项目报告

摘　要

随着人工智能的高速发展,计算机视觉技术成为了当今研究的热点,而深度学习的卷积神经网络已被验证是在计算机视觉领域非常有效的方法。图像语义分割是计算机视觉的一项基石性技术,图像语义分割算法好坏直接决定了对图像识别和理解的优劣,因此有效的图像语义分割算法的实现和应用具有非常重要的实际意义。本文首先总结了自动驾驶领域内语义分割的一些前沿成果。主要引用深度学习的方法进行图像的语义分割。利用数据集,以全卷积网络 FCN 进行语义分割的实验。

关键词: 深度学习、计算机视觉、语义分割、全卷积网络 FCN

一、自动驾驶领域内语义分割的一些前沿成果

（一）基于纹理特征的非结构化道路分割算法

主要思想：使用 Gabor 算子提取路面纹理特征,利用得到的方向特征值找出灭点,划分出可行驶区域。

方案实施：在建立可行区域中，对图像分 2 个频率、8 个方向进行 Gabor 变换并融合后得到其纹理特征，包括纹理强度和纹理方向，通过投票机制确定消失点，然后对有效投票区域进行 Hough 变换选取直线斜率，建立通过消失点的直线方程来划分出可行的道路区域。

1. 纹理特征提取

在对图像进行 Gabor 变换之前，为了提高计算效率，尽可能地摒除与车道无关的信息和干扰，所以要对图像进行有效区域选择。

（1） Gabor 变换

由于 Gabor 函数可以在频域不同尺度、不同方向上提取相关特征，其表达方式与人类视觉系统很相似，所以可以最大限度地凸显纹理方向特征。

二维 Gabor 滤波器表示为式：

$$G(x, y) = \frac{1}{2\pi\delta_u\delta_v}\exp\left\{-\frac{1}{2}\left(\frac{u^2}{\delta_u^2}+\frac{v^2}{\delta_v^2}\right)\right\}\cos(\omega u)$$

其中:$u = xcos\theta + ysin\theta$;$v = -ysin\theta + xcos\theta$;$\theta$表示滤波器的方向;$\delta_u$和$\delta_v$分别为高斯函数在$u$轴和$v$轴上的标准差，$u$轴平行于$\theta$，v 轴垂直于$\theta$;$\omega$表示正弦函数的频率，使 Gabor 滤波器有了频率选择特性。

经过实验验证，选取 2 个频率、8 个方向，一共 16 个模板进行滤波计算

（2） 纹理强度及方向

对于 8 个不同方向下的图像响应图，若该区域纹理方向与图像的纹理方向相同，则响应较强，反之则较弱。定义响应强度最大的方向为该点的纹理方向，并确定原图像中所对应点的方向索引值,即该点的纹理方向由 8 个 Gabor 方向特征最大值的索引值来确定。

2. 消失点检测

在非结构化的道路上，车辆行驶的痕迹或者道路边缘的纹理方向会在图像中相交于一点，即消失点。利用消失点可以对提取的道路边界起到约束作用，对道路的分割具有重要的意义

通过计算，可以得到图像中每一点的纹理强度以及纹理方向。为了计算出消失点，先对选取的候选消失点进行投票，候选消失点用 v 表示，投票点为图像中有效区域内的像素点，投票点用 p 表示，T(p)表示该点的纹理方向，d(p，v)表示 p 点与 v 点的连线方向，Vote(p，v)表示 p 点对 v 点的投票结果，Votes(v)为点 v 得到的票数，PVote 表示经过此投票机制得到的消失点。

3. 道路区域分割

通过建立两个直线方程来划分出车可行的道路区域。对有效的投票区域提取直线斜率，有效的投票区域即该点的纹理方向与该点投票方向的差值小于阈值的点的集合。

优点：无论是在强光下还是在光线不足的夜晚中，对于可行的道路区域分割有很好的效果，并且不受阴影的影响。

（二）基于小波变换和 K-means 的非结构化道路检测

主要思想：采用基于小波变换的方法提取图像边缘,结合高斯函数设定阈值去除非道路边缘点,使用基于斜率和截距的 K-means 聚类算法实现道路拟合。

方案实施：该方法将单目摄像机采集的彩色图像通过高斯金字塔进行降采样，金

字塔由高斯平滑和向下降采样构成，并使图像信息尽量不丢失，转成灰度图像处理，以提高实时性；利用既可有效降低图像加性噪声又可保持图像边缘细节的双边滤波对图像进行滤波，更好地保留道路边缘，同时去除噪声；通过小波变换求模极大值方法得到道路边缘，结合高斯函数设定阈值去除非道路边缘点，并根据基于斜率与截距的 K-means 聚类算法实现道路的拟合。

1. 基于小波变换模极大值的道路边缘提取

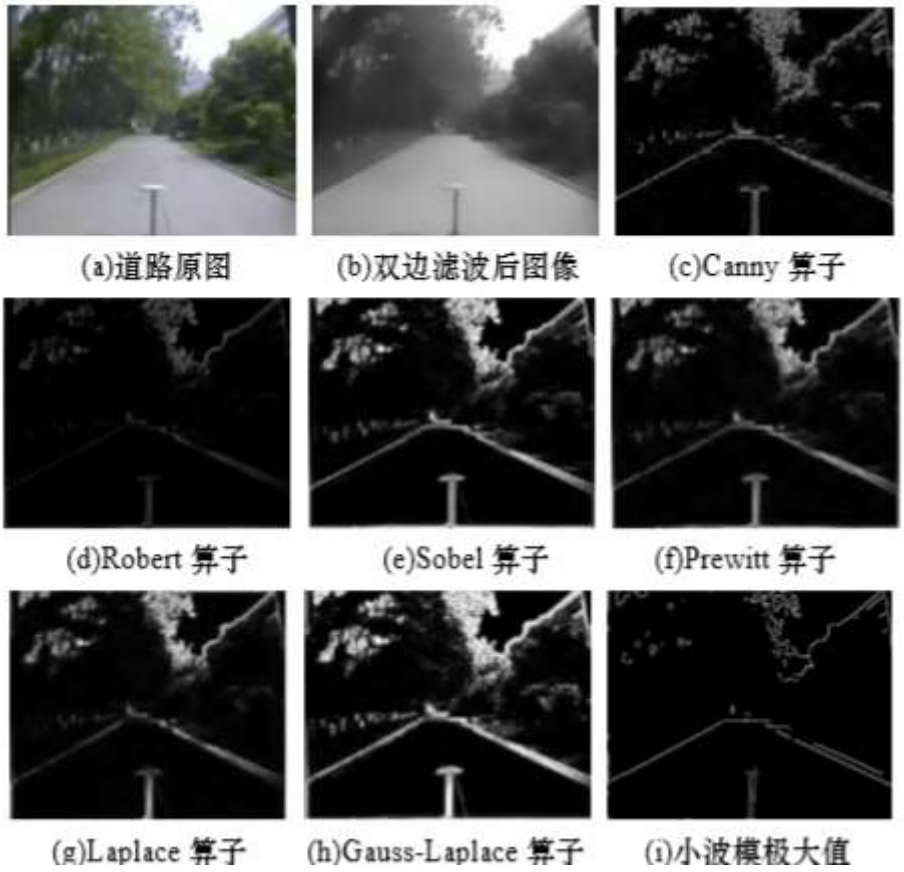
主要思想为沿梯度矢量方向就是梯度模极大值的方向，沿着梯度矢量方向检测小波系数模的局部极大值点便可得到图像的边缘点。求小波变换模极大值的一个关键步骤是构造小波函数，小波变换实质上是信号被平滑函数低通滤波后的导数，通过对二维高斯函数沿 x 和 y 方向求偏导获得小波函数，通过小波函数与信号进行卷积，获得梯度矢量的幅度和方向，求得小波变换模极大值。

2. 道路边缘点获取

在一幅图像中边缘方向可能是 0°~360°，对于道路检测，只对道路边界上的线感兴趣，因此，可以根据道路边缘方向的不同，去除边缘图像中的非道路边缘点，有效保留道路边缘点。

3. 基于斜率与截距的 K-means 聚类

K-means 聚类算法是一种无监督的分类方法，该方法的基本思想是通过最小平方误差准则或迭代方法，找出 x 个聚类中心 c1,c2,…,ck，使得每一个数据点 xi 与其最近的聚类中心 cv 的距离平方和最小。



(a)道路原图　(b)双边滤波后图像　(c)Canny 算子

(d)Robert 算子　(e)Sobel 算子　(f)Prewitt 算子

(g)Laplace 算子　(h)Gauss-Laplace 算子　(i)小波模极大值

(a)本文方法道路拟合结果

(b)最小二乘法道路拟合结果

（三）基于 RGB 熵和改进区域生长的非结构化道路识别方法

主要思想：首先生成道路 RGB 图像的熵图像并计算出熵图像直方图的最小差值，然后以此差值作为阈值分割道路图像，并使用区域生长方法提取出道路可行驶区域。

（四）后续发展

加州伯克利大学的 LongJ 等人提出了全卷积神经网络(FCN,fully convolutional networks),将 CNN 结构末端的全连接层用 1x1 的卷积层替代,得到低分辨率的预测结果，之后采用上采样的方式,将分辨率低的粗略图像恢复到原图的分辨率,首次实现了端到端的语义分割。后来的学者基于 FCN 的思想提出了更加优秀的网络结构，提出了在整个网络的后端加入条件随机场(CRF,conditional random fields),以细化粗糙的卷积输出。后来的学者又提出了空洞卷积理论,通过扩大卷积核,提取到更多的语义信息。加入了基于随机森林算法构成的卷积核,提高了语义分割的准确率。基于不断改进的 FCN 语义分割方法能够在完成城市道路图像多维特征的提取,实现道路的语义分割。

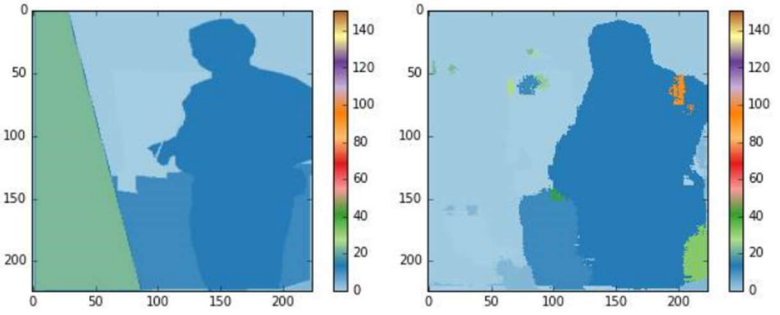
二、项目过程

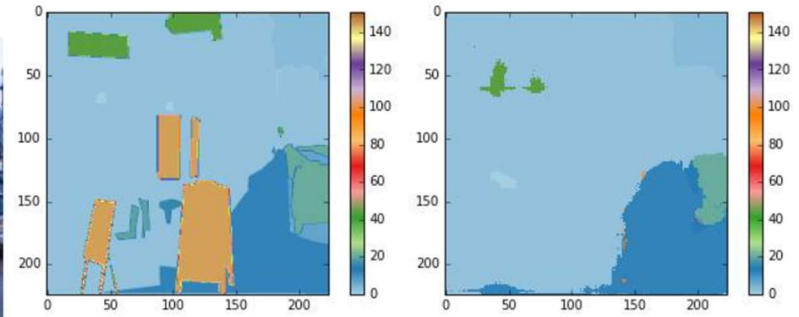
（一）代码流程图

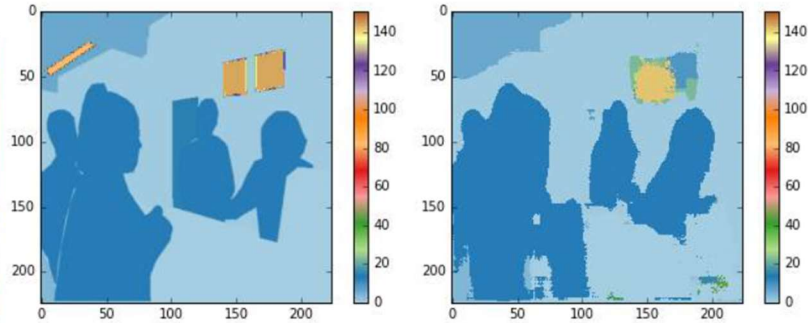
主要分几个程序：

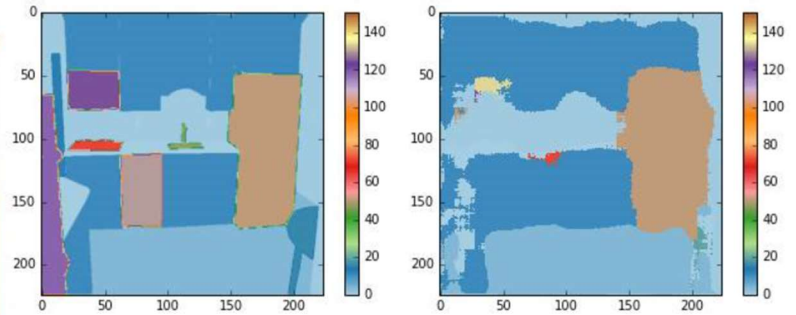
1.config.py：保存数据集的大部分参数；
2.read_data.py：读取数据集中用于的语义分割的训练集及其标签图片；
3.random_crop.py：随机裁剪和缩放图片；
4.fcn_vgg.py：定义 FCN 的网络结构并训练模型；
5.fcn_vgg_restore.py：用于加载预训练模型；

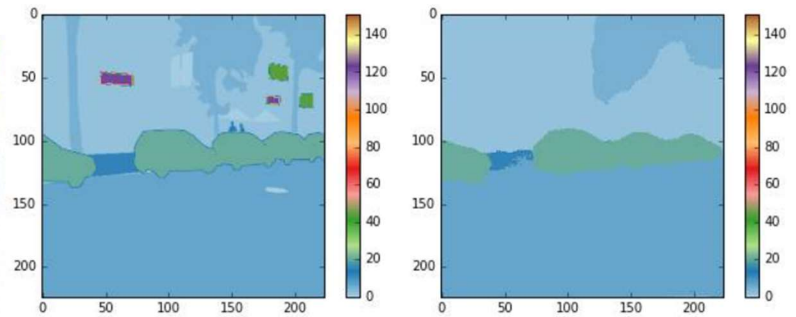
（二）运行结果



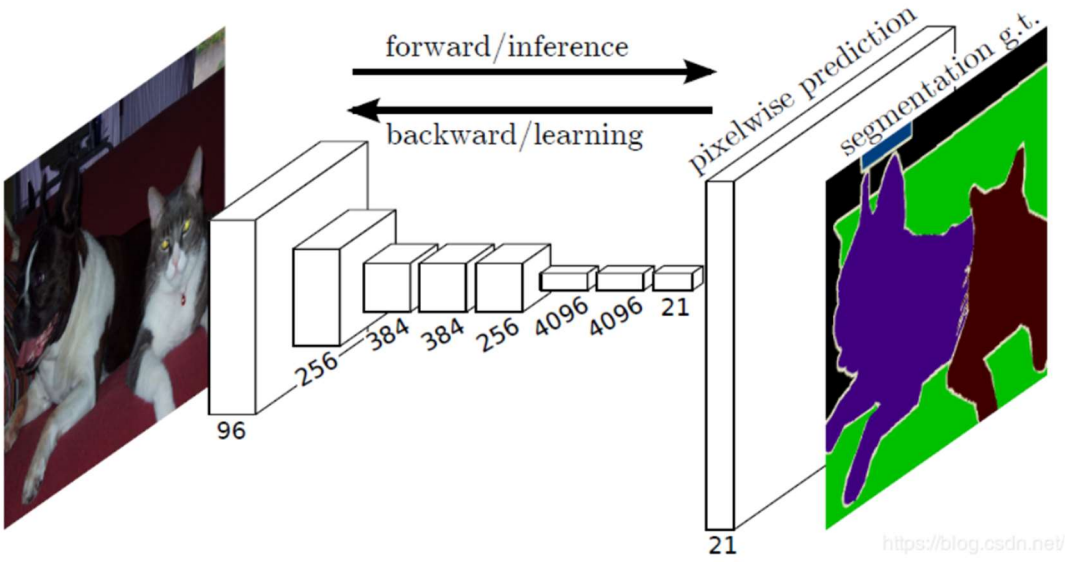






（三）模型结构图-FCN

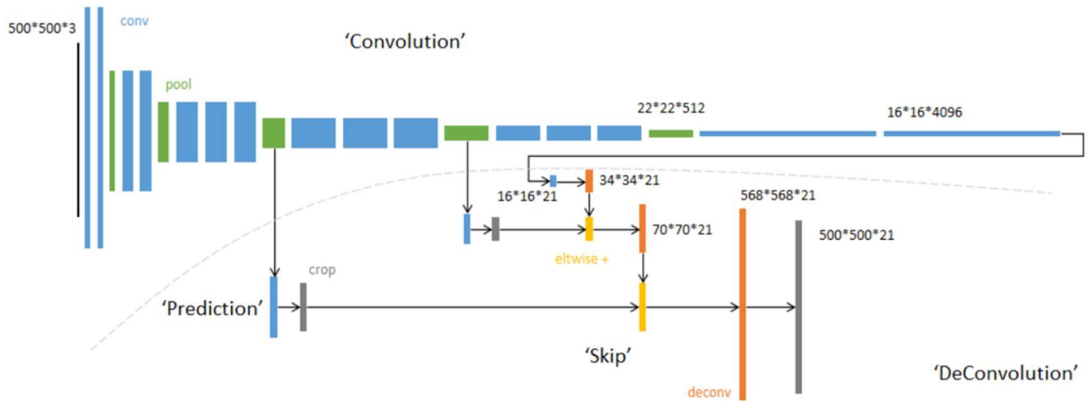
1.主要框架



2.算法详解

网络结构如下：

输入可为任意尺寸图像彩色图像；输出与输入尺寸相同，深度为：20 类目标+背景=21。



2.1 全卷积-提取特征

即将经典网络 AlexNet 最后两个全连接层改成了卷积层。

2.2 逐像素预测

图1

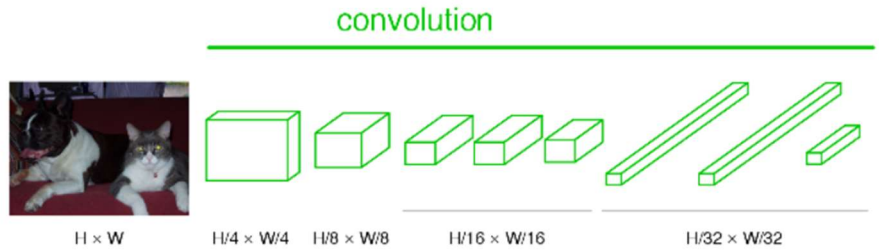


图2

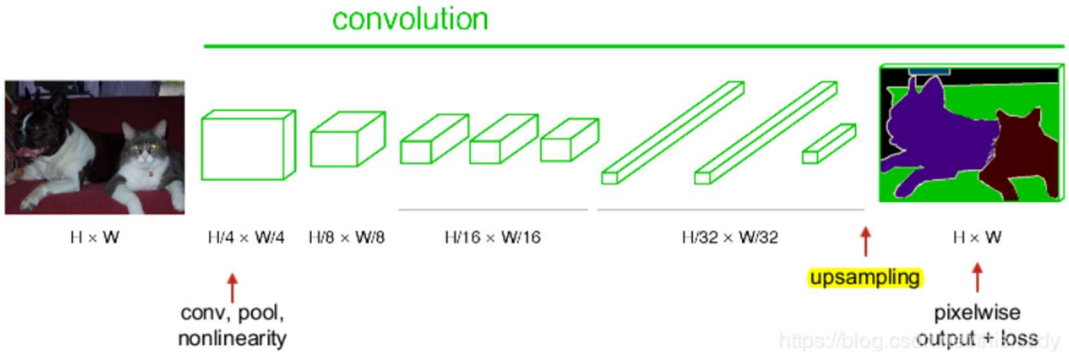
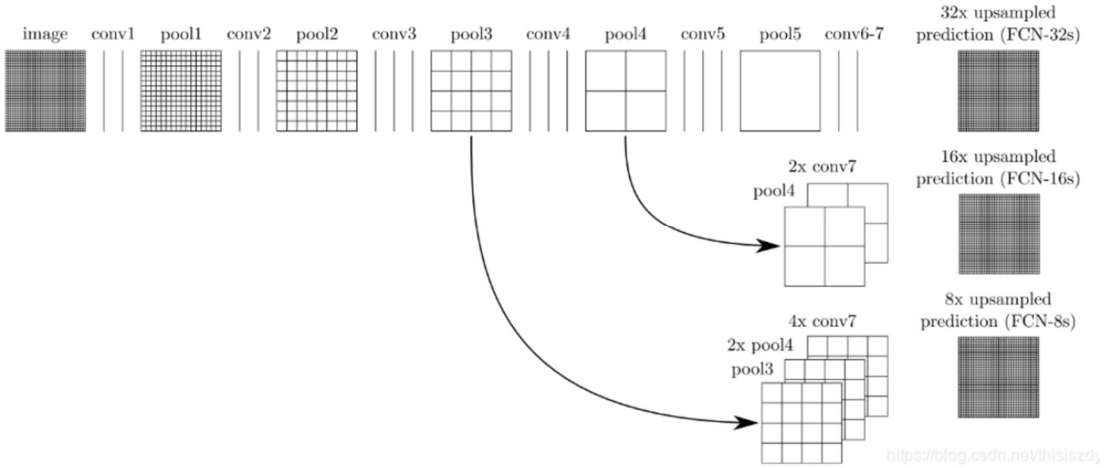


图 1 经过多次卷积和 pooling 以后，得到的图像越来越小，分辨率越来越低；图 2 采用反卷积层对最后一个卷积层的 feature map 进行上采样，使它恢复到输入图像相同的尺寸，从而可以对每个像素都产生了一个预测，同时保留了原始输入图像中的空间信息，最后在上采样的特征图上进行逐像素分类。

2.3 跳级结构



- 如上图所示：对原图进行卷积 conv1、pool1 后图像缩小为 1/2；对图像进行第二次卷积 conv2、pool2 后图像缩小为 1/4；对图像进行第三次卷积 conv3、pool3 后图像缩小为 1/8，此时保留 pool3 的 feature map；对图像进行第四次卷积 conv4、pool4 后图像缩小为 1/16，此时保留 pool4 的 feature map；对图像进行第五次卷积 conv5、pool5 后图像缩小为 1/32，然后把原来 CNN 操作过程中的全连接变成卷积操作的 conv6、conv7，图像的 feature map 的大小依然为原图的 1/32，此时图像不再叫 feature map 而是叫 heat map。

- 其实直接使用前两种结构就已经可以得到结果了，这个上采样是通过反卷积（deconvolution）实现的，对第五层的输出（32 倍放大）反卷积到原图大小(FCN-32s)。但是得到的结果还不够精确，一些细节无法恢复。于是将第四层的输出和第三层的输出也依次反卷积，分别需要 16 倍 (FCN-16s) 和 8 倍 (FCN-8s) 上采样，结果过也更精细一些了。这种做法的好处是兼顾了 local 和 global 信息。

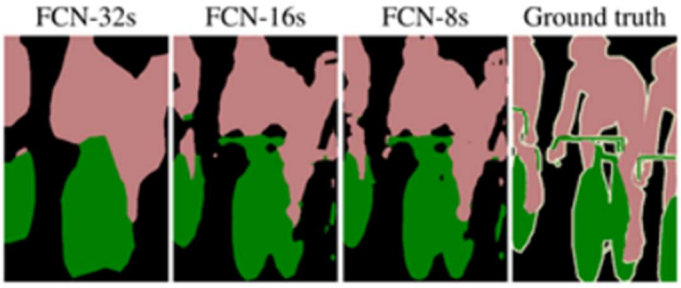
2.4 网络 loss 层

FCN 网络的输入 batchsize 是 1，因为分割 loss 的计算在每一个像素点都一个标签，相当于每一个像素点的都是一个分类任务 softmax，一个图像就有对应像素点个样本。所以分割任务的 batch 是一个图片，将一个图片最后在所有像素点上的分类 loss 加起来计算一次梯度的更新。

3.实验

3.1 FCN based 各种 CNN 性能对比

	FCN-AlexNet	FCN-VGG16	FCN-GoogLeNet[4]
mean IU	39.8	**56.0**	42.5
forward time	50 ms	210 ms	59 ms
conv. layers	8	16	22
parameters	57M	134M	6M
rf size	355	404	907
max stride	32	32	32



3.2FCN-32s/16s/8s 性能对比

	pixel acc.	mean acc.	mean IU	f.w. IU
FCN-32s-fixed	83.0	59.7	45.4	72.0
FCN-32s	89.1	73.3	59.4	81.4
FCN-16s	90.0	75.7	62.4	83.0
FCN-8s	**90.3**	**75.9**	**62.7**	**83.2**

从图和实验结果可以看出，使用了跳连接结构的 FCN 分割结果更加精细，其中 FCN-8s 效果最优。

参考文献

1.《基于纹理特征的非结构化道路分割算法》　刘富　袁雨桐　李洋

2.《基于小波变换和 K-means 的非结构化道路检测》　熊思　李磊民　黄玉清

3.《基于 RGB 熵和改进区域生长的非结构化道路识别方法》　吴骅跃　段里仁

4.《DeepLab:semantic image segmentation with deep convolutional nets,atrous convolution,and fully connected CRFs》　Chen L C　Papandreou G　Kokkinos I

5.《Learning deconvolution network for semantic segmentation》　Noh H　Hong S　Han B

6.《车道线检测和语义分割模型在自动驾驶中的应用》　陈必科

7.《基于 SegNet 的非结构道路可行驶区域语义分割》　张凯航　冀杰　蒋骆　周显林

8.《自动驾驶中的视频语义分割技术研究》　王蒲

9.《基于 FCN 的图像语义分割研究及可行驶区域分割应用》　姚森

10. https://blog.csdn.net/shareviews/article/details/83028038

11. https://blog.csdn.net/weixin_40446557/article/details/85624579

12. https://www.jianshu.com/p/71cda17bd6ef

13.https://blog.csdn.net/weixin_34370347/article/details/88722893?utm_medium=distribute.pc_aggpage_search_result.none-task-blog-2~all~first_rank_v2~rank_v25-10-88722893.nonecase

14.https://blog.csdn.net/m0_37862527/article/details/79843963?utm_medium=distribute.pc_aggpage_search_result.none-task-blog-2~all~sobaiduend~default-1-79843963.nonecase

15. https://www.cnblogs.com/wujianming-110117/p/12901960.html

16. https://blog.csdn.net/qq583083658/article/details/86321987

17. https://blog.csdn.net/solomon1558/article/details/70173223

18. https://blog.csdn.net/m0_37862527/article/details/79842547

19.https://blog.csdn.net/qq_37274615/article/details/78228087?utm_medium=distribute.pc_relevant.none-task-blog-BlogCommendFromMachineLearnPai2-3.channel_param&depth_1-utm_source=distribute.pc_relevant.none-task-blog-BlogCommendFromMachineLearnPai2-3.channel_param

20. https://blog.csdn.net/zz937211040/article/details/103571321

21. https://blog.csdn.net/thisiszdy/article/details/89080854