

Leveraging the Principal Component Analysis of Functional Magnetic Resonance Imaging Features for Attention Deficit Hyperactivity Disorder Classification

J. Tanner Koza
Mechanical Engineering
Auburn University
Auburn, United States
jtk0018@auburn.edu

Abstract—This paper proposes using principal component analysis (PCA) to transform raw features of functional magnetic resonance imaging (fMRI) data obtained through classical digital signal processing (DSP) techniques into new features for classifying ADHD populations with a support vector machine (SVM). The classification accuracy of three classes (healthy, ADHD-combined, ADHD-inattentive) was 94.25% for the collected set of data.

Index Terms—principal component analysis, support vector machine, ADHD, power spectral density analysis

I. INTRODUCTION

Since the creation of magnetic resonance imaging (MRI) in the 1970s, many disciplines of medicine have greatly reduced the need for invasive diagnostic practices. In addition, the diagnosis of injuries and diseases with internal physical symptoms has become more accessible because of MRI's flexibility in viewing anatomic structures in a variety of imaging planes [1]. However, despite its effectiveness in classifying physical ailments, MRI struggles to diagnose neurodevelopmental disorders as the majority of them do not possess any definitive structural or functional imaging patterns [2]. This fact currently limits the diagnosis of these disorders to behavioral evaluation by a medical professional in the absence of quantitative medical tests.

Among these seemingly imperceptible disorders is attention deficit hyperactivity disorder, otherwise known as ADHD. Given the general prevalence of ADHD among the global population, it is of great interest to improve the methodologies used in its diagnosis. Specifically, it is estimated that 7.2% of children and 3.4% percent of adults are affected by this spectrum disorder globally [3], [4], further motivating the need for accurate diagnosis to allow for effective treatment. Currently, ADHD and its sub-types are diagnosed using guidelines from the American Psychiatric Association's Diagnostic and Statistical Manual, Fifth Edition (DSM-5). The DSM-5 criteria for ADHD focus on identifying inattention and hyperactivity/impulsivity in prospective individuals [5]; however, no medical tests exist to quantitatively do the same. In this paper,

we investigate the use of a support vector machine (SVM) to improve the effectiveness of using functional magnetic resonance imaging (fMRI) data in classifying populations of ADHD subjects. Specifically, the classes include healthy, ADHD-combined, and ADHD-inattentive individuals.

A. Functional Magnetic Resonance Imaging (fMRI)

fMRI is a type of MRI that measures the small changes in blood flow that result from activity in the brain. It is primarily used to attempt to classify physical and mental brain states/diseases [6], [7], but is not necessarily successful in most cases. However, fMRI data has shown recent promise in classifying spectrum disorders as advances in machine and deep learning techniques allow for increasing performance in pattern detection among seemingly uncorrelated data. For example, [8] uses an artificial neural network (ANN) to classify ADHD subjects similarly to the proposed method.

In general, fMRI scans are collected while asking a patient to perform particular tasks. These particular tasks involve using different parts of the brain which are then seen to be active in such scans. An example of an fMRI scan can be seen in Figure 1. This work uses fMRI images that have been converted to time series signals to be processed with classical digital signal processing (DSP) techniques.

B. Support Vector Machine (SVM)

A support vector machine is a type of supervised machine learning model that uses additional learning algorithms to perform classification on sets of data containing multiple classes. They work by maximizing the distance in "class space" between the labeled sets of classes they have been trained on. Once trained, they can classify new members of the training classes based on their prior training knowledge. Practical applications of SVMs include text and hypertext categorization, image segmentation, and, in our case, image classification. An example of using SVMs to classify data from scikit-learn in Python can be seen in Figure 2. This particular example uses four different configurations of an SVM.

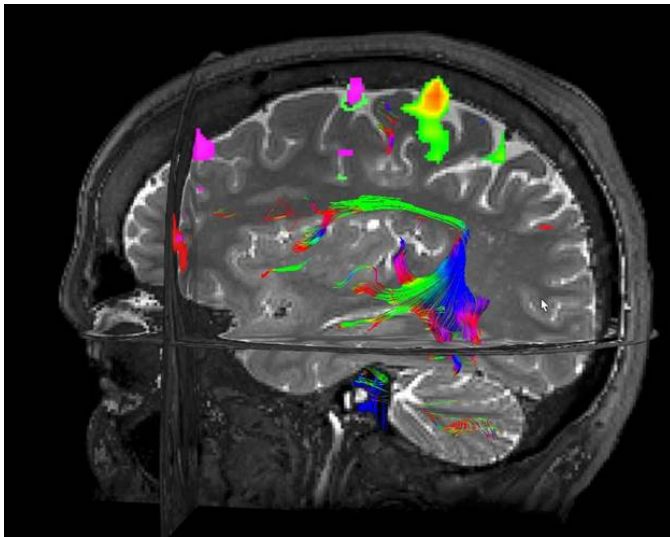


Fig. 1. Functional Magnetic Resonance Imaging (fMRI) [9].

In general, an SVM operates with the following goals [10]:

- **Hyperplane Separation** Determining a hyperplane in higher dimensional space that separates SVM input features (or classes) in a generalized manner. This hyperplane determination is performed using the support vectors of the features provided to the SVM.
- **Margin Maximization** Maximization of the boundaries between the input features is performed using the kernel trick. This effectively maximizes the hyperplane boundaries in the higher dimensional space by being able to account for nonlinear boundaries in the original “class space”. This ensures the robustness of classification. An example of this can be seen in the bottom left subplot of Figure 2.

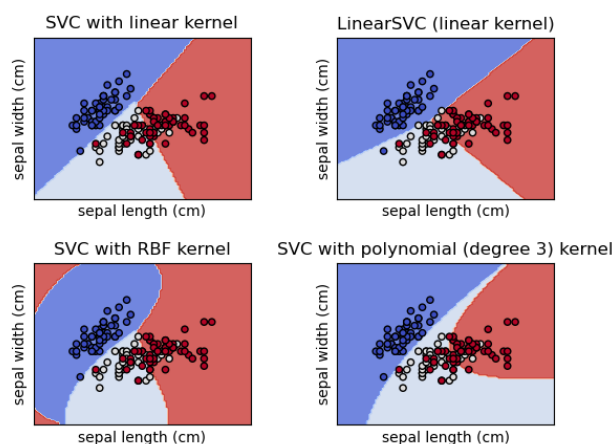


Fig. 2. Support Vector Classification with SVM using scikit-learn in Python [11].

The robustness of the SVM makes it attractive for classi-

fying ADHD populations and this is why it is used in this work.

The rest of this work outlines the prior art associated with this problem and expands upon the innovation of this paper’s approach to ADHD classification. The methodology behind the proposed approach is explained and classification results are presented. Finally, the results of the proposed methodology are discussed and conclusions are highlighted.

II. PRIOR ART & INNOVATION

The prior art associated with this problem is fairly extensive. Many have attempted to apply machine and deep learning methods to the classification of ADHD, but this work will only highlight a couple. The first is [8] from Deshpande et al., where a fully connected cascade ANN is used to classify the same ADHD sub-types presented in this work. The features used for classification use both directional and non-directional brain connectivity methods such as Correlation-Purged Granger Causality and the correlation between probabilities of recurrences. The methods used in this paper produced a classification accuracy of 90% and 95% for healthy and ADHD sub-types, respectively. It should be noted that this work also takes advantage of the principal component analysis of these features which has heavily inspired this work.

Another notable work is [12] from Eloyan et al., which achieves an overall classification accuracy of 61%. This work takes advantage of resting-state functional connectivity features and other biomarkers for feature extraction. The methodologies used to obtain these include but are not limited to CUR decomposition, gradient boosting, and voxel-based morphometry. Also, the singular value decomposition of the features was used in conjunction with an SVM.

The innovation of this work lies in the use of fairly simple features in conjunction with principal component analysis to achieve similar and better results than the aforementioned works. The prior works use a multitude of causality and connectivity features that are avoided in this work for simplicity. Despite using less complex features, the performance of the proposed method is still respectable. The methodology behind our approach is outlined next.

III. METHODOLOGY

Before describing the methodology, information about the data and SVM in question should be provided. We were supplied with three classes of ADHD populations with varying sizes. The three classes were healthy, ADHD-combined, and ADHD-inattentive. The aforementioned classes contained 748, 260, and 175 individuals, respectively. For each individual, 190 brain regions were scanned and converted to time series data of varying lengths and varying, unknown sampling rates. We were also provided with some MATLAB code that employs an SVM to classify these datasets based on features we generate.

The methodology for the proposed classification of the data involves a two-step feature extraction approach performed using open-source libraries in Python. Initially, the time series fMRI data was filtered before feature extraction. A low-pass

Butterworth filter (5^{th} order, $\omega_c = 0.2\pi$ rad/sample) was applied but provided no additional benefit as the data was already filtered. Classification results using this filtered data are not included as the additional filtering significantly reduced the classification performance. Additionally, no other filtering or processing was performed before feature extraction once the Butterworth filter was removed.

A. Raw Feature Extraction

The first step in feature extraction involves employing classical DSP techniques on the raw data to extract what are called *raw features*. Two methods were used to transform the fMRI time series samples to samples where features could be extracted. Specifically, the normalized Power Spectral Density (PSD) of each brain region and the autocorrelation of each individual were computed. An example of the PSD from a single region can be seen in Figure 3. The normalized PSD was computed because the sampling rate of the data was unknown and supposedly variable for each subject. It should be noted that the aforementioned Butterworth filter was designed based on this PSD. From here, four raw features were extracted

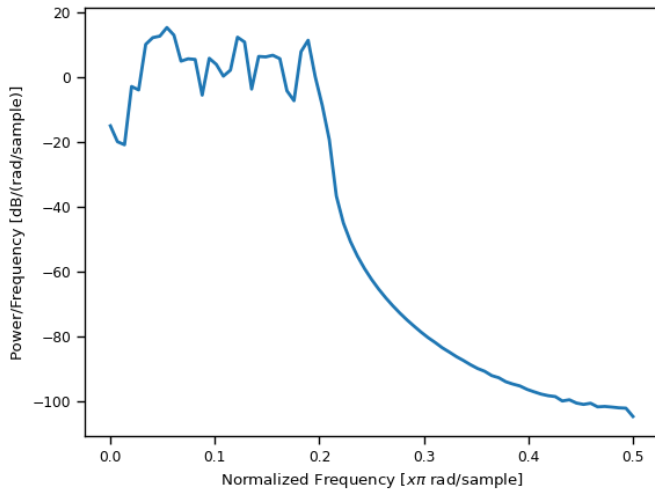


Fig. 3. Normalized Power Spectral Density (Healthy Population, Subject 1).

from the normalized PSD of each brain region and one was extracted from the autocorrelation of each individual. They are as follows:

- 1) **Total PSD Energy** The total PSD “energy” is formulated as the sum of the normalized power across each region’s PSD. This feature was chosen as it is seemingly distinct for each region, but similar for each individual in their respective classes. In total, an “energy” was produced for each region resulting in 190 raw features per individual.
- 2) **Maximum and Minimum Normalized Power** Although technically two features, the maximum and minimum normalized power of each region’s PSD were included together as they serve a similar purpose. They were chosen to provide the effective range in power of

each region’s PSD as this is seemingly different for each. An example of these two features for a single region can be seen in Figure 4. In total, the maximum and minimum power features resulted in 380 additional raw features per individual.

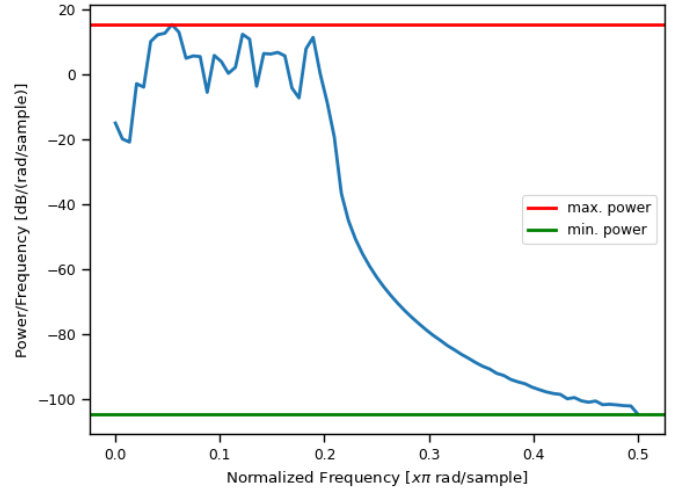


Fig. 4. Maximum and Minimum Power/Frequency Values (Healthy Population, Subject 1).

- 3) **Frequency of Maximum Normalized Power** The frequency of the maximum normalized power serves to highlight where the majority of power lies on a region-to-region basis. An example of this feature can be seen in Figure 5. This choice simply provides another rudimentary feature to use in the second feature extraction step. In total, an additional 190 raw features are added per individual when including the frequency of the maximum normalized power.

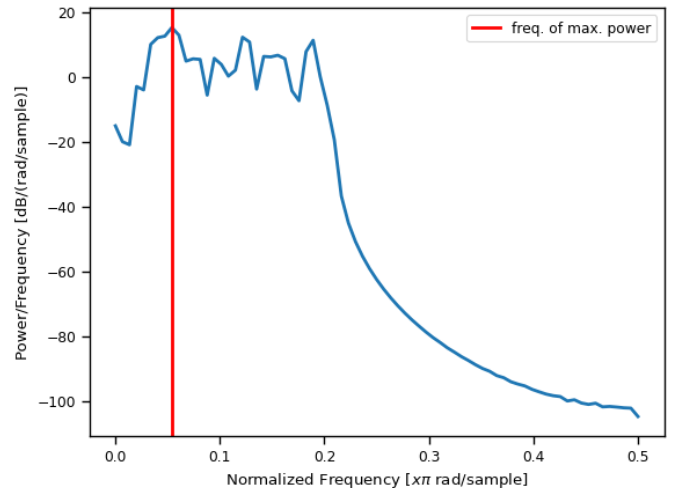


Fig. 5. Frequency of Maximum Power/Frequency Values (Healthy Population, Subject 1).

- 4) **Summed Power of PSD Peaks** The summed power of the PSD peaks was chosen as a raw feature because it is

heavily correlated with each region across individuals. Each region across each individual has a similar frequency response and, subsequently, a similar number of peaks. Therefore, the sum of these powers is incredibly indicative of which region is which in each population. An example of the peaks for a single region can be seen in Figure 6. The addition of this feature adds another 190 features per individual.

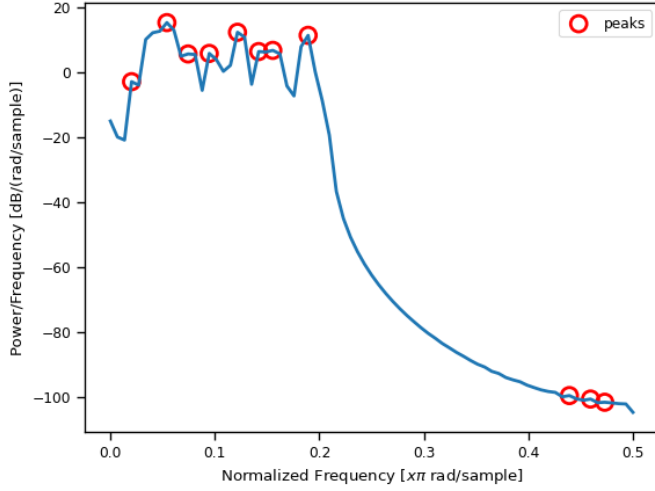


Fig. 6. Power Peaks of Power Spectral Density (Healthy Population, Subject 1).

- 5) **Sum of Autocorrelation Function Across Regions** The autocorrelation of each individual's data was computed and summed across regions to provide another raw feature. This feature is similar to the total PSD energy in that it distinguishes between each region. However, the scale of the resulting features is much different. The addition of this feature adds another 190 raw features per individual.

Using these five raw features results in a total 1140 raw features per individual, or 1,348,640 features among the whole dataset. This large number of features provides a lot of data to the SVM. However, principal component analysis can be performed on these features to better analyze their variation and reduce the total number of features needed for classification.

B. Principal Component Analysis (PCA) Feature Extraction

Inspired by [8], the principal component analysis of the raw features was taken. PCA works by computing the eigenvectors of the covariance matrix for a group of data. These components are ranked (using their eigenvalues) by how much they maximize the variance of the data when it is projected onto each component. Specifically, these components are calculated by performing singular value decomposition (SVD) on the data matrix.

It should be noted that the scikit-learn library in Python was used to perform the PCA of the raw features. The resulting features used in the SVM were data transformed by the 10

most effective principal components. This was empirically determined to produce the best results in classification through hand-tuning. This reduction of the raw features using PCA resulted in a total of 11,830 features among the whole dataset.

IV. RESULTS

The results for the ADHD classification using the methodology introduced in this paper are fairly straightforward. Across the entire dataset, the resulting classification accuracy was 94.25% for all 1183 individuals. The breakdown of these classifications can be seen in Table I.

TABLE I
TOTAL DATASET CLASSIFICATION ACCURACY

ADHD Type	Classification Accuracy (%)
Healthy	98.93
ADHD-combined	83.46
ADHD-inattentive	90.29
Total	94.25

To further validate the performance of the proposed methodology, 25 random subsets of balanced features (175 individuals from each population) were passed to the SVM and the average classification accuracy was calculated. These averages can be seen in Table II.

TABLE II
AVERAGE CLASSIFICATION ACCURACY FOR 25 SUBSETS OF 175 INDIVIDUALS PER CLASS

ADHD Type	Average Classification Accuracy (%)
Healthy	95.19
ADHD-combined	91.84
ADHD-inattentive	92.32
Total	93.19

These results indicate the proposed methodology is fairly successful in classifying ADHD. However, there is room for improvement.

V. DISCUSSION

The results of the proposed classification methodology indicate that ADHD can be accurately classified. However, more work needs to be done to ensure this methodology is robust in doing so. Currently, this methodology has only been tested on 1183 individuals so more testing is warranted. The use of PCA to generate features from raw features might be heavily dependent on the current relationships between the raw data of each population that may not persist in different datasets. Unfortunately, there was no access to additional data to validate this assumption. Although there is no other validation data, it is assumed that performance will degrade when classifying a different dataset. One way to combat this might involve calculating additional raw features to further generalize the PCA of those features. Also, additional causality criteria can be used in conjunction with PCA to further reduce the dependence this method has on its training dataset. Overall, this methodology is a decent foray into the world of classification that can be further refined.

VI. CONCLUSION

In this paper, we have demonstrated the ability to accurately classify ADHD sub-types based on the PCA of raw features extracted from fMRI data. Our primary contribution is the use of PCA on raw features we've extracted from the data using classical DSP techniques. This provides an easy entry into accurate ADHD classification that can be further built upon.

REFERENCES

- [1] Y. C. Chang, K. M. Huang, J. H. Chen, and C. T. Su, "Impact of magnetic resonance imaging on the advancement of medicine," *Journal of the Formosan Medical Association = Taiwan Yi Zhi*, vol. 98, no. 11, pp. 740–748, Nov. 1999, Available: <https://pubmed.ncbi.nlm.nih.gov/10705690/>
- [2] F. D. Firouzabadi, S. Ramezanpour, M. D. Firouzabadi, I. J. Yousem, N. A. J. Puts, and D. M. Yousem, "Neuroimaging in Attention-Deficit/Hyperactivity Disorder: Recent Advances," *American Journal of Roentgenology*, pp. 1–12, Dec. 2021, doi: <https://doi.org/10.2214/ajr.21.26316>.
- [3] R. Thomas, S. Sanders, J. Doust, E. Beller, and P. Glasziou, "Prevalence of attention-deficit/hyperactivity disorder: A systematic review and meta-analysis," *Pediatrics*, vol. 135, no. 4, pp. 994–1001, Mar. 2015, doi: <https://doi.org/10.1542/peds.2014-3482>.
- [4] J. Fayyad et al., "Cross-national prevalence and correlates of adult attention-deficit hyperactivity disorder," *British Journal of Psychiatry*, vol. 190, no. 05, pp. 402–409, May 2007, doi: <https://doi.org/10.1192/bjp.bp.106.034389>.
- [5] CDC, "Symptoms and diagnosis of ADHD," Centers for Disease Control and Prevention, Aug. 09, 2022. <https://www.cdc.gov/ncbddd/adhd/diagnosis.html>
- [6] J. Han, C. Chen, L. Shao, X. Hu, J. Han, and T. Liu, "Learning Computational Models of Video Memorability from fMRI Brain Imaging," *IEEE Transactions on Cybernetics*, vol. 45, no. 8, pp. 1692–1703, Aug. 2015, doi: <https://doi.org/10.1109/TCYB.2014.2358647>.
- [7] G. Deshpande et al., "Recursive Cluster Elimination Based Support Vector Machine for Disease State Prediction Using Resting State Functional and Effective Brain Connectivity," *PLOS ONE*, vol. 5, no. 12, pp. e14277–e14277, Dec. 2010, doi: <https://doi.org/10.1371/journal.pone.0014277>.
- [8] G. Deshpande, P. Wang, D. Rangaprakash, and B. M. Wilamowski, "Fully Connected Cascade Artificial Neural Network Architecture for Attention Deficit Hyperactivity Disorder Classification From Functional Magnetic Resonance Imaging Data," *IEEE transactions on cybernetics*, vol. 45, no. 12, pp. 2668–2679, Jan. 2015, doi: <https://doi.org/10.1109/tcyb.2014.2379621>.
- [9] "Functional Magnetic Resonance Imaging (fMRI) Department of Neurology - College of Medicine - University of Florida." <https://neurology.ufl.edu/divisions/epilepsy/epilepsy-surgery-program/the-epilepsy-management-conference-emc/functional-magnetic-resonance-imaging-fmri/>
- [10] W. S. Noble, "What is a support vector machine?," *Nature Biotechnology*, vol. 24, no. 12, pp. 1565–1567, Dec. 2006, doi: <https://doi.org/10.1038/nbt1206-1565>.
- [11] scikit learn, "1.4. Support Vector Machines — scikit-learn 0.20.3 documentation," Scikit-learn.org, 2018. <https://scikit-learn.org/stable/modules/svm.html>
- [12] A. Eloyan et al., "Automated diagnoses of attention deficit hyperactive disorder using magnetic resonance imaging," *Frontiers in Systems Neuroscience*, vol. 6, 2012, doi: <https://doi.org/10.3389/fnsys.2012.00061>.