Reverse-Inference for Meta-Analysis: Predicting Study Type with a Fully Bayesian Spatial Model

Jian Kang[1], Timothy Johnson[1], Thomas Nichols[2], Tor Wager[3], Lisa Barrett[4]

1. University of Michigan Biostatistics, Ann Arbor, MI. 2. University of Warwick, Dept. of Statistics, Coventry, United Kingdom. 3. University of Colorado Psychology, Boulder, CO. 4. Northeastern University Psychology & Mass General Hospital, Harvard Medical School, Boston, MA

Introduction

Meta-analysis for functional brain imaging is growing in importance as the number of available studies grows. Methods for meta-analysis, like ALE (Eickohoff, 2009) or MKDA (Wager, 2009), use reported peak activation foci to localize consistent effects over studies. However, these methods are mass-univariate and require fixed tuning parameters; they also do not provide an interpretable fitted model, and cannot produce spatial confidence intervals on location of activation.

In recent work (Kang, 2012) we produced a fully Bayesian Hierarchical Spatial model for neuroimaging meta-analysis. We explicitly model the clustering of study foci about latent activation centers (accounting for multiple peaks within a study describing a single anatomical region), which in turn cluster about latent population centers. In particular, our model can dissociate between inter-study spread of foci, and spatial uncertainty in population response location.

In the present work we use our fully Bayesian model to build a classifier of study type based on reported foci. We demonstrate our method on a meta-analysis of emotion, classifying different sub-types of emotion, and compare to a Naive Bayesian Classifier.

Methods

Model. See (Kang, 2012) for our single-group model. In brief, we have a fully 3D hierarchical spatial model that accounts for: intra-study variation in location (important when many foci are reported for essentially the same region), and inter-study variation about latent population centers. The key parameters are intensity maps; e.g. the posterior population intensity map shows locations about which studies' foci cluster; the intensity map is not a density, and this particular map integrates to the posterior expected number of population centers.

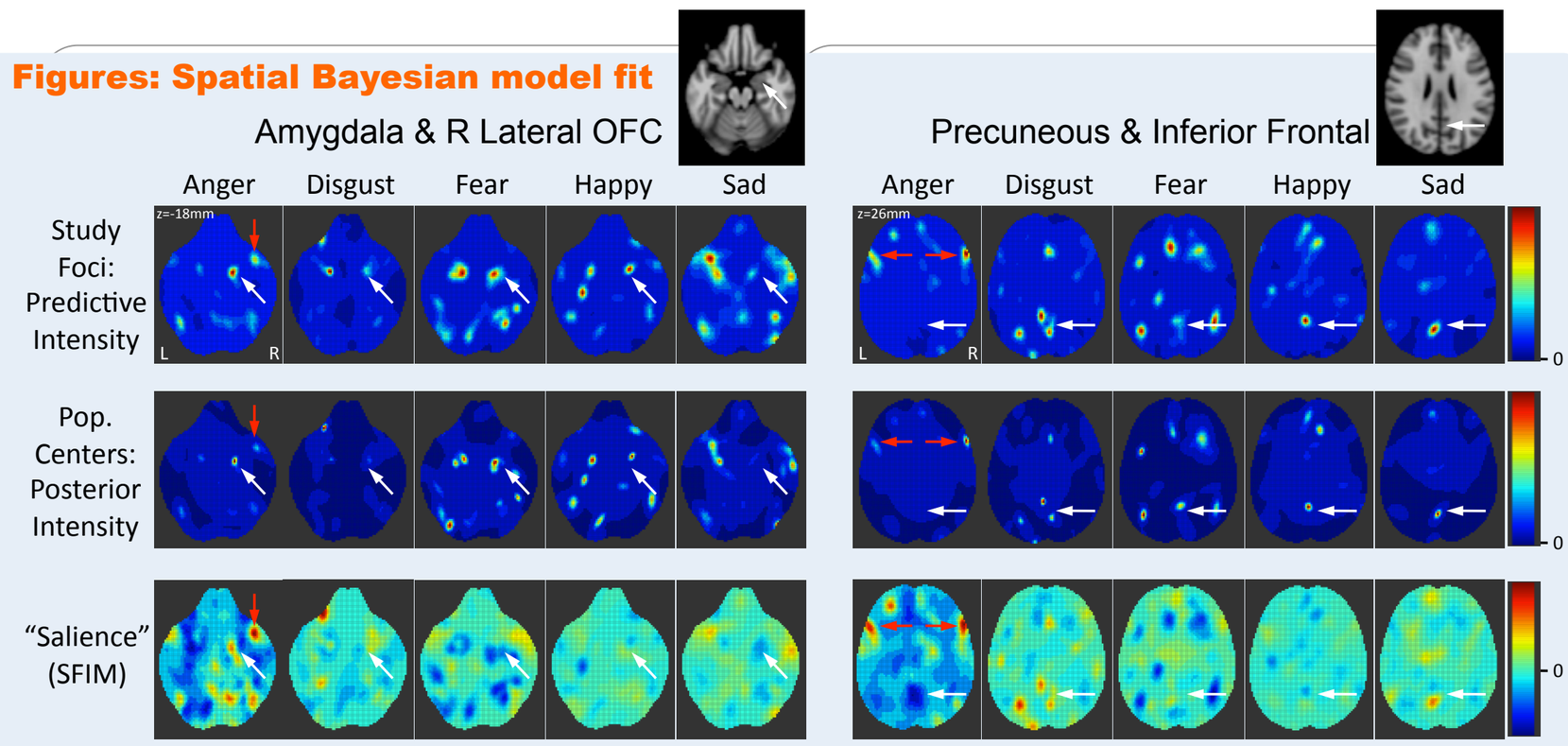
We have extended our model to fit K groups of studies, which allows us to predict the group label of a new study. Let $\phi(\boldsymbol{x}_{\text{new}},k)$ be the predictive probability that a new study's points $\boldsymbol{x}_{\text{new}}$ arise from category k; the predicted category is the k that maximizes $\phi(\boldsymbol{x}_{\text{new}},k)$. We use leave one out cross validation (LOOCV) to measure the accuracy; an importance sampling approach is used to avoid multiple posterior simulations.

We compare our method to a Naive Bayesian Classifier (NBC), where the point pattern $\boldsymbol{x}$ is converted to a voxel-wise binary map as in MKDA (kernel radius 10mm).

Salience & SFIM's. Since our model is multivariate and considers the entire point pattern $\boldsymbol{x}$ simultaneously, there is no unique way to measure the contribution of a single voxel. To understand the role of individual locations in our classification we compute a Single Foci Impact Map (SFIM) as follows: At voxel i the SFIM is the change in ϕ with the addition of an activation focus at voxel i. Specifically, for one randomly generated $\boldsymbol{x}_{\text{new}}$ we compute ϕ twice, before and after adding a point at i, and subtract; the final SFIM is averaged over many $\boldsymbol{x}_{\text{new}}$ simulated from the fitted model. Thus voxels with activations occurring uniquely for a category will have positive SFIM, while if a category is associated with a unique absence of points at a location, the SFIM will be negative.

Data. We use a meta-analysis of emotion (Kober, 2008), comprised of a total of 437 studies. We considered only the subset of 219 studies classified as: Anger, Disgust, Fear, Happy or Sad.

Figures: Spatial Bayesian model fit



Spatial Bayesian model fit for two locations, for each of the five study types. **Top**: Posterior predictive intensity, a prediction of a new study's foci for each emotion. **Middle:** Population center intensity, showing where there is evidence for latent population centers. **Bottom**: Single Focus Impact Map, a version of salience. **Left:** Every emotion has some amygdala foci, and so that area is not particularly salient for this 5-way classification (white arrow), while R lateral OFC is unique to Anger and thus is very salient for that study type (red arrow). **Right**: Precuneous is seen in all emotions except Anger, and thus this region is negatively salient for Anger (white arrow), while bilateral IFG is unique and positively salient (red arrow).

Tables: LOOCV Results – Naïve Bayes Classifier vs. Spatial Bayesian Model

Spatial Bayesian Model

Truth ↓	Prediction				
	sad	happy	anger	fear	disgust
sad	0.78	0.00	0.11	0.04	0.07
happy	0.06	0.92	0.00	0.03	0.00
anger	0.08	0.08	0.69	0.15	0.00
fear	0.13	0.01	0.00	0.85	0.00
disgust	0.05	0.02	0.02	0.07	0.84

Average Accuracy:

Spat. Bayes. Mod. **82%**

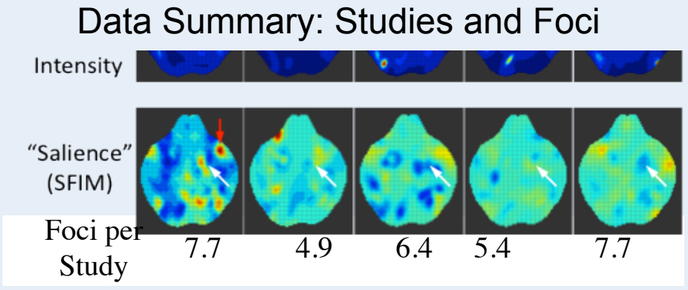
Naïve Bayes. Class. **36%**

Naïve Bayesian Classifier

Truth ↓	Prediction				
	sad	happy	anger	fear	disgust
sad	0.38	0.11	0.07	0.40	0.04
happy	0.11	0.25	0.03	0.56	0.06
anger	0.12	0.23	0.00	0.50	0.15
fear	0.06	0.06	0.01	0.81	0.06
disgust	0.09	0.16	0.05	0.32	0.39

Leave one out cross-validation (LOOCV) results for each method. Accuracy averaged over the five categories was 82% for our model, 36% for the Naïve Bayesian Classifier (NBC). While our model never has per-category accuracy less than 69%, NBC has 0% accuracy for Anger studies, misclassifying them mostly as Fear. NBC is perhaps challenged by the sparse data, as there are fewer than 8 foci per study on average. Changing the kernel radius to 15mm or 20mm when creating the voxel-wise maps only worsened NBC's accuracy.

Data Summary: Studies and Foci



Results

See Figures and Tables.

Conclusions

We have extended our single-group meta-analysis method to consider multiple groups and to allow classification of new studies. Our method obtains accuracy rates twice those of the Naive Bayesian Classifier. This difference is likely due to our model's more accurate representation of the true data generating process, relative to the MKDA kernel mapping of points to voxel-wise images. Also, our fully Bayesian model captures more sources of variation, and appropriately conveys the certainty (or lack there of) in the computation of the predictive probabilities that determine the classification outcome.

References

Kang et al. (2011). *J. Am. Stat. Assoc.*, 106(493): 124-134, 2011.
Eickhoff et al. (2009). *Hum. Brain Map.,* 30(9), 2907-26.
Wager et al (2009). *NeuroImage*, 45(1S1), 210-221.
Kober et al. (2008). *NeuroImage*, 42, 998–1031.

t.e.nichols@warwick.ac.uk

http://go.warwick.ac.uk/tenichols