**Model Recommendation: Generating Object Detectors from Few Samples**

Yu-Xiong Wang and Martial Hebert
Email: {yuxiongw, hebert}@cs.cmu.edu

---

**Motivation**

Learning from supervised data → Learning from experience
- A large annotated dataset → A large library of pre-trained models
- Expensive training from scratch → Selecting most suitable store models

Recommender systems → Vision domain
- Items → Models
- Users → Tasks

---

**Collaborative Filtering**

Joint latent factor space for task-model interactions
- Estimate of rating $r_{ij}$ of model $M_i$ on task $T_j$: $\hat{r}_{ij} = u_i^T v_j$

Low-rank matrix factorization
- SVD with biases $\hat{r}_{ij} = \mu + q_i + p_j + u_i^T v_j$, $B = E F^T = (E_d S_d) F_d = UV$
- NMF $R \approx UV = \sum_{i=1}^{d} U_i V_i$

Estimate of ratings $r$ for an input task given $k$ probe ratings $\left(\hat{U}_k\right)^T v = r_k$, $\hat{r} = U v$

---

**Pipeline**

- Store Tasks
- Input Tasks
- Ratings GAP
- Predicted Ratings
- Probe Ratings
- NA. 0.6, 0.01, 0.001, NA, 0.5

---

**Unsupervised Hyper-Training: PBC-CNN**

- **Feature Space:** Pre-trained CNNs on ILSVRC 2012
- **Processing Units:** Region proposals
- **Unsupervised Discovery:** Predictable Discriminative Binary Codes (PBCs [Kastogan et al.])

---

**Qualitative Visualization**

- Continuous category space discovery by model sharing
- Collaborative detection of different models
- Same model informative across different tasks

---

**Recommender System Analysis**

- **Model Library:** ESVMs
- **Store Tasks:** Random detection tasks on PASCAL 2007 trainval
- **Ratings Store:** 12,608 × 10,000
- **Target Tasks:** Detection of 20 categories on PASCAL 2007 test

---

**Ensemble Model Recommendation**

- **Model Library:** Unsupervised PBCs hyper-trained on PASCAL 2007
- **Store Tasks:** Random detection tasks on PASCAL 2007
- **Ratings Store:** 10,000 × 10,000
- **Target Tasks:** Detection of 107 categories on SUN 09 test
- **Input Tasks:** 10 random images per category on SUN 09 training

---

**Conclusions**

- A large collection of universal and expressive models in an unsupervised manner (pre-trained features + models)
- A far smaller set of annotated samples
- An efficient way to generate new models by recommending
- A promising mechanism for a broader range of vision tasks