

Supplemental Material: Basic-Level Categories using Visual and Language Context

1. Performance on SBU-1KA and SBU-1KB

Table 1 shows the performance of our method, *BasicName-Visual*, as compared to three different baselines on the SBU-1KA and SBU-1KB datasets. This table is included because it makes the numeric values of the performance clear and it matches the reporting style of Ordonez *et al.* From this table we can see that *BasicName-Visual* method outperforms the *Direct-to-noun* baseline and the results of Ordonez *et al.*, which are shown in the table as *Ngram-biased+SVM*.

	SBU-1KA		SBU-1KB	
	Precision	Recall	Precision	Recall
Ngram-biased+SVM	19.9 ± 1.2	10.4 ± 0.7	25.1 ± 2.4	14.4 ± 1.4
Direct-to-noun	20.0 ± 1.6	10.5 ± 0.9	21.5 ± 1.6	12.2 ± 0.7
Frequency+described	26.1 ± 1.7	12.8 ± 1.0	28.1 ± 1.3	15.2 ± 1.5
BasicName-Visual	26.3 ± 1.8	12.9 ± 1.1	28.9 ± 1.3	15.6 ± 1.2

Table 1. The precision and recall at 5 for different methods, evaluated on the SBU-1KA and the SBU-1KB datasets. This clearly shows that our method outperforms the Ordonez *et al.* approach, *Ngram-biased+SVM*, in terms of both precision and recall.

2. Noun prediction performance on strongly improved synsets

To further illustrate that using visual context leads to a performance improvement for some synsets we consider predicting nouns only for synsets which show *notable improvement* in Figure 3 from the attached paper. The image to noun prediction results for this restricted concept domain are shown as a precision recall curve in Figure 1. We consider this figure in conjunction with the small improvement of the *BasicName-Visual* method over the *Frequency+described* seen on the SBU-148K dataset. This leads to the conclusion that in the general domain the gains from using visual context to choose nouns is partially hidden by the semantic concepts for which visual context is not useful.

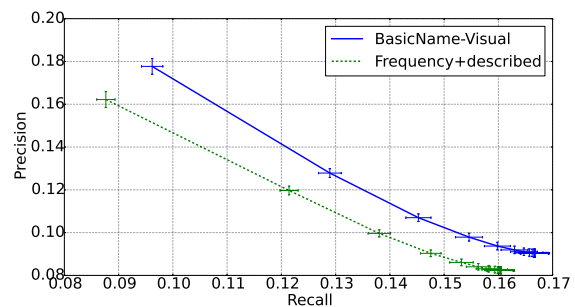


Figure 1. The precision recall curve for our method and the frequency method, considering only synsets that we show *notable improvement* in Figure 3 from the attached paper