

Fully Convolutional Neural Network with Relation-Aware Context Information for Image Parsing

Presenting Author

Basim Azam

b.azam@cqu.edu.au

Central Queensland University , Australia

Co-Authors

Ranju Mandal , Brijesh Verma

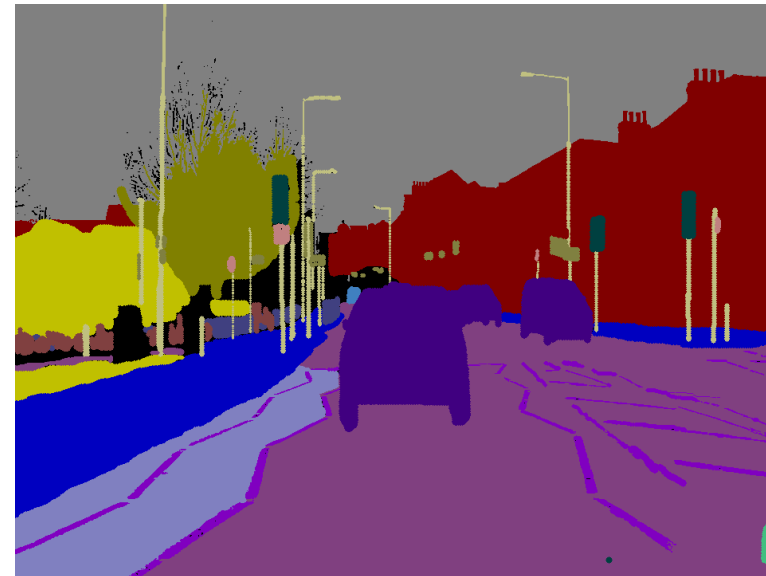
Central Queensland University, Australia

Digital Image Computing: Techniques and Applications, Gold Coast, Queensland, Australia



Introduction

Image parsing refers to segmentation of an image into regions with object category labels such as tree, building, car and road.

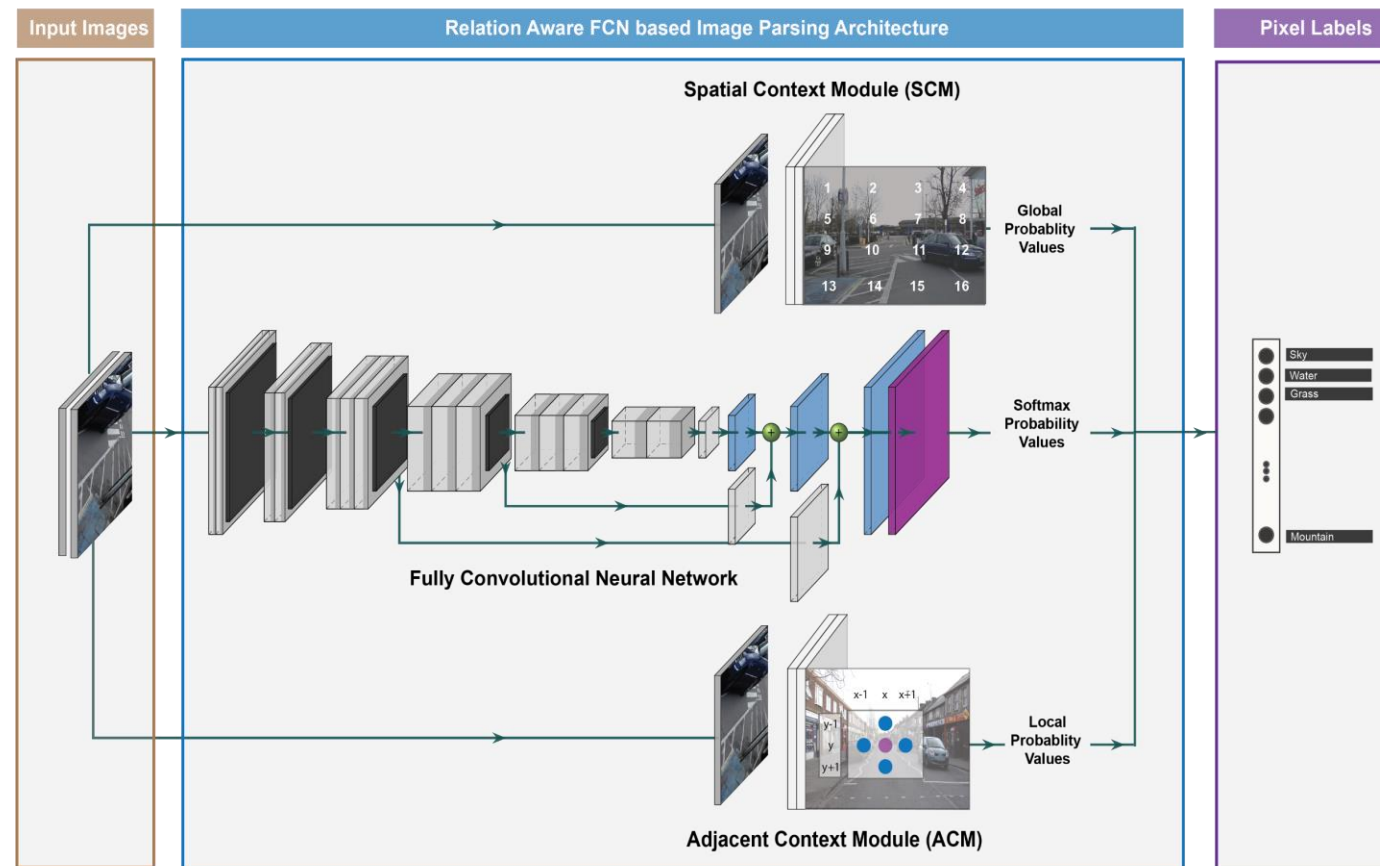


Motivation

- The modern frameworks have shown capability for producing accurate pixel-wise labelling.
- CNN based approaches obtain a coarse label map by applying pixelwise convolution operations on input images.
- It is highly desirable to learn critical contextual features only in such high-dimensional feature space.



Proposed Approach

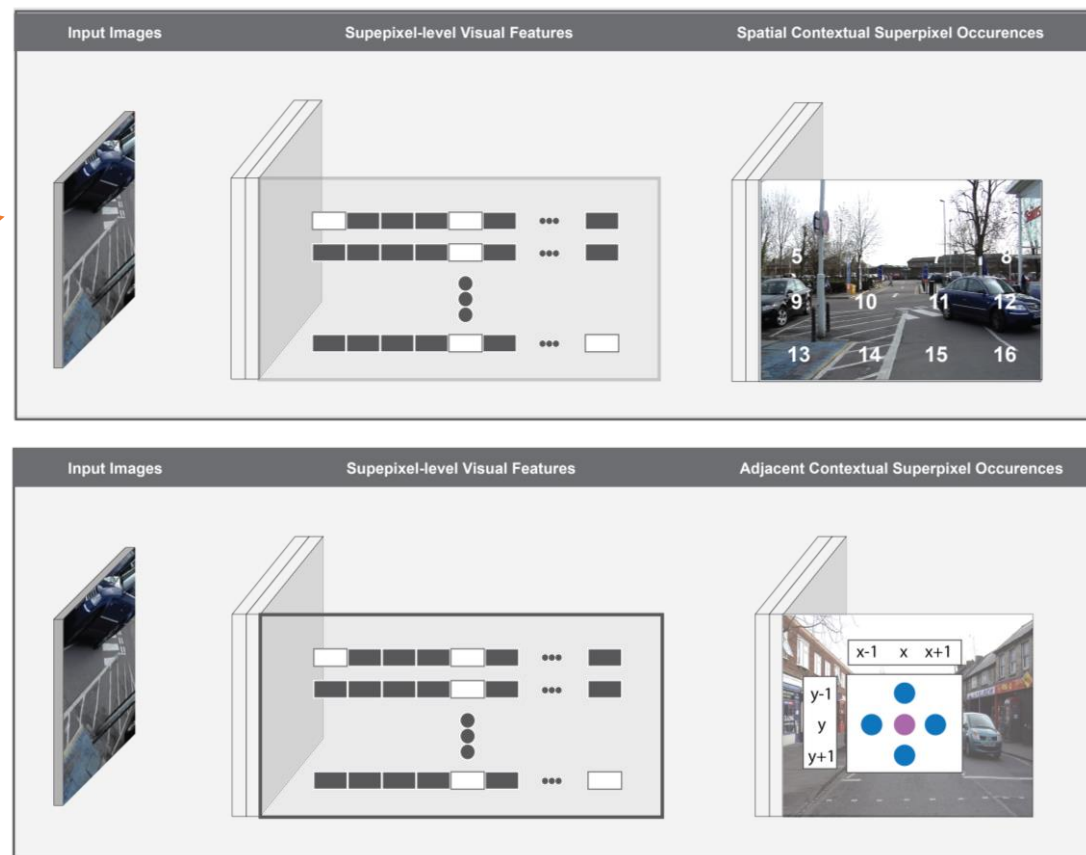
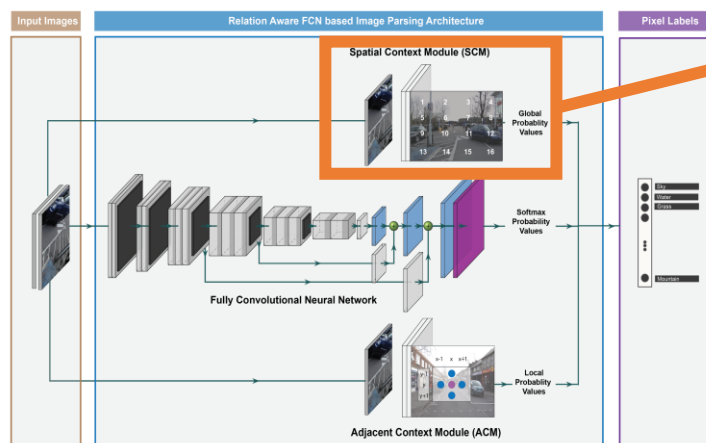


THE OVERVIEW OF THE RELATION-AWARE FCN BASED IMAGE PARSING FRAMEWORK WITH INTEGRATED SPATIAL CONTEXT MODULE AND ADJACENT CONTEXT MODULES.

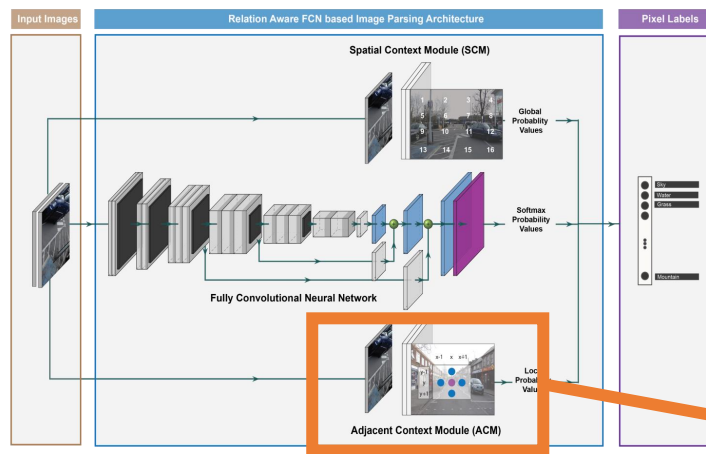


Relation-Aware Context Modules

Spatial Context Module



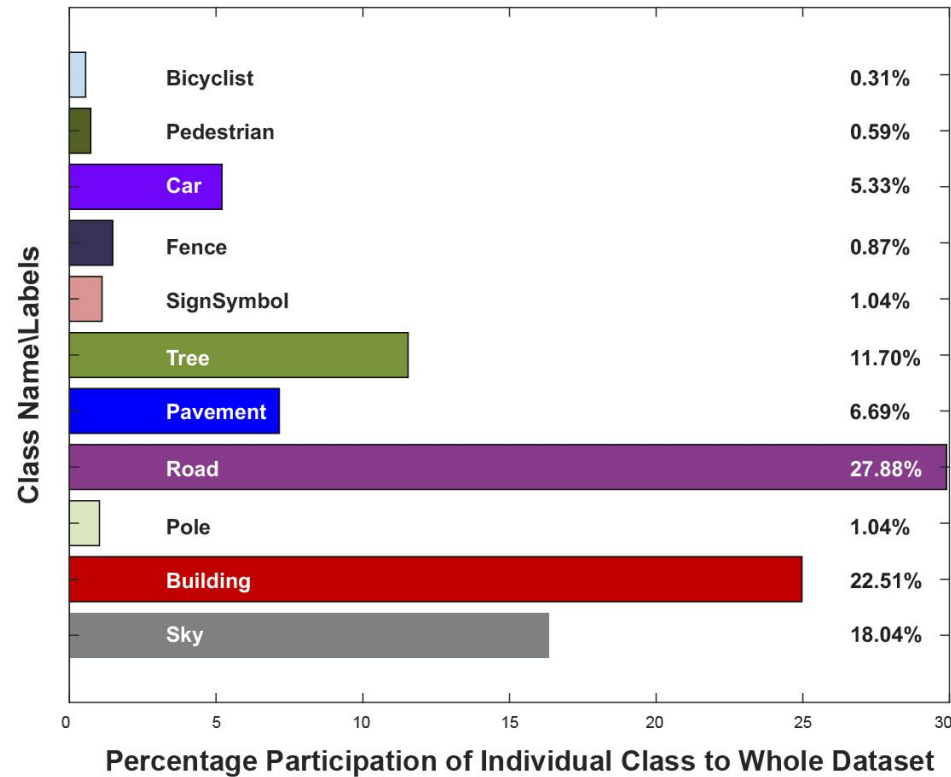
Relation-Aware Context Modules



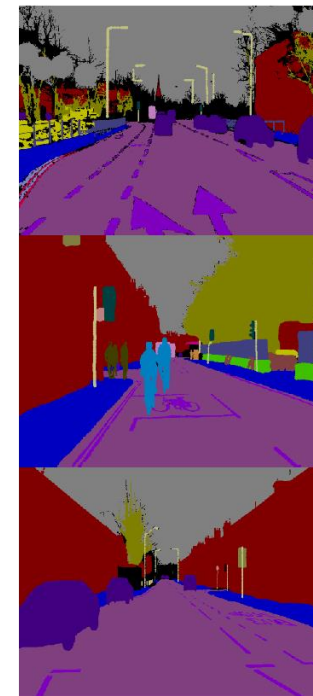
Adjacent Context Module



Dataset



Images



Labels

The overview of the CamVid Database used in the study. Left: Percentage breakdown by category, Right: Visual sample images with labels [1].



Results

RESULTS ON THE CAMVID DATASET IN TERMS OF CLASS-WISE ACCURACY, MEAN JACCARD SCORE AND GLOBAL ACCURACY IN COMPARISON WITH THE STATE-OF-THE ART SEGMENTATION TECHNIQUES.

Approach / Class	Building	Tree	Sky	Car	Sign	Road	Pedestrian	Fence	Pole	Sidewalk	Cyclist	Mean IOU	Global Accuracy
ReSeg [1]	86.6	84.7	93.0	87.3	48.6	98.0	63.3	20.9	35.6	87.3	43.5	58.8	88.9
SegNet[2]	68.7	52.0	87.0	58.5	13.4	86.2	25.3	17.9	16.0	60.5	24.8	46.4	62.5
B. Segnet [3]	-											63.1	86.9
FCN-8 [4]	77.8	71.0	88.7	76.1	32.7	91.2	41.7	24.4	19.9	72.7	31.0	57.0	88.0
Dilation-8 [5]	82.6	76.2	89.0	84.0	46.9	92.2	56.3	35.8	23.4	75.3	55.5	65.3	79.0
DeepLab-L [6]	81.5	74.6	89.0	82.2	42.3	92.2	48.4	27.2	14.3	75.4	50.1	61.6	-
FCN-Comb[7]	79.7	77.2	85.7	86.1	45.3	94.9	45.9	69.0	25.2	86.2	57.9	-	88.8
Proposed Approach	93.2	86.4	94.5	86.7	31.9	96.3	26.9	71.5	27.1	83.7	62.2	64.1	89.8



Results

RESULTS ON THE CAMVID DATASET IN TERMS OF CLASS-WISE ACCURACY, MEAN JACCARD SCORE AND GLOBAL ACCURACY IN COMPARISON WITH THE STATE-OF-THE ART SEGMENTATION TECHNIQUES.

Approach / Class	Building	Tree	Sky	Car	Sign	Road	Pedestrian	Fence	Pole	Sidewalk	Cyclist	Mean IOU	Global Accuracy
ReSeg [1]	86.6	84.7	93.0	87.3	48.6	98.0	63.3	20.9	35.6	87.3	43.5	58.8	88.9
SegNet[2]	68.7	52.0	87.0	58.5	13.4	86.2	25.3	17.9	16.0	60.5	24.8	46.4	62.5
B. Segnet [3]	-											63.1	86.9
FCN-8 [4]	77.8	71.0	88.7	76.1	32.7	91.2	41.7	24.4	19.9	72.7	31.0	57.0	88.0
Dilation-8 [5]	82.6	76.2	89.0	84.0	46.9	92.2	56.3	35.8	23.4	75.3	55.5	65.3	79.0
DeepLab-L [6]	81.5	74.6	89.0	82.2	42.3	92.2	48.4	27.2	14.3	75.4	50.1	61.6	-
FCN-Comb[7]	79.7	77.2	85.7	86.1	45.3	94.9	45.9	69.0	25.2	86.2	57.9	-	88.8
Proposed Approach	93.2	86.4	94.5	86.7	31.9	96.3	26.9	71.5	27.1	83.7	62.2	64.1	89.8



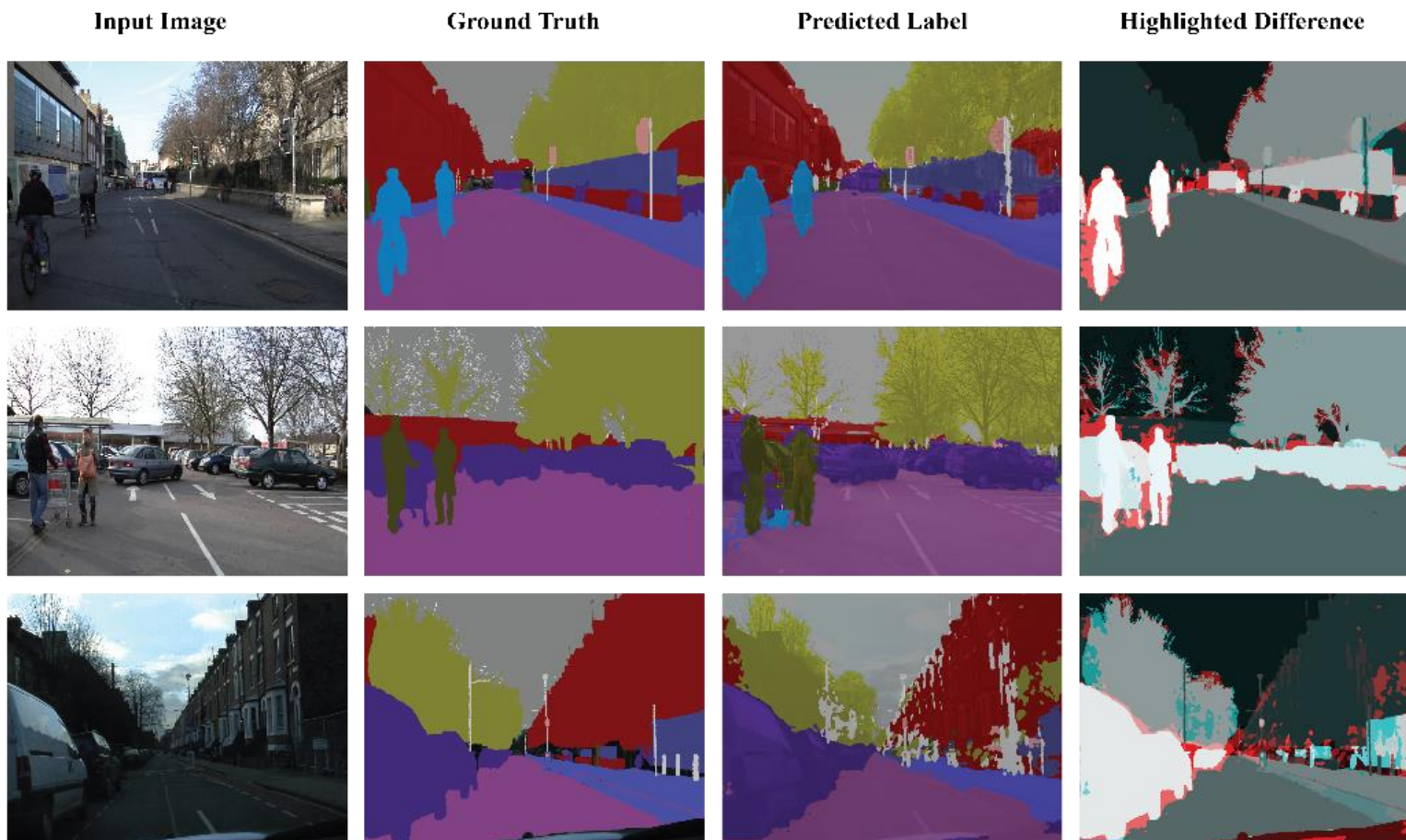
Results

RESULTS ON THE CAMVID DATASET IN TERMS OF CLASS-WISE ACCURACY, MEAN JACCARD SCORE AND GLOBAL ACCURACY IN COMPARISON WITH THE STATE-OF-THE ART SEGMENTATION TECHNIQUES.

Approach / Class	Building	Tree	Sky	Car	Sign	Road	Pedestrian	Fence	Pole	Sidewalk	Cyclist	Mean IOU	Global Accuracy
ReSeg [1]	86.6	84.7	93.0	87.3	48.6	98.0	63.3	20.9	35.6	87.3	43.5	58.8	88.9
SegNet[2]	68.7	52.0	87.0	58.5	13.4	86.2	25.3	17.9	16.0	60.5	24.8	46.4	62.5
B. Segnet [3]	-											63.1	86.9
FCN-8 [4]	77.8	71.0	88.7	76.1	32.7	91.2	41.7	24.4	19.9	72.7	31.0	57.0	88.0
Dilation-8 [5]	82.6	76.2	89.0	84.0	46.9	92.2	56.3	35.8	23.4	75.3	55.5	65.3	79.0
DeepLab-L [6]	81.5	74.6	89.0	82.2	42.3	92.2	48.4	27.2	14.3	75.4	50.1	61.6	-
FCN-Comb[7]	79.7	77.2	85.7	86.1	45.3	94.9	45.9	69.0	25.2	86.2	57.9	-	88.8
Proposed Approach	93.2	86.4	94.5	86.7	31.9	96.3	26.9	71.5	27.1	83.7	62.2	64.1	89.8



Results



EXAMPLES OF PIXEL-WISE SEGMENTATION RESULTS ON THE CAMVID DATASET. THE DIFFERENCES BETWEEN THE GROUND TRUTH LABELS AND PREDICTED LABELS ARE HIGHLIGHTED USING CYAN AND RED AREAS.



Results

RESULTS OF THE ABLATION STUDY ON CAMVID DATASET.

	Conv. Architecture	ACM	SCM	Proposed Architecture
Building	0.78	✓	✓	0.93
Tree	0.87	✓	✓	0.86
Sky	0.93	✓	✓	0.94
Car	0.90	✓	✓	0.87
Sign	0.79	✓	✓	0.32
Road	0.93	✓	✓	0.96
Pedestrian	0.87	✓	✓	0.27
Fence	0.81	✓	✓	0.72
Pole	0.71	✓	✓	0.40
Side Walk	0.88	✓	✓	0.83
Cyclist	0.88	✓	✓	0.62
Dice Score	-			0.65
Jaccard Score	-			0.63
Weighted Jaccard Score	-			0.80
Global Accuracy	-			0.87
Dice Score	-	✓	✓	0.67
Jaccard Score	-	✓	✓	0.64
Weighted Jaccard Score	-	✓	✓	0.81
Global Accuracy	-	✓	✓	0.89



Results

RESULTS OF THE ABLATION STUDY ON CAMVID DATASET.

	Conv. Architecture	ACM	SCM	Proposed Architecture
Building	0.78	✓	✓	0.93
Tree	0.87	✓	✓	0.86
Sky	0.93	✓	✓	0.94
Car	0.90	✓	✓	0.87
Sign	0.79	✓	✓	0.32
Road	0.93	✓	✓	0.96
Pedestrian	0.87	✓	✓	0.27
Fence	0.81	✓	✓	0.72
Pole	0.71	✓	✓	0.40
Side Walk	0.88	✓	✓	0.83
Cyclist	0.88	✓	✓	0.62
Dice Score	-			0.65
Jaccard Score	-			0.63
Weighted Jaccard Score	-			0.80
Global Accuracy	-			0.87
Dice Score	-	✓	✓	0.67
Jaccard Score	-	✓	✓	0.64
Weighted Jaccard Score	-	✓	✓	0.81
Global Accuracy	-	✓	✓	0.89



Thank You.

Day 1 – Monday 29 November

Poster Session 1 – 3-Minutes Spotlight

14:10-14:50 QLD Standard Time

