Object Classification using Convolutional Neural Network(CNN) for Advanced Driver Assistance Systems (ADAS).

Sai Sravan Manne
EDM13B018

Guided by:
Dr. Binsu J Kailath

Points discussed in previous review:

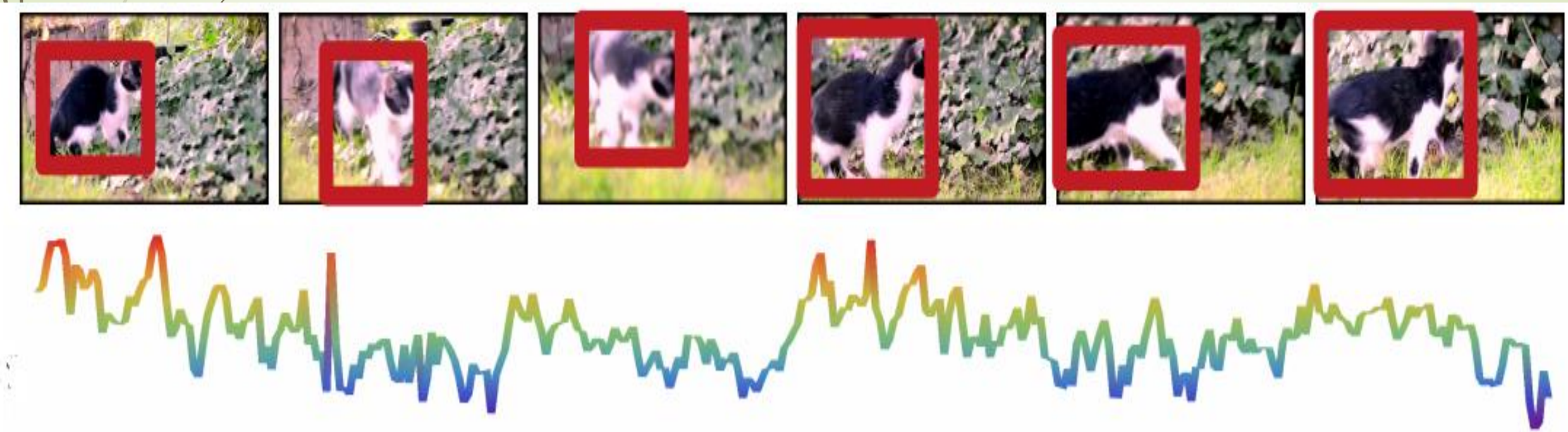
- What is ADAS?
- Which one is better, camera or LIDAR (Light Detection and Ranging) ?
- Why CNN?
- Architecture of CNN?

Today's Discussion:

- Performance issues associated with applying still image object detection CNN's on videos.
- Execution time analysis of Alexnet and VGG Net, and how to reduce it ?
- Problem with current method being used for object detection and classification from videos.
- Proposed solution for optimizing the execution time.
- Future work.

Performance issues associated with applying still image object detection CNN's on videos:

- Variable Confidence Level



Performance issues associated with applying still image object detection CNN's on videos:

- False Positives

Execution time analysis of Alexnet and VGG Net:

- Implemented on a 2GB GPU(Graphics Processing Unit)

Profile Summary

Generated 04-Oct-2016 08:59:41 using cpu time.

Function Name	Calls	**Total Time**	Self Time*	Total Time Plot (dark band = self time)
matconv	1	2.023 s	1.582 s	
run	1	0.282 s	0.005 s	
vl_setupnn	1	0.272 s	0.002 s	
addpath	5	0.262 s	0.007 s	
path	5	0.254 s	0.203 s	
vl_simplenn	1	0.069 s	0.009 s	
general\private\parsedirs	10	0.051 s	0.050 s	
graphics\private\clo	2	0.040 s	0.005 s	
vl_nnconv (MEX-file)	8	0.039 s	0.039 s	

- Execution time for Alexnet for a 384*512*3 image

Execution time analysis of Alexnet and VGG Net:

- Implemented on a 2GB GPU(Graphics Processing Unit)

Profile Summary

Generated 04-Oct-2016 08:24:33 using cpu time.

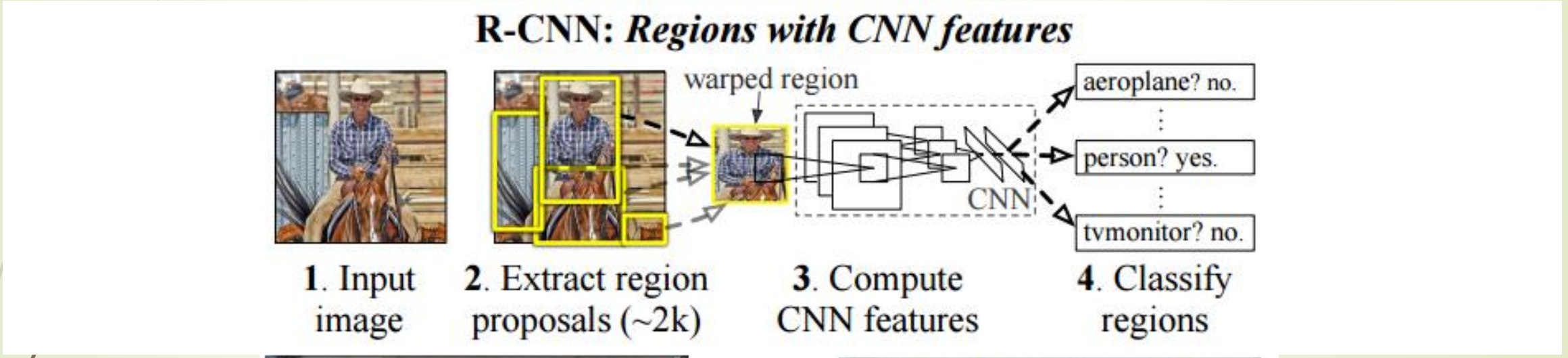
Function Name	Calls	**Total Time**	Self Time*	Total Time Plot (dark band = self time)
matconv	1	1.936 s	1.477 s	
run	1	0.289 s	0.010 s	
vl_setupnn	1	0.279 s	0.000 s	
addpath	5	0.279 s	0.010 s	
path	5	0.269 s	0.209 s	
vl_simplenn	1	0.070 s	0.000 s	
general\private\parsedirs	10	0.060 s	0.060 s	
graphics\private\clo	2	0.060 s	0.000 s	
setdiff	3	0.060 s	0.020 s	
setdiff>setdifflegacy	3	0.040 s	0.010 s	
vl_nnconv (MEX-file)	8	0.040 s	0.040 s	

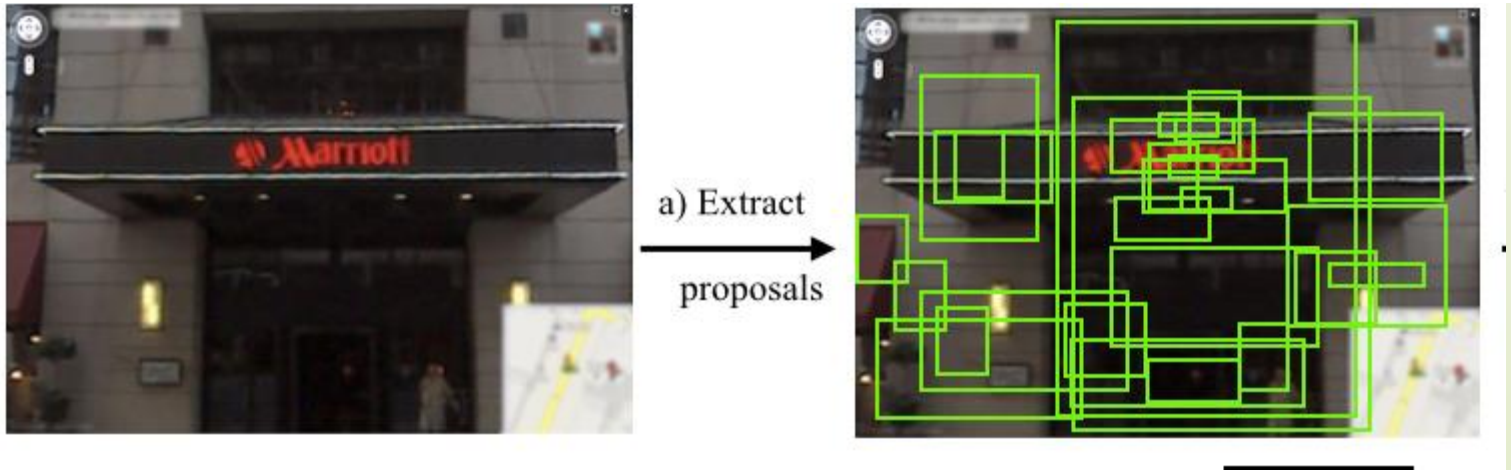
- Execution time for VGG Net for a 384*512*3 image

How to reduce the execution time period ?

- Using a powerful GPU(Graphics Processing Unit), but cost of a 2GB GPU is Rs 6,000 ~ 10,000.
- In-order to completely avoid or predict accidents, required FPS (frames per second) execution is 60~100, so each image needs to classified within 0.01~0.016th second.
- FPS of human eye is 24~40.

Current method being used for object detection and classification from videos:



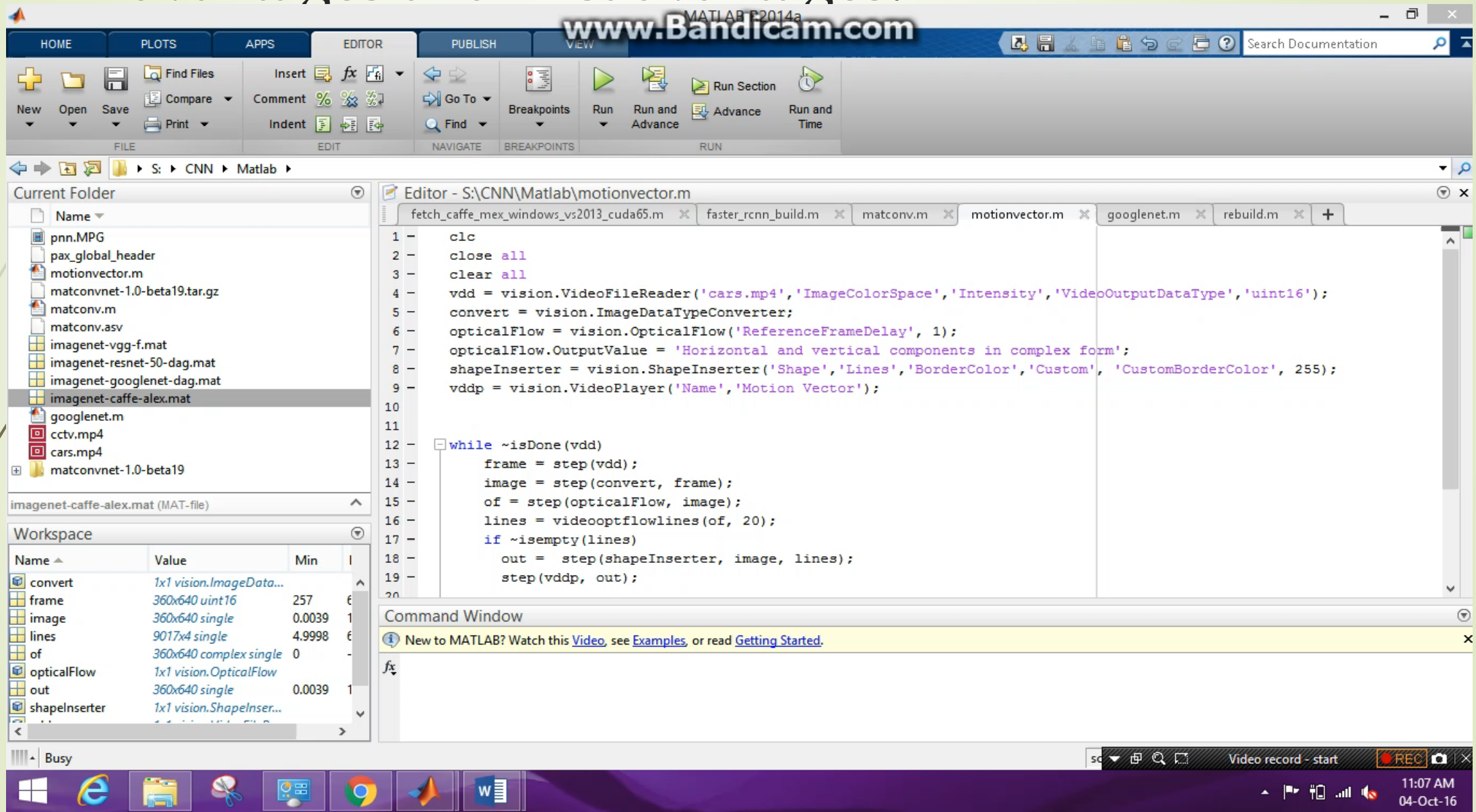


Proposed Solution:

- Velocity vector calculation for individual pixels.
- Can it be used for object detection in non-stationary environmental conditions ?

Result:

Advantages and Disadvantages:



Work to be done

1. The above mentioned method needs to be thoroughly tested to check its reliability.
2. Design, development and training of a new CNN for Indian road conditions, with incorporation of recent advancements in cost and back propagation algorithms, and also it can be applied to videos.
3. Path estimation for detected vehicles.

References

[1] Kai Kang, Hongsheng Li, Junjie Yan, Xingyu Zeng, Bin Yang, Tong Xiao, Cong Zhang, Zhe Wang, Ruohui Wang, Xiaogang Wang, and Wanli Ouyang, “T-CNN: Tubelets with Convolutional Neural Networks for Object Detection from Videos,” published in CVPR 2016.

[2] Andrej Karpathy, George Toderici, and Sanketh Shetty, “Large-scale Video Classification with Convolutional Neural Networks,” published in CVPR 2014.

[3] Sayanan Sivaraman, and Mohan Manubhai Trivedi, “Looking at Vehicles on the Road: A Survey of Vision-Based Vehicle Detection, Tracking, and Behavior Analysis,” IEEE Transactions on Intelligent Transportation Systems, vol. 14, no. 4, December 2013.

Thank You