Lane Detection using CNN

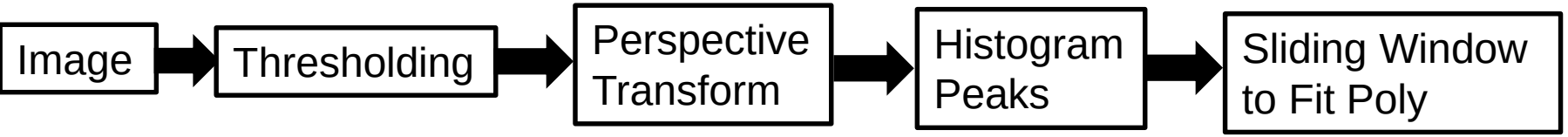
Mohammad Karami

Supervisor: Taufiq Rahman, Stephen Sweeny

June 2020

Lane Detection Techniques

1. Computer vision techniques:



Note: CV algorithm is challenging as the perspective transform needs dst pixel cords

2. Convolution neural network (Shallow network):

Lane Detection by CNN

- Labeling image data sets can be done in two different ways.
- Providing a set of coefficients that fit a polynomial equation on the lanes.
- The problem in this technique is that a small loss in the starting or ending point can have significant loss on the final model accuracy.
- Another technique is fitting a set of pixel coordinates (x,y) as a label for each image
- Here, the second method is selected.
- The CNN is applied on the available TuSimple DataSet. Link

Lane Detection by CNN

- Image data sets captured on the highway
- Challenging data set since not all images contain lanes!
- According to the paper, the current accuracy is 67.2!

Shallow Regression CNN

- Although a shallow network is used, the model was still able to learn the lane patterns in the image.

- The accuracy in regression CNN is different from what is defined for classification problem and needs to be defined manually as a call back function

Deep Neural Network

AlexNet Architecture

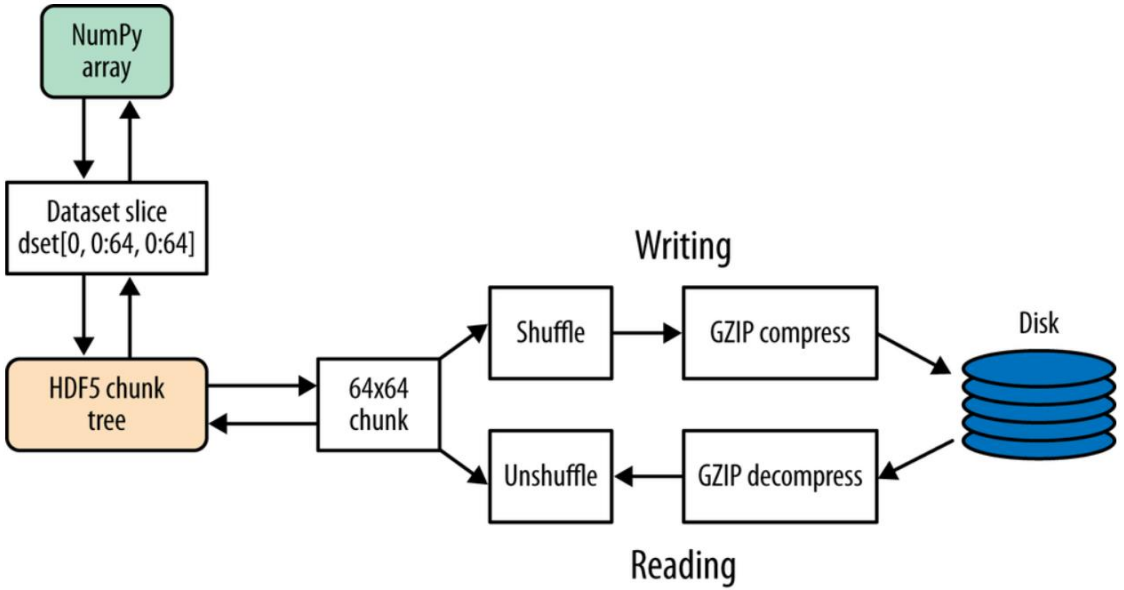
- AlexNet won the ImageNet competition in 2012
- Accepts only images with size 227*227
- Last layer consists of 64 nodes corresponding to the pixel coordinates of lanes
- AlexNet has more layers than the Shallownet, thus, we are not able to feed all images into the RAM. Solution?!

Layer Type	Output Size	Filter Size / Stride
INPUT IMAGE	$227 \times 227 \times 3$	
CONV	$57 \times 57 \times 96$	$11 \times 11/4 \times 4, K = 96$
ACT	$57 \times 57 \times 96$	
BN	$57 \times 57 \times 96$	
POOL	$16 \times 16 \times 96$	$3 \times 3/2 \times 2$
DROPOUT	$28 \times 28 \times 96$	
CONV	$28 \times 28 \times 256$	$5 \times 5, K = 256$
ACT	$28 \times 28 \times 256$	
BN	$28 \times 28 \times 256$	
POOL	$13 \times 13 \times 256$	$3 \times 3/2 \times 2$
DROPOUT	$13 \times 13 \times 256$	
CONV	$13 \times 13 \times 384$	$3 \times 3, K = 384$
ACT	$13 \times 13 \times 384$	
BN	$13 \times 13 \times 384$	
CONV	$13 \times 13 \times 384$	$3 \times 3, K = 384$
ACT	$13 \times 13 \times 384$	
BN	$13 \times 13 \times 384$	
CONV	$13 \times 13 \times 256$	$3 \times 3, K = 256$
ACT	$13 \times 13 \times 256$	
BN	$13 \times 13 \times 256$	
POOL	$13 \times 13 \times 256$	$3 \times 3/2 \times 2$
DROPOUT	$6 \times 6 \times 256$	
FC	4096	
ACT	4096	
BN	4096	
DROPOUT	4096	
FC	4096	
ACT	4096	
BN	4096	
DROPOUT	4096	

Large Data Sets and HDF5 Format

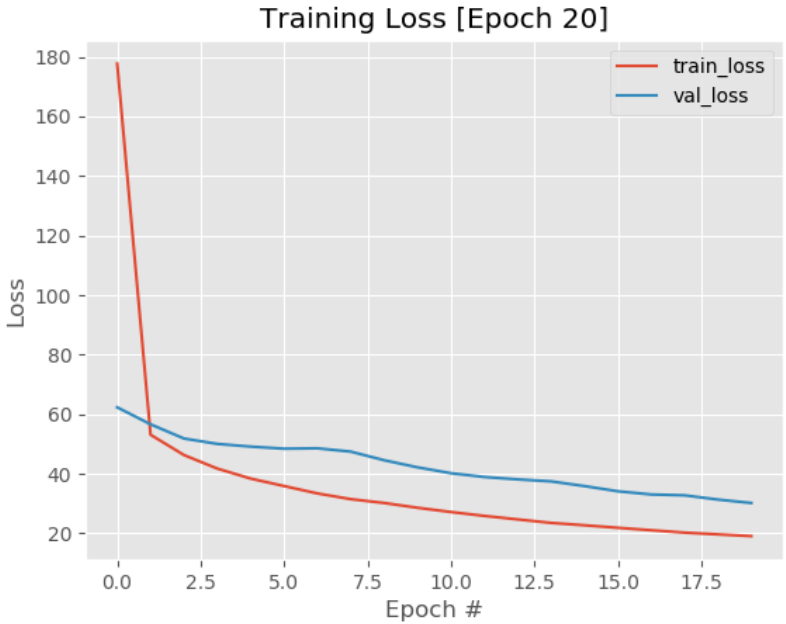
- When large data sets can not fed into the RAM, we convert them into HDF5 format and saved them on the hard disk.

- Optimized for I/O operations as well as sliced of data is reachable



Results

- Source of error: CNN is learned by the pixel coordinate rather than the actual feature
- Lanes are barely visible even by our eyes!

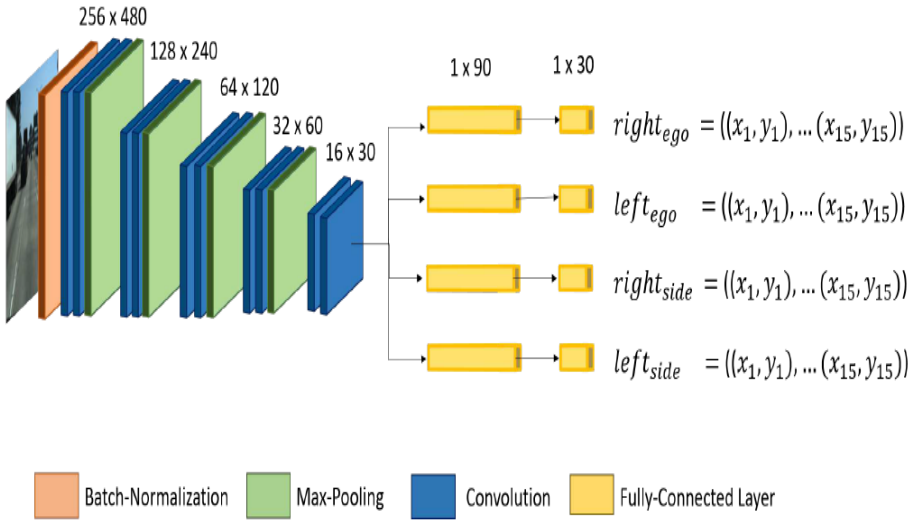


Data as of June 2019

Updated Work

Classification and Regression

- Courtesy of Chougule et al. (2018)
- Architecture consists of encoder, classification, and finally regression



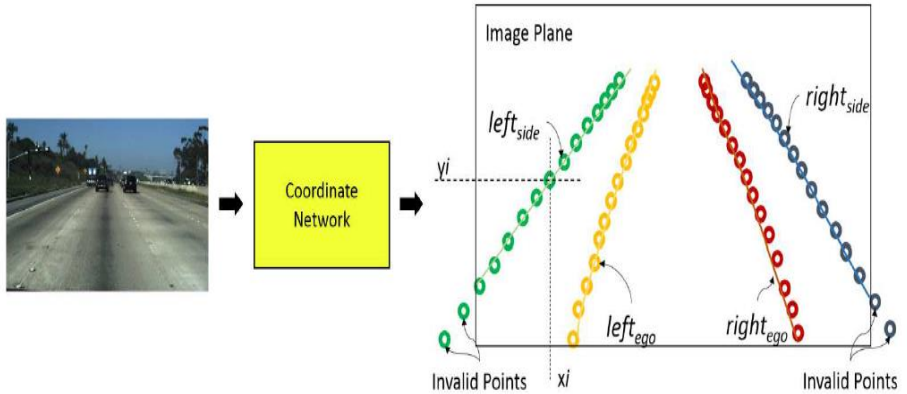
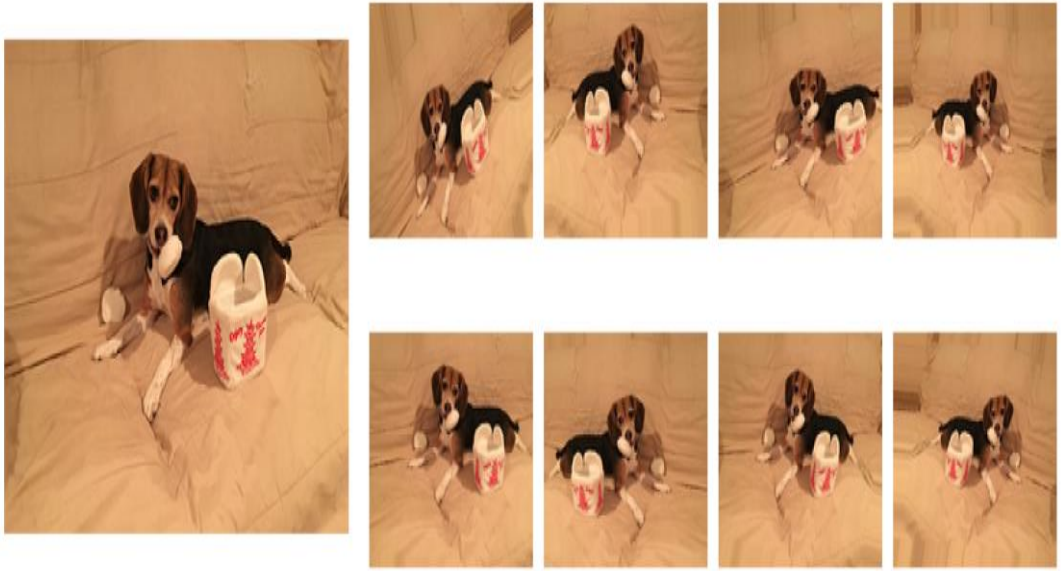


Fig. 3. Coordinate network outputs four lane vectors corresponding to four lane boundary types, each containing 15 locations along the predicted lane boundary

Data Augmentation

- Generate new data set by rotation, cropping, zooming or …
- Keras has a built-in function for data augmentation, but image label should remain the same!
- I had to write down a custom function for transforming labels for each image

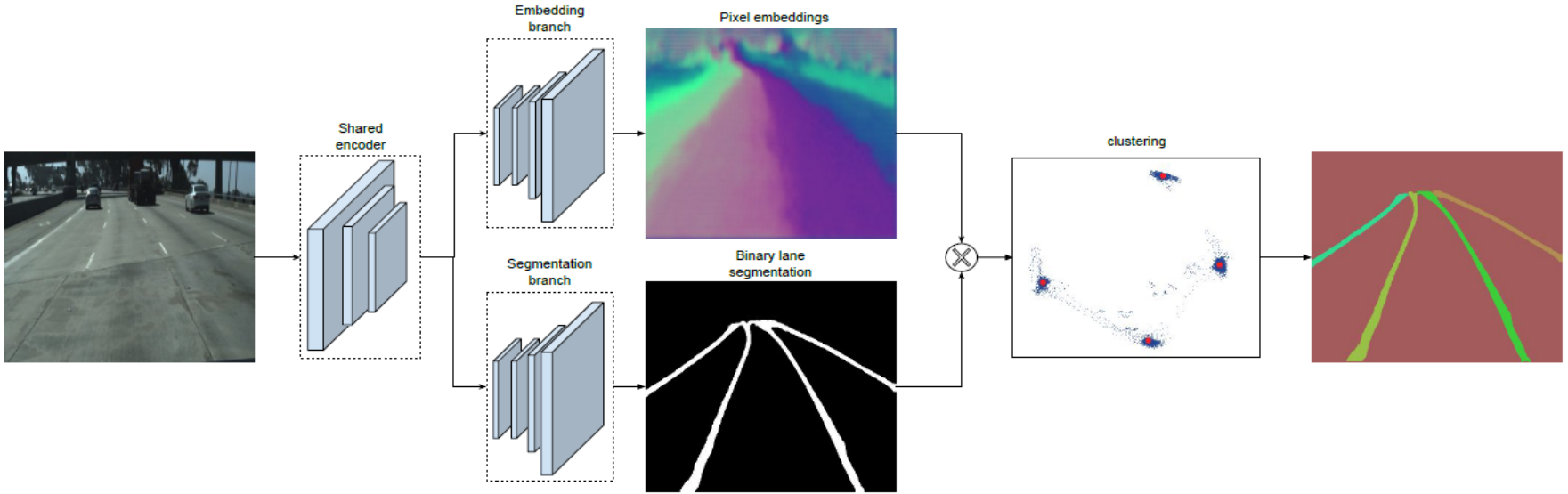


Limitation of Classification/Regression

- Cannot handle NAN values or different numbers of lanes
- Common limitation in many deep learning models

LaneNet Architecture

- Neven et al. (2018)
- Can handle: handles lane changes and allows the inference of an arbitrary number of lanes



THANK YOU

Mohammad Karami • COOP Student, NRC

Mohammad.karami@nrc-cnrc.gc.ca