

浙江大学

本科课程论文

课程名称： 计算机科学思想史

姓 名： 张溢弛

学 院： 计算机科学与技术学院

专 业： 软件工程 1801

学 号： 3180103772

指导教师： 巫英才等

2021 年 6 月 15 日

目录

一 引言	III
二 研究意义与发展现状	IV
2.1 研究意义	IV
2.2 发展现状	IV
三 图神经网络模型	IV
3.1 卷积图神经网络模型	IV
3.1.1 基于谱分解的方法	IV
3.1.2 谱分解方法的优缺点分析	V
3.1.3 基于空间结构的方法	V
3.2 循环图神经网络	VI
3.2.1 信息融合机制	VI
3.2.2 门控单元和 LSTM	VII
3.3 图自编码器	VII
3.3.1 图嵌入	VII
3.3.2 图生成	VIII
3.4 图注意力网络	VIII
四 图神经网络的应用	IX
4.1 结构化场景下的应用	IX
4.1.1 化学、生物学与药学	IX
4.1.2 知识图谱	X
4.1.3 推荐系统	X
4.2 非结构化场景中的应用	X
4.2.1 图像分类	X
4.2.2 语义分割	XI
4.2.3 文本分类	XI
4.2.4 信息抽取	XI
五 总结与展望	XI

图神经网络文献综述

摘要：图结构作为一种经典的非欧式数据结构在生产生活中都非常常见，社交网络、蛋白质分子结构、图像和文本中隐含的语义信息都可以构成图，而近几年十分火热的深度学习和神经网络模型（如 CNN, RNN, LSTM），虽然在图像等领域取得了不错的表现，但是面对图结构的时候表现仍然不尽人意。图神经网络（GNN）是一种在处理图结构数据中表现非常突出的神经网络架构，可以分为图卷积网络，图循环网络，图自编码器，图注意力网络等多种类型，这类神经网络可以提取图的邻域信息和全局特征，并在节点分类，链接预测，图分类等图学习经典任务中取得了非常好的表现，本文将围绕神经网络的发展，介绍各种类型的图神经网络架构的特点和优缺点，对图神经网络在各个领域的应用展开广泛而深入的探索，并展望其未来的发展方向。

关键词：深度学习、图神经网络、机器学习、图嵌入、图表示学习

一 引言

机器学习 (Machine Learning) 是人工智能 (Artificial Intelligence) 的核心，旨在研究计算机如何模拟或者实现人类的学习行为，让计算机可以从数据中“学习”到解决特定问题的能力。近年来，随着计算机算力的提高，一类被称为深度学习 (Deep Learning) 的方法逐渐取代了传统的统计机器学习成为研究热点。

深度学习通过使用一种被称为神经网络 (Neural Network) 的结构组合成深度学习模型来进行，许多曾经严重依赖于手工提取特征的机器学习任务（如目标检测、机器翻译和语音识别），都已被各种端到端的深度学习范式彻底改变。其中比较有代表性的是卷积神经网络 (CNN) 和循环神经网络以及由它们衍生出的一系列神经网络模型。卷积神经网络衍生出的一系列深度模型在处理图像 [1]，语音等领域的任务中取得了非常好的效果，而以 LSTM 模型为代表的循环神经网络在处理序列化的数据（如文本，时间序列）中表现良好而在自然语言处理等领域 [2] 产生了非常大的影响。

这些经典的神经网络架构都是在欧氏空间中表示图像，音频，文本等不同形式的欧式数据结构（即可以定义出有效的“几何距离”的数据结构），并通过神经网络层挖掘出其中隐含的模式，而图结构作为日常生活中出现非常多的结构，其本质是一种由“实体”（图中的节点）与“对应关系”（图中的边）组成的非欧式数据结构，因而传统的神经网络架构在图数据中的表现就比较差。图神经网络 (Graph Neural Network) 的提出 [3] 有效地解决了深度神经网络在图结构的学习中的应用，本文将围绕图神经网络的现有技术，综述图神经网络的发展和存在的不足，并为探索其未来的发展方向。

二 研究意义与发展现状

2.1 研究意义

图神经网络的主要建模对象是图结构，而图结构在我们的日常生活和科学研究中随处可见，比如人类的人际关系和社交网络，互联网中的各个网页，科研论文中的引用情况，化学中的分子结构等等，都可以构成一系列的图，因而图神经网络拥有广泛的应用场景，研究图神经网络也具有的重要意义，可以帮助我们进一步挖掘图数据结构中隐含的信息。

与图相关的机器学习任务主要包括节点分类、链接预测、图嵌入、图补全、图分类等等，图神经网络在这些与图相关的机器学习任务中都将发挥重要的作用。

2.2 发展现状

而图神经网络从最早提出后经过一系列发展，目前已经发展出了多种不同的模型架构，主要可以分为循环图神经网络，图卷积网络（Graph Convolution Networks, GCN）、图注意力网络（Graph Attention Networks）、图自编码器（Graph Autoencoders）和图时空网络（Graph Spatial-temporal Networks）等多种类型，并且在推荐系统、知识图谱、图像和文本等多种领域中都取得了非常多的应用。

这些图神经网络架构虽然各不相同，各有优缺点，但其核心目标始终是“如何提取图结构中的特征”，下面我们将重点介绍不同类型的图神经网络的特点和不足，以及图神经网络在各个领域取得的应用。

三 图神经网络模型

3.1 卷积图神经网络模型

卷积图神经网络的提出收到了卷积神经网络 CNN 的启发，CNN 中通过卷积层这样一种特殊的神经网络层，使用一系列卷积核来提取图像等结构中的局部特征，取得了非常好的效果，而图卷积神经网络就是将“卷积”操作推广到了图结构中的神经网络，而图作为一种非欧式数据结构，无法直接定义有效的卷积操作，这就衍生出了图卷积神经网络的两种不同的发展方向，分别是基于谱分解的方法和基于空间结构的办法。

3.1.1 基于谱分解的方法

基于谱分解的方法有着非常坚实的数学背景和可解释性，这种方法最早在论文 [4] 中被提出。基于谱分解的方法从图形信号处理的角度引入滤波器，并在傅立叶域中定义图的卷积运算。这里面用到了一种非常重要的图表示方式——拉普拉斯矩阵，图的拉普拉斯矩阵被

定义为:

$$L = I_n - D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \quad (1)$$

其中 D 是图的度数对焦矩阵而 A 是图的邻接矩阵, 这是一种重要的图的特征表示方式, 且是一个实对称的半正定矩阵, 因此可以被分解为 $L = U \Lambda U^T$ 的形式, 并且矩阵 U 是 L 的特征值所构成的对角矩阵。关于谱分解的数学证明这里就不再赘述, 仅作简要介绍。在此基础上, 基于谱分解的图卷积网络将由节点特征表示组成的图信号输入 x 进行图傅立叶变换 $\mathcal{F}(x) = U^T x$, 这种图傅立叶变换将输入的图信号映射到了以图的拉普拉斯矩阵的特征值为基的正交空间中。这样一来图上的卷积操作就很自然地被定义为:

$$\begin{aligned} x *_G g &= \mathcal{F}^{-1}(\mathcal{F}(x) \odot \mathcal{F}(g)) \\ &= U (U^T x \odot U^T g) \end{aligned} \quad (2)$$

这里的 g 表示卷积核, 如果我们定义 $g_\theta = \text{diag}(U^T g)$ 那么这样一来图卷积可以简化成:

$$x *_G g_\theta = U g_\theta U^T x \quad (3)$$

所有的图卷积网络都遵循这种形式的卷积定义, 区别在于卷积核的选取方式。

2011 年, David Hammond 等人 [5] 提出可以用切比雪夫多项式的前 K 阶来逼近 $g_\theta(\Lambda)$ 从而提出了 ChebNet, 2017 年的时候 Thomas Kipf 等人 [6] 在 ChebNet 的基础上将层级卷积运算的 K 限制为 1, 缓解了 ChebNet 中存在的过拟合问题, 并对滤波器的参数进行了简化, 最终提出了 GCN, 而在此基础上又出现了可以挖掘不同节点之间隐式关系的自适应图卷积网络 (AGCN)[7], 这种形式的图卷积网络会从图邻接矩阵中学习出“残差”拉普拉斯矩阵并将其添加到原始的拉普拉斯矩阵中, 并通过超参数的调节实现图结构中的自适应, 可以挖掘出图中的隐藏关系。

3.1.2 谱分解方法的优缺点分析

谱分解方法虽然将卷积操作引入了图结构中, 是一种重大突破, 然而基于谱分解的卷积定义都依赖于图的拉普拉斯矩阵的特征基向量, 而这完全取决于图本身的结构, 这意味着基于谱分解的方法存在一个严重的缺点, 那就是模型必须针对不同结构的图进行针对性的训练, 而不能直接使用到其他的图中, 一言蔽之就是基于谱分解定义的图卷积的泛化能较差。

3.1.3 基于空间结构的方法

和谱分解方法相反, 基于空间结构的方法直接在图上定义卷积运算, 并针对空间上相邻的邻域 (即图中某个节点附近的若干节点组成的区域) 进行卷积运算, 这种方法的主要挑战是如何针对不同大小的邻域来定义卷积并且保持卷积操作的局部不变性 (谱分解的方法通过对不同的图定义不同的卷积核解决了这个问题)。

David Duvenaud 等人在 2015 年 [8] 的时候就提出了针对不同度数的节点使用不同度

数的权重矩阵 $W_t^{|N_v|}$ (N_v 表示节点 v 的邻居节点构成的集合), 这样一来图节点的嵌入表示的更新方式就变成了:

$$\begin{aligned} x &= h_v^{t-1} + \sum_{i=1}^{|N_v|} h_i^{t-1} \\ h_v^t &= \sigma(xW_t^{|N_v|}) \end{aligned} \quad (4)$$

这种模型现将节点自己和所有邻居节点的嵌入表示相加再使用一个权重矩阵进行投影得到新的嵌入表示, 这种模型的主要缺陷在于, 在大规模的图结构中节点的度数往往有很多不同取值, 这就会导致模型的参数过多而难以应用。

而在 2017 年, William Hamilton 等人提出了 GraphSAGE[9] 结构, 这种模型是一种通用的归纳推理框架, 通过采样和聚合相邻节点的特征来生成节点的嵌入表示, 其传播过程可以描述为:

$$\begin{aligned} h_{N_v}^t &= \text{AGGREGATE}(h_u^{t-1}, u \in N_v) \\ h_v^t &= \sigma(W^t[h_v^{t-1} || h_{N_v}^t]) \end{aligned} \quad (5)$$

即将邻居节点的嵌入表示进行一定的聚合 (通过聚合函数 AGGREGATE), 然后和当前节点的嵌入进行组合并投影到嵌入空间中生成新的节点嵌入, 和其他的基于空间结构的模型不同的是, GraphSAGE 并不会使用所有的相邻节点, 而是均匀采样固定数量的邻居节点, 而聚合函数通常有均值聚合, LSTM 聚合和最大池化聚合等多种选择, 分别对应不同的特征提取需求而选用。我们可以将 GraphSAGE 中的聚合操作理解为是对 CNN 中池化操作的一种推广。

而 Federico Monti 等人在论文 [10] 中提出的空间域模型 MoNet 使用一种伪坐标机制实现了非欧式数据结构中的卷积操作, 并衍生出了诸如 GCNN, ACNN 和 DCNN 等多种模型, 而 GCN 也可以被看成是 MoNet 的一种特例。

3.2 循环图神经网络

图神经网络的另一种发展方向是循环图神经网络, 如果说卷积图神经网络是受到卷积神经网络的启发, 那么循环神经网络就是在以 RNN 和 LSTM 等变体为代表的循环神经网络影响下产生的。

与卷积图神经网络相比, 循环图神经网络在不同的层之间共享参数, 这使得参数在训练过程中可以逐渐收敛, 同时也减少了模型训练过程中的参数量, 而卷积神经网络在不同的层之间使用了不同的参数, 可以更好地从不同尺度提取图中的特征信息。

3.2.1 信息融合机制

这种图神经网络架构基于一种信息融合机制 [11], 在更新每个节点对应的隐状态的时候会 and 邻近的节点互相交换信息, 并将自己接受到的邻近节点的信息和自己本身的节点特

征进行融合，更新到下一个隐藏状态中，其核心的状态更新方程可以表示为：

$$h_v^{(t)} = \sum_{u \in N(v)} f(x_v, x_{(v,u)}^e, x_u, h_u^{(t-1)}) \quad (6)$$

这里的函数 f 就是一个用节点的嵌入向量，节点邻居的嵌入向量，节点自身的隐状态，节点接受到的邻居信息作为参数的函数，为了保证隐状态可以收敛，函数 f 必须是一个收缩映射 (Contraction Mapping，一种将向量投影到潜在空间后使其距离收缩的函数)。

信息融合的机制使得循环图神经网络在有向图和无向图中都有较好的表现，同时也可以使用前向传播和反向传播算法进行模型的训练和参数的更新。

3.2.2 门控单元和 LSTM

就好像 RNN 发展成具有门控单元的 GRU 和 LSTM 一样，循环图神经网络也引入门控单元并产生了更多拥有捕捉长距离节点依赖关系的模型，比如门控图神经网络 (Gated GNN)[12] 将门控循环单元 (GRU) 引入了普通的循环图神经网络中，并将循环的次数减少到了一个固定的常数，其优点在于不再需要保证参数最终会收敛，一个节点的隐状态由上一层的自身隐状态和从别的节点中传递过来的隐状态通过门控单元函数计算而成，即：

$$h_v^{(t)} = \text{GRU} \left(h_v^{(t-1)}, \sum_{u \in N(v)} W h_u^{(t-1)} \right) \quad (7)$$

而在基于树和图结构的传播过程中，LSTM 也可以起到和 GRU 相似的作用，Kai Sheng Tai 等人于 2015 年提出了由基本的 LSTM 单元扩展而来的 Tree-LSTM 来聚合节点的特征信息，而 2018 年由 Victoria Zayats 等人提出的图结构 LSTM 对其进行了一定的简化，在弱化的数据集上取得了更好的效果，除此之外还有将文本转化成一个图结构并进行文本编码的 S-LSTM 等工作也是对基于 LSTM 的循环图神经网络的进一步发展。

我们可以看到在引入了循环函数和门控单元等循环神经网络中的组件之后，图神经网络拥有了更强的处理序列化数据的能力，可以捕捉到距离更远的节点之间的依赖关系，并且常常与自然语言、文本等序列化数据的处理工作相结合，取得了比较好的表现。

3.3 图自编码器

图自编码器 (GAE) 是一种将节点映射到低维特征空间并通过潜在表示来解码图结构信息的神经网络架构，图自编码器可以被用于学习图嵌入和生成新的图，下面我们将围绕这两方面的工作展开介绍图自编码器的研究现状。

3.3.1 图嵌入

图自编码器通过一个编码器来学习出图嵌入表示，而通过一个解码器从图嵌入表示中解析出图的拓扑结构。早期的图自编码器通过多层感知机来学习图节点的特征，而图表示

深度神经网络 (DNNGR)[16] 使用一系列降噪自编码器对多层感知机生成的 PPMI 矩阵进行编解码。结构化深度网络嵌入 (SDNE)[17] 通过一系列自编码器来保留节点的一阶和二阶逼近特征。

然而 DNNGR 和 SDNE 两种自编码器只考虑了图的结构信息，忽略了节点自身包含的特征信息，而图自编码器 (GAE)[18] 由卷积图神经网络发展而来，将节点的结构信息和自身的特征信息进行融合，这种模型的架构包含两个图卷积层，其编码方式可以表示为：

$$Z = \text{enc}(X, A) = \text{Gconv}(f(\text{Gconv}(A, X; \Theta_1)); \Theta_2) \quad (8)$$

这里的函数 f 往往使用 ReLU 函数作为激活函数，而 Θ 则表示每个卷积层的可学习参数。而在 GAE 的基础上又发展出了 VGAE, ARVGA, DRNE 等多种新的模型。

3.3.2 图生成

GAE 可以从多个图中通过将图进行编码学习到图的总体分布模式，并且可以用于生成新的图。这种功能主要被用在分子结构预测和药物产物预测等任务中，图生成的方法主要包括序列化方式和全局化方式两种。

序列化方式通过从一步步处理节点和边来生成一张图，比如图深度生成模型 (Deep-GMG)[19]，这种模型通过所有节点可能的排列对应的概率来预测最终生成的图，并挑选出可能性最大的：

$$p(G) = \sum_{\pi} p(G, \pi) \quad (9)$$

而全局化方式直接一次性生成整张图的结构，比较著名的模型有图变分编码器 (Graph-GAE)[20]，GraphGAE 将图中已经存在的节点和边看成独立的随机变量，先得到后验概率分布 $q_{\phi}(z|G)$ ，然后对于输入的图嵌入 z 用一个生成概率分布 $p_{\theta}(G|z)$ 并根据这个概率生成最合适的图，这个模型需要优化的目标函数可以表示为：

$$L(\phi, \theta; G) = E_{q_{\phi}(z|G)} [-\log p_{\theta}(G | z)] + KL[q_{\phi}(z | G) || p(z)] \quad (10)$$

一般来说 GraphGAE 输出的结果包含图的邻接矩阵以及节点和边的特征，而在药物生成推断等具体的应用领域，也有 MolGAN 等图生成模型取得了比较好的效果。

3.4 图注意力网络

注意力机制在基于序列的各项任务，如自然语言处理的机器翻译中取得了巨大的成功，从 Seq2Seq 模型中的注意力机制到最近几年大热的 Transformer 模型都用到了不同的注意力机制，注意力机制可以为序列分配不同的权重以达到对“注意力”的模拟，就像卷积和循环门控单元可以引入到 GNN 中一样，注意力机制也可以引入 GNN 中，形成图注意力网络模型，为不同的节点分配不同的注意力分数，从而识别出比较重要的节点。图注意力网络可

以看成是卷积图神经网络的一种衍生类别。

Petar 等人于 2018 年提出一种被称为 GAT 的图注意力网络 [21]，这种网络在信息传播的过程中引入了自注意力机制，节点的隐状态是通过针对邻域计算注意力分数得到的：

$$h_v^{(k)} = \sigma \left(\sum_{u \in \mathcal{N}(v) \cup v} \alpha_{vu}^{(k)} W^{(k)} h_u^{(k-1)} \right) \quad (11)$$

而注意力分数通过堆叠简单的图注意力层计算而来，并采用 LeakyReLU 作为激活函数，其计算方式是：

$$\alpha_{vu}^{(k)} = \text{softmax} \left(g \left(a^T [W^{(k)} h_v^{(k-1)} \| W^{(k)} h_u^{(k-1)}] \right) \right) \quad (12)$$

GAT 模型具有如下的特点：

- 计算可并行化，计算效率高
- 可以处理不同度数的节点并为其分配相应的权重
- 可以很容易滴应用在归纳学习中

因此 GAT 在半监督节点分类和链接预测等多项任务中胜过 GCN，而门控注意力网络 (GaAN) 在多头注意力的机制上引入了门控单元，并使用基于 key-value 的注意力机制，而 GAT 计算注意力系数的时候使用的是全连接层，在归纳式的节点分类任务中，GaAN 的表现比 GAT 和其他使用了不同聚合函数的 GNN 模型都要好。

四 图神经网络的应用

图神经网络在很多领域都有着广泛的应用，在结构化的场景中，数据本身就是按照图结构来组织的，比如社交网络预测，交通流量预测，推荐系统和传统的图表示学习，而在非结构化的场景中，数据本身不是图的结构，我们通过一定的方式将数据构造成图的结构，并将图神经网络应用在这样的图结构中以实现一定的目的。

4.1 结构化场景下的应用

4.1.1 化学、生物学与药学

化学和生物学中的分子和蛋白质是可以用来表示的结构化尸体，用原子或者残基为节点，化学键为边，通过 GNN 的表示学习学习到的向量可以帮助进行化学反应预测和相互作用预测。

比如在化学反应预测中，Kien Do 等人的论文 [129] 中将化学反应看成图转换的过程并设计了 GTPN 模型，使用 GNN 来学习反应物和产物的表示向量，然后使用强化学习以键变化的形式以此来预测反应物转化成产物的最佳反应路径。

4.1.2 知识图谱

知识图谱是一种用有向图来表示的知识库，用节点表示实体，用边来表示关系，知识图谱在推荐、搜索、问答等领域取得了广泛的应用。而 GNN 也在知识图谱补全、知识图谱归纳式嵌入和知识图谱对齐等任务。比如 Chao shang 等人提出的端到端的结构感知卷积网络 SCAN 可以使用堆叠的 GCN 层来学习实体和关系的嵌入表示，在知识图谱的补全中取得了非常好的写过

4.1.3 推荐系统

推荐系统往往采用基于用户-商品评分图的模型，将用户和商品和属性看成是图中的节点，并将其中的联系看成是边，边的值就表示交互的结果，译者张方式，商品推荐问题被转换成了图的连接预测问题，而 GNN 凭借着强大的表示能力和比较好的可解释性在推荐系统中取得了成功。

比如可分离的循环多图卷积网络 (sRMGCNN) 考虑了多图 CNN 和 RNN 的组合，使用用户-商品图的行列编码的相似度信息，从评分矩阵中提取局部信息，然后将这些局部特征输入 RNN 中进行评分值的传播和评分矩阵的重建，同时采用了矩阵分解模型，通过将评分矩阵分解成两个低维的矩阵来提高计算的效率。

而在社交推荐领域基于 GNN 的神经扩散模型和 GraphRec 模型也取得了非常好的效果。

4.2 非结构化场景中的应用

GNN 应用于非结构化的场景主要有两种方式，第一是结合其他领域的结构信息来提高性能，比如使用知识图谱中的信息来缓解图像任务中的零样本问题，二是先对场景中的关系结构进行推断，然后再应用 GNN 模型来解决定义在图上的问题，比如将文本建模为图。

4.2.1 图像分类

图像分类是深度学习中最经典的任务之一，基于 CNN 及其衍生模型的深度学习方法已经在图像分类中的多个图像分类数据集中取得了 SOTA 的效果，而图神经网络在图像分类中的作用主要体现到零样本 (Zero-Shot) 学习和少样本 (Few-Shot) 学习任务中，知识图谱可以用来指导零样本识别分类的额外信息，使用具有较大领域的单层 GCN 构造的零样本分类器比现有的深度模型更加高效。

除了知识图谱，图像之间的相似性也有助于少样本学习，2018 年提出的基于像思想构建加权的全连接图像网络通过在图结构中进行消息传递，达到少样本识别的目的。

4.2.2 语义分割

语义分割也是计算机视觉中的经典任务，是“理解”图像的关键步骤，语义分割指的是将图像上的每个像素都分配唯一的标签或者类别，因而这是一种稠密分类的问题，但是图像中的区域通常不是网格状的，并且需要非局部的信息，这导致传统的 CNN 在语义分割问题上的表现并不好，因为 CNN 的核心操作卷积只能提取出图像局部的特征。

而循环图神经网络中的 Graph-LSTM 等模型通过基于距离的超像素映射形式构建图结构并把 LSTM 应用到全局范围内进行邻域信息的扩散传播 [13]，可以在语义分割任务中取得比较好的效果，此外诸如三维 GNN，PointNet 等模型也在语义分割的不同子任务中取得了比较好的效果。

4.2.3 文本分类

文本分类是自然语言处理领域中一个重要的问题，经典的 GCN 模型和 GAT 模型被应用在文本分类问题中，比如通过将文本转化成“词图”，然后对词图进行卷积运算从而提取文本的特征并以此为特征进行文本的分类 [14]。

这种方式将文本看成是由一系列单词节点组成的图，或者依靠文档中的引用关系来构建图，而 Liang Yao 等提出的构建异构语料图并使用 Text-LSTM 进行学习则是另一种 GNN 应用到文本分类的实例。

情感分析是一类具体的文本分类任务，针对这类任务，使用含有 LSTM 单元的图神经网络可以取得较好的表现。

4.2.4 信息抽取

信息抽取是从文本中抽取出一定的信息，比如关系抽取、事件抽取等等，一些信息抽取模型将信息抽取分为两个部分，一部分是命名实体的识别 (NER)，另一部分是关系抽取，即需要先从文本中识别出一系列实体，再进一步挖掘其具体的关系。

而 2018 年 Yuhao Zhang 等人提出了一种针对关系抽取任务的卷积图神经网络扩展模型，将剪枝策略应用到了输入的树状 LSTM 结构中。而 Nanyun Peng 等人提出了基于图 LSTM 的跨句 N 元关系抽取通用框架，是循环图神经网络结构在信息抽取中的一个成功实践。

五 总结与展望

通过对已有的神经网络架构的迁移和拓展，图神经网络不断引入新的组件 (卷积，循环，门控，注意力等等)，实现了对图数据结构的建模和特征提取，在多种领域的多种问题上取得了不俗的表现。论文 [22] 在总结了图神经网络的发展现状之后提出图神经网络未来的发展方向和仍然需要解决的问题主要有：

- 模型的深度：理论上来说，只要卷积层的层数足够大，图卷积网络可以提取到任意精确度的特征，而图节点的特征表示也都将收敛到某个特定的值上，CNN 等神经网络已经在深度的问题上取得了突破，一些复杂的，层数非常大的 CNN 模型在图像任务中取得了非常好的效果，而图神经网络是否要从一两个卷积层走向深度，仍然是一个值得思考的感觉。
- 可扩展性的权衡：图神经网络在可扩展性上也需要有更长远的思考，许多 GNN 模型在特定问题上可以提取到更多的特征，拥有更好的表现，但是可扩展性比较差，而可扩展性好的模型也未必能在所有任务中都取得最好的效果。在特征提取的过程中，不管是使用采样还是聚类，都会失去图中存储的部分信息，如果使用采样的方式进行节点嵌入的学习可能会丢失一部分邻居节点的信息，而如果采用聚类的方式，一种模式往往只能应用在一张图里面，这也是需要进行权衡的。
- 异构图问题：目前的很多 GNN 模型都是用来解决同构图的，这也导致这些模型难以直接应用在异构图的问题上（异构图就是有多种不同类型的节点和边的图，比如文本和图像），因此，探索在异构图中效果好的 GNN 模型也是 GNN 未来发展的一个大方向。
- 动态图问题：动态图是一种节点和边会随着时间空间的变化而变化的图，在生产生活中非常常见，目前的图神经网络的研究对动态图的研究还比较少，未来也将是 GNN 研究的一个大发展方向。
- 低资源学习：低资源学习即使用非常少的样本进行图结构的学习，并完成一系列零样本和少样本的学习任务，低资源学习的场景这在现实世界中也是非常常见的，需要我们对未知的特征/类别作出判断，这也是 GNN 的一个未来发展方向。

综上所述，图神经网络虽然已经发展出了许多种各有所长的模型，解决了现实世界中的一些与图结构有关的机器学习问题，但是也依然在诸多方面存在着或者或那的不足，仍然有非常大的发展空间，相信图神经网络在未来会继续发展并取得更多成就。

参考文献

- [1] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2012. ImageNet classification with deep convolutional neural networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'12). Curran Associates Inc., Red Hook, NY, USA, 1097-1105.
- [2] Chen, D. 2017. Enhanced LSTM for Natural Language Inference. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 1657-1668). Association for Computational Linguistics.

- [3] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, “The graph neural network model,” *IEEE Transactions on Neural Networks*,
- [4] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, “The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains,” *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 83–98, 2013.
- [5] M. Defferrard, X. Bresson, and P. Vandergheynst, “Convolutional neural networks on graphs with fast localized spectral filtering,” in *Proc. of NIPS*, 2016, pp. 3844–3852.
- [6] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” in *Proc. of ICLR*, 2017.
- [7] R. Li, S. Wang, F. Zhu, and J. Huang, “Adaptive graph convolutional neural networks,” in *Proc. of AAAI*, 2018, pp. 3546–3553.
- [8] D. K. Duvenaud, D. Maclaurin, J. Iparraguirre, R. Bombarell, T. Hirzel, A. Aspuru-Guzik, and R. P. Adams, “Convolutional networks on graphs for learning molecular fingerprints,” in *Proc. of NIPS*, 2015, pp. 2224–2232.
- [9] W. Hamilton, Z. Ying, and J. Leskovec, “Inductive representation learning on large graphs,” in *Proc. of NIPS*, 2017, pp. 1024–1034.
- [10] F. Monti, D. Boscaini, J. Masci, E. Rodola, J. Svoboda, and M. M. Bronstein, “Geometric deep learning on graphs and manifolds using mixture model cnns,” in *Proc. of CVPR*, 2017, pp. 5115–5124.
- [11] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, “The graph neural network model,” *IEEE Transactions on Neural Networks*, vol. 20, no. 1, pp. 61–80, 2009.
- [12] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, “Learning phrase representations using rnn encoder-decoder for statistical machine translation,” in *Proc. of EMNLP*, 2014, pp. 1724–1734.
- [13] D. Xu, Y. Zhu, C. B. Choy, and L. Fei-Fei, “Scene graph generation by iterative message passing,” in *Proc. of CVPR*, vol. 2, 2017.
- [14] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” in *Proc. of ICLR*, 2017.
- [15] L. Song, Y. Zhang, Z. Wang, and D. Gildea, “A graph-to-sequence model for amr-to-text generation,” in *Proc. of ACL*, 2018.

- [16] S. Cao, W. Lu, and Q. Xu, “Deep neural networks for learning graph representations,” in Proc. of AAAI, 2016, pp. 1145–1152.
- [17] D. Wang, P. Cui, and W. Zhu, “Structural deep network embedding,” in Proc. of KDD. ACM, 2016, pp. 1225–1234
- [18] T. N. Kipf and M. Welling, “Variational graph auto-encoders,” NIPS Workshop on Bayesian Deep Learning, 2016.
- [19] Y. Li, O. Vinyals, C. Dyer, R. Pascanu, and P. Battaglia, “Learning deep generative models of graphs,” in Proc. of ICML, 2018.
- [20] M. Simonovsky and N. Komodakis, “Graphvae: Towards generation of small graphs using variational autoencoders,” in ICANN. Springer, 2018, pp. 412–422.
- [21] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, “Graph attention networks,” in Proc. of ICLR, 2017.
- [22] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang and P. S. Yu, “A Comprehensive Survey on Graph Neural Networks,” in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 4–24, Jan. 2021