A Probabilistic Framework for Multi-view Feature Learning with Many-to-many Associations via Neural Networks

Akifumi Okuno, Tetsuya Hada, Hidetoshi Shimodaira

Presented by: Zhun Liu

What is Multi-view Data?

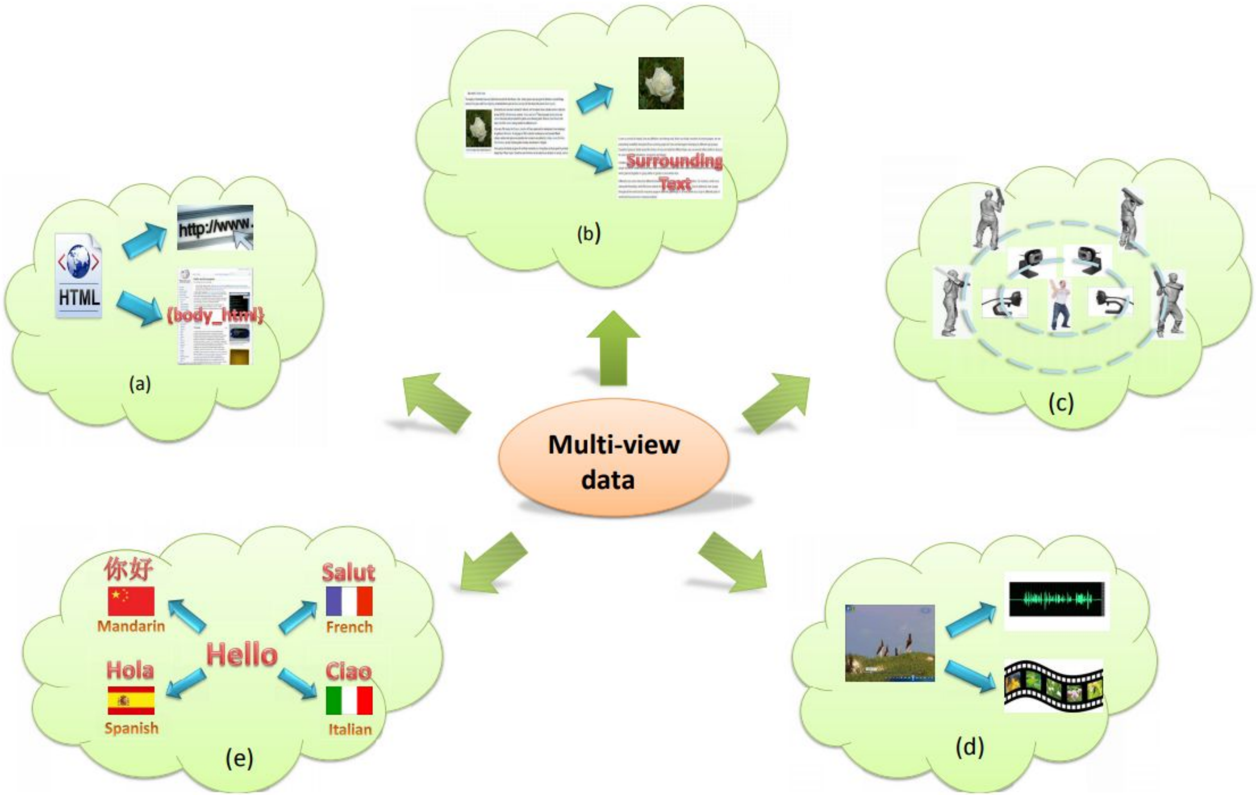


Image excerpted from: Xu, C., Tao, D. and Xu, C., 2013. A survey on multi-view learning. arXiv preprint arXiv:1304.5634.

Multi-view Representation Learning

- CCA
- DCCA
- Cycle-translation
- Multimodal skip-gram
- Graph Convolution
- ...

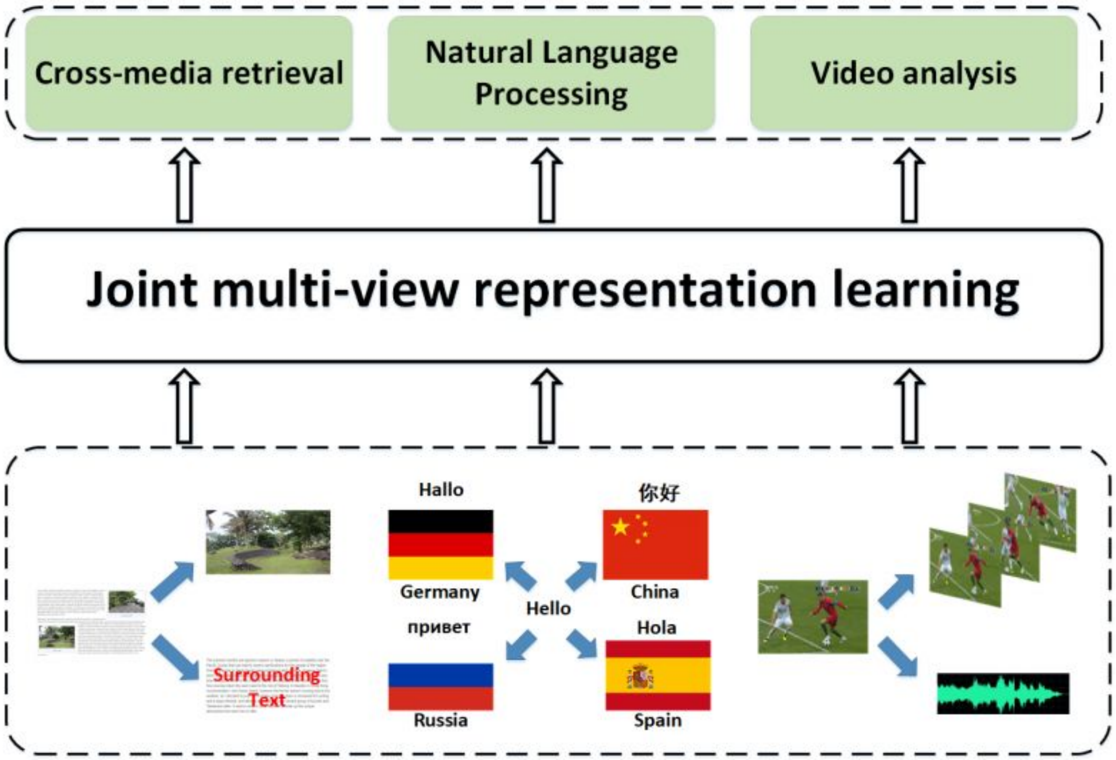
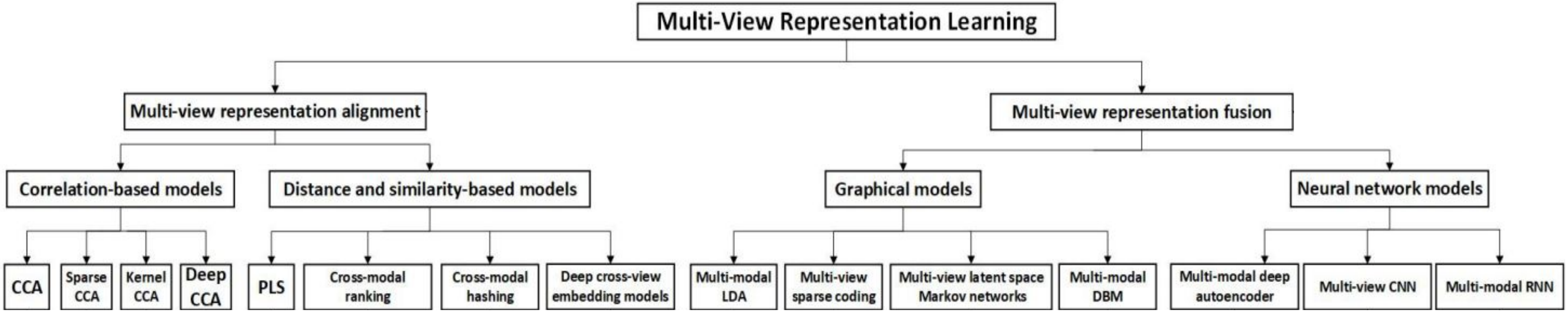


Image excerpted from: Li, Y., Yang, M. and Zhang, Z.M., 2018. A Survey of Multi-View Representation Learning. IEEE Transactions on Knowledge and Data Engineering.

Multi-view Representation Learning



A nice survey if you want to dig into it:

A Survey of Multi-View Representation Learning

https://arxiv.org/pdf/1610.01206.pdf

Image excerpted from: Li, Y., Yang, M. and Zhang, Z.M., 2018. A Survey of Multi-View Representation Learning. IEEE Transactions on Knowledge and Data Engineering.

What previous models achieve

CCA: Paired-data, one-to-one associations, linear relationships

DCCA: Paired-data, one-to-one associations, non-linear relationships

CDMCA: Graph data, many-to-many associations, linear relationships

…

Authors of this paper: We need a many-to-many, non-linear model!

PMvGE: the Model Input

Data points from D different views:

$$\mathbf{X} = \{(x_1, d_1), (x_2, d_2), ..., (x_n, d_n)\}$$

$$d_i \in \{1, 2, ..., D\}$$

Observed links (link weights) between the data points

$$\mathbf{W} = \{w_{ij}, (i, j) \in [n] \times [n]\}$$

PMvGE: the Model Parameters

We first define a similarity score parametrized by neural nets and a symmetric matrix α to regulate the connections between views

$$\mu_{ij} = \alpha^{(d_i,d_j)} \cdot \exp(\langle f_{\psi_i}^{(d_i)}(x_i), f_{\psi_j}^{(d_j)}(x_j)\rangle)$$

These scores are normalized to get association probabilities

$$P(e_t = (x_i, x_j)|\{x_k\}_{k=1}^{N}) := \frac{\mu_{ij}}{\sum_{i\neq j} \mu_{ij}}$$

PMvGE: the Generative Modelling Process

First we drawn a random non-negative integer T from a particular Poisson process

$$T \sim Poisson(\lambda), \lambda := \sum_{i \neq j} \mu_{ij}$$

We then draw from the link distribution we have defined previously T times

$$s_t \sim P(e_t = (x_i, x_j) | \{x_k\}_{k=1}^{N}), t = 1, ..., T$$

Then w_{ij} is defined as the number of times link (x_i, x_j) appeared

$$w_{ij} := \sum_{i=1}^{T} \mathbb{I}_{e_t = (x_i, x_j)}$$

PMvGE: the Objective Function

From previous generative process, it turns out w_{ij} follows a Poisson distribution:

$$w_{ij} \sim Poisson(\mu_{ij})$$

Hence it has the PMF

$$P(w_{ij}) = \mu_{ij}^{w_{ij}} \cdot \frac{\exp(-\mu_{ij})}{w_{ij}!}$$

Taking the log likelihood for the entire dataset we have

$$\log P(\{w_{ij}\}, (i,j) \in \mathcal{I}_n) = \log \prod_{(i,j)\in\mathcal{I}_n} P(w_{ij})$$

$$= \sum_{(i,j)\in\mathcal{I}_n} w_{ij} \log \mu_{ij} - \mu_{ij} - \log(w_{ij}!)$$

PMvGE: the Objective Function - continued

For optimization purposes, we could throw away the last term that is not related to parameters

$$\ell_n(\{x_k\}, \{w_{ij}\} | \alpha, \psi) = \sum_{(i,j) \in \mathcal{I}_n} w_{ij} \log \mu_{ij} - \mu_{ij}$$

Then we can maximize this objective to perform MLE of model parameters.

Yay! Mission accomplished!

```
loss.backward()
optimizer.step()
```

PMvGE: Optimization

- Can we go ahead and perform SGD on the log likelihood?

  No, because that way we cannot enforce symmetry of α

We need an alternating iterative algorithm that optimizes α, ψ .

First fix α and do SGD on ψ : some mini-batch sampling details involved

Then update α analytically: see equation (11) in the paper

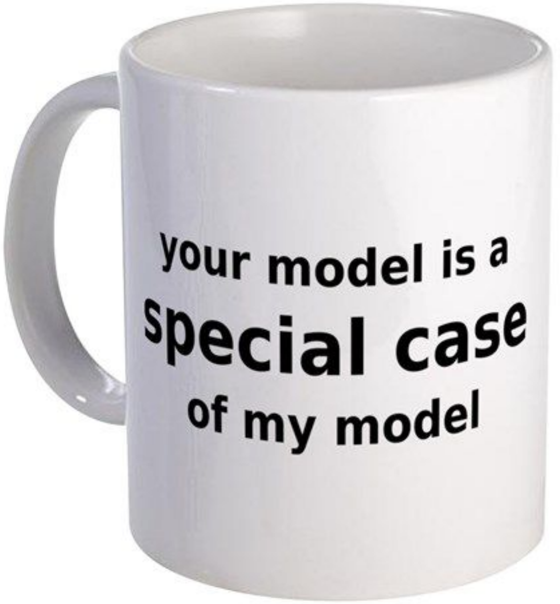
PMvGE: “Your model is a special case of my model”

CCA…

MLP...

CDMCA…

DCCA…



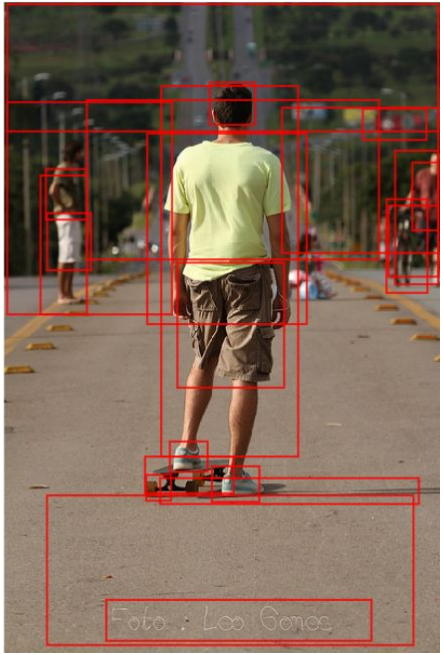
Multi-Entity Variational Autoencoder

Charlie Nash, S. M. Ali Eslami, Chris Burgess, Irina Higgins, Daniel Zoran, Theophane Weber, Peter Battaglia

Presented by: Zhun Liu

What if our views are not naturally separated?

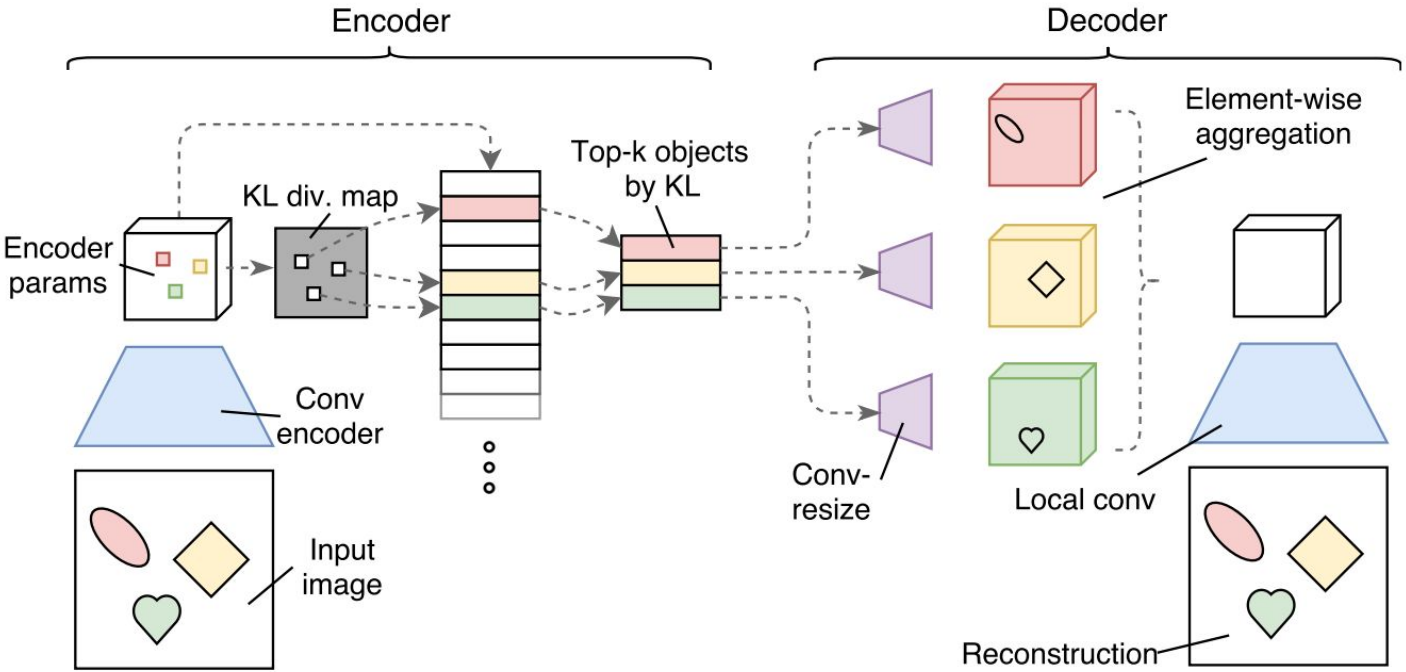
E.g. We have bounding boxes of objects and we treat them each as a view



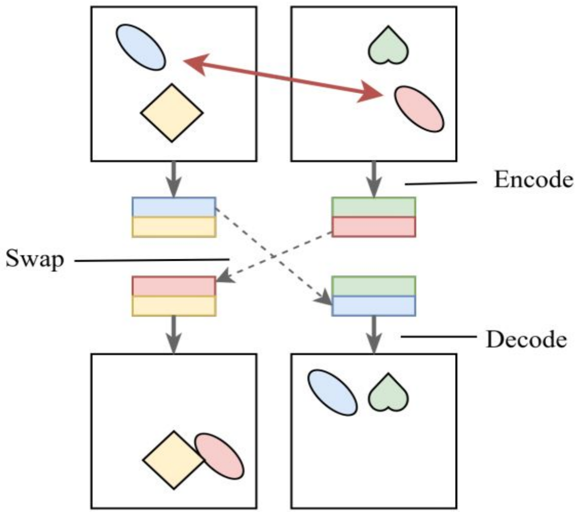
What if we don’t have bounding boxes?

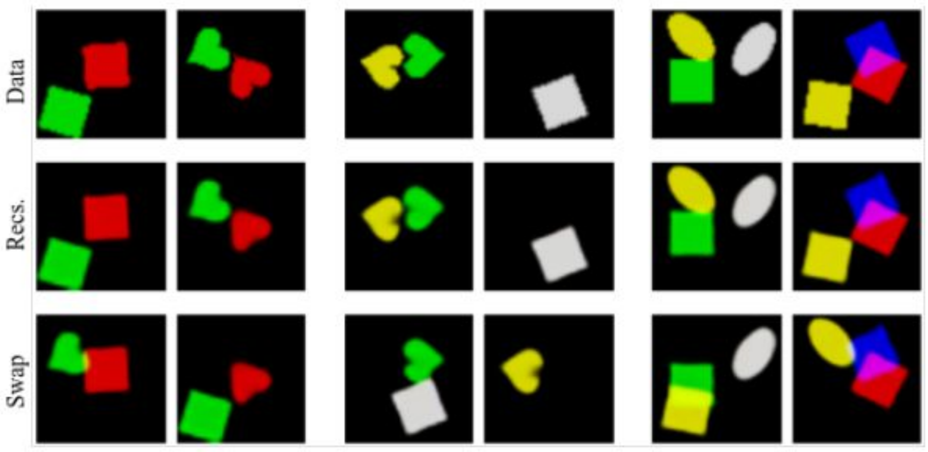
What if our views are just mixed together?

Multi-Entity VAE: Picking Multiple Representations

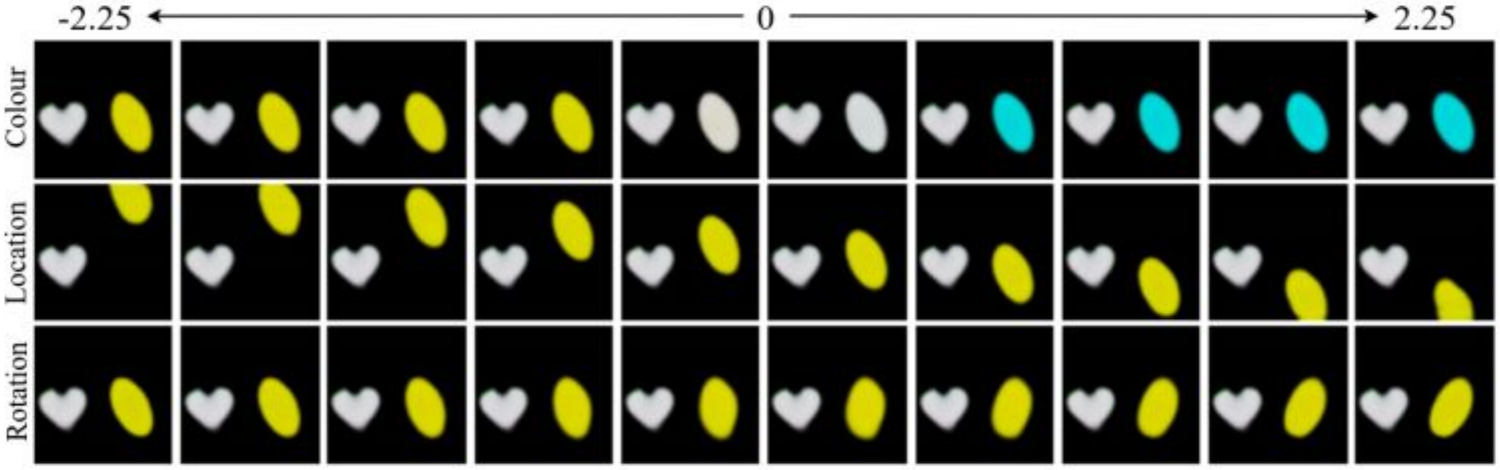


Multi-Entity VAE: Disentangling Between Objects





Multi-Entity VAE: Disentangling Within Objects



Thanks!