Can neural-based weights improve image classification via support vector machines?

Alex Mulrooney

Introduction

•In the last few years, object classification and identification in images by computers has improved greatly with the use of more advanced machine learning algorithms and convolutional neural networks
•Even with this much improved performance, humans still outperform computers in some tasks
•Recent works by Fong et al. showed promising improvement in a support vector machine (SVM) -based image classifier by introducing a modified hinge-loss function that took into account the level of activation in certain regions of the brains of human subjects that each image elicited to more heavily weight images easily classified by the brain [1].
•In this research, I aim to replicate and extend the work done by Fong on a different dataset, with somewhat different image categories as well as a different mechanism of weighting images

Research Questions

•Does human brain activity recorded while viewing images have the ability to enhance image classification by traditional machine learning methods?
•What effect do different standard image features (HOG, CNN, etc.) have on how much neural features can improve a classifier?

Methods & Analysis

A. Data was gathered from the BOLD5000 dataset, which provides about 5000 unique images from three different image recognition categories as well as fMRI data gathered from 3 different subjects while they viewed the images [2].
B. I used a subset of the BOLD5000 images that included images representing "scenes" that depicted a wider focus and various interacting components for class 1 and images from the ImageNet database, which tend to focus on a single object, animal, or person, as class 2 (see fig. 1 for examples). It is important to note that there is a wide variety of image characteristics in each class and that there may be some overlap between the two classes.

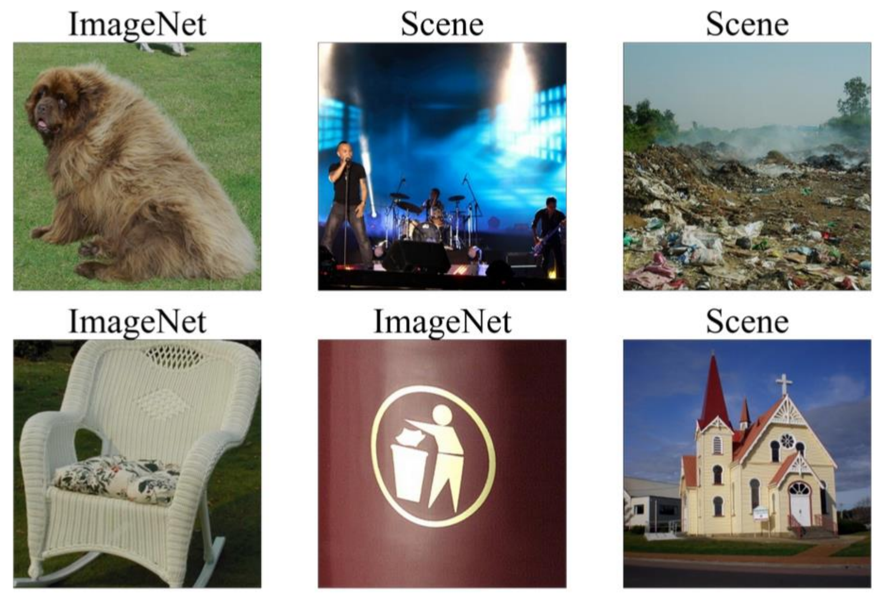

Figure 1: Labeled representative images from scene and ImageNet categories

C. As recommended by the creators of the BOLD5000 dataset an average of the region of interest (ROI) fMRI data from the 4-6s and 6-8s periods was used for each subject. 10 ROIs were available, each hypothesized to engage in some level of human visual processing and/or object/place recognition: Early Visual, LOC, OPA, PPA, and RSC for both left and right hemispheres of the brain.
D. To mimic the ROI classifier built by Fong et al., a support vector machine with a radial-basis function kernel was trained on the ROI data. Although many combinations of the various ROIs were tested for their accuracy, I found the most accurate model included all 10 ROIs. 5-fold cross validation was used to find the best hyperparameters C and γ for the classifier for each subject. See fig. 2 in the results sections for a confusion matrix for each classifier.
E. In order to get inputs to be put into the function that generates the image weights, the trained ROI classifiers were used to predict probabilities of each class for every image using Platt scaling.
F. The distribution of the probabilities for the correct class can be visualized with an empirical CDF (figure 3 in results). To assess if the model was producing overconfident probabilities, the expected calibration error was computed (figure 4).
G. To convert the probabilities to weights that can be used in the image SVM classifier, for each instance (image), a varied mathematical function was applied to the probability that the ROI classifier assigned to the image being in the *correct* class.
H. HOG features of all the images in the dataset were generated to be used as input for the SVM image classifier.
I. The control SVM image classifiers were built for each subject, which did not incorporate the instance weighting for the images. 5-fold cross validation was used to determine the best hyperparameter C for each classifier.
J. Weighted SVMs were built for each subject that incorporated the instance weighting (using multiple different functions to get the weights). 5-fold cross validation was used to find the best C for each classifier.
K. The average accuracy over 5-fold cross validation was compared for the control and weighed classifiers for each subject (fig. 5).

Results

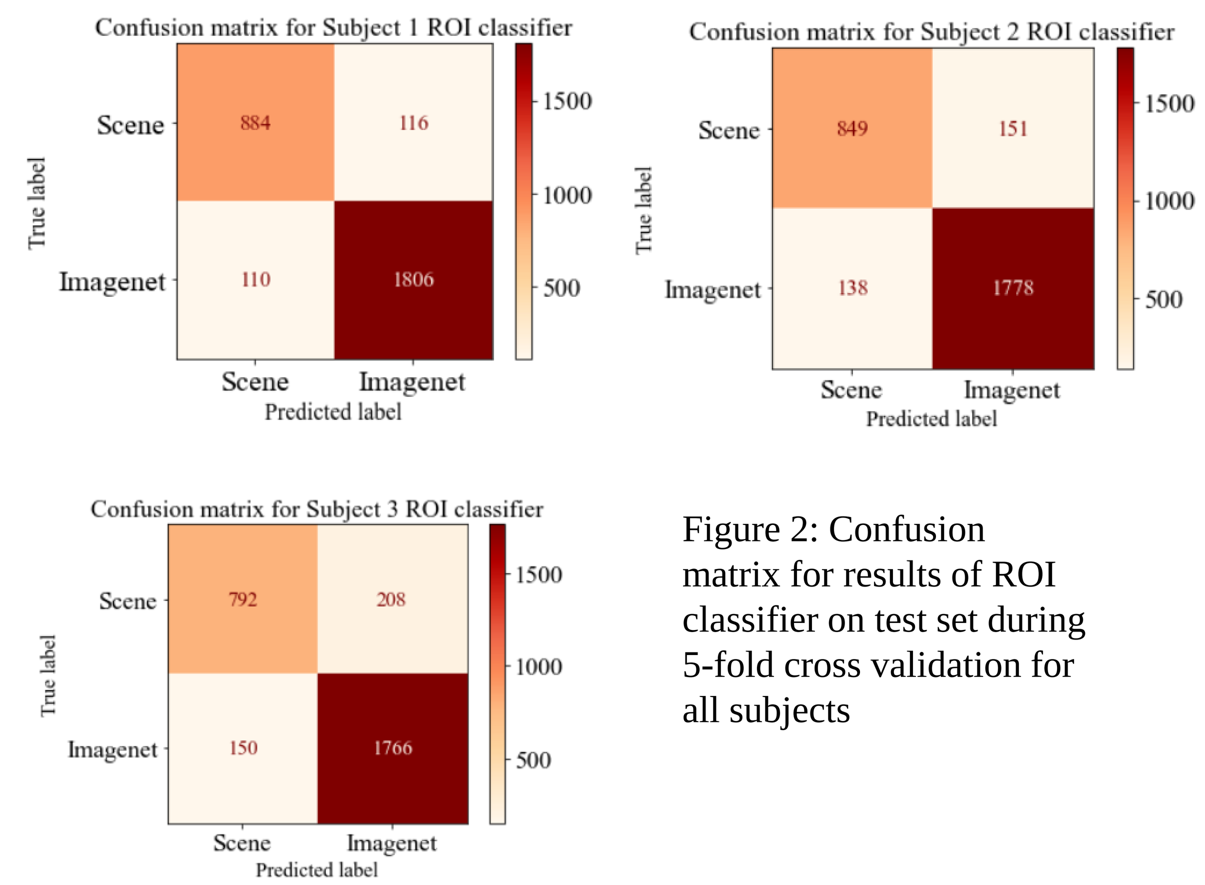

Figure 2: Confusion matrix for results of ROI classifier on test set during 5-fold cross validation for all subjects

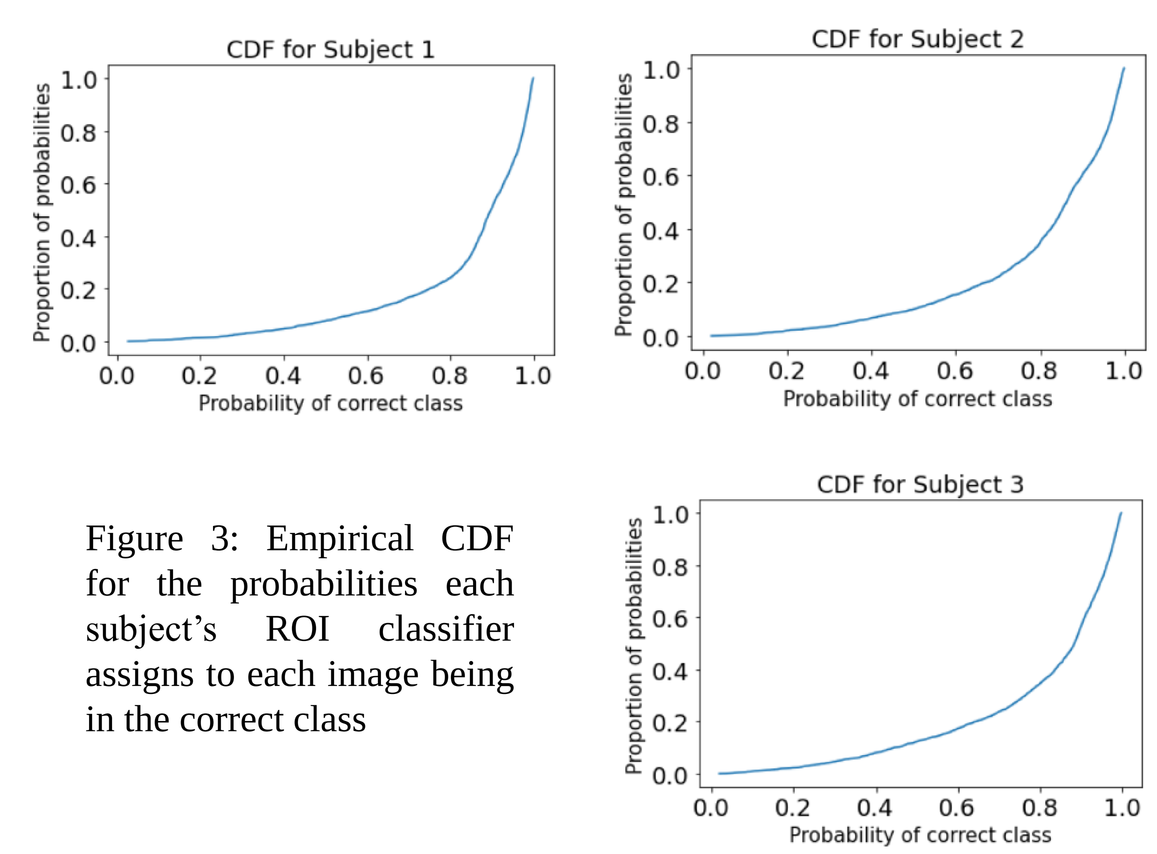

Figure 3: Empirical CDF for the probabilities each subject's ROI classifier assigns to each image being in the correct class

	Subject 1	Subject 2	Subject 3
ECE	5.91%	6.52%	4.13%

Figure 4: Expected calibration error (ECE) for all three subjects. The ECE can be interpreted roughly as the average percentage error in the confidence prediction of a model as compared to the proportion of times that instance is indeed from the predicted class. Since the ECE is fairly low for all three subjects, I considered the confidence predictions to be well-calibrated and took no further action to modify them.

	Control	Linear	Squared	Cubed
Accuracy	81.27%	81.31%	81.48%	81.62%

Figure 5: Average accuracy over 5-fold cross validation for best model using no instance weights (control), linear weights, the square of the weights, and the cube of the weights. The best hyperparameter C for each model was found via 5-fold cross validation.

Conclusions

•Given data from relevant regions of the brain, support-vector machines can, with reasonable accuracy, predict whether the data is from one of two varying image classes
•For this application, probabilities calculated using Platt scaling produced appropriately calibrated confidence estimates for each image
•Minor improvements may be able to be made to classifier accuracy by using instance weighting
•Further research should be done using different methods of weighting, different image categories, etc. to assess the impact of the neural weighting

Directions for Future Research

•Still to be analyzed is the effects of these neural weights on SVM image classifiers using different image features as input, such as features from a CNN, FID scores, or other models
•Other relevant areas of the brain may be used for other image classification task. For example, the Fusiform face area in the brain helps with human face recognition, so it is possible that fMRI data from there could aid computer facial recognition.
•The idea of guiding machine learning with human data can be advanced beyond just using brain data with image recognition. One example is work by Chen et al. that showed that the performance of a SVM used to distinguish different stimulus categories in fMRI data could be improved by using subject behavioral data as a constraint when training the SVM [3].

References

[1] Fong, R.C., Scheirer, W.J. & Cox, D.D. Using human brain activity to guide machine learning. *Sci Rep* **8,** 5397 (2018). https://doi.org/10.1038/s41598-018-23618-6
[2] Chang, N., Pyles, J.A., Marcus, A. *et al.* BOLD5000, a public fMRI dataset while viewing 5000 visual images. *Sci Data* **6,** 49 (2019). https://doi.org/10.1038/s41597-019-0052-3
[3] Chen, Danmei & Li, Sheng & Kourtzi, Zoe & Wu, Si. (2010). Behavior-Constrained Support Vector Machines for fMRI Data Analysis. IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council. 21. 1680-5. 10.1109/TNN.2010.2060353.