



Sense Discovery via Co-Clustering on Images and Text

Xinlei Chen¹, Alan Ritter², Abhinav Gupta¹, Tom Mitchell¹

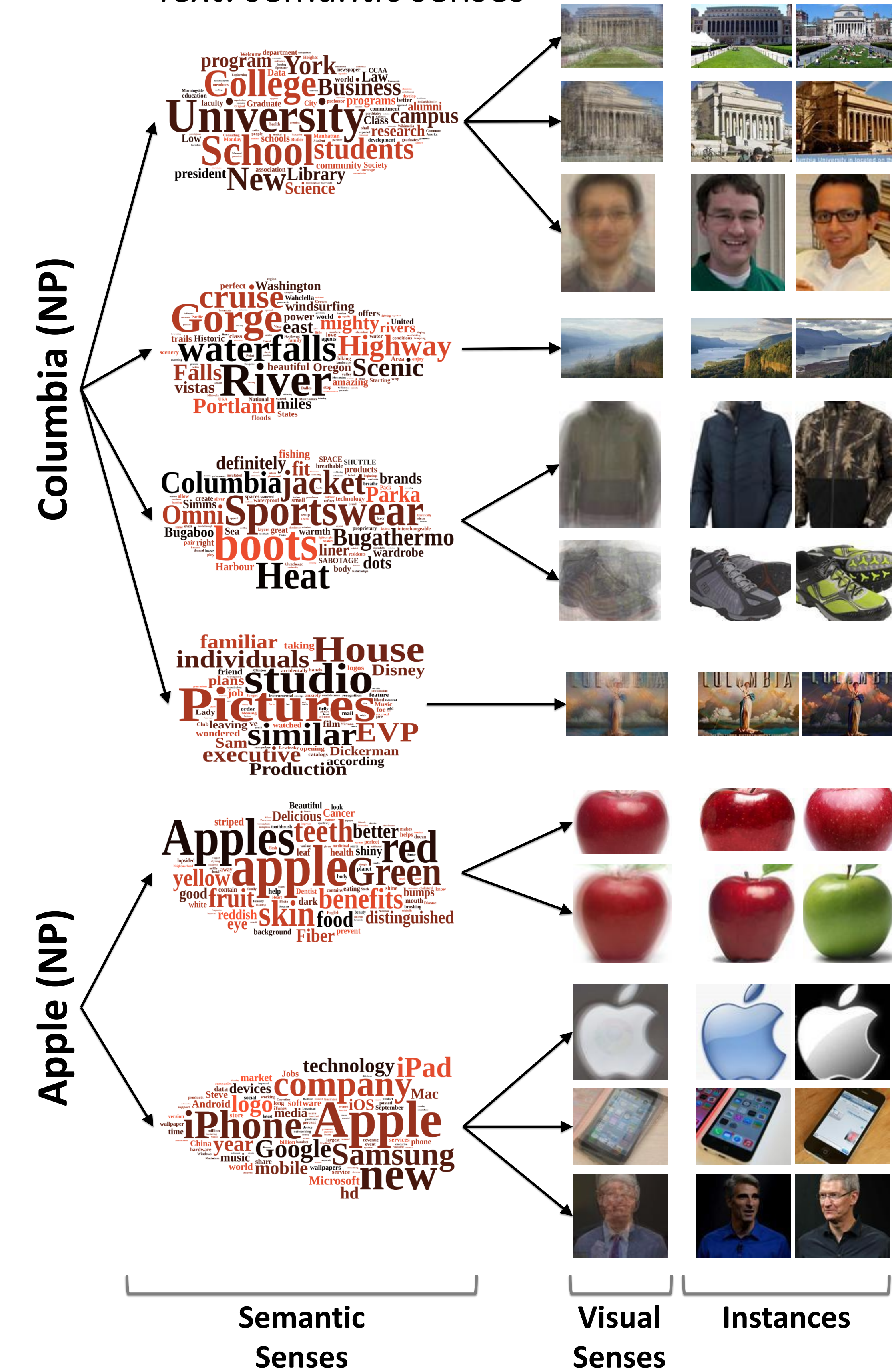
¹Carnegie Mellon University ²Ohio State University



Overview

Goal: Co-clustering images and text for a Noun Phrase (NP) to discover senses

- Images: visual senses
- Text: semantic senses



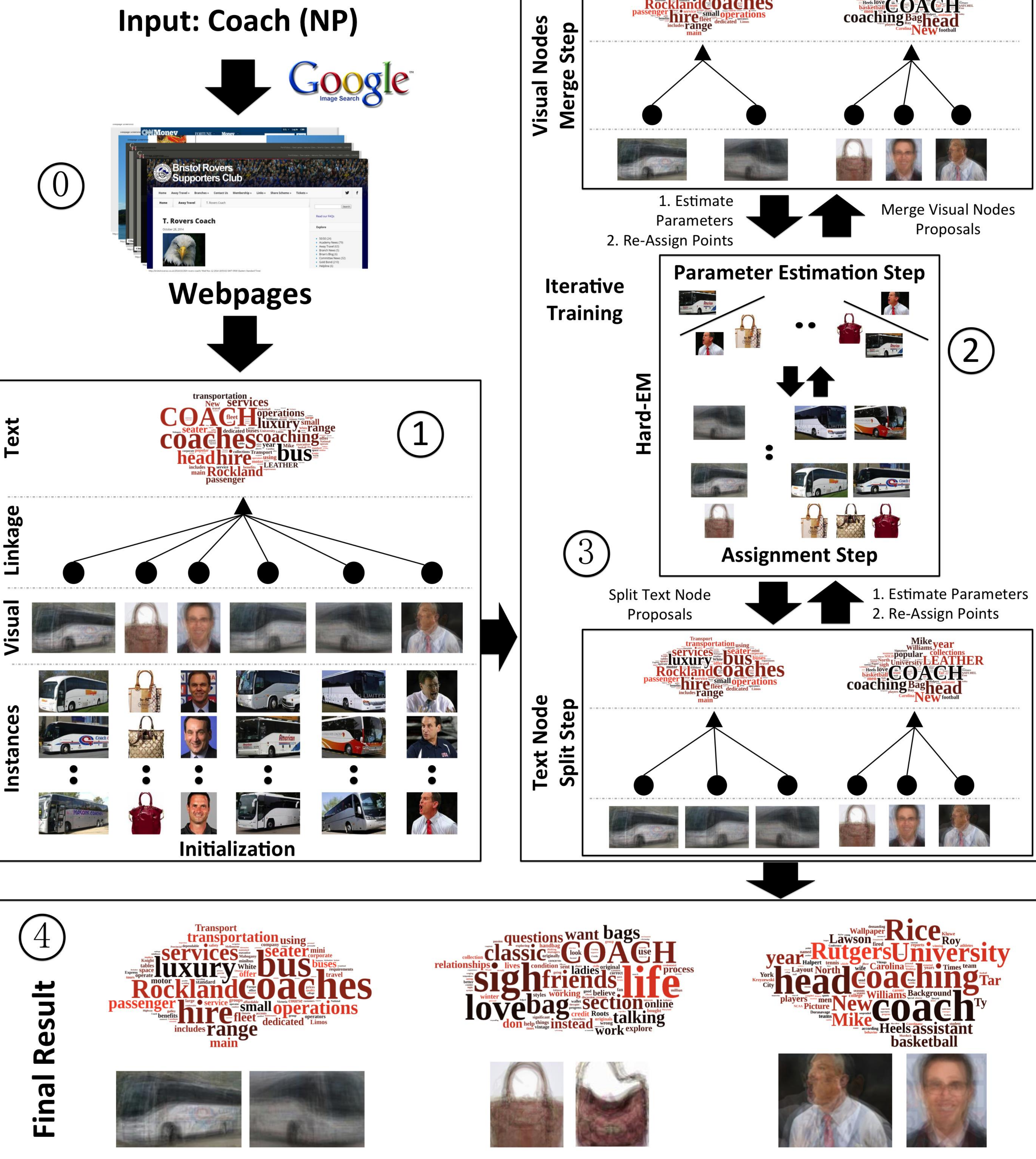
Observation: One-to many mapping between the two domain. I.e., One visual sense can only belong to one semantic sense

Formulation:

- Data point: $X_i = \langle x_i^1, x_i^2 \rangle$
- Cluster index: $Y_i = \langle y_i^1, y_i^2 \rangle$
- Energy function:

$$\sum_{d \in \{\mathcal{D}_1, \mathcal{D}_2\}} \left(\sum_i \Psi^d(x_i^d, y_i^d, \theta^d) + \sum_{i,j} \Phi^d(x_i^d, x_j^d, y_i^d, y_j^d) \right) + \sum_{i,j} \Phi^{12}(x_i^1, x_j^1, y_i^2, y_j^2)$$

Approach

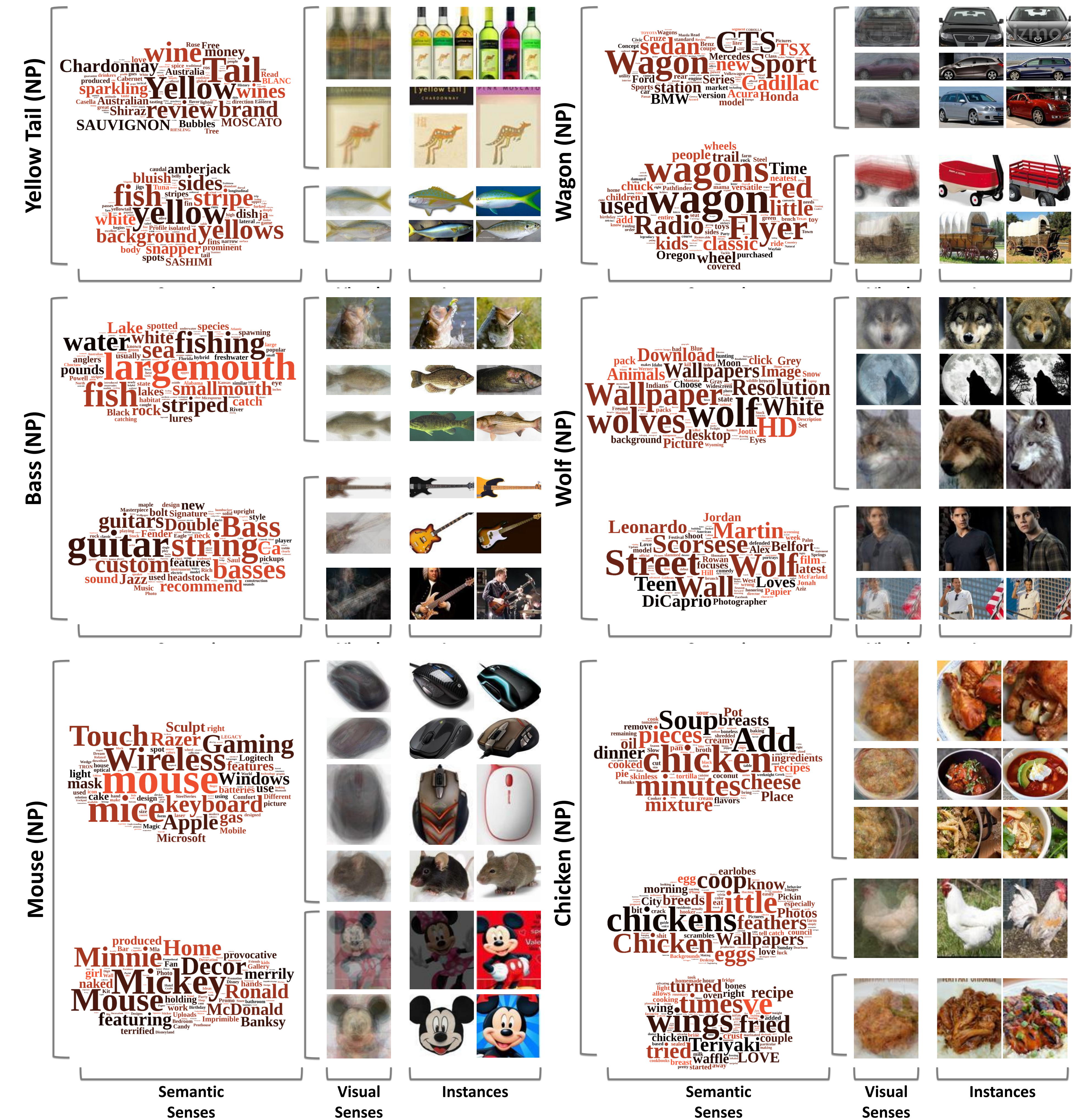
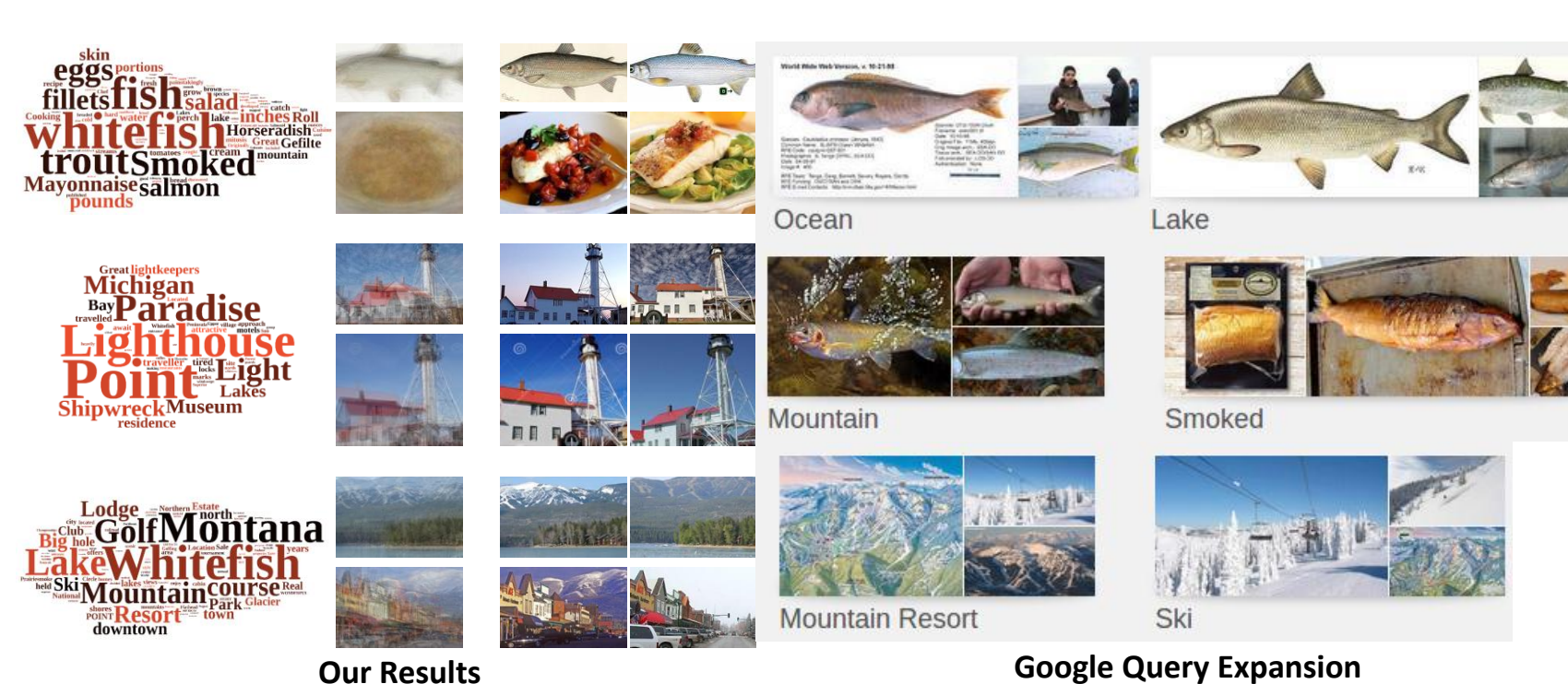


Results

CMU Poly-30: semantic sense annotations

Airbus A380	732	AK 47	654	Apple	635	Bass	804
Bean	902	Black Swan	507	Chicken	845	Coach	754
Columbia	779	Corolla	667	Daybed	763	F18	611
Falcon	831	Football	559	Jordan	662	LA Lakers	651
M16	664	Mouse	838	M Lancer	663	Motorbike	850
Note	766	Robin	830	Sofa	585	Sparrow	859
Subway	803	Tuna	814	Wagon	824	Whitefish	817
Wolf	779	Yellow Tail	682				

Comparison with Query Expansion



- Pseudo-NP:** Pseudo polysemous NPs created by combining together multiple single-sense NPs; we used accord, boding, tire, cricket ball for our evaluation
- Also evaluated on **NEIL** and **MIT ISD** (see paper)!

	AC		NMI	
	Fixed	Eigen	Fixed	Eigen
TF-IDF	79.31	76.66	11.04	19.65
BOW	77.94	70.79	15.97	12.64
I-BOW	77.86	70.17	6.84	8.04
[37]	80.52	73.21	18.69	17.75
[24]	78.39	79.89	8.96	27.00
[45]	80.62	73.13	19.03	18.26
Our Approach	86.70		34.11	

CMU Poly-30 Dataset

[37]: Saenko and Darrel 2008 [24]: Loeff et al 2006 [45]: Wan et al 2009

	AC		NMI	
	Fixed	Eigen	Fixed	Eigen
TF-IDF	53.69	49.12	10.32	6.07
BOW	58.98	68.95	15.97	31.64
I-BOW	61.66	57.47	22.03	16.03
[37]	70.04	77.20	46.55	44.52
[24]	54.83	55.52	12.80	18.72
[45]	69.08	72.83	32.80	43.33
Our Approach	83.89		53.81	

Pseudo-NP Dataset