



(12) 发明专利

(10) 授权公告号 CN 112534447 B

(45) 授权公告日 2024. 05. 24

(21) 申请号 201980051564.9

克里斯托弗·蒂茨

(22) 申请日 2019.07.29

(74) 专利代理机构 北京康信知识产权代理有限公司 11240

(65) 同一申请的已公布的文献号
申请公布号 CN 112534447 A

专利代理师 李海霞

(43) 申请公布日 2021.03.19

(51) Int.Cl.

(30) 优先权数据

G06N 3/08 (2023.01)

18187335.7 2018.08.03 EP

G06N 20/00 (2019.01)

G06N 3/047 (2023.01)

(85) PCT国际申请进入国家阶段日
2021.02.02

(56) 对比文件

(86) PCT国际申请的申请数据

PCT/IB2019/056435 2019.07.29

CN 105531724 A, 2016.04.27

CN 106462797 A, 2017.02.22

CN 106980822 A, 2017.07.25

(87) PCT国际申请的公布数据

W02020/026107 DE 2020.02.06

CN 107492090 A, 2017.12.19

US 2006248031 A1, 2006.11.02

US 2010063948 A1, 2010.03.11

US 2018107926 A1, 2018.04.19

US 2018144466 A1, 2018.05.24

(73) 专利权人 西门子医疗保健诊断公司
地址 美国纽约

审查员 裴强

(72) 发明人 斯特凡·德佩韦格

马库斯·迈克尔·盖佩尔

加比·马夸特 丹尼拉·塞德尔

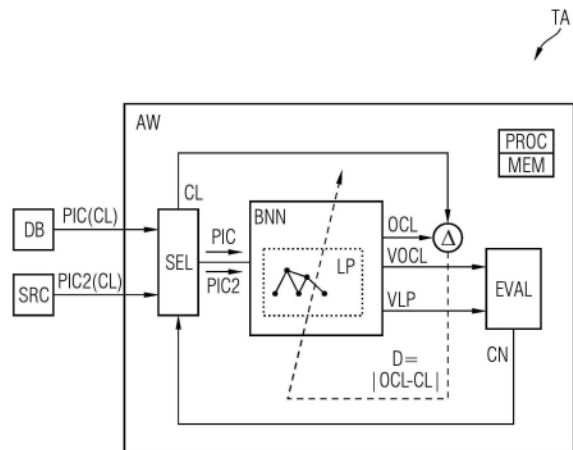
权利要求书2页 说明书7页 附图2页

(54) 发明名称

训练用于控制工程系统的机器学习例程的方法和装置

(57) 摘要

为了训练机器学习例程 (BNN), 通过机器学习例程 (BNN) 读取第一训练数据 (PIC) 的序列。根据第一训练数据 (PIC) 训练机器学习例程 (BNN), 其中, 通过训练调节机器学习例程 (BNN) 的多个学习参数 (LP)。另外, 得出学习参数 (LP) 的在训练中出现的值分布 (VLP), 并且根据学习参数 (LP) 的得出的值分布 (VLP) 来产生继续消息 (CN)。然后, 根据继续消息 (CN) 利用第一训练数据 (PIC) 的另一序列继续训练, 或者对于训练请求其他训练数据 (PIC2)。



1. 一种用于训练用于控制工程系统 (TS) 的机器学习例程 (BNN) 的以计算机执行的方法, 其中

a) 通过机器学习例程 (BNN) 读取具有工程系统 (TS) 的第一传感器数据的第一训练数据 (PIC) 的序列,

b) 根据所述第一训练数据 (PIC) 训练所述机器学习例程 (BNN), 其中, 通过所述训练调节所述机器学习例程 (BNN) 的多个学习参数 (LP),

c) 得出所述学习参数 (LP) 的在所述训练中出现的值分布 (VLP),

d) 根据所述学习参数 (LP) 的所得出的所述值分布 (VLP) 产生继续消息 (CN), 并且

e) 根据所述继续消息 (CN) 利用所述第一训练数据 (PIC) 的另一序列继续所述训练, 或者为所述训练请求其他训练数据 (PIC2), 其中, 将所述第一训练数据的训练数据集分别与特定的识别类别 (CL) 相关联, 并且对于相应的识别类别 (CL), 根据与该识别类别 (CL) 相关联的所述第一训练数据集:

- 得出所述学习参数的类别特定的值分布,

- 根据所述学习参数的所述类别特定的值分布, 产生类别特定的继续消息, 并且

- 根据所述类别特定的继续消息, 利用与该识别类别相关联的第一训练数据继续所述训练, 或者请求与该识别类别 (CL) 相关联的其他训练数据。

2. 根据权利要求1所述的方法, 其特征在于, 所述其他训练数据 (PIC2)

- 具有与所述第一训练数据 (PIC) 不同的类型,

- 以与所述第一训练数据 (PIC) 不同的方式被检测,

- 包括与所述第一传感器数据不同的传感器数据, 和/或

- 来自与所述第一训练数据 (PIC) 不同的来源。

3. 根据权利要求1或2所述的方法, 其特征在于,

根据所述继续消息 (CN) 驱控传感器或传感器系统, 从而

- 修改用于检测训练数据的所述传感器的传感器参数,

- 促使所述传感器系统的传感器检测训练数据, 和/或

- 选择所述传感器系统的传感器数据作为训练数据。

4. 根据权利要求1或2所述的方法, 其特征在于,

得出所述学习参数 (LP) 的所述值分布 (VLP) 的分布宽度和/或熵值, 并且

随着分布宽度的增加和/或熵值的增加, 优选地利用所述第一训练数据 (PIC) 的另一序列继续所述训练。

5. 根据权利要求1或2所述的方法, 其特征在于,

得出所述学习例程 (BNN) 的输出数据 (OCL) 的值分布 (VOCL), 并且

根据所述输出数据 (OCL) 的值分布 (VOCL) 形成所述继续消息 (CN)。

6. 根据权利要求5所述的方法, 其特征在于, 得出所述输出数据 (OCL) 的值分布 (VOCL) 的分布宽度和/或熵值, 并且

随着分布宽度的增加和/或熵值的增加, 优选请求所述其他训练数据 (PIC2)。

7. 根据权利要求1或2所述的方法, 其特征在于, 由分布宽度代表相应的值分布 (VLP、VOCL)。

8. 根据权利要求1或2所述的方法, 其特征在于,

所述机器学习例程实施人工神经网络、递归神经网络、卷积神经网络、贝叶斯神经网络(BNN)、自编码器、深度学习架构、支持向量机、数据驱动的可训练回归模型、k最近邻分类器、物理模型和/或决策树。

9. 根据权利要求1或2所述的方法,其特征在于,

通过所述机器学习例程借助于贝叶斯神经网络(BNN)得出所述学习参数(LP)的所述值分布(VLP),其中,所述学习参数(LP)是贝叶斯神经网络(BNN)的神经权重。

10. 根据权利要求1或2所述的方法,其特征在于,

通过所述机器学习例程实施子机器学习例程的整体,通过所述子机器学习例程分别调节所述学习参数的集,并且经由学习参数集从所述学习参数(LP)的分布推导出所述学习参数(LP)的所述值分布(VLP)。

11. 根据权利要求1或2所述的方法,其特征在于,

所述机器学习例程实施神经网络,其中,所述学习参数(LP)是所述神经网络的神经权重,

多次变化所述神经网络的神经结构,并且为相应的变化得出所述神经网络的输出数据的集,并且

经由输出数据集从所述输出数据(OCL)的分布推导出所述学习参数(LP)的所述值分布(VLP)。

12. 一种用于训练用于控制工程系统(TS)的机器学习例程(BNN)的设备(TA),所述设备设置用于执行根据前述权利要求中任一项所述的方法。

13. 一种计算机程序产品,设置用于实施根据权利要求1至11中任一项所述的方法。

14. 一种计算机可读的存储介质,具有根据权利要求13所述的计算机程序产品。

训练用于控制工程系统的机器学习例程的方法和设备

技术领域

[0001] 训练用于控制工程系统的机器学习例程的方法和设备。

背景技术

[0002] 在控制复杂的工程系统、例如医学诊断装置、图像分析装置、机器人、自治系统、风力涡轮机、燃气轮机或制造设施时,越来越多地使用机器学习方法,以便关于预设的标准优化工程系统的能力。在此,尤其能够通过训练过的机器学习例程来控制工程系统的传感器数据的评估。

[0003] 提供大量已知的学习方法、即例如监督学习方法来训练机器学习例程。在此,尤其能够将要控制的工程系统或类似的工程系统的历史或当前的运行数据或传感器数据用作训练数据。

[0004] 因此,例如,在医学诊断仪器中,能够训练机器学习例程以在显微图像上专门对细胞或组织类型进行分类或识别或在断层扫描照片上专门分类和识别肿瘤。对于这样的训练,经常使用大量先前分类过的医学图像,该医学图像作为训练数据输送给机器学习例程。因此,能够训练机器学习例程来尽可能好地复现预设的分类,或者以与预设的分类尽可能小的偏差识别细胞类型、组织类型或肿瘤。与此类似,能够训练用于燃气轮机、制造设施或另一工程系统的机器学习例程来根据传感器数据或其他运行数据识别或分类特定的运行状态和/或以优化的方式控制工程系统。

[0005] 与此同时,在实践中经常发生,某些要识别的类别比其他识别类别显著更罕见地出现。相应地,相比于对于其他的识别类别的训练数据,对于极其罕见的识别类别通常存在显著更少的训练数据。

[0006] 但是,刚好在医学领域,正确识别罕见的病理模式或变化通常是重要的。在这种情况下,在评估有足够多的罕见类别训练数据从而也足够可靠地识别该类别之前,会持续非常长时间。

[0007] 与上述无关,根据预设的训练数据,与其他识别类别相比,会更难于将某些识别类别彼此区分。在这种情况下,经常发生:即使进行持续训练,也不再能够显著地改进识别可靠性。

[0008] 根据迄今为止的现有技术,经常通过如下方式考虑上述情况:逐步增加训练数据量并且在此观察机器学习例程的成功率。一旦成功率不再显著改进,尤其在罕见识别类别中也是如此的话,则能够结束训练。如果成功率始终保持较低,则这还能够被评价为以下指示:训练所使用的传感器数据不够良好地适于可靠识别。在这种情况下,通常利用其他的传感器数据继续或重复训练。

[0009] 通过上述处理方式,所需的训练耗费(尤其在存在罕见的识别类别的情况下)会显著增加。

发明内容

[0010] 发明的目的是提供一种用于训练用于控制工程系统的机器学习例程的方法和设备,该方法和设备允许更有效的训练。

[0011] 该目的通过具有本发明的特征的方法、通过具有权利要求13的特征的设备、通过具有本发明的特征的计算机程序产品以及通过具有本发明的特征的计算机可读的存储介质来实现。

[0012] 为了训练用于控制工程系统的机器学习例程,通过机器学习例程读入包括工程系统的第一传感器数据的第一训练数据的序列。在此,该工程系统尤其能够是医学诊断装置、图像分析装置、机器人、自主系统、风力涡轮机、燃气轮机或制造设施。根据第一训练数据来训练机器学习例程,其中,通过训练来调节机器学习例程的多个学习参数。另外,得出学习参数的在训练中出现的值分布,并根据学习参数的得出的值分布来产生继续消息。然后,根据继续消息,利用第一训练数据的另一序列继续训练,或者对于训练请求其他训练数据。

[0013] 证明了:学习参数的在训练中出现的值分布并且特别是其分布宽度通常是是否能够通过输送其他第一训练数据来改进训练成功与否的早期指标。在后一种情况下,已经能够提早请求质量上不同的训练数据,并用于进一步训练。以该方式,通常能够避免输送不再显著改进训练成效的其他第一训练数据。

[0014] 通过本发明,能够以有利的方式控制训练,并且尤其控制合适的训练数据的输送,进而通常能够显著地改进或缩短训练。

[0015] 为了执行根据本发明的方法,提出相应的设备、计算机程序产品以及计算机可读的存储介质。通过根据本发明的方法、根据本发明的设备或根据本发明的计算机程序产品能够例如训练辅助系统的、分析装置的、评估装置的或控制装置的机器学习例程。然后,通过训练过的机器学习例程能够执行工程系统的其他的运行数据或传感器数据的评估、分析、分类或预测。根据此,能够以优化的方式控制工程系统。

[0016] 能够例如借助于一个或多个处理器、专用集成电路(ASIC)、数字信号处理器(DSP)和/或所谓的“现场可编程门阵列”(FPGA)来执行或实施根据本发明的方法、根据本发明的设备以及根据本发明的计算机程序产品。

[0017] 本发明的有利的实施方式和改进形式在各个实施例中说明。

[0018] 根据本发明的一个有利的实施方式,其他训练数据能够具有与第一训练数据不同的类型,能够以与第一训练数据不同的方式被检测,包括与第一传感器数据不同的传感器数据和/或来自与第一训练数据不同的来源。特别地,其他训练数据能够包括更高分辨率,或更细化的传感器数据或包括其他物理或化学变量的测量数据。

[0019] 此外,根据继续消息,能够驱控传感器或传感器系统,使得修改用于检测训练数据的传感器的传感器参数,促使传感器系统的传感器检测训练数据,和/或选择传感器系统的传感器数据作为训练数据。

[0020] 根据本发明的一个有利的实施方式,能够得出学习参数的值分布的分布宽度和/或熵值。然后,随着分布宽度的增加和/或熵值的增加,能够优选地以第一训练数据的另一序列继续训练。这是有利的,因为在许多情况下能够将学习参数的值分布的较大的分布宽度或较大的熵值评价为尚未找到学习参数的最佳调节的指示。在这种情况下,通常能够期望通过利用第一训练数据进行进一步训练来改进训练成效。特别地,能够从各个学习参数

的值分布、从学习参数的组的值分布或从所有学习参数的值分布来得出相应的分布宽度。

[0021] 根据本发明的另一有利的实施方式,能够得出学习例程的输出数据的值分布,并且能够根据输出数据的值分布来形成继续消息。

[0022] 特别地,能够得出输出数据的值分布的分配宽度和/或熵值。随着分布宽度的增加和/或熵值的增加,然后能够优选请求其他训练数据。这是有利的,因为在许多情况下能够将输出数据的值分布的较大的分布宽度或较大的熵值评价为训练过的学习例程只能提供不可靠结果的指示。在这种情况下,通常能够通过切换到质量上不同的训练数据来改进训练成效。

[0023] 特别地,能够得出学习参数的值分布、或输出数据的值分布的离散度、方差或标准偏差作为分布宽度。

[0024] 有利地,相应的值分布能够通过识别类别之上的优选统计高斯分布或分类分布的分布宽度和/或平均值来代表、图示或建模。以该方式,能够节约空间地存储以及有效地处理值分布。

[0025] 根据本发明的一个有利的改进形式,第一训练数据的训练数据集能够分别与特定的识别类别相关联。优选地,通过存储在相应的训练数据集中的类别数据来实现与识别类别的关联。这种类别数据例如能够将生物细胞的显微图像与作为识别类别的特定的细胞类型相关联。对于相应的识别类别,能够根据与该识别类别相关联的第一训练数据集得出学习参数的类别特定的值分布,根据学习参数的类别特定的值分布来产生类别特定的继续消息,并根据类别特定的继续消息利用与该识别类别相关联的第一训练数据继续进行训练,或者请求与该识别类别相关联的其他训练数据。以此方式,能够类别特定地控制促进训练的训练数据的输送。特别地,能够针对相应的识别类别个体地得出:以另外的第一训练数据进行训练或切换到另外的训练数据预计是否是更有利的。相应地,能够类别特定地选择还要用于相应的识别类别的训练数据。

[0026] 根据本发明的一个特别有利的实施方式,能够通过机器学习例程借助于贝叶斯神经网络来得出学习参数的值分布,其中,通过贝叶斯神经网络的神经权重来实现学习参数。例如能够借助于马尔可夫链蒙特卡罗方法、拉普拉斯近似和/或推论方法尤其通过参数拟合来得出值的分布。这种贝叶斯神经网络特别适合于对统计分布、特别是学习参数的统计分布建模。

[0027] 根据本发明的另一实施方式,能够通过机器学习例程实施机器学习例程的整体,通过该子机器学习例程分别调节学习参数的集。然后,能够从学习参数在学习参数集上的分布推导出学习参数的值分布。

[0028] 根据本发明的另一实施方式,机器学习例程能够实施神经网络,其中,学习参数是神经网络的神经权重。在此,神经网络的神经结构能够多次变化,其中,为相应的变化得出神经网络的输出数据的集。然后,能够从输出数据在输出数据集上的分布推导出学习参数的值分布。为了变化神经结构,能够以不同的方式、优选随机地断开神经网络的神经元。这种断开通常也称为丢弃。不同的断开相应地引起输出数据的变化。从变化的输出数据的分布中以及从断开的模式中能够推断出学习参数的值分布。

附图说明

[0029] 下面根据附图更详细地阐述本发明的实施例。在此,分别以示意图示出:

[0030] 图1示出具有训练过的评估装置的医学诊断仪器,和

[0031] 图2示出根据本发明的用于训练评估装置的设备。

具体实施方式

[0032] 图1示例性地示出作为工程系统TS的医学诊断仪器的示意图,该医学诊断仪器具有至少一个传感器S和借助于机器学习方法训练的评估装置AW。后者用于评估传感器S的传感器数据。评估装置AW能够实施为诊断仪器TS的一部分,或者部分地在诊断仪器TS的外部实施。评估装置AW耦联至传感器S和耦联至诊断仪器TS的控制装置CTL。处理器控制或计算机控制的控制装置CTL用于控制诊断仪器TS。

[0033] 在当前的实施例中,示例性地关注诊断仪器TS,该诊断仪器应当根据生物组织或生物细胞的显微图像照片PIC自动识别组织或细胞的类型或状态和/或相应的病理类别。然后,根据识别,能够通过诊断仪器TS或其控制装置CTL采取其他措施,例如对所拍摄的生物组织或生物细胞进行特定处理或检查,即例如进行药物的施用和/或配量。

[0034] 相应地,在本实施例中的传感器S优选地被设计为显微镜,该显微镜拍摄生物组织或生物细胞的显微图像PIC并传输至评估装置AW。替代地或附加地,也能够设有其他的成像传感器,例如照相机或成像方法,例如断层摄影或超声检查方法,以便创建二维和/或三维图像照片。

[0035] 训练过的评估装置AW根据预设的识别类别CL、即例如根据它们的类型、状态或根据其病理类别,对在图像PIC上示出的组织或细胞进行分类。分别在图像PIC上识别的组织或细胞的类别CL通过评估装置AW分别以分类标识符的形式传输至控制装置CTL。如上所述,根据所传输的识别类别CL,通过控制装置CTL控制诊断仪器TS的另外的行为。

[0036] 替代地或附加地,机器人、自主系统、风力涡轮机、燃气轮机或生产设施作为工程系统TS也能够根据专门训练过的评估装置AW来控制。在此,评估装置AW能够被训练为根据工程系统TS的传感器数据或其他运行数据来识别或分类工程系统TS的特定运行状态和/或借助于控制装置CTL以优化的方式控制工程系统TS。

[0037] 在此,也将工程系统的控制理解为与控制相关的、即例如有助于控制工程系统的数据和控制信号的输出和使用。这种与控制相关的数据尤其能够包括分类数据、分析数据、预测数据和/或监控数据,这些数据尤其可用于监控工程系统和/或用于识别磨损识别和/或损坏识别。

[0038] 图2示出根据本发明的用于训练评估装置AW的设备TA的示意图。评估装置AW具有一个或多个用于实施评估装置AW的所有方法步骤的计算机或处理器PROC以及具有一个或多个与处理器PROC耦联的、用于存储要由评估装置AW处理的数据的存储器MEM。

[0039] 根据多个分类的训练数据集来训练评估装置AW,该训练数据集在当前的实施例中还从耦联至评估装置AW的数据库DB中读取。

[0040] 从数据库DB中读取显微图像照片PIC作为分类的第一训练数据PIC (CL),其中,分别将识别类别CL与图像照片相关联。在此,相应的识别类别CL例如说明在相应的图像照片PIC上示出的细胞或组织的类型、状态和/或病理类别。相应的识别类别CL的关联能够事先

例如通过专家进行。除了图像数据PIC之外,相应的识别类别CL优选以鉴别相关的识别类别CL的类别标识符的形式存储在相关的分类的图像照片PIC(CL)中。图像照片PIC例如能够作为第一传感器数据来自作为成像传感器的显微镜。

[0041] 另外,评估装置AW还能够耦合至分类的训练数据PIC2(CL)的另一来源SRC。那里可用的训练数据PIC2(CL)包含与第一训练数据PIC(CL)不同的传感器数据PIC2。在其他分类的训练数据PIC2(CL)中,将其他传感器数据同样分别与识别类别CL相关联。

[0042] 其他传感器数据PIC2能够与第一传感器数据PIC尤其在其类型、其质量、其检测方式方面和/或在进行检测的传感器和/或其传感器设置方面不同。特别地,其他传感器数据PIC2能够包含医学图像照片,该图像照片以更高的图像分辨率、不同的放大、利用不同的显微镜、利用不同的相机和/或在不同的光谱范围中被拍摄。另外,其他传感器数据PIC2还能够包括和量化其他的物理或化学变量。通常,其他传感器数据PIC2应该优选地涉及或包含识别类别CL的与第一传感器数据PIC不同的识别特征或不同的识别方面。

[0043] 通过评估装置AW能够从训练数据源SRC中尤其识别类别特定地请求其他训练数据PIC2(CL)。训练数据源SRC能够包括另外的数据库、另外的传感器和/或传感器网络。

[0044] 为了训练评估装置AW,首先通过评估装置AW的选择模块SEL将分类的图像照片PIC(CL)作为第一训练数据的序列读入。选择模块SEL将图像照片PIC传输至评估装置AW的数据驱动的机器学习例程。

[0045] 在本实施例中,机器学习例程通过与选择模块SEL耦联的贝叶斯神经网络BNN实施。贝叶斯神经网络BNN被设计为卷积神经网络。这种卷积神经网络能够众所周知地有利地用于图像处理和图像模式识别。神经网络BNN或由此实施的机器学习例程能够以数据驱动的方式训练并且具有在训练期间形成的训练结构。通过学习参数LP对神经网络BNN或其训练结构进行参数化,该学习参数通过训练来设置。

[0046] 在此,训练通常被理解为对参数化的系统模型、例如神经网络的输入参数到一个或多个输出参数、在此为要识别的类别CL的映射进行优化。在训练阶段期间,根据预设的、学习的和/或要学习的标准对该映射进行优化。在分类模型中,尤其能够将分类错误、分析错误和/或预测错误考虑作为标准。另外,能够设有另外的、涉及工程系统TS的控制的标准。训练结构例如能够包括神经网络的神经元的联网结构和/或神经元之间的连接的权重,该连接通过训练形成,使得尽可能良好地满足预设的标准。

[0047] 学习参数LP是系统模型的要通过训练优化的模型参数,机器学习例程基于该系统模型。在当前的实施例中,学习参数LP通过神经网络BNN的神经连接的权重实现。

[0048] 通过训练致力于使得神经网络BNN将所输送的图像照片PIC尽可能频繁地利用正确的识别分类CL进行分类。为此目的,通过训练来调节学习参数LP,使得由神经网络BNN对于所输送的图像照片PIC输出的识别类别OCL尽可能良好地复现该图像照片PIC的预设的识别类别CL。例如,这能够通过如下方式实现:将由神经网络BNN的输出层作为输出数据输出的识别类别OCL与预设的识别类别CL比较,该预设的识别类别由选择装置SEL提供。在该比较的范畴内,在输出数据OCL与对此相应的预设的识别类别CL之间形成距离D。距离D在此代表神经网络BNN的分类错误。如图2中通过虚线箭头所表明的那样,距离D被反馈到神经网络BNN。根据反馈的距离D,训练神经网络BNN被训练为将该距离D最小化,即通过输出的识别类别OCL尽可能良好地复现预设的识别类别CL。为此,能够通过标准优化方法变化学习参数

LP,直至距离D被最小化。例如,为了最小化能够使用梯度下降法。

[0049] 此外,通过神经网络BNN得出学习参数LP的在训练中出现的值分布VLP。在得出值分布VLP时,有利地假设学习参数LP根据可能多维的统计高斯分布来分布,并且各个学习参数的分布彼此独立。于是,值分布VLP通过分布宽度(在此为该高斯分布的方差或标准偏差)来代表。网络内部的学习参数LP的这种值分布能够通过贝叶斯神经网络(此处为BNN)借助标准方法有效地建模和得出。替代于或者除了值分布VLP的分布宽度,能够得出值分布VLP的熵值。

[0050] 证实了能够将学习参数LP的较小的分布宽度评价为学习参数LP已经相对精确调节的指示,即根据第一训练数据PIC(CL)已经尽可能地优化了神经网络BNN。在这种情况下,通过以第一训练数据PIC(CL)进行进一步的训练仅可预期到少量的改进。相反,能够将学习参数LP的更大的分布宽度评价为尚未找到最佳值,并且能够通过以相同类型的训练数据、在此为PIC(CL)进行进一步的训练来改进分类精度的指示。

[0051] 能够通过以更多相同类型的训练数据进行进一步训练来减少的分类不确定性通常也被称为认知不确定性。

[0052] 还借助于神经网络BNN得出神经网络BNN的输出数据OCL的值分布VOCL。在得出值分布VOCL时,有利地假设:输出数据OCL根据可能多维的统计高斯分布或分类分布而分布在识别类别之上,并且各个输出数据的分布彼此独立。然后,值分布VOCL通过分布宽度(在此为高斯分布或分类分布的方差或标准偏差)来代表。替代于或除了分布宽度,也能够得出值分布VOCL的熵值。

[0053] 证实了能够将输出数据OCL的较大的分布宽度评价为训练过的神经网络BNN仅可执行不确定分类的指示。如果同时学习参数LP的值分布VLP相对较小,这代表训练有素的神经网络BN,则几乎无法预期通过用相同类型的训练数据、在此为PIC(CL)进行进一步训练显著地改进分类质量。但是,在这种情况下,通过以其他训练数据、此处为PIC2(CL)继续进行训练,通常显著地改进了分类质量。

[0054] 通过以相同类型的训练数据进行进一步训练没有显著改进的这种分类不确定性通常被称为无意识不确定性。例如,在尝试预测随机抛硬币的情况下或在尝试根据不相关的或与分类不关联的传感器数据进行分类时出现无意识不确定性。这种分类无法通过增加(不相关的)训练数据的量来改进。

[0055] 值分布VLP以及VOCL从神经网络BNN传输至评估装置AW的评价模块EVAL。评价模块EVAL耦联至神经网络BNN,并且用于评价训练进度。评价模块EVAL根据输出数据OCL的值分布VOCL和学习参数LP的值分布VLP量化无意识不确定性以及认知不确定性。取决于此,评价模块EVAL通过以当前的第一训练数据、此处为PIC(CL)进行进一步训练形成关于分类的可预期的改进的量化的改进说明。改进说明例如能够根据分布宽度VLP和VOCL的加权的和来形成。如上面已经提到,改进说明优选形成为,使得其随分布宽度VLP的减小和分布宽度VOCL的增加而下降。

[0056] 此外,根据改进说明,生成继续消息CN并从评价模块EVAL传输至选择模块SEL。继续消息CN尤其能够包含值分布VLP和VOCL以及改进说明。

[0057] 通过选择模块SEL,根据接收到的继续消息CN和其中包含的改进说明,为了进一步训练,或者从数据库DB中选择另外的第一训练数据PIC(CL),或者从训练数据源SRC中请求

其他训练数据PIC2(CL)。在此,能够预设改进说明的阈值,使得在阈值之上还以第一训练数据PIC(CL)进行训练,并且在阈值之下,请求其他训练数据PIC2(CL)并用于进一步训练。

[0058] 替代地或附加地,继续消息CN也能够被传输至训练数据源SRC,以便在那里例如促使传感器网络的传感器检测训练数据,修改传感器的调节和/或专门选择传感器数据。

[0059] 优选地,识别类别特定地、即个体地针对训练数据的相应的识别等级CL执行要用于进一步训练的训练数据的之前描述的选择。在此,为相应的识别类别CL得出识别类别特定的值分布VLP和VOCL,产生识别类别特定的继续消息CN,并且通过选择模块SEL根据识别类别特定的继续消息CN为进一步训练或者从数据库DB中选择其他的第一训练数据PIC(CL)或者从训练数据源SRC中选择相关的识别类别CL的其他训练数据PIC2(CL)。

[0060] 通过根据本发明的方法,能够有效地控制神经网络BNN或评估装置AW的训练。代替可能在较长时间后才识别到,分类质量对于特定识别类别没有显著改进,已经能够及早地并且类别特定地请求质量上不同的训练数据并且用于进一步训练。因此,通常能够避免:输送尤其对于个别罕见的识别类别没有显著改进分类成效的其他训练数据。借此,通常能够显著缩短训练持续时间和/或显著改进学习成效。

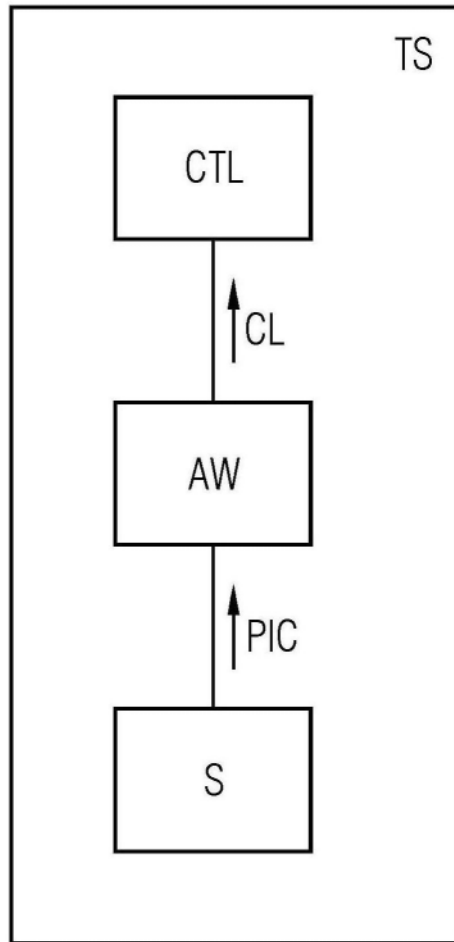


图1

TA
↙

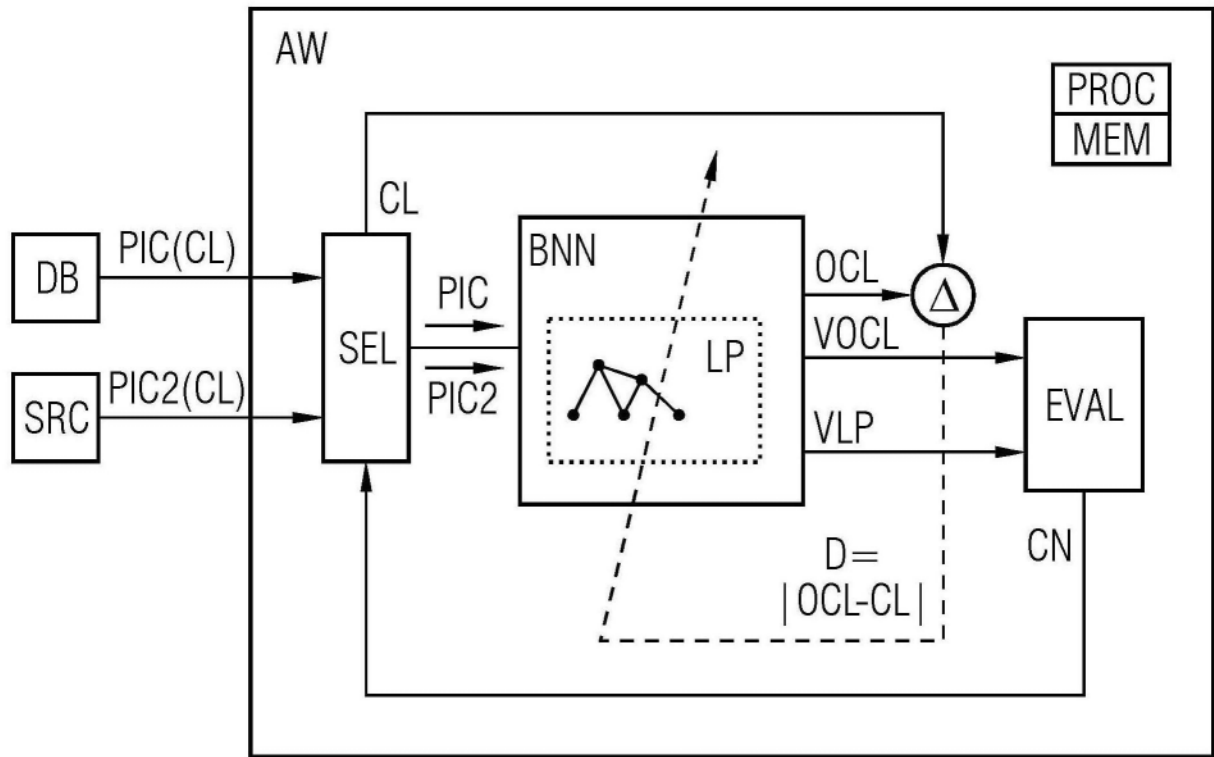


图2