Wheelchair Detection and Pose Estimation in Monocular Images

Ingrid Navarro Anaya, *ITESM*, Kevin Brian Kwan Chong-Loo, *ITESM*, Dr. Luis E. Navarro-Serment, *CMU*

Introduction

- **Wheelchairs** are the most important auxiliary instrument for people with **mobility impairments**. Accurate **detection** and **tracking** of these devices could bring a number of improvements in **automated services**.
- Our goal is to present a **deep learning-based wheelchair detection and tracking system** along with an algorithm that can **estimate its pose**.

Challenges

- Extract features from **different types** of wheelchairs.
- Detection from **different viewing angles** and in **cluttered environments.**
- Avoid **pre-calibration.**

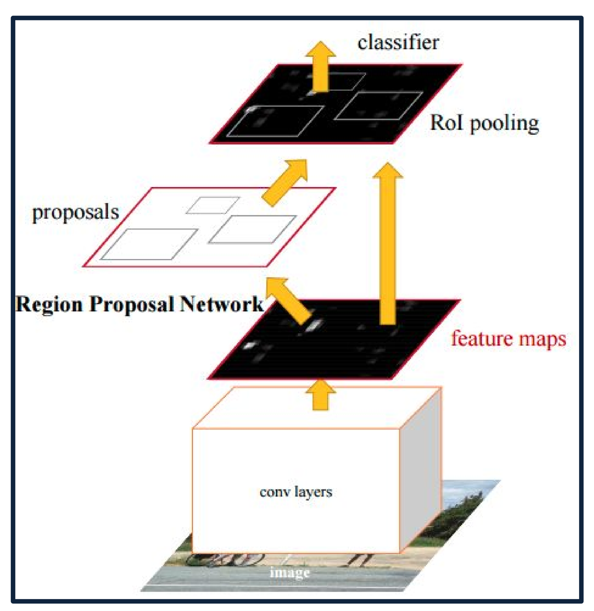

Fig. 1: Faster R-CNN [1] network.

Our Approach

- We used object recognition framework **Faster R-CNN** to perform **detection** and **tracking**.
- We evaluated the performance of the framework using the **approximate joint training** scheme.
- We used **ResNet50** as our baseline architecture.

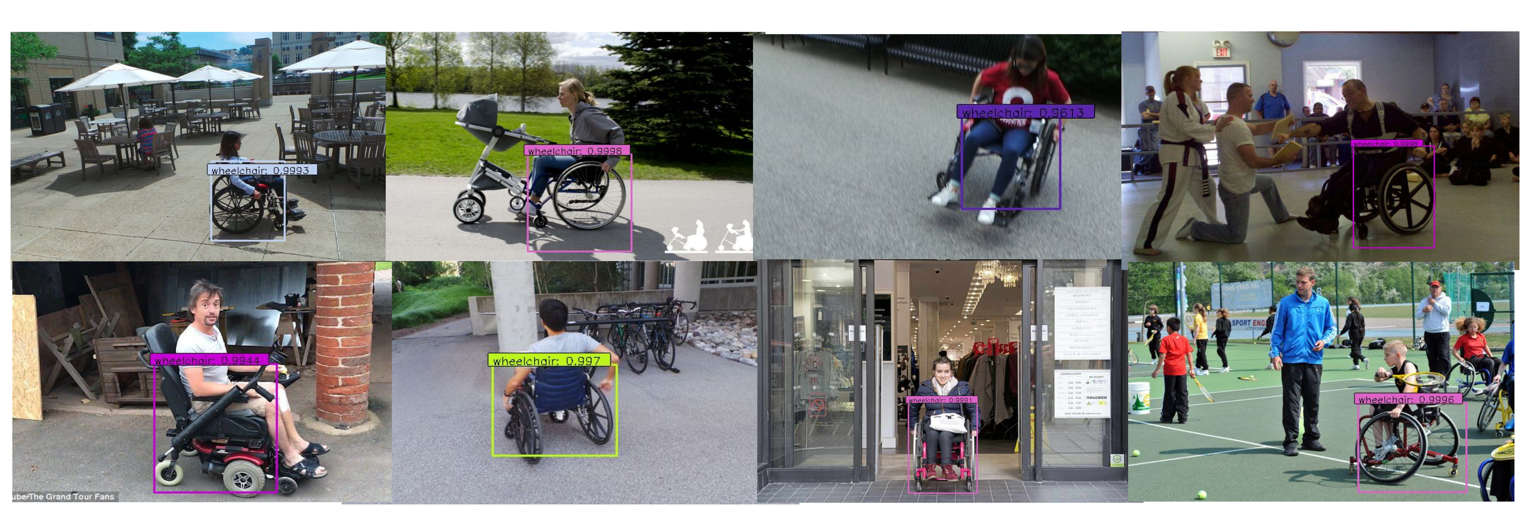

Fig. 2: Selected examples of wheelchair detection results from **Approximate joint training.**

- We used **data augmentation** and **bin-specific separation** techniques to increase robustness and accuracy.
- Generated **bounding boxes** are used for wheelchair **pose estimation.**
- **Wheelchair pose** is estimated using **contours** and wheel search using **least mean squares fitting ellipse algorithm**.

Fig. 3: (1) All Contours are obtained. (2) Fit ellipse algorithm applied to contours. (3) Pose estimated along with ground points.

Results

Wheelchair Dataset	No Data Augmentation			Data Augmentation		
	mAP@0.50 (%)	mAP@0.70 (%)	mAP@0.85 (%)	mAP@0.50 (%)	mAP@0.70 (%)	mAP@0.85 (%)
Non-Motorized wheelchair bin	92.71	**78.73**	**19.90**	**92.58**	75.23	14.09
Complete dataset	91.21	75.47	15.21	91.59	82.81	14.62

Table 1: mAP results of the **Approximate Joint Training.**

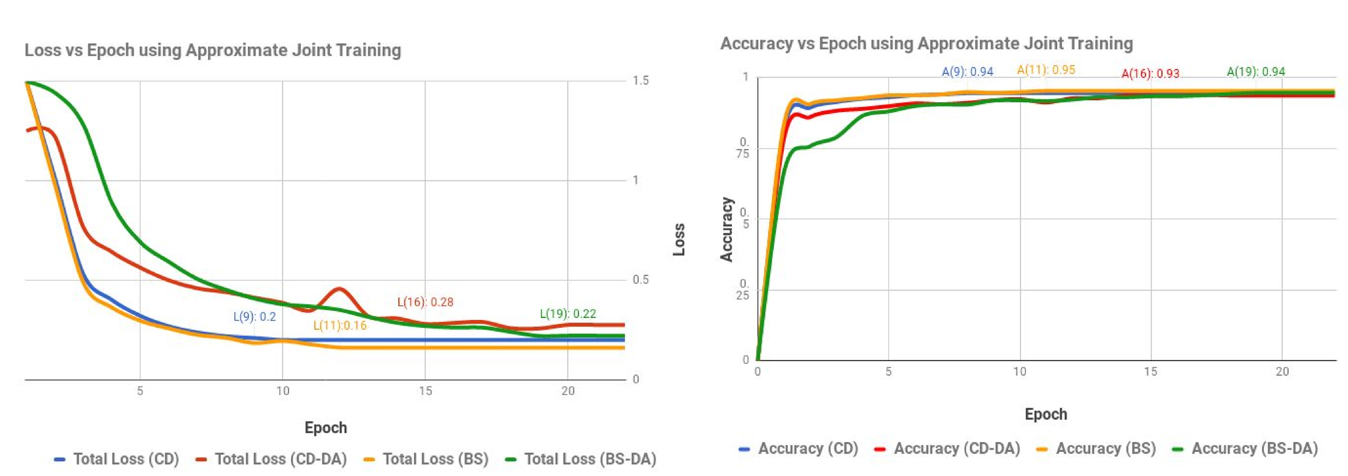

Fig. 4: Loss (left) and accuracy (right) of tests from **table 1.**

Case	Description	Success Rate	Big Wheel Error (Pixels)		Small Wheel Error (Pixels)		Angle Error (Degrees)	
			Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
1	2 Same Side Ground Points	**28.93 %**	12.01	8.21	11.03	9.88	4.17	3.95
2	2 Opposite Ground Points	**3.65%**	13.16	8.40	8.64	6.86	-	-
3	1 Big Wheel Ground Point	**39.04%**	15.32	11.19	-	-	-	-
4	1 Small Wheel Ground Point	**9.27%**	-	-	7.51	6.24	-	-
5	No Ground Points Detected	**19.10%**	-	-	-	-	-	-

Table 2: Wheelchair **Pose Estimation** results

Discussion

- Our implementation robustly and accurately addressed the challenges previously presented.

Fig. 5: feature map visualization

- The approach for the **wheelchair pose** estimation is **simple** to apply, and does not require previous **training data sets**.

Future Work

- Implement Mask R-CNN to increase the robustness of detection and to facilitate pose estimation by segmenting the regions of interest.

Acknowledgments

- We thank **Dr. Luis E. Navarro** for his support and mentoring in the development of the project as well as **The Robotics Institute Carnegie Mellon University.**
- We also thank **Professor Luis Ricardo Salgado** and **Tomasa Valdivieso** for their support and **Tecnológico de Monterrey** for sponsoring our research.

References

[1] S. Ren, K. He, R. Girshick. *Faster R-CNN: Towards real-time object detection with region proposal networks.* NIPS, 2015.

[2] A. Myles, N. Vitoria, M. Shah, Wheelchair Detection in a Calibrated Environment, 2002.