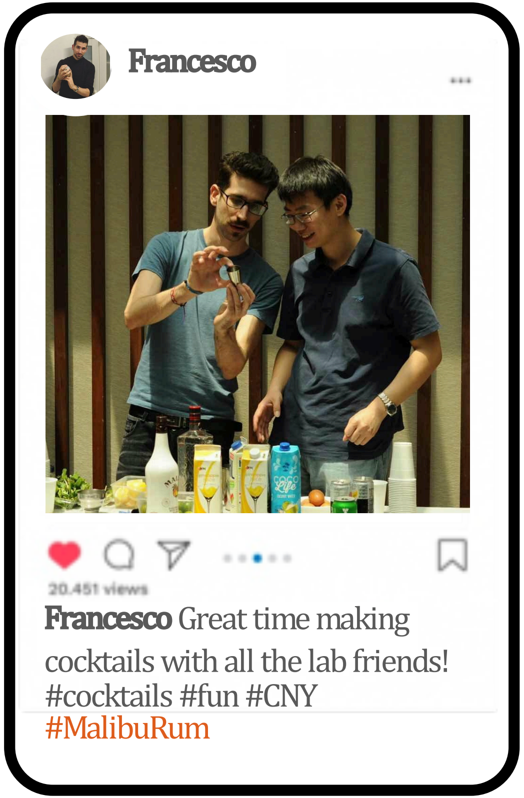
Beyond the Product: Discovering Image Posts for Brands in Social Media

Francesco Gelli*, Tiberio Uricchio†, Xiangnan He*, Alberto Del Bimbo†, Tat-Seng Chua*
*National University of Singapore, †Università degli Studi di Firenze

Content Discovery for Brands

- Recent trend: discovering actionable UGC (User Generated Content) for a brand
- Current solutions solely rely on brand-defined hashtags
- Can we discover actionable UGC by visual content only?



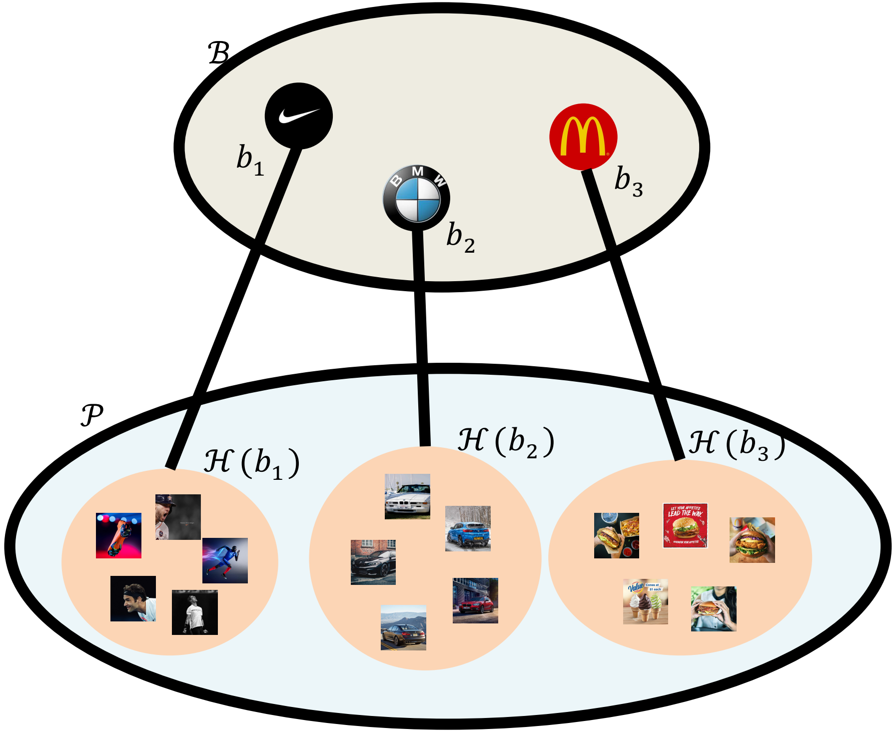
Problem Formulation

- $\mathcal{B} = \{b_1, \dots, b_N\}$: set of brands
- $\mathcal{P} = \{p_1, \dots, p_M\}$: set of posts
- $\mathcal{H}(b)$: posting history of brand b
- Goal: learn $f: \mathcal{B} \times \mathcal{P} \to \mathbb{R}$ s.t. for post p_x of brand $b \in \mathcal{B}$:

$$f(b, p_x) > f(b, p_y)$$

  where p_y is a new post of any other brand $\hat{b} \neq b$

- For example: $f($ M ,  $) > f($ M ,  $)$



Challenges

Two challenges make this problem different from traditional retrieval applications.

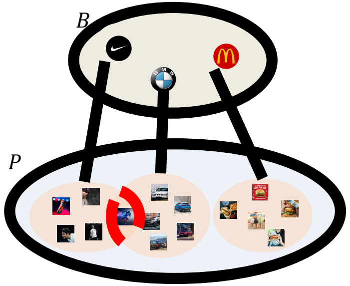
- **Inter-brand similarity**: subtle differences between posts by competitor brands

Timberland
Carlsberg Carlsberg Timberland

Red Bull
Coca Cola Coca Cola Red Bull


Emirates
Emirates Air France Air France


- **Brand-post sparsity**: posts are rarely shared among different brands. Different from recommendation tasks



Personalized Content Discovery (PCD)

Inputs:

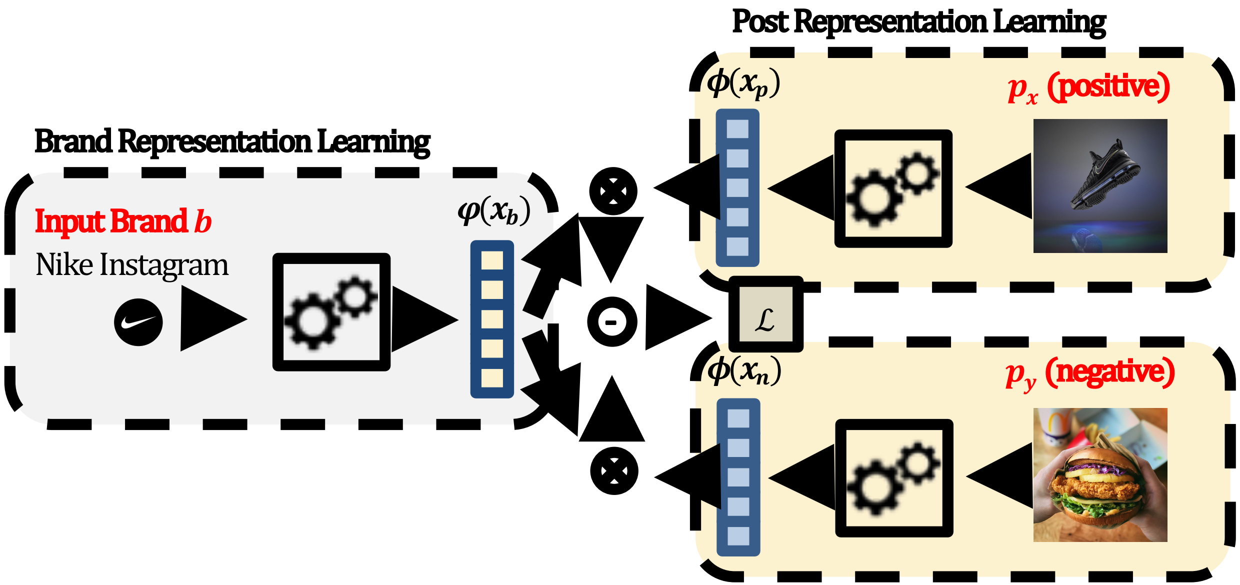
- Brand b
- Image Post p

Output:

- $f(b,p) = \cos\_sim(\varphi(\boldsymbol{x}_b), \phi(\boldsymbol{x}_p))$

Loss Function:

- $\mathcal{L} = \max\left(0, f(b,p_y) - f(b,p_x)\right) + margin + \|\theta\|_2$



Brand Representation Learning

- Brand Associations: images and symbols associated with a brand.
- Examples:
  - BMW: sophistication, fun driving and superior engineering
  - Apple: Steve Jobs, luxury design
- Brand associations are reflected in Web photos (Kim, WSDM'14)
- A brand identity is determined by the unique combination of the brand associations

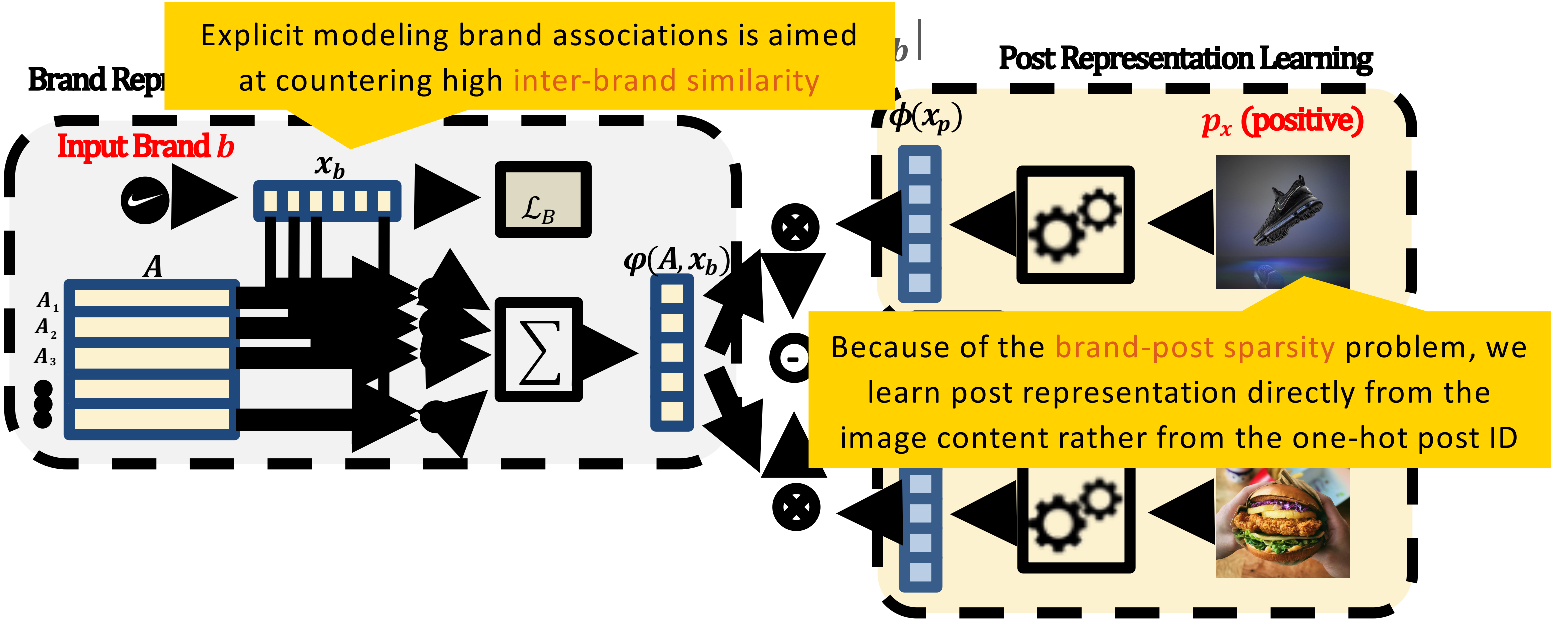
Brand Representation Learning

Brand Representation Learning:

- $\varphi(\boldsymbol{A}, \boldsymbol{x}_b) = \sum_{i=1}^{N} \boldsymbol{A}_i \circ \boldsymbol{x}_b$

Loss Function:

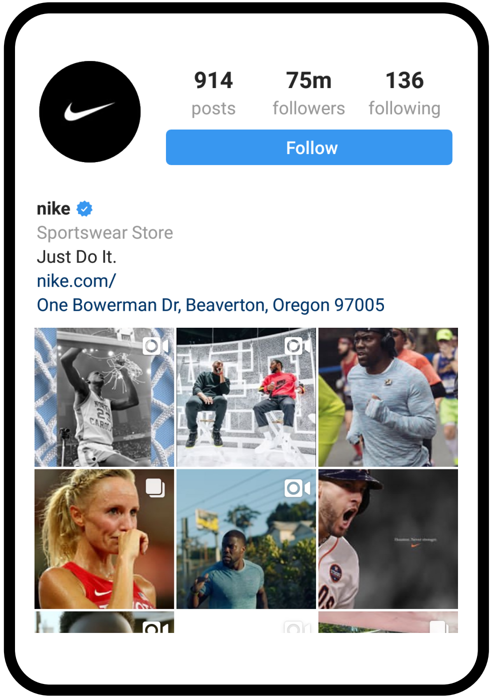
- $\mathcal{L} = \mathcal{L}_1 + \alpha \mathcal{L}_2 + \|\theta\|_2$
- $\mathcal{L}_A = \max(0, f(b, p_y) - f(b, p_x)) + margin$



Dataset

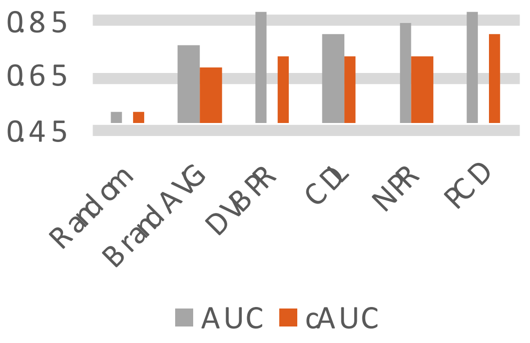
- Need large-scale dataset with brand visual history
- Instagram posting history for 927 brands from 14 verticals (1,158,474 posts in total)
- Testing set: brand's 10 most recent posts (1,149,204 training + 9,270 testing)

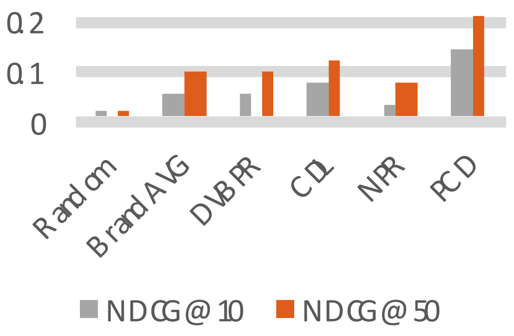
Alcohol 69	Airlines 57	Auto 83	Fashion 98	Food 85
Furnishing 49	Electronics 79	Nonprofit 71	Jewelry 71	Finance 37
Services 69	Entertainment 88	Energy 4	Beverages 67	**Total** **927**



PCD vs Others

- We evaluate the performance of PCD versus state-of-the-art baselines
- AUC: prob. of ranking a randomly chosen positive sample higher than a randomly chosen negative sample
- cAUC: prob. of ranking a randomly chosen positive sample higher than a randomly chosen negative sample **from a competitor brand**





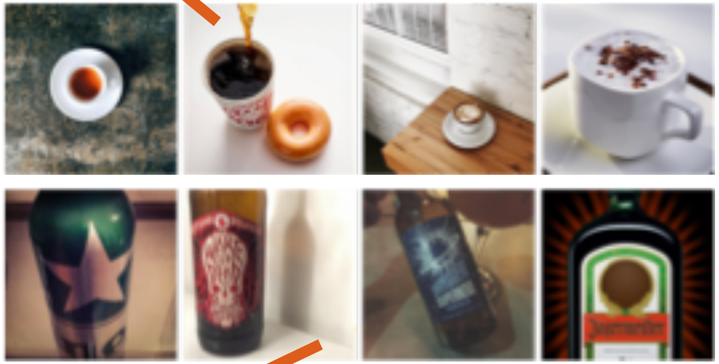
	MedR
Random	568
BrandAVG	29
DVBPR [ICDM'17]	20
CDL [CVPR'16]	19
NPR [WSDM'18]	33
PCD	**5**

- cAUC results are consistently lower than AUC
- PCD has the highest score for all metrics
- MedR for PCD is ~4 times smaller than CDL

Visualizing Brand Associations

Four nearest neighbors images from the dataset

Costa Coffee, Starbucks, Salt Spring Coffee



Dom Pérignon, Moët & Chandon

Rolls-Royce,
Tesla,
Cadillac, Volvo

Conclusions

- We formulate the problem of Content Discovery for Brands
- We propose and evaluate Personalized Content Discovery (PCD), which explicitly models brand associations
- A large scale dataset with the Instagram history of more than 900 brands was released
- As future studies, we plan to integrate temporal context and investigate on which high level attributes make images and videos actionable

PCD vs Others

Metrics:

- AUC: probability of ranking a randomly chosen positive example higher than a randomly chosen negative one
- cAUC: probability of ranking a randomly chosen positive example higher than a randomly chosen negative sample **from a competitor**
- NDCG: quality of a ranking list based on the post position in the sorted result list
- MedR: the median position of the first relevant document

Baselines:

- Random: generate a random ranking
- BrandAVG: nearest neighbor with respect to mean feature vector
- DVBPR: pairwise model inspired by VPR, which excludes non-visual latent factors. ICDM'17
- CDL: Comparative Deep Learning, pure content based pairwise architecture. CVPR'16
- NPR: Neural Personalized Ranking, recent pairwise architecture. WDSM'18

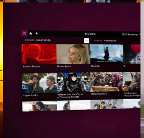
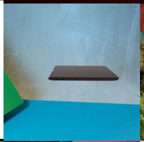
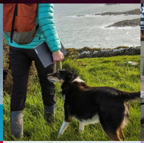
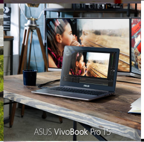
PCD vs Others, Results

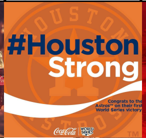
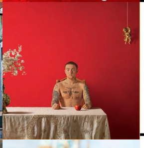
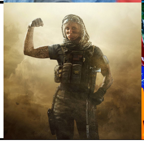
	AUC	cAUC	NDCG@10	NDCG@50	MedR
Random	0.503	0.503	0.001	0.003	568
BrandAVG	0.769	0.687	0.068	0.105	29
DVBPR	0.862	0.734	0.059	0.102	20
CDL	0.807	0.703	0.079	0.119	19
NPR	0.838	0.716	0.040	0.076	33
PCD	**0.880**	**0.785**	**0.151**	**0.213**	**5**

- cAUC results are consistently lower than AUC → Competitor brands have subtle differences
- PCD has the highest score for all metrics → PCD learns finer-grained brand representations
- MedR for PCD is ~4 times smaller than CDL → PCD is more likely to discover a single relevant UGC

Case Studies

True Positive, False Negative and False Positive are shown for eight example brands

Brand	TP	FN	FP	
Carlsberg				from: Astra
Qatar Airways				from: United
Lenovo				from: Asus
Ford				from: Allianz

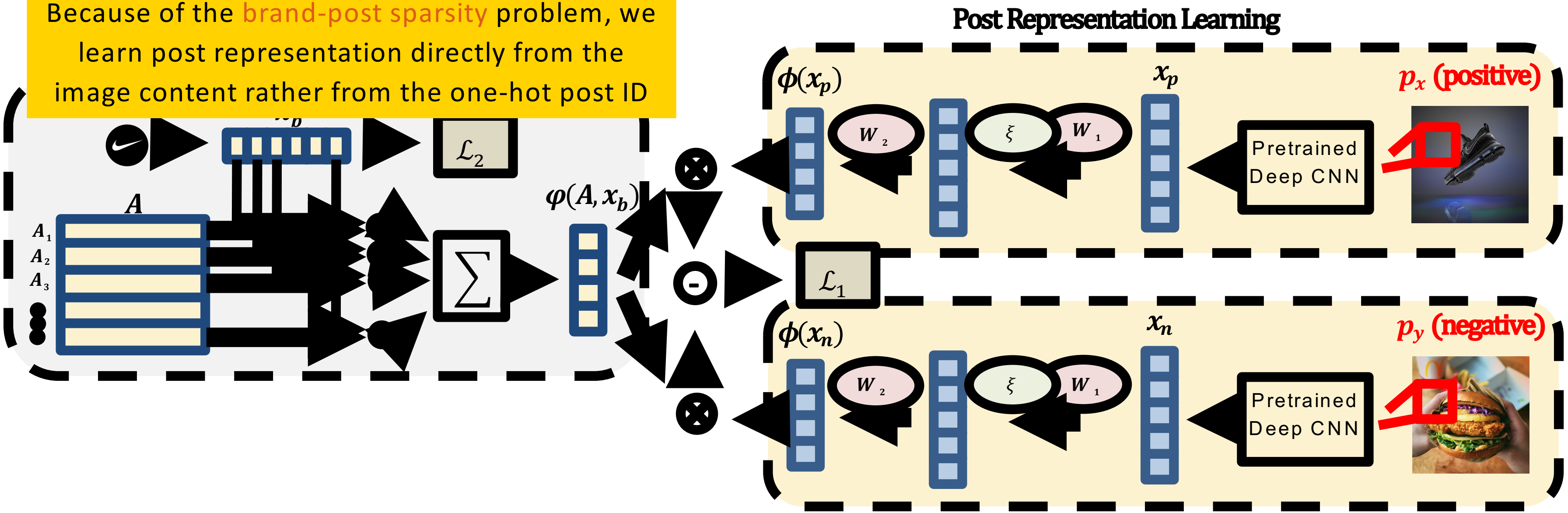
Brand	TP	FN	FP	
Coca Cola				from: Vodacom
Gucci				from: Google
Nintendo				from: Disney
Ubisoft				from: Marvel

Post Representation Learning

Post Representation Learning:

- $\phi(x_p) = W_2(\xi(W_1 x_p + \gamma_1)) + \gamma_2$
- $\xi(x) = \begin{cases} x, & if\ x > 0 \\ 0.01x, & otherwise \end{cases}$

Because of the brand-post sparsity problem, we learn post representation directly from the image content rather from the one-hot post ID



Brand Associations: Ablation Study

- What is the impact of brand associations?
- Ablation study, comparing:
  - PCD: our method, with explicit brand association learning
  - PCD1H: direct brand embedding learning from one-hot ID
- We compare the two methods in terms of NDCG, for different cut-off values
- PCD consistently exhibits a higher NDCG