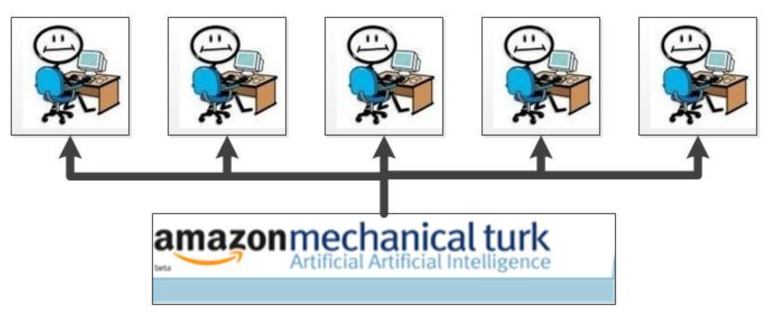
Active Visual Recognition with Expertise Estimation in Crowdsourcing

Chengjiang Long, Gang Hua — Stevens Institute of Technology, Hoboken, NJ 07030, {clong, ghua}@stevens.edu

Ashish Kapoor — Microsoft Research, Redmond, WA 98052, akapoor@microsoft.com

Background and motivation



Crowd-sourcing labeling

- ➢Pros: cheap and fast to obtain large quantity of label data.
- ➢Cons: the obtained labels can be very noisy.

Previous work

- ➢Majority voting based confidence. [Donez et al 2009-2010]
- ➢Incremental relabeling mechanism. [Zhao et al 2011]

Disadvantage

- ➢Cannot handle label noise during the labeling process.
- ➢The label quality will be heavily affect if the malicious labelers occur at the early stage.
- ➢Only investigate the case where a single copy of labels is engaged.

Motivation

- ➢We introduce the active learning strategy into the online framework.
- ➢We want to enable the collected labels are got by the quality labelers.
- ➢We want to handle the label noise during the labeling process.
- ➢We also want to make full use of multiple copies of labels.

Datasets

ImageNet dataset (10 categories, LLC features)
Gender face dataset (9441 face images)
CMU-MMAC dataset (14 category of actions)

Comparisons

Method	Label treatment	Flip noise	Sample	labelers
JGPC-ASAL(our)	Joint processing	With	Active	Active
JGPC-ASRL(our)	Joint processing	With	Active	Random
JGPC-RSAL(our)	Joint processing	With	Random	Active
JGPC-RSRL(our)	Joint processing	With	Random	Random
GPC-MVAS-F	Majority voting	With	Active	-
GPC-MVRS-F	Majority voting	With	Random	-
GPC-MVAS-K	Majority voting	Without	Active	-
GPC-MVRS-K	Majority voting	Without	Random	-

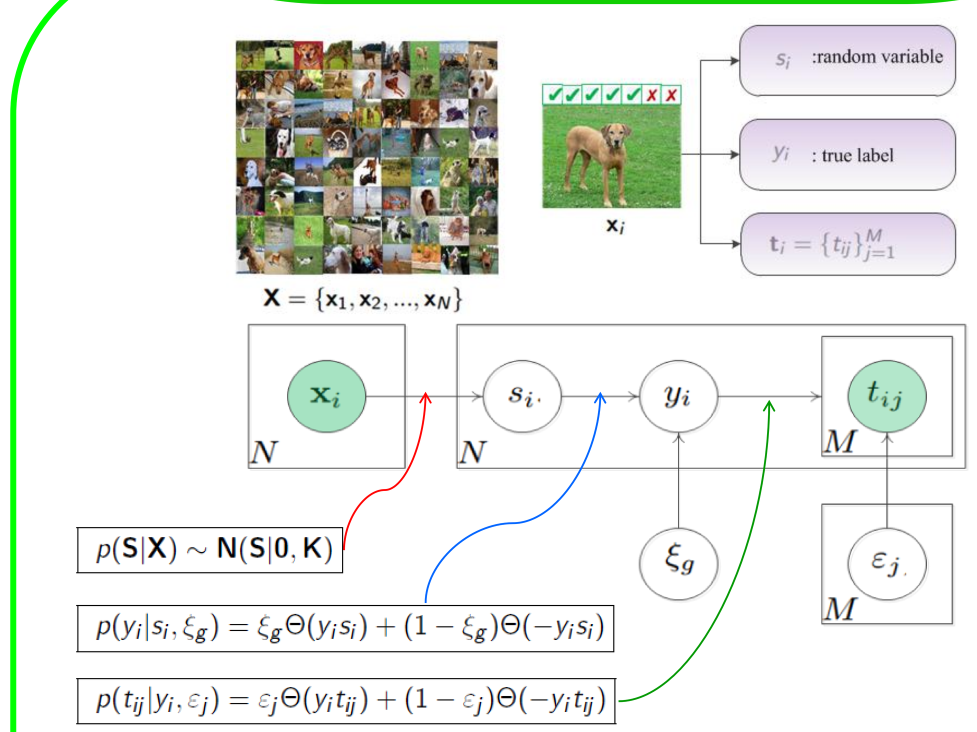
(Note: GPC-MVAS-K/GPC-MVRS-K/GPC are proposed by Ashish Kapoor et al [ICCV 2009)

Additional baselines:

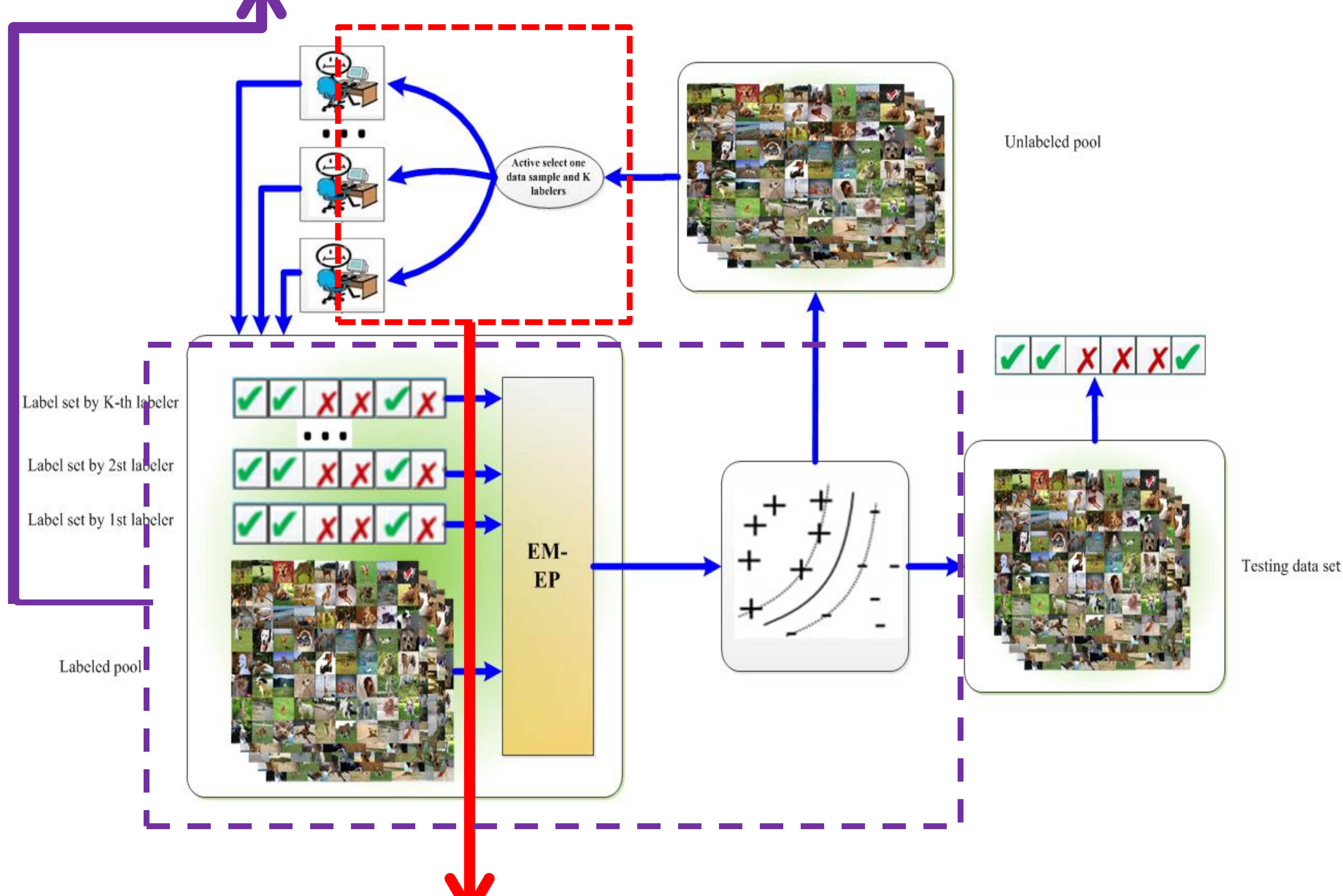
- JGPC-AS: joint learning GPC with active selection of samples.
- JGPC-RS: joint learning GPC with random selection of samples..
- ML-Bernoulli-AL: active learning with multiple labelers (Bernoulli version) proposed by Yan Yan et al. [ICML 2011]
- ML-Gaussian-AL: active learning with multiple labelers (Gaussian version) proposed by Yan Yan et al. [ICML 2011]

Sponsors

Graphical model



$$p(\mathbf{T}, \mathbf{Y}, \mathbf{S}|\mathbf{X}, \vec{\varepsilon}) \propto \frac{1}{Z} p(\mathbf{S}|\mathbf{X}) \prod_i p(y_i|s_i, \xi_g) \prod_{i,j} p(t_{ij}|y_i, \epsilon_j)$$



Expectation propagation

Integrating y_i out, we obtain:

$$\begin{aligned} p(\mathbf{t}_i|s_i, \vec{\varepsilon}) &= p(+1|s_i, \xi_g) \textstyle\prod_j p(t_{ij}| +1, \epsilon_j) \\ &+ p(-1|s_i, \xi_g) \textstyle\prod_j p(t_{ij}| -1, \epsilon_j) \end{aligned}$$

Rewrite joint probability function:

$$p(\mathbf{T}, \mathbf{S}|\mathbf{X}, \vec{\varepsilon}) = p(\mathbf{S}|\mathbf{X}) \prod_i p(\mathbf{t}_i|s_i, \vec{\varepsilon})$$

Assuming $\tilde{F}_i(s_i)$ follows Gaussian distribution,

$$\tilde{F}_i(s_i) \approx p(\mathbf{t}_i|s_i, \vec{\varepsilon})$$

Using Expectation Propagation [Thomas Minka, MIT PhD thesis, 2001], we get the posterior distribution $Q(\mathbf{S}_L)$.

$$\tilde{\mathbf{m}} = [\tilde{m}_1, \tilde{m}_2, \dots, \tilde{m}_L]^T$$
$$\tilde{\mathbf{v}} = [\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_L]^T$$

Inference

To any data sample $\mathbf{x}_u$,

$$\begin{aligned} &p(y_u|\mathbf{x}_u, \mathbf{D}_L) \\ &\doteq (2\xi_g - 1)\Phi\left(\frac{y_u \mathbf{k}_u^T(\mathbf{K}+\mathbf{\Lambda})^{-1}\tilde{\mathbf{m}}}{\sqrt{k(\mathbf{x}_u, \mathbf{x}_u) - \mathbf{k}_u^T(\mathbf{K}+\mathbf{\Lambda})^{-1}\mathbf{k}_u}}\right) \\ &+1 - \xi_g \end{aligned}$$

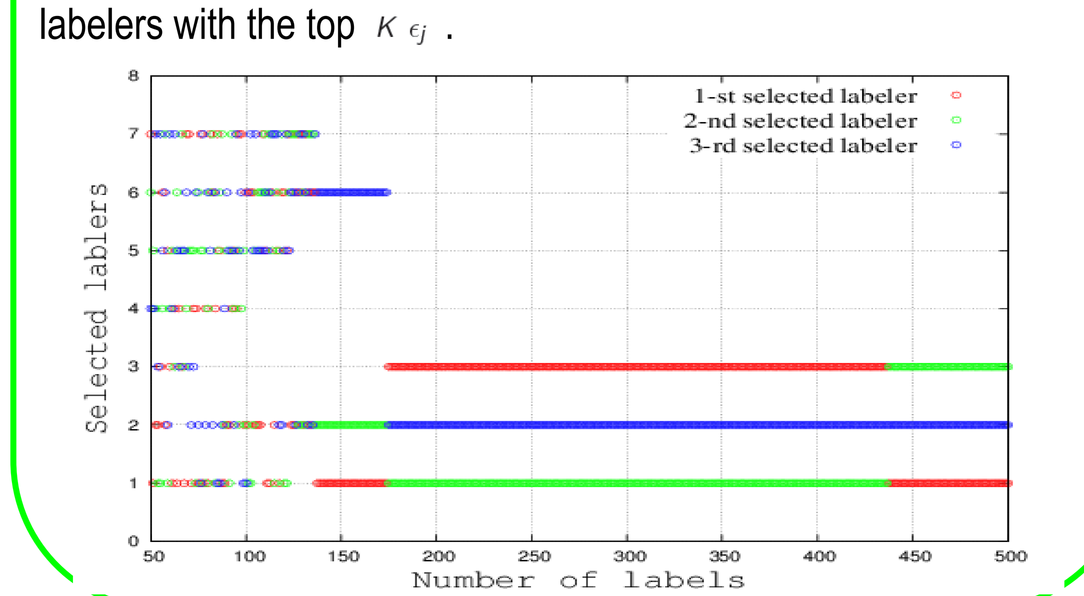
The diagonal matrix formed by the elements in $\tilde{\mathbf{v}}$

Active learning strategy

Sample selection: based on the entropy $H(y_u)$

$$\mathbf{x}_u^* = \arg\max_{\mathbf{x}_u \in \mathbf{X}_U} H(y_u)$$

Labeler selection: based on ϵ_j , select the top $K < M$ labelers with the top K ϵ_j .



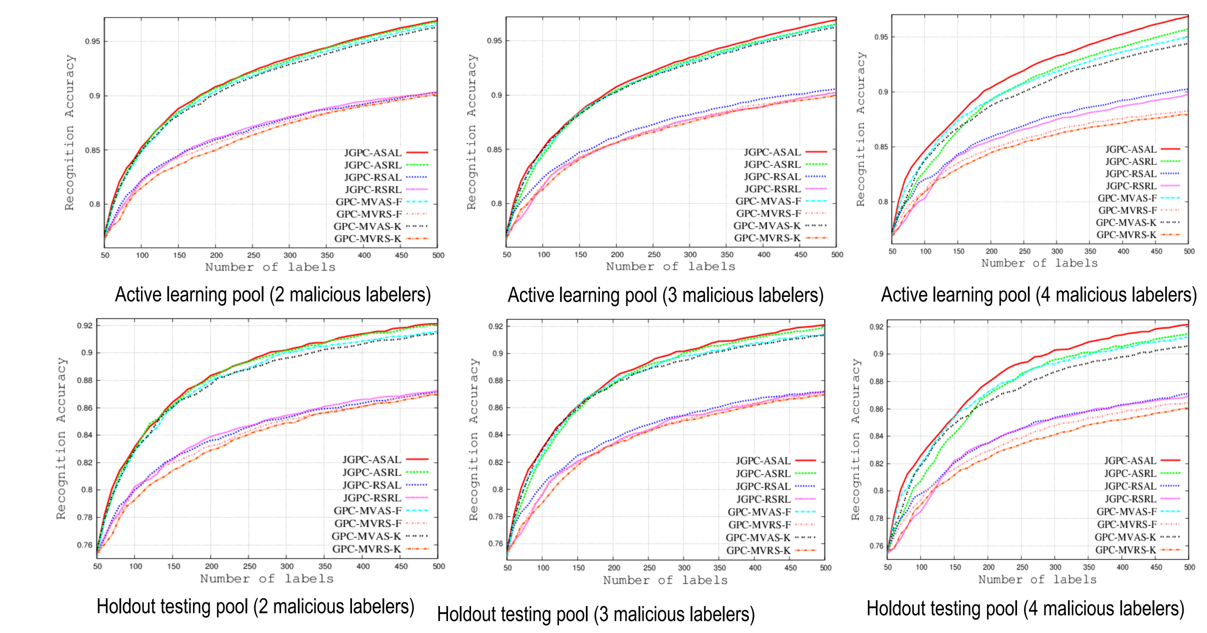
Parameter estimation

$$\begin{aligned} &\log p(\mathbf{T}_L, \mathbf{S}_L|\mathbf{X}_L, \vec{\varepsilon}) \\ &\geq \int_{\mathbf{S}_L} Q(\mathbf{S}_L) \log \frac{p(\mathbf{T}_L, \mathbf{S}_L|\mathbf{X}_L, \vec{\varepsilon})}{Q(\mathbf{S}_L)} \\ &= C + \sum_{i=1}^{L} \int_{s_i} q(s_i) \log p(\mathbf{t_i}|s_i, \vec{\varepsilon}) ds_i \end{aligned}$$

E-Step: Given the current parameter $\vec{\varepsilon}_p$, conduct EP inference to obtain and approximate inference of $Q(\mathbf{S}_L)$.
M-Step: Maximize the lower bound of $\log p(\mathbf{T}_L, \mathbf{S}_L|\mathbf{X}_L, \vec{\varepsilon})$over $\vec{\varepsilon}$ to obtain a new parameter $\vec{\varepsilon}$. $\vec{\varepsilon}_p \leftarrow \vec{\varepsilon}$, go to the **E-Step** and iterate until convergence.

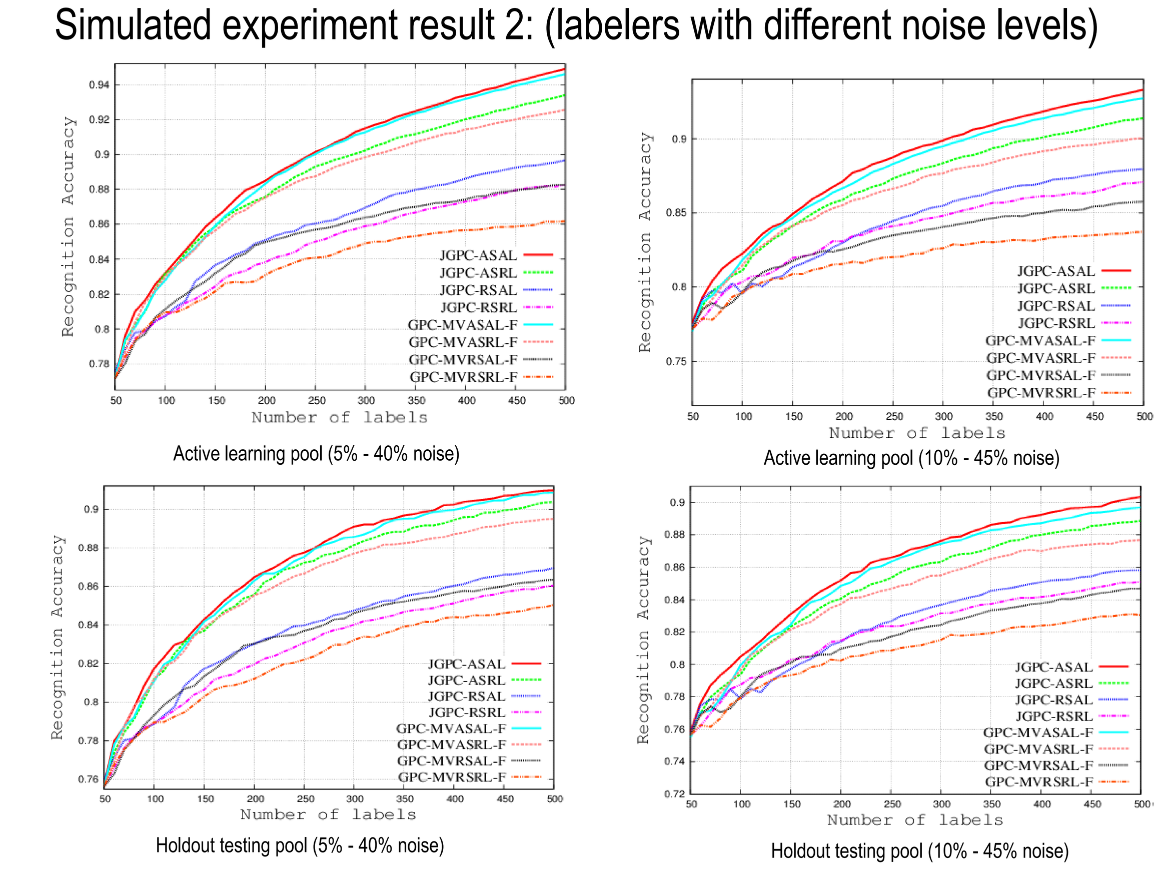
Simulated experiments

Simulated experiment result 1: (with 2, 3, 4 malicious labelers)



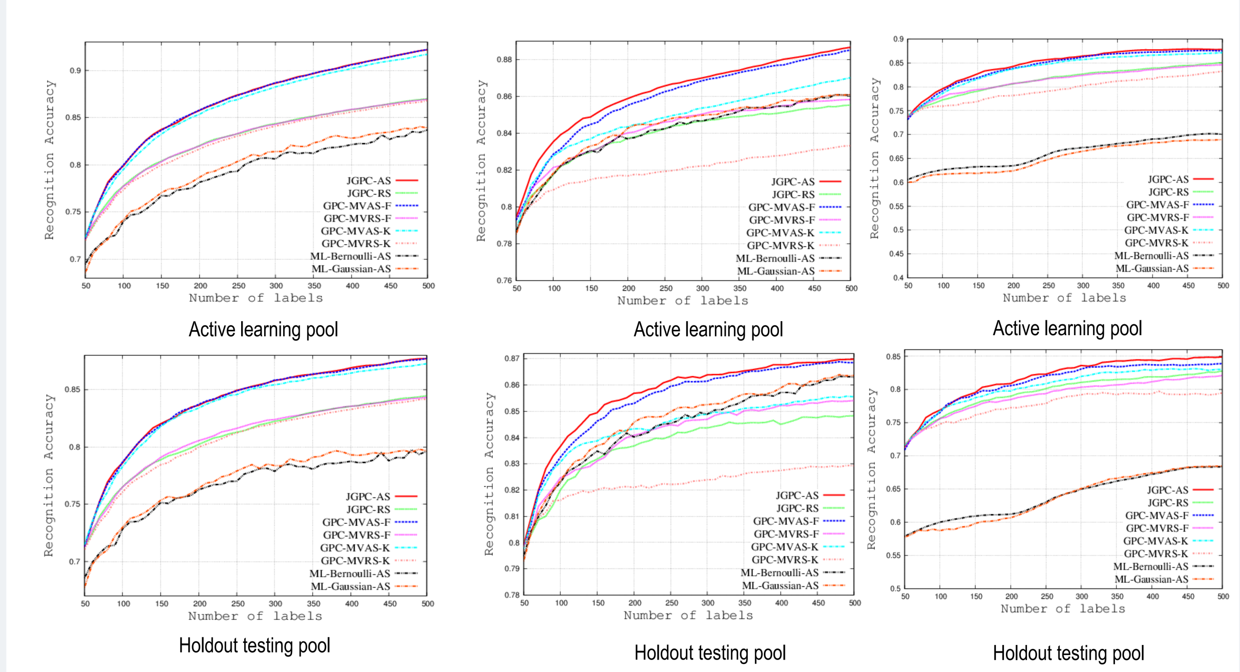
Our JGPC-ASAL is constantly ranks on the top.

Simulated experiment result 2: (labelers with different noise levels)



Our JGPC-ASAL is more robust to label noise than the naive majority voting criterion.

Experiments with real labels



ImageNet　　Gender face　　CMU-MMAC

Our JGPC-AS again shows superior recognition accuracy in both the active learning pool and the holdout testing pool.

Conclusion

We Present a hierarchical Bayesian model to learn a GPC from crowd-sourced labels by jointly processing multiple labels. Our two-level flip model enables active selection of both data sample and quality labelers. Our joint treatment of multiple labels is proven to be superior to the online majority voting scheme.