

武汉大学

本科毕业论文（设计）

跨分辨率遥感影像
精准信息提取方法研究

姓 名： 张 铄 琦
学 号： 2020300004045
专 业： 计算机科学与技术
学 院： 计 算 机 学 院
指导教师： 肖 晶

二〇二四年三月

原创性声明

本人郑重声明：所呈交的论文（设计），是本人在指导教师的指导下，严格按照学校和学院有关规定完成的。除文中已经标明引用的内容外，本论文（设计）不包含任何其他个人或集体已发表及撰写的研究成果。对本论文（设计）做出贡献的个人和集体，均已在文中以明确方式标明。本人承诺在论文（设计）工作过程中没有伪造数据等行为。若在本论文（设计）中有侵犯任何方面知识产权的行为，由本人承担相应的法律责任。

作者签名：

指导教师签名：

日 期： 年 月 日

版权使用授权书

本人完全了解武汉大学有权保留并向有关部门或机构送交本论文（设计）的复印件和电子版，允许本论文（设计）被查阅和借阅。本人授权武汉大学将本论文的全部或部分内容编入有关数据进行检索和传播，可以采用影印、缩印或扫描等复制手段保存和汇编本论文（设计）。

作者签名：

指导教师签名：

日 期： 年 月 日

摘 要

论文的摘要是对论文研究内容和成果的高度概括。摘要应对论文所研究的问题及其研究目的进行描述，对研究方法和过程进行简单介绍，对研究成果和所得结论进行概括。摘要应具有独立性和自明性，其内容应包含与论文全文同等量的主要信息。使读者即使不阅读全文，通过摘要就能了解论文的总体内容和主要成果。

论文摘要的书写应力求精确、简明。切忌写成对论文书写内容进行提要的形式，尤其要避免“第 1 章……；第 2 章……；……”这种或类似的陈述方式。

关键词是为了文献标引工作、用以表示全文主要内容信息的单词或术语。关键词不超过 5 个，每个关键词中间用分号分隔。

关键词：关键词 1；关键词 2；关键词 3；关键词 4；一个非常非常，非常非常长——的关键词 5

ABSTRACT

An abstract of a dissertation is a summary and extraction of research work and contributions. Included in an abstract should be description of research topic and research objective, brief introduction to methodology and research process, and summary of conclusion and contributions of the research. An abstract should be characterized by independence and clarity and carry identical information with the dissertation. It should be such that the general idea and major contributions of the dissertation are conveyed without reading the dissertation.

An abstract should be concise and to the point. It is a misunderstanding to make an abstract an outline of the dissertation and words “the first chapter”, “the second chapter” and the like should be avoided in the abstract.

Keywords are terms used in a dissertation for indexing, reflecting core information of the dissertation. An abstract may contain a maximum of 5 keywords, with semi-colons used in between to separate one another.

Key words: key word 1; key word 2; key word 3; key word 4; and a very very, very very long key word—the key word 5

目 录

1	绪论	1
1.1	研究背景与意义	1
1.1.1	高分辨率和低分辨率遥感图像简介	1
1.1.1.1	高分辨率图像	1
1.1.1.2	低分辨率图像	2
1.1.1.3	重访周期与成本	2
1.1.2	研究意义	3
1.2	国内外研究现状	3
1.2.1	跨分辨率遥感影像目标计数现状	3
1.2.1.1	目标检测	4
1.2.1.2	语义分割	4
1.2.1.3	基于回归的计数方法	4
1.2.1.4	基于密度图的计数方法	4
1.2.1.5	跨分辨率车辆计数数据集	4
1.3	研究目标与内容	4
1.3.1	研究目标	4
1.3.2	研究内容	5
1.4	章节安排	5
2	目标计数方法技术框架	6
2.1	目标计数	6
2.2	语义分割模型	6
2.2.1	卷积层	6
2.2.1.1	卷积模型	7
2.2.1.2	多通道输入与多通道输出	7
2.2.1.3	填充和步幅	7
2.2.2	激活函数	8
2.2.3	池化层	8
2.2.4	权重衰减	9

2.2.5	暂退法	9
2.2.6	批量归一化	9
2.2.7	损失函数	10
2.2.8	梯度下降	10
2.3	U-Net 网络	10
2.4	注意力机制	10
2.5	Attention UNet 网络	10
2.6	光流估计	10
2.7	跨分辨率遥感图像目标计数技术路线	10
2.8	插图的使用	11
2.9	表格的使用	12
2.9.1	普通表格	12
2.9.2	跨页表格	13
2.10	列表的使用	13
2.10.1	有序列表	13
2.10.2	不计数列表	14
2.11	数学环境的使用	14
2.12	单位	15
2.13	物理符号	15
3	跨分辨率车辆计数数据集	16
3.1	脚注	16
3.2	脚注	16
3.3	脚注	16
3.4	脚注	16
3.5	脚注	16
3.6	脚注	16
3.7	引用文中小节	16
3.8	引用参考文献	16
3.9	链接相关	17
4	基于注意力机制的跨分辨率遥感影像计数	18

4.1	18
4.2	代码	18
4.3	代码	18
4.4	代码	18
4.4.1	原始代码	18
4.4.2	算法描述/伪代码	18
4.5	绘图	19
4.6	写在最后	19
5	跨分辨率遥感影像计数实验及分析	20
5.1	AB 测试	20
5.2	代码	20
5.3	代码	20
5.4	代码	20
6	总结与展望	21
6.1	结论	21
6.2	代码	21
6.3	代码	21
6.4	代码	21
	参考文献	22
	致谢	25
	附录 A 数据	26
A.1	第一个测试	26

1 绪论

随着遥感技术的快速发展，我们能获得越来越丰富的遥感观测数据。既有高分辨率的清晰图像，也包括低分辨率的鸟瞰图。这些图像从更宏观的视角为我们提供了许多有价值的数据。然而如何合理的使用这些图像数据，特别是从低分辨率图像中找出更多难以通过人类肉眼直接识别出的有效信息成为了一个重要的研究课题。在车辆计数等真实场景中，高分辨率图像虽然能提供丰富的细节，但有着获取成本高昂以及难以获取稳定连续的高频率数据的问题。本文在 CRVC 跨分辨率车辆计数数据集上，设计了一种基于注意力机制的深度学习 U 型网络模型，旨在提高跨分辨率遥感图像中隐含的空间和时间信息进行更丰富全面的表示，以进一步提升目标计数的准确性。

1.1 研究背景与意义

1.1.1 高分辨率和低分辨率遥感图像简介

在遥感领域，分辨率是用来描述遥感图像细节程度的一个重要指标，根据遥感卫星搭载的传感器不同，可以从多个维度进行描述，包括空间分辨率、时间分辨率、光谱分辨率和辐射分辨率。在讨论高分辨率（High Resolution, HR）图像和低分辨率（Low Resolution, LR）图像时，通常指的是空间分辨率的差异。空间分辨率的详细定义可以参照^[1]：(i) 低或粗分辨率定义为地面采样距离（GSD）为 30 m 或更大的像素，(ii) 中分辨率是 GSD 范围为 2.0–30 m，(iii) 高分辨率为 GSD 0.5–2.0 m，以及 (iv) 极高分辨率为像素大小 <0.5 m

本文中将该定义简化，只进行高分辨率与低分辨图像的区分。

1.1.1.1 高分辨率图像

高分辨率图像通常指具有较高空间分辨率的图像，即图像中单个像素所代表的地面面积较小，能够显示更加精细的地面特征。高分辨率图像使得用户可以观察到较小的地面对象，例如单个车辆、道路标线甚至是行人。虽然“高分辨率”这个术语没有绝对的定义，但在遥感领域，我们把空间分辨率小于 1 米（通常在 0.3 米到 1 米之间）的图像常被认为是高分辨率图像。目前我们可以使用的高分辨率遥感图像来源主要有：航空摄影（搭载高分辨率摄像机或低空高分辨率无人机拍摄的

数据)和某些高性能的卫星遥感仪器,例如 WorldView 系列、GeoEye-1、QuickBird 等。高分辨率图像中的精细地面特征信息,在城市规划、交通监控、农业监测(如作物健康分析)、详细的地物分类、灾害评估等领域都有广泛的使用。特别是在目标识别领域中,高分辨率的图像可以使用深度学习中多种模型和方法,具有很高的应用价值。

1.1.1.2 低分辨率图像

低分辨率图像指的是空间分辨率较低的图像,即图像中单个像素所代表的地面面积较大,只能显示较为粗糙的地面特征。低分辨率图像难以分辨较小的地面对象,但适合于观察大范围的地表变化。通常空间分辨率大于 10 米(如 10 米、30 米或更大)的图像被认为是低分辨率图像。低分辨率图像主要来源于具有宽幅覆盖能力的卫星遥感仪器,如 MODIS(具有数百米的空间分辨率)、Landsat 系列(15 米到 30 米分辨率)、Sentinel-2(10 米到 60 米分辨率)等。图像中的大范围地表特征,在气候变化研究、大范围土地覆盖变化监测、环境监测、城市发展规划、海洋和大气研究等领域有着很高的应用价值。对于目标识别以及目标计数领域来说,模糊的图像质量大大加大了现有深度学习模型的识别难度。

1.1.1.3 重访周期与成本

地球卫星的轨道半径和周期可以根据开普勒第三定律和牛顿的万有引力定律推导出来的。下面是公式推导的过程:牛顿的万有引力定律描述了两个物体之间的引力,公式为:

$$F = G \frac{m_1 m_2}{r^2} \quad (1.1)$$

其中, F 是两个物体之间的引力, G 是万有引力常数, m_1 和 m_2 是两个物体的质量, r 是两个物体的中心之间的距离。对于在圆轨道上运动的卫星,向心加速度 a_c 提供了必要的向心力,可由下式给出:

$$F_c = m \frac{v^2}{r} \quad (1.2)$$

这里, F_c 是向心力, m 是卫星的质量, v 是卫星的轨道速度, r 是卫星绕行天体的轨道半径。由于卫星在轨道上的向心力正是由地球对卫星的万有引力提供的,可得到:

$$G \frac{Mm}{r^2} = m \frac{v^2}{r} \quad (1.3)$$

轨道周期 T 是卫星完成一圈轨道所需的时间，与轨道速度和轨道半径相关，可表示为：

$$T = \frac{2\pi r}{v} \quad (1.4)$$

在一般情况下，当考虑两个物体质量都不能忽略时， $M + m$ 代替上式中的 M ，最终得到两个物体之间的轨道周期 T 与轨道半径 r 之间的关系：

$$T = 2\pi \sqrt{\frac{r^3}{G(M + m)}} \quad (1.5)$$

这就是天体绕另一天体旋转周期的计算公式，它揭示了轨道周期与轨道半径之间的关系。这个公式在天文学和航天工程中有广泛应用，用于预测天体运动及设计人造卫星的轨道。高分辨率图像一般由低轨卫星拍摄，具有较小的轨道半径和周期。同时也因此具有较小的视场（the field of view, FOV）较小。高分辨率图像分辨率卫星通常使用任务驱动模式进行地球观测。这意味着，如果不提前提交观测任务，一颗卫星需要 6 个月才能完成全球覆盖，获得特定区域的图像的重访周期相当长。此外，高分辨率图像非常昂贵，例如 WorldView-3 的价格为 34 美元/km²。作为对比，低分辨率卫星的往往运行在更高的轨道上，虽然空间分辨率有所降低，但视场较大。因此获得同一地点重访周期要短得多，例如 PlanetScope 卫星每天重访一次，价格也低得多，为 1.8 美元/平方公里。

1.1.2 研究意义

单独依靠低分辨率图像进行目标计数是十分困难的，而仅通过高分辨率图像进行目标计数，不仅花费巨大，同时还需要面对连续监控数据的缺失。如何通过低成本且具有时间连续性的低分辨率图像进行目标计数及实时监测就成为解决问题的关键。本文提出的方法通过少量高分辨率图像的辅助，在时间连续的低分辨率图像上实现目标计数，具有很高的应用价值。该方法不仅局限于目标计数，更好地利用了低分辨率图像中蕴含的模糊信息，在稠密车流人流识别、智慧城市设计等领域也有很高的应用潜力。

1.2 国内外研究现状

1.2.1 跨分辨率遥感影像目标计数现状

高分辨率图像计数主要有两类方法。一类是检测方法，通过识别出具体的物体位置来进一步计数。另一种是计数类方法，直接估计图像中对应物体的数目。

1.2.1.1 目标检测

基于检测的目标计数是目标检测下的一个分支。目标检测作为计算机视觉的一个主要研究方向，主要研究的是如何识别出特定的对象，并确定它的大小和位置。目前随着深度学习技术的发展，利用神经网络进行目标检测已经成为主流方案。主要的目标检测算法包括 RCNN 系列、SSD、YOLO 系列等

1.2.1.2 语义分割

语义分割是一类特殊的目标检测任务，需要给出每个像素的分类。主要的语义分割模型包括全卷积网络（FCN）、U-Net、SegNet、DeepLab 系列、Mask R-CNN 和 Attention U-Net。

1.2.1.3 基于回归的计数方法

这类方法从深度神经网络提取的特征中直接进行数目的回归估计，常常用于处理目标密集、相互遮挡严重的场景，如人群计数。

1.2.1.4 基于密度图的计数方法

基于密度图的计数方法通过估计目标区域的目标密度，从而计算出数量估计。目前主要的密度图估计方法包括 CSRNet、MCNN 和 TEDNet。

1.2.1.5 跨分辨率车辆计数数据集

该数据集有少量高分辨率图像和拍摄于这些高分辨率图像日期之间的相同位置的低分辨率图像组成。该数据集主要关注低分辨率下的此车辆计数问题。在低分辨率下，汽车的边界极难分辨，人类在标注识别的过程中也并不容易。该数据集提供了少量拍摄了同位置的高分辨率图像作为参照。目前，已经有一些工作在这个数据集上取得了不错的效果。在本文的第三部分将进行进一步说明。

1.3 研究目标与内容

1.3.1 研究目标

1. 跨分辨率车辆计数算法的开发：开发一种新的基于深度学习的车辆计数算法，该算法能够有效利用有限的高分辨率图像来指导低分辨率图像中的车辆计数。

2. 探究高效利用数据集中空间一致性和时间连续性信息的方式：探究同一时刻高分辨率与低分辨率图像间的空间一致性和连续低分辨率图像间的时间连续性的高效利用方式。研究不同分辨率图像对算法效果的具体影响。
3. 探究注意力机制在跨分辨率目标计数的应用：探索注意力机制在提高跨分辨率车辆计数准确性中的应用，尤其是如何通过注意力机制来增强模型综合多种特征表示的能力。

1.3.2 研究内容

1. 研究背景与意义分析：分析相关遥感技术的发展背景，高分辨率与低分辨率遥感图像的特点及其在车辆计数中的面临困难和挑战。梳理了当前目标检测、语义分割、基于回归和密度图的计数方法等方面的研究进展，特别关注跨分辨率图像处理及车辆计数领域的最新研究成果。
2. 跨分辨率车辆计数数据集分析：对 CRVC 数据集进行分析调研，了解其数据组成分布以及数据中隐含的性质的分析及建模。
3. 基于注意力机制的车辆计数模型设计：设计并实现一种新的基于注意力机制的深度学习模型，用于提高跨分辨率遥感图像车辆计数的准确性。
4. 算法的性能评估及优化：在 CRVC 数据集上评估设计算法的性能，进行消融实验，并与现有的车辆计数方法进行比较。进一步优化算法，以达到更高的计数准确性和更好的泛化能力。

1.4 章节安排

第一章为绪论，整体介绍研究背景及研究内容。第二章为目标计数方法计数框架，详细解释了目标计数领域的方法细节。第三章为跨分辨率车辆计数数据集介绍，包括了数据的组成、分布以及对数据性质的初步分析。第四章介绍了基于注意力机制的跨分辨率遥感影像计数方法，详细解释了基于注意力机制的深度学习网络设计细节。

2 目标计数方法技术框架

2.1 目标计数

目前目标计数领域主要有三类方法。一类是检测方法，通过目标检测模型识别出具体的物体位置，之后根据结果来进一步计数。但这类方法对于输入图像的分辨率有着较高的要求，往往需要物体具有明确清晰的边缘特征。在低分辨率下往往表现效果较差。一种是基于回归的方法，直接拟合出图像特征和目标数目之间的回归模型得到图像中对应物体的数目。但这种方法未能完整利用图像中的空间，及序列信息。当输入图像的大小和分布有变化的情况下，往往不具有很强的泛化能力。另一类方法是基于密度图的目标计数方法。此类方法通常先得出一个目标物体在区域内的一个分部，之后就可以通过密度分布来估计总体的数量。该方法在稠密计数的场景下往往具有较好的效果。在本文使用的跨分辨率车辆计数数据集上，可以把车辆计数视为一个稠密计数场景。使用基于密度图的计数方法相较其余两类方法有着更好的表现。受上述方法启发，本文将跨分辨率车辆计数问题转换为两个子问题，即综合跨分辨率图像信息的图像分割网络和映射分割结果和最终计数目标的回归模型。

2.2 语义分割模型

语义分割作为计算机视觉的一个核心研究方向，目前已经有了较为成熟的解决方法。它的目标是对图像中的每个像素进行细致的分类，从而实现对图像的像素级理解。像素级的输出能力使得该领域的很多方法在密度图的估计上也有着不错的表现。

2.2.1 卷积层

对于图像数据，常使用卷积层而不是全连接层来进行特征提取。卷积层具有的平移不变性和局部性非常适合处理图像数据，可以掌握图像的空间特征。下面给出卷积层的基本定义。

2.2.1.1 卷积模型

$$[\mathbf{H}]_{i,j} = u + \sum_{a=-\Delta}^{\Delta} \sum_{b=-\Delta}^{\Delta} [\mathbf{V}]_{a,b} [\mathbf{X}]_{i+a,j+b} \quad (2.1)$$

通过使用系数 $[\mathbf{V}]_{a,b}$ 对位置 (i,j) 附近的像素 $(i+a,j+b)$ 进行加权得到 $[\mathbf{H}]_{i,j}$ 。其中 $|a| > \Delta$ 或 $|b| > \Delta$ 约束条件使得该式满足局部性，即只关注于在位置像素 $(i+a,j+b)$ 的小领域范围内的参数，大大减少了参数量。 \mathbf{V} 被称为卷积核（convolution kernel）或者滤波器（filter），也是该卷积层的权重，通常该权重是可学习的参数。参数 a, b 也对应着卷积核的尺寸 k_h, k_w 。

2.2.1.2 多通道输入与多通道输出

上式（2.1）是单通道情况下卷积层的数学表示，当输入图像的通道数为 c_i 时，那么我们需要构造一个形状为 $c_i \times k_h \times k_w$ 的卷积核。由于输入和卷积核都有 c_i 个通道，我们可以对每个通道输入的二维张量和卷积核的二维张量进行互相关运算，再对通道求和得到一个二维张量。这就是一个输出通道的结果。如果我们需要输出通道数为 c_o 时，只需创建一个卷积核的形状是 $c_o \times c_i \times k_h \times k_w$ 。通道数量可以视作对于不同特征的描述，随着神经网络层数的加深，通常的做法是减少空间分辨率的同时增加通道数量。

2.2.1.3 填充和步幅

在应用多层卷积时，我们常常丢失边缘像素。填充（padding）可以解决这个问题。在输入图像的边界填充一定数量的元素（通常填充元素是 0）。通常，如果我们添加 p_h 行填充（大约一半在顶部，一半在底部）和 p_w 列填充（左侧大约一半，右侧一半），则输出形状将为

$$(n_h - k_h + p_h + 1), (n_w - k_w + p_w + 1) \quad (2.2)$$

这意味着输出的高度和宽度将分别增加 p_h 和 p_w 。在许多情况下，我们可以设置 $p_h = k_h - 1$ 和 $p_w = k_w - 1$ ，这样使得输入和输出具有相同的高度和宽度。假设 k_h 是奇数，我们将在高度的两侧填充 $p_h/2$ 行。如果 k_h 是偶数，通常会在输入顶部填充 $\lceil p_h/2 \rceil$ 行，在底部填充 $\lfloor p_h/2 \rfloor$ 行。同理，我们填充宽度的两侧。

感受野是指卷积网络中某一层输出特征图上的一个元素所对应的输入图像上

的区域大小。它表征着特征图能“看到”的区域的大小。我们可以通过连续的卷积来增加感受野，但这会增加参数量。我们还可以通过调整步幅来增大感受野。步幅是卷积操作中卷积核移动的步长。在对图像进行卷积时，卷积核从图像的一个角落开始，按照指定的步幅在图像上滑动，每次移动指定的像素数，直到覆盖整个图像。当步幅大于 1 时，卷积核每次移动多个像素，输出的特征图的尺寸也会随之减小。具体公式如下：

通常，当垂直步幅为 s_h ，水平步幅为 s_w 时，输出形状为

$$\lfloor (n_h - k_h + p_h + s_h)/s_h \rfloor \times \lfloor (n_w - k_w + p_w + s_w)/s_w \rfloor \quad (2.3)$$

如果我们设置了 $p_h = k_h - 1$ 和 $p_w = k_w - 1$ ，则输出形状将简化为 $\lfloor (n_h + s_h - 1)/s_h \rfloor \times \lfloor (n_w + s_w - 1)/s_w \rfloor$ 。更进一步，如果输入的高度和宽度可以被垂直和水平步幅整除，则输出形状将为 $(n_h/s_h) \times (n_w/s_w)$ 。

2.2.2 激活函数

卷积神经网络中常用的激活函数包括 ReLU（线性整流单元）、Sigmoid、Tanh（双曲正切）等。这些激活函数的目的是在网络中引入非线性特性，使得网络能够学习到更加复杂的数据表示。本文用到的是 ReLU (Rectified Linear Unit) 函数。对于给定元素 x ，ReLU 函数被定义为该元素与 0 的最大值。它是目前最常用的激活函数之一。因为它的导数在大于 0 时为 1，小于 0 时为 0，这使得它可以用来缓解梯度消失的问题。

$$f(x) = \max(0, x) \quad (2.4)$$

2.2.3 池化层

池化（pooling）是卷积神经网络中常见的一种方法，主要用于减少特征图的维度，减少计算量的同时保留重要的一致性信息。与卷积层类似，池化运算也是通过一个固定形状的窗口滑动来实现的。与之不同的是，池化通过对邻近像素进行统计学操作（如取最大值或平均值）来实现，因此也不包含参数。主要有两种类型的池化：最大池化（Max Pooling）和平均池化（Average Pooling）。池化操作通常有两个参数：池化核的大小（ $K \times K$ ）和步幅（ S ）。池化核指定了池化操作的邻域范围，步幅定义了池化操作的移动间隔。对于输入大小为 $W \times H$ 的特征图，池化操作后的输出大小 $W' \times H'$ 可以通过以下公式计算：

$$W' = \left\lfloor \frac{W - K}{S} + 1 \right\rfloor \quad (2.5)$$

$$H' = \left\lfloor \frac{H - K}{S} + 1 \right\rfloor \quad (2.6)$$

在卷积网络的实践中，池化层通常有降低特征维度、引入不变性、增加鲁棒性和防止过拟合的作用。

2.2.4 权重衰减

在模型训练时，可能会遇到过拟合的问题，使得模型在已有数据上有着较好的性能，而在测试数据上表现不佳。我们可以使用多种正则化技术来缓解过拟合的问题。权重衰减（weight decay）是最广泛使用的正则化的技术之一，它通常也被称为 L_2 正则化。 L_2 正则化在损失函数中添加模型权重的平方之和作为惩罚项。同时通过一个非负的超参数 λ 来控制正则化的强度。 L_2 正则化修正后的损失函数如下式：

$$L(\mathbf{w}, b) = \frac{1}{n} \sum_{i=1}^n \frac{1}{2} (\mathbf{w}^\top \mathbf{x}^{(i)} + b - y^{(i)})^2 + \frac{\lambda}{2} \|\mathbf{w}\|^2 \quad (2.7)$$

L_2 正则化的目的是鼓励模型学习到更小更分散的权重值，从而提高模型的泛化能力。它对大的权重值施加较大的惩罚，从而防止模型依赖于少数几个可能具有高噪声的特征。

2.2.5 暂退法

Dropout 在训练过程中以一定几率随机“丢弃”（即暂时移除）网络中的一部分神经元（包括其连接），这有助于模型学习到更加鲁棒的特征，减少神经元间复杂的共适应关系。需要注意的是，在测试时，我们通常不使用 dropout。

2.2.6 批量归一化

批量归一化（Batch Normalization）是通过对每个小批量数据进行归一化处理，调整神经网络中间层的输出，使其均值接近 0，标准差接近 1。这可以通过减去它们的均值除以它们的标准差得到。这有助于稳定和加速深度网络的训练过程，同时也具有一定的正则化效果。批量归一化（Batch Normalization，简称 BN）是一种在深度神经网络中广泛使用的技术，用于加速训练过程并提高模型的稳定性。其基本思想是在网络的每层之后添加一个归一化步骤，这个步骤会对每个小批量数据（mini-batch）进行归一化处理，以确保网络中间层的激活分布保持稳定。批量归一化的公式如下：

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (2.8)$$

其中， ϵ 是一个很小的数，用来防止除以零。归一化后的 \hat{x}_i 具有零均值和单位方差。

2.2.7 损失函数

2.2.8 梯度下降

2.3 U-Net 网络

U-Net 是一个广泛被应用的语义分割模型，U-Net 是一个具有对称结构的网络，通过使用跳跃连接（Skip Connection）来结合低层次的位置信息和高层次的语义信息，从而在细节上进行更准确的预测。

2.4 注意力机制

2.5 Attention UNet 网络

Attention UNet 网络

2.6 光流估计

2.7 跨分辨率遥感图像目标计数技术路线

在文中引用公式可以这么写： $a^2 + b^2 = c^2$ 。这是勾股定理，它还可以表示为 $c = \sqrt{a^2 + b^2}$ 。还可以让公式单独一段并且加上编号：

$$\sin^2 \theta + \cos^2 \theta = 1 \quad (2.9)$$

注意，公式前请不要空行。

还可以通过添加标签在正文中引用公式，如式（2.9）。

我们还可以轻松打出一个漂亮的矩阵：

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 11 & 22 & 33 & 44 \end{bmatrix} \times \begin{bmatrix} 22 & 24 \\ 32 & 34 \\ 42 & 44 \\ 52 & 54 \end{bmatrix} \quad (2.10)$$



图 2.1 插图示例

或者多行对齐的公式:

$$\begin{aligned} f_1(x) &= (x + y)^2 \\ &= x^2 + 2xy + y^2 \end{aligned} \quad (2.11)$$

模板使用了 `unicode-math` 包更改数学字体。所以在使用数学字体时, 尽量使用 `unicode-math` 包提供的 `\sym` 接口, 详情请阅读 `unicode-math` 文档。

2.8 插图的使用

\LaTeX 环境下可以使用常见的图片格式: JPEG、PNG、PDF 等。当然也可以使用 \LaTeX 直接绘制矢量图形, 可以参考 `pgf/tikz` 等包中的相关内容。需要注意的是, 无论采用什么方式绘制图形, 首先考虑的是图片的清晰程度以及图片的可理解性, 过于不清晰的图片将可能会浪费很多时间。

`[htbp]` 选项分别是此处、页顶、页底、独立一页。`[width=\textwidth]` 让图片占满整行, 或 `[width=2cm]` 直接设置宽度。可以随时在文中进行引用, 如图 2.1, 建议缩放时保持图像的宽高比不变。

如果一个图由两个或两个以上分图组成时, 各分图分别以 (a)、(b)、(c)..... 作为图序, 并须有分图题。模板使用 `subcaption` 宏包来处理, 比如图 2.2a 和图 2.2b。



(a) 武汉大学校徽



(b) 武汉大学

图 2.2 多个分图的示例

2.9 表格的使用

表格的输入可能会比较麻烦，可以使用在线的工具，如 Tables Generator 能便捷地创建表格，也可以使用离线的工具，如 Excel2LaTeX 支持从 Excel 表格转换成 LaTeX 表格。LaTeX/Tables 上及 Tables in LaTeX 也有更多的示例能够参考。

2.9.1 普通表格

下面是一些普通表格的示例：

表 2.1 简单表格

我是	一只	普通
的	表格	呀

也可以使用 booktabs 包创建三线表。

表 2.2 一般三线表

姓名	学号	性别
张三	001	男
李四	002	女

三线表中三条横线分别使用 `\toprule`、`\midrule` 与 `\bottomrule`。若要添加 $m-n$ 列的横线，可使用 `\cmidrule{m-n}`。

要创建占满给定宽度的表格需要使用到 tabularx 包提供的 tabularx 环境。引用表格与其它引用一样，只需要如表 2.3。

表 2.3 占满文字宽度的三线表

序号	年龄	身高	体重
1	14	156	42
2	16	158	45
3	14	162	48
4	15	163	50
平均	15	159.75	46.25

- (b) 第一项中的第二项
- 2. 第二项
 - (i) 第一项中的第一项
 - (ii) 第一项中的第二项
- 3. 第三项

2.10.2 不计数列表

这是一个不计数的列表

- 第一项
 - 第一项中的第一项
 - 第一项中的第二项
- 第二项
- 第三项

2.11 数学环境的使用

模板简单定义了 8 种数学环境，具体见表 2.5，使用方法如下所示。

定理 2.11.1：设向量 $\mathbf{a} \neq \mathbf{0}$ ，那么向量 $\mathbf{b} \parallel \mathbf{a}$ 的充分必要条件是：存在唯一的实数 λ ，使 $\mathbf{b} = \lambda \mathbf{a}$ 。

定义 2.11.1：这是一条定义。

引理 2.11.1：这是一条引理。

推论 2.11.1：对数轴上任意一点 P ，轴上有向线段 \overline{OP} 都可唯一地表示为点 P 的坐标与轴上单位向量 \mathbf{e}_u 的乘积： $\overline{OP} = ue_u$ 。

性质 2.11.1：这是一条性质。

例 2.11.1：这是一条例。

注 2.11.1：这是一条注。

证明：留作练习。

□

表 2.5 模板定义的数学环境

theorem 定理	definition 定义	lemma 引理	corollary 推论
proposition 性质	example 例	remark 注	proof 证明

若要定义自己的数学环境，可通过如下代码实现：

```
\newtheorem{nonsense}{胡说}
\newtheorem*{bullshit}{八道}
```

其中，带星号 * 的命令不会自动编号。

胡说 1：啊吧啊吧啊吧。

八道：不啦不啦不啦。

2.12 单位

单位的输入请使用 `siunitx` 包中提供的 `\si` 与 `\SI` 命令，可以方便地处理希腊字母以及数字与单位之间的空白。在以前， \LaTeX 中输入角度需要使用 `\circ` 的奇技淫巧，现在只需要 `\ang` 命令解决问题。当然 `siunitx` 包中还提供了不少其他有用的命令，有需要的可以自行阅读 `siunitx` 文档。

示例： $6.4 \times 10^6 \text{ m}$ ， $9 \mu\text{m}$ ， kg m s^{-1} ， $104^\circ 28'$ 。

2.13 物理符号

`physics` 宏包可以让用户更加方便、简洁地使用、输入物理符号，具体也请自行阅读 `physics` 文档。示例如下

$$\begin{aligned}
 \int_0^{2\pi} |\sin x| \, dx &= 2 \int_0^{\pi} \sin x \, dx \\
 &= -2 \cos x \Big|_0^{\pi} \\
 &= 4
 \end{aligned}
 \tag{2.12}$$

3 跨分辨率车辆计数数据集

3.1 脚注

3.2 脚注

3.3 脚注

3.4 脚注

3.5 脚注

3.6 脚注

注释是对论文中特定名词或新名词的注解。注释可用页末注或篇末注的一种。选择页末注的应在注释与正文之间加细线分隔，线宽度为 1 点，线的长度不应超过纸张的三分之一宽度。同一页类列出多个注释的，应根据注释的先后顺序编排序号。字体为宋体 5 号，注释序号以“①、②”等数字形式标示在被注释词条的右上角。页末或篇末注释条目的序号应按照“①、②”等数字形式与被注释词条保持一致，脚注序号每面更新。示例：这里有个注释^①。

3.7 引用文中小节

如引用小节 3.7

3.8 引用参考文献

这是一个参考文献引用的范例：“^[2] 提出……”。还可以引用多个文献：“^[2-4] 提出……”。不同的引用方法：“江泽民^[2]”“^[4]”更多引用命令请参阅 `natbib` 文档或 `biblatex` 文档。

文献引用需要配合 `BibTeX` 使用，很多工具可以直接生成 `BibTeX` 文件（如 `EndNote`、`NoteExpress`、百度学术、谷歌学术等），此处不作介绍

①我是解释注释的

3.9 链接相关

模板使用了 `hyperref` 包处理相关链接，使用 `\href` 可以生成超链接，默认不显示链接颜色。如果需要输出网址，可以使用 `\url` 命令，示例：`https://github.com`。

4 基于注意力机制的跨分辨率遥感影像计数

4.1

4.2 代码

4.3 代码

4.4 代码

4.4.1 原始代码

朴实的代码块：

使用 `verbatim` 环境可以得到如下原样的输出。

```
print("Hello world!")
```

使用 `listings` 包提供的 `lstlisting` 环境可以对代码进行进一步的格式化，`minted` 包所提供的 `minted` 环境还可以对代码进行高亮。更多定制功能请自行参照文档配置。

4.4.2 算法描述/伪代码

参考 `Algorithms` 与 `algorithm2e` 文档，给出一个简单的示例，见算法 1。

Data: this text

Result: how to write algorithm with $\text{\LaTeX 2}_{\epsilon}$

initialization;

while *not at end of this document* **do**

 read current;

if *understand* **then**

 go to next section;

 current section becomes this one;

else

 go back to the beginning of current section;

end

end

算法 1 如何写算法

4.5 绘图

关于使用 \LaTeX 绘图的更多例子，请参考 `Pgfplots package`。一般建议使用如 Photoshop、PowerPoint 等制图，再转换成 PDF 等格式插入。

4.6 写在最后

工具不重要，对工具的合理运用才重要。希望本模板对大家的论文写作有所帮助。

5 跨分辨率遥感影像计数实验及分析

5.1 AB 测试

5.2 代码

5.3 代码

5.4 代码

6 总结与展望

6.1 结论

6.2 代码

6.3 代码

6.4 代码

参考文献

- [1] NAVULUR K. Multispectral image analysis using the object-oriented paradigm[M].
- [2] 江泽民. 能源发展趋势及主要节能措施[J]. 上海交通大学学报, 1989, 23(3): 1-16.
- [3] KUHN R L. The man who changed china: The life and legacy of jiang zemin[M]. Crown Publishers, 2004.
- [4] 江泽民. 新时期我国信息技术产业的发展[J]. 上海交通大学学报, 2008, 42(10): 1589-1607.
- [5] 张昆, 冯立群, 余昌钰, 等. 机器人柔性手腕的球面齿轮设计研究[J]. 清华大学学报: 自然科学版, 1994, 34(2): 1-7.
- [6] 竺可桢. 物理学论[M]. 北京: 科学出版社, 1973: 56-60.
- [7] DUPONT B. Bone marrow transplantation in severe combined immunodeficiency with an unrelated mhc compatible donor[C]//WHITE H J, SMITH R. Proceedings of the third annual meeting of the International Society for Experimental Hematology. Houston: International Society for Experimental Hematology, 1974: 44-46.
- [8] 郑开青. 通讯系统模拟及软件[D]. 北京: 清华大学无线电系, 1987.
- [9] 姜锡洲. 一种温热外敷药制备方案: 中国, 88105607.3[P]. 1980-07-26.
- [10] 中华人民共和国国家技术监督局. GB3100-3102. 中华人民共和国国家标准-量与单位[S]. 北京: 中国标准出版社, 1994.
- [11] MERKT F, MACKENZIE S R, SOFTLEY T P. Rotational autoionization dynamics in high rydberg states of nitrogen[J]. J Chem Phys, 1995, 103: 4509-4518.
- [12] MELLINGER A, VIDAL C R, JUNGEN C. Laser reduced fluorescence study of the carbon monoxide nd triplet rydberg series - experimental results and multichannel quantum defect analysis[J]. J Chem Phys, 1996, 104: 8913-8921.
- [13] BIXON M, JORTNER J. The dynamics of predissociating high Rydberg states of NO[J]. J Chem Phys, 1996, 105: 1363-1382.
- [14] 马辉, 李俭, 刘耀明, 等. 利用 REMPI 方法测量 BaF 高里德堡系列光谱[J]. 化学物理学报, 1995, 8: 308-311.
- [15] CARLSON N W, TAYLOR A J, JONES K M, et al. Two-step polarization-labeling

- spectroscopy of excited states of Na₂[J]. Phys Rev A, 1981, 24: 822-834.
- [16] TAYLOR A J, JONES K M, SCHAWLOW A L. Scanning pulsed-polarization spectrometer applied to Na₂[J]. J Opt Soc Am, 1983, 73: 994-998.
- [17] TAYLOR A J, JONES K M, SCHAWLOW A L. A study of the excited 1Σ_g⁺ states in Na₂[J]. Opt Commun, 1981, 39: 47-50.
- [18] SHIMIZU K, SHIMIZU F. Laser induced fluorescence spectra of the a 3Π_u-X 1Σ_g⁺ band of Na₂ by molecular beam[J]. J Chem Phys, 1983, 78: 1126-1131.
- [19] ATKINSON J B, BECKER J, DEMTRÖDER W. Experimental observation of the a 3Π_u state of Na₂[J]. Chem Phys Lett, 1982, 87: 92-97.
- [20] KUSCH P, HESSEL M M. Perturbations in the a 1Σ_u⁺ state of Na₂[J]. J Chem Phys, 1975, 63: 4087-4088.
- [21] 广西壮族自治区林业厅. 广西自然保护区[M]. 北京: 中国林业出版社, 1993.
- [22] 霍斯尼. 谷物科学与工艺学原理[M]. 李庆龙, 译. 2 版. 北京: 中国食品出版社, 1989: 15-20.
- [23] 王夫之. 宋论[M]. 刻本. 金陵: 曾氏, 1865 (清同治四年) .
- [24] 赵耀东. 新时代的工业工程师[M/OL]. 台北: 天下文化出版社, 1998[1998-09-26]. <http://www.ie.nthu.edu.tw/info/ie.newie.htm>.
- [25] 全国信息与文献工作标准化技术委员会出版物格式分委员会. GB/T 12450-2001 图书书名页[S]. 北京: 中国标准出版社, 2002: 1.
- [26] 全国出版专业职业资格考试办公室. 全国出版专业职业资格考试辅导教材: 出版专业理论与实务·中级[M]. 2014 版. 上海: 上海辞书出版社, 2004: 299-307.
- [27] World Health Organization. Factors regulating the immune response: Report of WHO Scientific Group[R]. Geneva: WHO, 1970.
- [28] PEEBLES P Z, Jr. Probability, random variables, and random signal principles[M]. 4th ed. New York: McGraw Hill, 2001.
- [29] 白书农. 植物开花研究[M]//李承森. 植物科学进展. 北京: 高等教育出版社, 1998: 146-163.
- [30] WEINSTEIN L, SWERTZ M N. Pathogenic properties of invading microorganism [M]//SODEMAN W A, Jr, SODEMAN W A. Pathologic physiology: mechanisms of disease. Philadelphia: Saunders, 1974: 745-772.
- [31] 韩吉人. 论职工教育的特点[C]//中国职工教育研究会. 职工教育研究论文集.

- 北京: 人民教育出版社, 1985: 90-99.
- [32] 中国地质学会. 地质评论[J]. 1936, 1(1)-. 北京: 地质出版社, 1936-.
- [33] 中国图书馆学会. 图书馆学通讯[J]. 1957(1)-1990(4). 北京: 北京图书馆, 1957-1990.
- [34] American Association for the Advancement of Science. Science[J]. 1883, 1(1)-. Washington, D.C.: American Association for the Advancement of Science, 1883-.
- [35] 傅刚, 赵承, 李佳路. 大风沙过后的思考[N/OL]. 北京青年报, 2000-04-12(14)[2002-03-06]. <http://www.bjyouth.com.cn/Bqb/20000412/B/4216%5ED0412B1401.htm>.
- [36] 萧钰. 出版业信息化迈入快车道[EB/OL]. (2001-12-19)[2002-04-15]. <http://www.creader.com/news/20011219/200112190019.htm>.
- [37] Online Computer Library Center, Inc. About OCLC: History of cooperation [EB/OL]. 2000[2000-01-08]. <http://www.oclc.org/about/cooperation.en.htm>.
- [38] Scitor Corporation. Project scheduler[CP/DK]. Sunnyvale, Calif.: Scitor Corporation, 1983.
- [39] SOUDANI K, FRANÇOIS C, le Maire G, et al. Comparative analysis of ikonos, spot, and etm+ data for leaf area index estimation in temperate coniferous and deciduous forest stands[J]. Remote Sensing of Environment, 2006, 102(1): 161-175.
- [40] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention u-net: Learning where to look for the pancreas: arXiv:1804.03999[M]. arXiv, 2018.
- [41] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation: arXiv:1505.04597[M]. arXiv, 2015.

致谢

以简短的文字表达作者对完成论文和学业提供帮助的老师、同学、领导、同事及亲属的感激之情。

附录 A 数据

A.1 第一个测试

测试公式编号

$$1 + 1 = 2.$$

(A.1)

表格编号测试

表 A.1 测试表格

11	13	13	13	13
12	14	13	13	13