(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 7 月 23 日 (23.07.2020)

WIPO | PCT

(10) 国际公布号
WO 2020/147612 A1

(51) 国际专利分类号:
G06N 3/08 (2006.01) *G06N 3/04* (2006.01)

(21) 国际申请号: PCT/CN2020/070584

(22) 国际申请日: 2020 年 1 月 7 日 (07.01.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910041725.8 2019年1月16日 (16.01.2019) CN

(71) 申请人: 阿里巴巴集团控股有限公司 (**ALIBABA GROUP HOLDING LIMITED**) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。

(72) 发明人: **杨斯然(YANG, Siran)**; 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 **任毅(REN, Yi)**; 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 **陈根宝(CHEN, Genbao)**; 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 **魏源(WEI, Yuan)**; 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 **张研(ZHANG, Yan)**; 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 **田旭(TIAN, Xu)**; 中国浙江省杭州市余杭区文一西路969号3号楼5楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

(54) **Title:** GRAPH-BASED CONVOLUTION NETWORK TRAINING METHOD, APPARATUS AND SYSTEM

(54) **发明名称:** 一种基于图的卷积网络训练方法、装置及系统

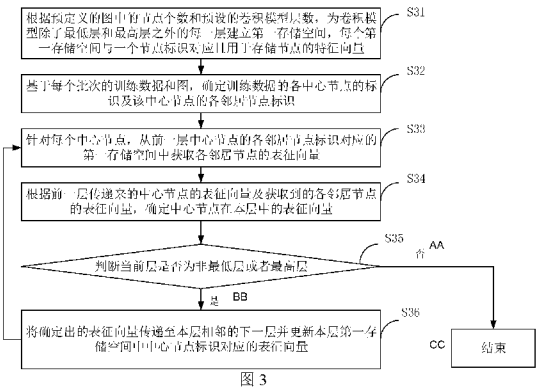


图 3

S31 Establish a first storage space for each layer, except the lowest layer and the highest layer, of a convolutional model according to the number of nodes in a predefined graph and the preset number of layers of the convolutional model, wherein each first storage space corresponds to a node identifier and is used for storing a characterization vector of a node

S32 Based on each batch of training data and graphs, determine identifiers of various central nodes of the training data and identifiers of various neighboring nodes of the central nodes

S33 With regard to each of the central nodes, acquire characterization vectors of various neighboring nodes from the first storage space corresponding to identifiers of various neighboring nodes of the central node in the previous layer

S34 Determine a characterization vector of the central node in the present layer according to the characterization vector of the central node transferred from the previous layer and the acquired characterization vectors of various neighboring nodes

S35 Determine whether the current layer is not the lowest layer or the highest layer

S36 Transfer the determined characterization vector to the next layer adjacent to the present layer and update the characterization vector corresponding to the identifier of the central node in the first storage space of the present layer

AA No

BB Yes

CC End

(57) **Abstract:** Provided are a graph-based convolution network training method, apparatus and system. The method comprises: establishing a first storage space for each layer, except the lowest layer and the highest layer, of a convolutional model; based on each batch of training data and graphs, determining identifiers of various central nodes and of various neighboring nodes of the central nodes of the training data (S32); with regard to each of the central nodes, acquiring characterization vectors of various neighboring nodes from the first storage space corresponding to identifiers of various neighboring nodes of the central node in the previous layer (S33); determining

WO 2020/147612 A1

[见续页]

(74) 代理人：北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.)；中国北京市金融街35号国际企业大厦A座16层，Beijing 100033 (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, 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))。

a characterization vector of the central node in the present layer according to the characterization vector of the central node transferred from the previous layer and the acquired characterization vectors of various neighboring nodes (S34); and when the current layer is not the lowest layer or the highest layer, transferring the determined characterization vector to the next layer adjacent to the present layer and updating the characterization vector corresponding to the identifier of the central node in the first storage space of the present layer (S36), until the characterization vector of the central node at the highest layer is obtained. According to the method, the calculation amount and training time of training can be effectively reduced.

(57) 摘要：一种基于图的卷积网络训练方法、装置及系统。所述方法包括：为卷积模型除了最低层和最高层之外的每层建立第一存储空间；基于每个批次的训练数据和图，确定训练数据的各中心节点及中心节点的各邻居节点的标识（S32）；针对每个中心节点，从前一层中心节点的各邻居节点标识对应的第一存储空间中获取各邻居节点的表征向量（S33）；根据前一层传递来的中心节点的表征向量及获取到的各邻居节点的表征向量，确定中心节点在本层中的表征向量（S34），当当前层并非最低层或者最高层时，将确定出的表征向量传递至本层相邻的下一层并更新本层第一存储空间中中心节点标识对应的表征向量（S36）；直至得到中心节点在最高层的表征向量。该方法可有效降低训练的计算量和训练时间。

一种基于图的卷积网络训练方法、装置及系统

本申请要求2019年01月16日递交的申请号为201910041725.8、发明名称为“一种基于图的卷积网络训练方法、装置及系统”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

本发明涉及机器学习技术领域，特别涉及一种基于图的卷积网络训练方法、装置及系统。

背景技术

随着移动终端及应用软件的普及，在社交、电商、物流、出行、外卖、营销等领域的服务提供商沉淀了海量业务数据，基于海量业务数据，挖掘不同业务实体（实体）之间的关系成为数据挖掘领域一个重要的技术研究方向。而随着机器处理能力的提升，越来越多技术人员开始研究如何通过机器学习技术进行挖掘。

本发明的发明人发现：目前，通过机器学习技术，对海量业务数据进行学习，得到用于表达实体及实体之间关系的图（Graph），即，对海量业务数据进行图学习，成为一个优选的技术方向。简单理解，图由节点和边构成，图中的每个序号代表一个节点，一个节点用于表示一个实体，节点与节点之间的边用于表示节点之间的关系。一张图一般会包括两个以上的节点和一条以上的边，因此，图也可以理解为由节点的集合和边的集合组成，通常表示为：G（V，E），其中，G表示图，V表示图G中节点的集合，E是图G中边的集合。图可以分为同构图和异构图，其中，异构图指的是一张图中的节点的类型不同（边的类型可以相同或者不同），或者一张图中边的类型不同（节点的类型可以相同或者不同）。图1所示则为一张异构图，同样类型的边用同样的线形表示，同样类型的点用同样的几何图形表示。

图卷积网络（GCN，Graph Convolution Network）是一种有效的图表征的学习方法，在很多关键任务上取得了超过以往的方法的效果。

为了获取图中每个节点的表征向量，图卷积网络通常的做法是对于每个节点聚合其邻居节点的表征向量（表征向量通常是需要训练的GCN各节点的属性特征值），进而得到更高一层节点的结果，依次进行层迭代计算。

容易看到此类方法的计算量随着层数的增多而指数级递增，随着应用的深入，对于节点数量庞大的图来说，可导致模型的整体训练时间不可接受。

高昂的计算力的需求增加了 GCN 模型训练的时间，制约了 GCN 模型在各种实际应用例如搜索、广告投放、推荐、社交网络挖掘等方面中的应用。

发明内容

鉴于上述问题，提出了本发明以便提供一种克服上述问题或者至少部分地解决上述问题的一种基于图的卷积网络训练方法、装置及系统。

第一方面，本发明实施例提供一种基于图的卷积网络训练方法，包括：

根据预定义的图中的节点个数和预设的卷积模型层数，为卷积模型除了最低层和最高层之外的每一层建立第一存储空间，每个第一存储空间与一个节点标识对应且用于存储所述节点的特征向量；

基于每个批次的训练数据和所述图，确定所述训练数据的各中心节点的标识及所述中心节点的各邻居节点标识；

针对每个中心节点，从前一层所述中心节点的各邻居节点标识对应的第一存储空间中获取所述各邻居节点的表征向量；

根据前一层传递来的所述中心节点的表征向量及获取到的所述各邻居节点的表征向量，确定所述中心节点在本层中的表征向量，当当前层并非最低层或者最高层时，将确定出的表征向量传递至本层相邻的下一层并更新本层第一存储空间中所述中心节点标识对应的表征向量；

直至得到所述中心节点在最高层的表征向量。

在一个实施例中，邻居节点标识对应的第一存储空间中邻居节点的表征向量为：预先写入所述存储空间中各邻居节点表征向量的初始值；或者为所述邻居节点在其他批次的训练数据中作为中心节点时更新的表征向量。

在一个实施例中，根据前一层传递来的所述中心节点的表征向量及获取到的所述各邻居节点的表征向量，确定所述中心节点在本层中的表征向量，包括：

将所述前一层传递来的所述中心节点的表征向量与获取到的所述各邻居节点的表征向量中各邻居节点的特征向量值进行聚合，得到所述中心节点在本层中的表征向量。

在一个实施例中，上述方法还包括：

根据预定义的图中的节点个数和预设的卷积模型层数，为卷积模型除了最低层和最

高层之外的每一层建立第二存储空间，每个第二存储空间与一个节点标识对应且用于存储所述节点的梯度值；

在得到所述中心节点在最高层的表征向量之后，所述方法还包括：

将中心节点在最高层的表征向量和所述中心节点在所述图中的标签 label 输入预设的损失函数，输出所述中心节点在最高层的梯度值；

通过所述中心节点在最高层的梯度值，更新所述最高层相邻的前一层的各邻居节点的梯度值和所述中心节点的梯度值；当所述前一层并非最低层时，将更新后的所述前一层的各邻居节点的梯度值写入前一层所述邻居节点标识对应的第二存储空间中；

从前一层的中心节点标识对应的第二存储空间中获取所述中心节点的梯度值；

根据从第二存储空间中获取到的中心节点的梯度值及更新后的中心节点的梯度值，得到所述中心节点在所述前一层的梯度值；

循环执行上述流程，直至得到最低层各邻居节点和中心节点的梯度值。

在一个实施例中，上述第二存储空间中所述中心节点的梯度值为：预设的初始梯度值；或者为所述中心节点在其他批次的训练数据中作为邻居节点时更新的梯度值。

在一个实施例中，所述从前一层的中心节点标识对应的第二存储空间中获取所述中心节点的梯度值之后，还包括：

将所述前一层的中心节点标识对应的第二存储空间中的梯度值还原为所述初始梯度值。

第二方面，本发明实施例提供一种基于图的卷积网络训练装置，包括：

存储模块，用于提供卷积模型除了最低层和最高层之外的每一层中节点标识对应的第一存储空间，每个第一存储空间用于存储所述节点的特征向量；所述第一存储空间是根据预定义的图中的节点个数和预设的卷积模型层数预先建立的；

节点确定模块，用于基于每个批次的训练数据和所述图，确定所述训练数据的各中心节点的标识及所述中心节点的各邻居节点标识；

获取模块，用于针对每个中心节点，从前一层所述中心节点的各邻居节点标识对应的第一存储空间中获取所述各邻居节点的表征向量；

表征向量确定模块，用于根据所述中心节点在前一层中的表征向量及获取到的所述各邻居节点的表征向量，确定所述中心节点在本层中的表征向量；直至得到所述中心节点在最高层的表征向量；

向量传递模块，用于当当前层并非最低层或者最高层时，将确定出的表征向量传递

至本层相邻的下一层；

更新模块，用于当当前层并非首层或者最后一层时，更新本层第一存储空间中所述中心节点标识对应的表征向量。

第三方面，本发明实施例提供一种计算节点装置，包括：处理器、用于存储处理器可执行命令的存储器；其中，处理器被配置为可执行如第一方面所述的基于图的卷积网络训练方法。

第四方面，本发明实施例提供一种基于图的卷积网络训练系统，包括：至少一个前述计算节点装置和至少两个存储节点装置；

所述计算节点装置，用于从所述至少一个存储节点装置中获取所述图的数据和基于所述图的各批次的训练数据；

所述至少两个存储节点装置，用于分布式地存储所述图，以及提供所述第一存储空间和/或第二存储空间。

第五方面，本发明实施例提供一种非临时性计算机可读存储介质，当所述存储介质中的指令由处理器执行时，能够实现前述基于图的卷积网络训练方法。

本发明实施例提供的上述技术方案的有益效果至少包括：

本发明实施例提供的上述基于图的卷积网络训练方法、装置及系统中，根据预定义的图中的节点个数和预设的卷积模型层数，为卷积模型除了首层和最后一层之外的每一层建立第一存储空间，该第一存储空间用于存储所述节点的特征向量，这样，针对每批次的训练数据中的每个中心节点，在计算其下一层的表征向量时，对于其前一层的邻居节点的表征向量，可直接从对应层的第一存储空间中获取，进而确定出中心节点在本层的表征向量，并更新至本层中心节点标识对应的第一存储空间中，以便在后续其他批次的训练数据中该节点作为其他中心节点的邻居节点时直接进行复用，这样，针对数据量庞大的训练数据而言，每批次训练数据，只需要计算每批次训练数据中中心节点的卷积结果，而无需对扩散后的所有邻居节点再进行卷积计算，可实现计算量与层数呈线性关系，降低了训练的计算量和训练时间，避免了现有训练过程中计算量与层数呈指数关系所带来的计算量庞大、整体训练时间不可接受的问题。

另外，由于在卷积网络训练过程的前向计算中采用了对于其前一层的邻居节点的表征向量，直接从对应层的第一存储空间中获取的方式，而未实际对每层的邻居节点向下进行扩展，因此，在对应的反向计算过程中，为了解决邻居节点的梯度反向传递问题，本发明还根据预定义的图中的节点个数和预设的卷积模型层数，为卷积模型除了最低层

和最高层之外的每一层建立了第二存储空间，该第二存储空间用于存储节点的梯度值；这样，对最高层和非最低层而言，可根据本层中心节点的梯度，更新前一层中心节点和各邻居节点的梯度值，并将前一层邻居节点的梯度保存于前一层预设的第二存储空间中。再根据更新后的中心节点的梯度值与从该前一层该中心节点标识对应的第二存储空间获取该中心节点的梯度值，最终得到用于向下传递的该中心节点的梯度值。如此循环，使用前一层中心节点的梯度值继续向前一层相邻的再前一层传递，直至最低层。这样，在该邻居节点在其他批次训练数据中再次作为中心节点时，可以顺利地将梯度传递至前一层的中心节点和各邻居节点，以实现整个卷积网络训练中梯度值的反向传递。

本发明的其它特征和优点将在随后的说明书中阐述，并且，部分地从说明书中变得显而易见，或者通过实施本发明而了解。本发明的目的和其他优点可通过在所写的说明书、权利要求书、以及附图中所特别指出的结构来实现和获得。

下面通过附图和实施例，对本发明的技术方案做进一步的详细描述。

附图说明

附图用来提供对本发明的进一步理解，并且构成说明书的一部分，与本发明的实施例一起用于解释本发明，并不构成对本发明的限制。在附图中：

图 1 为异构图的一个例子的示意图；

图 2 为现有技术中图卷积网络模型的训练方式的示意图；

图 3 为本发明实施例提供的基于图的卷积网络训练方法的流程图之一；

图 4 为本发明实施例提供的实例中图卷积网络的模型的前向计算的过程示意图；

图 5 为本发明实施例提供的基于图的卷积网络训练方法的流程图之二；

图 6 为本发明实施例提供的基于图的卷积网络训练方法的流程图之三；

图 7 为本发明实施例提供的实例中图卷积网络的模型的后向计算的过程示意图；

图 8 为本发明实施例提供的基于图的卷积网络训练装置的结构示意图。

具体实施方式

下面将参照附图更详细地描述本公开的示例性实施例。虽然附图中显示了本公开的示例性实施例，然而应当理解，可以以各种形式实现本公开而不应被这里阐述的实施例所限制。相反，提供这些实施例是为了能够更透彻地理解本公开，并且能够将本公开的范围完整的传达给本领域的技术人员。

图学习面向业务场景才是有意义的，所以，当基于业务场景确定了图中的节点对应的实体和边对应的实体间关系后，图则被赋予了业务含义和技术含义，按照该业务场景要解决的技术问题和业务问题执行相应的图学习任务，则可得到解决相应问题的结果。比如，图表示学习可将复杂的图表示成低维、实值、稠密的向量形式，使其具有表示及推理能力，可以方便执行其他机器学习任务。

现有技术中，为了获取图中每个节点的表征向量，可能采用多种学习方法，图卷积网络（GCN，Graph Convolution Network）就是其中之一，图卷积网络通常的做法是对于每个节点聚合其邻居节点的表征向量（表征向量通常是需要训练的 GCN 各节点的属性特征值），进而得到更高一层节点的结果，然后依次进行层迭代计算。

此类方法的计算量可随着层数的增多而指数级递增，参照图 2 所示的图卷积网络的模型，假设使用每层度为 2 的邻居节点采样来实现 $h^{(0)}$~$h^{(3)}$ 层的卷积网络，则对于每个节点来说，其每层的表征向量都依赖于前一层的包括自身和其 2 个邻居节点在内的 3 个节点的表征向量；比如，每个 $h^{(1)}$ 的表征向量，都依赖于前一层即 3 个 $h^{(0)}$ 的表征向量来计算，以此类推，为了获取最后一层 $h^{(3)}$ 节点的表征向量，从计算量上来说，则需要计算 15 个 $h^{(0)}$、7 个 $h^{(1)}$、3 个 $h^{(2)}$ 和 1 个 $h^{(3)}$ 的节点的相关表征向量。而为了在规模图上降低采样带来的精度损失，需要更宽的邻居节点采样，也就是每个节点其前一层需要参与计算的邻居节点的数量更多，例如为 5 个或 10 个，这样话，各节点的表征向量的计算量膨胀会更为显著，导致模型的整体训练时间不可接受。

为了解决上述基于图的卷积网络的训练问题中存在的计算量庞大所带来的训练时间不可接受的问题，本发明实施例提供了一种基于图的卷积网络训练的方法，该方法可以适用于任何种类的图，包括各种不同类型的同构图和异构图。在这些图中，包含若干彼此相连的节点，每个节点存储有自身与其他节点的连接关系、自身属性和与其他节点相连的边的属性等信息，这些节点可以根据图的类型和用途的不同，而可以呈现为各种实体对象的表现形式，例如为商品购买网络中的商品、查询关键词、用户、搜索广告等；或者为社交网络中的用户、用户偏好等。本发明实施例对图和节点的所表征的是何种对象并不限定。

下面结合附图，对本发明实施例提供一种基于图的卷积网络训练方法的具体实施方式进行详细说明。

参照图 3 所示，本发明实施例提供的基于图的卷积网络训练方法，包括下述步骤：

S31、根据预定义的图中的节点个数和预设的卷积模型层数，为卷积模型除了最低层和最高层之外的每一层建立第一存储空间，每个第一存储空间与一个节点标识对应且用于存储节点的特征向量。

本发明的发明人发现，由于考虑到不同批次中心节点在不同批次的训练数据被大量复用的情形，本发明实施例对 GCN 的训练过程中为图卷积网络模型引入了第一存储空间，在缓存中直接查询前一层邻居节点的表征向量，参与本层中心节点的表征向量的运算，而不用对前一层邻居节点进行实际的扩散得到其表征向量，因此实现计算量与层数呈线性关系，大大降低 GCN 的训练时间。

本步骤 S31 可以在下述步骤 S32~S36 之前预先完成。

在本发明实施例中，步骤 S31~ S36 对图卷积网络学习的前向运算进行了改进。前向运算是指从 GCN 的最低层，逐步学习到最高层的中心节点的表征向量的过程。

在本发明实施例中，仅为了区分存储空间所存储的内容的不同，将与卷积模型中除了最低层和最高层之外的每一层所对应的存储空间分为两类，分别称呼为第一存储空间和第二存储空间，第一存储空间用于存储在图卷积网络训练过程前向运算过程中需要使用的节点的特征向量；而第二存储空间用于存储后向运算过程中需要使用的节点的梯度值，随后会在 GCN 后向运算过程中说明。

在具体实施时，上述第一存储空间和第二存储空间在物理上可以是一体的，也就是可以位于同一个物理存储区域上，也可以是在物理上分隔的不同区域，即分属于不同的物理存储区域。

具体实施时，第一存储空间和第二存储空间可以采用各类缓存等，方便根据节点的数量和图卷积网络模型的层数灵活设置空间大小，并在图卷积网络模型学习结束后及时释放存储空间，具有较强的灵活性，本发明实施例对采用何种缓存并不做限定。

图卷积网络模型的层数，决定了需要几层的第一存储空间。举例来说，假设图卷积网络模型的层数为 n，则除了最低层和最高层之外，一共有 n-2 层需要设置对应的第一存储空间。

对于每一层来说，预定义的图中有多少节点，不论是同构节点还是异构的节点，就需要有多少个第一存储空间，每个第一存储空间都分别与一个节点标识对应。

假设图中一共有 10 万个节点，需要设置第一存储空间的层数是 2 层，那么，则一共需要 20 万个第一存储空间。每个层的第一存储空间都是独立的。

针对一个预定义的图来说，每批次的训练数据中包含了该批次中各中心节点的标识，

各个批次的中心节点则涵盖了整个图中的所有节点，举个简单例子来说，假设图中有 100 个节点，分为 10 个批次的训练数据，则每个批次的训练数据中就包含了 10 个中心节点。

S32、基于每个批次的训练数据和图，确定训练数据的各中心节点的标识及该中心节点的各邻居节点标识；

对于包含数量庞大的节点的图来说，由于计算能力有限，通常可以采用分批次的训练数据进行训练的方式。例如划分为若干个 min-batch，每个 min-batch 作为一个批次的训练数据。每个批次的训练数据中，都包含了本批次的各中心节点的标识。

根据每批训练数据中中心节点的标识以及图，就可以确定每批次训练数据中每个中心节点的各邻居节点的标识。

S33、针对每个中心节点，从前一层中心节点的各邻居节点标识对应的第一存储空间中获取各邻居节点的表征向量；

在本步骤 S33 中，每个中心节点的计算方式与现有技术不同，是直接从前一层的第一存储空间中，找到该中心节点的各邻居节点的标识所对应的那些第一存储空间，从其中读取这些邻居节点在前一层的表征向量。

在本发明实施例中，在 GCN 开始训练之前，每个第一存储空间中都存有该节点在该层的表征向量的一个初始值，这个预设的初始值可以是个固定的数值，也可以是个随机生成的随机数。随着训练的开始，不同批次的中心节点在训练时，作为中心节点的节点对应的第一存储空间的数据逐渐被更新。

因此，从第一存储空间读出表征向量，可以是下面两种情形，一种是，这个第一存储空间的数据从未被更新过，还是预设的初始值，换言之，这个邻居节点在本批次数据之前未作为中心节点更新过其自身的表征向量。另一种是，这个邻居节点在本批次之前已作为其他批次的中心节点，将初始值更新为计算后的表征向量值。

S34、根据前一层传递来的中心节点的表征向量及获取到的各邻居节点的表征向量，确定中心节点在本层中的表征向量；

在本步骤 S34 来说，可以采用将前一层传递来的所述中心节点的表征向量与获取到的所述各邻居节点的表征向量中各邻居节点的特征向量值进行聚合的方式，得到中心节点在本层中的表征向量。

用数学的方式例如可以表达为：

对于一个 GCN 模型来说，定义某批次的训练数据中，中心节点 v 在第 l 层的表征向量为 h_v^l，其中 h_v^0 为节点的原始输入特征。那么 GCN 每层的计算过程可以如下：

$$a_v^{l+1} = Agg_l(\{h_u^l, \forall u \in \mathcal{N}(v)\}) \quad (1)$$

$$h_v^{l+1} = Comb_l(a_v^{l+1}, h_v^l) \quad (2)$$

要计算中心节点v在$l+1$层的表征向量，上述公式 1 是将该中心节点v在第l层的所有邻居节点的表征向量（从对应的第一存储空间中获取）先进行预设的 Aggregation 算法的聚合运算，然后再通过公式 2，再将前面的聚合的结果再与该中心节点在第l层的表征向量再次进行另一类型的 combine 算法的聚合运算。

当然，还可以一次性将中心节点v在第l层的所有邻居节点的表征向量和中心节点在第l层的表征向量进行一次聚合运算，本发明实施例对采用何种聚合的方式不做限定。

聚合算法可以采用现有任何一种算法，本发明实施例对此不做限定，例如采用 mean（平均）、加权平均等等。

S35、判断当前层是否为非最低层或者最高层，当判断为非最低层或者非最高层时，执行下述步骤 S36；否则，结束流程；

由于步骤 S31~S36 是前向计算过程，即最低层逐层向最高层运算的过程，所以在本步骤 S35 中，当判断当前层为最高层时，其实就是步骤 S34 中已计算得到了中心节点在最高层中的表征向量，此时前向计算过程就结束了。

S36、将确定出的表征向量传递至本层相邻的下一层并更新本层第一存储空间中中心节点标识对应的表征向量；然后再转向执行步骤 S33，循环步骤 S33~S36，直到流程结束。

中心节点在非最低层的表征向量，都可以从上一层计算后直接传递过来，而不需要从上一层对应的第一存储空间中读取，在本步骤 S26 中，将计算出来的中心节点传递至相邻的下一层时，可以同时更新本层第一存储空间中的存储的表征向量。

为了更好地说明上述步骤 S31~S36，下面以一个简单的图卷积网络模型的训练过程中的前向计算过程进行说明：

参照图 4 所示，该图卷积网络的模型共包含$h^0 \sim h^3$层，位于每层的中间的黑色节点为中心节点，在每层中，与黑色节点连接的灰色的 3 个节点为其邻居节点。非最低层和非最高层的h^1层和h^2层分别预设有对应的缓存空间$\widetilde{h}_k^1$和$\widetilde{h}_k^2$。缓存空间$\widetilde{h}_k^1$和$\widetilde{h}_k^2$中包含若干小的缓存，分别存储每个节点在本层的表征向量。

最低层h^0的中心节点和邻居节点的表征向量，是原始值，可以直接从图中得到。

对于h^1层的中心节点的表征向量，是通过h^0层的 3 个邻居节点和中心节点的表征向

量进行聚合得到，然后将h^1层的中心节点的表征向量写入$\overrightarrow{f_k^1}$对应的缓存中（见图 4 中h^1层中心节点指向缓存空间$\overrightarrow{f_k^1}$的箭头），以便在下次该中心节点作为其他批次的训练数据中的邻居节点时使用，同时，得到的h^1层的中心节点的表征向量继续向相邻的下一层传递（见图 4 中h^1层中心节点指向缓存空间下一层h^2中心节点的箭头）。

h^2层的中心节点的表征向量的得到方式与前面方式类似，即：先是从缓存空间$\overrightarrow{f_k^1}$中读取中心节点的 3 个邻居节点在h^1层的表征向量，然后将读取到的 3 个邻居节点在h^1层的表征向量与从h^1层传递上来的该中心节点在h^1层的表征向量进行聚合，得到中心节点在h^2层的表征向量，然后再将中心节点在h^2层的表征向量写入到缓存空间$\overrightarrow{f_k^2}$对应的位置。

h^3层的中心节点的表征向量的得到方式与前面的方式类似，不同的在于，在得到中心节点在h^3层的表征向量后，不需要将其更新到对应的缓存空间中。

本发明实施例提供的上述基于图的卷积网络训练方法中，根据预定义的图中的节点个数和预设的卷积模型层数，为卷积模型除了首层和最后一层之外的每一层建立第一存储空间，该第一存储空间用于存储所述节点的特征向量，这样，针对每批次的训练数据中的每个中心节点，在计算其下一层的表征向量时，对于其前一层的邻居节点的表征向量，可直接从对应层的第一存储空间中获取，进而确定出中心节点在本层的表征向量，并更新至本层中心节点标识对应的第一存储空间中，以便在后续其他批次的训练数据中该节点作为其他中心节点的邻居节点时直接进行复用，这样，针对数据量庞大的训练数据而言，每批次训练数据，只需要计算每批次训练数据中中心节点的卷积结果，而无需对扩散后的所有邻居节点再进行卷积计算，可实现计算量与层数呈线性关系，降低了训练的计算量和训练时间，避免了现有训练过程中计算量与层数呈指数关系所带来的计算量庞大、整体训练时间不可接受的问题。

对于每批次训练数据来说，图卷积网络模型的训练通常会包含前向计算过程，即获得最高层的中心节点的表征向量，后续还需要按照最高层中心节点的表征向量进行反向梯度的传递，也就是后向计算（Back Propagation）的过程。

由于在卷积网络训练过程的前向计算中采用了对于其前一层的邻居节点的表征向量，直接从对应层的第一存储空间中获取的方式，而未实际对每层的邻居节点向下进行扩展，所以，现有 GCN 后向计算过程无法适用于本发明实施例，为了解决后向计算的过程的上述问题，本发明实施例提供的卷积网络训练方法，在上述 S31~S36 之后，参照图

5 所示，还可以执行下述流程：

S51、将中心节点在最高层的表征向量和中心节点在图中的标签 label 输入预设的损失函数，输出中心节点在最高层的梯度值；

与步骤 S31 类似，根据预定义的图中的节点个数和预设的卷积模型层数，为卷积模型除了最低层和最高层之外的每一层建立第二存储空间，每个第二存储空间与一个节点标识对应且用于存储节点的梯度值。

S52、从最高层开始，在每层，通过中心节点在本层的梯度值，更新本层相邻的前一层的各邻居节点的梯度值和中心节点的梯度值；

S53、判断前一层是否为最低层，若是，结束流程，若否，继续执行下述步骤 S54；

S54、将更新后的前一层的各邻居节点的梯度值写入前一层邻居节点标识对应的第二存储空间中；

S55、从前一层的中心节点标识对应的第二存储空间中获取中心节点的梯度值；

S56、根据从第二存储空间中获取到的中心节点的梯度值及更新后的中心节点的梯度值，最终得到中心节点在最高层的前一层的梯度值，这个梯度值用于继续向再前一层传递；然后跳转至步骤 S52，再次循环，直至得到最低层的各邻居节点的梯度值和中心节点的梯度值。

由于后向计算过程中，中心节点的梯度是可以直接向下传递的，而作为邻居节点，因为没有实际扩展，没办法向下传递，所以在上述步骤 S51~S56 中，将邻居节点在每层（非最高层和非最低层）的梯度进行缓存，以便这些邻居节点后续在其他批次中作为中心节点中可以直接使用这些梯度数据，实现梯度向下传递。

在卷积模型除了最低层和最高层之外的每一层建立第二存储空间中，在训练初始时，其中每个节点标识对应的第二存储空间存储的是各节点梯度的初始值。随着每批次的训练数据的训练过程的进行，在后向计算过程中，作为邻居节点的节点对应的第二存储空间中的初始值逐渐被更新为其他数值，从而可以在其他批次的训练中作为中心节点时，其梯度值可以进行后向传递。

节点梯度的初始值，较佳地，可以设置为 0。

参照图 6 所示，在上述步骤 S55 中从前一层的中心节点标识对应的第二存储空间中获取中心节点的梯度值之后，上述流程还可以包括下述步骤：

S57、将前一层的中心节点标识对应的第二存储空间中的梯度值还原为初始梯度值。

还是以一个图 4 所示的 GCN 模型的例子来说明，参照图 7 所示，该图卷积网络的模

型共包含h^0~h^3层，黑色节点是中心节点，灰色节点是其邻居节点。非最低层和非最高层的h^1层和h^2层分别预设有对应的缓存空间$\hat{h}^1$和$\hat{h}^2$。缓存空间$\hat{h}^1$和$\hat{h}^2$中包含若干小的缓存，分别存储每个节点在本层的梯度值。

1、将中心节点在最高层的表征向量和中心节点在图中的标签 label 输入预设的损失函数，输出中心节点在最高层的梯度值即图 7 所示的$\nabla_{h^3} loss$。

2、然后使用中心节点的$\nabla_{h^3} loss$，更新 h2 层 3 个邻居节点的梯度值和中心节点自己的梯度值（通过调整权重）；并把 h2 层的 3 个邻居节点的梯度值存到 h2 层的缓存中（见图 7 中 h2 层 3 个邻居节点指向$\hat{h}^2$缓存空间的箭头）。

3、从 h2 层中心节点对应的$\hat{h}^2$缓存空间中读取其中的梯度值（见图 7 中$\hat{h}^2$缓存空间指向 h2 层中心节点的箭头），该梯度值要么是初始值（例如为 0），要么是在之前批次中作为邻居节点更新过的；

4、将 h2 层中心节点对应的$\hat{h}^2$缓存空间中的梯度值，再次还原为初始值 0。

5、将从$\hat{h}^2$缓存空间中读取的中心节点的梯度值，与步骤 2 中中心节点更新后的梯度值再次进行运算，得到 h2 层中心节点最终的梯度值（见图 7 中 h2 层中心节点指向 Back Prop 的箭头），也就是用于继续向下传递的$\nabla_{h^2} loss$。

6、再次循环，按照与以上类似方式得到 h1 层邻居节点的梯度值，以及中心节点向下传递的梯度值$\nabla_{h^1} loss$，以及 h1 层邻居节点的梯度值和中心节点的梯度值。

上述从 h2 层中心节点对应的$\hat{h}^2$缓存空间中读取其中的梯度值，该梯度值是中心节点的累计未消费梯度，实际上是该中心节点在作为邻居节点时的梯度进行延迟的后向计算，在消费掉该梯度值后（也就是读取了该梯度值后），将其值重置为初始值（例如为 0）。

采用本发明实施例提供的上述基于图的卷积网络训练方法，本发明的发明人经过试验发现，相对于现有技术，本发明实施例可大大缩短 GCN 模型的训练时间。

试验中使用 512 个 min-batch 训练 20 个神经网络模型，并在开源数据集上验证上述方法的执行效率。

采用本发明实施例提供的上述训练方法称为 ScalableGCN，现有技术中 GCN 训练的方式选择了 GraphSAGE。

两种方式训练的结果相差很小，因此采用本发明实施例提供的上述方法可以取得多层卷积网络模型的收益，但是在训练时间上，下表 1 是以 Reddi 数据集（23 万个节点）上每个 min-batch 两种方式所需训练的时间的对比：

表 1 单位秒

GCN 模型类型	ScalableGCN	GraphSAGE
2 layer（2 层）	0.026	0.120
3 layer（3 层）	0.035	1.119

注意到，采用本发明实施例的上述方法，可以大大每批次压缩训练的时间，对 GCN 训练的时间相对于卷积模型层数是呈线性的。

基于同一发明构思，本发明实施例还提供了一种基于图的卷积网络训练装置、计算节点装置和基于图的卷积网络训练系统，由于这些装置和系统所解决问题的原理与前述基于图的卷积网络训练方法相似，因此该装置和系统的实施可以参见前述方法的实施，重复之处不再赘述。

本发明实施例提供的基于图的卷积网络训练装置，参照图 8 所示，包括：

存储模块 81，用于提供卷积模型除了最低层和最高层之外的每一层中节点标识对应的第一存储空间，每个第一存储空间用于存储所述节点的特征向量；所述第一存储空间是根据预定义的图中的节点个数和预设的卷积模型层数预先建立的；

节点确定模块 82，用于基于每个批次的训练数据和所述图，确定所述训练数据的各中心节点的标识及所述中心节点的各邻居节点标识；

获取模块 83，用于针对每个中心节点，从前一层所述中心节点的各邻居节点标识对应的第一存储空间中获取所述各邻居节点的表征向量；

表征向量确定模块 84，用于根据所述中心节点在前一层中的表征向量及获取到的所述各邻居节点的表征向量，确定所述中心节点在本层中的表征向量；直至得到所述中心节点在最高层的表征向量；

向量传递模块 85，用于当当前层并非最低层或者最高层时，将确定出的表征向量传递至本层相邻的下一层；

更新模块 86，用于当当前层并非首层或者最后一层时，更新本层第一存储空间中所述中心节点标识对应的表征向量。

在一个实施例中，上述邻居节点标识对应的第一存储空间中邻居节点的表征向量为：预先写入所述存储空间中各邻居节点表征向量的初始值；或者为所述邻居节点在其他批

次的训练数据中作为中心节点时更新的表征向量。

在一个实施例中，表征向量确定模块 84，具体用于将所述前一层传递来的所述中心节点的表征向量与获取到的所述各邻居节点的表征向量中各邻居节点的特征向量值进行聚合，得到所述中心节点在本层中的表征向量。

在一个实施例中，上述基于图的卷积网络训练装置，参照图 8 所示，还可以包括：梯度反向传递模块 87；

相应地，所述存储模块 81，还用于提供卷积模型除了最低层和最高层之外的每一层中节点标识对应的第二存储空间，每个第二存储空间用于存储所述节点的梯度值；所述第二存储空间是根据预定义的图中的节点个数和预设的卷积模型层数预先建立的；

梯度反向传递模块 87，用于将中心节点在最高层的表征向量和所述中心节点在所述图中的标签 label 输入预设的损失函数，输出所述中心节点在最高层的梯度值；通过所述中心节点在最高层的梯度值，更新所述最高层相邻的前一层的各邻居节点的梯度值和所述中心节点的梯度值；当所述前一层并非最低层时，将更新后的所述前一层的各邻居节点的梯度值写入前一层所述邻居节点标识对应的第二存储空间中；从前一层的中心节点标识对应的第二存储空间中获取所述中心节点的梯度值；根据从第二存储空间中获取到的中心节点的梯度值及更新后的中心节点的梯度值，得到所述中心节点在所述前一层的梯度值；循环执行上述流程，直至得到最低层各邻居节点和中心节点的梯度值。

本发明实施例还提供了一种计算节点装置，包括：处理器、用于存储处理器可执行命令的存储器；其中，处理器被配置为可执行前述基于图的卷积网络训练方法。

本发明实施例还提供了一种基于图的卷积网络训练系统，包括：至少一个前述的计算节点装置和至少两个存储节点装置；

所述计算节点装置，用于从所述至少一个存储节点装置中获取所述图的数据和基于所述图的各批次的训练数据；

所述至少两个存储节点装置，用于分布式地存储所述图，以及提供所述第一存储空间和/或第二存储空间。

本发明实施例还提供了一种非临时性计算机可读存储介质，当所述存储介质中的指令由处理器执行时，能够实现前述基于图的卷积网络训练方法。

本领域内的技术人员应明白，本发明的实施例可提供为方法、系统、或计算机程序产品。因此，本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本发明可采用在一个或多个其中包含有计算机可用程序代码的

计算机可用存储介质（包括但不限于磁盘存储器和光学存储器等）上实施的计算机程序产品的形式。

本发明是参照根据本发明实施例的方法、设备（系统）、和计算机程序产品的流程图和／或方框图来描述的。应理解可由计算机程序指令实现流程图和／或方框图中的每一流程和／或方框、以及流程图和／或方框图中的流程和／或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能的步骤。

显然，本领域的技术人员可以对本发明进行各种改动和变型而不脱离本发明的精神和范围。这样，倘若本发明的这些修改和变型属于本发明权利要求及其等同技术的范围之内，则本发明也意图包含这些改动和变型在内。

权 利 要 求 书

1. 一种基于图的卷积网络训练方法，其特征在于，包括：

根据预定义的图中的节点个数和预设的卷积模型层数，为卷积模型除了最低层和最高层之外的每一层建立第一存储空间，每个第一存储空间与一个节点标识对应且用于存储所述节点的特征向量；

基于每个批次的训练数据和所述图，确定所述训练数据的各中心节点的标识及所述中心节点的各邻居节点标识；

针对每个中心节点，从前一层所述中心节点的各邻居节点标识对应的第一存储空间中获取所述各邻居节点的表征向量；

根据前一层传递来的所述中心节点的表征向量及获取到的所述各邻居节点的表征向量，确定所述中心节点在本层中的表征向量，当当前层并非最低层或者最高层时，将确定出的表征向量传递至本层相邻的下一层并更新本层第一存储空间中所述中心节点标识对应的表征向量；

直至得到所述中心节点在最高层的表征向量。

2. 如权利要求 1 所述的方法，其特征在于，邻居节点标识对应的第一存储空间中邻居节点的表征向量为：预先写入所述存储空间中各邻居节点表征向量的初始值；或者为所述邻居节点在其他批次的训练数据中作为中心节点时更新的表征向量。

3. 如权利要求 1 所述的方法，其特征在于，根据前一层传递来的所述中心节点的表征向量及获取到的所述各邻居节点的表征向量，确定所述中心节点在本层中的表征向量，包括：

将所述前一层传递来的所述中心节点的表征向量与获取到的所述各邻居节点的表征向量中各邻居节点的特征向量值进行聚合，得到所述中心节点在本层中的表征向量。

4. 如权利要求 1-3 任一项所述的方法，其特征在于，还包括：

根据预定义的图中的节点个数和预设的卷积模型层数，为卷积模型除了最低层和最高层之外的每一层建立第二存储空间，每个第二存储空间与一个节点标识对应且用于存储所述节点的梯度值；

在得到所述中心节点在最高层的表征向量之后，所述方法还包括：

将中心节点在最高层的表征向量和所述中心节点在所述图中的标签 label 输入预设的损失函数，输出所述中心节点在最高层的梯度值；

通过所述中心节点在最高层的梯度值，更新所述最高层相邻的前一层的各邻居节点

的梯度值和所述中心节点的梯度值；当所述前一层并非最低层时，将更新后的所述前一层的各邻居节点的梯度值写入前一层所述邻居节点标识对应的第二存储空间中；

从前一层的中心节点标识对应的第二存储空间中获取所述中心节点的梯度值；

根据从第二存储空间中获取到的中心节点的梯度值及更新后的中心节点的梯度值，得到所述中心节点在所述前一层的梯度值；

循环执行上述流程，直至得到最低层各邻居节点和中心节点的梯度值。

5. 如权利要求 4 所述的方法，其特征在于，第二存储空间中所述中心节点的梯度值为：预设的初始梯度值；或者为所述中心节点在其他批次的训练数据中作为邻居节点时更新的梯度值。

6. 如权利要求 5 所述的方法，其特征在于，所述从前一层的中心节点标识对应的第二存储空间中获取所述中心节点的梯度值之后，还包括：

将所述前一层的中心节点标识对应的第二存储空间中的梯度值还原为所述初始梯度值。

7. 一种基于图的卷积网络训练装置，其特征在于，包括：

存储模块，用于提供卷积模型除了最低层和最高层之外的每一层中节点标识对应的第一存储空间，每个第一存储空间用于存储所述节点的特征向量；所述第一存储空间是根据预定义的图中的节点个数和预设的卷积模型层数预先建立的；

节点确定模块，用于基于每个批次的训练数据和所述图，确定所述训练数据的各中心节点的标识及所述中心节点的各邻居节点标识；

获取模块，用于针对每个中心节点，从前一层所述中心节点的各邻居节点标识对应的第一存储空间中获取所述各邻居节点的表征向量；

表征向量确定模块，用于根据所述中心节点在前一层中的表征向量及获取到的所述各邻居节点的表征向量，确定所述中心节点在本层中的表征向量；直至得到所述中心节点在最高层的表征向量；

向量传递模块，用于当当前层并非最低层或者最高层时，将确定出的表征向量传递至本层相邻的下一层；

更新模块，用于当当前层并非首层或者最后一层时，更新本层第一存储空间中所述中心节点标识对应的表征向量。

8. 一种计算节点装置，其特征在于，包括：处理器、用于存储处理器可执行命令的存储器；其中，处理器被配置为可执行如权利要求 1-6 任一项所述的基于图的卷积网络

训练方法。

9. 一种基于图的卷积网络训练系统，其特征在于，包括：至少一个如权利要求 8 所述的计算节点装置和至少两个存储节点装置；

所述计算节点装置，用于从所述至少一个存储节点装置中获取所述图的数据和基于所述图的各批次的训练数据；

所述至少两个存储节点装置，用于分布式地存储所述图，以及提供所述第一存储空间和/或第二存储空间。

10. 一种非临时性计算机可读存储介质，其特征在于，当所述存储介质中的指令由处理器执行时，能够实现如权利要求上述 1-6 任一项基于图的卷积网络训练方法。

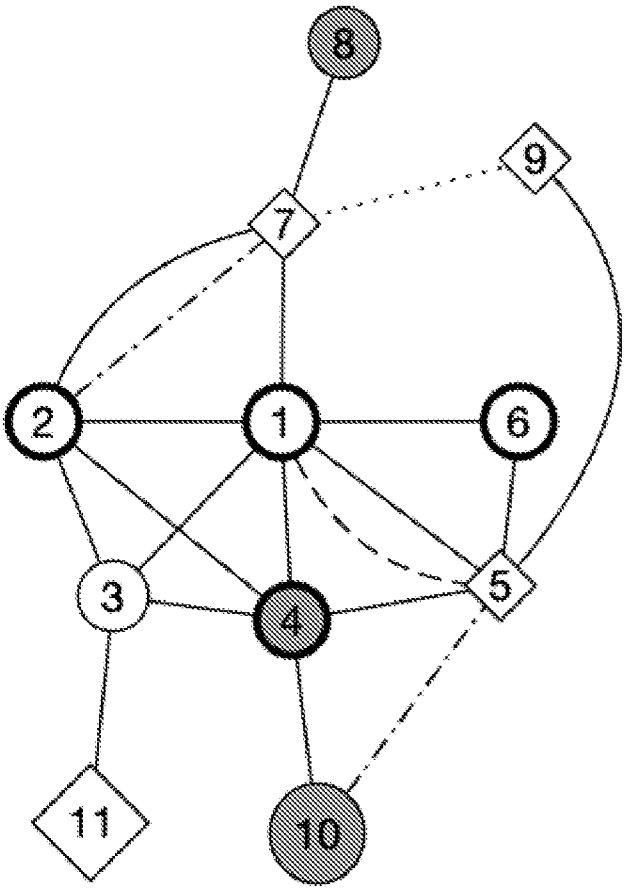


图 1

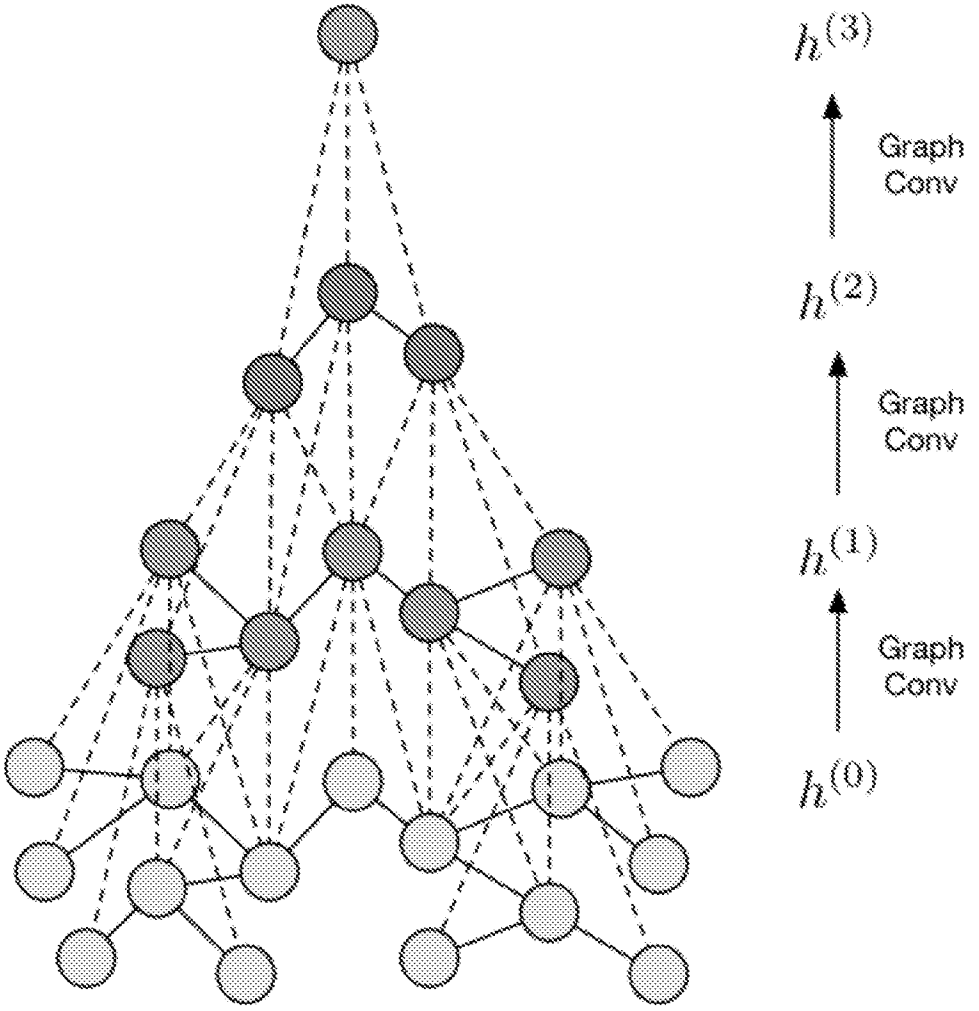


图 2

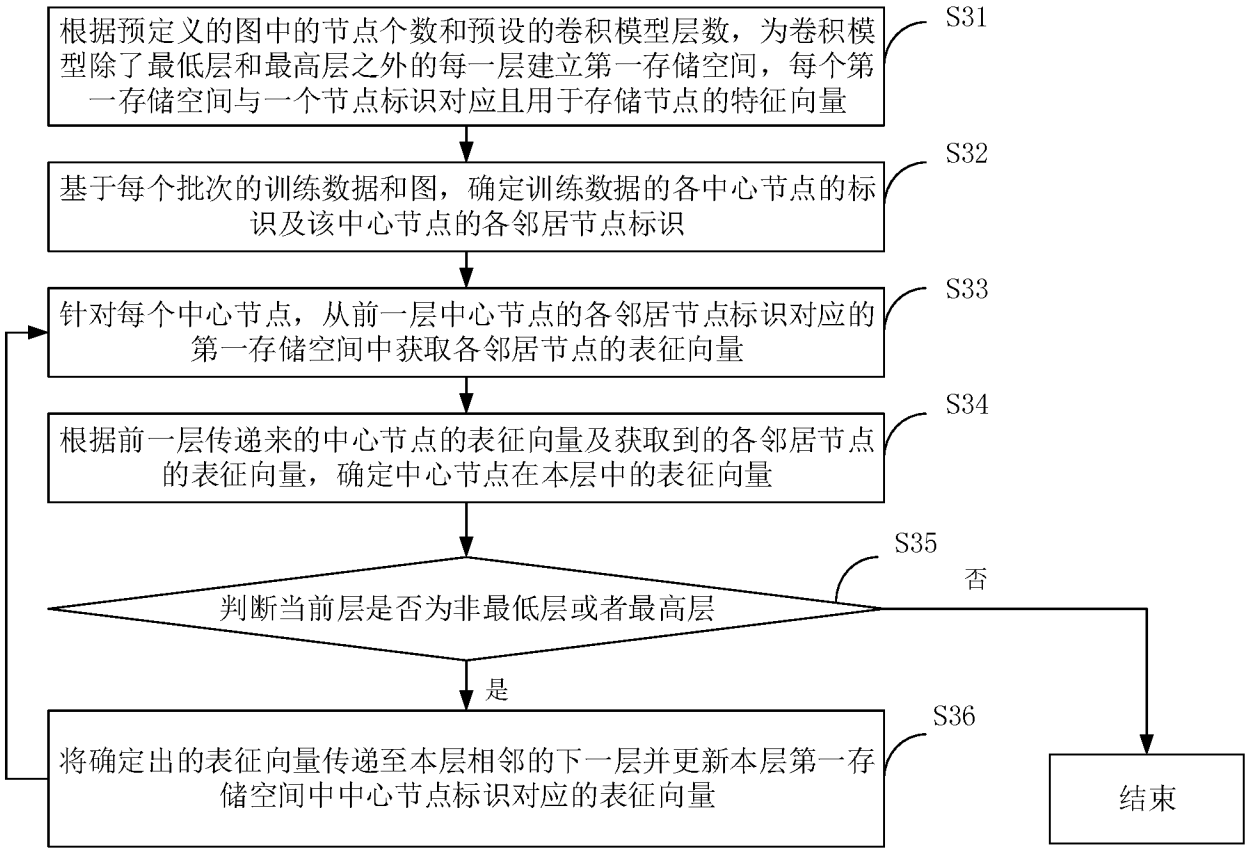

根据预定义的图中的节点个数和预设的卷积模型层数，为卷积模型除了最低层和最高层之外的每一层建立第一存储空间，每个第一存储空间与一个节点标识对应且用于存储节点的特征向量
S31
基于每个批次的训练数据和图，确定训练数据的各中心节点的标识及该中心节点的各邻居节点标识
S32
针对每个中心节点，从前一层中心节点的各邻居节点标识对应的第一存储空间中获取各邻居节点的表征向量
S33
根据前一层传递来的中心节点的表征向量及获取到的各邻居节点的表征向量，确定中心节点在本层中的表征向量
S34
判断当前层是否为非最低层或者最高层
S35
否
是
将确定出的表征向量传递至本层相邻的下一层并更新本层第一存储空间中中心节点标识对应的表征向量
S36
结束

图 3

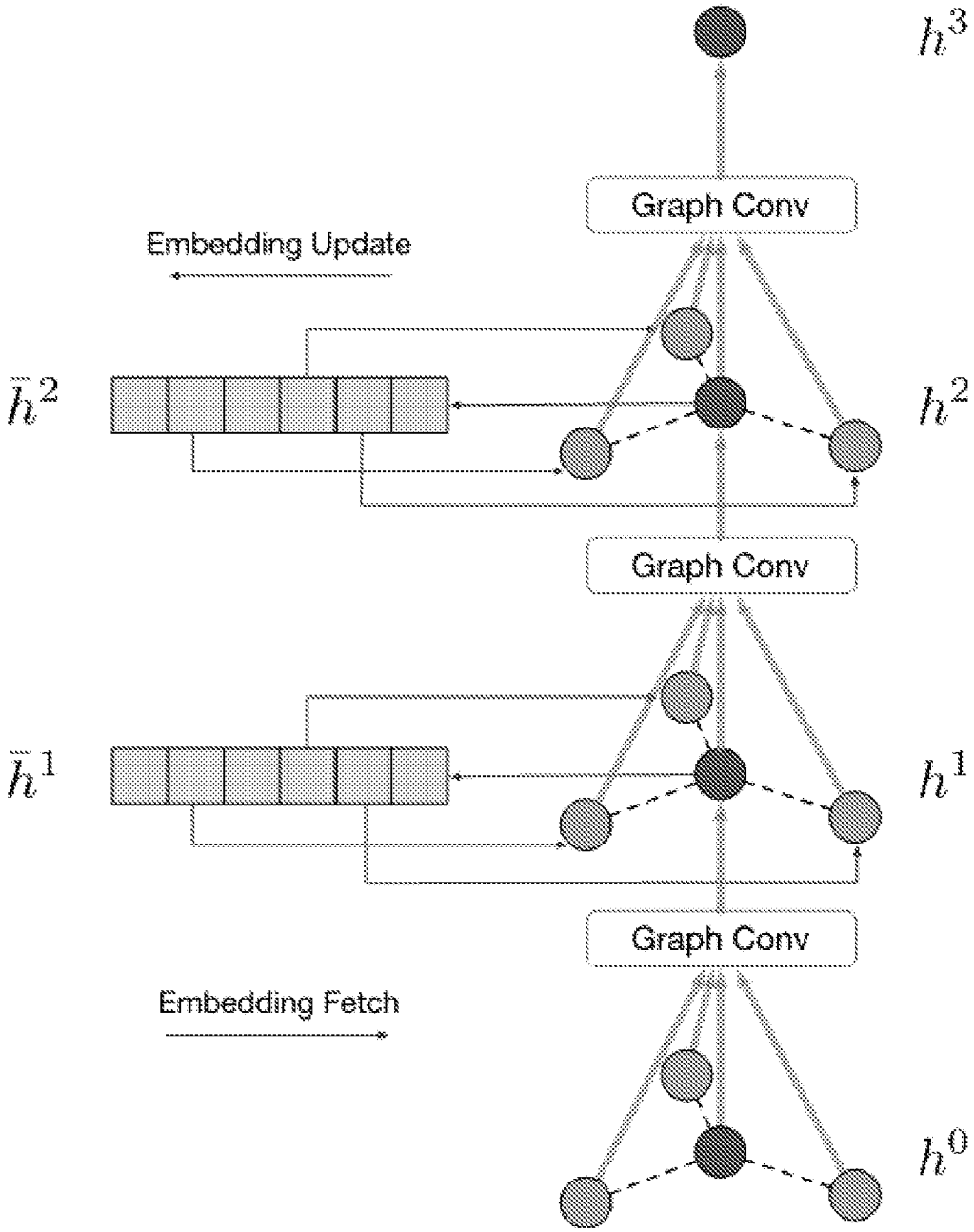

h^3
Graph Conv
Embedding Update
$\bar{h}^2$
h^2
Graph Conv
$\bar{h}^1$
h^1
Graph Conv
Embedding Fetch
h^0

图 4

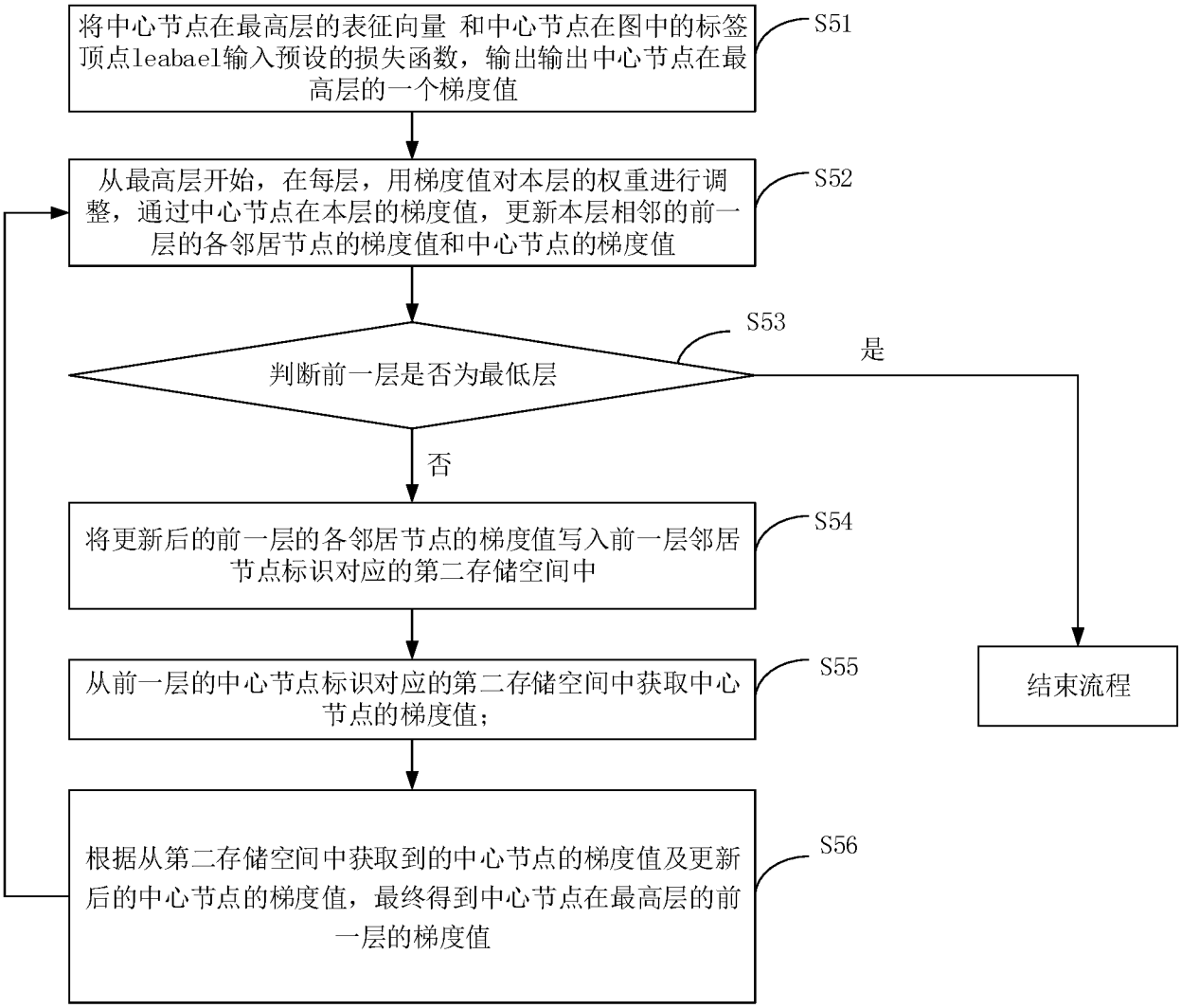

将中心节点在最高层的表征向量 和中心节点在图中的标签顶点leabael输入预设的损失函数，输出输出中心节点在最高层的一个梯度值
S51
从最高层开始，在每层，用梯度值对本层的权重进行调整，通过中心节点在本层的梯度值，更新本层相邻的前一层的各邻居节点的梯度值和中心节点的梯度值
S52
判断前一层是否为最低层
S53
是
否
将更新后的前一层的各邻居节点的梯度值写入前一层邻居节点标识对应的第二存储空间中
S54
从前一层的中心节点标识对应的第二存储空间中获取中心节点的梯度值；
S55
根据从第二存储空间中获取到的中心节点的梯度值及更新后的中心节点的梯度值，最终得到中心节点在最高层的前一层的梯度值
S56
结束流程

图 5

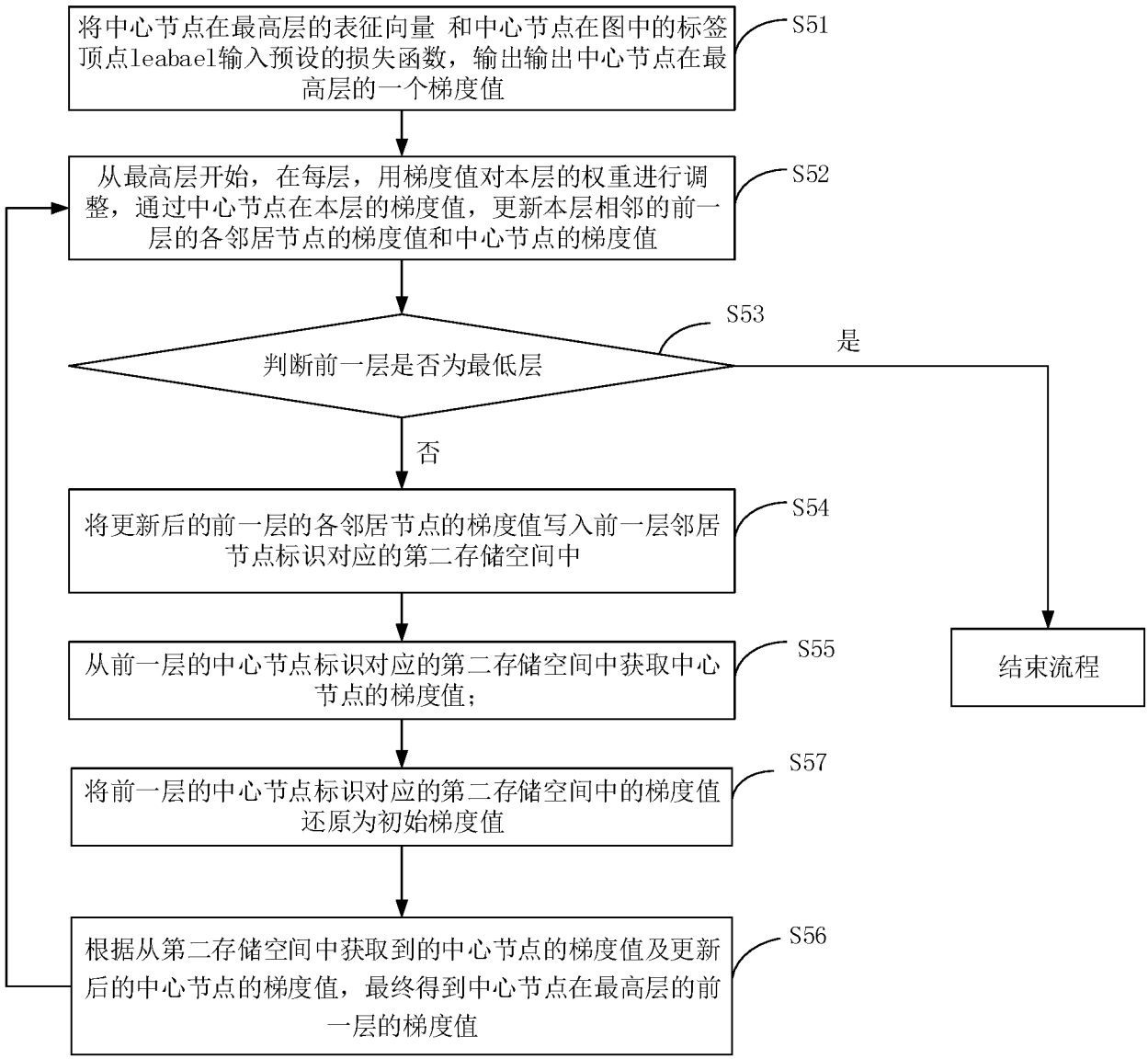

将中心节点在最高层的表征向量 和中心节点在图中的标签顶点leabael输入预设的损失函数，输出输出中心节点在最高层的一个梯度值
S51
从最高层开始，在每层，用梯度值对本层的权重进行调整，通过中心节点在本层的梯度值，更新本层相邻的前一层的各邻居节点的梯度值和中心节点的梯度值
S52
判断前一层是否为最低层
S53
是
否
结束流程
将更新后的前一层的各邻居节点的梯度值写入前一层邻居节点标识对应的第二存储空间中
S54
从前一层的中心节点标识对应的第二存储空间中获取中心节点的梯度值；
S55
将前一层的中心节点标识对应的第二存储空间中的梯度值还原为初始梯度值
S57
根据从第二存储空间中获取到的中心节点的梯度值及更新后的中心节点的梯度值，最终得到中心节点在最高层的前一层的梯度值
S56

图 6

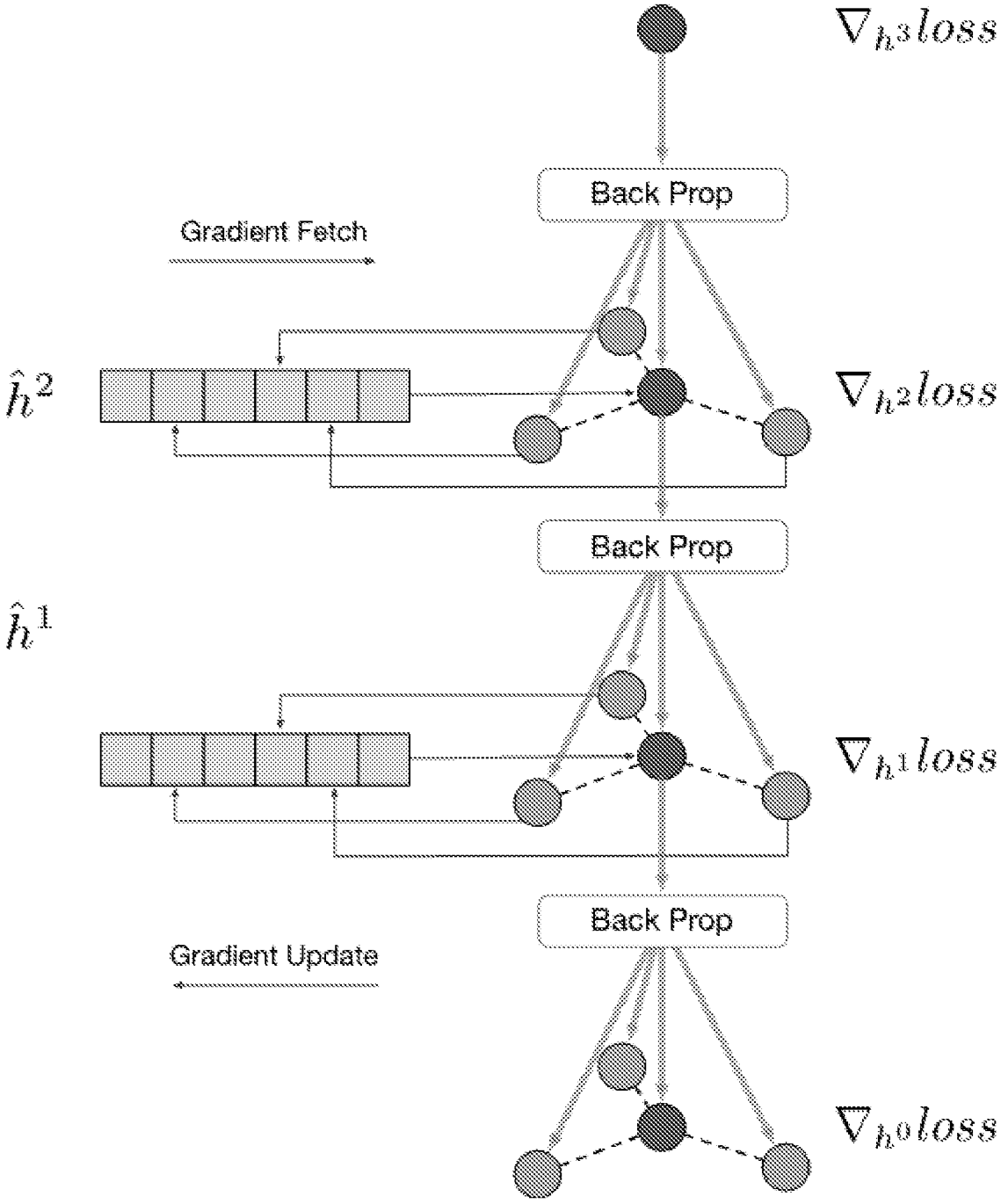

$\nabla_{h^3} loss$
Back Prop
Gradient Fetch
$\hat{h}^2$
$\nabla_{h^2} loss$
Back Prop
$\hat{h}^1$
$\nabla_{h^1} loss$
Back Prop
Gradient Update
$\nabla_{h^0} loss$

图 7

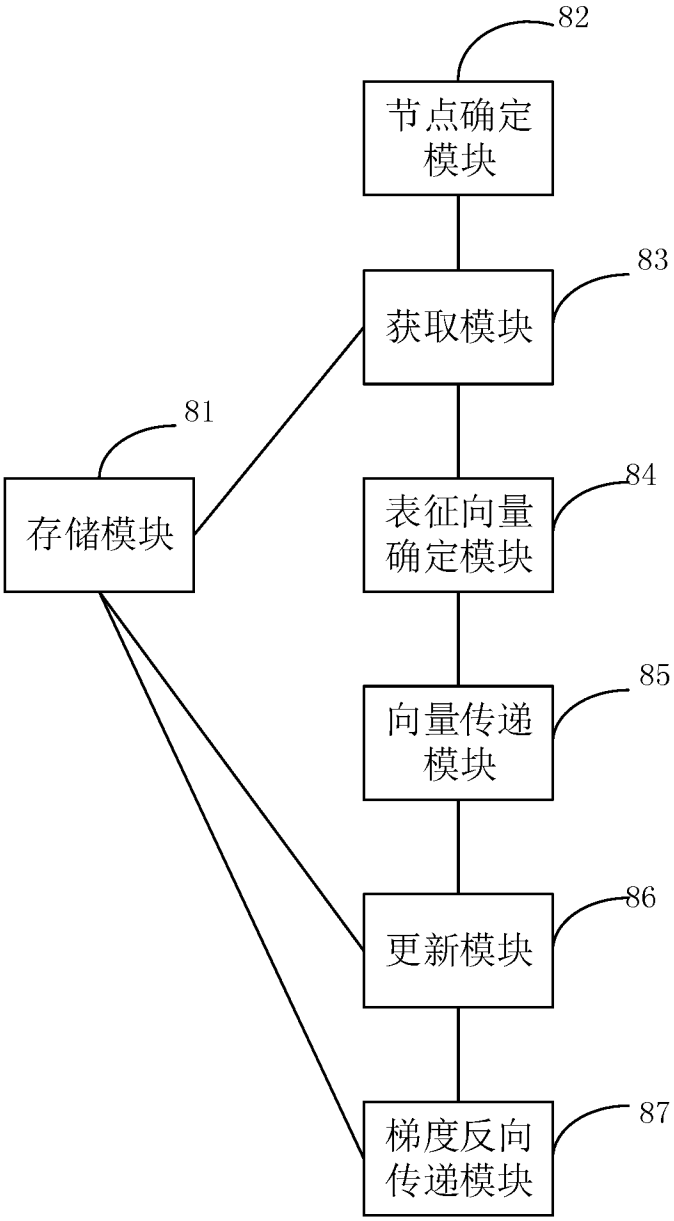

82
节点确定模块
83
获取模块
81
存储模块
84
表征向量确定模块
85
向量传递模块
86
更新模块
87
梯度反向传递模块

图 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2020/070584

A. CLASSIFICATION OF SUBJECT MATTER

G06N 3/08(2006.01)i; G06N 3/04(2006.01)n

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, EPODOC, WPI, CNKI, ISI, IEEE: 图卷积网络, 训练, 邻, 节点, 存储, 内存, 层, 向量, GCN, graph convolutional network, train, adjacent, node, store, layer, vector

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 108664687 A (ZHEJIANG UNIVERSITY OF TECHNOLOGY) 16 October 2018 (2018-10-16) description, paragraphs [0037]-[0048]	1-10
A	CN 108648095 A (ZHEJIANG UNIVERSITY OF TECHNOLOGY) 12 October 2018 (2018-10-12) entire document	1-10
A	CN 107533667 A (GOOGLE INC.) 02 January 2018 (2018-01-02) entire document	1-10
A	US 2003163437 A1 (LEE, S. J. J.) 28 August 2003 (2003-08-28) entire document	1-10
A	KIPF, T. N. "Semi-supervised Classification with Graph Convolutional Networks" *ICLR 2017*, 22 February 2017 (2017-02-22), entire document	1-10

☐ 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	Date of mailing of the international search report
17 March 2020	**26 March 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/CN2020/070584

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 108664687 A	16 October 2018	None	
CN 108648095 A	12 October 2018	None	
CN 107533667 A	02 January 2018	DE 112016002296 T5	15 February 2018
		KR 20170126997 A	20 November 2017
		US 2017103315 A1	13 April 2017
		TW 201729124 A	16 August 2017
		WO 2016186813 A1	24 November 2016
		EP 3298545 A1	28 March 2018
		TW 201706873 A	16 February 2017
		JP 2018518731 A	12 July 2018
		US 2016342889 A1	24 November 2016
		GB 2553055 A	21 February 2018
		US 2019228301 A1	25 July 2019
		JP 2020017281 A	30 January 2020
		DE 202016107442 U1	23 February 2017
US 2003163437 A1	28 August 2003	None	

国际检索报告

国际申请号
PCT/CN2020/070584

A. 主题的分类

G06N 3/08(2006.01)i; G06N 3/04(2006.01)n

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, EPODOC, WPI, CNKI, ISI, IEEE: 图卷积网络，训练，邻，节点，存储，内存，层，向量，GCN, graph convolutional network, train, adjacent, node, store, layer, vector

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 108664687 A (浙江工业大学) 2018年 10月 16日 (2018-10-16) 说明书第[0037]-[0048]段	1-10
A	CN 108648095 A (浙江工业大学) 2018年 10月 12日 (2018-10-12) 全文	1-10
A	CN 107533667 A (谷歌公司) 2018年 1月 2日 (2018-01-02) 全文	1-10
A	US 2003163437 A1 (LEE, Shih-jong J.) 2003年 8月 28日 (2003-08-28) 全文	1-10
A	KIPF, Thomas N. "Semi-supervised Classification with Graph Convolutional Networks" ICLR 2017, 2017年 2月 22日 (2017-02-22), 全文	1-10

☐ 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 3月 17日	2020年 3月 26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 欧晓丹 电话号码 86-(10)-53961416

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2020/070584

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	108664687	A	2018年 10月 16日	无			
CN	108648095	A	2018年 10月 12日	无			
CN	107533667	A	2018年 1月 2日	DE	112016002296	T5	2018年 2月 15日
				KR	20170126997	A	2017年 11月 20日
				US	2017103315	A1	2017年 4月 13日
				TW	201729124	A	2017年 8月 16日
				WO	2016186813	A1	2016年 11月 24日
				EP	3298545	A1	2018年 3月 28日
				TW	201706873	A	2017年 2月 16日
				JP	2018518731	A	2018年 7月 12日
				US	2016342889	A1	2016年 11月 24日
				GB	2553055	A	2018年 2月 21日
				US	2019228301	A1	2019年 7月 25日
				JP	2020017281	A	2020年 1月 30日
				DE	202016107442	U1	2017年 2月 23日
US	2003163437	A1	2003年 8月 28日	无			