HIGH PERFORMANCE CAMERA-BASED MODULES FOR AUTONOMOUS VEHICLES

Alaa Ali⋆ *Magdy A. Bayoumi*†

The Center for Advanced Computer Studies (CACS)
University of Louisiana at Lafayette, LA, USA
⋆ ama8030@louisiana.edu † mab@cacs.louisiana.edu

ABSTRACT

Vision based modules are essential to realize artificial intelligence robotics, smart traffic and driverless cars. Our research work focus on traffic sign recognition and object detection in context of vehicle and pedestrian detection. Several detection techniques are based on deformable part models (DPMs) and convolutional neural network (CNN) for its foreseen precision. We proposed multiple approaches to improve the performance of DPM and the more accurate DeepPyramid DPM (a mapped model into convolutional layers). Multiple techniques are employed to build our frameworks including but not limited to Fast Fourier transform, feature approximation, edge boxes, neighborhood aware cascade and region proposal networks Our fastest technique results show that we process Pascal VOC 2010 at 47HZ and VGA image at 40HZ with loss of 2% in precision.

***Index Terms*—** Object Detection, DPM, DeepPyramid, Autonomous Vehicles, ADAS

1. INTRODUCTION

Automobile manufacturers are in race nowadays towards fully autonomous vehicles and trucks. Self-driving encompass various challenges to interact with physical world using cameras, LIDAR, RADAR and GPS. High performance object detection, traffic sign recognition and lane detection are primary components for self-driving technologies.

Our main research problem is object detection for autonomous vehicles. We focus on one of the accurate state-of-the-art detectors. The deformable part model (DPM) [1] may includes multi-components of object category plus the main root model and multiple parts based on Histogram of Oriented Gradients (HOG). In addition, DeepPyramid DPM [2] presents a convolutional network mapping design to examine a significant power of DPM with better performance and outperform R-CNN detector (running 20x faster).

Our research work contribution includes a proposal for three approaches based on complementary modules to achieve a real-time DPM detector without sacrificing its precision. We present a comprehensive performance evaluation on different detection benchmarks like PASCAL VOC 2010 and KITTI car benchmarks. Our proposed approaches are designed to be generic and to detect class-independent objects.

2. RELATED WORK

Here, we discuss some recent object detection approaches related to DPM [1] and DeepPyramid DPM [2]. DPM is one of the most competitive detectors for its accuracy but with intensive computations. Fast Fourier Transform (FFT)[2] is employed by Dubout to speedup convolution step. Yan et al. introduced a faster DPM approach [3] which runs at 3.3 FPS (about 4x Cascade DPM) on Pascal VOC. On the other hand, Girshick introduced DeepPyramid DPM [2] based on convolutional layers to outperform original DPM [1] in terms of precision. Li Wan integrates ConvNets and DPM in a new end-to-end model to boost DPM's precision by $+13.2$.

3. METHODOLOGICAL APPROACH

Our proposed work is complementary and related to recent speedup techniques of DPM and DeepPyramid DPM as explained in the following three approaches.

First approach. We build a framework here to enhance the computation bottleneck in building feature pyramid and classification steps [4]. we integrate a famous fast feature pyramid technique, FFT scheme to replace convolutions and lookup table HOG features with an early classification method.

Second approach. We introduce a novel region based DPM pipeline [5] to create object region proposals and then enhance the feature map complexity and finally speedup the parts score computation step. Edge Boxes are employed for region proposals and integrated with fast feature map technique plus neighborhood aware cascade [3]. Thereafter, we introduce an optimized implementation for these two proposed approaches based on SIMD instructions, assembly programming and multi-threading to overcome time performance bottlenecks with maintained results in terms of precision.

[1]DeepDPM [2] code: https://github.com/rbgirshick/DeepPyramid.
[2]FFT-DPM code: http://www.idiap.ch/scientific-research/resources.

Table 1. Average-Precision (AP in %) of several approaches on Pascal VOC 2010.

	aero	bike	bird	boat	botl	bus	car	cat	chair	cow	table	dog	horse	motor	person	plant	sheep	sofa	train	tv	mAP	FPS
HOG-DPMv5 [1]	49.2	53.8	13.1	15.3	35.5	53.4	49.7	27	17.2	28.8	14.7	17.8	46.4	51.2	47.7	10.8	34.2	20.7	43.8	38.3	33.43	0.1
FF-DPM[2]	48.2	53.2	12.9	13.7	34.6	53.9	48.7	26.5	17.1	28.3	14.5	18.1	45.6	50.4	46.7	11.0	33.9	20.8	43.2	36.3	32.88	0.7
FastDPM [3][3]	47.2	54.4	12.9	14.7	34.8	52.8	49.2	27.1	16.8	28.7	14.3	17	45.9	50.6	46.9	10.8	34.1	20.7	43.9	37.6	33.02	3.3
DPM1 [4] (Ours)	48.1	53.4	12.2	13.8	34.2	52.4	48.7	26.2	16.1	27.8	13.2	16.4	45.1	50.2	46.3	9.9	33.4	19.8	42.6	37.2	32.35	45.5
R-DPM2 [5] (Ours)	47.9	53.3	12.3	13.5	34.1	52.1	48.3	26.3	16.2	27.5	12.9	16.2	45.3	50	46.1	10.1	33.2	19.5	42.4	37.4	32.23	47.6
DeepDPM [2][1]	44.6	65.3	32.7	24.7	35.1	54.3	56.5	40.4	26.3	49.4	43.2	41.0	61.0	55.7	53.7	25.5	47.0	39.8	47.9	59.2	45.20	1.6
R-CNN FT fc_7[5]	67.1	64.1	46.7	32.0	30.5	56.4	57.2	65.9	27.0	47.3	40.9	66.6	57.8	65.9	53.6	26.7	56.5	38.1	52.8	50.2	50.20	0.1
R-DPDPM3 (Ours)	44.2	65.1	32.6	24.7	35.2	54.1	56.2	39.9	26.4	48.0	43.1	40.7	60.5	55.1	53.3	24.9	46.8	39.6	47.5	58.7	44.83	17.2

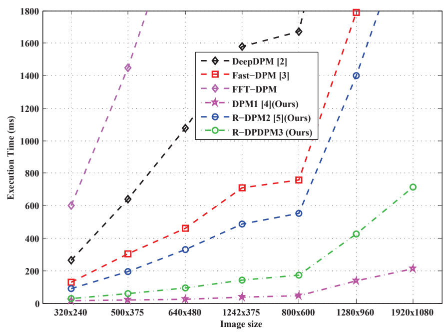

Fig. 1. Time performance based on input image resolution.

Third approach. Our work is related to DeepPyramid DPM [2] approach using two modules to accelerate DeepPyramid pipeline, we integrate RPN strategy inspired from Faster R-CNN[4] to exploit its shared computations feature for almost cost-free region proposals with another optimized FFT module as a replacement for convolution. Then we employ a GPU implementation for all steps to obtain better performance.

4. RESULTS AND DISCUSSION

Our research work is evaluated and discussed in this section, Table 1 shows a comparison of our proposed three methods in terms of average precision on 20 categories of Pascal VOC 2010 dataset. Overall, our proposed methods 'DPM1 [4]' and 'R-DPM2 [5]' produce comparable accuracy to original DPM and outperforms the other accelerated versions of HOG based DPM. Meanwhile, our third proposed method 'R-DPDPM3' attains a very similar precision to DeepDPM [2].

Figure 1 shows time performance of our work with respect to image resolution compared to state-of-the-art approaches. We extended our evaluation results to KITTI vehicle benchmark as shown in Table 2. Our third framework 'R-DPDPM3' has comparable accuracy to DeepDPM [2] (while running 10 times faster) and outperforms original DPM besides its other acclereated versions like 'FastDPM [3]'. We experience a loss of 2% in mAP using the evaluation settings (Easy, Moderate, Hard).

[3] Our FastDPM [3] implementation on CPU.

[4] Faster R-CNN code: https://github.com/shaoqingren/faster_rcnn.

[5] R-CNN code: https://github.com/rbgirshick/rcnn

Table 2. Average Precision (AP in %) and time (in seconds) on KITTI vehicle dataset with three test settings.

	Easy	Moderate	Hard	Time(s)
HOG-DPMv5 [1]	77.35	60.84	47.52	28.01
FFT-DPM[2]	76.42	58.48	45.82	3.52
FastDPM [3][3]	77.19	60.69	47.18	0.71
DPM1 [4] (Ours)	75.90	59.14	45.63	0.039
R-DPM2 [5] (Ours)	75.72	59.13	45.53	0.036
DeepDPM [2][1]	86.79	70.27	56.91	1.57
R-DPDPM (ours)	85.64	69.44	55.23	0.14

5. CONCLUSION

In this paper, we discussed problem domain of autonomous vehicles and our specific problem towards high performance object detection. we presented two frameworks based to achieve a real-time deformable part model detector without sacrificing its accuracy and another framework as a synthesis of convolutional network and structured learning to build a faster DeepPyramid DPM detector. This framework improves feasibility of DPM to driver assistance and self-driving technologies.

6. REFERENCES

[1] R. B. Girshick, P. F. Felzenszwalb, and D. McAllester, "Discriminatively trained deformable part models, release 5," `http://people.cs.uchicago.edu/~rbg/latent-release5/`, 2012.

[2] Ross Girshick, Forrest Iandola, Trevor Darrell, and Jitendra Malik, "Deformable part models are convolutional neural networks," in *CVPR*, 2015, pp. 437–446.

[3] Junjie Yan, Zhen Lei, Longyin Wen, and S.Z. Li, "The fastest deformable part model for object detection," in *CVPR*, 2014, pp. 2497–2504.

[4] A. Ali and M. A. Bayoumi, "Towards real-time dpm object detector for driver assistance," in *ICIP*, 2016, pp. 3842–3846.

[5] A. Ali, O. Olaleye, and M. Bayoumi, "Fast region-based dpm object detection for autonomous vehicle," in *MWSCAS*, 2016.