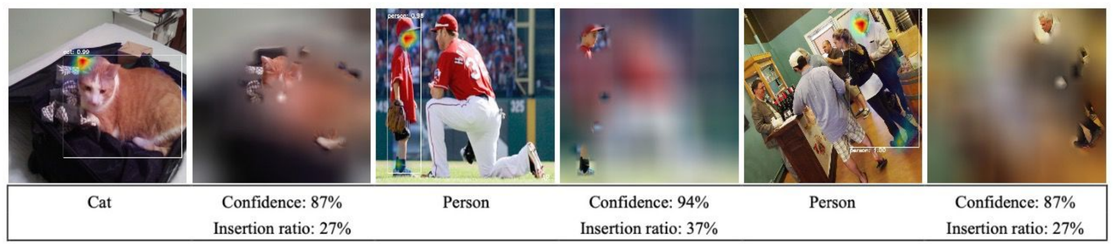
Diverse Explanations for Object Detectors with Nesterov-Accelerated iGOS++

Mingqi Jiang, Saeed Khorram, Li Fuxin

Collaborative Robotics and Intelligent Systems (CoRIS) Institute, Oregon State University

Introduction

- **Motivation**
  - Most existing work focus solely on image classification
  - Most heatmap approaches generate only a single heatmap for each image.



Examples generated by our proposed NAG-iGOS++

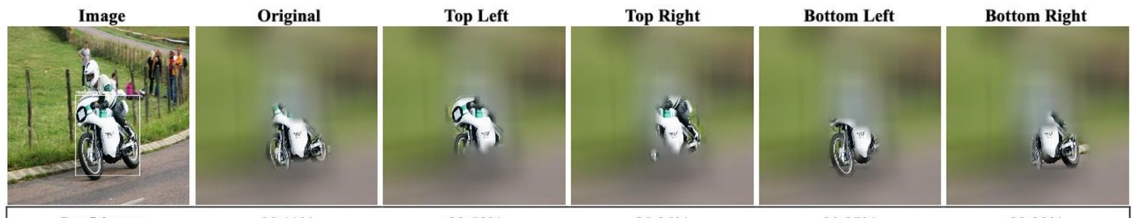
- **Contributions**
  - We extend iGOS++ to object detection tasks.
  - We propose to use Nesterov Accelerated Gradient (NAG) in iGOS++ to replace the line search, which speeds up the algorithm by 2× and improves the performance.
  - We propose a scheme that initializes iGOS++ with multiple starting masks, which further improved performance and makes the algorithm capable of generating multiple explanations for a single detection.

Method

- **Nesterov Accelerated Gradient** Using Nesterov Accelerated Gradient (NAG) in iGOS++ to replace the line search (LS). We update the mask with NAG as:

$$\omega^{k+1} = M^k - \alpha \cdot TG(M^k)$$
$$M^{k+1} = \omega^{k+1} + \varepsilon(\omega^{k+1} - \omega^k)$$

  - M is the mask, $\mathcal{E} = k / (k+3)$. α is the learning rate.
- **Diverse Initialization**
  - Generating K^2 different initializations on a $K \times K$ grid
  - Initializing the mask with nonzero values in only one cell for each initialization



Examples generated by NAG-iGOS++ without and with initialization using different regions of predicted mask in insertion tasks using Mask R-CNN (*K*=*2*)

Experiments

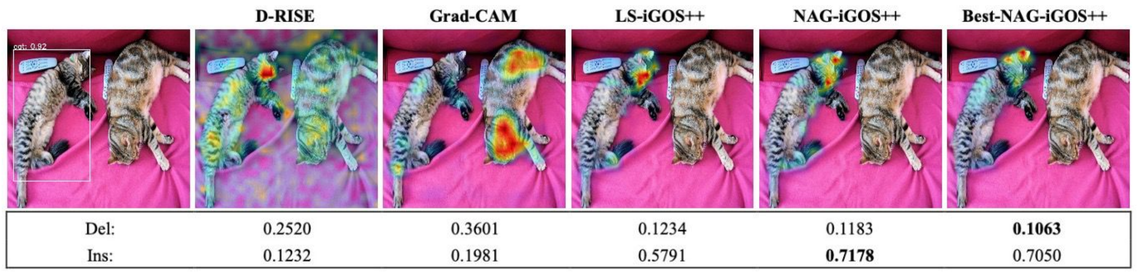
- **Insertion, Deletion Scores and Runtime on MSCOCO**

Resolution / Method	16 × 16 Del ↓	16 × 16 Ins ↑	16 × 16 Time(s)	25 × 25 Del ↓	25 × 25 Ins ↑	25 × 25 Time(s)	100 × 100 Del ↓	100 × 100 Ins ↑	100 × 100 Time(s)
D-RISE	0.6422	0.6322	220	– –	– –	– –	– –	– –	– –
Grad-CAM	– –	– –	– –	0.7839	0.3048	**7**	– –	– –	– –
LS-iGOS++	0.5630	0.6692	146	0.4685	0.6210	138	0.2370	0.5455	115
NAG-iGOS++	0.5577	0.6760	**62**	0.4641	0.6285	62	0.2380	0.5478	**62**
Best-NAG-iGOS++	**0.5388**	**0.6952**	248	**0.4399**	**0.6500**	248	**0.2048**	**0.5950**	248

Quantitative comparison using Mask R-CNN

Resolution / Method	16 × 16 Del ↓	16 × 16 Ins ↑	16 × 16 Time(s)	64 × 64 Del ↓	64 × 64 Ins ↑	64 × 64 Time(s)
D-RISE	0.4985	0.4953	70	– –	– –	– –
Grad-CAM	0.6980	0.2210	**2**	– –	– –	– –
LS-iGOS++	0.4804	0.4833	40	0.2475	0.2968	41
NAG-iGOS++	0.4688	0.4922	14	0.2384	0.3201	**14**
Best-NAG-iGOS++	**0.4015**	**0.5403**	56	**0.1704**	**0.3846**	56

Quantitative comparison using YOLOv3-SPP



Visualization of heatmap generated from different methods

- **Ablation Study**

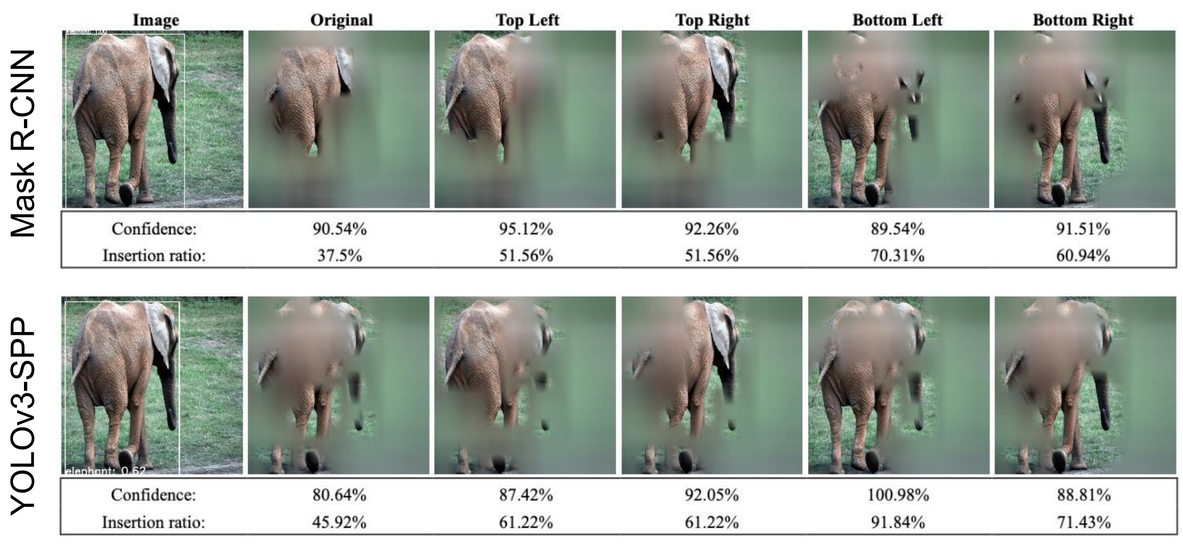
Optimization	16 × 16 Del ↓	16 × 16 Ins ↑	16 × 16 Time(s)	25 × 25 Del ↓	25 × 25 Ins ↑	25 × 25 Time(s)	100 × 100 Del ↓	100 × 100 Ins ↑	100 × 100 Time(s)
LS	0.5630	0.6692	146	0.4685	0.6210	138	0.2370	0.5455	115
Adam	0.5747	0.6439	62	0.5040	0.5877	62	0.2997	0.4329	62
Nadam	0.6272	0.5895	62	0.5843	0.5280	62	0.4621	0.2951	62
NAG	0.5577	0.6760	**62**	0.4641	0.6285	**62**	0.2380	0.5478	**62**

Ablation study on optimization methods using Mask R-CNN

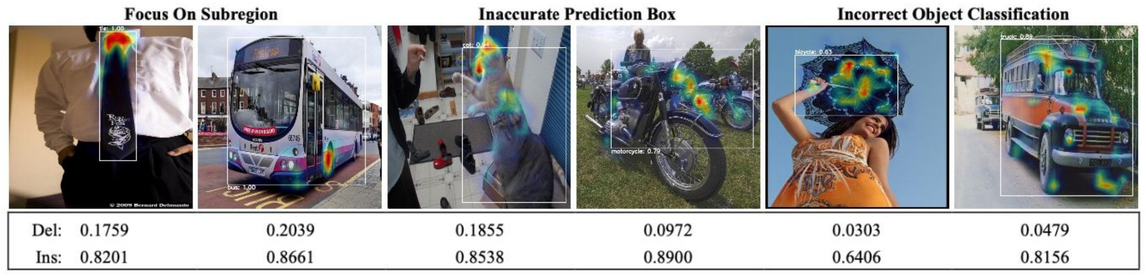
Best-NAG-iGOS++	16 × 16 Del ↓	16 × 16 Ins ↑	16 × 16 Time(s)	25 × 25 Del ↓	25 × 25 Ins ↑	25 × 25 Time(s)	100 × 100 Del ↓	100 × 100 Ins ↑	100 × 100 Time(s)
K = 2	0.5388	0.6952	**248**	0.4399	0.6285	**248**	0.2048	0.5950	**248**
K = 3	**0.5237**	**0.7044**	558	**0.4248**	**0.6601**	558	**0.1936**	**0.6036**	558

Ablation study with multiple initialization K using Mask R-CNN

- **Multiple initializations**



- **More Visualizations using Mask R-CNN**



Conclusion

- This paper presents NAG-iGOS++, an algorithm that extends the iGOS++ algorithm to explain object detection networks
- We proposed to use Nesterov Accelerated Gradient which improved the efficiency and accuracy of explanations
- Our multiple initializations provided a more complete picture of the object detection and segmentation networks to be explained.

The work is partially supported by NSF-1751402, NSF-1927564, ONR- UWSC12017-BPO49408 and ONR-N00014-21-1-2052. The authors would also like to thank Dr. Qing Tao for helpful discussions.