

# 基于深度学习的无人机航拍图像道路检测方法

## A Road Detection Method Based on Deep Learning for Uav Aerial Image

孙卫东 孙成雨

Weidong Sun Chengyu Sun

66133 部队  
中国·北京 100144  
66133 Troops,  
Beijing, 100144, China

**【摘要】**论文概述了图像卷积神经网络和语义分割的基本概念,通过采用基于深度学习的语义分割算法实现了航拍图像地面道路的分析与检测,进一步拓宽了无人机航拍图像的应用范围。

**【Abstract】**This paper summarizes the basic concepts of image convolution neural network and semantic segmentation, and realizes the analysis and detection of ground roads in aerial images by using the semantic segmentation algorithm based on deep learning, which further expands the application range of aerial images taken by uav.

**【关键词】**路面检测;无人机;航拍图像

**【Keywords】**road detection; uav; aerial images

**【DOI】**10.36012/etr.v1i3.396

## 1 引言

近年来,无人机技术获得了长足的发展。在国防、工业、安全等领域及消费市场都开始广泛应用无人机技术,并辅以自动控制及图像处理技术,使得基于图像(视频)采集处理技术的火灾预警、国土资源监控等技术获得了突破性的进展。

## 2 卷积神经网络与语义分割方法概述

常用的语义分割方法可以基于区域特征,也可以基于深度卷积神经网络。基于深度卷积神经网络的方法不需要对区域进行划分、对区域特征进行提取和对区域进行分类,而是采用神经网络算法直接提取图像中每个像素的特征,然后依据这些像素特征对像素进行分类识别,从而实现图像进行分割的目的。基于深度卷积神经网络提取的方法与基于区域特征提取的方法相比,具有更强的图像特征表示能力,更加简洁,准确率也更高,因而具有更好的效果。论文的运算即采用基于深度卷积神经网络的方法。

## 3 卷积神经网络的搭建

论文采用的卷积神经网络包括输入层、卷积层、池化层、激活函数层和全连接层,图像进入卷积神经网络后,依次经过各层。含有卷积层和池化层的神经网络利用卷积层提取目标

的特征,实现了权值共享,降低了模型过拟合,减少了模型参数,降低了计算复杂度;通过池化层提取具有空间变化鲁棒性的特征,减少了网络神经元的数量,提高了网络训练速度<sup>[12]</sup>。该类卷积神经网络运算速度较快、应用广泛。

## 4 基于深度学习的图像语义分割

基于深度卷积神经网络的图像语义分割方法也是建立在对图像像素进行分类识别的基础上的,该类方法主要利用包含卷积层和池化层的神经网络来实现,不同之处在于不再对图像进行区域划分、区域特征识别和区域分类,而是直接对图像像素进行运算,因而具有更高的准确率,且算法更简单。

## 5 道路检测应用

### 5.1 采集数据

进行神经网络的训练需要采集大量图片数据,论文利用无人机实地拍摄了部分图片和视频,如图1所示。



图1 无人机航拍图像

## 5.2 人工标注

基于卷积神经网络的语义分割是监督学习的过程,需要带有语义标注的图像对网络进行训练<sup>[3]</sup>。论文运算使用的数据来源对现实场景拍摄的视频,是不带标签的,因此在运算开始之前需要人工对视频场景目标进行标注。

论文使用的标注工具为 Python-Labelme。Labelme 支持颗粒度的像素级标注,在语义分割的标注中被广泛应用,程序界面如图 2 所示,标记过程如图 3 所示。



图 2 Labelme 程序界面



图 3 人工标注过程

## 5.3 网络数据集

由于样本及时间有限,论文大约使用了 200 张网络数据集 Scene Benchmark 有关道路的图片。

## 5.4 训练

训练过程分为 4 个阶段。

第一阶段是对分类网络进行初始化,如图 4 所示。



图 4 训练第一阶段

第二阶段采用 FCN-32s 网络,其中反卷积(图中橙色部分)步长为 32,实现两个功能:一是从特征小图(16x16x4096)预测分割小图(16x16x21),二是将预测出的分割小图升采样为大图,具体过程如图 5 所示。

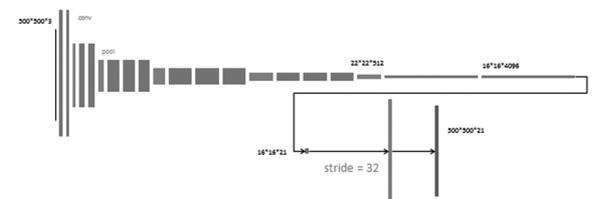


图 5 训练第二阶段

第三阶段采用 FCN-16s 网络,其中反卷积步长为 16,通过跳级结构提升精确性,实现两次升采样功能,并在第二次升采样前,把第 4 个 pooling 层的预测结果融合进来,具体过程如图 6 所示。

第四阶段采用 FCN-8s 网络,反卷积步长为 8,进一步融

合第 3 个 pooling 层的预测结果,完成 3 次升采样,具体过程如图 7 所示。

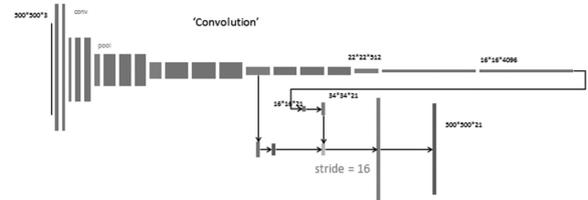


图 6 训练第三阶段

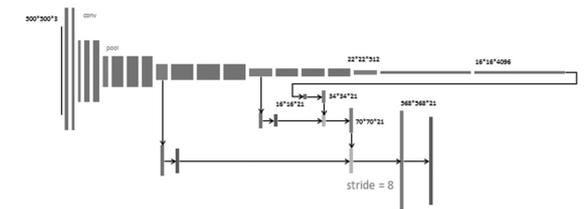


图 7 第四阶段训练

## 5.5 成果图展示

从结果上看,计算机对道路的检测没有问题,基本没有受到车道线的干扰,但对于车辆的识别精确度不高。一方面,因为标记的手法不够到位,标记的图像分割不清晰;另一方面,样本数量不够,导致训练量不足,对行人的检测精度也不够。以下是几张成果图及原图对比,如图 8 所示。



图 8 原图及成果图对比(左图为原图,右图为成果图)

## 6 结语

论文利用基于卷积神经网络的语义分割算法,实现了无人机航拍图像地面道路的分析与检测,通过仿真结果可以看出这种方法具有较好的效果。

### 参考文献

- [1]Hubel D H,Wiesel T N.Receptive Fields, Binocular Interaction and Functional Architecture Inthe Cat's Visual Cortex [J]. Journal of Physiology, 1962, 60(1):106.
- [2]Fukushima K.Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Patten Recognition Unaffected by Shift in Position[J]. Biological Cybernetics, 1980, 36(4):193-202.
- [3]Lecun Y, Bottou L, Bengio Y, et al. Gradient-based Learning Applied to Documentre Cognition [J].Proceedings of the IEEE,1998,86 (11):2278-2324.