



(12)发明专利申请

(10)申请公布号 CN 108985456 A

(43)申请公布日 2018.12.11

(21)申请号 201810823422.7

(22)申请日 2018.07.25

(71)申请人 大国创新智能科技(东莞)有限公司

地址 523000 广东省东莞市松山湖高新技术产业开发区创意生活城商场B二楼商场2部份场地(201号)

(72)发明人 朱定局

(74)专利代理机构 广州市华学知识产权代理有限公司 44245

代理人 李君

(51)Int.Cl.

G06N 3/08(2006.01)

G06N 3/04(2006.01)

权利要求书2页 说明书10页 附图4页

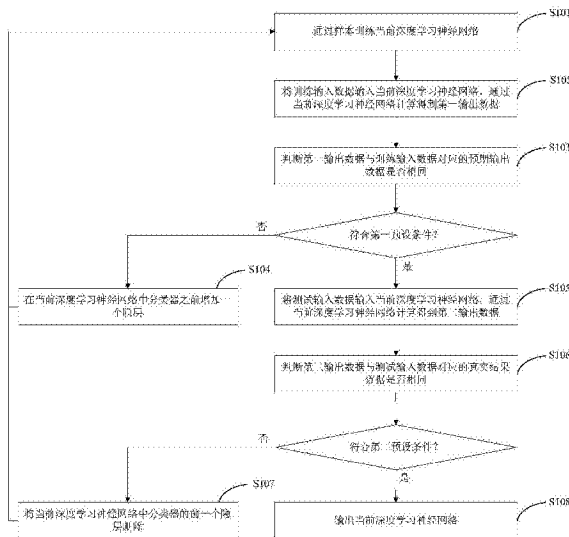
(54)发明名称

层数增减深度学习神经网络训练方法、系统、介质和设备

(57)摘要

本发明公开了一种层数增减深度学习神经网络训练方法、系统、介质和设备,所述方法包括:将训练输入数据输入当前深度学习神经网络,通过当前深度学习神经网络计算得到第一输出数据;判断第一输出数据与预期输出数据是否相同;若不符合第一预设条件,则在当前深度学习神经网络中分类器之前增加一个隐层;否则,将测试输入数据输入当前深度学习神经网络,通过深度学习神经网络计算得到第二输出数据;判断第二输出数据与真实结果数据是否相同;若不符合第二预设条件,则将当前深度学习神经网络中分类器的前一个隐层删除;否则,输出当前深度学习神经网络。本发明能达到充分拟合时顶层概念就是刚好足以与输出数据进行充分拟合的概念。

CN 108985456 A



1. 层数增减深度学习神经网络训练方法,其特征在于:所述方法包括:

通过样本训练当前深度学习神经网络;其中,所述当前深度学习神经网络包括输入层、隐层、分类器和输出层;

将训练输入数据输入当前深度学习神经网络,通过当前深度学习神经网络计算得到第一输出数据;

判断第一输出数据与训练输入数据对应的预期输出数据是否相同;

当第一输出数据与训练输入数据对应的预期输出数据不相同的数量不符合第一预设条件时,在当前深度学习神经网络中分类器之前增加一个隐层;

当第一输出数据与训练输入数据对应的预期输出数据不相同的数量符合第一预设条件时,将测试输入数据输入当前深度学习神经网络,通过深度学习神经网络计算得到第二输出数据;

判断第二输出数据与测试输入数据对应的真实结果数据是否相同;

当第二输出数据与测试输入数据对应的真实结果数据不相同的数量不符合第二预设条件时,将当前深度学习神经网络中分类器的前一个隐层删除;

当第二输出数据与测试输入数据对应的真实结果数据不相同的数量符合第二预设条件时,输出当前深度学习神经网络。

2. 根据权利要求1所述的层数增减深度学习神经网络训练方法,其特征在于:所述在当前深度学习神经网络中分类器之前增加一个隐层,具体为:以最后一个隐层的输出与新插入的隐层的输入通过编解码网络相连,以新插入的隐层的输出作为当前深度学习神经网络中分类器的输入。

3. 根据权利要求2所述的层数增减深度学习神经网络训练方法,其特征在于:所述新插入的隐层的节点数小于或等于最后一个隐层的节点数。

4. 根据权利要求1所述的层数增减深度学习神经网络训练方法,其特征在于:所述将当前深度学习神经网络中分类器的前一个隐层删除,具体为:以倒数第二个隐层的节点为当前深度学习神经网络中分类器的输入节点。

5. 根据权利要求4所述的层数增减深度学习神经网络训练方法,其特征在于:所述倒数第二个隐层的节点数大于最后一个隐层的节点数。

6. 根据权利要求1-5任一项所述的层数增减深度学习神经网络训练方法,其特征在于:所述第一预设条件包括:第一输出数据与训练输入数据对应的预期输出数据的误差率小于或等于第一预设阈值;

所述第一输出数据与训练输入数据对应的预期输出数据的误差率的计算过程为:第一输出数据与训练输入数据对应的预期输出数据不相同的数量除以训练输入数据进行测试的总数。

7. 根据权利要求1-5任一项所述的层数增减深度学习神经网络训练方法,其特征在于:所述第二预设条件包括:第二输出数据与测试输入数据对应的真实结果数据的误差率小于或等于第二预设阈值;

所述第二输出数据与测试输入数据对应的真实结果数据的误差率的计算过程为:第二输出数据与测试输入数据对应的真实结果数据不相同的数量除以测试输入数据进行测试的总数。

8. 层数增减深度学习神经网络训练系统,其特征在于:所述系统包括:

训练模块,用于通过样本训练当前深度学习神经网络;其中,所述当前深度学习神经网络包括输入层、隐层、分类器和输出层;

第一输入模块,用于将训练输入数据输入当前深度学习神经网络,通过当前深度学习神经网络计算得到第一输出数据;

第一判断模块,用于判断第一输出数据与训练输入数据对应的预期输出数据是否相同;

隐层增加模块,用于当第一输出数据与训练输入数据对应的预期输出数据不相同的数量不符合第一预设条件时,在当前深度学习神经网络中分类器之前增加一个隐层;

第二输入模块,用于当第一输出数据与训练输入数据对应的预期输出数据不相同的数量符合第一预设条件时,将测试输入数据输入当前深度学习神经网络,通过深度学习神经网络计算得到第二输出数据;

第二判断模块,用于判断第二输出数据与测试输入数据对应的真实结果数据是否相同;

隐层删除模块,用于当第二输出数据与测试输入数据对应的真实结果数据不相同的数量不符合第二预设条件时,将当前深度学习神经网络中分类器的前一个隐层删除;

输出模块,用于当第二输出数据与测试输入数据对应的真实结果数据不相同的数量符合第二预设条件时,输出当前深度学习神经网络。

9. 存储介质,存储有程序,其特征在于:所述程序被处理器执行时,实现权利要求1-7任一项所述的层数增减深度学习神经网络训练方法。

10. 计算设备,包括处理器以及用于存储处理器可执行程序存储器,其特征在于:所述处理器执行存储器存储的程序时,实现权利要求1-7任一项所述的层数增减深度学习神经网络训练方法。

## 层数增减深度学习神经网络训练方法、系统、介质和设备

### 技术领域

[0001] 本发明涉及一种深度学习神经网络训练方法,尤其是一种层数增减深度学习神经网络训练方法、系统、介质和设备,属于神经网络训练领域。

### 背景技术

[0002] 现有深度学习技术能通过输入数据得到输出标签(例如通过头像得到该人身份证号,又如通过语音得到该人身份证号),在自顶向下的监督训练阶段必须要通过带标签数据的监督训练(例如带有身份证号标签的头像,又如带有身份证号标签的语音)。

[0003] 但是,现有深度学习技术自顶向下监督训练,要么只是调节输出层与隐层之间的网络权值,要么调节所有层的网络权值。当顶层概念的类别比标签类别要多时,如果只调节输出层与隐层之间的分类器网络权值,如果分类器的网络结构比较简单,反复调节分类器的网络参数的结果往往是符合了这个输出标签又无法符合那个输出标签,也就是说无法实现充分的拟合。如果把分类器的网络结构设计得非常复杂,例如以层次复杂的BP神经网络作为分类器,那么又会出现过拟合的情况,使得某些关键特征在拟合中被舍弃,从而使得对样本而言分类结果完全正确,但到应用时就会发现不对。

[0004] 可见,仅仅对输出层与隐层之间层次进行监督训练要么会无法充分拟合,要么会过拟合,都会导致深度学习在应用时失败。如果调节所有层的网络权值,那么又会破坏隐层中的认知权重和生成权重,使得调节之后得到的概念和景象不再完全是源于输入数据的特征和景象,而是为了输出标签的需要而被扭曲的特征和景象,同样会出现过拟合的现象,从而使得对样本而言分类结果完全正确,但到应用时就会发现不对。

### 发明内容

[0005] 本发明的第一个目的是为了解决上述现有技术的缺陷,提供了一种层数增减深度学习神经网络训练方法,该方法使得顶层概念刚好具备足够区分不同样本数据的特征信息,从而与预期输出数据、真实结果数据能够完全对应,能达到充分拟合时顶层概念就是刚好足以与输出数据进行充分拟合的概念。

[0006] 本发明的第二个目的在于提供一种层数增减深度学习神经网络训练系统。

[0007] 本发明的第三个目的在于提供一种存储介质。

[0008] 本发明的第四个目的在于提供一种计算设备。

[0009] 本发明的第一个目的可以通过采取如下技术方案达到:

[0010] 层数增减深度学习神经网络训练方法,所述方法包括:

[0011] 通过样本训练当前深度学习神经网络;其中,所述当前深度学习神经网络包括输入层、隐层、分类器和输出层;

[0012] 将训练输入数据输入当前深度学习神经网络,通过当前深度学习神经网络计算得到第一输出数据;

[0013] 判断第一输出数据与训练输入数据对应的预期输出数据是否相同;

[0014] 当第一输出数据与训练输入数据对应的预期输出数据不相同的数量不符合第一预设条件时,在当前深度学习神经网络中分类器之前增加一个隐层;

[0015] 当第一输出数据与训练输入数据对应的预期输出数据不相同的数量符合第一预设条件时,将测试输入数据输入当前深度学习神经网络,通过深度学习神经网络计算得到第二输出数据;

[0016] 判断第二输出数据与测试输入数据对应的真实结果数据是否相同;

[0017] 当第二输出数据与测试输入数据对应的真实结果数据不相同的数量不符合第二预设条件时,将当前深度学习神经网络中分类器的前一个隐层删除;

[0018] 当第二输出数据与测试输入数据对应的真实结果数据不相同的数量符合第二预设条件时,输出当前深度学习神经网络。

[0019] 进一步的,所述在当前深度学习神经网络中分类器之前增加一个隐层,具体为:以最后一个隐层的输出与新插入的隐层的输入通过编解码网络相连,以新插入的隐层的输出作为当前深度学习神经网络中分类器的输入。

[0020] 进一步的,所述新插入的隐层的节点数小于或等于最后一个隐层的节点数。

[0021] 进一步的,所述将当前深度学习神经网络中分类器的前一个隐层删除,具体为:以倒数第二个隐层的节点为当前深度学习神经网络中分类器的输入节点。

[0022] 进一步的,所述倒数第二个隐层的节点数大于最后一个隐层的节点数。

[0023] 进一步的,所述第一预设条件包括:第一输出数据与训练输入数据对应的预期输出数据的误差率小于或等于第一预设阈值;

[0024] 所述第一输出数据与训练输入数据对应的预期输出数据的误差率的计算过程为:第一输出数据与训练输入数据对应的预期输出数据不相同的数量除以训练输入数据进行测试的总数。

[0025] 进一步的,所述第二预设条件包括:第二输出数据与测试输入数据对应的真实结果数据的误差率小于或等于第二预设阈值;

[0026] 所述第二输出数据与测试输入数据对应的真实结果数据的误差率的计算过程为:第二输出数据与测试输入数据对应的真实结果数据不相同的数量除以测试输入数据进行测试的总数。

[0027] 本发明的第二个目的可以通过采取如下技术方案达到:

[0028] 层数增减深度学习神经网络训练系统,所述系统包括:

[0029] 训练模块,用于通过样本训练当前深度学习神经网络;其中,所述当前深度学习神经网络包括输入层、隐层、分类器和输出层;

[0030] 第一输入模块,用于将训练输入数据输入当前深度学习神经网络,通过当前深度学习神经网络计算得到第一输出数据;

[0031] 第一判断模块,用于判断第一输出数据与训练输入数据对应的预期输出数据是否相同;

[0032] 隐层增加模块,用于当第一输出数据与训练输入数据对应的预期输出数据不相同的数量不符合第一预设条件时,在当前深度学习神经网络中分类器之前增加一个隐层;

[0033] 第二输入模块,用于当第一输出数据与训练输入数据对应的预期输出数据不相同的数量符合第一预设条件时,将测试输入数据输入当前深度学习神经网络,通过深度学习

神经网络计算得到第二输出数据；

[0034] 第二判断模块,用于判断第二输出数据与测试输入数据对应的真实结果数据是否相同；

[0035] 隐层删除模块,用于当第二输出数据与测试输入数据对应的真实结果数据不相同的数量不符合第二预设条件时,将当前深度学习神经网络中分类器的前一个隐层删除；

[0036] 输出模块,用于当第二输出数据与测试输入数据对应的真实结果数据不相同的数量符合第二预设条件时,输出当前深度学习神经网络。

[0037] 本发明的第三个目的可以通过采取如下技术方案达到：

[0038] 存储介质,存储有程序,所述程序被处理器执行时,实现上述的层数增减深度学习神经网络训练方法。

[0039] 本发明的第四个目的可以通过采取如下技术方案达到：

[0040] 计算设备,包括处理器以及用于存储处理器可执行程序存储器,所述处理器执行存储器存储的程序时,实现上述的层数增减深度学习神经网络训练方法。

[0041] 本发明相对于现有技术具有如下的有益效果：

[0042] 1、本发明将训练输入数据输入深度学习神经网络,通过深度学习神经网络计算得到第一输出数据,如果第一输出数据与训练输入数据对应的预期输出数据不相同的数量不符合预设条件,则增加隐层的层数;如果第一输出数据与训练输入数据对应的预期输出数据不相同的数量符合预设条件,则将测试输入数据输入深度学习神经网络,通过深度学习神经网络计算得到第二输出数据,如果第二输出数据与测试输入数据对应的测试输出数据不相同的数量不符合预设条件,则减少隐层的层数,通过增加或减少隐层的层数,直到刚好达到充分拟合,即实现了拟合的纠正,使得顶层概念刚好具备足够区分不同样本数据的特征信息,从而与预期输出数据、真实结果数据能够完全对应,能达到充分拟合时顶层概念就是刚好足以与输出标签进行充分拟合的概念。

[0043] 2、本发明在增加深度学习神经网络的隐层时,以最后一个隐层的输出与新插入的隐层的输入通过编解码网络相连,以新插入的隐层的输出作为分类器的输入,新插入的隐层的节点数小于或等于最后一个隐层的节点数,可以使得输入分类器的顶层概念更为抽象,忽略掉无法与输出标签进行映射的特征,抽象出能与输出数据能充分对应的特征。

[0044] 3、本发明在减少深度学习神经网络的隐层时,以倒数第二个隐层的节点为当前深度学习神经网络中分类器的输入节点,倒数第二个隐层的节点数一般都大于最后一个隐层的节点数,这样可以使得输入分类器的顶层概念更为具体,具体出被忽略的能与输出数据能充分对应的特征。

## 附图说明

[0045] 图1为本发明实施例1的层数增减深度学习神经网络训练方法的流程图。

[0046] 图2为本发明实施例1的层数增减深度学习神经网络训练方法中深度学习神经网络的模型图。

[0047] 图3为本发明实施例1的层数增减深度学习神经网络训练方法中深度学习神经网络插入新隐层的示意图。

[0048] 图4为本发明实施例1的层数增减深度学习神经网络训练方法中深度学习神经网络

络删除隐层的示意图。

[0049] 图5为本发明实施例2的层数增减深度学习神经网络训练系统的结构框图。

### 具体实施方式

[0050] 下面结合实施例及附图对本发明作进一步详细的描述,但本发明的实施方式不限于此。

[0051] 实施例1:

[0052] 深度学习的说明和建立过程如下:

[0053] 从一个输入中产生一个输出所涉及的计算可以通过一个流向图 (flow graph) 来表示:流向图是一种能够表示计算的图,在这种图中每一个节点表示一个基本的计算以及一个计算的值,计算的结果被应用到这个节点的子节点的值。考虑这样一个计算集合,它可以被允许在每一个节点和可能的图结构中,并定义了一个函数族。输入节点没有父节点,输出节点没有子节点。

[0054] 这种流向图的一个特别属性是深度 (depth):从一个输入到一个输出的最长路径的长度。

[0055] 把学习结构看作一个网络,则深度学习的核心思路如下:

[0056] 第一步:采用自下而上的无监督训练

[0057] 1) 逐层构建单层神经元。

[0058] 2) 每层采用wake-sleep算法进行调优。每次仅调整一层,逐层调整。

[0059] 这个过程可以看作是一个feature learning的过程,是和传统神经网络区别最大的部分。wake-sleep算法:1) wake阶段:认知过程,通过下层的输入特征 (Input) 和向上的认知 (Encoder) 权重产生每一层的抽象表示 (Code),再通过当前的生成 (Decoder) 权重产生一个重建信息 (Reconstruction),计算输入特征和重建信息残差,使用梯度下降修改层间的下行生成 (Decoder) 权重。也就是“如果现实跟我想象的不一样,改变我的生成权重使得我想象的东西变得与现实一样”。

[0060] 2) sleep阶段:

[0061] 生成过程,通过上层概念 (Code) 和向下的生成 (Decoder) 权重,生成下层的状态,再利用认知 (Encoder) 权重产生一个抽象景象。利用初始上层概念和新建抽象景象的残差,利用梯度下降修改层间向上的认知 (Encoder) 权重。也就是“如果梦中的景象不是我脑中的相应概念,改变我的认知权重使得这种景象在我看来就是这个概念”。

[0062] 第二步:自顶向下的监督训练

[0063] 这一步是在第一步学习获得各层参数进的基础上,在最顶的编码层添加一个分类器 (例如罗杰斯特回归、SVM等),而后通过带标签数据的监督训练,利用梯度下降法去微调整个网络参数。

[0064] 深度学习的第一步实质上是一个网络参数初始化过程。区别于传统神经网络初值随机初始化,深度学习神经网络是通过无监督训练输入数据的结构得到的,因而这个初值更接近全局最优,从而能够取得更好的效果。

[0065] 对深度学习神经网络进行监督训练的过程中,如果层数过多,那么会使得顶层概念过于抽象、缺少判别的细节,而在监督训练过程中加入了刚好能区分输出标签的“噪音”

细节,从而使得顶层概念与标签过于拟合,在训练中为了拟合而在顶层概念中加入了非区别性特征的噪音,必然导致在后面的测试中会加大错误率。例如标签包括“白色男人”、“白色女人”、“黑色男人”、“黑色女人”,而顶层概念抽象到了“男人”、“女人”特征,忽略了“黑色”、“白色”特征,此时所有样本数据都能拟合到“男人”、“女人”这2个标签上,因为深度学习采用的是自下而上的无监督训练和自顶向下的监督训练,在自下而上的无监督训练中,“白色男人”、“黑色男人”这几类样本数据,显然都能与“男人”顶层概念对应,“白色女人”、“黑色女人”这几类样本数据,显然都能与“女人”顶层概念对应,在自顶向下的监督训练中,“白色男人”、“黑色男人”这几类标签,显然都能与“男人”顶层概念对应,“白色女人”、“黑色女人”这几类标签,显然都能与“女人”顶层概念对应,通过监督训练,深度学习神经网络会自动调整网络权值,最终会使得“男人+噪音1”顶层概念对应“黑色男人”,“男人+噪音2”顶层概念对应“白色男人”,“女人+噪音3”顶层概念对应“黑色女人”,“女人+噪音4”顶层概念对应“白色女人”,由于在监督训练过程中进行了反复拟合,从而对训练数据而言是达到了充分拟合的效果。但在应用时,输入“白色男人”测试数据时,得到的顶层概念“男人+噪音2”,但因为噪音2并不是区分白色男人与黑色男人的区别性特征,此时通过分类器对应到的标签可能是“白色男人”,也可能是“黑色男人”,甚至会由于噪音2的干扰得到的输出标签可能是“白色女人”或“黑色女人”,从而就会使得测试时错误率加大。

[0066] 如果层数过少,那么会使得顶层概念过于具体、与输出标签判别无关甚至矛盾的细节,从而使得顶层概念与标签不能拟合,也就是说无法使得顶层概念与标签一一对应起来。例如标签包括“男人”、“女人”,而顶层概念的特征除了包括区别男女的必要特征外,还包括头发特征、肤色特征。因为深度学习采用的是自下而上的无监督训练和自顶向下的监督训练,在自顶向下的监督训练中,如果刚开始有很多“短发男人”、“黑色男人”、“白色女人”、“长发女人”的训练样本,那么就会使得顶层概念中形成“短发男人”、“黑色男人”、“白色女人”、“长发女人”的概念,其中“短发男人”、“黑色男人”通过分类器对应“男人”输出标签,“白色女人”、“长发女人”通过分类器对应“女人”输出标签,但如果后来又有大量“长发男人”、“黑色女人”的训练样本,就会使得顶层概念调整为“长发男人”、“黑色男人”、“黑色女人”、“长发女人”的概念,但调整之后显然又无法拟合“短发男人”、“白色女人”那些样本,从而导致深度学习神经网络随着样本的变化不变的调整网络权值,却始终无法充分拟合,是由于网络层数不够,抽象程度不足,没有通过足够的抽象把非区别特征的细节剔除掉,所以需要增加网络层数。

[0067] 因此,本实施例提供了一种层数增减深度学习神经网络训练方法,该方法使得顶层概念刚好具备足够区分不同样本数据的特征信息,从而与预期输出数据、真实结果数据能够完全对应,能达到充分拟合时顶层概念就是刚好足以与输出数据进行充分拟合的概念。

[0068] 如图1所示,本实施例的层数增减深度学习神经网络训练方法包括以下步骤:

[0069] S101、通过样本训练当前深度学习神经网络。

[0070] 当前深度学习神经网络获取:初始化深度学习神经网络或获取预设深度学习神经网络,作为当前深度学习神经网络,如图2所示,深度学习神经网络包括输入层、隐层、分类器和输出层,其中隐层与输出层之间是分类器,该隐层是分类器的输入,输出层是分类器的输出。



[0071] 通过样本训练当前深度学习神经网络,例如:通过人脸图像及姓名标签,训练得到当前深度学习神经网络。

[0072] S102、将训练输入数据输入当前深度学习神经网络,通过当前深度学习神经网络计算得到第一输出数据。

[0073] 训练输入数据可以为人脸图像,具体地,训练输入数据可以是通过采集获取,例如:通过摄像头采集获取人脸图像;训练输入数据也可以是从数据库查找获取,例如:预先在数据库内存储人脸图像,从数据库中搜索人脸图像即可得到训练输入数据。

[0074] 第一输出数据可以为姓名标签,具体地,通过每一个人脸训练图像输入当前深度学习神经网络,通过深度学习神经网络计算得到输出的姓名标签。

[0075] S103、判断第一输出数据与训练输入数据对应的预期输出数据是否相同。

[0076] 具体地,判断输出的姓名标签与人脸训练图像对应的预期姓名标签是否相同,每一次姓名标签的输出,都记录在人脸训练图像进行测试的总数中,也即第一次输出姓名标签,人脸训练图像进行测试的总数记为1,第二次输出姓名标签,人脸训练图像进行测试的总数记为2,依次类推,第M次输出姓名标签,人脸训练图像进行测试的总数记为M,统计M次姓名标签中与人脸训练图像对应的预期姓名标签不相同的数量,将该数量记为A,将输出的姓名标签与人脸训练图像对应的预期姓名标签的误差率记为第一误差率,第一误差率利用姓名标签与人脸训练图像对应的预期姓名标签不相同的数量除以人脸训练图像进行测试的总数得到,即 $A/M$ ,若第一误差率小于或等于第一预设阈值,说明姓名标签与人脸训练图像对应的预期姓名标签不相同的数量符合第一预设条件,进入步骤S105,否则,即第一误差率大于第一预设阈值,说明姓名标签与人脸训练图像对应的预期姓名标签不相同的数量不符合第一预设条件,进入步骤S104。

[0077] 假设第一预设阈值为30%,第一误差率例如90%,此时第一误差率大于第一预设阈值,则说明无法充分拟合,即拟合程度不够,说明当前深度学习神经网络无法将训练数据有效映射到输出标签,说明隐层层数不够,无法蕴含足够的映射关系,所以需要增加隐层的层数;第一误差率例如20%,此时第一误差率小于或等于第一预设阈值,则说明拟合程度够了,所以接下来需要判断是否过分拟合了。

[0078] S104、在当前深度学习神经网络中分类器之前增加一个隐层。

[0079] 在样本学习的过程中深度学习神经网络如果未充分拟合,那会导致连对训练样本数据进行测试都不一定能得到正确的预期输出标签,从而导致了深度学习的失败。而在深度学习中,如果未充分拟合,说明样本数据的特征可能在认知过程中被加入了不必要的特性细节,而这特征细节往往会对分类造成干扰,因此就需要增加一个隐层,直到刚好达到充分拟合。

[0080] 该步骤是通过增加隐层的层数,来增加拟合的程度,具体地,以最后一个隐层的输出与新插入的隐层的输入通过编解码网络相连,以新插入的隐层的输出作为分类器的输入,插入新隐层的示意图如图3所示,图中圆圈标记的隐层为插入的新隐层,插入新隐层返回步骤S101继续执行。

[0081] 优选地,新插入的隐层的节点数可以小于或等于最后一个隐层的原先节点数,之所以让新插入的隐层的节点数可以小于或等于最后一个隐层的原先节点数,而不是大于最后一个隐层的原先节点数,是因为这样可以使得输入分类器的顶层概念更为抽象,忽略掉

无法与输出标签进行映射的特征,抽象出能与输出标签能充分对应的特征。

[0082] S105、将测试输入数据输入当前深度学习神经网络,通过当前深度学习神经网络计算得到第二输出数据。

[0083] 测试输入数据也可以为人脸图像,第二输出数据也可以为姓名标签,具体地,通过每一个人脸测试图像输入当前深度学习神经网络,通过深度学习神经网络计算得到输出的姓名标签。

[0084] S106、判断第二输出数据与测试输入数据对应的真实结果数据是否相同。

[0085] 具体地,判断输出的姓名标签与人脸测试图像对应的预期姓名标签是否相同,每一次姓名标签的输出,都记录在人脸测试图像进行测试的总数中,也即第一次输出姓名标签,人脸测试图像进行测试的总数记为1,第二次输出姓名标签,人脸测试图像进行测试的总数记为2,依次类推,第N次输出姓名标签,人脸测试图像进行测试的总数记为N,统计N次姓名标签中与人脸测试图像对应的真实姓名标签不相同的数量,将该数量记为B,将输出的姓名标签与人脸测试图像对应的真实姓名标签的误差率记为第二误差率,第二误差率利用姓名标签与人脸训练图像对应的真实姓名标签不相同的数量除以人脸测试图像进行测试的总数得到,即 $B/N$ ,若第二误差率小于或等于第二预设阈值,说明符合姓名标签与人脸测试图像对应的真实姓名标签不相同的数量第二预设条件,进入步骤S108,否则,即第二误差率大于第二预设阈值,说明姓名标签与人脸测试图像对应的真实姓名标签不相同的数量不符合第二预设条件,进入步骤S107。

[0086] 假设第二预设阈值为20%,第二误差率例如80%,此时第二误差率大于第二预设阈值,则说明过分拟合,说明测试输入数据的特征可能在认知过程中被忽略了必要的特性细节,而这特征细节恰恰是进行分类时所不可缺少了,因此就需要减少隐层的层数,直到刚好达到充分拟合;第二误差率例如10%,此时第二误差率小于或等于第二预设阈值,则说明刚好达到充分拟合。此外,第二误差率一般都小于第一误差率,因为即使刚好充分拟合的情况下,使用测试数据的准确度总是低于训练数据的准确度,所以第二误差率一般都会比第一误差率小。

[0087] S107、将当前深度学习神经网络中分类器的前一个隐层删除。

[0088] 在样本学习的过程中深度学习神经网络如果过分拟合,那会导致通过训练后样本数据输入后都得到正确的预期输出标签,但测试数据输入后很多得不到正确的输出标签,从而导致了深度学习的失败。而在深度学习中,如果过分拟合,说明样本数据的特征可能在认知过程中被忽略了必要的特性细节,而这特征细节恰恰是进行分类时所不可缺少了,因此就需要向下递减一个层,直到刚好达到充分拟合。

[0089] 该步骤是通过减少隐层的层数,来减少拟合的程度,具体地,将当前深度学习神经网络中分类器的前一个隐层删除,具体为:以倒数第二个隐层的节点为当前深度学习神经网络中分类器的输入节点,删除隐层的示意图如图4所示,图中圆圈标记的隐层为需要删除的隐层,删除隐层后返回步骤S101继续执行。

[0090] 优选地,倒数第二个隐层的节点数一般都大于最后一个隐层的节点数,这样可以使得输入分类器的顶层概念更为具体,具体出被忽略的能与输出标签能充分对应的特征。

[0091] S108、输出当前深度学习神经网络。

[0092] 第二误差率小于或等于第二预设阈值时,结束监督训练,输出当前深度学习神经

网络,作为刚好能够充分拟合的深度学习神经网络,能够充分拟合的深度学习神经网络指的是既不会不充分拟合,又不会过分拟合,而是恰到好处的拟合。此时,如果测试输入数据输入到当前深度学习神经网络,通过无监督学习得到的顶层概念,则该概念能最大程度地反映出测试输入数据与预期输出标签对应的特征,该特征输入分类器,自然能最大概率地得到与预期标签输出。

[0093] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分步骤是可以通程序来指令相关的硬件来完成,相应的程序可以存储于一计算机可读取存储介质中,所述的存储介质,如ROM/RAM、磁盘或光盘等。

[0094] 实施例2:

[0095] 如图5所示,本实施例提供了一种层数增减深度学习神经网络训练系统,该系统包括训练模块501、第一输入模块502、第一判断模块503、隐层增加模块504、第二输入模块505、第二判断模块506、隐层删除模块507和输出模块508,各个模块的具体功能如下:

[0096] 所述训练模块501,用于通过样本训练当前深度学习神经网络;其中,所述当前深度学习神经网络包括输入层、隐层、分类器和输出层。

[0097] 所述第一输入模块502,用于将训练输入数据输入当前深度学习神经网络,通过当前深度学习神经网络计算得到第一输出数据。

[0098] 所述第一判断模块503,用于判断第一输出数据与训练输入数据对应的预期输出数据是否相同。

[0099] 所述隐层增加模块504,用于当第一输出数据与训练输入数据对应的预期输出数据不相同的数量不符合第一预设条件时,在当前深度学习神经网络中分类器之前需要增加一个隐层,具体为:以最后一个隐层的输出与新插入的隐层的输入通过编解码网络相连,以新插入的隐层的输出作为当前深度学习神经网络中分类器的输入;其中,第一预设条件包括:第一输出数据与训练输入数据对应的预期输出数据的误差率小于或等于第一预设阈值。

[0100] 所述第二输入模块505,用于当第一输出数据与训练输入数据对应的预期输出数据不相同的数量符合第一预设条件时,将测试输入数据输入当前深度学习神经网络,通过深度学习神经网络计算得到第二输出数据。

[0101] 所述第二判断模块506,用于判断第二输出数据与测试输入数据对应的真实结果数据是否相同。

[0102] 所述隐层删除模块507,用于当第二输出数据与测试输入数据对应的真实结果数据不相同的数量不符合第二预设条件时,将当前深度学习神经网络中分类器的前一个隐层删除,具体为:以倒数第二个隐层的节点为当前深度学习神经网络中分类器的输入节点;其中,第二预设条件包括:第二输出数据与测试输入数据对应的真实结果数据的误差率小于或等于第二预设阈值。

[0103] 所述输出模块508,用于当第二输出数据与测试输入数据对应的真实结果数据不相同的数量符合第二预设条件时,输出当前深度学习神经网络。

[0104] 在此需要说明的是,上述实施例提供的系统仅以上述各功能模块的划分进行举例说明,在实际应用中,可以根据需要而将上述功能分配由不同的功能模块完成,即将内部结构划分成不同的功能模块,以完成以上描述的全部或者部分功能。

[0105] 可以理解,上述实施例的机器人系统所使用的术语“第一”、“第二”等可用于描述各种单元,但这些单元不受这些术语限制。这些术语仅用于将第一个模块与另一个模块区分。举例来说,在不脱离本发明的范围的情况下,可以将第一判断模块称为第二判断模块,且类似地,可将第二判断模块称为第一判断模块,第一判断模块和第二判断模块两者都是判断模块,但其不是同一判断模块。

[0106] 实施例3:

[0107] 本实施例提供了一种存储介质,该存储介质存储有一个或多个程序,所述程序被处理器执行时,实现上述实施例1的层数增减深度学习神经网络训练方法,如下:

[0108] 通过样本训练当前深度学习神经网络;其中,所述当前深度学习神经网络包括输入层、隐层、分类器和输出层;

[0109] 将训练输入数据输入当前深度学习神经网络,通过当前深度学习神经网络计算得到第一输出数据;

[0110] 判断第一输出数据与训练输入数据对应的预期输出数据是否相同;

[0111] 当第一输出数据与训练输入数据对应的预期输出数据不相同的数量不符合第一预设条件时,在当前深度学习神经网络中分类器之前增加一个隐层;

[0112] 当第一输出数据与训练输入数据对应的预期输出数据不相同的数量符合第一预设条件时,将测试输入数据输入当前深度学习神经网络,通过深度学习神经网络计算得到第二输出数据;

[0113] 判断第二输出数据与测试输入数据对应的真实结果数据是否相同;

[0114] 当第二输出数据与测试输入数据对应的真实结果数据不相同的数量不符合第二预设条件时,将当前深度学习神经网络中分类器的前一个隐层删除;

[0115] 当第二输出数据与测试输入数据对应的真实结果数据不相同的数量符合第二预设条件时,输出当前深度学习神经网络。

[0116] 本实施例中所述的存储介质可以是ROM、RAM、磁盘、光盘等介质。

[0117] 实施例4:

[0118] 本实施例提供了一种计算设备,该计算设备包括处理器和存储器,存储器存储有一个或多个程序,处理器执行存储器存储的程序时,实现上述实施例1的层数增减深度学习神经网络训练方法,如下:

[0119] 通过样本训练当前深度学习神经网络;其中,所述当前深度学习神经网络包括输入层、隐层、分类器和输出层;

[0120] 将训练输入数据输入当前深度学习神经网络,通过当前深度学习神经网络计算得到第一输出数据;

[0121] 判断第一输出数据与训练输入数据对应的预期输出数据是否相同;

[0122] 当第一输出数据与训练输入数据对应的预期输出数据不相同的数量不符合第一预设条件时,在当前深度学习神经网络中分类器之前增加一个隐层;

[0123] 当第一输出数据与训练输入数据对应的预期输出数据不相同的数量符合第一预设条件时,将测试输入数据输入当前深度学习神经网络,通过深度学习神经网络计算得到第二输出数据;

[0124] 判断第二输出数据与测试输入数据对应的真实结果数据是否相同;

[0125] 当第二输出数据与测试输入数据对应的真实结果数据不相同的数量不符合第二预设条件时,将当前深度学习神经网络中分类器的前一个隐层删除;

[0126] 当第二输出数据与测试输入数据对应的真实结果数据不相同的数量符合第二预设条件时,输出当前深度学习神经网络。

[0127] 本实施例中所述的计算设备可以是台式电脑、笔记本电脑、智能手机、PDA手持终端、平板电脑或其他具有显示功能的终端设备。

[0128] 综上所述,本发明将训练输入数据输入深度学习神经网络,通过深度学习神经网络计算得到第一输出数据,如果第一输出数据与训练输入数据对应的预期输出数据不相同的数量不符合预设条件,则增加隐层的层数;如果第一输出数据与训练输入数据对应的预期输出数据不相同的数量符合预设条件,则将测试输入数据输入深度学习神经网络,通过深度学习神经网络计算得到第二输出数据,如果第二输出数据与测试输入数据对应的测试输出数据不相同的数量不符合预设条件,则减少隐层的层数,通过增加或减少隐层的层数,直到刚好达到充分拟合,即实现了拟合的纠正,使得顶层概念刚好具备足够区分不同样本数据的特征信息,从而与预期输出数据、真实结果数据能够完全对应,能达到充分拟合时顶层概念就是刚好足以与输出标签进行充分拟合的概念。

[0129] 以上所述,仅为本发明专利较佳的实施例,但本发明专利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发明专利所公开的范围内,根据本发明专利的技术方案及其发明构思加以等同替换或改变,都属于本发明专利的保护范围。

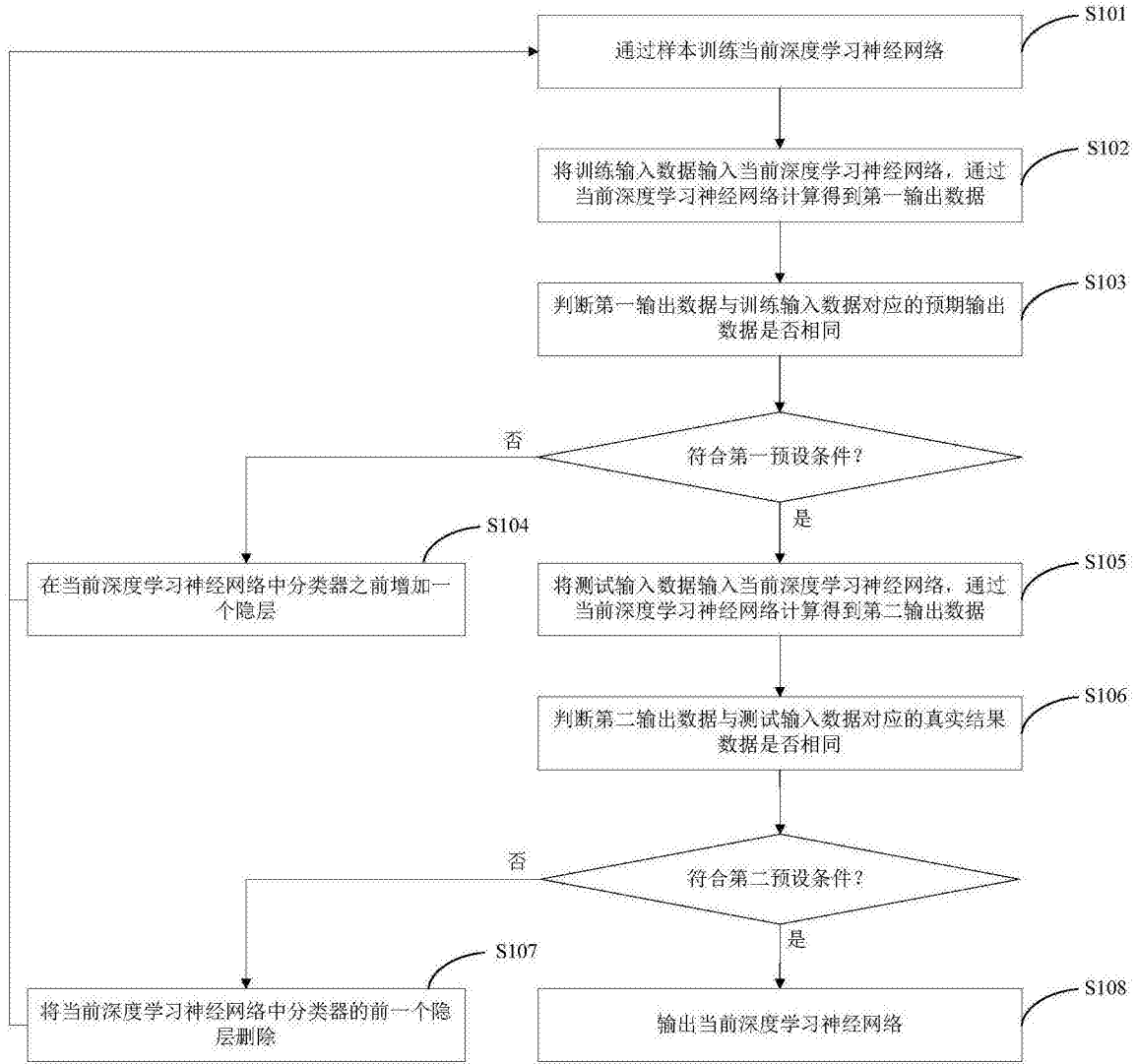


图1

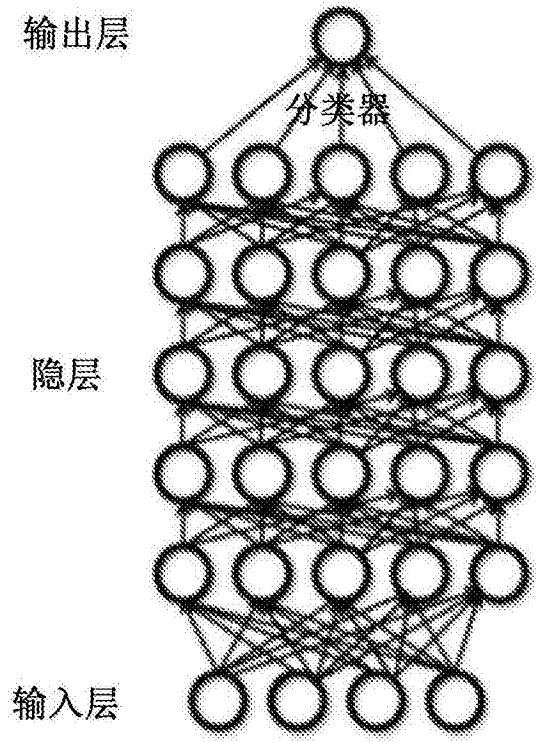


图2

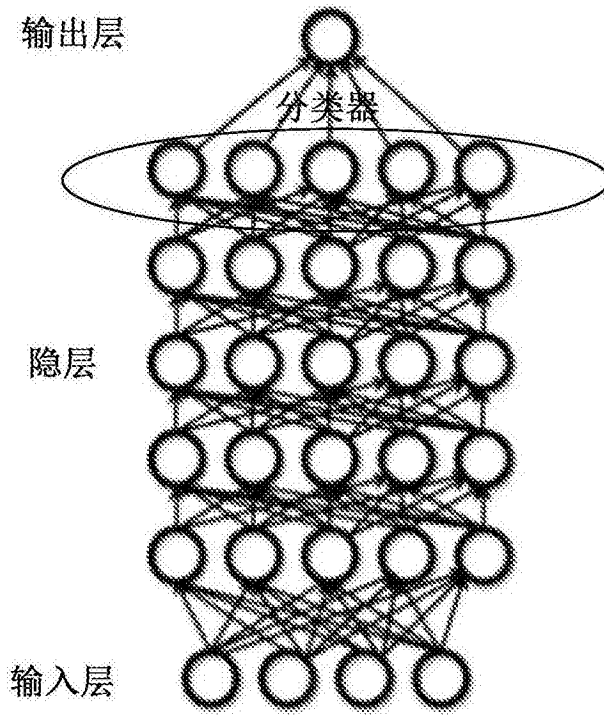


图3

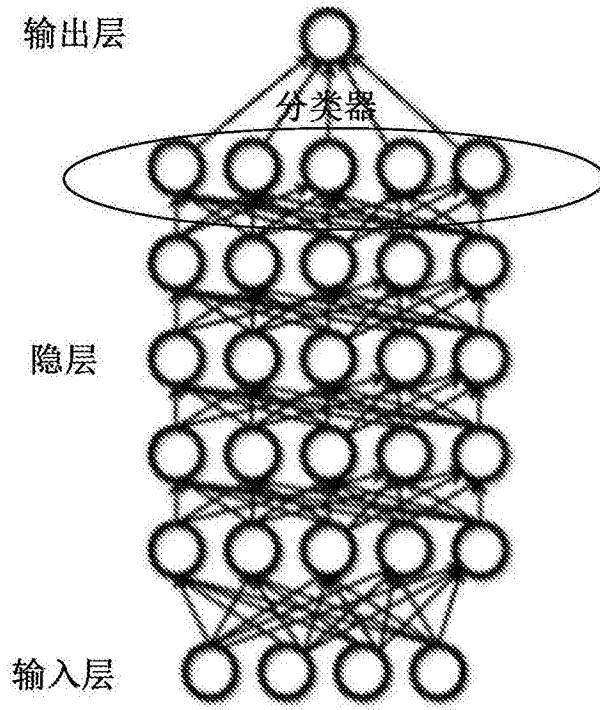


图4



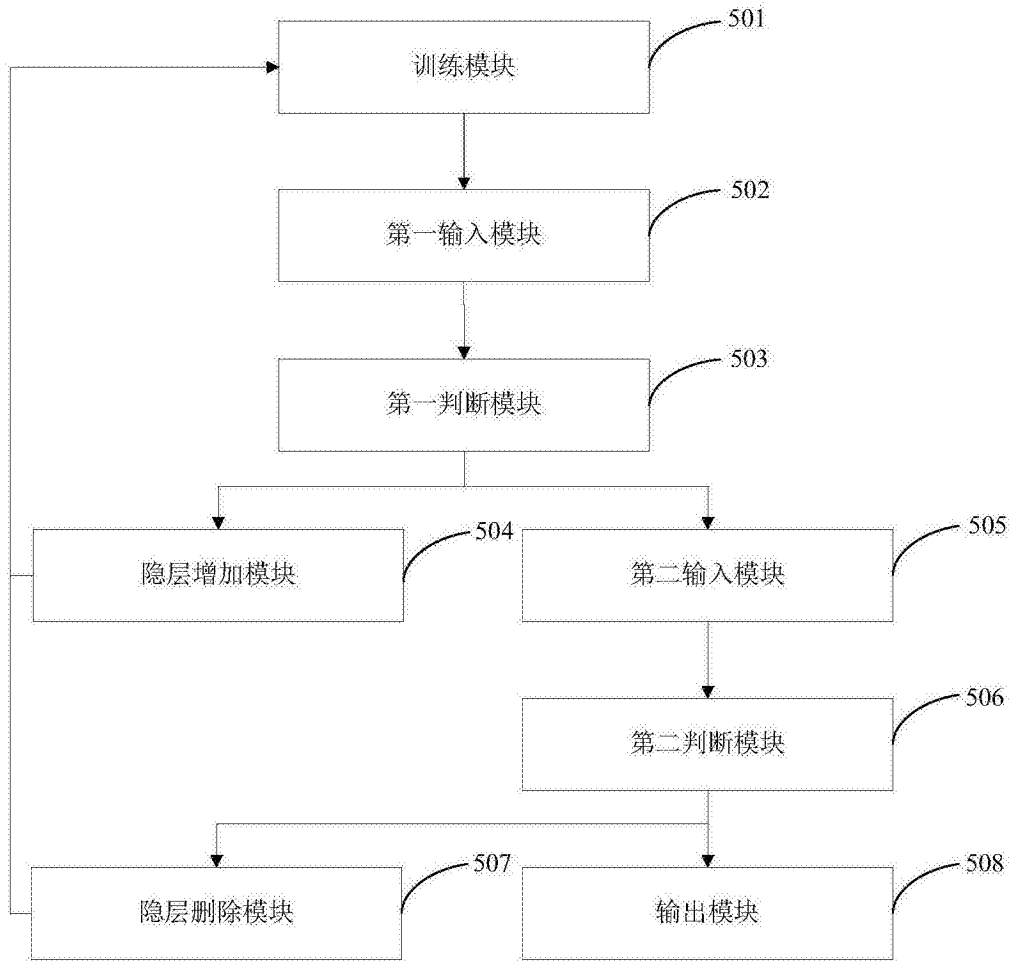


图5