

Rakshit Naidu^{†‡}, Haofan Wang[‡], Soumya Snigdha Kundu[§],
Ankita Ghosh[†], Yash Maurya[†], Shamanth R Nayak K[†], Joy Michael[†]

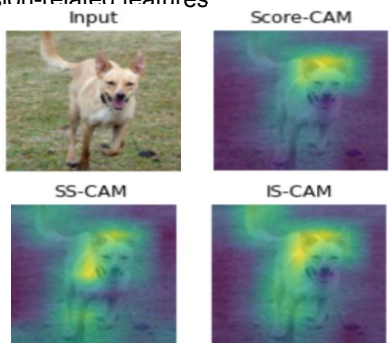
[†]Manipal Institute of Technology, [‡]Carnegie Mellon University, [§]SRM Institute of Science and Technology

INTRODUCTION

- Class Activation Mappings (CAMs) highlight the features that contribute to the output of the model.

- We present two new variants of Score-CAM:-
> First, by a smoothing function to generate localised features.
> Second, through an Integration function to furnish sharper axiomatic-based attribution maps.

- We visually demonstrate that our methods significantly assist in interpreting models by providing concentrated heatmaps and concrete decision-related features



METHODOLOGY



SS-CAM Pipeline (IS-CAM Involves the integration function)

$$L^c = ReLU \left(\sum_k \alpha_k^c A_l^k \right)$$

where

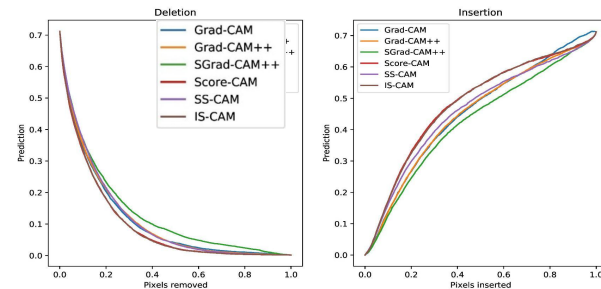
$$\alpha_k^c = \frac{\sum_{i=1}^N (C(M_i))}{N} \quad \text{for SS-CAM}$$

$$\alpha_k^c = \frac{\sum_{i=1}^N (C(M_i))}{N} \quad \text{for IS-CAM}$$

$$M_{i+1} \leftarrow M_i + \left((X_0 * A_l^k) * \frac{i}{N} \right)$$

SS-CAM & IS-CAM Equations

EXPERIMENTS



Average AUC Insertion and Deletion curves

CAM Techniques	VGG-16		Resnet		SqueezeNet	
	Avg Drop%	Avg Inc%	Avg Drop%	Avg Inc%	Avg Drop%	Avg Inc%
Score-CAM	66.03	51.85	64.23	53.55	13.42	60.85
SS-CAM	79.15	51.30	64.53	54.80	12.06	64.85
IS-CAM	63.30	52.35	64.85	53.50	13.00	62.15

Average Drop and Average Increase % Scores

CONCLUSION

- Generated concentrated heatmaps with concrete decision related features

References:

- 1) Wang, H., Wang, Z., Du, M., Yang, F., Zhang, Z., Ding, S., Mardziel, P. and Hu, X., 2020. Score-CAM: Score-weighted visual explanations for convolutional neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 24-25).

