

(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 13 日 (13.01.2022)

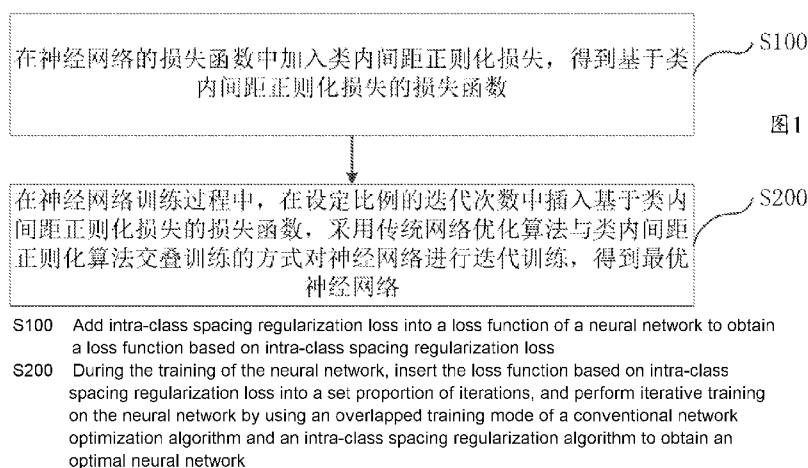


(10) 国际公布号
WO 2022/007349 A1

- (51) 国际专利分类号:
G06N 3/08 (2006.01) *G06N 3/02* (2006.01)
- (21) 国际申请号: PCT/CN2020/139297
- (22) 国际申请日: 2020 年 12 月 25 日 (25.12.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202010657269.2 2020年7月9日 (09.07.2020) CN
- (71) 申请人: 中国科学院深圳先进技术研究院 (SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY CHINESE ACADEMY OF SCIENCES) [CN/CN]; 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。
- (72) 发明人: 赵宝新(ZHAO, Baoxin); 中国广东省深圳市南山区深圳大学城学苑大道1068号, Guangdong 518055 (CN)。 须成忠(XU, Chengzhong); 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。 赵娟娟(ZHAO, Juanjuan); 中国广东省深圳市南山区深圳大学城学苑大道1068号, Guangdong 518055 (CN)。
- (74) 代理人: 北京中巡通大知识产权代理有限公司 (BEIJING ZHONG XUN TONG DA INTELLECTUAL PROPERTY AGENCY CO., LTD.); 中国北京市海淀区杏石口路 50 号中间建筑 1 区 148 号, Beijing 100195 (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, IT,

(54) Title: NEURAL NETWORK TUNING METHOD AND SYSTEM, TERMINAL, AND STORAGE MEDIUM

(54) 发明名称: 一种神经网络调优方法、系统、终端以及存储介质



(57) Abstract: The present application relates to a neural network tuning method and system, a terminal, and a storage medium. The method comprises: adding intra-class spacing regularization loss into a loss function of a neural network to obtain a loss function based on intra-class spacing regularization, the intra-class spacing regularization loss indicating that the intra-class spacing between feature maps of the same class of data outputted by the neural network is close; and during the training of the neural network, inserting the loss function based on intra-class spacing regularization into a set proportion of iterations, and performing iterative training on the neural network by using an overlapped training mode of a network optimization algorithm and an intra-class spacing regularization algorithm to obtain an optimal neural network. According to embodiments of the present application, the intra-class spacing regularization algorithm is inserted into a certain proportion of iterations to perform iterative training on the neural network, so that the generalization ability and the anti-interference performance of the network structure are improved in the case that the structure of the neural network is not changed, and thus the additional time overhead is relatively small.

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: 本申请涉及一种神经网络调优方法、系统、终端以及存储介质。所述方法包括: 在神经网络的损失函数中加入类内间距正则化损失, 得到基于类内间距正则化的损失函数; 所述类内间距正则化损失表示所述神经网络输出的同一类别的数据的特征图之间的类内间距接近; 在所述神经网络训练过程中, 在设定比例的迭代次数中插入所述基于类内间距正则化的损失函数, 采用网络优化算法与类内间距正则化算法交叠训练的方式对所述神经网络进行迭代训练, 得到最优神经网络。本申请实施例通过在一定比例的迭代次数中插入类内间距正则化算法对神经网络进行迭代训练, 在不改变神经网络结构的情况下, 提升网络结构的泛化能力和抗干扰性, 且由此带来的额外时间开销较小。

一种神经网络调优方法、系统、终端以及存储介质

技术领域

本申请属于人工智能技术领域，特别涉及一种神经网络调优方法、系统、终端以及存储介质。

背景技术

随着人工智能的发展，深度神经网络以其强大的拟合能力在众多领域取得了惊人的成绩。然而神经网络的强大拟合能力依旧没有一个强有力的理论支持，深度神经网络的能力依旧具有很强的开发空间。为了得到更好的神经网络泛化能力，现有的网络结构变得越来越复杂，网络的参数量也呈现爆炸式增长的趋势，复杂的神经网络的训练过程变得非常缓慢，同时训练网络所带来的能量消耗也成倍增长。另外，神经网络也存在抗干扰性差、对输入数据敏感等问题。为了改善上述问题，需要对神经网络泛化的能力进行提升。

目前，提升神经网络泛化能力的主要方法分为三大类，具体为：

一、从网络结构出发，设计更好的网络结构。这种方法主要从深度神经网络的深度、宽度和广度三个角度进行网络结构的修正，复杂的网络结构可以带来更好的性能收益，但是也会导致网络的参数量巨大，训练时间较长，能源消耗较大。

二、提供更多的训练数据，使得神经网络能够识别更多的数据类型。该方法需要更多优质的数据，对数据进行处理需要消耗非常多的人力和算力。

三、从训练的过程出发，寻找更好的超参数或者优化方法。该方法可以在不改变网络结构的条件下，提高网络性能。然而超参搜索也需要迭代非常多次的网络训练过程，对算力和能源的消耗非常大。

发明内容

本申请提供了一种神经网络调优方法、系统、终端以及存储介质，旨在至少在一定程度上解决现有技术中的上述技术问题之一。

为了解决上述问题，本申请提供了如下技术方案：

一种神经网络调优方法，包括以下步骤：

在神经网络的损失函数中加入类内间距正则化损失，得到基于类内间距正则化的损失函数；其中，所述类内间距正则化损失表示所述神经网络输出的同一类别的数据的特征图之间的类内间距接近；给定一个神经网络结构 $z(x_i, \omega)$ ， ω 是神经网络的参数， x_i 是输入数据，所述基于类内间距正则化的损失函数为：

$$loss = \min_{\omega} \sum_{i=1}^n L(z(x_i, \omega), y_i) + \lambda \cdot \Omega_{intra}(\omega, F_C)$$

上式中， y_i 是输入数据 x_i 对应的标签， $L(z(x_i, \omega), y_i)$ 是经验损失函数， λ 是一个超参数， $\Omega_{intra}(\omega, F_C)$ 为类内间距正则化损失， F_C 是所有类别的特征图集合；

在所述神经网络训练过程中，在设定比例的迭代次数中插入所述基于类内间距正则化的损失函数，采用网络优化算法与类内间距正则化算法交叠训练的方式对所述神经网络进行迭代训练，得到最优神经网络。

本申请实施例采取的技术方案还包括：所述基于类内间距正则化的损失函数用于提取各个类别的数据的特征图的类别中心，所述类别中心计算方式为：

将所述基于类内间距正则化的损失函数展开，得：

$$\Omega_{intra} = \min_{\omega} \sum_{i=1}^n d(fm_{x_i}, fm_c), c \in Y, fm_c \in F_c$$

上述公式中， fm_{x_i} 代表输入数据 x_i 经过神经网络后得到的特征图，数据 x_i 的标签为 c ，记作 $l(x_i) \in c$ ， fm_c 为标签为 c 的所有数据的特征图的类别中心， $d(fm_{x_i}, fm_c)$ 代表数据 x_i 的特征图与所属类别中心之间的距离；

所述类别中心的特征图计算公式为：

$$f_c = \frac{1}{|X_c|} \sum_{i=1}^{|X_c|} fm_{x_i} \text{ if } l(x_i) \in c$$

上式中， $|X_c|$ 代表所有样本 X 中类别属于 c 的样本数目。

本申请实施例采取的技术方案还包括：所述网络优化算法采用网络优化算法与类内间距正则化算法交叠训练的方式对所述神经网络进行迭代训练包括：

将所述神经网络的一次完整训练记为一次epoch，假设整个训练过程共有 T 个epoch，首先采用网络优化算法进行迭代训练，每当所述网络优化算法的迭代次数达到设定次数阈值时，在最后一次网络优化算法训练的epoch之

后插入所述设定次数阈值的10%的迭代次数进行类内间距正则化算法迭代训练，并在最后一次类内间距正则化算法训练的epoch之后再次采用网络优化算法进行迭代训练；如此循环，直到完成T个epoch。

本申请实施例采取的技术方案还包括：所述网络优化算法迭代次数的设定次数阈值为20次。

本申请实施例采取的技术方案还包括：所述类内间距正则化算法的损失函数为：

$$loss_{ICR} = \min_{\omega} \sum_{i=1}^n L(z(x_i, \omega), y_i) + \lambda \cdot \Omega_{intra}(\omega, F_C) \quad \circ$$

本申请实施例采取的技术方案还包括：所述网络优化算法为随机梯度下降法。

本申请实施例采取的技术方案还包括：所述随机梯度下降法的损失函数为：

$$loss_{default} = \min_{\omega} \sum_{i=1}^n L(z(x_i, \omega), y_i) \quad \circ$$

本申请实施例采取的另一技术方案为：一种神经网络调优系统，包括：

损失函数优化模块：用于在神经网络的损失函数中加入类内间距正则化损失，得到基于类内间距正则化的损失函数；其中，所述类内间距正则化损失表示所述神经网络输出的同一类别的数据的特征图之间的类内间距接近；给定一个神经网络结构 $z(x_i, \omega)$ ， ω 是神经网络的参数， x_i 是输入数据，所述基于类内间距正则化的损失函数为：

$$loss = \min_{\omega} \sum_{i=1}^n L(z(x_i, \omega), y_i) + \lambda \cdot \Omega_{intra}(\omega, F_C)$$

上式中， y_i 是输入数据 x_i 对应的标签， $L(z(x_i, \omega), y_i)$ 是经验损失函数， λ 是一个超参数， $\Omega_{intra}(\omega, F_C)$ 为类内间距正则化损失， F_C 是所有类别的特征图集合；

模型训练模块：用于在所述神经网络训练过程中，在设定比例的迭代次数中插入所述基于类内间距正则化的损失函数，采用网络优化算法与类内间距正则化算法交叠训练的方式对所述神经网络进行迭代训练，得到最优神经网络。

本申请实施例采取的又一技术方案为：一种终端，所述终端包括处理器、与所述处理器耦接的存储器，其中，

所述存储器存储有用于实现所述神经网络调优方法的程序指令；

所述处理器用于执行所述存储器存储的所述程序指令以控制神经网络调优。

本申请实施例采取的又一技术方案为：一种存储介质，存储有处理器可运行的程序指令，所述程序指令用于执行所述神经网络调优方法。

相对于现有技术，本申请实施例产生的有益效果在于：本申请实施例的神经网络调优方法、系统、终端及存储介质通过在神经网络的损失函数中加入类内间距正则化损失，在网络训练的过程中，在一定比例的迭代次数中插入类内间距正则化算法对神经网络进行迭代训练，类内间距正则化算法只需提取各个类别的样本数据的特征图的类别中心，而其余迭代过程与原始网络

的训练方法一致，从而在不改变神经网络结构的情况下，提升网络结构的泛化能力和抗干扰性，且由此带来的额外时间开销较小。

附图说明

图1是本申请实施例的神经网络调优方法的流程图；

图2为神经网络结构示意图；

图3为本申请实施例的神经网络训练示意图；

图4为本申请实施例的神经网络调优系统结构示意图；

图5为本申请实施例的终端结构示意图；

图6为本申请实施例的存储介质的结构示意图。

具体实施方式

为了使本申请的目的、技术方案及优点更加清楚明白，以下结合附图及实施例，对本申请进行进一步详细说明。应当理解，此处所描述的具体实施例仅用以解释本申请，并不用于限定本申请。

为了解决现有技术的不足，本申请实施例的神经网络调优方法通过在神经网络的损失函数中加入类内间距正则化损失，在网络训练的过程中，在一定比例的迭代次数中插入类内间距正则化算法对神经网络进行迭代训练，而其余迭代过程与原始网络的训练方法一致，在不改变神经网络结构的情况下，提升网络结构的泛化能力，且由此带来的额外时间开销较小。

请参阅图1，是本申请实施例的神经网络调优方法的流程图。本申请实施例的神经网络调优方法包括以下步骤：

步骤100：在神经网络的损失函数中加入类内间距正则化损失，得到基于类内间距正则化的损失函数；

步骤100中，神经网络通常包括卷积层和全连接层，具体如图2所示，为神经网络结构示意图。其中梯形为卷积层，三角形为全连接层。以图片分类任务为例，当输入一张图片后，首先经过若干层卷积层进行特征提取，并输出特征图(Feature Map)，然后经过若干全连接层得到输出的图片分类结果。

假设输入数据记作 $X=\{x_1, x_2, \dots, x_n\}$ ，其中 n 为样本数；对于每一个输入 x_i ，其特征图输出记作 FM_i ，输入数据的标签记作 $Y=\{1, 2, \dots, C\}$ ，数据的分类类别一共包括 C 类。为了使神经网络可以对输入数据进行非常精确的识别分类，假设其输出的特征图也具有非常好的聚类效果，即同一类别的数据的特征图之间的类内间距尽可能的小，基于此假设本申请实施例提出类内间距正则化损失。

给定一个神经网络结构 $z(x_i, \omega)$ ， ω 是神经网络的参数， x_i 是输入数据，其损失函数包括经验损失函数和正则化损失函数，损失函数可以归纳为：

$$loss = \min_{\omega} \sum_{i=1}^n L(z(x_i, \omega), y_i) + \lambda \cdot \Omega(\omega) \quad (1)$$

式(1)中， y_i 是输入数据 x_i 对应的标签， $L(z(x_i, \omega), y_i)$ 是经验损失函数， $\Omega(\omega)$ 是正则化损失函数， λ 是一个超参数，用以平衡经验损失函数与正则化损失函数。

基于上式，将类内间距正则化损失加入到损失函数中，得到基于类内间距正则化的损失函数：

$$loss = \min_{\omega} \sum_{i=1}^n L(z(x_i, \omega), y_i) + \lambda \cdot \Omega_{intra}(\omega, F_C) \quad (2)$$

式(2)中, $\Omega_{intra}(\omega, F_C)$ 代表类内间距正则化损失, F_C 是所有类别的特征图集合。将其展开为:

$$\Omega_{intra} = \min_{\omega} \sum_{i=1}^n d(fm_{x_i}, fm_c), c \in Y, fm_c \in F_C \quad (3)$$

式(3)中, fm_{x_i} 代表输入数据 x_i 经过神经网络后得到的特征图, 数据 x_i 的标签为 c , 记作 $l(x_i) \in c$, fm_c 为标签为 c 的所有样本的特征图的聚类中心, 即类别中心, $d(fm_{x_i}, fm_c)$ 代表样本 x_i 的特征图与所属类别中心之间的距离。

其中, 类别中心的特征图通过自学习得到:

$$f_c = \frac{1}{|X_c|} \sum_{i=1}^{|X_c|} fm_{x_i} \text{ if } l(x_i) \in c \quad (4)$$

式(4)中, $|X_c|$ 代表所有样本 X 中类别属于 c 的样本数目。

步骤200: 在神经网络训练过程中, 在设定比例的迭代次数中插入基于类内间距正则化的损失函数(ICR), 采用传统网络优化算法与类内间距正则化算法交叠训练的方式对神经网络进行迭代训练, 得到最优神经网络;

步骤200中, 神经网络训练过程中, 使用训练集的全部数据样本对模型进行一次完整训练记为一次epoch (时期), 整个训练过程总共有 T 个epoch, 所有epoch的索引记作序列 $[0, 1, 2, \dots, T]$ 。传统网络优化算法包括但不限于随机梯度下降法 (Stochastic gradient descent, SGD)、小批量梯度下降法 (Mini-

batch gradient descent, MBGD)等。以SGD为例, 随机梯度下降法的损失函数为:

$$loss_{default} = \min_{\omega} \sum_{i=1}^n L(z(x_i, \omega), y_i) \quad (5)$$

类内间距正则化算法的损失函数为:

$$loss_{ICR} = \min_{\omega} \sum_{i=1}^n L(z(x_i, \omega), y_i) + \lambda \cdot \Omega_{intra}(\omega, F_C) \quad (6)$$

本申请实施例中, 随机梯度下降法与类内间距正则化方法的交叠方式为: 在神经网络训练过程中, 首先采用SGD进行迭代训练, 每当SGD迭代次数达到设定次数阈值时, 在最后一次SGD训练的epoch之后插入该设定次数阈值的10%的迭代次数进行ICR迭代训练, 然后在最后一次ICR训练的epoch之后再次采用SGD进行迭代训练。如此循环, 直到完成T个epoch。具体如图3所示, 为本申请实施例的神经网络训练示意图。其中灰色代表SGD的训练方法, 黑色代表ICR的训练方法, SGD与ICR交叠训练。优选地, 通过实验对比, SGD迭代训练的次数阈值为20次时, 神经网络的性能提升效果最好。即: 每间隔20个SGD训练的epoch之后插入两次ICR训练, 可以理解, 该次数阈值以及ICR方法的插入比例可根据实际操作进行调节。

基于上述, 本申请实施例通过在训练过程中插入较小比例的类内间距正则化方法, 类内间距正则化算法只需提取各个类别的样本数据的特征图的类别中心, 从而在不改变神经网络的结构的情况下, 提升神经网络的泛化能力和抗干扰性, 对神经网络的训练时间带来的额外负担也可以忽略不计。

为了验证本申请实施例的可行性和有效性，使用图片识别的公用数据集 CIFAR100 进行了实验。实验结果表明，本申请实施例在 VGG（Visual Geometry Group，深度卷积神经网络）、ResNet（Residual Network，残差网络）等都取得了非常好的性能提升和对白噪声的抗干扰能力，同时训练过程带来的额外时间开销小于原有训练时间的5%。

请参阅图4，是本申请实施例的神经网络调优系统的结构示意图。本申请实施例的神经网络调优系统包括：

损失函数优化模块：用于在神经网络的损失函数中加入类内间距正则化损失，得到基于类内间距正则化的损失函数；其中，给定一个神经网络结构 $z(x_i, \omega)$ ， ω 是神经网络的参数， x_i 是输入数据，其损失函数包括经验损失函数和正则化损失函数，损失函数可以归纳为：

$$loss = \min_{\omega} \sum_{i=1}^n L(z(x_i, \omega), y_i) + \lambda \cdot \Omega(\omega) \quad (1)$$

式（1）中， y_i 是输入数据 x_i 对应的标签， $L(z(x_i, \omega), y_i)$ 是经验损失函数， $\Omega(\omega)$ 是正则化损失函数， λ 是一个超参数，用以平衡经验损失函数与正则化损失函数。

基于上式，将类内间距正则化损失加入到损失函数中，得到基于类内间距正则化的损失函数：

$$loss = \min_{\omega} \sum_{i=1}^n L(z(x_i, \omega), y_i) + \lambda \cdot \Omega_{intra}(\omega, F_C) \quad (2)$$

式（2）中， $\Omega_{intra}(\omega, F_C)$ 代表类内间距正则化损失， F_C 是所有类别的特征图集合。将其展开为：

$$\Omega_{intra} = \min_{\omega} \sum_{i=1}^n d(fm_{x_i}, fm_c), c \in Y, fm_c \in F_c \quad (3)$$

式(3)中, fm_{x_i} 代表输入数据 x_i 经过神经网络后得到的特征图, 数据 x_i 的标签为 c , 记作 $l(x_i) \in c$, fm_c 为标签为 c 的所有样本的特征图的聚类中心, 即类别中心, $d(fm_{x_i}, fm_c)$ 代表样本 x_i 的特征图与所属类别中心之间的距离。

其中, 类别中心的特征图通过自学习得到:

$$f_c = \frac{1}{|X_c|} \sum_{i=1}^{|X_c|} fm_{x_i} \text{ if } l(x_i) \in c \quad (4)$$

式(4)中, $|X_c|$ 代表所有样本 X 中类别属于 c 的样本数目。

模型训练模块: 用于在神经网络训练过程中, 在设定比例的迭代次数中插入基于类内间距正则化的损失函数(ICR), 采用传统网络优化算法与类内间距正则化算法交叠训练的方式对神经网络进行迭代训练, 得到最优神经网络;

具体的, 神经网络训练过程中, 使用训练集的全部数据样本对模型进行一次完整训练记为一次epoch (时期), 整个训练过程总共有 T 个epoch, 所有epoch的索引记作序列 $[0, 1, 2, \dots, T]$ 。传统网络优化算法包括但不限于随机梯度下降法 (Stochastic gradient descent, SGD)、小批量梯度下降法 (Mini-batch gradient descent, MBGD)等。以SGD为例, 随机梯度下降法的损失函数为:

$$loss_{default} = \min_{\omega} \sum_{i=1}^n L(z(x_i, \omega), y_i) \quad (5)$$

类内间距正则化算法的损失函数为:

$$\text{loss}_{\text{ICR}} = \min_{\omega} \sum_{i=1}^n L(z(x_i, \omega), y_i) + \lambda \cdot \Omega_{\text{intra}}(\omega, F_C) \quad (6)$$

本申请实施例中，随机梯度下降法与类内间距正则化方法的交叠方式为：在神经网络训练过程中，首先采用SGD进行迭代训练，每当SGD迭代次数达到设定次数阈值时，在最后一次SGD训练的epoch之后插入该设定次数阈值的10%的迭代次数进行ICR迭代训练，然后在最后一次ICR训练的epoch之后再次采用SGD进行迭代训练。如此循环，直到完成T个epoch。优选地，通过实验对比，SGD迭代训练的次数阈值为20次时，神经网络的性能提升效果最好。即：每间隔20个SGD训练的epoch之后插入两次ICR训练，可以理解，该次数阈值以及ICR方法的插入比例可根据实际操作进行调节。

请参阅图5，为本申请实施例的终端结构示意图。该终端50包括处理器51、与处理器51耦接的存储器52。

存储器52存储有用于实现上述神经网络调优方法的程序指令。

处理器51用于执行存储器52存储的程序指令以控制神经网络调优。

其中，处理器51还可以称为CPU（Central Processing Unit，中央处理单元）。处理器51可能是一种集成电路芯片，具有信号的处理能力。处理器51还可以是通用处理器、数字信号处理器（DSP）、专用集成电路（ASIC）、现成可编程门阵列（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。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

请参阅图6，为本申请实施例的存储介质的结构示意图。本申请实施例的存储介质存储有能够实现上述所有方法的程序文件61，其中，该程序文件61

可以以软件产品的形式存储在上述存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等等）或处理器（processor）执行本发明各个实施方式方法的全部或部分步骤。而前述的存储介质包括：U盘、移动硬盘、只读存储器（ROM，Read-Only Memory）、随机存取存储器（RAM，Random Access Memory）、磁碟或者光盘等各种可以存储程序代码的介质，或者是计算机、服务器、手机、平板等终端设备。

对所公开的实施例的上述说明，使本领域专业技术人员能够实现或使用本申请。对这些实施例的多种修改对本领域的专业技术人员来说将是显而易见的，本申请中所定义的一般原理可以在不脱离本申请的精神或范围的情况下，在其它实施例中实现。因此，本申请将不会被限制于本申请所示的这些实施例，而是要符合与本申请所公开的原理和新颖特点相一致的最宽的范围。

权利要求书

1、一种神经网络调优方法，其特征在于，包括以下步骤：

在神经网络的损失函数中加入类内间距正则化损失，得到基于类内间距正则化的损失函数；其中，所述类内间距正则化损失表示所述神经网络输出的同一类别的数据的特征图之间的类内间距接近；给定一个神经网络结构 $z(x_i, \omega)$ ， ω 是神经网络的参数， x_i 是输入数据，所述基于类内间距正则化的损失函数为：

$$loss = \min_{\omega} \sum_{i=1}^n L(z(x_i, \omega), y_i) + \lambda \cdot \Omega_{intra}(\omega, F_C)$$

上式中， y_i 是输入数据 x_i 对应的标签， $L(z(x_i, \omega), y_i)$ 是经验损失函数， λ 是一个超参数， $\Omega_{intra}(\omega, F_C)$ 为类内间距正则化损失， F_C 是所有类别的特征图集合；

在所述神经网络训练过程中，在设定比例的迭代次数中插入所述基于类内间距正则化的损失函数，采用网络优化算法与类内间距正则化算法交叠训练的方式对所述神经网络进行迭代训练，得到最优神经网络。

2、根据权利要求1所述的神经网络调优方法，其特征在于，所述基于类内间距正则化的损失函数用于提取各个类别的数据的特征图的类别中心，所述类别中心计算方式为：

将所述基于类内间距正则化的损失函数展开，得：

$$\Omega_{intra} = \min_{\omega} \sum_{i=1}^n d(fm_{x_i}, fm_c), c \in Y, fm_c \in F_C$$

上述公式中, fm_{x_i} 代表输入数据 x_i 经过神经网络后得到的特征图, 数据 x_i 的标签为 c , 记作 $l(x_i) \in c$, fm_c 为标签为 c 的所有数据的特征图的类别中心, $d(fm_{x_i}, fm_c)$ 代表数据 x_i 的特征图与所属类别中心之间的距离;

所述类别中心的特征图计算公式为:

$$f_c = \frac{1}{|X_c|} \sum_{i=1}^{|X_c|} fm_{x_i} \text{ if } l(x_i) \in c$$

上式中, $|X_c|$ 代表所有样本 X 中类别属于 c 的样本数目。

3、根据权利要求2所述的神经网络调优方法, 其特征在于, 所述网络优化算法采用网络优化算法与类内间距正则化算法交叠训练的方式对所述神经网络进行迭代训练包括:

将所述神经网络的一次完整训练记为一次epoch, 假设整个训练过程共有 T 个epoch, 首先采用网络优化算法进行迭代训练, 每当所述网络优化算法的迭代次数达到设定次数阈值时, 在最后一次网络优化算法训练的epoch之后插入所述设定次数阈值的10%的迭代次数进行类内间距正则化算法迭代训练, 并在最后一次类内间距正则化算法训练的epoch之后再次采用网络优化算法进行迭代训练; 如此循环, 直到完成 T 个epoch。

4、根据权利要求3所述的神经网络调优方法, 其特征在于, 所述网络优化算法迭代次数的设定次数阈值为20次。

5、根据权利要求1至4任一项所述的神经网络调优方法, 其特征在于, 所述类内间距正则化算法的损失函数为:

$$loss_{ICR} = \min_{\omega} \sum_{i=1}^n L(z(x_i, \omega), y_i) + \lambda \cdot \Omega_{intra}(\omega, F_c) \quad \circ$$

6、根据权利要求1至4任一项所述的神经网络调优方法，其特征在于，所述网络优化算法为随机梯度下降法。

7、根据权利要求6所述的神经网络调优方法，其特征在于，所述随机梯度下降法的损失函数为：

$$loss_{default} = \min_{\omega} \sum_{i=1}^n L(z(x_i, \omega), y_i) \quad \circ$$

8、一种神经网络调优系统，其特征在于，包括：

损失函数优化模块：用于在神经网络的损失函数中加入类内间距正则化损失，得到基于类内间距正则化的损失函数；其中，所述类内间距正则化损失表示所述神经网络输出的同一类别的数据的特征图之间的类内间距接近；给定一个神经网络结构 $z(x_i, \omega)$ ， ω 是神经网络的参数， x_i 是输入数据，所述基于类内间距正则化的损失函数为：

$$loss = \min_{\omega} \sum_{i=1}^n L(z(x_i, \omega), y_i) + \lambda \cdot \Omega_{intra}(\omega, F_C)$$

上式中， y_i 是输入数据 x_i 对应的标签， $L(z(x_i, \omega), y_i)$ 是经验损失函数， λ 是一个超参数， $\Omega_{intra}(\omega, F_C)$ 为类内间距正则化损失， F_C 是所有类别的特征图集合；

模型训练模块：用于在所述神经网络训练过程中，在设定比例的迭代次数中插入所述基于类内间距正则化的损失函数，采用网络优化算法与类内间距正则化算法交叠训练的方式对所述神经网络进行迭代训练，得到最优神经网络。

9、一种终端，其特征在于，所述终端包括处理器、与所述处理器耦接的存储器，其中，

所述存储器存储有用于实现权利要求1-7任一项所述的神经网络调优方法的程序指令；

所述处理器用于执行所述存储器存储的所述程序指令以控制神经网络调优。

10、一种存储介质，其特征在于，存储有处理器可运行的程序指令，所述程序指令用于执行权利要求1至7任一项所述神经网络调优方法。

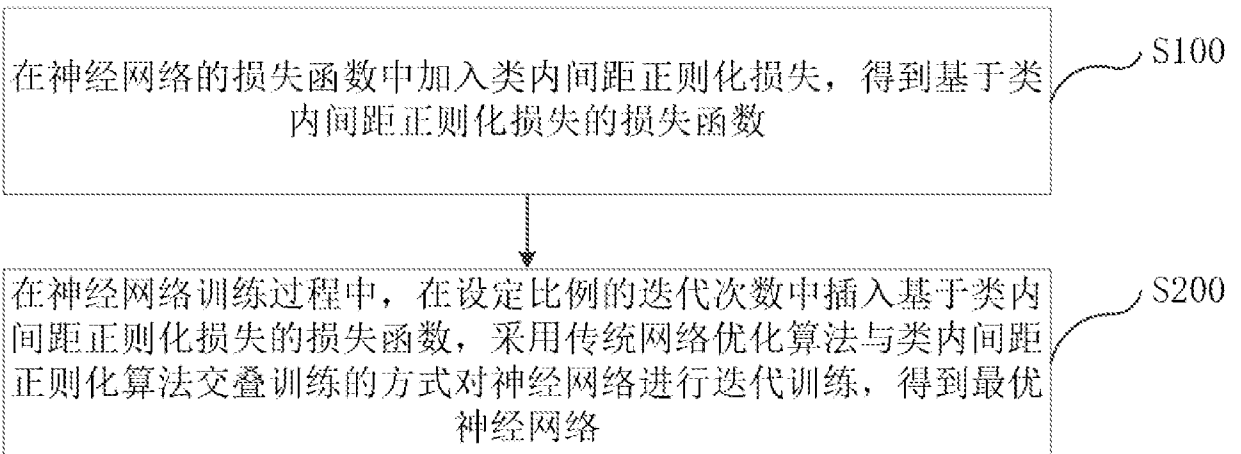


图1

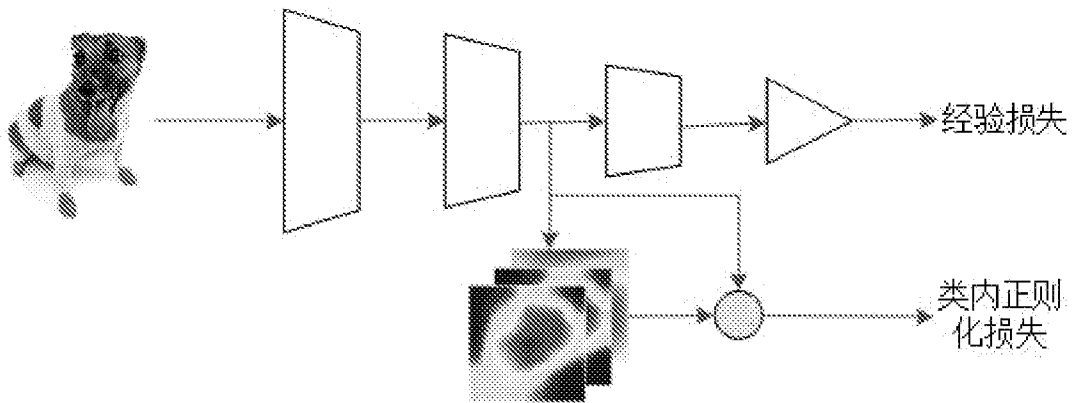


图2

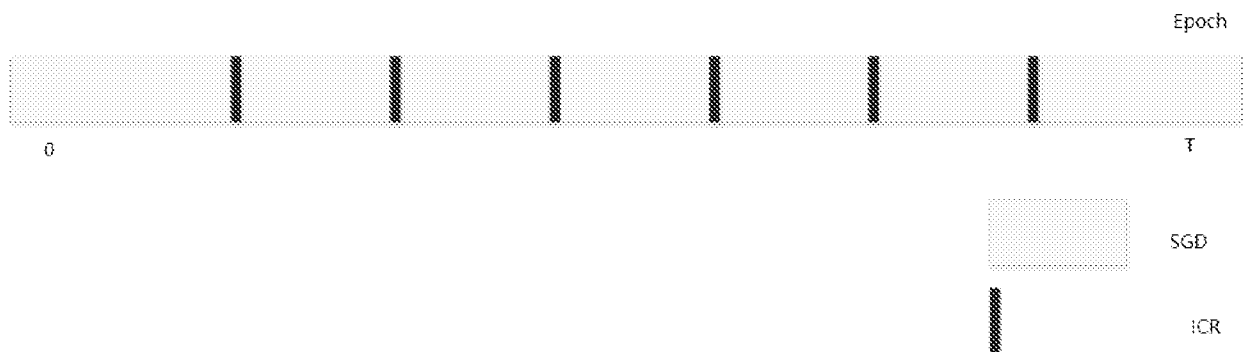


图3

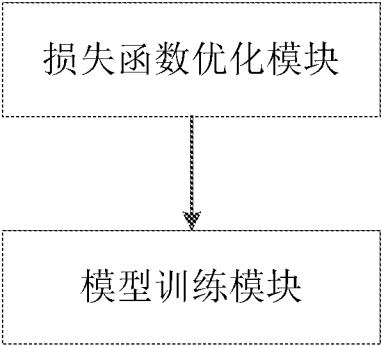


图4

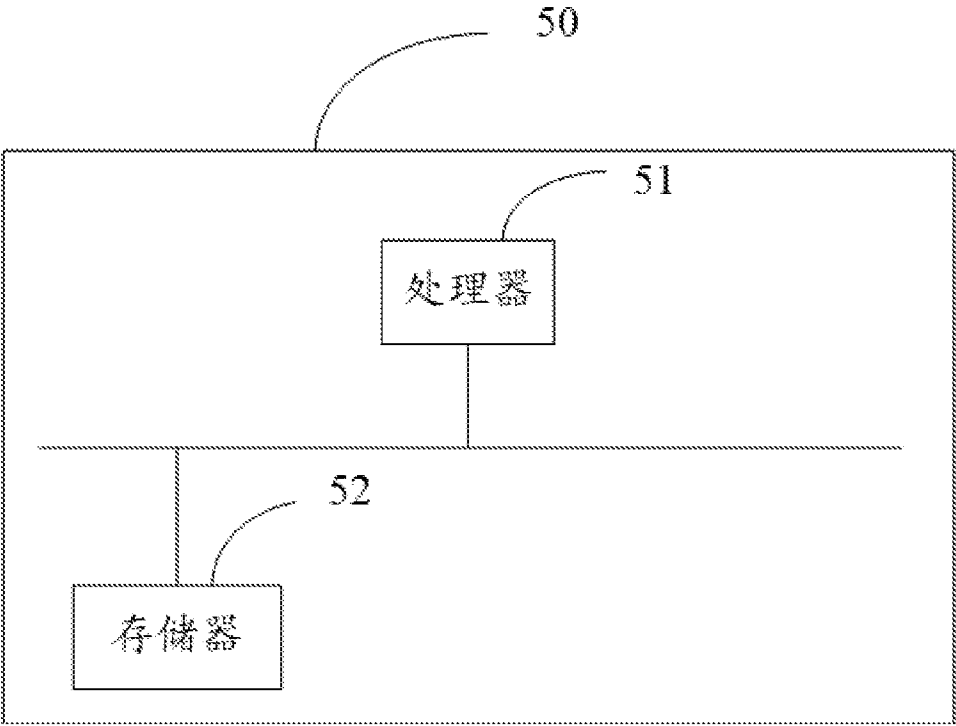


图5

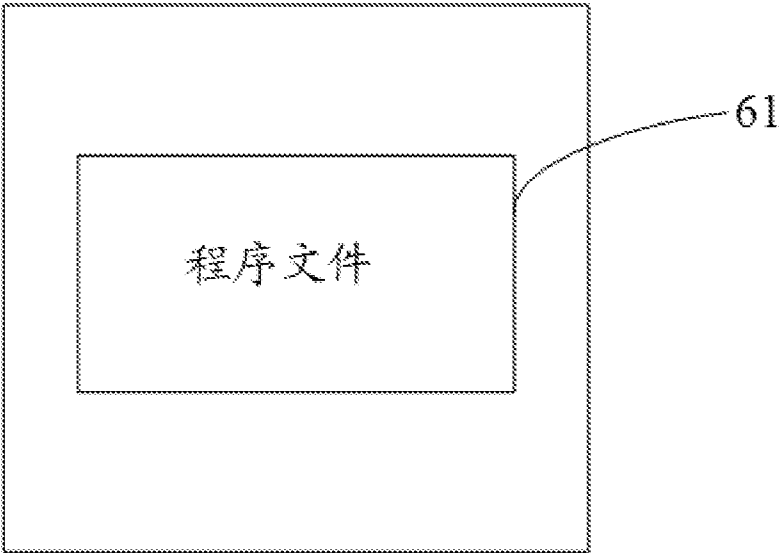


图6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/139297

A. CLASSIFICATION OF SUBJECT MATTER

G06N 3/08(2006.01)i; G06N 3/02(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)

CNABS; CNTXT; CNKI; SIPOABS; DWPI; USTXT; WOTXT; EPTXT: 神经网络, 调优, 优化, 正则化, 损失函数, 类内, 间距, 接近, neural network, optimize, regularization, loss function, clustering

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 111860834 A (SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY, CHINESE ACADEMY OF SCIENCES) 30 October 2020 (2020-10-30) claims 1-10	1-10
Y	摒除杂念 (non-official translation: Bing Chu Za Nian). "“神经网络优化---正则化 (正则化损失函数) (non-official translation: "Neural Network Optimization: Regularization (Regularized Loss Function)")" https://blog.csdn.net/congcong7267/article/details/81324467 , 01 August 2018 (2018-08-01), pp. 1-7	1, 5-10
A	摒除杂念 (non-official translation: Bing Chu Za Nian). "“神经网络优化---正则化 (正则化损失函数) (non-official translation: Neural Network Optimization: Regularization (Regularized Loss Function)")" https://blog.csdn.net/congcong7267/article/details/81324467 , 01 August 2018 (2018-08-01), pp. 1-7	2-4
Y	新垣结衣 (Yui Arakagi). "“神经网络与深度学习chap-网络优化与正则化” (non-official translation: Neural Network and Deep Learning CHAP-Network Optimization and Regularization)" http://www.doc88.com/p-8711736075598.html , 04 May 2019 (2019-05-04), pp. 1-3	1, 5-10



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

01 March 2021

Date of mailing of the international search report

02 April 2021

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/139297

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	新垣结衣 (Yui Arakagi). "“神经网络与深度学习chap-网络优化与正则化” (non-official translation: Neural Network and Deep Learning CHAP-Network Optimization and Regularization)" http://www.doc88.com/p-8711736075598.html , 04 May 2019 (2019-05-04), pp. 1-3	2-4
A	CN 110598552 A (JILIN UNIVERSITY) 20 December 2019 (2019-12-20)	1-10
A	WO 2019208564 A1 (NIPPON TELEGRAPH & TELEPHONE CORPORATION) 31 October 2019 (2019-10-31) entire document	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/139297

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	111860834	A	30 October 2020	None	
CN	110598552	A	20 December 2019	None	
WO	2019208564	A1	31 October 2019	None	

国际检索报告

国际申请号

PCT/CN2020/139297

A. 主题的分类

G06N 3/08(2006.01)i; G06N 3/02(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))

CNABS;CNTXT;CNKI;SIPOABS;DWPI;USTXT;WOTXT;EPTXT: 神经网络, 调优, 优化, 正则化, 损失函数, 类内, 间距, 接近, neural network, optimize, regularization, loss function, clustering

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 111860834 A (中国科学院深圳先进技术研究院) 2020年 10月 30日 (2020 - 10 - 30) 权利要求第1-10项	1-10
Y	摒除杂念. " "神经网络优化----正则化 (正则化损失函数) " " https://blog.csdn.net/congcong7267/article/details/81324467 , 2018年 8月 1日 (2018 - 08 - 01), 第1-7页	1, 5-10
A	摒除杂念. " "神经网络优化----正则化 (正则化损失函数) " " https://blog.csdn.net/congcong7267/article/details/81324467 , 2018年 8月 1日 (2018 - 08 - 01), 第1-7页	2-4
Y	新垣结衣. " "神经网络与深度学习chap-网络优化与正则化" " http://www.doc88.com/p-8711736075598.html , 2019年 5月 4日 (2019 - 05 - 04), 第1-3页	1, 5-10
A	新垣结衣. " "神经网络与深度学习chap-网络优化与正则化" " http://www.doc88.com/p-8711736075598.html , 2019年 5月 4日 (2019 - 05 - 04), 第1-3页	2-4

☒ 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

2021年 3月 1日

国际检索报告邮寄日期

2021年 4月 2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

授权官员

李娇

电话号码 (86-512) 88995726

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
A	CN 110598552 A (吉林大学) 2019年 12月 20日 (2019 - 12 - 20)	1-10
A	WO 2019208564 A1 (NIPPON TELEGRAPH & TELEPHONE) 2019年 10月 31日 (2019 - 10 - 31) 全文	1-10

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/139297

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	111860834	A	2020年 10月 30日	无	
CN	110598552	A	2019年 12月 20日	无	
WO	2019208564	A1	2019年 10月 31日	无	