



# Top-Down Saliency Detection Driven by Visual Classification

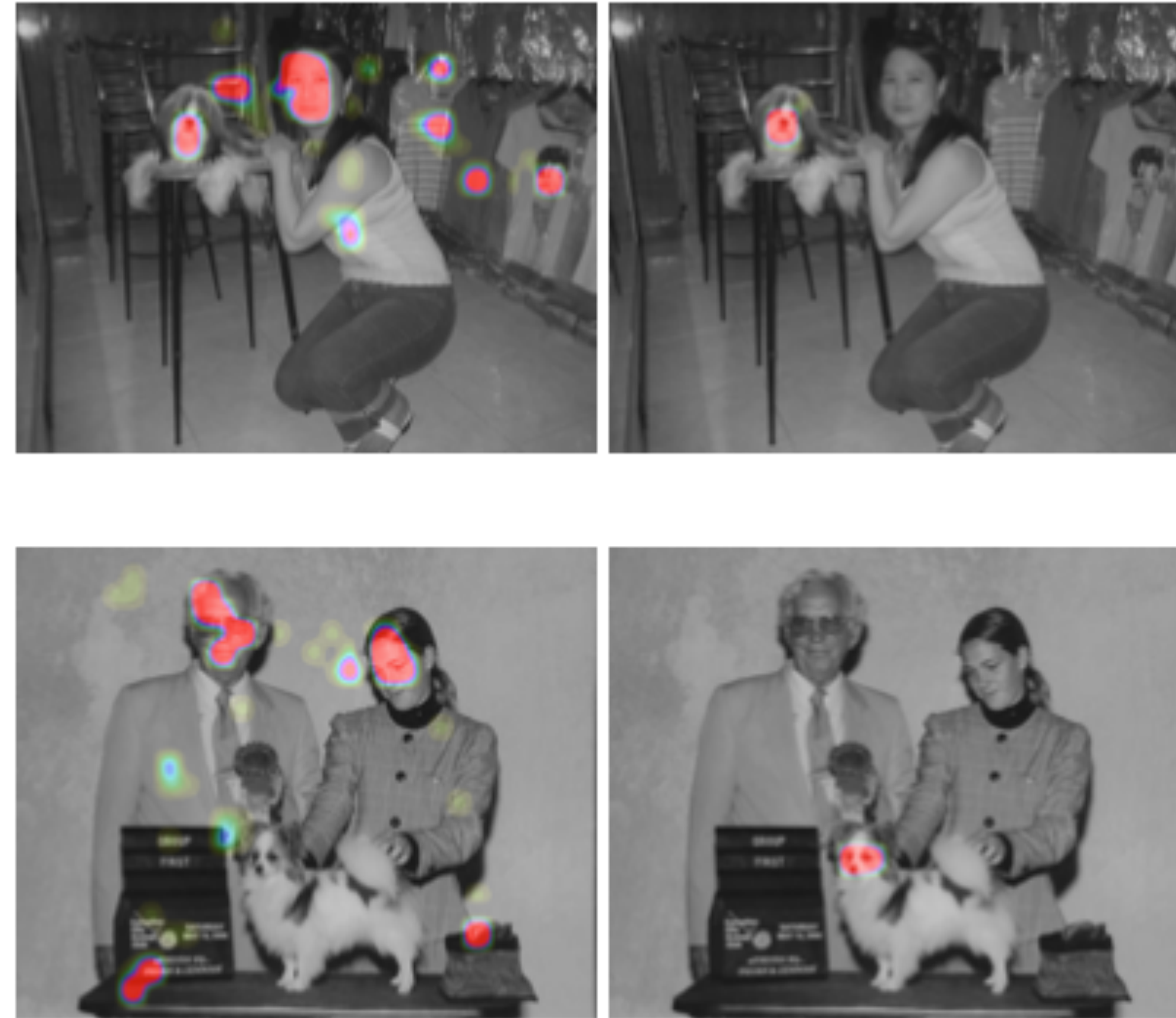
Ph.D Student: Francesca Murabito – Tutor: Prof. C. Spampinato  
PeRCeiVeLab



Dottorato di Ricerca in Ingegneria dei Sistemi, Energetica, Informatica e delle Telecomunicazioni XXXI CICLO

## Rationale

- Bottom-up saliency methods[1] might miss objects of interest in highly cluttered backgrounds.
- Low and middle level visual descriptors sometimes do not represent the most salient image parts.
- When guided by a task eye fixation may shift.



## Contributions

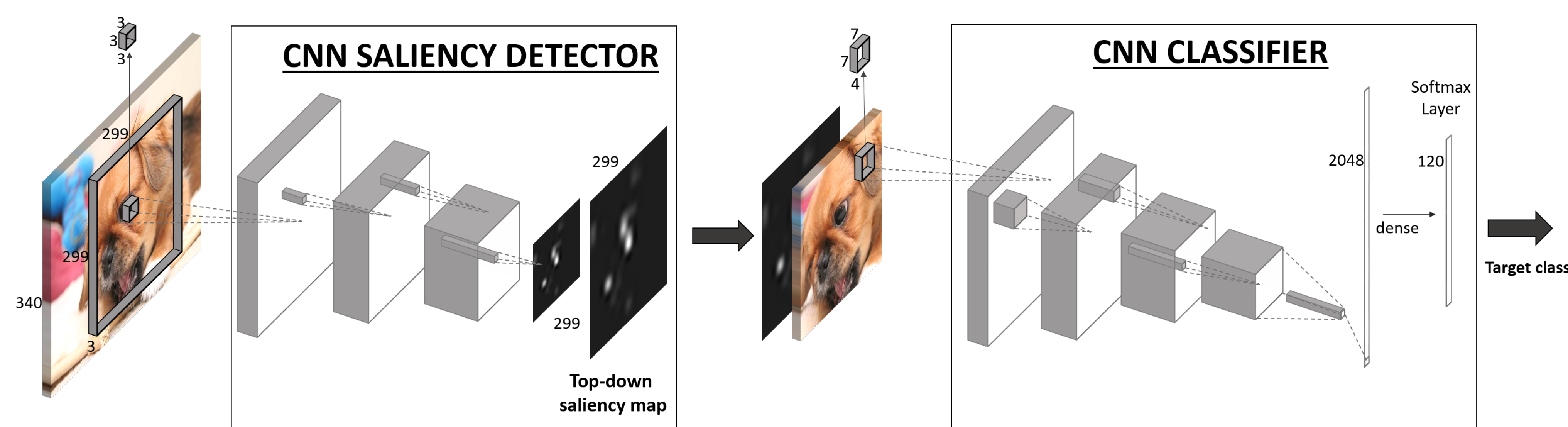
1. A saliency dataset containing about 10 K maps recorded from multiple users when performing a visual classification task on the 120 Standard Dogs[2] fine-grained classes.
2. A saliency detection method guided by a visual classification task.
3. A classification network able to exploit such task-based saliency maps in order to improve the fine-grained recognition performance.

## Top-Down Saliency Dataset

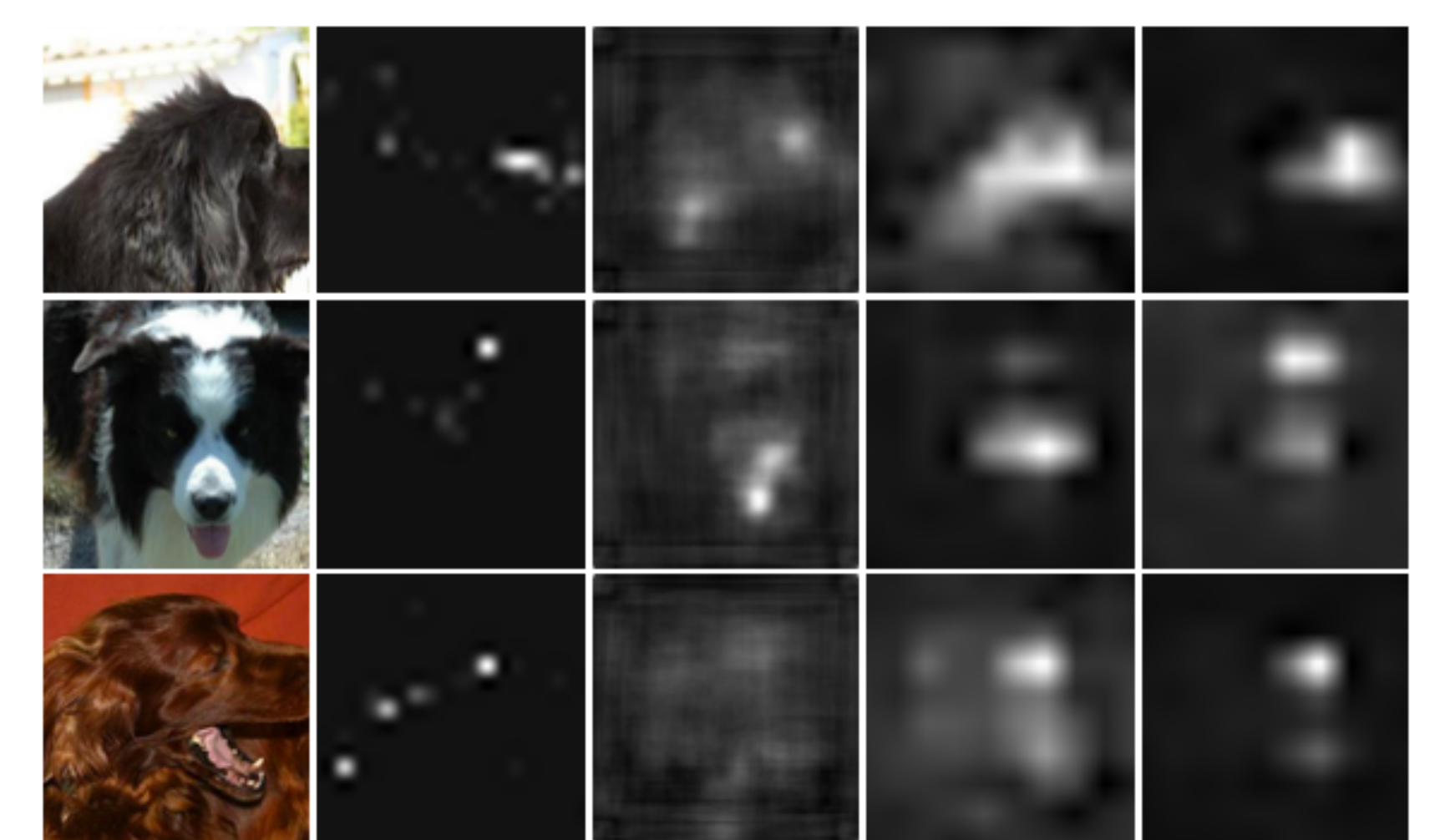
Our Dataset	
Number of images	9,861
Number of classes	120
Avg. number of images per class	82.2
Avg. number of fixation points per image	6.2

Dataset	Task	Viewers	Number of Images
SALICON[3]	Free-viewing	Crowd	20,000
iSUN[4]	Free-viewing	Crowd	8,926
MIT300[5]	Free-viewing	39	-
POET[6]	Basic classification	28	6,270
OUR	Fine-grained classification	12	9,861

## SalClassNet Architecture



Dataset	Method	s-AUC	NSS	CC
OUR	SALICON[7]	0.720	2.106	0.408
OUR	SalNet[8]	0.703	2.182	0.400
OUR	SalClassNet	0.771	2.364	0.434
POET	SALICON[7]	0.634	1.232	0.289
POET	SalNet[8]	0.649	1.093	0.300
POET	SalClassNet	0.653	1.405	0.315



## Generalization



### CUB-200-2001[9]

Method	Accuracy
VGG19[11]	50.1%
Inception[12]	64.7%
SalClassNet	65.4%

### Oxford Flower 102[10]

Method	Accuracy
VGG19[11]	64.7%
Inception[12]	77.6%
SalClassNet	78.5%

## References

- [1] Itti, L., Koch, C.. A saliency-based search mechanism for overt and covert shifts of visual attention. Vision Research 2000 .
- [2] Khosla, A., Jayadevaprakash, N., Yao, B., Fei-Fei, L.. Novel dataset for fine-grained image categorization, in: CVPRW 2011.
- [3] Jiang, M., Huang, S., Duan, J., Zhao, Q.. Salicon: Saliency in context, in: CVPR 2015.
- [4] Xu, P., Ehinger, K.A., Zhang, Y., Finkelstein, A., Kulkarni, S.R., Xiao, J.. Turkergaze: crowdsourcing saliency with webcam based eye tracking. arXiv preprint arXiv:1504.06755 2015
- [5] Bylinskii, Z., Judd, T., Borji, A., Itti, L., Durand, F., Oliva, A., Torralba, A., . Mit saliency benchmark
- [6] Papadopoulos, D.P., Clarke, A.D., Keller, F., Ferrari, V.. Training object class detectors from eye tracking data, in: ECCV 2014
- [7] Huang, X., Shen, C., Boix, X., Zhao, Q.. Salicon: Reducing the semantic gap in saliency prediction by adapting deep neural networks, in: ICCV 2015
- [8] Pan, J., Sayrol, E., Giro-i Nieto, X., McGuinness, K., O'Connor, N.E., 2016. Shallow and deep convolutional networks for saliency prediction, in: CVPR 2016.
- [9] Wah, C., Branson, S., Welinder, P., Perona, P., Belongie, S.. The Caltech- UCSD Birds-200-2011 Dataset .
- [10] Nilsback, M.E., Zisserman, A.. Automated flower classification over a large number of classes, in: ICVGIP 2008.
- [11] Simonyan, K., Zisserman, A.. Very deep convolutional networks for large-scale image recognition. ICLR 2015
- [12] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.. Going deeper with convolutions, in: CVPR 2015