Exploiting Spatial Information to Improve fMRI Pattern Classification

Melissa K. Carroll[1], Kenneth A. Norman[2], James V. Haxby[2], Robert E. Schapire[1]

Princeton University Departments of Computer Science[1] and Psychology[2]

12th Annual Meeting of the Organization for Human Brain Mapping, Florence, Italy, June 14, 2006

1 Introduction

- Classification methods have been successfully applied to pattern extraction from fMRI (e.g. [1,2]).
- Most classification approaches have treated individual voxels as features, ignoring the spatial correlation of activity between voxels.
- The present method, adapted from computer vision, incorporates spatial information via:
  1. **Richer features that capture correlation between adjacent regions**
  2. **AdaBoost as a multivariate feature selector**
- This method can improve classification accuracy and has the potential for discerning which types of neural features are most useful for discriminating between cognitive states.

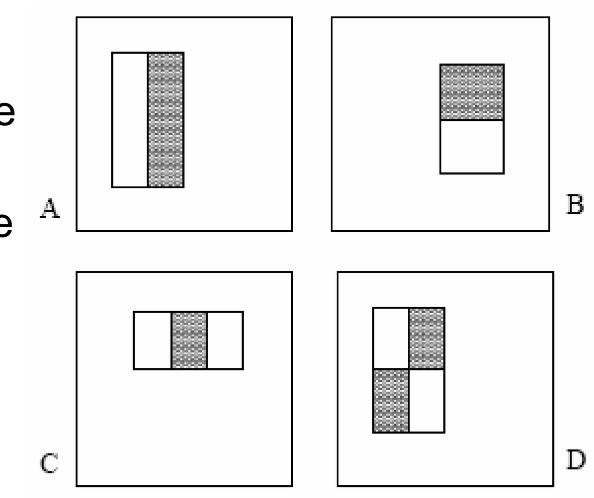
Methods

2 Neuroimaging Methods

- Two subjects underwent fMRI studies on a 3.0 Tesla scanner while performing a 1-back recognition task of images from seven categories:
  - female and male human faces
  - monkey and dog faces
  - houses, shoes, and chairs
- 8 runs of 10 2-second TR intervals for each of the 7 stimuli classes were obtained.
  - First 10 TRs out of 17 were selected due to *adaptation effect*
- Cortical surface mapping was performed to produce a 2D image reflecting spatial adjacencies.
- Analyses were confined to the Ventral Temporal region.

3 Image Features

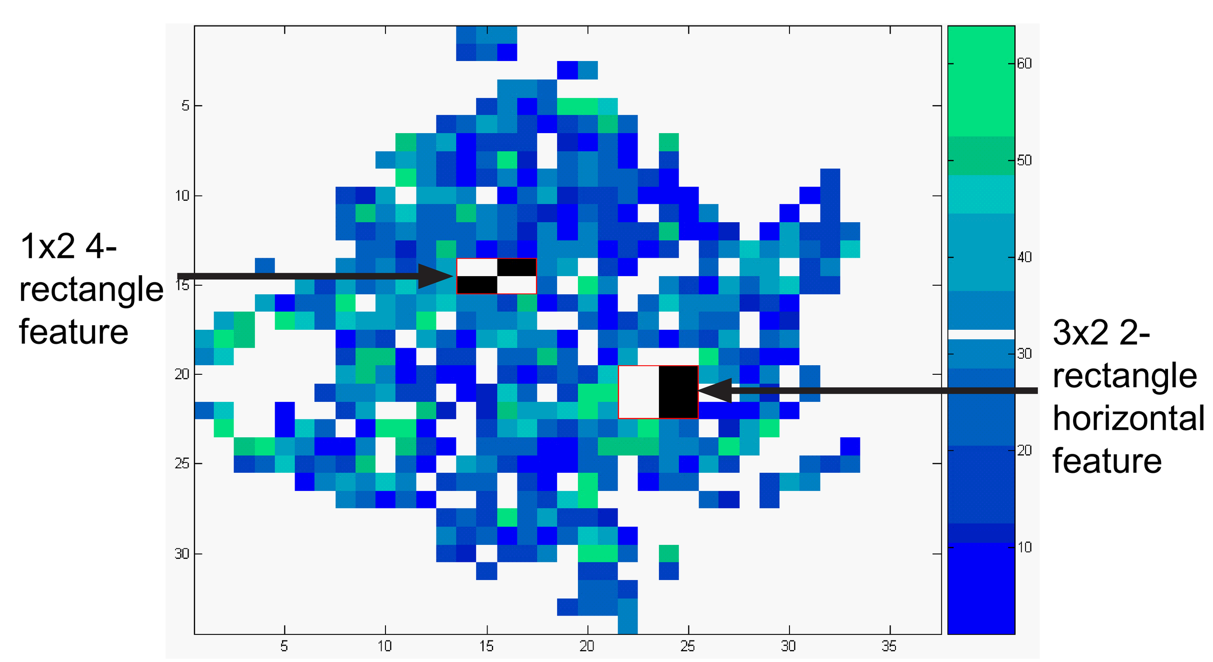
- Instead of individual pixels, richer features are used for classification.
- Mean activity in the white regions are subtracted from mean activity in the gray regions.
- Features are characterized by:
  - number of rectangles (1-4)
  - orientation
  - size (vert. and hoz.)
  - position within image



a) 2-rectangle horizontal
b) 2-rectangle vertical
c) 3-rectangle horizontal
d) 4-rectangle
Not shown: (1-rectangle and 3-rectangle vertical)

Features derived from Viola and Jones, CVPR, 2001

4 Example Features on Actual Image



One subject's left hemisphere ventral temporal region at one TR

5 AdaBoost as Feature Selector and Learning Algorithm

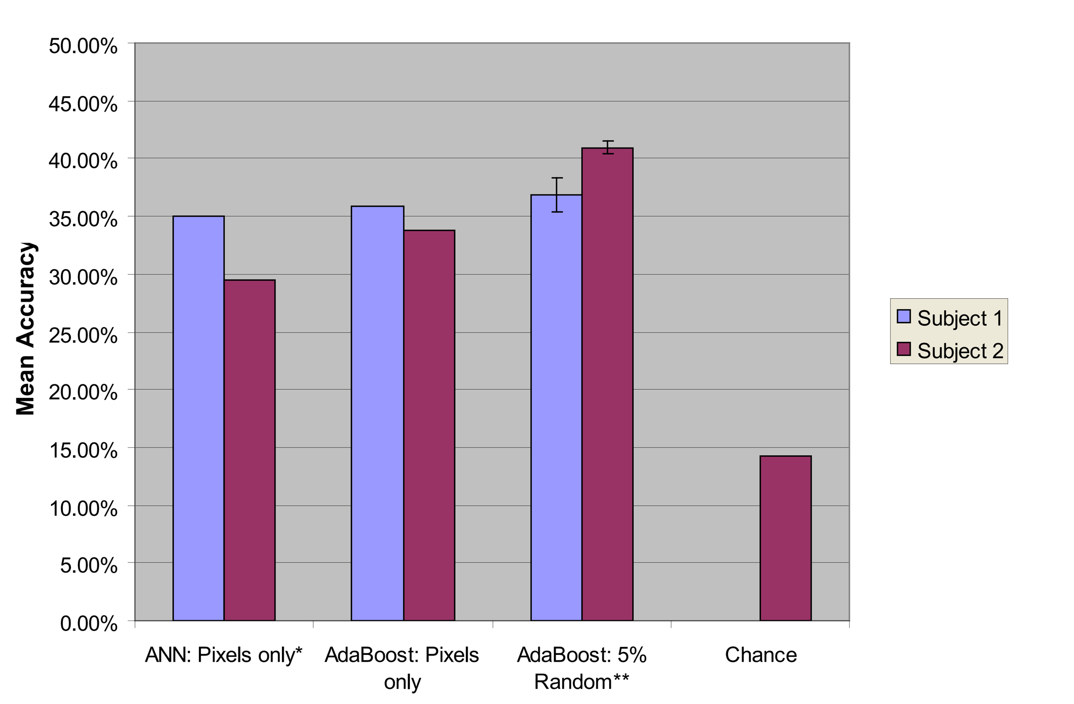
- Performs multivariate feature selection
- Theoretically and empirically less sensitive to large feature set sizes
- Weak classifier: binary classifier that is slightly better than random guessing
- Basic idea: combine many weak classifiers into a strong classifier
- Algorithm (Freund and Schapire, 1996)
  - Assign uniform weights to training instances
  - On each of *T* rounds
    - Select new weak binary classifier based on learning algorithm
    - Increase relative weights of instances incorrectly classified
    - Assign weight to the weak classifier based on training accuracy
  - Final classifier: weighted vote over weak classifier outputs

6 Algorithm Evaluation Experiments

- Instances: individual TRs from same subject
- Target classes: 7 object categories
- Training: 1000 AdaBoost rounds over 8 "leave one run out" cross-validation runs
- Weak classifiers: thresholded features from all permutations of:
  - 4 numbers of rectangles + 2 orientations if applicable = 6 types
  - all 100 size combinations between 1x1 and 10x10
  - all positions in image
- Due to enormous feature space (1.2 million features), random feature selection was performed
  - Chose percentage of total possible feature set size
  - Selected feature type and size permutations randomly
  - Computed for all positions in both hemispheres

Results

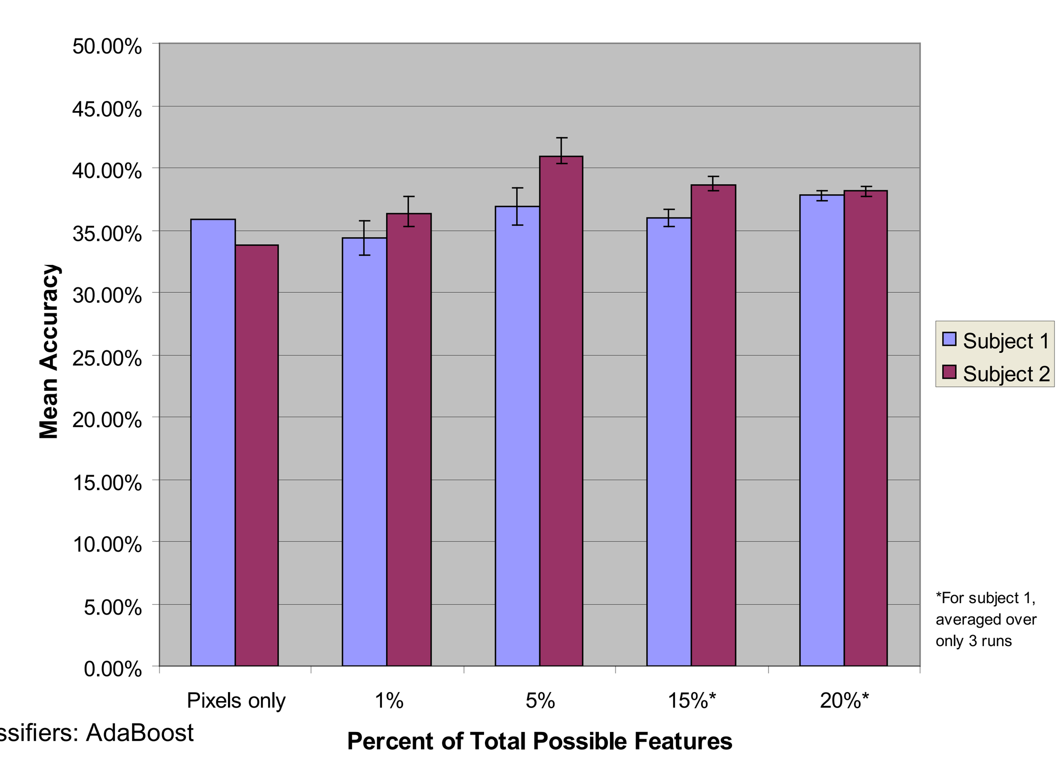
7 Comparisons: AdaBoost vs. Artificial Neural Networks, Richer Features vs. Single Pixels



*ANOVA with p = 0.01 was pre-applied

**averaged over 5 randomly generated feature sets

8 Performance by Feature Set Size



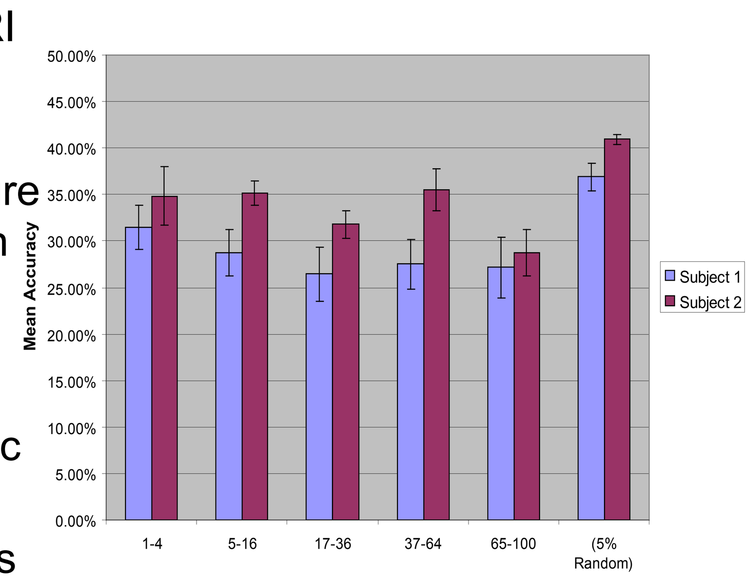
All classifiers: AdaBoost

9 EBC Competition Data

- Raw pixels with *no feature selection*: AdaBoost versus ANN
  - Subject 1: AdaBoost better on 29 of 30 regressors; average relative accuracy change: 30.6%
  - Subject 2: AdaBoost better on 28 of 30 regressors; average relative accuracy change: 24.2%
- 5 random 5% AdaBoost runs versus raw pixels
  - Subject 1: Richer features better on 57% of *runs;* average relative accuracy change: .3%
  - Subject 2: Richer features better on 67% of *runs*; average relative accuracy change: 1%

10 Accuracy by Rectangle Area

- The optimal spatial frequency for fMRI classification is unknown
- The random feature set classifiers with AdaBoost outperform classifiers using features of specific sizes
- AdaBoost appears to "find" the important features



11 Conclusions

- Novel approach for classifying fMRI images
  - Use of features that capture spatial information
  - Multivariate feature selection
- Potential benefits
  - Improve classification accuracy directly
  - Improve classification indirectly by revealing important features
  - Useful test-bed for exploring neuroscientific questions
    - e.g. What is the optimal spatial frequency for classification?
- Generally a moderate classification improvement
  - Varies across subjects
  - Not clear if justifies added computation
- More exploration of feature importance is warranted

12 References

1. Mitchell, Tom M., Hutchinson, Rebecca, Niculescu, Radu S., Pereira, Francisco, Wang, Xuerui, Just, Marcel, Newman, Sharlene. (2004). *Learning to Decode Cognitive States from Brain Images*. Machine Learning (57 1-2: 145 – 175).
2. Polyn, Sean M., Natu, Vaidehi S., Cohen, Jonathan D., Norman, Kenneth A. (2005) *Category-Specific Cortical Activity Precedes Retrieval During Memory Search*. Science (310: 1963-1966).
3. Viola, Paul, Jones, Michael. (2001). *Rapid object detection using a boosted cascade of simple features*. CVPR.
4. Freund, Yoav, Schapire, Robert E. (1996). *Experiments with a New Boosting Algorithm*, in Proceedings of the Thirteenth International Conference on Machine Learning (148-156). Morgan Kaufmann

An electronic copy of this poster is available at http://compmem.princeton.edu/publications.html