Using Brain MRI Images to predict BMI, Memory and Age

Chhavi Yadav [1]

Abstract

In recent years, a lot of research has been done on applying deep learning to medical problems. A significant chuck of this work focuses on using computer vision for medical imaging tasks like classification and segmentation. In this paper, we use simple 3D neural network architectures to predict parameters like memory scores, body mass index (BMI) and age of a person. We model the problem as both single and multitask predictions. We compare the performance of the two approaches. We also use visualizations to find biomarkers for these parameters.

1. Introduction

It is a matter of fact that using brain images to predict any parameter is an extremely useful task, even if it can only supplement the doctor in making a decision. Machine Learning has been since decades used as a tool by scientists worldwide to do so. Previously traditional methods like Support Vector Machines and Regression were being used. Recently, a lot of advances have been made in computer vision that can be adapted to the prediction and classification tasks in neuroimaging as well. Datasets like ADNI(Jack Jr et al., 2008), HCP(Van Essen et al., 2013) and OASIS(Marcus et al., 2007) are publicly available to aid such research. In our study, we use the HCP dataset and a 3D CNN architecture to predict memory, age and BMI.

A recent paper by (LePort et al., 2016) suggests structural differences in the brains of people with Highly Superior Autobiographical Memory as compared to normal people. Others like (Wingfield et al., 1988) have discussed the deterioration of memory performance with increasing age. (Bobinski et al., 1999), (Vereecken et al., 1994) and (Ball et al., 1985) give a detailed analysis of the shrinkage of the Hippocampus, a region in the brain responsible for memory, in people suffering from Alzheimer's disease. Some proof has also been found by (Raji et al., 2010) and (Jagust et al., 2005) for similar changes caused by obesity. All this work makes for a compelling case for a relationship between brain structure and Age, Memory, BMI (which is one of the measures of obesity). We wanted to investigate if the presence of such relationships can help our CNN model make better predictions. To do this, we modeled our prediction problem as a multitask problem. We compared the performance of a single task model with a multitask model.

Structural brain alterations are important biomarkers. They can help predict or confirm the presence of a particular disease like in the previously mentioned case of Alzheimer's. They can also tell us more about the inherent characteristics of humans as a species. For example (Srinivasan et al., 2015) reason that primates have better visual acuity than non-primates due to larger neuronal densities in their primary visual cortices. Another interesting study by (Erickson et al., 2011) and (Thomas et al., 2012) discovered a relationship between aerobic exercise and an increase in size of the Hippocampus. However, there are many relationships that are speculative. For example, it has been argued that a straight forward relationship between Extraversion, Neuroticism, Agreeableness, and Conscientiousness and some brain regions like lateral prefrontal cortex, medial orbitofrontal cortex and others does not exist and more research is needed to verify or deny such claims[2]. It will be interesting to see if our models can at all support studies of this nature and which side of the argument they take. We generate heatmaps for our model using occlusion (Zeiler & Fergus, 2014) and try to find regions of interest focused on by the model.

To summarize, we

- model Memory, BMI and Age predictions as separate tasks and predict using separate 3D CNN models,
- model Memory, BMI and Age predictions as a single multitask prediction problem and predict using one 3D CNN model,
- compare the performance of the above approaches,
- visualize the regions of interest for the models using occlusion in the hope of finding biomarkers.

[1]Department of Computer Science, New York University. Correspondence to: Chhavi Yadav <chhavi@nyu.edu>.

[2]See article on Psychology Today here.

Table 1. Input Data Statistics

PARAMETER	MIN	MAX	MEAN	SD
MEMORY	81	145	110.9	11.3
BMI	16	56	27.1	5.9
AGE	22	37	28.8	3.7

2. Data

The dataset used in our study was obtained from the Human Connectome Project (Van Essen et al., 2013) database. It consists of 1113 healthy young individuals in the age group of 20-40 years. The MRIs contained include diffusion imaging (dMRI), resting-state fMRI (R-fMRI), task-evoked fMRI (T-fMRI), T1- and T2-weighted MRI for structural and myelin mapping, and combined magnetoencephalography and electroencephalography (MEG/EEG) for all subjects. The dataset has demographics information like age, race, gender, education, income and so on. There is also data regarding substance usage, physical health (BMI, height, weight, menstrual history, thyroid, blood pressure, family history of psychiatric and neurological disorders), sleep patterns, alertness, cognition, emotional health and sensory information.

2.1. Data Pre-Processing

We use the T1-weighted MRIs, memory, age and BMI scores. The working memory score in the dataset was calculated using a list sorting task: size order different visually- and orally-presented stimuli (foods and animals). This measures both information processing and storage. High scores on each of these indicate high levels of working memory.

Each MRI volume is axially cut into 256 slices of size 256 * 256. We do a 60-20-20 train, validate, test split on the dataset. The training set has 668 subjects, the validation set has 223 subjects and the testing set has 222 subjects. The slices which are completely blank or have very little grey area are removed, leaving us with 110 * 256 * 256 images for each subject. These are then min-max normalized. Some of the randomly chosen training images are rotated and translated slightly to improve generalization. Memory scores ranged from 81-145, BMI from 16-56 and Age from 22-37 years. The respective histograms are as shown in figures 1, 2 and 3. Other statistics can be found in table 1. SD there stands for standard deviation.

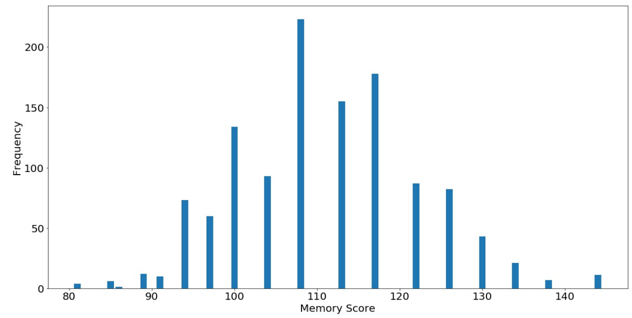

Figure 1. Memory Histogram

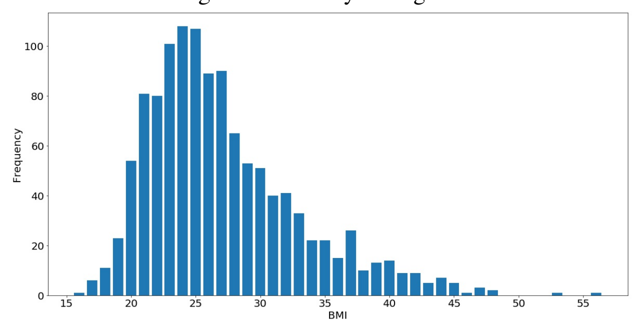

Figure 2. BMI Histogram

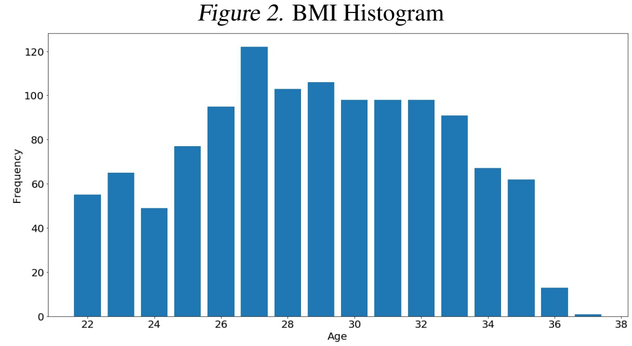

Figure 3. Age Histogram

3. Methods

We started by building a 2D CNN architecture. Each of the 256 images for a subject was considered a separate data sample. This model did not learn well and overfit a lot. Next we tried the 3D approach. Here one subject with all its 256 * 256 * 256 images is considered a single data sample. We also tested 3D Resnet(He et al., 2016) architectures vs plain 3D architectures. The Resnet ones were as good as the plain ones. This was reiterated by the age prediction study done by (Korolev et al., 2017), where the plain CNN and Resnet architectures perform almost at par. Hence, going by Occam's Razor, we decided to stick to the simpler plain 3D CNN architecture for further analysis.

Architectures with varying number of layers and filters were validated upon. We finally chose a 5 layer 3D CNN which was the best performing one. The architecture is as shown in figure 4. FC stands for fully connected layer.

We use the Adam optimizer with a learning rate of 0.0001, a slight weight decay of 3e-6 and the Mean Squared Error loss function. The batchsize is set to 1 due to limited computing resources. In the multitask approach the losses in the forward pass for all three parameters are added together. The losses are then jointly optimized.

After the training, we generate heatmaps for each of the models. We do this using the occlusion(Zeiler & Fergus, 2014) technique because it is pretty stable as compared to other gradient based techniques. Heatmaps are generated for each of the 110 images and then stacked on top of each other to give the final heatmap. All the 2D image slices are also stacked on top of each other for displaying.

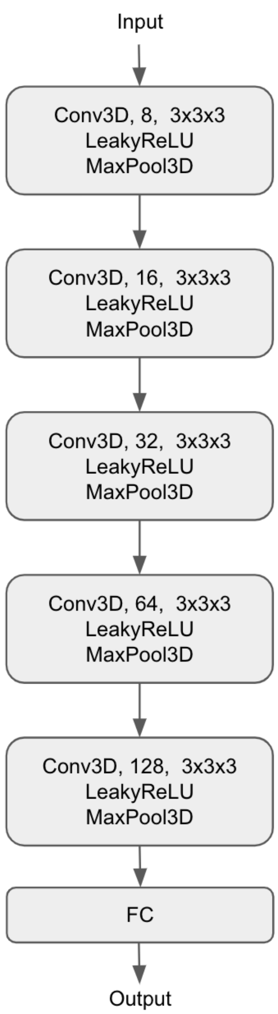

Figure 4. Model Architecture

4. Results

In this section we will summarize the results of our experiments. We do this in 2 parts. In the first part, we summarize the numbers while in the second part, we discuss about the heatmaps.

Table 2. RMSE Results

METHOD	MEMORY	BMI	AGE
SINGLE TASK	11.30	**4.47**	3.87
MULTITASK	**11.00**	4.79	**3.68**
TRADITIONAL	–	–	15.10
DUMMY	11.33	5.28	3.695

4.1. Summarizing the numbers

The results of the experiments using the 5 layer 3D plain CNN architecture have been summarized in Table 2. These are the root mean square error (RMSE) values on the test set.

The "TRADITIONAL" row in the table refers to predicting parameters using traditional machine learning methods like Regression and Support Vector Machines. The value for traditional age was obtained from the paper (Cole et al., 2017). Here the authors predicted age using Gaussian Processes Regression (GPR). Their best results have been mentioned in the table. Such results could not be found in the literature in a comparable form for Memory and BMI, explaining the absence from the table. The "DUMMY" row in the table is obtained by predicting the mean value of the parameter for every subject and then calculating the RMSE.

As can be seen from the table, there is not a significant difference between the RMSE values of the Single task and Multitask approaches. Also neither of the models performs significantly better than predicting the just the mean values. However, both the approaches for Age perform significantly better than the traditional approach. It is also worth mentioning that (Cole et al., 2017) get an RMSE of 15.10 when they use raw images as input which is what we use as well. But they get an RMSE of 5.43 when they feed the grey and white matter volumteric maps of the MRIs together as the two inputs of the models instead of just one input. For more details please refer to the paper. In either case our models perform better. e

4.2. Heatmaps

The heatmaps for the 4 models (predicting Age, BMI, Memory separately and then together) have been shown in figure 5.

Pleasantly surprisingly the heatmap for Memory does point towards the region where the Hippocampus and Amygdala reside as shown in figure 5b . Both of these regions correspond to memory. We cannot say much about the regions highlighted in the other heatmaps. A thorough analysis for

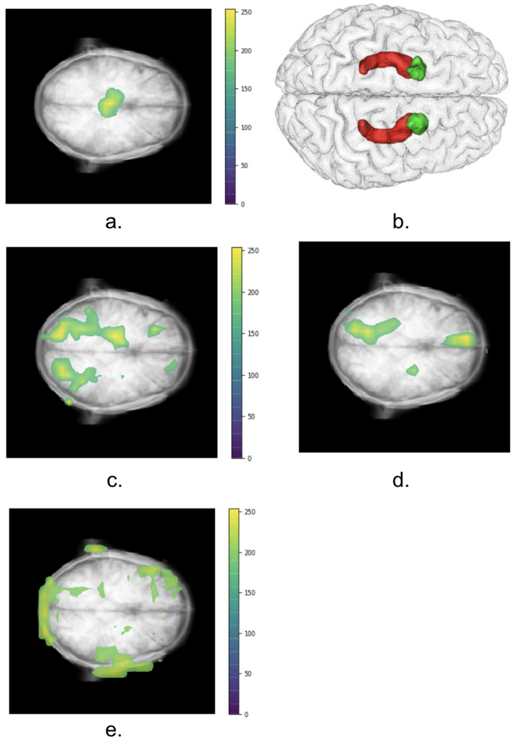

Figure 5. Heatmaps a.Memory Heatmap b.Brain showing Hippocampus (Red) and Amyglada (Green) c.Age Heatmap d.BMI Heatmap e.All Heatmap

all the heatmaps needs to be done with a neuroscientist. Nothing conclusive can be said before that.

5. Conclusion

We proposed a 3D CNN architecture for predicting Memory, Age and BMI from Brain MRI images. We modeled this problem as both single and multitask. We did not find any significant difference between the performance of the 2 approaches. We also found that our model did not perform better than predicting the mean values for all parameters. However, our model did quite well as compared to the traditional methods in predicting age. We also visualized the heatmaps for all models. We did find some very clear regions of interest that the models focused on. However, more investigation needs to be done to see if the heatmaps are medically sound.

In our future work, we would like to investigate if using representations of the brain can be of any help in making better predictions. We would also like to send in segmented brain regions to the model. We suspect if better models for example a modified version of U-Net (Ronneberger et al., 2015) can improve predictions. A distinct approach to this problem can be to turn it into a classification problem by binning the output labels. We would also like to explore this way.

References

Ball, M., Hachinski, V., Fox, A., Kirshen, A., Fisman, M., Blume, W., Kral, V., Fox, H., and Merskey, H. A new definition of alzheimer's disease: a hippocampal dementia. *The Lancet*, 325(8419):14–16, 1985.

Bobinski, M., De Leon, M., Wegiel, J., Desanti, S., Convit, A., Saint Louis, L., Rusinek, H., and Wisniewski, H. The histological validation of post mortem magnetic resonance imaging-determined hippocampal volume in alzheimer's disease. *Neuroscience*, 95(3):721–725, 1999.

Cole, J. H., Poudel, R. P., Tsagkrasoulis, D., Caan, M. W., Steves, C., Spector, T. D., and Montana, G. Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage*, 163:115–124, 2017.

Erickson, K. I., Voss, M. W., Prakash, R. S., Basak, C., Szabo, A., Chaddock, L., Kim, J. S., Heo, S., Alves, H., White, S. M., et al. Exercise training increases size of hippocampus and improves memory. *Proceedings of the National Academy of Sciences*, 108(7):3017–3022, 2011.

He, K., Zhang, X., Ren, S., and Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

Jack Jr, C. R., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P. J., L. Whitwell, J., Ward, C., et al. The alzheimer's disease neuroimaging initiative (adni): Mri methods. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 27(4):685–691, 2008.

Jagust, W., Harvey, D., Mungas, D., and Haan, M. Central obesity and the aging brain. *Archives of neurology*, 62 (10):1545–1548, 2005.

Korolev, S., Safiullin, A., Belyaev, M., and Dodonova, Y. Residual and plain convolutional neural networks for 3d brain mri classification. In *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on*, pp. 835–838. IEEE, 2017.

LePort, A. K., Stark, S. M., McGaugh, J. L., and Stark, C. E. Highly superior autobiographical memory: Quality and quantity of retention over time. *Frontiers in psychology*, 6:2017, 2016.

Marcus, D. S., Wang, T. H., Parker, J., Csernansky, J. G., Morris, J. C., and Buckner, R. L. Open access series of imaging studies (oasis): cross-sectional mri data in young, middle aged, nondemented, and demented older adults. *Journal of cognitive neuroscience*, 19(9):1498–1507, 2007.

Raji, C. A., Ho, A. J., Parikshak, N. N., Becker, J. T., Lopez, O. L., Kuller, L. H., Hua, X., Leow, A. D., Toga, A. W., and Thompson, P. M. Brain structure and obesity. *Human brain mapping*, 31(3):353–364, 2010.

Ronneberger, O., Fischer, P., and Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241. Springer, 2015.

Srinivasan, S., Carlo, C. N., and Stevens, C. F. Predicting visual acuity from the structure of visual cortex. *Proceedings of the National Academy of Sciences*, 112(25): 7815–7820, 2015.

Thomas, A., Dennis, A., Bandettini, P. A., and Johansen-Berg, H. The effects of aerobic activity on brain structure. *Frontiers in psychology*, 3:86, 2012.

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., Consortium, W.-M. H., et al. The wu-minn human connectome project: an overview. *Neuroimage*, 80:62–79, 2013.

Vereecken, T. H., Vogels, O., and Nieuwenhuys, R. Neuron loss and shrinkage in the amygdala in alzheimer's disease. *Neurobiology of aging*, 15(1):45–54, 1994.

Wingfield, A., Stine, E. A., Lahar, C. J., and Aberdeen, J. S. Does the capacity of working memory change with age? *Experimental aging research*, 14(2):103–107, 1988.

Zeiler, M. D. and Fergus, R. Visualizing and understanding convolutional networks. In *European conference on computer vision*, pp. 818–833. Springer, 2014.