

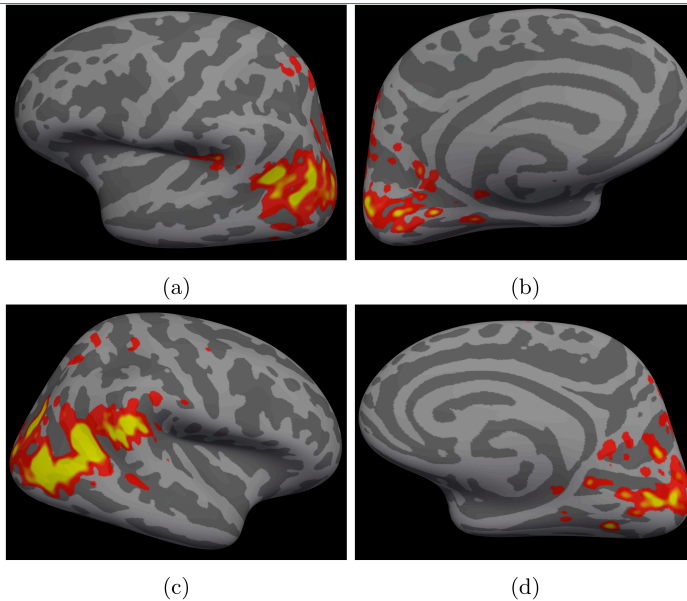
[Print this Page for Your Records](#)[Close Window](#)**Control/Tracking Number:** 2013-S-2742-SfN**Activity:** Scientific Abstract**Current Date/Time:** 5/8/2013 3:45:44 PM**Using sensitivity analysis to interpret machine learning applied to fMRI data****AUTHOR BLOCK:** \*A. FLOREN, B. NAYLOR, D. RESS;  
Univ. of Texas at Austin, Austin, TX**Abstract:**

Recently, machine learning has been employed as an analysis tool in fMRI. Algorithms are trained to classify or predict various stimuli from voxel time series. These flexible algorithms can potentially interpret complex and distributed patterns of activation in the brain. However, relating their results to brain functional anatomy is difficult.

Classifier performance is bounded by the mutual information between the voxels and stimulus. If some set of voxels can be used to predict the stimulus then some, but not necessarily all, of those voxels must be driven by that stimulus. However, poor classification performance of a single voxel is insufficient to disregard that voxel because of potential conditional dependencies; two voxels may be able to better predict the stimulus when taken jointly rather than independently. This is consistent with the concept of population encoding.

Starting with an input set of voxels, we want to isolate the minimum distributed network that best encodes the stimulus without evaluating all permutations. Here, we measure the sensitivity of the classifier outputs to its inputs. In feed-forward neural networks, this can be solved analytically. We train the network, calculate the sensitivity of the network to each voxel, and remove the voxels with low sensitivity. We repeat this process until the network performance begins to degrade.

We tested our approach on a complex visual stimulus. Whole-brain fMRI data was collected while subjects viewed a virtual world that contained a variable number (1—6) of animated characters. We trained a feed-forward neural network to predict the number of characters present in each frame; classification performance was >50%, significantly above chance (~17%). Sensitivity analysis and voxel culling were performed as described above. Sensitivity values were then smoothed and averaged across subjects and projected onto the cortical surface (Figure). The network of classification voxels is almost entirely clustered in visual cortex, indicating that character number is largely encoded by early sensory and association



The sensitivity analysis averaged across all subjects. Brighter regions indicate voxels that were important for successful classification by the feed-forward neural network.

:

**Presentation Preference (Complete):** No Preference

**Linking Group (Complete):** None selected

**Nanosymposium Information (Complete):**

**Theme and Topic (Complete):** G.07.a. Data analysis and statistics

**Keyword (Complete):** FUNCTIONAL MRI ; MACHINE LEARNING ; VISION

**Support (Complete):**

**Support:** Yes

**Grant/Other Support :** ARO W911NF-12-1-0160

**Special Requests (Complete):**

**Would you be interested in being considered for a dynamic poster?:** Yes, I am interested in presenting a Dynamic Poster

**Is the first (presenting) author of this abstract a high school or undergraduate student?:** None

**Religious Conflict?:** No Religious Conflict

**Additional Conflict?:** No

**Status:** Complete

[OASIS Helpdesk](#)

[Leave OASIS Feedback](#)

Powered by [OASIS](#), The Online Abstract Submission and Invitation System <sup>SM</sup>

© 1996 - 2013 [Coe-Truman Technologies, Inc.](#) All rights reserved.