Fine-tuning for one-look regression vehicle counting in low-shot aerial datasets

Aneesh Rangnekar[1], Yi Yao[2], Matthew Hoffman[1] and Ajay Divakaran[2]

[1]Rochester Institute of Technology, NY

[2]SRI International, Princeton, NJ

This research is based upon work supported in part by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via 2018- 18050400004. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein

Outline

- Project overview
- Existing approaches
- Methods
- Experiments and results

Project Overview

- Regress the count of vehicles present in PUCPR+ and CARPK

- PUCPR+:
  - Training: 100 images
  - Test: 25 images

- CARPK:
  - Training: 989 images
  - Test: 459 images

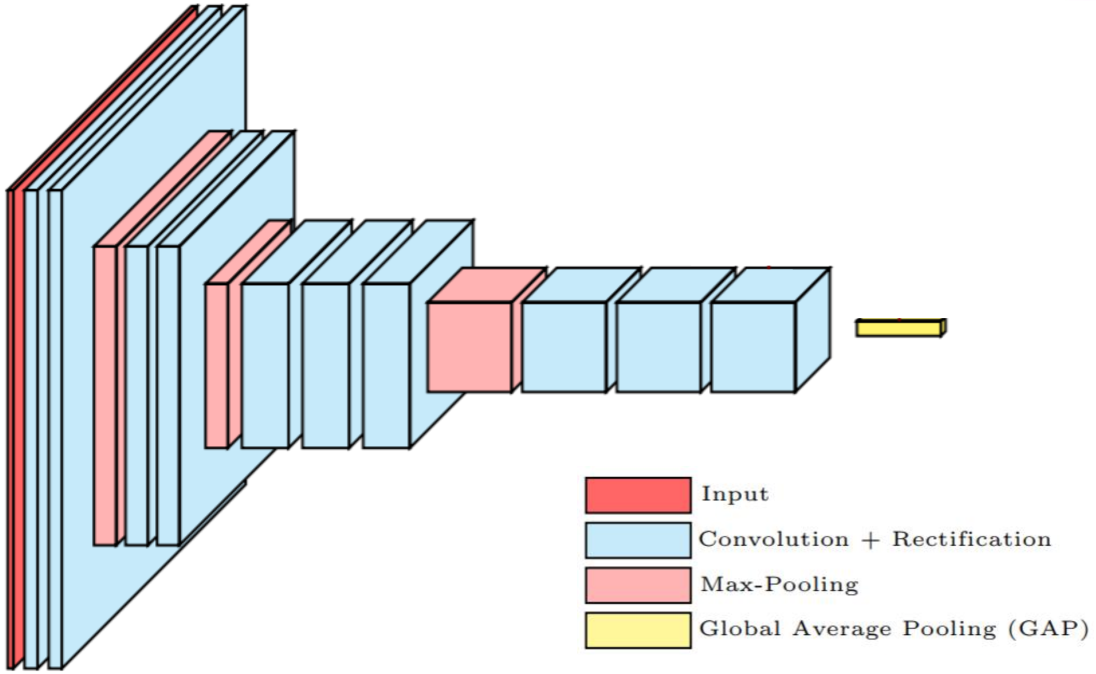
Hsieh, Meng-Ru, Yen-Liang Lin, and Winston H. Hsu. "Drone-based object counting by spatially regularized regional proposal network." Proceedings of the IEEE International Conference on Computer Vision. 2017.

Existing Approaches

- Broadly classified into two categories:
  - Counting by detection
  - Counting by regression

Existing Approaches

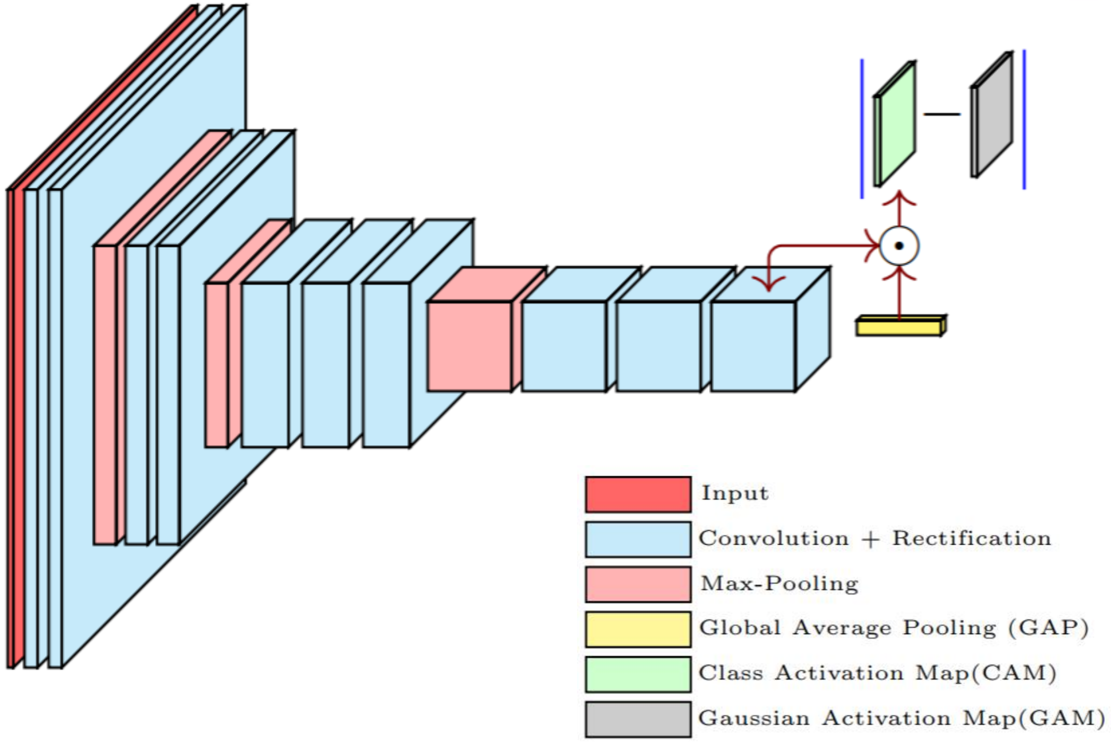
- Broadly classified into two categories:
  - Counting by detection
  - Counting by regression
- We focus on counting by regression:
  - Heatmap Regulation



Aich, Shubhra, and Ian Stavness. "Improving object counting with heatmap regulation." European conference on computer vision. 2018..

Existing Approaches

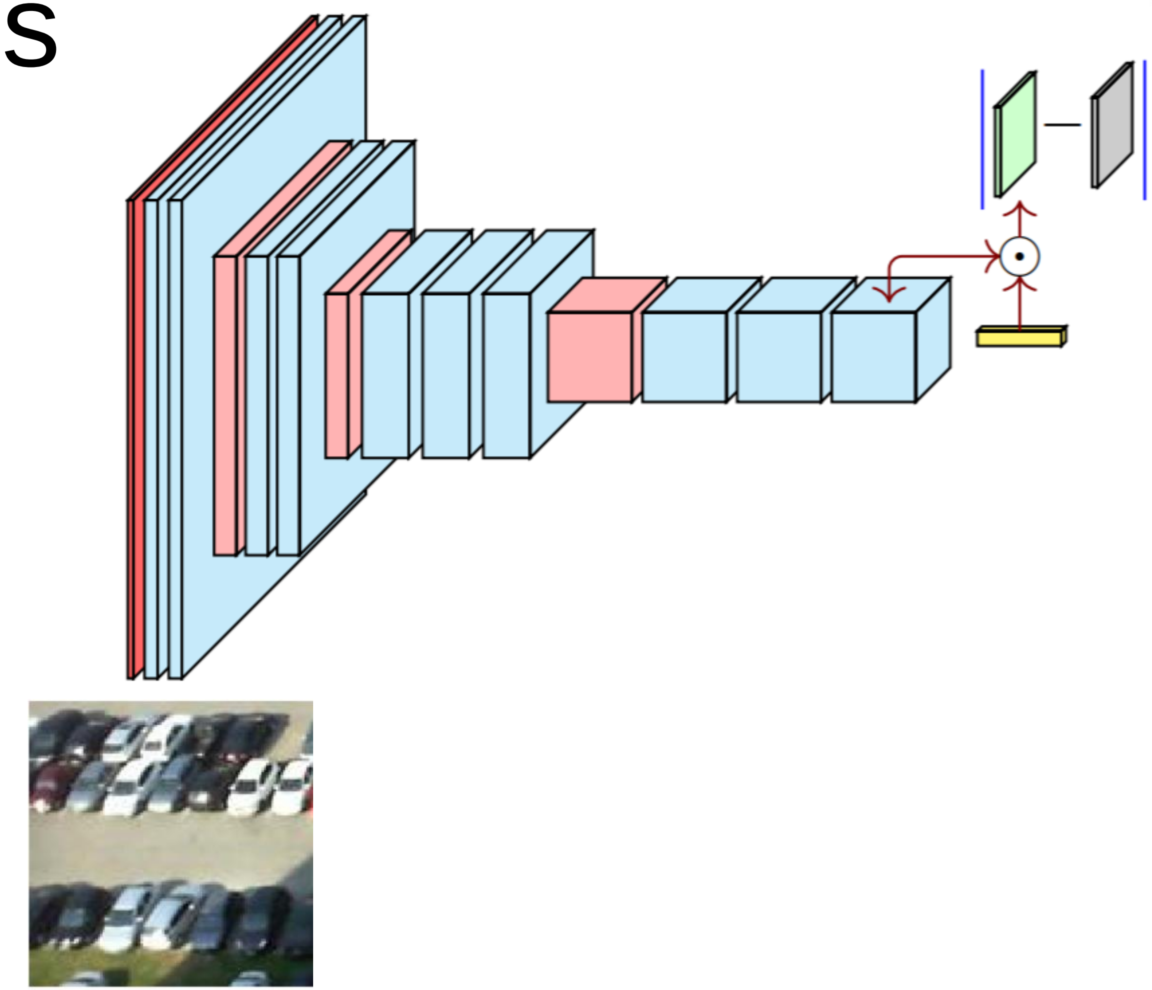
- Broadly classified into two categories:
  - Counting by detection
  - Counting by regression
- We focus on counting by regression:
  - Heatmap Regulation



Aich, Shubhra, and Ian Stavness. "Improving object counting with heatmap regulation." European conference on computer vision. 2018..

Existing Approaches

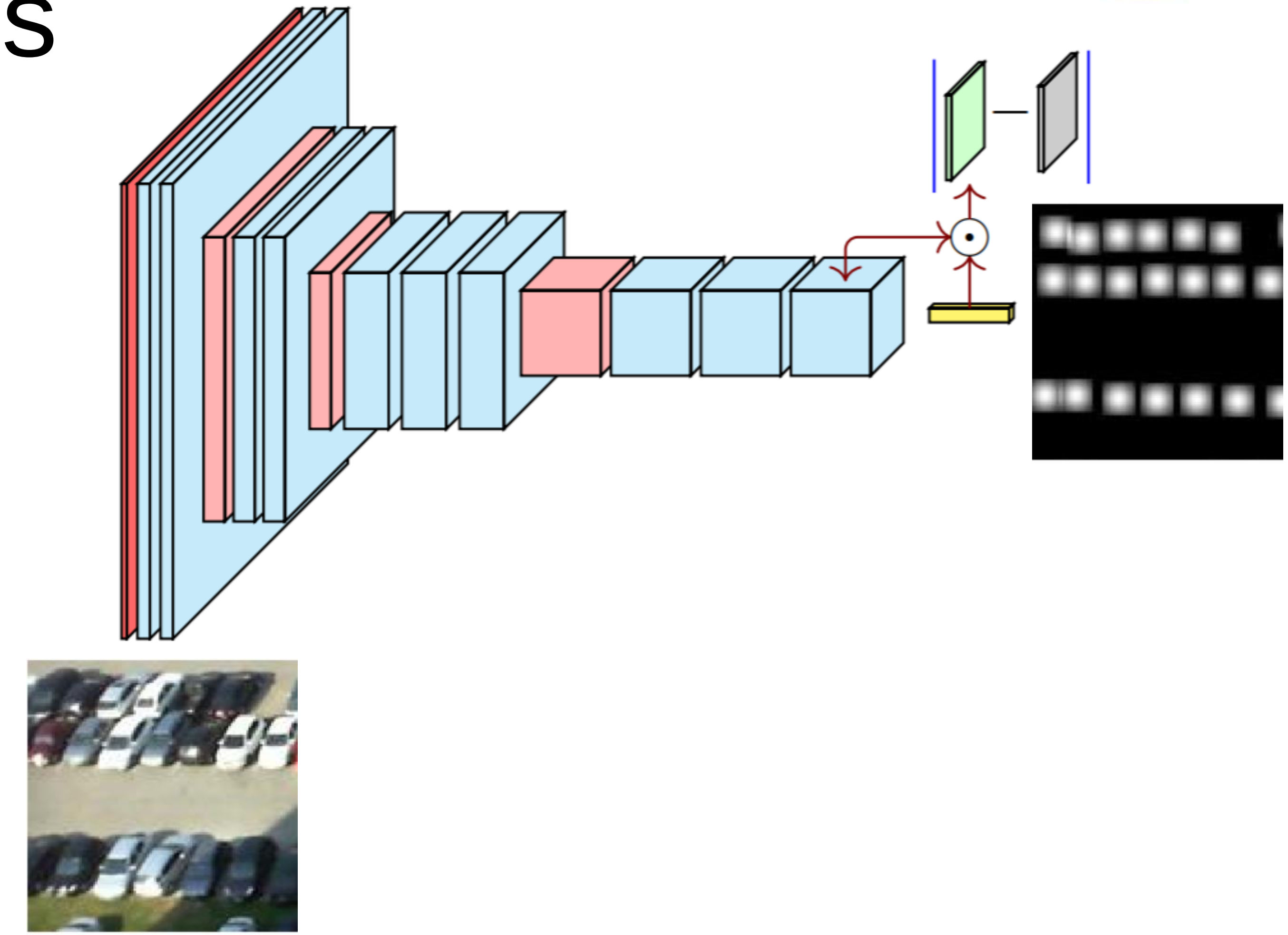
- Broadly classified into two categories:
  - Counting by detection
  - Counting by regression
- We focus on counting by regression:
  - Heatmap Regulation



Aich, Shubhra, and Ian Stavness. "Improving object counting with heatmap regulation." European conference on computer vision. 2018..

Existing Approaches

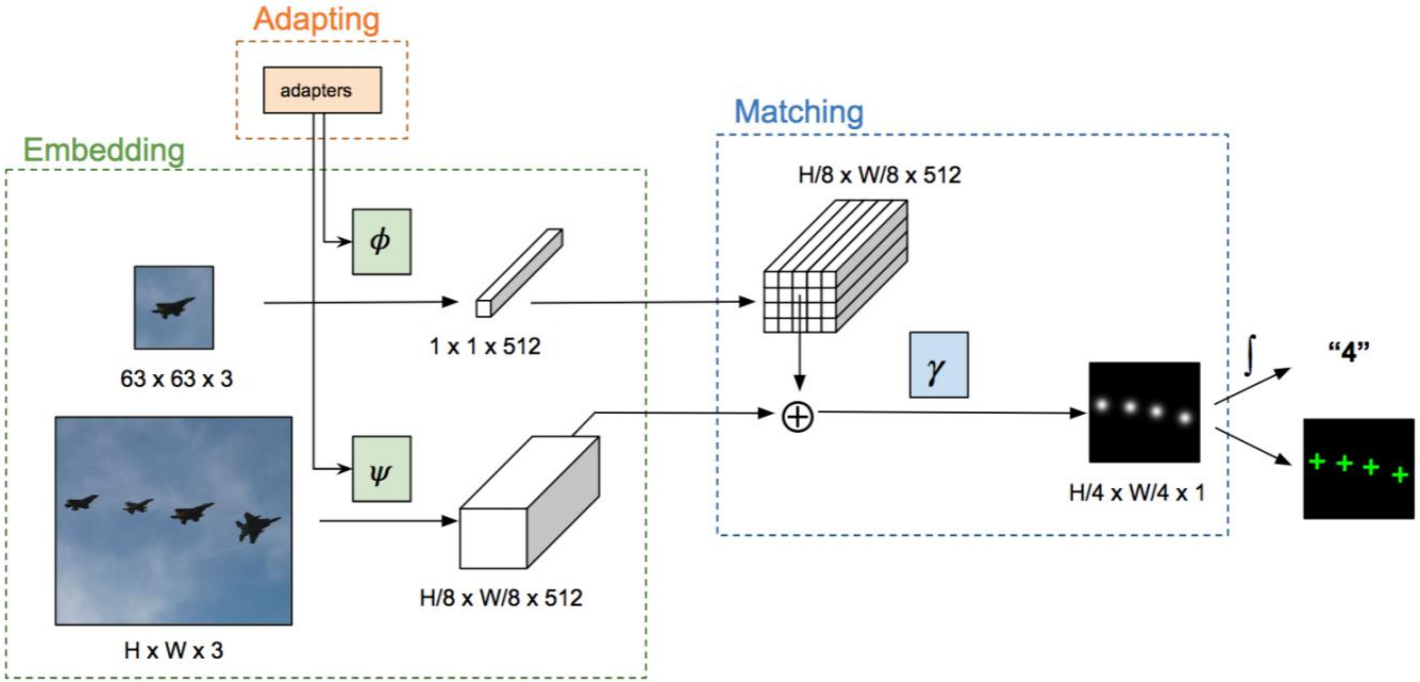
- Broadly classified into two categories:
  - Counting by detection
  - Counting by regression
- We focus on counting by regression:
  - Heatmap Regulation



Aich, Shubhra, and Ian Stavness. "Improving object counting with heatmap regulation." European conference on computer vision. 2018..

Existing Approaches

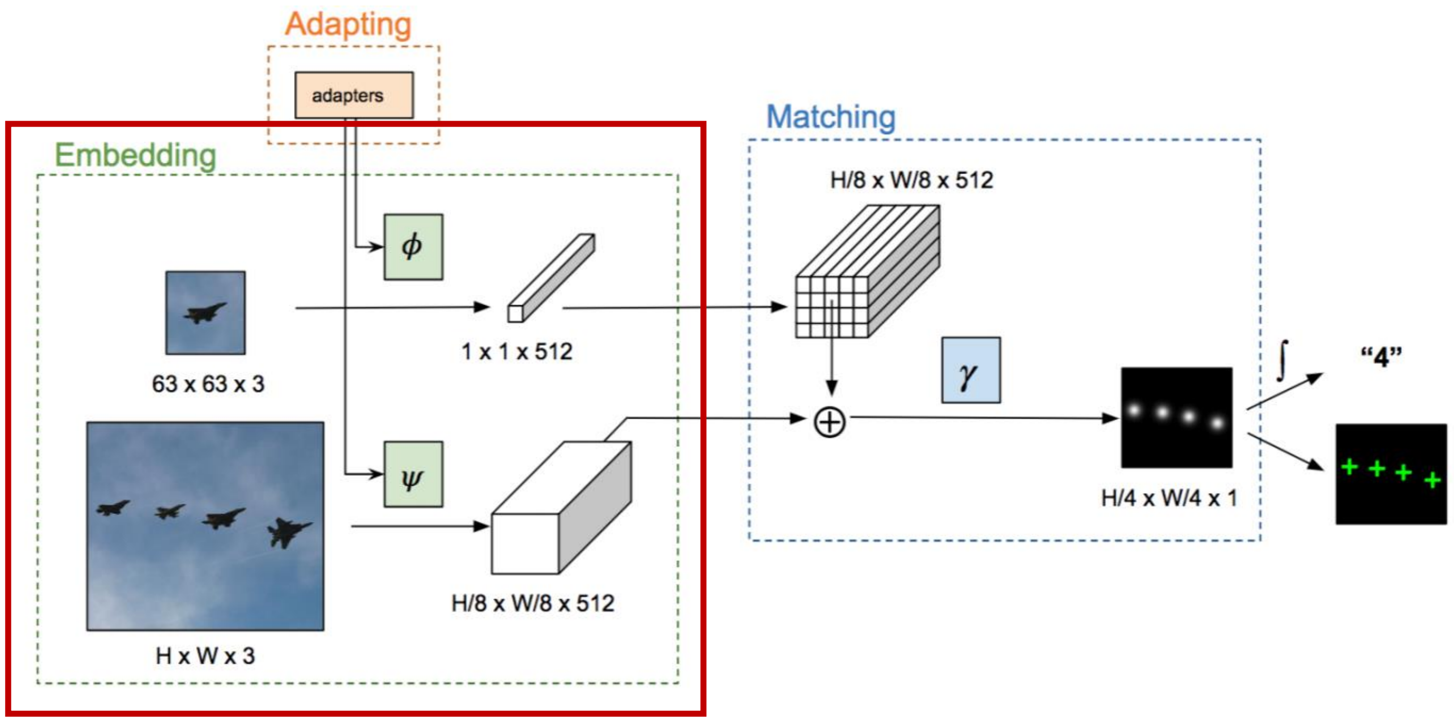
- Broadly classified into two categories:
  - Counting by detection
  - Counting by regression
- We focus on counting by regression:
  - Heatmap Regulation
  - Class Agnostic Counting



Lu, Erika, Weidi Xie, and Andrew Zisserman. "Class-agnostic counting." Asian conference on computer vision. Springer, Cham, 2018..

Existing Approaches

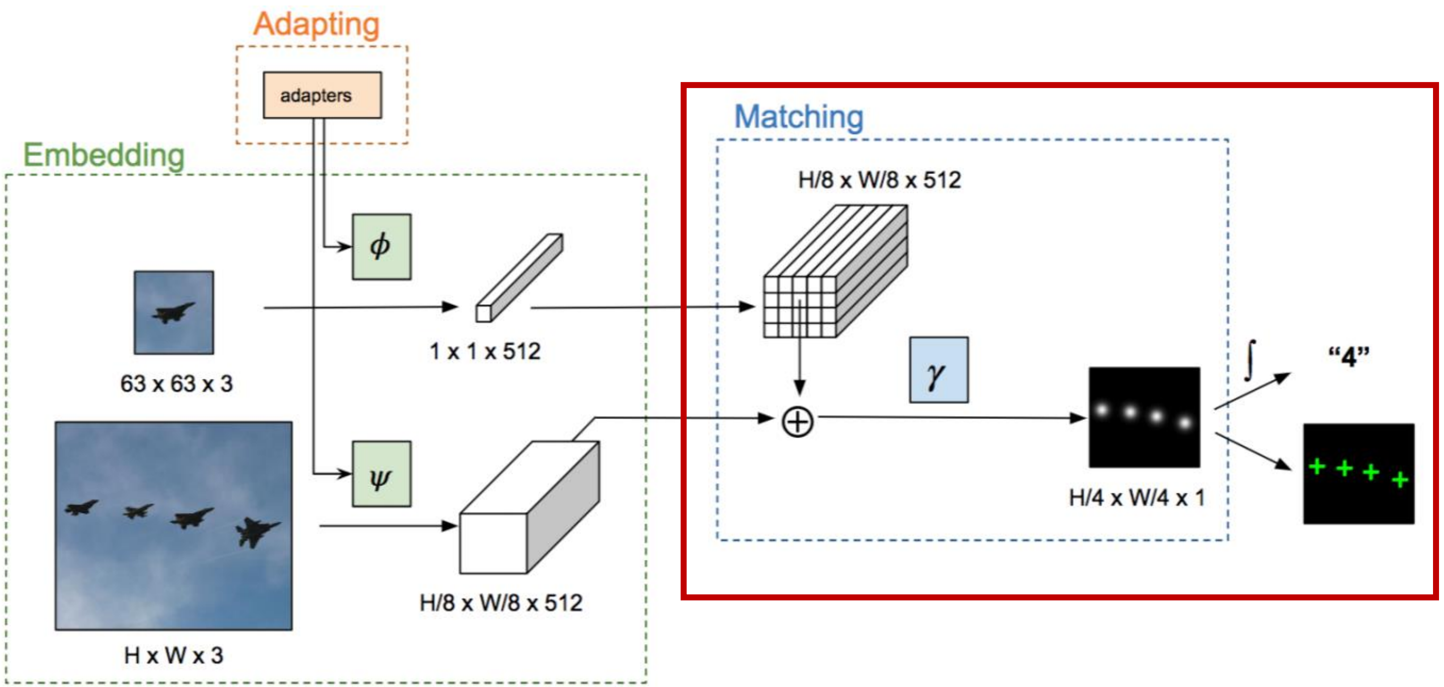
- Broadly classified into two categories:
  - Counting by detection
  - Counting by regression
- We focus on counting by regression:
  - Heatmap Regulation
  - Class Agnostic Counting



Lu, Erika, Weidi Xie, and Andrew Zisserman. "Class-agnostic counting." Asian conference on computer vision. Springer, Cham, 2018..

Existing Approaches

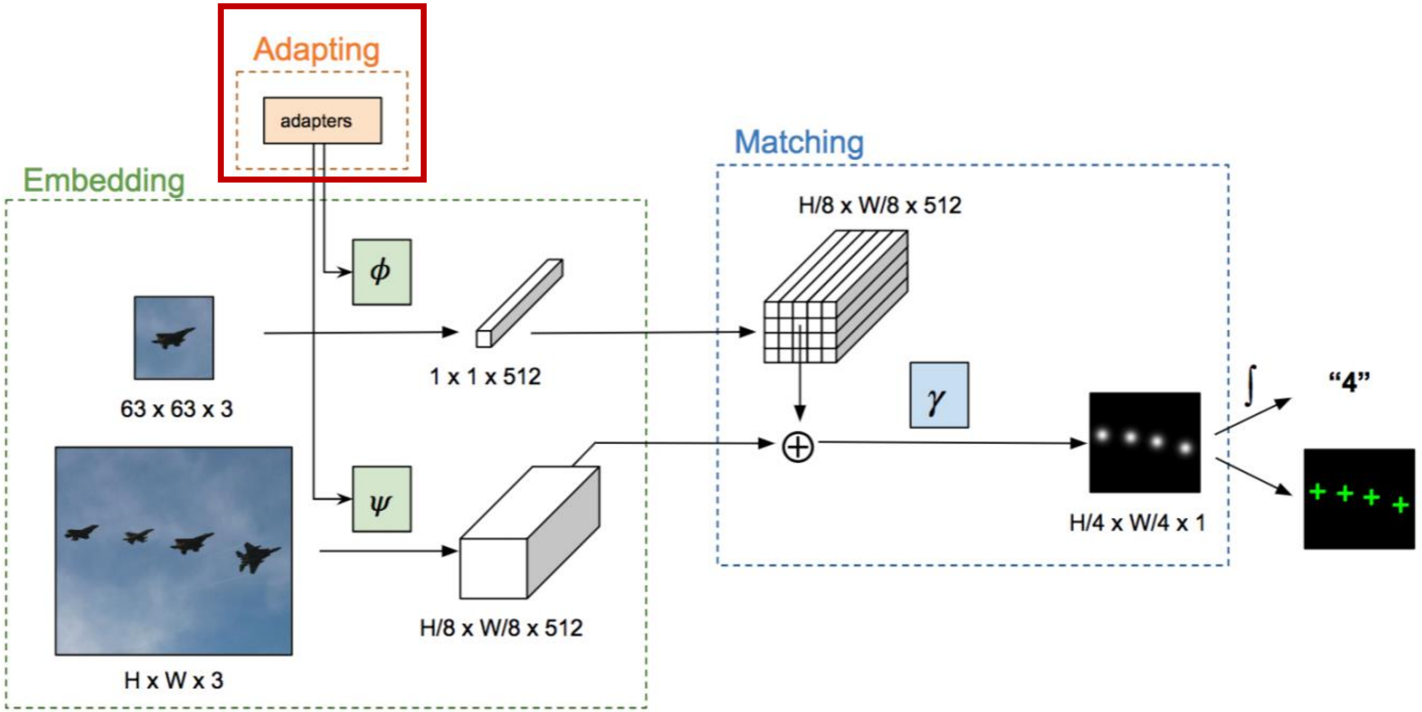
- Broadly classified into two categories:
  - Counting by detection
  - Counting by regression
- We focus on counting by regression:
  - Heatmap Regulation
  - Class Agnostic Counting



Lu, Erika, Weidi Xie, and Andrew Zisserman. "Class-agnostic counting." Asian conference on computer vision. Springer, Cham, 2018..

Existing Approaches

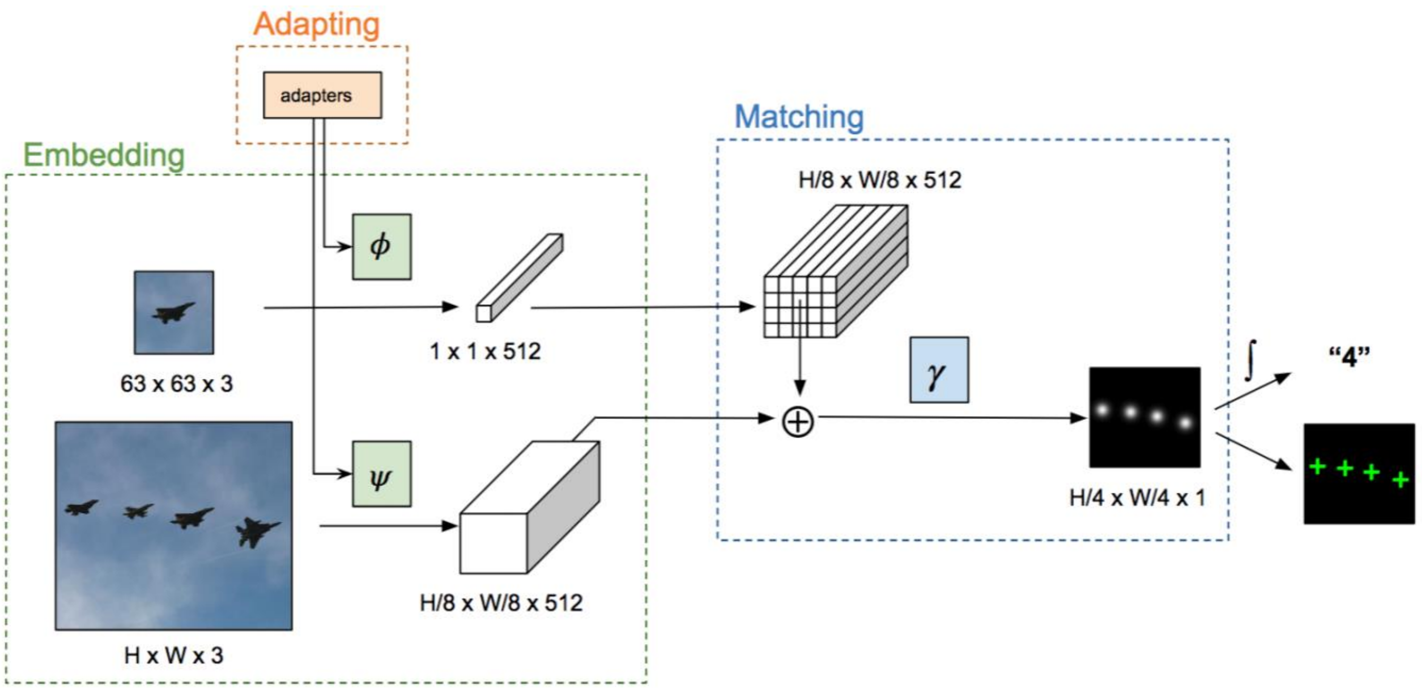
- Broadly classified into two categories:
  - Counting by detection
  - Counting by regression
- We focus on counting by regression:
  - Heatmap Regulation
  - Class Agnostic Counting



Lu, Erika, Weidi Xie, and Andrew Zisserman. "Class-agnostic counting." Asian conference on computer vision. Springer, Cham, 2018..

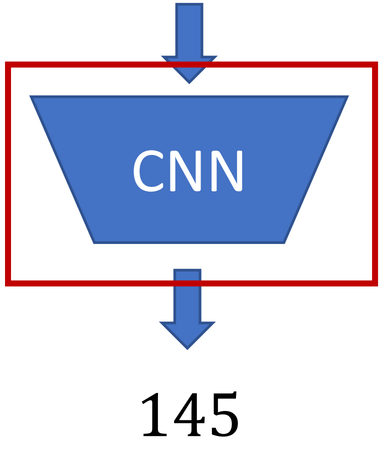
Existing Approaches

- Broadly classified into two categories:
  - Counting by detection
  - Counting by regression
- We focus on counting by regression:
  - Heatmap Regulation
  - Class Agnostic Counting
- However, both approaches rely on using localization information



Lu, Erika, Weidi Xie, and Andrew Zisserman. "Class-agnostic counting." Asian conference on computer vision. Springer, Cham, 2018..

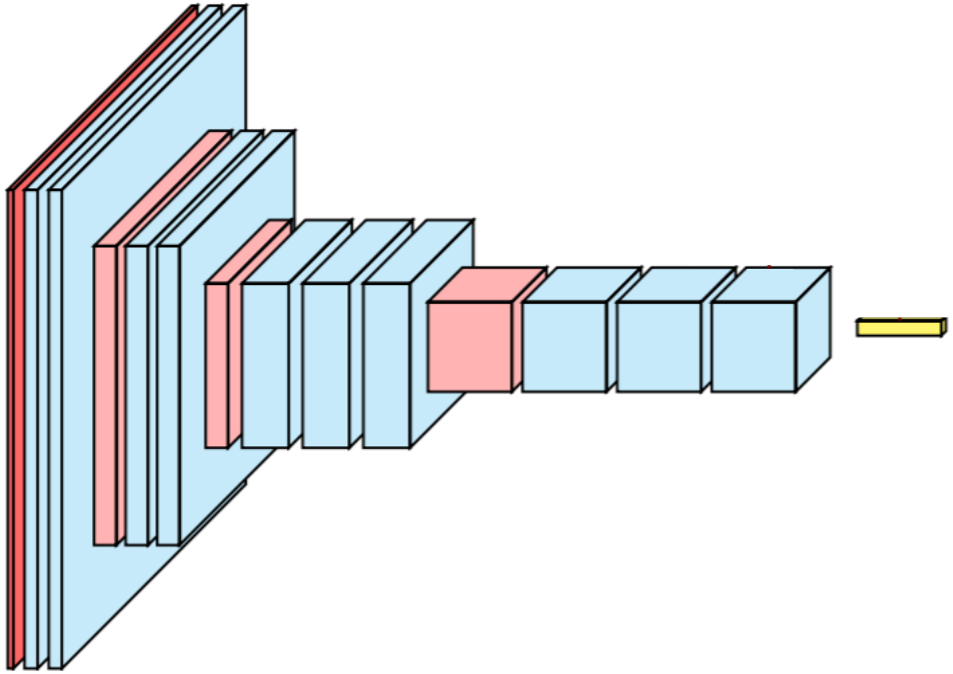
Methods



Methods

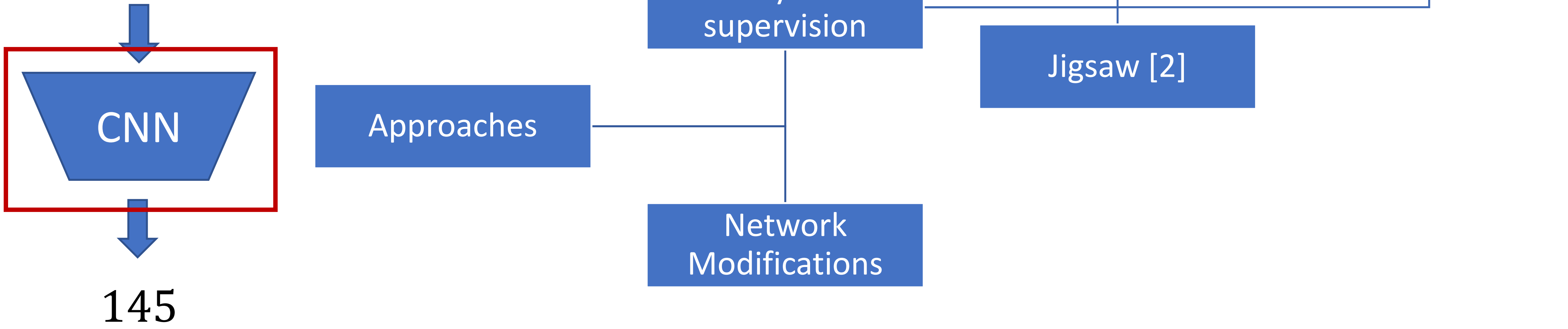
CNN

145



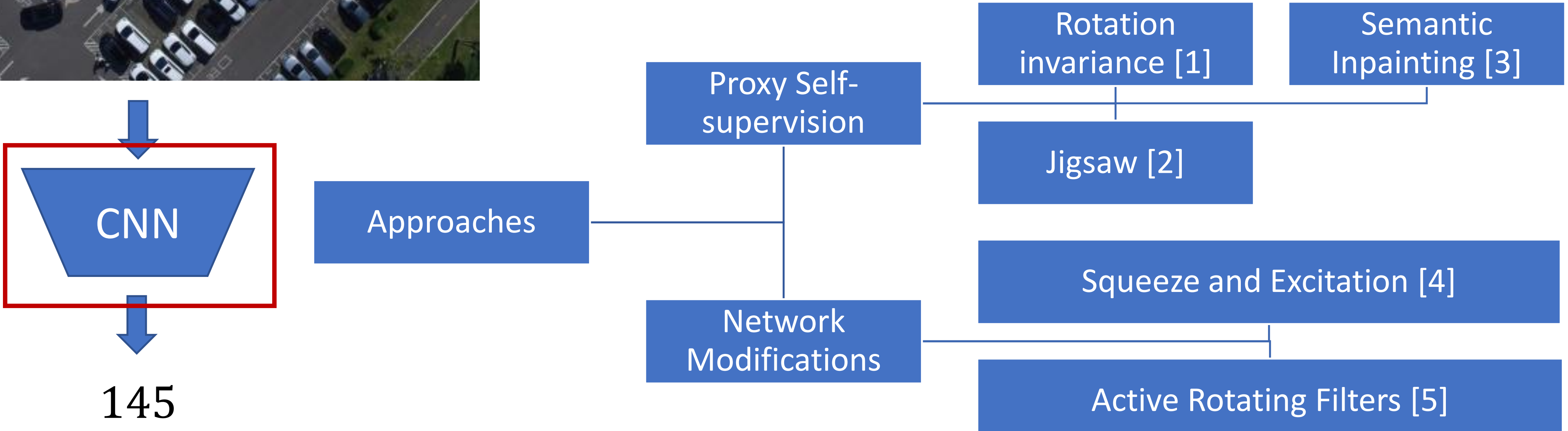
Aich, Shubhra, and Ian Stavness. "Improving object counting with heatmap regulation." European conference on computer vision. 2018..

Methods


CNN
145
Approaches
Proxy Self-supervision
Network Modifications
Rotation invariance [1]
Jigsaw [2]
Semantic Inpainting [3]

RIT

Methods


CNN
145
Approaches
Proxy Self-supervision
Rotation invariance [1]
Semantic Inpainting [3]
Jigsaw [2]
Network Modifications
Squeeze and Excitation [4]
Active Rotating Filters [5]

Methods

- Proxy self-supervision tasks

Methods

- Proxy self-supervision tasks
  - Rotation invariance

0|90|180|270?

Methods

- Proxy self-supervision tasks
  - Rotation invariance

0|90|180|270?

RIT

Methods

- Proxy self-supervision tasks
  - Rotation invariance

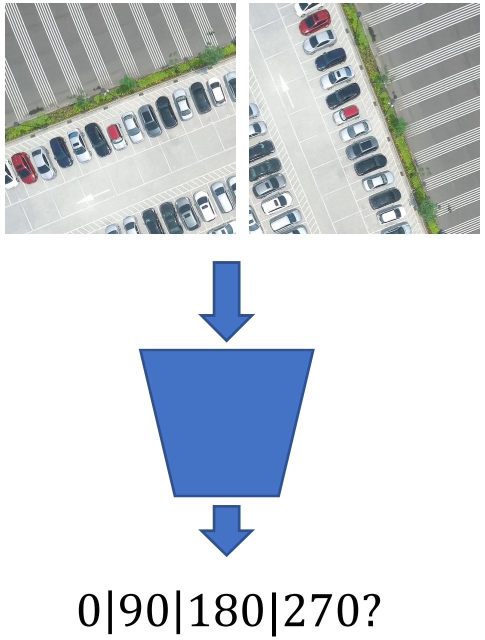
0|90|180|270?

Methods

- Proxy self-supervision tasks
  - Rotation invariance

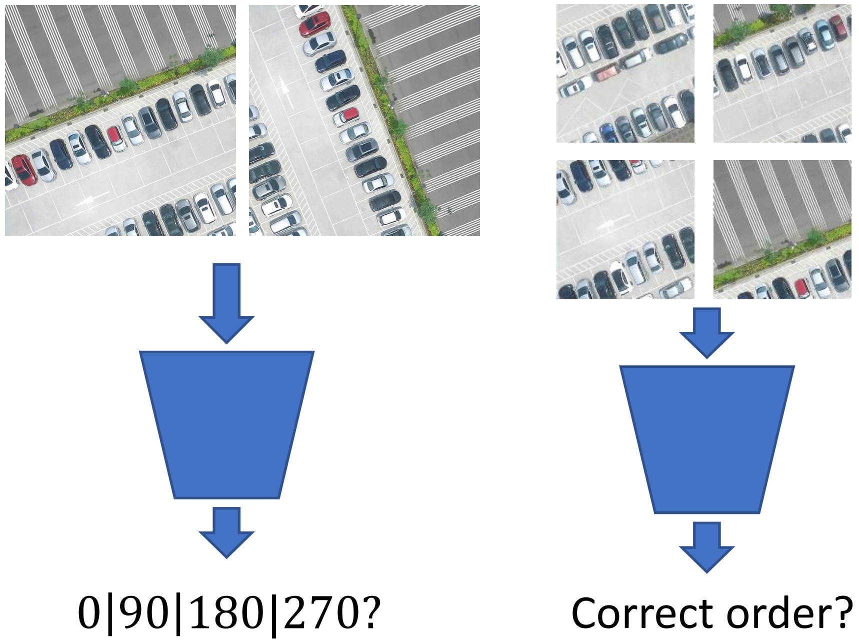
Methods

- Proxy self-supervision tasks
  - Rotation invariance



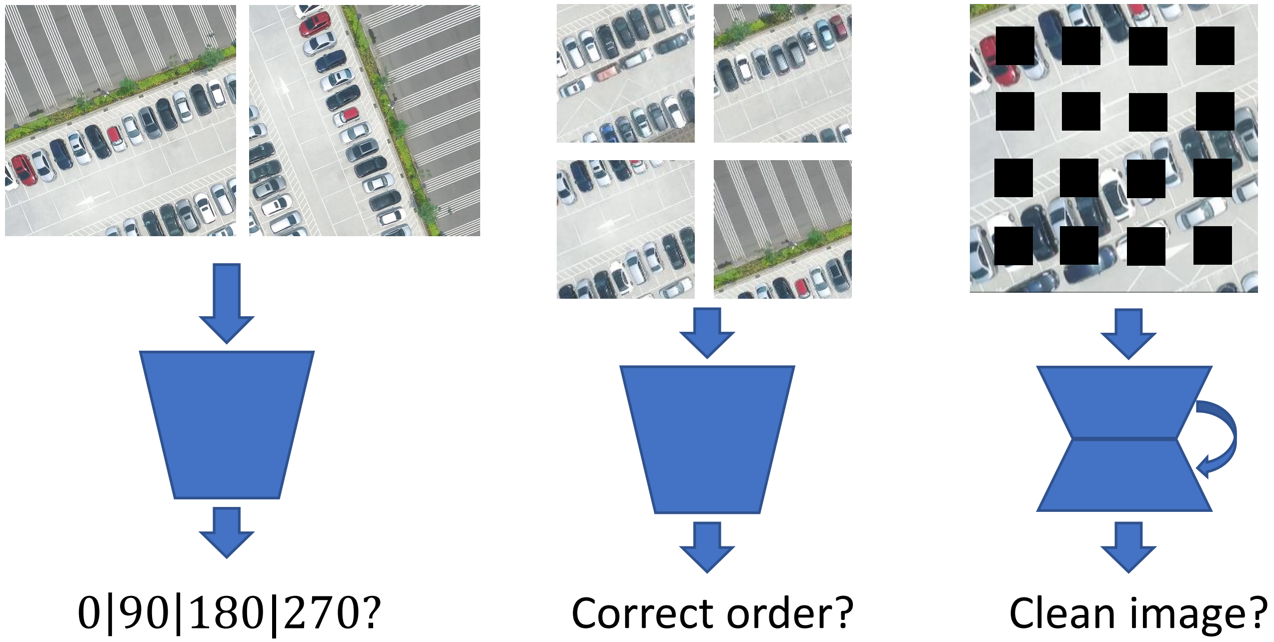
Methods

- Proxy self-supervision tasks
  - Rotation invariance
  - Jigsaw



Methods

- Proxy self-supervision tasks
  - Rotation invariance
  - Jigsaw
  - Semantic Inpainting

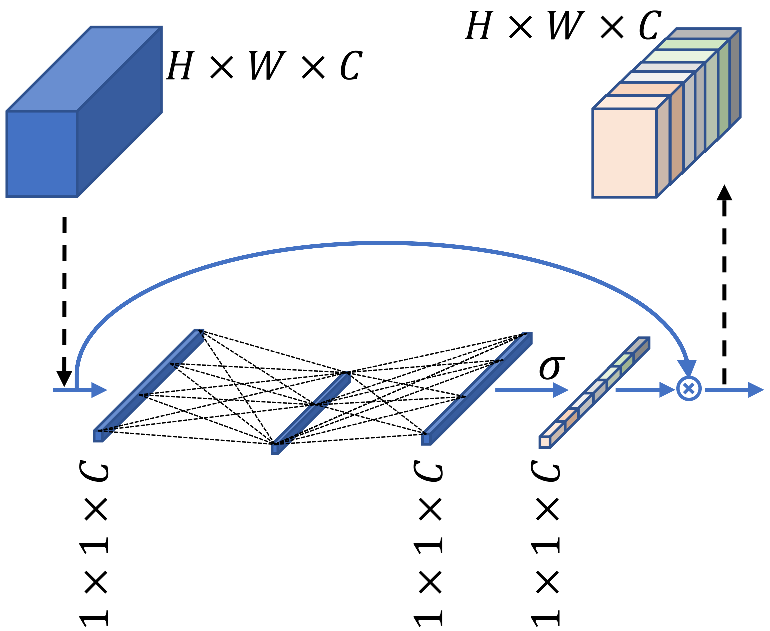


Methods

- Network modifications

Methods

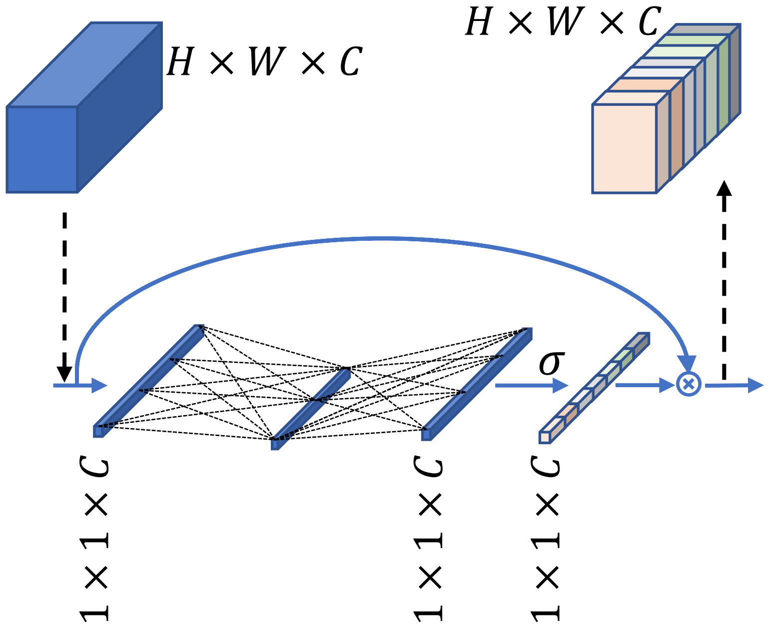
- Network modifications
  - Squeeze and excitation block network

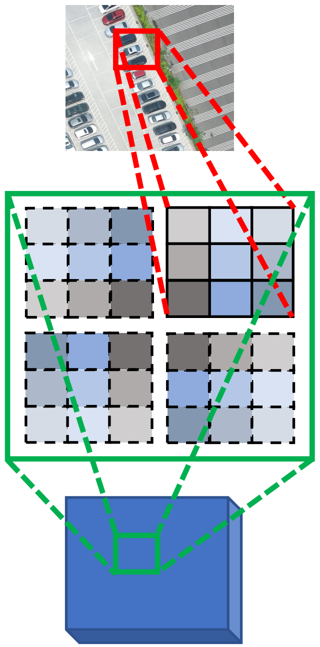


RIT

Methods

- Network modifications
  - Squeeze and excitation block networks
  - Active Rotating Filter





Results

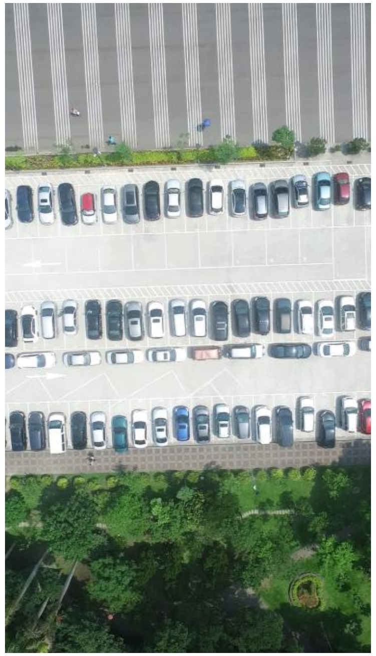
Task/Blocks VGG-16	PUCPR+				CARPK			
	MAE	RMSE	Over-est (%)	Under-est (%)	MAE	RMSE	Over-est (%)	Under-est (%)
Pretrained	5.84	8.51	2.58	1.15	6.88	9.40	0.85	5.80
Rotation	**3.72**	**6.32**	**1.00**	**1.38**	7.81	10.02	2.10	5.45
Jigsaw	4.56	6.07	2.35	0.56	9.05	11.64	3.35	5.40
Inpainting	4.16	6.33	1.96	0.69	6.31	8.25	2.70	3.40
SE Block	5.76	9.50	1.45	2.22	**5.93**	**7.90**	**1.79**	**3.94**
ARF	4.12	5.84	1.89	0.74	12.38	15.83	0.77	11.99

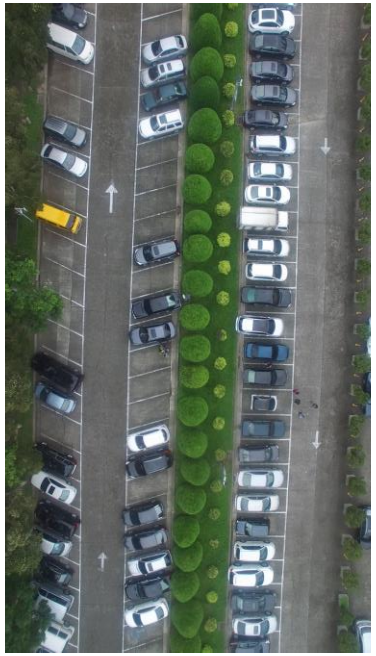
Results

- Transposed Augmentations for CARPK dataset

- Similar to Random Rotation

- Pretrained VGG:
  - Without: 11.5 ± 1.4
  - With: 6.2 ± 0.6

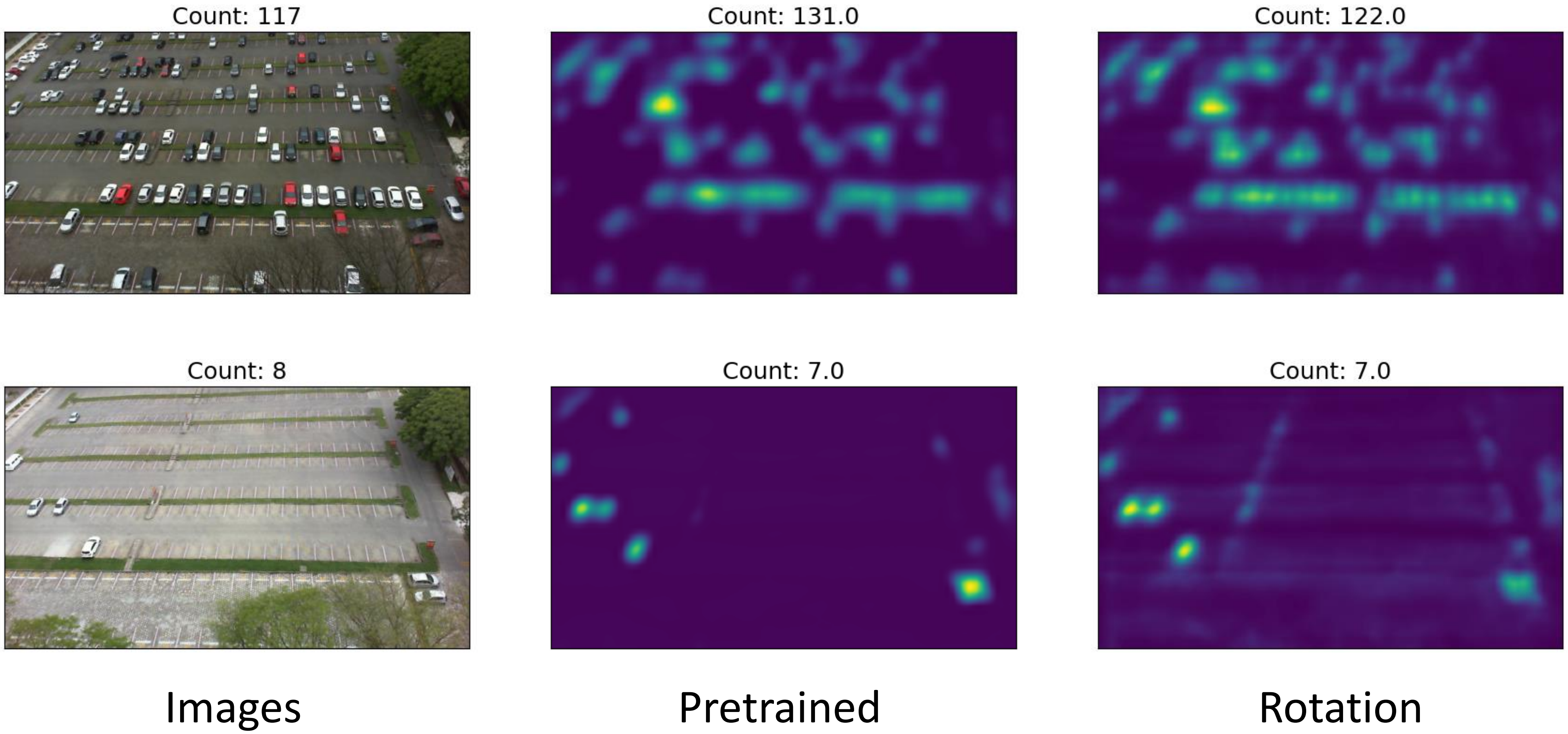




RIT

SRI International

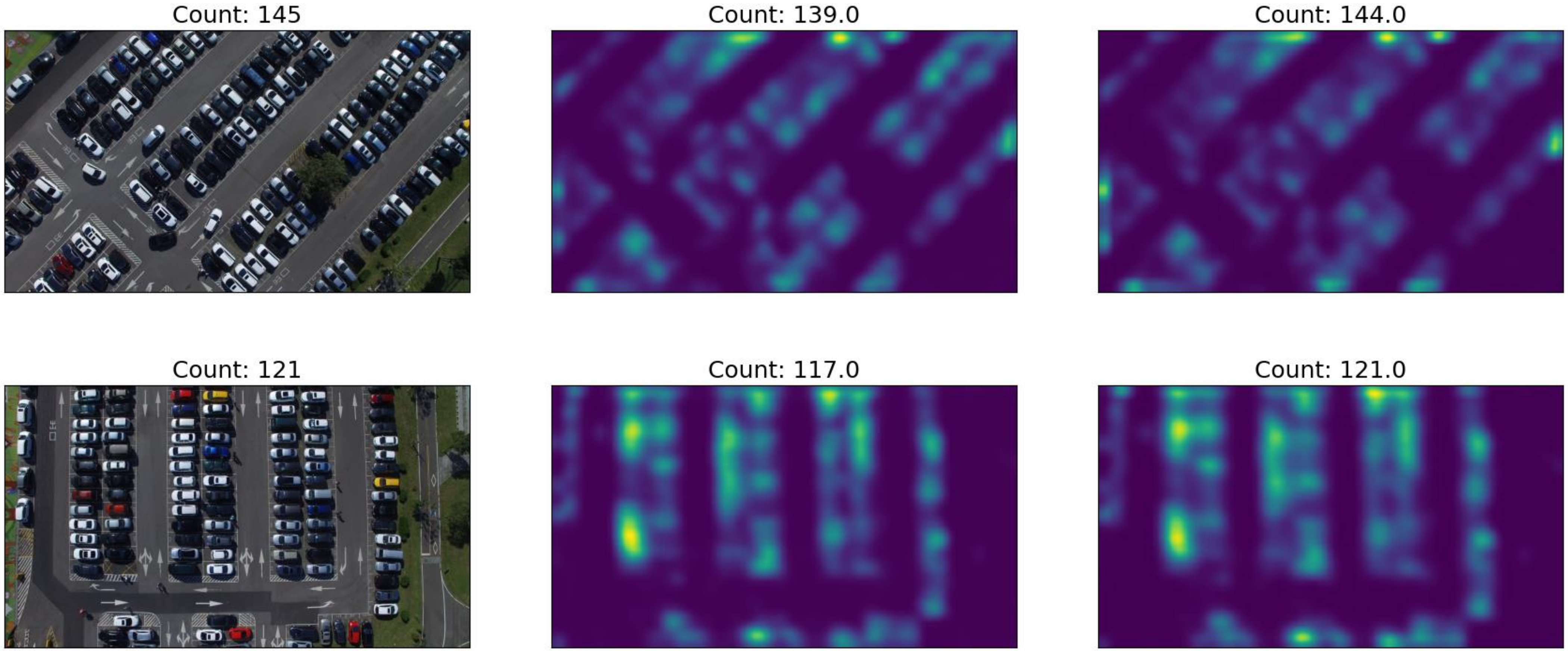
Results PUCPR+



RIT

SRI International

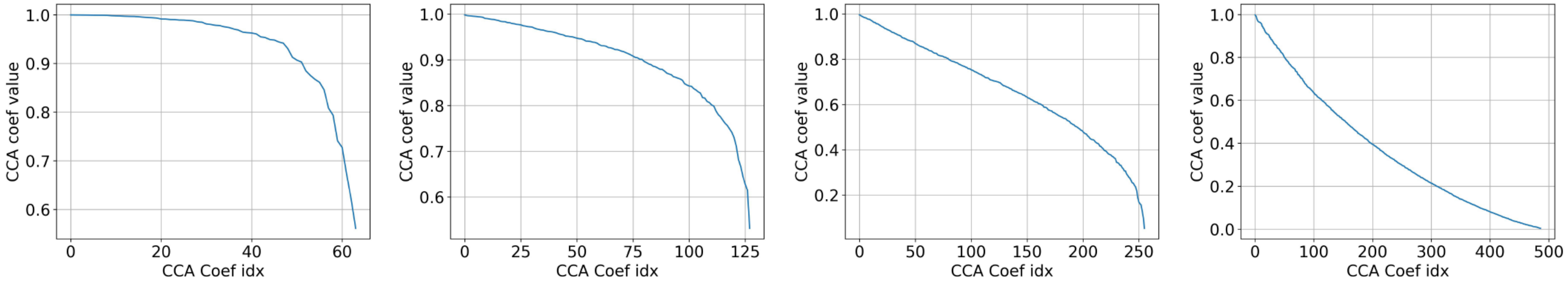
Results CARPK



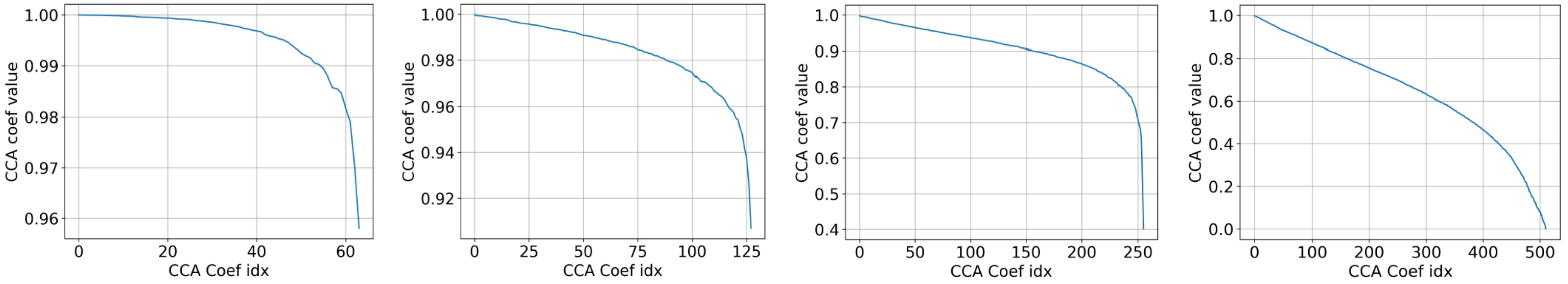
Images Pretrained SE blocks

Results

- SVCCA[6] – PUCPR+



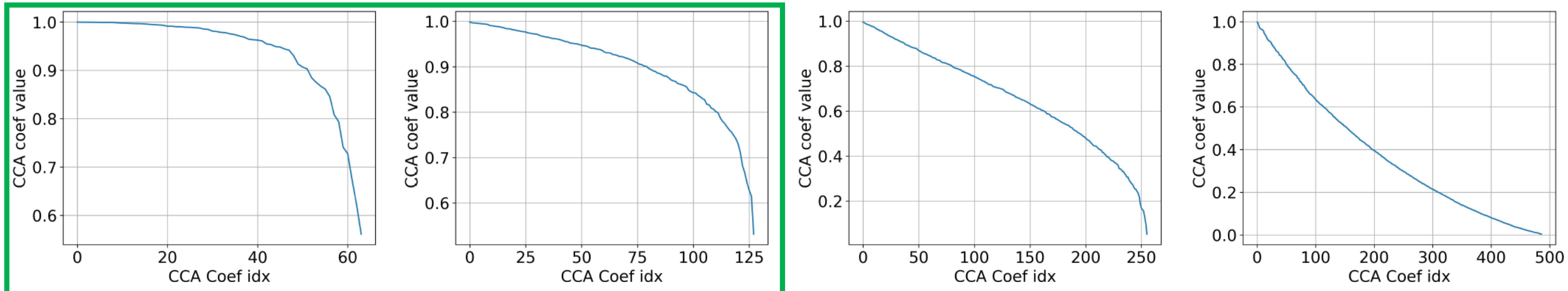
- SVCCA – CARPK



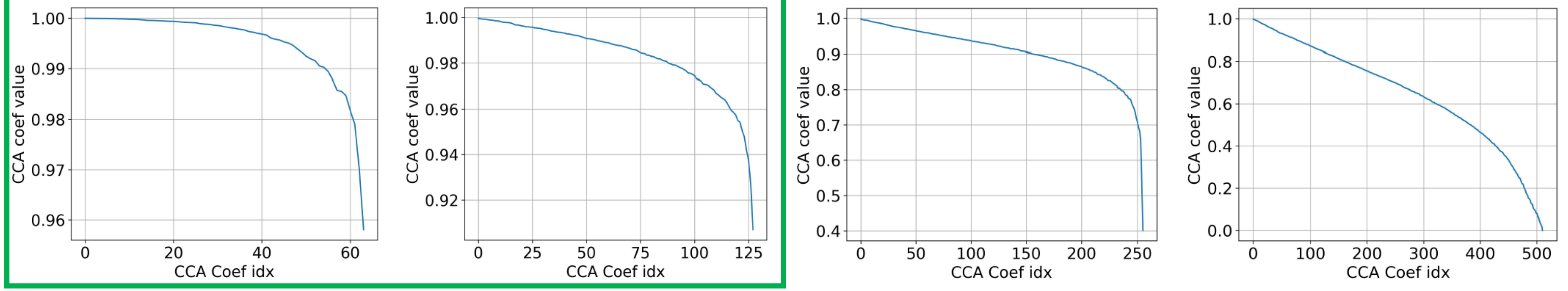
MaxPool1 | MaxPool2 | MaxPool3 | PreGAP

Results

- SVCCA – PUCPR+



- SVCCA – CARPK

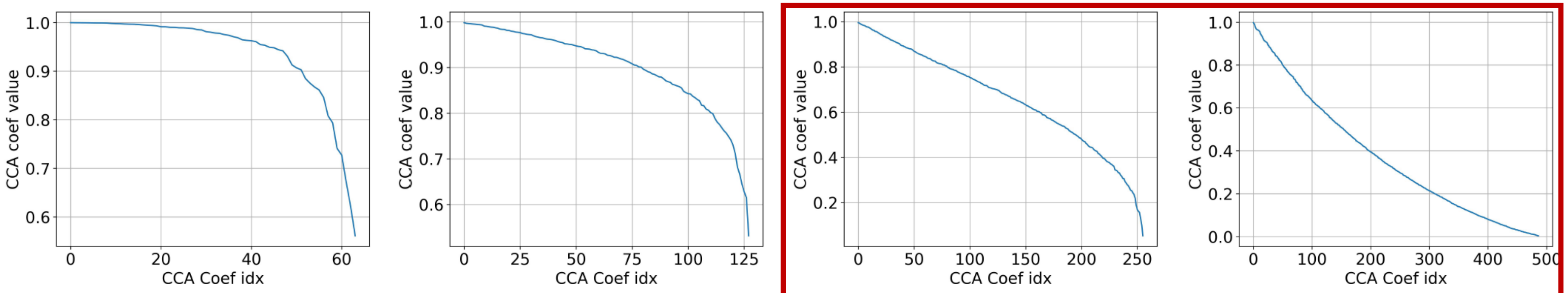


High Corr.

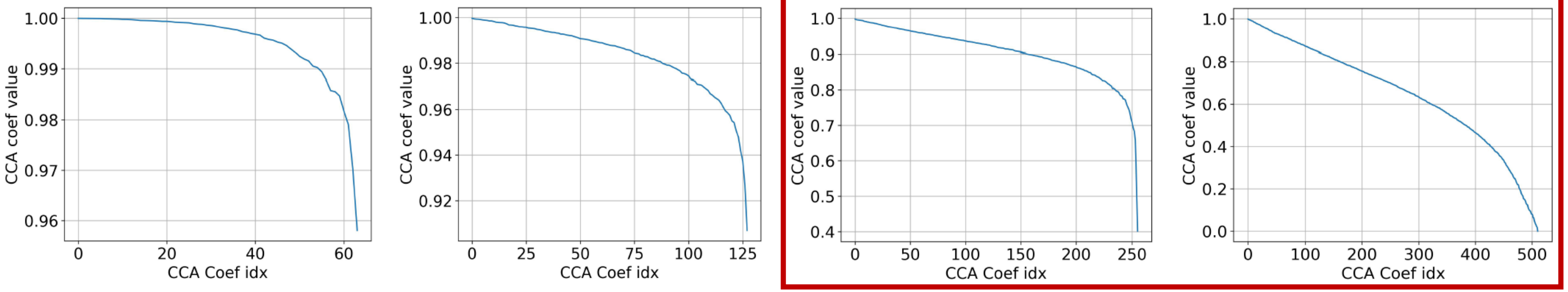
MaxPool1 | MaxPool2 | MaxPool3 | PreGAP

Results

- SVCCA – PUCPR+



- SVCCA – CARPK

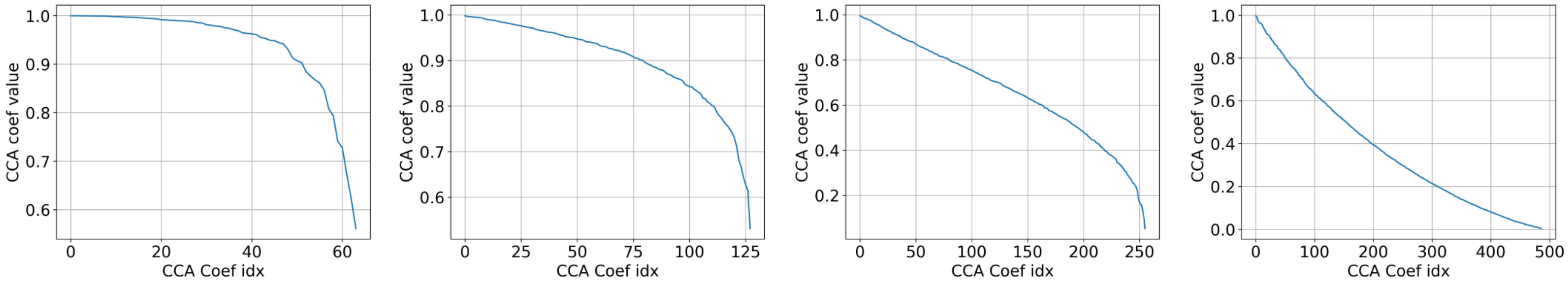


Low Corr.

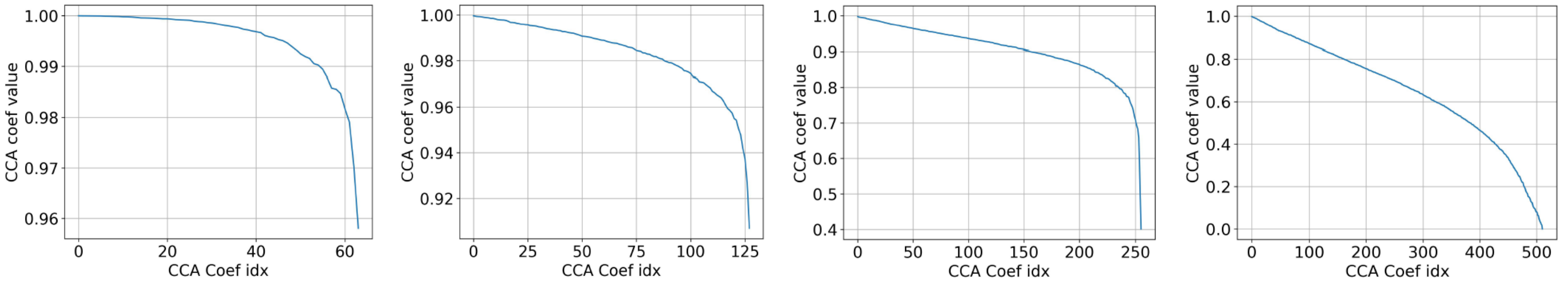
MaxPool1 | MaxPool2 | MaxPool3 | PreGAP

Results

- SVCCA – PUCPR+



- SVCCA – CARPK



MaxPool1 | MaxPool2 | MaxPool3 | PreGAP

Indicates similar early level features and relatively different mid-deep features

References

[1] Gidaris, Spyros, Praveer Singh, and Nikos Komodakis. "Unsupervised Representation Learning by Predicting Image Rotations." *ICLR 2018*. 2018.

[2] Noroozi, Mehdi, and Paolo Favaro. "Unsupervised learning of visual representations by solving jigsaw puzzles." European Conference on Computer Vision. Springer, Cham, 2016.

[3] Singh, Suriya, et al. "Self-Supervised Feature Learning for Semantic Segmentation of Overhead Imagery." BMVC. Vol. 1. No. 2. 2018.

[4] Hu, Jie, Li Shen, and Gang Sun. "Squeeze-and-excitation networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.

[5] Zhou, Yanzhao, et al. "Oriented response networks." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017.

[6] Raghu, Maithra, et al. "Svcca: Singular vector canonical correlation analysis for deep learning dynamics and interpretability." Advances in neural information processing systems 30 (2017): 6076-6085.

Thank you

Questions?