# Introduction

这一章主要对整个毕设项目的整体做一个简单的介绍。其中1.1节给出了项目的背景，和一些基本概念；1.2节则并对本毕设的问题进行了定义；1.3节则对提出的新模型进行了简单介绍；1.4节给出了本文的组织架构以方便读者查阅。

## Background

航拍图像指的是……，（给例子）。它具有很大的应用。因为现实生活在不断更新，所展现的地图和也需要因此而更新。它能够对电子地图地图的绘制提供参考。它还有……的用途。这一节主要介绍毕设项目的背景。给出了什么是航拍图，什么是图像分割，以及展现我们所希望的到的几何图形。

### Image Segmentation

传统的图像分割技术是为图像中的每一个像素打label。如用0表示，1表示，2表示。航拍图像的图像分割指的是在航拍图像中提取出房屋、道路。（给例子）然而这样的表示方式有很多的局限性——如无法对房屋进行定位、存储上对每个像素，很废空间什么的。

### Geometrical Shape

我们想要脱离每像素标签，并直接通过向量表示学习多边形。当我们偏离标注像素的标准范例并旨在直接利用和学习建筑物的几何形状时  
-多边形表示包括比像素方式标签少得多的冗余  
-在存储上有优势  
-可以矢量表示，即在不同层面都可以用 ，建筑物等对象可以更自然地建模  
-学习到的多边形可以直接标注在电子地图中，而像素的不行  
因此多边形表示是感兴趣对象的更紧凑和有用的表示。因此我们希望能够从航拍图像中获取多边形

最近，深度学习方法在诸如图像分割和对象检测等许多计算机视觉任务中展现出了显着的成就。而相关方法已经被应用于航拍图像中[10,11,12,13,14]，但是，这些传统的CNN架构仍然局限于常规的网格结构图。这些网络的僵化使得很难开发关于图像内容的高级先验。他们输出的结果仍然是pixel wise的。那么，如何快速地从航拍图像中检测出矢量图变成了一大问题。

总之，这个项目的目的是为图形和任意结构的几何形状开发新颖的深度学习方法，我们试图将对象拓扑和形状引入深度学习技术。我们想要脱离标注像素的标准范例，而是直接利用并学习对象的几何图形。

## Problem Definition

（给图）给定一张航拍图，为航拍图中的每一个房屋，寻找它的多边形shape。输入是什么……输出是什么……

## Focus of This Work

为了解决这个问题，本文提出的模型叫region-based PolygonRNN。它是结合了Mask R-CNN中的FPN部分和PolygonRNN的产物。其中，FPN用于进行定位，即检测图像中的RoI，PolygonRNN部分用于find geometric shape。

## Thesis Organization

本论文组织如下。首先，我们回顾相关工作2 并着眼于像素级别的方法和bbox回归的方法（previous thesis）。并且看一下最近的新模型——mask r-cnn和单个polygon的

在第3章中，两个新模型的具体架构将会被讨论。成功的对象检测算法Faster R-CNN将在第4章中解释，因为它作为我们在第5章中描述的第一种建筑物分割方法的基础。我们的第二种建筑物分割方法基于建筑物角点提取，其后是连体网络来检测属于同一建筑物的匹配角落。这个方法的描述可以在第6章中找到。我们在第7章中对这两种方法进行了评估。在上一章中对结论进行了描述并对未来的工作进行了讨论。

# Related Work

在这一章中，我们XXX

## Previous Theses

在这一节中

## Recent Models

### PolygonRNN

### Mask R-CNN

## Motivation

# Model Architecture

这一章主要描述PolygonRNN、Mask R-CNN的模型架构，考虑到问题需要，结合PolygonRNN和FPN，提出了新的模型——region based PolygonRNN（R-PolygonRNN）

## PolygonRNN

这一张主要讲述寻找几何形状的网络——PolygonRNN

它是由XXX率先在哪里提出，用于汽车标记，更快annotate

### CNN Part

在介绍PolygonRNN的CNN部分之前，我们先介绍一下VGG-16

（VGG-16的部分）

在PolygonRNN中，CNN部分的主要作用是XXX，它为RNN提供了

需要注意的是，不同于图像分类，特征是从VGG-16的不同层次提取的，具体如下图所示

### RNN Part

ConvLSTM

## Faster/Mask R-CNN

### Region Proposal Network

### Feature Pyramid Network

## R-PolygonRNN

### Two-step Model

### Hybrid Model

# Experiments and Results

## Ground Truth

Our ground truth dataset consists of satellite images and buildings’ coordinates, which are used as inputs and labels respectively when training. This section describes how the ground truth dataset is generated. In detail, Subsection $ describes two different types of ground truth.

Generally, all of the high-resolution satellite images are collected from Google Static Maps API and all of the latitude and longitude coordinates of the polygon vertices of buildings are collected from OpenStreetMap. Subsection $ and subsection $

4.1.1小节描述了训练。4.1.2小节描述了从OpenStreetMap获取的数据类型。4.1.4小节描述了如何从经纬坐标投影到图像的像素坐标。

Since the dataset is collected from two different sources, the problem of inconsistency may exist. Subsection $ describes the details of the inconsistency and other problems of the ground true dataset.，这个问题的解决请参考4.1.5小节。

### Buildings and Areas

As mentioned in the section $, {model name} is formed by FPN and PolygonRNN.

根据3模型里面提到的，训练需要两种数据

RPN需要Area和bbox的gt，PolygonRNN需要Buildings和polygon的gt

虽然Buildings可以从area里提取，但是在area边缘处可能不完整，而且polygon可能会被图像边缘裁剪，导致不一样，产生新的顶点，这一块比较难计算

（提供边缘被裁减的例子）

所以是分别下载的。

而且独立可以保证训练时的batch\_size一致，如果每次训练的Building从area提取，可能会导致每次训练PolygonRNN的batch\_size不一致。

### Google Maps APIs

可以下载不同类型的数据，除了卫星图像，还有roadmap

然是roadmap只有房屋边界，所以最多用于4.1.5节的对准

还有PolygonRNN 里CNN part作为gt使用，但本毕设并没有采用

需要给定center坐标和resolution

scale

本毕设的zoom值在19, scale设为1

### OpenStreetMap

OSM能够在给定bbox的情况下，用XML的方式描述区域内房屋和道路的坐标

可以读

给出一个文件的例子

从OSM上获取的是房屋顶点坐标的经纬度，我们需要转化为像素坐标

在从下载Opensteet

给公式

将坐标化为整个地球的像素坐标

给ref

相减再相加即可

### Problems训练集

不一致的问题

（1）本身polygon偏大或偏小的问题，肉眼观察偏小的居多，偏大的比较少。这是由于osm测量房屋底部数据，航拍图则是从顶上拍的

(2)shift的问题

本身采集的问题

（3）高层房屋的角度问题，不能很好的cover到

（4）独栋房屋，虽然底下分割，但是航拍图看起来连续，bbox不准，polygon边缘的地方没有边缘也不准

4.1.6 Adjustment of inconsistency

数据存在一定偏差，因此我们需要对数据调整，使其一致

（1）var

（2）移动后edge与边缘的契合度

(3)corner是否都在角点

在一定程度上提升了准确度，但还是不太准

## Implementation Details

### Configuration

### Training

### Prediction

Beam Search

## Experiment Results

### Single Building Segmentation

### Buildings Localization

### Region-based PolygonRNN

# Problems and Future Work

## Problems

道路检测

## Future Work

### Improvement of Ground Truth

GT的提升，比如房屋的合并问题bbox，还有如何从单一源4.1.3说的，更准确

### Model Generalization

模型的泛化，Region-based PolygonRNN不仅可以做别的东西

5.2.3 采用alternative 训练方式而不是一起训练

5.2.4 采用RoI Align