A Novel Context-Aware Topic Model for Category Discovery in Natural Scenes

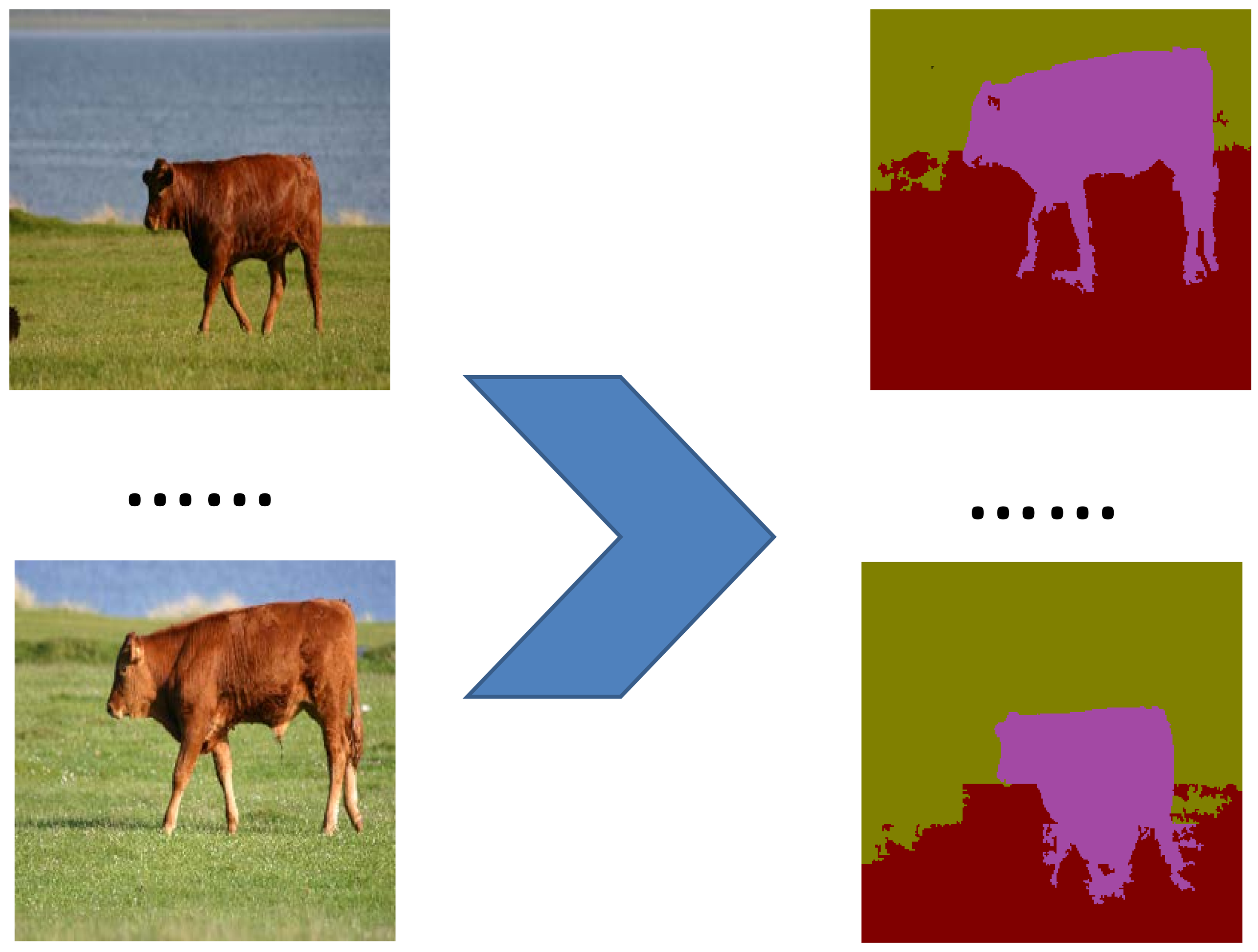
Zehuan Yuan, Tong Lu*,
National Key Lab for Novel Software Technology, Nanjing University, China

Task

Our target is to recognize visually similar categories and segment out their various instances by directly mining an unlabeled image set.

Contribution

1. We put forward a novel context-aware topic model (NCA-TM) by integrating multi-level image features, and
2. Spatial preference of categories is characterized in a more flexible way for assisting category discovery from complex natural scene images.



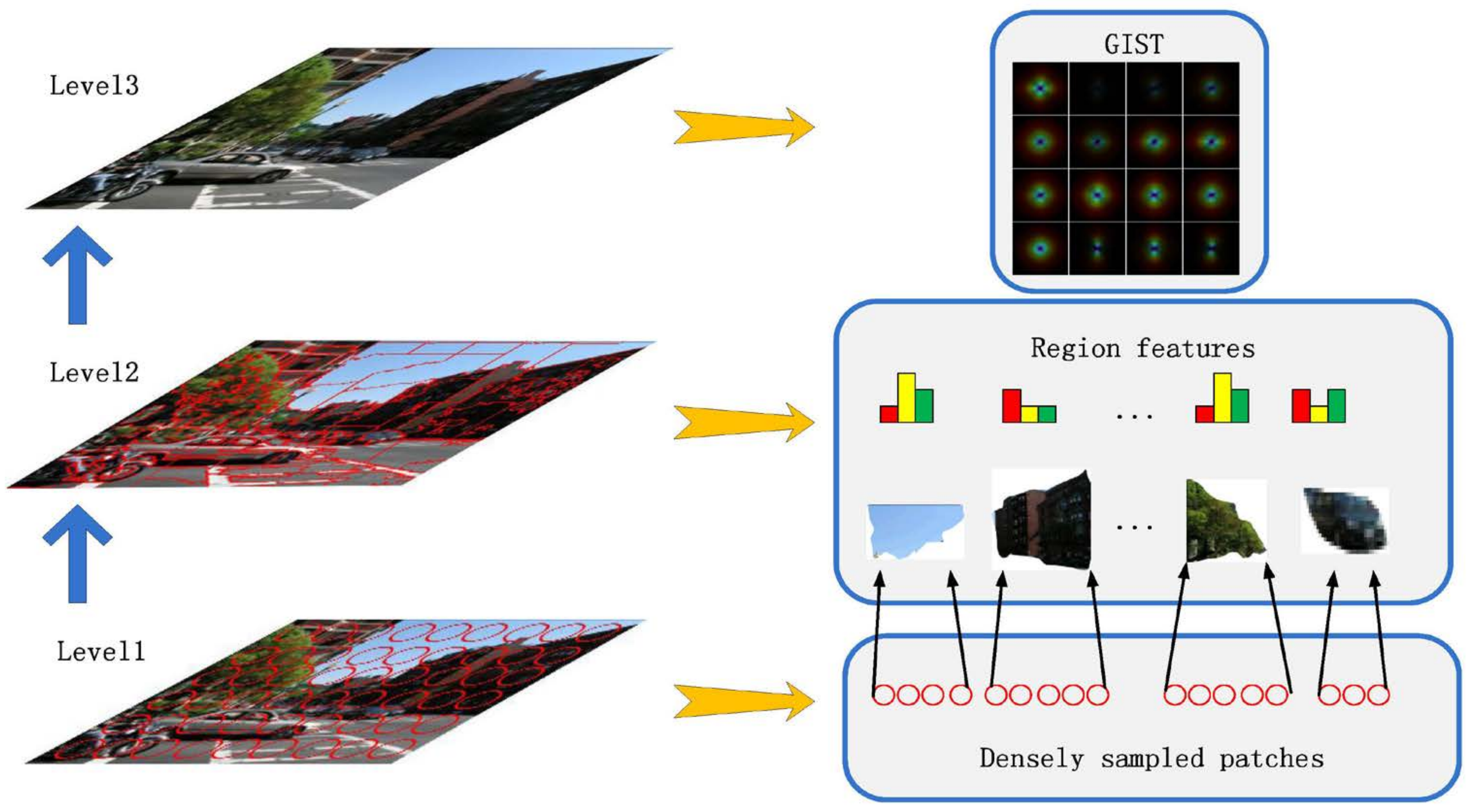


Image description

The first level are the features of dense image patches. The middle level corresponds to regions features and the top level has only one GIST feature of this image.

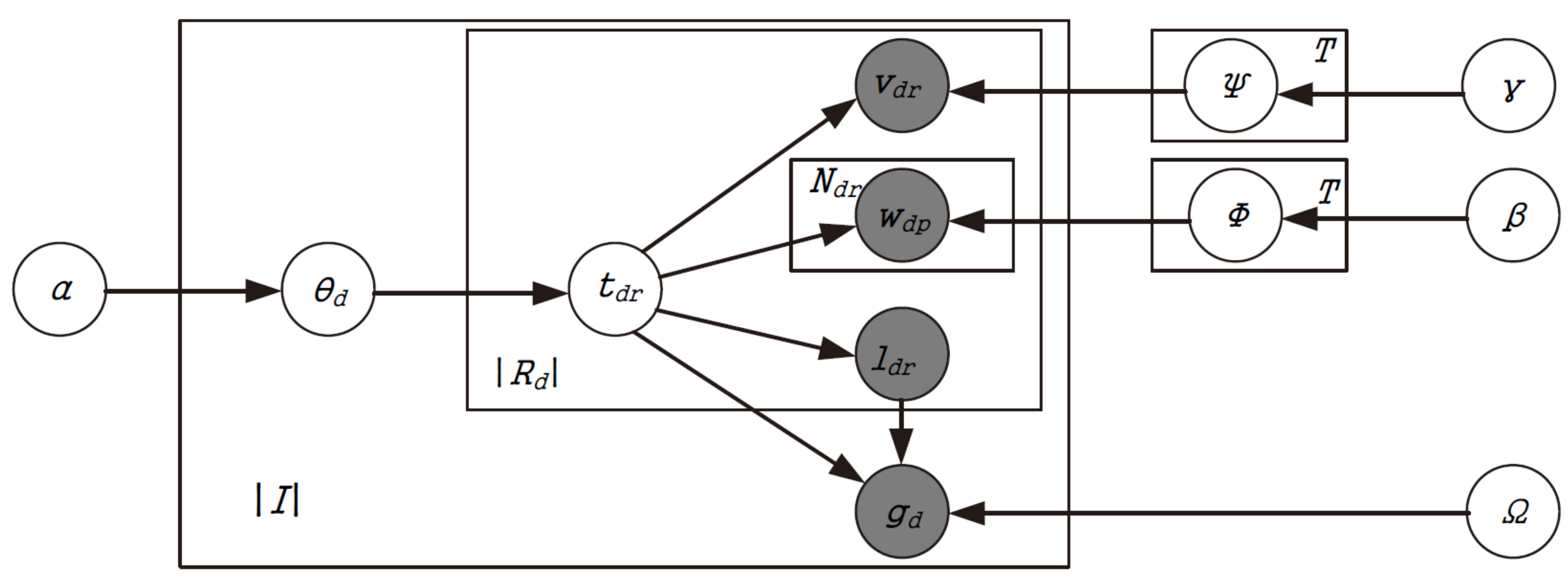
Table 1. The generative process of our NCA-TM

For each topic t, sample $\Phi_t \sim \mathrm{Dir}(\beta)$ and $\psi_t \sim \mathrm{Dir}(\gamma)$;
For each image I_d,sample its topic distribution $\theta_d \sim \mathrm{Dir}(\alpha)$;
For each region $R_r \in I_d$, sample $t_{dr} \sim \mathrm{Multi}(\theta_d)$,$v_{dr} \sim \mathrm{Multi}(\psi_{t_{dr}})$;
For each patch $P_p \in R_r$, sample its visual word $\omega_{dp} \sim \mathrm{Multi}(\Phi_{t_{dr}})$;
Given all sampled $\boldsymbol{t_d}$, sample $g_d \sim P(g_d|F_d(\boldsymbol{t_d}, \boldsymbol{l_d}))$.

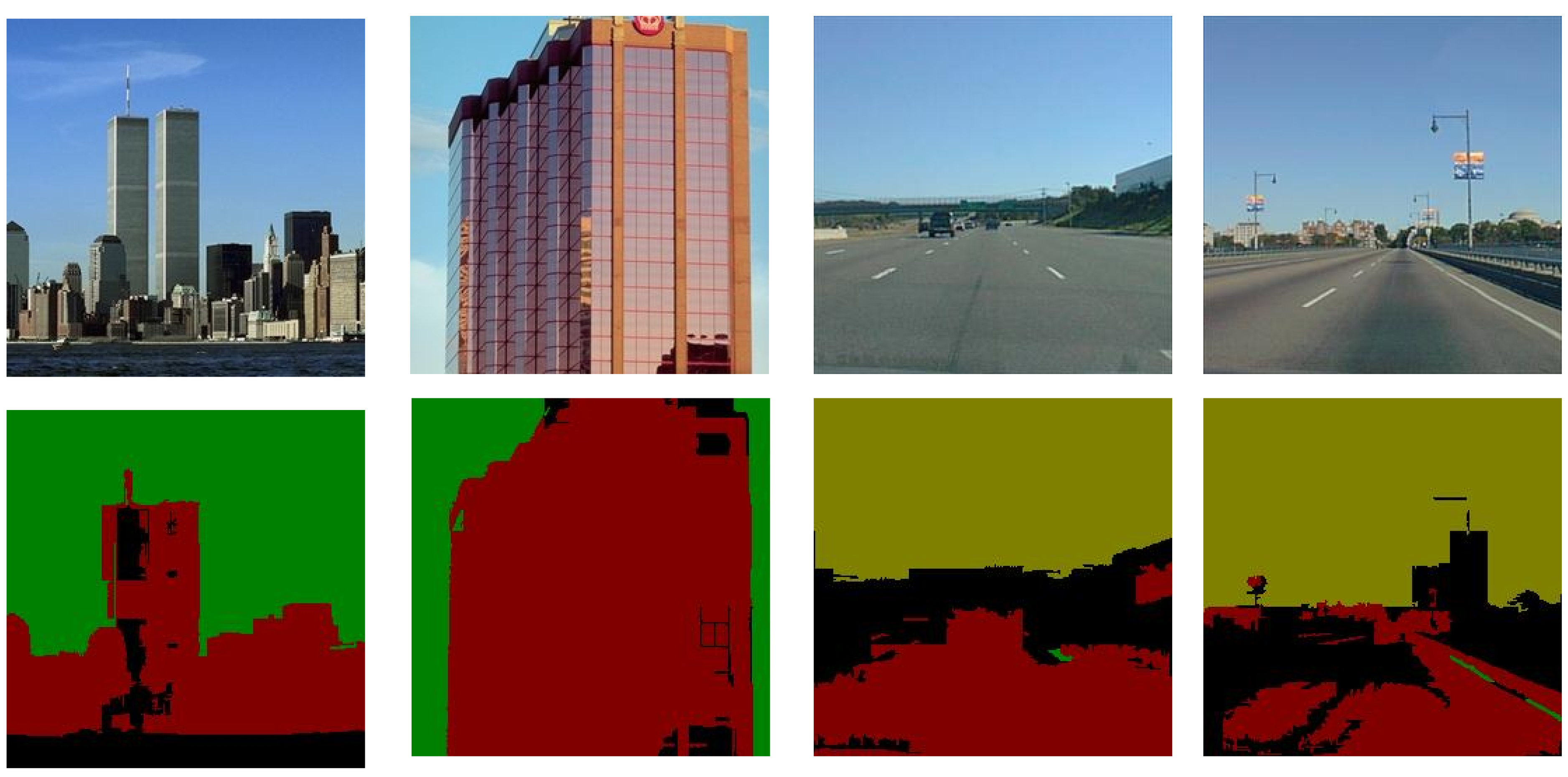
The proposed generative model

After the three-level representation of any image is generated, we further construct a generative probabilistic model (NCA-TM) to derive these observations

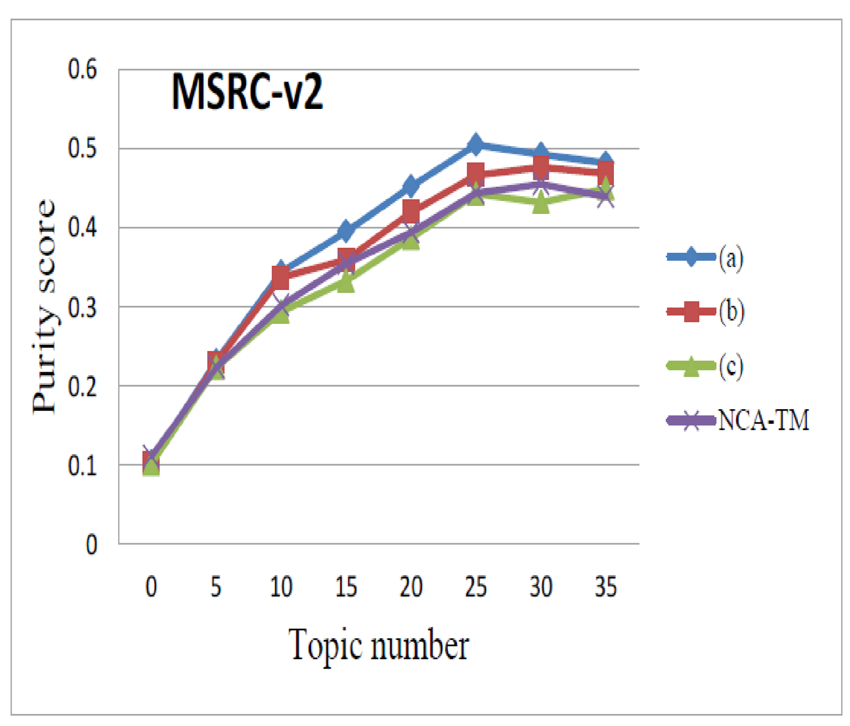
$$P(\boldsymbol{\omega}, \boldsymbol{v}, \boldsymbol{l}, \boldsymbol{g}, \boldsymbol{t}) = \iint \prod_{k} P(\psi_k|\gamma)P(\Phi_k|\beta) \prod_{d}^{|I|} P(\theta_d|\alpha) \prod_{r}^{|R|} P(t_{dr}|\theta_d)$$
$$P(v_{dr}|\psi_{t_{dr}})\, P(l_{dr}|t_{dr}) \prod_{p\in r} P(\omega_{dp}|\Phi_{t_{dr}})\, P(g_d|F_d(\boldsymbol{t_d}, \boldsymbol{l_d}))\, d\theta\, d\Phi\, d\psi$$

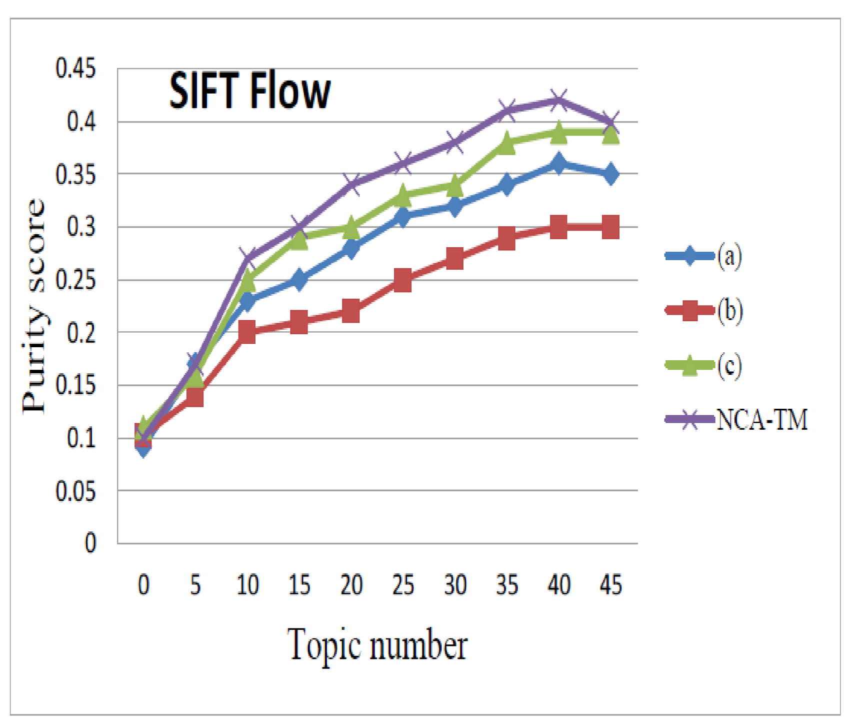


We adopt the simple generalized linear model to generate $P(g_d|F_d(\boldsymbol{t_d}, \boldsymbol{l_d}))$. The goal of category discovery corresponds to the inference of the graphical model, namely, maximizing the posterior distribution of latent variables given all observations $P(\boldsymbol{t}|\boldsymbol{\omega}, \boldsymbol{l}, \boldsymbol{g}, \boldsymbol{v}; \Omega, \alpha, \beta, \gamma)$. We adopt a Gibbs EM algorithm [1] to the model inference and parameter learning.



Results





[1]. Andrieu, C., de Freitas, N., Doucet, A., Jordan, M.I.: An introduction to mcmc for machine learning. Machine Learning **50** (2003) 5–43