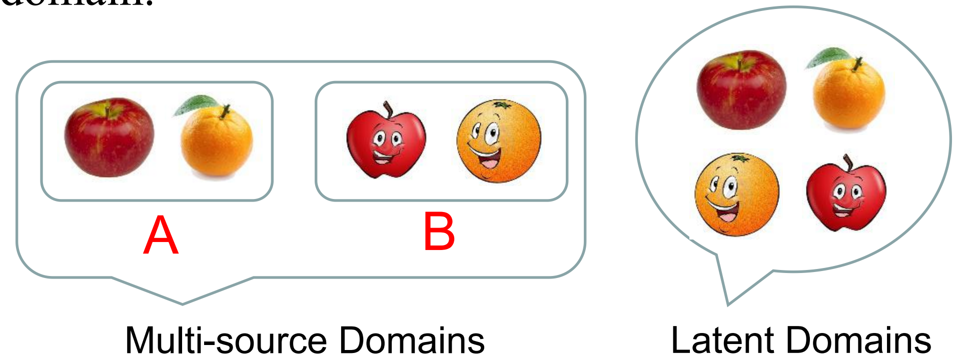
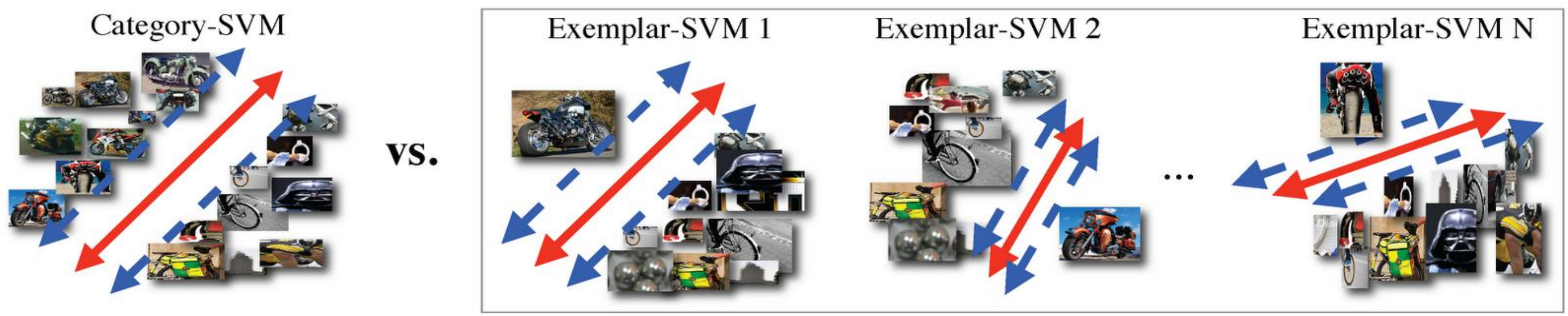
Multi-view Domain Generalization for Visual Recognition

Li Niu presents

- Latent domains are characterized by different hidden factors (e.g., pose, illumination). Domain generalization is to generalize latent source domains to unknown target domain.



- Since each sample can be treated as an atomic domain, exemplar classifiers [1] can be readily incorporated into our proposed method.



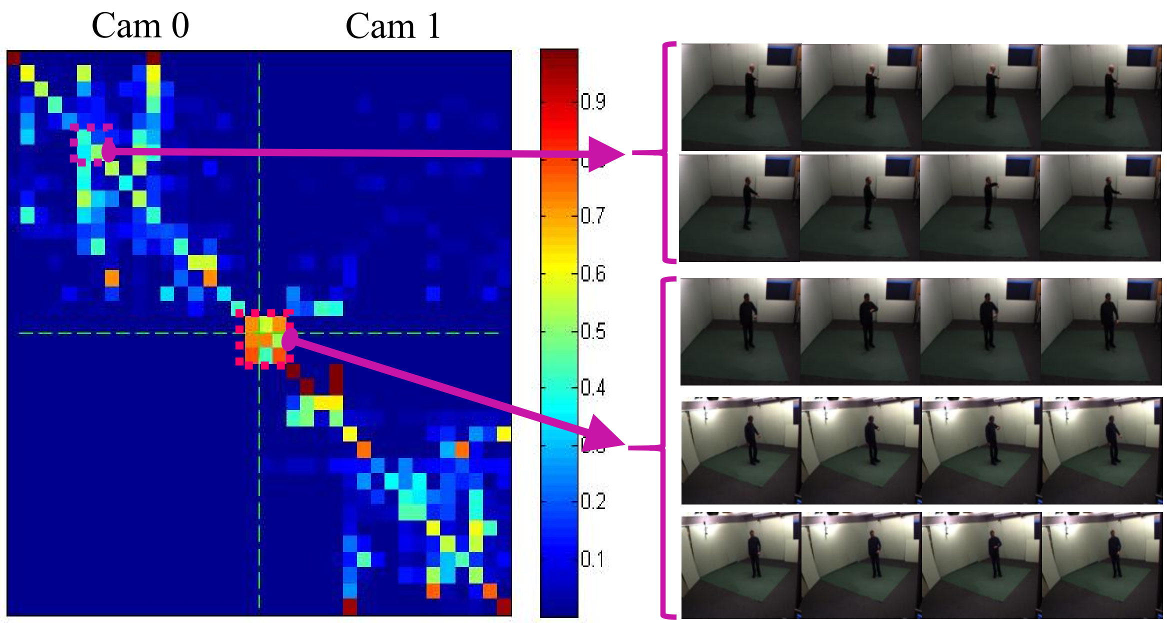
NANYANG TECHNOLOGICAL UNIVERSITY

[1] Malisiewicz, T., Gupta, A., Efros, A.A.: Ensemble of exemplar-svms for object detection and beyond. In: ICCV. (2011)

- Low rank assumption (Single-View): [2]

Low rank likelihood matrix $\mathbf{G} = [g_{ij}] \in \mathbb{R}^{n \times n}$,

where g_{ij} is the likelihood of the i-th positive training sample by using the j-th exemplar exemplar classifier.



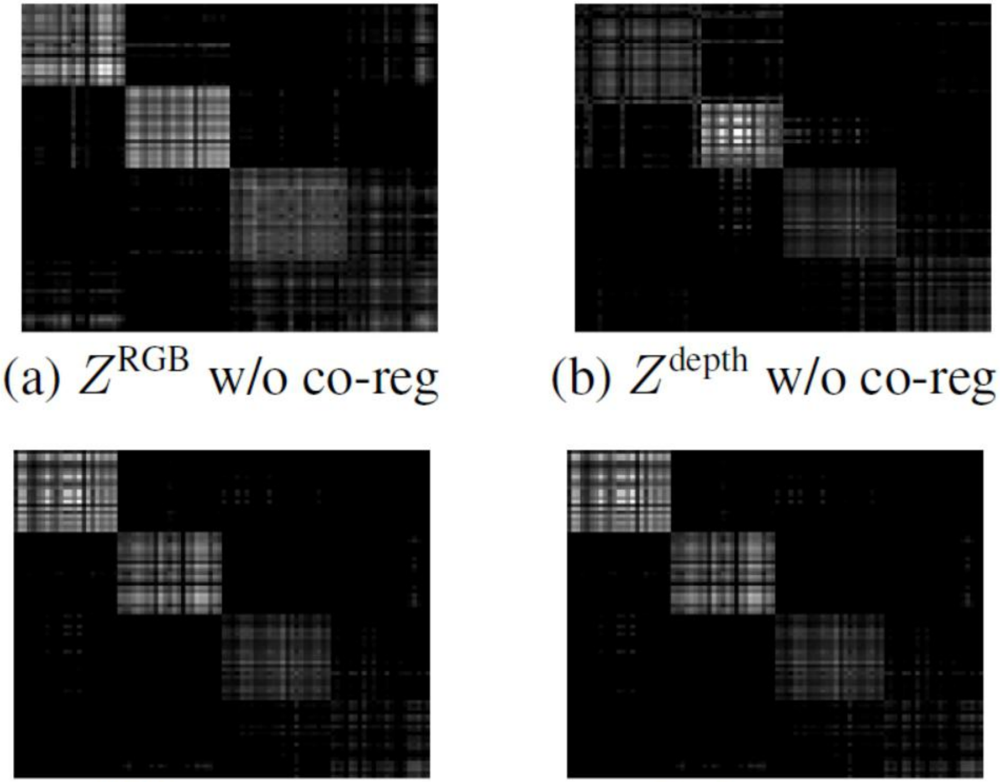
(a) Likelihood matrix **G** (b) Video frames

NANYANG TECHNOLOGICAL UNIVERSITY

[2] Xu, Zheng, Wen Li, Li Niu, and Dong Xu. "Exploiting low-rank structure from latent domains for domain generalization." In ECCV, 2014.

- Low rank assumption (Multi-View):

Low rank representation matrix $\mathbf{Z}^v \in \mathbb{R}^{n \times n}$



(a) Z^{RGB} w/o co-reg (b) Z^{depth} w/o co-reg

(c) Z^{RGB} with co-reg (d) Z^{depth} with co-reg

NANYANG TECHNOLOGICAL UNIVERSITY

➢ Multi-View Domain Generalization: Formulation

$$\min_{\substack{\mathbf{Z}^v, \mathbf{W}^v, \mathbf{E}^v \\ \xi_i^v, \epsilon_{ij}^v}} \sum_{v=1}^{V} \left(\frac{1}{2} \|\mathbf{W}^v\|_F^2 + C \sum_{i=1}^{n} \xi_i^v + C \sum_{i=1}^{n} \sum_{j=1}^{m} \epsilon_{ij}^v \right)$$

low-rank

$$+ \sum_{v=1}^{V} \left(\lambda_2 \|\mathbf{E}^v\|_F^2 + \boxed{\lambda_3 \|\mathbf{Z}^v\|_*} \right) + \boxed{\frac{\gamma}{2} \sum_{v, \tilde{v}: v \neq \tilde{v}} \|\mathbf{Z}^v - \mathbf{Z}^{\tilde{v}}\|_F^2}$$

co-regularizer

$$\text{s.t.} \quad \boxed{\begin{array}{ll} \mathbf{w}_i^{v\prime} \mathbf{x}_i^{v+} \geq 1 - \xi_i^v, & \xi_i^v \geq 0, \\ \mathbf{w}_i^{v\prime} \mathbf{x}_j^{v-} \leq -1 + \epsilon_{ij}^v, & \epsilon_{ij}^v \geq 0 \end{array}} \quad \begin{array}{l} \forall v, \forall i, \\ \forall v, \forall i, \forall j, \end{array}$$

Exemplar SVM

$$\boxed{\mathbf{W}^v = \mathbf{W}^v \underbrace{\mathbf{Z}^v}_{\text{representaton matrix}} + \underbrace{\mathbf{E}^v}_{\text{reconstruction error}},} \quad \forall v$$

Low-Rank Representation (LRR)

NANYANG TECHNOLOGICAL UNIVERSITY

➢ Multi-View Domain Generalization: Optimization

$$\min_{\substack{\mathbf{Z}^v,\mathbf{W}^v,\mathbf{G}^v\\ \mathbf{E}^v,\xi_i^v,\epsilon_{ij}^v}} \sum_{v=1}^{V}\left(\frac{1}{2}\|\mathbf{W}^v\|_F^2 + C\sum_{i=1}^{n}\xi_i^v + C\sum_{i=1}^{n}\sum_{j=1}^{m}\epsilon_{ij}^v\right)$$

$$+\sum_{v=1}^{V}\left(\boxed{\lambda_1\|\mathbf{W}^v-\mathbf{G}^v\|_F^2} + \lambda_2\|\mathbf{E}^v\|_F^2 + \lambda_3\|\mathbf{Z}^v\|_*\right)$$

intermediate variable
for ease of optimization

$$+\frac{\gamma}{2}\sum_{v,\tilde{v}:v\neq\tilde{v}}\|\mathbf{Z}^v-\mathbf{Z}^{\tilde{v}}\|_F^2$$

$$\text{s.t.}\quad \mathbf{w}_i^{v\prime}\mathbf{x}_i^{v+} \geq 1-\xi_i^v,\quad \xi_i^v \geq 0,\quad \forall v,\forall i,$$

$$\mathbf{w}_i^{v\prime}\mathbf{x}_j^{v-} \leq -1+\epsilon_{ij}^v,\quad \epsilon_{ij}^v \geq 0,\quad \forall v,\forall i,\forall j,$$

$$\mathbf{G}^v = \mathbf{G}^v\mathbf{Z}^v + \mathbf{E}^v,\quad \forall v$$

(1) Update Z and E (ADM)

(2) Update G and W
1. Update W (dual form)
2. Update G (closed-form)

NANYANG
TECHNOLOGICAL
UNIVERSITY

➢ Domain Adaptation: Formulation

$$
\begin{aligned}
\min_{\substack{\mathbf{Z}^v,\mathbf{W}^v,\mathbf{G}^v\\ \mathbf{E}^v,\xi_i^v,\epsilon_{ij}^v}} \sum_{v=1}^{V} \Big(& \frac{1}{2}\|\mathbf{W}^v\|_F^2 + C\sum_{i=1}^{n}\xi_i^v + C\sum_{i=1}^{n}\sum_{j=1}^{m}\epsilon_{ij}^v \\
& + \lambda_1\|\mathbf{W}^v - \mathbf{G}^v\|_F^2 + \lambda_2\|\mathbf{E}^v\|_F^2 + \lambda_3\|\mathbf{Z}^v\|_* \Big) \\
& + \frac{\gamma}{2}\sum_{v,\tilde{v}:v\neq\tilde{v}} \|\mathbf{Z}^v - \mathbf{Z}^{\tilde{v}}\|_F^2 + \boxed{\theta\sum_{v=1}^{V}\Omega(\mathbf{W}^v,\mathbf{L}^v,\mathbf{U}^v)} \\
\text{s.t.} \quad & \mathbf{w}_i^{v\prime}\mathbf{x}_i^{v+} \geq 1 - \xi_i^v, \quad \xi_i^v \geq 0, \quad \forall v, \forall i, \\
& \mathbf{w}_i^{v\prime}\mathbf{x}_j^{v-} \leq -1 + \epsilon_{ij}^v, \quad \epsilon_{ij}^v \geq 0, \quad \forall v, \forall i, \forall j, \\
& \mathbf{G}^v = \mathbf{G}^v\mathbf{Z}^v + \mathbf{E}^v, \quad \forall v,
\end{aligned}
$$

Laplacian Regularizer

where $\Omega(\mathbf{W}^v,\mathbf{L}^v,\mathbf{U}^v) \;=\; \mathrm{tr}(\mathbf{W}^{v\prime}\mathbf{U}^v\mathbf{L}^v\mathbf{U}^{v\prime}\mathbf{W}^v)$

NANYANG TECHNOLOGICAL UNIVERSITY

Experiments: Domain Generalization

- Dataset && Features
  - ACT 4^2
    - RGB and Depth (2 Views)
    - 4 Camera-viewpoints (4 Domains)
    - 6000-dim IDT-BOW
  - ORGBD
    - RGB and Depth (2 Views)
    - 2 different environments (2 Domains)
    - 6000-dim IDT-BOW
  - Office-Caltech
    - 2 different features: 4096-dim Caffe-6 and Decaf-6 features (2 Views)
    - Amazon, Webcam, Dslr, and Caltech (4 Domains)

- Experimental setting

We mix several domains as the source domain for training classifiers and use the remaining domains as the target domain for testing.

NANYANG TECHNOLOGICAL UNIVERSITY

Experiments: Domain Generalization

- Baselines

Basic	SVM E-SVM
Multi-View	LRCS SVM-2K KCCA
Domain Generalization	DICA LRESVM
Discovering Latent Domains	Gong Latent Domain Hoffman Latent Domain
subcategory	Sub-category

NANYANG TECHNOLOGICAL UNIVERSITY

Experiments: Domain Generalization

➢ Results

Recognition accuracy of domain generalization

Dataset	ACT4^2	ORGBD	Office
SVM [10]	56.40	48.67	84.52
ESVM [29]	58.60	51.79	86.14
LRCS [11]	59.72	52.68	85.28
SVM-2K [16]	59.68	50.00	86.10
KCCA [21]	57.72	51.34	86.33
DICA [30]	59.10	47.32	86.12
LRESVM [38]	62.61	53.57	87.04
[18](match)	57.83	50.00	86.47
[18](ensemble)	58.42	51.79	86.06
[23](match)	55.21	44.65	85.75
[23](ensemble)	57.78	50.45	84.81
Sub-Cate [22]	59.71	52.68	86.64
MVDG	**66.16**	**55.81**	**88.13**

NANYANG TECHNOLOGICAL UNIVERSITY

Experiments: Domain Adaptation

- Baselines

Domain Adaptation
- DASVM
- KMM
- TCA
- SA
- DIP
- GFK
- SGF

Multi-View Semi-Supervised Learning
- Co-training
- Co-LapSVM

Multi-View Domain Adaptation
- Coupled
- MVTL_LM
- MDT

Experiments: Domain Adaptation

- Results

Recognition accuracy of domain adaptation

Dataset	ACT4[2]	ORGBD	Office
SVM [10]	56.40	48.66	84.52
DASVM [6]	60.22	50.45	85.60
KMM [24]	59.46	52.12	86.34
TCA [32]	59.12	48.66	85.79
SA [17]	63.42	52.24	86.79
DIP [1]	58.86	54.46	86.58
GFK [19]	60.61	53.13	86.22
SGF [20]	56.17	52.23	85.78
Co-training [4]	62.15	53.13	87.96
Co-LapSVM [34]	61.57	52.68	88.20
Coupled [3]	64.79	54.02	86.48
MVTL_LM [41]	63.70	55.36	87.76
MDT [39]	64.97	54.46	86.87
LRCS [11]	62.07	55.81	86.12
MVDA	**68.67**	**58.04**	**91.04**

NANYANG TECHNOLOGICAL UNIVERSITY

Thanks for your attention!

NANYANG TECHNOLOGICAL UNIVERSITY