(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2022 年 1 月 20 日 (20.01.2022)

WIPO | PCT

(10) 国际公布号
WO 2022/012668 A1

(51) 国际专利分类号:
G06N 7/00 (2006.01)

(21) 国际申请号: PCT/CN2021/106758

(22) 国际申请日: 2021 年 7 月 16 日 (16.07.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202010692947.9 2020年7月17日 (17.07.2020) CN

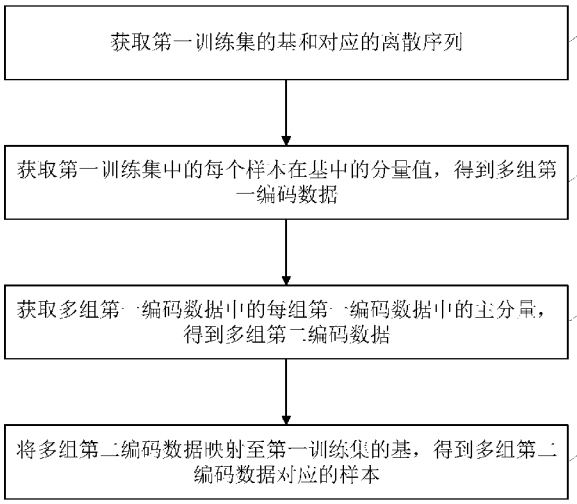
(71) 申请人: 华为技术有限公司 (**HUAWEI TECHNOLOGIES CO., LTD.**) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 清华大学(**TSINGHUA UNIVERSITY**) [CN/CN]; 中国北京市海淀区清华大学, Beijing 100084 (CN)。

(72) 发明人: 李玥儒(**LI, Yueru**); 中国山西省孝义市颐泰苑小区, Shanxi 032300 (CN)。 程书宇(**CHENG, Shuyu**); 中国上海市闵行区黎明路155弄星丰苑, Shanghai 201199 (CN)。 苏航(**SU, Hang**); 中国北京市海淀区清华大学FIT楼1-508, Beijing 100084 (CN)。 朱军(**ZHU, Jun**); 中国北京市海淀区清华大学FIT楼4-513, Beijing 100084 (CN)。 戴挺(**DAI, Ting**); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 时杰(**SHI, Jie**); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

(74) 代理人: 深圳市深佳知识产权代理事务所（普通合伙）(**SHENPAT INTELLECTUAL PROPERTY AGENCY**); 中国广东省深圳市罗湖区南湖街道春风路庐山大厦B座18C2、18D、18E、18E2, Guangdong 518001 (CN)。

(54) **Title:** TRAINING SET PROCESSING METHOD AND APPARATUS

(54) 发明名称: 一种训练集处理方法和装置



601 Acquire a basis and a corresponding discrete sequence of a first training set
602 Acquire a component value of each sample in the first training set in the basis, to obtain multiple groups of first encoded data
603 Acquire a primary component in each of the multiple groups of first encoded data, to obtain multiple groups of second encoded data
604 Map the multiple groups of second encoded data to the basis of the first training set, to obtain samples corresponding to the multiple groups of second encoded data

图 6

(57) **Abstract:** A training set processing method, used for training a neural network by extracting a primary component of a training set. More robust features can be used to train the neural network, thereby improving the robustness of the obtained neural network without decreasing the output accuracy of the neural network. Said method comprises: first, acquiring a basis and a discrete sequence of a first training set, the discrete sequence comprising discrete values corresponding to each basis vector in the basis on a one-to-one basis, and the discrete sequence being used for indicating the discrete degree of the basis; then, acquiring a component value of each sample on each basis vector so as to obtain multiple groups of first encoded data; subsequently, acquiring a primary component in each group of first encoded data, to obtain multiple groups of second encoded data, a discrete value corresponding to the primary component being higher than a preset discrete value; and performing mapping according to the multiple groups of second encoded data, to obtain new samples so as to form a second training set, the second training set being used for training a neural network.

WO 2022/012668 A1

[见续页]

(81) 指定国(除另有指明，要求每一种可提供的国家保护)：AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

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

本国际公布：

— 包括国际检索报告(条约第21条(3))。

(57) 摘要：一种训练集处理方法，用于通过提取训练集的主分量训练神经网络，可以使用更鲁棒的特征来进行神经网络的训练，从而在不降低神经网络的输出准确率的基础上，提升得到的神经网络的鲁棒性。该方法包括：首先，获取第一训练集的基和离散序列，离散序列包括与基中的每个基向量一一对应的离散值，离散序列用于表示基的离散程度；然后，获取每个样本在每个基向量上的分量值，以得到多组第一编码数据；随后，获取每组第一编码数据中的主分量，得到多组第二编码数据，主分量对应的离散值高于预设离散值；根据多组第二编码数据进行映射，得到新的样本，以组成第二训练集，第二训练集用于训练神经网络。

一种训练集处理方法和装置

本申请要求于 2020 年 07 月 17 日提交中国专利局、申请号为“202010692947.9”、申请名称为“一种训练集处理方法和装置”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

本申请涉及人工智能领域，尤其涉及一种训练集处理方法和装置。

背景技术

在人工智能领域中，神经网络近年来在计算机视觉、自然语言处理、网络安全等任务中的一系列突破，展现了智能时代的全新可能。然而神经网络普遍存在难以解释的对对抗样本的脆弱性，其有可能被加在其输入数据上的非常微小但精心设计的噪声所欺骗而得出错误的结果，这种精心构造的样本被称为对抗样本。

面对对抗具有脆弱性的人工智能系统显然无法使用于安全要求高的领域如医疗，智能驾驶等领域。因此分析导致人工神经网络脆弱的因素，获得具有对抗鲁棒性的人工神经网络至关重要。在现有方案中，可以在原样本的基础上增加较多的对抗样本，将原样本和对抗样本都加入训练集中进行训练，从而使训练得到的神经网络可以识别出对抗样本中的扰动。然而，对抗训练的效果受模型的影响较大，可能针对不同的模型产生了完全不同的效果，且对抗训练需要在训练过程中不断产生对抗样本，大大降低了训练效率，且可能造成训练得到的神经网络的输出准确率下降，即降低了神经网络的鲁棒性。因此，如何得到鲁棒性更优的神经网络，成为亟待解决的问题。

发明内容

本申请提供一种训练集处理方法，用于通过提取训练集的主分量训练神经网络，可以使用更鲁棒的特征来进行神经网络的训练，从而在不降低神经网络的输出准确率的基础上，提升得到的神经网络的鲁棒性。

有鉴于此，第一方面，本申请提供一种训练集处理方法，包括：首先，根据第一训练集进行分解，得到第一训练集的基和该基的离散序列，该第一训练集中包括多个样本，该第一训练集的基可以理解为多个方向的数据组成的空间，该基中包括至少一个基向量，离散序列包括与基中的每个基向量一一对应的离散值，离散序列可以用于表示基的离散程度；然后，获取第一训练集中的每个样本在每个基向量上的分量值，得到多组第一编码数据，其中，每组第一编码数据对应一个样本，每个样本在每个基向量中具有一个分量值，该第一编码数据中包括的每个基向量对应一个分量，第一编码数据包括的分量与至少一个基向量一一对应；获取多组第一编码数据中的每组第一编码数据中的主分量，得到多组第二编码数据，主分量的离散值高于预设离散值，或者说第一编码数据中的离散值高于预设值的分量组成了主分量；将该多组第二编码数据映射至第一训练集的，得到多组第二编码数据对应的新的样本，多组第二编码数据对应的样本组成第二训练集，第二训练集用于训练神经网络，从而得到鲁棒性更优的神经网络。

因此，在本申请实施方式中，在提取到每个样本在基中的第一编码数据中的主分量之后，得到降低了干扰之后的第二编码数据，可以根据第二编码数据得到新的样本，并使用新的样本进行神经网络训练。而新的样本中减少了除主要分量外的其他分量的干扰，从而使得训练得到的神经网络可以通过主分量进行训练，提高神经网络的鲁棒性。并且，本申请提供的数据集处理方法，针对训练集进行了处理，处理方式仅与训练集本身的主分量的分布有关，因此，即使在不同的场景中训练不同的模型，也可以使用本申请提供的训练集处理方法对训练集进行处理，泛化能力强，不依赖于被训练模型的通用算法，从而具有较高的通用性。

在一种可能的实施方式中，根据第一训练集获得所述第一训练集的基和所述基的离散序列，可以包括：对第一训练集进行主成分分析（principal component analysis，PCA）处理，得到正交基和正交基的离散序列。

通常，对第一训练集进行分解的方式可以有多种， 如PCA处理或者稀疏编码等方式对第一训练集进行分解，从而得到第一训练集的基和离散序列，进而获取到第一训练集的每个样本的分量。

在一种可能的实施方式中，对第一训练集进行 PCA 处理，得到正交基和正交基的离散序列，可以包括：对第一训练集进行中心化处理，得到中心化后的第一训练集，中心化后的第一训练集包括的数据的均值为 0；对中心化后的第一训练集进行 PCA 处理，得到正交基，以及正交基的离散序列，离散序列包括中心化后的第一训练集在正交基中的方差组成的序列。通常，离散序列所包括的方差呈递减排列，且离散序列所包括的方差按顺序和正交基中的基向量按顺序一一对应。

本申请实施方式中，可以对第一训练集进行中心化处理，得到中心化后的第一训练集，然后对中心化后的训练集进行 PCA 处理，从而完成对第一训练集的快速分析，得到正交基和正交基的离散序列。

在一种可能的实施方式中，获取第一训练集中的每个样本在基中的分量值，可以包括：通过预设算法，计算中心化后的第一训练集中的每个样本在正交基中的每个基向量上的分量值，得到多组第一编码数据。

本申请实施方式中，若对第一训练集进行 PCA 处理，相应地，可以基于 PCA 处理后的正交基计算每个样本在每个基向量上的分量值，从而得到第一编码数据。

在一种可能的实施方式中，预设算法包括内积运算或者稀疏编码运算。本申请实施方式中，可以通过内积运算或者稀疏编码运算计算每个样本在每个基向量上的分量值，实现了对每个样本在每个基向量上的投影。

在一种可能的实施方式中，获取每组第一编码数据中的主分量，得到多组第二编码数据，可以包括：保留每组第一编码数据中的主分量，将每组第一编码数据中除主分量外的其他分量替换为预设的值，得到多组第二编码数据，其中，每组第一编码数据的主分量组成至少一组第二编码数据。本申请实施方式中，可以将每组第一编码数据中除主分量外的其他分量替换为预设的值，从而使每组第一编码数据的分量可以组成一组或者多组第二编码数据。

在一种可能的实施方式中，预设的值包括 0 或者预设的噪声向量，预设的噪声向量包括高斯噪声或者均匀分布噪声。

因此，本申请实施方式中，可以将第一编码数据中除主分量以外的值替换为 0 或者预设的噪声向量，从而得到一组或者多组第二编码数据，以降低第一训练集中每个样本中除主分量以外的值对神经网络的训练的影响，提高最终得到的神经网络的鲁棒性。

在一种可能的实施方式中，预设的噪声向量与被换的分量的离散值成比例。因此，本申请实施方式中，替换的噪声向量与第一编码数据中原有的数据的分布方式类似，在降低了每个样本中除主分量以外的值对神经网络的训练的影响的基础上，降低对神经网络的训练的干扰，得到的神经网络的输出更准确。

第二方面，本申请提供一种训练集处理装置，包括：

分解单元，用于根据第一训练集获得第一训练集的基和该基的离散序列，第一训练集中包括多个样本，基中包括至少一个基向量，离散程度包括与基中的每个基向量一一对应的离散值，离散序列用于表示基的离散程度；

获取单元，用于获取第一训练集中的每个样本在基中的分量值，得到多组第一编码数据，每组第一编码数据对应一个样本；

获取单元，还用于获取多组第一编码数据中的每组第一编码数据中的主分量，得到多组第二编码数据，主分量的离散值高于预设离散值；

映射单元，用于将多组第二编码数据映射至第一训练集的基，得到多组第二编码数据对应的样本，多组第二编码数据对应的样本组成第二训练集，第二训练集用于训练神经网络。

第二方面及第二方面任一种可能的实施方式产生的有益效果可参照第一方面及第一方面任一种可能实施方式的描述。

在一种可能的实施方式中，基为正交基，分解单元，具体用于对第一训练集进行主成分分析 PCA 处理，得到正交基和正交基的离散序列。

在一种可能的实施方式中，分解单元，具体用于：对第一训练集进行中心化处理，得到中心化后的第一训练集，中心化后的第一训练集包括的数据的均值为 0；对中心化后的第一训练集进行 PCA 处理，得到正交基，以及正交基的离散序列，离散序列包括中心化后的第一训练集在正交基中的方差组成的序列。

在一种可能的实施方式中，分解单元，具体用于通过预设算法，计算中心化后的第一训练集中的每个样本在正交基中的每个基向量上的分量值，得到多组第一编码数据。

在一种可能的实施方式中，预设算法包括内积运算或者稀疏编码运算。

在一种可能的实施方式中，获取单元，具体用于保留每组第一编码数据中的主分量，将每组第一编码数据中除主分量外的其他分量替换为预设的值，得到多组第二编码数据，其中，每组第一编码数据的主分量组成至少一组第二编码数据。

在一种可能的实施方式中，预设的值包括 0 或者预设的噪声向量，预设的噪声向量包括高斯噪声或者均匀分布噪声。

在一种可能的实施方式中，预设的噪声向量与被替换的分量的离散值成比例。

第三方面，本申请实施例提供一种训练集处理装置，该训练集处理装置具有实现上述第一方面训练集处理方法的功能。该功能可以通过硬件实现，也可以通过硬件执行相应的软件实现。该硬件或软件包括一个或多个与上述功能相对应的模块。

第四方面，本申请实施例提供一种训练集处理装置，包括：处理器和存储器，其中，处理器和存储器通过线路互联，处理器调用存储器中的程序代码用于执行上述第一方面任一项所示的训练集处理方法中与处理相关的功能。可选地，该训练集处理装置可以是芯片。

第五方面，本申请实施例提供了一种训练集处理装置，该训练集处理装置也可以称为数字处理芯片或者芯片，芯片包括处理单元和通信接口，处理单元通过通信接口获取程序指令，程序指令被处理单元执行，处理单元用于执行如上述第一方面或第一方面任一可选实施方式中与处理相关的功能。

第六方面，本申请实施例提供了一种计算机可读存储介质，包括指令，当其在计算机上运行时，使得计算机执行上述第一方面或第一方面任一可选实施方式中的方法。

第七方面，本申请实施例提供了一种包含指令的计算机程序产品，当其在计算机上运行时，使得计算机执行上述第一方面或第一方面任一可选实施方式中的方法。

附图说明

图 1 本申请应用的一种人工智能主体框架示意图；

图 2 本申请提供的一种系统架构示意图；

图 3 为本申请实施例提供的一种卷积神经网络结构示意图；

图 4 为本申请实施例提供的另一种卷积神经网络结构示意图；

图 5 本申请提供的另一种系统架构示意图；

图 6 为本申请实施例提供的一种训练集处理方法的流程示意图；

图 7 为本申请实施例提供的一种对抗样本的示意图；

图 8 为本申请提供的一种训练集处理装置的一种结构示意图；

图 9 为本申请提供的一种训练集处理装置的另一种结构示意图；

图 10 为本申请实施例提供的一种芯片的结构示意图。

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做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。

本申请提供的训练集处理方法可以应用于人工智能（artificial intelligence, AI）场景中。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分支，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作出反应的智能机器。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。人工智能领域的研究包括机器人，自然语言处理，计算机视觉，决策与推理，人机交互，推荐与搜索，AI 基础理论等。

图 1 示出一种人工智能主体框架示意图，该主体框架描述了人工智能系统总体工作流程，适用于通用的人工智能领域需求。

下面从“智能信息链”（水平轴）和“IT 价值链”（垂直轴）两个维度对上述人工智能主题框架进行阐述。

“智能信息链”反映从数据的获取到处理的一列过程。举例来说，可以是智能信息感知、智能信息表示与形成、智能推理、智能决策、智能执行与输出的一般过程。在这个过程中，数据经历了“数据—信息—知识—智慧”的凝练过程。

“IT 价值链”从人智能的底层基础设施、信息（提供和处理技术实现）到系统的产业生态过程，反映人工智能为信息技术产业带来的价值。

（1）基础设施：

基础设施为人工智能系统提供计算能力支持，实现与外部世界的沟通，并通过基础平台实现支撑。通过传感器与外部沟通；计算能力由智能芯片，如中央处理器（central processing unit，CPU）、网络处理器（neural-network processing unit，NPU）、图形处理器（英语：graphics processing unit，GPU）、专用集成电路（application specific integrated circuit，ASIC）或现场可编程逻辑门阵列（field programmable gate array，FPGA）等硬件加速芯片）提供；基础平台包括分布式计算框架及网络等相关的平台保障和支持，可以包括云存储和计算、互联互通网络等。举例来说，传感器和外部沟通获取数据，这些数据提供给基础平台提供的分布式计算系统中的智能芯片进行计算。

（2）数据

基础设施的上一层的数据用于表示人工智能领域的数据来源。数据涉及到图形、图像、语音、视频、文本，还涉及到传统设备的物联网数据，包括已有系统的业务数据以及力、位移、液位、温度、湿度等感知数据。

（3）数据处理

数据处理通常包括数据训练，机器学习，深度学习，搜索，推理，决策等方式。

其中，机器学习和深度学习可以对数据进行符号化和形式化的智能信息建模、抽取、预处理、训练等。

推理是指在计算机或智能系统中，模拟人类的智能推理方式，依据推理控制策略，利用形式化的信息进行机器思维和求解问题的过程，典型的功能是搜索与匹配。

决策是指智能信息经过推理后进行决策的过程，通常提供分类、排序、预测等功能。

（4）通用能力

对数据经过上面提到的数据处理后，进一步基于数据处理的结果可以形成一些通用的能力，比如可以是算法或者一个通用系统，例如，翻译，文本的分析，计算机视觉的处理（如图像识别、目标检测等），语音识别等等。

（5）智能产品及行业应用

智能产品及行业应用指人工智能系统在各领域的产品和应用，是对人工智能整体解决

方案的封装，将智能信息决策产品化、实现落地应用，其应用领域主要包括：智能制造、智能交通、智能家居、智能医疗、智能安防、自动驾驶，平安城市，智能终端等。

参见附图 2，本申请实施例提供了一种系统架构 200。该系统架构中包括数据库 230、客户设备 240。数据采集设备 260 用于采集数据并存入数据库 230，训练模块 211 基于数据库 230 中维护的数据生成目标模型/规则 201。下面将更详细地描述训练模块 211 如何基于数据得到目标模型/规则 201，目标模型/规则 201 即本申请以下实施方式中构建得到的神经网络，具体参阅以下图 6 的相关描述。

计算模块可以包括训练模块 211，训练模块 211 得到的目标模型/规则可以应用不同的系统或设备中。在附图 2 中，执行设备 210 配置收发器 212，该收发器 212 可以是无线收发器、光收发器或有线接口（如 I/O 接口）等，与外部设备进行数据交互，“用户”可以通过客户设备 240 向收发器 212 输入数据，例如，本申请以下实施方式，客户设备 240 可以向执行设备 210 发送目标任务，请求执行设备构建神经网络，并向执行设备 210 发送用于训练的数据库。

执行设备 210 可以调用数据存储系统 250 中的数据、代码等，也可以将数据、指令等存入数据存储系统 250 中。

计算模块 211 使用目标模型/规则 201 对输入的数据进行处理。具体地，计算模块 211 用于：首先，获取第一训练集，该第一训练集中包括多个样本；然后，根据第一训练集得到第一训练集的基和基的离散序列，第一训练集中包括多个样本，基中包括至少一个基向量，离散序列包括与基中的每个基向量一一对应的离散值，该离散序列可以用于表示基的离散程度；获取第一训练集中的每个样本在每个基向量上的分量值，以得到多组第一编码数据，每组第一编码数据对应一个样本；获取多组第一编码数据中的每组第一编码数据中的主分量，得到多组第二编码数据，主分量的离散值高于预设离散值；将多组第二编码数据映射至第一训练集的基，得到多组第二编码数据对应的样本，多组第二编码数据对应的样本组成第二训练集，第二训练集用于训练神经网络。

最后，收发器 212 将构建得到的神经网络返回给客户设备 240，以在客户设备 240 或者其他设备中部署该神经网络。

更深层地，训练模块 211 可以针对不同的任务，基于不同的数据得到相应的目标模型/规则 201，以给用户提供更佳的结果。

在附图 2 中所示情况下，可以根据用户的输入数据确定输入执行设备 210 中的数据，例如，用户可以在收发器 212 提供的界面中操作。另一种情况下，客户设备 240 可以自动地向收发器 212 输入数据并获得结果，若客户设备 240 自动输入数据需要获得用户的授权，用户可以在客户设备 240 中设置相应权限。用户可以在客户设备 240 查看执行设备 210 输出的结果，具体的呈现形式可以是显示、声音、动作等具体方式。客户设备 240 也可以作为数据采集端将采集到与目标任务关联的数据存入数据库 230。

需要说明的是，附图 2 仅是本申请实施例提供的一种系统架构的示例性的示意图，图中所示设备、器件、模块等之间的位置关系不构成任何限制。例如，在附图 2 中，数据存储系统 250 相对执行设备 210 是外部存储器，在其它场景中，也可以将数据存储系统 250

置于执行设备 210 中。

在本申请所提及的训练或者更新过程可以由训练模块 211 来执行。可以理解的是，神经网络的训练过程即学习控制空间变换的方式，更具体即学习权重矩阵。训练神经网络的目的是使神经网络的输出尽可能接近期望值，因此可以通过比较当前网络的预测值和期望值，再根据两者之间的差异情况来更新神经网络中的每一层神经网络的权重向量（当然，在第一次更新之前通常可以先对权重向量进行初始化，即为深度神经网络中的各层预先配置参数）。例如，如果网络的预测值过高，则调整权重矩阵中的权重的值从而降低预测值，经过不断的调整，直到神经网络输出的值接近期望值或者等于期望值。具体地，可以通过损失函数（loss function）或目标函数（objective function）来衡量神经网络的预测值和期望值之间的差异。以损失函数举例，损失函数的输出值（loss）越高表示差异越大，神经网络的训练可以理解为尽可能缩小 loss 的过程。本申请以下实施方式中更新起点网络的权重以及对串行网络进行训练的过程可以参阅此过程，以下不再赘述。

本申请提及的神经网络可以包括多种类型，如深度神经网络（deep neural network，DNN）、卷积神经网络（convolutional neural network，CNN）、循环神经网络（recurrent neural networks，RNN)或残差网络其他神经网络等。

示例性地，下面以 CNN 为例。

CNN 是一种带有卷积结构的深度神经网络。CNN 是一种深度学习（deep learning）架构，深度学习架构是指通过机器学习的算法，在不同的抽象层级上进行多个层次的学习。作为一种深度学习架构，CNN 是一种前馈（feed-forward）人工神经网络，该前馈人工神经网络中的各个神经元对输入其中的图像中的重叠区域作出响应。卷积神经网络包含了一个由卷积层和子采样层构成的特征抽取器。该特征抽取器可以看作是滤波器，卷积过程可以看作是使用一个可训练的滤波器与一个输入的图像或者卷积特征平面(feature map)做卷积。卷积层是指卷积神经网络中对输入信号进行卷积处理的神经元层。在卷积神经网络的卷积层中，一个神经元可以只与部分邻层神经元连接。一个卷积层中，通常包含若干个特征平面，每个特征平面可以由一些矩形排列的神经单元组成。同一特征平面的神经单元共享权重，这里共享的权重就是卷积核。共享权重可以理解为提取图像信息的方式与位置无关。这其中隐含的原理是：图像的某一部分的统计信息与其他部分是一样的。即意味着在某一部分学习的图像信息也能用在另一部分上。所以对于图像上的所有位置，我们都能使用同样的学习得到的图像信息。在同一卷积层中，可以使用多个卷积核来提取不同的图像信息，一般地，卷积核数量越多，卷积操作反映的图像信息越丰富。

卷积核可以以随机大小的矩阵的形式初始化，在卷积神经网络的训练过程中卷积核可以通过学习得到合理的权重。另外，共享权重带来的直接好处是减少卷积神经网络各层之间的连接，同时又降低了过拟合的风险。

卷积神经网络可以采用误差反向传播（back propagation，BP）算法在训练过程中修正初始的超分辨率模型中参数的大小，使得超分辨率模型的重建误差损失越来越小。具体地，前向传递输入信号直至输出会产生误差损失，通过反向传播误差损失信息来更新初始的超分辨率模型中参数，从而使误差损失收敛。反向传播算法是以误差损失为主导的反向

传播运动，旨在得到最优的超分辨率模型的参数，例如权重矩阵。

如图 3 所示，卷积神经网络（CNN）100 可以包括输入层 110，卷积层/池化层 120，其中池化层为可选的，以及神经网络层 130。

如图 3 所示卷积层/池化层 120 可以包括如示例 121-126 层，在一种实现中，121 层为卷积层，122 层为池化层，123 层为卷积层，124 层为池化层，125 为卷积层，126 为池化层；在另一种实现方式中，121、122 为卷积层，123 为池化层，124、125 为卷积层，126 为池化层。即卷积层的输出可以作为随后的池化层的输入，也可以作为另一个卷积层的输入以继续进行卷积操作。

以卷积层 121 为例，卷积层 121 可以包括很多个卷积算子，卷积算子也称为核，其在图像处理中的作用相当于一个从输入图像矩阵中提取特定信息的过滤器，卷积算子本质上可以是一个权重矩阵，这个权重矩阵通常被预先定义。在对图像进行卷积操作的过程中，权重矩阵通常在输入图像上沿着水平方向一个像素接着一个像素（或两个像素接着两个像素……这取决于步长 stride 的取值）的进行处理，从而完成从图像中提取特定特征的工作。该权重矩阵的大小应该与图像的大小相关。需要注意的是，权重矩阵的纵深维度（depth dimension）和输入图像的纵深维度是相同的，在进行卷积运算的过程中，权重矩阵会延伸到输入图像的整个深度。因此，和一个单一的权重矩阵进行卷积会产生一个单一纵深维度的卷积化输出，但是大多数情况下不使用单一权重矩阵，而是应用维度相同的多个权重矩阵。每个权重矩阵的输出被堆叠起来形成卷积图像的纵深维度。不同的权重矩阵可以用来提取图像中不同的特征，例如一个权重矩阵用来提取图像边缘信息，另一个权重矩阵用来提取图像的特定颜色，又一个权重矩阵用来对图像中不需要的噪点进行模糊化等。该多个权重矩阵维度相同，经过该多个维度相同的权重矩阵提取后的特征图维度也相同，再将提取到的多个维度相同的特征图合并形成卷积运算的输出。

通常，权重矩阵中的权重值在实际应用中需要经过大量的训练得到，通过训练得到的权重值形成的各个权重矩阵可以从输入图像中提取信息，从而帮助卷积神经网络 100 进行正确的预测。

当卷积神经网络 100 有多个卷积层时，初始的卷积层（例如 121）往往提取较多的一般特征，该一般特征也可以称之为低级别的特征；随着卷积神经网络 100 深度的加深，越往后的卷积层（例如 126）提取到的特征越来越复杂，比如高级别的语义之类的特征，语义越高的特征越适用于待解决的问题。

池化层：

由于常常需要减少训练参数的数量，因此卷积层之后常常需要周期性的引入池化层，即如图 3 中 120 所示例的 121-126 各层，可以是一层卷积层后面跟一层池化层，也可以是多层卷积层后面接一层或多层池化层。在图像处理过程中，池化层的唯一目的就是减少图像的空间大小。池化层可以包括平均池化算子和/或最大池化算子，以用于对输入图像进行采样得到较小尺寸的图像。平均池化算子可以在特定范围内对图像中的像素值进行计算产生平均值。最大池化算子可以在特定范围内取该范围内值最大的像素作为最大池化的结果。另外，就像卷积层中用权重矩阵的大小应该与图像大小相关一样，池化层中的运算符也应

该与图像的大小相关。通过池化层处理后输出的图像尺寸可以小于输入池化层的图像的尺寸，池化层输出的图像中每个像素点表示输入池化层的图像的对应子区域的平均值或最大值。

神经网络层 130：

在经过卷积层/池化层 120 的处理后，卷积神经网络 100 还不足以输出所需要的输出信息。因为如前所述，卷积层/池化层 120 只会提取特征，并减少输入图像带来的参数。然而为了生成最终的输出信息（所需要的类信息或别的相关信息），卷积神经网络 100 需要利用神经网络层 130 来生成一个或者一组所需要的类的数量的输出。因此，在神经网络层 130 中可以包括多层隐含层（如图 3 所示的 131、132 至 13n）以及输出层 140。在本申请中，该卷积神经网络为：对选取的起点网络进行至少一次变形得到串行网络，然后根据训练后的串行网络得到。该卷积神经网络可以用于图像识别，图像分类，图像超分辨率重建等等。

在神经网络层 130 中的多层隐含层之后，也就是整个卷积神经网络 100 的最后层为输出层 140，该输出层 140 具有类似分类交叉熵的损失函数，具体用于计算预测误差，一旦整个卷积神经网络 100 的前向传播（如图 3 由 110 至 140 的传播为前向传播）完成，反向传播（如图 3 由 140 至 110 的传播为反向传播）就会开始更新前面提到的各层的权重值以及偏差，以减少卷积神经网络 100 的损失及卷积神经网络 100 通过输出层输出的结果和理想结果之间的误差。

需要说明的是，如图 3 所示的卷积神经网络 100 仅作为一种卷积神经网络的示例，在具体的应用中，卷积神经网络还可以以其他网络模型的形式存在，例如，如图 4 所示的多个卷积层/池化层并行，将分别提取的特征均输入给全神经网络层 130 进行处理。

参见附图 5，本申请实施例还提供了一种系统架构 300。执行设备 210 由一个或多个服务器实现，可选的，与其它计算设备配合，例如：数据存储、路由器、负载均衡器等设备；执行设备 210 可以布置在一个物理站点上，或者分布在多个物理站点上。执行设备 210 可以使用数据存储系统 250 中的数据，或者调用数据存储系统 250 中的程序代码实现本申请以下图 6 对应的训练集处理方法的步骤。

用户可以操作各自的用户设备（例如本地设备 301 和本地设备 302）与执行设备 210 进行交互。每个本地设备可以表示任何计算设备，例如个人计算机、计算机工作站、智能手机、平板电脑、智能摄像头、智能汽车或其他类型蜂窝电话、媒体消费设备、可穿戴设备、机顶盒、游戏机等。

每个用户的本地设备可以通过任何通信机制/通信标准的通信网络与执行设备 210 进行交互，通信网络可以是广域网、局域网、点对点连接等方式，或它们的任意组合。具体地，该通信网络可以包括无线网络、有线网络或者无线网络与有线网络的组合等。该无线网络包括但不限于：第五代移动通信技术（5th-Generation，5G）系统，长期演进（long term evolution，LTE）系统、全球移动通信系统（global system for mobile communication，GSM）或码分多址（code division multiple access，CDMA）网络、宽带码分多址（wideband code division multiple access，WCDMA）网络、无线保真（wireless fidelity，WiFi）、蓝牙（bluetooth）、紫蜂协议（Zigbee）、射频识别技术（radio frequency identification，

RFID)、远程(Long Range，Lora)无线通信、近距离无线通信(near field communication，NFC)中的任意一种或多种的组合。该有线网络可以包括光纤通信网络或同轴电缆组成的网络等。

在另一种实现中，执行设备 210 的一个方面或多个方面可以由每个本地设备实现，例如，本地设备 301 可以为执行设备 210 提供本地数据或反馈计算结果。

需要注意的，执行设备 210 的所有功能也可以由本地设备实现。例如，本地设备 301 实现执行设备 210 的功能并为自己的用户提供服务，或者为本地设备 302 的用户提供服务。

在一些场景中，可以在原样本的基础上增加较多的对抗样本，将原样本和对抗样本都加入训练集中进行训练，从而使训练得到的神经网络可以识别出对抗样本中的扰动。然而，对抗训练的效果受模型的影响较大，可能针对不同的模型产生了完全不同的效果，且对抗训练需要在训练过程中不断产生对抗样本，大大降低了训练效率，且可能造成训练得到的神经网络的输出准确率下降，即降低了神经网络的鲁棒性。鲁棒性可以理解为神经网络在面对输入的微小变化能保持输出值不变的能力。或者，在另一些场景中，可以针对对抗样本进行去噪，在神经网络前添加一个去噪网络，通过梯度下降算法减小成对的对抗样本和原样本产生的顶层表示的差距，从而使去噪网络达到对抗鲁棒效果。然而，该去噪网络需要有针对性地添加，在不同的训练场景中可能需要设定不同的去噪网络，泛化能力弱，导致训练效率低。

通常，神经网络的脆弱性源于在进行神经网络的训练的过程中，更倾向于学习到哪些在数据分布流行之外的变化比较剧烈的函数，即神经网络主要依靠在线性空间中变化比较小，但同样包含可以用来分类的信息的那些分量施行分类，该分量即前述的除主分量外的分量。为便于理解，可以理解为，训练集的样本中所包括的主分量之外的其他分量，对神经网络的训练产生了较大影响，其往往具有较高的线性可分性而更容易由神经网络学得并泛化，这些主分量之外的其他分量影响神经网络的鲁棒性的主要原因，降低了神经网络的鲁棒性，该现象在以下称为梯度泄漏现象。

因此，本申请提供一种训练集处理方法，通过在训练之前，对训练集进行处理，提取每个样本的主分量，得到新的样本，组成第二训练集，该第二训练集中所包括的样本包括了原样本中的主分量，从而使通过第二训练集训练神经网络，可以提升神经网络的鲁棒性。

参阅图 6，本申请提供的一种训练集处理方法的流程示意图，如下所述。

601、获取第一训练集的基和对应的离散序列。

其中，可以对第一训练集包括的所有数据进行分解，得到第一训练集的基和该基对应的离散序列，该第一训练集中包括多个样本，该基可以包括一个或者多个基向量，该离散序列用于表示基的离散程度，该离散序列可以包括与基中的每个基向量一一对应的离散值。

为便于理解，前述的基可以理解为空间，该空间中可以包括一个或者多个基向量，离散序列用于表示第一训练集所包括的数据在基中的分散程度。

应理解，第一训练集所包括的样本的类型，与需要训练的神经网络相关。例如，若需要训练分类网络，则第一训练集中可以包括大量的已进行分类的图像，若需要训练人脸识别网络，则第一训练集中可以包括大量的人脸图像等。

在一种可选的实施方式中，前述的离散值可以是方差、标准差或者极差等用于表示离散程度的值。下面示例性地，以方差来表示本申请实施例中的基的离散程度为例进行示例性说明，应理解，以下所提及的方差也可以替换为标准差或者极差等用于表示离散程度的值，以下不再赘述。

具体地，对第一训练集进行分解的方式可以有多种，如 PCA 或者稀疏编码算法等。

例如，可以通过稀疏编码算法对第一训练集进行分解。针对数据集 X，解优化问题 $\min\sum_i(\|AS_i - X_i\|_2^2 + \lambda\|S_i\|_0)$, s. t. $\|A_i\|_2 \leq 1$ 其中A_i是A的第i列，S_i，X_i同理；$\|.\|_p$代表 L_p范数，λ是正则项；从而得到X_i的稀疏编码S_i和字典矩阵（即基）A。

又例如，对第一训练集进行 PCA 处理，得到正交基和正交基对应的离散序列。PCA 是对通过输入数据的协方差矩阵进行正交谱分解，以确定数据点之间主要变化的线性方向，从而实现维数缩减的算法。为便于理解，该 PCA 处理可以理解为对输入数据，即第一数据集的协方差矩阵进行正交谱分解以确定数据点之间主要变化的线性方向，从而实现降维的算法。为便于理解，下面示例性地，以对第一训练集进行 PCA 处理为例，对后续的步骤进行示例性说明。

在一种可能的实施方式中，前述的对第一训练集进行主成分分析 PCA 处理，得到正交基和正交基对应的离散序列，具体可以包括：对第一训练集进行中心化处理，得到中心化处理后的第一训练集。然后对中心化后的第一训练集进行 PCA 处理，得到正交基和该正交基对应的离散序列，该离散序列为包括了中心化后的第一训练集在正交基中的方差组成的序列，离散序列所包括的方差呈递减排列，且离散序列所包括的方差按顺序和正交基中的基向量按顺序一一对应。

前述的中心化处理可以理解为零均值化，即检测第一训练集所包括的所有数据的均值，在空间中平移第一训练集，如将第一训练集中的所有数据减均值，使第一训练集所包括的所有数据的均值为 0，得到中心化后的第一训练集。

在一种可能的实施方式中，在计算得到每个基向量对应的离散值，即方差之后，可以对计算得到的方差进行排序，按照递减的顺序进行排列，得到离散序列。然后对正交基所包括的基向量进行排序，且正交基中的基向量的排列方式与离散序列中的排序方式对应。例如，正交基可以表示为$\{v_1, v_2, \cdots.., v_N\}$，离散序列可以表示为$\{\lambda_1, \lambda_2, \cdots.., \lambda_N\}$，基向量$v_1$对应的方差为$\lambda_1$，基向量$v_2$对应的方差为$\lambda_2$等，以此类推。

当然，离散序列中的也可以按照其他方式进行排序，例如，可以按照递增的顺序排列、或者按照其他设定的规律等，相应地，正交基中的基向量也可以按照与离散序列对应的排序方式进行排列。

602、获取第一训练集中的每个样本在基中的分量值，得到多组第一编码数据。

其中，在对第一训练集进行分解得到基和对应的离散序列之后，获取第一训练集中的每个样本在基中的每个基向量中的分量值，得到多组第一编码数据，每个样本在每个基向量中的分量值组成一组第一编码数据。

为便于理解，前述的基可以理解为第一训练集所包括的数据组成的空间，基向量即表示该空间内的方向，步骤 602 即获取每个样本投影到每个方向上的分量值，从而得到每个

样本的第一编码数据。

可选地，若对第一训练集进行 PCA 处理，得到正交基和对应的离散序列，则可以获取第一训练集中的每个样本在正交基中的分量，得到每个样本对应的第一编码数据。

可选地，若在步骤 601 中，对第一训练集进行了中心化，则步骤 602 中的第一训练集可以替换为中心化后的第一训练集。

可选地，可以通过预设算法计算第一训练集中的每个样本在基中的分量值，得到多组第一编码数据。该预设算法可以包括内积运算或者稀疏编码运算等。例如，以内积运算为例，可以对第一训练集中的每个样本和基中的每个基向量进行内积运算，得到每个样本的第一编码数据。

603、获取多组第一编码数据中的每组第一编码数据中的主分量，得到多组第二编码数据。

其中，在得到每个样本的第一编码数据之后，提取该多组第一编码数据中每组第一编码数据中的主分量，从而得到多组第二编码数据。该主分量可以理解为对应的离散值大于预设离散值的分量，或者说，该主分量可以理解为第一编码数据中对应的离散值最大的预设数量的分量。

例如，前述步骤 601 中的基可以是正交基，该正交基可以表示为$\{v_1, v_2, \cdots.., v_N\}$，离散序列可以表示为$\{\lambda_1, \lambda_2, \cdots.., \lambda_N\}$，且正交基中的基向量与离散序列中的方差按顺序一一对应。第一编码数据可以表示为 a:$\{a_1, a_2, \cdots\cdots, a_N\}$，该第一编码数据所包括的分量与离散序列中的方差按顺序一一对应，因此，可以根据确定出离散序列中大于预设离散值的方差，并从第一编码数据中确定出与该大于预设离散值的方差对应的分量作为主分量，得到第二编码数据。例如，当离散值以递减的顺序排列时，从 a:$\{a_1, a_2, \cdots\cdots, a_N\}$可以提取得到 a′ :$\{a_1, a_2, \cdots\cdots, a_\mathrm{K}\}$，a′ 即一组第二编码数据，K 小于或者等于 N。

在一种可能的实施方式中，可以保留每组第一编码数据中的主分量，将每组第一编码数据中除主分量外的其他分量替换为预设的值，得到多组第二编码数据。其中，每组第一编码数据可以对应一个或者多组第二编码数据。例如，可以将每组第一编码数据中除主分量外的其他分量通过多种不同的方式替换，得到多组第二编码数据，从而得到每组第一编码数据对应的多组第二编码数据。

可选地，该预设的值可以包括 0 或者预设的噪声向量。该预设的噪声向量可以包括与样本的真实值独立的噪声，或者与样本的真实值独立，但与第一训练集中的数据分布类似的噪声，或者特定分布的噪声等。例如，该预设的噪声向量可以是高斯噪声或者均匀分布特征等。因此，在本申请实施方式中，通过将除主要分量之外的其他分量替换为 0 或者噪声向量的方式，消除该部分分量对训练神经网络时的影响，使后续在训练神经网络时，可以以主分量为主，得到鲁棒性更优的神经网络。

在一种可能的实施方式中，若将每组第一编码数据中除主分量外的其他分量替换为噪声，则替换的噪声可以与替换前的分量对应的方差成比例。例如，噪声可以是原分量对应的方差的 5 倍、10 倍等。因此，本申请实施方式中，替换后的噪声的分布规律可以与原分量的方差的分布规律类似，从而降低对后续训练神经网络的影响，使得在训练神经网络的

过程中，在使用主分量进行训练，得到鲁棒性更优的神经网络的前提下，避免对神经网络的输出准确率的影响，得到鲁棒性和输出准确率更平衡的神经网络。可以理解为，若前述步骤 601 中对第一训练集采用 PCA 处理，本申请实施方式中，可以在 PCA 子空间的法方向增加一定大小的噪声，以抑制梯度泄漏现象。

在一种可能的场景中，若离散序列中的方差值按照递减的顺序排列，相应地，第一编码数据所包括的分量的排列顺序与离散序列中的方差的排列顺序对应，确定大于预设离散值的方差的数量 d，该 d 为正整数，保留第一编码数据中的前 d 个分量，将第 d+1 至第 D 个分量替换为与分量对应的方差成比例的独立高斯噪声，并将第 D 个分量之后的分量替换为 0，得到第二编码数据。因此，在此场景中，可以将第一编码数据中除主分量外的其他分量替换为 0 或者噪声，从而使后续进行神经网络训练时，可以使用主分量进行训练，可以避免对抗样本中的扰动对神经网络造成的影响，提高神经网络的鲁棒性。

604、将多组第二编码数据映射至第一训练集的基，得到多组第二编码数据对应的样本。

在得到多组第二编码数据之后，将多组第二编码数据映射至第一训练集的基中，得到多组第二编码数据中每组第二编码数据对应的新的样本，该多组第二编码数据对应的样本组成第二训练集。为便于理解，第一训练集的基可以理解为一个空间，第二编码数据中包括了该空间内的一个或者多个方向上的分量值，将一组第二编码数据映射至该空间中，即可得到新的样本。

具体地，在得到第二编码数据之后，可以根据第二编码数据对前述步骤 601 中的基进行映射，得到每组第二编码数据对应的新样本。该多组第二编码数据对应的多个新样本组成第二训练集。为便于理解，可以将前述步骤 601 中提及的基理解为空间，第二编码数据中即包括新样本在每个基向量的方向上投影的值，对第二编码数据所包括的分量在基中进行映射，即可得到新的样本。

可选地，在得到新样本组成的第二训练集之后，即可使用该第二训练集进行神经网络训练，得到鲁棒性更优的神经网络。

因此，在本申请实施方式中，在提取到每个样本在基中第一编码数据中的主分量之后，得到降低了干扰之后的第二编码数据，可以根据第二编码数据得到新的样本，并使用新的样本进行神经网络训练。而新的样本中减少了除主要分量外的其他分量的干扰，从而使得训练得到的神经网络可以通过主分量进行训练，提高神经网络的鲁棒性。并且，本申请提供的训练集处理方法，针对训练集进行了处理，处理方式仅与训练集本身的主分量的分布有关，因此，即使在不同的场景中训练不同的模型，也可以使用本申请提供的训练集处理方法对训练集进行处理，泛化能力强，不依赖于被训练模型的通用算法，从而具有较高的通用性。

通常，神经网络模型对于对抗样本容易误分类。对抗样本即包括了扰动信息的样本，对抗样本所包括的扰动将降低神经网络的输出准确率。该对抗样本即包括了扰动信息的样本，该对抗样本所包括的扰动将影响后续训练得到的神经网络的鲁棒性，降低神经网络的输出准确率。示例性地，对抗样本的示意图可以如图 7 所示，图像 A1 为未包括扰动的图像，在图像 A1 的基础上叠加扰动信息 A2，0.007 标识添加的扰动信息的系数，从而得到增加了

扰动信息的图像 A3，该图像 A3 即对抗样本。通常，对抗样本中包括了扰动信息，该扰动信息容易造成神经网络的识别错误，如图像 A1 的实际分类为熊猫，添加了扰动信息之后的图像 3 则可能被识别为分类为狗，因此，若使用该对抗样本进行训练，则可能导致神经网络的鲁棒性差。而本申请实施方式中，在使用训练集进行神经网络训练之前，对训练集进行了处理，减少了训练集中每个样本中的扰动部分，从而使后续进行神经网络训练时，可以使用减少了扰动的训练集进行训练，避免样本中的扰动对神经网络的影响，提高神经网络的鲁棒性。

而本申请实施方式中，以通过消除干扰的方式，使后续训练神经网络时，以样本的主分量为主，训练得到鲁棒性更优的神经网络。且可以以样本的主分量为主，训练得到神经网络，在提升神经网络的鲁棒性的同时，减少对神经网络的输出准确率的影响，得到鲁棒性和输出准确率均衡的神经网络。

可以理解为，在本申请提供的训练集处理方法中，保留了原样本对应的编码数据中的主分量，将除主分量外的其他分量替换为噪声或者 0，从而得到消除了除主分量外的其他分量的新样本。在使用包括了新样本的训练集进行神经网络训练时，可以以主分量为主进行训练，减少了样本中所包括的扰动。相当于在训练的过程中，引入了数据流形来对对抗样本进行防御，以提升神经网络的鲁棒性。数据流形可以理解为，假设中的一个嵌入在数据真实维数的线性空间，但维度远低于该空间的流形，观察得到的输入数据点都分布在其附近，类似自然图像等神经网络擅长分类的数据基本都符合该分布假设。通常，不同的样本的主分量的差异较大，对主分量成功分类，可以得到较高的鲁棒性的神经网络。从而在不降低神经网络的输出准确率的基础上，提升了神经网络的鲁棒性，抑制了神经网络训练过程中的梯度泄漏的问题。

前述对本申请提供的训练集处理方法的流程进行了详细说明，为便于理解，下面结合具体的应用场景，对本申请提供的训练集处理方法进行更详细的说明。

第一训练集中的样本可以表示为一个数据点x_i，该第一训练集可以是 N*D 维的数据集，该第一数据集可以表示为$\mathrm{X} = [x_1, \ldots x_N]^T$。对训练集进行 PCA 投影之后，得到正交基$\{v_1, v_2, \cdots.., v_N\}$，和对应的离散序列$\{\lambda_1, \lambda_2, \cdots.., \lambda_N\}$，且离散序列所包括的方差呈递减排列。确定主分量所包括的分量维数 d，增加的噪声维数 m，以及噪声大小 c。

利用协方差矩阵$\mathrm{C} = \frac{1}{N}\sum_{n=1}^{N}(x_n - \bar{x})(x_n - \bar{x})^T$进行谱分解，其中，$\bar{x} = \frac{1}{N}\sum_{n=1}^{N} x_n$。得到$\mathrm{C} = \sum_{i=d}^{D} \lambda_i v_i {v_i}^T$，$\lambda_i \geq \lambda_2 \geq \cdots \geq \lambda_D$。

计算$x - \bar{x}$的均值在正交基 V 中的分量，$\mathrm{V} = [v_1, \ldots v_D]$，得到$V^T(x - \bar{x}) \rightarrow \mathrm{a}$，a即前述的第一编码数据。

把x投影到 PCA 子空间，将a的第i个分量替换为噪声，其中，$c\sqrt{\lambda_i}\xi_i a_i$，$\xi_i \overset{i.i.d}{\sim} \mathcal{N}(0,1)$，且$i = d+1, d+2 \ldots d+m$，$d+m$之后的分量替换为 0，得到更新后的$a$，$c\sqrt{\lambda_i}\xi_i a_i$表示噪声，$\xi_i$服从相互独立同分布的标准正态（高斯）分布。

然后进行重构，得到新样本 $Va + \bar{x} \rightarrow x'$，多个新样本即组成第二训练集。

随后，即可使用新的第二训练集进行神经网络训练。

因此，在本场景中，将主分量为的其他分量替换为了与对应的方差成比例的噪声，在不影响神经网络的训练过程的基础上，提升神经网络的鲁棒性。

为便于理解，对另一些具体的场景进行示例性说明。

首先，对训练集$\{(x_i, y_i)\}, i = 1,2, \ldots, n$中的输入数据（设为行向量）进行中心化，得到$x_i^{'} = x_i - \frac{1}{n}\sum_j x_j$，然后计算器协方差矩阵：$\Sigma = \sum_j x_j^{T'} x_j$，设其正交对角化结果为$U^T diag(B)U$，其中，$B$为一个维度与输入数据（即训练集中的样本）相同的非负向量，表示输入数据在这些方向上的变化量的大小。

然后取样本的公式可以包括：$x_{new} = U^T(s \odot Ux_{old} + c(1-s) \odot \sqrt{B} \odot \epsilon)$，$Ux_{old}$即表示前述的第一编码数据，$s$为尺度缩放向量，$\epsilon$为噪声向量，$c$为大于 1 的常数。通常，可以取$s$为截断向量$\{1,1, \ldots, 1,1,0,0, \ldots, 0\}$, 1 的数量为截断超参数，即前述的预设值 d。为了达到鲁棒目的，一般需要令对应尺度较小的分量进一步被压制，ϵ可以取在截断值附近的标准正态分布向量（即在截断值附近的维度为标准正态分布，其他维度为 0）。

可选地，前述的尺度缩放向量可以替换为平滑向量，如该平滑向量中的某一个点可以表示为$s_i = sigmoid(-i + d)$，d表示软截断阈值，i表示维度的下标。

或者，可选地，前述的尺度缩放向量也可以替换为随机向量，如该随机向量中的某一个点可以表示为$s_i = Bernoulli(sigmoid(-i + d))$。

或者，前述的噪声向量也可以替换为非高斯分布，但与样本本身的分布独立的其他方差水平类似的向量。

因此，在本场景中，提出了一种检测并改善训练集中非鲁棒部分导致神经网络对抗脆弱性的方法，本申请提供的方法具有通用性和高效性，可以直接适用于大多数任务和网络架构，泛化能力强，在提高神经网络的鲁棒性的基础上，可以训练得到输出准确率更高的神经网络，得到鲁棒性和输出准确率达到均衡的神经网络。

通常，以分类网络为例，若采用在测试阶段进行投影的防御方法，即在测试阶段对对抗样本中的扰动进行训练，由于投影到的流形相比于真实的数据流形存在误差，训练时使用的是未经过投影的数据来训练，则模型对被投影到的流形上的数据的分类效果是未知的。一方面会导致明显的分类精度下降，另一方面由于分类网络没有在该流形上良好训练，较容易找到该流形上的对抗样本。而本申请提供的方法中，在训练中就进行投影操作，可以避免前述的问题。此外，通过本申请提供的方法得到的训练集，训练得到的分类网络在测试时可不必对输入数据进行投影，提升了训练效率。

若采用对抗样本进行训练，往往需要数倍的时间和更高的计算资源以计算输入样本的对抗样本。而本申请提供的方法中，将相对本质的流形外对抗方向分开进行加噪声增广，避免了在模型训练时产生的流形内，误分类样本训练造成的难度提升和原测试集准确率下降。因此本发明能够在鲁棒准确率和原测试集准确率之间有更好和更灵活的权衡取舍。并且，本申请可以根据应用场景需要，通过调节参数 d 和所加噪声具体的分布及幅度，可以灵活调节模型训练最终的准确率，调整神经网络的输出准确率和鲁棒性的均衡。

前述对本申请提供的方法进行了详细介绍，下面基于前述的方法，对本申请提供的装

置进行介绍。

参阅图 8，本申请提供的一种训练集处理装置的结构示意图，该训练集处理装置用于执行前述图 6 中的方法的步骤。

该训练集处理装置可以包括：

分解单元 801，用于根据第一训练集进行分解获得第一训练集的基和该基对应的离散序列，第一训练集中包括多个样本，基中包括至少一个基向量，离散程度包括与基中的每个基向量一一对应的离散值，该离散序列可以用于表示基的离散程度；

获取单元 802，用于获取第一训练集中的每个样本在基中的分量值，得到多组第一编码数据，每组第一编码数据对应一个样本；

获取单元 802，还用于获取多组第一编码数据中的每组第一编码数据中的主分量，得到多组第二编码数据，主分量对应的离散值高于预设离散值，每个第一；

映射单元 803，用于将多组第二编码数据映射至第一训练集的基中，得到多组第二编码数据对应的样本，多组第二编码数据对应的样本组成第二训练集，第二训练集用于训练神经网络。

在一种可能的实施方式中，基为正交基，分解单元 801，具体用于对第一训练集进行主成分分析 PCA 处理，得到正交基和正交基中对应的离散序列。

在一种可能的实施方式中，分解单元 801，具体用于：对第一训练集进行中心化处理，得到中心化后的第一训练集，中心化后的第一训练集包括的数据的均值为 0；对中心化后的第一训练集进行 PCA 处理，得到正交基，以及正交基的离散序列，离散序列包括中心化后的第一训练集在正交基中的方差组成的序列。

在一种可能的实施方式中，分解单元 801，具体用于通过预设算法，计算中心化后的第一训练集中的每个样本在正交基中的每个基向量上的分量值，得到多组第一编码数据。

在一种可能的实施方式中，预设算法包括内积运算或者稀疏编码运算。

在一种可能的实施方式中，获取单元 802，具体用于保留每组第一编码数据中的主分量，将每组第一编码数据中除主分量外的其他分量替换为预设的值，得到多组第二编码数据，其中，每组第一编码数据的主分量组成至少一组第二编码数据。

在一种可能的实施方式中，预设的值包括 0 或者预设的噪声向量，预设的噪声向量包括高斯噪声或者均匀分布噪声。

在一种可能的实施方式中，预设的噪声向量与替换的分量对应的方差成比例。

请参阅图 9，本申请提供的另一种训练集处理装置的结构示意图，如下所述。

该训练集处理装置可以包括处理器 901 和存储器 902。该处理器 901 和存储器 902 通过线路互联。其中，存储器 902 中存储有程序指令和数据。

存储器 902 中存储了前述图 6 中的步骤对应的程序指令以及数据。

处理器 901 用于执行前述图 6 中任一实施例所示的训练集处理装置执行的方法步骤。

可选地，该训练集处理装置还可以包括收发器 903，用于接收或者发送数据。

本申请实施例中还提供一种计算机可读存储介质，该计算机可读存储介质中存储有用于生成车辆行驶速度的程序，当其在计算机上行驶时，使得计算机执行如前述图 6 所示实

施例描述的方法中的步骤。

可选地，前述的图 9 中所示的训练集处理装置为芯片。

本申请实施例还提供了一种训练集处理装置，该训练集处理装置也可以称为数字处理芯片或者芯片，芯片包括处理单元和通信接口，处理单元通过通信接口获取程序指令，程序指令被处理单元执行，处理单元用于执行前述图 6 中任一实施例所示的训练集处理装置执行的方法步骤。

本申请实施例还提供一种数字处理芯片。该数字处理芯片中集成了用于实现上述处理器 901，或者处理器 901 的功能的电路和一个或者多个接口。当该数字处理芯片中集成了存储器时，该数字处理芯片可以完成前述实施例中的任一个或多个实施例的方法步骤。当该数字处理芯片中未集成存储器时，可以通过通信接口与外置的存储器连接。该数字处理芯片根据外置的存储器中存储的程序代码来实现上述实施例中训练集处理装置执行的动作。

本申请实施例中还提供一种包括计算机程序产品，当其在计算机上行驶时，使得计算机执行如前述图 6 所示实施例描述的方法中训练集处理装置所执行的步骤。

本申请实施例提供的训练集处理装置可以为芯片，芯片包括：处理单元和通信单元，所述处理单元例如可以是处理器，所述通信单元例如可以是输入/输出接口、管脚或电路等。该处理单元可执行存储单元存储的计算机执行指令，以使服务器内的芯片执行上述图 6 所示实施例描述的训练集处理方法。可选地，所述存储单元为所述芯片内的存储单元，如寄存器、缓存等，所述存储单元还可以是所述无线接入设备端内的位于所述芯片外部的存储单元，如只读存储器（read-only memory，ROM)或可存储静态信息和指令的其他类型的静态存储设备，随机存取存储器（random access memory，RAM)等。

具体地，前述的处理单元或者处理器可以是中央处理器（central processing unit，CPU)、网络处理器（neural-network processing unit，NPU)、图形处理器（graphics processing unit，GPU)、数字信号处理器（digital signal processor，DSP)、专用集成电路（application specific integrated circuit，ASIC）或现场可编程逻辑门阵列（field programmable gate array，FPGA）或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者也可以是任何常规的处理器等。

示例性地，请参阅图 10，图 10 为本申请实施例提供的芯片的一种结构示意图，所述芯片可以表现为神经网络处理器 NPU 100，NPU 100 作为协处理器挂载到主 CPU（Host CPU）上，由 Host CPU 分配任务。NPU 的核心部分为运算电路 1003，通过控制器 1004 控制运算电路 1003 提取存储器中的矩阵数据并进行乘法运算。

在一些实现中，运算电路 1003 内部包括多个处理单元（process engine，PE)。在一些实现中，运算电路 1003 是二维脉动阵列。运算电路 1003 还可以是一维脉动阵列或者能够执行例如乘法和加法这样的数学运算的其它电子线路。在一些实现中，运算电路 1003 是通用的矩阵处理器。

举例来说，假设有输入矩阵 A，权重矩阵 B，输出矩阵 C。运算电路从权重存储器 1002 中取矩阵 B 相应的数据，并缓存在运算电路中每一个 PE 上。运算电路从输入存储器 1001

中取矩阵 A 数据与矩阵 B 进行矩阵运算，得到的矩阵的部分结果或最终结果，保存在累加器（accumulator）1008 中。

统一存储器 1006 用于存放输入数据以及输出数据。权重数据直接通过存储单元访问控制器（direct memory access controller，DMAC）1005，DMAC 被搬运到权重存储器 1002 中。输入数据也通过 DMAC 被搬运到统一存储器 1006 中。

总线接口单元（bus interface unit，BIU）1010，用于 AXI 总线与 DMAC 和取指存储器（instruction fetch buffer，IFB）1009 的交互。

总线接口单元 1010（bus interface unit，BIU），用于取指存储器 1009 从外部存储器获取指令，还用于存储单元访问控制器 1005 从外部存储器获取输入矩阵 A 或者权重矩阵 B 的原数据。

DMAC 主要用于将外部存储器 DDR 中的输入数据搬运到统一存储器 1006 或将权重数据搬运到权重存储器 1002 中或将输入数据数据搬运到输入存储器 1001 中。

向量计算单元 1007 包括多个运算处理单元，在需要的情况下，对运算电路的输出做进一步处理，如向量乘，向量加，指数运算，对数运算，大小比较等等。主要用于神经网络中非卷积/全连接层网络计算，如批归一化(batch normalization)，像素级求和，对特征平面进行上采样等。

在一些实现中，向量计算单元 1007 能将经处理的输出的向量存储到统一存储器 1006。例如，向量计算单元 1007 可以将线性函数和/或非线性函数应用到运算电路 1003 的输出，例如对卷积层提取的特征平面进行线性插值，再例如累加值的向量，用以生成激活值。在一些实现中，向量计算单元 1007 生成归一化的值、像素级求和的值，或二者均有。在一些实现中，处理过的输出的向量能够用作到运算电路 1003 的激活输入，例如用于在神经网络中的后续层中的使用。

控制器 1004 连接的取指存储器（instruction fetch buffer）1009，用于存储控制器 1004 使用的指令；

统一存储器 1006，输入存储器 1001，权重存储器 1002 以及取指存储器 1009 均为 On-Chip 存储器。外部存储器私有于该 NPU 硬件架构。

其中，循环神经网络中各层的运算可以由运算电路 1003 或向量计算单元 1007 执行。

其中，上述任一处提到的处理器，可以是一个通用中央处理器，微处理器，ASIC，或一个或多个用于控制上述图 6 的方法的程序执行的集成电路。

另外需说明的是，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。另外，本申请提供的装置实施例附图中，模块之间的连接关系表示它们之间具有通信连接，具体可以实现为一条或多条通信总线或信号线。

通过以上的实施方式的描述，所属领域的技术人员可以清楚地了解到本申请可借助软件加必需的通用硬件的方式来实现，当然也可以通过专用硬件包括专用集成电路、专用 CPU、

专用存储器、专用元器件等来实现。一般情况下，凡由计算机程序完成的功能都可以很容易地用相应的硬件来实现，而且，用来实现同一功能的具体硬件结构也可以是多种多样的，例如模拟电路、数字电路或专用电路等。但是，对本申请而言更多情况下软件程序实现是更佳的实施方式。基于这样的理解，本申请的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，该计算机软件产品存储在可读取的存储介质中，如计算机的软盘、U 盘、移动硬盘、只读存储器（read only memory，ROM）、随机存取存储器（random access memory，RAM）、磁碟或者光盘等，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，或者网络设备等）执行本申请各个实施例所述的方法。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。

所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时，全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一计算机可读存储介质传输，例如，所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例如同轴电缆、光纤、数字用户线（DSL））或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存储的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质，（例如，软盘、硬盘、磁带）、光介质（例如，DVD）、或者半导体介质（例如固态硬盘（solid state disk，SSD））等。

本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”、“第三”、“第四”等（如果存在）是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便这里描述的实施例能够以除了在这里图示或描述的内容以外的顺序实施。此外，术语“包括”和“具有”以及他们的任何变形，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、系统、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

最后应说明的是：以上，仅为本申请的具体实施方式，但本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本申请的保护范围之内。因此，本申请的保护范围应以权利要求的保护范围为准。

权 利 要 求

1、一种训练集处理方法，其特征在于，包括：

根据第一训练集获得所述第一训练集的基和所述基的离散序列，所述第一训练集中包括多个样本，所述基中包括至少一个基向量，所述离散序列包括与所述基中的每个基向量一一对应的离散值；

获取所述第一训练集中的每个样本在所述每个基向量上的分量值，以得到多组第一编码数据，每组第一编码数据对应一个样本；

获取所述多组第一编码数据中的每组第一编码数据中的主分量，得到多组第二编码数据，所述主分量的离散值高于预设离散值；

将所述多组第二编码数据映射至所述第一训练集的基，得到所述多组第二编码数据对应的样本，所述多组第二编码数据对应的样本组成第二训练集，所述第二训练集用于训练神经网络。

2、根据权利要求 1 所述的方法，其特征在于，所述第一训练集的基为正交基，所述根据第一训练集获得所述第一训练集的基和所述基的离散序列，包括：

对所述第一训练集进行主成分分析 PCA 处理，得到所述正交基和所述正交基的离散序列。

3、根据权利要求 2 所述的方法，其特征在于，所述对所述第一训练集进行主成分分析 PCA 处理，得到正交基和所述正交基的离散序列，包括：

对所述第一训练集进行中心化处理，得到中心化后的第一训练集，所述中心化后的第一训练集包括的数据的均值为 0；

对所述中心化后的第一训练集进行所述 PCA 处理，得到所述正交基，以及所述正交基的离散序列，所述离散序列包括所述中心化后的第一训练集在所述正交基中的方差组成的序列。

4、根据权利要求 3 所述的方法，其特征在于，所述获取所述第一训练集中的每个样本在所述每个基向量上的分量值，包括：

通过预设算法，计算所述中心化后的第一训练集中的每个样本在所述正交基中的每个基向量的分量值，得到所述多组第一编码数据。

5、根据权利要求 4 所述的方法，其特征在于，所述预设算法包括内积运算或者稀疏编码运算。

6、根据权利要求 1-5 中任一项所述的方法，其特征在于，所述获取所述每组第一编码数据中的主分量，得到多组第二编码数据，包括：

保留所述每组第一编码数据中的所述主分量，将所述每组第一编码数据中除所述主分量外的其他分量替换为预设的值，得到所述多组第二编码数据，其中，所述每组第一编码数据的主分量组成至少一组第二编码数据。

7、根据权利要求 6 所述的方法，其特征在于，所述预设的值包括 0 或者预设的噪声向量，所述预设的噪声向量包括高斯噪声或者均匀分布噪声。

8、根据权利要求 7 所述的方法，其特征在于，所述预设的噪声向量与被替换的分量的

离散值成比例。

9、一种训练集处理装置，其特征在于，包括：

分解单元，用于根据第一训练集获得所述第一训练集的基和所述基的离散序列，所述第一训练集中包括多个样本，所述基中包括至少一个基向量，所述离散程度包括与所述基中的每个基向量一一对应的离散值；

获取单元，用于获取所述第一训练集中的每个样本在所述每个基向量上的分量值，以得到多组第一编码数据，每组第一编码数据对应一个样本；

所述获取单元，还用于获取所述多组第一编码数据中的每组第一编码数据中的主分量，得到多组第二编码数据，所述主分量的离散值高于预设离散值，其中，所述每组第一编码数据的主分量组成至少一组第二编码数据；

映射单元，用于将所述多组第二编码数据映射至所述第一训练集的基，得到所述多组第二编码数据对应的样本，所述多组第二编码数据对应的样本组成第二训练集，所述第二训练集用于训练神经网络。

10、根据权利要求 9 所述的装置，其特征在于，所述基为正交基，

所述分解单元，具体用于对所述第一训练集进行主成分分析 PCA 处理，得到所述正交基和所述正交基的离散序列。

11、根据权利要求 10 所述的装置，其特征在于，所述分解单元，具体用于：

对所述第一训练集进行中心化处理，得到中心化后的第一训练集，所述中心化后的第一训练集包括的数据的均值为 0；

对所述中心化后的第一训练集进行所述 PCA 处理，得到所述正交基，以及所述正交基的离散序列，所述离散序列包括所述中心化后的第一训练集在所述正交基中的方差组成的序列。

12、根据权利要求 11 所述的装置，其特征在于，

所述分解单元，具体用于通过预设算法，计算所述中心化后的第一训练集中的每个样本在所述正交基中的每个基向量的分量值，得到所述多组第一编码数据。

13、根据权利要求 12 所述的装置，其特征在于，所述预设算法包括内积运算或者稀疏编码运算。

14、根据权利要求 9-13 中任一项所述的装置，其特征在于，

所述获取单元，具体用于保留所述每组第一编码数据中的所述主分量，将所述每组第一编码数据中除所述主分量外的其他分量替换为预设的值，得到所述多组第二编码数据。

15、根据权利要求 14 所述的装置，其特征在于，所述预设的值包括 0 或者预设的噪声向量，所述预设的噪声向量包括高斯噪声或者均匀分布噪声。

16、根据权利要求 15 所述的装置，其特征在于，所述预设的噪声向量与被替换的分量的离散值成比例。

17、一种训练集处理装置，其特征在于，包括处理器，所述处理器和存储器耦合，所述存储器存储有程序，当所述存储器存储的程序指令被所述处理器执行时实现权利要求 1

至 8 中任一项所述的方法。

18、一种计算机可读存储介质，包括程序，当其被处理单元所执行时，执行如权利要求 1 至 8 中任一项所述的方法。

19、一种训练集处理装置，其特征在于，包括处理单元和通信接口，所述处理单元通过所述通信接口获取程序指令，当所述程序指令被所述处理单元执行时实现权利要求 1 至 8 中任一项所述的方法。

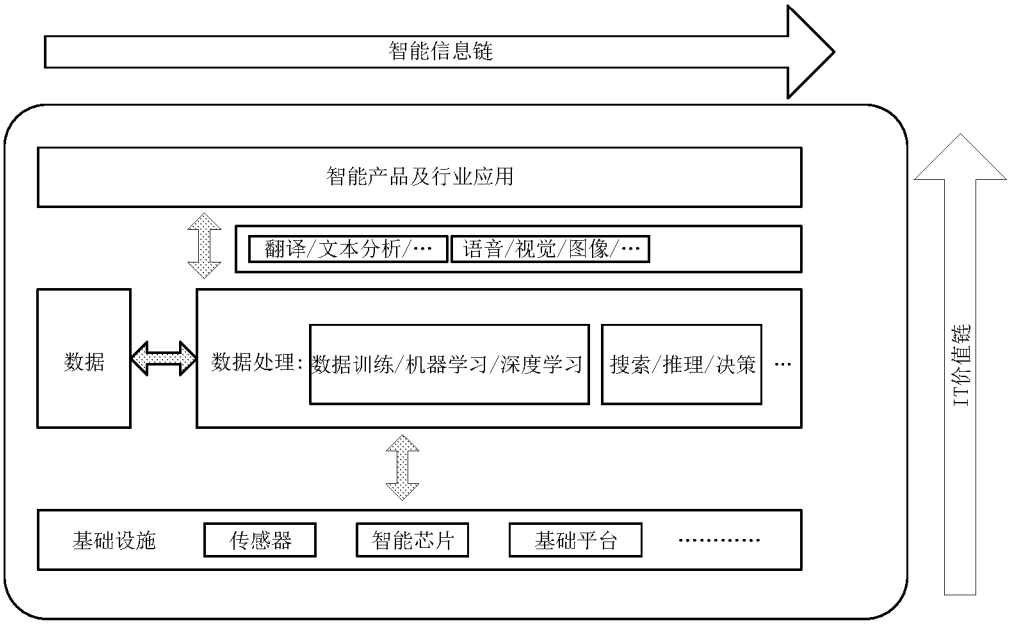

智能信息链
智能产品及行业应用
翻译/文本分析/…
语音/视觉/图像/…
数据
数据处理:
数据训练/机器学习/深度学习
搜索/推理/决策
…
基础设施
传感器
智能芯片
基础平台
……
IT价值链

图 1

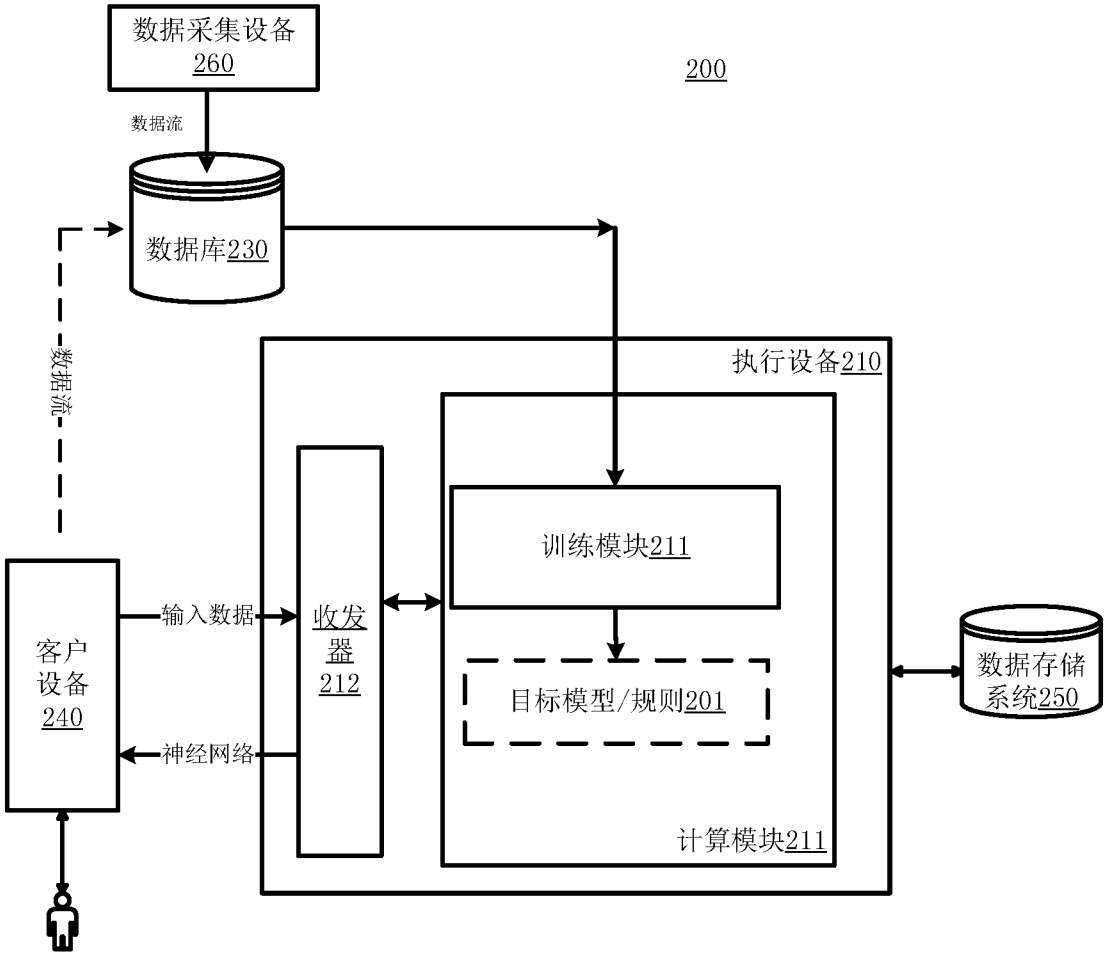

数据采集设备
260
200
数据流
数据库230
数据流
执行设备210
训练模块211
客户
设备
240
输入数据
收发
器
212
神经网络
目标模型/规则201
数据存储
系统250
计算模块211

图 2

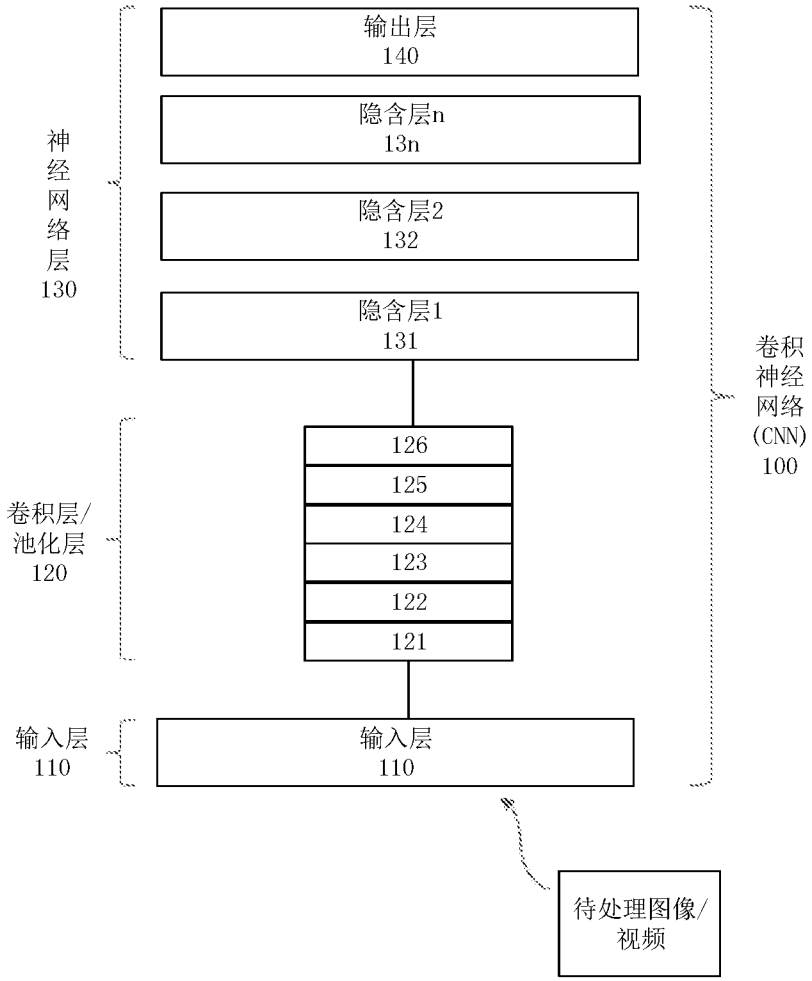

输出层
140
隐含层n
13n
隐含层2
132
隐含层1
131
神经网络层
130
126
125
124
123
122
121
卷积层/池化层
120
输入层
110
输入层
110
卷积神经网络(CNN)
100
待处理图像/视频

图 3

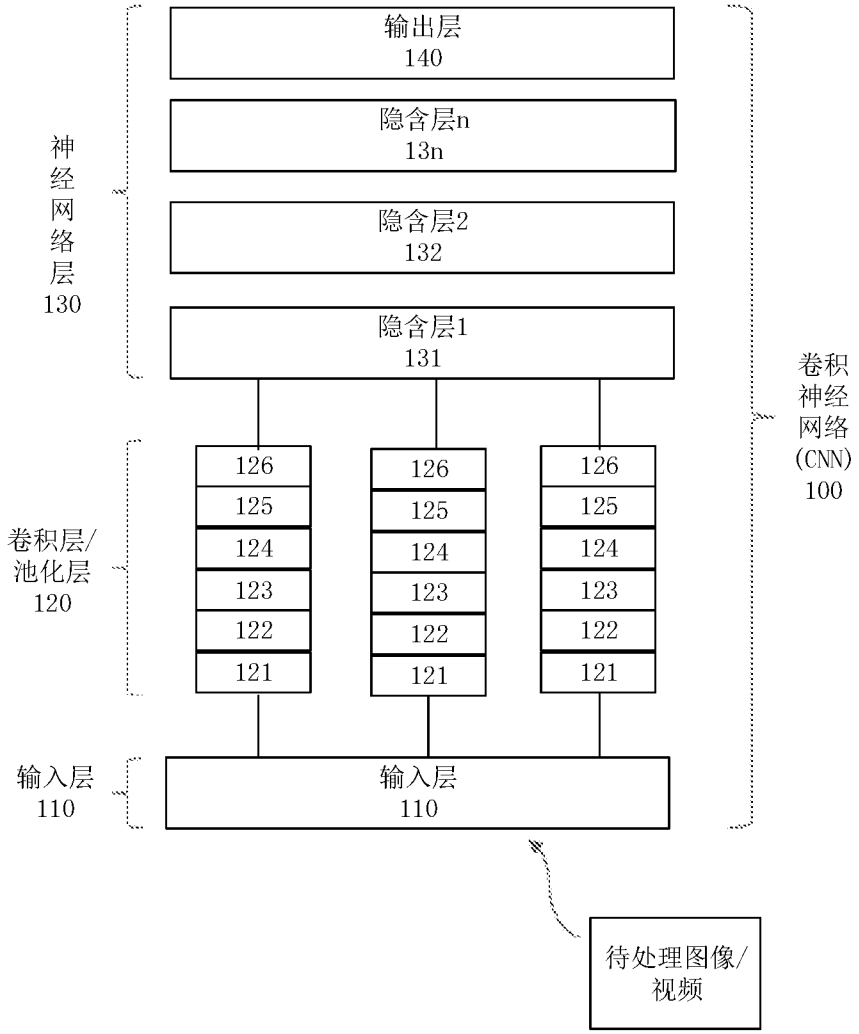

输出层
140
隐含层n
13n
隐含层2
132
隐含层1
131
神经网络层
130
126
125
124
123
122
121
卷积层/池化层
120
输入层
110
卷积神经网络(CNN)
100
待处理图像/视频

图 4

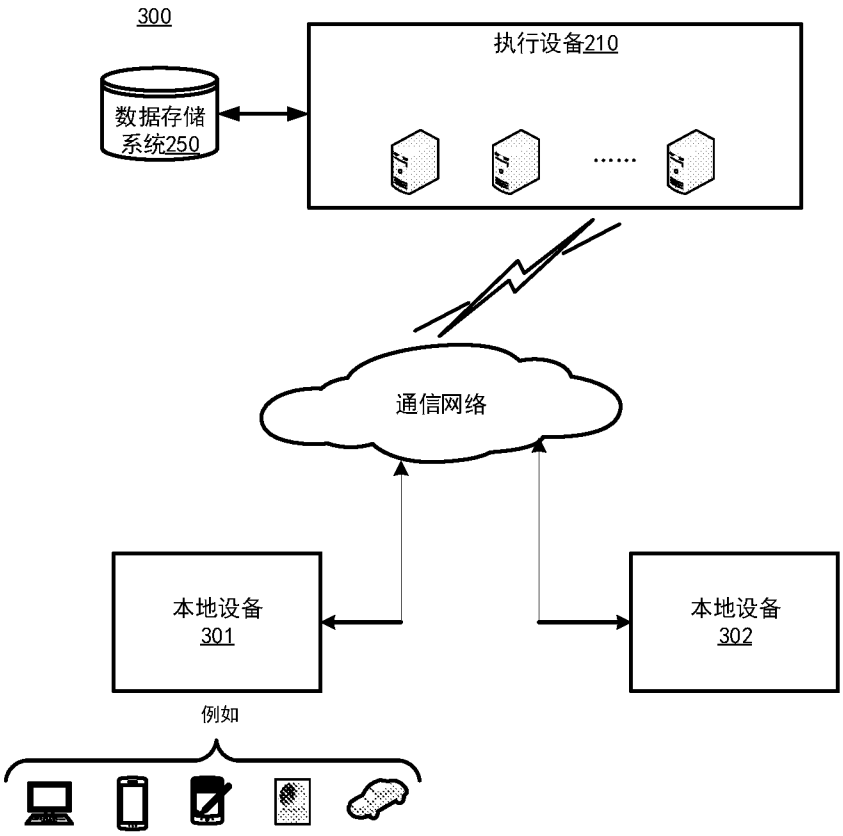


图 5

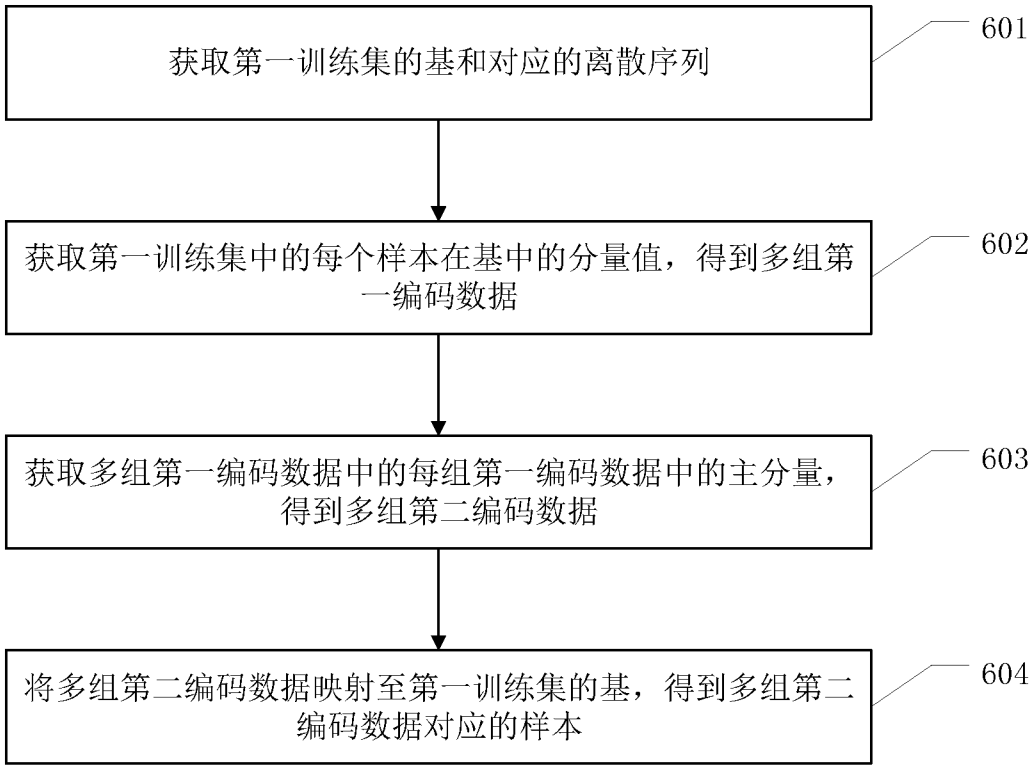

获取第一训练集的基和对应的离散序列
601
获取第一训练集中的每个样本在基中的分量值，得到多组第一编码数据
602
获取多组第一编码数据中的每组第一编码数据中的主分量，得到多组第二编码数据
603
将多组第二编码数据映射至第一训练集的基，得到多组第二编码数据对应的样本
604

图 6

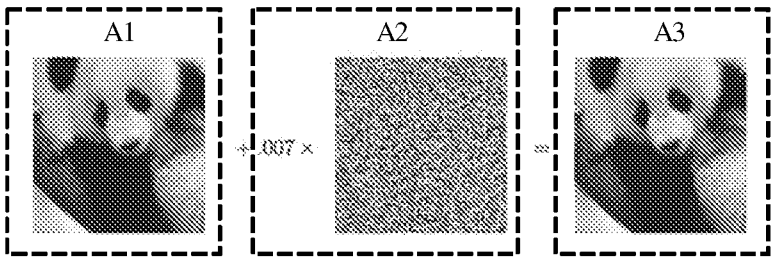

A1
A2
A3
+0.007×
≈

图 7

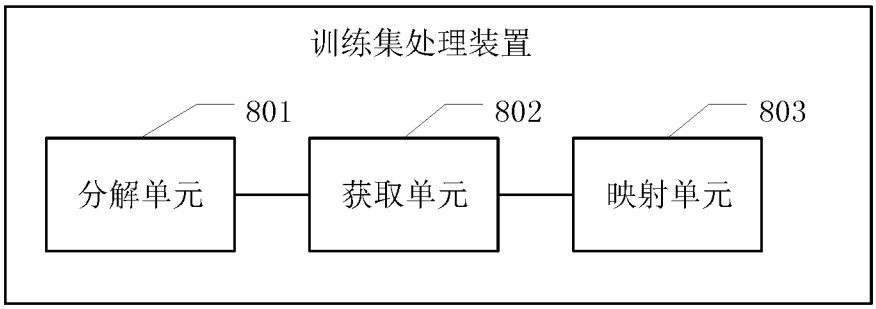

训练集处理装置
801
802
803
分解单元
获取单元
映射单元

图 8

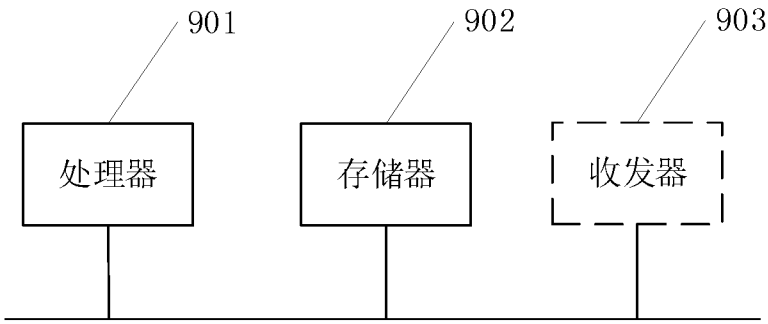

901
902
903
处理器
存储器
收发器

图 9

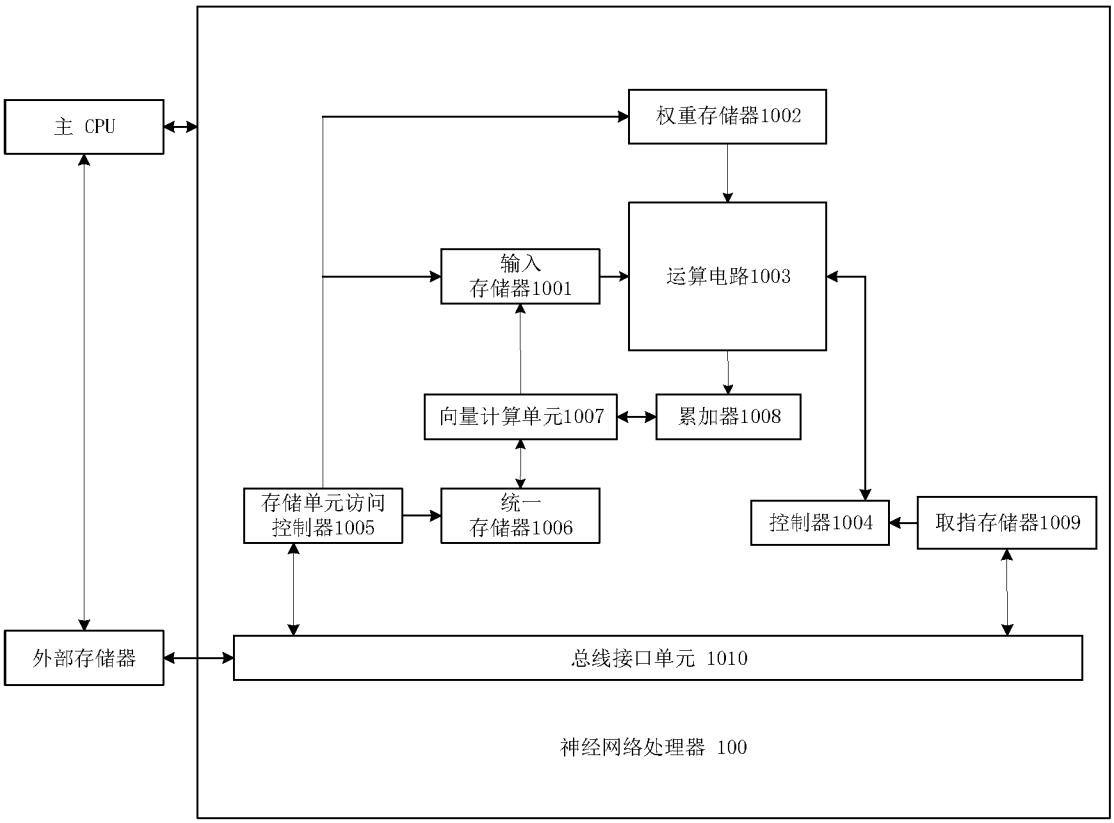

主 CPU
权重存储器1002
输入
存储器1001
运算电路1003
向量计算单元1007
累加器1008
存储单元访问
控制器1005
统一
存储器1006
控制器1004
取指存储器1009
外部存储器
总线接口单元 1010
神经网络处理器 100

图 10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2021/106758

A. CLASSIFICATION OF SUBJECT MATTER

G06N 7/00(2006.01)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)

G06N

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)

CNPAT, IEEE, CNKI, WPI, EPODOC: 训练, 样本, 数据集, 基向量, 主成分, 主向量, PCA, 投影, 方差, train, sample, dataset, basis vector, principal, projection, variance

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	(non-official translation: Some hints on Machine Learning Algorithms). "(non-official translation: Summary of Principal Components Analysis (PCA))" *https://tianchi.aliyun.com/forum/postDetail?postId=79395,* 22 October 2019 (2019-10-22), full text on the webpage	1-5, 9-13, 17-19
X	(non-official translation: Penguin Media Content Platform – AI Classroom). "(non-official translation: Face Recognition Based on PCA for Dimensionality Reduction and BP Neural Network)" *https://cloud.tencent.com/developer/news/193912,* 27 April 2018 (2018-04-27), full text on the webpage	1-5, 9-13, 17-19
A	Indranil Chakraborty et al. "Constructing Energy-efficient Mixed-precision Neural Networks through Principal Component Analysis for Edge Intelligence" *https://arxiv.org/pdf/1906.01493.pdf,* 03 December 2019 (2019-12-03), entire document	1-19
A	CN 108896499 A (XI'AN UNIVERSITY OF ARCHITECTURE AND TECHNOLOGY) 27 November 2018 (2018-11-27) entire document	1-19

☐ 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
"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
30 September 2021	**18 October 2021**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088** **China**	
Facsimile No. **(86-10)62019451**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2021/106758

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 108896499 A	27 November 2018	None	

国际检索报告

国际申请号
PCT/CN2021/106758

A. 主题的分类

G06N 7/00(2006.01)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)

G06N

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）)

CNPAT,IEEE,CNKI,WPI,EPODOC:训练，样本，数据集，基向量，主成分，主向量，PCA，投影，方差，train，sample，dataset，basis vector，principal，projection，variance

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
X	机器学习算法那些事. "主成分分析（PCA）原理总结" https://tianchi.aliyun.com/forum/postDetail?postId=79395, 2019年 10月 22日 (2019-10-22), 网页正文	1-5，9-13，17-19
X	企鹅号-AI讲堂. "基于PCA降维和BP神经网络的人脸识别" https://cloud.tencent.com/developer/news/193912, 2018年 4月 27日 (2018-04-27), 网页正文	1-5，9-13，17-19
A	Indranil Chakraborty 等. "Constructing Energy-efficient Mixed-precision Neural Networks through Principal Component Analysis for Edge Intelligence" https://arxiv.org/pdf/1906.01493.pdf, 2019年 12月 3日 (2019-12-03), 全文	1-19
A	CN 108896499 A (西安建筑科技大学) 2018年 11月 27日 (2018-11-27) 全文	1-19

☐ 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

“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
2021年 9月 30日	2021年 10月 18日

ISA/CN的名称和邮寄地址	受权官员
中国国家知识产权局(ISA/CN) 中国 北京市海淀区蓟门桥西土城路6号 100088	王平
传真号 (86-10)62019451	电话号码 86-(10)-53961372

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2021/106758

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 108896499 A	2018年 11月 27日	无	