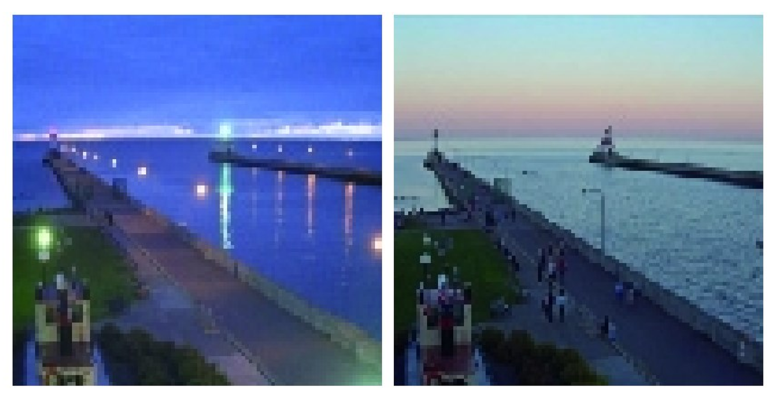
Sunrise or *Sunset* : Selective Comparison Learning for Subtle Attribute Recognition

Hong-Yu Zhou, Bin-Bin Gao, Jianxin Wu
National Key Laboratory for Novel Software Technology, Nanjing University

Have a Guess

(a) Monet's paintings



(b) Images from webcams

Motivation

Image Classification

Is this a **car** or **plane**? ✓

What is this **scenario**? ✓

What is the **brand**? ✓✗

Where, When, Temp? ✗

$\underbrace{\textbf{Where, When, Temp}}_{\textbf{Subtle Attributes}}$

Subtle Attributes

Subtle & Transient Attributes

	Transient	Subtle
Is it taken at night?	✓	✗
Which month?	✗	✓
Where?	✗	✓
Does it feel warm?	✓	✗
What is the temperature?	✗	✓
Is it cloudy?	✓	✗
Morning or afternoon?	✗	✓
Which Season?	✓	✗

Contributions

- Subtle attributes recognition
- SoS: a new dataset for sunrise or sunset
- A new comparison learning approach
- New STOA results on temperature predictions

SoS: A Benchmark

Two sets of tasks are defined for SoS: an **easy** task and a **difficult** task.

- **Easy** task: images are randomly separated into training set (10,488 images) and test set (2,522)
- **Hard** task: we use images from 104 webcams (10,448) for training, and the images from the rest 24 cameras are for testing

Deep ConvNets vs. Humans

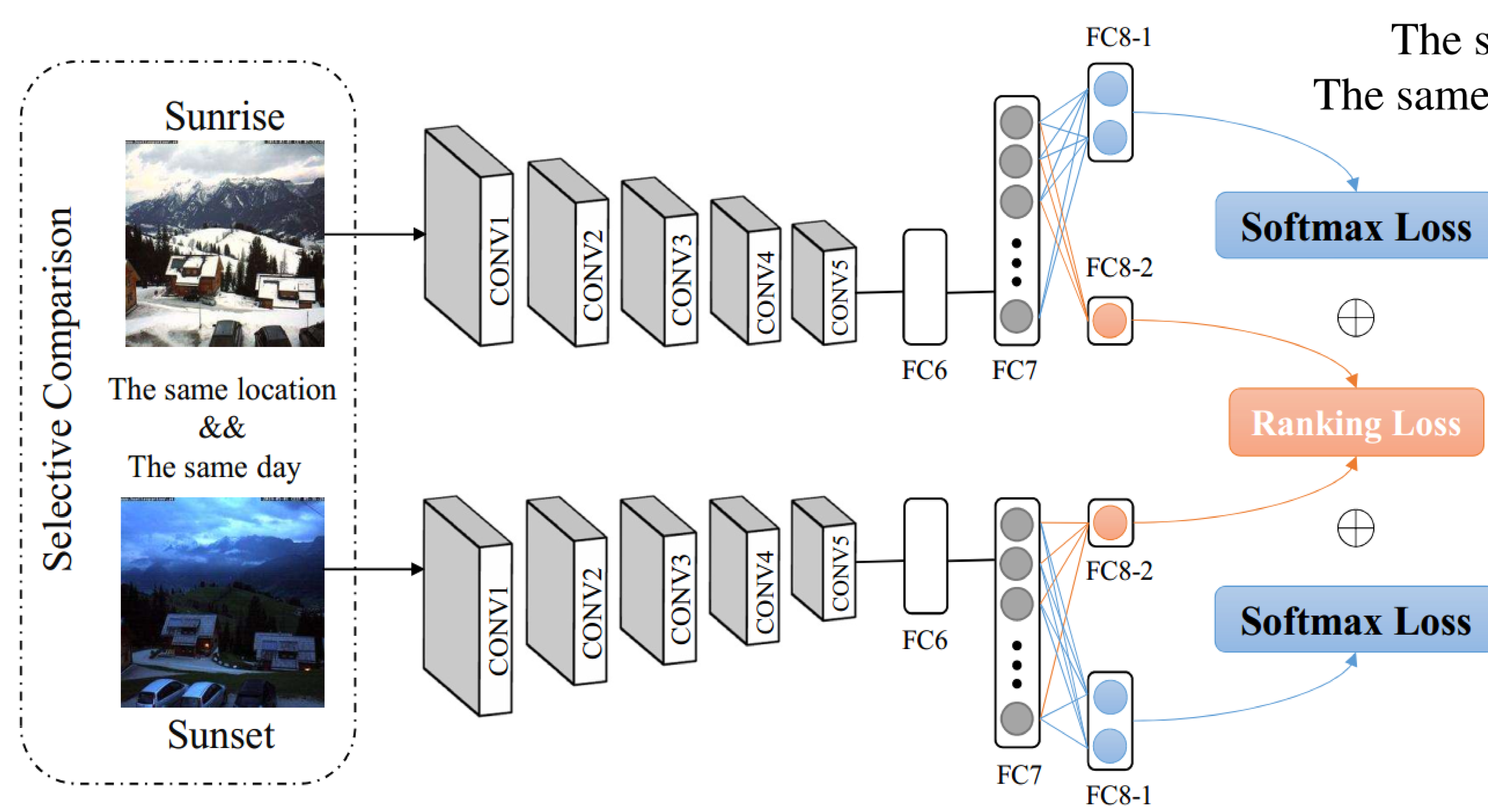
Method	Task	Sunrise	Sunset	mAcc
FT-VGG-16	Easy	79.6	80.0	79.8
FT-VGG-16	Hard	53.3	52.9	53.1
Human	Hard	64.3	63.7	64.0

More Details

- 12,970 images from 128 webcams over 30 countries

Selective Comparison Learning

Network Architecture



Loss Function

$$\ell(X_R^n, X_S^n) = \frac{1}{N}\sum_{n=1}^{N}\left(\ell_{softmax}(f_{\theta_1}(X_R^n), y_R^n) + \ell_{softmax}(f_{\theta_1}(X_S^n), y_S^n) + \lambda \ell_{ranking}(f_{\theta_2}(X_R^n), f_{\theta_2}(X_S^n))\right), \quad (1)$$

in which

$$\ell_{ranking}(f_{\theta_2}(X_R^n), f_{\theta_2}(X_S^n)) = \frac{1}{1+\exp(f_{\theta_2}(X_R^n) - f_{\theta_2}(X_S^n))}, \quad (2)$$

where $\ell_{softmax}$ is softmax loss, N is the number of pairs, $\{y_R^n, y_S^n\}$ $(n = 1, 2, \ldots, N)$ represent image-level labels and λ is a balancing parameter. This loss function is differentiable with respective to its parameters.

Experimental Results

Sunrise or Sunset Acc.

	FV(s.p.)	FV	VLAD(aug.)	BOVW(aug.)	AlexNet	NiN	VGG-16	ResNet-101	SiameseNet	SoSNet-rand	SoSNet
sunrise	54.1	53.4	50.6	56.6	52.2	52.1	53.3	53.8	54.2	58.6	**70.9**
sunset	52.1	54.0	52.8	56.0	52.0	52.5	52.9	53.2	54.8	59.0	**71.6**
mAcc	53.1	53.7	51.7	56.3	52.1	52.3	53.1	53.5	54.5	58.8	**71.2**

Ablation Studies

Give a pair of input images, the performance of humans on the SoS dataset. Note that each volunteer is shown 5 groups of paired images corresponding to 5 different settings, and each group contains 50 pairs. SS is another restriction on each pair which requires that one image is sunrise, the other is sunset.

Pair Constraint	SS	Sunrise	Sunset	mAcc
Random pair	w/o	65.3	64.9	65.1
Random pair	w	67.3	66.9	67.1
The same day	w	67.8	66.8	67.3
The same location	w	70.7	70.0	70.3
The same location and day	w	**72.4**	**72.4**	**72.3**

Temperature Estimation Results

The **first three** use simple **pixel intensities** as Features while the **last two** use more sophisticated **global image features**. The first five methods are proposed by Glasner et al.

R^2 (the higher the better)/ *RMSE* (the lower the better)

	Local Regression	LR Temporal Win.	Global Ridge Reg.	CNN	Transient Attrib.	FC6	Pool4	Ours
(a)	0.67/6.85	0.61/7.52	0.00/18.16	0.49/8.55	0.36/9.60	0.52/8.28	0.58/7.79	**0.73/6.26**
(b)	0.65/7.24	0.69/6.86	0.78/5.74	0.79/5.59	0.70/6.69	0.80/5.46	0.84/4.87	**0.89/4.57**
(c)	0.70/6.03	0.72/5.82	0.00/35.02	0.71/5.96	0.58/7.20	0.61/6.89	0.79/5.03	**0.83/4.92**
(d)	0.59/4.53	0.64/4.23	0.00/11.37	0.24/6.17	0.55/4.75	0.56/4.72	0.60/4.45	**0.70/3.80**
(e)	0.76/5.77	0.79/5.39	0.00/43.51	0.61/7.30	0.68/6.62	0.80/5.30	0.87/4.22	**0.90/3.98**
(f)	0.38/3.19	0.53/2.77	0.10/3.84	0.48/2.90	0.21/3.59	0.21/3.60	0.40/3.14	**0.58/2.53**
(g)	0.50/7.63	0.54/7.35	0.74/5.54	0.39/8.48	0.58/7.03	0.54/7.34	0.63/6.61	**0.80/5.20**
(h)	0.77/5.09	0.76/5.22	0.00/13.86	0.79/4.88	0.65/6.31	0.79/4.90	0.80/4.72	**0.86/3.95**
(i)	0.10/3.68	0.11/3.67	0.23/3.41	0.43/2.93	0.16/3.56	0.06/3.78	0.52/2.70	**0.55/2.48**
(j)	0.59/7.77	0.58/7.85	0.46/8.91	0.66/7.12	0.67/7.00	0.59/7.80	0.76/6.01	**0.78/5.81**