Demo Proposals

ICCV23 PARIS

CA-SpaceNet: Counterfactual Analysis for 6D Pose Estimation in Space

(Accepted as a regular paper by IROS 2022.)

Shunli Wang[1,2†], Shuaibing Wang[1,2†], Bo Jiao[1,2], Dingkang Yang[1,2], Liuzhen Su[1,2], Peng Zhai[1,2], Chixiao Chen[1,2*], Lihua Zhang[1,2,3,4*]

Academy for Engineering & Technology, Fudan University[1] Engineering Research Center of AI and Robotics, Ministry of Education, China[2]

Jilin Provincial Key Laboratory of Intelligence Science & Engineering, China[3]

Artifical Intelligence and Unmanned Systems Engineering Research Center of Jilin Province, China[4]

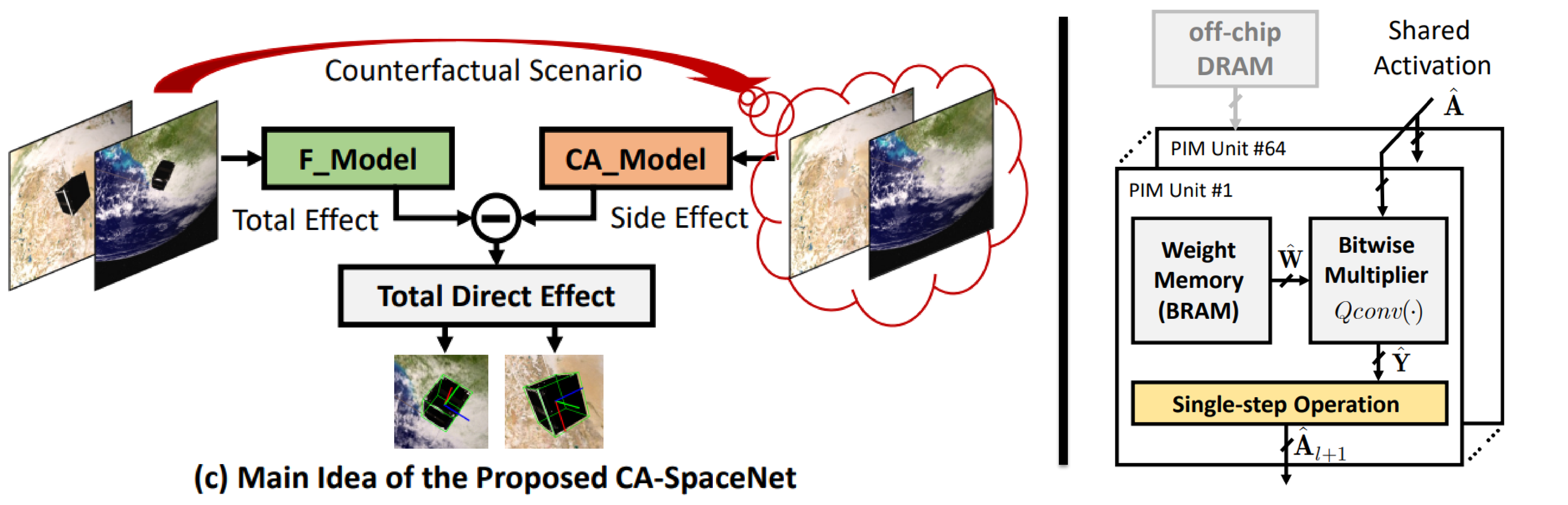
Code: https://shunli-wang.github.io/

Introduction

(a) SpaceX Crew-2 Docking Mission

(b) Preview of the ClearSpace-1

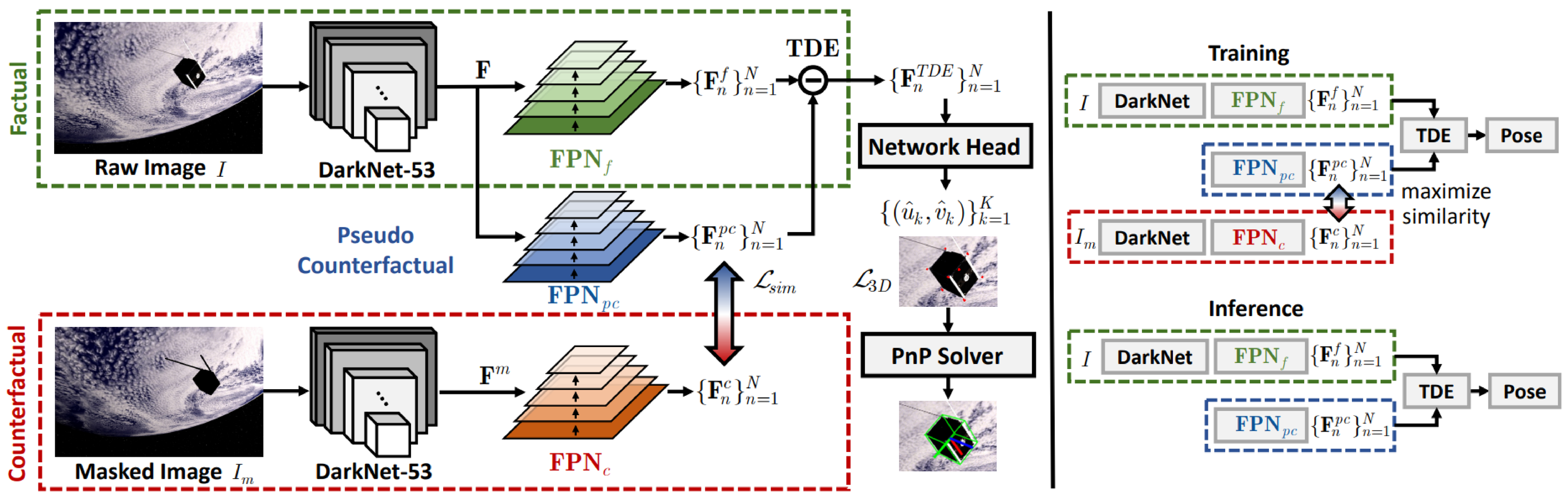
- Practical applications of the 6D pose estimation in many space missions. The complicated background of aerial images will interfere with the stability of the 6D pose estimator.



(c) Main Idea of the Proposed CA-SpaceNet

- This paper introduces counterfactual analysis to the 6D pose estimation task in space and proposes the CA-SpaceNet framework.
- We quantize the CA-SpaceNet into a low-bit-width model and deploy a part of the quantized network into a Processing-In-Memory chip on FPGA.

Method



(a) Overview of the proposed CA-SpaceNet

(b) Diagram of training and inference phases

Feature Extraction:

$$\begin{cases} \mathcal{F}^f = \mathbf{FPN}_f(\mathbf{F}) \\ \mathcal{F}^c = \mathbf{FPN}_c(\mathbf{F}^m) \\ \mathcal{F}^{pc} = \mathbf{FPN}_{pc}(\mathbf{F}) \end{cases} \mathcal{L}_{sim}$$

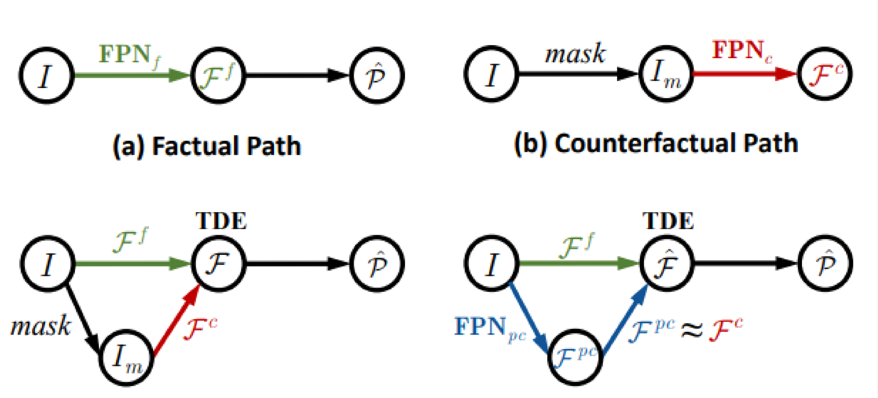
Total Direct Effect (TDE):

$$\begin{cases} \mathcal{F} = \mathcal{F}^f - \mathcal{F}^c \\ \hat{\mathcal{F}} = \mathcal{F}^f - \mathcal{F}^{pc} \end{cases}$$

Total Loss: $\mathcal{L} = \lambda_{3D}\mathcal{L}_{3D} + \lambda_c\mathcal{L}_{cls} + \lambda_s\mathcal{L}_{sim}$

- **Factual Path**: The factual path is designed to simulate the phenomenon of background interference.
- **Counterfactual Path**: The idea of counterfactual analysis is to imagine a non existent path, that is, to study the effect under the ***What If*** scenario.
- **Pseudo Counterfactual Path**: As its name implies, pseudo means that this path is a fake path, which aims to imitate the counterfactual path.

CA Module & Quantization



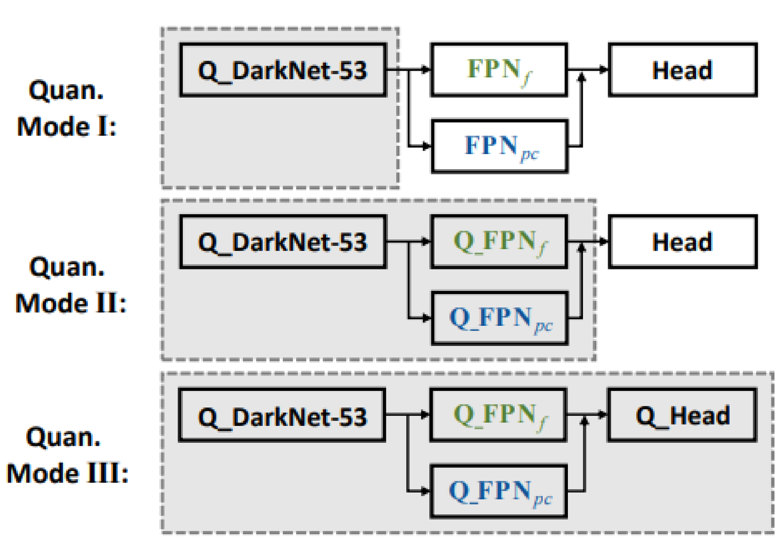
(a) Factual Path

(b) Counterfactual Path

(c) Ideal Counterfactual Analysis

(d) Real Counterfactual Analysis

- Simplified causal graphs of CA-SpaceNet in four situation.
- These causal graphs consist of four types of nodes: image node, feature node, TDE node, and pose results node.



- Three quantization modes are set up:
  - only quantizing the backbone,
  - quantizing the backbone and FPN,
  - quantizing all modules.

Experiments

Quantitative Analysis on the Swisscube and SPEED Datasets

Comparison with SOTAs on Swisscube.

Method	Near ↑	Medium ↑	Far ↑	All ↑
SegDriven [39]	41.1	22.9	7.1	21.8
SegDriven-Z [39]	52.6	45.4	29.4	43.2
DLR [5]	63.8	47.8	28.9	46.8
WDR [4]	65.2	48.7	31.9	47.9
WDR* [4]	**92.37**	84.16	61.27	78.78
CA-SpaceNet	91.01	**86.32**	**61.72**	**79.39**

Comparison with SOTAs on SPEED.

Method	$\mathbf{e}_q + \mathbf{e}_t$ ↓
SLAB Baseline [3]	0.0626
pedro-fairspace [42]	0.0571
WDR [4]	**0.0180**
WDR* [4]	0.0400
CA-SpaceNet	0.0385

Results on 3 different quantization modes of 8-bit and 3-bit CA-SpaceNet on SwissCube.

#Bits	Quan. Mode	ADI-0.1d ↑	OPs & FLOPs	Perc.(%)
8	I	76.21	36.91 GOPs + 33.79 GFLOPs	52.21
	II	75.04	44.51 GOPs + 26.19 GFLOPs	62.96
	III	74.65	70.47 GOPs + 0.23 GFLOPs	99.67
3	I	75.10	36.91 GOPs + 33.79 GFLOPs	52.21
	II	74.47	44.51 GOPs + 26.19 GFLOPs	62.96
	III	68.68	70.47 GOPs + 0.23 GFLOPs	99.67

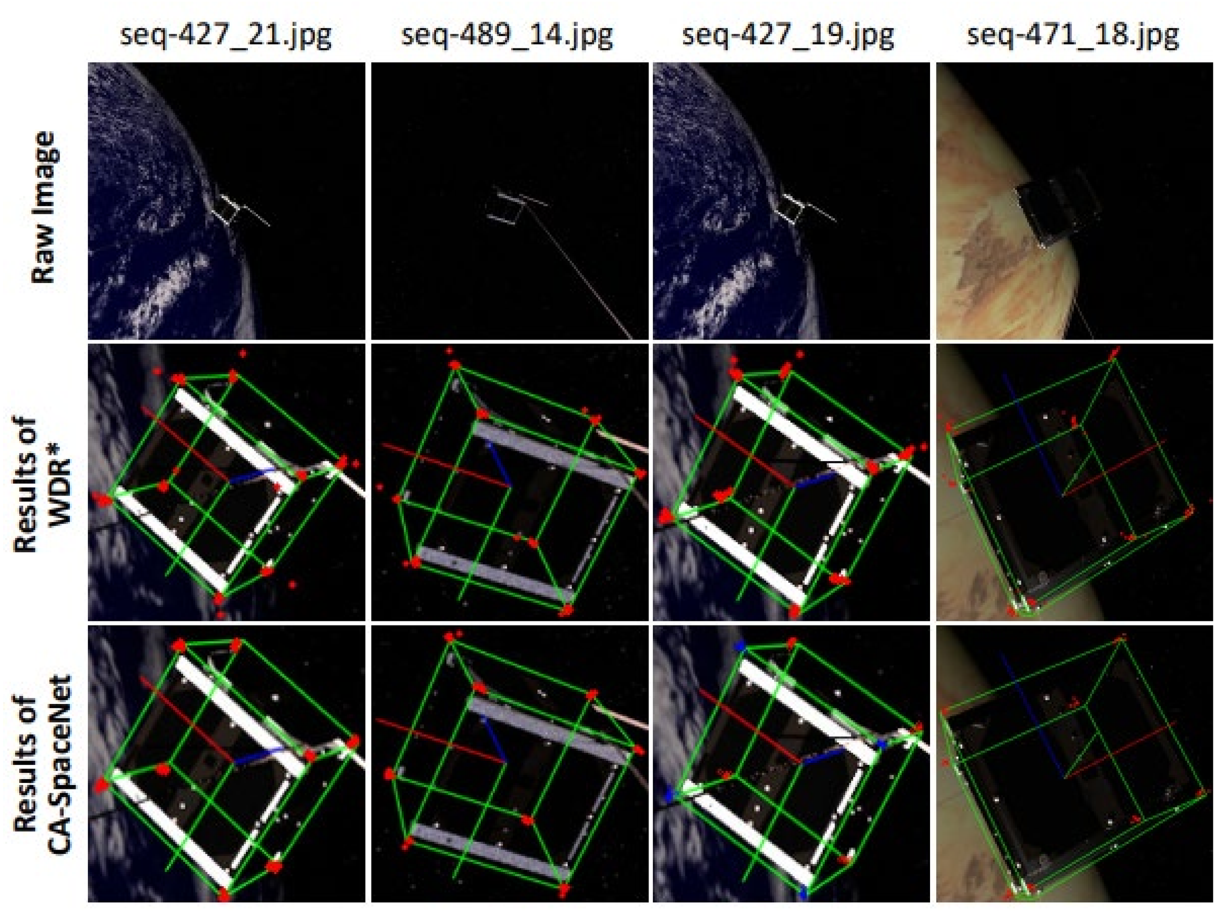
Summary of #parameters and storage size.

Format	#Para.	Model Size	Stor. Saving (%) ↑
FP32	51.29 M	205.17 MB	0.00
8-bit	51.29 M	51.29 MB	75.00
3-bit	51.29 M	19.23 MB	90.63

Measured latency on different hardware

Device	Latency (ms) ↓
ARM v8.2 64-bit CPU (Nvidia Xavier)	26.16
Intel Core i7-8700K CPU	10.25
PIM Arch. on Ultra96v2 FPGA	**5.99**

Visualization on Swisscube

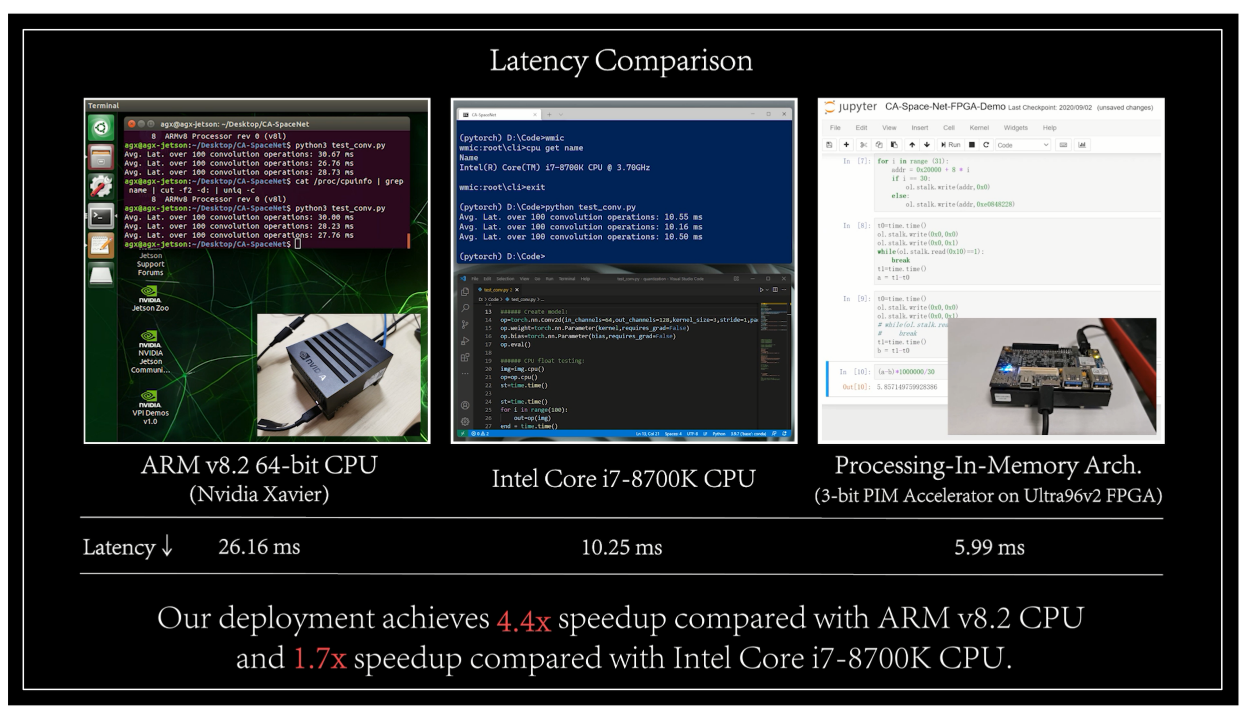


The CA-SpaceNet significantly reduces the background interference and generates robust pose estimation results.

Ground-truth boxes In Green.

Predicted corners In Red.

Measured Latency Comparison: PIM *v.s.* CPU



Our real deployment of the **PIM Architecture** on the Ultra96v2 FPGA achieves:

4.4x speedup than ARM v8.2 CPU,

&

1.7x speedup than Intel Core i7-8700K CPU.

Conclusion

- In this paper, We propose CA-SpaceNet based on counterfactual analysis to weaken the interference of background from the mixed features.
- Experimental results on SwissCube and SPEED datasets show that the proposed framework achieves robust performance.
- Further, we quantize the CA-SpaceNet into 3-bit and 8-bit and deploy part of the quantized network to a neural network accelerator on FPGA.

Acknowledgements

This work is supported by Shanghai Municipal Science and Technology Major Project (2021SHZDZX0103) and NSFC Grant (61974033).

Contact

Shunli Wang
slwang19@fudan.edu.cn
Lihua Zhang*
lihuazhang@fudan.edu.cn