ELANENET: USING LANE PARAMETERS FOR BETTER DETECTION OF LANES IN AUTONOMOUS DRIVING SYSTEMS

12/12/2023

Elikem Buertey, Kshirasagar Naik, Nitin Naik, Sriram Sivaraman

Presenter: Elikem Buertey

Affiliation: University of Waterloo
Contact email: ebuertey@uwaterloo.ca

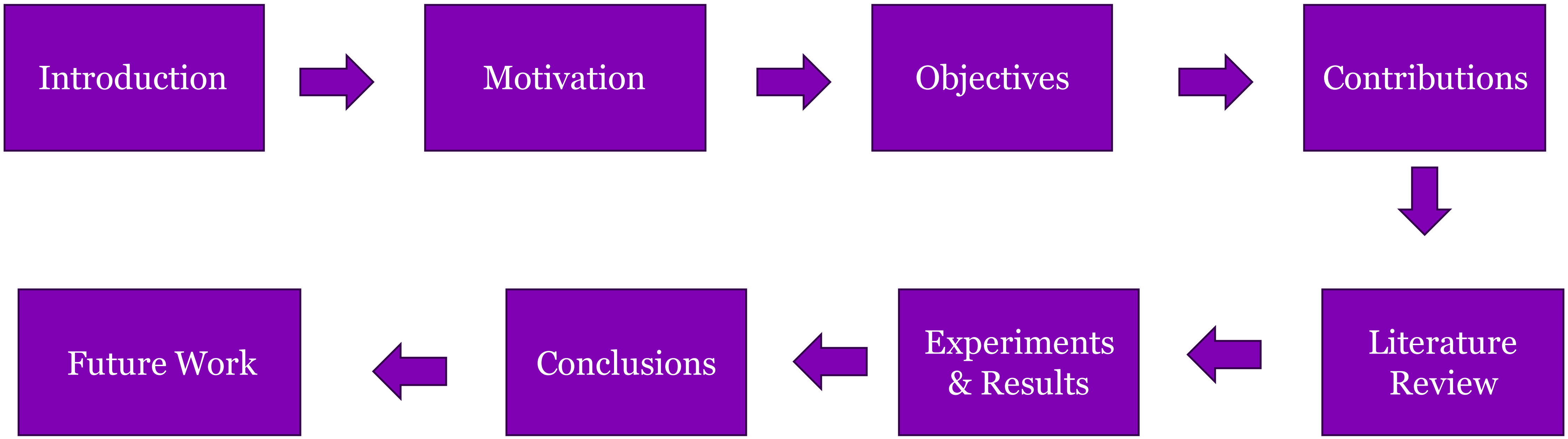
UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

IARIA

ELIKEM BUERTEY

- **Elikem Buertey** received his bachelor's degree in Electrical and Electronic Engineering from the University of Mines and Technology, Ghana in 2019. He is currently a masters student specializing in AI and pattern matching at the University of Waterloo, Ontario, Canada.
- His research interest lies in the intersection of artificial intelligence (particularly, computer vision), and autonomous vehicles.

OUTLINE


Introduction
Motivation
Objectives
Contributions
Literature Review
Experiments & Results
Conclusions
Future Work

INTRODUCTION: MOTIVATION

- According to the World Health Organization, approximately 1.3 million people lose their lives due to road traffic crashes, and between 20 and 50 million more suffer non-fatal injuries, with many enduring disabilities[1].
- 94% of these accidents are caused by human error, highlighting the potential for significant reduction if human error could be minimized [2].
- The anticipated benefits of autonomous vehicles include crash prevention, reduced travel times, improved fuel efficiency, and parking benefits, with estimated savings of up to $2000 per year per autonomous vehicle and potentially reaching nearly $4000 when considering comprehensive crash costs [3].

INTRODUCTION: AUTONOMOUS VEHICLE SYSTEM MODEL

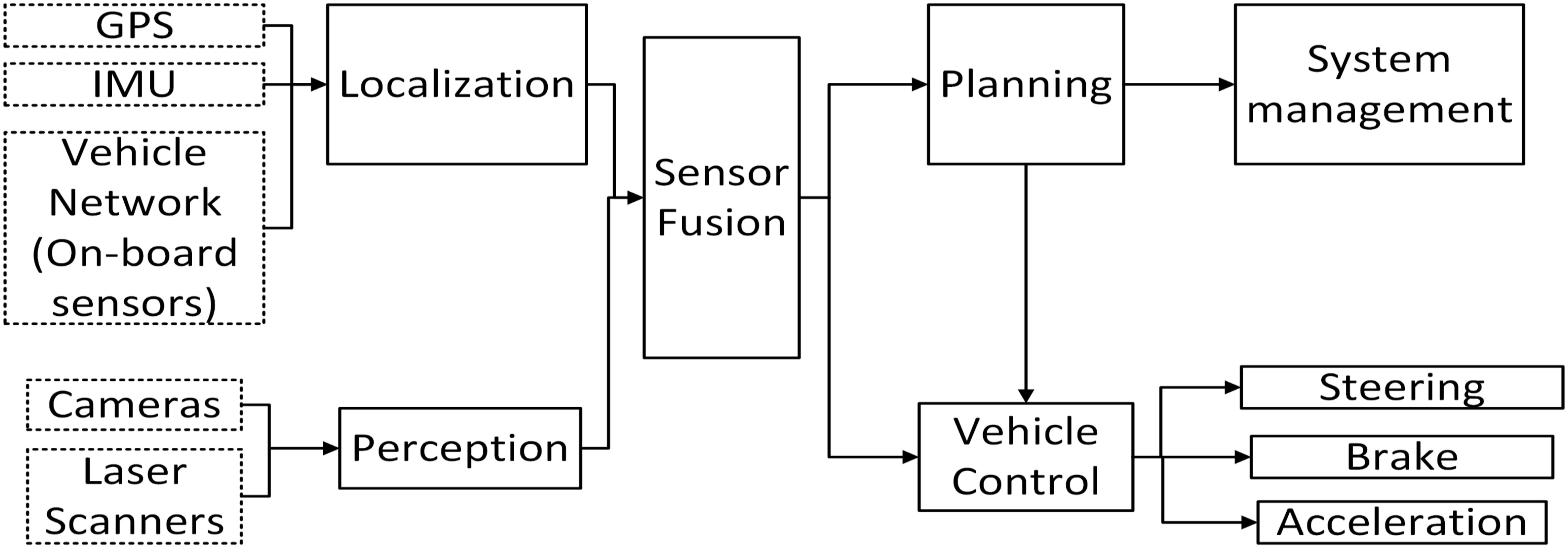


Figure 1: Autonomous Vehicle System Model

INTRODUCTION: CONTRIBUTIONS

- We propose ELaneNet as a new approach to address lane detection challenges.
- We introduce capacity, lost capacity and unsafe driving measure as performance metrics since they are more specific to lane detection than general metrics such as recall.
- We propose that a lane abstracting method instead of the conventional line abstracting method should be used to assess the performance of lane detection algorithms.

INTRODUCTION: OBJECTIVES

- Implement and Understand the LaneNet Model
- Improve on LaneNet's Detection Capabilities
- Evaluate the performance of LaneNet and ELaneNet
- Visualize and Analyze the Lane Detection Results
- Introduce new lane detection evaluation metrics

LITERATURE REVIEW: LANE DETECTION METHODS

- Traditional Methods
- Deep-Learning Methods

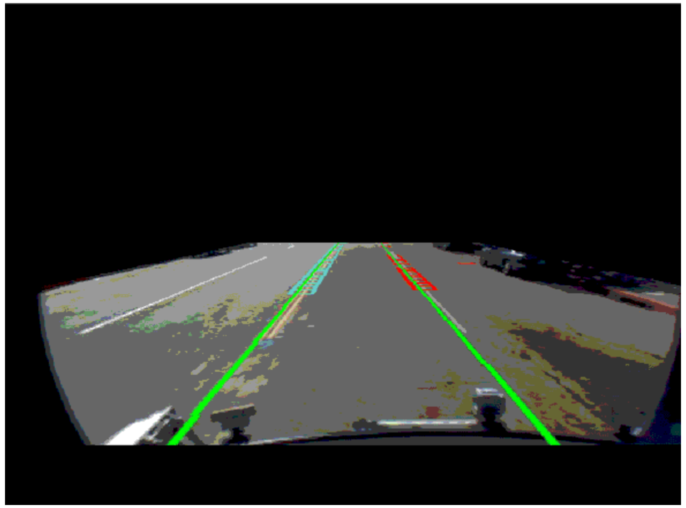
LITERATURE REVIEW: TRADITIONAL METHODS

- The lane detection procedure involves four main stages:
  - Image preprocessing
  - Feature extraction
  - Lane fitting
  - Lane tracking

LITERATURE REVIEW: IMAGE PREPROCESSING

GRAY SCALE CONVERSION

BLURRING OF IMAGE

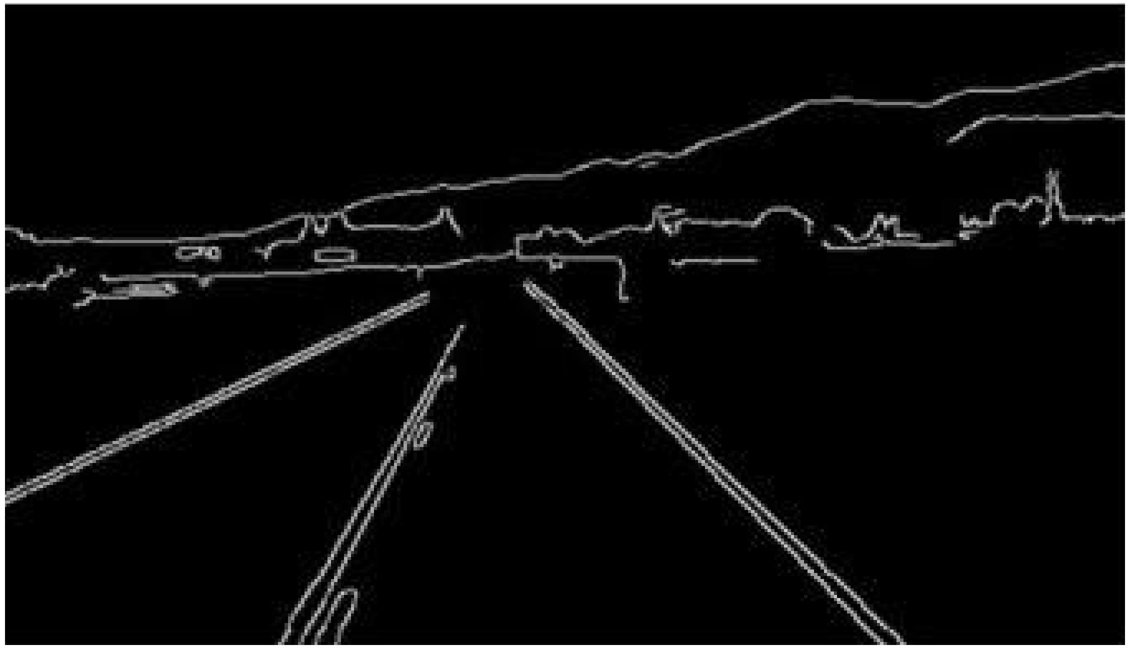


ROI SELECTION

Figure 2: Sample image preprocessing techniques

LITERATURE REVIEW: FEATURE EXTRACTION

INPUT IMAGE



EDGE DETECTION

Figure 3: Feature extraction using Edge detection

LITERATURE REVIEW: LANE TRACKING

Figure 4: Advantage of lane tracking in poor lighting

LITERATURE REVIEW: DEEP-LEARNING METHODS

- Encoder-decoder CNN
- FCN with optimization algorithms
- CNN+RNN
- GAN model

LITERATURE REVIEW: ENCODER DECODER CNN

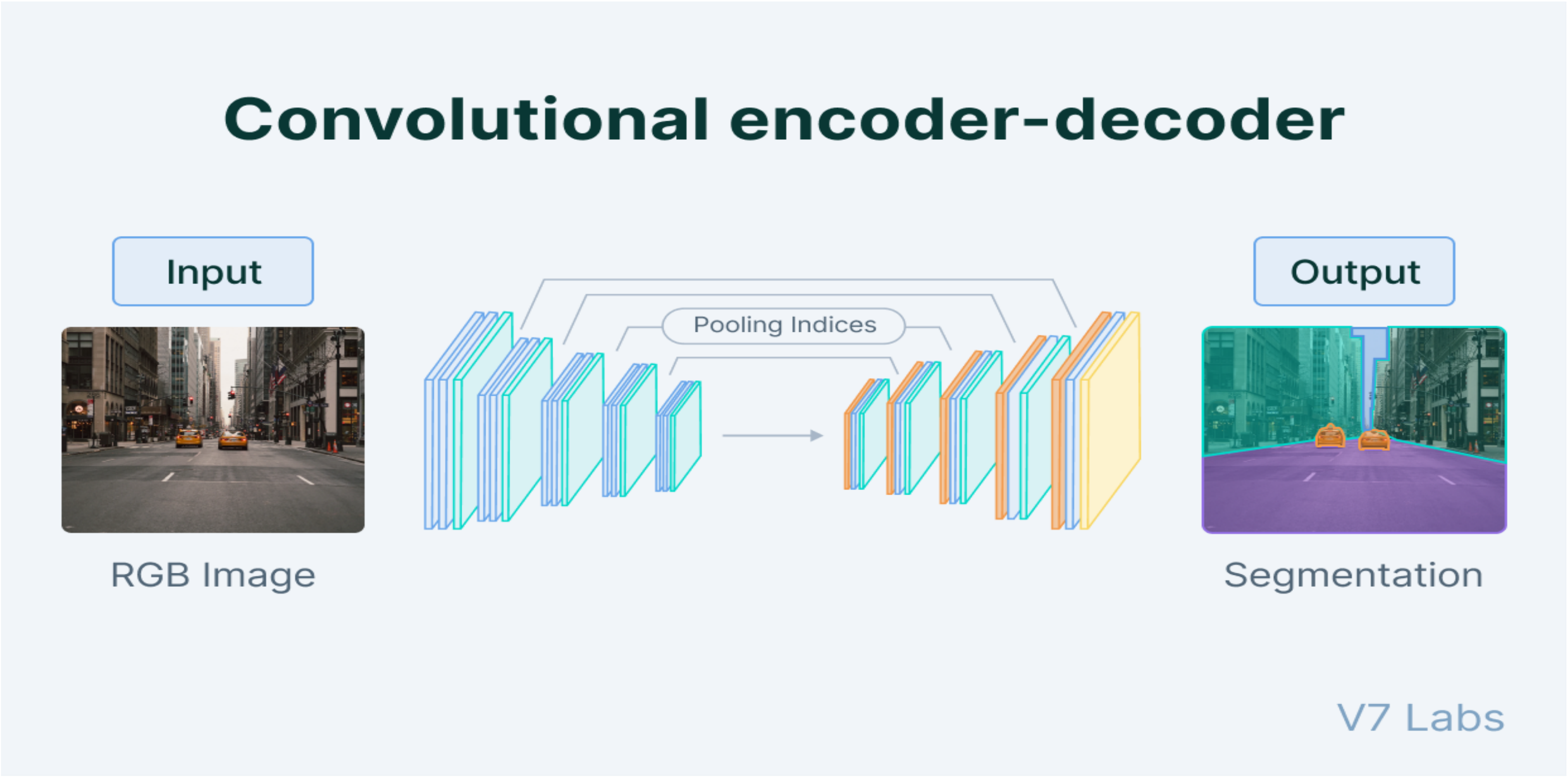


Figure 5: Convolutional encoder-decoder

LITERATURE REVIEW: DATASETS

- TuSimple
- BDD100K
- CULane
- Unsupervised LLAMAS

METHODOLOGY: LaneNet MODEL

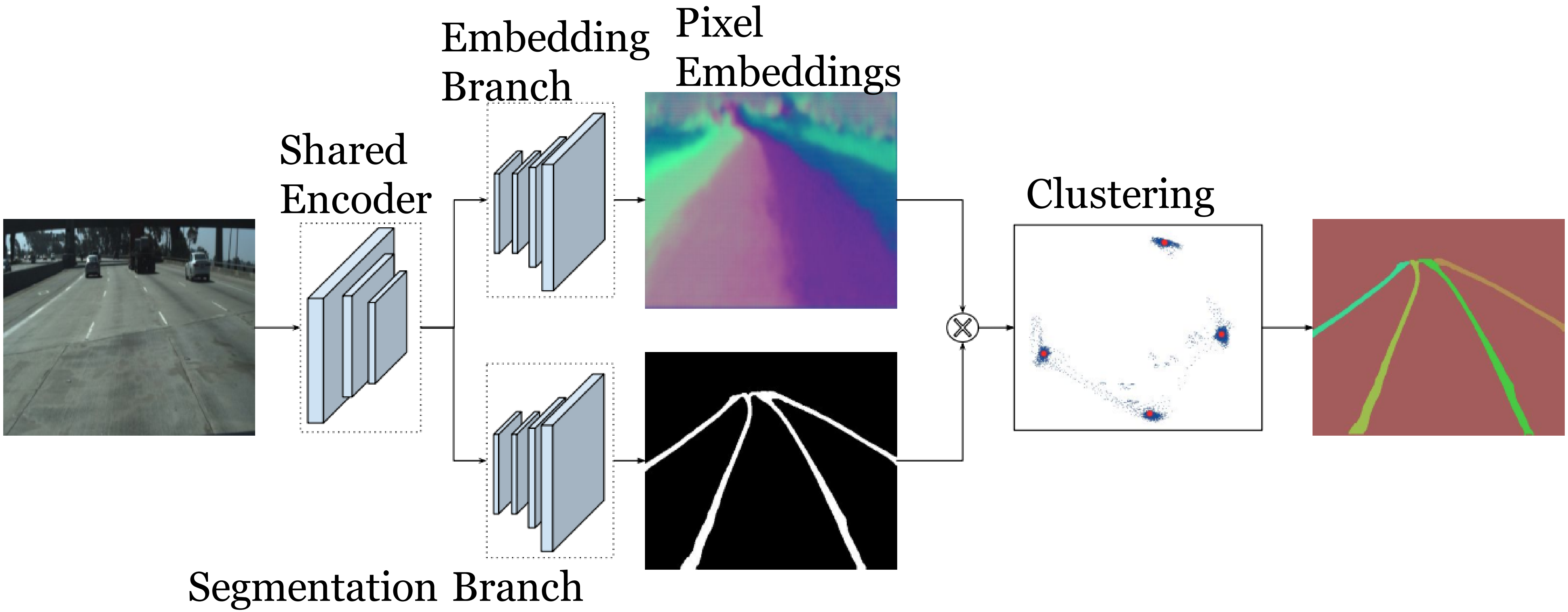


Figure 6: LaneNet model

METHODOLOGY: LaneNet MODEL

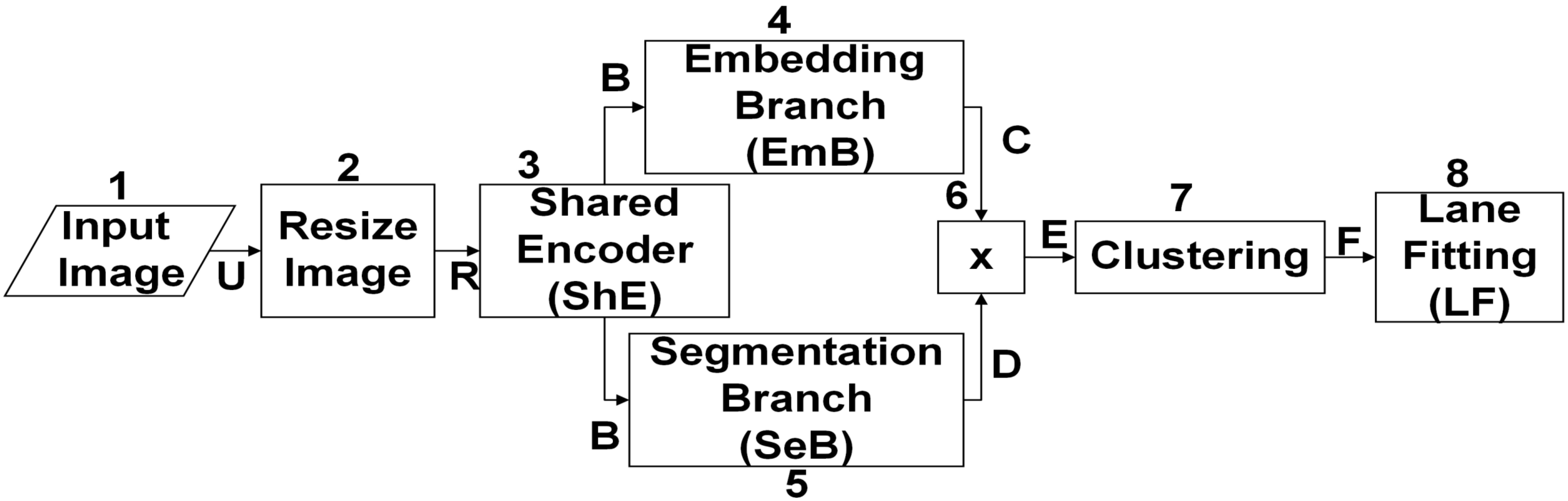


Figure 7: LaneNet model

METHODOLOGY: LaneNet MODEL

U = Input Image ($\alpha m \times \beta n \times c$)

R = Resized image ($m \times n \times c$)

B = Encoded image ($m \times n \times c$)

C = Pixel embeddings ($m \times n \times N$)

D = Binary lane segmentation ($m \times n$)

E = Lane embeddings ($m \times n \times N$)

F = Lane instance embeddings ($p \times 2$)

W = Splines ($q \times 1$)

Where $\alpha m, \beta n, m, n, N, p, q, c \in \mathbb{N}$

METHODOLOGY: INPUT IMAGE

- Input Image: In the image processing pipeline, input images are resized from their original resolution of $\alpha m \times \beta n \times c$ pixels to a reduced resolution of $m \times n \times c$ where $\alpha m, \beta n, m, n, c \in \mathbb{N}$.

METHODOLOGY: SHARED ENCODER

- **Shared Encoder:** Two modifications to ENet's architecture was introduced in LaneNet's shared encoder.
- Firstly, the output of ENet was adapted to create a two-branched network, accommodating both binary segmentation and instance segmentation.
- Secondly, in LaneNet, only the first two stages (stages 1 and 2) of ENet's encoder are shared between the two branches, while the full ENet decoder (stages 4 and 5) serves as the backbone for each separate branch. This means that stage 3 of ENet's encoder is not used in LaneNet.

METHODOLOGY: SHARED ENCODER

- **Shared Encoder:**

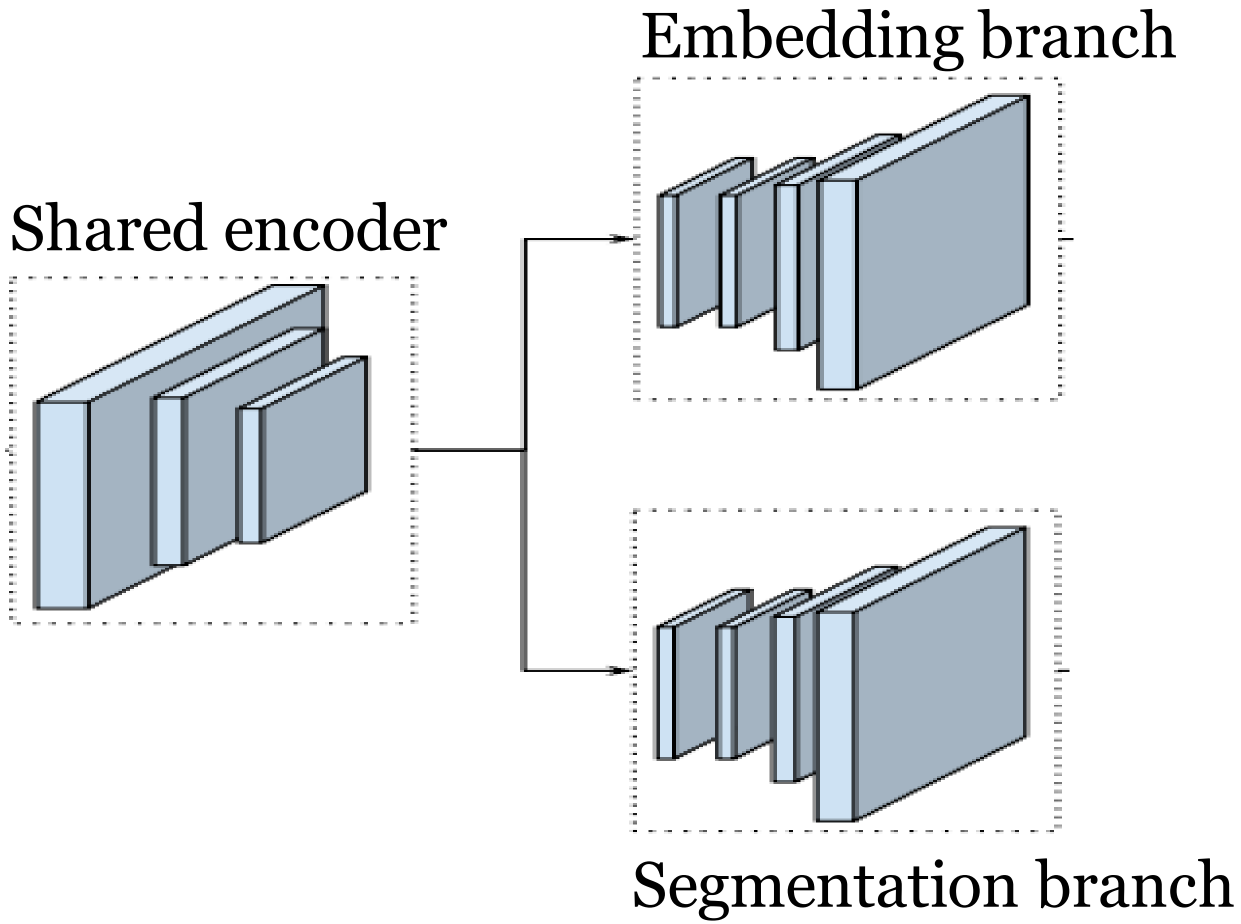


Figure 8: Shared Encoder

METHODOLOGY: LaneNet MODEL

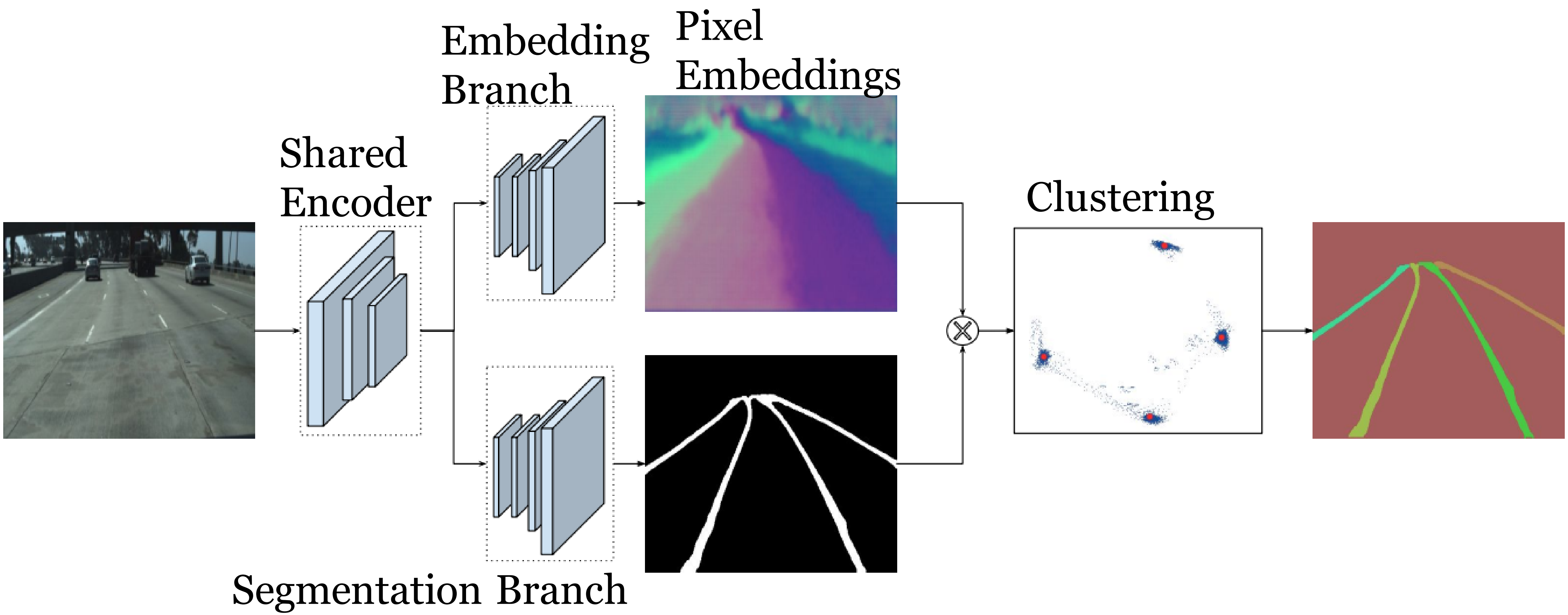


Figure 9: LaneNet model

METHODOLOGY: EMBEDDING VISUALISATION

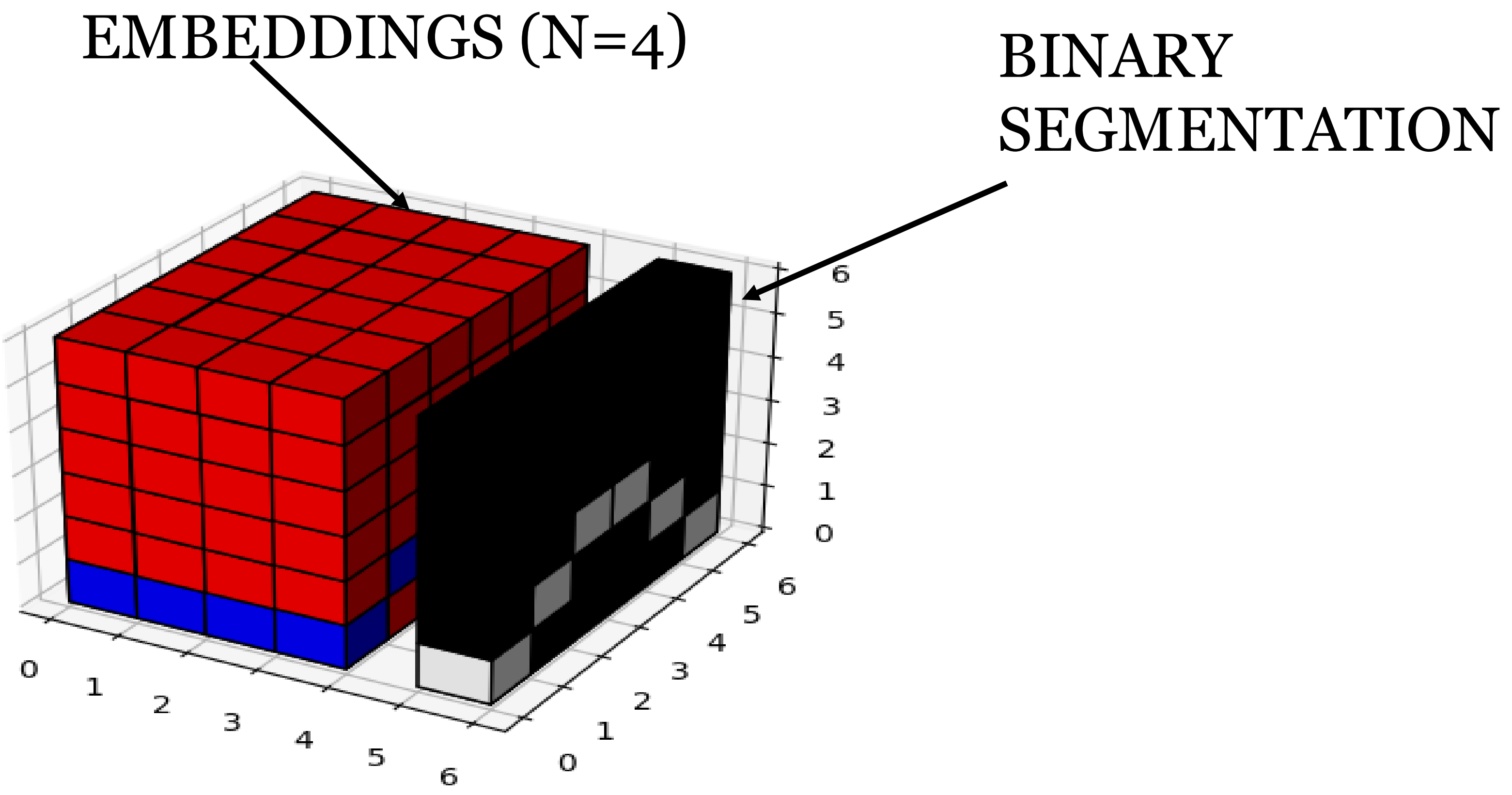


Figure 10: Visualization of embeddings

METHODOLOGY: SEGMENTATION BRANCH

- **Segmentation Branch**: Segmentation branch of the network is designed to produce a binary segmentation map which classifies the pixels into either lane or background categories. The class weighted cross entropy loss[4] is used to account for imbalance between the lane pixels and the background pixels. As stated earlier, the output of the segmentation branch is a binary segmentation map which classifies pixels into either lane or background. Since the background pixels far exceed the lane pixels, there is an imbalance between the lane pixels and the background pixels. To address this, the class weighted cross entropy loss[4] is used to account for the imbalance between the lane pixels and the background pixels.

METHODOLOGY: PRODUCT AND CLUSTERING

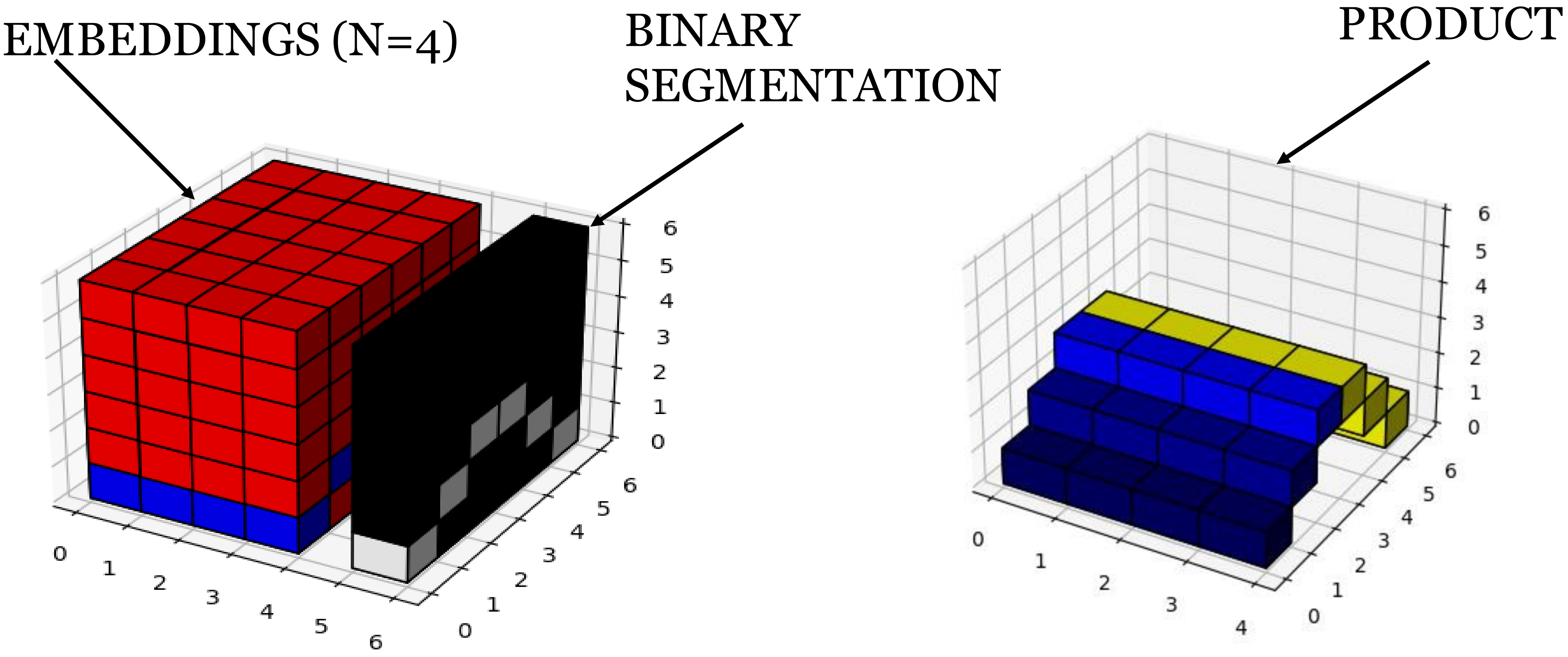


Figure 11: Visualization of lane embeddings

METHODOLOGY: EMBEDDING BRANCH

- Embedding Branch: Embeddings produced by the embedding branch have the characteristic that lane pixels belonging to the same lane have similar embeddings while lane pixels belonging to different lanes have different embeddings.

$$L_v = \frac{1}{K} \sum_{k=1}^{K} \frac{1}{N_k} \sum_{i=1}^{N} [|| \mu_k - x_i || - \delta_v]_+^2$$

$$L_d = \frac{1}{K(K-1)} \sum_{k_A=1}^{K} \sum_{k_B=1 \, k_A \neq k_B}^{K} ||\sigma_d - ||\mu_{kA} - \mu_{kB}||$$

METHODOLOGY: EMBEDDING BRANCH

δ_d is the minimum distance allowed between cluster centers.

δ_v is the maximum distance allowed between an embedding and the mean embedding of its corresponding cluster.

K denote the number of clusters (lanes)

N_k the number of elements in cluster k where $1 \leq k \leq K$

x_i is the i^{th} pixel embedding in cluster k

μ_k denotes the embedding of cluster k

$[x]_+$= max (0, x) the hinge

The total loss L is equal to $L_v + L_d$

IARIA

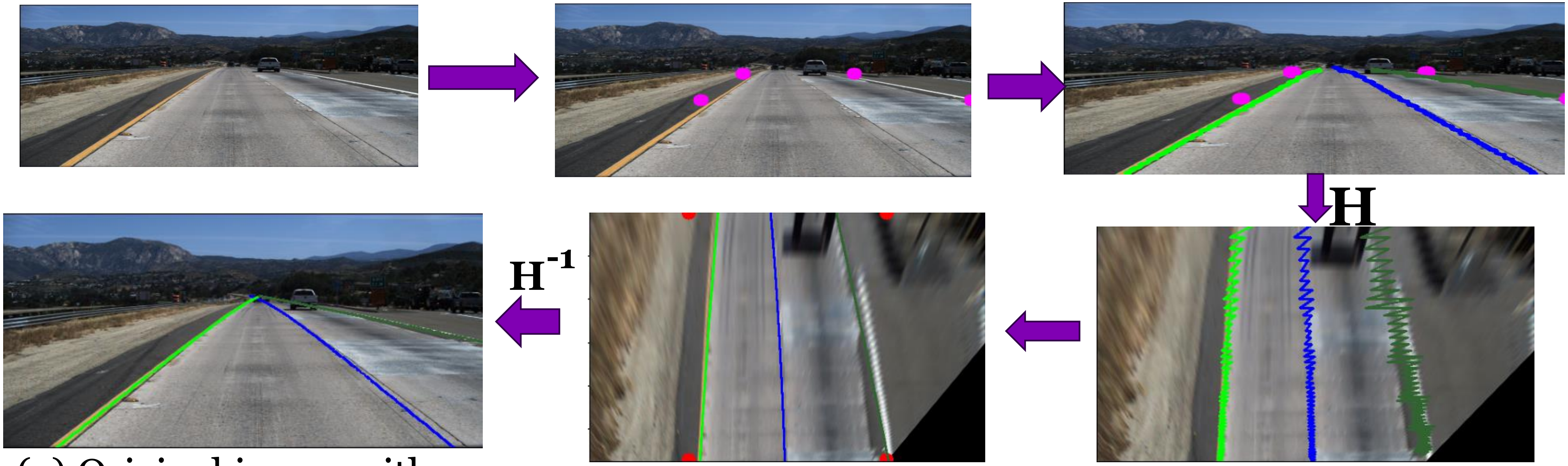
UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

METHODOLOGY: LANE FITTING

(a) Input Image

(b) Selection of ROI in image

(c) Original image with predicted lanes



(g) Original image with fitted lanes

(e) Transformed image with lane fitted

(d) Image transformed to BEV

Figure 12: Lane Fitting

IARIA

UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

METHODOLOGY: LANE FITTING

- Given N ground-truth lane points $p_i = [x_i,\ y_i,\ 1]^T \in P$
- Assume H is the transformation matrix, we transform the points using $P' = HP$
- Through these projected points we fit a polynomial $f(y) = \alpha y'^2 + \beta y' + \gamma$

$$\mathbf{w} = (\mathbf{Y}^T\mathbf{Y})^{-1}\mathbf{Y}^T\mathbf{x}'$$

with $\mathbf{w} = [\alpha, \beta, \gamma]^T$, $\mathbf{x}' = [x'_1, x'_2, ..., x'_N]^T$ and

$$\mathbf{Y} = \begin{bmatrix} y_1'^2 & y_1' & 1 \\ \vdots & \vdots & \vdots \\ y_N'^2 & y_N' & 1 \end{bmatrix}$$

METHODOLOGY: ENHANCED LANENET

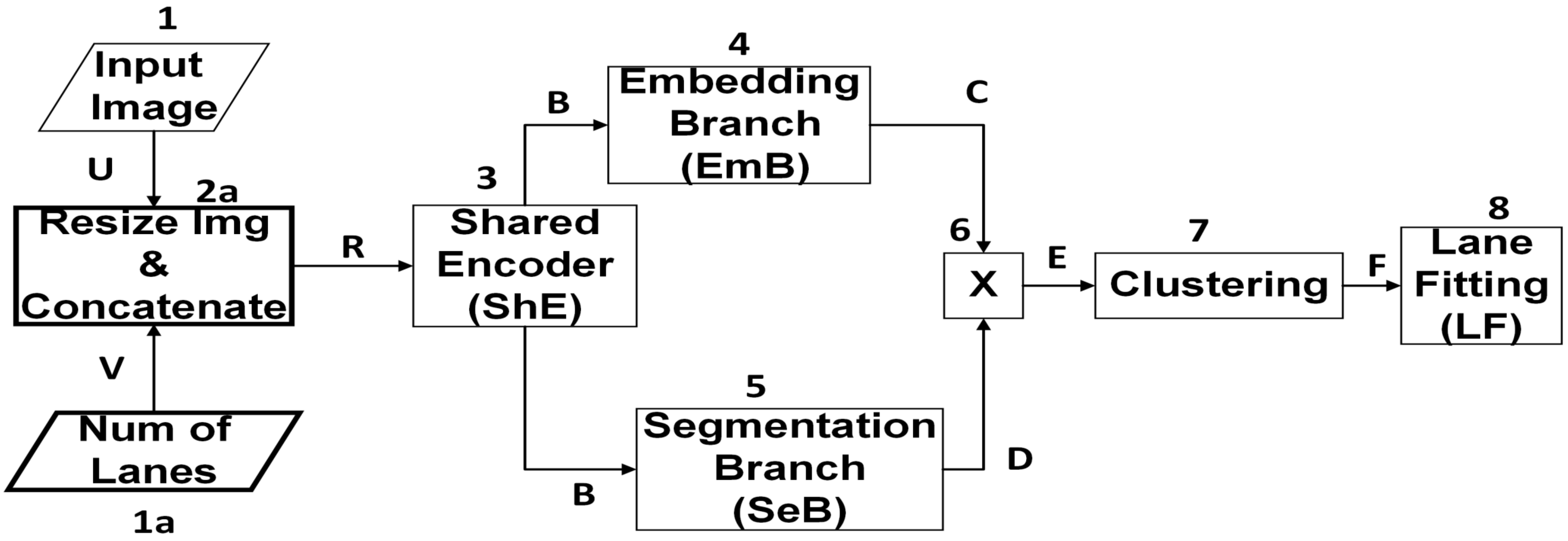


Figure 13: Enhanced LaneNet Network

METHODOLOGY: ENHANCED LANENET

U = Input Image ($\alpha m \times \beta n \times c$)
V = One-hot encoded number of lanes ($1 \times mc$)
B = Encoded image ($m \times n \times c$)
C = Pixel embeddings ($m \times n \times N$)
D = Binary lane segmentation ($m \times n$)
R = Resize Img & Concatenate ($m \times (n-1) \times c$)
E = Lane embeddings ($m \times n \times N$)
F = Lane instance embeddings ($p \times 2$)
W = Splines ($q \times 1$)
Where $\alpha m, \beta n, m, n, N, V, p, q, c \in \mathbb{N}$

METHODOLOGY: LANE ENCODING

INDEX	1	2	3	4	5	6
1	1	0	0	0	0	0
2	0	1	0	0	0	0
3	0	0	1	0	0	0
4	0	0	0	1	0	0
5	0	0	0	0	1	0
6	0	0	0	0	0	1

Table 1: Examples of One-Hot Encoding

METHODOLOGY: RESIZE IMAGE AND CONCATENATE

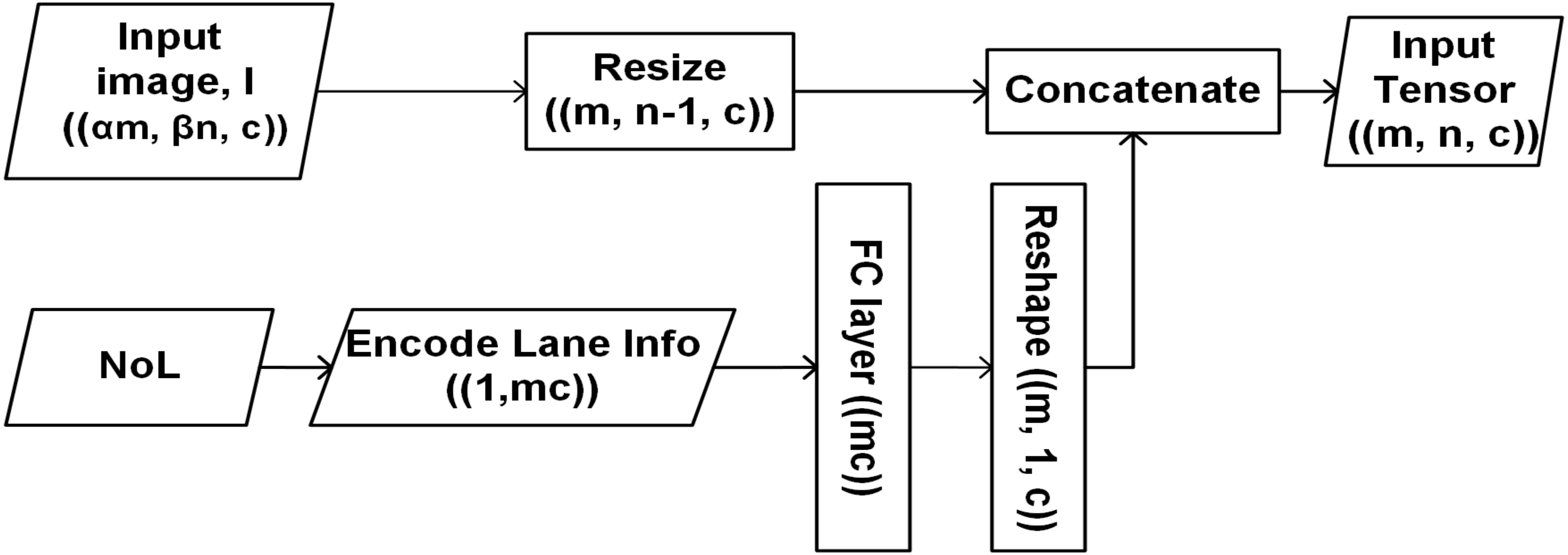


Figure 14: Concatenating image with lanes

IARIA

UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

METHODOLOGY: RESIZE IMAGE AND CONCATENATE

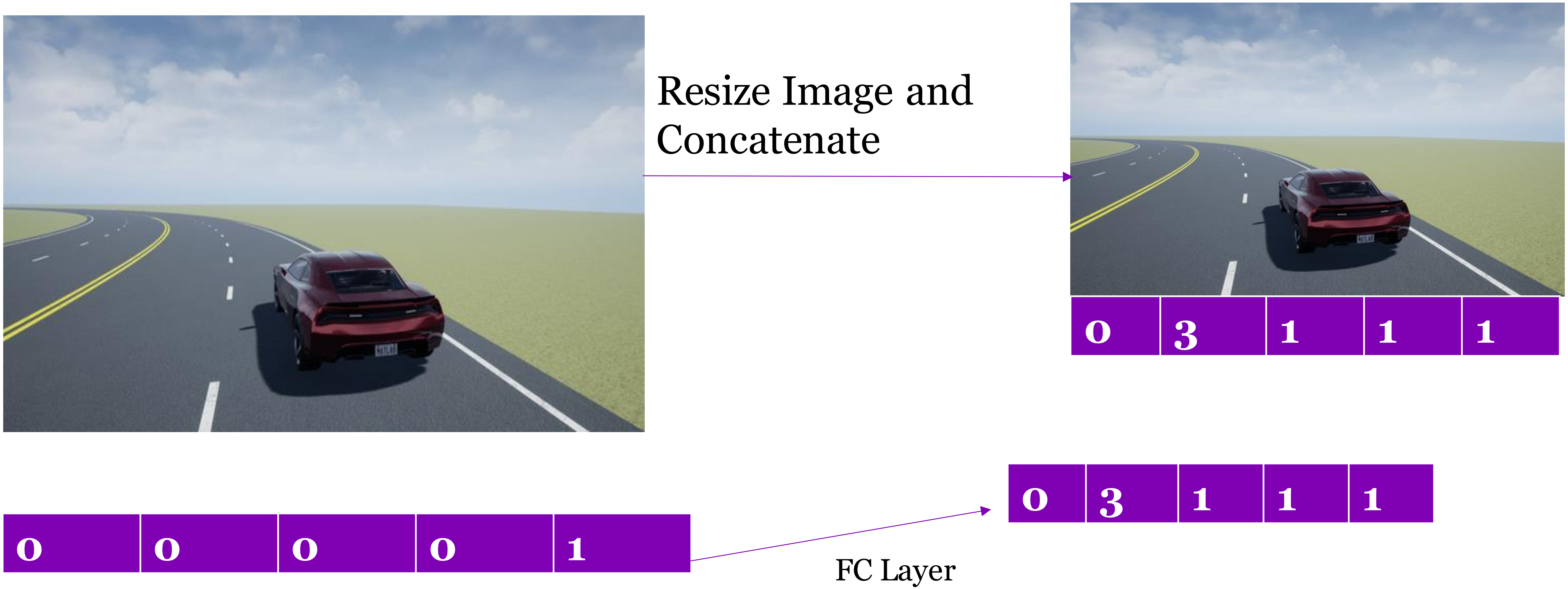


Figure 15: Sample concatenating of Images

METHODOLOGY: FULLY CONNECTED LAYER

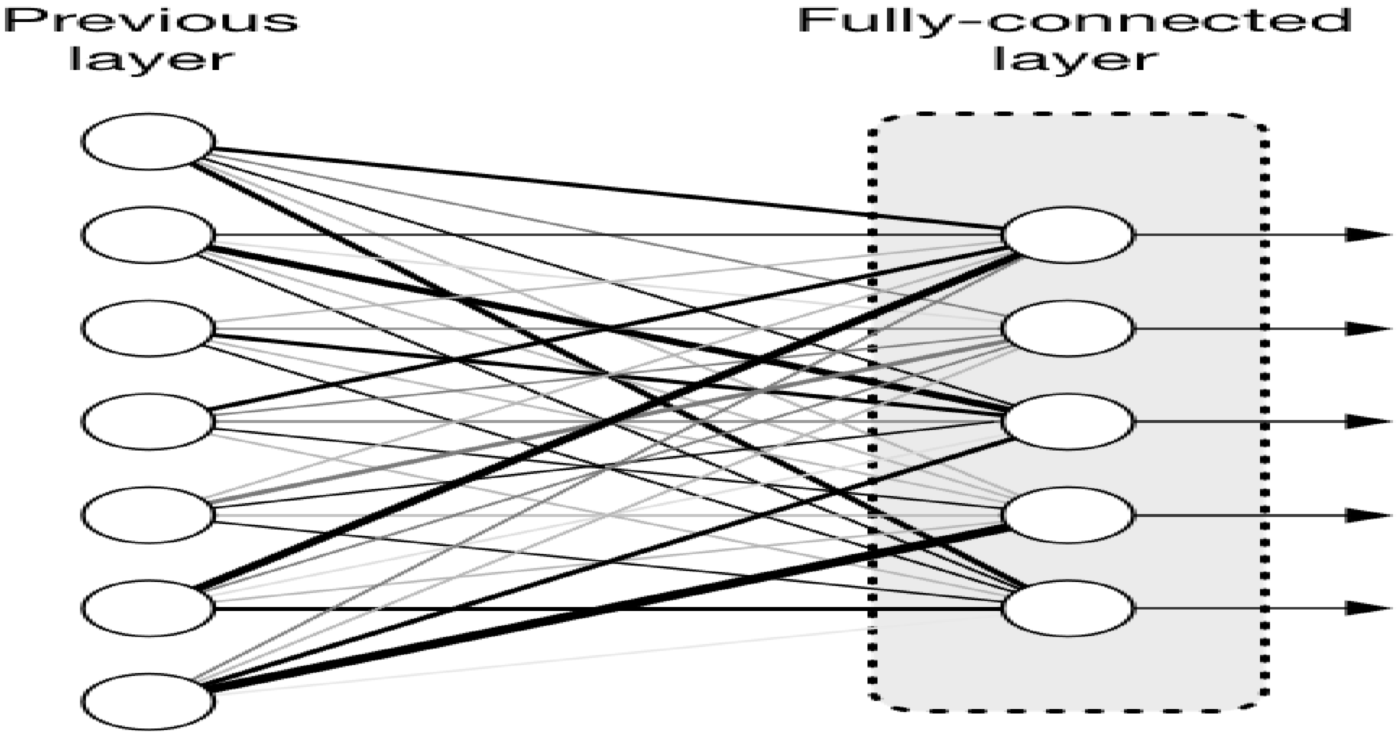


Figure 16: Fully connected layer

METHODOLOGY: CAPACITY

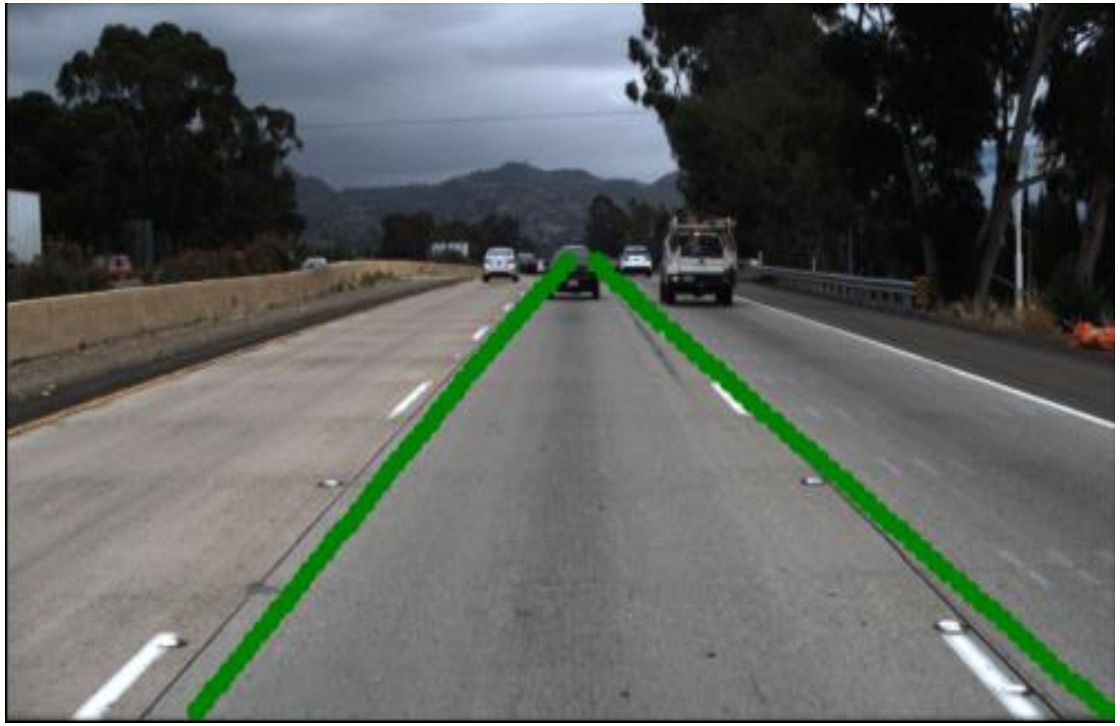


Figure 1: Autonomous system with lower capacity

Figure 2: Autonomous system with higher capacity

Figure 17: Capacity illustration

METHODOLOGY: CAPACITY

$$Capacity = \frac{Detected\ lane\ lines}{Num\ of\ ground\ truth\ lane\ lines}$$

$$\boldsymbol{Capacity = \frac{TP}{TP + FN}}$$

$$Lost\ Capacity = 1 - Capacity$$

METHODOLOGY: UNSAFE DRIVING MEASURE

Figure 18: False positive lane detection causes unsafe driving

METHODOLOGY: UNSAFE DRIVING MEASURE

$$Unsafe\ Driving\ Measure = \frac{total\ number\ of falsely\ predicted\ lane\ lines}{total\ number\ of\ correctly\ predicted\ lane\ lines}$$

$$Unsafe\ Driving\ Measure = \frac{F_{pred}}{N_{pred}}$$

METHODOLOGY: LANE ABSTRACTION APPROACH

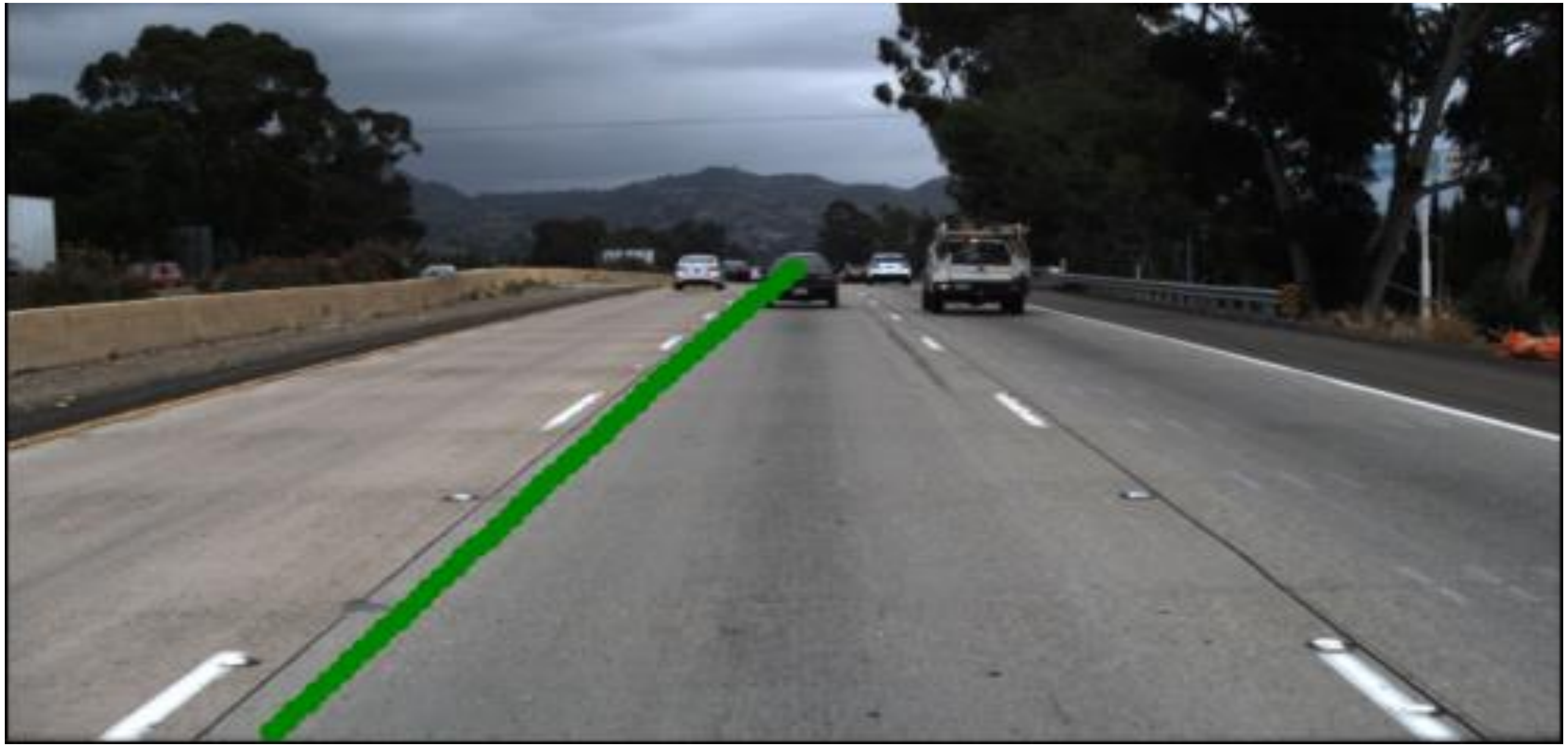


Figure 19: Driving scene illustrating the problem with the lane abstraction approach

METHODOLOGY: UPDATED EXPRESSIONS

$$Capacity^{L} = \frac{TP^{L}}{TP^{L} + FN^{L}}$$

$$Unsafe\ Driving\ Measure^{L} = \frac{F_{pred}^{L}}{N_{pred}^{L}}$$

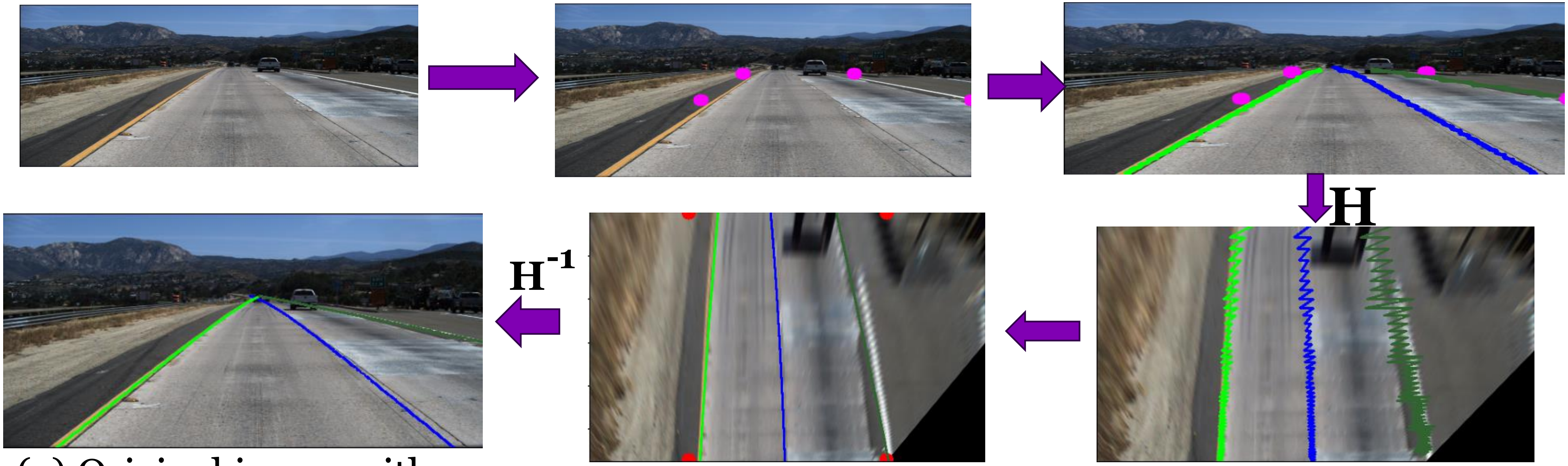
$$Lost\ Capacity^{L} = 1 - Capacity^{L}$$

METHODOLOGY: LANE FITTING

(a) Input Image

(b) Selection of ROI in image

(c) Original image with predicted lanes



(g) Original image with fitted lanes

(e) Transformed image with lane fitted

(d) Image transformed to BEV

Figure 20: Lane Fitting

EXPERIMENTS AND RESULTS: LANE ABSTRACTION

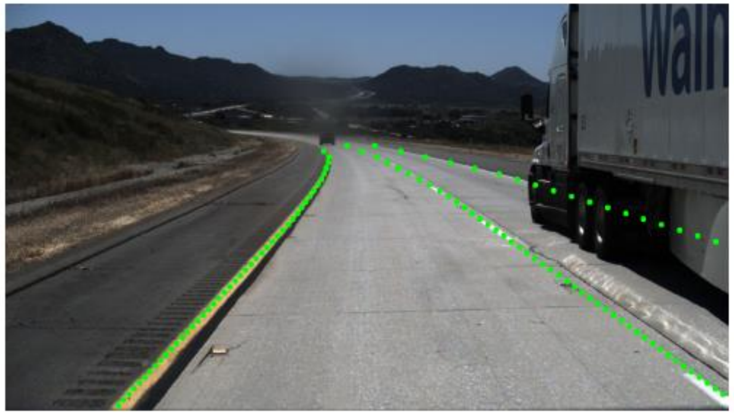
LANE ABSTRACTION			
NETWORK	USED CAPACITY	LOST CAPACITY	UNSAFE DRIV.
ELaneNet	87.5 %	12.5 %	27.3 %
LaneNet	80.4 %	19.6 %	38.5 %

LINE ABSTRACTION				
NETWORK	USED CAPACITY	LOST CAPA CITY	UNSAFE DRIV.	ACCURACY
ELaneNet	93.1 %	6.9 %	13.9 %	94.5 %
LaneNet	88.9 %	11.1 %	23.0 %	92.3 %

EXPERIMENTS AND RESULTS: SPEED METRICS

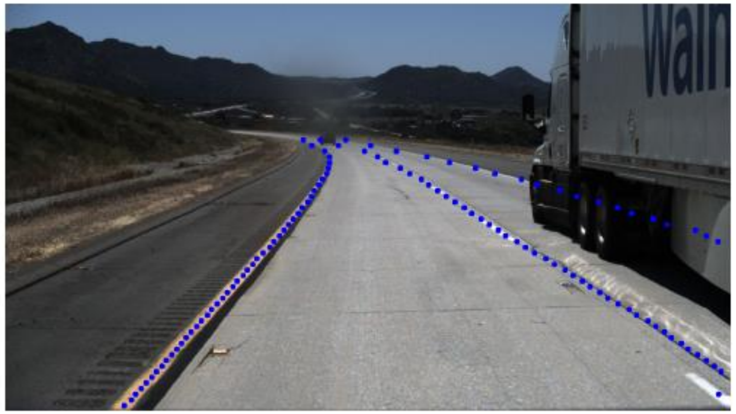
Metric	LaneNet	ELaneNet
Forward pass time per image (ms)	43.5	51.6
Clustering time per image (ms)	231.8	232.8
Total time per image (ms)	275.34	284.4

EXPERIMENTS AND RESULTS: VISUALISATION



(a) Ground truth

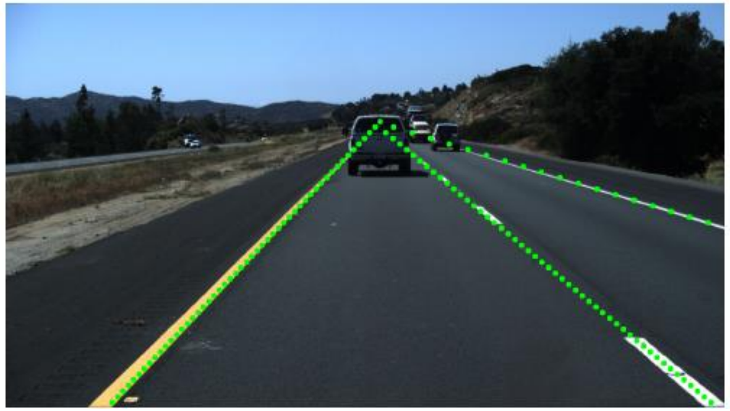
(b) LaneNet detects false positive



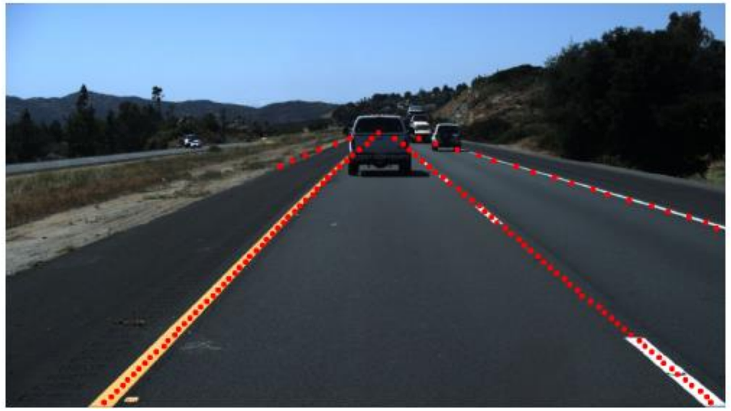
(c) ELaneNet does not detect false positive

Figure 16: Visualization of results

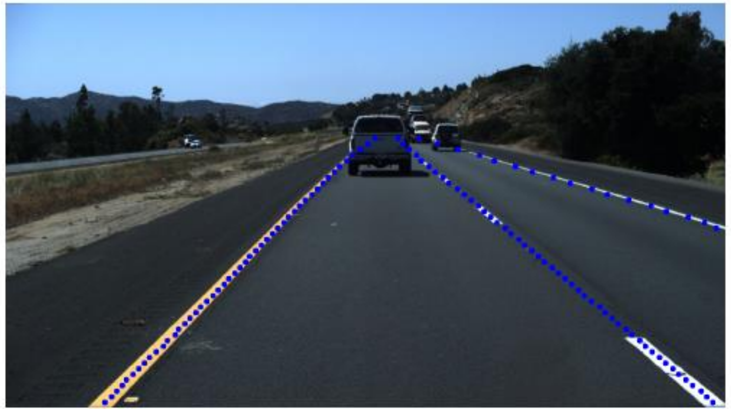
EXPERIMENTS AND RESULTS: VISUALISATION



a) Ground truth



b) LaneNet detects false positive



c) ELaneNet does not detect false positive

Figure 17: Visualization of results

EXPERIMENTS AND RESULTS: VISUALISATION

a) Ground truth

b) LaneNet misses a lane

c) ELaneNet identifies missed lane

Figure 18: Visualization of results

IARIA

UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

FUTURE WORK

- We plan to further enhance eLaneNet by using the NoL to extrapolate missing lanes and eliminate false positives.
- Also, we plan on using other datasets to evaluate the effectiveness of eLaneNet.

UNIVERSITY OF WATERLOO

FACULTY OF ENGINEERING

YOU+WATERLOO

Our greatest impact happens together.

REFERENCES

- National Highway Traffic Safety Administration, “Traffic safety facts 2015 data: Pedestrians,” https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812115, U.S. Department of Transportation, Tech. Rep., 2015
- World Health Organization. Road traffic injuries. https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries, 2021
- Daniel J. Fagnant and Kara Kockelman. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. Transportation Research Part A: Policy and Practice, 77:167–181, 2015.

IARIA UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING

REFERENCES

- A. Paszke, A. Chaurasia, S. Kim, and E. Culurciello, “Enet: A deep neural network architecture for real-time semantic segmentation,” CoRR, vol. abs/1606.02147, 06 2016.

UNIVERSITY OF WATERLOO | FACULTY OF ENGINEERING