Filter Recommendation through Analyzing Objects, Scenes and Aesthetics

Yi-Ning Chen and Mei-Chen Yeh
Department of Computer Science and Information Engineering
National Taiwan Normal University

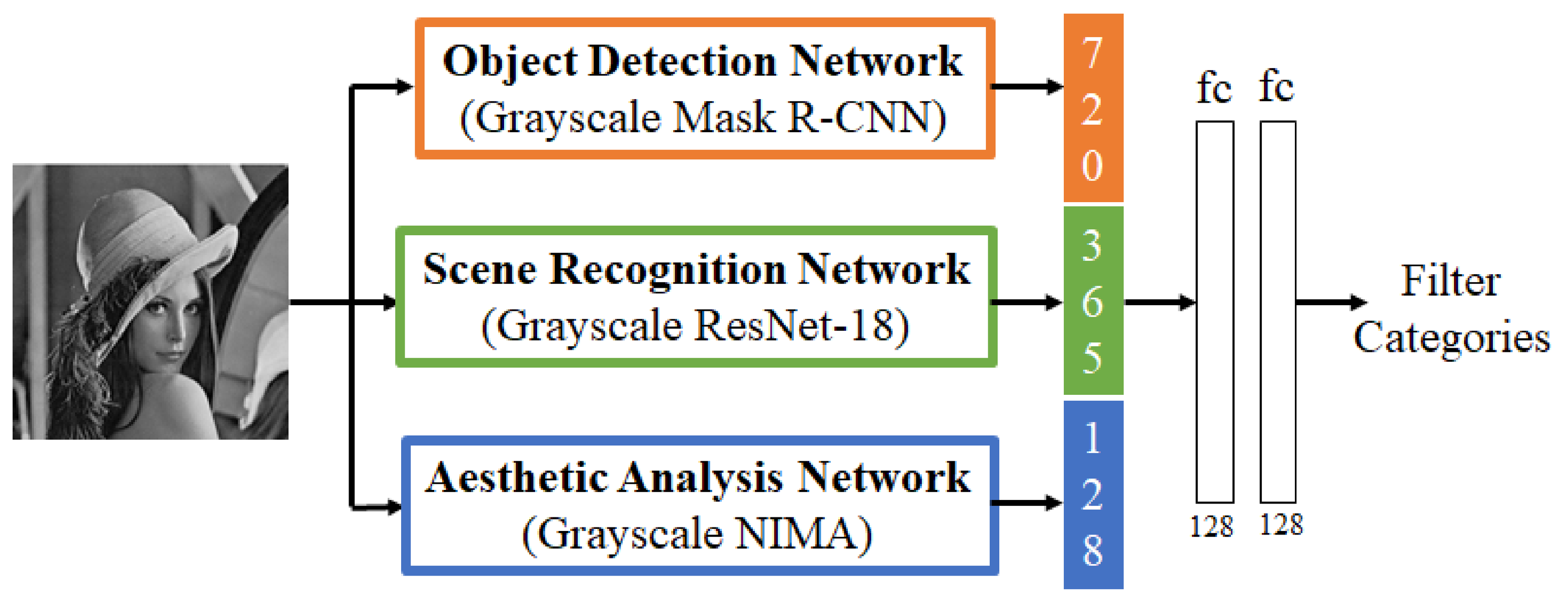
Goal: Automatic selection of photo filters based on photo content

- Demonstrate how filtered images on Instagram can be used for training a filter recommendation model
- Use high level image representation for predicting proper filters

APPROACH

Filter Recommendation Network

- ✔ Objects: Mask R-CNN + MS COCO
- ✔ Scenes: Res-18 + Places365
- ✔ Aesthetic attributes: NIMA + AVA



Style invariant representation

- ✔ Extract from grayscale images
- ✔ Fine-tune the networks using grayscale images

Object detection performance on MSCOCO

Model	AP^{bb}	AP^{bb}_{50}	AP^{bb}_{75}
Mask R-CNN	38.2	60.3	41.7
Grayscale Mask R-CNN	33.9	51.4	37.4

Scene recognition performance on Places365

Model	Top-1 acc.	Top-5 acc.
ResNet-18	54.74%	85.08%
Grayscale ResNet-18	51.00%	82.00%

Aesthetic quality estimation performance on AVA

Model	Top-1 acc.
NIMA	57.84%
Grayscale NIMA	56.43%

Mei-Chen Yeh (myeh@csie.ntnu.edu.tw)
http://www2.csie.ntnu.edu.tw/~myeh

DATASET

- 68,400 Instagram filtered photos (19 filters, 3600 photos per filter type)

	Pros	Cons
Instagram	• Low-cost collection • High diversity	• **No original photo** • One filter per photo (decided by user)
FACD [Sun 17]	• Original photo available • Multiple filters per photo (decided by AMT)	• High cost of ground truth construction • Low quantity

[Sun 17] W. -T. Sun, T. -H. Chao, Y. -H. Kuo, Winston Hsu. Photo Filter Recommendation by Category-Aware Aesthetic Learning. *IEEE Trans. on Multimedia*, 2017.

RESULTS

Recommendation results on FACD

Model	Top-1 Accuracy (%)	Top-3 Accuracy (%)
AlexNet [3]	33.13	70.63
RAPID net [4]	37.50	72.50
Category-aware learning (AlexNet) [2]	41.25	**80.00**
Category-aware learning (RAPID) [2]	41.88	79.50
Ours (Scene + Aesthetics)	**51.25**	**80.00**

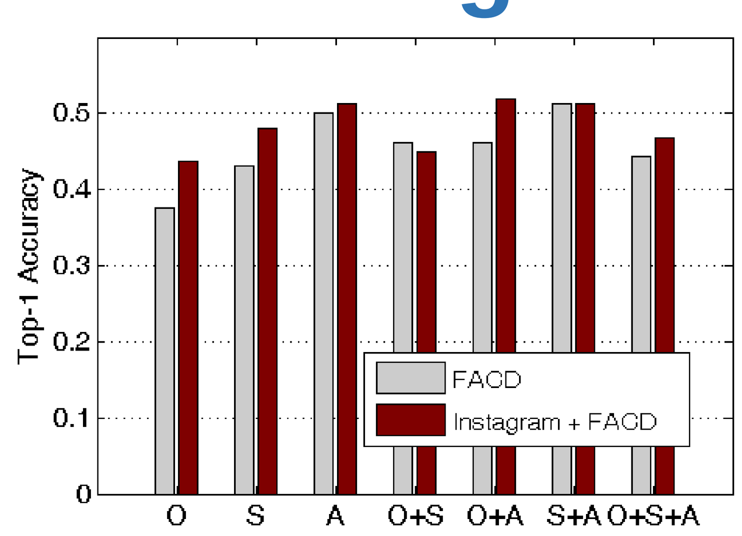
The proposed method achieved the best performances in both of top-1 and top-3 predictions. The gain was significant on the top-1 accuracy (over 9%).

Effects of different features

Model	Top-1 Accuracy (%)	Top-3 Accuracy (%)
Object	43.75	76.25
Scene	48.12	79.37
Aesthetics	51.25	75.62
Object + Scene	45.00	75.62
Object + Aesthetics	**51.87**	77.50
Scene + Aesthetics	51.25	**80.00**
Object + Scene + Aesthetics	46.87	75.00

Our approach using a single type of feature outperformed previous approaches; scenes and the aesthetics outperformed objects.

Does Instagram filtered images help?



The use of Instagram filtered images alleviated the filter recommendation task.