# AN ANALYSIS OF PREDICTIVE IDENTIFICATION FOR MAJOR DEPRESSIVE DISORDER USING TRANSFER LEARNING

*Rohith Ravindranath*

Columbia University
Department of Computer Science

*Sharmi Mathur, Sarosh Sopariwalla*

Columbia University
Department of Data Science

## ABSTRACT

Major Depressive Disorder (MDD) is a common disorder that often goes undiagnosed. Most diagnoses of MDD are evaluated through a series of structured interviews. However, with new computer vision techniques, the use of sMRI scans have proven to be a useful tool in diagnosing diseases related to the brain. Problems such as identifying Schizophrenia and Alzheimer's have shown great results when using neural networks. In this paper, we analyze the results of using transfer learning to identify MDD in MRI scans[1]. Our results show that the architectures used are exceptional and show meaningful insights to other problems such as Schizophrenia, but perform poorly on the MDD problem due to lack of a well curated dataset.

***Index Terms*—** Transfer Learning, ResNet, VGG19, Major Depressive Disorder, Schizophrenia

## 1. INTRODUCTION

In the past decade, mental health has become more prominent within society and less neglected within the medical industry. Researchers and hospitals have begun to put more resources to develop methodologies to identify such illnesses effectively. Recently, the most common mental illnesses that have been explored are Schizophrenia [2, 3, 4] and Alzheimer's [5, 6, 7, 8, 9, 10, 11]. However, in this paper we focus on a more challenging and minimally explored problem - Major Depressive Disorder (MDD).

MDD [12] is a persistent low mood and/or anhedonia and is a common condition disorder that can often go undiagnosed. Currently, most diagnoses of MDD are evaluated through a series of structured interviews that look at clinical parameters. Even more so, around two-thirds of all cases of depression, including Major Depressive Disorder (MDD), are undiagnosed due to unreliable symptom-based criteria and lack of a widely accepted quantifiable diagnostic tool.

Within the medical community, there is a huge push and involvement to leverage machine learning in this task. We are interested in using Structural Magnetic Resonance Imaging (sMRI), which can potentially contain important features such as biomarkers to help diagnosis MDD.

In this paper, we explore the use of Transfer Learning [13] to build a robust deep learning model to predict MDD based on sMRIs. Transfer learning is a methodology that uses a model that has been pre-trained on a different dataset and is used as starting point for a model on a second task. The intuition is that the knowledge learned from the pre-trained model might have insights that can be used on our current problem. Formally, a machine uses the knowledge learned from a prior assignment to increase prediction about a new task in transfer learning.

We aim to explore the use of transfer learning and analyze how various pre-trained models perform on our MDD dataset. We first show that our models performs well on other classification problems such as the Schizophrenia dataset, and then we compare results on the MDD dataset.

## 2. DATA

We use two different datasets in this paper, the Schizophrenia dataset and the MDD dataset. We use the Schizophrenia dataset as a baseline to show that the transfer learning models are a good architecture to aid MDD predictions. The MDD dataset is directly related to our application. Both datasets contain structural magnetic resonance neuroimaging (sMRI) scans. These are a subset of MRIs using a non-invasive technique for examining the anatomy and pathology of the brain. The following sub-sections go into detail regarding the breakdown of the datasets.

| MDD Dataset Breakdown | | | |
|---|---|---|---|
| | Cog Normal | Cog Impaired | |
| Train | 7060 | 980 | S/P [1] |
| | 353 | 49 | Patients |
| Validation | 1000 | 100 | S/P [1] |
| | 50 | 5 | Patients |
| Test | 1040 | 100 | S/P [1] |
| | 52 | 5 | Patients |

**Table 1**. Breakdown of MDD dataset.

[1] Slices per Perspective (Axial, Coronal, Sagittal)

### 2.1. Major Depression Disorder Data

The MDD dataset is a private dataset given to us by Professor Jia Guo. The dataset contains sMRI scans from 514 patients. These 3D scans are then sliced into three different perspectives - Axial (divides the brain into top and bottom halves, see Figure 1), Coronal (perpendicular, see Figure 2), and Sagittal

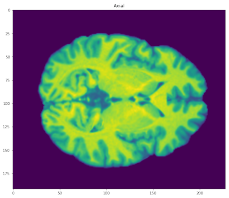

**Fig. 1**. Axial Perspective of MRI Image

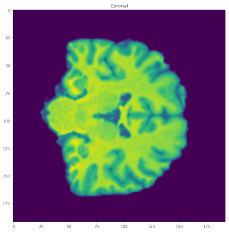

**Fig. 2**. Coronal Perspective of MRI Image

(midline of the brain, see Figure 3 ).

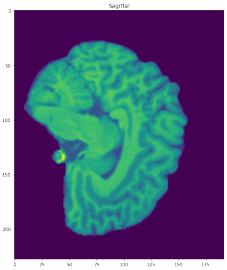

**Fig. 3**. Sagittal Perspective of MRI Image

Table 1 shows the breakdown of the datasets for each perspective. It should be noted that there is an extreme imbalance in the classes, as 91.22% of the dataset is Cognitive Normal MRI scans, while only 8.78% of the dataset contains MDD MRI scans.

### 2.2. Schizophrenia Data

The Schizophrenia dataset is a private dataset given to us by Professor Jia Guo. The dataset contains sMRI scans from 2218 patients. Table 2 shows the breakdown of the datasets for each perspective. Contrary to the MDD dataset, the train and validation data are well balanced, with a larger proportion of Schizophrenia MRI scans over Cognitive Normal. The test data is more equally balanced, having only one more Cognitive Normal patient over Schizophrenia.

| Schizophrenia Dataset Breakdown | | | |
|---|---|---|---|
| | Cog Normal | Cog Impaired | |
| Train | 12600 | 13540 | S/P [2] |
| | 1358 | 728 | Patients |
| Validation | 380 | 740 | S/P [2] |
| | 19 | 37 | Patients |
| Test | 520 | 500 | S/P [2] |
| | 26 | 25 | Patients |

**Table 2**. Breakdown of Schizophrenia dataset.

## 3. METHODS

We implemented various pre-trained neural networks on both datasets to showcase that (i) our chosen model performs well on a well-established problem that has previously shown great results and (ii) see how our MDD dataset performs on various architectures. This section is split up into two sections. The first subsection discusses the various transfer learning models we used, and the second subsection discusses our training settings for reproducibility.

### 3.1. Transfer Learning Models

All transfer learning models that we use are from the Pytorch library. These models are pretrained on the 1000-class ImageNet[14] dataset. When loading the pre-trained model, we freeze the pre-trained weights and substitute the final output layer of each model with a dense layer of 2 hidden units.

The first pre-trained model we used is the ResNet [15] model. This model is a residual learning framework to ease the training of networks that are substantially deeper than those used previously. The author's explicitly reformulate the layers as learning residual functions with reference to the layer inputs, instead of learning unreferenced functions [15]. ResNet is a common and powerful architecture that has shown great results on various image classification problems.

Our second pre-trained model is one previously discussed in class VGG-19[16]. This model is also very popular in both industry and academia. We have chosen the 19 weight layers configuration of this model. The main contribution of this architecture is that they use a 3x3 convolution filter throughout their whole net.

We train both full datasets (all perspectives) on both pre-trained models. Furthermore, we also train each model on specific perspectives from each dataset. We did this so we can compare how the model performs with all perspectives vs. trained individually.

[2] Slices per Perspective (Axial, Coronal, Sagittal)

| | Axial | | Coronal | | Sagittal | | All Perspectives | |
|---|---|---|---|---|---|---|---|---|
| | MDD | Schizophrenia | MDD | Schizophrenia | MDD | Schizophrenia | MDD | Schizophrenia |
| Accuracy | 0.9122 | 0.8300 | 0.9122 | 0.5300 | 0.9122 | 0.5300 | 0.9122 | 0.5300 |
| Sensitivity | 0.0000 | 0.8000 | 0.0000 | 0.5300 | 0.0000 | 0.5500 | 0.0000 | 0.5700 |
| Specificity | 1.0000 | 0.7900 | 1.0000 | 0.6400 | 1.0000 | 0.6600 | 1.0000 | 0.8500 |
| Precision | 0.0000 | 0.8700 | 0.0000 | 0.4300 | 0.0000 | 0.2300 | 0.0000 | 0.1900 |

**Table 3**. Performance of VGG-19 Network on MDD and Schizophrenia Datasets after Transfer Learning from ImageNet

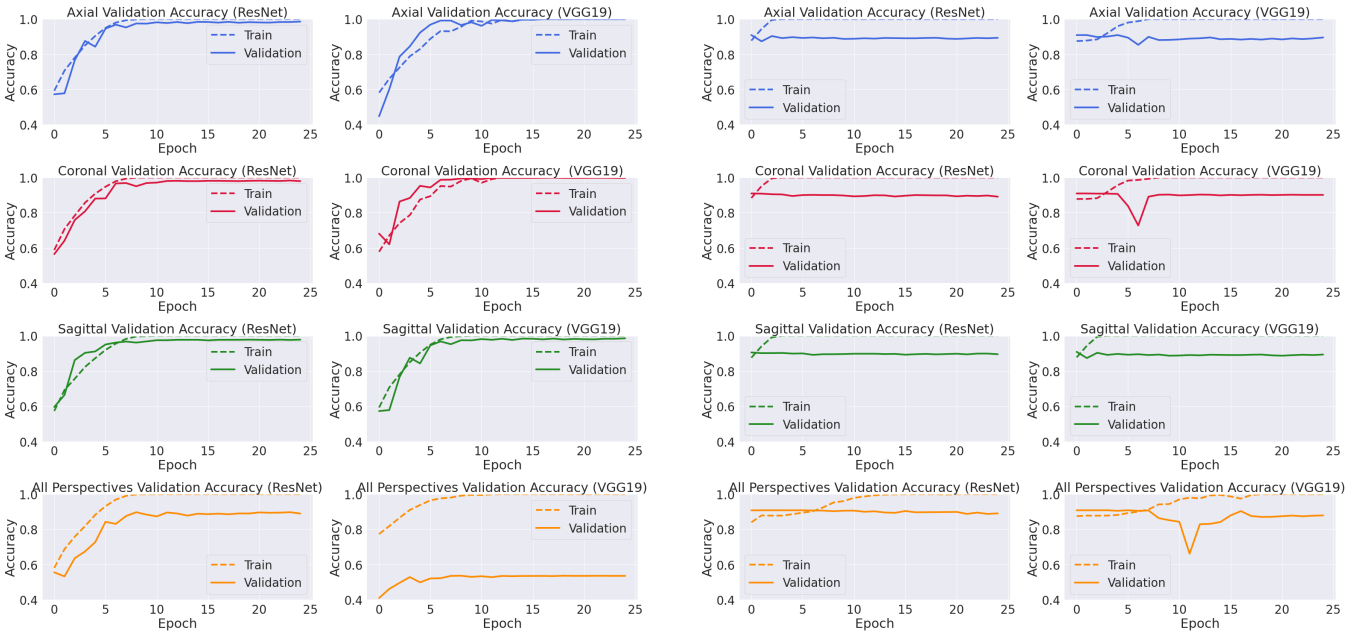


**Fig. 4**. Training and Validation Curves for ResNet and VGG-19 on Schizophrenia (left) and MDD (right) 2D MRI-slices.

| | Axial | | Coronal | | Sagittal | | All Perspectives | |
|---|---|---|---|---|---|---|---|---|
| | MDD | Schizophrenia | MDD | Schizophrenia | MDD | Schizophrenia | MDD | Schizophrenia |
| Accuracy | 0.9122 | 0.9900 | 0.9105 | 0.6400 | 0.9061 | 0.7100 | 0.9122 | 0.8800 |
| Sensitivity | 0.000 | 0.9800 | 0.000 | 0.6100 | 0.000 | 0.6600 | 0.000 | 0.8500 |
| Specificity | 1.000 | 0.9500 | 0.9981 | 0.5200 | 0.9933 | 0.5900 | 1.000 | 0.8400 |
| Precision | 0.000 | 0.9900 | 0.000 | 0.7700 | 0.000 | 0.8200 | 0.000 | 0.9330 |

**Table 4**. Performance of ResNet on MDD and Schizophrenia Datasets after Transfer Learning from ImageNet

### 3.2. Model Training Settings

There were several considerations when applying transfer learning on our MRI data. Firstly, all our models were pre-trained on the open-source dataset, ImageNet [14], which consists of three channel RGB images. However, MRI data contains only one channel, so it cannot be fed directly into any of our networks. We chose to solve this problem by replicating the one channel we were given three times. Aside from replicating the given channel, we applied a center crop of size 192 on the Axial slices. This allowed us to reduce the image size while keeping all the information in the MRI scan. Center crop could not be applied on Sagittal or Coronal slices, since that would remove relevant information. Finally, the scans were normalized according to the well-known means and standard deviations of the ImageNet dataset. The means for the three channels are 0.480, 0.456, and 0.406. The corresponding standard deviations are 0.229, 0.224, and 0.225.

With the pre-processing complete, we only needed to re-train the last fully-connected layer of the networks. This is called fine-tuning within the transfer learning paradigm. We used PyTorch's SGD optimizer with a dynamic learning rate that decreased by one-tenth every seven epochs. (After several epochs, we expect the algorithm to near the global minimum, so it makes sense to decrease the learning rate to prevent our algorithm from diverging). Each model was trained (and saved) for twenty-five epochs, without early stopping. After the train-validation loop, we selected the saved model which maximized the validation accuracy and minimized the validation loss.

## 4. RESULTS

We successfully ran the two pre-trained neural networks, ResNet and VGG-19, on both the Schizophrenia and MDD datasets. We trained both networks on each of the three perspectives (Axial, Coronal, Sagittal), as well as a combination of "All" perspectives. We produced predictions of whether or not the MRI is MDD or Schizophrenia, respective to the dataset. Table 3 and Table 4 show the performance of the VGG-19 and ResNet models respectively. The training and validation curves can be seen in Figure 4.

### 4.1. Results on Schizophrenia Data

The VGG-19 models trained on each of the Coronal, Sagittal, and All perspectives all resulted in an accuracy of about 53%. However, the Axial perspective had a noticeably greater accuracy of of about 83%. The Axial perspective had a visibly greater sensitivity and precision compared to the other perspectives as well. The All perspective had the highest specificity of 85%.

The ResNet models all had a higher performance than the VGG-19 models. Specifically, training on the Axial perspective greatly outshined the other perspectives, with an accuracy of 99%, sensitivity of 98%, specificity of 95%, and precision of 99%. Training on the All perspective, which includes Axial in its dataset, had the second best performance, although Axial still had a much greater accuracy, sensitivity, and specificity. The Coronal perspective had the worst performance, with a noticeably lower accuracy of 64% and specificity of 52%.

In Figure 4 we can see that the training and validation accuracies converge to around 100% for the Axial, Coronal, and Sagittal perspectives for both VGG-19 and ResNet. However, the validation accuracy for the All perspective only converged to around 85%.

### 4.2. Results on Major Depression Disorder Data

The VGG-19 models trained on each of the Axial, Coronal, Sagittal, and All perspectives all resulted in an accuracy of 91.22%. Every model also reported a sensitivity of 0%, precision of 0%, and specificity of 100%.

The ResNet models had slightly more variance than the VGG-19 models, where the Axial and All perspectives also resulted in an accuracy of 91.22%. Coronal and Sagittal perspectives had a slightly worse performance of 91.05% and 90.61% respectively. This is a very similar performance to the VGG-19 models. Once again, all of the ResNet models reported 0% for both sensitivity and precision. The specificities for the Axial and All perspectives were both 100%, while Coronal had a specificity of 99.81% and Sagittal 99.33%.

In Figure 4 we can see that the training accuracies converge to around 100% for the Axial, Coronal, Sagittal, and All perspectives for both VGG-19 and ResNet. However, the validation accuracy for the Axial, Coronal, Sagittal, and All perspectives only converge to about 87% for VGG-19 and ResNet. In the All perspective VGG-19 model, we can even see a large dip in accuracy at epoch 11.

## 5. DISCUSSION

From our training and validation plots, we observe that the all models on all the various datasets converge and have a relatively high accuracy. However, these metrics are misleading with respect to the MDD dataset.

We notice that in both Table 3 and Table 4, the test accuracy is high. However, the other metrics - sensitivity and specificity - showcase highly unbalanced metrics. With respect to the MDD dataset, we report that the specificity is almost almost a 100%. In contrast, the sensitivity is always 0%. As a matter of fact, some of the MDD models end up with an accuracy of 91.22%, which is the exact percentage of Cognitive Normal scans in the dataset. This indicates that the model is not learning anything and is just predicting Cognitive Normal for all examples in the test set. This can be attributed to (i) the imbalance of classes in the MDD datasets and (ii) not having enough distinct patients in the datasets. Having an

imbalance in the dataset leads to the neural network not being able to properly distinguish patterns and characteristics between the two classes. Not having enough distinct patients in the data results in the model not able to generalize the MDD problem to unknown distinct patients. Even if we perform data augmentation on our current dataset, we would only be augmenting data on a small patient sample, which would not help in generalizing the data.

In contrast, for the schizophrenia dataset, we see much better results across all metrics. Both Table 3 and Table 4 show that the schizophrenia dataset is able to effectively distinguish healthy and non-healthy classes based on the sensitivity and specificity metrics across all models and subsets of data.

We also observe that between both datasets, the ResNet model shows better results. This can be attributed to the architectural differences between the two models. VGG-19 has more layers and thus has a tendency to overfit, leading to a relatively worse performance on the test set.

We also notice that between both datasets and models, the Axial subset of the data tends to have better metrics compared to other data subsets. We see that in Figure 1, this axis of the scan shows more information of the brain compared to the other perspectives. This allows the models to extract more patterns and characteristics to distinguish the classes.

## 6. CONCLUSION

In this paper, we explore the use of transfer learning on the MDD application. We run the MDD dataset on two pretrained models - ResNet and VGG-19. Both these architectures have proven to report great results on the 100-class ImageNet dataset. Furthermore, we also ran these two models on a schizophrenia dataset to show that these models perform well on similar problems and sMRI scans. From our results and discussion, we observed that the models perform well on schizophrenia but not on MDD. Two reasons arise from our analysis, (i) a high class imbalance in the MDD data, and (ii) not enough distinct patients in the MDD data to allow for generalization. We can conclude that for the MDD task, the problem arises not from the model choice but rather the dataset. It is known that neural networks rely heavily on a well curated dataset to alleviate bias and to increase generalization. As future work, we would like to focus on curating a better data set for MDD. This would entail in acquiring more distinct patients with MDD. With a well curated MDD dataset, based on our insights from the schizophrenia results, the models should effectively be able to distinguish the two classes.

## 7. REFERENCES


[1] Marzieh Mousavian, Jianhua Chen, and Steven Greening, "Depression detection using feature extraction and deep learning from smri images," in *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, 2019, pp. 1731–1736.

[2] J. Oh, B. L. Oh, K. U. Lee, J. H. Chae, and K. Yun, "Identifying Schizophrenia Using Structural MRI With a Deep Learning Algorithm," *Front Psychiatry*, vol. 11, pp. 16, 2020.

[3] J. A. Cortes-Briones, N. I. Tapia-Rivas, D. C. D'Souza, and P. A. Estevez, "Going deep into schizophrenia with artificial intelligence," *Schizophr Res*, Jun 2021.

[4] G. B. Chand, D. B. Dwyer, G. Erus, A. Sotiras, E. Varol, D. Srinivasan, J. Doshi, R. Pomponio, A. Pigoni, P. Dazzan, R. S. Kahn, H. G. Schnack, M. V. Zanetti, E. Meisenzahl, G. F. Busatto, B. Crespo-Facorro, C. Pantelis, S. J. Wood, C. Zhuo, R. T. Shinohara, H. Shou, Y. Fan, R. C. Gur, R. E. Gur, T. D. Satterthwaite, N. Koutsouleris, D. H. Wolf, and C. Davatzikos, "Two distinct neuroanatomical subtypes of schizophrenia revealed using machine learning," *Brain*, vol. 143, no. 3, pp. 1027–1038, 03 2020.

[5] J. H. So, N. Madusanka, H. K. Choi, B. K. Choi, and H. G. Park, "Deep Learning for Alzheimer's Disease Classification using Texture Features," *Curr Med Imaging Rev*, vol. 15, no. 7, pp. 689–698, 2019.

[6] D. Ushizima, Y. Chen, M. Alegro, D. Ovando, R. Eser, W. Lee, K. Poon, A. Shankar, N. Kantamneni, S. Satrawada, E. A. Junior, H. Heinsen, D. Tosun, and L. T. Grinberg, "Deep learning for Alzheimer's disease: Mapping large-scale histological tau protein for neuroimaging biomarker validation," *Neuroimage*, vol. 248, pp. 118790, 03 2022.

[7] M. F. Romano and V. B. Kolachalama, "Deep learning for subtyping the Alzheimer's disease spectrum," *Trends Mol Med*, vol. 28, no. 2, pp. 81–83, 02 2022.

[8] T. Jo, K. Nho, and A. J. Saykin, "Deep Learning in Alzheimer's Disease: Diagnostic Classification and Prognostic Prediction Using Neuroimaging Data," *Front Aging Neurosci*, vol. 11, pp. 220, 2019.

[9] A. K. Ambastha and T. Y. Leong, "A Deep Learning Approach to Neuroanatomical Characterisation of Alzheimer's Disease," *Stud Health Technol Inform*, vol. 245, pp. 1249, 2017.

[10] C. L. Saratxaga, I. Moya, A. Picón, M. Acosta, A. Moreno-Fernandez-de Leceta, E. Garrote, and

A. Bereciartua-Perez, "MRI Deep Learning-Based Solution for Alzheimer's Disease Prediction," *J Pers Med*, vol. 11, no. 9, Sep 2021.

[11] S. Koga, A. Ikeda, and D. W. Dickson, "Deep learning-based model for diagnosing Alzheimer's disease and tauopathies," *Neuropathol Appl Neurobiol*, vol. 48, no. 1, pp. e12759, 02 2022.

[12] C. Otte, S. M. Gold, B. W. Penninx, C. M. Pariante, A. Etkin, M. Fava, D. C. Mohr, and A. F. Schatzberg, "Major depressive disorder," *Nat Rev Dis Primers*, vol. 2, pp. 16065, 09 2016.

[13] Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, and Qing He, "A comprehensive survey on transfer learning," 2019.

[14] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision (IJCV)*, vol. 115, no. 3, pp. 211–252, 2015.

[15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," 2015.

[16] Karen Simonyan and Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014.