Analyzing Deep-Learning Methods for Power Line Component Detection in Unmanned Aircraft System Imagery with Few Data

International Conference on Unmanned Aircraft Systems 2025

Guillaume Fourret[1,2], Marc Chaumont[1,3], Christophe Fiorio[2], Gérard Subsol[1,2], Samuel Brau[2]

[1]ICAR, LIRMM, University of Montpellier, CNRS, France
[2]Drone Geofencing, Nîmes, France
[3]University of Nîmes, France

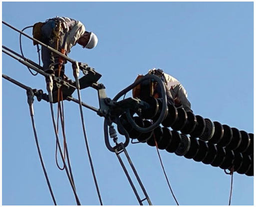
Contact : `guillaume.fourret@lirmm.fr`

1. UAS monitoring

Objective: Prevent power line shutdowns by analyzing their components.

Traditional inspection methods include:

Manual inspection



Robotic systems

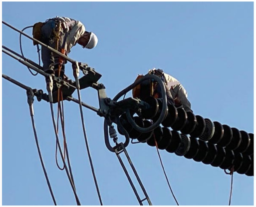
Helicopter-assisted inspection

1. UAS monitoring

Objective: Prevent power line shutdowns by analyzing their components.

Traditional inspection methods include:

Manual inspection



Robotic systems

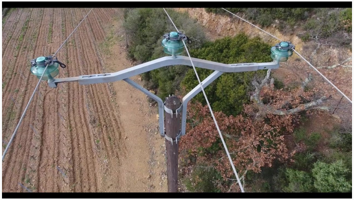
Helicopter-assisted inspection

Unmanned **A**ircraft **S**ystems (**UAS**) have revolutionized the field:

- Minimize risk to technicians
- Significantly lower operational costs
- Efficiently cover large and hard-to-reach areas
- Capture high-quality views of components from multiple angles

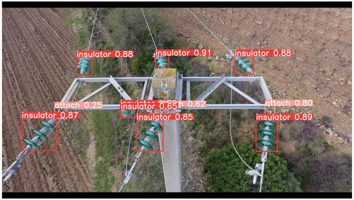
1. Monitoring automation

Technicians using UAS provide us with numerous RGB inspection videos:



drone_inspection.mp4

To automate defect monitoring on components, the first step is to apply **object detection** using deep learning:



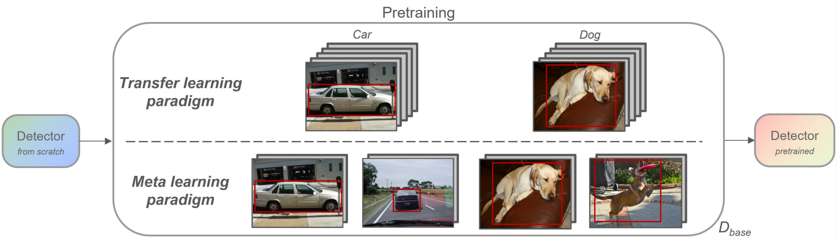
drone_detection.mp4

1. Interest of FSOD

In real-world scenarios: a wide variety of components must be detected!

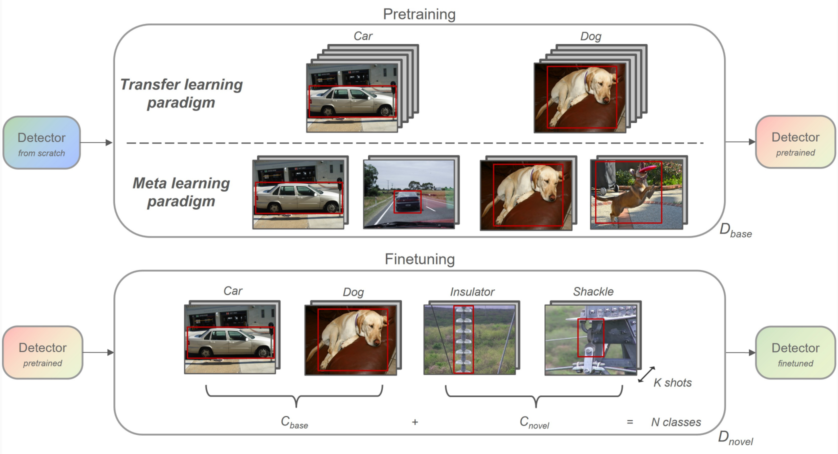
Few-Shot Object Detection (FSOD): Training a detector to recognize new classes using as few as 1 to 30 annotated examples.

2. General FSOD setting



- **Pretraining**: Use a large dataset with many classes and examples

2. General FSOD setting



- **Pretraining**: Use a large dataset with many classes and examples
- **Finetuning**: Add new classes using very few examples (shots)

2. Evaluated FSOD methods

Methods selected for our study:

- DeFRCN[1]: A modified Faster R-CNN based method for FSOD

[1] Qiao, Limeng et al. 2021. "Defrcn: Decoupled faster r-cnn for few-shot object detection" *Proceedings of the IEEE/CVF International Conference on Computer Vision*.

[2] Fu, Yuqian et al. 2024. "Cross-Domain Few-Shot Object Detection via Enhanced Open-Set Object Detector" *arXiv preprint arXiv:2402.03094*.

[3] Jocher, Glenn, Ayush Chaurasia, and Jing Qiu 2023. *YOLO by Ultralytics*.

2. Evaluated FSOD methods

Methods selected for our study:

- DeFRCN[1]: A modified Faster R-CNN based method for FSOD
- CD-ViTO[2]: Most performant method, using **S**elf-**S**upervised **L**earning (**SSL**) for high quality feature extraction for FSOD

[1] Qiao, Limeng et al. 2021. "Defrcn: Decoupled faster r-cnn for few-shot object detection" *Proceedings of the IEEE/CVF International Conference on Computer Vision*.

[2] Fu, Yuqian et al. 2024. "Cross-Domain Few-Shot Object Detection via Enhanced Open-Set Object Detector" *arXiv preprint arXiv:2402.03094*.

[3] Jocher, Glenn, Ayush Chaurasia, and Jing Qiu 2023. *YOLO by Ultralytics*.

2. Evaluated FSOD methods

Methods selected for our study:

- DeFRCN[1]: A modified Faster R-CNN based method for FSOD
- CD-ViTO[2]: Most performant method, using **S**elf-**S**upervised **L**earning (**SSL**) for high quality feature extraction for FSOD
- Yolov8[3] (baseline): Widely adopted one-stage detector
    - Lightweight models suited for embedded devices
    - High inference speed
    - Strong performance on MSCOCO
    - However, **NOT** designed for FSOD

[1] Qiao, Limeng et al. 2021. "Defrcn: Decoupled faster r-cnn for few-shot object detection" *Proceedings of the IEEE/CVF International Conference on Computer Vision*.

[2] Fu, Yuqian et al. 2024. "Cross-Domain Few-Shot Object Detection via Enhanced Open-Set Object Detector" *arXiv preprint arXiv:2402.03094*.

[3] Jocher, Glenn, Ayush Chaurasia, and Jing Qiu 2023. *YOLO by Ultralytics*.

2. Evaluated FSOD methods

Methods selected for our study:

- DeFRCN[1]: A modified Faster R-CNN based method for FSOD
- CD-ViTO[2]: Most performant method, using **S**elf-**S**upervised **L**earning (**SSL**) for high quality feature extraction for FSOD
- Yolov8[3] (baseline): Widely adopted one-stage detector
  - Lightweight models suited for embedded devices
  - High inference speed
  - Strong performance on MSCOCO
  - However, **NOT** designed for FSOD
- Yolov8_DeFRCN: Our hybrid approach, combining DeFRCN components with YOLOv8

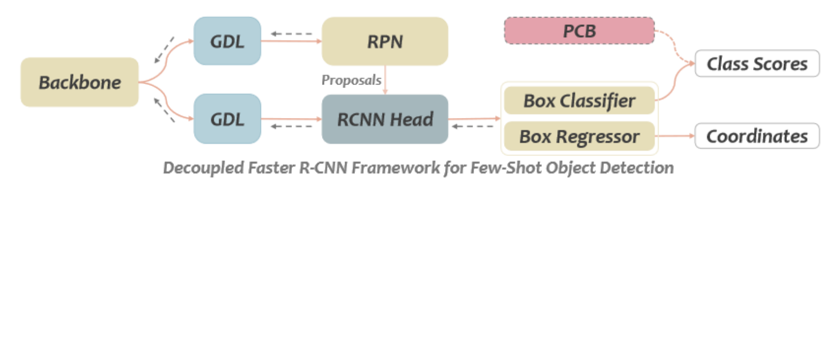
---

[1] Qiao, Limeng et al. 2021. "Defrcn: Decoupled faster r-cnn for few-shot object detection" *Proceedings of the IEEE/CVF International Conference on Computer Vision.*

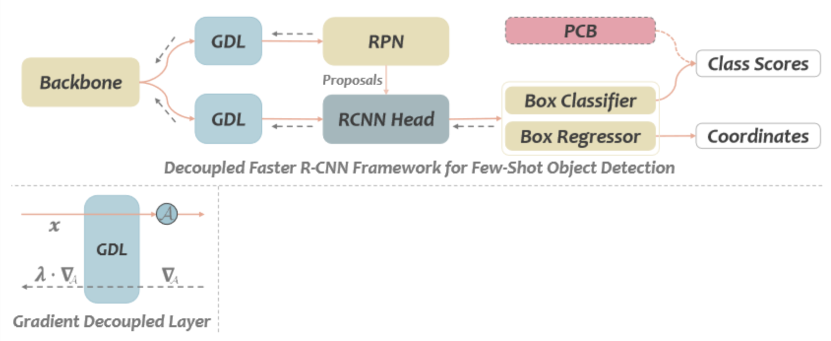
[2] Fu, Yuqian et al. 2024. "Cross-Domain Few-Shot Object Detection via Enhanced Open-Set Object Detector" *arXiv preprint arXiv:2402.03094.*

[3] Jocher, Glenn, Ayush Chaurasia, and Jing Qiu 2023. *YOLO by Ultralytics.*

2. DeFRCN, a Faster R-CNN based method

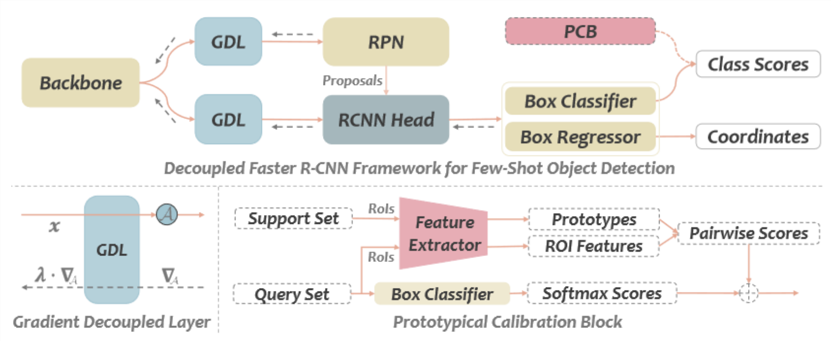


GDL and PCB modules in DeFRCN.



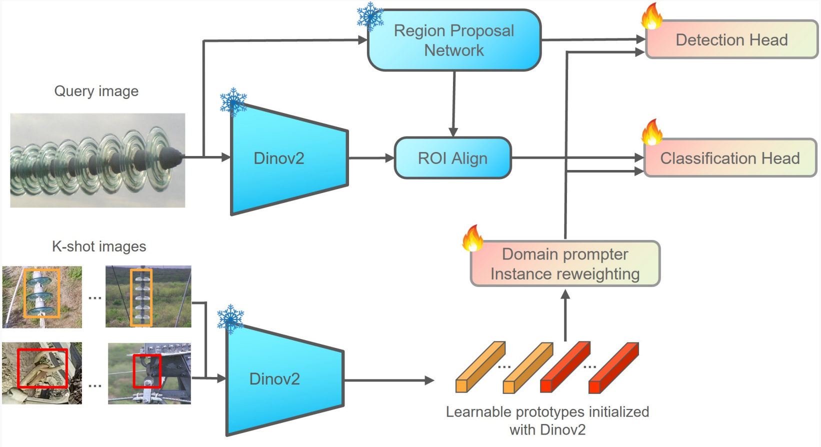
GDL and PCB modules in DeFRCN.

2. DeFRCN, a Faster R-CNN based method



GDL and PCB modules in DeFRCN.

2. CD-ViTO, a SSL based method



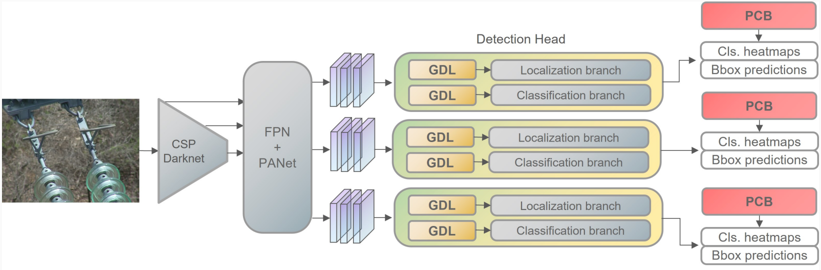
CD-ViTO uses DINOv2[4] pretrained with SSL on 142 million images.

[4]Oquab, Maxime et al. 2023. "DINOv2: Learning Robust Visual Features without Supervision" *arXiv:2304.07193*.

2. Our adaptation of Yolov8

Traditional YOLO: Shared weights for localization and classification.

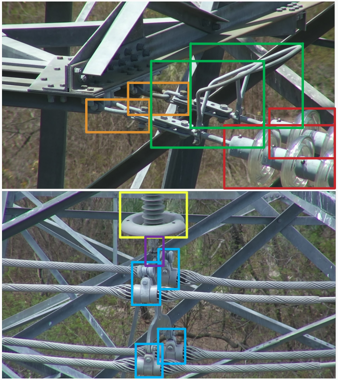
⇒ **YOLOv8_DeFRCN** combines DeFRCN with YOLOv8's decoupled head to achieve faster and more accurate FSOD:



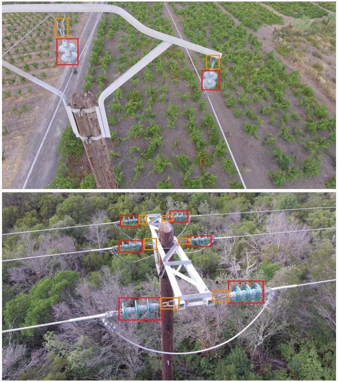
- YOLOv8's decoupled head enables the use of GDL.
- PCB can be computed from the classification branch outputs.

3. Two power line components detection datasets

Insplad[5]



A private dataset from Enedis



⇒ Insplad offers close-up views of components, while our private dataset provides a broader, aerial perspective.

[5] (Silva, et al Andre Luiz Buarque Vieira e 2023. "InsPLAD: A Dataset and Benchmark for Power Line Asset Inspection in UAV Images" *International Journal of Remote Sensing*)

3. Real world FSOD

Insplad

- 18 components
- Evaluated on 2,627 RGB images (1920×1080)

Enedis dataset

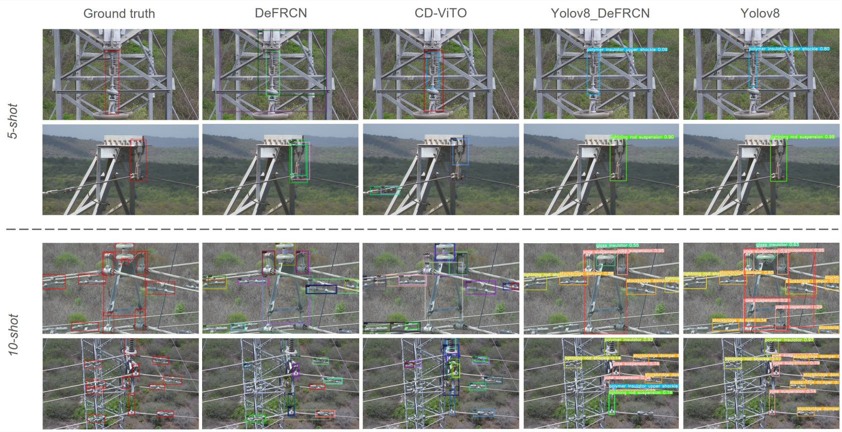
- 2 components
- Evaluated on 122 RGB images (4096×2160)

In real-world FSOD scenarios, classes rarely have the same number of instances, we conducted our experiments on:

- 5-shot and 10-shot splits on Insplad, balanced as much as possible
- 12-shot split on the Enedis dataset

4. Qualitative results

We performed few-shot fine-tuning of all methods using their base weights pretrained on the FSOD-adapted version of MSCOCO[6]:



[6] Wang, Xin et al. 2020. "Frustratingly Simple Few-Shot Object Detection" *Proceedings of the 37th International Conference on Machine Learning*.

4. Results on Insplad

Mean average precision on novel classes (nAP) in our 5-shot and 10-shot splits of the Insplad dataset

Method	Shot	nAP50	nAP50-95
Yolov8	5-shot	**52.5**	**38.6**
Yolov8_DeFRCN	5-shot	45.6	33.0
DeFRCN	5-shot	34.2	20.5
CD-VITO	5-shot	50.5	32.4
Yolov8	10-shot	**62.6**	**45.7**
Yolov8_DeFRCN	10-shot	59.6	44.1
DeFRCN	10-shot	53.5	31.0
CD-VITO	10-shot	59.0	37.6

4. Results on Enedis dataset

Mean average precision on novel classes (nAP) in our 12-shot split of our private dataset

12-shot	**Insulator**		**Shackle**	
	nAP50	**nAP50-95**	**nAP50**	**nAP50-95**
Yolov8	**89.8**	64.7	41.4	**20.7**
Yolov8_DeFRCN	89.2	**67.6**	**42.8**	19.6
DeFRCN	84.8	57.6	17.0	6.0
CD-ViTO	84.9	56.8	29.5	8.3

4. Results on Enedis dataset

Mean average precision on novel classes (nAP) in our 12-shot split of our private dataset

12-shot	**Insulator**		**Shackle**	
	nAP50	**nAP50-95**	**nAP50**	**nAP50-95**
Yolov8	**89.8**	64.7	41.4	**20.7**
Yolov8_DeFRCN	89.2	**67.6**	**42.8**	19.6
DeFRCN	84.8	57.6	17.0	6.0
CD-ViTO	84.9	56.8	29.5	8.3

$\Rightarrow$ Surprisingly, FSOD-specialized methods do **NOT** deliver the best performance on either dataset.

4. Why do FSOD methods underperform here?

Surprisingly, Yolov8_DeFRCN underperforms on both datasets despite outperforming Yolov8 on FSOD-adapted MS COCO:

	3-shot				5-shot			
	bAP50	bAP50-95	nAP50	nAP50-95	bAP50	bAP50-95	nAP50	nAP50-95
Yolov8	19.1	13.4	9.3	5.6	19.3	13.6	11.7	7.4
Yolov8_DeFRCN	**55.9**	**40.4**	15.9	10.8	**53.0**	**38.1**	23.0	15.0
DeFRCN	49.8	32.5	**24.2**	**13.4**	50.6	33.1	**28.4**	**15.3**

Hypothesis: The few-shot performance of some FSOD methods is heavily influenced by their pretraining, leading to strong biases.

4. Why do FSOD methods underperform here?

Surprisingly, Yolov8_DeFRCN underperforms on both datasets despite outperforming Yolov8 on FSOD-adapted MS COCO:

	3-shot				**5-shot**			
	bAP50	bAP50-95	nAP50	nAP50-95	bAP50	bAP50-95	nAP50	nAP50-95
Yolov8	19.1	13.4	9.3	5.6	19.3	13.6	11.7	7.4
Yolov8_DeFRCN	**55.9**	**40.4**	15.9	10.8	**53.0**	**38.1**	23.0	15.0
DeFRCN	49.8	32.5	**24.2**	**13.4**	50.6	33.1	**28.4**	**15.3**

Hypothesis: The few-shot performance of some FSOD methods is heavily influenced by their pretraining, leading to strong biases.

PCB gain	**MSCOCO**		**Insplad**	
Splits	5-shot	10-shot	5-shot	10-shot
nAP50	+1.7	+1.7	-3.4	-3.1
nAP50-95	+1.0	+1.0	-0.1	-0.2

5. Conclusion

To conclude, this work:

- FSOD methods often perform poorly in real-world scenarios and are highly sensitive to bias, unlike their performance on standard benchmarks.
- Cross-Domain Few-Shot Object Detection (CD-FSOD) is a relatively new field, with growing interest through initiatives like the NTIRE workshop at CVPR2025.
- Another emerging approach is the use of **V**ision-**L**anguage **M**odels (VLMs) for zero-shot object detection, with promising methods like YOLO-World[7] and Grounding-DINO[8], which present interesting directions for future work.

---

[7] Cheng, Tianheng et al. 2024. "Yolo-world: Real-time open-vocabulary object detection".

[8] Liu, Shilong et al. 2024. "Grounding dino: Marrying dino with grounded pre-training for open-set object detection".

Thanks for your attention!
Questions?

References i

- Cheng, Tianheng et al. (2024). **"Yolo-world: Real-time open-vocabulary object detection".** In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 16901–16911.
- Fu, Yuqian et al. (2024). **"Cross-Domain Few-Shot Object Detection via Enhanced Open-Set Object Detector".** In: *arXiv preprint arXiv:2402.03094*.
- Jocher, Glenn, Ayush Chaurasia, and Jing Qiu (Jan. 2023). **YOLO by Ultralytics.** Version 8.0.0. URL: https://github.com/ultralytics/ultralytics.
- Liu, Shilong et al. (2024). **"Grounding dino: Marrying dino with grounded pre-training for open-set object detection".** In: *European Conference on Computer Vision*. Springer, pp. 38–55.
- Oquab, Maxime et al. (2023). **"DINOv2: Learning Robust Visual Features without Supervision".** In: *arXiv:2304.07193*.

References ii

- Qiao, Limeng et al. (2021). **"Defrcn: Decoupled faster r-cnn for few-shot object detection".** In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 8681–8690.
- Silva, et al Andre Luiz Buarque Vieira e (2023). **"InsPLAD: A Dataset and Benchmark for Power Line Asset Inspection in UAV Images".** In: *International Journal of Remote Sensing* 44.23, pp. 1–27. DOI: 10.1080/01431161.2023.2283900. eprint: https://doi.org/10.1080/01431161.2023.2283900. URL: https://doi.org/10.1080/01431161.2023.2283900.
- Wang, Xin et al. (Apr. 2020). **"Frustratingly Simple Few-Shot Object Detection".** In: *Proceedings of the 37th International Conference on Machine Learning* 119. Ed. by Hal Daumé III and Aarti Singh, pp. 9919–9928. URL: https://proceedings.mlr.press/v119/wang20j.html.