论文翻译：

基于视觉图神经网络的星地链路的语义通信框架

## 摘要

像地球观测任务这类新兴在轨应用的发展，将导致星地链路的数据流量迎来爆炸性增长。然而，当前的星地通信链路带宽受到严重的限制，星地链路的时间窗口持续时间非常有限。为解决这个问题，我们提出了一种新的用于星地数据处理的语义通信框架——SCViG（Semantic Communication Visual Graph neural network）。SCViG使用视觉图神经网络 (ViG) 作为语义提取器，利用图神经网络(GNN)在遥感图像处理中的优势，在保持遥感图像推理可用性的同时，有效压缩遥感图像的数据量。与现有的语义通信框架相比，SCViG不再需要卫星和地面站之间的频繁通信，从而减轻了星地链路的通信负担。我们使用基于真实世界的数据集进行的实验表明，我们的架构在多分类任务和多标签分类场景的数据处理应用中优于基线方法，包括Vision Transformer (ViT)和ResNet。此外，它还具有良好的抗干扰能力，能够有效地增加窗口期的数据传输量。最后的实验结果表明，在给定的时间窗口约束下，我们的方法对遥感图像的处理性能相比原始方法提升接近4倍。

## 介绍

如今，卫星技术的快速发展加上部署成本的显著降低，使得卫星部署更加可行和经济。这一趋势推动了低地球轨道 (LEO) 卫星的大规模增长\cite{xie1}。与过去传统的高轨道卫星或地球静止轨道(GEO)卫星相比，低轨道卫星由于距离地球较近而具有竞争力。低轨道卫星具有较低的传输路径损耗，这点优势使它们能够比地球同步轨道卫星更快地传输卫星图像和传感器数据。这一点为卫星智能技术的发展提供了新的机遇，可以更好地应对各种全球挑战，包括但不限于碳排放监测\cite{buchwitz1}、灾害监测\cite{Lhermitte1,Filipponi1}、气候变化\cite{Yang1}和对各种天气事件的快速响应\cite{risdiyanto2001weather}。它甚至可以促进军事场景中哪些需要快速响应和低延迟的应用程序的开发。

但由于地球观测任务所需的数据量很大，低轨道卫星也有需要解决的挑战。首先，有限的电池容量、计算资源和通信资源制约了在轨数据处理的功能和性能。因此，我们不能简单地将现有的地面智能部署工作简单复制到空间层中。其次，与传统的高轨道或地球同步轨道卫星不同，低轨道卫星每天只能与地面站建立有限次数的通信会话\cite{Denby1}。特别是，两个小时内只能建立15分钟的连接 \cite{ge2022leo}。此外，地面站的分布不仅受地理区域的限制并且成本高，进一步限制了下行带宽\cite{Vasisht1}。最后，卫星无线通信经常受到大气吸收\cite{giggenbach2022atmospheric}、电离层、降雨\cite{yeo2014rain}等因素的影响，造成信号的损失和衰减。这些限制给现有的卫星遥感图像推断任务带来了持续时间短、信号衰减和传输延迟等方面的挑战。

目前最先进的星地通信技术已经接近香农物理容量极限。语义通信被认为是香农范式的突破，其目的在于成功地传递源所传递的语义信息，而不是准确地接收每一个单独的符号或比特\cite{luo2022semantic}。它是一种新兴的通信范式，利用深度学习模型对数据进行编码以提高通信效率。最近在语义通信领域的一些工作已经展示了在弱带宽和不利信道条件下的的潜力\cite{Gunduz1,Jiang2}。例如，Bourtsoulatze等人\cite{Bourtsoulatze1}提出了一种基于DeepLearning的无线图像联合信源信道编码(DeepJSCC)技术，该方案在低信噪比(SNR)和信道带宽下，性能优于级联JPEG或JPEG2000压缩的数字传输。近期的进展还有将语义通信应用于非陆地网络(NTN)的特定场景，比如无线无人机控制\cite{Xu1}和遥感频谱感知\cite{Yi1}。然而，大多数语义通信方法要求发送端和接收端共享一部分数据以控制压缩率并保证数据的有效性。众所周知，频繁的卫星地面站通信是不合理的。除此以外，卫星与地面站之间的数据通信属于长距离传输。传统的遥感图像语义压缩方法会在物理通道上造成不可忽视的损失，影响数据的可用性。

为解决上述挑战，本文提出一种语义通信框架，利用可视图神经网络(ViG) \cite{Han1}作为语义提取器，即SCViG。该架构不仅可以加速遥感图像推理过程，而且可以在较短的时间窗口内传输尽可能多的数据。可以通过修改初始视觉GNN模型，调整参数大小和层数，使计算更轻量。不仅如此，这些操作还帮助模型在处理低分辨率图像时避免了小图像块的崩溃。并且，我们创新性地在最终的Grapher层之前引入了残差卷积块(RCB)。残差结构有助于增强信源信道编码能力，提高模型训练过程的收敛速度，还可以作为抵抗模块来增强抗噪能力。具体来说，系统将编码部分设置在卫星上，解码部分设置在地面站上，实现了模型的解耦。这将允许地面站在接收到数据后立即进行推断。为了验证本文所提出的语义通信方法在对地观测场景中的有效性，我们将该框架应用于大规模多标签的土地覆盖基准数据集BigEarthNet \cite{Sumbul1,Sumbul2}以及RESISC45 \cite{Cheng1}和Patternet \cite{Zhou1}。此外，为了进一步研究星地信道天气因素的影响，考虑了雨衰落在慢瑞利衰落信道中的影响，将其作为影响因素加入到测试中，检验所提出的通信框架性能。为了证明图像预处理能够有效地增加可传输的数据量，我们在文章的最后使用线性优化策略来计算，本文的方法在窗口周期内到底可以增加多少数据量。

总之，主要贡献总结如下:

本文提出了一种增强卫星地面链路数据传输量的语义通信框架ViGSC，打破了传统的传输瓶颈，在推进对地观测任务方面具有巨大潜力。据我们所知，这是第一个为卫星和地面站之间周期性通信断开的场景下的遥感图像传输设计的框架。

与现有的语义通信方法需要卫星与地面站之间频繁通信以进行知识共享或oracle-net相比，该架构降低了星地链路上的通信压力。

我们还设计了一个编码器框架来适应遥感图像数据集。在实验中，基于真实数据集比较了不同通道的推理精度，表明所提出方法优于基线。为了更真实地模拟通信，我们还测试了该框架在降雨衰减效果下的性能。最后，在ViGSC上提出了一种多任务优化策略，实验结果表明，在给定的时间窗口约束下，ViGSC可以处理4倍于原始数据量的数据。

## 相关工作

在本节中，我们探索了与卫星语义通信和图神经网络相关的现有文献和研究。语义通信利用了深度学习模型强大的数据压缩和噪声恢复能力。与传统的数字通信模型不同，基于深度学习的通信模型不会受到“悬崖效应”的影响\cite{Kokalj1,Bourtsoulatze1}。通过共享知识库\cite{Gunduz1,Lu1}、有效的语义提取和重构能力\cite{Du1,Jiang2}，语义通信已经展示了其在带宽减少和从不利信道条件中恢复方面的潜力。最近的进展已将语义通信应用于解决非陆地网络(NTN)的特定挑战，包括卫星卸载\cite{Zheng1}、无线无人机(UAV)控制\cite{Xu1}和遥感频谱感知\cite{Yi1}。星地领域利用语义通信的优势，这个领域在近年来也有比较大的进展。

其中，Zheng等人提出了一种新的增强卫星-边缘云通信的框架SemCom-SEC \cite{Zheng1}。他们采用自适应剪枝-分裂联合学习(PSFed)方法在SemCom-SEC框架内更新语义编码器。针对个性化服务，提出了一个值得关注的框架，即基于频率调制的语义卫星通信(Frequency Modulation-based Semantic Satellite Communication, FMSAT) \cite{Jiang1}。该框架利用基于fm的分割和重建技术，可以显著降低带宽需求，同时在高噪声和干扰等挑战性条件下准确地恢复语义特征。与卫星场景类似的无人机通信也面临着数据传输压力，以及单个设备有限的计算资源和能量的挑战。Song等人提出了一种面向无人机目标检测的认知语义通信系统\cite{Song2}。他们的方法利用知识图谱，并结合多尺度压缩网络进行语义压缩，确保检测精度和减少数据传输。

然而，现有的方法往往需要在接收端和发送端同时构建先验模型\cite{Zhang1}，这在星地通信模型中是不切实际的。这种方法可能导致传输图像的精度损失或丢失。这种模式虽然减少了每次传输的数据量，但也因为传输了无用的数据而浪费了通信资源。与之相反，语义交流强调传达信息的意义，只传输必要的信息，在提高效率的同时显著减少了数据通信量。

此外，视觉图神经网络结构\cite{Han1}在从图像块中提取特征时，可以有效地保留图像的语义信息。图结构能够在图像内部语义相关的部分之间建立联系。随着更多数据的可用性和大规模基准数据集\cite{Sumbul2,Shen1}的存在，相关领域的专家已经开始将计算机视觉模型适应航空图像，这些模型显示出了优越的性能。应用于遥感领域的基于transformer的模型在数据丰富的场景中表现出了比CNN更好的性能\cite{Dosovitskiy1}。本文工作中，图像推理任务中采用的特征金字塔结构可以被视为一种压缩形式，其中每一层特征提取都减少了数据量。这在卫星通信场景中尤其重要，因为在星地通信场景中迫切需要在短时间内将图像数据传输到地面。

鉴于这些考虑，本文提出了一种利用GNN（图神经网络）作为语义提取器的语义通信框架。该方法不再需要卫星和地面站之间的频繁通信，并有效地压缩了数据的传输量，使得图像传输任务可以在较短的时间窗口内简单完成，同时也能保持较高的推理精度。我们的方法在复杂的信道环境下（即使受到降雨衰减的影响），也能取得较好的性能。

## 方法

在本节中，我们首先介绍所提出的框架，然后构建整个系统模型，最后对每个子模块进行详细介绍。

该框架如图\ref{SemComArt}所示。首先，卫星对遥感图像进行切片分块。之后，将这些分块被投影到高维嵌入空间，这个过程称为位置编码。随后，ViG语义编码器对这些分块进行特征提取和更新。该模块的描述见3.3节。为了提升数据通过物理通道后的可用性，我们将残差卷积块应用在通道前后的压缩和恢复。我们称之为RCB和RTCB，将在3.4节介绍。最后，由3.5节中介绍的ViG语义解码器模块对卫星传输的数据进行推断，完成对地观测任务。

我们选择图神经网络结构有几个原因。与网格或序列相比，图更灵活，更适合建模复杂对象。在卫星遥感图像数据中，内容往往由山脉、湖泊、海洋和森林等自然元素的混合物组成。这些图像中的物体通常形状不规则，而且不是网格状。其次，遥感图像也用于观测人造物体，如船舶和飞机。这些独立的物体可以看作是部件的组成(例如，一架飞机可以大致分为机身、机翼和尾翼三个部分)，利用图形结构可以更好地描述这些部件之间的联系。通过部署图神经网络，所提出的语义通信框架旨在从卫星图像中提取有意义的信息，并在卫星语义通信场景的背景下实现高效通信。

\subsection{模型}

\， \， \;\textbf{信道模型:如图}\ref{SemComArt}所示，我们的框架由一个LEO卫星和一个地面站组成。链路的信道质量由噪声功率$\sigma^2$决定，其主要是由物理信道中的自然噪声和干扰引起的。我们假设物理信道噪声为AWGN和慢瑞利衰落信道\cite{torrance1999interference}，定义信噪比如下

\begin{equation}

\gamma = 10 \log\_{10}{\frac{P\_z}{\sigma^2}}.

\end{equation}

其中$P\_z$表示源的平均功率。

\textbf{遥感任务模型:}我们在两种类型的推理任务上评估了所提出的框架。

\begin{itemize}

\item 在“多分类”任务中:对于给定的输入图像$\boldsymbol X$和一组$\mathcal N$表示潜在类别的可能标签，预测模型$f\_{pre}$在标记的图像上进行训练，以产生一个验证向量，由$N$概率值$v\_j$组成，范围从0到1。每个值$v\_j$表示输入图像$\boldsymbol X$只属于$j-th$类的概率。向量中的概率之和为1。因此，根据模型$f\_{pre}$的预测，输入图像$\boldsymbol X$与具有最高概率的类相关联。

\item 在“多标签分类”任务中:

基本定义与“多分类”任务相同，区别在于概率向量。特别地，向量中的概率之和不是1。每个卫星采集图像$\boldsymbol X$与预测概率$v\_j$\textgreater\space$t$的所有类相关联。在我们的例子中，$t$是一个预定义的阈值，设置为0.5。

\end{itemize}

\subsection{ViG语义编码器}

ViG编码器和Grapher模块共同组成ViG语义编码器，如图\ref{vigencoder}所示。

\textbf{ViG编码器: }

视觉图神经网络框架引入的主要创新点是将输入图像分割为多个图像块，并将其投影到高维嵌入空间。这将图像的规则网格结构与现有框架(如cnn和基于transformer的架构)解耦。对于大小为$H (height) \times W (weight) \times Channel$的图像，它分为$N$图像块。通过将每个patch转换为特征向量$x\_i$$\in$${R}^D$，我们得到$\boldsymbol X$ = [$x\_1$, $x\_2$，…]， $x\_N$]，其中$D$是特征维度，$i$ = 1, 2，…， $N$表示图像补丁的索引。

计算图作为图神经网络中用于消息传递的邻接矩阵，是根据嵌入之间的相似度动态计算的，可以配置所需的邻居数量。每个图像块被视为有向图中的一个节点。根据它们的嵌入，在patch本身和它的$K$最相似的邻近patch之间建立有向边。

为了表示节点的位置信息，要为每个节点特征添加一个位置编码向量:

\begin{equation}

x\_i + e\_i \xrightarrow{} x\_i,

\end{equation}

其中$e\_i$是一个d维向量。如公式2所述的绝对位置编码应用于金字塔结构。一旦构建了图，信息就通过神经消息传递机制进行处理，由Grapher层实现，该层由两个具有ReLU非线性激活函数的多层感知器(MLP)层和最大相对图卷积\cite{Li1}组成。Grapher层的描述如下:

我们从特征开始$\boldsymbol X$$\in$${R}^D$。我们首先基于这些特征构建一个图:${G} = {G}(\boldsymbol X)$。图卷积层通过聚合相邻节点的特征来交换节点之间的信息。具体来说，图卷积操作表示为:

\begin{equation}

G = F(G, W) = W\_{update}(G,W\_{agg}),

\end{equation}

其中$W\_{agg}$和$W\_{update}$分别是聚合和更新操作的可学习权重。聚合操作通过聚合其邻近节点的特征来计算节点的表示。随后，更新操作进一步合并聚合的特征。我们采用最大相对图卷积，因为它简单高效:

\begin{equation}

{x\_i}' = W\_{update} [x\_i,\{max( x\_j - x\_i,\{j\in N(x\_i)\})\},W\_{agg} ],

\end{equation}

其中${N}(x\_i)$为节点$x\_i$的相邻节点集合。所以我们将这个操作定义为Grapher($\boldsymbol X$)，(对于$\boldsymbol X$中的所有$x\_i$)。

更详细地说，我们还在聚合过程中使用了类似原始实现的多头机制。首先将聚合特征${x\_i}'$拆分为$h$头，然后分别对这些头进行不同的权重更新。所有头可以并行更新，并连接起来作为最终值。

\begin{equation}

{x\_i}' = [h^1{W^1}\_{update}, h^2{W^2}\_{update},..,h^n{W^n}\_{update}].

\end{equation}

为了缓解深度GCNs中的过平滑现象，我们在ViG模块中引入了额外的特征变换和非线性激活。在图卷积之后插入非线性激活函数，以避免层坍缩。

在实践中具体为，对于一个给定的输入特征$\boldsymbol X$$\in$${R}^D$, Grapher模块可以表示为:

\begin{equation}

\boldsymbol Y = FW\_b (\sigma(Grapher(FW\_{nb}(\boldsymbol X)))).

\end{equation}

为了进一步增强特征变换能力并缓解过平滑现象，在每个节点上使用前馈网络(FFN)。

\begin{equation}

\boldsymbol Z = FW\_b(\sigma(FW\_{nb}(Y))).

\end{equation}

$\sigma$表示ReLU非线性激活函数。$FW\_{nb}$和$FW\_{b}$分别代表两个带权重的全连接层，没有和有偏差项。

$\boldsymbol {Y}$和$\boldsymbol {Z}$表示由Grapher在前馈神经网络(FFN)模块的预投影和后投影状态下学习到的块嵌入，其中$\boldsymbol X$是输入图像块的嵌入。

\textbf{ViG Grapher模块:}

ViG编码器由一系列的Grapher模块和卷积层组成。每个卷积层将输入图像的分辨率降低一半，导致每个Grapher模块和convolutional模块之后的图像块数量减少4倍。ViG基于输入图像和块嵌入动态构建图。每个图像块被视为图中的一个节点，由其块嵌入描述。在每个Grapher层中，ViG通过计算每个patch与所有其他patch之间的距离，自动构建一个不同的图。给定预定义的邻居数量$K$，每个补丁以定向的方式连接到其$K$最近邻补丁。

值得注意的是，对于给定的一对输入图像$\boldsymbol X\_i$和$\boldsymbol X\_j$，以及相同的补丁嵌入${\boldsymbol X\_i}$和${\boldsymbol X\_j}$，与补丁${\boldsymbol X\_i}$相邻的补丁的邻居可能位于不同的位置，与补丁${\boldsymbol X\_j}$的邻居相比。块的位置信息不直接参与边缘的创建，而是通过块嵌入间接考虑。金字塔架构逐渐缩小了特征图的空间大小，被广泛应用于图像处理任务中，以利用尺度不变性并生成多尺度特征(如ResNet \cite{He1}和PVT \cite{Wang1})。这种架构提高了性能和准确性。在本文实现的语义通信模型中，ViG编码模块也采用了金字塔结构，称为ViG金字塔，如图\ref{vigencoder}所示。一方面，金字塔结构提高了性能和准确性;另一方面，分层压缩过程中数据量减少，有利于卫星地面站的语义通信框架。

我们的工作修订了PyramidViG的架构。在特征提取的第一步，将图像块的高度和宽度缩小到原始大小的1 / 4; 在每个Grapher层之后，中间表示的高度和宽度减半，每一步图像块的数量减少4倍。它可能会在最后一层造成一个非常小的分块。因此，为了避免这种情况，我们修改了ViG的编码器结构，使其具有与原始实现相似的可学习参数数量，防止模型在图像块方面崩溃。具体来说，与原始的PyramidViG实现不同，我们将Grapher阶段的数量从4个减少到3个，使用维度[64,128,256]的嵌入空间而不是[48,96,240,384](vigi - tiny)。我们的工作使用的数据集之一是bigearthnet，输入图像分辨率较低(120 $\times$ 120像素)。为了防止模型坍缩为单个图像块，在这个数据集的第一步中的宽度减少不是四倍，而是减半的图像块数量，类似于在Grapher层之后的处理。由于数据集的差异，实现聚合更新头的数量为16，以保持图片特征的多样性。在为每个Grapher层创建图形时，我们还设置邻居数$K$ = 9。

语义信道压缩过程的架构如图\ref{RCB&RTCB}所示，由两个模块组成:残差卷积块(RCB)和残差转置卷积块(RTCB)。残差结构有助于增强信源信道联合编码能力。残差网络由卷积层和带有PReLU激活函数的转置卷积层组成。在卷积或转置卷积层之后应用的归一化操作是基于广义分裂归一化(GDN)的，它已经被证明在图像压缩和密度建模任务\cite{Balle1}中是有效的。卷积或转置卷积层由参数$m \times n | s$指定，其中$m$和$n$对应于使用的卷积核的宽度和高度。然后在RCB模块中进行下采样操作，在RTCB模中进行上采样操作。参数$s$表示步长。在这项工作中，使用了5 × 5卷积核和2的采样步幅。该处理在有效压缩卫星传输数据量的同时，保证地面站能够在一定程度上从接收到的数据重建图像的结构特征。然后再利用像素上采样模块对特征信息进行上采样。最后，将与图像结构相同维度的特征张量输入预训练推理模块进行图像推理。

\subsection{ViG解码器模块}

在ViG解码器一侧，使用分类头执行图像分类推理任务，该分类头由池化层和MLP分类器组成。池化层将图像特征聚合为固定长度的向量表示。然后，MLP分类器将这个向量作为输入并执行分类预测。架构如图\ref{VIGDECODER}所示。

综上所述，ViG编码器通过一系列的Grapher模块和卷积层从输入图像中提取特征，同时利用金字塔结构逐步减小特征图的空间大小。它动态地构建一个图来捕获不同图像块之间的关系。最后，由接收端来完成图像分类任务。

\section{性能评估}

本节介绍了实验设置、基线数据集、选择的超参数，以及后续对ViG框架与其竞争对手相比的性能的评估。

\subsection{数据集和基线模型}

\textbf{数据集:}我们使用了三个数据集进行实验，分别是BigEarthNet \cite{Sumbul1,Sumbul2}, RESISC45 \cite{Cheng1}和PatternNet \cite{Zhou1}。这些数据集都是遥感图像数据集。 RESISC45图像的分辨率为每像素0.2到30米，而PatternNet图像的分辨率为每像素6到50厘米。 RESISC45图像收集自100多个国家，在天气、地形形态和照明方面能够明显表现出不同。RESISC45和PatternNet都是多分类数据集。图像的分辨率为256x256像素，由3个通道的RGB采集组成。 这两个RGB数据集完全平衡。RESISC45包含31,500张可用图像，分为45个不同的类，而PatternNet包含30,400张可用图像，有38个不同的类。 对于每个类，RESISC45有700个样本，PatternNet有800个样本。我们使用的BigEarthNet数据集是一个大规模的多标签图像场景分类数据集。它由590,326个Sentinel-2 L2A和Sentinel-1收购组成，共有14个通道。BigEarthNet的最高分辨率为120x120像素。使用双线性插值对较低分辨率的采集进行上采样，使得数据集中所有输入图像的分辨率为120x120像素。我们从BigEarthNet数据集中提供的两组标签中选择了一个更具挑战性的任务，该数据集由43个不同的类组成。我们将这三个数据集分为训练集、验证集和测试集。对于BigEarthNet数据集，遵循原作者的划分方式。对于RESISC45，我们使用60/20/20的随机非重叠划分。对于PatternNet，我们采用70/15/15的随机非重叠划分。我们的实验展示了这三个数据集的微平均F1分数、召回率以及精度。

\textbf{基线模型:}

基线模型方面我们的选择是ResNet18和ViT-Tiny。

\begin{itemize}

\item ResNet \cite{he2016deep}:它是一种传统的基于CNN的用于计算机视觉任务的网络，已经是一种成熟且广泛使用的方法。

\item ViT \cite{dosovitskiy2020image}:与传统的基于CNN的网络结构不同，它是基于transformer架构的计算机视觉网络。

\end{itemize}

由于ViT还会进行图像分割，然后嵌入和编码以进行特征提取，考虑到这两个模型的参数数量与我们的模型比较接近，我们选择它们作为基线模型。

我们的ViG模型是基于原始论文\cite{Han1}中的vigtiny的规模进行修改的。一方面，这种选择考虑了卫星有限的计算能力。另一方面，它解决了能源消耗问题，使我们选择了一个更小的模型。

为了在所有三种方法中实现相同级别的图像压缩，并确保每种方法的信道负载相同，我们将自动编码器应用于ResNet18和ViT-Tiny，如图\ref{autoen}所示。

我们为三个模型和数据集选择了不同的可训练参数。在BigEarthNet数据集中，ViT-Tiny的可训练参数为5.54M, ViG和ResNet18的可训练参数分别为6.98M和11.22M。在另外两个数据集上，由于它们的图像通道比BigEarthNet小，因此相同的模型在这两个数据集上具有相同的参数数量。我们将ViT-Tiny的参数设置为5.58M, ViG设置为8.6M, ResNet18设置为11M。由于RESISC45和PatternNet数据集的较高输入分辨率，该方法中可学习位置编码向量的参数数量显著增加。此外，BigEarthNet输入信道数量的差异也导致了这种增加。

对于所有数据集和模型的训练设计，我们使用AdamW优化器，学习率参数设置为$10^{-4}$。权重衰减范围从$10^{-2}$到$10^{-4}$，最多执行100个epoch。我们使用交叉熵损失函数优化了模型。训练超参数的摘要如表\ref{hyperparameters}所示。

此外，我们采用了一种基于监测验证损失的早期停止机制。如果连续10个epoch的验证损失没有改善，则学习率降低10倍，并设置5个epoch的等待期。

在本节中，我们将在两个用于多分类的数据集和一个用于多标签分类的数据集上比较上述模型。用F1分数、召回率和精度对每个实验进行测试，聚焦在模型的推理性能。对于中等大小的数据集，本文考虑的多分类问题相对简单。在PatterNet数据集上获得的结果表明，所考虑的图像分类任务相对于RESISC45更容易。所有模型的分数均达到98.0以上。RESISC45在解决分类任务时表现出更高的复杂性，从实验结果来看，ViG模型的表现始终优于其他模型，取得了最高的平均F1分数87.34，分别超过ViT和ResNet约12.3和5.18个百分点。ViG也取得了86.19％的最高准确率。对于多标签分类任务，在BigEarthNet数据集上的实验证实，ViG架构在RGB以外的图像方面也表现良好。考虑到准确性指标，ViG表现出更好的性能，其次是ViT(提高了3.09个百分点)，而ResNet表现最差。这可以归因于图像中块的数量显著减少，导致与位置编码张量相关的参数数量减少。ViG和ViT都对图像采用基于块的编码。表\ref{basic performance}展示了宏平均F1、精确率和召回率分数。总的来说，我们的方法在基于图的学习在中型数据集上表现出良好的性能。

\subsection{信道干扰性能}

本文提出的方法基于语义通信框架，在传输过程中噪声干扰的影响下，对比了三种模型的性能。在所有实验结果中加入高斯白噪声，利用信噪比控制噪声强度。从性能评估的角度，我们选择平均准确率(Average Precision, AP)指标来评估框架的推理性能。图\ref{fig:2a}展示了在AWGN信道中随着信噪比的增加，PatternNet数据集上各种模型的平均精度性能。在这个实验中，由于PatternNet任务的简单性，所有模型都比前两组表现更好。当信噪比小于20时，ViG具有显著的优势。

图\ref{fig:2b}说明了在AWGN信道中随着信噪比的增加，各种模型在RESISC45数据集上的平均精度性能。同样，在SNR = 0时，模型性能并不理想，但由于图像比BigearthNet数据集更大，可以保留更多的特征，从而提高AP。可以观察到，无论是在之前的无干扰测试中，还是在噪声测试中，ViT在该数据集上的表现都弱于ResNet18。这可能是因为基于transformer的架构需要更大的数据量来学习，但RESISC45数据集的样本最少。

图\ref{fig:2c}显示了在AWGN信道中随着信噪比的增加，各种模型在BigearthNet数据集上的平均精度性能。可以观察到，当SNR为0时，由于数据集图像相对较小(120 $\times$ 120)，所有模型的AP都低于50％，导致提取特征后的信息较少。而当信噪比增加到20时，ViG(我们的)的AP可以达到70％，并且随着信噪比的增加，AP还在继续增加。这表明该方法在保留图像语义信息方面表现良好。

综合上述3个实验，ViG虽然没有展现出压倒性的优势，但仍然表现最好。所提框架可以容忍相同AP下较低的信噪比，这在信道容量受限的场景中尤为重要。总体而言，该框架能够为卫星遥感图像推理任务节省计算资源和通信资源。

目前有许多卫星使用ku波段\cite{Han2}与地面站进行无线通信，这是指12至18 GHz之间的微波频率范围。在ku波段(10GHz以上)卫星通信中，地表降雨引起的雨衰减\cite{ojo1}严重影响信号的传输质量。

我们的工作采用了ITU-R P.838建议和相关研究成果，并加入了瑞利衰落信道，建立了计算降雨衰减的数学模型。利用该模型在真实的卫星通信场景中模拟本文所提出的语义通信，评估它是否能保持预想的功能。降雨衰减的具体模型如下:

\begin{gather}

\gamma\_R(dB/km)=kR^\alpha, \\

k=[k\_{H}+k\_{V}+(k\_{H}-k\_{V})\cos^{2}\theta\cos2 \tau]/2, \\

a=[k\_{H}a\_{H}+k\_{V}a\_{V}+(k\_{H}a\_{H}-k\_{V}a\_{V})\mathrm{cos}^{2}\theta\cos2 \tau]/2k.

\end{gather}

方程8展示了如何描述降雨衰减效应，参数$k$和$a$紧随方程9和10。$k\_H$和$a\_H$是水平极化因子。相反，$k\_V$和$a\_V$是垂直极化因子。在固定的通信频率下，它们都是常数。具体数值可在ITU-R P.838建议书中找到。本文采用圆极化方案，卫星与地面站之间的通信角度选择为45°，$\tau$。

我们的工作主要考虑了Ku频段，所以我们在12 GHz和15 GHz微波频段进行了测试，瑞利信道信噪比(SNR)值固定为20dB和30dB。我们使用BigEarthnet数据集评估了它们的性能，该数据集具有更大的通道数量和更小的分辨率。我们还选择了PatternNet数据集，它代表了最简单级别的任务难度。图\ref{fig:3}的实验结果表明，较高的信号频率表现出更强的降雨衰减效应，当降雨量超过50 mm/hr时，信号损失变得有些不可接受。然而，该模型在不同降雨强度下仍优于其他模型。此外，无论在15 GHz还是12 GHz的信号频率下，ViG语义通信都呈现出最慢的下降趋势，即使在小雨条件下也能保持良好的推理精度。

\subsection{批处理图像处理性能}

最后，我们还考虑到低轨卫星在绕地球轨道运行期间捕获了大量图像的情况。理想状况是能够处理所有捕获的图像，但需要考虑卫星上有限的能量和计算资源无法达到理想状态。因此，我们对所提方法进行了批量图像处理的性能评估。给定窗口时间阈值$T\_{window}$，意味着卫星可以与地面站通信的时间。假设卫星的计算能力为$P\_{comp}$，其传输能力为$P\_{trans}$，图像数据的总大小为$M$，卫星对图像数据的处理能力为$N$，经过我们的语义通信处理后转换为$AN$，满足如下等式。我们设置卫星计算能力为1000 GFLOPS \cite{cao1}，发射功率为200Mbps \cite{liu1}，窗口时间为15分钟\cite{periyapatna1}。在信噪比= 20的瑞利衰落信道下，不同降雨强度下进行了计算仿真，结果如表\ref{trans\_under\_rain}所示。可以观察到，下行链路在给定时间15分钟内可以传输22 GB的数据量。但经过我们的方法处理后，有效地传输了等同于原来90.6 GB的数据量。这对于减轻卫星上的工作量，加快地面观测任务的完成具有重要意义。即使雨衰效应导致本可以传输的数据量减少，也可以通过语义通信实现更多的数据传输。

\section{结论}

文中利用视觉GNN模型在处理卫星图像方面的优势，提出了一种基于视觉图神经网络的星地链路的语义通信框架。将该框架应用于遥感图像分类任务，包括多分类和多标签分类问题。即使在AWGN信道和不同干扰强度的瑞利衰落信道中，所提方法也能保持最优性能。不仅如此，针对星地通信链路的情况，模拟了降雨衰减对框架稳定性的影响。实验结果表明，该方法能够容忍较高的信道噪声和天气引起的无线通信干扰，同时保持良好的推理精度。