(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
2020 年 8 月 20 日 (20.08.2020)

WIPO | PCT

(10) 国际公布号
WO 2020/164271 A1

(51) 国际专利分类号:
G06K 9/62 (2006.01) *G06N 3/04* (2006.01)

(21) 国际申请号: PCT/CN2019/117863

(22) 国际申请日: 2019 年 11 月 13 日 (13.11.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910113187.9 2019年2月13日 (13.02.2019) CN

(71) 申请人: 平安科技（深圳）有限公司 (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN]; 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

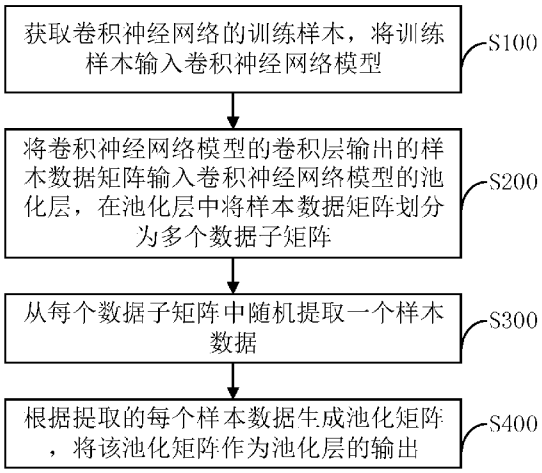
(72) 发明人: 房树明 (FANG, Shuming); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。 王健宗 (WANG, Jianzong); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(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, 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, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK,

(54) **Title:** POOLING METHOD AND DEVICE FOR CONVOLUTIONAL NEURAL NETWORK, STORAGE MEDIUM AND COMPUTER DEVICE

(54) **发明名称**: 卷积神经网络的池化方法、装置及存储介质、计算机设备



S100 OBTAIN A CONVOLUTIONAL NEURAL NETWORK TRAINING SAMPLE, AND INPUTTING THE TRAINING SAMPLE INTO A CONVOLUTIONAL NEURAL NETWORK MODEL

S200 INPUT A SAMPLE DATA MATRIX OUTPUTTED BY A CONVOLUTIONAL LAYER OF THE CONVOLUTIONAL NEURAL NETWORK MODEL INTO A POOLING LAYER OF THE CONVOLUTIONAL NEURAL NETWORK MODEL, AND DIVIDE THE SAMPLE DATA MATRIX INTO A PLURALITY OF DATA SUB-MATRICES IN THE POOLING LAYER

S300 RANDOMLY EXTRACT SAMPLE DATA FROM EACH DATA SUBMATRIX

S400 GENERATE A POOLING MATRIX ACCORDING TO EACH EXTRACTED SAMPLE DATA, AND USE THE POOLING MATRIX AS THE OUTPUT OF THE POOLING LAYER

图 4

(57) **Abstract:** A pooling method and device for a convolutional neural network, a storage medium and a computer device, the method comprising: obtaining a convolutional neural network training sample, and inputting the training sample into a convolutional neural network model (S100); inputting a sample data matrix outputted by a convolutional layer of the convolutional neural network model into a pooling layer of the convolutional neural network model, and dividing the sample data matrix into a plurality of data sub-matrices in the pooling layer (S200); randomly extracting sample data from each data submatrix (S300); and generating a pooling matrix according to said extracted sample data, and using the pooling matrix as the output of the pooling layer (S400). According to the present method, pooled data samples may be expanded, and the input information of the convolutional neural network is enriched.

WO 2020/164271 A1

[见续页]

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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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：一种卷积神经网络的池化方法、装置及存储介质、计算机设备，所述方法包括：获取卷积神经网络的训练样本，将训练样本输入卷积神经网络模型（S100）；将卷积神经网络模型的卷积层输出的样本数据矩阵输入卷积神经网络模型的池化层，在池化层中将样本数据矩阵划分为多个数据子矩阵（S200）；从每个数据子矩阵中随机提取一个样本数据（S300）；根据提取的每个样本数据生成池化矩阵，将该池化矩阵作为池化层的输出（S400）。上述方法可以扩充池化后的数据样本，丰富了卷积神经网络的输入信息。

卷积神经网络的池化方法、装置及存储介质、计算机设备

本申请要求于 2019 年 02 月 13 日提交中国专利局、申请号为201910113187.9、申请名称为“卷积神经网络的池化方法、装置及存储介质、计算机设备”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

本申请涉及卷积神经网络模型技术领域，具体而言，本申请涉及一种卷积神经网络的池化方法、装置及存储介质、计算机设备。

背景技术

卷积神经网络（Convolutional Neural Network，CNN）是一种前馈神经网络。卷积神经网络采用大量的人工神经元，人工神经元可以响应一部分覆盖范围内的周围单元，其常用于大型图像处理。卷积神经网络采用包括卷积层(convolutional layer)和池化层(pooling layer)。

目前通用的池化函数采用最大池化(max-pooling)或平均池化(mean-pooling)。最大池化的缺点是，每次池化都会丢失最大值位置以外的信息；平均池化的缺点是，不能体现出一些离平均值较远但比较重要的信息。因此，当前卷积神经网络采用的池化方式，减少了卷积神经网络的输入信息，导致卷积神经网络模型的输出结果准确性降低。

发明内容

本申请提出一种卷积神经网络的池化方法、装置及存储介质、计算机设备，以扩充池化后的数据样本，丰富了卷积神经网络的输入信息。

本申请提供以下方案：

一种卷积神经网络的池化方法，包括：获取卷积神经网络的训练样本，将训练样本输入卷积神经网络模型；将卷积神经网络模型的卷积层输出的样本数据矩阵输入卷积神经网络模型的池化层，在池化层中将样本数据矩阵划分为多个数据子矩阵；从每个数据子矩阵中随机提取一个样本数据；根据提取的每个样本数据生成池化矩阵，将该池化矩阵作为池化层的输出。

一种卷积神经网络的池化装置，包括：获取模块，用于获取卷积神经网络的训练样本，将训练样本输入卷积神经网络模型；划分模块，用于将卷积神经网络模型的卷积层输出的样本数据矩阵输入卷积神经网络模型的池化层，在池化层中将样本数据矩阵划分为多个数据子矩阵；提取模块，用于从每个数据子矩阵中随机提取一个样本数据；生成模块，用于根据提取的每个样本数据生成池化矩阵，将该池化矩阵作为池化层的输出。

一种计算机非易失性存储介质，其上存储有计算机程序；所述计算机程序适于由处理器加载并执行上述任一实施例所述的卷积神经网络的池化方法。

一种计算机设备，其包括：一个或多个处理器；存储器；一个或多个应用程序，其中所述一个或多个应用程序被存储在所述存储器中并被配置为由所述一个或多个处理器执行，所述一个或多个应用程序配置用于执行根据上述任一实施例所述的卷积神经网络的池化方法。

上述实施例提供的卷积神经网络的池化方法，通过在池化层中对样本数据的每个数据子矩阵进行数据随机采集后池化，一个 M*M 的样本数据矩阵可得到 M*M*M*M，与传统的最大池化或平均池化的方式相比，扩充了池化后的数据样本，丰富了卷积神经网络的输入信息。

本申请附加的方面和优点将在下面的描述中部分给出，这些将从下面的描述中变得明显，或通过本申请的实践了解到。

附图说明

本申请上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解，其中：

图 1 为本申请提供的卷积神经网络模型的一实施例中的内部结构示意图；

图 2 为本申请提供的最大池化方法的一实施例中的示意图；

图 3 为本申请提供的平均池化方法的一实施例中的示意图；

图 4 为本申请提供的一种卷积神经网络的池化方法的一实施例中的方法流程图；

图 5 为本申请提供的步骤 S200 的一实施例中的方法流程图；

图 6 为本申请提供的随机池化方法的一实施例中的示意图；

图 7 为本申请提供的一种卷积神经网络的池化装置的一实施例中的结构框图；

图 8 为本申请提供的一种计算机设备的一实施例中的结构示意图。

具体实施方式

下面详细描述本申请的实施例，所述实施例的示例在附图中示出。

本申请提供一种卷积神经网络的池化方法，应用于卷积神经网络的池化层中。以下先对本申请所述的卷积神经网络的池化方法的背景技术做先导性说明：

参见图 1 所示，卷积神经网络在输入层到输出层之间，包含有两个卷积层、两个池化层、一个全连接隐层。一般来说，池化层的输入来源于上一个卷积层，主要提供了很强的鲁棒性。在减少了深度学习参数的数量的同时，还能防止过拟合现象的发生。参数数量的减少意味着卷积神经网络训练的速度会更快，训练后的模型文件更小，预测样本时候推理速度更快。因而池化函数对于卷积神经网络有着重要的意义。

卷积神经网络中池化函数一般采用最大池化(max-pooling)和平均池化(mean-pooling)方法。对于最大池化方法，如图 2 所示。一个 4 乘 4 的矩阵经过 2 乘 2 的最大池函数池化后，变为一个 2 乘以 2 的矩阵。池化计算步骤为：将 4 乘 4 的矩阵拆分为 4 个 2 乘 2 的子矩阵，分别求取每个子矩阵的最大值，构成一个新的矩阵，即为最大池化。对于平均池化方法，如图 3 所示。一个 4 乘 4 的矩阵经过 2 乘 2 的平均池函数池化后，变为一个 2 乘以 2 的矩阵。池化计算步骤：将 4 乘 4 的矩阵拆分为 4 个 2 乘 2 的子矩阵，分别求取每个子矩阵的平均值，构成一个新的矩阵，即为平均池化。然而，采用最大池化方法，每次池化都会丢失最大值位置以外的信息。采用平均池化方法，不能体现出一些离平均值较远但比较重要的信息。

本申请提供一种卷积神经网络的池化方法，以扩充池化后的数据样本，丰富卷积神经网络的输入信息。在一实施例中，如图 4 所示，该卷积神经网络的池化方法，包括以下步骤：

S100，获取卷积神经网络的训练样本，将训练样本输入卷积神经网络模型。

在本实施例中，在进行卷积神经网络结构的数据运算之前，服务器先获取

特征数据的训练样本。该训练样本为服务器通过卷积神经网络进行计算的目标特征。进一步地，将该训练样本输入卷积神经网络模型中。其中，卷积神经网络模型包括卷积层、采样层、激活层、池化层和全连接层。卷积层用于提取输入数据的空间特征。卷积层可以包括多个卷积核，以提取输入数据的多个空间特征。激活层可采用非线性激活函数。池化层用于避免卷积过程的过拟合现象。全连接层用于网络中神经元与神经元之间的邻接相连，可通过 softmax 函数计算后输出，得到不同的概率值。

在一实施例中，所述训练样本为图像样本数据。步骤 S100，包括：获取所述图像样本数据作为所述卷积神经网络模型的训练样本。

在该实施例中，对于模型训练的样本数据可以是图像样本数据，模型输出则为图像样本数据的分类。进一步地，所述获取所述图像样本数据作为所述卷积神经网络模型的训练样本，包括：获取所述图像样本数据中的图片样本数据作为所述卷积神经网络模型的训练样本。

其中，所述将卷积神经网络模型的卷积层输出的样本数据矩阵输入卷积神经网络模型的池化层，在池化层中将样本数据矩阵划分为多个数据子矩阵，从每个数据子矩阵中随机提取一个样本数据，根据提取的每个样本数据生成池化矩阵，将该池化矩阵作为池化层的输出，包括：将卷积神经网络模型的第一卷积层输出的图片样本数据矩阵经过激活后，输入卷积神经网络模型的第一池化层，在第一池化层中将样本数据矩阵划分为多个数据子矩阵，从每个数据子矩阵中随机提取一个图片样本数据，根据提取的图片样本数据生成池化矩阵进行第一池化层输出后，再将所述第一池化层输出的数据输入第二卷积层进行卷积，对所述第二卷积层输出的数据进行随机激活，并将随机激活后的数据输入第二池化层进行池化训练，将该池化训练结果作为系统池化层的输出。

S200，将卷积神经网络模型的卷积层输出的样本数据矩阵输入卷积神经网络模型的池化层，在池化层中将样本数据矩阵划分为多个数据子矩阵。

在本实施例中，服务器将训练样本输入到卷积神经网络模型时，首先训练样本输入到卷积神经网络模型的卷积层。卷积层对训练样本进行卷积训练，提取输入的训练样本的不同空间特征，输出包含多个特征的样本数据矩阵。通常在卷积层之后得到维度较大的特征的样本数据矩阵。此时，将特征切成几个区

域，取其最大值或平均值，得到新的、维度较小的特征。也即是卷积神经网络的池化运算。在进行卷积神经网络的池化层运算时，将卷积层输出的样本数据矩阵划分为子矩阵。具体可以是，设置预设大小（小于样本数据矩阵的大小）的卷积核。卷积核窗口为长和宽值大小相同的窗口。当然，也不排除长和宽值大小不相同的窗口的卷积核。通过卷积核可将样本数据矩阵划分为多个子矩阵。

在一实施例中，所述多个数据子矩阵包括行数与列数相同的数据子矩阵。如图 5 所示，步骤 S200 中，在池化层中将样本数据矩阵划分为多个数据子矩阵，包括：

S210，获取所述样本数据矩阵的行数和列数。

S230，根据所述样本数据矩阵的行数和列数，在池化层中将所述样本数据矩阵划分为多个行数与列数相同的数据子矩阵。

卷积神经网络中的卷积核一般取小的正方形矩阵。主流的深度学习框架一般支持的也是长宽一样的卷积核。例如，可采用 3 乘 3 的卷积核，或者 2 乘 2 的卷积核。在该实施例中，获取样本数据矩阵的行数 M 和列数 N，根据行数 M 和列数 N 将样本数据矩阵划分为多个行数与列数相同的数据子矩阵。其中，行数 M 和列数 N 可以相同。

在其他实施例中，多个数据子矩阵包括行数与列数相同的第一数据子矩阵以及行数与列数不相同的第二数据子矩阵。也即是，进行样本数据矩阵划分时，可同时划分出多个行数与列数相同的第一数据子矩阵，以及多个行数与列数不相同的第二数据子矩阵。具体地，步骤 S200 中，在池化层中将样本数据矩阵划分为多个数据子矩阵，包括：获取样本数据矩阵的行数和列数；根据样本数据矩阵的行数和列数在池化层中将样本数据矩阵划分为多个所述第一数据子矩阵以及多个所述第二数据子矩阵。

S300，从每个数据子矩阵中随机提取一个样本数据。

在本实施例中，服务器在卷积神经网络模型中进行池化运算时，从每个数据子矩阵中随机提取一个样本数据，从而可根据每个数据子矩阵中提取的数据构成新的矩阵。此处，每次随机提取的数据可以是数据子矩阵中任何一个数据值，每个数据值还可被重复提取。

在一实施例中，所述数据子矩阵为多行及多列的矩阵。步骤 S300，包括：

从所述每个数据子矩阵的多行中随机选取任一行数据；从该行数据中，随机选取任一列对应的数据作为所述样本数据；或，从所述每个数据子矩阵的多列中随机选取任一列数据；从该列数据中，随机选取任一行对应的数据作为所述样本数据。

在该实施例中，服务器先从多行及多列的数据子矩阵中随机筛选出任意一行的数据，再从该行数据中筛选任意一列，该列对应的数据则为所述样本数据。或者，服务器先从多行及多列的数据子矩阵中随机筛选出任意一列的数据，再从该列数据中筛选任意一行，该行对应的数据则为所述样本数据。

在一实施例中，所述数据子矩阵为多行及多列的矩阵。步骤S300，包括：从所述每个数据子矩阵的多行及多列中，随机选取任一行和任一列对应的数据作为所述样本数据。在该实施例中，服务器随机选出任意行值和列值，该行值和列值对应的数据即为所述样本数据。如服务器随机选取出行值和列值为（4、3），则（4、3）对应数据子矩阵中的数据即为所述样本数据。

S400，根据提取的每个样本数据生成池化矩阵，将该池化矩阵作为池化层的输出。

在本实施例中，可根据每个数据子矩阵中提取的样本数据生成新的矩阵，也即是池化矩阵，将该池化矩阵作为池化层的输出，以输出到卷积神经网络模型的全连接隐层。在一具体实施方式中，如图6所示，一个4乘4的样本数据矩阵经过划分为2乘2数据子矩阵的随机池函数池化后，变为多个2乘以2的矩阵。具体地，随机池化的计算步骤为：将4乘4的矩阵拆分为4个2乘2的子矩阵，分别在每个子矩阵的里面随机取一个值，构成一个新的矩阵，即为随机池化矩阵。通过采用随机池化，一个4乘4的矩阵将产生4*4*4*4 = 1024种不同的池化结果，相当于把样本扩充了1024倍，极大地丰富了卷积神经网络的输入信息。

上述实施例提供的卷积神经网络的池化方法，将训练样本输入卷积神经网络模型中，经过卷积神经网络模型的卷积层进行卷积处理后，将卷积层输出的样本数据矩阵输入卷积神经网络模型的池化层。在池化层中，将样本数据矩阵划分为多个数据子矩阵，并且从每个数据子矩阵随机提取一个数据样本以生成池化矩阵。因此，通过在池化层中对样本数据的每个数据子矩阵进行数据随机

采集后池化，一个 M*M 的样本数据矩阵可得到 M*M*M*M，与传统的最大池化或平均池化的方式相比，扩充了池化后的数据样本，丰富了卷积神经网络的输入信息。

在一实施例中，上述卷积神经网络的池化方法可用于图像分类中。具体地，应用该卷积神经网络的池化方法的图像分类方法包括：获取目标图像的图像数据；将所述图像数据输入卷积神经网络模型，得到所述目标图像的待分类结果数据；所述卷积神经网络模型用于对所述图像数据进行图像特征类别分析，输出待分类结果数据；其中，所述卷积神经网络模型中的池化层，用于将所述卷积神经网络模型中的卷积层输出的图像数据划分为多个数据矩阵，并从每个数据子矩阵中随机提取一个图像样本数据，根据提取的图像样本数据生成池化矩阵，将该池化矩阵作为池化输出；根据所述待分类结果数据，对目标图像进行分类。

进一步地，所述卷积神经网络模型还包括第一激活单元和第二激活单元；所述卷积层包括第一卷积层和第二卷积层；所述池化层包括第一池化层和第二池化层；其中，所述第一池化层和所述第二池化层均用于将所述卷积神经网络模型中的卷积层输出的图像数据划分为多个数据矩阵，并从每个数据子矩阵中随机提取一个图像样本数据，根据提取的图像样本数据生成池化矩阵，将该池化矩阵作为池化输出；所述第一激活单元用于将所述第一卷积层输出的图像数据进行非随机激活后，将结果值输入到所述第一池化层；所述第一池化层用于将输入的图像数据进行池化后输出到所述第二卷积层；所述第二卷积层用于将输入的图像数据进行卷积后输出到所述第二激活单元；所述第二激活单元用于将输入的图像数据进行随机激活后输出到所述卷积神经网络模型的全连接层，所述全连接层用于输出所述待分类结果数据。

具体地，一个典型的卷积神经网络有卷积层、激活层、池化层、全连接层，一个三层的卷积神经网络可以表示为：输入图片->卷积->激活->池化->卷积->激活->池化->卷积->激活->池化->全连接层->待分类输出。如果将随机池化应用到该卷积神经网络，网络的结构变成：输入图片->卷积->激活->随机池化->卷积->随机激活->池化->卷积->随机激活->池化->全连接层->待分类输出。

本申请还提供一种卷积神经网络的池化装置。在一实施例中，如图 7 所示，

该卷积神经网络的池化装置包括获取模块10、划分模块20、提取模块30和生成模块40。

获取模块10用于获取卷积神经网络的训练样本，将训练样本输入卷积神经网络模型。在本实施例中，在进行卷积神经网络结构的数据运算之前，服务器先获取特征数据的训练样本。该训练样本为服务器通过卷积神经网络进行计算的目标特征。进一步地，将该训练样本输入卷积神经网络模型中。其中，卷积神经网络模型包括卷积层、采样层、激活层、池化层和全连接层。卷积层用于提取输入数据的空间特征。卷积层可以包括多个卷积核，以提取输入数据的多个空间特征。激活层可采用非线性激活函数。池化层用于避免卷积过程的过拟合现象。全连接层用于网络中神经元与神经元之间的邻接相连，可通过softmax函数计算后输出，得到不同的概率值。

划分模块20用于将卷积神经网络模型的卷积层输出的样本数据矩阵输入卷积神经网络模型的池化层，在池化层中将样本数据矩阵划分为多个数据子矩阵。在本实施例中，服务器将训练样本输入到卷积神经网络模型时，首先训练样本输入到卷积神经网络模型的卷积层。卷积层对训练样本进行卷积训练，提取输入的训练样本的不同空间特征，输出包含多个特征的样本数据矩阵。通常在卷积层之后得到维度较大的特征的样本数据矩阵。此时，将特征切成几个区域，取其最大值或平均值，得到新的、维度较小的特征。也即是卷积神经网络的池化运算。在进行卷积神经网络的池化层运算时，将卷积层输出的样本数据矩阵划分为子矩阵。具体可以是，设置预设大小（小于样本数据矩阵的大小）的卷积核。卷积核窗口为长和宽值大小相同的窗口。当然，也不排除长和宽值大小不相同的窗口的卷积核。通过卷积核可将样本数据矩阵划分为多个子矩阵。

提取模块30用于从每个数据子矩阵中随机提取一个样本数据。在本实施例中，服务器在卷积神经网络模型中进行池化运算时，从每个数据子矩阵中随机提取一个样本数据，从而可根据每个数据子矩阵中提取的数据构成新的矩阵。此处，每次随机提取的数据可以是数据子矩阵中任何一个数据值，每个数据值还可被重复提取。

生成模块40用于根据提取的每个样本数据生成池化矩阵，将该池化矩阵作为池化层的输出。在本实施例中，可根据每个数据子矩阵中提取的样本数据

生成新的矩阵，也即是池化矩阵，将该池化矩阵作为池化层的输出，以输出到卷积神经网络模型的全连接隐层。

在其他实施例中，本申请提供的卷积神经网络的池化装置中的各个模块还用于执行本申请所述的卷积神经网络的池化方法中，对应各个步骤执行的操作，在此不再做详细的说明。

本申请还提供一种存储介质。该存储介质上存储有计算机程序；所述计算机程序被处理器执行时，实现上述任一实施例所述的卷积神经网络的池化方法。该存储介质可以是存储器。例如，内存储器或外存储器，或者包括内存储器和外存储器两者。内存储器可以包括只读存储器(ROM)、可编程 ROM(PROM)、电可编程 ROM(EPROM)、电可擦写可编程 ROM(EEPROM)、快闪存储器、或者随机存储器。外存储器可以包括硬盘、软盘、ZIP 盘、U 盘、磁带等。本申请所公开的存储介质包括但不限于这些类型的存储器。本申请所公开的存储器只作为例子而非作为限定。

本申请还提供一种计算机设备。一种计算机设备包括：一个或多个处理器；存储器；一个或多个应用程序。其中所述一个或多个应用程序被存储在所述存储器中并被配置为由所述一个或多个处理器执行，所述一个或多个应用程序配置用于执行上述任一实施例所述的卷积神经网络的池化方法。

图 8 为本申请一实施例中的计算机设备的结构示意图。本实施例所述计算机设备可以是服务器、个人计算机以及网络设备。如图 8 所示，设备包括处理器 803、存储器 805、输入单元 807 以及显示单元 809 等器件。本领域技术人员可以理解，图 8 示出的设备结构器件并不构成对所有设备的限定，可以包括比图示更多或更少的部件，或者组合某些部件。存储器 805 可用于存储应用程序 801 以及各功能模块，处理器 803 运行存储在存储器 805 的应用程序 801，从而执行设备的各种功能应用以及数据处理。存储器可以是内存储器或外存储器，或者包括内存储器和外存储器两者。内存储器可以包括只读存储器(ROM)、可编程 ROM(PROM)、电可编程 ROM(EPROM)、电可擦写可编程 ROM(EEPROM)、快闪存储器、或者随机存储器。外存储器可以包括硬盘、软盘、ZIP 盘、U 盘、磁带等。本申请所公开的存储器包括但不限于这些类型的存储器。本申请所公开的存储器只作为例子而非作为限定。

输入单元807用于接收信号的输入，以及接收用户输入的关键字。输入单元807可包括触控面板以及其它输入设备。触控面板可收集用户在其上或附近的触摸操作（比如用户使用手指、触笔等任何适合的物体或附件在触控面板上或在触控面板附近的操作），并根据预先设定的程序驱动相应的连接装置；其它输入设备可以包括但不限于物理键盘、功能键（比如播放控制按键、开关按键等）、轨迹球、鼠标、操作杆等中的一种或多种。显示单元809可用于显示用户输入的信息或提供给用户的信息以及计算机设备的各种菜单。显示单元809可采用液晶显示器、有机发光二极管等形式。处理器803是计算机设备的控制中心，利用各种接口和线路连接整个电脑的各个部分，通过运行或执行存储在存储器805内的软件程序和/或模块，以及调用存储在存储器内的数据，执行各种功能和处理数据。

在一实施方式中，设备包括一个或多个处理器803，以及一个或多个存储器805，一个或多个应用程序801。其中所述一个或多个应用程序801被存储在存储器805中并被配置为由所述一个或多个处理器803执行，所述一个或多个应用程序801配置用于执行以上实施例所述的卷积神经网络的池化方法。

此外，在本申请各个实施例中的各功能单元可以集成在一个处理模块中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用时，也可以存储在一个计算机可读取存储介质中。

本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成，也可以通过程序来指令相关的硬件完成，该程序可以存储于一计算机可读存储介质中，存储介质可以包括存储器、磁盘或光盘等。

应该理解的是，在本申请各实施例中的各功能单元可集成在一个处理模块中，也可以各个单元单独物理存在，也可以两个或两个以上单元集成于一个模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。以上所述仅是本申请的部分实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本申请原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也应视为本申请的保护范围。

权利要求

1、一种卷积神经网络的池化方法，其特征在于，包括：

获取卷积神经网络的训练样本，将训练样本输入卷积神经网络模型；

将卷积神经网络模型的卷积层输出的样本数据矩阵输入卷积神经网络模型的池化层，在池化层中将样本数据矩阵划分为多个数据子矩阵；

从每个数据子矩阵中随机提取一个样本数据；

根据提取的每个样本数据生成池化矩阵，将该池化矩阵作为池化层的输出。

2、根据权利要求 1 所述的方法，其特征在于，所述多个数据子矩阵包括行数与列数相同的数据子矩阵；所述在池化层中将样本数据矩阵划分为多个数据子矩阵，包括：

获取所述样本数据矩阵的行数和列数；

根据所述样本数据矩阵的行数和列数，在池化层中将所述样本数据矩阵划分为多个行数与列数相同的数据子矩阵。

3、根据权利要求 1 所述的方法，其特征在于，所述多个数据子矩阵包括行数与列数相同的第一数据子矩阵以及行数与列数不相同的第二数据子矩阵；所述在池化层中将样本数据矩阵划分为多个数据子矩阵，包括：

获取样本数据矩阵的行数和列数；

根据样本数据矩阵的行数和列数在池化层中将样本数据矩阵划分为多个所述第一数据子矩阵以及多个所述第二数据子矩阵。

4、根据权利要求 1 所述的方法，其特征在于，所述数据子矩阵为多行及多列的矩阵；所述从每个数据子矩阵中随机提取一个样本数据，包括：

从所述每个数据子矩阵的多行中随机选取任一行数据；从该行数据中，随机选取任一列对应的数据作为所述样本数据；或，

从所述每个数据子矩阵的多列中随机选取任一列数据；从该列数据中，随机选取任一行对应的数据作为所述样本数据。

5、根据权利要求 1 所述的方法，其特征在于，所述数据子矩阵为多行及多列的矩阵；所述从每个数据子矩阵中随机提取一个样本数据，包括：

从所述每个数据子矩阵的多行及多列中，随机选取任一行和任一列对应的数据作为所述样本数据。

6、根据权利要求 1 所述的方法，其特征在于，所述训练样本为图像样本数据；所述获取卷积神经网络的训练样本，将训练样本输入卷积神经网络模型，包括：

获取所述图像样本数据作为所述卷积神经网络模型的训练样本。

7、根据权利要求 6 所述的方法，其特征在于，所述获取所述图像样本数据作为所述卷积神经网络模型的训练样本，包括：获取所述图像样本数据中的图片样本数据作为所述卷积神经网络模型的训练样本；

所述将卷积神经网络模型的卷积层输出的样本数据矩阵输入卷积神经网络模型的池化层，在池化层中将样本数据矩阵划分为多个数据子矩阵，从每个数据子矩阵中随机提取一个样本数据，根据提取的每个样本数据生成池化矩阵，将该池化矩阵作为池化层的输出，包括：

将卷积神经网络模型的第一卷积层输出的图片样本数据矩阵经过激活后，输入卷积神经网络模型的第一池化层，在第一池化层中将样本数据矩阵划分为多个数据子矩阵，从每个数据子矩阵中随机提取一个图片样本数据，根据提取的图片样本数据生成池化矩阵进行第一池化层输出后，再将所述第一池化层输出的数据输入第二卷积层进行卷积，对所述第二卷积层输出的数据进行随机激活，并将随机激活后的数据输入第二池化层进行池化训练，将该池化训练结果作为系统池化层的输出。

8、一种卷积神经网络的池化装置，其特征在于，包括：

获取模块，用于获取卷积神经网络的训练样本，将训练样本输入卷积神经网络模型；

划分模块，用于将卷积神经网络模型的卷积层输出的样本数据矩阵输入卷积神经网络模型的池化层，在池化层中将样本数据矩阵划分为多个数据子矩阵；

提取模块，用于从每个数据子矩阵中随机提取一个样本数据；

生成模块，用于根据提取的每个样本数据生成池化矩阵，将该池化矩阵作为池化层的输出。

9、根据权利要求 8 所述的装置，其特征在于，所述多个数据子矩阵包括行数与列数相同的数据子矩阵；所述划分模块执行所述在池化层中将样本数据矩阵划分为多个数据子矩阵时，具体用于：

获取所述样本数据矩阵的行数和列数；

根据所述样本数据矩阵的行数和列数，在池化层中将所述样本数据矩阵划分为多个行数与列数相同的数据子矩阵。

10、根据权利要求 8 所述的装置，其特征在于，所述多个数据子矩阵包括行数与列数相同的第一数据子矩阵以及行数与列数不相同的第二数据子矩阵；所述划分模块执行所述在池化层中将样本数据矩阵划分为多个数据子矩阵时，具体用于：

获取样本数据矩阵的行数和列数；

根据样本数据矩阵的行数和列数在池化层中将样本数据矩阵划分为多个所述第一数据子矩阵以及多个所述第二数据子矩阵。

11、根据权利要求 8 所述的装置，其特征在于，所述数据子矩阵为多行及多列的矩阵；所述提取模块在执行所述从每个数据子矩阵中随机提取一个样本数据时，具体用于：

从所述每个数据子矩阵的多行中随机选取任一行数据；从该行数据中，随机选取任一列对应的数据作为所述样本数据；或，

从所述每个数据子矩阵的多列中随机选取任一列数据；从该列数据中，随机选取任一行对应的数据作为所述样本数据。

12、根据权利要求 8 所述的装置，其特征在于，所述数据子矩阵为多行及多列的矩阵；所述提取模块在执行所述从每个数据子矩阵中随机提取一个样本数据时，具体用于：

从所述每个数据子矩阵的多行及多列中，随机选取任一行和任一列对应的数据作为所述样本数据。

13、根据权利要求 8 所述的装置，其特征在于，所述训练样本为图像样本数据；所述获取模块具体用于：

获取所述图像样本数据作为所述卷积神经网络模型的训练样本。

14、根据权利要求 13 所述的装置，其特征在于，所述获取模块在执行所述获取所述图像样本数据作为所述卷积神经网络模型的训练样本时，具体用于：获取所述图像样本数据中的图片样本数据作为所述卷积神经网络模型的训练样本；

所述划分模块在执行所述将卷积神经网络模型的卷积层输出的样本数据矩阵输入卷积神经网络模型的池化层，在池化层中将样本数据矩阵划分为多个数据子矩阵时，具体用于：将卷积神经网络模型的第一卷积层输出的图片样本数据矩阵经过激活后，输入卷积神经网络模型的第一池化层，在第一池化层中将样本数据矩阵划分为多个数据子矩阵；

所述提取模块在执行所述从每个数据子矩阵中随机提取一个样本数据时，具体用于：从每个数据子矩阵中随机提取一个图片样本数据；

所述生成模块在执行根据提取的每个样本数据生成池化矩阵，将该池化矩阵作为池化层的输出时，具体用于：根据提取的图片样本数据生成池化矩阵进行第一池化层输出后，再将所述第一池化层输出的数据输入第二卷积层进行卷积，对所述第二卷积层输出的数据进行随机激活，并将随机激活后的数据输入第二池化层进行池化训练，将该池化训练结果作为系统池化层的输出。

15、一种计算机非易失性存储介质，其特征在于，其上存储有计算机程序；所述计算机程序适于由处理器加载并执行上述权利要求1至7中任一项所述的卷积神经网络的池化方法。

16、一种计算机设备，其特征在于，其包括：

一个或多个处理器；

存储器；

一个或多个应用程序，其中所述一个或多个应用程序被存储在所述存储器中并被配置为由所述一个或多个处理器执行，所述一个或多个应用程序配置用于执行以下步骤：

获取卷积神经网络的训练样本，将训练样本输入卷积神经网络模型；

将卷积神经网络模型的卷积层输出的样本数据矩阵输入卷积神经网络模型的池化层，在池化层中将样本数据矩阵划分为多个数据子矩阵；

从每个数据子矩阵中随机提取一个样本数据；

根据提取的每个样本数据生成池化矩阵，将该池化矩阵作为池化层的输出。

17、根据权利要求16所述的计算机设备，其特征在于，所述多个数据子矩阵包括行数与列数相同的数据子矩阵；所述在池化层中将样本数据矩阵划分为多个数据子矩阵时，所述一个或多个应用程序被配置用于执行以下步骤：

获取所述样本数据矩阵的行数和列数；

根据所述样本数据矩阵的行数和列数，在池化层中将所述样本数据矩阵划分为多个行数与列数相同的数据子矩阵。

18、根据权利要求 16 所述的计算机设备，其特征在于，所述多个数据子矩阵包括行数与列数相同的第一数据子矩阵以及行数与列数不相同的第二数据子矩阵；所述在池化层中将样本数据矩阵划分为多个数据子矩阵时，所述一个或多个应用程序被配置用于执行以下步骤：

获取样本数据矩阵的行数和列数；

根据样本数据矩阵的行数和列数在池化层中将样本数据矩阵划分为多个所述第一数据子矩阵以及多个所述第二数据子矩阵。

19、根据权利要求 16 所述的计算机设备，其特征在于，所述数据子矩阵为多行及多列的矩阵；所述从每个数据子矩阵中随机提取一个样本数据时，所述一个或多个应用程序被配置用于执行以下步骤：

从所述每个数据子矩阵的多行中随机选取任一行数据；从该行数据中，随机选取任一列对应的数据作为所述样本数据；或，

从所述每个数据子矩阵的多列中随机选取任一列数据；从该列数据中，随机选取任一行对应的数据作为所述样本数据。

20、根据权利要求 16 所述的计算机设备，其特征在于，所述训练样本为图像样本数据；所述获取卷积神经网络的训练样本，将训练样本输入卷积神经网络模型时，所述一个或多个应用程序被配置用于执行以下步骤：

获取所述图像样本数据中的图片样本数据作为所述卷积神经网络模型的训练样本；

所述将卷积神经网络模型的卷积层输出的样本数据矩阵输入卷积神经网络模型的池化层，在池化层中将样本数据矩阵划分为多个数据子矩阵，从每个数据子矩阵中随机提取一个样本数据，根据提取的每个样本数据生成池化矩阵，将该池化矩阵作为池化层的输出时，所述一个或多个应用程序被配置用于执行以下步骤：

将卷积神经网络模型的第一卷积层输出的图片样本数据矩阵经过激活后，输入卷积神经网络模型的第一池化层，在第一池化层中将样本数据矩阵划分为

多个数据子矩阵，从每个数据子矩阵中随机提取一个图片样本数据，根据提取的图片样本数据生成池化矩阵进行第一池化层输出后，再将所述第一池化层输出的数据输入第二卷积层进行卷积，对所述第二卷积层输出的数据进行随机激活，并将随机激活后的数据输入第二池化层进行池化训练，将该池化训练结果作为系统池化层的输出。

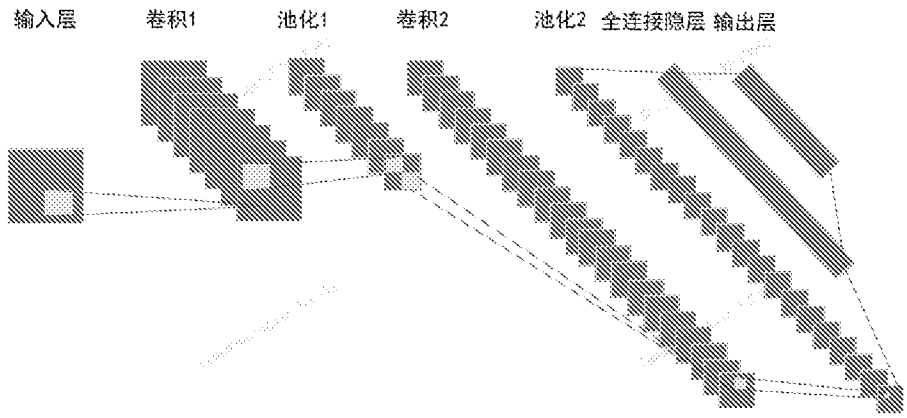

输入层
卷积1
池化1
卷积2
池化2
全连接隐层
输出层

图 1

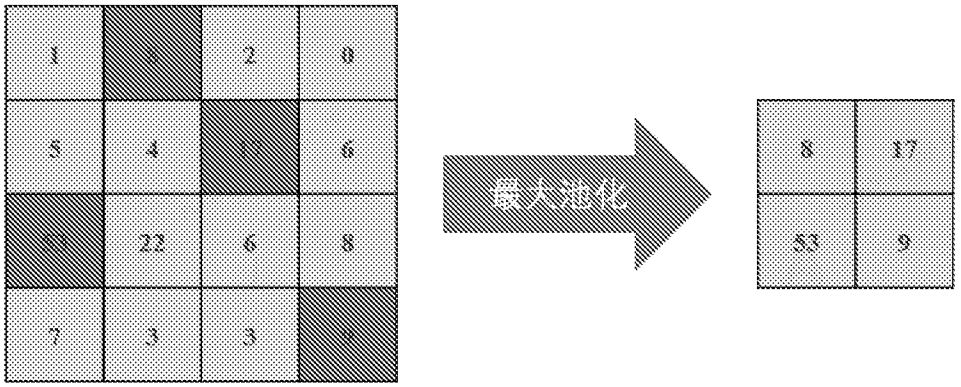

1
8
2
9
5
4
17
6
53
22
6
8
7
3
3
9
最大池化
8
17
53
9

图 2

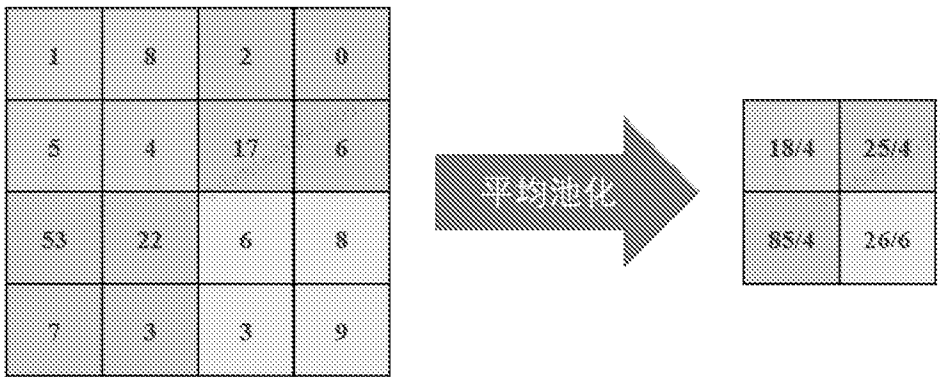

1
8
2
9
5
4
17
6
53
22
6
8
7
3
3
9
平均池化
18/4
25/4
85/4
26/6

图 3

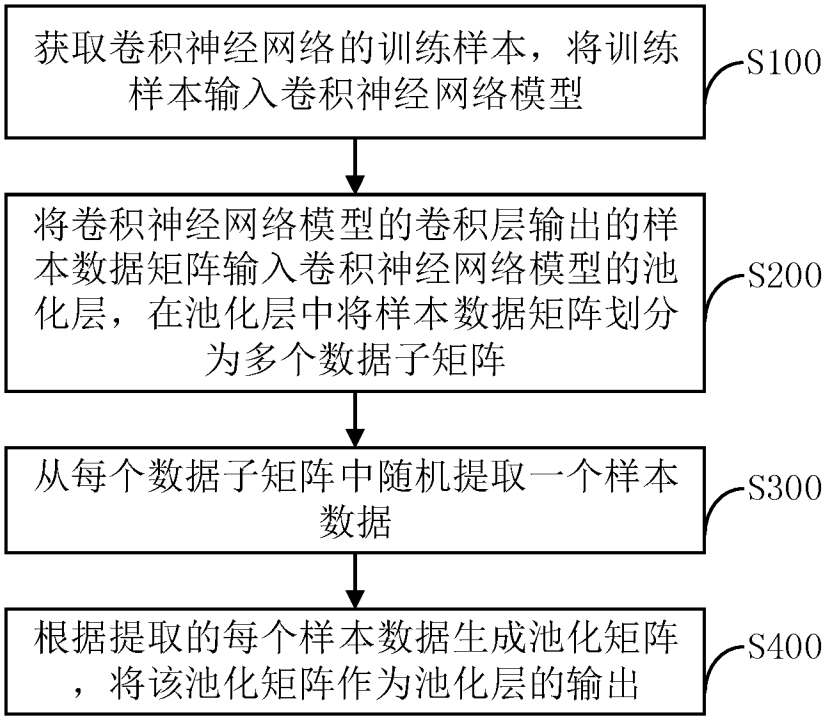

获取卷积神经网络的训练样本，将训练样本输入卷积神经网络模型
S100
将卷积神经网络模型的卷积层输出的样本数据矩阵输入卷积神经网络模型的池化层，在池化层中将样本数据矩阵划分为多个数据子矩阵
S200
从每个数据子矩阵中随机提取一个样本数据
S300
根据提取的每个样本数据生成池化矩阵，将该池化矩阵作为池化层的输出
S400

图 4

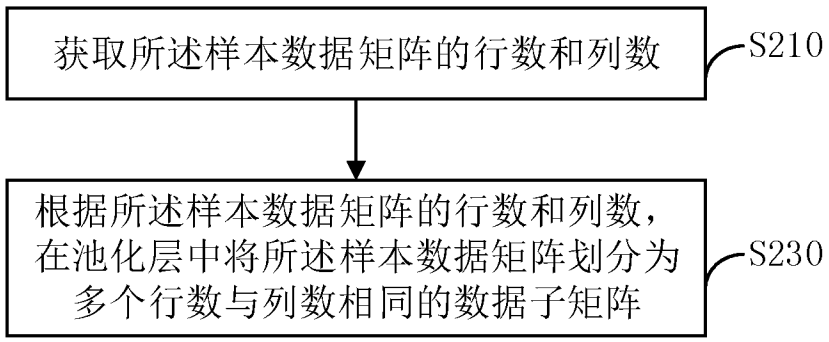

获取所述样本数据矩阵的行数和列数
S210
根据所述样本数据矩阵的行数和列数，在池化层中将所述样本数据矩阵划分为多个行数与列数相同的数据子矩阵
S230

图 5

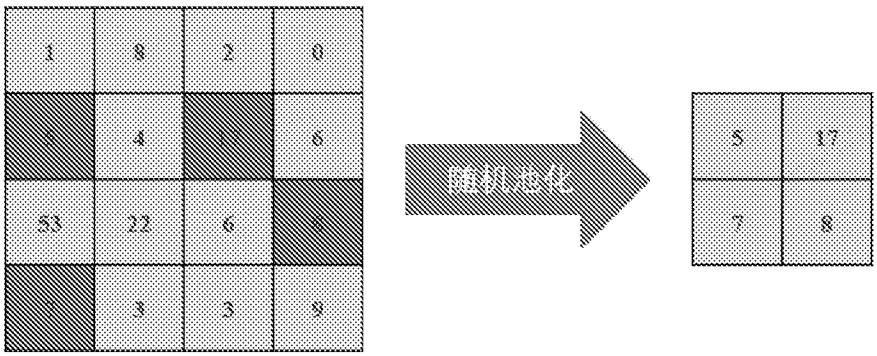

1
8
2
8
4
6
53
22
6
3
3
9
随机池化
5
17
7
8

图 6

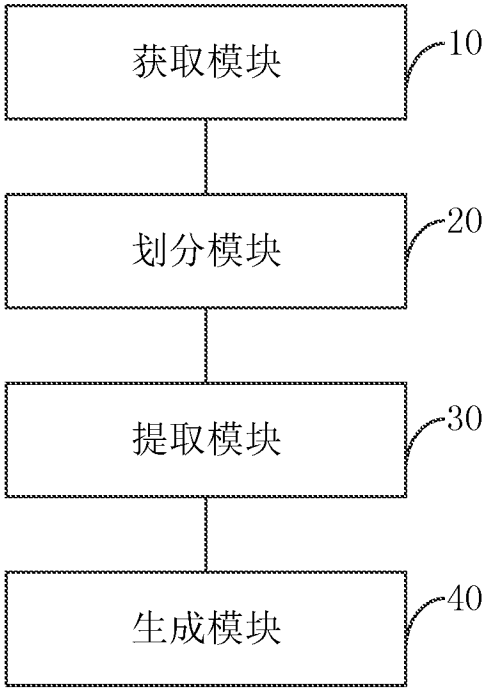

获取模块
10
划分模块
20
提取模块
30
生成模块
40

图 7

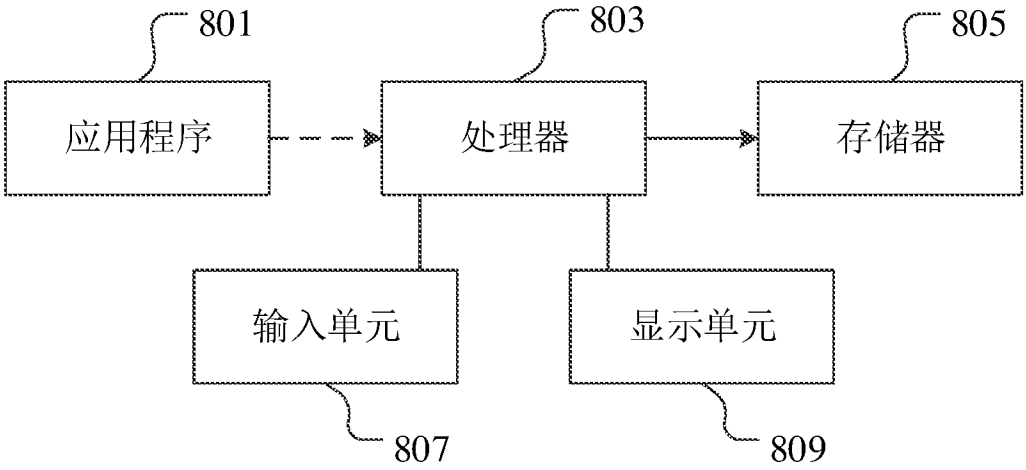

801
803
805
应用程序
处理器
存储器
输入单元
显示单元
807
809

图 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2019/117863

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/62(2006.01)i; G06N 3/04(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)

G06K,; 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; WPI; EPODOC; BAIDU; CNKI; ISI: 卷积神经网络, 训练, 池化, 随机, 任意, 任一; CNN, convolutional neural network, pooling, train, random

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109948651 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 28 June 2019 (2019-06-28) claims 1-10	1-20
X	CN 104408435 A (ZHEJIANG UNIVERSITY) 11 March 2015 (2015-03-11) description, paragraphs [0029]-[0044], and figures 4-6	1-20
A	CN 107871136 A (SUN YAT-SEN UNIVERSITY) 03 April 2018 (2018-04-03) entire document	1-20
A	CN 107506722 A (CHINA UNIVERSITY OF GEOSCIENCES WUHAN) 22 December 2017 (2017-12-22) entire document	1-20
A	US 2017323202 A1 (FUJITSU LIMITED) 09 November 2017 (2017-11-09) entire document	1-20

☐ 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
06 January 2020	**31 January 2020**

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/CN2019/117863

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109948651	A	28 June 2019	None			
CN	104408435	A	11 March 2015	None			
CN	107871136	A	03 April 2018	None			
CN	107506722	A	22 December 2017	None			
US	2017323202	A1	09 November 2017	KR	20170125720	A	15 November 2017
				CN	107346448	A	14 November 2017
				EP	3242255	A1	08 November 2017
				JP	2017201526	A	09 November 2017

国际检索报告

国际申请号
PCT/CN2019/117863

A. 主题的分类

G06K 9/62(2006.01)i; G06N 3/04(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)

G06K,; 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;WPI;EPODOC;BAIDU;CNKI;ISI: 卷积神经网络，训练，池化，随机，任意，任一；CNN, convolutional neural network, pooling, train, random

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 109948651 A (平安科技深圳有限公司) 2019年 6月 28日 (2019-06-28) 权利要求1-10	1-20
X	CN 104408435 A (浙江大学) 2015年 3月 11日 (2015-03-11) 说明书第[0029]-[0044]段、图4-6	1-20
A	CN 107871136 A (中山大学) 2018年 4月 3日 (2018-04-03) 全文	1-20
A	CN 107506722 A (中国地质大学武汉) 2017年 12月 22日 (2017-12-22) 全文	1-20
A	US 2017323202 A1 (FUJITSU LIMITED) 2017年 11月 9日 (2017-11-09) 全文	1-20

☐ 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
2020年 1月 6日	2020年 1月 31日
ISA/CN的名称和邮寄地址 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 传真号 (86-10)62019451	受权官员 张琳琳 电话号码 86-(10)-53961404

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2019/117863

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	109948651	A	2019年 6月 28日	无			
CN	104408435	A	2015年 3月 11日	无			
CN	107871136	A	2018年 4月 3日	无			
CN	107506722	A	2017年 12月 22日	无			
US	2017323202	A1	2017年 11月 9日	KR	20170125720	A	2017年 11月 15日
				CN	107346448	A	2017年 11月 14日
				EP	3242255	A1	2017年 11月 8日
				JP	2017201526	A	2017年 11月 9日