

# Tag Selection and Propagation for Improved Large-Scale Visual Landmark Recognition

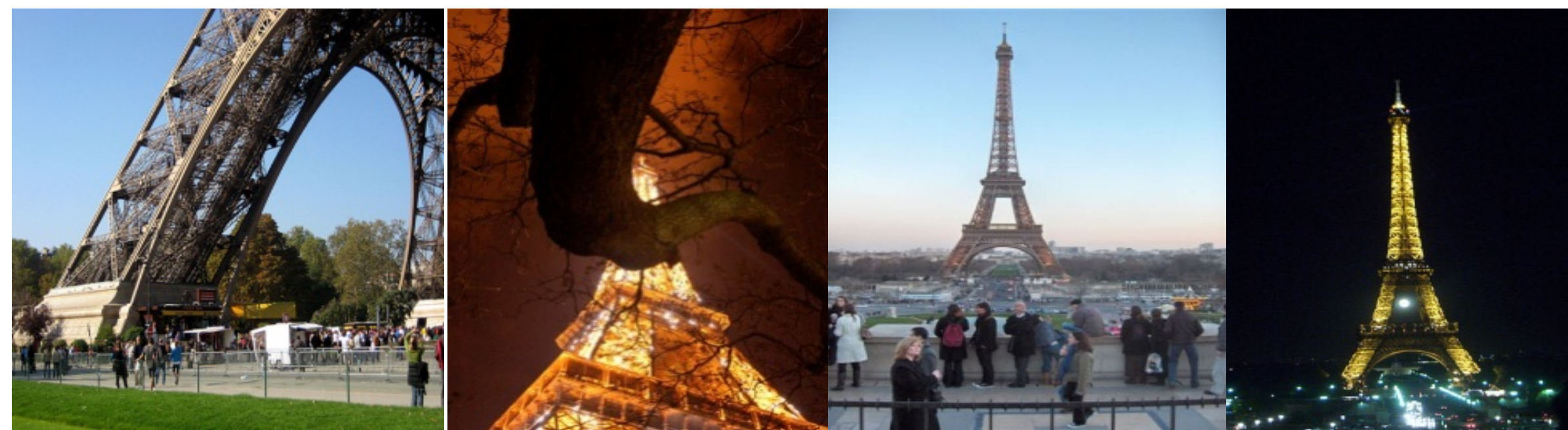


Manu Singh and David Crandall, Indiana University Bloomington

## Introduction

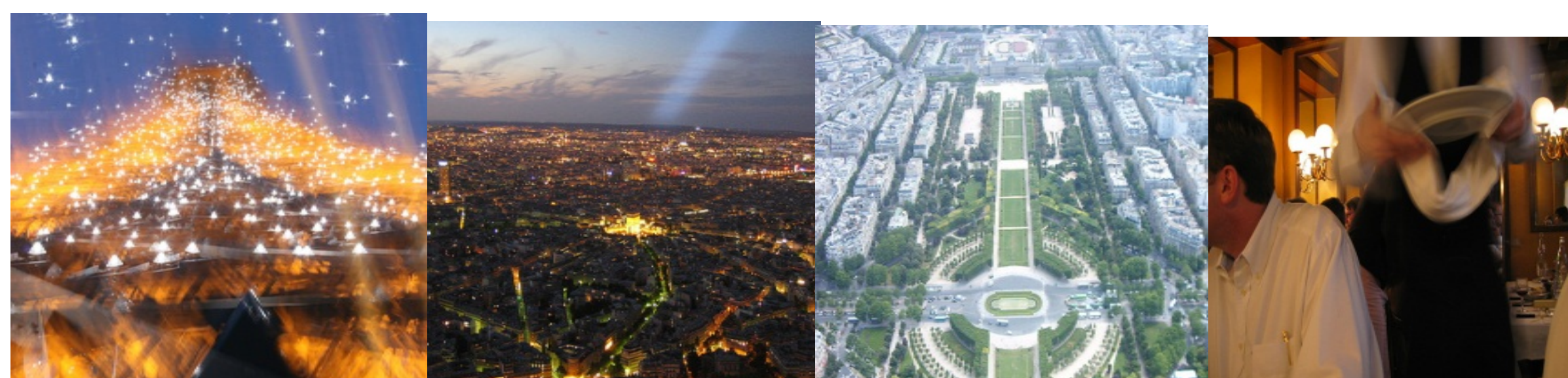
The task of landmark recognition is to locate places or landmarks in query photos. The ever increasing coverage of the world's landmarks through public available web platforms such as Flickr provides a good source of data to build these systems. The query photo may have metadata information such as textual tags which can provide additional cues for classification. The process consists of 1) clustering geotag images 2) determining landmarks through peaks and labels from tags and 3) building a robust recognition system.

Eiffel Tower - High confidence images and tags.



Paris, france, eiffel, eiffeltower, tower, toureiffel, europe, tour, night, travel

Eiffel Tower - Low confidence images and tags.



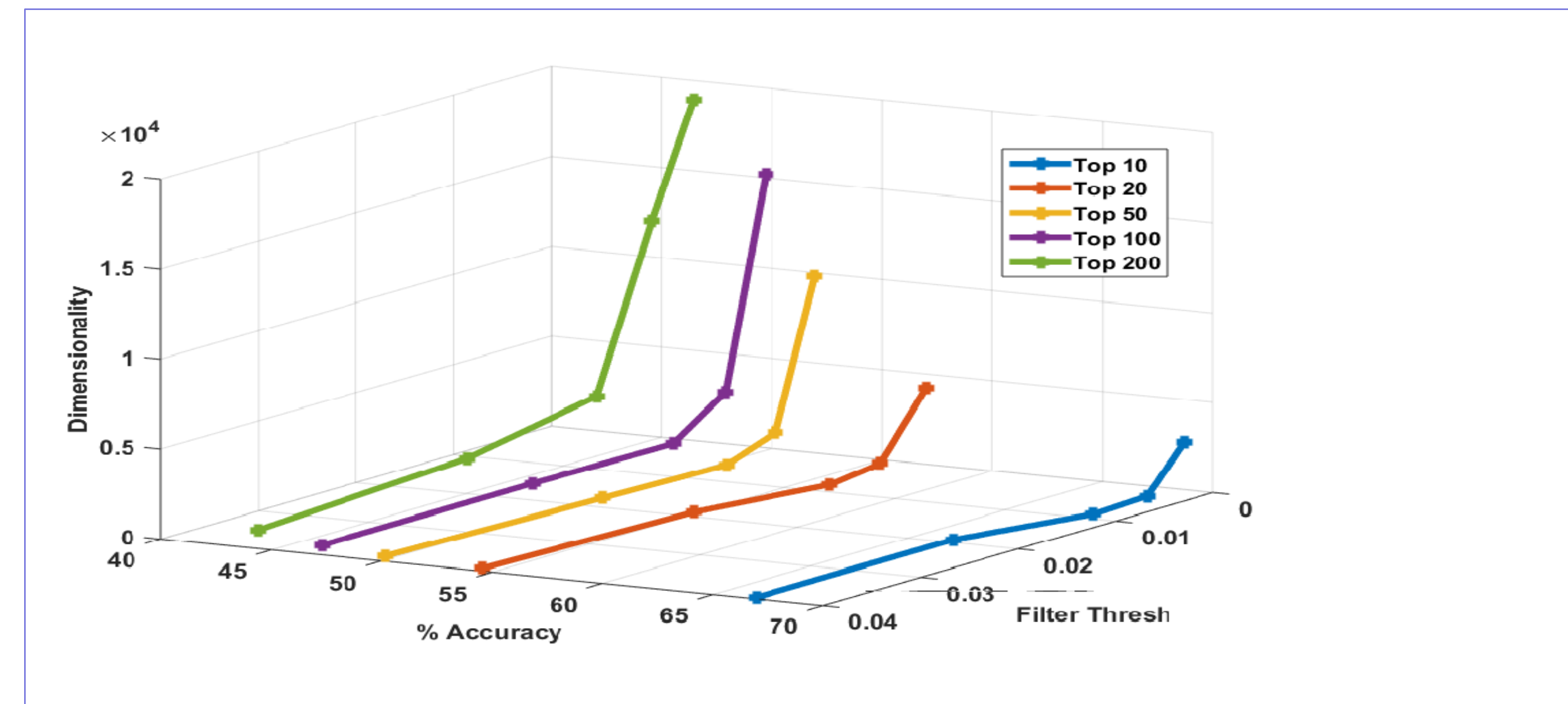
leaves, beauty, spain, espanha, bike, mexico, picture, nophotoshop, amateur, andrea

## Motivation and Objective

Li et al's [1] work on large scale image classification shows empirically that the non-visual data (Textual Tags) provide strong cues and even outperform the visual features on all categories of classification tasks.

Here we propose two enhancements:

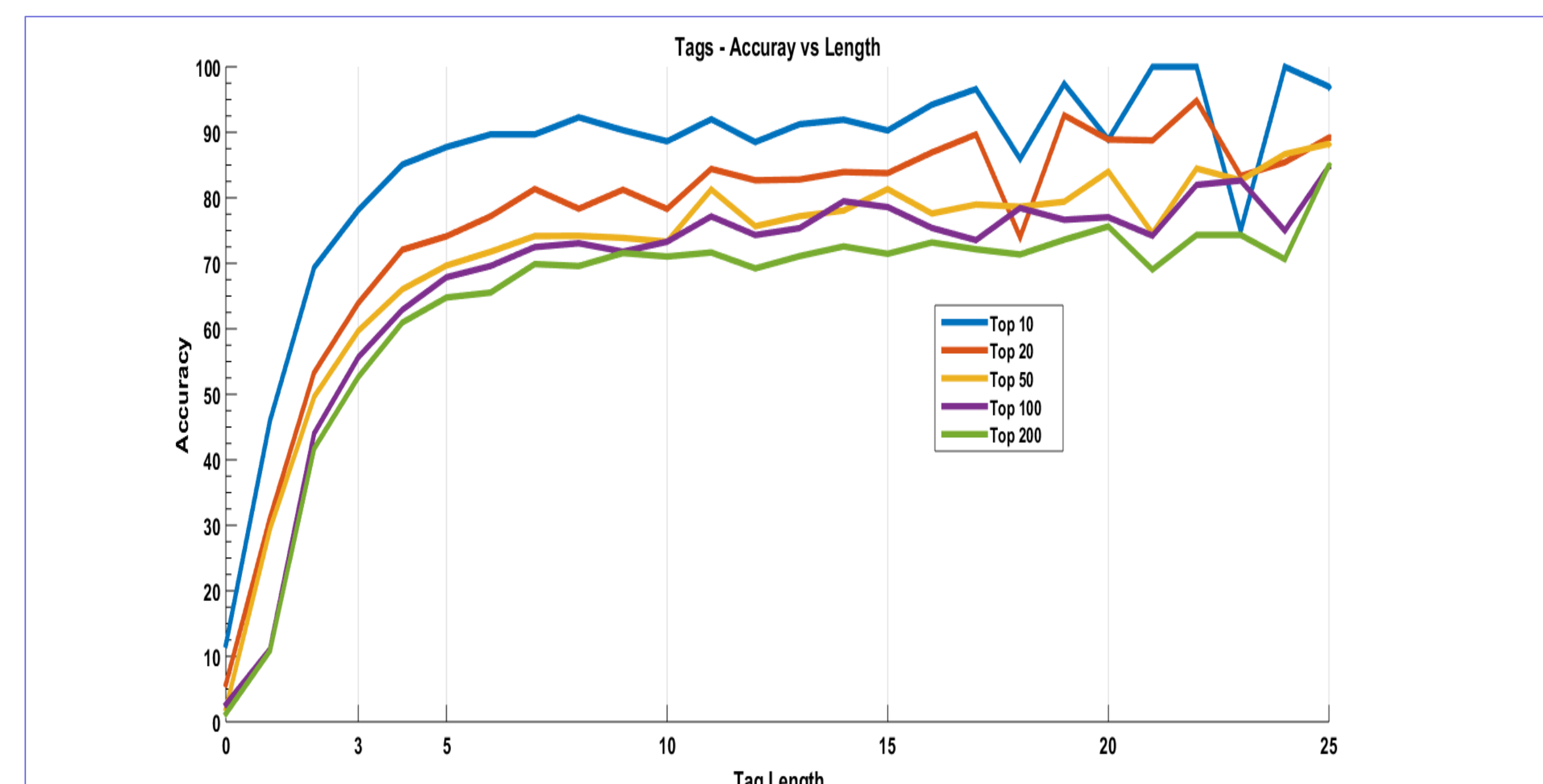
- Reduce the dimensional/sparsity of the textual features rendering a faster system while maintaining the baseline accuracy.
- Use tag propagation [2] to further boost the accuracy for missing/few textual tag features.



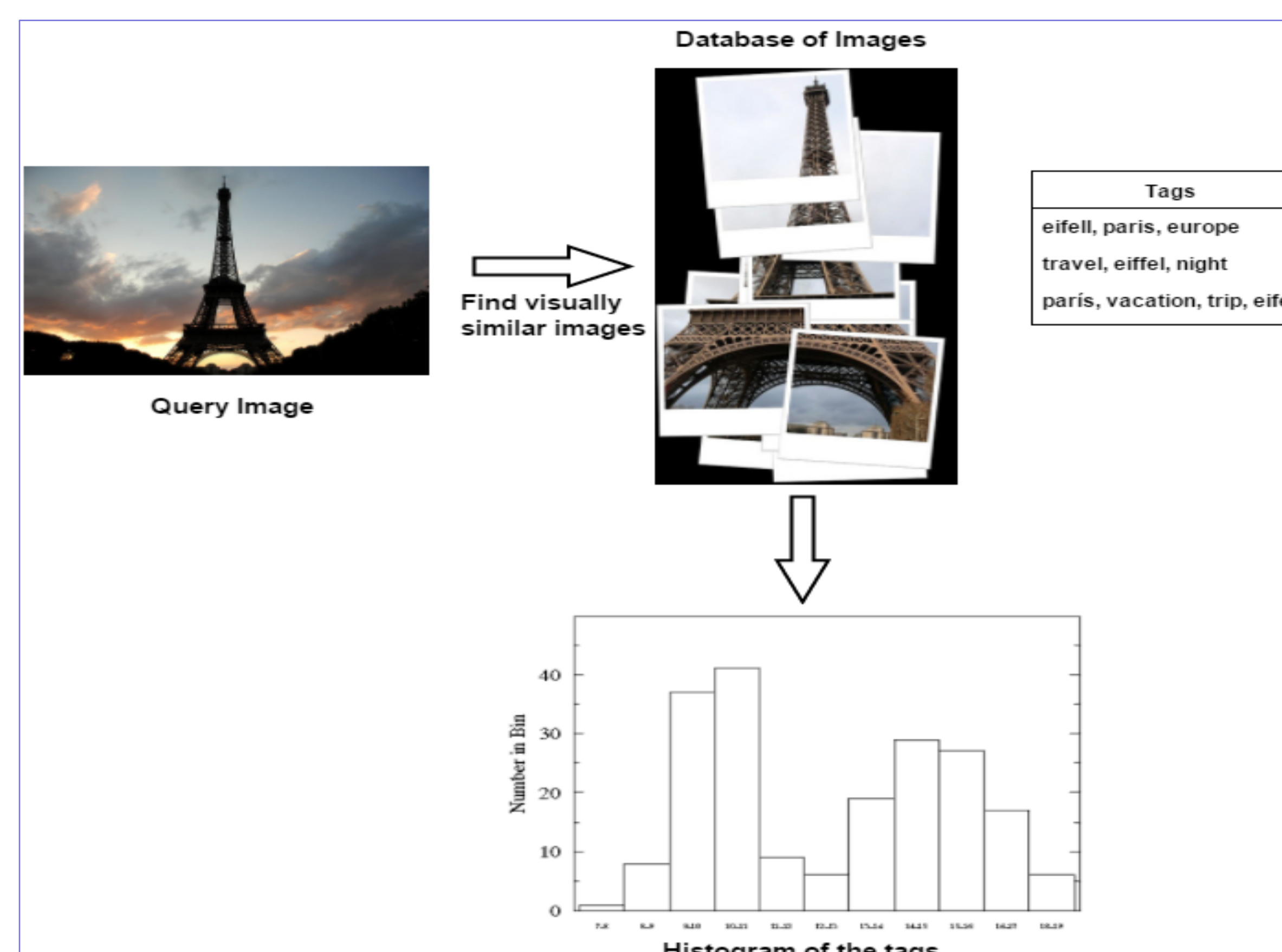
Filter threshold effect on accuracy and dimensionality

## Method

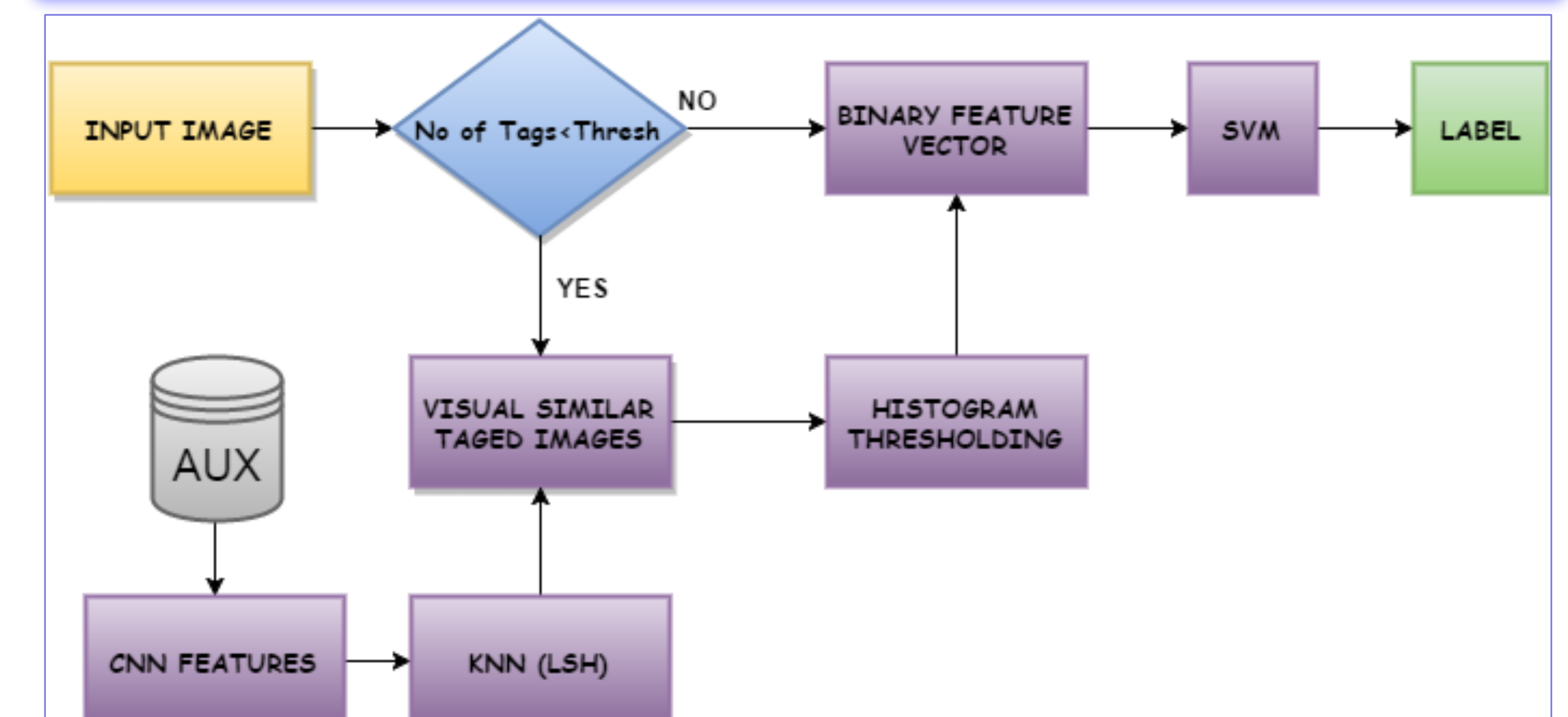
- Filter noisy tags based on ranking. (dimensionality + sparsity reduction)
- Match filtered tags based on context and spelling. (sparsity reduction)
- Tag propagate missing/few textual tag lengths. (boost accuracy)



Accuracy per tag length



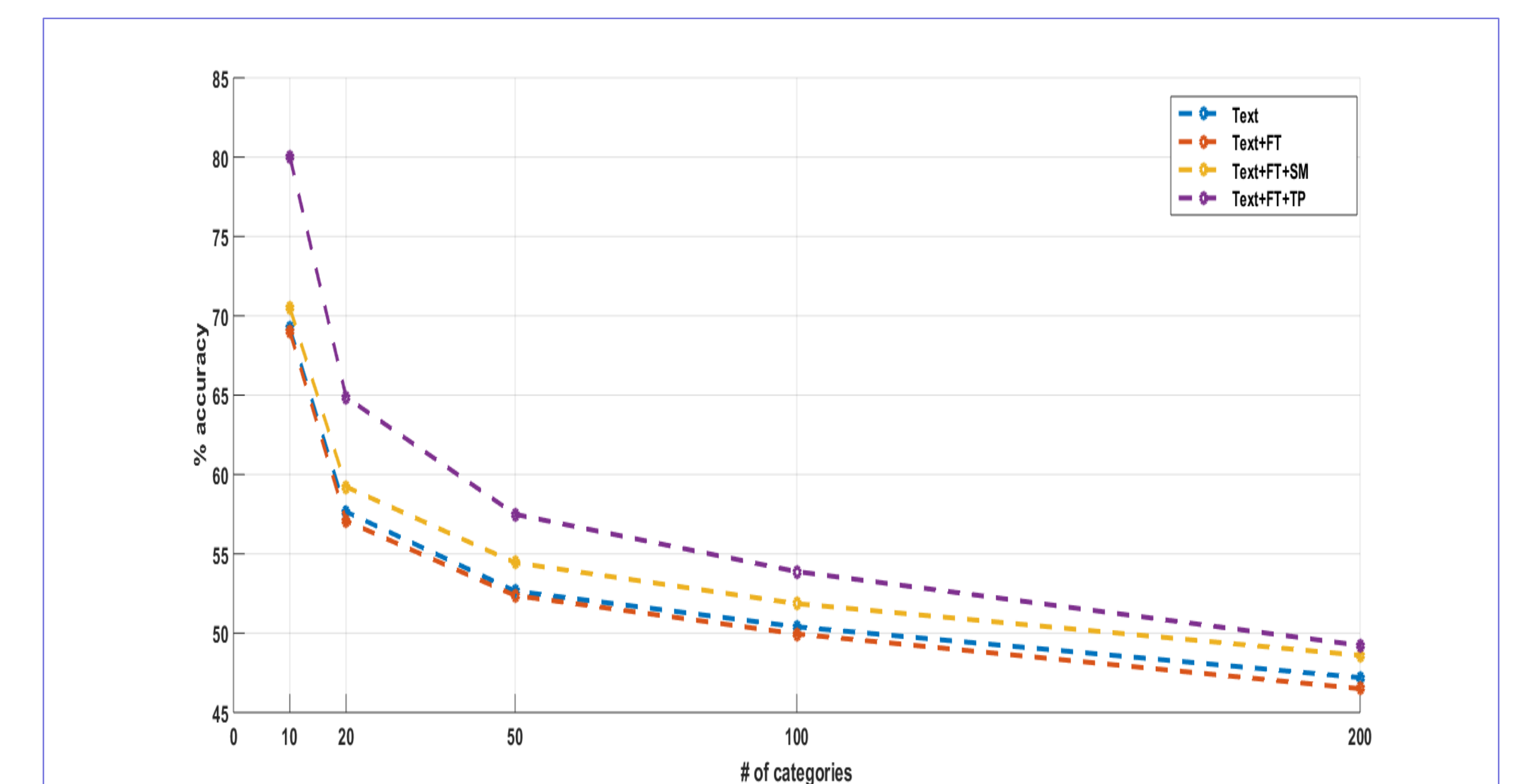
## Workflow



## Results

Context Matching(Tags)		Similarity matching(Tags)	
horizon	View	venexia	venexia
lighting	light	fontaine	fountain
nighttime	night	venizia	venezia
pedal	wheel	blackwhite	blackandwhite
circus	carnival	thecolosseum	colosseum

Landmark	Dimensionality	Dimensionality (Filter)	reduction
Top 10	6659	256	96%
Top 20	8552	422	95%
Top 50	10976	948	91%
Top 100	16976	1817	90%
Top 200	19197	4015	79%



## References

1. Y. Li, D. J. Crandall, and D. P. Huttenlocher. "Landmark classification in large-scale image collections", in ICCV 2009
2. G Wang, D Hoiem, and D Forsyth, "Building text features for object image classification," in Computer Vision and Pattern Recognition, 2009