Top Conferences' Paper Analysis

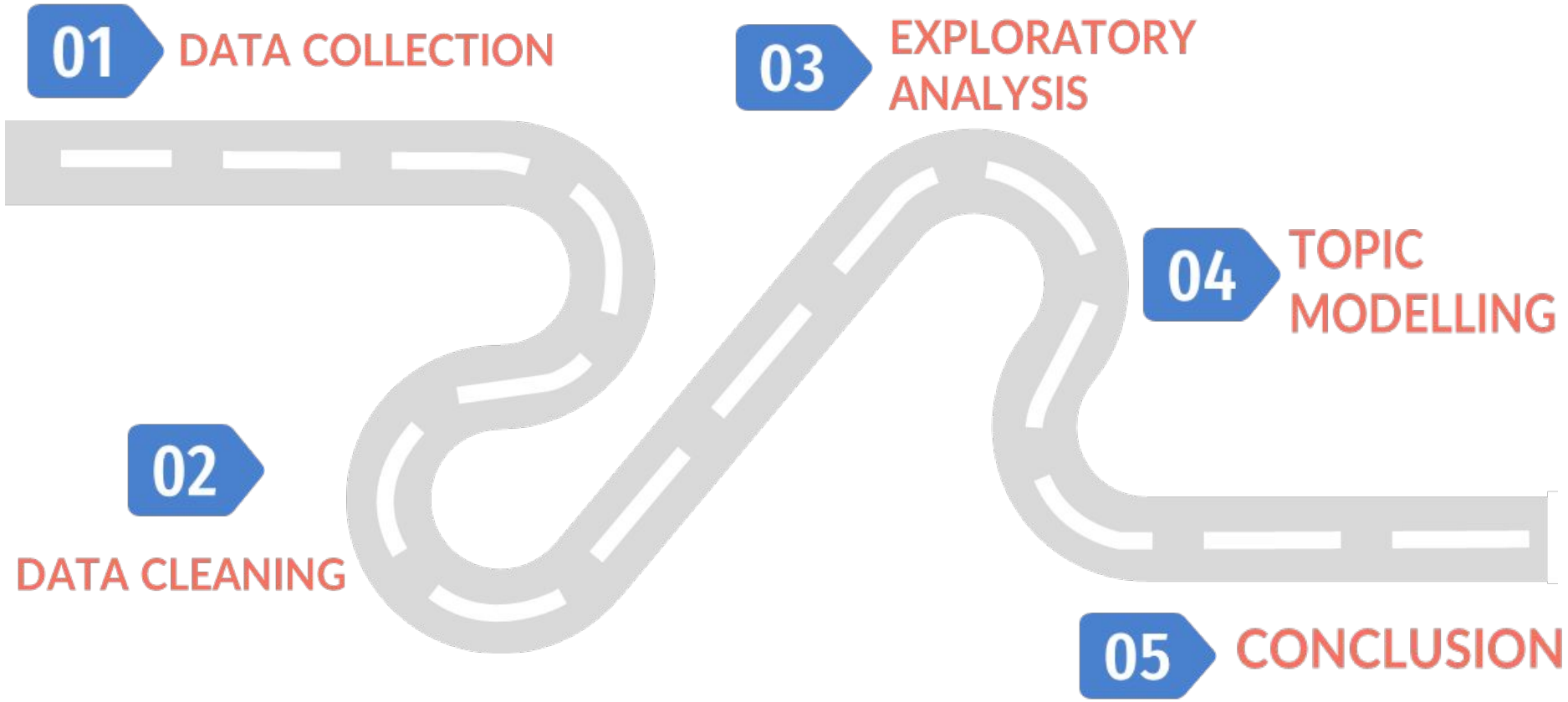
Group 16

Debaditya Basu
Rahul Rajaraman
Sambaran Ghosal
Sananya Majumder
Yuwei Zhang

Aim of the Project

- Collect and analyze data from top Machine Learning conferences.
- The main analyses that we focussed on are:
  1. Popularity of conferences based on citations
  2. Most active authors and institutes
  3. Trending research areas
  4. Recommend related papers

Project Roadmap



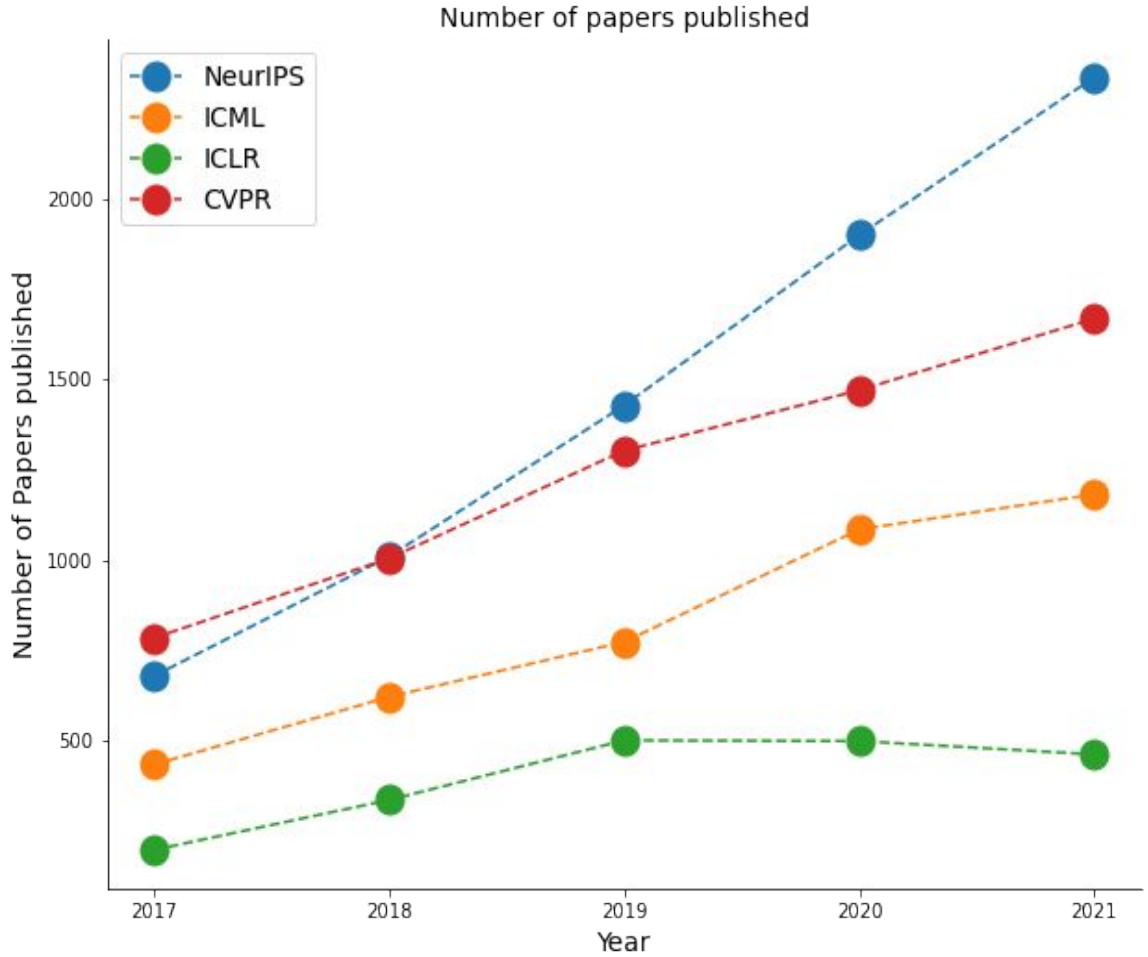
Dataset Collection

- We collected ~21,000 data from the following 4 conferences :
  CVPR (Conference on Computer Vision and Pattern Recognition)
  ICML (International Conference on Machine Learning)
  ICLR (International Conference on Learning Representations)
  NeurIPS (Neural Information Processing Systems)
- The dataset was constructed using Web Scraping.
- Citations of respective papers were collected using Scholarly library.
- The dataset obtained consists of the following fields:

YEAR	TITLE	AUTHORS	ABSTRACT	AFFILIATIONS	CITATIONS
2017	Decoupled Neural Interfaces using Synthetic Gradients	Max Jaderberg, Wojciech Czarnecki, Simon Osindero, Oriol Vinyals, Alex Graves, David Silver, koray kavukcuoglu	Training directed neural networks typically requires forward-propagating data through a computation graph, followed by backpropagating error signal, to produce	DeepMind', 'DeepMind', 'DeepMind', 'DeepMind', 'DeepMind', 'Google DeepMind', 'DeepMind'	261

Let’s get to the analysing part

Number of Publications per Year



- CVPR and NeurIPS have most number of accepted papers
- ICML has shown a surge in the number of papers accepted 2020 onwards

Citation Analysis

The more number of citations, the more impactful a paper is.

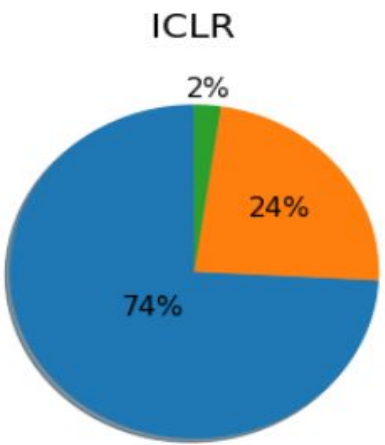
Neurips 2017

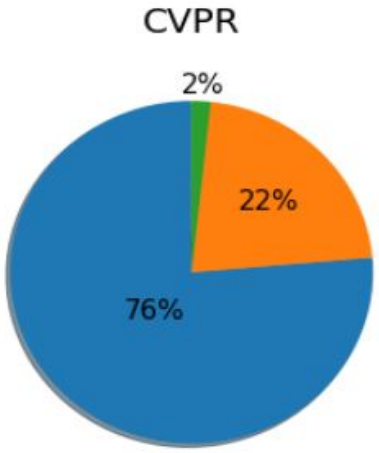
TITLE	citations
Attention is All you Need	31339
Improved Training of Wasserstein GANs	5834
GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium	3971
Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles	2058
Unsupervised Image-to-Image Translation Networks	1882
Self-Normalizing Neural Networks	1733
Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments	1663
Hindsight Experience Replay	1258
A simple neural network module for relational reasoning	1201
Deep Sets	1189

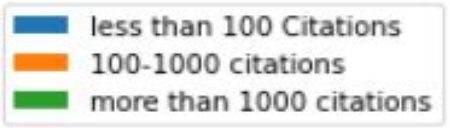
- **Attention is All you Need** was a groundbreaking work in Deep Learning which is supported by the large number of citations

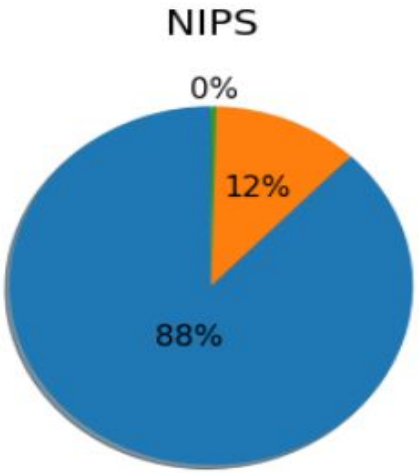
- Most of the papers with high citations are in the areas of **GAN** (Generative Adversarial Network), **neural networks, reinforcement learning**

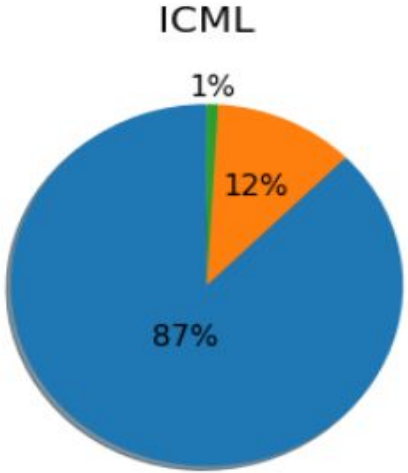
Citation Distributions of Conferences





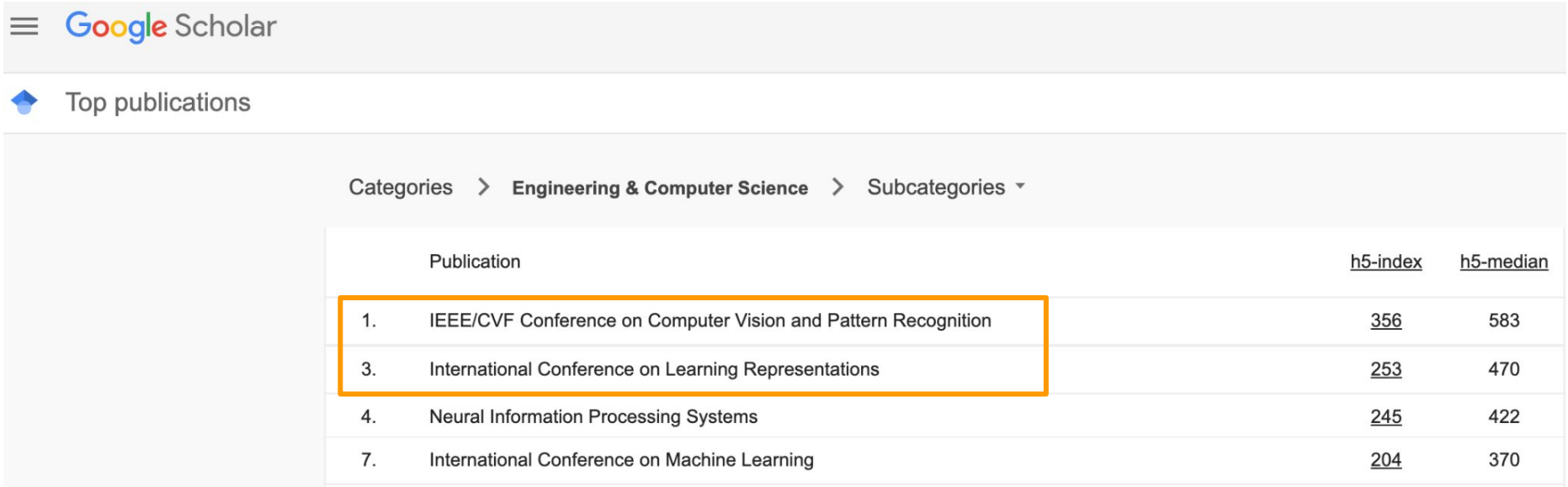




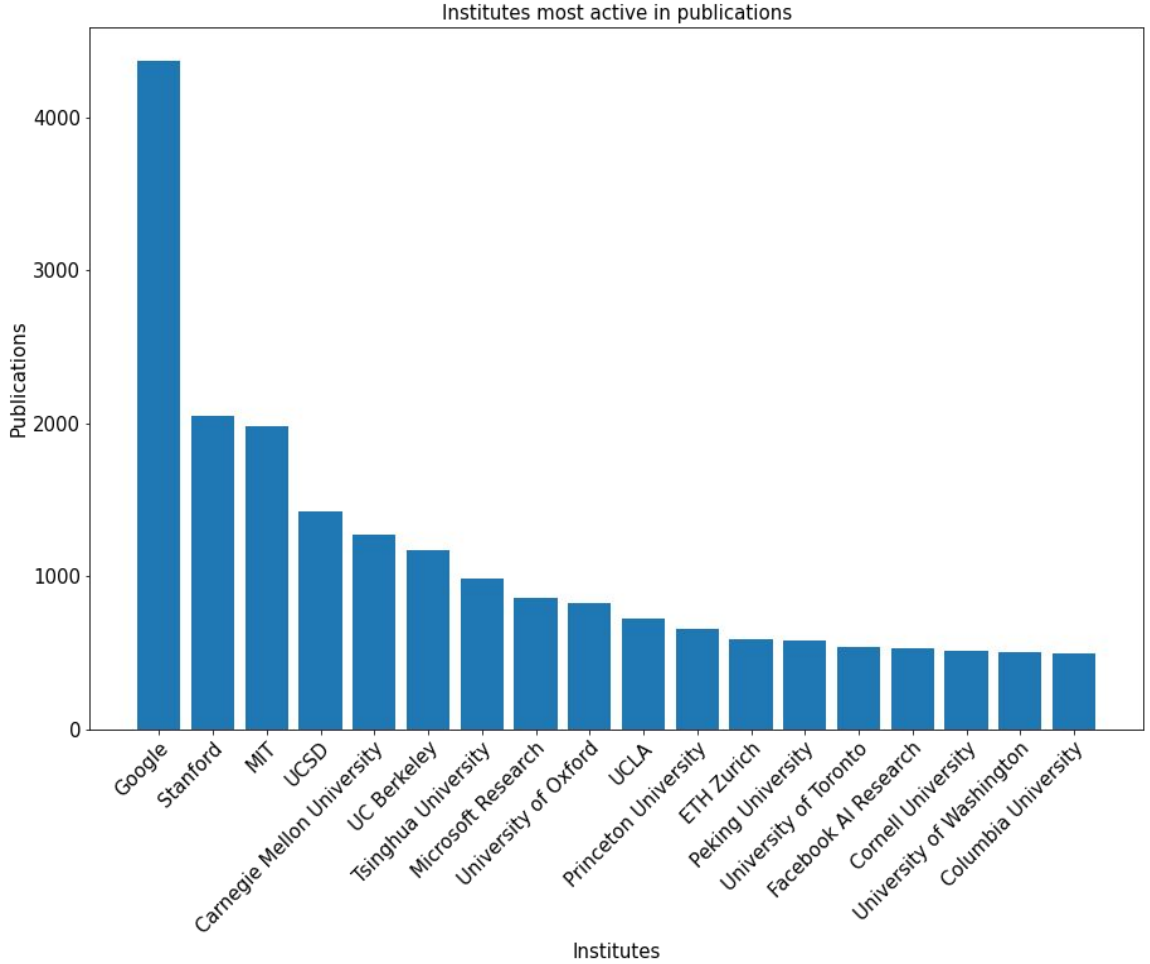


- ICLR and CVPR have the highest number of cited papers

Supported by Google Scholar



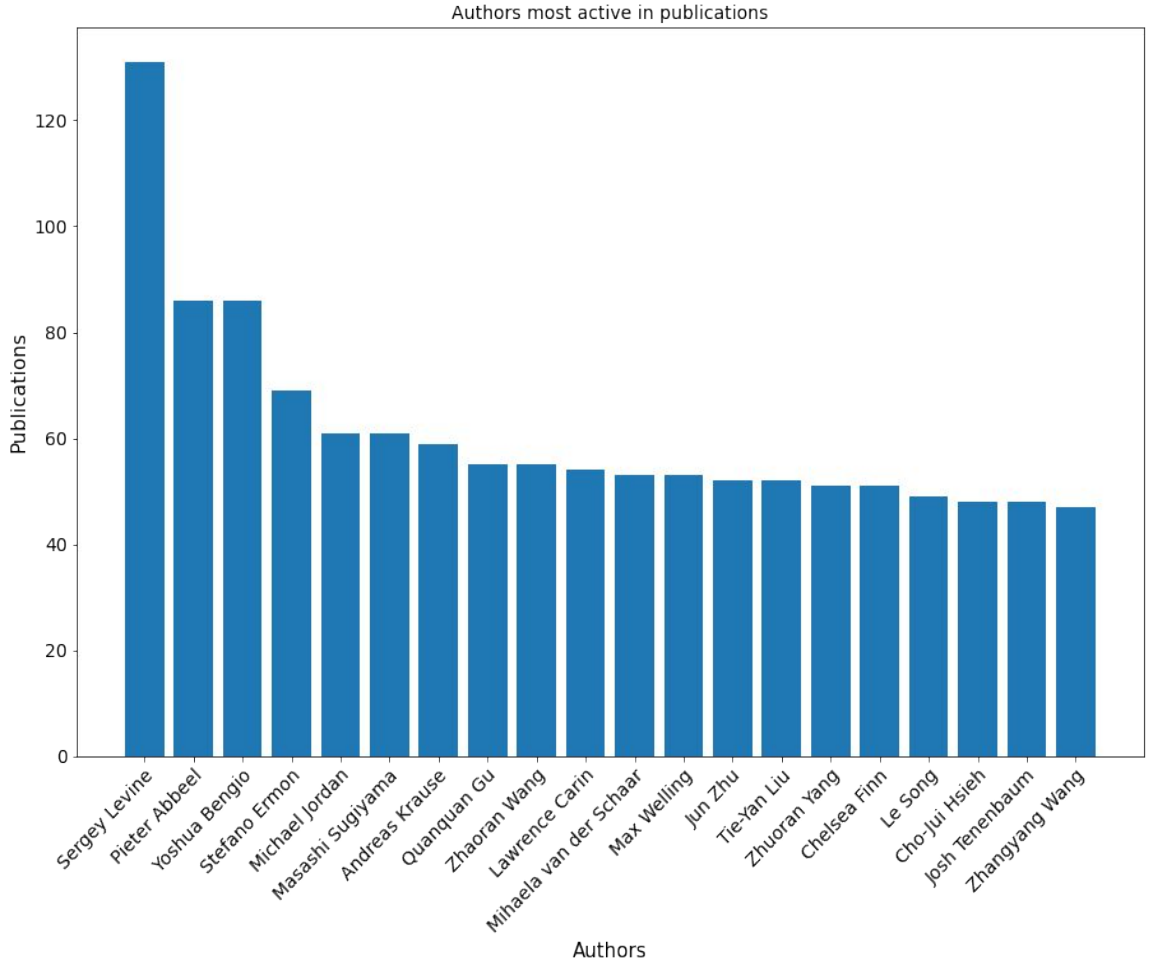
Which institute is publishing more ?



- Google leads the way
- UC system doing good
- Top ranked universities are up there in research

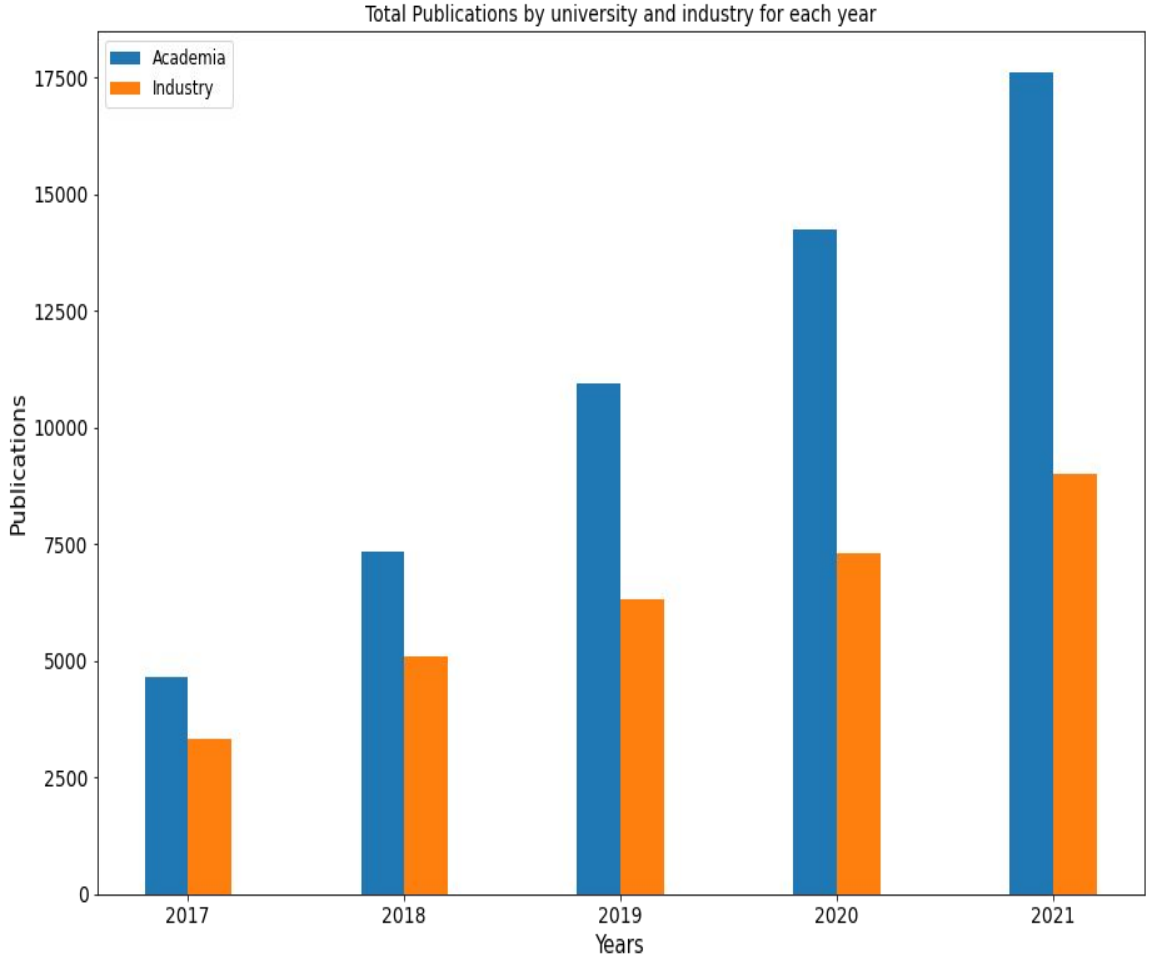
Institute	Publications
Google	4371
Stanford	2048
MIT	1984
UCSD	1427
CMU	1269

Which author is publishing more ?



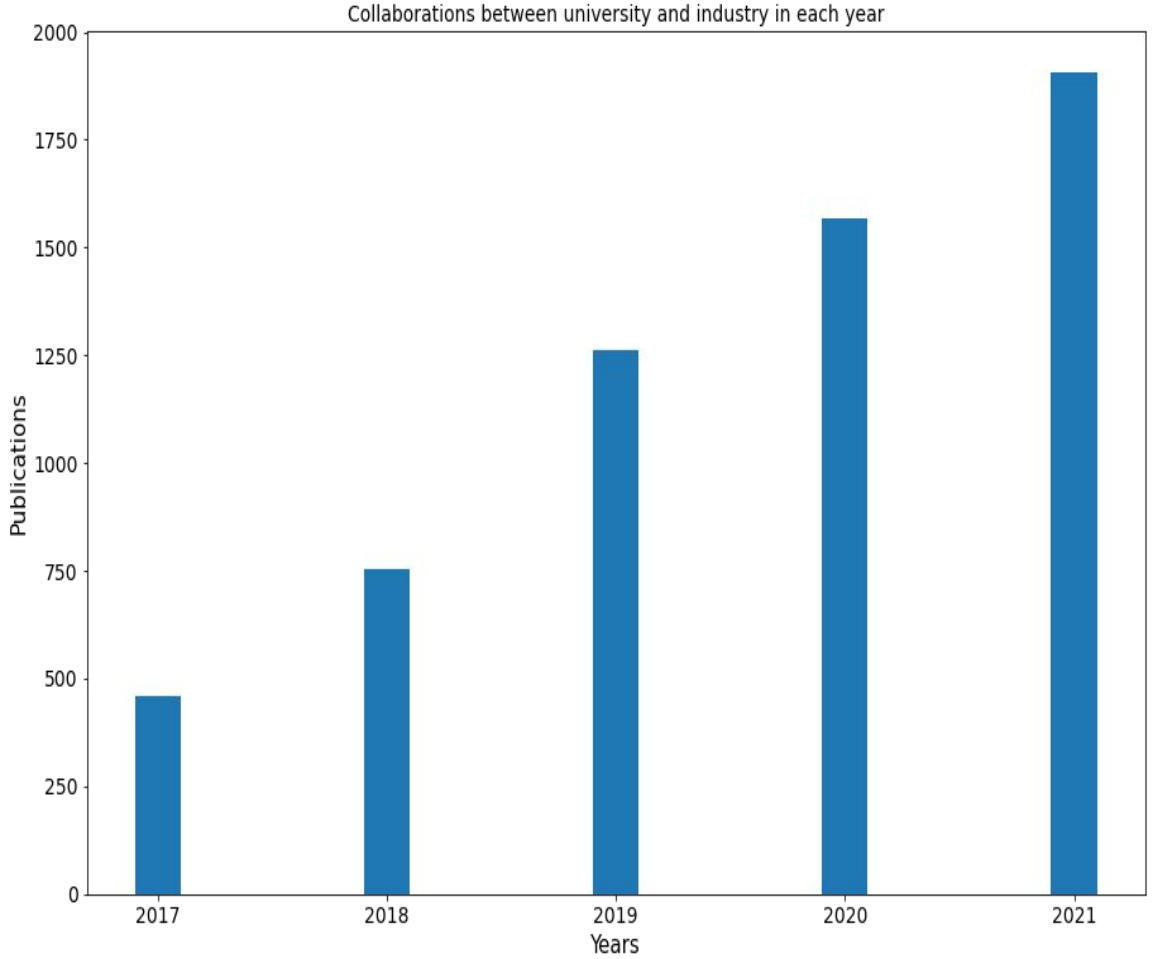
Author	Publications
Sergey Levine	131
Pieter Abbeel	86
Yoshua Bengio	86
Stefano Ermon	69
Michael Jordan	61

University vs Industry



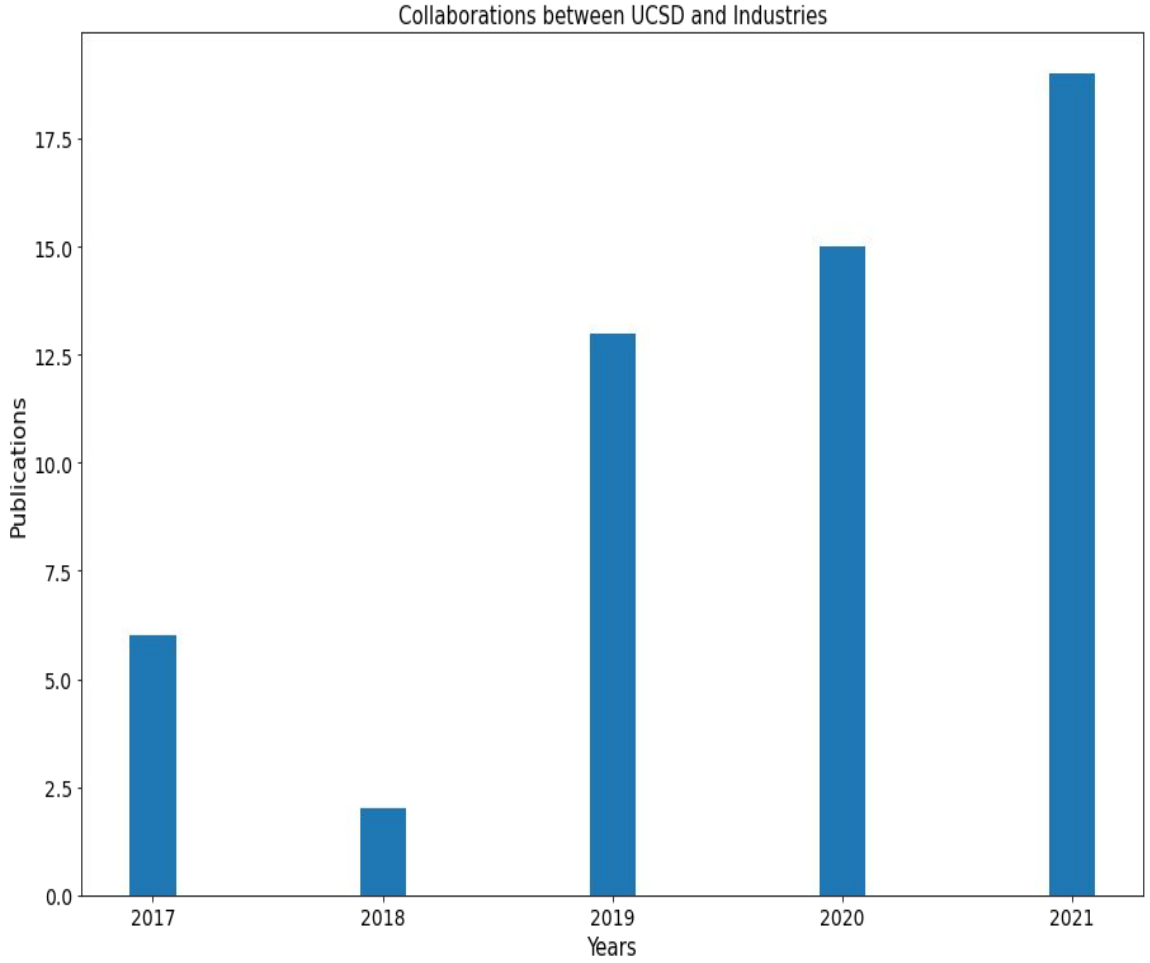
- Academia is showing an increase in research.
- Industry research also increasing but less compared to universities

Let's work together?

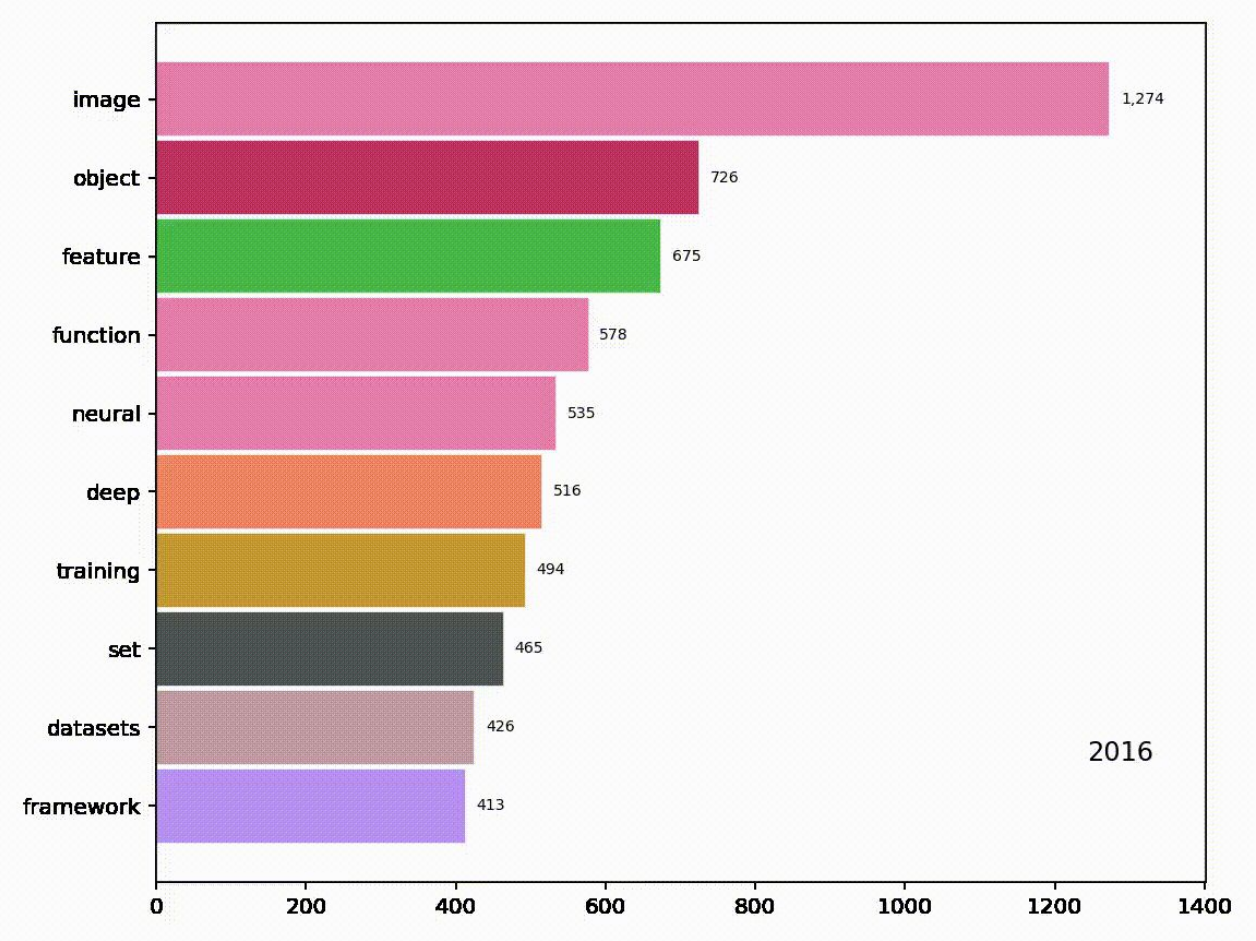


- Increasing collaborative work between universities and industry

UCSD Shows Increase in Collaboration

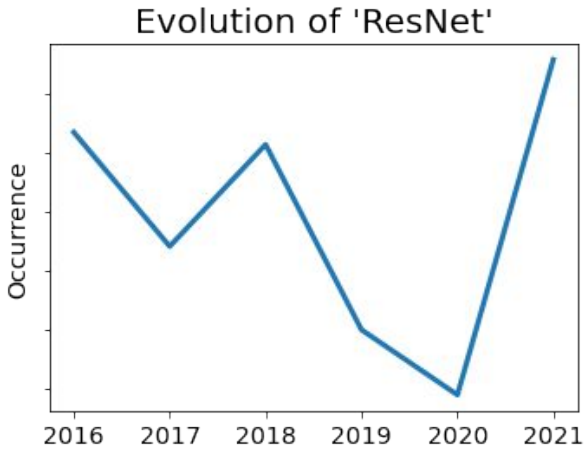


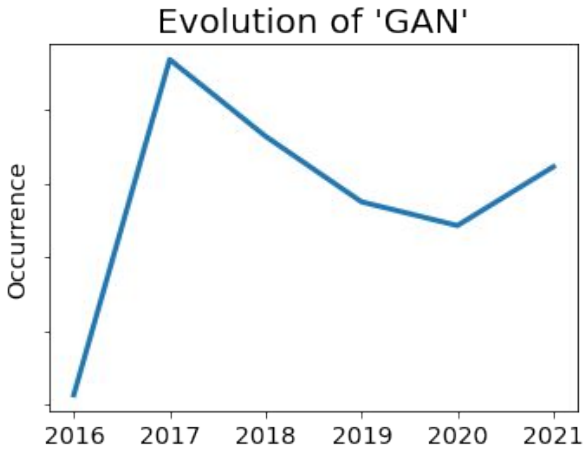
Trending Phrases in Abstracts (all conferences)

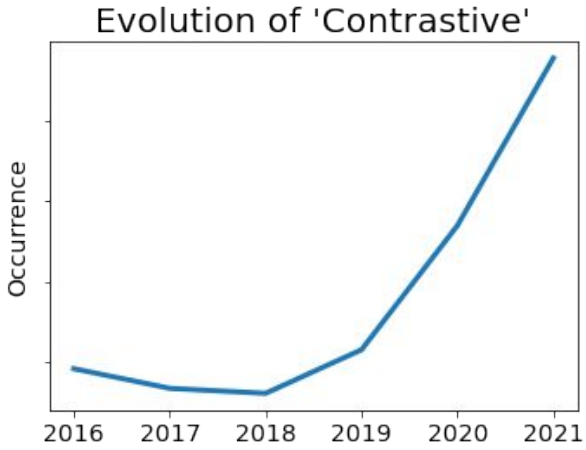


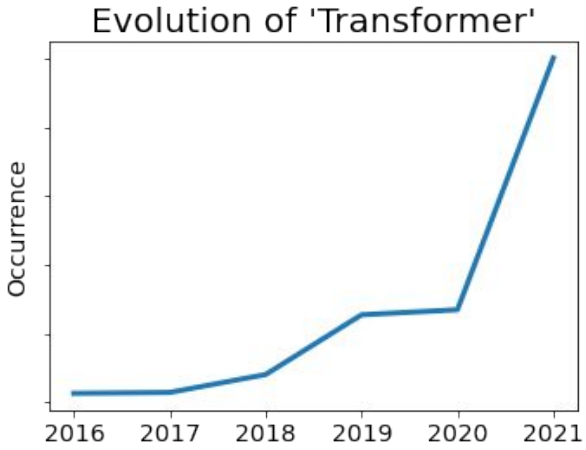
- "**neural**", "**image**" and "**training** " are the most common words
- "**deep**" is receiving decreasing attention
- Good quality of learning accepts more focus ("**representation**", "**feature**", "**distribution**")

Trending Phrases in Abstracts (all conferences)









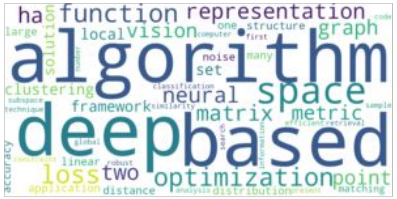
- "**ResNet**" and "**GAN**" are no longer the hottest topics
- More occurences of "**Contrastive**" and "**Transformer**"
- "**ResNet**" increases from 2020 to 2021 probably due to comparison to "**Transformer**"

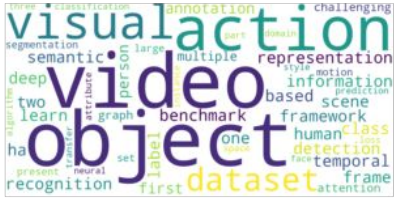
Topic Analysis of Abstracts (CVPR)

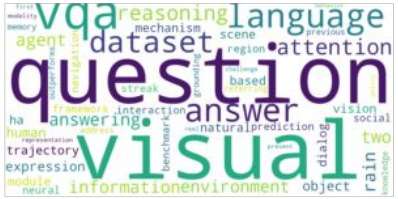
Abstracts → (input) → LDA → (generate) → Topics

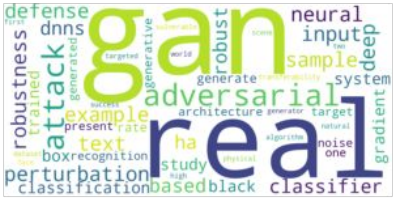
Topic = A bag of most representative words

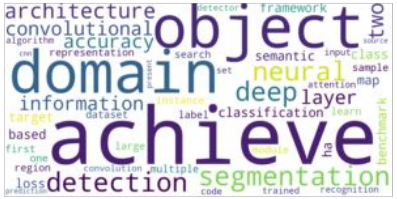
8 word clouds on CVPR, each corresponding to a topic

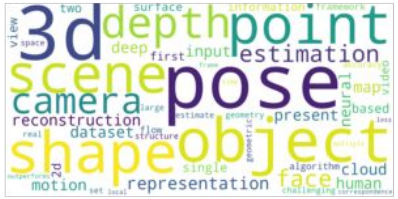












Topic Analysis of Abstracts (CVPR)

In LDA:
Article = Distribution over topics

Coherent Parametric Contours for Interactive Video Object Segmentation

Yao Lu[1], Xue Bai[2], Linda Shapiro[1], and Jue Wang[2]

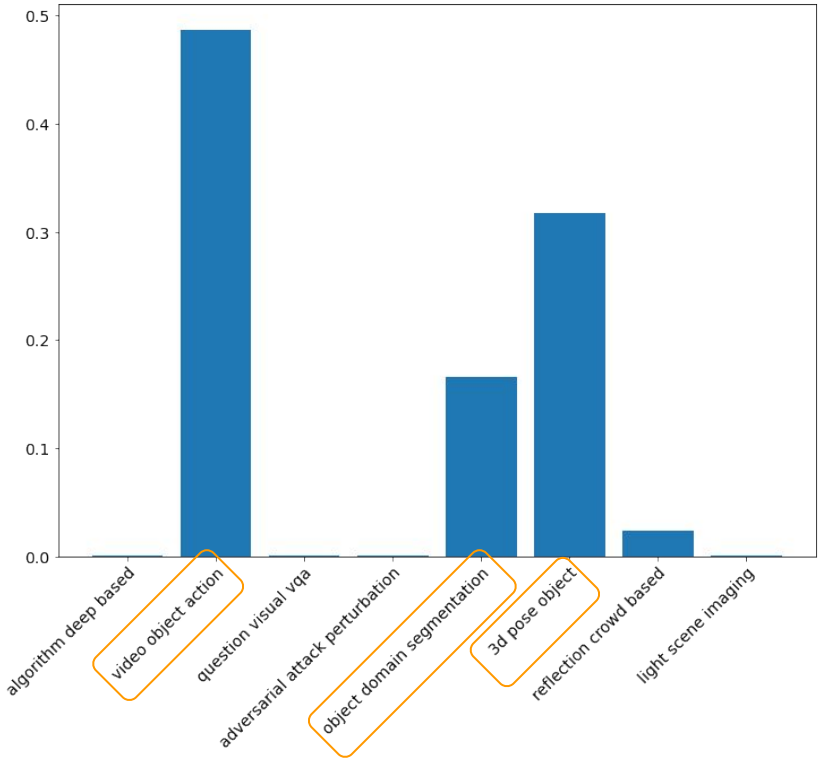
[1]University of Washington , {luyao, shapiro}@cs.washington.edu
[2]Adobe , {xubai, juewang}@adobe.com

Abstract

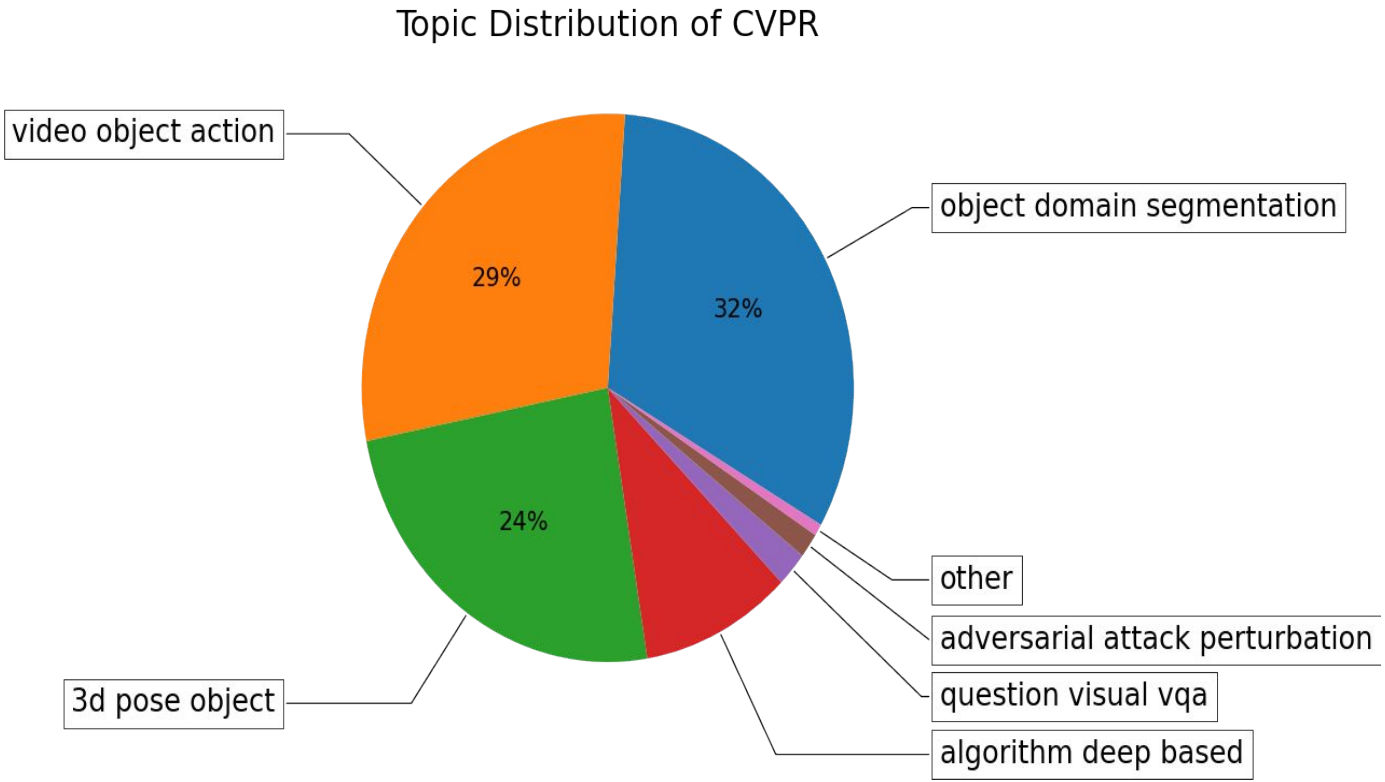
Interactive video segmentation systems aim at producing sub-pixel-level object boundaries for visual effect applications. Recent approaches mainly focus on using sparse user input (i.e. scribbles) for efficient segmentation; however, the quality of the final object boundaries is not satisfactory for the following reasons: (1) the boundary on each frame is often not accurate; (2) boundaries across adjacent frames wiggle around inconsistently, causing temporal flickering; and (3) there is a lack of direct user control for fine tuning.

We propose Coherent Parametric Contours, a novel video segmentation propagation framework that addresses all the above issues. Our approach directly models the object boundary using a set of parametric curves, providing direct user controls for manual adjustment. A spatio-temporal optimization algorithm is employed to produce object boundaries that are spatially accurate and temporally stable. We show that existing evaluation datasets are limited and demonstrate a new set to cover the common cases in professional rotoscoping. A new metric for evaluating temporal consistency is proposed. Results show that our approach generates higher quality, more coherent segmentation results than previous methods.

Distribution over topics for this paper

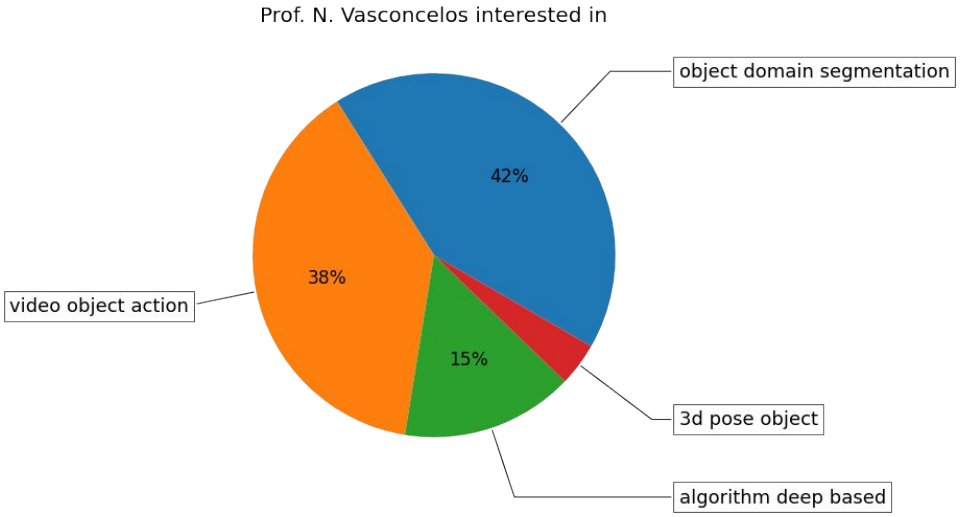


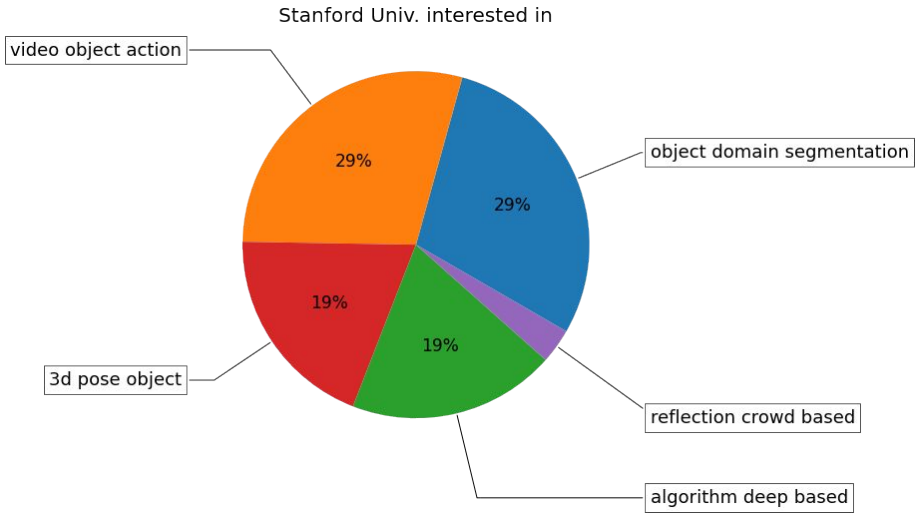
Topic Analysis of Abstracts (CVPR)



- "**3d**", "**video**" and "**object detection**" are the most popular topics
- The other major topics include "**robustness**", "**multi-modal**" and "**representation learning**"

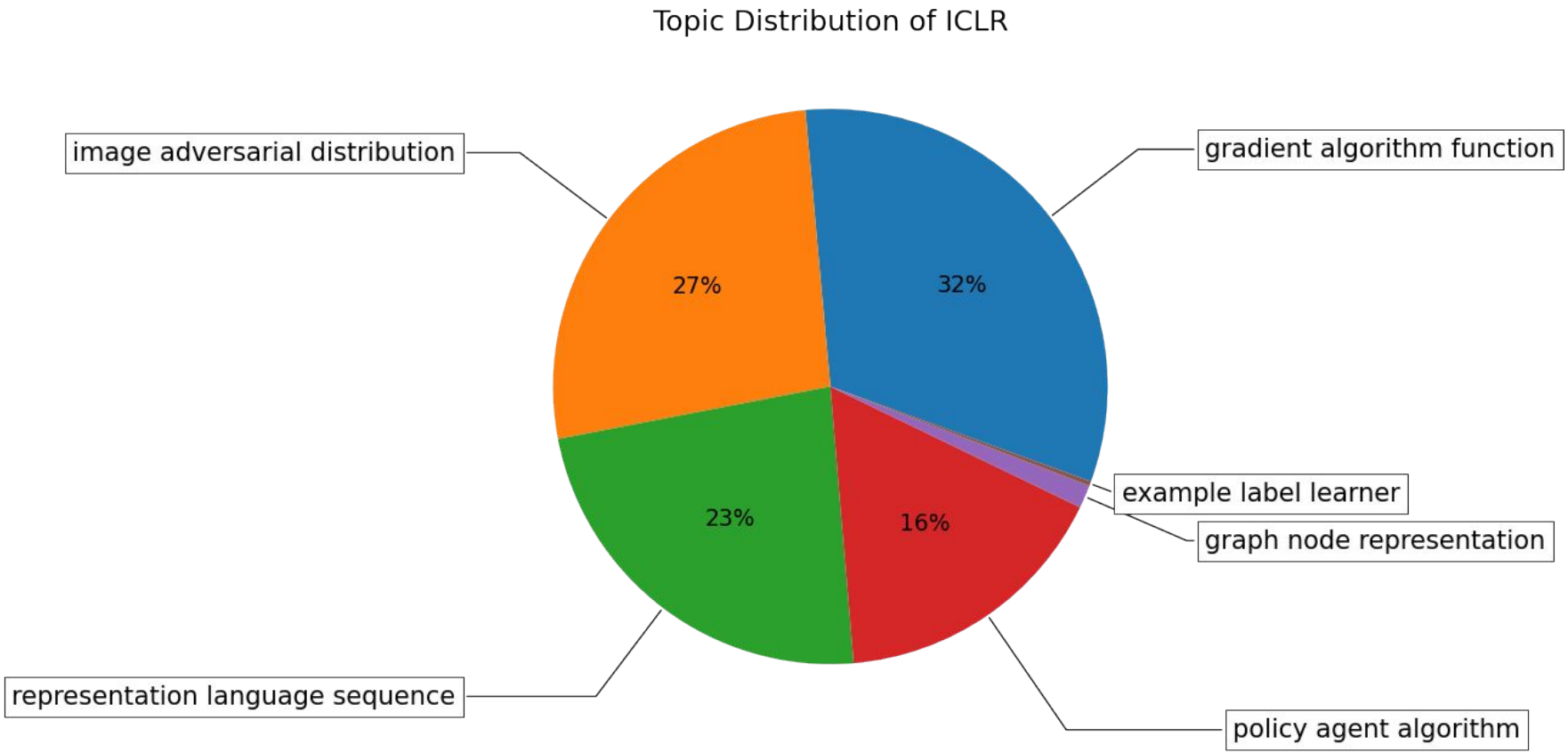
Find What They are Interested in (CVPR)...



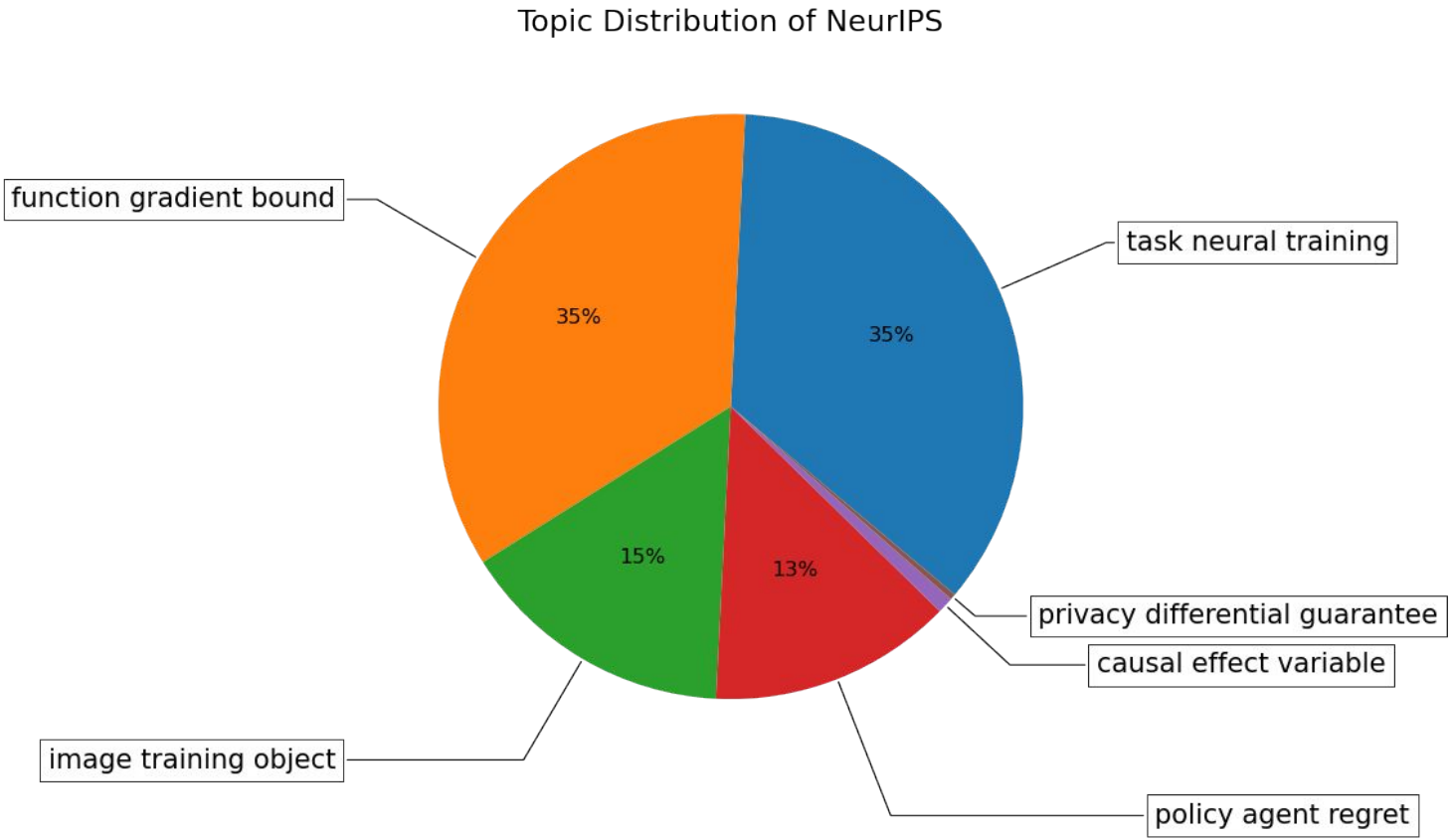


- Professor Nuno Vasconcelos is interested in "**object domain segmentation**".
- Stanford University is interested in "**object domain segmentation**" and "**video object action**".

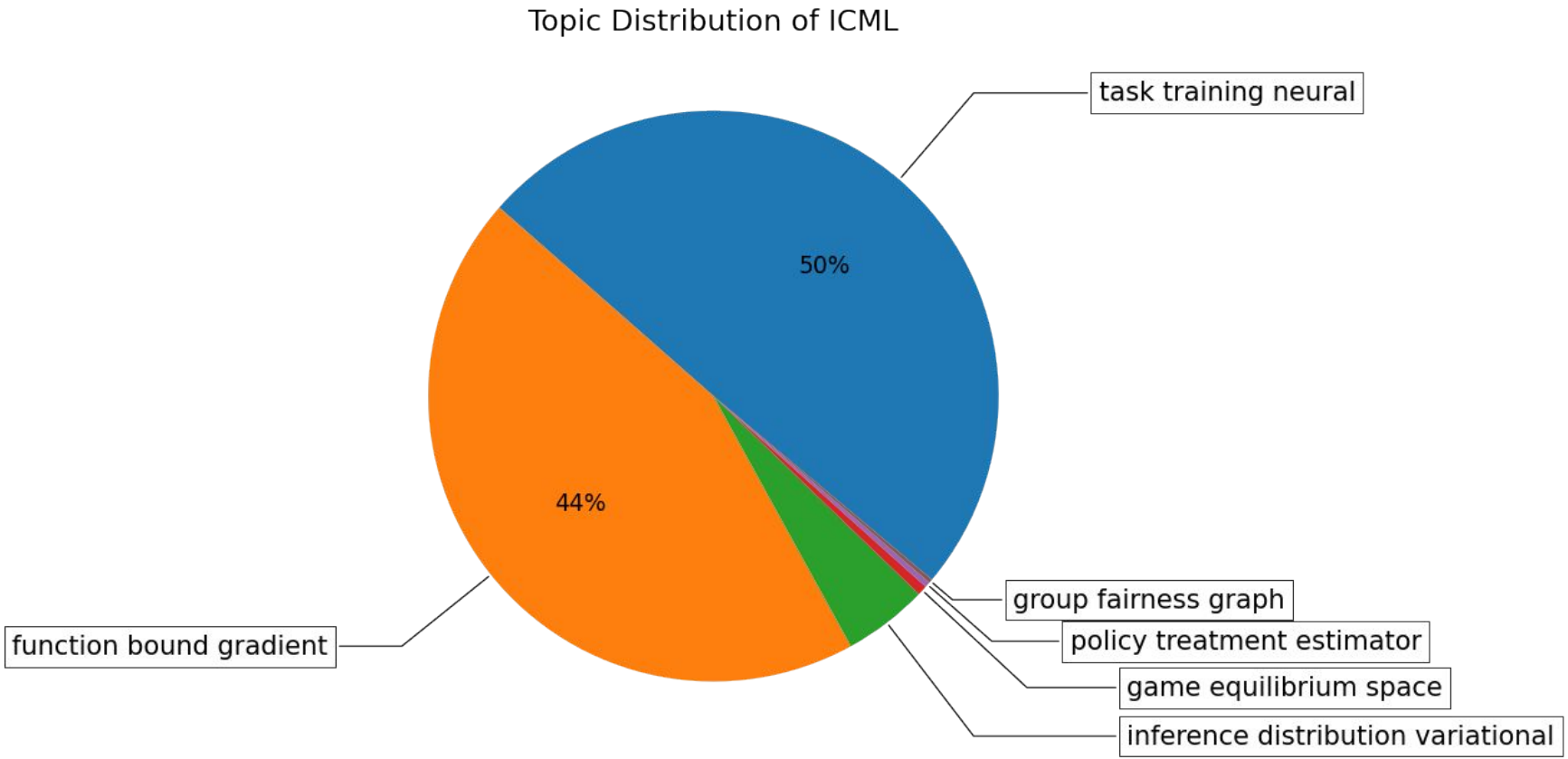
ICLR


Topic Distribution of ICLR
image adversarial distribution
gradient algorithm function
example label learner
graph node representation
representation language sequence
policy agent algorithm
27%
32%
23%
16%

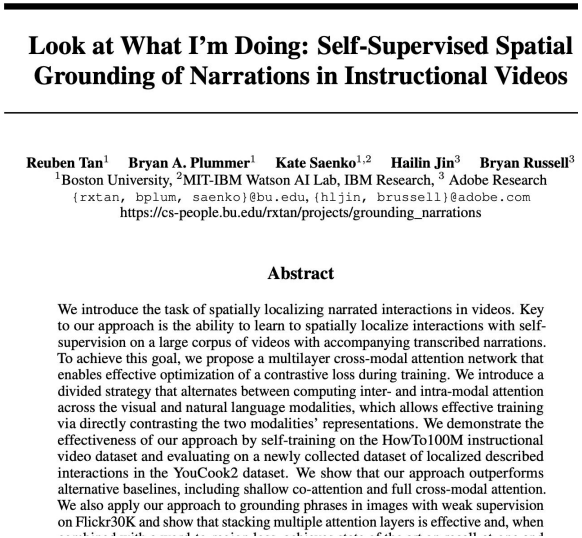
NeurIPS


Topic Distribution of NeurIPS
function gradient bound
task neural training
privacy differential guarantee
causal effect variable
image training object
policy agent regret
35%
35%
15%
13%

ICML


Topic Distribution of ICML
task training neural
50%
44%
function bound gradient
group fairness graph
policy treatment estimator
game equilibrium space
inference distribution variational

Recommender System



Look at What I'm Doing: Self-Supervised Spatial Grounding of Narrations in Instructional Videos

Reuben Tan[1] **Bryan A. Plummer**[1] **Kate Saenko**[1,2] **Hailin Jin**[3] **Bryan Russell**[3]
[1]Boston University, [2]MIT-IBM Watson AI Lab, IBM Research, [3] Adobe Research
{rxtan, bplum, saenko}@bu.edu, {hljin, brussell}@adobe.com
https://cs-people.bu.edu/rxtan/projects/grounding_narrations

Abstract

We introduce the task of spatially localizing narrated interactions in videos. Key to our approach is the ability to learn to spatially localize interactions with self-supervision on a large corpus of videos with accompanying transcribed narrations. To achieve this goal, we propose a multilayer cross-modal attention network that enables effective optimization of a contrastive loss during training. We introduce a divided strategy that alternates between computing inter- and intra-modal attention across the visual and natural language modalities, which allows effective training via directly contrasting the two modalities' representations. We demonstrate the effectiveness of our approach by self-training on the HowTo100M instructional video dataset and evaluating on a newly collected dataset of localized described interactions in the YouCook2 dataset. We show that our approach outperforms alternative baselines, including shallow co-attention and full cross-modal attention. We also apply our approach to grounding phrases in images with weak supervision on Flickr30K and show that stacking multiple attention layers is effective and, when

Word Embedding → Title Similarity

+

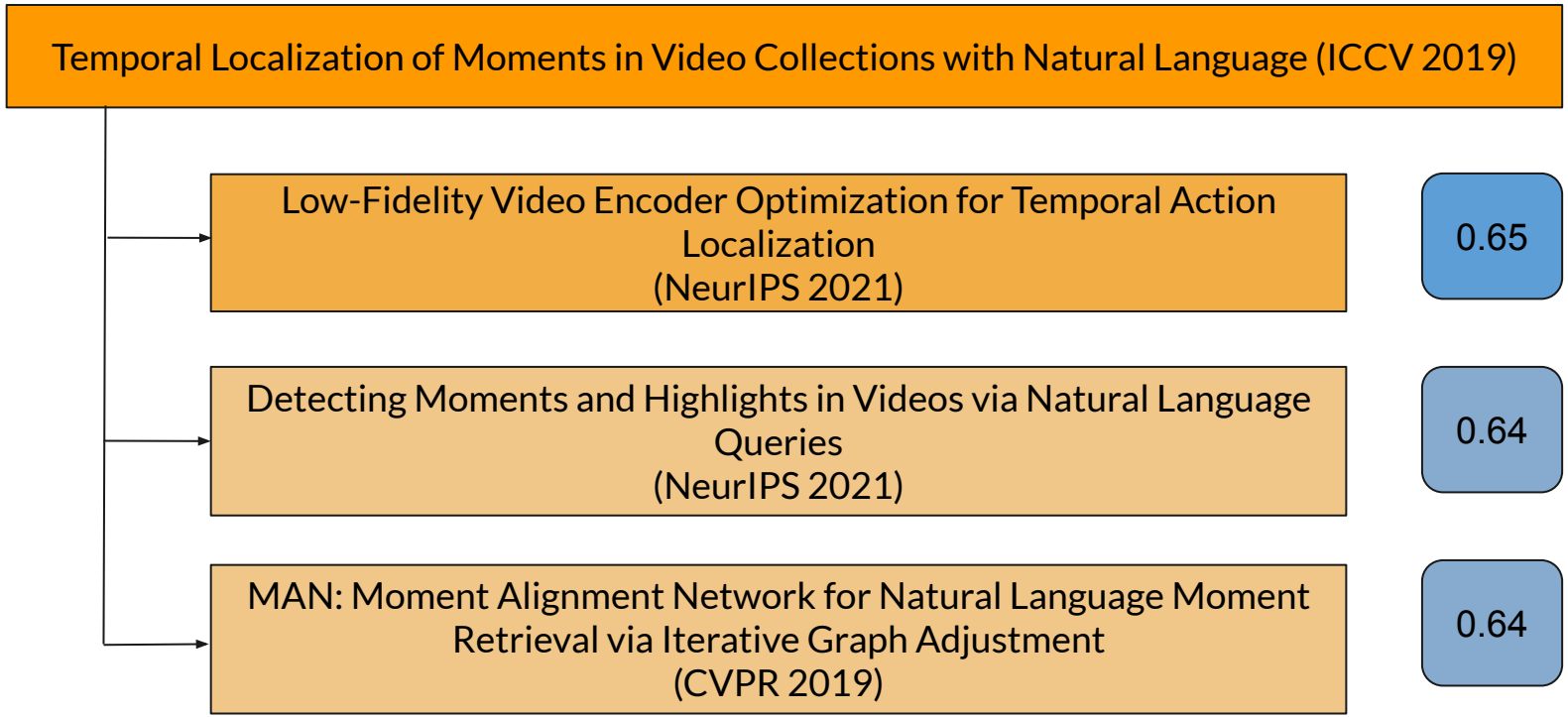
LDA → Abstract Similarity

=

Similarity Score

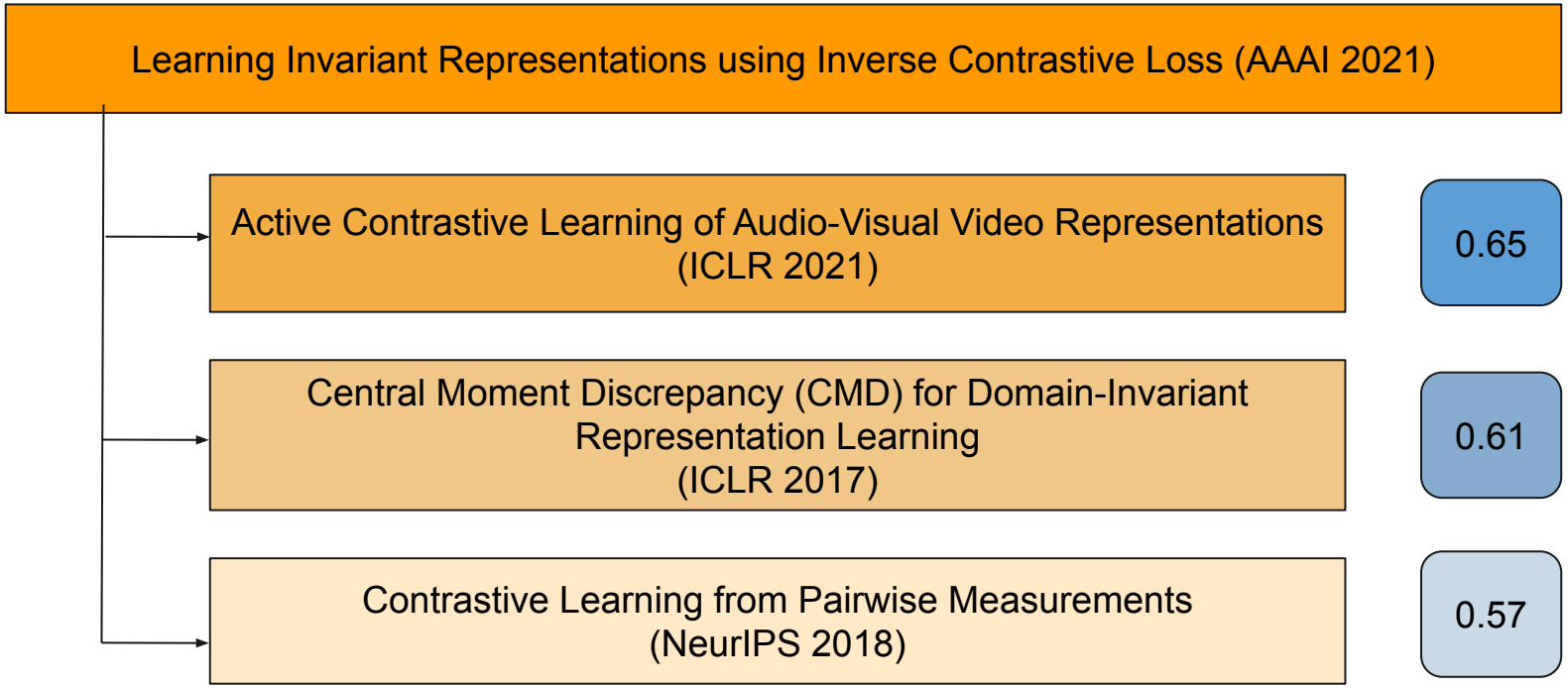
Recommender System

Example 1:



Recommender System

Example 2:



Conclusion

- Machine learning research is becoming more popular through years.
- CVPR and ICLR provide more relevant papers with more average citations.
- Collaborations between academia and industry are increasing.
- Better representation learning is attracting more attention in recent years (e.g. Contrastive and Transformer).
- We propose a method to recommend closely related works for researchers.

Thank you!

Questions?