

Pierre Le Jeune^{1,2}, Anissa Mokraoui¹

¹L2TI, Université Sorbonne Paris Nord, ²COSE

I – Détection d'objets *Few-Shot*

La détection d'objets *few-shot* (FSOD) consiste à localiser et classifier tous les objets appartenant à un ensemble de classes définies un nombre limité d'exemples. On parle de détection *n-ways k-shots* lorsque l'on dispose de *k* exemples pour chacune des *n* classes du problème.

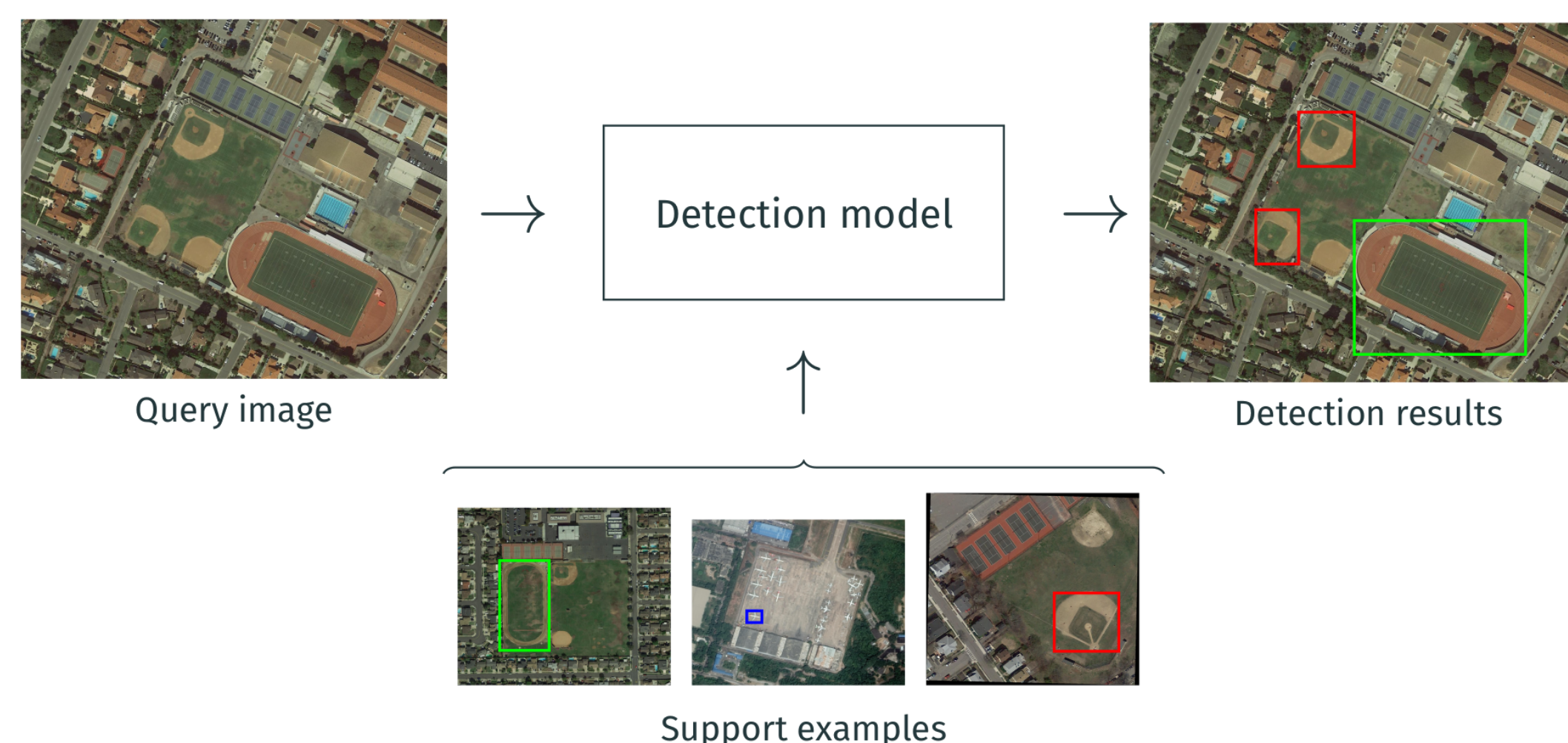


Fig. 1: Principe général de la détection *few-shot*.

II – Spécificité des images aériennes

Les objets présents dans les images aériennes sont beaucoup plus petits que ceux présents dans les images naturelles. Dans les images aériennes, certaines classes contiennent uniquement de petits objets alors que dans les images naturelles, la taille des objets varie plus, y compris au sein d'une même classe.

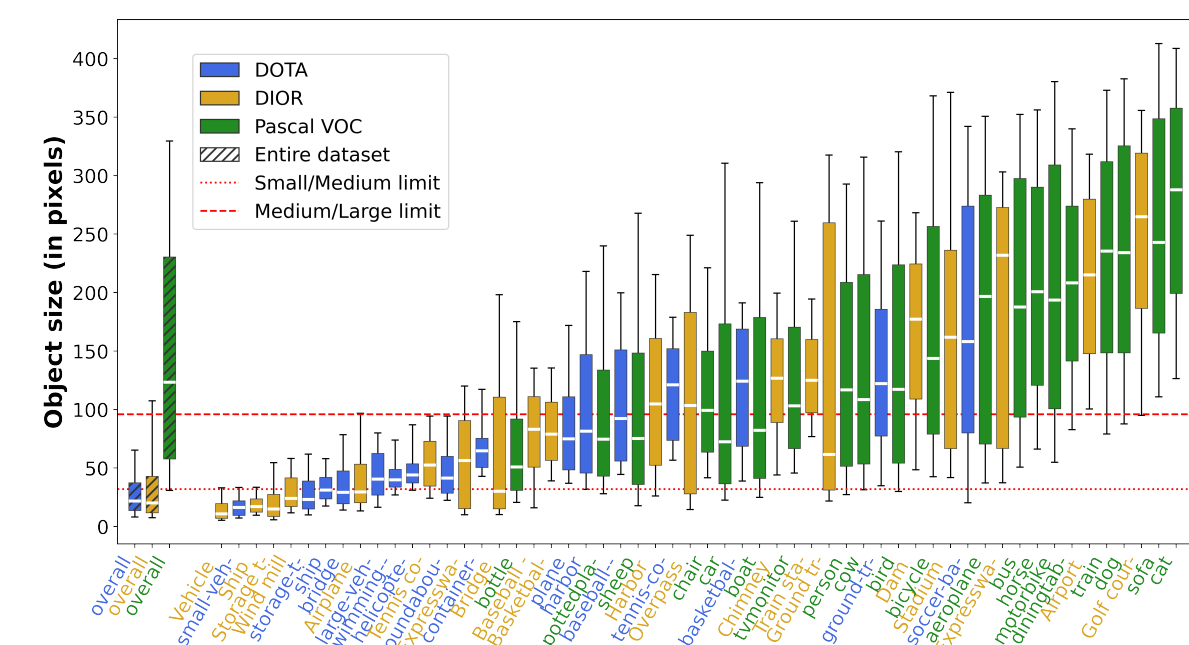


Fig. 2: Distribution des tailles d'objets pour chacune des classes des jeux de données DOTA, DIOR et Pascal VOC.

III – Analyse des performances

Les méthodes existantes (ex. FRW [2]) pour la détection *few-shot* fonctionnent relativement bien sur des images naturelles. Mais les performances baissent considérablement lorsque celles-ci sont appliquées à des images aériennes. Cela s'observe principalement sur les classes nouvelles (en rouge sur la Figure 3) mais aussi sur les classes de base dans une moindre mesure. Pour notre analyse, trois jeux de données ont été sélectionnés: DOTA [6], DIOR [3] (images aériennes) et Pascal VOC [1] (images naturelles).

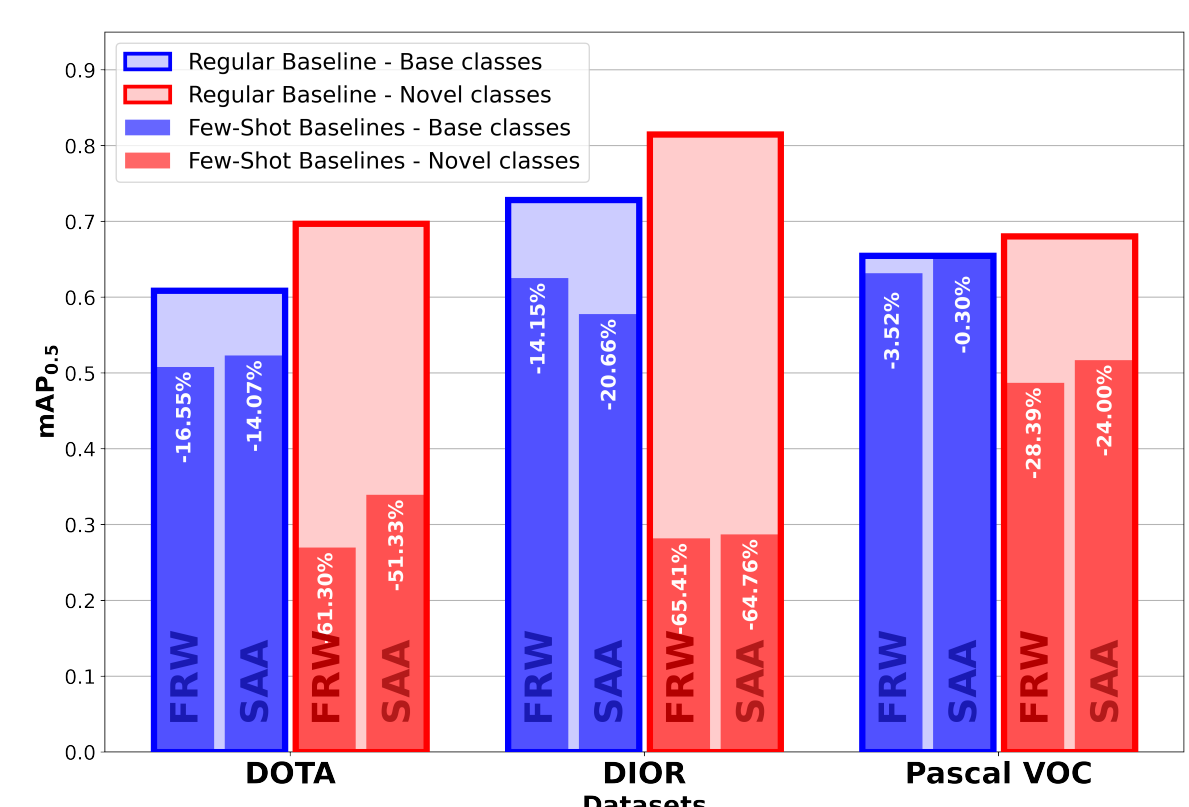


Fig. 3: Comparaison des performances de deux méthodes de détection *few-shot*: FRW [2] et SAA [7] sur 3 jeux de données: DOTA, DIOR et Pascal VOC. Les histogrammes transparents représentent les performances de FCOS [5] (c.a.d. sans *few-shot*).

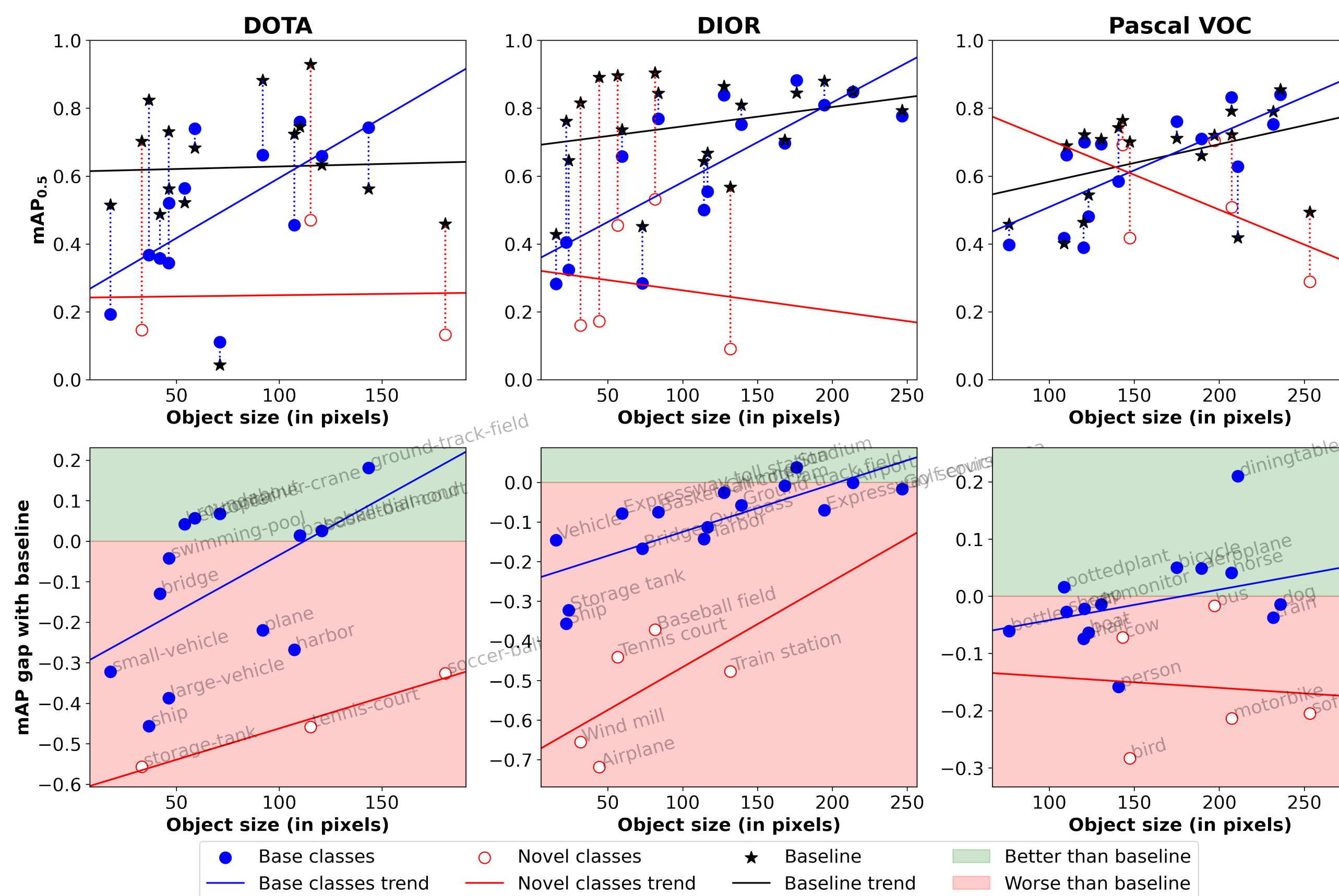


Fig. 4: Comparaison des performances de FRW [2] par classe et sur DOTA, DIOR et Pascal VOC.

En comparant l'écart entre les performances *few-shot* et la baseline FCOS (cf. [5]) par classe, on observe une tendance claire. L'écart avec la baseline diminue avec la taille des objets. Plus un objet est grand, plus il est facile de le détecter. C'est une conclusion déjà connue pour la détection d'objet, mais ici, puisque l'on analyse l'écart de performance avec la baseline, il s'agit d'un phénomène nouveau. Il est plus difficile de conditionner un modèle à partir d'exemples pour détecter des petits objets que des grands.

IV – Améliorations pour les petits objets

Afin d'améliorer la qualité des détections sur les objets de petite taille, nous proposons une nouvelle méthode d'extraction de caractéristiques dans les images de support. L'idée principale est de découper la zone autour de l'objet de support et de redimensionner le patch à une taille fixe. Les petits objets sont agrandis et leurs caractéristiques plus faciles à reconnaître, facilitant le conditionnement du réseau. En complément, un processus d'augmentation de données est défini pour améliorer la robustesse du modèle. Les techniques classiques d'augmentation ne sont pas toutes compatibles avec la détection d'objets, car elles peuvent masquer l'ensemble des objets. C'est notamment le cas des méthodes *random cut-out* et *random crop-resize*. Afin de résoudre ce problème et de bénéficier des avantages de ces méthodes, celles-ci sont modifiées pour être appliquées au niveau des objets et non de l'image entière (voir Figure 5).

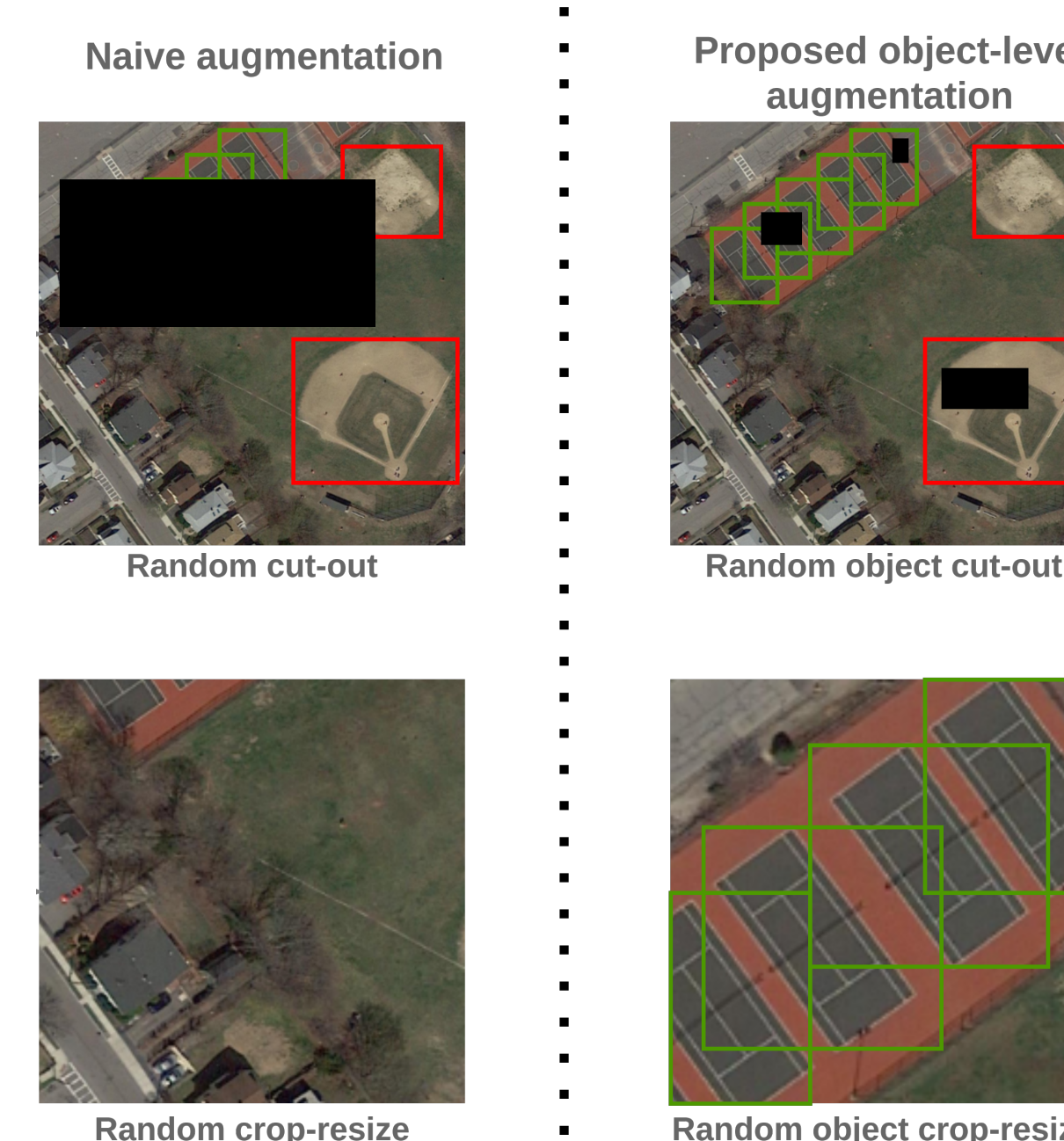


Fig. 5: Augmentations au niveau des objets.

V – Résultats

Le choix de la méthode d'extraction des caractéristiques des exemples ainsi que le processus d'augmentation mis en place permettent d'améliorer sensiblement les résultats pour les classes nouvelles des trois jeux de données étudiés: DOTA, DIOR et Pascal VOC. Ce gain de performance est plus important pour les images aériennes (DOTA et DIOR) que pour les images naturelles. Cela suggère que notre méthode gère mieux les objets de plus petite taille. En revanche, cette amélioration s'accompagne d'une baisse de performance sur les classes de base. Cela apparaît lorsque l'augmentation est appliquée aux exemples de support. Une étude plus poussée est nécessaire pour comprendre les détails de ce phénomène.

	DOTA		DIOR		Pascal VOC	
	FRW	SAA	FRW	SAA	FRW	SAA
Base classes	0.495	0.485	0.523	0.467	0.625	0.615
Novel classes	0.283	0.371	0.339	0.351	0.282	0.356

Tab. 1: Comparaison des performances obtenues avec notre nouvelle méthode d'extraction des supports. Celle-ci est testée sur DOTA, DIOR et Pascal VOC et avec deux méthodes différentes de détection *few-shot* FRW et SAA [7]

VI – Conclusions

- La détection *few-shot* est très **difficile sur des images aériennes** car celles-ci contiennent majoritairement des petits objets.
- Les **petits objets sont problématiques** de deux manières: 1) ils sont **difficiles à repérer** dans les images car leurs caractéristiques sont très localisées; 2) ils sont **difficiles à classifier** et sont ainsi de mauvais exemples pour conditionner la détection.
- Les méthodes **d'extraction des caractéristiques** des exemples et **d'augmentation** sont cruciales pour la détection *few-shot*.

Remerciements

Les auteurs remercient l'entreprise COSE pour leur collaboration étroite et le financement de ce projet.

Références

- [1] Mark EVERINGHAM et al. "The pascal visual object classes (voc) challenge". In: *International journal of computer vision* 88.2 (2010), p. 303-338.
- [2] Bingyi KANG et al. "Few-shot object detection via feature reweighting". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019, p. 8420-8429.
- [3] Ke LI et al. "Object detection in optical remote sensing images: A survey and a new benchmark". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 159 (2020), p. 296-307.
- [4] Xiang LI, Jingyu DENG et Yi FANG. "Few-Shot Object Detection on Remote Sensing Images". In: *IEEE Transactions on Geoscience and Remote Sensing* (2021), p. 1-14.
- [5] Zhi TIAN et al. "Fcos: Fully convolutional one-stage object detection". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019, p. 9627-9636.
- [6] Gui-Song XIA et al. "DOTA: A large-scale dataset for object detection in aerial images". In: *Proceedings of the IEEE Conference on CVPR*. 2018, p. 3974-3983.
- [7] Zixuan XIAO et al. "Few-Shot Object Detection With Self-Adaptive Attention Network for Remote Sensing Images". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14 (2021), p. 4854-4865.