

Supplementary Material for Image Captioning with Sentiment Terms via Weakly-Supervised Sentiment Dataset

Andrew Shin
andrew@mi.t.u-tokyo.ac.jp
Yoshitaka Ushiku
ushiku@mi.t.u-tokyo.ac.jp
Tatsuya Harada
harada@mi.t.u-tokyo.ac.jp

Graduate School of
Information Science and Technology,
The University of Tokyo
Tokyo, Japan

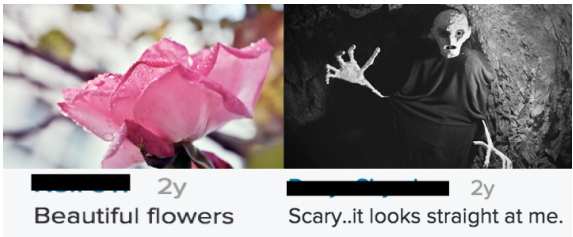


Figure 1: Actual examples of comments on social network services describing the sentiments of the images

Table 1: Size of vocabulary, parameters, dimensions of sentiment features after PCA for each dataset

Dataset	V	V _{senti}	Param	Param _{senti}	Dim _{senti}
Flickr8k	4,448	104	3,855,968	4,186,208	1,290
Flickr30k	11,794	399	7,624,466	7,824,658	782
MSCOCO	14,141	512	8,828,990	9,260,350	1,685

Table 2: Performances of the captions generated by each model on Flickr 8k and Flickr 30k

Dataset	Model	BLEU1	BLEU2	BLEU3	BLEU4	METEOR	Cider
Flickr8k	ImNet	51.1	33.0	20.1	12.6	11.4	39.4
	ImNet+	47.8	29.3	16.6	9.6	10.9	30.6
	Bigram	48.2	29.4	16.7	9.7	10.9	30.5
	Style	44.6	26.3	14.6	8.3	10.4	29.6
	Ours	46.6	28.0	15.6	9.0	10.5	30.3
Flickr30k	ImNet	55.0	35.7	22.9	14.8	10.5	33.0
	ImNet+	51.6	31.9	19.2	11.6	10.4	25.9
	Bigram	51.8	31.9	19.1	11.5	10.4	27.1
	Style	49.3	30.1	17.9	10.8	9.9	22.0
	Ours	51.1	31.0	18.5	11.1	10.1	23.0



Figure 2: More examples of comment-generated labels.



GT: pedestrians walk by outside a cafe
ImNet+: a man in a blue shirt is standing in **excess** front of a store
BiGram: a man in a blue shirt is standing in **united** front of a store
Style: a group of **lean** people are walking down a street
Senti: a man in a blue shirt and blue jeans is standing in **hollywood** front of a building



GT: two woman wearing similar shirts walk to the left
ImNet+: a woman in a white shirt and black pants is walking down the **littered** street
BiGram: a woman in a white shirt and black pants is walking down the **side** street
Style: a woman in a white shirt and black pants is walking down the **cleared** street
Senti: a **casual** woman in a white shirt and a woman in a white dress are dancing



GT: people playing a game in the pool
ImNet+: a man in a blue shirt is jumping into the **hot** air
BiGram: a man in a blue shirt is jumping into the **hot** air
Style: a **young** man in a blue shirt is surfing on a wave
Senti: a man in a blue shirt is jumping into the **drenched** air



ImNet+: a man is skiing **dirty** down a snowy hill
BiGram: a man is skiing **gone** down a snowy hill
Style: a man riding skis down a snow covered **protected** slope
Senti: a man in a blue **colored** jacket is skiing down a hill



ImNet+: a man in a blue **colored** jacket is walking through a forest
BiGram: a man in a blue **leather** jacket is walking through a forest
Style: a man in a blue shirt and a black hat is walking down a **snowy** hill
Senti: a man in a blue shirt and blue jeans is walking down a **rocky** path



ImNet+: a girl **shallow** in a pink shirt is playing with a toy
BiGram: a girl **defined** in a pink shirt is playing with a toy
Style: a little girl in a pink swimsuit is jumping into a swimming **overlooked** pool
Senti: a **young** girl in a pink dress is playing with a toy

Figure 3: More examples of generated captions with sentiment terms from all models.