DEEP CNN WITH COLOR LINES MODEL FOR UNMARKED ROAD SEGMENTATION

ICIP 2017 Paper #2702

Shashank Yadav Suvam Patra Chetan Arora Subhashis Banerjee

Computer Vision Group

INDIAN INSTITUTE OF TECHNOLOGY DELHI

INDRAPRASTHA INSTITUTE OF INFORMATION TECHNOLOGY DELHI

❖*This work is partly supported by grant from Continental India Pvt Ltd*

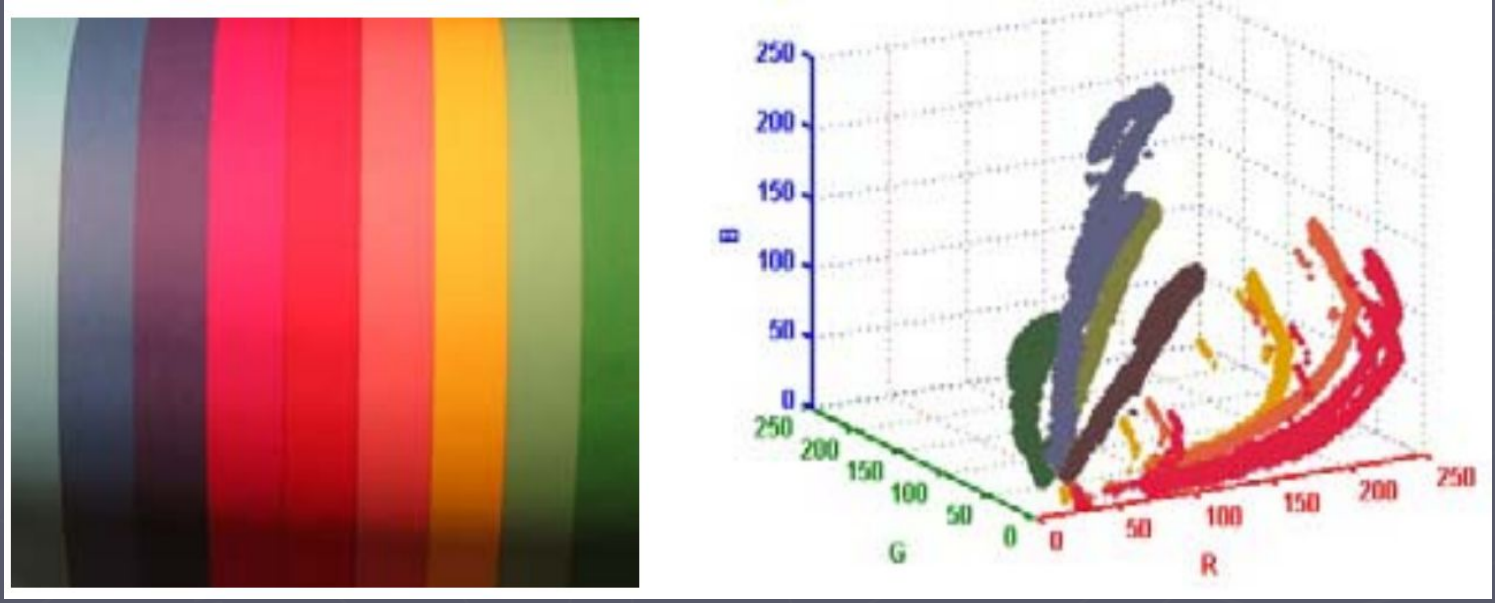
Motivation

- Road segmentation has important applications in the area of autonomous driving
- A lot of learning based methods have been developed owing to the recent advances in deep learning
- All of these methods suffer from *dataset bias*, a network trained on urban roads won't work well in rural scenario
- Our aim is to counter this bias and make learning based approach work on unmarked and unmaintained roads.

Approach

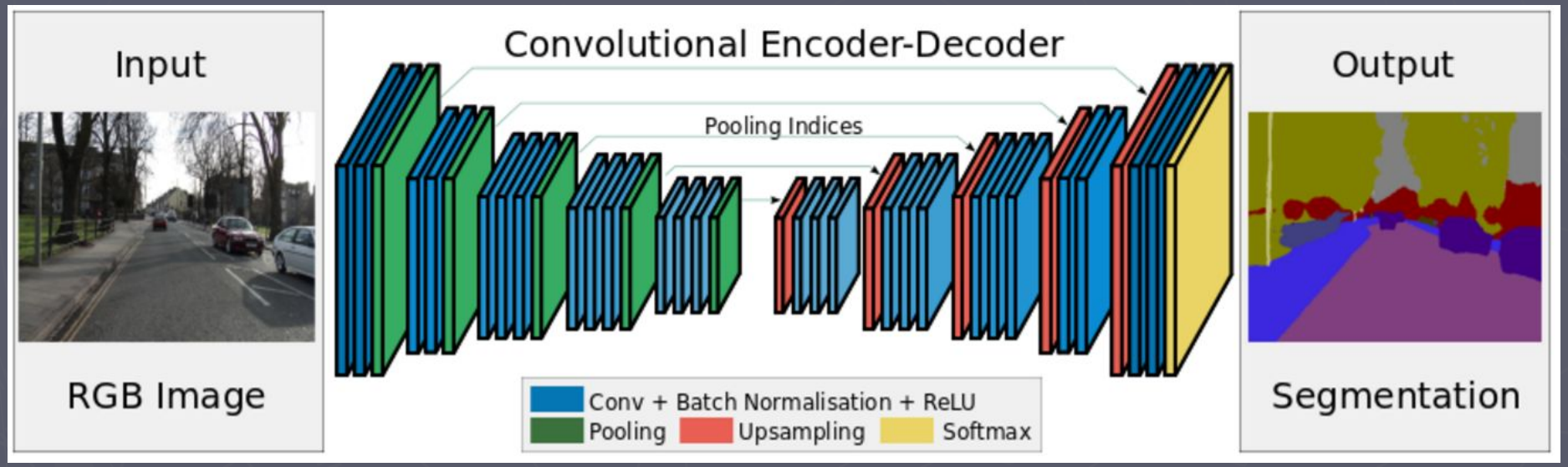
- Output from a trained CNN, albeit on a different dataset, contains useful information that can be used as a prior
- Intensity based segmentation approaches are more principled, but illumination and shadows cause problems
- We use both the information and solve road segmentation as a joint optimization using both the cues

Color Lines



An illumination invariance representation of image in the RGB space
[Omer et al., CVPR 2004]

SegNet

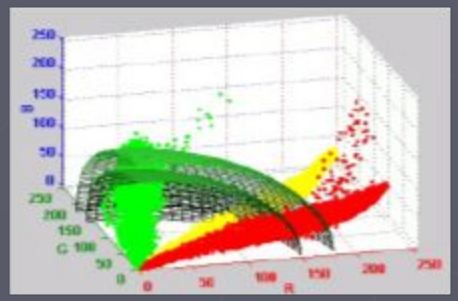


A Deep Convolutional Encoder-Decoder Architecture for Robust Semantic Pixel-Wise Labelling
[Kendall et al., PAMI 2017]

Approach

Input Frame

CNN



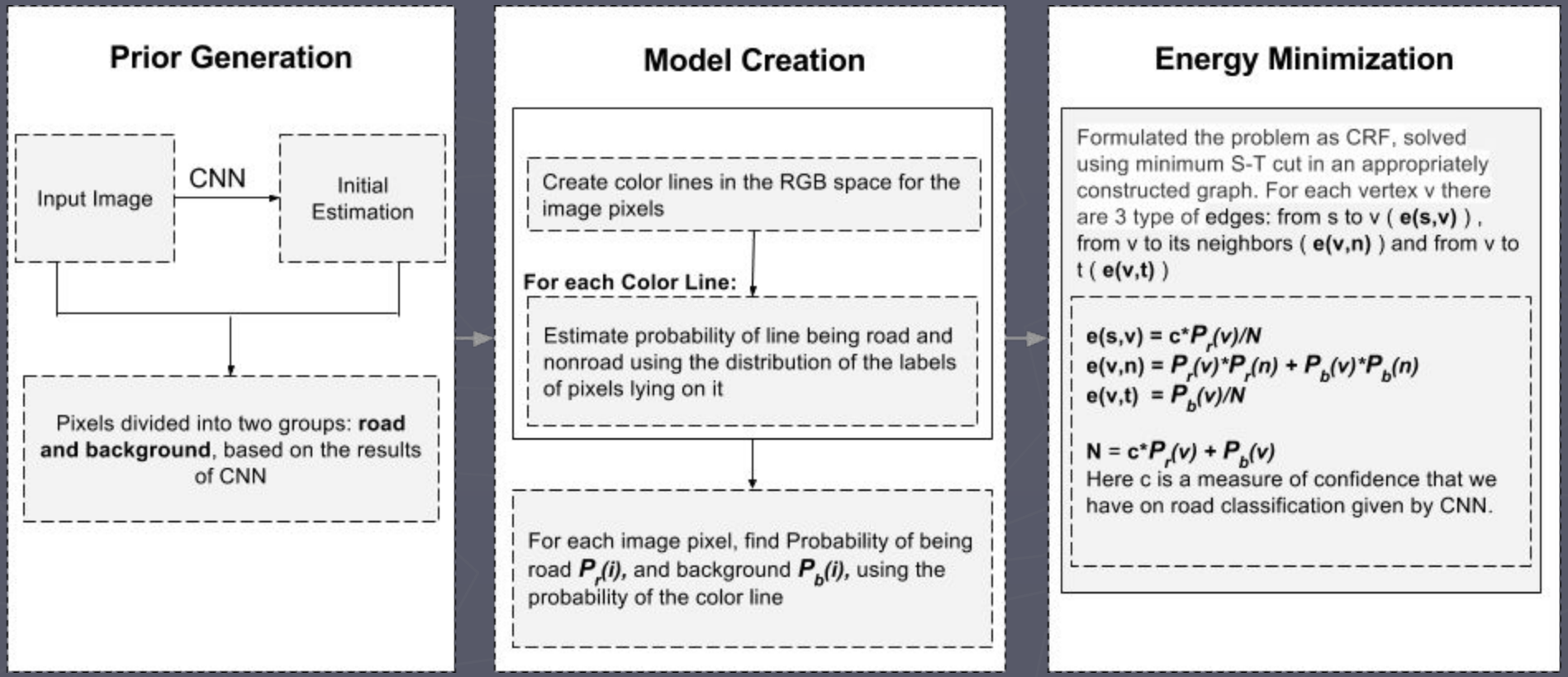
Color Lines Model

Prior from CNN

CRF

Output

Framework


Prior Generation
Input Image
CNN
Initial Estimation
Pixels divided into two groups: road and background, based on the results of CNN
Model Creation
Create color lines in the RGB space for the image pixels
For each Color Line:
Estimate probability of line being road and nonroad using the distribution of the labels of pixels lying on it
For each image pixel, find Probability of being road $P_r(i)$, and background $P_b(i)$, using the probability of the color line
Energy Minimization
Formulated the problem as CRF, solved using minimum S-T cut in an appropriately constructed graph. For each vertex v there are 3 type of edges: from s to v (e(s,v)) , from v to its neighbors (e(v,n)) and from v to t (e(v,t))
$e(s,v) = c*P_r(v)/N$
$e(v,n) = P_r(v)*P_r(n) + P_b(v)*P_b(n)$
$e(v,t) = P_b(v)/N$
$N = c*P_r(v) + P_b(v)$
Here c is a measure of confidence that we have on road classification given by CNN.

Qualitative Results

Results on some rural images:

Segmentation using state of the art CNN (SegNet) in first row vs our method in second row

Quantitative Results

Method	Benchmark	MaxF	AP	PRE	REC	FPR	FNR
Segnet	UM_ROAD	82.17 %	76.46 %	84.03 %	80.40 %	6.97 %	19.60 %
	UMM_ROAD	88.59 %	83.54 %	88.35 %	88.84 %	12.88 %	11.16 %
	UU_ROAD	77.23 %	69.23 %	82.29 %	72.76 %	5.10 %	27.24 %
	URBAN_ROAD	84.04 %	78.76 %	85.50 %	82.63 %	7.72 %	17.37 %
Ours	UM_ROAD	83.50 %	72.28 %	76.37 %	92.09 %	12.98 %	7.91 %
	UMM_ROAD	90.30 %	83.33 %	86.43 %	94.53 %	16.32 %	5.47 %
	UU_ROAD	79.89 %	67.48 %	77.01 %	82.99 %	8.07 %	17.01 %
	URBAN_ROAD	85.73 %	76.89 %	81.02 %	91.01 %	11.74 %	8.99 %

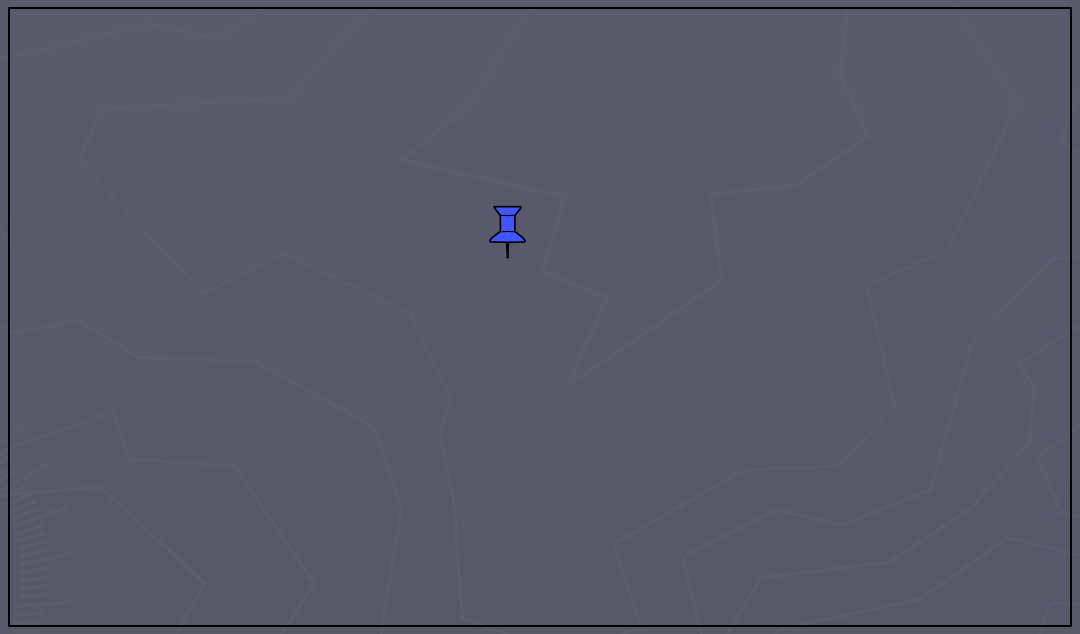
Table 1: Segmentation Results on the KITTI [8] dataset. MaxF: Maximum F1-measure, AP: Average precision as used in PASCAL VOC [24] challenges, PRE: Precision, REC: Recall, FPR: False Positive Rate, FNR: False Negative Rate (the four latter measures are evaluated at the working point MaxF). This is as mentioned in the KITTI Benchamrk Suite [8]

Results on KITTI and CamVid datasets as shown in the paper. SegNet has been trained on CamVid hence we see no significant improvements. KITTI on the other hand shows improvements in all sequences

Method	PRE	REC	F val
Segnet	93.07 %	94.86 %	93.95 %
Ours	93.31 %	94.99 %	94.14 %

Table 2: Segmentation Results on Camvid [9] Testing dataset

Results on different datasets



Summary

- Proposed a simple method to segment unmarked and unmaintained roads common in rural and developing world scenarios
- Introduced the idea of augmenting color lines using CNN prior and creating a probabilistic model
- Validated the proposed technique on various benchmark and self shot videos in Indian road conditions.

Thank You!!!