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

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 [Dore et al. 2009-2010]
- Incremental relabeling mechanism [Chao et al. 2011]

Disadvantage
- Cannot handle label noise during the labeling process.
- The label quality will be heavily affected 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)
- CMU-MMAC dataset (14 category of actions)

Method | Label treatment | Flip noise | Sample | Labelers
---|---|---|---|---
JGPC-ASAL (our) | Joint processing | With | Active | ASAL
JGPC-ASAS (our) | Joint processing | With | Active | ASAL
JGPC-ASRL (our) | Joint processing | With | Random | ASAL
JGPC-ASRF (our) | Joint processing | With | Random | ASAL
GPC-NVLS-F | Majority voting | With | Active | -
GPC-NVLS-R | Majority voting | With | Random | -
GPC-NVLS-M | Majority Voting | Without | Active | -
GPC-NVLS-RM | Majority Voting | Without | Random | -

(Note: GPC-NVLS-K/GPC-NVLS-K/GPC are proposed by Ashish Kapoor et al [ICCV 2009]

Comparison

Active learning strategy

E-step: Given the current parameter $\theta$, conduct EP inference to obtain and approximate inference of $P(S|X)$. M-step: Maximize the lower bound of $\log (P(X|T_S)P(S))$ over $\theta$ to obtain a new parameter $\theta^{*}$ and go to the E-step and iterate until convergence.

Parameter estimation

$\hat{\theta}^{*} = \arg\max_{\theta} \{ \log P(X|T_S)P(S) \}$

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

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

Experiments with real labels

Our JGPC-ASAL is constantly ranks on the top.

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

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.

Sponsors

Google, Research, NEC, Stevens Institute of Technology