

基于深度学习的大棚及地膜农田无人机航拍监测方法

孙 钰<sup>1</sup> 韩京冶<sup>1</sup> 陈志泊<sup>1</sup> 史明昌<sup>2</sup> 付红萍<sup>1</sup> 杨 猛<sup>1</sup>

(1. 北京林业大学信息学院, 北京 100083; 2. 北京林业大学水土保持学院, 北京 100083)

**摘要:** 随着精准农业技术的发展,快速获取大棚和地膜农田面积及地理分布的需求越来越大,但沿用面向卫星遥感影像的解译方法处理无人机航拍影像,存在特征选择复杂、识别精度较低、处理时间长等问题。基于此,本文提出一种基于深度学习的大棚及地膜农田无人机航拍监测方法,即采用六旋翼无人机搭载索尼 NEX-5k 相机进行航拍作业,对采集到的 558 幅赤峰市王爷府镇地区的无人机航片进行正射校正与拼接,构建全卷积神经网络(Fully convolutional network, FCN),通过多尺度融合的方法实现了 FCN 的 5 个变种模型:FCN-32s、FCN-16s、FCN-8s、FCN-4s、FCN-2s,使用带动量的随机梯度下降算法端到端训练模型,自动提取并分类影像特征。FCN 模型与 ENVI 商用遥感软件的基于像素的分类方法、eCognition 软件的面向对象的分类方法对比后表明:FCN-4s 模型为识别大棚和地膜农田的最佳模型,对于测试区域的平均整体正确率为 97%,而基于像素的分类方法平均整体正确率为 74.1%,面向对象的分类方法平均整体正确率为 81.78%。FCN-4s 模型平均运行时间为 16.85 s,是基于像素的分类方法运行时间的 0.06%,是面向对象的分类方法运行时间的 5.62%。本方法可快速准确获取大棚和地膜农田的地理分布及面积,满足设施农业对无人机航拍监测的需求。

**关键词:** 农业监测; 无人机; 深度学习; 语义分割; 全卷积神经网络

**中图分类号:** TP391.41      **文献标识码:** A      **文章编号:** 1000-1298(2018)02-0133-08

Monitoring Method for UAV Image of Greenhouse and Plastic-mulched Landcover Based on Deep Learning

SUN Yu<sup>1</sup> HAN Jingye<sup>1</sup> CHEN Zhibo<sup>1</sup> SHI Mingchang<sup>2</sup> FU Hongping<sup>1</sup> YANG Meng<sup>1</sup>

(1. School of Information Science and Technology, Beijing Forestry University, Beijing 100083, China

2. School of Soil and Water Conservation, Beijing Forestry University, Beijing 100083, China)

**Abstract:** With the development of precision agriculture, the demand on rapidly obtaining the area and geographical distribution of greenhouses, plastic-mulched landcover is increased. However, using the interpretation method for satellite remote sensing images to process unmanned aerial vehicle (UAV) images is not ideal, due to the complex feature extraction, low recognition accuracy, long processing time and so on. To circumvent this issue, a UAV aerial monitoring method was proposed based on deep learning for greenhouses and plastic-mulched landcover monitoring. The six-rotor UAV equipped with Sony NEX-5k camera captured aerial photographs in the Wangyefu town of Chifeng City. The 558 UAV images were orthographically corrected and stitched. The five fully convolutional network (FCN) variants, i. e. the FCN-32s, FCN-16s, FCN-8s, FCN-4s and FCN-2s models were built by multi-scale fusion. The modes were trained end-to-end by the stochastic gradient descent algorithm with momentum. The features were extracted and learned from the photographs automatically. The FCN models were compared with two economic softwares, i. e. the pixel-based classification method of ENVI and the object-oriented classification method of eCognition. The results showed that the FCN-4s was the best model on the identification of greenhouses and plastic-mulched landcover. The average overall accuracy of test area was 97%, while that of pixel-based classification method and the object-oriented classification method was 74.1% and 81.78%, respectively. The average runtime of the FCN-4s was 16.85 s, which was 0.06% and 5.62% of those of pixel-based classification method and the object-

收稿日期: 2017-07-07    修回日期: 2017-10-13

**基金项目:** 中央高校基本科研业务费专项资金交叉创新项目(2017JC02)和国家自然科学基金项目(61402038)

**作者简介:** 孙钰(1985—),男,副教授,主要从事林业物联网研究,E-mail: sunyv@bjfu.edu.cn

**通信作者:** 陈志泊(1967—),男,教授,博士生导师,主要从事数据库和农业及林业信息工程研究,E-mail: zhibo@bjfu.edu.cn

oriented classification method, respectively. The proposed method demonstrated high recognition accuracy and fast speed, which can meet the demand on UAV monitoring of facilities agriculture.

**Key words:** agricultural monitoring; unmanned aerial vehicle; deep learning; semantic segmentation; fully convolutional network

0 引言

近年来,随着农业现代化的快速发展,设施农业技术逐渐成熟,我国设施农业面积迅速扩大,已成为全球设施农业大国。农业大棚作为新兴的农业设施,具有价格低廉、抵御病虫害、降低气候环境的影响、提高单位面积产量等突出优点<sup>[1-2]</sup>。同样,地膜作为现代农业重要的生产资料,其蓄水、保温、保墒、保药等作用显著,利于作物早熟,提高产量和品质<sup>[2]</sup>。2014 年我国大棚和地膜覆盖面积近 20 万 km<sup>2</sup><sup>[3]</sup>,大棚和地膜农田已经成为重要的土地覆盖类型(地类)。

传统的大棚和地膜覆盖面积信息获取以人工实地测定为主,该方式需要耗费大量的人力、物力,且无法获得准确地理分布情况。准确获取大棚和地膜覆盖面积对于农作物监测、农业产值估算以及政府合理发放补贴具有重要的意义。

随着遥感技术的发展,基于卫星遥感影像的大棚和地膜农田监测方法成为研究热点,AGUERA 等<sup>[4]</sup>提出一种利用纹理分析提高大棚像素分类精度的方法。黄振国等<sup>[5]</sup>针对山东寿光大棚菜地,提出了一种基于 HJ-1 影像的知识分类法,该方法综合了遥感影像的光谱特征、空间结构特征、地物分布规律以及来自专家的经验,其分类总体精度为 92.01%。王志盼等<sup>[6]</sup>基于 Landsat 影像提出一种基于增强型水体指数的大棚遥感检测方法,分类总体精度在 75% 以上。沙先丽<sup>[7]</sup>提出一种决策树模型,实现了地膜农田信息的提取方法,分类总体精度达到 90% 以上。尽管方法众多,但卫星遥感影像价格昂贵、获取周期时间长、容易受到自然环境(如云层)影响、分辨率较低,进而影响大棚和地膜农田的识别精度,难以满足小区域、精度要求高的农业监测需求。

无人机(Unmanned aerial vehicle, UAV)是利用无线电遥控设备和自备的程序控制装置操纵的无人驾驶飞机<sup>[8]</sup>,具有机动性强、获取影像分辨率高、时效性高、成本低的优点,越来越多的学者将无人机用于资源调查与监测领域<sup>[9-12]</sup>。这些基于无人机影像的监测方法,大部分依旧沿用传统卫星遥感解译方法,人工选取特征参数如光谱、纹理、形状等进行分类,不仅需要领域专家的先验知识,而且工程量

大。自 2012 年 KRIZHEVSKY 等<sup>[13]</sup>利用深度学习技术刷新了 ImageNet 大规模视觉识别竞赛的世界纪录以来,深度学习技术在图像分类、图像语义分割等方面已经取得了令人瞩目的成就。图像语义分割是将像素按照图像中表达的含义进行分组,属于像素级别的分类。利用卷积神经网络(Convolution neural network, CNN)的分割方法<sup>[14]</sup>,在 PASCAL VOC 2012 图像分割数据集上的 mIoU(Mean intersection-over-union)达到 51.6%,达到当时最佳的性能。但 CNN 对于小目标识别精度不够且计算效率低下。基于全卷积神经网络的图像语义分割方法<sup>[15]</sup>,将 mIoU 提高到 62.2%。文献[16]将全连接的条件随机场(Conditional random field, CRF)作为 FCN 的后处理的方法,将 mIoU 提高到 66.4%。文献[17]实现端到端训练 FCN 和 CRF 的分割方法,将 mIoU 提高到 74.7%。但 CRF 具有训练代价大、复杂度高的缺点,对于高分辨率的无人机影像,存在计算成本大、处理耗时长等问题。

针对现有无人机航拍监测研究存在特征选择复杂、识别精度较低、处理时间较长等不足,本文基于深度学习图像语义分割算法,提出一种端到端的无人机影像大棚和地膜农田识别方法。选用赤峰市王爺府镇大棚和地膜农田无人机航拍影像,其中大棚与该区域的广场石砖颜色相近,地膜农田间隙中的土地与普通耕地特征相近,同时影像分辨率高、细节丰富、数据量大,使研究具有典型性和代表性。针对数据集的特点,提出一套完整的无人机航拍监测方法,构建全卷积神经网络,从数据中自动学习特征,将像素按照影像中表达语义的不同进行分割,获得更为准确的大棚和地膜覆盖面积及地理分布信息。

1 数据材料

1.1 研究区概况

实验点位于内蒙古自治区赤峰市喀喇沁旗王爺府镇(41°50′34.28″N, 118°30′49.78″E),如图 1 所示。王爺府镇的土地总面积为 517 km<sup>2</sup>,其中耕地面积 47.42 km<sup>2</sup>。

1.2 数据获取与预处理

采用六旋翼无人机,搭载索尼 NEX-5k 相机,详细参数见表 1、2。无人机按照规划好的航线飞行,飞行高度 300 m,飞行速度 5 m/s,镜头垂直航

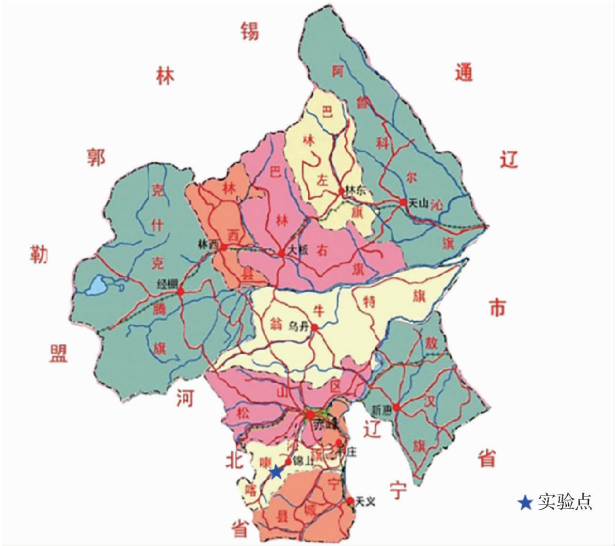


图 1 研究区位置

Fig. 1 Location of study area

表 1 无人机主要参数

Tab.1 Main specifications of UAV

| 参数            | 数值/类型       |
|---------------|-------------|
| 起飞质量/kg       | 2.5         |
| 长×宽×高/(m×m×m) | 0.7×0.7×0.5 |
| 最大飞行高度/m      | 1 000       |
| 最大飞行距离/m      | 3 000       |
| 起降方式          | 垂直起降        |

表 2 索尼 NEX-5k 相机主要参数

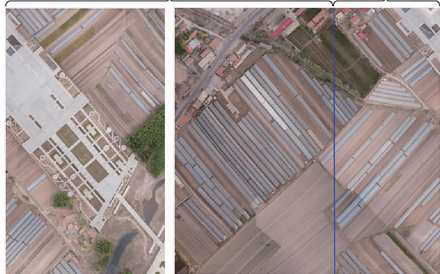
Tab.2 Main specifications of Sony NEX-5k camera

| 参数               | 数值/型号             |
|------------------|-------------------|
| 长×宽×高/(mm×mm×mm) | 110.8×58.8×38.2   |
| 质量/g             | 287               |
| 传感器              | Exmor APS HD CMOS |
| 最高分辨率/(像素×像素)    | 4 592×3 056       |
| 镜头型号             | Sigma 30mm/f 2.8  |

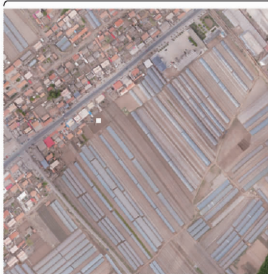
拍,航拍影像为 RGB 图像,其空间分辨率为 8 cm。

无人机航片通过商用无人机图像拼接软件 Pix4Dmapper 进行正射校正和拼接处理:首先导入筛选后的无人机航片并添加地理位置信息,通过空中三角测量法自动实现控制点加密,结合地面控制

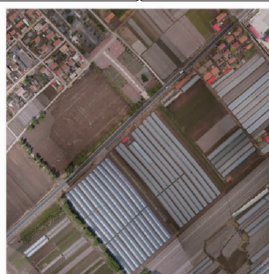
训练集                      验证集



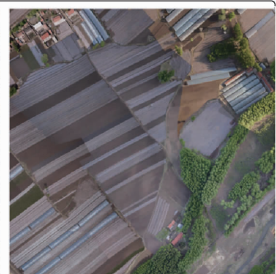
(a) 区域1                      (b) 区域2



(c) 测试区域1



(d) 测试区域2



(e) 测试区域3

图 2 无人机航拍影像

Fig. 2 UAV aerial images

点数据实现航片的几何校正,最终生成正射影像。本研究选择 5 块典型区域,区域 1 为 3 000 像素×5 000 像素(图 2a),覆盖面积为 0.096 km<sup>2</sup>。区域 2 为 5 000 像素×5 000 像素(图 2b),覆盖面积为 0.16 km<sup>2</sup>。测试区域 1、2、3 均为 5 000 像素×5 000 像素(图 2c、2d、2e),覆盖面积均为 0.16 km<sup>2</sup>,各区域均无交集。

## 2 识别方法

### 2.1 全卷积神经网络(FCN)

采用全卷积神经网络实现像素级别的分类。FCN 具有和 CNN 相同的构建块,即卷积层和下采样层。卷积层是深度神经网络的核心,用于提取图像的颜色、纹理、形状等特征,下采样层用于减小特征的大小并减少网络所需的计算量,也用于防止过拟合。目前 FCN 主要是基于现有卷积神经网络如 AlexNet<sup>[13]</sup>、VGGNet<sup>[18]</sup>、GoogleNet<sup>[19]</sup> 和 ResNet<sup>[20]</sup>, 然后进行微调操作来实现语义分割任务。

FCN 与 CNN 区别在于,CNN 在卷积层之后使用全连接层得到固定长度的特征向量进行分类,但这样会丢失空间信息;而 FCN 把全连接层替换成卷积层后,对最后一个卷积层的特征图进行上采样,使它恢复到输入图像相同的大小,从而每个像素都有一个预测值,并保留了原始输入图像中的空间信息,实现逐像素分类,如图 3 所示。

### 2.2 多尺度特征融合

由于高分辨率无人机影像具有很多细节信息,如低矮的农作物、车辆等。如果直接将全卷积后的结果进行反卷积,会丢失很多细节特征,得到的结果比较粗糙。因此为了提取影像中细节信息,本文在反卷积时采取融合一部分较浅层的细节特征信息的方法优化输出。如图 4 所示,考虑大棚和地膜农田无人机影像与 PASCAL VOC 2012 数据集有较大差异,因此在原有的 FCN-32s、FCN-16s、FCN-8s 3 种网络结构模型的基础上,融合了更浅层的特征图实现了 FCN-4s、FCN-2s 2 种网络结构模型,然

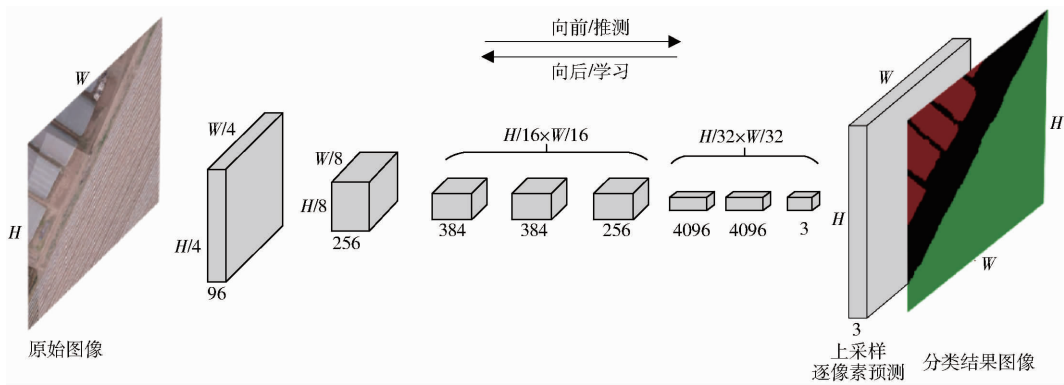


图3 FCN 基本结构

Fig.3 Basic structure of FCN

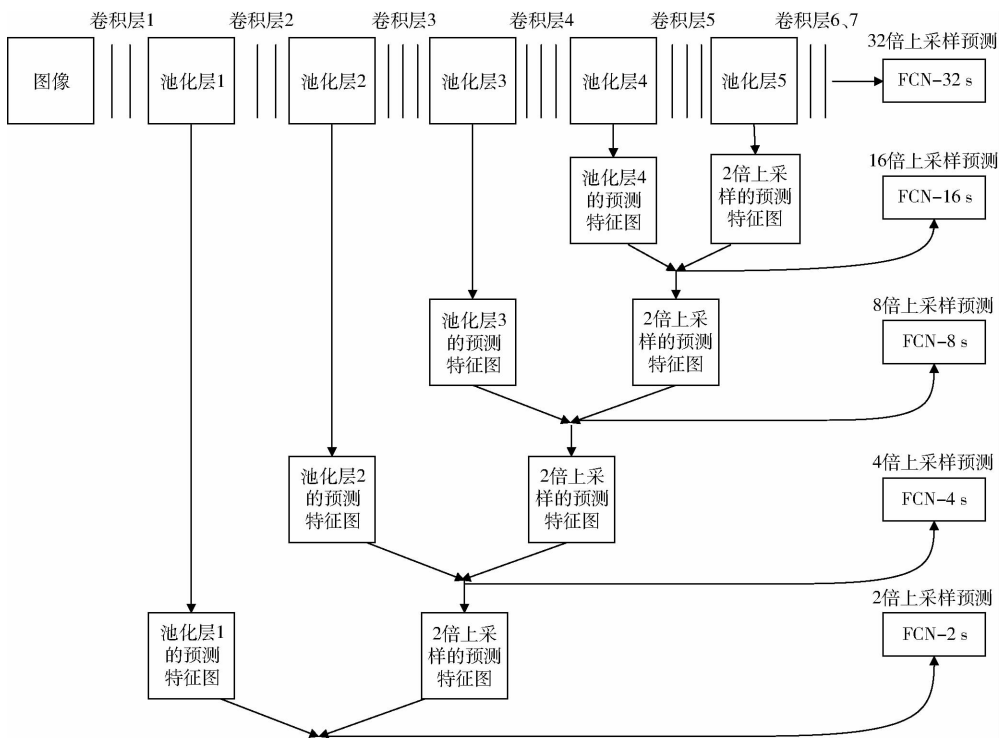


图4 多尺度特征融合示意图

Fig.4 Multi-scale feature illustration

后通过实验确定最适合大棚和地膜农田识别任务的FCN模型。

3 实验

分别对融合不同尺度细节特征的FCN-32s、FCN-16s、FCN-8s、FCN-4s、FCN-2s网络结构模型进行训练,完整实验流程如图5所示。

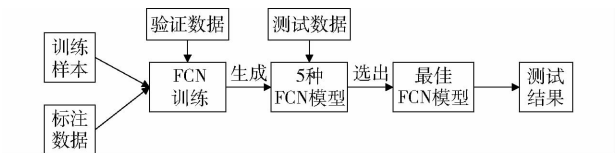


图5 实验流程图

Fig.5 Experiment flow chart

3.1 数据集划分与数据标注

考虑到实验运行时间和配合网络结构等因素,

将5块区域的影像分割成若干个500像素×500像素的图像,如图2所示,训练集为120幅,验证集为40幅,测试集为300幅。

由于无人机影像分辨率高、细节清晰,本文采用人工目视分类为主、地面调查的结果为辅的方式对影像进行像素级语义标注,作为实验分割标准。首先使用零参数的简单线性迭代聚类的超像素图像分割算法<sup>[21]</sup>,自适应地选择每个超像素的紧凑度参数,将无人机影像(图6a)分成10000块紧凑、均匀尺寸的超像素(图6b),得到的超像素边缘更加规整,提高了标注效果。通过标注每个超像素的类不仅能确定出整幅图的类别区域,而且会大幅度地提升标记速度。然后进行颜色填充,其中背景类的RGB值为(0,0,0),大棚类的RGB值为(128,0,0),地膜农田类的RGB值为(0,128,0),生成标注图像

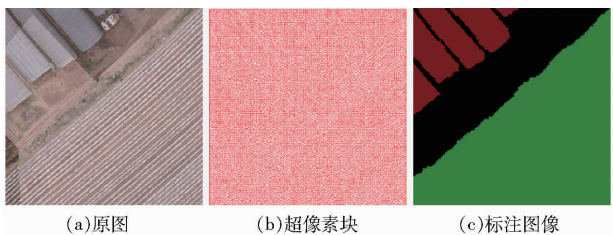


图 6 数据标注流程  
Fig. 6 Data annotation procedure

(图 6c)。

3.2 模型训练

本实验采取 Linux 下的 Chainer 深度学习平台<sup>[22]</sup>,考虑到无人机影像数据量大,计算时间长,因此采用 NVIDIA Titan X GPU 加速模型训练,其中 FCN 的初始模型为 VGG-16,使用在 ImageNet 数据集<sup>[23]</sup>上的预训练模型进行微调,选用随机梯度下降优化算法(Stochastic gradient descent,SGD),并引入动量进行优化,其中动量为 0.99,学习率为  $1 \times 10^{-10}$ 。原始的 SGD 算法将训练集分成多个批量,每次利用一个批量的样本数据更新参数,公式定义为

$$x_{t+1} = x_t + \Delta x_t \tag{1}$$

其中

$$\Delta x_t = -\eta g_t \tag{2}$$

式中, $\eta$  是学习率, $g_t$  是  $x$  在  $t$  时刻的梯度,利用一个批量数据更新可以减少计算量并提升收敛速率。但 SGD 最大的缺点在于每次更新依赖当前的批量,可能按照错误的方向进行。因此本文引入动量改进的 SGD 算法,更新时在一定程度上保留之前的更新方向,并微调当前批量梯度的更新方向,公式定义为

$$\Delta x = \rho \Delta x_{t-1} - \eta g_t \tag{3}$$

其中  $\rho$  为动量,表示在多大程度上保留原有的更新方向。因此带动量的 SGD 算法增加了稳定性、学习速度更快。本实验选用交叉熵作为损失函数,用于衡量神经网络的预测值与实际值差异,公式定义为

$$C = -\frac{1}{n} \sum_x [y \ln a + (1 - y) \ln (1 - a)] \tag{4}$$

其中

$$a = \sigma \left( \sum_i w_i x_i + b \right) \tag{5}$$

- 式中
- $n$ ——训练数据个数
  - $y$ ——期望的输出
  - $\sigma$ ——激活函数
  - $a$ ——神经元实际输出
  - $w_i$ ——权重向量
  - $x_i$ ——输入向量
  - $b$ ——偏置向量

4 结果与分析

为验证本文方法的有效性以及泛化能力,选择无交集的测试区域 1、2、3 作为测试集进行实验验证。对 5 种 FCN 网络结构模型进行训练和评估,并与目前广泛应用的 ENVI 软件和 eCognition 软件的

识别效果进行比较。同时采取整体正确率和混淆矩阵两种方法来评判识别精度效果。

4.1 大棚、地膜农田识别结果

利用 ENVI 软件进行基于像素的分类方法的对比实验,通过软件中 ROI 工具在测试区域影像上添加感兴趣区(Regions of interest, ROI),通过 ENVI 软件中分类器自带的支持向量机(Support vector machine, SVM)方法进行分类。然后利用 eCognition 软件进行面向对象的分类方法的对比实验,进行影像分割,并通过试误法确定多尺度分割的最佳参数:分割尺度为 200、形状权重为 0.4、紧致度权重为 0.5。影像中大棚、地膜农田在几何和纹理等特征方面与其它地物有明显区别,因此对分割后的影像进行光谱、纹理和形状的特征提取,其中光谱特征包括各波段均值、标准差、比率、亮度等信息,在高分辨率无人机影像中能够提取丰富纹理特征,弥补光谱数据少的缺点,形状特征为长宽比、形状指数等,并采用 SVM 方法进行分类。最后对 5 种 FCN 模型进行分类实验,统计各方法的整体正确率,如表 3 所示。5 种 FCN 模型中,FCN-4s 整体识别率最高,为 96.10%、97.52%、97.40%,而 FCN-32s 的整体识别率最低。FCN-4s 相比 FCN-32s 融合了更多细节特征进而优化识别效果。从表 3 中可以观察到随着融合细节特征的增加,FCN-32s、FCN-16s、FCN-8s、FCN-4s 整体正确率逐渐提升,但 FCN-2s 获取的过多浅层细节特征无助于正确率的进一步提升。

表 3 各方法的整体正确率

| Tab.3 Overall accuracy of each method % |        |        |        |
|-----------------------------------------|--------|--------|--------|
| 分类方法                                    | 测试区域 1 | 测试区域 2 | 测试区域 3 |
| FCN-32s                                 | 94.12  | 95.23  | 95.84  |
| FCN-16s                                 | 95.52  | 97.01  | 96.86  |
| FCN-8s                                  | 96.09  | 97.51  | 97.39  |
| FCN-4s                                  | 96.10  | 97.52  | 97.40  |
| FCN-2s                                  | 96.09  | 97.50  | 97.39  |
| 面向对象的分类方法                               | 80.50  | 83.50  | 81.35  |
| 基于像素的分类方法                               | 77.34  | 76.20  | 68.75  |

图 7 为识别效果图,通过对比可知,FCN-4s 在 3 种方法中识别效果最好,基于像素的分类方法识别结果较差,出现椒盐现象和碎片化,面向对象的分类方法虽然明显改善椒盐现象,但对于地膜农田和背景的分类存在大量错分情况。

4.2 实验结果分析

4.2.1 识别精度评价

采用混淆矩阵的方法进行分析,混淆矩阵的行代表分类器预测出的类别,列代表实际类别。由表 4~6 可知,对于测试区域 1、2、3,FCN-4s 对背景的识别精度为 97.06%、98.55%、97.94%,对大

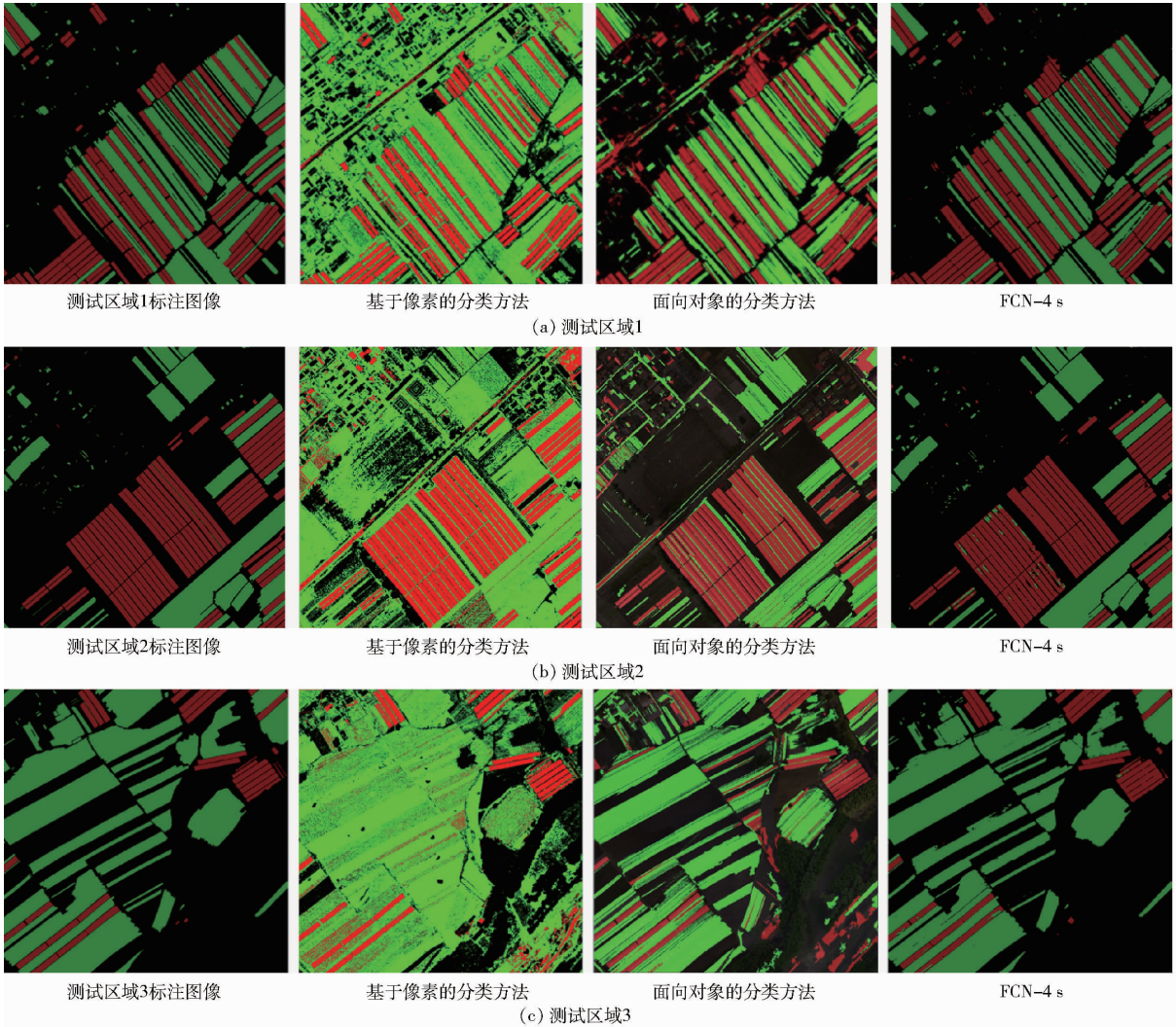


图 7 测试集识别结果对比

Fig. 7 Comparisons of experimental results on test set

棚识别精度为 96.74%、93.91%、95.71%，对地膜农田识别精度为 93.08%、98.18%、96.80%。而基于像素的分类方法对于大棚的识别效果较好，但对于背景和地膜农田效果较差，两者存在大量错分现象。面向对象的分类方法相比基于像素的分类方法对地膜农田识别精度有明显提升，但对背景类识别效果依旧较差，主要原因是背景类（包含裸地和未覆盖地膜的农田）与地膜农田特征相近，增加了错分的可能。实验结果证明，本文 FCN-4s 模型识别效果明显优于上述两种方法，对背景和地膜农田的错分情况有显著改善。

4.2.2 识别速度评价

为验证模型的识别效率，将 5 种 FCN 模型分别对测试集进行 100 次实验，统计运行时间，并与基于像素的分类方法、面向对象的分类方法进行比较。相比 ENVI 和 eCognition 遥感软件，本文方法采用 GPU 加速对模型的训练和识别。由表 7 可知，5 种 FCN 模型运行时间均明显小于面向对象的分类方

法和基于像素的分类方法的运行时间。通过对比 5 种 FCN 模型的平均运行时间，FCN-32s 模型用时最少，为 16.43 s，FCN-4s 次之，为 16.85 s。同时统计了整体正确率最高的 FCN-4s 模型不进行 GPU 加速的平均运行时间，为 536.80 s。因此利用 GPU

表 4 测试区域 1 实验结果的混淆矩阵  
Tab. 4 Confusion matrix of experiment results in test region 1 %

| 分类方法      | 预测类别 | 实际类别  |       |       |
|-----------|------|-------|-------|-------|
|           |      | 背景    | 大棚    | 地膜农田  |
| 基于像素的分类方法 | 背景   | 57.90 | 0.59  | 23.13 |
|           | 大棚   | 3.42  | 98.07 | 34.12 |
|           | 地膜农田 | 38.68 | 1.34  | 38.68 |
| 面向对象的分类方法 | 背景   | 54.62 | 4.47  | 4.38  |
|           | 大棚   | 29.63 | 91.00 | 1.56  |
|           | 地膜农田 | 15.75 | 4.53  | 94.06 |
| FCN-4s    | 背景   | 97.06 | 3.01  | 6.63  |
|           | 大棚   | 0.87  | 96.74 | 0.29  |
|           | 地膜农田 | 2.07  | 0.25  | 93.08 |

表 5 测试区域 2 实验结果的混淆矩阵

Tab.5 Confusion matrix of experiment results

| in test region 2 |      | %     |       |       |
|------------------|------|-------|-------|-------|
| 分类方法             | 预测类别 | 实际类别  |       |       |
|                  |      | 背景    | 大棚    | 地膜农田  |
| 基于像素的分类方法        | 背景   | 45.06 | 1.82  | 28.12 |
|                  | 大棚   | 43.52 | 93.01 | 21.32 |
|                  | 地膜农田 | 11.42 | 5.17  | 50.56 |
| 面向对象的分类方法        | 背景   | 81.25 | 7.02  | 14.79 |
|                  | 大棚   | 4.19  | 79.50 | 6.08  |
|                  | 地膜农田 | 10.56 | 13.48 | 79.13 |
| FCN-4s           | 背景   | 98.55 | 4.12  | 1.79  |
|                  | 大棚   | 0.83  | 93.91 | 0.03  |
|                  | 地膜农田 | 0.62  | 1.97  | 98.18 |

表 6 测试区域 3 实验结果的混淆矩阵

Tab.6 Confusion matrix of experiment results

| in test region 3 |      | %     |       |       |
|------------------|------|-------|-------|-------|
| 分类方法             | 预测类别 | 实际类别  |       |       |
|                  |      | 背景    | 大棚    | 地膜农田  |
| 基于像素的分类方法        | 背景   | 44.61 | 0.62  | 12.86 |
|                  | 大棚   | 8.02  | 89.83 | 15.02 |
|                  | 地膜农田 | 47.37 | 9.55  | 72.12 |
| 面向对象的分类方法        | 背景   | 84.82 | 13.69 | 20.49 |
|                  | 大棚   | 8.63  | 85.28 | 9.40  |
|                  | 地膜农田 | 6.55  | 1.03  | 70.11 |
| FCN-4s           | 背景   | 97.94 | 2.42  | 3.13  |
|                  | 大棚   | 0.49  | 95.71 | 0.07  |
|                  | 地膜农田 | 1.57  | 1.87  | 96.80 |

加速能够有效地提升模型的识别速度。

5 结论

(1) 利用无人机采集高分辨率、低成本的航拍

表 7 各方法的平均运行时间

Tab.7 Average runtime of each method

| 实验方法           | 运行时间/s    |
|----------------|-----------|
| FCN-32s(GPU)   | 16.43     |
| FCN-16s(GPU)   | 16.92     |
| FCN-8s(GPU)    | 16.98     |
| FCN-4s(GPU)    | 16.85     |
| FCN-2s(GPU)    | 16.95     |
| FCN-4s(CPU)    | 536.80    |
| 面向对象的分类方法(CPU) | 299.82    |
| 基于像素的分类方法(CPU) | 29 906.23 |

影像弥补了卫星遥感影像分辨率低、成本高、获取周期长的缺点。提出一种基于深度学习的大棚及地膜农田无人机监测方法,利用全卷积神经网络自动提取和学习图像的颜色、纹理、形状等特征,避免了传统分类方法复杂的特征选择过程,实现模型端到端训练。

(2) 采用多尺度的特征融合方法,实现了 5 种 FCN 模型,通过实验选择出识别效果最好的 FCN 模型:FCN-4s。该模型通过融合适量底层细节信息特征,得到了较好的识别效果,对于测试区域 1、2、3,本方法的平均整体正确率为 97%,明显高于基于像素分类方法和面向对象的分类方法,对大棚及地膜农田有较好的识别效果。

(3) 本文方法识别速度较快,利用 GPU 加速等手段优化了模型的识别速度,运行时间明显小于基于像素的分类方法和面向对象的分类方法的运行时间。

(4) 研究结果表明,本方法能够满足对大棚和地膜农田监测的要求,具有广泛的应用前景。

参 考 文 献

1 王勉. 高效设施农业发展现状及趋势分析[J]. 经济论坛, 2012, 2(2): 108-111.  
WANG Mian. The development status and trend of high efficiency facility agriculture[J]. Economic Forum, 2012, 2(2): 108-111. (in Chinese)

2 李中华, 王国占, 齐飞. 我国设施农业发展现状及发展思路[J]. 中国农机化, 2012, 34(1): 7-10.  
LI Zhonghua, WANG Guozhan, QI Fei. Current situation and thinking of development of protected agriculture in China[J]. Chinese Agricultural Mechanization, 2012, 34(1): 7-10. (in Chinese)

3 赵爱琴, 魏秀菊, 朱明. 基于 Meta-analysis 的中国马铃薯地膜覆盖产量效应分析[J]. 农业工程学报, 2015, 31(24): 1-7.  
ZHAO Aiqin, WEI Xiuju, ZHU Ming. Meta-analysis on impact of plastic film on potato yield in China[J]. Transactions of the CSAE, 2015, 31(24): 1-7. (in Chinese)

4 AGUERA F, AGUILAR F J, AGUILAR M A. Using texture analysis to improve per-pixel classification of very high resolution images for mapping plastic greenhouses[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2008, 63(6): 635-646.

5 黄振国, 陈仲新, 刘芳清, 等. 基于 HJ-1 影像的大棚菜地遥感监测技术研究[J]. 中国农业资源与区划, 2013, 34(5): 102-106.  
HUANG Zhen'guo, CHEN Zhongxin, LIU Fangqing, et al. Monitoring of greenhouse vegetables land using HJ-1 remotely-sensed imagery[J]. Chinese Journal of Agricultural Resources and Regional Planning, 2013, 34(5): 102-106. (in Chinese)

6 王志盼, 张清凌, 钱静, 等. 基于增强型水体指数的大棚遥感检测研究——以广东江门地区为例[J]. 集成技术, 2017, 6(2): 12-21.

- WANG Zhipan,ZHANG Qingling,QIAN Jing, et al. Greenhouse extraction based on the enhanced water index—a case study in Jiangmen of Guangdong[J]. Journal of Integration Technology,2017,6(2):12–21. (in Chinese)
- 7 沙先丽. 地膜农田遥感信息提取及覆膜地表温度反演[D]. 杭州:浙江大学,2012.  
SHA Xianli. A study on extraction of plastic-mulched landcover and in retrieval of mulched land surface temperature using remote sensing data[D]. Hangzhou:Zhejiang University,2012. (in Chinese)
- 8 WATTS A C, AMBROSIA V G, HINKLEY E A. Unmanned aircraft systems in remote sensing and scientific research: classification and considerations of use[J]. Remote Sensing, 2012, 4(6): 1671–1692.
- 9 刘峰,刘素红,向阳. 园地植被覆盖度的无人机遥感监测研究[J/OL]. 农业机械学报,2014,45(11):250–257. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20141139&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20141139&journal_id=jcsam). DOI:10.6041/j.issn.1000-1298.2014.11.039.
- LIU Feng,LIU Suhong,XIANG Yang. Study on monitoring fractional vegetation cover of garden plots by unmanned aerial vehicles [J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2014,45(11):250–257. (in Chinese)
- 10 HUNG C, XU Z, SUKKARIEH S. Feature learning based approach for weed classification using high resolution aerial images from a digital camera mounted on a UAV[J]. Remote Sensing, 2014, 6(12): 12037–12054.
- 11 鲁恒,付萧,贺一楠,等. 基于迁移学习的无人机影像耕地信息提取方法[J/OL]. 农业机械学报,2015,46(12):274–279. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20151237&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20151237&journal_id=jcsam). DOI:10.6041/j.issn.1000-1298.2015.12.037.
- LU Heng,FU Xiao,HE Yi'nan, et al. Cultivated land information extraction from high resolution UAV images based on transfer learning[J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2015,46(12):274–279. (in Chinese)
- 12 韩文霆,郭聪聪,张立元,等. 基于无人机遥感的灌区土地利用与覆被分类方法[J/OL]. 农业机械学报,2016,47(11):270–277. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20161137&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20161137&journal_id=jcsam). DOI:10.6041/j.issn.1000-1298.2016.11.037.
- HAN Wenting,GUO Congcong,ZHANG Liyuan, et al. Classification method of land cover and irrigated farm land use based on UAV remote sensing in irrigation[J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2016,47(11):270–277. (in Chinese)
- 13 KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//Advances in Neural Information Processing Systems, 2012: 1097–1105.
- 14 HARIHARAN B, ARBELAEZ P, GIRSHICK R, et al. Simultaneous detection and segmentation[C]//European Conference on Computer Vision. Springer International Publishing, 2014: 297–312.
- 15 LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3431–3440.
- 16 CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets and fully connected crfs[C]//International Conference on Learning Representations (ICLR), 2015.
- 17 ZHENG S, JAYASUMANA S, ROMERAP B, et al. Conditional random fields as recurrent neural networks[C]//Proceedings of the IEEE International Conference on Computer Vision,2015: 1529–1537.
- 18 SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C]//International Conference on Learning Representations (ICLR),2015.
- 19 SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1–9.
- 20 HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770–778.
- 21 ACHANTA R, SHAJI A, SMITH K, et al. SLIC Superpixels compared to state-of-the-art superpixel methods[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2274–2282.
- 22 TOKUI S, OONO K, HIDO S, et al. Chainer: a next-generation open source framework for deep learning[C]//Proceedings of workshop on machine learning systems (LearningSys) in the twenty-ninth annual conference on neural information processing systems (NIPS), 2015:1–6.
- 23 RUSSAKOVSKY O, DENG J, SU H, et al. ImageNet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211–252.