F-CAM: Full Resolution Class Activation Maps via Guided Parametric Upscaling

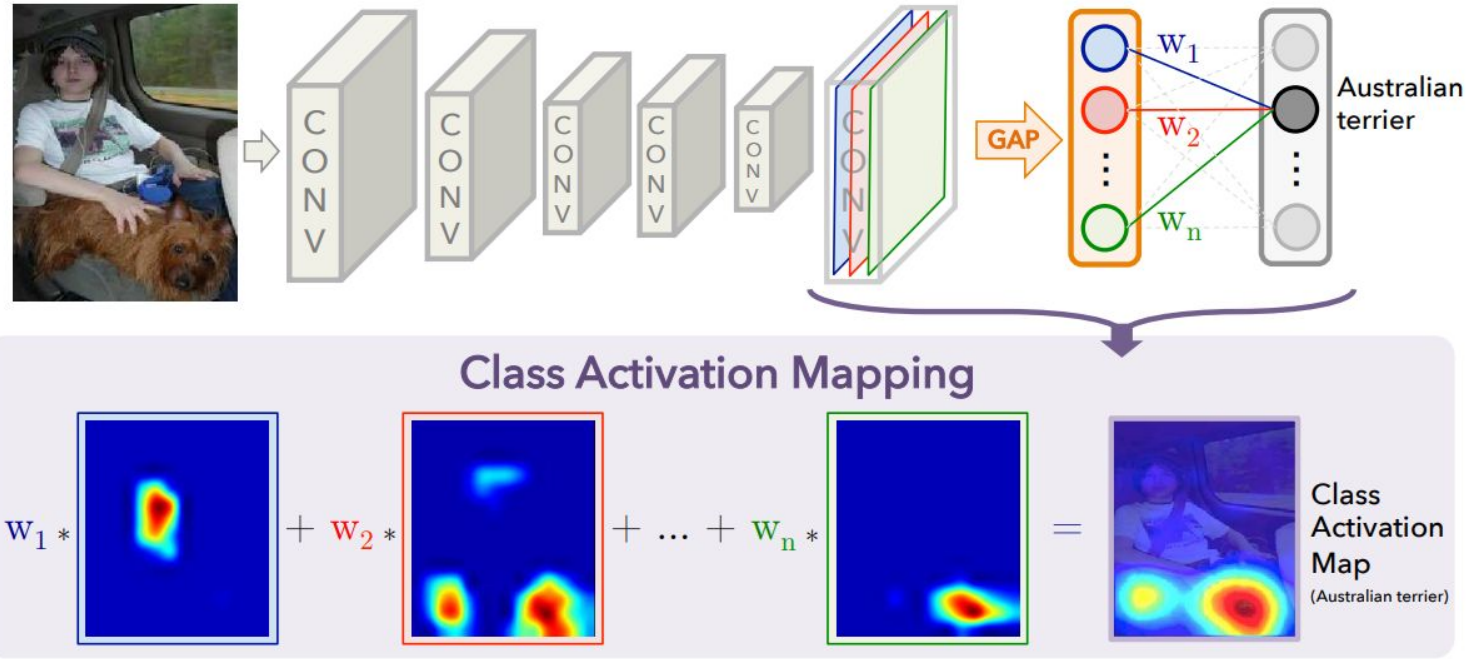
#1177

Soufiane Belharbi, Aydin Sarraf, Marco Pedersoli, Ismail Ben Ayed, Luke McCaffrey, Eric Granger

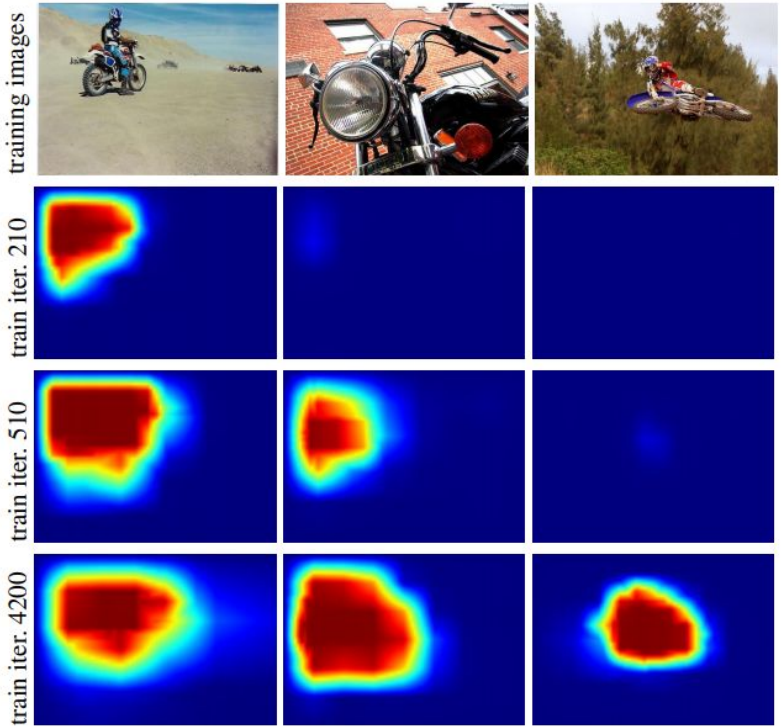
WACV 2022

1. **Challenges:** WSOL with CAMs

Application: Localizing objects using global annotation



1. **Challenges:** WSOL with CAMs (issues)



Standard cams interpolated

CAMs: low resolution (convolution, downsampling)

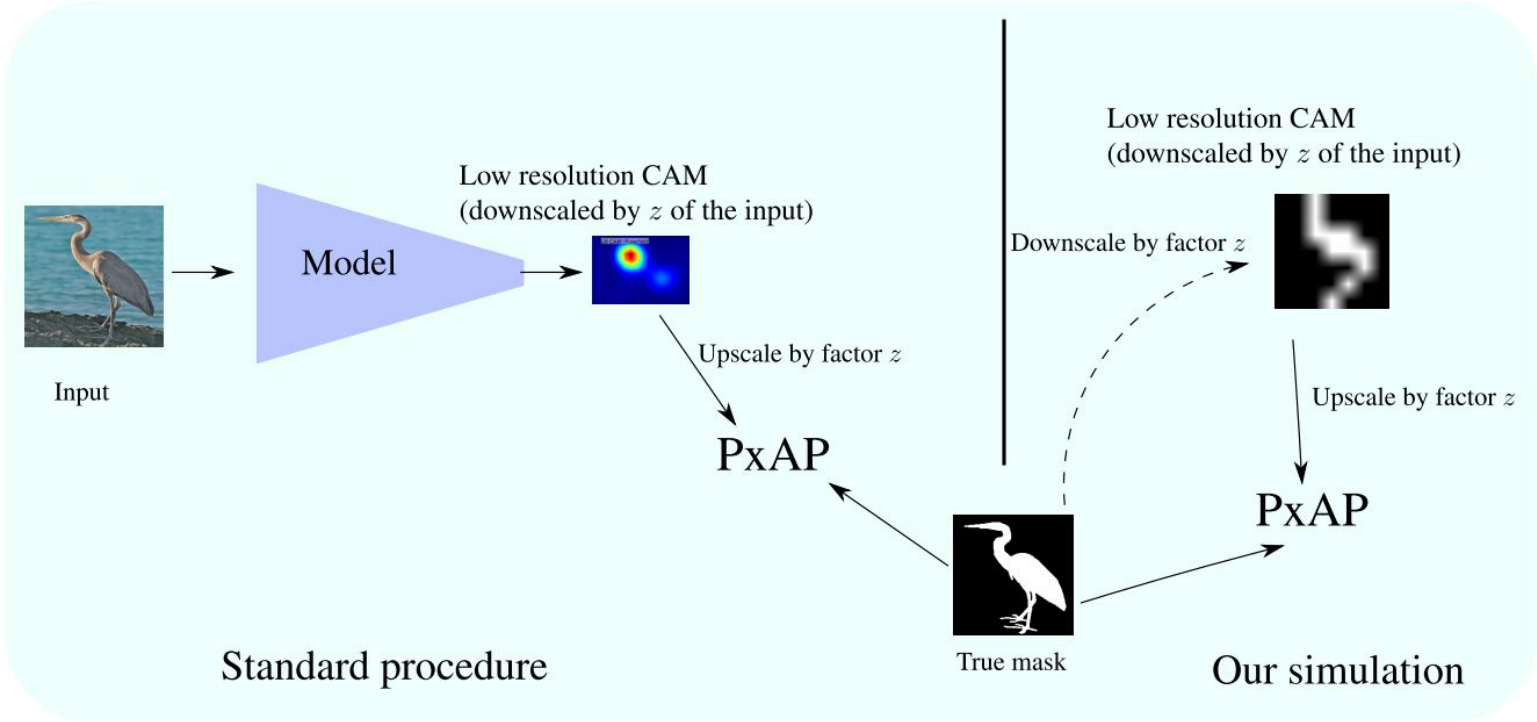
↓

Negative impact on localization performance

Standard cams 8x8 resolution
Downscale factor 32.

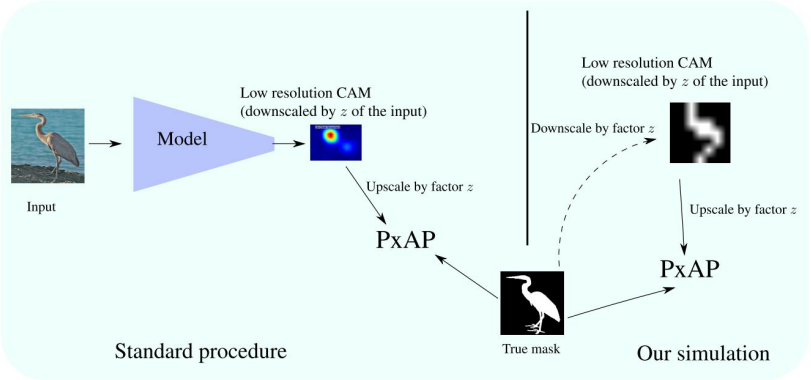
1. Challenges:

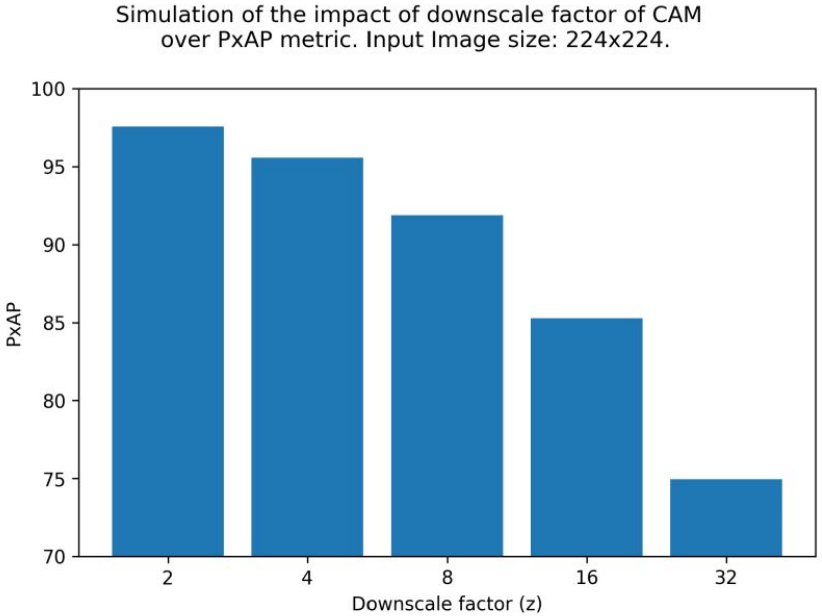
Impact of CAMs size over localization performance (issues)



1. Challenges:

Impact of CAMs size over localization performance (issues)





2. **Proposal:** Improve CAM resolution

- Avoid interpolation

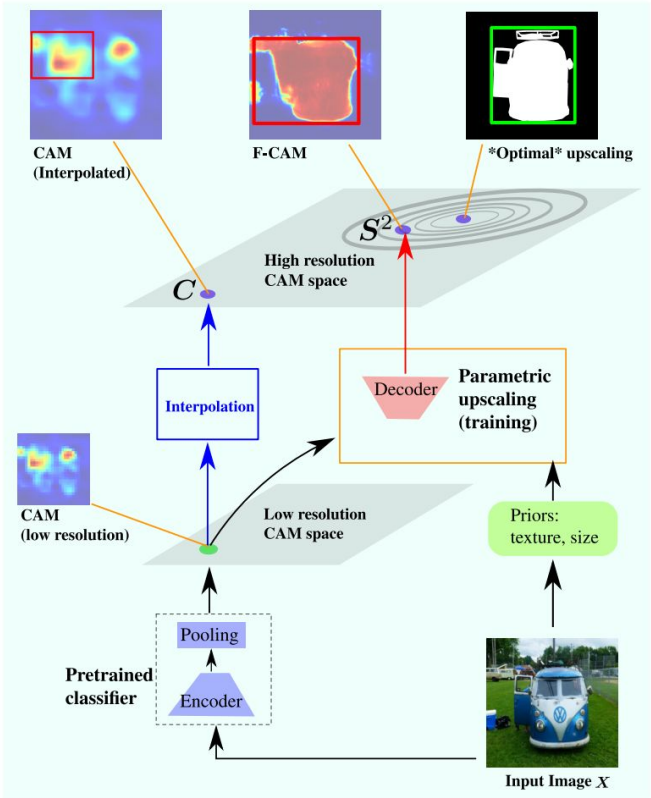
  → use learnable upscaling function

- Guide the upscaling

Training loss := SR (CAM) + CRF (Image) + ASC (Size)

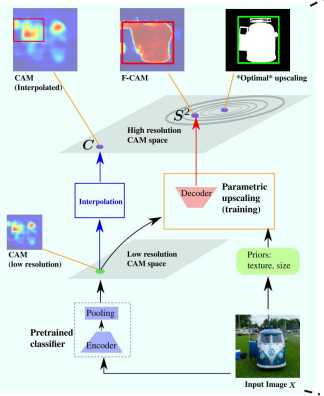
SR: seeds
CRF: image properties
ASC: unsupervised size constraint

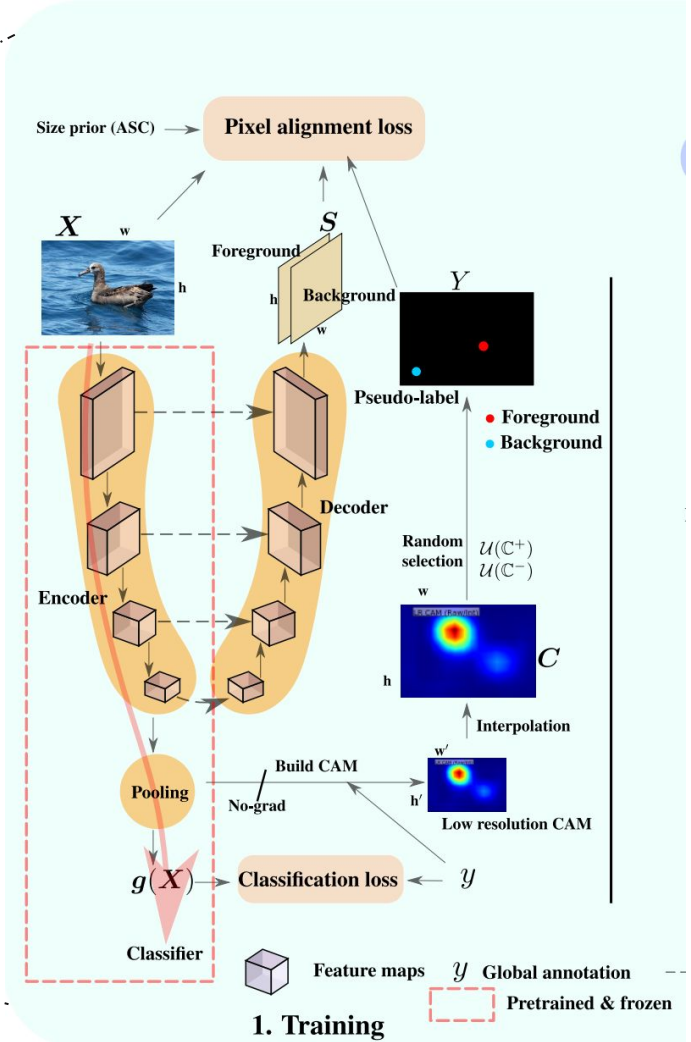
Exploit a trained classifier



2. Proposal:

Improve CAM resolution

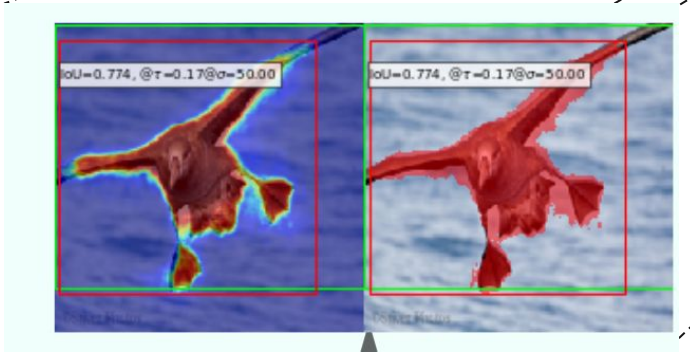


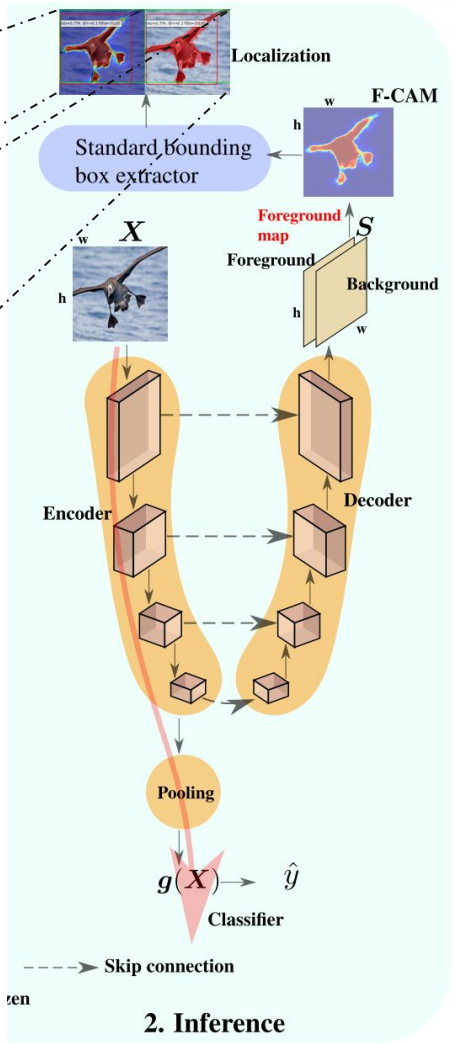


Model foreground and background

2. Proposal:

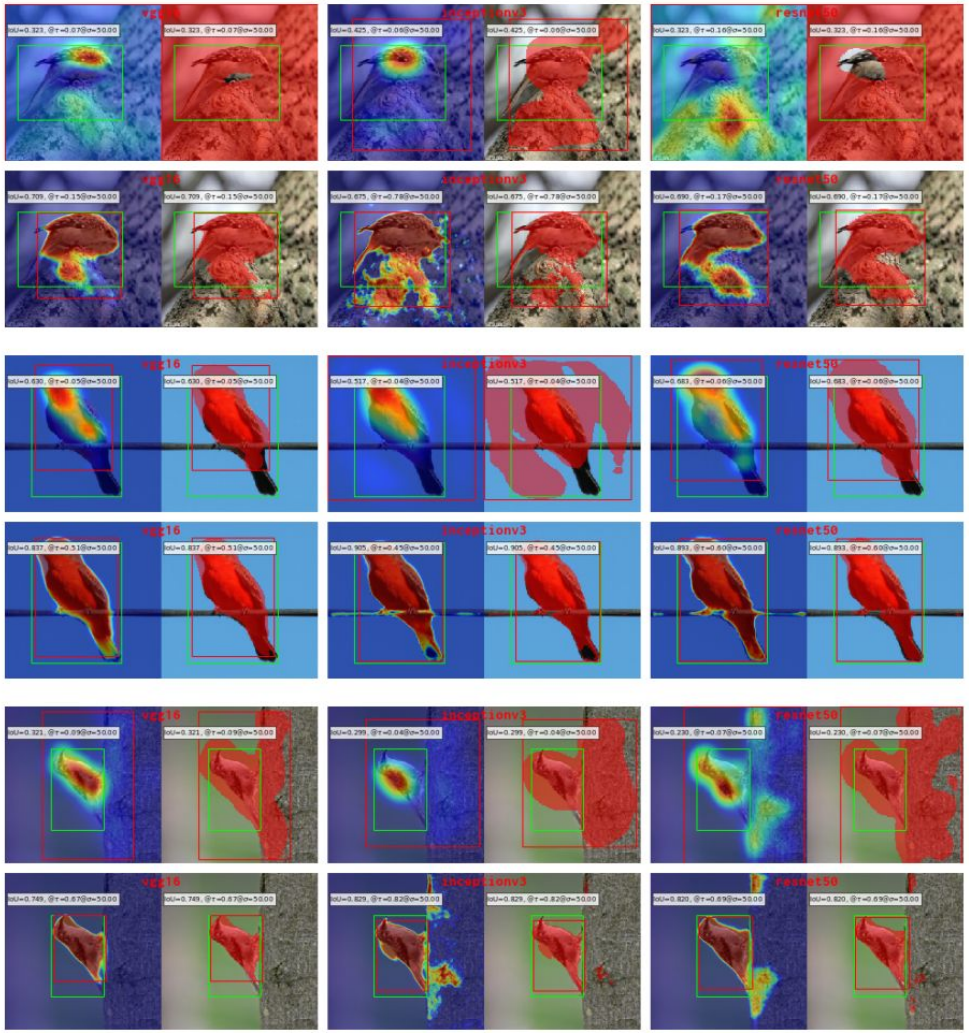
Improve CAM resolution





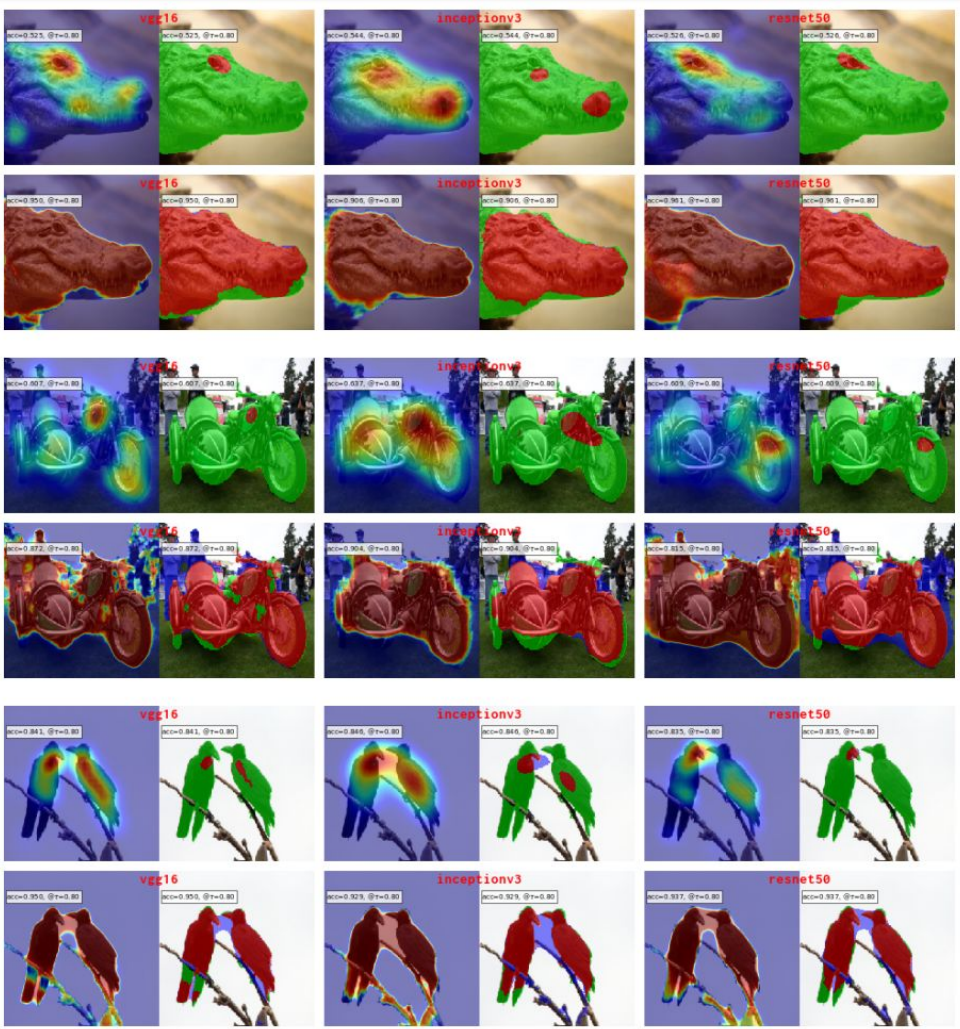
3. Experiments:
Results

CUB dataset



3. Experiments:
Results

OpenImages dataset



3. Experiments:
Results

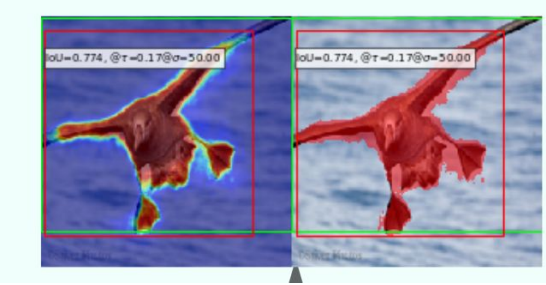
Finetune simple baselines

Methods	CUB (MaxBoxAcc) VGG	Inception	ResNet	Mean	OpenImages (PxAP) VGG	Inception	ResNet	Mean
CAM [57] *(cvpr,2016)*	71.1	62.1	73.2	68.8	58.1	61.4	58.0	59.1
HaS [34] *(iccv,2017)*	76.3	57.7	78.1	70.7	56.9	59.5	58.2	57.8
ACoL [53] *(cvpr,2018)*	72.3	59.6	72.7	68.2	54.7	63.0	57.8	58.4
SPG [54] *(eccv,2018)*	63.7	62.8	71.4	66.0	55.9	62.4	57.7	58.6
ADL [9] *(cvpr,2019)*	75.7	63.4	73.5	70.8	58.3	62.1	54.3	58.2
CutMix [51] *(eccv,2019)*	71.9	65.5	67.8	68.4	58.2	61.7	58.7	59.5
Best WSOL	76.3	65.5	78.1	70.8	58.3	63.0	58.7	59.5
FSL baseline	86.3	94.0	95.8	92.0	61.5	70.3	74.4	68.7
Center baseline	59.7	59.7	59.7	59.7	45.8	45.8	45.8	45.8
CSTN [22] *(icpr,2020)*	Resnet101 [14]: 76.0				–	–	–	–
TS-CAM [13] *(corr,2021)*	Deit-S [39]: 83.8				–	–	–	–
MEIL [21] *(cvpr,2020)*	73.8	–	–	–	–	–	–	–
DANet [47] *(iccv,2019)*	67.7	67.03	–	–	–	–	–	–
SPOL [44] *(cvpr,2021)*	–	–	96.4	–	–	–	–	–
CAM* [57] *(cvpr,2016)*	61.6	58.8	71.5	63.9	53.0	62.7	56.8	57.5
GradCAM [32] *(iccv,2017)*	69.3	62.3	73.1	68.2	59.6	63.9	60.1	61.2
GradCAM++ [7] *(wacv,2018)*	84.1	63.3	81.9	76.4	60.5	64.0	60.2	61.5
Smooth-GradCAM++ [25] *(corr,2019)*	69.7	66.9	76.3	70.9	52.2	61.7	54.3	56.0
XGradCAM [12] *(bmvc,2020)*	69.3	60.9	72.7	67.6	59.0	63.9	60.2	61.0
LayerCAM [15] *(ieee,2021)*	84.3	66.5	85.2	78.6	59.5	63.5	61.1	61.3
CAM* [57] + ours	87.3	82.0	90.3	86.5	67.8	71.9	72.1	70.6
GradCAM [32] + ours	87.5	84.4	90.5	87.4	68.6	70.0	70.9	69.8
GradCAM++ [57] + ours	91.5	84.6	91.0	89.0	64.8	67.1	66.3	66.0
Smooth-GradCAM++ [57] + ours	89.1	86.8	90.7	88.8	60.3	65.4	64.4	63.3
XGradCAM [57] + ours	86.8	84.4	90.4	88.8	68.7	71.3	70.4	70.1
LayerCAM [57] + ours	91.0	85.3	92.4	89.7	64.3	64.9	65.3	64.8
Best WSOL + ours	91.5	86.8	92.4	89.7	68.7	71.9	72.1	70.6

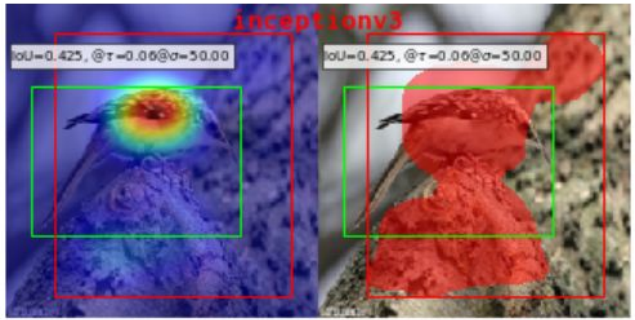
Table 1: Performance on MaxBoxAcc and PxAP metrics.

top1/5, classification accuracy are in the paper

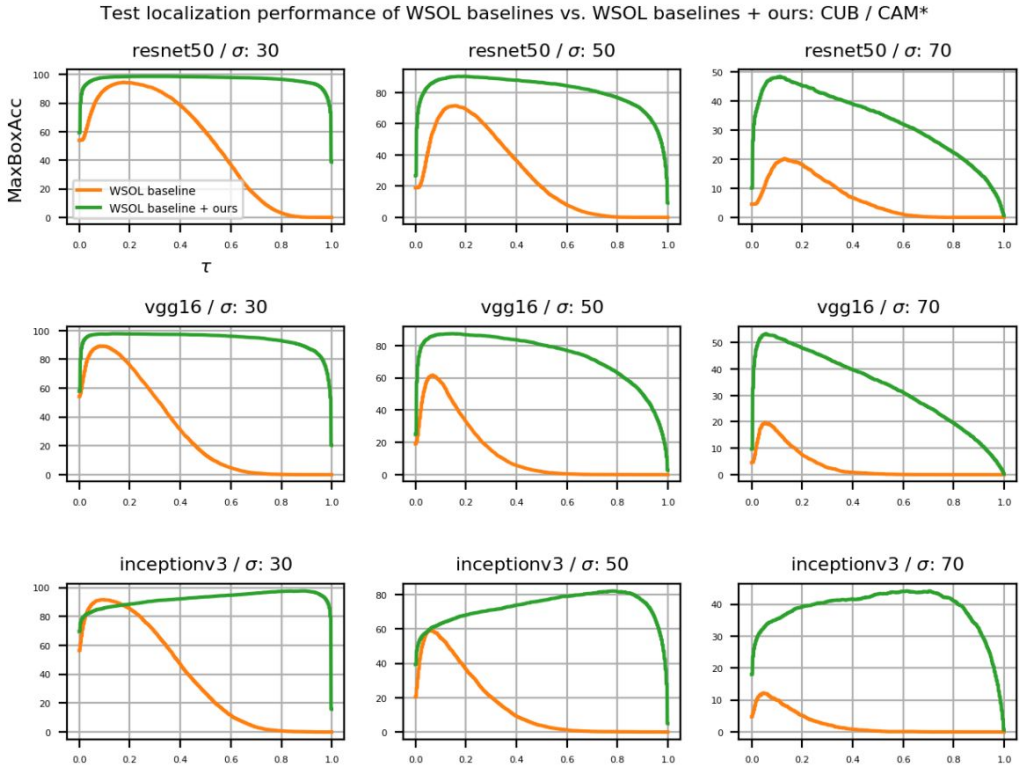
5. Experiments:
Sensitivity to threshold



CAM + ours

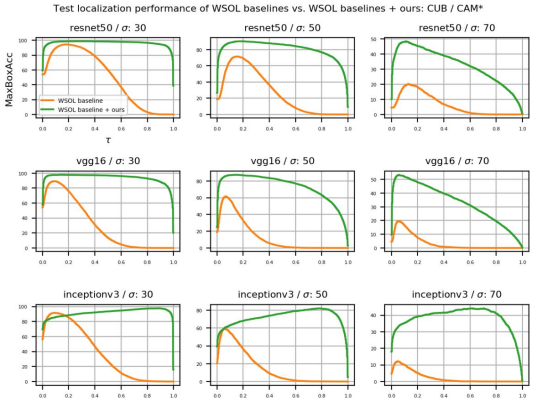


CAM

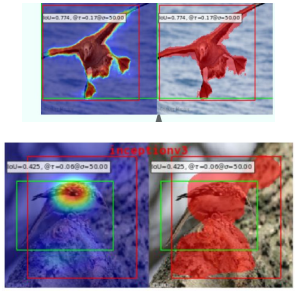


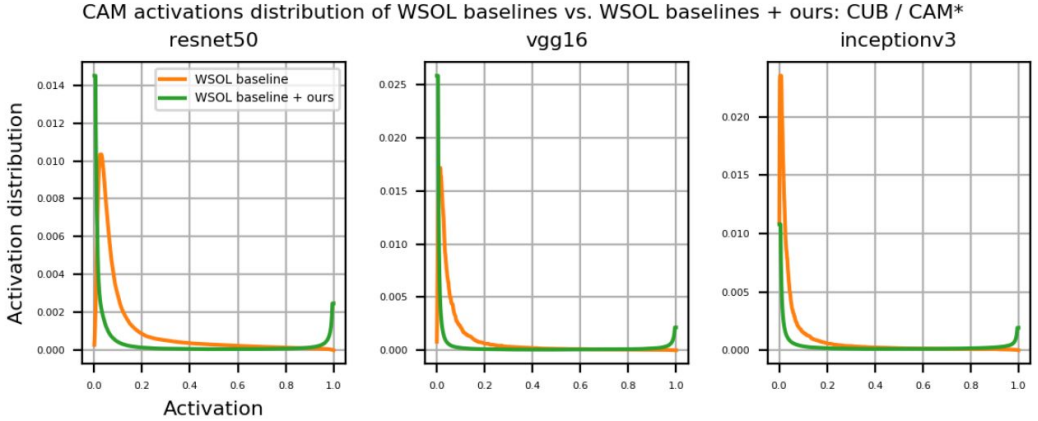
(a) MaxBoxAcc metric.

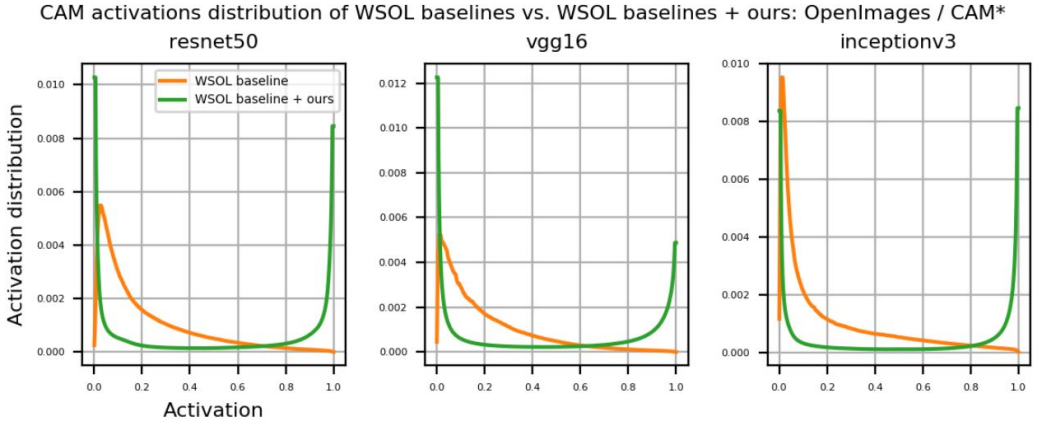
3. Experiments:
Sensitivity to threshold



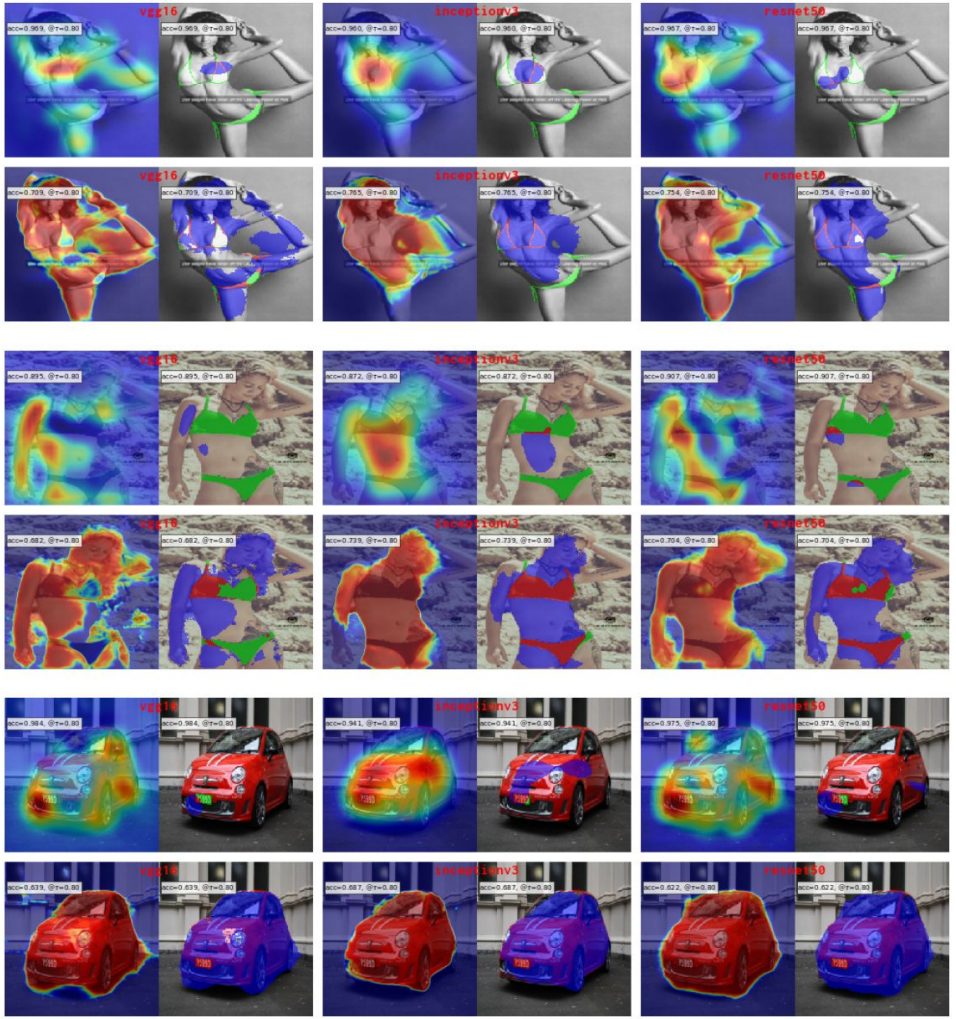
(a) `MaxBoxAcc` metric.







3. Experiments:
Failed cases



3. Experiments:
Runtime (inference)

Backbones (encoders)	VGG16				Inception				ResNet50			
Methods	#PCL	#NFM	SFM	#PDEC	#PCL	#NFM	SFM	#PDEC	#PCL	#NFM	SFM	#PDEC
Details	≈19.6	1024	28x28	≈23.1	≈25.6	1024	28x28	≈5.7	≈23.9	2048	28x28	≈9
CAM* [58]	.2ms				.2ms				.3ms			
GradCAM [32]	7.7ms				21.1ms				27.8ms			
GradCAM++ [7]	23.5ms				23.7ms				28.0ms			
Smooth-GradCAM [25]	62.0ms				150.7ms				136.2ms			
XGradCAM [12]	2.9ms				19.2ms				14.2ms			
LayerCAM [15]	3.2ms				18.2ms				17.9ms			
Mean	16.6ms				38.8ms				37.4ms			
ours + STDCL	6.2ms				25.5ms				18.5ms			
ACoL [55]	12.0ms				19.2ms				24.9ms			
SPG [56]	11.0ms				18ms				23.9ms			
ADL [9]	6.4ms				16.0				14.4ms			
ScoreCAM [44]	1.9sec				3.4sec				9.3sec			
SSCAM [24]	1min45sec				2min16sec				5min49sec			
IS-CAM [23]	30.1sec				39.0sec				1min39sec			

Thanks! Questions?

Please visit us at #1177

Code: https://github.com/sbelharbi/fcam-wsol