Detecting Chronic Pain with Structural MRIs

Hoameng Ung
hoameng.ung@stanford.edu
CS229, Autumn 2009, Stanford University

Abstract

The subjective nature of pain makes an objective measurement difficult to produce. If such a measurement is possible, then this knowledge would assist in the management of patients whose self report of pain is either unavailable or called into question. Past studies have demonstrated that chronic pain is associated with gray matter (GM) atrophy, detectable by magnetic resonance imaging. We investigate to see if this decline of GM density can be used to distinguish between healthy and chronic low back pain (LBP) subjects with greater than chance accuracy. Structural MR scans from 94 age and gender matched subjects were used. Several feature selection and reduction methods were performed, including grid and ROI parcellation, principal component analysis (PCA), and ranking by mutual information. Accuracy, assessed with both 1-norm and 2-norm SVM using striated 10-fold cross-validation, while significant, were low and corresponding features driving the classifier were inconsistent. Parcellation gave a high of 64% accuracy, while PCA and MI gave 95% and 97%, respectively.

1. Introduction

An estimated 50 million Americans suffer from chronic pain, and the costs of low back pain in the United States exceed $100 billion per year (1). Since pain is subjective, self-reported pain cannot be verified. Because of this, an objective measure of pain could assist in the management of patients who are suspected of malingering or drug seeking and of patients unable to communicate such as the critically ill. Studies have shown that regional changes in gray matter density resulting from chronic pain are detectable by MRI (2) (3).

Recently, machine learning techniques have been adopted in the field of neuroimaging in an attempt to learn about cognitive science and to detect conditions based on imaging data. Some examples include the detection of Alzheimer's and Huntington's disease using SVMs and gray matter density as features (1) (2). However, applications to detection of diseases such as Alzheimer's and Huntington's are fairly trivial tasks due to the etiology of the disease and its tremendous effect on the structure of the brain with time. The etiology of chronic pain is relatively more complicated. The differences in the type of pain have been shown to affect the underlying biological changes. For instance, Apkarian demonstrated a larger decline in gray matter density with neuropathic chronic pain patients versus nonneuropathic patients through Voxel Based Morphometry (VBM) analysis techniques (3).

VBM is currently the standard method used to study differences in tissue classes between subjects with structural MRI data (4). It essentially allows voxel-by-voxel comparison of subjects by transforming each brain to a template and then performing statistical tests on the voxel level. The advantage of using an SVM as opposed to traditional techniques of VBM is that an SVM is can capture higher dimensional relationships within the data, rather than comparing between individual voxels as in VBM. In addition, SVMs allow for classification of one sample, whereas VBM

analyses are generally limited to comparisons between groups of subjects.

2. Materials and Methods

2.1. Dataset

The dataset includes 94 gender and age-matched subjects (47 controls, 47 patients). Chronic pain patients were defined as having greater than 6 months of pain. Each scan was segmented into gray matter and normalized to the MNI template. Each example contained roughly 300,000 features, with each feature corresponding to the gray matter density at a particular voxel. Each feature was normalized to a Euclidean distance of 1.

2.2. Feature Selection and Reduction

Four different feature reduction/selection methods were employed: Two types of parcellation, principal components, and ranking based on mutual information.

2.2.1. Grid parcellation

Each gray matter map was divided into 5x5x5 voxel volumes, and the gray matter density averaged within each. Each volume represented 1 cm^3 cubes. To prevent skewed volumes, volumes with fewer than 10 voxels were averaged into the nearest volume with greater than or equal to 10 voxels. This method reduced the number of features to roughly 2100.

2.2.2. AAL Parcellation

The Automatic Anatomical Labeling template divides the brain into 116 different ROIs based on anatomical structure. Each gray matter map was divided into these ROIs and the gray matter density averaged within each, producing 116 features.

2.2.3. Principal components

Principal component analysis (PCA) was used to determine the directionality of the training data. The dataset was transformed into principal component space, consisting of 93 PCs. The weights of the feature vector, corresponding to weights of each principal component, was transformed back into brain space by a multiplication by the eigenvectors (principal components) and the most highly weighted individual voxels were found. This method is similar to that employed in (5).

2.2.4. Mutual Information

Mutual information (MI) was calculated for each of the original voxels with the labels. Intensity values were discretized to 100 values in the range [0, 1] and label {-1, 1}. Ranking was performed using a forward search wrapper method, adding the highest MI valued features at a time until the error from leave-one-out cross-validation (LOOCV) increased. This resulted in the selection of 55 individual voxels.

2.3. Classification

Both an L1 and L2 regularized SVM (with a linear kernel) was used with each feature reduction method. The MATLAB SVM implementation was developed by Anton Schwaighofer and other code was based on in-house implementations. C parameter selection was done using a standard grid search. 10-fold striated cross-validation was used to assess the performance of the classifier. Striated k-fold cross-validation involves selecting an equal ratio of examples from each class representative of the dataset on a whole and has been shown to be more accurate at assessing performance (7).

The significance of the performance was assessed using a Monte Carlo permutation test, iterated 1000 times. This test involves

permuting the labels and rerunning the training and testing to determine if the separation was based on meaningful patterns between the two classes.

2.4. Regression

Linear regression was performed using elastic net regularization with grid parcellated features. A MATLAB implementation of Least Angle Regression (LARS) by Sjöstrand was used. The dependent variable was the duration in pain (years). The model was trained on 70% of the data and tested on the remaining 30%. The resulting beta weights were used to determine the volumes driving the classifier.

3. Results

3.1. Classification Accuracies

Accuracies for grid and AAL parcellation were marginally better than chance. Feature reduction by PCA resulted in high accuracies, as well as selection of voxels by computation of the MI values (Table 1). With PCA, 100% accuracy was achieved based on the first 93 principal components, and 97% with MI. The linear regression by elastic net resulted in a sum of squared errors of 5.29 (years).

Differences in classification accuracy between 1-norm and 2-norm SVMs were marginal at best.

3.2. Neuronal correlates

The weights obtained by training with the SVM indicate the features most important in determining whether an example belongs to a specific class. For AAL parcellation, highly weighted features include the right pallidum (Figure 1a). For grid parcellation, the highly weighted features includes a section of the left thalamus, a part of the ventromedial prefrontal cortex (VMPC), and a part of the dorsolateral prefrontal cortex (DLPC) (Figure 1b). PCA and MI gave individual voxels that were delocalized throughout the brain (not shown). Regression by elastic net performed on grid parcellated features have similarities to the weights given by training with the SVM (Figure 1c).

Feature Selection Method	1-norm SVM	2-norm SVM
Grid (1700)	64%	61%
AAL (116)	62%	64%
PCA (93)	100%	100%
MI (55)	97%	96%

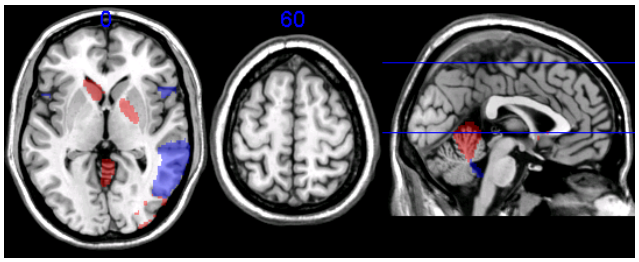
Table 1. Classification accuracies based on feature selection methods (with its corresponding number of features) and 1, 2-norm SVMs. All values are significant $p<0.05$

4. Discussion

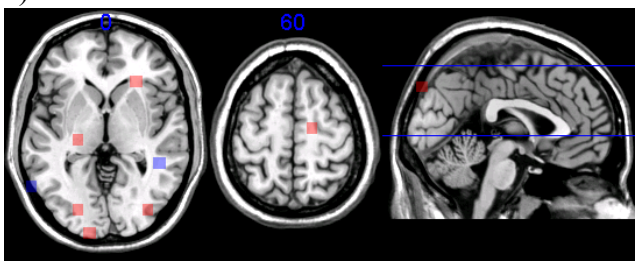
The accuracies obtained for parcellation are poor, but because it is significant, this indicates that something can be learned about the dataset provided.

The AAL template is commonly used to select predefined ROIs, but the lack of specificity may explain the low accuracy. The large divisions of the brain will mask any regional differences within each ROI. For instance, different parts of the cingulate cortex are known to have different functions, so grouping them all and averaging their gray matter densities may smooth out any existing differences. Nonetheless, areas that are much smaller are captured by the AAL divisions, such as the pallidum. Since the pallidum is involved in thalamic pathways, such as inhibiting the relaying of sensory stimuli from specific nuclei, overstimulation may lead to changes.

a)



b)



c)

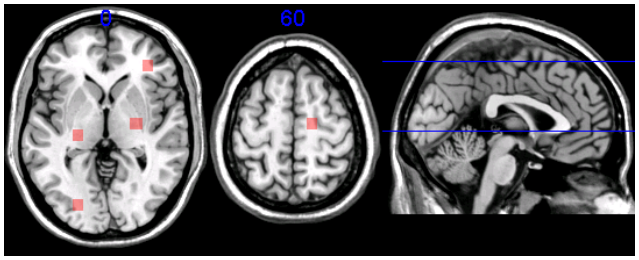


Figure 1. Highly weighted features by a) AAL parcellation b) grid parcellation and c) elastic net regression

Grid parcellation was motivated by prior knowledge about the functions of anatomical parts in the brain. With chronic pain, certain structural areas of the brain are expected to be affected. Thus, parcellation into 1 cm^3 volumes is small enough to capture areas of interest. The thalamus is believed to be a site that relays sensory information received from the body towards different cortices in the brain and is believed to play a role in mediating pain signals (7) (8) (9). Nociceptive stimuli synapse at the thalamus before being sent to parts of the prefrontal cortex. For this reason, an overload of sensory stimulus at the thalamus can cause a decrease in gray matter density, a result of suppressed neuronal activity from a decrease in membrane receptors, neuronal cell degeneration, or neuron apoptosis. The prefrontal cortex has also been shown to mediate chronic pain. Decreases in DLPC gray matter density align with previous findings (3).

Principal component analysis enables very high accuracy when performed on the original features. Upon transformation of the weight vector back into brain space, the individual voxels that were highly weighted were delocalized and did not show any comprehensible pattern. The first principal component explains only 18% of the variance, and subsequent components much less. The need for many components to explain the data may indicate a low signal to noise ratio, and the components may not capture conditions inherent in the underlying condition.

Feature ranking by mutual information was performed on the original feature set, and, through cross-validation, resulted in the selection of 55 individual voxels. These voxels were, similar to the PCA results, scattered throughout the brain and were unhelpful in determining correlations with chronic pain and brain structure.

The use of a 1-norm SVM has been shown to give better performance in the presence of feature redundancies (10). Thus, if multiple areas of the brain undergo atrophy, then this would be a redundant feature and performance should improve with the use of an L1 norm penalty. Although a 1-norm SVM gave higher accuracies in certain cases, the improvement was marginal at best.

Since duration of chronic pain also helps determine the amount of gray matter loss (3), a regression where the dependent variable is duration of pain may help determine areas of the brain affected by LBP. The elastic net is a regularization and variable selection method that represents the data as a sparse matrix where a limited number of features are given weights and is thus useful when the number of predictors is greater than the number of

examples (14). The resulting beta weights closely aligned with the results from grid parcellation, indicating that those features are important in determining both the duration of pain with regression and the presence of chronic pain with classification.

Future directions would be to obtain more data in order to increase the signal to noise ratio. Furthermore, manual feature selection by defining ROIs may prove helpful. For instance, ROIs of the left thalamus, DLPC, and VMPC and classifying based only on voxels in those areas may be interesting. In addition, LBP encompasses a fairly large array of conditions, and each of these conditions may have different effects on the brain. A narrower subset of patients may help in classification.

5. Conclusion

Although classification accuracy was low for parcellated features, the resulting weights indicate that decreased gray matter density in the left thalamus, DLPC, and VLPC help predict incidence of chronic pain. The low accuracy of parcellation methods and the high accuracy of PCA and MI performed on original voxels indicate a low signal to noise ratio in the data possibly a result of the variability in the underlying LBP conditions. Better feature reduction methods, such as manual selection of ROIs, may be performed and more data gathered in order to reduce this noise. Nonetheless, the significance of the values gives a promising outlook that gray matter density may be used to detect chronic pain.

Acknowledgements

I would like to thank Catie Chang and Quoc Le for providing guidance in this project, as well as Professor Andrew Ng for instructing CS229. In addition, I would like to thank Dr. Sean Mackey and Dr. Amit Etkin for providing data used in this analysis.

References

1. *Lumbar Disc Disorders and Low-Back Pain: Socioeconomic Factors and Consequences.* **Katz, J.** 2006, The Journal of Bone and Joint Surgery, pp. 21-24.
2. *Chronic Back Pain is Associated with Decreased Prefrontal and Thalamic Gray Matter Density.* **Apkarian, AV.** 2004, The Journal of Neuroscience, pp. 10410-10415.
3. *Chronic pain may change the structure of the brain.* **May, Arne.** 2008, Pain, pp. 7-15.
4. *Automatic classification of MR scans in Alzheimer's disease.* **Klöppel, Stefan.** 2008, Brain, pp. 681-689.
5. *Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI.* **Magnin, B.** 2009, Neuroradiology, pp. 72-83.
6. *Voxel-based morphometry -- the methods.* **Ashburner, John.** 2000, Neuroimage, pp. 805-21.
7. *Classifying brain states and determining the discriminating activation patterns: Support Vector Machine on functional MRI data.* **Mourao-Miranda, Janaina.** 2005, NeuroImage.
8. *A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection.* **Kohavi, Ron.** 1995, International Joint Conference on Artificial Intelligence.
9. *Functional MRI Study of Thalamic and Cortical Activations Evoked by Cutaneous Heat, Cold, and Tactile Stimuli* . **David, Karen.** 1988, The Journal of Neurophysiology, pp. 1533-1546.
10. *Affective components and intensity of pain correlate with structural differences in gray matter in chronic back pain patients.* **Schmidth-Wilcke.** 2006, Pain, pp. 89-97.
11. *Abnormal brain chemistry in chronic back pain: an in vivo proton magnetic resonance spectroscopy study.* **Grachey, Igor.** 2000, Pain, pp. 7-18.
12. *1-norm Support Vector Machines.* **Zhu, J.** 2003, NIPS.
13. *Regularization and variable selection via the elastic net.* **Zuo, H.** 2005, Journal of the Royal Statistical Society B.