**报告正文**

参照以下提纲撰写，要求内容翔实、清晰，层次分明，标题突出。**请勿删除或改动下述提纲标题及括号中的文字。**

**（一）立项依据与研究内容**（建议8000字以内）：

1．**项目的立项依据**

具有地域特色的特色农产品,相对于一般农产品来说,无论是在形态还是品质上都有很大优势,其具有明显的地域特色、优良品质和特殊功效。正是由于特色农产品的诸多优点,使其具有较好的经济效益。但是在经济利益的驱使下,部分无良商家产地造假、浑水摸鱼,以劣质农产品代替特色农产品,致使食品问题频发,对人民的生命安全造成了一定的威胁。因此对特色农产品产地鉴别研究具有重大经济应用价值和现实意义。近年来随着科技的突飞猛进,农产品溯源技术也得到了不断的发展与完善,逐步形成了技术体系。多种多样的农产品产地溯源技术为我国的食品安全及农产品溯源做出了重要贡献,有效保护了特色农产品的品牌效益。农产品产地溯源主要是通过分析表征不同地域来源农产品特征性指标,并以此特征指标来实现农产品产地溯源。目前主要采用质谱、光谱和分子生物学等技术,通过分析农产品的有机组成、挥发性成分、同位素含量与比率等特征成分或指标,结合化学计量学研究方法,建立起能区分农产品产地来源的特征性指纹图谱,从而对不同产地的农产品进行产地溯源。

高光谱图像技术是一种既可以获得待成像对象图片信息又可以获得其光谱图像的一种技术。这种技术具有图像和光谱信息合一的特点。像近红外、拉曼光谱等溯源技术都是只能获得待检测对象光谱信息和图像信息的一种,正因为高光谱图像技术所具有的这种发展潜力,近几年高光谱图像技术在农产品溯源中得到了越来越多的应用。

图卷积网络（Graph Convolutional Network，GCN）是近年来逐渐流行的一种神经网络结构。不同于只能用于网格结构（grid-based）数据的传统网络模型 LSTM 和 CNN，图卷积网络能够处理具有广义拓扑图结构的数据，并深入发掘其特征和规律，例如 PageRank、引用网络、社交网络、通信网络、蛋白质分子结构等一系列具有空间拓扑图结构的不规则数据。CNN 在处理图像数据时具有很强的特征抽取能力和整合能力，这得益于卷积核的参数共享机制和加权平均机制。卷积本质上就是一种加权求和的过程，而卷积核的参数就是不同像素点对应的权重，并且不同的图片都共享同一个卷积核，这使得CNN能够通过对卷积核参数的迭代更新来隐式的学习图像中具有的像素排列规律，进而学习到不同的形状特征和空间特征。但值得注意的一点是，CNN 所处理的数据都具有规则的网格结构，也就是排列很整齐的矩阵，具有 Euclidean Structure，例如 RGB 图片。如果要将CNN应用于非图像领域，就必须将数据组合为规整的网络结构，才能作为CNN的输入。

然而现实生活和科学研究中有很多数据都不具备完整的矩阵结构，相反，更多的是以一定的连接关系聚合在一起，社交网络，通信网络，互联网络等都具有类似的结构。

类似这样的网络结构就是图论中所定义的拓扑图。更一般的，图就是指图论中用顶点和边建立相应关系的拓扑图。可以用一个点和边的集合来表示图：G=(E,V)；其中E表示边的集合，V表示顶点的集合。

那么对于这种具有拓扑图结构的数据而言，CNN 处理起来是非常困难的，而且通常不能很好的抽取节点与节点之间的连接关系信息。根本的原因还是在于数据的多样性，广义上来讲，任何数据在赋范空间内都可以建立拓扑关联。所以说拓扑连接是一种广义的数据结构，因此GCN 有很大的应用空间。

CNN无法处理Non Euclidean structure data, 因为在Non Euclidean structure data上无法保持平移不变性。在拓扑图中，每个顶点相邻的邻居节点数目都不相同，那么也就无法用同一个尺寸的卷积核进行卷积。由于CNN无法在Non Euclidean structure data上做处理，可是人们又希望能够在这样的数据上有效提取空间特征来进行机器学习，所以就GCN成为了研究重点。

GCN能够通过对样本（或顶点）之间的关系进行建模来有效地处理图结构数据。因此，自然可以使用GCN来建模HS图像中的空间关系，而在CNN中则无法做到这一点。

2．**项目的研究内容、研究目标，以及拟解决的关键科学问题**（此部分为重点阐述内容）**；**

HS图像以其丰富而详细的光谱信息为特征，通过从与其像素相关联的光谱特征的连续形状捕获更细微的差异，使得能够更有效地辨别感兴趣的对象（特别是光谱相似类别中的对象）。与RGB和多光谱（MS）数据相比，HS图像能够以更精细和准确的水平检测和识别物质。然而，材料之间的高光谱混合[5]和光谱可变性以及复杂的噪声效应[6]给从此类数据中提取鉴别信息带来了困难。

相比而言，图卷积网络(graph convolutional networks, GCNs)[32]是一个热点和新兴的网络架构，它能够通过建模样本(或顶点)之间的关系来有效地处理图结构数据。目前，GCNs在HS图像分类中不如CNNs受欢迎。尽管有一些著作与GCNs在HSI分类中的使用有关。Shahraki et al。[33]提出将一维CNNs和GCNs级联用于HS图像分类。Qin等人[34]通过同时考虑空间邻域和光谱邻域，将原来的GCNs扩展到二阶版本。Wan等[35]对HS图像进行超像素分割，并将其输入GCN以降低计算成本，提高分类精度

然而，GCNs在以下方面存在一些潜在的局限性:

1. 计算成本高(由于构造邻接矩阵)是GCNs在HS图像分类任务中的一个重要瓶颈，特别是在使用大规模的HS图像数据时
2. GCNs只允许全批网络学习，即一次性将所有样本输入网络。这可能会导致较大的内存开销和缓慢的梯度下降，以及变量更新的负面影响。
3. 最后但并非最不重要的是，一个训练过的基于gcn的模型在没有重新训练网络的情况下无法预测新的输入样本(即样本外)，这对GCNs在实践中的使用有很大的影响。

为了克服这些困难，本课题旨在研究一种简单而有效的小批量GCN (miniGCN)。

1. 与CNNs类似，miniGCNs能够以小批量的方式在一个下采样图(或拓扑结构)上有效地训练网络进行分类，同时学习到的网络也能进行分类模型可以直接用于新数据的预测目的。
2. 本课题将系统地分析CNNs和GCNs，并重点分析农产品的HS图像分类。顾名思义，miniGCNs可以以小批量的方式进行训练，试图在计算复杂度一定的情况下找到更好、更稳健的局部最优。
3. 不同的网络架构能够提取不同的HS图像表示，例如CNNs中的空间-光谱特征，GCNs中样本之间的拓扑关系。一般来说，单个模型由于缺乏特征多样性，可能无法提供最优的性能结果。本课题将自然提出通过融合不同的模型或特征，通过联合训练CNNs和GCNs来增强特征识别能力。与传统的GCNs不同，本文提出的miniGCNs可以进行小批量学习，并可以与标准的CNN模型相结合。这就产生了一个端到端融合网络。考虑了三种融合策略:加性(A)、元素乘性(M)和连接(C)。三种融合模型(A、M、C)分别可表示为:

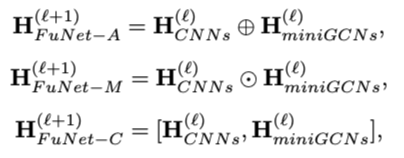


图 1

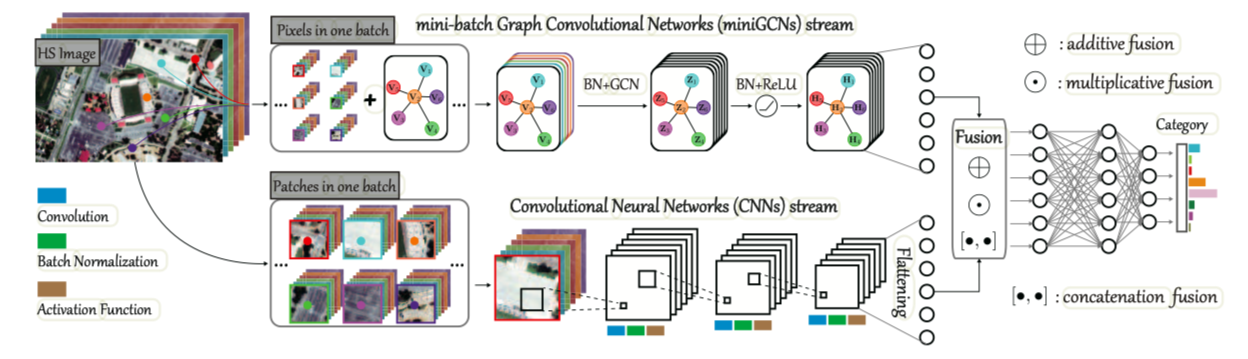


图 2

3．**拟采取的研究方案及可行性分析**（包括研究方法、技术路线、实验手段、关键技术等说明）；

1)数据准备:众所周知，在HS图像分类中，cnn的输入是逐patch的，输出是一组一次性编码的标签。与cnn不同的是，GCNs通过一个邻接矩阵将像素样例输入网络，该矩阵对样本之间的关系进行建模，需要在训练过程开始之前进行计算。

2)特征表示:CNNs能够从HS图像中提取近距离区域丰富的空间和光谱信息，而GCNs能够通过样本之间的图结构来建模样本之间的中远空间关系。

3)网络训练:cnn作为DL家族的主要成员，通常通过使用小批量策略进行训练。相反，GCNs只允许全批网络训练，因为所有样本都需要同时输入网络。

4)计算代价:一层CNNs和GCNs的计算代价主要由矩阵乘积决定，显然，对于较大的图，GCNs相比CNNs计算起来比较复杂，这是因为矩阵相乘的时间比较大。为此，一个可行的解决方案可能是在GCNs中执行的小批量策略。

4．**本项目的特色与创新之处；**

1)构建农产品检测的HS图像的邻接矩阵，用于图卷积分类

2）解决图卷积必须一次性进行全网计算的难题

3）降低图卷积在HS上计算的复杂度

4）将CNN和GCN的分类效果融合，以期获得更好的分类效果。

5．**年度研究计划及预期研究结果**（包括拟组织的重要学术交流活动、国际合作与交流计划等）。

**（二）研究基础与工作条件**

1．**研究基础**（与本项目相关的研究工作积累和已取得的研究工作成绩）；

2．**工作条件**（包括已具备的实验条件，尚缺少的实验条件和拟解决的途径，包括利用国家实验室、国家重点实验室和部门重点实验室等研究基地的计划与落实情况）；

3．**正在承担的与本项目相关的科研项目情况**（申请人正在承担的与本项目相关的科研项目情况，包括国家自然科学基金的项目和国家其他科技计划项目，要注明项目的名称和编号、经费来源、起止年月、与本项目的关系及负责的内容等）；

4．**完成国家自然科学基金项目情况**（对申请人负责的前一个已结题科学基金项目（项目名称及批准号）完成情况、后续研究进展及与本申请项目的关系加以详细说明。另附该已结题项目研究工作总结摘要（限500字）和相关成果的详细目录）。

**（三）其他需要说明的问题**

1. 申请人同年申请不同类型的国家自然科学基金项目情况（列明同年申请的其他项目的项目类型、项目名称信息，并说明与本项目之间的区别与联系）。

2. 具有高级专业技术职务（职称）的申请人是否存在同年申请或者参与申请国家自然科学基金项目的单位不一致的情况；如存在上述情况，列明所涉及人员的姓名，申请或参与申请的其他项目的项目类型、项目名称、单位名称、上述人员在该项目中是申请人还是参与者，并说明单位不一致原因。

3. 具有高级专业技术职务（职称）的申请人是否存在与正在承担的国家自然科学基金项目的单位不一致的情况；如存在上述情况，列明所涉及人员的姓名，正在承担项目的批准号、项目类型、项目名称、单位名称、起止年月，并说明单位不一致原因。

1. 其他。