

涉密论文 ☐ 公开论文 ☐

浙江大學

## 本科生毕业论文（设计）



题目 基于深度学习和卫星遥感的小麦制图

姓名与学号 曹徐韬 3180100230

指导教师 林涛研究员

年级与专业 农工 18 级

所在学院 生工食品学院

提交日期 \_\_\_\_\_



## 浙江大学本科生毕业论文（设计）承诺书（单面打印）

1. 本人郑重地承诺所呈交的毕业论文（设计），是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。

2. 本人在毕业论文（设计）中除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得浙江大学或其他教育机构的学位或证书而使用过的材料。

3. 与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

4. 本人承诺在毕业论文（设计）工作过程中没有伪造数据等行为。

5. 若在本毕业论文（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。

6. 本人完全了解浙江大学有权保留并向有关部门或机构送交本论文（设计）的复印件和磁盘，允许本论文（设计）被查阅和借阅。本人授权浙江大学可以将本论文（设计）的全部或部分内容编入有关数据库进行检索和传播，可以采用影印、缩印或扫描等复制手段保存、汇编本论文（设计）。

作者签名：

导师签名：

签字日期：      年      月      日      签字日期：      年      月      日



## 致谢

## 摘要

对于冬小麦制图问题，本文从数据集和模型两个方向入手提供解决方案。在数据集方面，本文利用美国中部小麦种植区十个地块构建了一个冬小麦制图数据集，为同类研究提供便利。同时，本项目构建了机器学习类模型 RandomForest, RNN 类模型 DCM, CNN 类模型 U-net, 以及 RNN+CNN 类模型 WMM。其中 WMM 是在本项目中第一次被提出的，也在冬小麦制图数据集上进行了训练和测试，结果表现出色。此外，本文也探究了多源数据对于冬小麦制图模型构建的影响，结果显示多源数据的使用有助于模型的表现。

**关键词：**冬小麦，深度学习，卫星，遥感

## Abstract

This article gives answers to the winter wheat mapping challenge from both the dataset and model perspectives. To aid comparable analyses, this paper creates a winter wheat cartographic dataset utilizing ten sites in the wheat-growing regions of the central United States. Simultaneously, this project develops the machine learning models RandomForest, RNN model DCM, CNN model U-net, and RNN+CNN WMM. WMM was proposed for the first time in this project and was also trained and tested on the winter wheat cartographic dataset, with outstanding results. Additionally, this article investigates the effect of multi-source data on the development of the winter wheat cartographic model. The results indicate that incorporating data from multiple sources improves the model's performance.

**Keywords:** Winter wheat, deep learning, satellite, remote sensing





## 目录（示例）

### 第一部分 毕业论文（设计）

<b>1</b>	<b>绪论 .....</b>	<b>1</b>
1.1	研究背景 .....	1
1.2	相关研究 .....	1
1.3	存在问题 .....	3
1.4	主要研究内容 .....	4
<b>2</b>	<b>卫星数据集获取与预处理 .....</b>	<b>5</b>
2.1	卫星及地理数据来源 .....	5
2.1.1	<i>Sentinel-1</i> .....	5
2.1.2	<i>Landsat</i> .....	6
2.1.3	<i>USDA NASS Cropland Data Layers</i> .....	6
2.2	数据获取 .....	7
2.3	数据预处理 .....	8
2.3.1	缺失值填补 .....	8
2.3.2	数据集制作 .....	9
<b>3</b>	<b>模型构建 .....</b>	<b>12</b>
3.1	机器学习类：RANDOMFOREST 模型 .....	12
3.1.1	模型简介 .....	12
3.1.2	<i>RandomForest</i> 模型构建 .....	12
3.2	RNN 类：AtBiLSTM 模型 (DCM) .....	13
3.2.1	模型简介 .....	14
3.2.2	模型构建 .....	14
3.3	CNN 类：U-NET 模型 .....	15
3.3.1	模型简介 .....	16
3.3.2	模型构建 .....	17
3.4	CNN+RNN 类：CONVLSTM 模型 (WHEATMAPPINGMODEL) .....	17
3.4.1	模型简介 .....	18

<b>4</b>	<b>实验设计 .....</b>	<b>23</b>
4.1	实验环境 .....	23
4.2	模型性能测试 .....	23
4.2.1	本地测试 .....	23
4.2.2	特征提取可视化 .....	24
4.3	数据特征测试 .....	24
<b>5</b>	<b>实验结果与分析 .....</b>	<b>25</b>
5.1	模型性能测试 .....	25
5.1.1	本地测试 .....	25
5.1.2	特征提取可视化 .....	26
5.2	数据特征测试 .....	27
<b>6</b>	<b>总结与展望 .....</b>	<b>29</b>
	参考文献 .....	30
<b>附录33</b>		
	代码（UPDATING）： <a href="https://github.com/CAOXUTAO7/WHEAT-MAPPING.GIT">HTTPS://GITHUB.COM/CAOXUTAO7/WHEAT-MAPPING.GIT</a> ....	33
	作者简历 .....	34
	《浙江大学本科生毕业论文（设计）任务书》 .....	（可不编页码）
	《浙江大学本科生毕业论文（设计）考核表》 .....	（可不编页码）

## 第二部分 文献综述和开题报告

	指导教师对文献综述和开题报告具体内容要求 .....	（可不编页码）
	一、文献综述 .....	1
	二、开题报告 .....	3
	三、外文翻译 .....	5
	四、外文原文 .....	（可不编页码）
	《浙江大学本科生文献综述和开题报告考核表》 .....	（可不编页码）

# 第一部分

毕业论文（设计）



# 1 绪论

## 1.1 研究背景

小麦是中国乃至世界的主要粮食作物之一<sup>[1]</sup>，及时准确的小麦制图对于粮食安全以及环境可持续性都非常重要。然而，由于一些国家的气候变化、农田转换和作物种植结构的调整，小麦种植正在发生变化。因此，及时准确地了解小麦的空间分布和时间变化信息对于农业管理、土地规划和环境可持续性至关重要<sup>[2]</sup>。而对小麦进行田间调查费时，劳动密集，昂贵，调查数据的质量容易受到人为错误的影响。在这种情况下高效自动的卫星遥感制图就成为了重要替代，它为从区域到局部的高时空分辨率作物制图提供了新的机会。

在此类任务中，所使用的遥感卫星主要分为两类，雷达卫星 (SAR卫星) 和高光谱卫星。雷达卫星具有穿透性强的特点，因此卫星数据不容易受天气的影响。而高光谱卫星能提供更多波段信息，但容易受天气的影响，通常只会选用云量<5%的数据。

获取了遥感数据后需要建立模型对冬小麦进行分类。有三种常见的分类方法：1、基于阈值的方法；2、基于机器学习的方法；3、基于深度学习的方法。

其中基于深度学习的方法是目前较为主流的方法，也是我们研究的重点方向。相比于另外两种方法，深度学习方法不需要依赖于人工进行特征工程，可以自动化的方式实现特征的提取，同时可能提取到人工难以提取到的深层次特征。在这个研究课题中，卫星数据十分丰富，我们可以构建起时序卫星序列图像，可以将多源卫星数据进行融合，甚至可以再融入统计数据。面对巨量数据，机器学习方法往往会陷入欠拟合的困境，而拥有合理模型架构的深度学习模型则会提取更多有效特征，使模型更具泛化能力。所以在这个课题中深度学习方法有着良好的应用前景，而如何寻找合理的模型架构也成为了研究者们重点关注的目标。

## 1.2 相关研究

表 1.1 小麦制图相关研究

模型方法	输入特征	卫星	规模	结果	参考来源
------	------	----	----	----	------

基于 阈值 的方法	构建特征变量 EVE, EVL	EVI	MODIS, Landsat	省级	OA: 92.19% 88.86%	Qiu, Luo, Tang, Chen, Lu, Huang, Chen, Chen and Xu <sup>[2]</sup>
	相似度度量阈值	NDVI	Sentinel-2	市级	OA: 94.5% K: 0.8894	Li, Ren, Wu, Zhao and Zhang <sup>[3]</sup>
	Phenology-Time Weighted Dynamic Time Warping (PT-DTW)	NDPI	Sentinel-2	省级	OA: 89.98% K: 0.7978	Dong, Chen, Chen, Zhang, Liu, Cao, Zang, Zhu and Cui <sup>[4]</sup>
	time-weighted dynamic time warping (TWDTW)	NDVI	Sentinel-2 Landsat	省级	PA: 89.30% UA:90.59%	Dong, Fu, Wang, Tian, Fu, Niu, Han, Zheng, Huang and Yuan <sup>[5]</sup>
机器 学习	RandomForest	VI	Sentinel-2	市级	K: 0.7-0.78	Zhang, Fang, She, Zhang, Jin, Xia, Yang and Ding <sup>[6]</sup>
	RandomForest	OP	GF-1 Landsat	市级	OA: 92.9% K: 0.858	Liu, Feng, Gong, Zhou, Liang and Li <sup>[7]</sup>
	RandomForest SVM	PM, OP	Sentinel-1 Landsat	市级	F1: 87.89% 98.06%	Zhou, Pan, Zhang, Wei and Han <sup>[1]</sup>
深度 学习	CNN	PM	MODIS	省级	OA: 95%+	Zhong, Hu, Zhou and Tao <sup>[8]</sup>

表格中缩略词解释。NPVI: 正则化差分时间指数 (normalized differential phenology index)。NDVI: 正则化植被指数。EVI: 强化植被指数。OP: 光谱图像。PM: 极化图像。VI: 植被指数。OA: 总体准确度。UA: 用户准确度。PA: 生产者准确度。K: kappa系数

其中使用深度学习方法对小麦进行制图的相关研究并不是很多，但是有很多利用其它作物进行深度学习制图的研究值得借鉴。

表 1.2 基于深度学习的作物制图研究

模型方法	输入特征	卫星	规模	作物	参考来源
Conv2d Conv1d	OP, PM	Landsat, Sentinel-1	市级	多种	Kussul, Lavreniuk, Skakun and Shelestov <sup>[9]</sup>

Conv1d LSTM	EVI	Landsat	市级	多种	Zhong, Hu and Zhou [10]
Conv3d	OP	OHS-2A	市级	多种	Meng, Wang, Hu, Luo and Zhong [11]
LSTM	OP	Landsat	市级	玉米和大豆	Xu, Zhu, Zhong, Lin, Xu, Jiang, Huang, Li and Lin [12]
2D U-Net + CLSTM, 3D U-Net	OP, PM	N/A	国家	多种	M Rustowicz, Cheong, Wang, Ermon, Burke and Lobell [13]
1D CNN, LSTM, GRU, LSTM-CNN, GRU-CNN	OP	Sentinel-2	市级	多种	Zhao, Duan, Liu, Sun and Reymondin [14]
convRNN	OP	Sentinel-2	市级	多种	Turkoglu, D'Aronco, Perich, Liebisch, Streit, Schindler and Wegner [15]
U-net 和其他语义分析模型	PM	Sentinel-1	省级	米	Wei, Chai, Lin, Tang, Du and Huang [16]
DeepLabv3	RGB	WorldView-2	市级	作物/非作物	Du, Yang, Ou and Zhang [17]

### 1.3 存在问题

遥感数据中存在样本不平衡的问题。往往遥感图像中是小麦的像素比不是小麦的像素要少很多，会影响模型判别小麦的准确率。针对这个问题，研究者也选择了几种解决方法：1、尽量选择小麦比例较大的卫星数据。2、改变损失函数，增加小麦像素的权重。但是这两种解决方法都不能普适性地解决样本不平衡的问题。

卫星图像数据会有缺失，天气（云量）也会影响图像数据的可用性，这对于利用时间序列图像构建模型的方法有较大影响。为了提高模型的泛化能力，将多源卫星数据融合是一种可行的<sup>[18]</sup>途径，这也对卫星图像的完整性提出了更高要求。

现有许多研究是基于 MODIS 卫星数据，分辨率仅为 250m，为了更精确的冬小麦制图工作，就对于卫星分辨率提出了更高要求。所以现有许多研究是基于10m 分辨率的 Sentinel 卫星和 Landsat 卫星。

部分区域缺少有地块标签的数据进行模型训练。由于该研究主流的模型是有监督模型，需要有标签数据进行训练，在缺少地块标签数据的区域进行冬小麦制图时则需要从其他区域训练模型并进行迁移，精确度并不能得到保证。

基于深度学习的作物制图的主要限制在于模型的可解释性。由于分层特征转换能力，深度神经网络通常优于人工规则设计的浅层分类器<sup>[19]</sup>。然而，深度学习模型通常被认为是“黑匣子”，具有不可见的特征学习过程，可产生更高级别的抽象特征<sup>[20]</sup>。解释深度学习方法的特征学习可能有利于我们解释复杂的判别模式并确保模型的可靠性<sup>[21]</sup>。所以还需要更多面向模型解释性的研究。

## 1.4 主要研究内容

首先，本项目将以冬小麦为研究主体，探究融合 Sentinel-1 卫星和 Landsat 卫星数据的基本方法，并构建相应冬小麦分类数据集。

其次，本项目会基于该数据集构建多种常见模型，并在本数据集上训练，评估其性能。同时，项目也会研究不同数据来源对模型精度的影响，比较并分析多源数据融合对于小麦制图任务的价值。



## 2 卫星数据集获取与预处理

### 2.1 卫星及地理数据来源

#### 2.1.1 Sentinel-1

Sentinel-1 卫星是欧洲航天局 (ESA) 和欧洲委员会 (EC) 联合启动的哥白尼计划 (GMES) 中的一颗地球观测卫星。哥白尼计划旨在基于从地球观测卫星接收到的观测数据和地面信息, 为环境和安全提供信息服务。

Sentinel-1 卫星包括以四种独特的成像模式运行的C波段 (5.405GHz) 成像, 具有不错的分辨率 (最低5m) 和覆盖范围 (高达400km)。它提供双极化能力、极短的重返周期和快速的产品交付服务。对于每次观测, 它都可以获得航天器位置和姿态的精确测量。

合成孔径雷达 (SAR) 具有在不受云层覆盖或阳光不足影响的优势, 并且可以在所有天气条件下在白天或黑夜获取站点上的数据。

Sentinel-1 及其 C-SAR 仪器可以提供可靠、重复的广域监测。该卫星由共享同一轨道平面的两颗卫星 Sentinel-1A 和 Sentinel-1B 组成。Sentinel-1 旨在以预编程、无冲突的操作模式工作, 以高分辨率对全球所有陆地、沿海地区和航线进行成像。这确保了运营服务所需的服务可靠性, 以及为基于长时间序列的应用构建一致的长期数据存档。<sup>[22]</sup>

Sentinel-1 可以通过使用其天线同时接收特定的极化信号, 从同一系列脉冲中收集多个不同的图像。Sentinel-1 是一种保相双极化 SAR 系统。它可以以水平 (H) 或垂直 (V) 极化传输信号, 然后以 H 和 V 极化接收。双极化 Level-1 Single Look Complex (SLC) 产品包含复数值。除了可以从每个单偏振测量的反向散射强度外, 通道间相位信息还允许对反向散射特性进行增强分析。

地面上的目标具有独特的偏振特征, 以不同的强度反映不同的偏振, 并将一种偏振转换为另一种。例如, 体积散射体 (例如森林树冠) 与表面散射体 (例如海面) 具有不同的偏振特性。极化技术, 如极化目标分解 (基于特征向量、基于模型等) 允许分离不同的散射贡献, 并可用于提取有关散射过程的信息。专用的双极化 SAR 分类技术可以改进点目标和分布式目标区域的分类<sup>[23]</sup>。

在本次实验中，我选取了 Sentinel-1A 卫星中的 VV 和 VH 通道作为数据集来源。

### 2.1.2 Landsat

Landsat 计划是运行时间最长的地球卫星图像采集计划。这是一项 NASA / USGS 联合计划。1972 年 7 月 23 日，地球资源技术卫星 (Earth Resources Technology Satellite) 发射。这最终在 1975 年更名为 Landsat 1。最新的 Landsat 9 于 2021 年 9 月 27 日发射。

Landsat 卫星上的仪器已经获取了数百万张图像。这些图像在美国和世界各地的 Landsat 接收站存档，是全球变化研究和农业、制图、地质、林业、区域规划、监测和教育应用的独特资源。Landsat 7 数据有 8 个光谱带，空间分辨率从 15 到 60 m; 时间分辨率为 16 天。Landsat 图像通常分为景 (scene) 以便于下载。每个 Landsat 景长约 185 公里，宽约 185 公里<sup>[24]</sup>。

此次实验我也同样获取了 Landsat-7 卫星与 Landsat-8 卫星高光谱数据作为数据集来源。获取的波段数据如下。

表 2.1 Landsat 选取波段

波段名称	波长范围
Blue	0.452-0.512 $\mu\text{m}$
Green	0.533-0.590 $\mu\text{m}$
Red	0.636-0.673 $\mu\text{m}$
Near infrared	0.851-0.879 $\mu\text{m}$
Shortwave infrared 1	1.566-1.651 $\mu\text{m}$
Shortwave infrared 2	2.107-2.294 $\mu\text{m}$

### 2.1.3 USDA NASS Cropland Data Layers

Cropland Data Layers (CDL) 是每年使用中等分辨率卫星图像和广泛的农业地面实况为美国大陆创建的特定作物土地覆盖数据层。CDL 由美国农业部、国家农业统计局 (NASS)、研发，地理空间信息和空间分析研究中心创建。

在本实验中，我选取了其中的冬小麦土地覆盖数据层作为数据集标签。

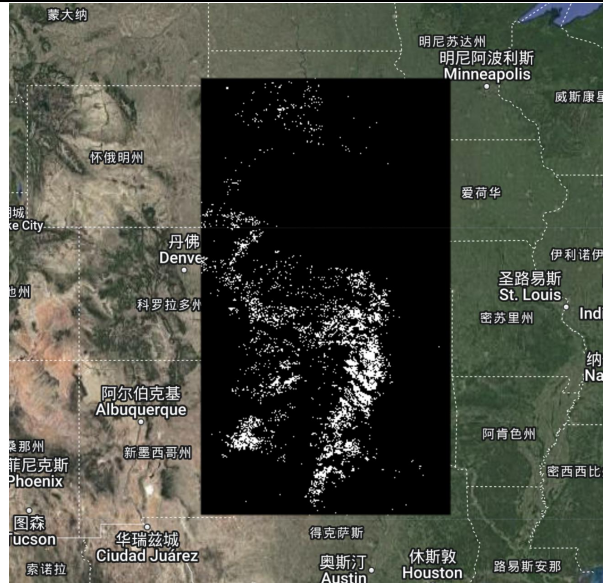


图 2.1 2020 年美国中部地区冬小麦种植土地覆盖数据层 (CDL), 其中白色像素代表冬小麦种植区域, 黑色像素代表非冬小麦种植区域

## 2.2 数据获取

数据获取通过 Google Earth Engine 平台完成。Google Earth Engine 中有大量卫星及地理数据集, 我在其中选取了 Sentinel-1 SAR GRD; USGS Landsat 7 Level 2, Collection 2, Tier 1; USGS Landsat 8 Level 2, Collection 2, Tier 1; USDA NASS Cropland Data Layers 四类数据集。获取 Sentinel-1 数据时要在该平台上对图像进行地形矫正、Refined Lee 滤波。对于 Landsat 图像则也要在该平台上进行去云处理。

对于实验的每一个地点, 都要获取每一年10月4日至次年5月8日之间的所有图像来构建时间序列图像。

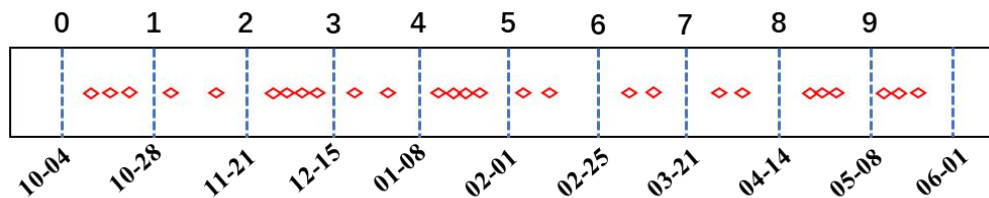


图 2.2 时间序列图像的序列分布



图 2.3 10 个取样数据点分布

## 2.3 数据预处理

### 2.3.1 缺失值填补

在使用 Google Earth Engine 完成图像获取之后，需要进行缺失值处理。因为我在后续的实验中除了序列模型还需要构建图像模型，所以单纯的去除缺失值并不可取，需要对缺失值进行填补。在本实验中，我一共采取了两种方法进行缺失值填补。

第一种是滑动窗口 (sliding window) 平均值法。如图2.3所示，假设当前需要填补的值为  $x$ ，窗口中数的集合为  $W$ ，可以用以下高斯分布估计  $x$

$$x \sim N(\text{nanmean}(W), \text{nanstd}(W))$$

再从获得的分布中对  $x$  随机取样，作为填补值。为了防止填补值失真严重，当  $5 \times 5$  的窗口中有大于等于给定阈值（在本次实验中设置为 6）个缺失像素时，会停止对当前缺失像素进行填补。该算法的复杂度仅为  $O(N)$ ， $N$  为缺失像素个数，相比 KNN 填补法、循环填补法更适合用于此类数据量较大的任务中。

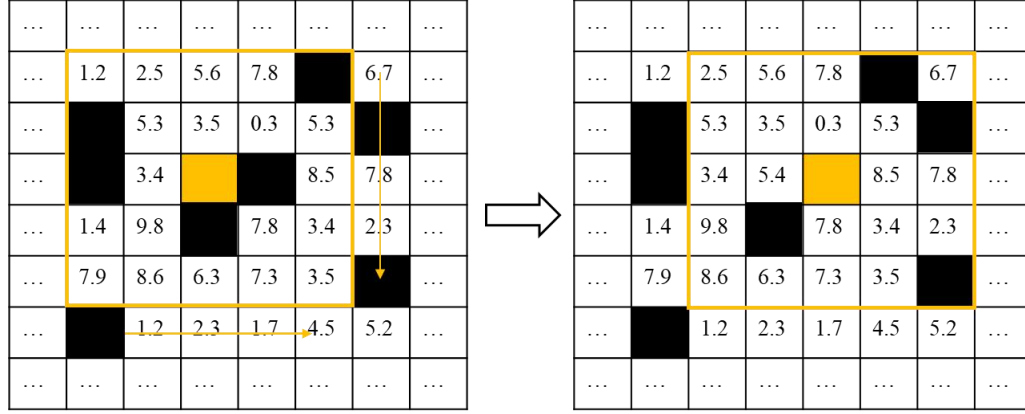


图 2.4 滑动窗口平均值法示意图。黑色格子处表示缺失值，黄色方框表示当前处理窗口，黄色格子处为当前正在填补的缺失值。黄色格子会通过相邻的  $5 \times 5$  窗口中的值计算出合适的值进行填补，然后滑动窗口，对下一个缺失值进行填补。

第二种是时序线性插值法，缺失值可以用上一个时间步和下一个时间步的平均值来填补（前提是上下两个时间步的对应像素不为缺失值）。同样，这也是一个  $O(N)$  的填补算法。

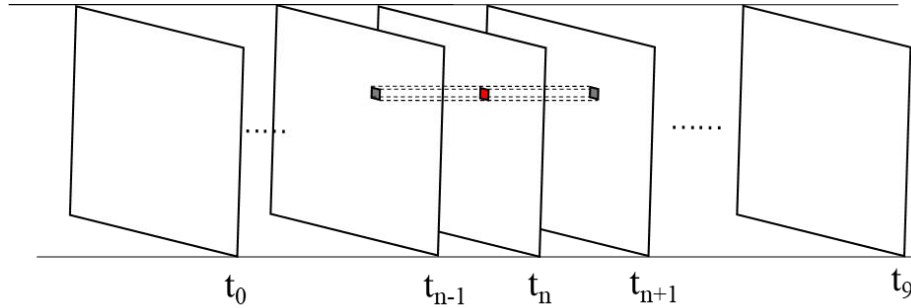


图 2.5 时序线性插值法示意图。红色格点表示当前缺失值，其填补值即为前后时间步的平均值

### 2.3.2 数据集制作

首先，将每一地点、每一年的所有图像以 24 天为时间间隔分开。如图 2.2，每个时间步的图像设为其之后 24 天所有图像的平均值。再将获取的 10 个时间步的图像堆叠起来，使图像增加新的时间维度。然后将同一地点的 Sentinel-1 卫星图像和 Landsat 卫星图像堆叠起来作为图像的特征维度。最终生成的图像



有4个维度：(H, W, C, T)。其中 H, W 分别代表图片的长、宽，C 代表图片的特征维度，T 代表图片的时间维度。

受GPU显存限制，大部分用于计算机视觉的深度学习模型的输入图片尺寸都不会太大，如 ImageNet<sup>[25]</sup> 的输入图片尺寸仅为 224\*224。因此，我需要将原始卫星图像裁剪以制作适合深度学习模型输入的数据集。

因为在上述处理后的图像中仍存有不少缺失值，而且图像有较多的维度与特征通道，用计算机视觉领域中常用的随机裁剪 (Random Cropping) 的方法的效果并不理想。随机裁剪得到的图像往往因为内含有缺失值而无法使用。如果随机裁剪后丢弃所有含有缺失值的图像，可能会因为缺失像素较为分散使得可用的随机裁剪后图像之间重合度较高，最终导致模型重复训练相似的数据集而形成过拟合。

因此，我采用了一种类似于移动窗口的方法进行图像裁剪。先将图片切割成 H\*W （本实验中选择尺寸为256\*256）大小的块，窗口的从每个块的初始位置开始向右或者向下移动直到寻找到没有缺失值的图片或者移动距离达到阈值。这种方法虽然复杂度为  $O(N^2)$ ，但是有效的避免了上述裁剪后图像重合度较高的问题。

表 2.2 数据集详情

地点	年份	尺寸 (px * px)	图片数量	冬小麦比例
0	2017	1095*1352	20	31.47%
	2018		20	28.01%
	2019		20	28.87%
	2020		20	24.27%
1	2017	1979*1999	49	41.73%
	2018		47	38.40%
	2019		49	38.02%
	2020		49	36.01%
2	2017	1915*2190	54	27.33%
	2018		54	31.01%
	2019		54	32.14%
	2020		54	20.30%
3	2017	1898*1915	49	40.13%
	2018		49	29.83%
	2019		49	35.84%
	2020		49	37.16%
4	2017	1959*2077	46	20.02%

基于深度学习和卫星遥感的小麦制图

	2018		46	21.19%
	2019		46	17.29%
	2020		46	18.82%
5	2017	2437*2464	67	23.12%
	2018		67	24.14%
	2019		67	18.64%
	2020		67	23.78%
6	2017	1963*1877	49	11.53%
	2018		49	14.34%
	2019		49	8.77%
	2020		49	11.14%
7	2017	1706*1942	42	44.89%
	2018		19	44.76%
	2019		42	44.33%
	2020		42	44.27%
8	2017	2350*2467	64	34.5%
	2018		68	32.36%
	2019		78	32.61%
	2020		78	35.83%
9	2017	1836*1732	42	44.99%
	2018		42	44.25%
	2019		42	42.50%
	2020		42	41.53%

### 3 模型构建

#### 3.1 机器学习类：RandomForest 模型

##### 3.1.1 模型简介

RandomForest<sup>[26]</sup> 模型由 Ho 于 1995 年首次提出，是一个在各类分类以及回归问题中常见的有监督机器学习算法。

RandomForest 由于其优秀的高维数据处理能力以及并行运算能力，十分适合在大数据集上进行训练。所以数据量普遍较大的遥感问题就十分适用这种方法<sup>[27]</sup>。

RandomForest 通过 Bagging 策略集成了多个决策树，并利用 Bootstrap 重抽样的方法限制每次树分类时进入计算的特征数和样本数量，使得其有较好的泛化能力以及较低的时间复杂度，而且具有较难过拟合的性质。所以在本项目中，RandomForest 被建立为基准模型，用于与其他模型对比。

##### 3.1.2 RandomForest 模型构建

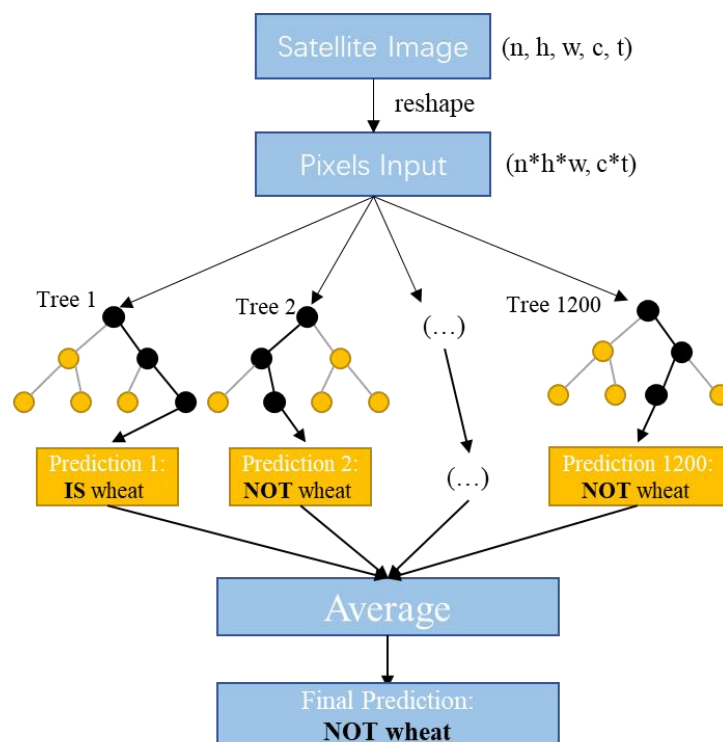


图 3.1 RandomForest 模型示意图。模型通过构建 1200 个决策树集成决策，预测输入像素是否为小麦。



该模型的构建使用了 Python Sklearn 库中的 RandomForestClassifier<sup>[28]</sup> 类。构建模型前需要将原始数据集中的5维图片向量变形为2维向量（如图3.1）。

### 3.2 RNN 类：AtBiLSTM 模型 (DCM)

RNN (Recurrent Neural Network) 是一种适用于序列数据的深度学习模型，它可以通过隐藏状态 (hidden state) 传递序列信息。但是RNN在处理较长序列时效果并不理想，容易产生梯度消失 (gradient vanish) 或者梯度爆炸 (gradient explode) 的问题。LSTM<sup>[29]</sup> 产生的初衷也就是为了解决 vanilla RNN 的这个问题问题，在 LSTM 中，第  $t$  个时间步的信息被选择性地传输到下一个时间步  $t+1$  中。LSTM采用了”门“结构来实现这种选择性，可以控制 LSTM 保留当前时间步的程度以及传入下一个时间步的程度，每次有新的输入输入时，如果输入门被激活，则其信息将基类到单元中，此外如果遗忘门  $f_t$  开启，则过去单元的状态会被遗忘。更进一步，是否将最新的单元输出  $c_t$  传播到最终状态  $h_t$  是通过输出门。这种设计使得 LSTM 相比于普通的 RNN 可以更好学习长依赖关系，更有利于处理时间序列信息。下列为 LSTM 的一些重要方程（其中 $\odot$ 代表Hadamard 乘积）：

$$\begin{aligned} i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci} \odot c_{t-1} + b_i) \\ f_t &= \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf} \odot c_{t-1} + b_f) \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \\ o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co} \odot c_t + b_o) \\ h_t &= o_t \odot \tanh(c_t) \end{aligned}$$

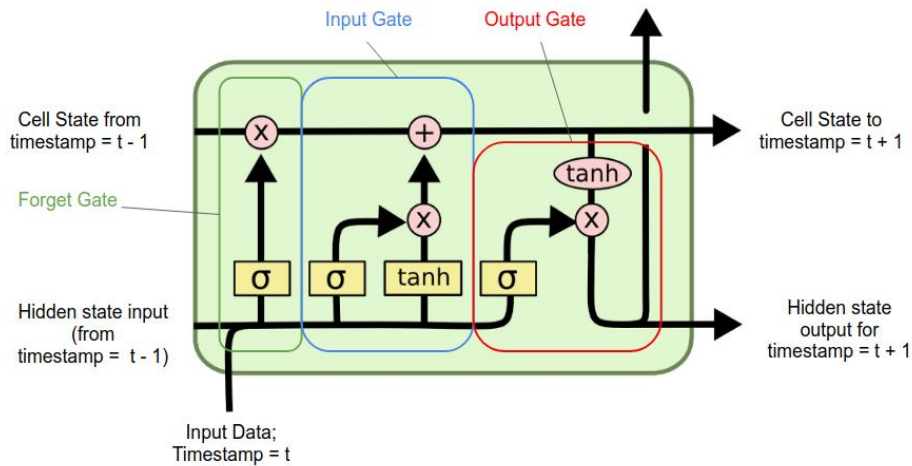


图 3.2 vanilla LSTM cell 示意图<sup>[30]</sup>

在遥感任务中，时序图像同样是一种序列数据，可以通过RNN类序列模型建模。

### 3.2.1 模型简介

DCM<sup>[12]</sup> 是 Xu 等在2020年提出的玉米和大豆制图深度学习模型，其模型简洁而有优良的性能，因此我将其模型迁移到本实验中。

DCM 模型由三个模块组成：输入模块，输出模块，AtBiLSTM 模块。在实验中，我保留了它前两个模块的设计，仅仅改变了输出层和相应的 loss function 来适配本次实验。其中输出层改变最终输出类别维度中向量长度为1，并将输出层激活函数改成 Sigmoid 函数。同时，我也将该模型中原先使用的 Cross Entropy Loss 改变为 Binary Cross Entropy Loss。修改部分具体公式如下所示。

$$\text{Sigmoid}(x) = \frac{e^x}{e^x + 1}$$

$$\text{Loss} = \frac{1}{N} \sum_{i=1}^N [p_c y_{n,c} \cdot \log \text{Sigmoid}(x_{n,c}) + (1 - y_{n,c}) \cdot \log(1 - \text{Sigmoid}(x_{n,c}))]$$

其中  $p_c$  值是用来平衡样本，增加在小样本在损失函数中权重，我们将  $p_c$  定义为  $\frac{\text{不是冬小麦的像素总数}}{\text{是冬小麦的像素总数}}$ ，其效果相当于将一个不平衡的数据集转化为正反例都是 1:1 的平衡数据集 DCM

### 3.2.2 模型构建

在实验中，DCM 模型利用 Pytorch 构建，输入DCM模型时同样要对图片进行变形，将原先的 (n,h,w,t,c) 的图片向量变形为 (n\*h\*w, t, c) 的向量输入。

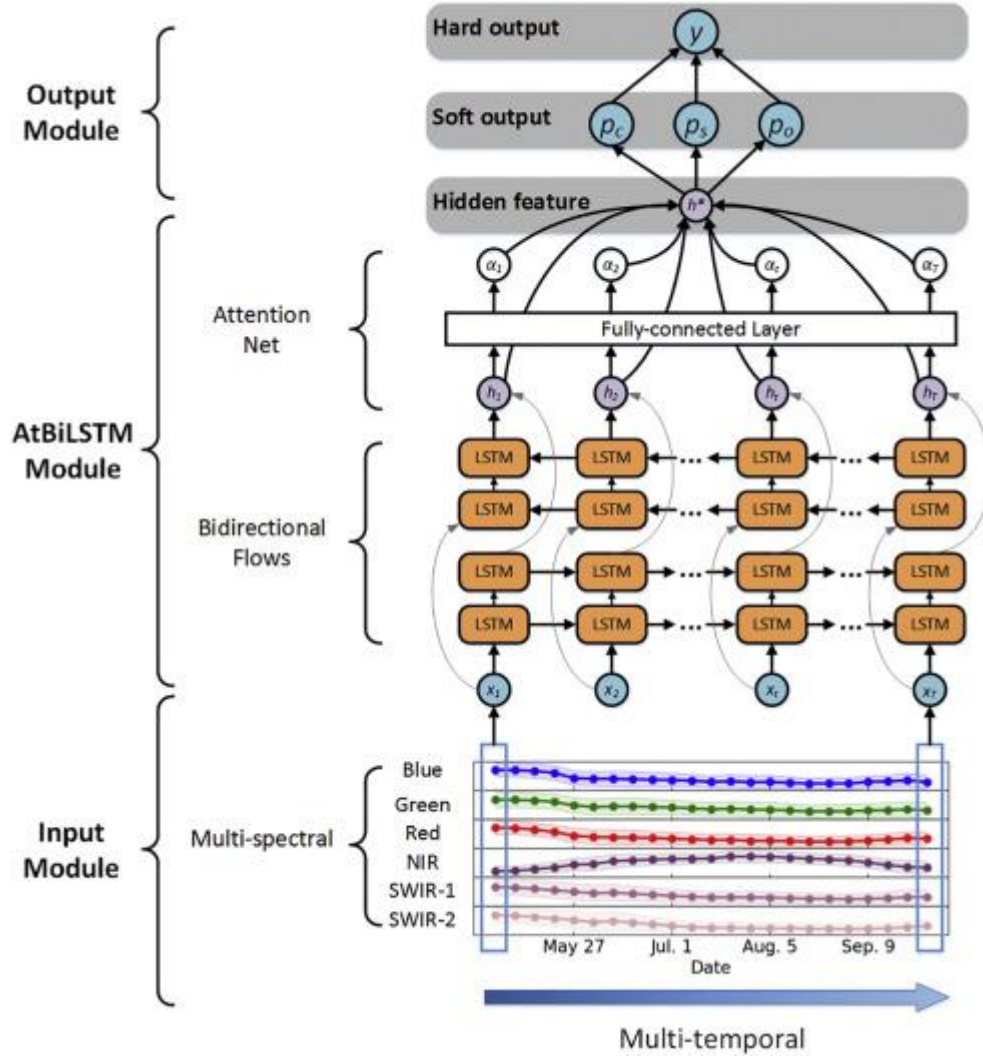


图 3.3 DCM 模型<sup>[12]</sup>

### 3.3 CNN 类：U-net 模型

另一种在作物制图中十分常见的模型就是基于卷积神经网络 (CNN) 的模型。不同于 RNN 类网络对时间序列信息的提取，CNN 类模型更关注空间信息，使用不同权重的卷积核将空间信息提取成不同的特征图 (feature map)，再用这些特征图进行分类。

遥感制图问题和计算机视觉中的语义分析 (semantic segmentation) 问题相似，最终结果都是要对图像的每一个像素输出相应分类。语义分析模型通常采用的是编码器-解码器 (Encoder-Decoder) 结构。编码器负责下采样 (down sample)，即增加特征图数量，减少图像大小来提取高维度的信息。解码器则负责上采样 (up sample)，减少特征图数量，输出原尺寸特征图像，从编码器提取

的高维信息中总结结果。计算机视觉领域中经典语义分析网络有 FCN<sup>[31]</sup>, Segnet<sup>[32]</sup>, U-net<sup>[33]</sup> 等，目前在经典语义分析数据集 ADE20k<sup>[34]</sup>上的 SOTA 网络为基于 Swin Transformer V2<sup>[35]</sup> 和 UperNet<sup>[36]</sup> 的网络 (SwinV2-G with UperNet)。但由于该网络的规模极大，无法迁移在本实验中，本实验选取了经典的 U-net 模型进行小麦制图的尝试。

### 3.3.1 模型简介

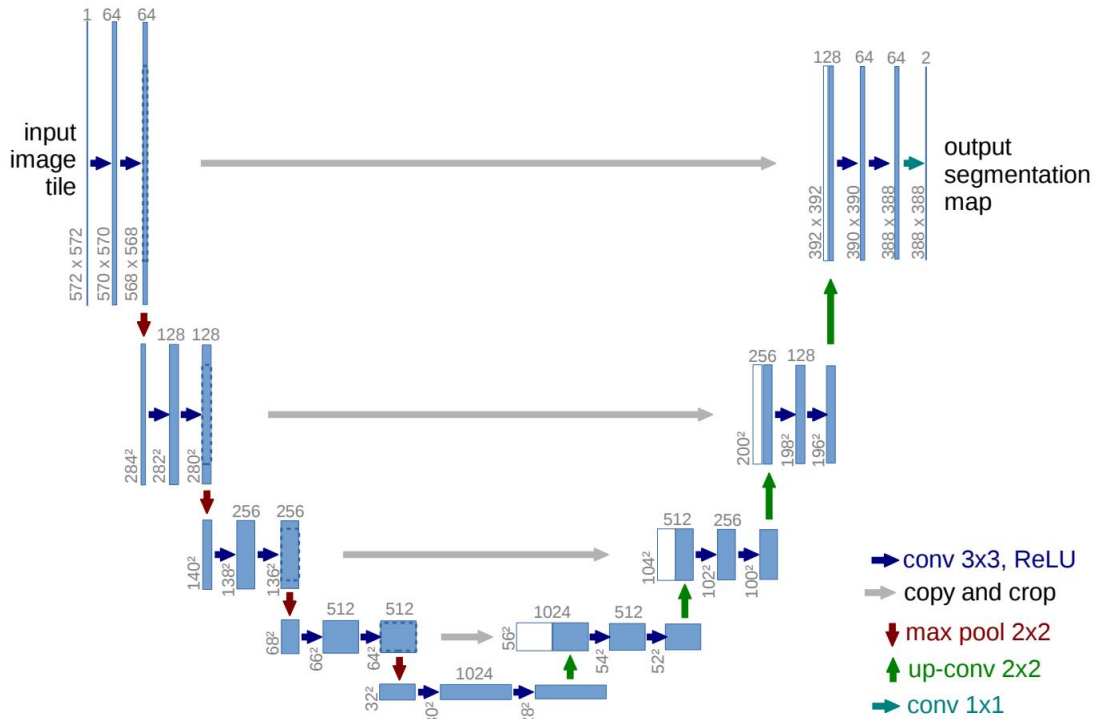


图 3.4 经典 U-net 结构<sup>[33]</sup>

U-net 最初被提出是应用于医学影像分割中，它是一个全卷积神经网络。它的主要思想是利用连续的上采样操作代替常用的池化操作来提取信息，同时提高图片的分辨率。同时，一个连续的卷积层可以学习这些信息组装一个更精确的输出。此外，由于 copy 操作的存在，U-net 既保留了传统 encoder-decoder 结构的信息传递路径，又保证了上采样部分有大量的特征通道，使得网络能够将信息传播到更高分辨率的层。该网络仅使用了卷积操作，没有任何全连接层。

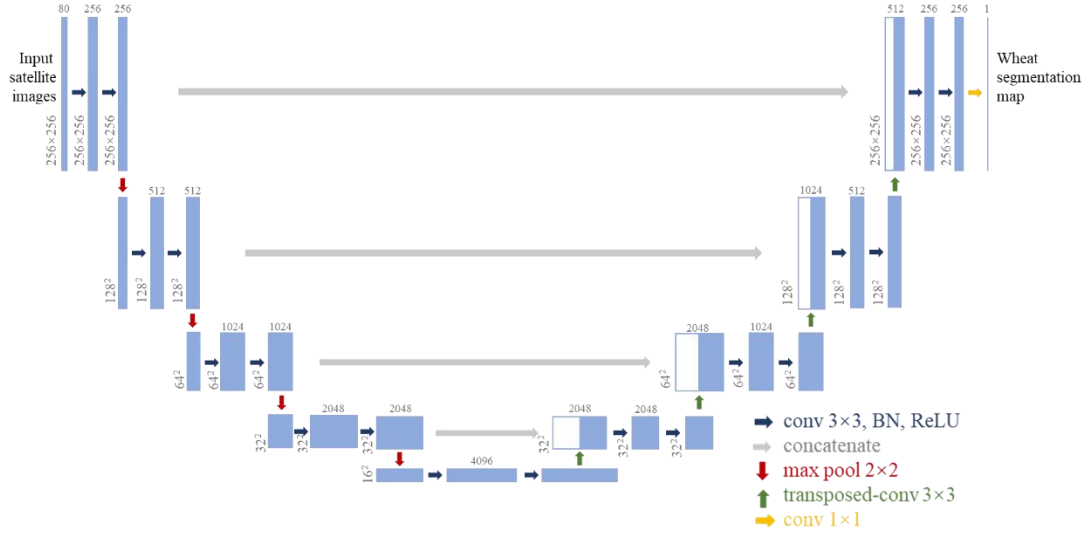


图 3.5 应用于本项目中的 U-net 模型

模仿经典 U-net 网络的架构，我适配本项目的数据构造了一个类似的模型。在经典 U-net 基础上，我增加了特征通道数，从原先的 64 个初始特征通道增加到了 256 个。这是因为经典 U-net 模型是应用于特征维度为 1 的灰度图像中的，而目前使用的卫星图像数据集有多达 80 个输入特征，所以需要更多的初始特征通道。在原先的 U-net 模型中，所有的卷积层都没有padding操作，为了提高边缘信息的提取能力，但是这也会对图片裁剪提出更高的要求。在本次项目中，我给模型的卷积层都加上了相应的 padding 操作，保证其输入输出图像尺寸的一致性，也为后续模型评估提供便利。

### 3.3.2 模型构建

在实验中，该模型也利用 Pytorch 构建，输入模型时同样要对图片进行变形，将原先的  $(n, h, w, t, c)$  的图片向量变形为  $(n, t \times c, h, w)$  的向量输入。

## 3.4 CNN+RNN 类：ConvLSTM 模型 (WheatMappingModel)

ConvLSTM<sup>[37]</sup> 最初被提出是应用于雷达图像降雨预测领域，它具有同时处理时序图像中的空间和时间信息，之后被广泛地应用于视频理解，气象、交通预测等领域。在ConvLSTM之后也有很多对它进行优化的模型被提出，如PredRNN<sup>[38-40]</sup>，MIM<sup>[41]</sup>，self attention ConvLSTM<sup>[42]</sup>等。在本项目中，我构建了

一个以经典 ConvLSTM 和 U-net 相结合的模型，将其命名为 WheatMappingModel (WMM)。

### 3.4.1 模型简介

WheatMappingModel(WMM) 模型的输入是预处理后的遥感图像，输出是图像每个像素是否种植冬小麦，输出1代表种植的是冬小麦，输出0代表种植的不是冬小麦。总的来说，模型分为六个模块：输入模块，下采样（Down Sample）模块、BConvLSTM 连接模块、Dense Block 模块、上采样（Up Sample）模块和输出模块。

#### 3.4.1.1 输入模块

输入模型的图像维度为 (batchsize, time, channels, height, width)，在进入输入模块后，要将图像变形为(batchsize, time\*channels, height, width) 以便于后续进行二维卷积操作。

输入模块由两个连续的3\*3卷积构成，并通过这两个卷积层将特征通道数增加到140个。

#### 3.4.1.2 下采样（Down Sample）模块

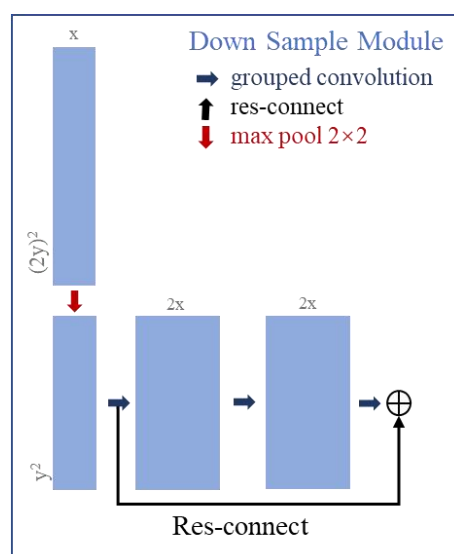


图 3.6 下采样模块示意图

下采样的主要目的是减小图像分辨率，增加特征通道，从而提取更多更高维度的特征。

如图所示，该模块主要涵盖了三种不同的运算。第一种是 **max pooling**，它是一种常见的池化手段，用于减少图像的分辨率。

第二种是 **grouped convolution**，即将 **channels** 分组分别进行卷积操作，这样能够增加特征通道的同时，不破坏时间维度的信息。

第三种是 **res-connection**，来源于 **ResNet**<sup>[43]</sup> 的一种操作，通过 **res-connection** 将输入输出相加，能避免梯度消失、爆炸，更好地保持梯度

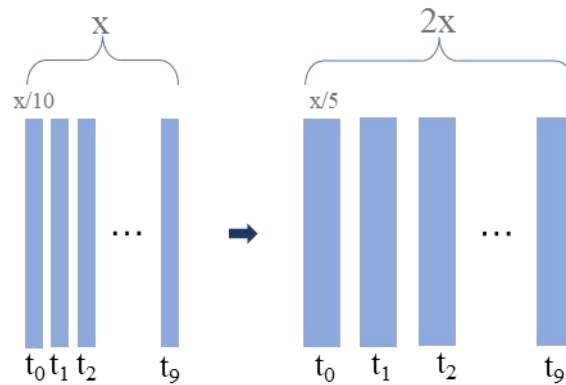


图 3.7 Grouped convolution 示意图

### 3.4.1.3 BiConvLSTM 连接模块

**ConvLSTM**: LSTM 在各个状态之间使用了全连接的设计，输入向量也是一维的时间序列，无法体现空间上的关联性。而 **ConvLSTM** 中所有的输入  $X_1, X_2, \dots, X_t$ ，单元输出  $C_1, C_2, \dots, C_t$ ，隐藏状态  $H_1, H_2, \dots, H_t$  都是三维向量，在输入到状态以及状态之间的转换中都使用了卷积操作。这种设计使得 **ConvLSTM** 网络中某个单元的未来状态，取决于其在图像中相邻单元的输入和其本身的过去状态，这种方式使得 **ConvLSTM** 能同时提取时间和空间的信息，下面是 **ConvLSTM** 中的一些重要方程（其中“\*”表示卷积算符，而“ $\circ$ ”像之前一样表示 Hadamard 乘积）：



$$\begin{aligned}
 i_t &= \sigma(W_{xi} * \mathcal{X}_t + W_{hi} * \mathcal{H}_{t-1} + W_{ci} \odot C_{t-1} + b_i) \\
 f_t &= \sigma(W_{xf} * \mathcal{X}_t + W_{hf} * \mathcal{H}_{t-1} + W_{cf} \odot C_{t-1} + b_f) \\
 C_t &= f_t \odot C_{t-1} + i_t \odot \tanh(W_{xc} * \mathcal{X}_t + W_{hc} * \mathcal{H}_{t-1} + b_c) \\
 o_t &= \sigma(W_{xo} * \mathcal{X}_t + W_{ho} * \mathcal{H}_{t-1} + W_{co} \odot C_t + b_o) \\
 \mathcal{H}_t &= o_t \odot \tanh(C_t)
 \end{aligned}$$

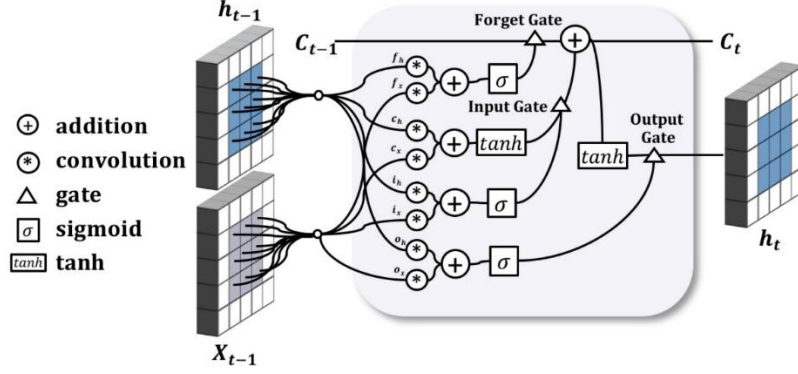


图 3.8 Conv-LSTM 单元示意图

在本模型中，我使用了双向 ConvLSTM 模块来取代原先U-net中的简单复制操作。通过双向 ConvLSTM 将模型 encoder 部分信息传递到 decoder。

#### 3.4.1.4 Dense Block 模块

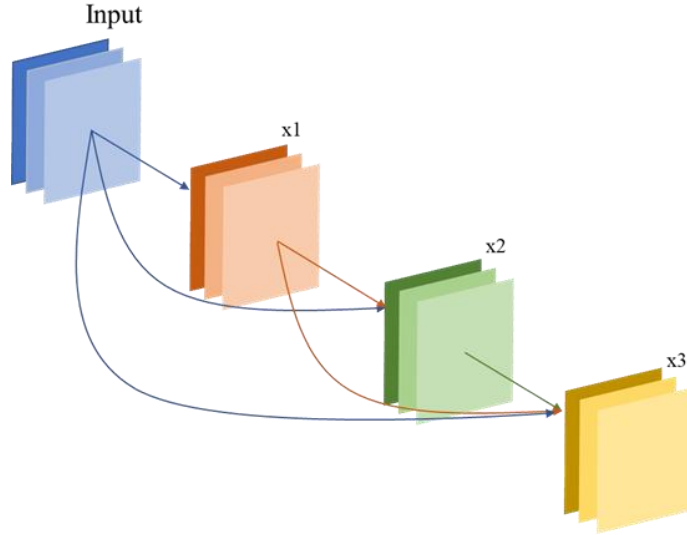


图 3.9 模型中使用的三层 Dense Block 示意图

该模块使用了 DenseNet<sup>[44]</sup> 中的 Dense Block。它的想法与 ResNet 中 res-connection 类似，加深网络的同时减少梯度消失的问题。在保证网络中层与层之间最大程度的信息传输的前提下，直接将所有层连接起来。为了能够保证



前馈的特性，每一层将之前所有层的输入进行拼接，之后将输出的特征图传递给之后的所有层。

经过调试，该模型使用了三层 Dense Block。

#### 3.4.1.5 上采样模块

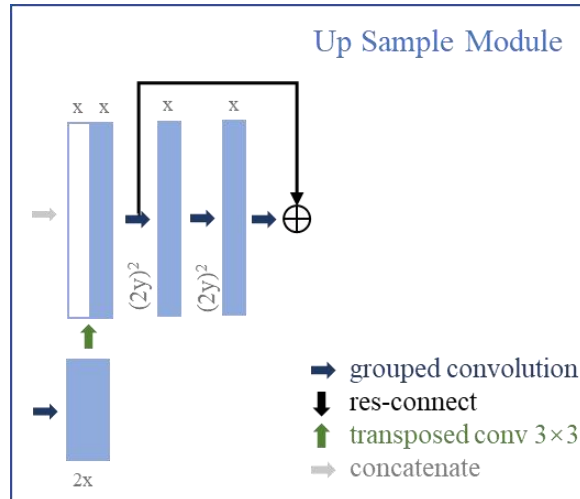


图 3.10 上采样模块示意图

与下采样模块类似，上采样也是用了 res-connect 和 grouped convolution 操作。与下采样模块相反的是，上采样模块采用了 transposed conv 来减小特征维度的同时，增加图像分辨率，起到解码的作用。

此外，上采样模块还会将 BiConvLSTM 连接模块的输出连到输入中，增加特征维度。

#### 3.4.1.6 输出模块

遵照经典 U-net 的结构，输出模块也没有采用全连接层，而是使用了  $1 \times 1$  的卷积将向量特征维度降为一维。

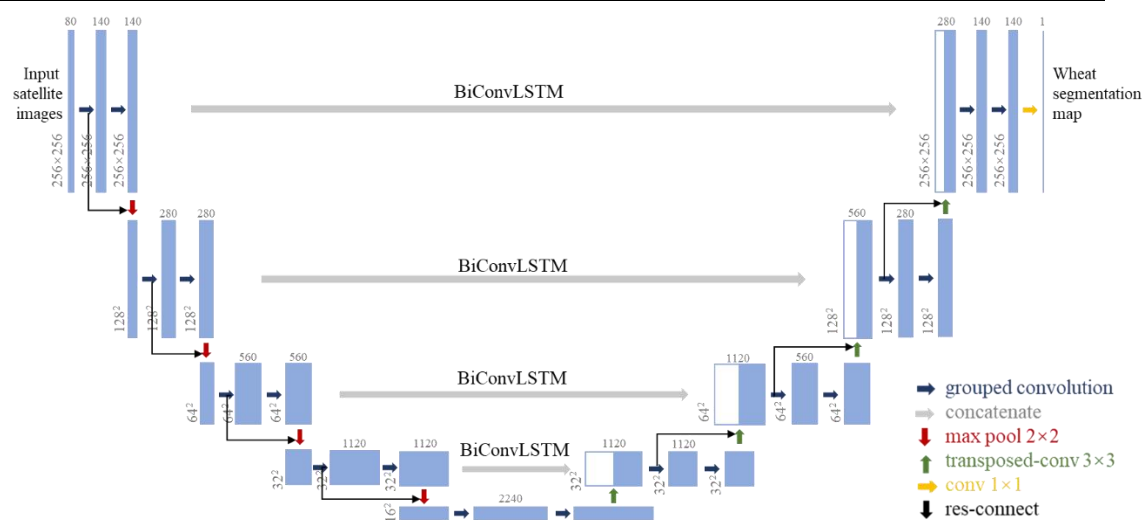


图 3.11 WMM 总体模型图

## 4 实验设计

### 4.1 实验环境

CPU: AMD EPYC™ 7003 系列14核处理器

GPU: NVIDIA Quadro RTX A6000 48GB GDDR6 Graphics Card

### 4.2 模型性能测试

模型精度评价首先要绘制混淆矩阵，然后根据混淆矩阵，我计算了 Accuracy 和 F1 两种评估指标。F1 和 Accuracy 的计算公式如下。

		Predicted Label	
		Not Wheat	Is Wheat
True Label	Not Wheat	True Negative(TN)	False Positive(FP)
	Is Wheat	False Negative(FN)	True Positive(TP)

图 4.1 混淆矩阵示意图

$$Accuracy = \frac{TN + TP}{TN + TP + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

其中 Accuracy 可以较好地衡量总的正确率，但是数据集中部分地点样本并不平衡，非小麦的像素个数比是小麦的像素个数多，在这个情况下 Accuracy 并不是一个很好的指标。因而要引入 F1 指标作为补充，F1 指标会更加考量正样本（小麦）的正确率。

#### 4.2.1 本地测试

本地测试中，将选取选定地点的 2017-2019 年度数据作为模型训练集，2020年的数据作为模型测试集。

此实验的目的是测试模型在本地的精度以及在时间维度的迁移能力。

### 4.2.2 特征提取可视化

t-Distributed Stochastic Neighbor Embedding (t-SNE)<sup>[45]</sup> 是一种无监督的非线性技术，主要用于数据探索和高维数据的可视化。t-SNE 的主要优点是能够保持局部结构，这意味着在高位空间中比较接近的点在 t-SNE 降维之后也会彼此接近。所以利用 t-SNE 我们可以把高维的数据降维并可视化。

在本实验中，我从训练集中随机选取6000个冬小麦样本点和6000个非冬小麦样本点，将其原先的80维降维到2维，绘制图像。再用模型预测这些样本点，获得模型输出层前隐藏层，该层即为模型提取的信息，将其使用 t-SNE 降维并可视化。

## 4.3 数据特征测试

在本文第二章提到的，本项目的数据集融合了两种不同类型的卫星数据，此实验的目的是探究多源数据能够帮助模型取得更高精度。

该测试将数据输入分为三类，仅 Sentinel 卫星数据、仅 Landsat 卫星数据和两种数据融合，考量不同模型在不同地点使用三种数据进行训练的模型精度。

## 5 实验结果与分析

### 5.1 模型性能测试

#### 5.1.1 本地测试

表 5.1 模型 Accuracy

地点 模型	1	3	5	7	9
RandomForest	0.8666	0.8120	0.9077	0.8577	0.8728
U-net	0.8574	0.8250	0.9714	0.8708	0.8633
DCM	0.8480	0.7535	0.9438	0.8511	0.8418
WMM	0.8620	0.8436	0.9630	0.8650	0.8753

表 5.2 模型 F1 得分

地点 模型	1	3	5	7	9
RandomForest	0.8364	0.7969	0.8443	0.8421	0.8604
U-net	0.823	0.7796	0.9443	0.8599	0.8525
DCM	0.8146	0.7403	0.8942	0.8376	0.8286
WMM	0.8244	0.8090	0.9298	0.8626	0.8627

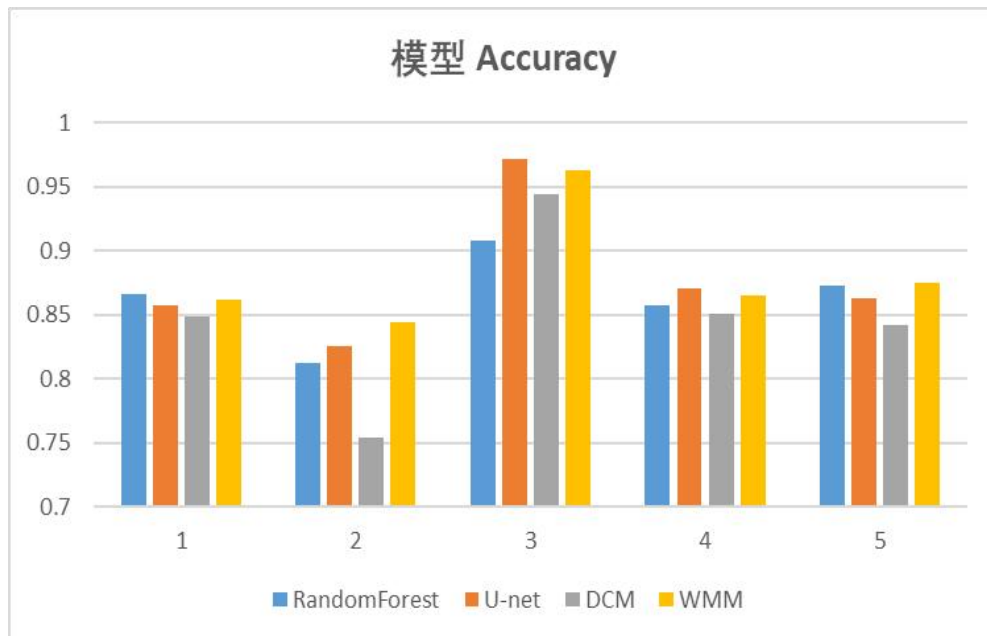


图 5.1 模型 Accuracy 统计图

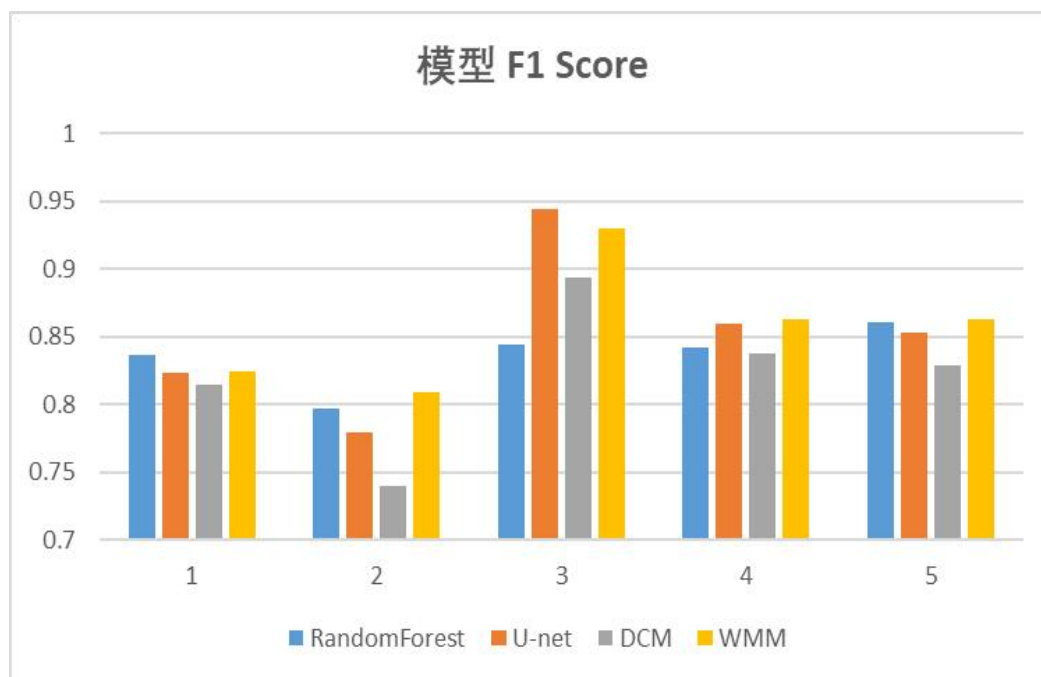


图 5.2 模型 F1 统计图

因为数据集样本点较多，所以随机选取了5个样本点进行测试，结果如上表所示。可以看出，U-net 模型和 WMM 模型在本数据集上效果较好。

### 5.1.2 特征提取可视化

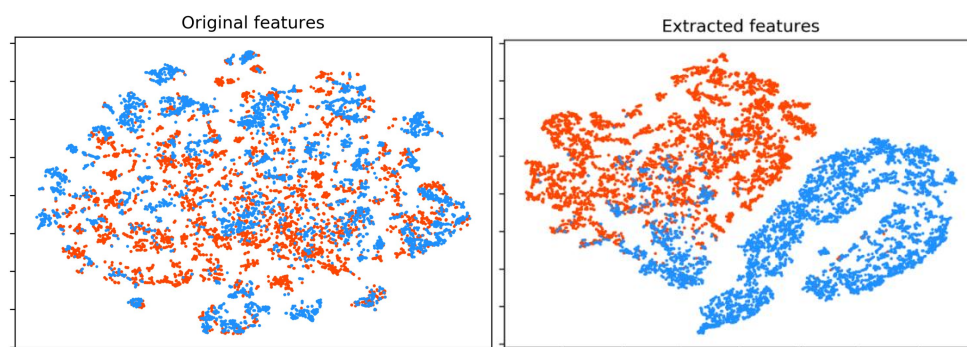


图 5.3 输入模型前后特征变化（红色为非小麦，蓝色为小麦）

上图为对原始特征和模型提取特征进行降维的结果，可以看出提取后非小麦和小麦像素的特征分离较为明显，体现出模型较强的特征提取能力。

## 5.2 数据特征测试

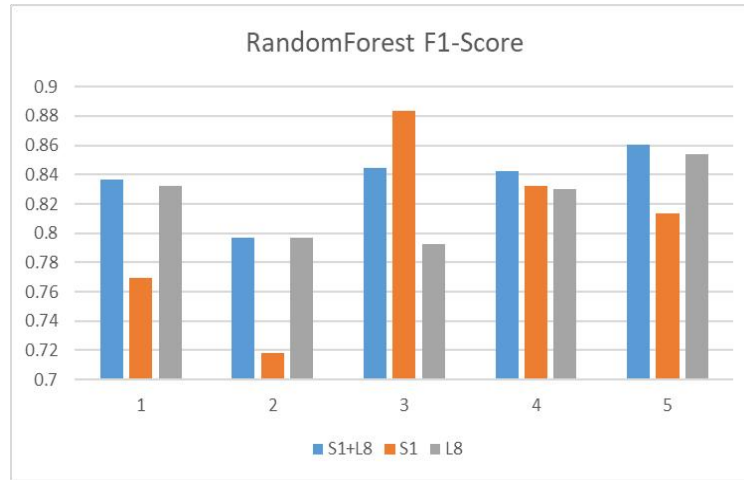


图 5.4 不同数据输入下，RandomForest F1 得分

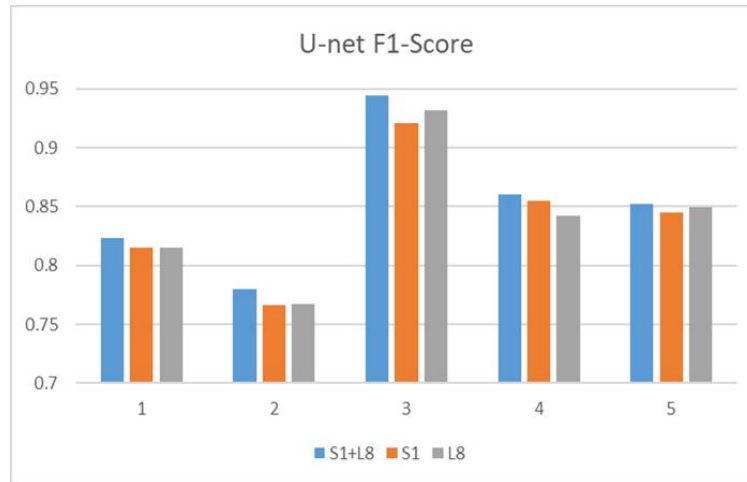


图 5.5 不同数据输入下，U-net F1 得分

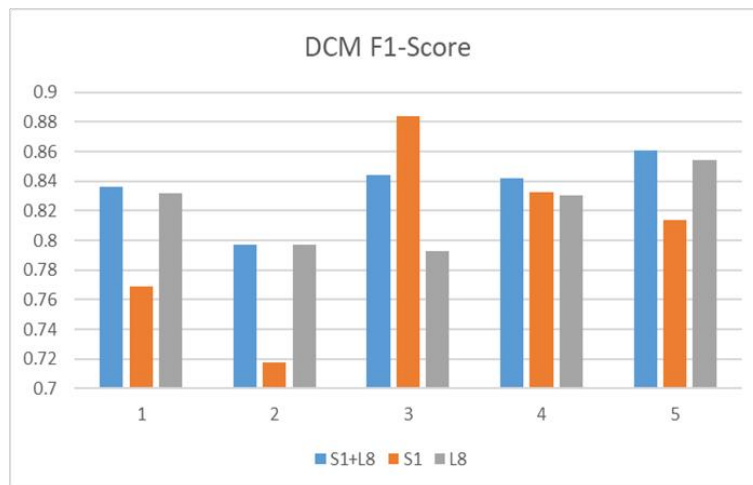


图 5.6 不同数据输入下 DCM F1 得分

可以看出，在绝大部分情况下，使用多源数据融合（Sentinel 加 Landsat）都比仅使用一种数据有利于模型训练。



## 6 总结与展望

本项目构建了一个涵盖美国中部冬小麦种植区的10个不同地点的用于冬小麦制图的时间序列数据集。该数据集可以为之后研究者做同类研究提供样本。同时基于该数据集，项目也构建了4种不同类型的模型，其中 WMM 模型是在本项目中被第一次提出，也经过测试证明在本数据集上有比较出色的表现。此外，本项目也分析了并证明了多源数据融合对于模型训练的积极影响。

对于未来的研究，我认为主要的改进方向还会是在数据源和模型方面。在数据源方面，现在较多的研究使用的是中分辨率卫星数据MODIS，为了进行更准确的小麦制图，使用更高分辨率的卫星数据是未来的趋势。Landsat卫星、Sentinel卫星都提供免费的 10m 分辨率的卫星数据。除此之外，多源数据融合也是未来研究的方向之一，随着模型性能越发强大，需要有更丰富的数据来源提供更丰富的特征信息。

在模型方面，随着深度学习的发展，可用的模型越来越多，也就有了更多可能的架构，不同模型架构的探究和比较也是未来研究的重中之重。

在研究地点方面，现在大部分使用有监督学习算法的研究者都倾向于选择美国作为数据来源。因为美国有CDL标签数据以供分类。作为全球最大的冬小麦生产国之一，中国的冬小麦产量占全球的18%以上。然而，中国却没有大空间范围内具有高空间分辨率的小麦分布图<sup>[5]</sup>。所以未来的研究也会很大部分集中于如何为这些原本没有小麦分布图的国家 and 地区制作小麦分布图。

## 参考文献

- [1] ZHOU T, PAN J, ZHANG P, et al. Mapping winter wheat with multi-temporal SAR and optical images in an urban agricultural region [J]. *Sensors*, 2017, 17(6): 1210.
- [2] QIU B, LUO Y, TANG Z, et al. Winter wheat mapping combining variations before and after estimated heading dates [J]. *ISPRS journal of photogrammetry and remote sensing*, 2017, 123: 35-46.
- [3] LI F, REN J, WU S, et al. Comparison of regional winter wheat mapping results from different similarity measurement indicators of NDVI time series and their optimized thresholds [J]. *Remote Sensing*, 2021, 13(6): 1162.
- [4] DONG Q, CHEN X, CHEN J, et al. Mapping winter wheat in North China using Sentinel 2A/B data: A method based on phenology-time weighted dynamic time warping [J]. *Remote Sensing*, 2020, 12(8): 1274.
- [5] DONG J, FU Y, WANG J, et al. Early-season mapping of winter wheat in China based on Landsat and Sentinel images [J]. *Earth System Science Data*, 2020, 12(4): 3081-95.
- [6] ZHANG D, FANG S, SHE B, et al. Winter wheat mapping based on Sentinel-2 Data in heterogeneous planting conditions [J]. *Remote Sensing*, 2019, 11(22): 2647.
- [7] LIU J, FENG Q, GONG J, et al. Winter wheat mapping using a random forest classifier combined with multi-temporal and multi-sensor data [J]. *International journal of digital earth*, 2018, 11(8): 783-802.
- [8] ZHONG L, HU L, ZHOU H, et al. Deep learning based winter wheat mapping using statistical data as ground references in Kansas and northern Texas, US [J]. *Remote Sensing of Environment*, 2019, 233: 111411.
- [9] KUSSL N, LAVRENIUK M, SKAKUN S, et al. Deep learning classification of land cover and crop types using remote sensing data [J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(5): 778-82.
- [10] ZHONG L, HU L, ZHOU H. Deep learning based multi-temporal crop classification [J]. *Remote sensing of environment*, 2019, 221: 430-43.
- [11] MENG S, WANG X, HU X, et al. Deep learning-based crop mapping in the cloudy season using one-shot hyperspectral satellite imagery [J]. *Computers and Electronics in Agriculture*, 2021, 186: 106188.
- [12] XU J, ZHU Y, ZHONG R, et al. DeepCropMapping: A multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping [J]. *Remote Sensing of Environment*, 2020, 247: 111946.
- [13] M RUSTOWICZ R, CHEONG R, WANG L, et al. Semantic segmentation of crop type in Africa: A novel dataset and analysis of deep learning methods; proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, F, 2019 [C].
- [14] ZHAO H, DUAN S, LIU J, et al. Evaluation of Five Deep Learning Models for Crop Type Mapping Using Sentinel-2 Time Series Images with Missing Information [J]. *Remote Sensing*, 2021, 13(14): 2790.
- [15] TURKOGLU M O, D'ARONCO S, PERICH G, et al. Crop mapping from image time series: Deep learning with multi-scale label hierarchies [J]. *Remote Sensing of Environment*, 2021, 264: 112603.
- [16] WEI P, CHAI D, LIN T, et al. Large-scale rice mapping under different years based on

- time-series Sentinel-1 images using deep semantic segmentation model [J]. ISPRS journal of photogrammetry and remote sensing, 2021, 174: 198-214.
- [17] DU Z, YANG J, OU C, et al. Smallholder crop area mapped with a semantic segmentation deep learning method [J]. Remote Sensing, 2019, 11(7): 888.
- [18] ORYNBAIKYZY A, GESSNER U, CONRAD C. Crop type classification using a combination of optical and radar remote sensing data: A review [J]. international journal of remote sensing, 2019, 40(17): 6553-95.
- [19] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. nature, 2015, 521(7553): 436-44.
- [20] HEAVEN D. Why deep-learning AIs are so easy to fool [J]. Nature, 2019, 574(7777): 163-6.
- [21] XU J, YANG J, XIONG X, et al. Towards interpreting multi-temporal deep learning models in crop mapping [J]. Remote Sensing of Environment, 2021, 264: 112599.
- [22] <https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-1/overview> [Z].
- [23] <https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-1-sar/product-overview/polarimetry> [Z].
- [24] [https://en.wikipedia.org/wiki/Landsat\\_program](https://en.wikipedia.org/wiki/Landsat_program) [Z].
- [25] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [J]. Advances in neural information processing systems, 2012, 25.
- [26] HO T K. Random decision forests; proceedings of the Proceedings of 3rd international conference on document analysis and recognition, F, 1995 [C]. IEEE.
- [27] BELGIU M, DRĂGUȚ L. Random forest in remote sensing: A review of applications and future directions [J]. ISPRS journal of photogrammetry and remote sensing, 2016, 114: 24-31.
- [28] <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html> [Z].
- [29] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural computation, 1997, 9(8): 1735-80.
- [30] <https://medium.com/analytics-vidhya/lstms-explained-a-complete-technically-accurate-conceptual-guide-with-keras-2a650327e8f2> [Z].
- [31] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation; proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, F, 2015 [C].
- [32] BADRINARAYANAN V, KENDALL A, CIPOLLA R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(12): 2481-95.
- [33] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation; proceedings of the International Conference on Medical image computing and computer-assisted intervention, F, 2015 [C]. Springer.
- [34] ZHOU B, ZHAO H, PUIG X, et al. Scene parsing through ade20k dataset; proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, F, 2017 [C].
- [35] LIU Z, HU H, LIN Y, et al. Swin Transformer V2: Scaling Up Capacity and Resolution [J]. arXiv preprint arXiv:211109883, 2021.
- [36] XIAO T, LIU Y, ZHOU B, et al. Unified perceptual parsing for scene understanding; proceedings of the Proceedings of the European Conference on Computer Vision (ECCV), F, 2018 [C].

- [37] SHI X, CHEN Z, WANG H, et al. Convolutional LSTM network: A machine learning approach for precipitation nowcasting [J]. Advances in neural information processing systems, 2015, 28.
- [38] WANG Y, LONG M, WANG J, et al. Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms [J]. Advances in neural information processing systems, 2017, 30.
- [39] WANG Y, WU H, ZHANG J, et al. PredRNN: A recurrent neural network for spatiotemporal predictive learning [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022.
- [40] WANG Y, GAO Z, LONG M, et al. Predrnn++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning; proceedings of the International Conference on Machine Learning, F, 2018 [C]. PMLR.
- [41] WANG Y, ZHANG J, ZHU H, et al. Memory in memory: A predictive neural network for learning higher-order non-stationarity from spatiotemporal dynamics; proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, F, 2019 [C].
- [42] LIN Z, LI M, ZHENG Z, et al. Self-attention convlstm for spatiotemporal prediction; proceedings of the Proceedings of the AAAI Conference on Artificial Intelligence, F, 2020 [C].
- [43] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition; proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, F, 2016 [C].
- [44] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks; proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, F, 2017 [C].
- [45] VAN DER MAATEN L, HINTON G. Visualizing data using t-SNE [J]. Journal of machine learning research, 2008, 9(11).

## 附录

代码（updating）：<https://github.com/CaoXutao7/Wheat-Mapping.git>

## 作者简历

姓名：曹徐韬 性别：男 民族：汉

出生年月：2000-07-25 籍贯：浙江省绍兴市

2012.09-2015.07 杭州文澜中学

2015.09-2018.07 浙江省杭州第二中学

2018.09-2022.07 浙江大学攻读学士学位

2021.08-2023.06 美国 University of Illinois at Urbana Champaign 攻读硕士学位

获奖情况：浙江大学二等奖学金，优秀学生

参加项目：

发表的学术论文：暂无

# 本科生毕业论文（设计）任务书

一、题目：

二、指导教师对毕业论文（设计）的进度安排及任务要求

（此页单面打印）

起讫日期 20 年 月 日至 20 年 月 日

指导教师（签名）\_\_\_\_\_ 职称\_\_\_\_\_

三、系或研究所审核意见

负责人（签名）\_\_\_\_\_

年 月 日





# 毕 业 论 文（设计） 考 核

## 一、指导教师对毕业论文（设计）的评语

（此页单面打印）

指导教师(签名) \_\_\_\_\_

年 月 日

## 二、答辩小组对毕业论文（设计）的答辩评语及总评成绩

成绩 比例	文献综述 占（10%）	开题报告 占（15%）	外文翻译 占（5%）	毕业论文(设计) 质量及答辩 占（70%）	总评成绩
分值					

答辩小组负责人（签名） \_\_\_\_\_

年 月 日



# 第二部分

## 文献综述和开题报告



使用《浙江大学本科生毕业论文（设计）文献综述和开题报告（模板）》