 A 个样本图像集中的样本图像集 C1 为例进行叙述，可以理解的是，A 个样本图像集中的剩余样本图像集的处理过程，与下文描述的样本图像集 C1 的处理过程相同。业务服务器 100 在样本图像集 C1 中获取采样样本图像（可以随机获取或者通过轮询的方式获取），另外再在 A 个类别标签中获取错误类别标签，错误类别标签属于 A 个类别标签中除了样本图像集 C1 的类别标签之外的类别标签（可以称为剩余类别标签），具体可以将在剩余类别标签中随机获取到的类别标签确定为错误类别标签，例如错误类别标签为样本图像集 C2 的类别标签。进一步可以将错误类别标签和采样样本图像输入至图像识别模型，其中，错误类别标签和采样样本图像可以分别作为 2 个相互独立的数据输入图像识别模型，且错误类别标签和采样样本图像之间存在映射关系；或者，可以将错误类别标签添加到采样样本图像的文件头或文件尾中，进而可以将携带错误类别标签的采样样本图像输入图像识别模型。进一步，业务服务器 100 通过图像识别模型，可以生成采样样本图像针对 A 个类别标签的第一概率向量；所述第一概率向量包括 A 个概率元素，所述 A 个概率元素分别指代所述 A 个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率。可以理解的是，第一概率向量的向量维度等于 A，即一个向量维度上的元素（本申请实施例也称为概率元素），可以表征一个类别标签的预测概率值，即某个概率元素的值越大，则表示采样样本图像属于该概率元素所指代的类别标签的可能性越大。因此，通常可以将第一概率向量中数值最大的概率元素所指代的类别标签确定为采用样本图像的类别预测结果。故在第一概率向量中，业务服务器 100 可以获取指代错误类别标签的第一概率元素，进一步，根据第一概率元素，可以对采样样本图像进行调整，得到采样样本图像对应的对抗样本图像。其中所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签，即图像识别模型无法准确地识别对抗样本图像，进而得到一个错误的类别分类。例如采样样本图像为一张包含猫的图像，其类别标签是猫对应的标签，通过对携带狗对应的标签（即错误类别标签）的采样样本图像进行调整，得到其对应的对抗样本图像，图像识别模型的优化人员可以准确地判断该对抗样本图像为一张包含猫的图像，但图像识别模型会误判断该对抗样本图像为一张包含狗的图像。

通过上述过程，业务服务器 100 可以生成对抗样本图像，后续，通过 A 个样本图像集、正确类别标签以及对抗样本图像，可以继续对图像识别模型进行优化训练，得到图像识别优化模型。其中，正确类别标签是指样本图像集 C1 的类别标签。可以理解的是，由于训练图像识别模型时，是通过针对对抗样本图像的预测类别标签和正确类别标签来计算模型损失，故训练好的图像识别优化模型，可以准确地识别该对抗样本图像，进而可以准确地识别出对抗样本图像的预测类别结果是正确类别标签，因此，图像识别优化模型可以具体更强的抗干扰能力。

后续，业务服务器 100 可以生成针对图像识别优化模型的模型优化完成消息，将模型优化完成消息发送至终端设备 200a，终端设备 200a 接收到业务服务器 100 发送的模型优化完成消息后，可以获取图像识别优化模型，本申请实施例不对终端设备 200a 获取图像优化识别模型的方式进行限定，可以根据实际应

用场景进行设定，例如模型优化完成消息中携带图像优化识别模型，例如从数据库中获取图像识别优化模型，例如从区块链网络中获取图像识别优化模型。

可选的，若终端设备 200a 的本地存储了上述图像识别模型以及 A 个样本图像集，且终端设备 200a 具有离线计算能力，则终端设备 200a 在接收到针对图像识别模型的模型优化指令时，可以在本地获取样本图像集 C1 中的样本图像，将获取到的样本图像作为采样样本图像。终端设备 200a 将错误类别标签和采样样本图像输入至图像识别模型，后续处理过程与上述描述的过程相同，故此处不进行赘述。

可以理解的是，本申请实施例对采样样本图像赋予错误类别标签，并通过图像识别模型，获取指代错误类别标签的第一概率元素，通过第一概率元素对采样样本图像进行调整，得到可让图像识别模型错误预测类别的对抗样本图像。本申请实施例通过对抗样本图像，对 A 个样本图像集（即训练数据）进行训练样本的扩增以及增强，故可以提升用于优化训练图像识别模型的样本训练集的丰富性，从而提升图像识别优化模型的识别准确度。

需要说明的是，上述业务服务器 100、终端设备 200a、终端设备 200b、终端设备 200c...、终端设备 200n 均可以为区块链网络中的区块链节点，全文叙述的数据（例如 A 个样本图像集以及图像识别模型）可以进行存储，存储方式可以是区块链节点根据数据生成区块，并将区块添加至区块链中进行存储的方式。

区块链是一种分布式数据存储、点对点传输、共识机制以及加密算法等计算机技术的新型应用模式，主要用于对数据按时间顺序进行整理，并加密成账本，使其不可被篡改和伪造，同时可进行数据的验证、存储和更新。区块链本质上是一个去中心化的数据库，该数据库中的每个节点均存储一条相同的区块链，区块链网络可以将节点区分为核心节点、数据节点以及轻节点。核心节点、数据节点以及轻节点共同组成区块链节点。其中核心节点负责区块链全网的共识，也就是说核心节点为区块链网络中的共识节点。对于区块链网络中的交易数据被写入账本的流程可以为，区块链网络中的数据节点或轻节点获取到交易数据，将交易数据在区块链网络中传递（也就是节点以接力棒的方式进行传递），直到共识节点收到该交易数据，共识节点再将该交易数据打包进区块，对该区块执行共识，待共识完成后将该交易数据写入账本。此处以 A 个样本图像集以及图像识别模型示例交易数据，业务服务器 100（区块链节点）在通过对交易数据的共识后，根据交易数据生成区块，将区块存储至区块链网络中；而对于交易数据（即 A 个样本图像集以及图像识别模型）的读取，则可以由区块链节点在区块链网络中，获取到包含该交易数据的区块，进一步，在区块中获取交易数据。

可以理解的是，本申请实施例提供的方法可以由计算机设备执行，计算机设备包括但不限于终端设备或业务服务器。其中，业务服务器可以是独立的物理服务器，也可以是多个物理服务器构成的服务器集群或者分布式系统，还可以是提供云数据库、云服务、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、CDN、以及大数据和人工智能平台等基础云计算服务的云服务器。终端设备包括但不限于手机、电脑、智能语音交互设备、智能家电、车载终端、飞行器等。其中，终端设备和业务服务器可以通过有线或无线方式进行直接或间接地连接，本申请实施例在此不做限制。

进一步地，请参见图 2a，图 2a 是本申请实施例提供的一种数据处理的场景示意图一。本申请实施例可应用于各种场景，包括但不限于云技术、人工智能、智慧交通、辅助驾驶、音视频等。本申请实施例可适用于图像搜索场景、图像生成场景、图像推荐场景、图像分发场景等业务场景，这里将不对具体的业务场景进行一一列举。其中，该数据处理场景的实现过程可以在业务服务器中进行，也可以在终端设备中进行，还可以在终端设备和业务服务器中交互进行，此处不做限制。其中，终端设备可以为上述图 1 所对应实施例的终端设备集群中的任意一个终端设备，业务服务器可以为上述图 1 所对应实施例的业务服务器 100。为了便于叙述以及理解，本申请实施例以在业务服务器中进行为例进行叙述。

如图 2a 所示，业务服务器 100 可以获取 A 个样本图像集 20a 以及图像识别模型 20d，其中，A 为大于 1 的正整数，本申请实施例为了便于叙述以及理解，将 A 示例为 4，即样本图像集的数量为 4 个，如图 2a 中所示例的样本图像集 201a、样本图像集 202a、样本图像集 203a 以及样本图像集 204a（样本图像集 201a、样本图像集 202a、样本图像集 203a 以及样本图像集 204a 分别包含的样本图像的数量可以相同，也可以不相同）。4 个样本图像集分别对应的类别信息互不相同，如图 2a 中所示例的样本图像集 201a 的类别信息为萝卜，样本图像集 202a 的类别信息为辣椒，样本图像集 203a 的类别信息为南瓜，样本图像集 204a 的类别信息为洋葱。因此，业务服务器 100 可以为上述 4 个样本图像集分别设置不同的类别标签，将萝卜的类别标签设置为 0，即样本图像集 201a 的类别标签为 0，将辣椒的类别标签设置为 1，即样本图像集 202a 的类别标签为 1，将南瓜的类别标签设置为 2，即样本图像集 203a 的类别标签为 2，将洋葱的类别标签设置为 3，即样本图像集 204a 的类别标签为 3。

可以理解的是，本申请实施例是为了叙述以及理解，以萝卜、辣椒、南瓜以及洋葱示例类别信息，而实际应用时，类别信息应当根据实际场景进行设定。同理，本申请实施例是为了叙述以及理解，以0-3数值示例类别标签，而实际应用时，类别标签应当根据实际场景进行设定。

进一步，业务服务器100在上述4个样本图像集中分别获取采样样本图像，如图2a所示例，在样本图像集201a中获取采样样本图像201b，在样本图像集202a中获取采样样本图像202b，在样本图像集203a中获取采样样本图像203b，在样本图像集204a中获取采样样本图像204b。本申请实施例不对采样样本图像201b的总数量进行限定，可以为一个或多个，同理，本申请实施例不对采样样本图像202b、采样样本图像203b以及采样样本图像204b分别对应的总数量进行限定，均可以为一个或多个。

业务服务器100将采样样本图像201b、采样样本图像202b、采样样本图像203b以及采样样本图像204b均输入至图像识别模型20d，其中，图像识别模型20d是根据上述4个样本图像集所预训练得到的。可以理解的是，业务服务器100通过图像识别模型20d，对每个采样样本图像的处理过程均相同，故下文以一个采样样本图像（如图2a所示例的采样样本图像201c）为例进行描述，剩余的采样样本图像的处理过程，均可以参见下文采样样本图像201c的处理过程。

采样样本图像201c和错误类别标签可以输入至图像识别模型20d，如图2a所示例，采样样本图像201c是一张类别信息为萝卜的图像，即采样样本图像201c属于样本图像集201a。可以理解的是，采样样本图像201c的正确类别标签为萝卜的类别标签，即0，图2a以辣椒的类别标签（即1）示例采样样本图像201c的错误类别标签（错误类别标签可以通过在辣椒、南瓜以及洋葱中随机选择得到）。需要说明的是，采样样本图像201c的错误类别标签是随机生成的。

进一步，业务服务器100通过图像识别模型20d，可以生成采样样本图像201c针对4个类别标签的第一概率向量201e，可以理解的是，第一概率向量201e的维度与类别标签的总数量相同，且每个维度上的元素值，可以表示对应的类别标签的预测概率值。故在图2a中，第一概率向量201e中的第一个维度上的0.9可以表示类别标签0的预测概率，第二个维度上的0.02可以表示类别标签1的预测概率，第三个维度上的0.03可以表示类别标签2的预测概率；第四个维度上的0.05可以表示类别标签3的预测概率，这4个维度上的预测概率也可以称之为4个概率元素。通过第一概率向量201e，图像识别模型20d可以确定指代类别标签0的预测概率的数值最大，因此可以确定采样样本图像201c的类别预测结果是类别标签0，这与输入图像识别模型20d的错误类别标签（即1）不相同。产生这样的结果是因为，图像识别模型20d是通过图2a中的样本图像集201a、样本图像集202a、样本图像集203a以及样本图像集204a所预训练得到的，且采样样本图像201c来自样本图像集201a，故图像识别模型20d识别采样样本图像201c的类别预测结果为类别标签0。

进一步，业务服务器100通过图像识别模型20d可以在第一概率向量201e中获取指代错误类别标签（如图2a中所示例的类别标签1）的第一概率元素，如图2a中所示例的0.02。业务服务器100可以根据第一概率元素对采样样本图像201c进行调整，得到采样样本图像201c对应的对抗样本图像。

其中，根据第一概率元素，对采样样本图像201c进行调整，得到采样样本图像201c对应的对抗样本图像的具体过程可以包括：业务服务器100可以根据第一概率元素对采样样本图像201c进行调整，得到采样样本图像201c对应的初始对抗样本图像。请一并参见图2b，图2b是本申请实施例提供的一种数据处理的场景示意图二。如图2b所示，业务服务器100生成第一概率元素（如图2a中所示例的0.02）对应的负概率元素，即图2b中的-0.02，对负概率元素以及第一概率向量201e中的最大概率元素（如图2a以及图2b所示例的0.9）进行求和处理，得到采样样本图像201c针对错误类别标签的标签损失值，即图2b中的0.88。进一步，业务服务器100根据标签损失值，对采样样本图像201c进行调整，得到采样样本图像201c对应的初始对抗样本图像202c。其中，根据标签损失值，对采样样本图像201c进行调整，得到采样样本图像201c对应的初始对抗样本图像202c的具体过程，本申请实施例暂不展开描述，请参见下文图3所对应的实施例中步骤S103中的描述。

请再参见图2b，业务服务器100将错误类别标签（如图2a以及图2b所示例的类别标签1）和初始对抗样本图像202c，输入至图像识别模型20d。业务服务器100可以通过图像识别模型20d生成初始对抗样本图像202c针对上述4个类别标签的第二概率向量202e。其中，第二概率向量202e的含义与第一概率向量201e的含义相同，故此处不对第二概率向量202e的含义进行解释，请参见上文第一概率向量201e的含义描述。

进一步，业务服务器100在第二概率向量202e中，获取指代错误类别标签（即图2a以及图2b中所示例的类别标签1）的第二概率元素，如图2b中所示例的0.08。业务服务器100可以根据第二概率元素，

确定采样样本图像 201c 对应的对抗样本图像。即业务服务器 100 可以分析第二概率元素是否为第二概率向量 202e 中具有最大数值的概率元素，若是，则可以将初始对抗样本图像 202c 确定为采样样本图像 201c 对应的对抗样本图像（即此时图像识别模型 20d 能够错误的识别出该对抗样本图像的类别预测结果为错误类别标签 1）；若分析出第二概率元素并不是第二概率向量 202e 中具有最大数值的概率元素，则会进一步对初始对抗样本图像 202c 进行调整，其调整目的就是希望图像识别模型 20d 输出针对调整后的对初始对抗样本图像的第三概率向量中指代错误类别标签的概率元素变大，直至图像识别模型 20d 能够错误的将多轮调整后的对抗样本图像识别为错误类别标签 1。

可以理解的是，图像识别模型 20d 针对对抗样本图像的类别预测结果为错误类别标签 1，即对抗样本图像仍然为一张包含萝卜的图像，但图像识别模型 20d 会错误地将对抗样本图像的类别预测结果识别为辣椒的类别标签 1。后续，业务服务器 100 可以通过正确类别标签（即萝卜的类别标签 0）和对抗样本图像，继续对图像识别模型 20d 进行优化训练，得到图像识别优化模型。可以理解的是，图像优化识别模型（即优化训练好的图像识别模型）可以准确地识别出对抗样本图像的类别预测结果为类别标签 0，即萝卜的类别标签。

上述可知，本申请实施例提出一种对抗样本图像的生成方法，通过该方法，可以生成图像识别模型错误识别的对抗样本图像，通过对抗样本图像，可以提高用于优化训练图像识别模型的训练集的样本多样性，通过具有样本多样性的训练集，可以提高图像识别模型的优化训练准确度，进而可以提高优化训练好的图像识别模型的类别标签识别准确度以及识别泛化性。而且对抗样本图像相比于采样样本图像是一种具有强扰动的图像（即会导致图像识别模型错误识别的图像），而图像优化识别模型能够克服该强扰动，即图像优化识别模型通过更强的抗干扰能力能够准确识别出对抗样本图像的类别预测结果为正确的类别标签（即类别标签 0）。

进一步地，请参见图 3，图 3 是本申请实施例提供的一种数据处理方法的流程示意图一。该数据处理方法可以由业务服务器（例如，上述图 1 所示的业务服务器 100）执行，也可以由终端设备（例如，上述图 1 所示的终端设备 200a）执行，还可以由业务服务器和终端设备交互执行。为便于理解，本申请实施例以该方法由业务服务器执行为例进行说明。如图 3 所示，该数据处理方法至少可以包括以下步骤 S101-步骤 S103。

步骤 S101，获取 A 个样本图像集以及图像识别模型，A 为大于 1 的正整数，所述 A 个样本图像集分别对应 A 个不同的类别标签中的一个，所述图像识别模型是根据所述 A 个样本图像集预训练得到的，所述 A 个类别标签包括错误类别标签。

具体的，业务服务器获取 A 个样本图像集，A 个样本图像集分别对应的类别信息互不相同，例如某个样本图像集的类别信息为萝卜（即该样本图像集中的每个样本图像均与萝卜相关），另一个样本图像集的类别信息为辣椒（即该样本图像集中的每个样本图像均与辣椒相关）。因此，业务服务器可以为每个样本图像集分别设置不同的类别标签，如将萝卜的类别标签设置为 0，即类别信息为萝卜的样本图像集的类别标签为 0；又如将辣椒的类别标签设置为 1，即类别信息为辣椒的样本图像集的类别标签为 1。

业务服务器将 A 个样本图像集分别输入至图像识别初始模型；通过图像识别初始模型，生成每个样本图像集分别对应的预测初始类别；根据每个样本图像集分别对应的预测初始类别以及每个样本图像集分别对应的类别标签，确定每个样本图像集分别对应的类别损失值；根据 A 个样本图像集分别对应的类别损失值，确定图像识别初始模型对应的总损失值；根据总损失值，对图像识别初始模型中的参数进行调整，得到图像识别模型；图像识别模型具有准确识别 A 个样本图像集的类别预测结果的能力，如图像识别模型可以准确识别出类别信息为萝卜的样本图像集的类别预测结果为类别标签为 0（该类别预测结果也可以理解为是针对该样本图像集中每个样本图像的类别预测结果）。

本申请实施例不对样本图像集中的样本图像的图像数量进行限定，可以为一个或多个，可以理解的是，同一个样本图像集中的样本图像对应的类别信息是相同的，故同一个样本图像集中的样本图像对应的类别标签是相同的。本申请实施例也不对样本图像集的数量进行限定，可以为两个或两个以上。

其中，本申请实施例中的类别标签是指用于标识类别信息的标签，本申请实施例不对类别信息进行限定，可以根据实际应用场景进行设定。例如，类别信息可以是对象信息，此时，每个样本图像集分别包含的对象信息不同，例如猫、狗、车等。例如，类别信息可以是对象载体信息，此时，每个样本图像分别包含的对象载体信息不同，例如面具对象、模具对象、照片对象、纸质对象、真实对象等。

本申请实施例不对图像识别初始模型的模型类型进行限定，可以由任意一种或多种神经网络组成，如卷积神经网络（Convolutional Neural Networks，CNN）、残差网络（Residual Network，ResNet）、基于

扩展通道数学习机制的卷积神经网络（Wide Residual Networks，Wide-ResNet）、高分辨率网络(High-Resolution Net，HRNet)等。

业务服务器对图像识别初始模型进行训练的具体过程可以为：将A个样本图像集分别输入至图像识别初始模型，可以理解的是，图像识别初始模型对A个样本图像集中的每个样本图像的处理过程相同，故下文以携带第一类别标签的第一样本图像为例进行描述，剩余样本图像的处理过程请参见下文的描述。其中，第一样本图像为A个样本图像集中的任意一个样本图像，若第一样本图像属于样本图像集C1，则第一类别标签为样本图像集C1对应的类别标签；若第一样本图像属于样本图像集C2，则第一类别标签为样本图像集C2对应的类别标签；其他情况如上描述理解，不一一赘述。业务服务器可以通过图像识别初始模型，获取第一样本图像对应的第一初始概率向量，第一初始概率向量的维度与A相同（即一个维度对应一个类别标签），第一初始概率向量中的每个维度上的元素值（也可以称之为概率元素），均表示对应类别标签的预测概率，且每个维度上的元素值的总和为1。第一初始概率向量中最大的元素值所对应的类别标签，即第一样本图像对应的预测初始类别，可以理解的是，在训练初期，第一样本图像对应的预测初始类别可能与第一类别标签不同，故业务服务器根据第一样本图像对应的预测初始类别以及第一类别标签，可以生成第一样本图像对应的类别损失值。通过上述过程，业务服务器可以获取A个样本图像集中每个样本图像分别对应的类别损失值，将获取到的类别损失值进行求和处理，得到图像识别初始模型对应的总损失值。进一步，业务服务器根据总损失值，对图像识别初始模型中的参数进行调整，得到完成预训练的图像识别模型，该图像识别模型能够准确识别A个样本图像集中每个样本图像的类别预测结果，如图像识别模型可以准确识别出第一样本图像的类别预测结果为第一类别标签。本申请实施例不对图像识别初始模型中的参数的调整方式进行限定，可以根据实际应用场景进行设定。此外，本申请实施例不对图像识别初始模型的迭代（epoch）次数进行限定，可以根据实际应用场景进行设定，也不对模型总损失阈值进行限定，可以根据实际应用场景进行设定。

步骤S102，在所述A个样本图像集中获取采样样本图像，将所述采样样本图像与所述错误类别标签输入至所述图像识别模型，以生成所述采样样本图像针对所述A个类别标签的第一概率向量，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同；所述第一概率向量包括A个概率元素，所述A个概率元素分别指代所述A个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率。

具体的，以A个样本图像集中的样本图像集业务服务器Cb为例进行说明，b为正整数且b小于或等于A。业务服务器可以确定采样比例，根据采样比例，对样本图像集Cb中的样本图像进行采样处理，得到采样样本图像；在A个类别标签中，获取与所述样本图像集Cb的类别标签不同的类别标签（可以通过随机方式获取），将获取到的类别标签确定为错误类别标签。

请再参见图2a，业务服务器100在上述4个样本图像集中分别获取采样样本图像，其中，一种可行的采样方式如下：确定采样比例，例如5%，则业务服务器100在样本图像集201a中进行图像随机采样，得到样本图像集201a中5%的样本图像，将其组成采样样本图像201b，例如样本图像集201a中的样本图像为10000张，则采样样本图像201b的数量为500张。其他的样本图像集对应的采样样本图像（包括图2a中的采样样本图像202b、采样样本图像203b以及采样样本图像204b）的生成过程，可以参见采样样本图像201b的生成过程，此处不进行赘述。

为了便于叙述以及理解，本申请实施例示例图2a中的采样样本图像201b包括第一采样样本图像以及第二采样样本图像，可以理解的是，第一采样样本图像以及第二采样样本图像均为包含萝卜的样本图像，故第一采样样本图像以及第二采样样本图像分别对应的正确类别标签，均为萝卜的类别标签，即图2a中所示例的类别标签0。本申请实施例示例图2a中的采样样本图像202b包括第三采样样本图像以及第四采样样本图像，可以理解的是，第三采样样本图像以及第四采样样本图像均为包含辣椒的样本图像，故第三采样样本图像以及第四采样样本图像分别对应的正确类别标签，均为辣椒的类别标签，即图2a中所示例的类别标签1。

业务服务器100在图2a所示例的类别标签1，类别标签2以及类别标签3中，随机抽取一个类别标签，例如类别标签1，将其设置为第一采样样本图像的错误类别标签。同理，在图2a所示例的类别标签1，类别标签2以及类别标签3中，业务服务器100随机抽取一个类别标签，例如类别标签2，将其设置为第二采样样本图像的错误类别标签。业务服务器100在图2a所示例的类别标签0，类别标签2以及类别标签3中，随机抽取一个类别标签，例如类别标签2，将其设置为第三采样样本图像的错误类别标签。同理，在图2a所示例的类别标签0，类别标签2以及类别标签3中，业务服务器100随机抽取一个类别标签，例如

类别标签 0，将其设置为第四采样样本图像的错误类别标签。

可以理解的是，每个采样样本图像的错误类别标签的设置过程均相同，故此处不一一进行赘述，请参见上述第一采样样本图像、第二采样样本图像、第三采样样本图像以及第四采样样本图像的错误类别标签的设置过程。

若在实际应用时，图像识别模型用于识别真实对象（例如人），则本申请实施例还可以通过如下过程为采样样本图像设置错误类别标签。为了便于叙述以及理解，将 A 示例为 3，即 3 个样本图像集，分别为样本图像集 3a、样本图像集 3b 以及样本图像集 3c，样本图像集 3a 的类别信息为真实对象，其类别标签为 30a；样本图像集 3b 的类别信息是纸片对象，其类别标签为 30b，样本图像集 3c 的类别信息是面具对象，其类别标签为 30c。此时，样本图像集 3b 以及样本图像集 3c，可以视为样本图像集 3a 的攻击样本图像集。

业务服务器获取样本图像集 3a、样本图像集 3b 以及样本图像集 3c 分别对应的采样样本图像的过程，与上述描述的过程相同，故此处不进行赘述。假设样本图像集 3a 的采样样本图像包括第五采样样本图像以及第六采样样本图像，样本图像集 3b 的采样样本图像包括第七采样样本图像以及第八采样样本图像，样本图像集 3c 的采样样本图像包括第九采样样本图像以及第十采样样本图像。业务服务器 100 在类别标签 30b 以及类别标签 30c 中，随机抽取一个类别标签，例如类别标签 30b，将其设置为第五采样样本图像的错误类别标签。同理，在类别标签 30b 以及类别标签 30c 中，业务服务器 100 随机抽取一个类别标签，例如类别标签 30c，将其设置为第六采样样本图像的错误类别标签。业务服务器将第七采样样本图像、第八采样样本图像、第九采样样本图像以及第十采样样本图像分别对应的错误类别标签，均设置为类别标签 30a。故本申请实施例提供了两种错误类别标签的设置方法，实际应用时，可以根据实际需要，在上述两种设置方法中选择一种。

进一步，业务服务器将 A 个样本图像集分别对应的采样样本图像，均输入至图像识别模型。可以理解的是，图像识别模型对每个采样样本图像的处理过程均相同，故请参见上文图 2a 所对应的实施例中采样样本图像 201c 的相关描述。

可以理解的是，图像识别模型可以得到每个采样样本图像分别对应的第一概率向量，每个第一概率向量均包括 A 个概率值（也可以称之为概率元素）。由于图像识别模型是通过 A 个样本图像集所预训练得到的，因此图像识别模型对于每个采样样本图像都能输出正确的概率分布（正确的概率分布指示所输出的第一概率向量中指代正确的类别标签的概率值最大）。为了便于理解，请再参见图 2a，针对采样样本图像 201b 中的每个采样样本图像（包括采样样本图像 201c），图像识别模型 20d 输出的第一概率向量（包括图 2a 所示例的第一概率向量 201e）的第一个元素（也可以称之为概率元素，该概率元素用于指代类别标签 0，该概率元素用于表征所述采样样本图像属于所指代的类别标签 0 的概率），可以无限接近于 1，而剩余三个元素会无限接近于 0。针对采样样本图像 202b 中的每个采样样本图像，图像识别模型 20d 输出的第一概率向量的第二个元素，可以无限接近于 1，而剩余三个元素（具体为第一个元素、第三个元素以及第四个元素）会无限接近于 0。针对采样样本图像 203b 中的每个采样样本图像，图像识别模型 20d 输出的第一概率向量的第三个元素，可以无限接近于 1，而剩余三个元素（具体为第一个元素、第二个元素以及第四个元素）会无限接近于 0。针对采样样本图像 204b 中的每个采样样本图像，图像识别模型 20d 输出的第一概率向量的第四个元素，可以无限接近于 1，而剩余三个元素（具体为第一个元素、第二个元素以及第四个元素）会无限接近于 0。

步骤 S103，在所述第一概率向量中的所述 A 个概率元素中，获取指代所述错误类别标签的第一概率元素，根据所述第一概率元素，对所述采样样本图像进行调整，得到所述采样样本图像对应的对抗样本图像，所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签。

具体的，业务服务器根据第一概率元素，对采样样本图像进行调整，得到采样样本图像对应的初始对抗样本图像；将错误类别标签和初始对抗样本图像输入至图像识别模型，其中，错误类别标签和初始对抗样本图像可以分别作为 2 个相互独立的数据输入图像识别模型，且错误类别标签和初始对抗样本图像之间存在映射关系；或者，可以将错误类别标签添加到初始对抗样本图像的文件头或文件尾中，进而可以将携带错误类别标签的初始对抗样本图像输入图像识别模型。通过图像识别模型，生成初始对抗样本图像针对 A 个类别标签的第二概率向量，该第二概率向量同样包含 A 个概率元素，所述 A 个概率元素分别指代所述 A 个类别标签中的一个，一个所述概率元素用于表征初始对抗样本图像属于所指代的类别标签的概率；在第二概率向量中，获取指代错误类别标签的第二概率元素，根据第二概率元素，确定采样样本图像对应的对抗样本图像。确定对抗样本图像的其中一种方式：业务服务器可以分析第二概率元素是否为第二概率向

量中具有最大数值的概率元素，若是，则可以将初始对抗样本图像确定为采样样本图像对应的对抗样本图像（即此时图像识别模型能够错误的识别出该对抗样本图像的类别预测结果为错误类别标签）；若分析出第二概率元素并不是第二概率向量中具有最大数值的概率元素，则会进一步对初始对抗样本图像进行调整，其调整目的就是希望图像识别模型输出针对调整后的对初始对抗样本图像的第三概率向量中指代错误类别标签的概率元素变大，直至图像识别模型能够错误的将多轮调整后的对抗样本图像识别为错误类别标签。

其中，根据第一概率元素，对采样样本图像进行调整，得到采样样本图像对应的初始对抗样本图像的具体过程可以包括：生成第一概率元素对应的负概率元素，对负概率元素以及第一概率向量中的最大概率元素进行求和处理，得到采样样本图像针对错误类别标签的标签损失值；根据标签损失值，对采样样本图像进行调整，得到采样样本图像对应的初始对抗样本图像。其中，若初始对抗样本图像能够直接作为采样样本图像对应的对抗样本图像，则也可以理解为根据标签损失值，对采样样本图像进行调整，直接得到采样样本图像对应的对抗样本图像。

其中，根据标签损失值，对采样样本图像进行调整，得到采样样本图像对应的初始对抗样本图像的具体过程可以包括：根据标签损失值，生成针对采样样本图像的初始梯度值，获取初始梯度值的数值符号，根据数值符号生成初始梯度值对应的符号单位数值；对初始学习率以及符号单位数值进行乘积处理，得到第一待裁剪数值，通过由攻击强度所生成的第一裁剪区间，对第一待裁剪数值进行裁剪处理，得到用于调整采样样本图像的梯度值；对采样样本图像以及梯度值进行求差处理（梯度值也可以称之为扰动值，求差处理可以理解为在采用样本图像中添加该扰动值，以得到包含噪声的图像），得到第二待裁剪数值，通过由像素等级所生成的第二裁剪区间，对第二待裁剪数值进行裁剪处理，得到采样样本图像对应的初始对抗样本图像（若初始对抗样本图像能够直接作为采样样本图像对应的对抗样本图像，则此处的初始对抗样本图像也可以直接称之为对抗样本图像）。

具体的，业务服务器得到采样样本图像的第一概率向量后，业务服务器将第一概率向量中得最大概率元素作为第一损失值，将第一概率向量中错误类别标签对应的第一概率元素作为第二损失值，将第一损失值减去第二损失值的差值，作为采样样本图像针对错误类别标签的标签损失值。进一步，业务服务器根据标签损失值，可以对采样样本图像进行调整，得到采样样本图像的初始对抗样本图像，具体实现方式可以采取 Project Gradient Descent（一种迭代攻击算法，简称 PGD）攻击算法。其中，采取 PGD 攻击算法的具体实现过程，可以通过如下公式（1）描述。

$$I_T^{adv^t} = clip\{I_T^{adv^{t-1}} - clip\left\{\alpha * sign\left(\nabla_{I_T^{adv^{t-1}}} L_{adv}^t\right), [-eps, eps]\right\}, [0, 255]\} \quad (1)$$

其中，公式（1）中的 t 表示当前的迭代，t 为正整数，即图像识别模型会经过至少一次的迭代，也即采样样本图像会经过至少一次的调整。$I_T^{adv^t}$表示第 t 次迭代的结果数据，例如第一次迭代后所生成的初始对抗样本图像，又如最后一次迭代后所生成的对抗样本图像。$I_T^{adv^{t-1}}$表示第 t 次迭代的输入数据，例如第一次迭代的输入数据为采样样本图像，可以理解的是，若存在第二次迭代，则第二次迭代的输入数据为初始对抗样本图像，即第一迭代的结果数据；同理，若存在更多的迭代次数，则下一次迭代的输入数据，为上一次迭代的结果数据。$\nabla_{I_T^{adv^{t-1}}} L_{adv}^t = \frac{\partial L_{adv}^t}{\partial I_T^{adv^{t-1}}}$，表示初始梯度值，$L_{adv}^t$表示第 t 次迭代的标签损失值，例如第一次迭代的标签损失值。α表示学习率。$sign(.)$表示 sign 函数，sign 函数可以把初始梯度值（map）的每一个元素都变为$\{-1, 0, 1\}$。eps表示攻击强度，即每次用于调整的梯度 map 的每个元素值的绝对值不能大于该攻击强度，攻击强度是一个可调整的参数，一般被设为 8/255，16/255，32/255 等。$clip(.,.)$函数表示裁剪，即$clip(.,.)$函数可以将所有的值控制在指定的闭区间范围内，$[-eps, eps]$表示第一裁剪区间，$[0, 255]$表示第二裁剪区间。

本申请实施例不对调整图像的方式进行限定，可以为上述的 PGD 攻击方法也可以采用其他攻击方法代替，比如 Fast Gradient Sign Method（一种基于梯度生成对抗样本的算法，简称 FGSM），Iterative Fast Gradient Sign Method（改进版的 FGSM，简称 I-FGSM），Momentum Iterative Fast Gradient Sign Method（又一改进版的 FGSM，简称 MI-FGSM）等。

进一步，业务服务器将错误类别标签和初始对抗样本图像输入至图像识别模型，通过图像识别模型，

生成初始对抗样本图像针对 A 个类别标签的第二概率向量。一般而言，要实现一个成功的对抗攻击，α和迭代次数 t 都要满足一定的数量，这样才能保证调整后的采样样本图像可以跨越过图像识别模型的分界面，成为能够被图像识别模型识别为错误类别标签的样本图像（为了与对抗样本图像区分，将此处的样本图像称为强对抗样本图像，即图像识别模型输出的针对强对抗样本图像的指代错误类别标签的概率元素趋近于1）。然而，选取这样的强对抗样本图像虽然可以提升图像识别模型的鲁棒性，但是并不能作为一个合适的数据增强手段用于提升图像识别模型的精度。因为强对抗样本图像完全处于错误类别标签的分布当中，如果利用强对抗样本图像对图像识别模型进行优化训练，就会让优化训练后的图像识别模型的分界面变成不柔和的曲线，例如锯齿状，请一并参见图 4 以及图 5，图 4 是本申请实施例提供的一种通过干净样本数据所训练的分类边界面示意图，图 5 是本申请实施例提供的一种通过第一混合样本数据所训练的分类边界面示意图。其中，干净样本数据是指上述的 A 个样本图像集，第一混合样本数据包括上述的 A 个样本图像集以及上述的强对抗样本图像。为了便于叙述以及理解，A 可以等于 2，即 A 个类别标签包括图 4 以及图 5 所示例的第一类别标签以及第二类别标签。或者，A 大于 2，但一个类别标签作为真实类别标签，剩余的一个或多个类别标签作为攻击类别标签。在图 4 中，第一类别标签对应的类别特征分布区域（即符号“＋”所处的区域），以及第二类别标签对应的类别特征分布区域（即符号“—”所处的区域）相互独立，即图像识别模型可以准确地识别第一类别标签以及第二类别标签。

在图 5 中，第一类别标签对应的类别特征分布区域（即符号“＋”所处的区域以及符号“△”所处的区域），以及第二类别标签对应的类别特征分布区域（即符号“—”所处的区域以及符号“〇”所处的区域）相互交互，图像识别模型为了更好的拟合强对抗样本图像，会把分界面变成锯齿状，虽然这让优化后的图像识别模型对于强对抗样本数据的鲁棒性提升了，但是对于某些靠近分界面的置信度不高的样本图像也被错误分类，例如图 5 中存在一些符号“—”位于第一类别标签对应的类别特征区域，存在一些符号“＋”位于第二类别标签对应的类别特征区域，此时图像识别模型不能准确地识别第一类别标签以及第二类别标签，即损害了优化后的图像识别模型的精度。若 A 大于 2，且不区分真实类别标签以及攻击类别标签，则类别特征分布区域的数量与 A 相同，例如 A=3，则存在 3 个类别特征分布区域。

经过上述分析，本申请实施例要达到的目的应该是在提升鲁棒性的同时，提升图像识别模型的精度和泛化性，因此，本申请实施例加入了边界约束条件对对抗样本图像进行筛选。换言之，用于数据增强的对抗样本图像不应该越过边界太远，因此，确定对抗样本图像的另一种方式可以为：业务服务器根据所述第一概率元素，对所述采样样本图像进行调整，得到所述采样样本图像对应的初始对抗样本图像；将所述错误类别标签和所述初始对抗样本图像输入至所述图像识别模型，通过所述图像识别模型，生成所述初始对抗样本图像针对所述 A 个类别标签的第二概率向量；在所述第二概率向量所包含的 A 个概率元素中，获取针对所述错误类别标签的第二概率元素；若所述第二概率元素不满足边界约束条件，或针对所述对抗样本图像进行类别预测所得到的类别预测结果不为所述错误类别标签，则根据所述第二概率元素继续对所述初始对抗样本图像进行调整（继续对所述初始对抗样本图像进行调整的过程与根据第一概率元素，对采样样本图像进行调整，得到初始对抗样本图像的具体实现过程相同，故此处不进行赘述，请参见上文的描述。可以理解的是，在后续的模型迭代过程（即图像调整过程）中，图像调整方式都与采样样本图像的调整方式相同）；若所述第二概率元素满足边界约束条件，且针对所述对抗样本图像进行类别预测所得到的类别预测结果为所述错误类别标签，则将所述初始对抗样本图像确定为所述采样样本图像对应的对抗样本图像；在所述类别预测结果为所述错误类别标签的情况下，所述第二概率元素为所述第二概率向量所包含的 A 个概率元素中的最大数值。可见，本申请所期望的对抗样本图像是既能满足边界约束条件，又能保证图像识别模型能够错误识别对抗样本图像。其中，业务服务器还可以将所述第二概率元素以及概率约束值进行对比；所述概率约束值为 A 的倒数；若所述第二概率元素小于所述概率约束值，则确定所述第二概率元素不满足边界约束条件；若所述第二概率元素等于或大于所述概率约束值，则确定所述第二概率元素满足边界约束条件。即本申请通过控制学习率和迭代次数，若概率向量（包括上述的第二概率向量）中的错误类别标签所对应的概率元素（包括上述的第二概率元素）恰好超过 1/A，且此时的错误类别标签所对应的概率元素为概率向量中的最大值（具有最大值的概率元素用于指示图像识别模型针对初始对抗样本图像的类别预测结果为该概率元素所指代的类别标签），就停止模型迭代（即图像调整），将当前的调整好的样本图像作为一个合理的对抗样本图像，而不是等错误类别标签所对应的概率元素趋近于 1。

基于上述，本申请实施例所得到的对抗样本图像可以处于分类边界面附近，而不是在错误类别标签的分布之中，进而通过对抗样本图像所优化训练的图像识别模型的分类边界面就会更加柔和，请一并参见图 6，图 6 是本申请实施例提供的一种通过第二混合样本数据所训练的分类边界面示意图。第二混合样本数

据包括上述的A个样本图像集以及上述的对抗样本图像。如图6所示，第一类别标签对应的类别特征分布区域（即符号“十”所处的区域以及符号“△”所处的区域），以及第二类别标签对应的类别特征分布区域（即符号“一”所处的区域以及符号“〇”所处的区域）基本不交互，且存在少数交互的类别特征也基本位于分类边界面，此时优化后的图像识别模型可以准确地识别第一类别标签以及第二类别标签，即能够保证优化后的图像识别模型（简称图像识别优化模型）保持或提高原有识别精度的基础上，还可以补充训练数据的不足之处，故可以提升图像识别优化模型的防御能力以及安全性。

本申请实施例提出了一种全新的基于边界约束的对抗数据增强，利用对抗数据在边界分布的特殊性，对训练数据进行扩增和增强，从而提升优化后的图像识别模型的防御能力和泛化性。

在本申请实施例中，计算机设备可以获取A个样本图像集以及图像识别模型；其中，所述A个样本图像集分别对应A个不同的类别标签中的一个，图像识别模型是根据A个样本图像集预训练得到的，所以图像识别模型可以准确地确定A个样本图像集中的样本图像的类别标签；进一步，计算机设备在A个样本图像集中获取采样样本图像，将错误类别标签和采样样本图像输入至图像识别模型，其中，所述错误类别标签属于所述A个类别标签，且所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同；通过图像识别模型可以生成采样样本图像针对A个类别标签的第一概率向量，所述第一概率向量包括A个概率元素，所述A个概率元素分别指代所述A个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率；由于图像识别模型可以准确地确定A个样本图像集中的样本图像的类别标签，故在第一概率向量中，指代正确类别标签的概率元素最大，或者说，指代正确类别标签的概率元素远远大于指代错误类别标签的第一概率元素，所述正确类别标签是指所述采样样本图像所属的样本图像集对应的类别标签；进一步，根据第一概率元素，计算机设备可以对采样样本图像进行调整，得到采样样本图像对应的对抗样本图像，其中，所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签，即通过对采样样本图像进行调整，使得图像识别模型可以错误地确定对抗样本图像的类别标签。上述可知，本申请实施例提出一种对抗样本图像的生成方法，通过该方法，可以生成图像识别模型错误识别的对抗样本图像，该对抗样本图像可以提高用于优化训练图像识别模型的训练集的样本多样性，通过具有样本多样性的训练集，可以提高图像识别模型的优化训练准确度，进而可以提高优化训练好的图像识别模型的类别标签识别准确度以及识别泛化性。而且由于所述图像识别模型针对所述对抗样本图像所预测的针对所述错误类别标签的第二概率元素满足边界约束条件，所以就可以使得优化训练后的图像识别模型的分类边界更加柔和，进一步保证了模型精度。

请参见图7，图7是本申请实施例提供的一种数据处理方法的流程示意图二。该方法可以由业务服务器（例如，上述图1所示的业务服务器100）执行，也可以由终端设备（例如，上述图1所示的终端设备200a）执行，还可以由业务服务器和终端设备交互执行。为便于理解，本申请实施例以该方法由业务服务器执行为例进行说明。如图7所示，该方法至少可以包括以下步骤。

步骤S201，获取A个样本图像集以及图像识别模型，A为大于1的正整数，所述A个样本图像集分别对应A个不同的类别标签中的一个，所述图像识别模型是根据所述A个样本图像集预训练得到的，所述A个类别标签包括错误类别标签。

具体的，业务服务获取A个原始图像集，将A个原始图像集均输入至对象检测模型；A个原始图像集分别对应A个不同的类别信息中的一个；A个原始图像集包括原始图像集D_b；原始图像集D_b包括原始图像E_f；f为正整数，且f小于或等于原始图像集D_b中原始图像的总数量；通过对象检测模型，确定原始图像E_f中关键区域的区域坐标，根据区域坐标，生成原始图像E_f对应的待标注样本图像；将原始图像集D_b中每个原始图像对应的待标注样本图像，确定为待标注样本图像集；根据原始图像集Db对应的类别信息，生成类别标签Hb，将标注有类别标签Hb的待标注样本图像集，确定为样本图像集Cb。

其中，根据区域坐标，生成原始图像Ef对应的待标注样本图像的具体过程可以包括：根据区域坐标，生成包括关键区域的初始检测框，对初始检测框进行扩张处理，得到包括目标对象的检测框；关键区域属于目标对象；在检测框中，获取包括目标对象的待缩放图像，获取第一图像尺寸，根据第一图像尺寸，对待缩放图像进行缩放处理，得到待裁剪图像；获取第二图像尺寸，根据第二图像尺寸，对待裁剪图像进行裁剪处理，得到原始图像Ef对应的待标注样本图像；第二图像尺寸小于第一图像尺寸。

业务服务器不对对象检测模型的模型类型进行限定，可以根据实际应用场景进行设定，例如Region with CNN feature（一种目标检测模型，简称Faster R-CNN），Single Shot MultiBox Detector（另一种目标检测模型，简称SSD），You Only Look Once（一种简易快捷的目标检测算法，简称YOLO）等。

其中，关键区域属于目标对象，本申请实施例不对目标对象进行限定，例如可以是猫、狗、人、桌子、

辣椒等，故不对关键区域进行限定，应当根据目标对象进行设定。其中，初始检测框的扩张倍数，应当根据关键区域进行设定，故本申请实施例不对扩张倍数进行限定。其中，第二图像尺寸是指图像识别模型的输入尺寸。

步骤 S202，在所述 A 个样本图像集中获取采样样本图像，将所述采样样本图像与所述错误类别标签输入至所述图像识别模型，以生成所述采样样本图像针对所述 A 个类别标签的第一概率向量，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同；所述第一概率向量包括 A 个概率元素，所述 A 个概率元素分别指代所述 A 个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率。

步骤 S203，在所述第一概率向量中的所述 A 个概率元素中，获取指代所述错误类别标签的第一概率元素，根据所述第一概率元素，对所述采样样本图像进行调整，得到所述采样样本图像对应的对抗样本图像，所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签。

其中，步骤 S202-步骤 S203 的具体实现过程，请参见上文图 3 所对应的实施例中的步骤 S102 步骤 S103，此处不进行赘述。

步骤 S204，将所述 A 个样本图像集、所述 A 个类别标签、正确类别标签以及所述对抗样本图像均输入至所述图像识别模型；所述正确类别标签是指所述采样样本图像所属的样本图像集对应的类别标签。

其中，所述 A 个样本图像集、所述 A 个类别标签、正确类别标签以及所述对抗样本图像可以分别作为 4 个相互独立的数据输入图像识别模型，且所述 A 个样本图像集与所述 A 个类别标签之间存在映射关系，正确类别标签与所述对抗样本图像之间存在映射关系；或者，可以将所属的类别标签添加到样本图像集中的样本图像的文件头或文件尾中，且将正确类别标签添加到对抗样本图像的文件头或文件尾中，进而可以将携带类别标签的样本图像以及携带正确类别标签的对抗样本图像输入图像识别模型。

步骤 S205，通过图像识别模型，生成 A 个样本图像集分别对应的第一预测类别，以及对抗样本图像对应的第二预测类别。

步骤 S206，根据所述第一预测类别和所述 A 个类别标签生成第一损失值，根据所述第二预测类别以及所述正确类别标签生成第二损失值，根据所述第一损失值和所述第二损失值对所述图像识别模型中的参数进行调整（具体可以将第一损失值和第二损失值进行加权求和，得到总损失值，然后基于该总损失值对图像识别模型中的参数进行调整），得到所述图像识别优化模型；所述图像识别优化模型针对所述对抗样本图像进行类别预测所得到的类别预测结果为所述正确类别标签，且图像识别优化模型针对样本图像集中的样本图像进行类别预测所得到的类别预测结果依然为样本图像集所对应的类别标签。

其中，业务服务器对图像识别模型的优化训练的过程，与业务服务器对图像识别初始模型的训练过程相同，故步骤 S204-步骤 S206 的具体实现过程，请参见上文图 3 所对应的实施例中的步骤 S101，此处不进行赘述。

本申请实施例提出的图像识别优化模型可以部署于终端设备，例如对真实对象识别（例如真人识别）输入的图像进行检测，如果是真实对象，则通过并进入到后续的识别流程中，如果不是真实对象，例如照片对象，则进行报错，并提示重试。本申请实施例可以应用于所有对象识别相关的应用，包括但不限于：线上支付，线下支付，门禁解锁系统，手机解锁识别，自动对象识别通关等。

上述可知，本申请实施例提出一种对抗样本图像的生成方法，通过该方法，可以生成图像识别模型错误识别的对抗样本图像，通过对抗样本图像，可以提高用于优化训练图像识别模型的训练集的样本多样性，通过具有样本多样性的训练集，可以提高图像识别模型的优化训练准确度，进而可以提高优化训练好的图像识别模型的类别标签识别准确度以及识别泛化性。

请参见图 8，图 8 是本申请实施例提供的一种数据处理方法的流程示意图三。该方法可以由业务服务器（例如，上述图 1 所示的业务服务器 100）执行，也可以由终端设备（例如，上述图 1 所示的终端设备 200a）执行，还可以由业务服务器和终端设备交互执行。为便于理解，本申请实施例以该方法由业务服务器执行为例进行说明。如图 8 所示，该方法至少可以包括以下步骤。

步骤 S301，获取目标图像，将所述目标图像输入至图像识别优化模型；所述图像识别优化模型是根据混合样本训练集继续对图像识别模型进行优化训练所得到的；所述混合样本训练集包括 A 个样本图像集、A 个类别标签、正确类别标签以及对抗样本图像；所述 A 个样本图像集分别对应 A 个不同的类别标签中的一个，所述 A 个样本图像集包括采集样本图像；所述正确类别标签属于所述 A 个类别标签；A 为大于 1 的正整数；所述图像识别模型是根据所述 A 个样本图像集预训练得到的；所述对抗样本图像是通过第一概率向量中的 A 个概率元素中指代错误类别标签的第一概率元素对采样样本图像进行调整得到，所述第一概率

向量是由所述图像识别模型针对所述采样样本图像所生成的，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同，所述正确类别标签为所述采样样本图像所属的样本图像集对应的类别标签；所述第一概率向量包括 A 个概率元素，所述 A 个概率元素分别指代所述 A 个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率；所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签。

步骤 S302，通过图像识别优化模型，生成目标图像针对 A 个类别标签的目标概率向量。

步骤 S303，将目标概率向量中的最大概率元素对应的类别标签，确定为目标图像的类别预测结果。

具体的，本申请实施例的图像识别模型可以用于识别真实对象，例如真实的人，此时，图像识别模型可以称为真实对象检测模型。目前智能对象识别支付业务中，越来越多的真实对象检测技术被应用到产品中，有力的保障了支付的安全。然而，在真实对象检测的实际应用场景中，非法分子带来的攻击种类繁多，攻击形式新颖，防不胜防，比如在纸片攻击当中，样式繁多的材质中总有一款模型没训练过的材质，故可以绕过真实对象检测模型，从而实现攻击。而在高精度的 3 维攻击当中也面临这个问题，一旦有新的材质出现，真实对象检测模型就会面临防不住的情况。究其原因，还是训练数据没有覆盖整个类别空间导致的。因此，本申请实施例提出了一种全新的基于边界约束的对抗数据增强的真实对象检测技术，通过设计一种特殊的对抗样本图像来对真实对象训练数据进行增强，从而整体提升真实对象检测的性能。对抗样本图像的具体生成过程，请参见上述图 3 以及图 7 所分别对应的实施例中的描述。

上述可知，本申请实施例提出一种对抗样本图像的生成方法，通过该方法，可以生成图像识别模型错误识别的对抗样本图像，通过对抗样本图像，可以提高用于优化训练图像识别模型的训练集的样本多样性，通过具有样本多样性的训练集，可以提高图像识别模型的优化训练准确度，进而可以提高优化训练好的图像识别模型的类别标签识别准确度以及识别泛化性。

进一步地，请参见图 9，图 9 是本申请实施例提供的一种数据处理装置的结构示意图一。上述数据处理装置 1 可以用于执行本申请实施例提供的方法中的相应步骤。如图 9 所示，该数据处理装置 1 可以包括：数据获取模块 11、第一输入模块 12 以及第一调整模块 13。

数据获取模块 11，用于获取 A 个样本图像集以及图像识别模型，A 为大于 1 的正整数，所述 A 个样本图像集分别对应 A 个不同的类别标签中的一个，所述图像识别模型是根据所述 A 个样本图像集预训练得到的，所述 A 个类别标签包括错误类别标签；

第一输入模块 12，用于在所述 A 个样本图像集中获取采样样本图像，将所述采样样本图像与所述错误类别标签输入至所述图像识别模型，以生成所述采样样本图像针对所述 A 个类别标签的第一概率向量，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同；所述第一概率向量包括 A 个概率元素，所述 A 个概率元素分别指代所述 A 个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率；

第一调整模块 13，用于在所述第一概率向量中的所述 A 个概率元素中，获取指代所述错误类别标签的第一概率元素，根据所述第一概率元素，对所述采样样本图像进行调整，得到所述采样样本图像对应的对抗样本图像，所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签。

其中，数据获取模块 11、第一输入模块 12 以及第一调整模块 13 的具体功能实现方式，可以参见上述图 3 对应实施例中的步骤 S101-你在 S103，这里不再进行赘述。

其中，第一调整模块 13，可以具体用于生成第一概率元素对应的负概率元素，对负概率元素以及第一概率向量中的最大概率元素进行求和处理，得到采样样本图像针对错误类别标签的标签损失值，根据标签损失值，对采样样本图像进行调整，得到采样样本图像对应的对抗样本图像。

其中，第一调整模块 13 用于执行根据标签损失值，对采样样本图像进行调整，得到采样样本图像对应的对抗样本图像时，具体用于根据标签损失值，生成针对采样样本图像的初始梯度值，获取初始梯度值的数值符号，根据数值符号生成初始梯度值对应的符号单位数值，对初始学习率以及符号单位数值进行乘积处理，得到第一待裁剪数值，通过由攻击强度所生成的第一裁剪区间，对第一待裁剪数值进行裁剪处理，得到用于调整采样样本图像的梯度值，对采样样本图像以及梯度值进行求差处理，得到第二待裁剪数值，通过由像素等级所生成的第二裁剪区间，对第二待裁剪数值进行裁剪处理，得到采样样本图像对应的对抗样本图像。

再请参见图 9，第一调整模块 13 可以包括：第一调整单元 131、第一生成单元 132 以及第一获取单元 133。

第一调整单元 131，用于根据第一概率元素，对采样样本图像进行调整，得到采样样本图像对应的初始对抗样本图像；

第一生成单元 132，用于将所述错误类别标签和所述初始对抗样本图像输入至所述图像识别模型，通过图像识别模型，生成初始对抗样本图像针对 A 个类别标签的第二概率向量；

第一获取单元 133，用于在所述第二概率向量所包含的 A 个概率元素中，获取针对所述错误类别标签的第二概率元素；

第一调整单元 131，还用于若所述第二概率元素不满足边界约束条件，或针对所述对抗样本图像进行类别预测所得到的类别预测结果不为所述错误类别标签，则根据所述第二概率元素继续对所述初始对抗样本图像进行调整；

第一获取单元 133，还用于若所述第二概率元素满足边界约束条件，且针对所述对抗样本图像进行类别预测所得到的类别预测结果为所述错误类别标签，则将所述初始对抗样本图像确定为所述采样样本图像对应的对抗样本图像；在所述类别预测结果为所述错误类别标签的情况下，所述第二概率元素为所述第二概率向量所包含的 A 个概率元素中的最大数值。

其中，第一调整单元 131、第一生成单元 132 以及第一获取单元 133 的具体功能实现方式，可以参见上述图 3 对应实施例中的步骤 S103，这里不再进行赘述。

其中，第一获取单元 133，还用于将所述第二概率元素以及概率约束值进行对比；所述概率约束值为 A 的倒数；

第一获取单元 133，还用于若所述第二概率元素小于所述概率约束值，则确定所述第二概率元素不满足边界约束条件；

第一获取单元 133，还用于若所述第二概率元素等于或大于所述概率约束值，则确定所述第二概率元素满足边界约束条件。

再请参见图 9，数据处理装置 1 还可以包括：第二输入模块 14 以及第二调整模块 15。

第二输入模块 14，用于将所述 A 个样本图像集、所述 A 个类别标签、正确类别标签以及所述对抗样本图像均输入至所述图像识别模型；所述正确类别标签是指所述采样样本图像所属的样本图像集对应的类别标签；

第二输入模块 14，还用于通过图像识别模型，生成 A 个样本图像集分别对应的第一预测类别，以及对抗样本图像对应的第二预测类别；

第二调整模块 15，用于根据所述第一预测类别和所述 A 个类别标签生成第一损失值，根据所述第二预测类别以及所述正确类别标签生成第二损失值，根据所述第一损失值和所述第二损失值对所述图像识别模型中的参数进行调整，得到所述图像识别优化模型；所述图像识别优化模型针对所述对抗样本图像进行类别预测所得到的类别预测结果为所述正确类别标签。

其中，第二输入模块 14 以及第二调整模块 15 的具体功能实现方式，可以参见上述图 7 对应实施例中的步骤 S204-步骤 S206，这里不再进行赘述。

再请参见图 9，第一输入模块 12 可以包括：第一确定单元 121 以及第二获取单元 122。

其中，所述 A 个样本图像集包括样本图像集 Cb；b 为正整数且 b 小于或等于 A；

第一确定单元 121，用于确定采样比例，根据采样比例，对样本图像集 Cb 中的样本图像进行采样处理，得到采样样本图像；

第二获取单元 122，用于在 A 个类别标签中，获取与所述样本图像集 Cb 的类别标签不同的类别标签，将获取到的类别标签确定为错误类别标签。

其中，第一确定单元 121 以及第二获取单元 122 的具体功能实现方式，可以参见上述图 3 对应实施例中的步骤 S102，这里不再进行赘述。

再请参见图 9，数据获取模块 11 可以包括：第三获取单元 111、第二生成单元 112、第二确定单元 113、第三确定单元 114 以及第二调整单元 115。

第三获取单元 111，用于获取 A 个样本图像集，将 A 个类别标签和所述 A 个样本图像集分别输入至所述图像识别初始模型；

第二生成单元 112，用于通过图像识别初始模型，生成所述 A 个样本图像集分别对应的预测初始类别；

第二确定单元 113，用于根据所述预测初始类别以及所述 A 个类别标签，确定所述 A 个样本图像集分别对应的类别损失值；

第三确定单元 114，用于根据 A 个样本图像集分别对应的类别损失值，确定图像识别初始模型对应的

总损失值；

第二调整单元 115，用于根据所述总损失值，对所述图像识别初始模型中的参数进行调整，得到图像识别模型；所述图像识别模型针对样本图像集 Cb 进行类别预测所得到的类别预测结果为所述样本图像集 Cb 的类别标签；所述 A 个样本图像集包括样本图像集 Cb；b 为正整数且 b 小于或等于 A。

其中，第三获取单元 111、第二生成单元 112、第二确定单元 113、第三确定单元 114 以及第二调整单元 115 的具体功能实现方式，可以参见上述图 3 对应实施例中的步骤 S101，这里不再进行赘述。

再请参见图 9，第三获取单元 111 可以包括：图像获取子单元 1111、第二确定子单元 1112、第三确定子单元 1113 以及第四确定子单元 1114。

图像获取子单元 1111，用于获取 A 个原始图像集，将 A 个原始图像集均输入至对象检测模型；A 个原始图像集分别对应的类别信息互不相同；A 个原始图像集包括原始图像集 Db；原始图像集 Db 包括原始图像 Ef；f 为正整数，且 f 小于或等于原始图像集 Db 中原始图像的总数量；

第二确定子单元 1112，用于通过对象检测模型，确定原始图像 Ef 中关键区域的区域坐标，根据区域坐标，生成原始图像 Ef 对应的待标注样本图像；

第三确定子单元 1113，用于将原始图像集 Db 中每个原始图像对应的待标注样本图像，确定为待标注样本图像集；

第四确定子单元 1114，用于根据原始图像集 Db 对应的类别信息，生成类别标签 Hb，将标注有类别标签 Hb 的待标注样本图像集，确定为样本图像集 Cb。

其中，图像获取子单元 1111、第二确定子单元 1112、第三确定子单元 1113 以及第四确定子单元 1114 的具体功能实现方式，可以参见上述图 7 对应实施例中的步骤 S201，这里不再进行赘述。

再请参见图 9，第二确定子单元 1112 可以包括：第三处理子单元 11121、第四处理子单元 11122 以及第五处理子单元 11123。

第三处理子单元 11121，用于根据区域坐标，生成包括关键区域的初始检测框，对初始检测框进行扩张处理，得到包括目标对象的检测框；关键区域属于目标对象；

第四处理子单元 11122，用于在检测框中，获取包括目标对象的待缩放图像，获取第一图像尺寸，根据第一图像尺寸，对待缩放图像进行缩放处理，得到待裁剪图像；

第五处理子单元 11123，用于获取第二图像尺寸，根据第二图像尺寸，对待裁剪图像进行裁剪处理，得到原始图像 Ef 对应的待标注样本图像；第二图像尺寸小于第一图像尺寸。

其中，第三处理子单元 11121、第四处理子单元 11122 以及第五处理子单元 11123 的具体功能实现方式，可以参见上述图 7 对应实施例中的步骤 S201，这里不再进行赘述。

上述可知，本申请实施例提出一种对抗样本图像的生成方法，通过该方法，可以生成图像识别模型错误识别的对抗样本图像，通过对抗样本图像，可以提高用于优化训练图像识别模型的训练集的样本多样性，通过具有样本多样性的训练集，可以提高图像识别模型的优化训练准确度，进而可以提高优化训练好的图像识别模型的类别标签识别准确度以及识别泛化性。而且由于所述图像识别模型针对所述对抗样本图像所预测的针对所述错误类别标签的第二概率元素满足边界约束条件，所以就可以使得优化训练后的图像识别模型的分类边界更加柔和，进一步保证了模型精度。

进一步地，请参见图 10 ，图 10 是本申请实施例提供的一种数据处理装置的结构示意图二。上述数据处理装置 2 可以用于执行本申请实施例提供的方法中的相应步骤。如图 10 所示，该数据处理装置 2 可以包括：图像获取模块 21、向量生成模块 22 以及类别确定模块 23。

图像获取模块 21，用于获取目标图像，将所述目标图像输入至图像识别优化模型；所述图像识别优化模型是根据混合样本训练集继续对图像识别模型进行优化训练所得到的；所述混合样本训练集包括 A 个样本图像集、A 个类别标签、正确类别标签以及对抗样本图像；所述 A 个样本图像集分别对应 A 个不同的类别标签中的一个，所述 A 个样本图像集包括采集样本图像；所述正确类别标签属于所述 A 个类别标签；A 为大于 1 的正整数；所述图像识别模型是根据所述 A 个样本图像集预训练得到的；所述对抗样本图像是通过第一概率向量中的 A 个概率元素中指代错误类别标签的第一概率元素对采样样本图像进行调整得到，所述第一概率向量是由所述图像识别模型针对所述采样样本图像所生成的，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同，所述正确类别标签为所述采样样本图像所属的样本图像集对应的类别标签；所述第一概率向量包括 A 个概率元素，所述 A 个概率元素分别指代所述 A 个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率；所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签；

向量生成模块22，用于通过图像识别优化模型，生成目标图像针对A个类别标签的目标概率向量；

类别确定模块23，用于将目标概率向量中的最大概率元素对应的类别标签，确定为目标图像的类别预测结果。

其中，图像获取模块21、向量生成模块22以及类别确定模块23的具体功能实现方式，可以参见上述图8对应实施例中的步骤S301-步骤S303，这里不再进行赘述。

上述可知，本申请实施例提出一种对抗样本图像的生成方法，通过该方法，可以生成图像识别模型错误识别的对抗样本图像，通过对抗样本图像，可以提高用于优化训练图像识别模型的训练集的样本多样性，通过具有样本多样性的训练集，可以提高图像识别模型的优化训练准确度，进而可以提高优化训练好的图像识别模型的类别标签识别准确度以及识别泛化性。

进一步地，请参见图11，图11是本申请实施例提供的一种计算机设备的结构示意图。如图11所示，该计算机设备1000可以包括：至少一个处理器1001，例如CPU，至少一个网络接口1004，用户接口1003，存储器1005，至少一个通信总线1002。其中，通信总线1002用于实现这些组件之间的连接通信。其中，在一些实施例中，用户接口1003可以包括显示屏（Display）、键盘（Keyboard），网络接口1004可选地可以包括标准的有线接口、无线接口（如WI-FI接口）。存储器1005可以是高速RAM存储器，也可以是非不稳定的存储器（non-volatile memory），例如至少一个磁盘存储器。存储器1005可选地还可以是至少一个位于远离前述处理器1001的存储装置。如图11所示，作为一种计算机存储介质的存储器1005可以包括操作系统、网络通信模块、用户接口模块以及设备控制应用程序。

在图11所示的计算机设备1000中，网络接口1004可提供网络通讯功能；而用户接口1003主要用于为用户提供输入的接口；而处理器1001可以用于调用存储器1005中存储的设备控制应用程序，以实现：

获取A个样本图像集以及图像识别模型，A为大于1的正整数，所述A个样本图像集分别对应A个不同的类别标签中的一个，所述图像识别模型是根据所述A个样本图像集预训练得到的，所述A个类别标签包括错误类别标签；

在所述A个样本图像集中获取采样样本图像，将所述采样样本图像与所述错误类别标签输入至所述图像识别模型，以生成所述采样样本图像针对所述A个类别标签的第一概率向量，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同；所述第一概率向量包括A个概率元素，所述A个概率元素分别指代所述A个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率；

在所述第一概率向量中的所述A个概率元素中，获取指代所述错误类别标签的第一概率元素，根据所述第一概率元素，对所述采样样本图像进行调整，得到所述采样样本图像对应的对抗样本图像，所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签。

或者，处理器1001可以用于调用存储器1005中存储的设备控制应用程序，以实现：

获取目标图像，将所述目标图像输入至图像识别优化模型；所述图像识别优化模型是根据混合样本训练集继续对图像识别模型进行优化训练所得到的；所述混合样本训练集包括A个样本图像集、A个类别标签、正确类别标签以及对抗样本图像；所述A个样本图像集分别对应A个不同的类别标签中的一个，所述A个样本图像集包括采集样本图像；所述正确类别标签属于所述A个类别标签；A为大于1的正整数；所述图像识别模型是根据所述A个样本图像集预训练得到的；所述对抗样本图像是通过第一概率向量中的A个概率元素中指代错误类别标签的第一概率元素对采样样本图像进行调整得到，所述第一概率向量是由所述图像识别模型针对所述采样样本图像所生成的，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同，所述正确类别标签为所述采样样本图像所属的样本图像集对应的类别标签；所述第一概率向量包括A个概率元素，所述A个概率元素分别指代所述A个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率；所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签；

通过所述图像识别优化模型，生成所述目标图像针对所述A个类别标签的目标概率向量；

将所述目标概率向量中的最大概率元素对应的类别标签，确定为所述目标图像的类别预测结果。

应当理解，本申请实施例中所描述的计算机设备1000可执行前文各实施例中对数据处理方法或装置的描述，在此不再赘述。另外，对采用相同方法的有益效果描述，也不再进行赘述。

本申请实施例还提供一种计算机可读存储介质，该计算机可读存储介质存储有计算机程序，该计算机程序被处理器执行时实现前文各实施例中对数据处理方法或装置的描述，在此不再赘述。另外，对采用相同方法的有益效果描述，也不再进行赘述。

上述计算机可读存储介质可以是前述任一实施例提供的数据处理装置或者上述计算机设备的内部存储单元，例如计算机设备的硬盘或内存。该计算机可读存储介质也可以是该计算机设备的外部存储设备，例如该计算机设备上配备的插接式硬盘，智能存储卡（smart media card，SMC），安全数字（secure digital，SD）卡，闪存卡（flash card）等。进一步地，该计算机可读存储介质还可以既包括该计算机设备的内部存储单元也包括外部存储设备。该计算机可读存储介质用于存储该计算机程序以及该计算机设备所需的其他程序和数据。该计算机可读存储介质还可以用于暂时地存储已经输出或者将要输出的数据。

本申请实施例还提供了一种计算机程序产品，该计算机程序产品包括计算机程序，该计算机程序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程序，处理器执行该计算机程序，使得该计算机设备可执行前文各实施例中对数据处理方法或装置的描述，在此不再赘述。另外，对采用相同方法的有益效果描述，也不再进行赘述。

本申请实施例的说明书和权利要求书及附图中的术语“第一”、“第二”等是用于区别不同对象，而非用于描述特定顺序。此外，术语“包括”以及它们任何变形，意图在于覆盖不排他的包含。例如包含了一系列步骤或单元的过程、方法、装置、产品或设备没有限定于已列出的步骤或模块，而是可选地还包括没有列出的步骤或模块，或可选地还包括对于这些过程、方法、装置、产品或设备固有的其他步骤单元。

本领域普通技术人员可以意识到，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，能够以电子硬件、计算机软件或者二者的结合来实现，为了清楚地说明硬件和软件的可互换性，在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行，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申请的范围。

以上所揭露的仅为本申请较佳实施例而已，当然不能以此来限定本申请之权利范围，因此依本申请权利要求所作的等同变化，仍属本申请所涵盖的范围。

权利要求书

1、一种图像处理方法，其特征在于，包括：

获取A个样本图像集以及图像识别模型，A为大于1的正整数，所述A个样本图像集分别对应A个不同的类别标签中的一个，所述图像识别模型是根据所述A个样本图像集预训练得到的，所述A个类别标签包括错误类别标签；

在所述A个样本图像集中获取采样样本图像，将所述采样样本图像与所述错误类别标签输入至所述图像识别模型，以生成所述采样样本图像针对所述A个类别标签的第一概率向量，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同；所述第一概率向量包括A个概率元素，所述A个概率元素分别指代所述A个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率；

在所述第一概率向量中的所述A个概率元素中，获取指代所述错误类别标签的第一概率元素，根据所述第一概率元素，对所述采样样本图像进行调整，得到所述采样样本图像对应的对抗样本图像，所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签。

2、根据权利要求1所述的方法，其特征在于，所述根据所述第一概率元素，对所述采样样本图像进行调整，得到所述采样样本图像对应的对抗样本图像，包括：

生成所述第一概率元素对应的负概率元素，对所述负概率元素以及所述第一概率向量中的最大概率元素进行求和处理，得到所述采样样本图像针对所述错误类别标签的标签损失值；

根据所述标签损失值，对所述采样样本图像进行调整，得到所述采样样本图像对应的对抗样本图像。

3、根据权利要求1或2所述的方法，其特征在于，所述根据所述标签损失值，对所述采样样本图像进行调整，得到所述采样样本图像对应的对抗样本图像，包括：

根据所述标签损失值，生成针对所述采样样本图像的初始梯度值，获取所述初始梯度值的数值符号，根据所述数值符号生成所述初始梯度值对应的符号单位数值；

对初始学习率以及符号单位数值进行乘积处理，得到第一待裁剪数值，通过由攻击强度所生成的第一裁剪区间，对所述第一待裁剪数值进行裁剪处理，得到用于调整所述采样样本图像的梯度值；

对所述采样样本图像以及所述梯度值进行求差处理，得到第二待裁剪数值，通过由像素等级所生成的第二裁剪区间，对所述第二待裁剪数值进行裁剪处理，得到所述采样样本图像对应的对抗样本图像。

4、根据权利要求1至3任一项所述的方法，其特征在于，所述根据所述第一概率元素，对所述采样样本图像进行调整，得到所述采样样本图像对应的对抗样本图像，包括：

根据所述第一概率元素，对所述采样样本图像进行调整，得到所述采样样本图像对应的初始对抗样本图像；

将所述错误类别标签和所述初始对抗样本图像输入至所述图像识别模型，通过所述图像识别模型，生成所述初始对抗样本图像针对所述A个类别标签的第二概率向量；

在所述第二概率向量所包含的A个概率元素中，获取针对所述错误类别标签的第二概率元素；

若所述第二概率元素不满足边界约束条件，或针对所述对抗样本图像进行类别预测所得到的类别预测结果不为所述错误类别标签，则根据所述第二概率元素继续对所述初始对抗样本图像进行调整；

若所述第二概率元素满足边界约束条件，且针对所述对抗样本图像进行类别预测所得到的类别预测结果为所述错误类别标签，则将所述初始对抗样本图像确定为所述采样样本图像对应的对抗样本图像；在所述类别预测结果为所述错误类别标签的情况下，所述第二概率元素为所述第二概率向量所包含的A个概率元素中的最大数值。

5、根据权利要求1至4任一项所述的方法，其特征在于，还包括：

将所述第二概率元素以及概率约束值进行对比；所述概率约束值为A的倒数；

若所述第二概率元素小于所述概率约束值，则确定所述第二概率元素不满足边界约束条件；

若所述第二概率元素等于或大于所述概率约束值，则确定所述第二概率元素满足边界约束条件。

6、根据权利要求1至5任一项所述的方法，其特征在于，所述方法还包括：

将所述A个样本图像集、所述A个类别标签、正确类别标签以及所述对抗样本图像均输入至所述图像识别模型；所述正确类别标签是指所述采样样本图像所属的样本图像集对应的类别标签；

通过所述图像识别模型，生成所述A个样本图像集分别对应的第一预测类别，以及所述对抗样本图像对应的第二预测类别；

根据所述第一预测类别和所述A个类别标签生成第一损失值，根据所述第二预测类别以及所述正确类

别标签生成第二损失值，根据所述第一损失值和所述第二损失值对所述图像识别模型中的参数进行调整，得到所述图像识别优化模型；所述图像识别优化模型针对所述对抗样本图像进行类别预测所得到的类别预测结果为所述正确类别标签。

7、根据权利要求 1 至 6 任一项所述的方法，其特征在于，所述 A 个样本图像集包括样本图像集 C_b；b 为正整数且 b 小于或等于 A；所述在所述 A 个样本图像集中获取采样样本图像，包括：

确定采样比例，根据所述采样比例，对所述样本图像集 C_b 中的样本图像进行采样处理，得到采样样本图像；

在所述 A 个类别标签中，获取与所述样本图像集 C_b 的类别标签不同的类别标签，将获取到的类别标签确定为所述错误类别标签。

8、根据权利要求 1 至 7 任一项所述的方法，其特征在于，所述获取 A 个样本图像集以及图像识别模型，包括：

获取 A 个样本图像集，将 A 个类别标签和所述 A 个样本图像集分别输入至所述图像识别初始模型；

通过所述图像识别初始模型，生成所述 A 个样本图像集分别对应的预测初始类别；

根据所述预测初始类别以及所述 A 个类别标签，确定所述 A 个样本图像集分别对应的类别损失值；

根据所述 A 个样本图像集分别对应的类别损失值，确定所述图像识别初始模型对应的总损失值；

根据所述总损失值，对所述图像识别初始模型中的参数进行调整，得到图像识别模型；所述图像识别模型针对样本图像集 C_b 进行类别预测所得到的类别预测结果为所述样本图像集 C_b 的类别标签；所述 A 个样本图像集包括样本图像集 C_b；b 为正整数且 b 小于或等于 A。

9、根据权利要求 8 所述的方法，其特征在于，所述获取 A 个样本图像集，包括：

获取 A 个原始图像集，将所述 A 个原始图像集均输入至对象检测模型；所述 A 个原始图像集分别对应 A 个不同的类别信息中的一个；所述 A 个原始图像集包括原始图像集 D_b；所述原始图像集 D_b 包括原始图像 E_f；f 为正整数，且 f 小于或等于所述原始图像集 D_b 中原始图像的总数量；

通过所述对象检测模型，确定所述原始图像 E_f 中关键区域的区域坐标，根据所述区域坐标，生成所述原始图像 E_f 对应的待标注样本图像；

将所述原始图像集 D_b 中每个原始图像对应的待标注样本图像，确定为待标注样本图像集；

根据所述原始图像集 D_b 对应的类别信息，生成类别标签 H_b，将标注有所述类别标签 H_b 的待标注样本图像集，确定为所述样本图像集 C_b。

10、根据权利要求 9 所述的方法，其特征在于，所述根据所述区域坐标，生成所述原始图像 E_f 对应的待标注样本图像，包括：

根据所述区域坐标，生成包括所述关键区域的初始检测框，对所述初始检测框进行扩张处理，得到包括目标对象的检测框；所述关键区域属于所述目标对象；

在所述检测框中，获取包括所述目标对象的待缩放图像，获取第一图像尺寸，根据所述第一图像尺寸，对所述待缩放图像进行缩放处理，得到待裁剪图像；

获取第二图像尺寸，根据所述第二图像尺寸，对所述待裁剪图像进行裁剪处理，得到所述原始图像 E_f 对应的待标注样本图像；所述第二图像尺寸小于所述第一图像尺寸。

11、一种数据处理方法，其特征在于，包括：

获取目标图像，将所述目标图像输入至图像识别优化模型；所述图像识别优化模型是根据混合样本训练集继续对图像识别模型进行优化训练所得到的；所述混合样本训练集包括 A 个样本图像集、A 个类别标签、正确类别标签以及对抗样本图像；所述 A 个样本图像集分别对应 A 个不同的类别标签中的一个，所述 A 个样本图像集包括采集样本图像；所述正确类别标签属于所述 A 个类别标签；A 为大于 1 的正整数；所述图像识别模型是根据所述 A 个样本图像集预训练得到的；所述对抗样本图像是通过第一概率向量中的 A 个概率元素中指代错误类别标签的第一概率元素对采样样本图像进行调整得到，所述第一概率向量是由所述图像识别模型针对所述采样样本图像所生成的，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同，所述正确类别标签为所述采样样本图像所属的样本图像集对应的类别标签；所述第一概率向量包括 A 个概率元素，所述 A 个概率元素分别指代所述 A 个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率；所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签；

通过所述图像识别优化模型，生成所述目标图像针对所述 A 个类别标签的目标概率向量；

将所述目标概率向量中的最大概率元素对应的类别标签，确定为所述目标图像的类别预测结果。

12、一种数据处理装置，其特征在于，包括：

数据获取模块，用于获取 A 个样本图像集以及图像识别模型，A 为大于 1 的正整数，所述 A 个样本图像集分别对应 A 个不同的类别标签中的一个，所述图像识别模型是根据所述 A 个样本图像集预训练得到的，所述 A 个类别标签包括错误类别标签；

第一输入模块，用于在所述 A 个样本图像集中获取采样样本图像，将所述采样样本图像与所述错误类别标签输入至所述图像识别模型，以生成所述采样样本图像针对所述 A 个类别标签的第一概率向量，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同；所述第一概率向量包括 A 个概率元素，所述 A 个概率元素分别指代所述 A 个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率；

第一调整模块，用于在所述第一概率向量中的所述 A 个概率元素中，获取指代所述错误类别标签的第一概率元素，根据所述第一概率元素，对所述采样样本图像进行调整，得到所述采样样本图像对应的对抗样本图像，所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签。

13、一种数据处理装置，其特征在于，包括：

图像获取模块，用于获取目标图像，将所述目标图像输入至图像识别优化模型；所述图像识别优化模型是根据混合样本训练集继续对图像识别模型进行优化训练所得到的；所述混合样本训练集包括 A 个样本图像集、A 个类别标签、正确类别标签以及对抗样本图像；所述 A 个样本图像集分别对应 A 个不同的类别标签中的一个，所述 A 个样本图像集包括采集样本图像；所述正确类别标签属于所述 A 个类别标签；A 为大于 1 的正整数；所述图像识别模型是根据所述 A 个样本图像集预训练得到的；所述对抗样本图像是通过第一概率向量中的 A 个概率元素中指代错误类别标签的第一概率元素对采样样本图像进行调整得到，所述第一概率向量是由所述图像识别模型针对所述采样样本图像所生成的，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同，所述正确类别标签为所述采样样本图像所属的样本图像集对应的类别标签；所述第一概率向量包括 A 个概率元素，所述 A 个概率元素分别指代所述 A 个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率；所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签；

向量生成模块，用于通过所述图像识别优化模型，生成所述目标图像针对所述 A 个类别标签的目标概率向量；

类别确定模块，用于将所述目标概率向量中的最大概率元素对应的类别标签，确定为所述目标图像的类别预测结果。

14、一种计算机设备，其特征在于，包括：处理器、存储器以及网络接口；

所述处理器与所述存储器、所述网络接口相连，其中，所述网络接口用于提供数据通信功能，所述存储器用于存储计算机程序，所述处理器用于调用所述计算机程序，以使得所述计算机设备执行权利要求 1 至 11 任一项所述的方法。

15、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质中存储有计算机程序，所述计算机程序适于由处理器加载并执行，以使得具有所述处理器的计算机设备执行权利要求 1-11 任一项所述的方法。

16、一种计算机程序产品，其特征在于，所述计算机程序产品包括计算机程序，所述计算机程序存储在计算机可读存储介质中，所述计算机程序适于由处理器读取并执行，以使得具有所述处理器的计算机设备执行权利要求 1-11 任一项所述的方法。

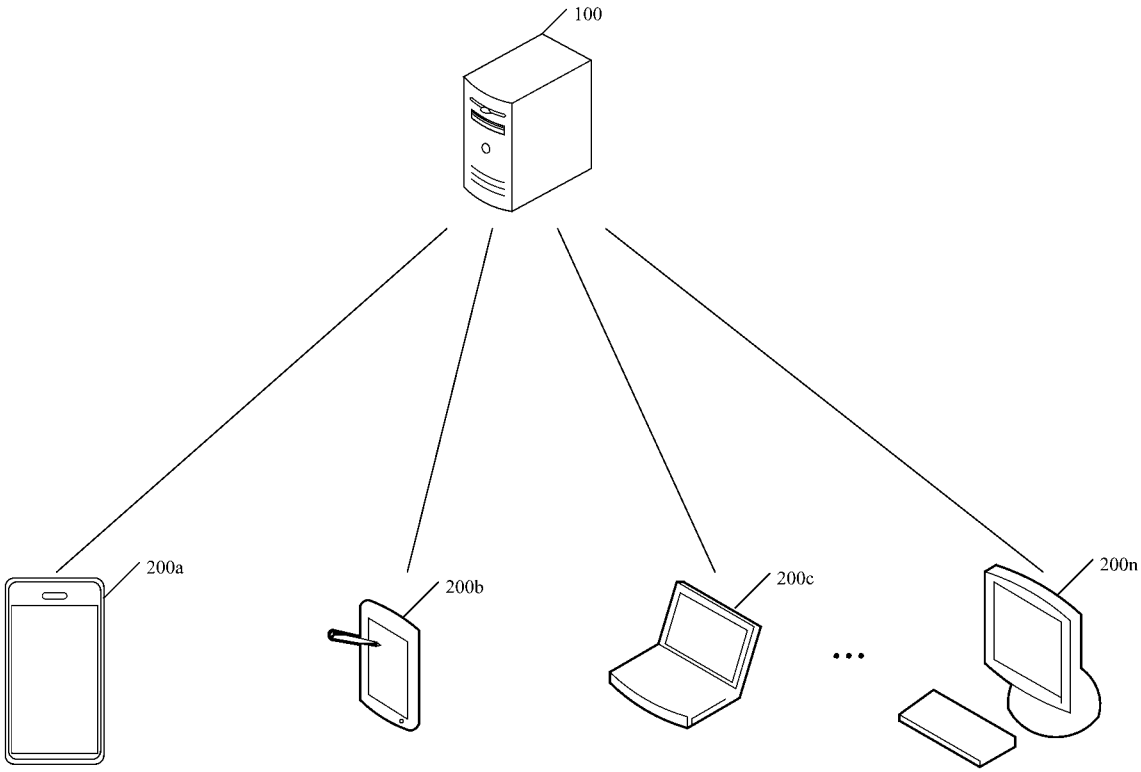


图 1

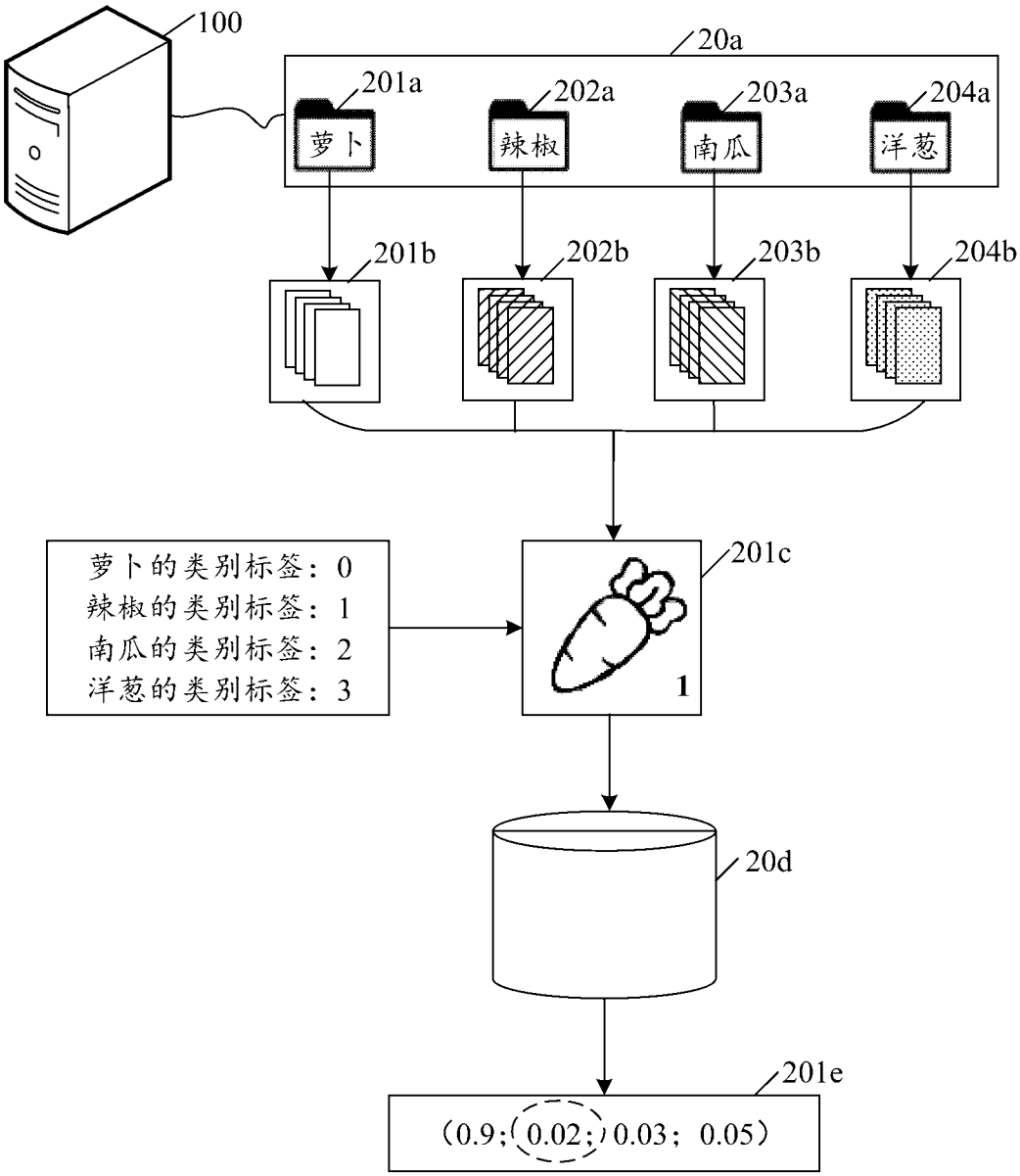

100
20a
201a
萝卜
202a
辣椒
203a
南瓜
204a
洋葱
201b
202b
203b
204b
萝卜的类别标签：0
辣椒的类别标签：1
南瓜的类别标签：2
洋葱的类别标签：3
201c
1
20d
201e
（0.9；0.02；0.03；0.05）

图 2a

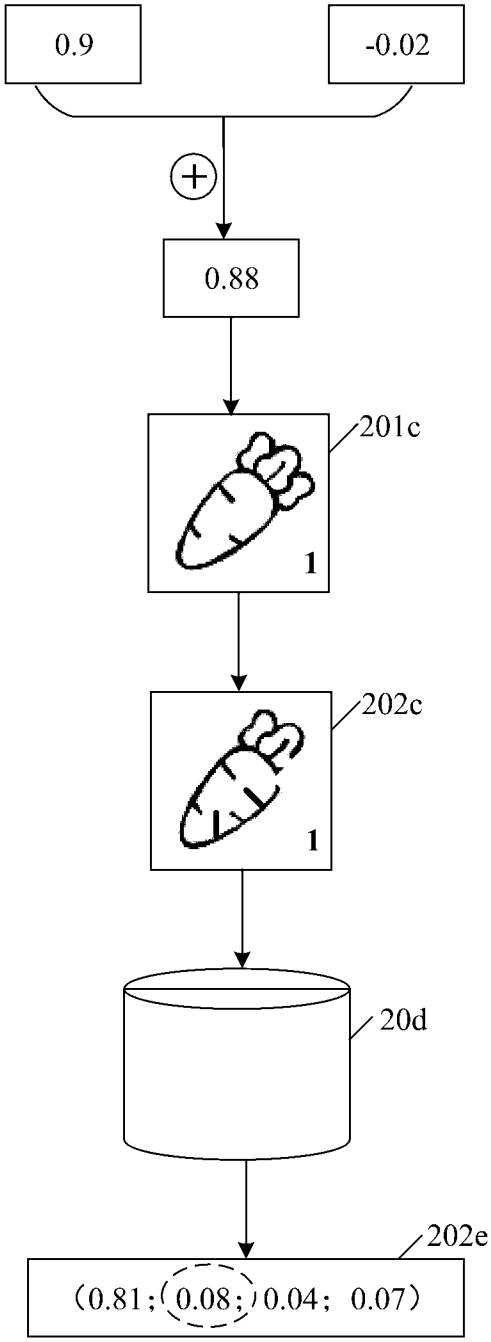


图 2b

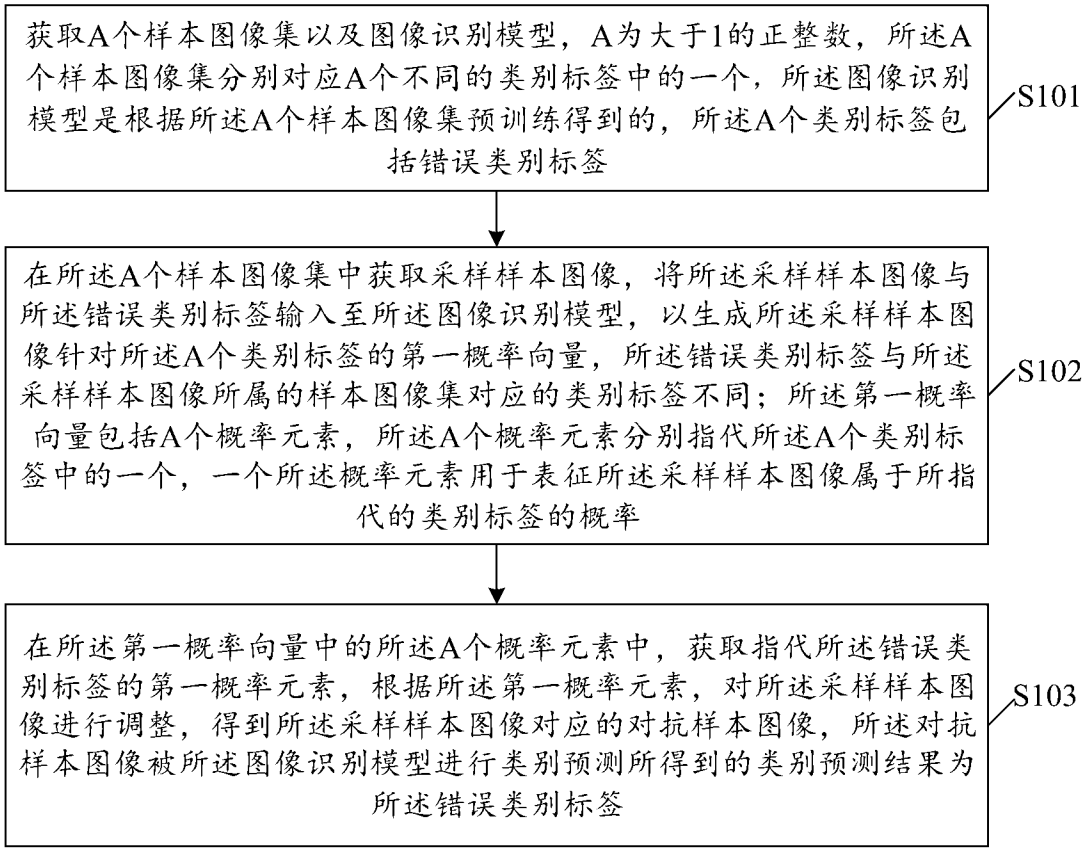

获取A个样本图像集以及图像识别模型，A为大于1的正整数，所述A个样本图像集分别对应A个不同的类别标签中的一个，所述图像识别模型是根据所述A个样本图像集预训练得到的，所述A个类别标签包括错误类别标签
S101
在所述A个样本图像集中获取采样样本图像，将所述采样样本图像与所述错误类别标签输入至所述图像识别模型，以生成所述采样样本图像针对所述A个类别标签的第一概率向量，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同；所述第一概率向量包括A个概率元素，所述A个概率元素分别指代所述A个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率
S102
在所述第一概率向量中的所述A个概率元素中，获取指代所述错误类别标签的第一概率元素，根据所述第一概率元素，对所述采样样本图像进行调整，得到所述采样样本图像对应的对抗样本图像，所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签
S103

图 3

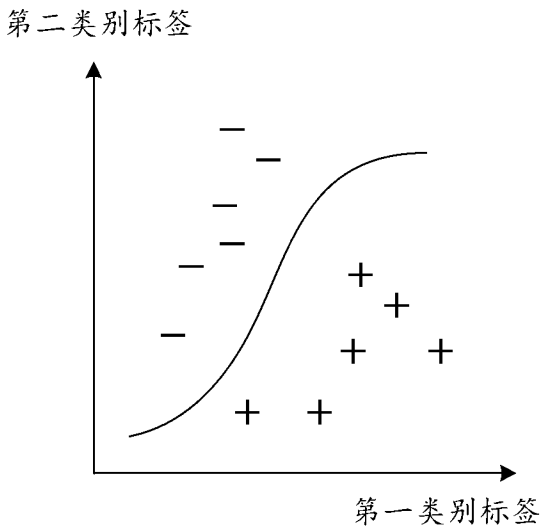


图 4

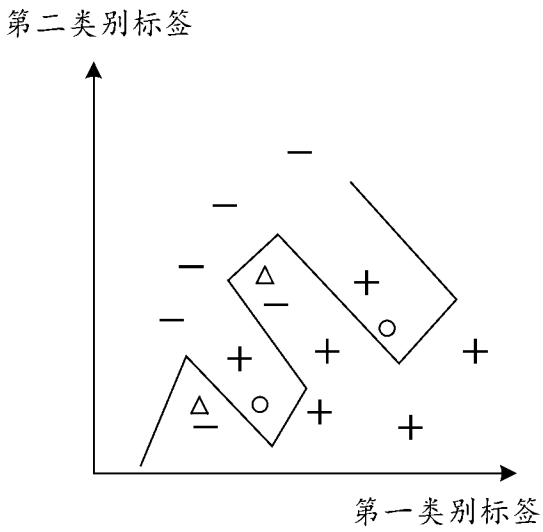


图 5

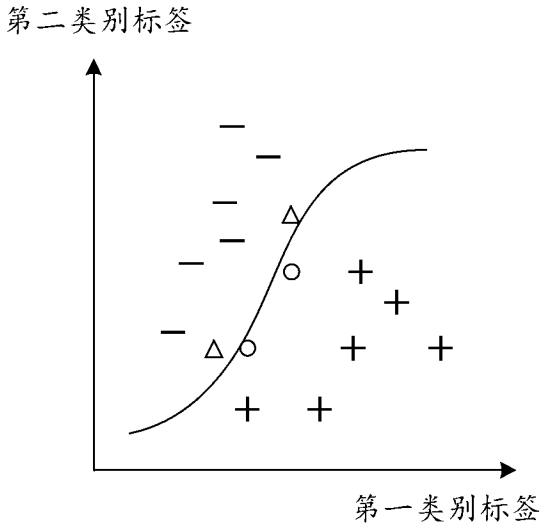


图 6

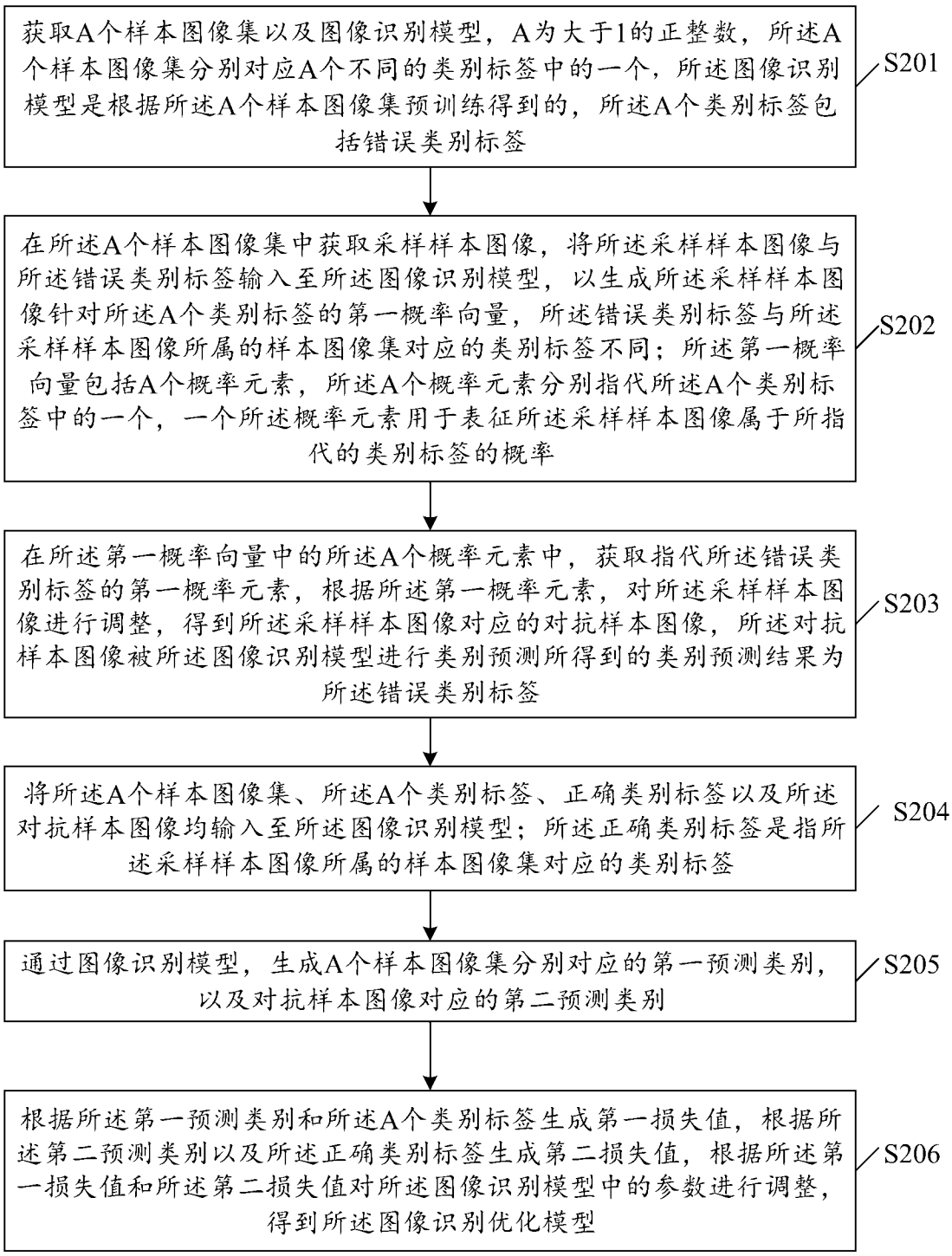

获取A个样本图像集以及图像识别模型，A为大于1的正整数，所述A个样本图像集分别对应A个不同的类别标签中的一个，所述图像识别模型是根据所述A个样本图像集预训练得到的，所述A个类别标签包括错误类别标签
S201
在所述A个样本图像集中获取采样样本图像，将所述采样样本图像与所述错误类别标签输入至所述图像识别模型，以生成所述采样样本图像针对所述A个类别标签的第一概率向量，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同；所述第一概率向量包括A个概率元素，所述A个概率元素分别指代所述A个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率
S202
在所述第一概率向量中的所述A个概率元素中，获取指代所述错误类别标签的第一概率元素，根据所述第一概率元素，对所述采样样本图像进行调整，得到所述采样样本图像对应的对抗样本图像，所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签
S203
将所述A个样本图像集、所述A个类别标签、正确类别标签以及所述对抗样本图像均输入至所述图像识别模型；所述正确类别标签是指所述采样样本图像所属的样本图像集对应的类别标签
S204
通过图像识别模型，生成A个样本图像集分别对应的第一预测类别，以及对抗样本图像对应的第二预测类别
S205
根据所述第一预测类别和所述A个类别标签生成第一损失值，根据所述第二预测类别以及所述正确类别标签生成第二损失值，根据所述第一损失值和所述第二损失值对所述图像识别模型中的参数进行调整，得到所述图像识别优化模型
S206

图 7

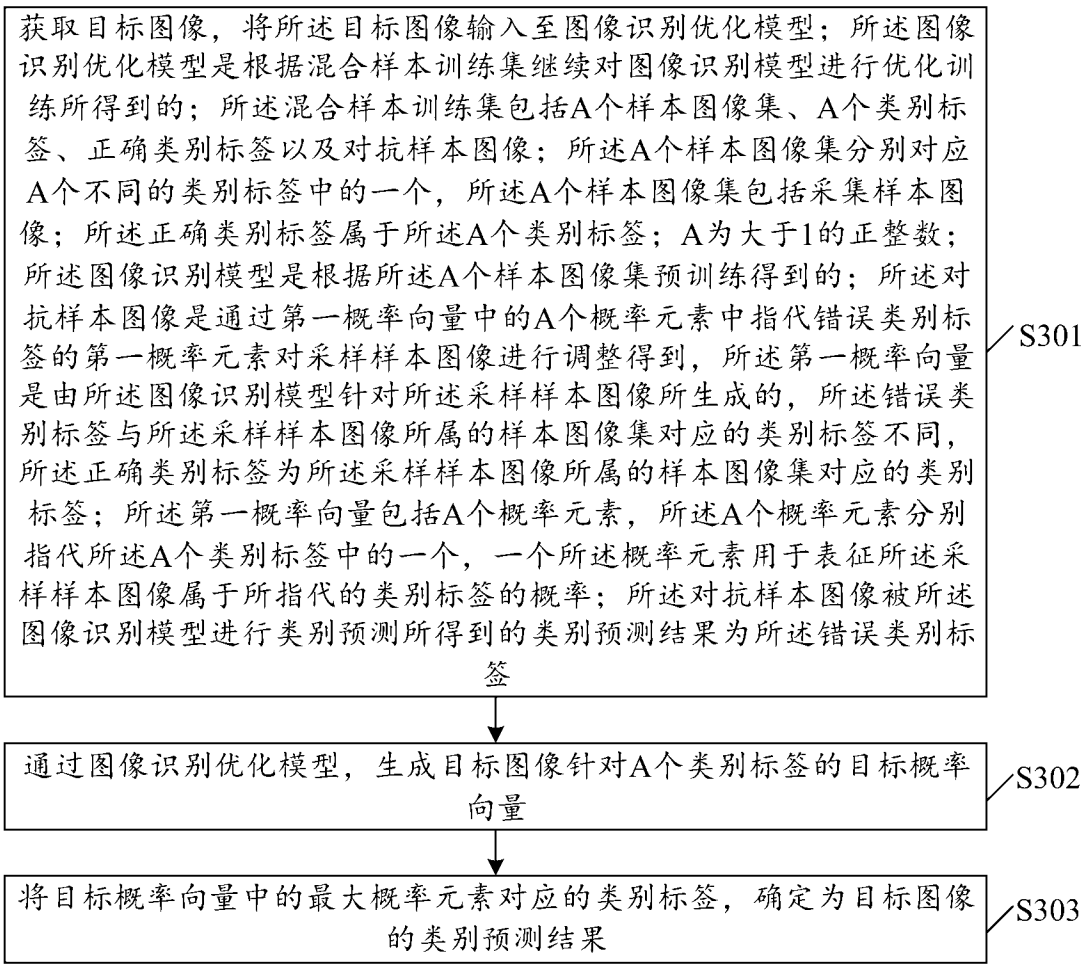

获取目标图像，将所述目标图像输入至图像识别优化模型；所述图像识别优化模型是根据混合样本训练集继续对图像识别模型进行优化训练所得到的；所述混合样本训练集包括A个样本图像集、A个类别标签、正确类别标签以及对抗样本图像；所述A个样本图像集分别对应A个不同的类别标签中的一个，所述A个样本图像集包括采集样本图像；所述正确类别标签属于所述A个类别标签；A为大于1的正整数；所述图像识别模型是根据所述A个样本图像集预训练得到的；所述对抗样本图像是通过第一概率向量中的A个概率元素中指代错误类别标签的第一概率元素对采样样本图像进行调整得到，所述第一概率向量是由所述图像识别模型针对所述采样样本图像所生成的，所述错误类别标签与所述采样样本图像所属的样本图像集对应的类别标签不同，所述正确类别标签为所述采样样本图像所属的样本图像集对应的类别标签；所述第一概率向量包括A个概率元素，所述A个概率元素分别指代所述A个类别标签中的一个，一个所述概率元素用于表征所述采样样本图像属于所指代的类别标签的概率；所述对抗样本图像被所述图像识别模型进行类别预测所得到的类别预测结果为所述错误类别标签
S301
通过图像识别优化模型，生成目标图像针对A个类别标签的目标概率向量
S302
将目标概率向量中的最大概率元素对应的类别标签，确定为目标图像的类别预测结果
S303

图 8

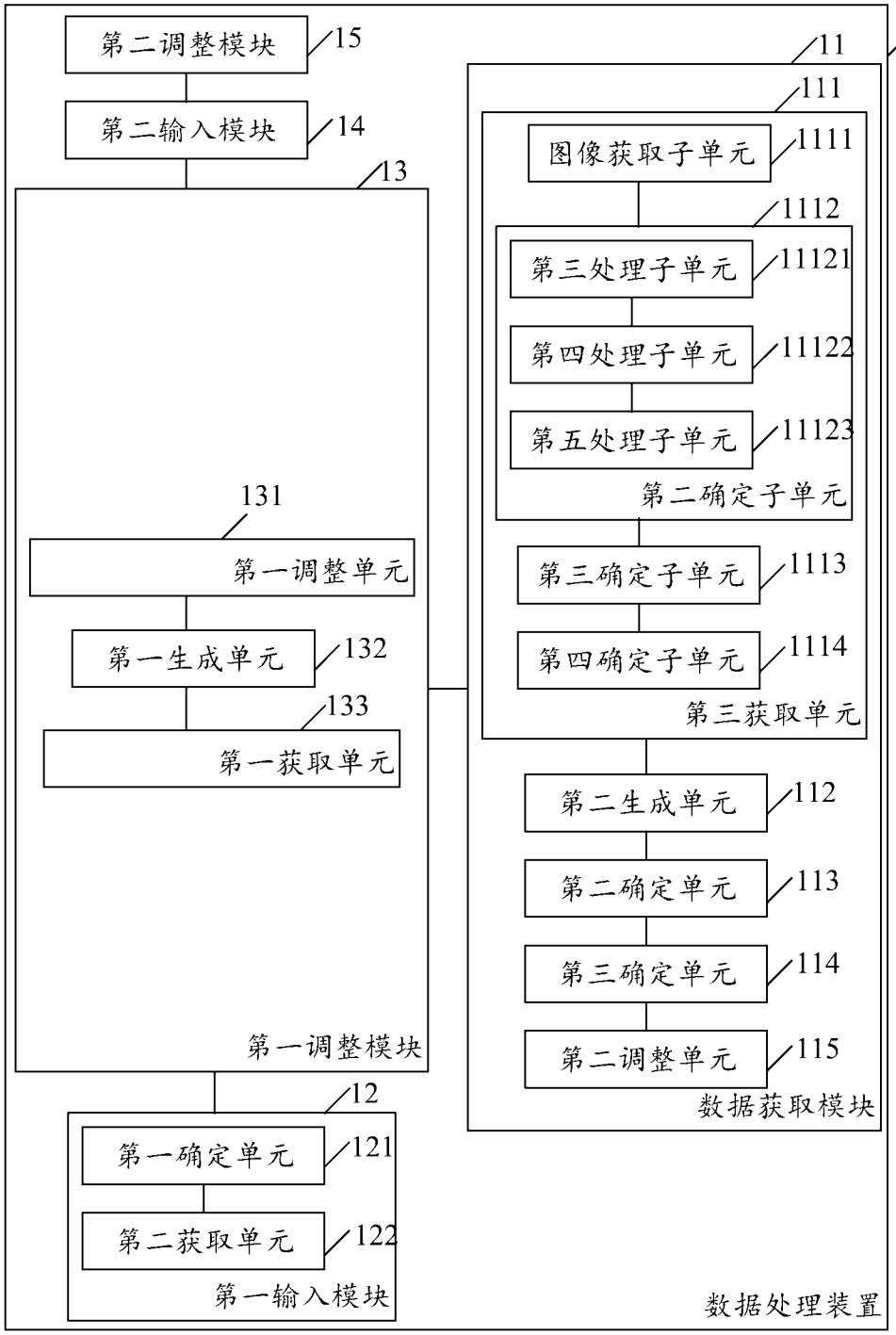

第二调整模块
15
第二输入模块
14
13
131
第一调整单元
第一生成单元
132
133
第一获取单元
第一调整模块
12
第一确定单元
121
第二获取单元
122
第一输入模块
1
11
111
图像获取子单元
1111
1112
第三处理子单元
11121
第四处理子单元
11122
第五处理子单元
11123
第二确定子单元
第三确定子单元
1113
第四确定子单元
1114
第三获取单元
第二生成单元
112
第二确定单元
113
第三确定单元
114
第二调整单元
115
数据获取模块
数据处理装置

图 9

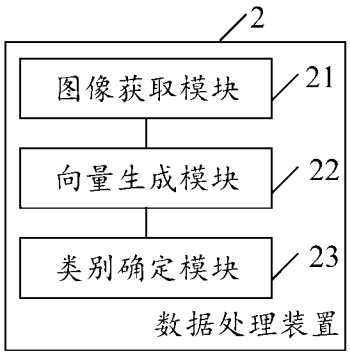

2
图像获取模块
21
向量生成模块
22
类别确定模块
23
数据处理装置

图 10

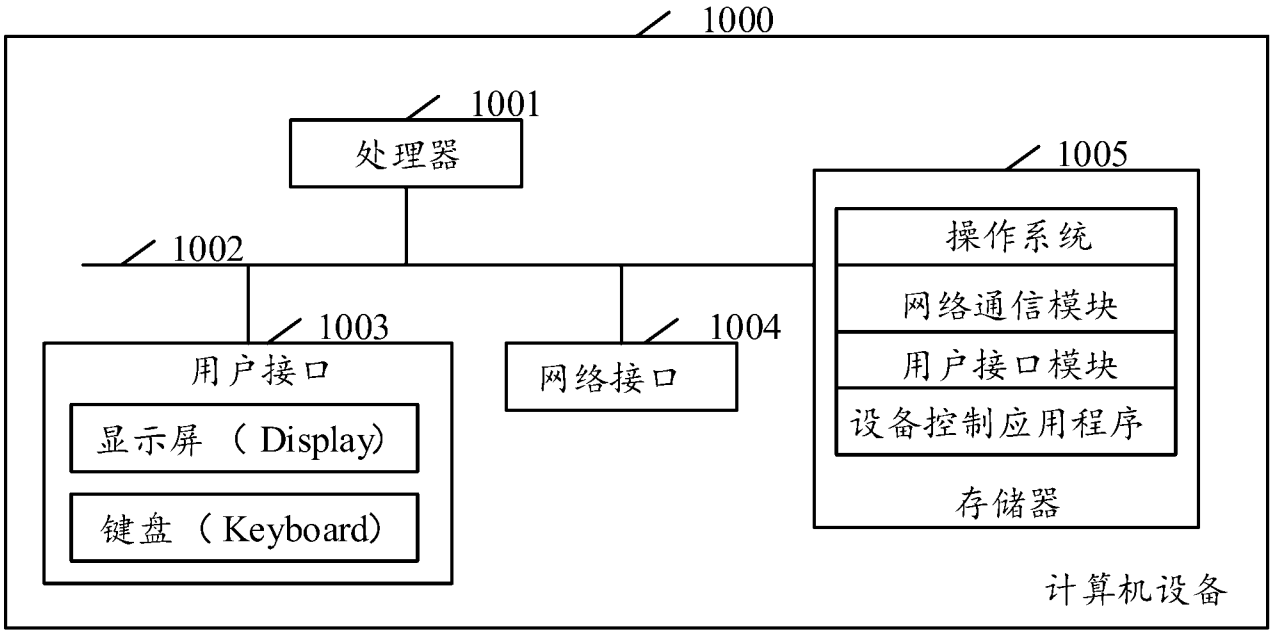

1000
1001
处理器
1005
操作系统
网络通信模块
用户接口模块
设备控制应用程序
存储器
1002
1003
用户接口
显示屏（Display）
键盘（Keyboard）
1004
网络接口
计算机设备

图 11

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2024/079587

A. CLASSIFICATION OF SUBJECT MATTER

G06V10/764(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)

IPC:G06V

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, DWPI, ENTXTC, VEN, WPABSC, CNKI, ACM: 图像, 图片, 分类, 标签, 概率, 矩阵, 对抗, 反, 样本, 向量, picture, image, classification, label, probability, matrix, contrast, anti, sample, vector

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 116977692 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.) 31 October 2023 (2023-10-31) description, paragraphs 2-244	1-16
A	EP 4075395 A2 (BEIJING BAIDU NETCOM SCIENCE & TECHNOLOGY CO., LTD.) 19 October 2022 (2022-10-19) description, pages 1-13	1-16
A	CN 111046380 A (ALIPAY (HANGZHOU) INFORMATION TECHNOLOGY CO., LTD.) 21 April 2020 (2020-04-21) entire document	1-16
A	CN 110210617 A (BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) 06 September 2019 (2019-09-06) entire document	1-16
A	CN 113569611 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.) 29 October 2021 (2021-10-29) entire document	1-16

☑ Further documents are listed in the continuation of Box C. ☑ 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
"D" document cited by the applicant in the international application
"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
11 May 2024	**28 May 2024**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088**	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2024/079587

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2021189364 A1 (SHENZHEN INSTITUTE OF ADVANCED TECHNOLOGY) 30 September 2021 (2021-09-30) entire document	1-16
A	US 2023022943 A1 (XIDIAN UNIVERSITY) 26 January 2023 (2023-01-26) entire document	1-16

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2024/079587

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	116977692	A	31 October 2023	None			
EP	4075395	A2	19 October 2022	EP	4075395	A3	22 February 2023
				KR	20220107120	A	02 August 2022
				JP	2022141931	A	29 September 2022
				JP	7414901	B2	16 January 2024
				CN	113705425	A	26 November 2021
				CN	113705425	B	16 August 2022
CN	111046380	A	21 April 2020	HK	40028423	A0	05 February 2021
				HK	40028423	A1	02 June 2022
				CN	111046380	B	15 February 2022
CN	110210617	A	06 September 2019	CN	110210617	B	18 June 2021
CN	113569611	A	29 October 2021	HK	40054013	A0	18 February 2022
WO	2021189364	A1	30 September 2021	US	2022092336	A1	24 March 2022
				AU	2020437435	A1	06 January 2022
				AU	2020437435	B2	20 July 2023
				GB	2607647	A	14 December 2022
				CN	111475797	A	31 July 2020
				CN	111475797	B	29 September 2023
US	2023022943	A1	26 January 2023	CN	113554089	A	26 October 2021

国际检索报告

国际申请号
PCT/CN2024/079587

A. 主题的分类

G06V10/764(2022.01)i

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

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC:G06V

包含在检索领域中的除最低限度文献以外的检索文献

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））

CNABS,CNTXT,DWPI,ENTXTC,VEN,WPABSC,CNKI,ACM:图像,图片,分类,标签,概率,矩阵,对抗,反,样本，向量, picture,image,classification,label,probability,matrix, contrast,anti,sample,vector

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
PX	CN 116977692 A (腾讯科技（深圳）有限公司) 2023年10月31日 (2023 - 10 - 31) 说明书第2-244段	1-16
A	EP 4075395 A2 (BEIJING BAIDU NETCOM SCI & TECH CO LTD) 2022年10月19日 (2022 - 10 - 19) 说明书第1-13页	1-16
A	CN 111046380 A (支付宝（杭州）信息技术有限公司) 2020年4月21日 (2020 - 04 - 21) 全文	1-16
A	CN 110210617 A (北京邮电大学) 2019年9月6日 (2019 - 09 - 06) 全文	1-16
A	CN 113569611 A (腾讯科技（深圳）有限公司) 2021年10月29日 (2021 - 10 - 29) 全文	1-16
A	WO 2021189364 A1 (SHENZHEN INST ADV TECH) 2021年9月30日 (2021 - 09 - 30) 全文	1-16
A	US 2023022943 A1 (UNIV XIDIAN) 2023年1月26日 (2023 - 01 - 26) 全文	1-16

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

“A” 认为不特别相关的表示了现有技术一般状态的文件

“D” 申请人在国际申请中引证的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“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）

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“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

“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

“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

“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

“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2024年5月11日	2024年5月28日
ISA/CN的名称和邮寄地址 中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088	受权官员 陈娜 电话号码 (+86) 027-59371796

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/079587

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	116977692	A	2023年10月31日	无			
EP	4075395	A2	2022年10月19日	EP	4075395	A3	2023年2月22日
				KR	20220107120	A	2022年8月2日
				JP	2022141931	A	2022年9月29日
				JP	7414901	B2	2024年1月16日
				CN	113705425	A	2021年11月26日
				CN	113705425	B	2022年8月16日
CN	111046380	A	2020年4月21日	HK	40028423	A0	2021年2月5日
				HK	40028423	A1	2022年6月2日
				CN	111046380	B	2022年2月15日
CN	110210617	A	2019年9月6日	CN	110210617	B	2021年6月18日
CN	113569611	A	2021年10月29日	HK	40054013	A0	2022年2月18日
WO	2021189364	A1	2021年9月30日	US	2022092336	A1	2022年3月24日
				AU	2020437435	A1	2022年1月6日
				AU	2020437435	B2	2023年7月20日
				GB	2607647	A	2022年12月14日
				CN	111475797	A	2020年7月31日
				CN	111475797	B	2023年9月29日
US	2023022943	A1	2023年1月26日	CN	113554089	A	2021年10月26日