ESA CleanSpace Industry Days (CSID) 2023 | Session: "In-Orbit Servicing & Debris Removal"

Deep Learning-based Spacecraft Pose Estimation

Opportunities and Challenges for Future In-orbit Servicing missions.

Dr. Arunkumar Rathinam

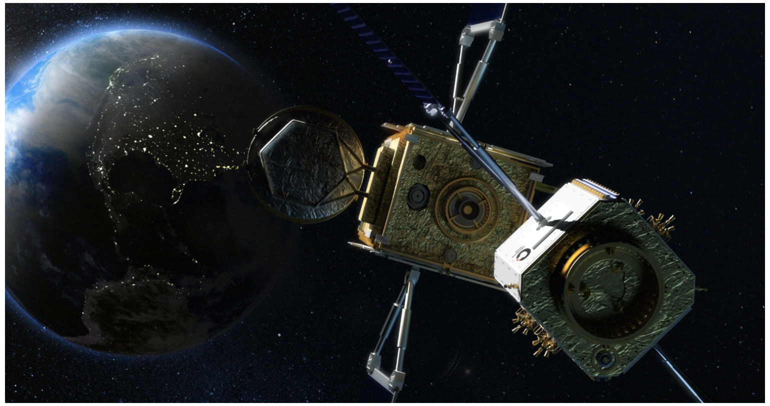
SnT, University of Luxembourg

Computer Vision, Imaging & Machine Intelligence Research Group

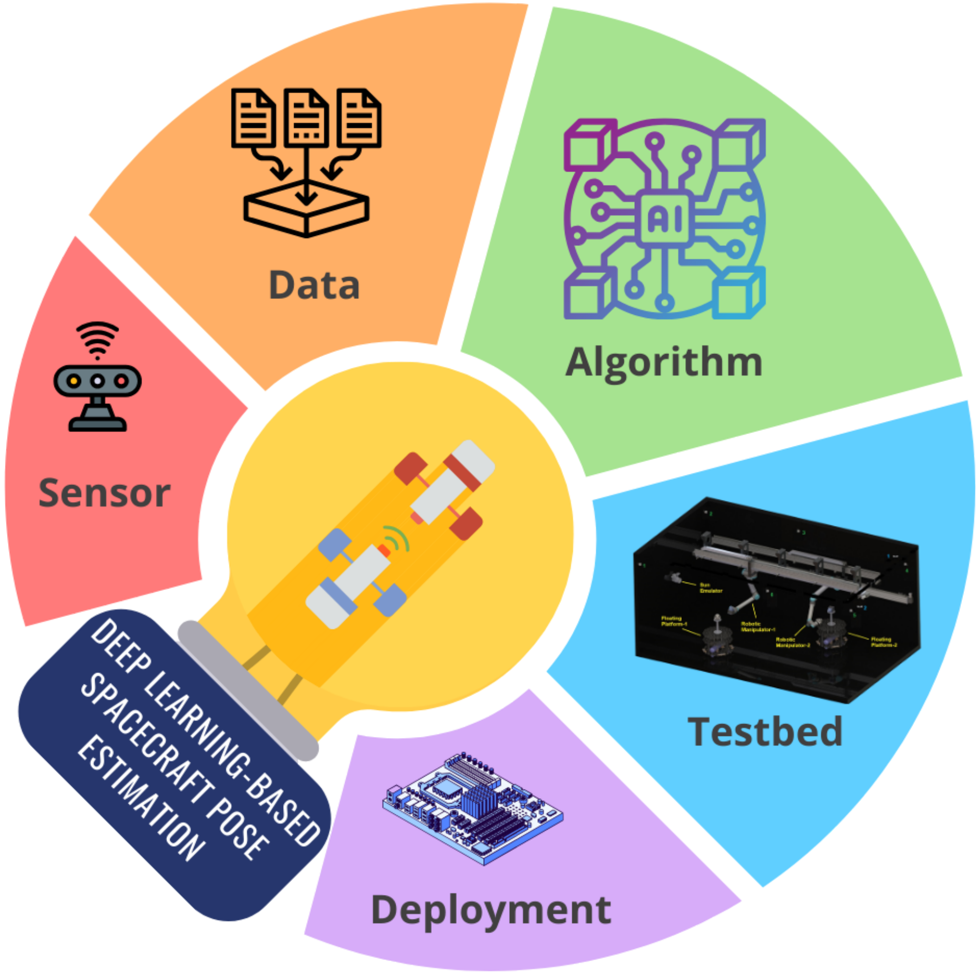
Luxembourg National Research Fund

Introduction

- **Orbital Robotic Missions**
  - significant level of autonomy is required for future orbital robotic missions
  - target orbits LEO, MEO, GEO
  - Mission Applications: **In-Orbit Servicing** and **Debris Removal**
- **Rendezvous and Proximity Operations**
  - Rendezvous: Far-range, Mid-range (100's m to 10 m), Close-range (< 10 m)
  - Proximity Ops require real-time estimation of a target position and orientation
  - Cameras are the preffered sensors for **Vision-based Navigation (VBN)**



Deep Learning-based Spacecraft Pose Estimation



Challenges:

- Sensors: RGB Camera limitations
- Data: Simulator dependent
- Algorithm:
  - Sim2Real or Domain Gap
- Testbed:
  - High-fidelity testbed
- Deployment:
  - large networks
  - network efficiency

DL Frameworks for Spacecraft Pose Estimation

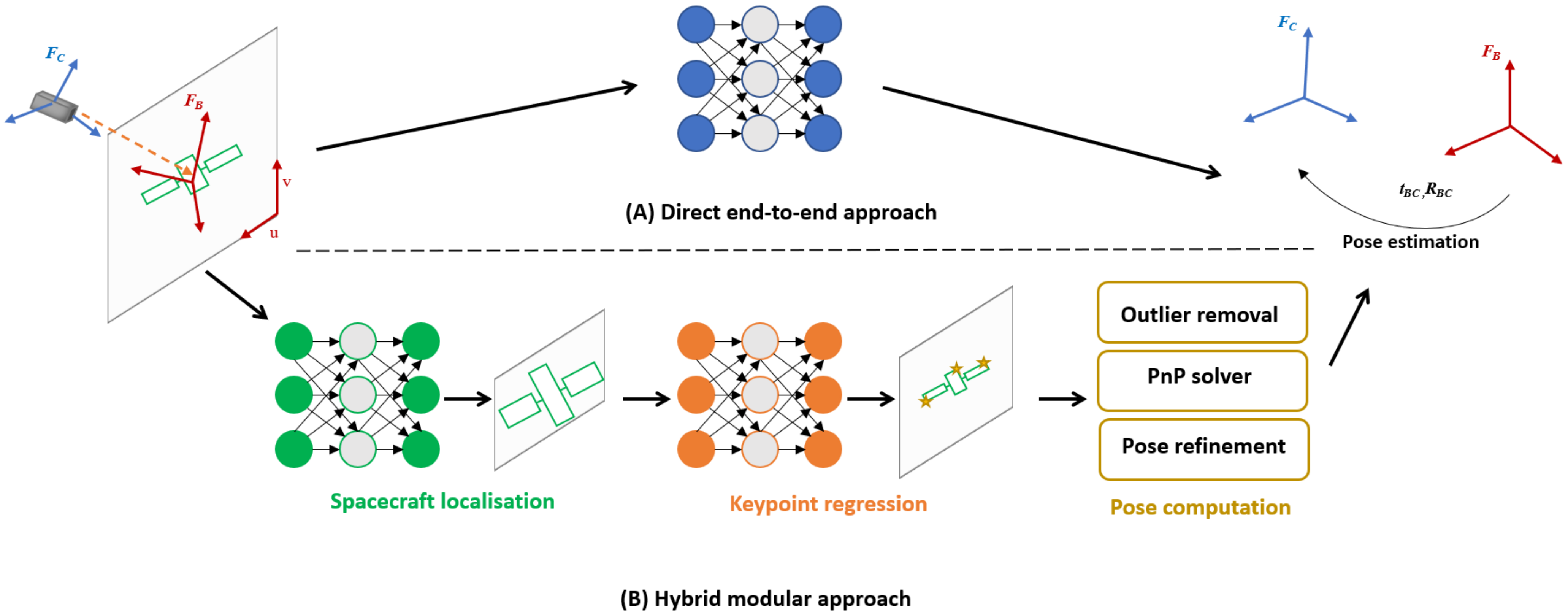
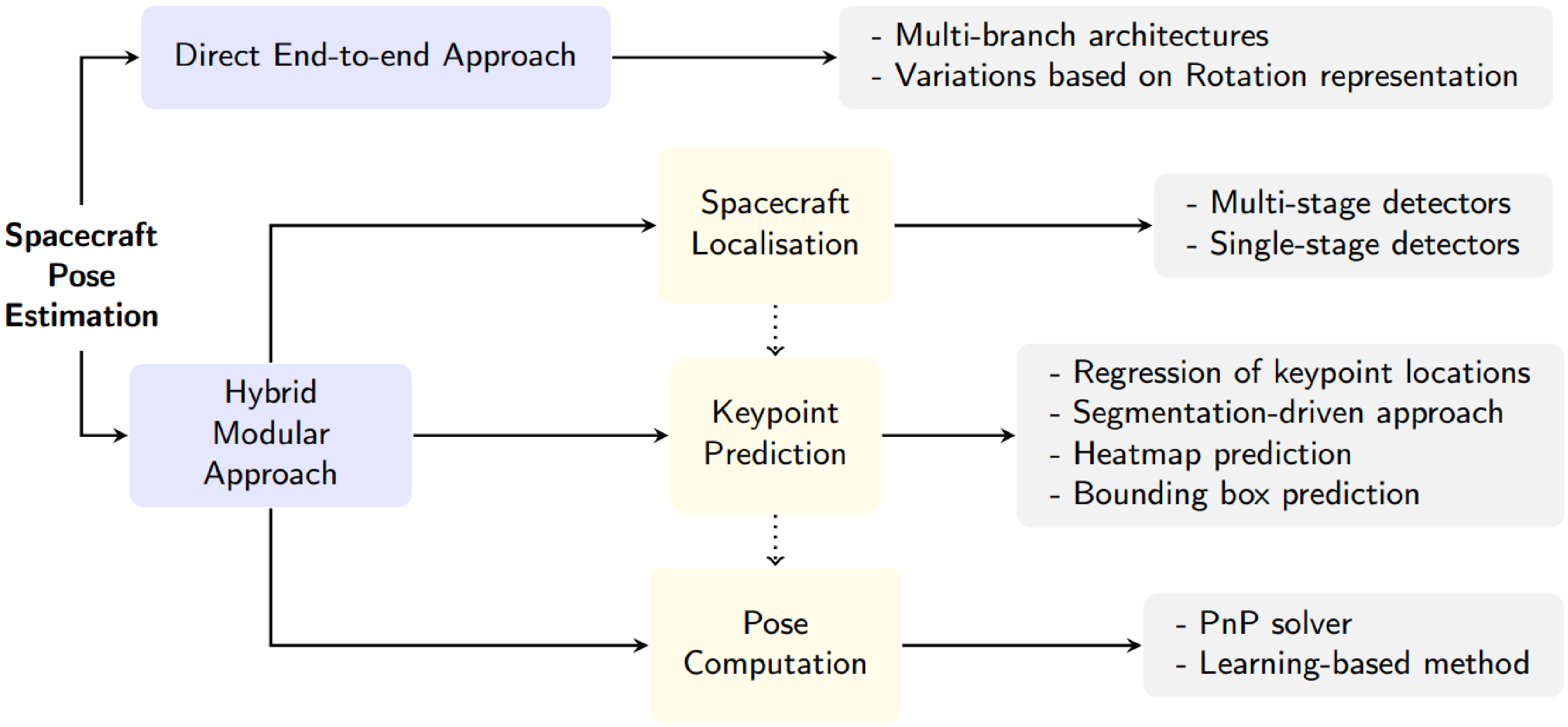


Illustration of different approaches for spacecraft pose estimation.

Pauly, Leo, et al. "A survey on deep learning-based monocular spacecraft pose estimation: Current state, limitations and prospects." Acta Astronautica (2023).

DL Frameworks for Spacecraft Pose Estimation



Overview of different approaches for spacecraft pose estimation.

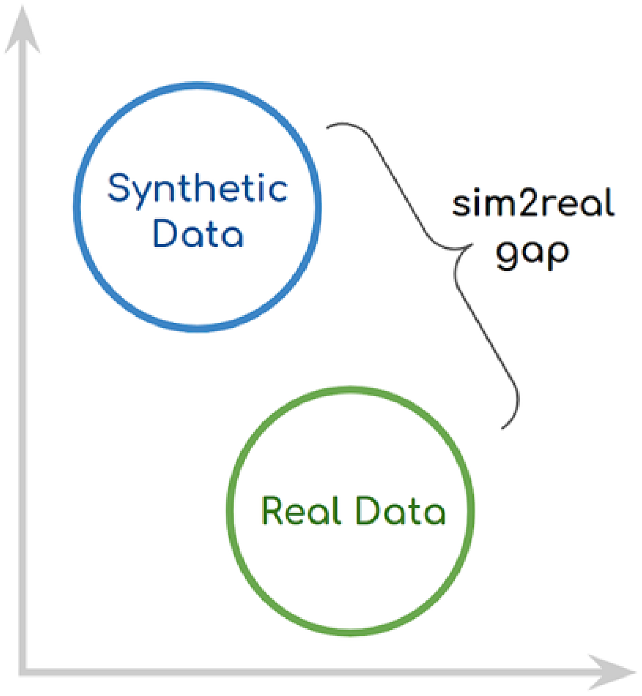
Spacecraft Pose Estimation Datasets

Dataset	Year	Syn	Lab	Space	Spacecraft	Resolution	I	Range (m)	Tools
SEENIC	2022	10k	5k	-	Hubble Telescope	640×480	E	[20, 45]	Blender
SHIRT	2022	5k	5k	-	Tango	1920×1200	G	≤ 8	OpenGL
SPARK2-Stream2	2022	30k	900	-	Proba-2	1440×1080	C	[1.5, 10]	Blender
COSMO	2022	15k	-	-	COSMO-SkyMed	1920×1200	C	[36, 70]	Blender
SwissCube	2021	50k	-	-	SwissCube	1024×1024	C	[0.1, 1]	Mitsuba 2
SPEED+	2021	60k	10k	-	Tango	1920×1200	G	≤ 10	OpenGL
Cygnus	2021	20k	-	540	Cygnus	1024×1024	C	[35, 75]	Blender
SPEED	2020	15k	305	-	Tango	1920×1200	G	[3, 40.5]	OpenGL
URSO	2019	15k	-	-	Dragon, Soyuz	1080×960	C	[10, 40]	UE 4
PRISMA12K	2019	12k	-	-	Tango	752×580	G	-	OpenGL

Overview of available spacecraft pose estimation datasets. RGB (C), grayscale (G) or event stream (E).

Domain Gap

- In general, satellite pose estimation models are trained using synthetic images.
- Leads to **domain gap** or **Sim2Real** problem, i.e. models trained in one domain (synthetic) face a drop in performance when tested on another (real data)
- overfitting features specific to training domain
- **How to achieve domain invariance?**



Domain Gap Visualization.

Domain Gap

- In general, satellite pose estimation models are trained using synthetic images.
- Leads to **domain gap** or **Sim2Real** problem, i.e. models trained in one domain (synthetic) face a drop in performance when tested on another (real data) overfitting features specific to training domain
- **How to achieve domain invariance?**
  - Domain Randomization
  - Domain Adaptation

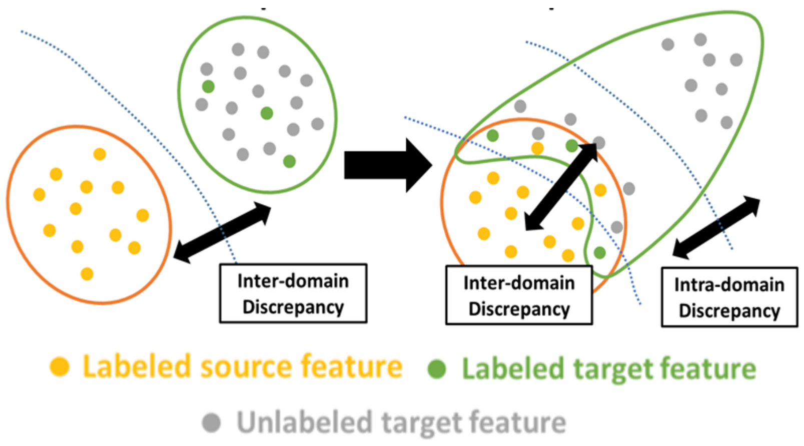
Domain Randomization.

Kim, T., & Kim, C. (2020). Attract, perturb, and explore: Learning a feature alignment network for semi-supervised domain adaptation. In Computer Vision–ECCV 2020: 16th European Conference, UK, August 23–28, 2020.

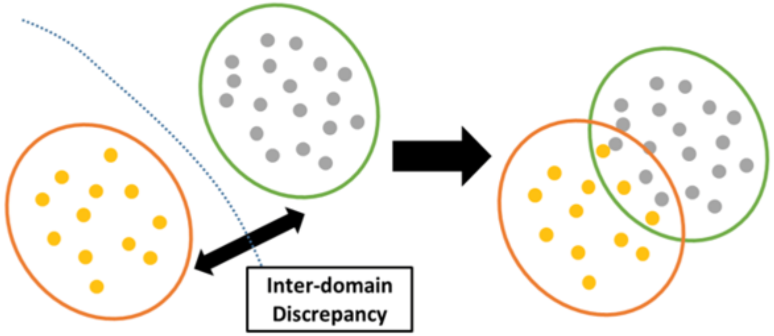
Domain Adaptation

- DA focus on mitigating the decrease in algorithms performance on the target domain for the same task
- **Supervised** and **Unsupervised** DA
- **Supervised:** access to both source and target (data + ground truth labels)
- **Unsupervised:** access to source (data + labels), but only to target (data) during training

Kim, T., & Kim, C. (2020). Attract, perturb, and explore: Learning a feature alignment network for semi-supervised domain adaptation. In Computer Vision–ECCV 2020: 16th European Conference, UK, 2020.

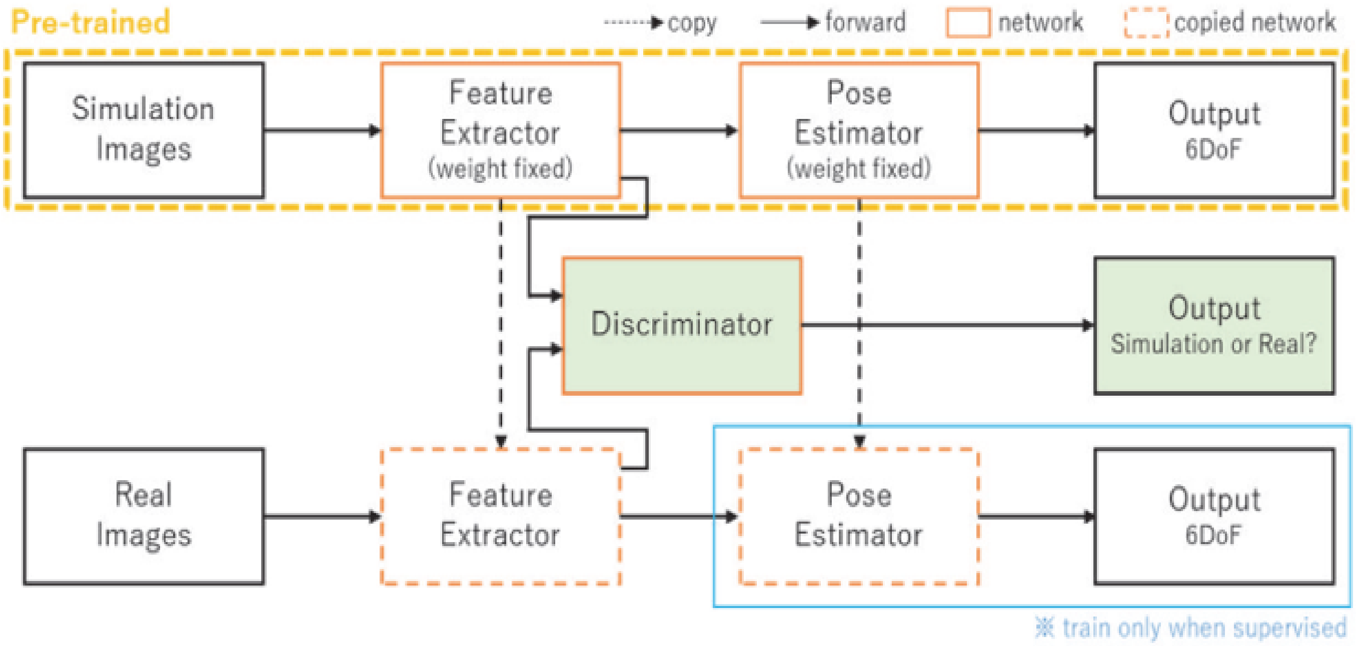


Supervised Domain Adaptation.

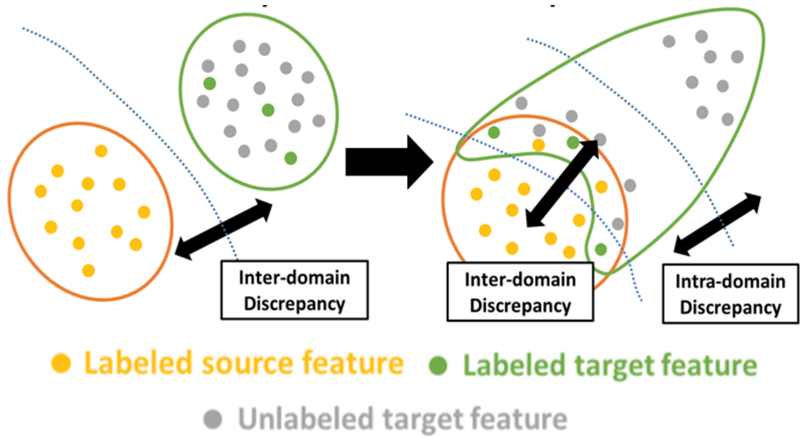


Unsupervised Domain Adaptation.

Supervised Domain Adaptation



CNN architecture for Supervised Domain Adaptation. [1]



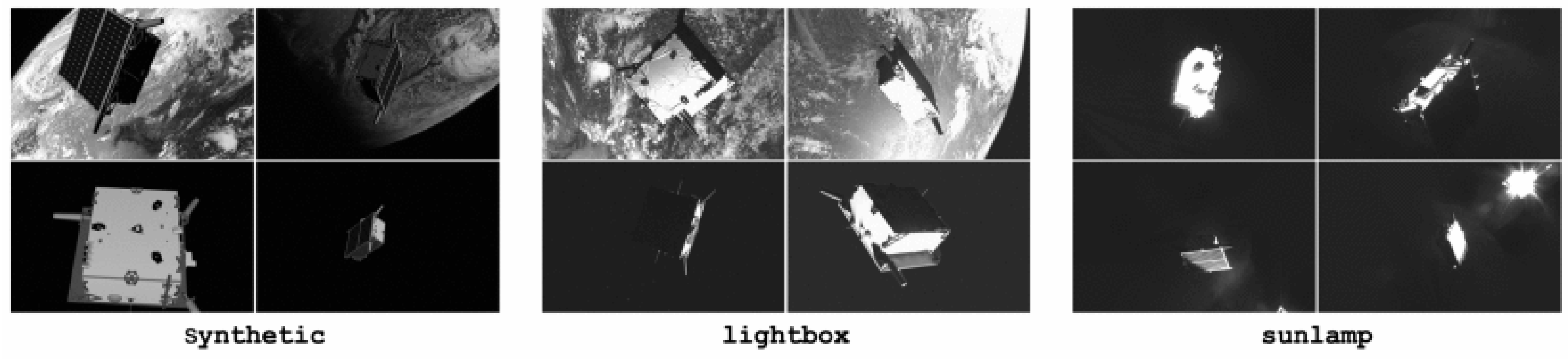
Supervised Domain Adaptation.

[1] S. Hashimoto et al., "Domain Adaptation for 6-DoF Pose Estimation: Deep Learning with Uncertainty," in 2022 IEEE Aerospace Conference (AERO), Big Sky, MT, USA: IEEE, Mar. 2022, pp. 1–12. doi: 10.1109/AERO53065.2022.9843646.

Unsupervised Domain Adaptation

- **Domain Invariant feature learning**
  - aims to align source and target at feature level, assuming that source and target come from same distribution
  - tries to aligns both domains with a discriminator
- **Input Alignment**
  - aims to align the source to the target domain at input level via generative adversarial networks
- **Self-training / Pseudo-labelling**
  - utilize target samples to train the model by generating pseudo-labels
  - possibility of noisy pseudo-labels

Datasets for Domain Gap: SPEED+



- `sunlamp` images have extreme surface glows and high contrast shadows due to direct light
- `lightbox` images are captured under extremely dim illumination conditions

M. Kisantal, S. Sharma, T. H. Park, D. Izzo, M. Märtens and S. D'Amico, "Satellite Pose Estimation Challenge: Dataset, Competition Design, and Results," in IEEE Transactions on Aerospace and Electronic Systems, vol. 56, no. 5, pp. 4083-4098, Oct. 2020, doi: 10.1109/TAES.2020.2989063.

Datasets for Domain Gap: SPEED+

Mean pose errors of the top-10 entries evaluated on the `lightbox` and `sunlamp` images.

Team	`lightbox` E_t [m]	E_q [°]	E^*_{pose} [–]	Team	`sunlamp` E_t [m]	E_q [°]	E^*_{pose} [–]
1. TangoUnchained	0.105	3.187	0.073	1. lava1302	0.065	2.728	0.059
2. VPU	0.131	4.577	0.101	2. VPU	0.076	2.828	0.061
SPNv2 [26]	0.150	5.577	0.122	3. TangoUnchained	0.086	4.299	0.090
3. lava1302	0.302	6.665	0.165	4. u3s_lab	0.181	6.241	0.141
4. haoranhuang_njust	0.180	8.131	0.173	5. haoranhuang_njust	0.163	8.406	0.175
5. u3s_lab	0.333	9.694	0.224	SPNv2 [26]	0.161	9.788	0.197
6. chusunhao	0.205	16.378	0.319	6. bbnc	0.542	21.955	0.465
7. for graduate	0.402	23.666	0.488	7. for graduate	0.455	22.970	0.487
8. Pivot SDA AI & Autonomy Sandbox	0.409	23.918	0.490	8. Pivot SDA AI & Autonomy Sandbox	0.854	36.445	0.766
9. bbnc	0.687	24.889	0.531	9. ItTakesTwoToTango	0.473	39.660	0.772
10. ItTakesTwoToTango	0.485	31.094	0.625	10. chusunhao	0.338	43.356	0.815

M. Kisantal, S. Sharma, T. H. Park, D. Izzo, M. Märtens and S. D'Amico, "Satellite Pose Estimation Challenge: Dataset, Competition Design, and Results," in IEEE Transactions on Aerospace and Electronic Systems, vol. 56, no. 5, pp. 4083-4098, Oct. 2020, doi: 10.1109/TAES.2020.2989063.

Domain Gap: Solutions

- Adversarial training as a crucial component to bridging the domain gap
- Object detection prior to landmark regression i.e. cropping the predicted 2D bounding box
  - help retain important visual features at far ranges
  - otherwise will be lost after image downsizing
- Generating accurate pseudo-labels during self-training (UDA) is critical for efficient learning
- multitask learning enables robust learning across domain gaps

Domain Gap: Alternative Prospects

- Limitations of RGB Sensors

Latency

Motion blur

Dynamic Range

Gallego, Guillermo, et al. "Event-based vision: A survey." IEEE transactions on pattern analysis and machine intelligence 44.1 (2020): 154-180.

Bridging Domain Gap: Sensor perspective

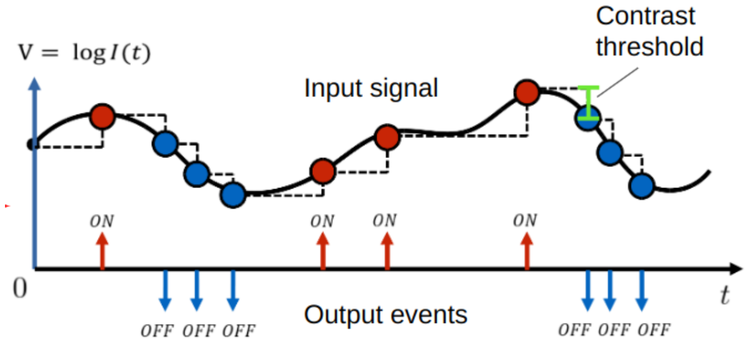
- Alternate vision sensors and fusion techniques are investigated.

Parameters	Visible (VIS)	Thermal Infrared (TIR)	Event
Illumination	Highly sensitive	Independent of illumination Rely on emitted radiance	Less Sensitive High Dynamic Range
Resolution & cost	High + Low	Low + High	Medium (720p) + High
Power	Low	Low + passive cooling	Lower than VIS
Simulation & Lab	Easy + Easy	Hard + Hard	Medium + Easy

[1] Jawaid, Mohsi, et al. "Towards bridging the space domain gap for satellite pose estimation using event sensing." 2023 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2023.

Event Camera : Introduction

- An Event Camera outputs asynchronous events.
- An event is generated each time a single pixel detects an intensity changes value.



Timestamp	x	y	polarity
160979782	293	100	1
160979789	483	200	0
....	...	...	...

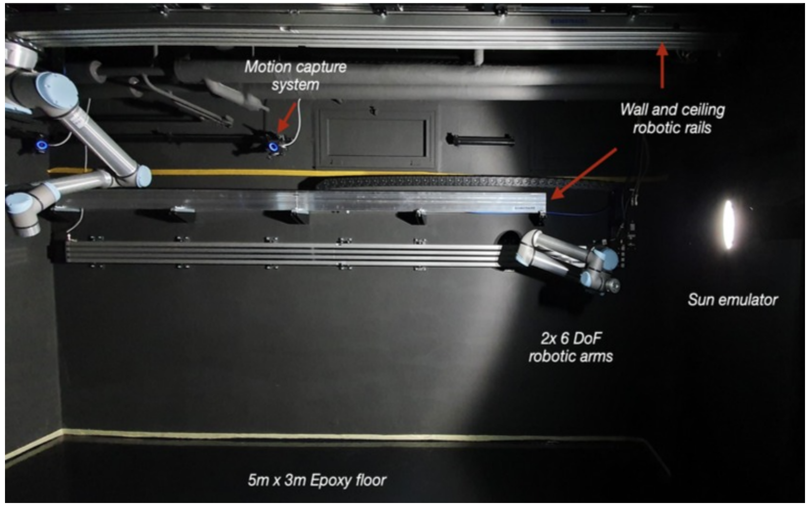
- Event Camera advantages
  - High Dynamic Range (140 dB compared to 60 dB for traditional cameras)
  - Low Latency (~ 1µs)
  - No motion blur
  - Low power consumption (1 mW compared to 1 W for traditional cameras)

Event Camera Dataset

- To investigate Event sensing for Spacecraft Pose Estimation, a new dataset was created.

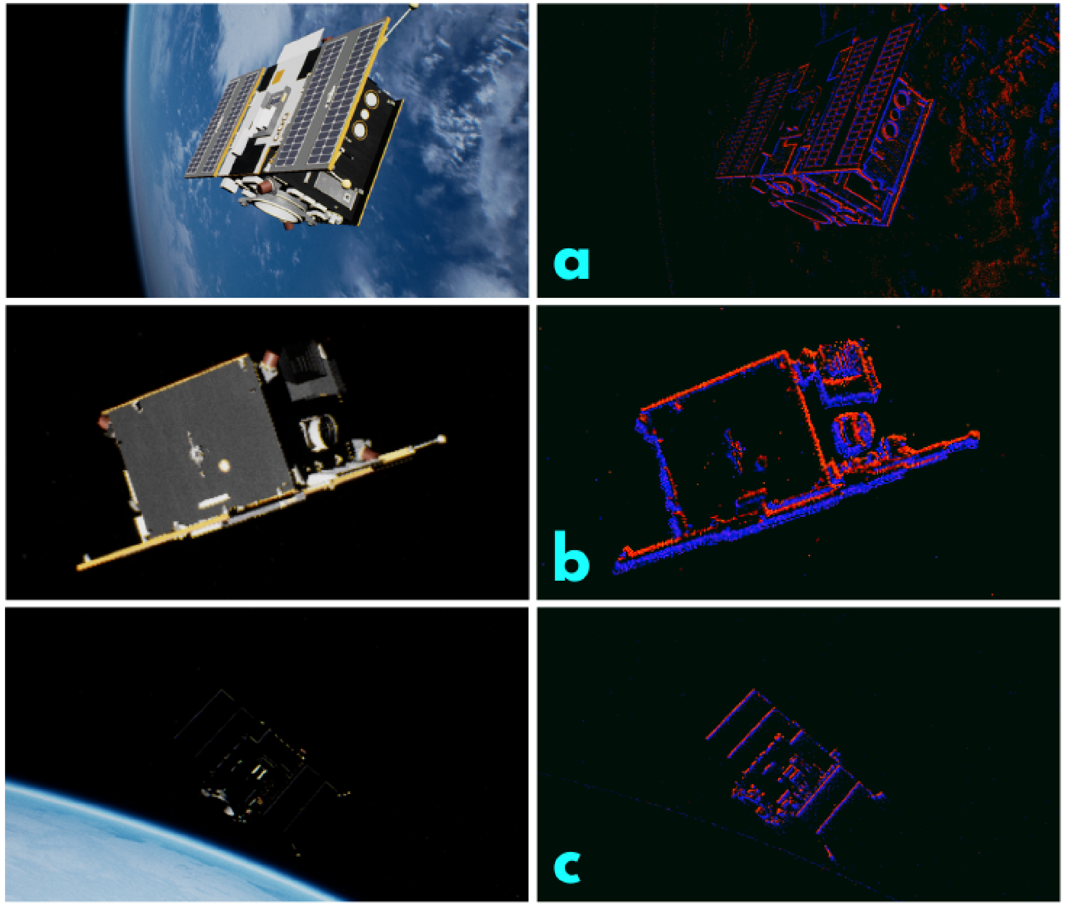
SPADES Dataset properties

	Synthetic	Real
Sensor resolution	1280x720	1280x720
Dataset size	179,400 (no. of poses)	15,500
No. Trajectories	300	31
No. poses/traj	598	500
Interpolation	80% spline & 20% Helix	-
Range	3.5 - 12 m	3.5 - 9 m
Range dist.	Close, Mid, Far, Limit	Close, Mid, Far
Lighting	Easy, Hard	L1, L2, L3, L4
Rendering	Unreal Engine (RGB)	-
Event Camera	ICNS Emulator	Prop. EVK4HD
Background	Earth	-
Filtering	Bbox/Mask	Min. event count

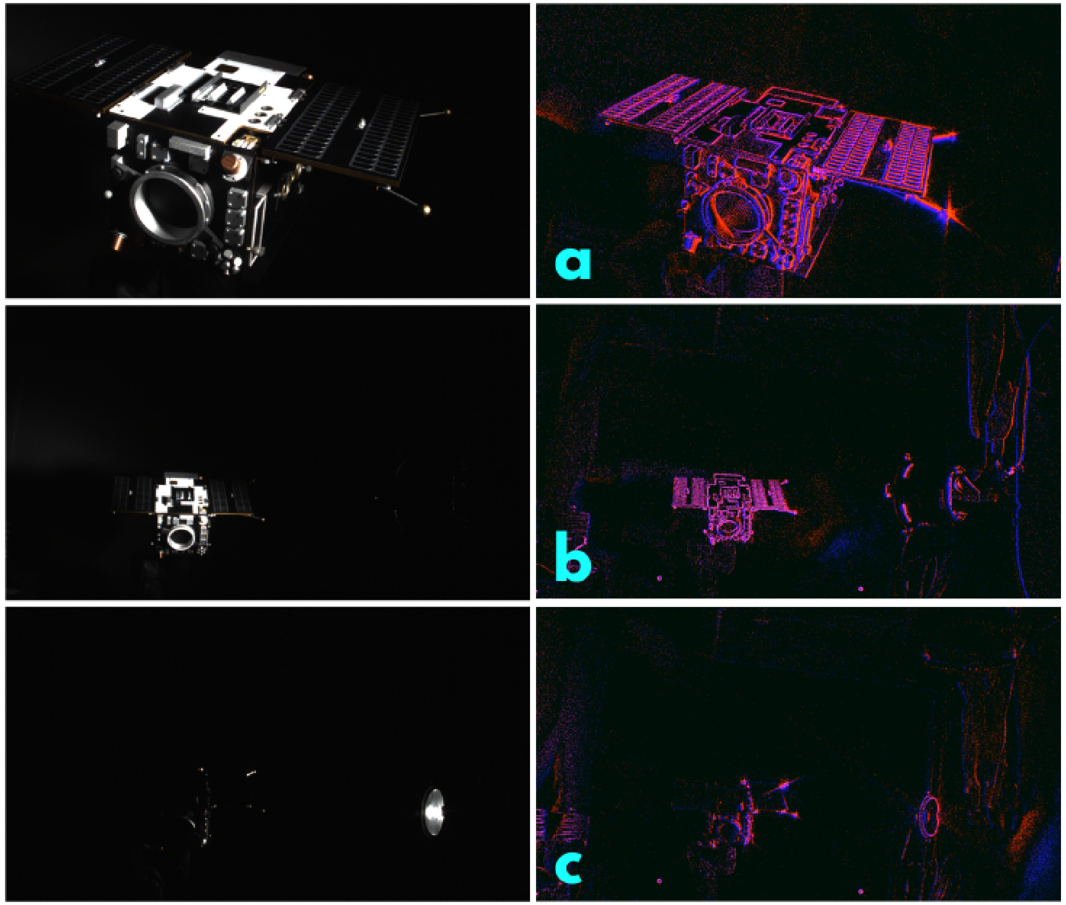


Zero-G Lab @ SnT

SPADES Dataset



Synthetic Samples



Real Samples

Event Pose Estimation Analysis

- Metrics

$$E_{\mathrm{T}} = \frac{\|\tilde{\mathbf{t}} - \mathbf{t}\|_2}{\|\mathbf{t}\|_2}; \quad E_{\mathrm{R}} = 2\, acos|\langle \tilde{q}, q \rangle|; \quad E_{\mathrm{P}} = E_{\mathrm{R}} + E_{\mathrm{T}}.$$

- Baseline Results on Simple Hybrid DL Framework.

Model	*Data* [%]	E_{T} [%]	E_{R} [°]	E_{P} [−]	*Data* [%]	E_{T} [%]	E_{R} [°]	E_{P} [−]
	Syn.				*Real*			
Direct	**97.32**	4.29	30.43	0.57	**73.32**	5.13	81.13	1.47
Hybrid	23.98	**3.23**	**6.69**	**0.15**	17.27	**3.34**	**78.98**	**1.41**

DL frameworks : Summary

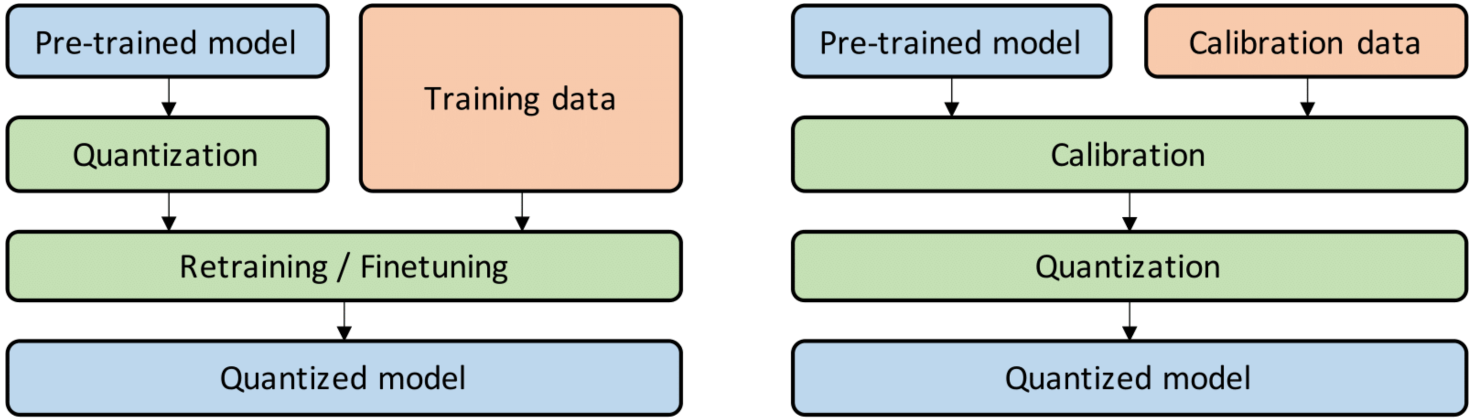
- Irrespective of data type or sensor, **domain gap** is imminent.
- The significance of domain gap and algorithm selection can make high impact in performance.
- **Sensor Fusion** will help balance the performance

Deployment in Edge Devices

- Real-time inference is key to orbital autonomy
- Edge devices have limited memory, computing resources, and power.
- Achieving efficient, real-time NNs with optimal accuracy requires rethinking the design, training, and deployment of NN models
- Methods to improve NNs efficiency
  - Network Pruning
  - Knowledge Distillation
  - Quantization
  - Neural Architecture Search

Quantization

- **Quantization** is the process of converting continuous values to discrete set of values using linear/non-linear scaling techniques. FP32/FP16 → FP8 / INT8
  - **Quantization-aware training** (QAT) is a fine-tuning process, where the model is further trained with quantization in mind.
  - **Post-training quantization** (PTQ) is a quantization technique where the model is quantized after it has been trained.



Gholami, Amir, et al. "A survey of quantization methods for efficient neural network inference." arXiv preprint arXiv:2103.13630 (2021).

Quantization

- Quantization-aware training the model is further trained with quantization in mind.

  - QAT aims at computing scale factors during training
  - Fine-Tuning: Once the network is fully trained, Quantize (Q) and Dequantize (DQ) nodes are inserted and further trained for few epochs
  - Q/DQ nodes simulate quantization loss

- Keeping performance before and after quantization is critical

Improving Quantization

- With current tools, it is straightforward to quantize and deploy different NN models to INT8, without losing accuracy.
  - software packages (e.g., Nvidia's TensorRT, TVM, etc.)
- the appropriate quantization method, identifying sensitive layers, and fine-tuning your models using QAT
- **Hardware and NN Architecture Co-Design:** useful for FPGA deployment, as one can explore many different possible hardware configurations
- changing the width of the NN architecture could reduce/remove generalization gap after quantization.
- **Quantized traning** to accelerate NN training with half-precision
  - with INT8 precision, the training can become unstable and diverge

Summary

Challenges	Opportunities
Algorithm: Sim2Real or Domain Gap	- Domain Adaptation Methods - Multi-task learning
Sensor limitations	- Potential for Neuromorphic Vision - Sensor Fusion with RGB cameras
Deployment in Edge Devices	- Quantization aware-training - Network Pruning - Neural Architecture Search
Data	- High-Fidelity Simulation - Closing Domain Gap via simulation tools

Luxembourg National Research Fund

LUXEMBOURG SPACE AGENCY

MEET-A

Multi-modal Fusion of Electro-optical Sensors for Spacecraft Pose Estimation Towards Autonomous in- Orbit Operations

AUREA

Space-based Space Domain Awareness (SDA) using small satellite platforms

AI4CC

Earth Observation

DIOSSA – ESA / LuxImpulse

Development of an In-Orbit SSA Payload for In-Orbit Servicing

ELITE

Enabling Learning and Inferring compact deep neural network Topologies on Edge devices

PARTNERS

LMO

Hydrosat

INFINITE ORBITS

SPARK CHALLENGE

2021 & 22

Computer Vision, Imaging & Machine Intelligence Research Group

CVI²

Luxembourg National Research Fund

uni.lu | SnT

Thank you for your attention!

Contact: arunkumar.rathinam@uni.lu

Computer Vision, Imaging
& Machine Intelligence
Research Group

Luxembourg
National
Research Fund