Using Machine Learning Classification Approaches for Prediction of Obstructive Sleep Apnea from Mean Diffusivity MRI Images

Sarthak Tiwari[1], Bhaswati Roy[2], Rajesh Kumar[2], Daniel Tward[2].
[1] University of Utah. [2] UCLA.

Introduction

- Obstructive sleep apnea (OSA) affects 10% of adults [1], and can increase risk of disease [2]
- Early diagnosis and treatment can lower the risk of diseases and brain damage [3]
- Current diagnosis methods are long and costly [4]
- Mean diffusivity (MD) is a promising alternative [5]

Methods

- MD images from 96 control, and 59 OSA patients
- 20-80% Test train split, 4-fold cross validation (val)
- Augmentations (deformation, noise, and affine)

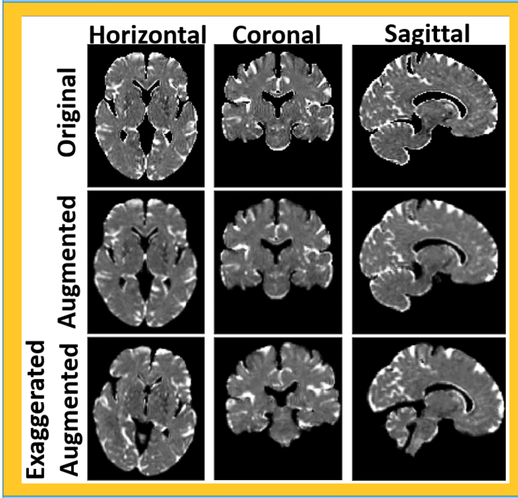

Figure 1: MD images of one patient. Views of the center of the 3 different planes are shown as the columns. The augmentations applied are also shown in the rows, both the actual augmentations, as well as an exaggerated version to better see what was applied

Models used:

1. Linear discriminant analysis (LDA)
2. Logistic regression network
3. Convolutional Neural network (CNN) based on Alzheimer's detection [5]

- Trained on combinations of weight decay rates and augmentations
- Receiver Operating Characteristic (ROC) Area Under the Curve (AUC) calculated. Model with highest AUC was applied to the test set
- Also compared val to test

Results

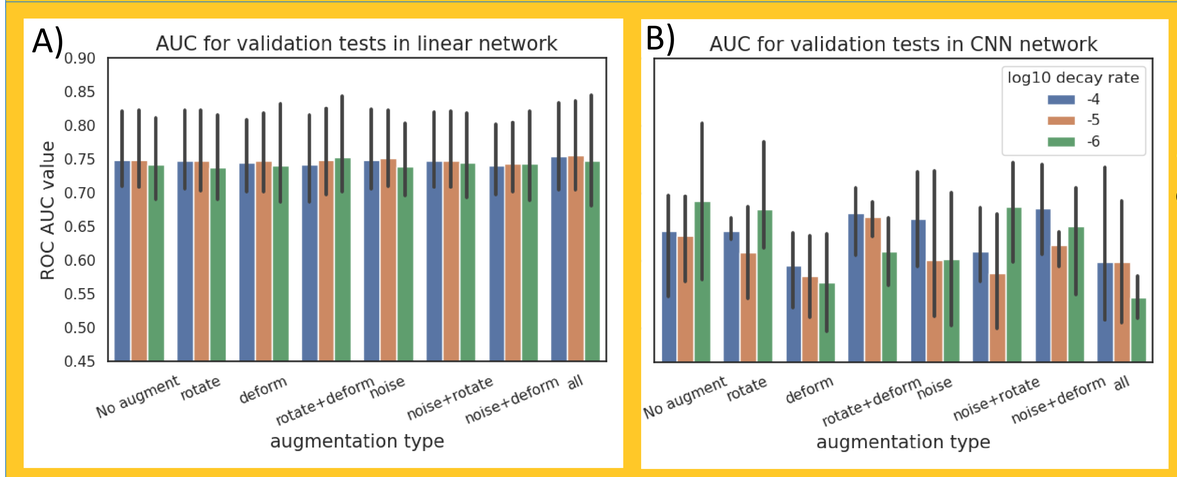

Figure 3A: the results of the linear classifier on the validation data. The mean validation AUC across the 4 folds for the different training conditions. The error bars are for the different folds. The linear network was quite consistent across augmentations and folds

Figure 3B: The same as figure 3A but for the CNN. The CNN varied a lot depending on the training conditions

The network with the best AUC was applied to the test data

- LDA had .64 AUC, and 63% accuracy, only predicted controls
- Logistic had .89 AUC and 75% accuracy on validation, .65 AUC and 63% accuracy on test, only predicted controls
- CNN had a best AUC of .84 with 83% accuracy on the val, AUC of .56 and accuracy of 63% on test, and predicted OSA and control

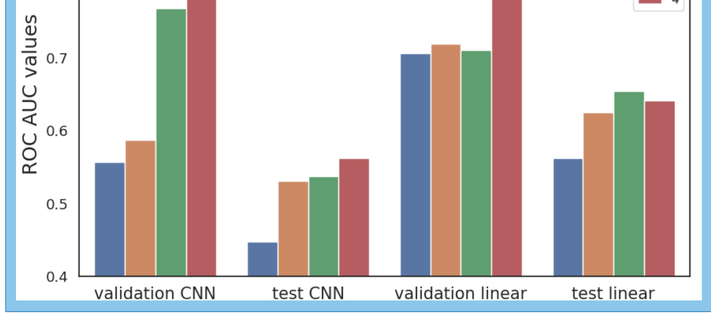

Figure 4: Comparison of validation with test data. All test sets performed quite a bit worse than the best validation sets. Shows that there are still issues with generalizing with a small dataset. But AUC is a better metric than accuracy for validation of our model.

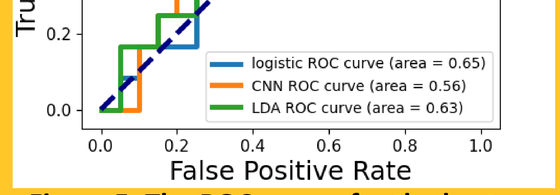

Figure 5: The ROC curves for the best classifier of each approach. CNN generalized the best since it didn't only predict controls

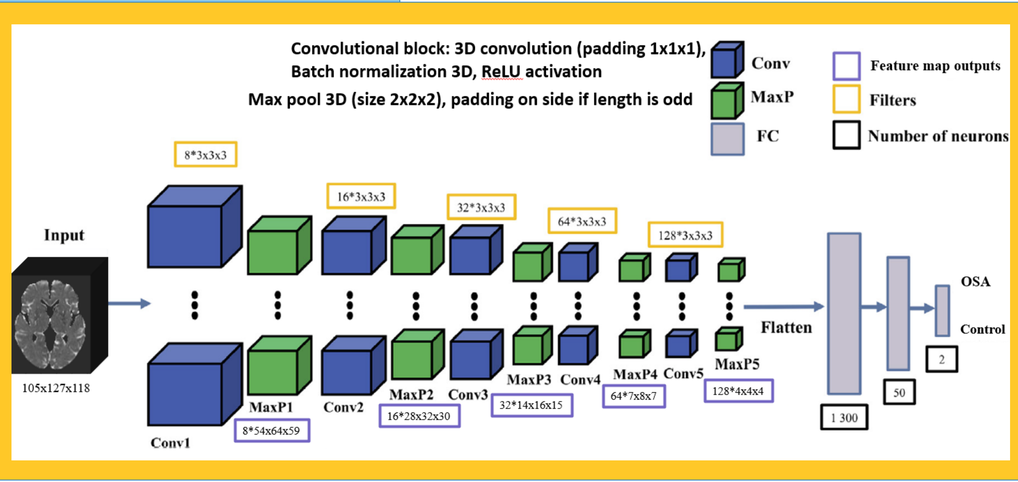

Figure 2: The architecture of our CNN based on [5]. Uses 5 convolutional blocks followed by max pooling before using linear layers to classify the data. Figure was inspired by review paper [6].

Conclusion

- We show the ability for different classification approaches to be used in OSA screening
- Generalization remains a big problem with small datasets, but only CNNs didn't predict only control
- Transfer learning approaches and increasing samples will improve results
- The need for cheaper and faster diagnosis of OSA remains, and our study shows that CNN models have promise to screen OSA from MRI

Acknowledgements: This work was conducted through the BIG summer program and funded by the Karen Toffler Charitable Trust
References: [1] Young, PT. Am. J. Respir. Crit. Care Med. 2002. 165(9) 1217-1239. [2] Lavie, P. Curr. Pharm. Des. 2008. 14(32), 3466–3473. [3] Kim, H. Hum Brain Mapp. 2015. 37(1), 395-409. [4] Kim, RD. Sleep. 2015. 38, 1027-1037. [5] Kumar, R. J. Neurosci. Res. 2012. 90(10), 2043-2052. [6] Wen, J. Med Image Anal. 2020. 63, 1005.

Sarthak Tiwari
University of Utah
Biomedical engineering
u1254813@utah.edu