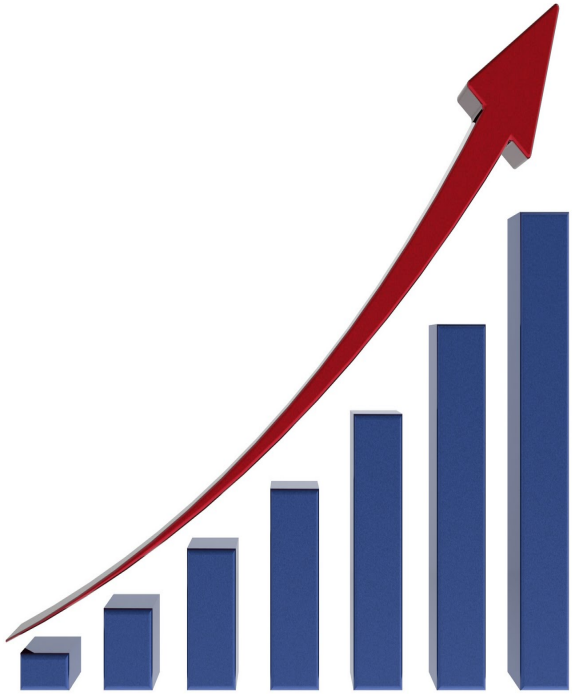
Introduction (Motivation)

- In 2015, a total of 728 millions of public pictures were uploaded to Flickr



- Such large amount of *user-generated data makes multimedia indexing and retrieval a more challenging task*

- *However, it also opens new opportunities for development of novel and more efficient tools*

Introduction (Motivation)

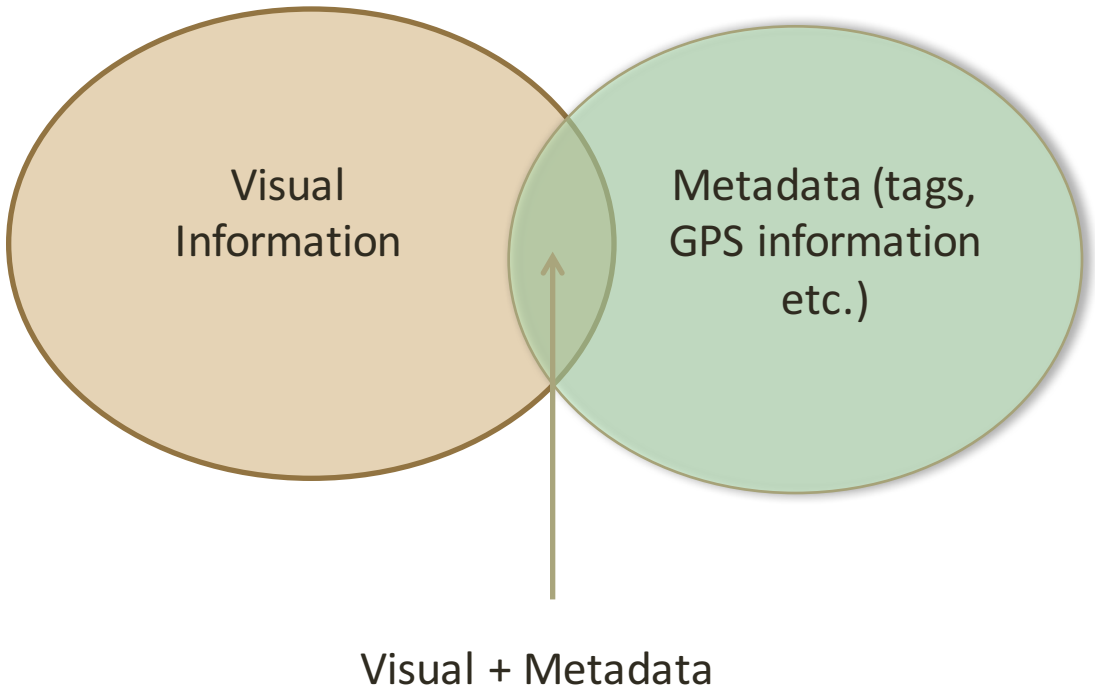
User-generated multimedia contents depict individual experiences or collective activities

- *What is an Event?*
  - A real world happening to Who?, What?, When? and Where?
  - An event is planned by people attended by people and related media are also captured by people

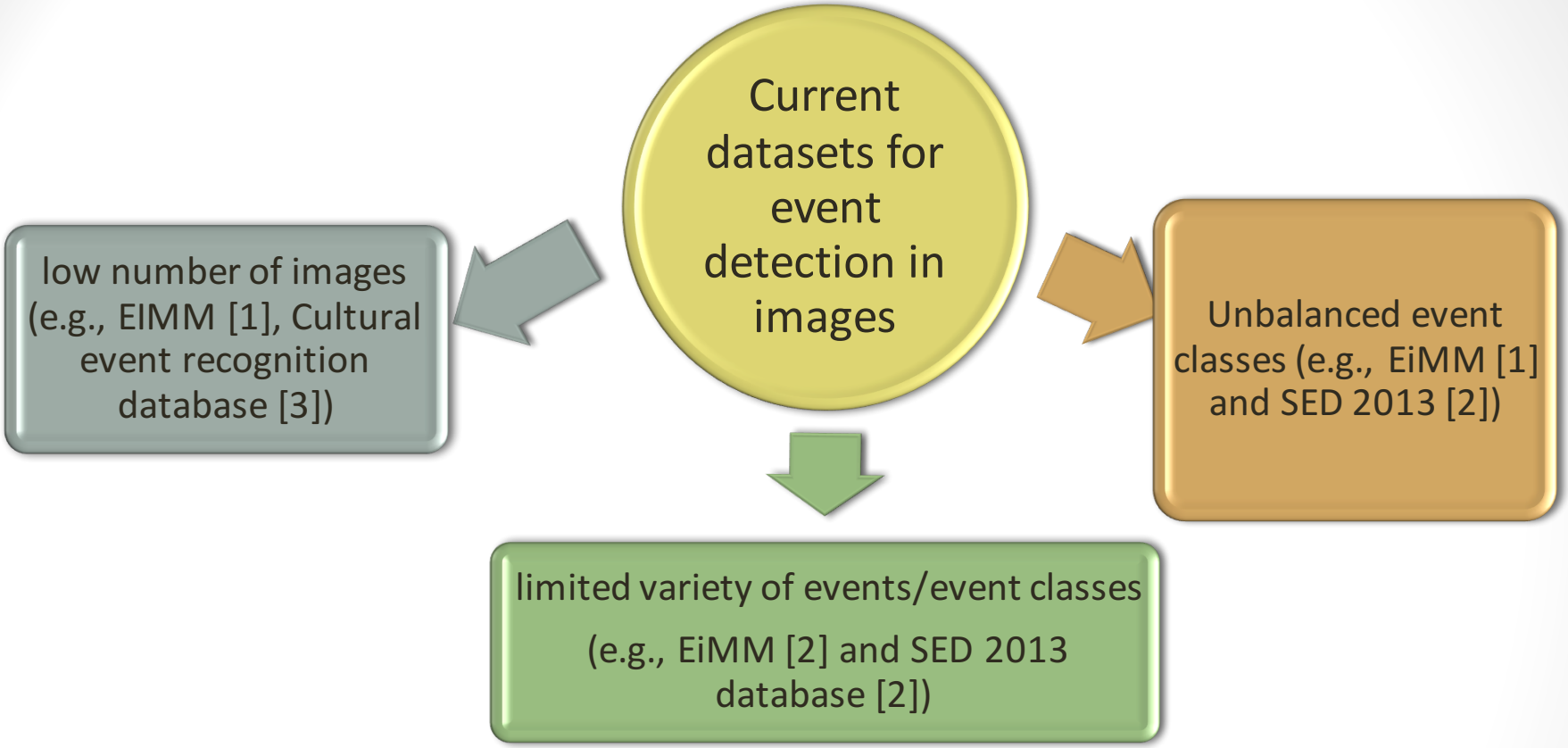
Personal experiences

Collective activities

Event Detection in Images: State-of-the-art



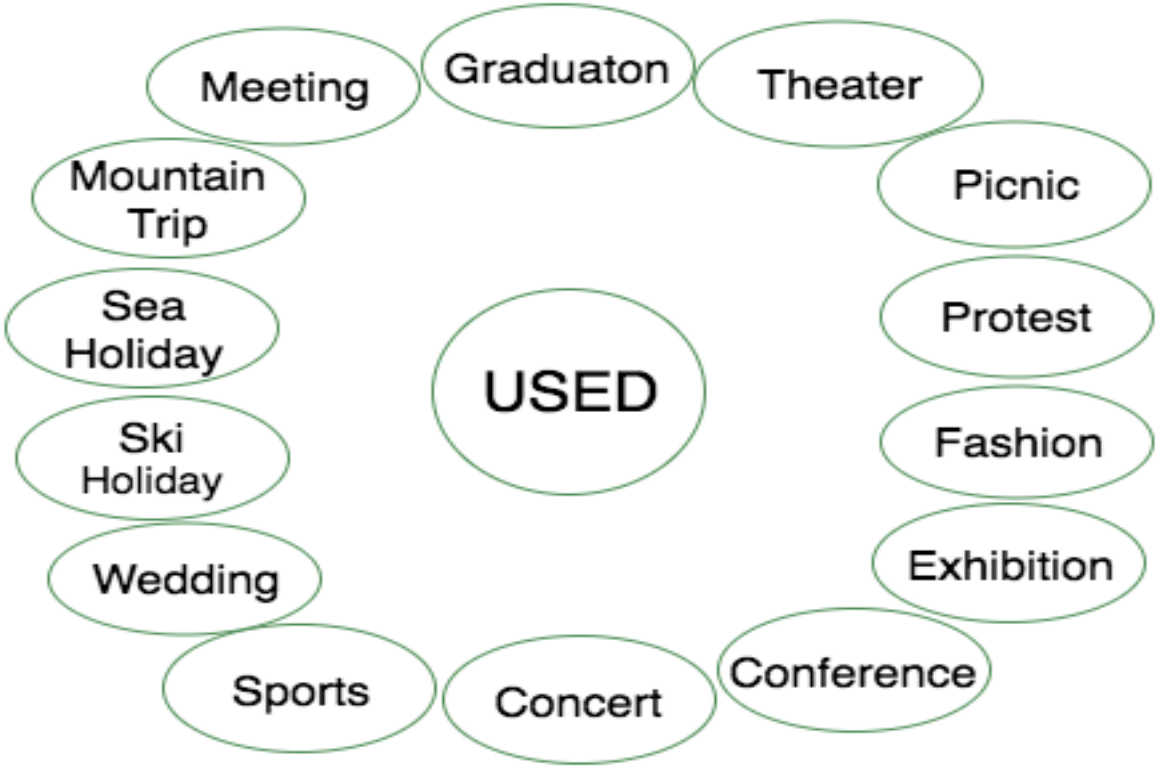
Benchmark Datasets: State-of-the-art



1. R. Mattivi et al. . Exploitation of time constraints for (sub-) event recognition. In Proceedings of the 2011 joint ACM workshop on Modeling and representing events, pages 7(12). ACM, 2011..
2. T. Reuter et al. . Social event detection at mediaeval 2013: Challenges, datasets, and evaluation. In MediaEval Workshop, 2013..
3. S. Escalera et al. . ChaLearn Looking at People 2015: Apparent Age and Cultural Event Recognition Datasets and Results, ICCV 2015

USED: A large Scale Social Event Detection Dataset

- A large collection of images
  - Covers 14 different events classes
- A balanced dataset
  - Equal number of images in each class (35,000)



Event-classes in USED Dataset

USED: A large Scale Social Event Detection Dataset

- Diversity in contents
  - Indoor Vs. outdoor
  - Group pictures Vs. Single portrait
  - Images of key-moments in an event
  - Multi-cultural
  - Outliers and borderline cases are manually removed

Some sample images from wedding class

USED: A large Scale Social Event Detection Dataset

Comparisons with state-of-the-art datasets

- Existing datasets for Event Detection
  - Cultural Event Detection Dataset
  - EiMM
  - SED

Dataset Name	# Event-classes	Total Images	Min images in a class	Max. images in a class
EiMM	8 (social events)	13219	795	2253
SED	7	82213	342	71556
Cultural Events	50	11776	180-200 (Avg.)	180-200 (Avg.)
USED	14	**490000**	**35000**	**35000**

Comparisons of USED with other Datasets

Experimental Validation of USED

DISCOVERING EVENTS FROM SINGLE PICTURES USING A CONVOLUTIONAL NEURAL NETWORK

Validation/Experimental Setup

Parameters of a CNN (Alex net) pre-trained on ImageNet dataset [NIPS 2012]

Pre-training

Fine-tuned on newly collected datasets
Reduced overall learning rate
Increased learning rate of new layer
Momentum = .9
Weight Decay = .0005

Fine-tuning CNN

Classification



Preliminary Results

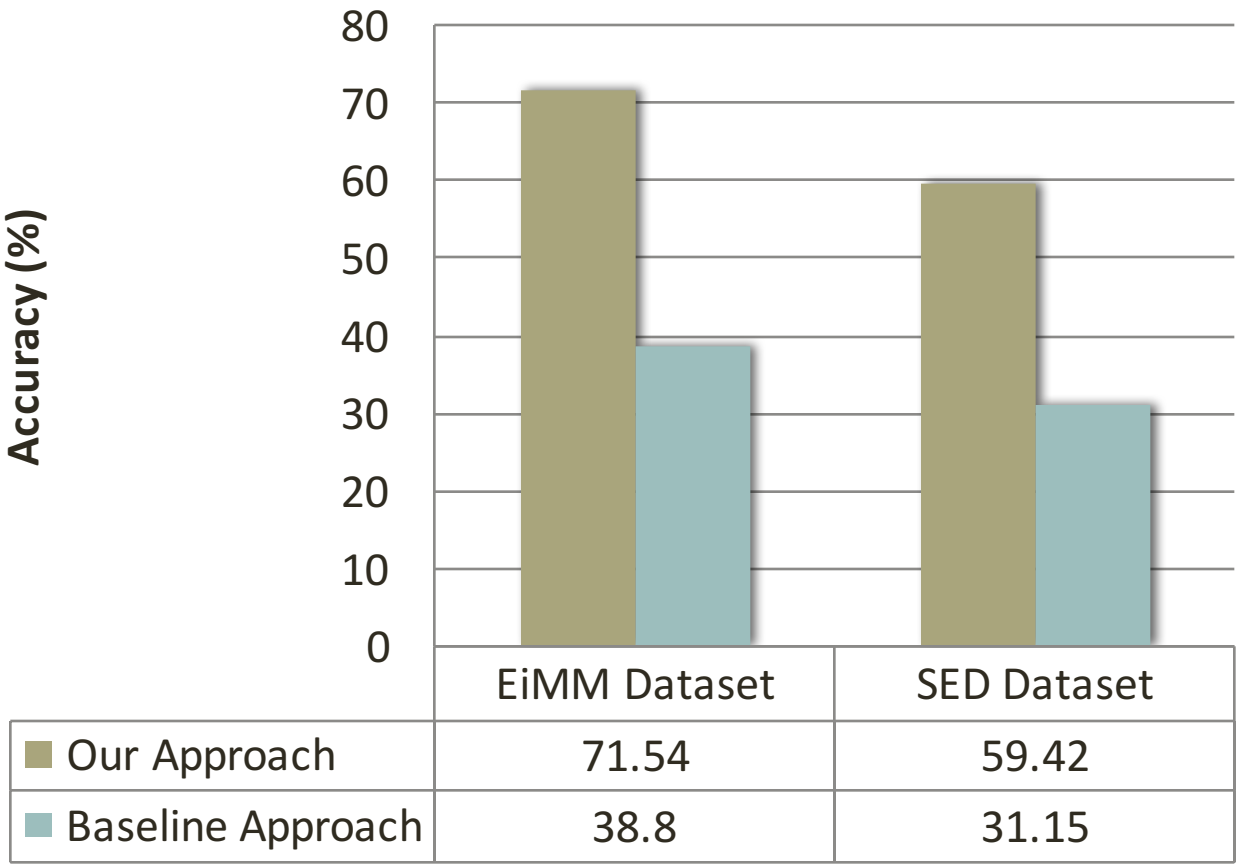
- Dataset
  - USED
- Data Assemblage
  - Training set = 20,000 images per class
  - Validation set = 7000 per class
  - Test set = 7000 images per class

Event Type	Accuracy	Event Type	Accuracy
Concert	74.20%	Conference	75.70%
Graduation	66.43%	Exhibition	58.54%
Meeting	78.70%	Fashion	65.43%
Mountain Trip	67.00%	Protest	74.58%
Picnic	54.42%	Sports	72.24%
Sea-holiday	74.24%	Theater	51.90%
Ski-holiday	48.00%		
Wedding	51.00%		

Results on USED dataset

Comparisons of a CNN trained on USED with Baseline Approaches

Comparison with Rosani et al., [IEEE TMM 2015]



	EiMM Dataset	SED Dataset
Our Approach	71.54	59.42
Baseline Approach	38.8	31.15

A. Rosani, G. Baoto, F. G.B. De Natale, "EventMask: a game-based framework for Event-saliency identification in Images", IEEE Transactions on Multimedia 2015

USED: A Large-scale Social Event Detection Dataset