Optimizing Schizophrenia Detection from MRI: The Role of Preprocessing and Data Augmentation in Machine Learning

Ilia Golub Pratit Kandel Prosperity Oguama

1.0 Introduction

Structural brain abnormalities such as gray and white matter alterations (Lee et al., 2020; Xiao et al., 2018; Zhang et al., 2015), cortical thinning (Goldman et al., 2009; Madre et al., 2020), and ventricular enlargement (Svancer & Spaniel, 2021) have been observed in patients with schizophrenia, making it possible for medical imaging modalities to identify biomarkers for schizophrenia (Gengeç & Andaç, 2023). Because of its high spatial resolution, magnetic resonance imaging (MRI) is used to understand the structural and functional abnormalities related to schizophrenia (Sadeghi et al., 2022). Studying brain anatomy using structural MRI (sMRI) has been shown to provide a reliable assessment of schizophrenia (SCZ) diagnosis (Verma et al.,2023).

Machine learning (ML) and deep learning (DL) algorithms have been employed for the past decade to detect differences in structure between SCZ and healthy brains (Verma et al.,2023). However, most previous work is focused on identifying the best algorithm or combination of algorithms to accurately classify between SCZ and healthy scans, with little attention being paid to the effect of preprocessing methods on the outcome of classification tasks.

This study aims to assess the various preprocessing techniques proposed in the literature for SCZ classification to develop an optimal and reproducible preprocessing pipeline for SCZ diagnosis from structural MRIs.

2.0 Literature Review

Numerous studies have been explored so far in the classification of schizophrenia using structural magnetic resonance imaging (sMRI) from publicly available neuroimaging datasets, employing a range of machine learning (ML) and deep learning (DL) techniques. The ML approach for this task typically involves feature extraction and principal component analysis (PCA), after which the features are classified with a support vector machine (SVM), logistic regression, extra-trees, AdaBoost, or XGBoost, with accuracy scores ranging from 50% to 80% (Benli & Andaç, 2023; Nimkar & Kubal, 2018; Rustam & Saragih, 2020). DL approaches, on the other hand, typically produce higher accuracies greater than 80%, with the most common architectures being pretrained models like VGG16, ResNet101, and EfficientNetB0 (Zhang et., 2023; Oh et al., 2020; Oh et al., 2019). However, schizophrenia detection from MRI scans remains challenging because the changes in brain structure are often subtle and may not manifest until years after the disease onset. Additionally, certain anatomical features found in schizophrenia are shared with other neurodegenerative diseases, making the application of imaging for diagnosis in the clinical setting limited (Oh et al., 2020).

The focus in literature is mainly on optimizing model architectures and feature extraction techniques, with little emphasis on preprocessing techniques. This is the gap we sought to fill in the course of this project.

3.0 Methods

3.1 Dataset and Exploratory Data Analysis

The data was obtained from the SchizConnect database (accessed January 1, 2025), which is a repository of structural and functional MRI data. The data obtained originates from the Center of Biomedical Research Excellence (COBRE) and Mind Clinical Imaging Consortium Share (MCICShare) datasets in NifTI format. It includes data from 242 individuals, 111 with schizophrenia and 131 healthy controls. The data consists of 242 three-dimensional (3D) MRI scans with 192 slices each, bringing the total number of slices to 11,904. We retrieved only T1-weighted images (MPRAGE sequence) acquired using 3T MRI scanners. This was done to ensure uniformity of the data obtained. Figure 1 presents an example of the raw data. Clinical data was made available in form of CSV files. Figure 2 summarizes the subject demographics, see supplementary data for a full description.

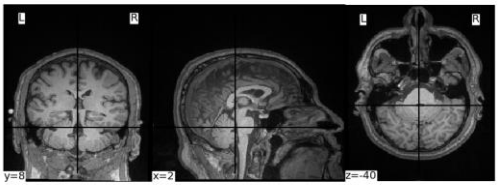

Figure 1: Three image slices representing the average of a 3D volume in three views: dorsal, lateral, and

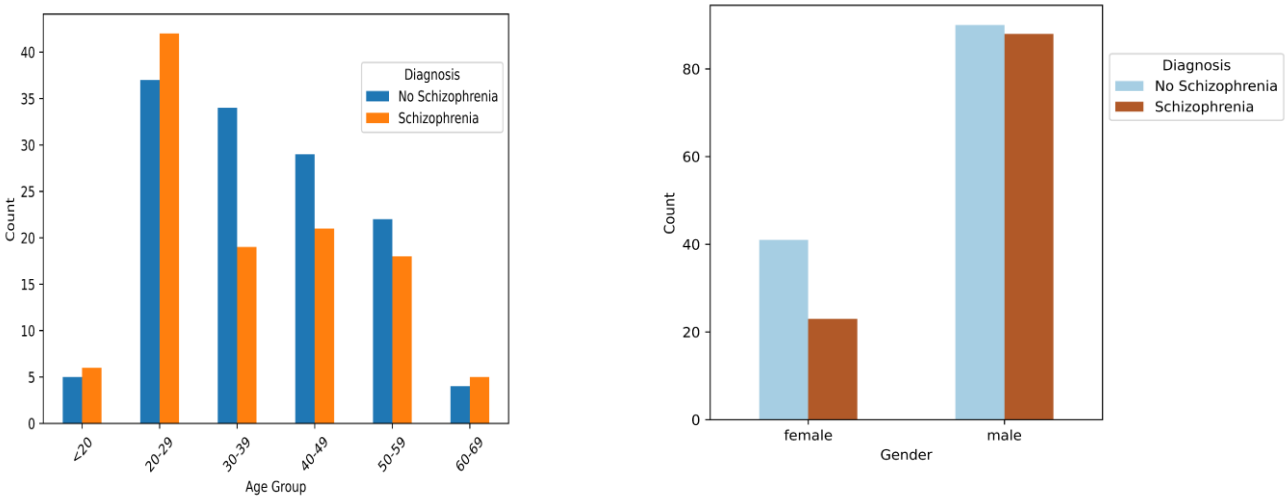

Figure 2: Demographic data. A) Age distribution of the dataset by class B) Gender distribution of the dataset by class.

3.2 Image Preprocessing and Augmentation

We experimented with different preprocessing and data augmentation techniques in isolation and in combinations to determine which pipeline yields the best performance by comparing accuracy, recall, precision, and F1 scores. Table 1 presents the preprocessing and augmentation techniques used, their respective hyperparameters, and tools used. Figure 3 shows a comparison of a raw image and a preprocessed image, and their corresponding metrics.

Table 1: Summary of preprocessing and augmentation techniques, hyperparameters and tools used

Preprocessing/ Augmentation Step	Hyperparameters	Tools
Resampling	Voxel size: 2x2x2	Nibabel library
Intensity normalization	Min-max scaling: [0, 1]	NumPy
Brain extraction	Modality: T1, U-net	ANTsPyNet
Cropping	Size: based on the largest bounding box among all images	Scikit-image, NumPy
Smoothing	Gaussian filter: sigma=0.5	SciPy
Translation	Mode: nearest	SciPy
Rotation	Mode: nearest	SciPy
Shearing	Mode: nearest	SciPy
Contrast enhancement	Method: gamma	Numpy

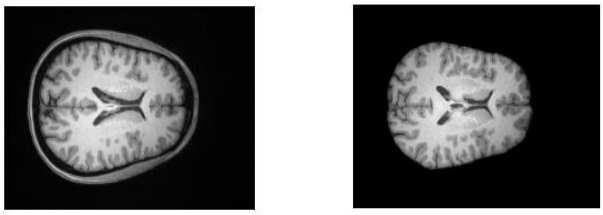

Figure 3: Comparison of A) raw scan and B) preprocessed scan. For preprocessed scans, the standard deviation of SNR values is relatively low (SNR = 25 ± 2), which ensures consistency in the image quality and prevents the model from making predictions based on possible artifacts.

3.3 Model Architecture and Training Process

We trained four model architectures on the raw images to decide on the best process to maximize accuracy and training speed for this task. An SVM used alongside principal component analysis (PCA) for dimensionality reduction outperformed combinations with pre-trained models as feature extractors, probably due to an inability of the SVM to separate the complex features from these models, even with an RBF kernel. Grid search was performed to select hyperparameters for the SVM pipeline, which yielded a radial basis function (RBF) kernel, C and gamma values of 100 and 0.0001 respectively, and PCA with 500 components as optimal. Figure 3 shows the final model training process. Additionally, we observed that non-informative MRI slices, typically located at the beginning and end of the scans, were introducing noise into the model. We therefore experimented

with removing different percentages of slices (10%, 15%, and 20%) from both ends and found that excluding 15% from the start and 15% from the end yielded the best performance.

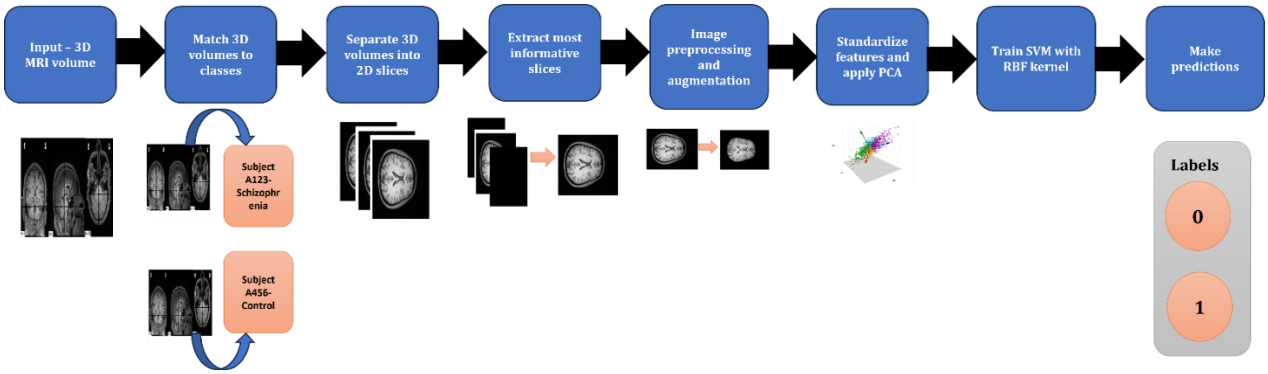

Figure 3: Overview of the training pipeline. The original scans are 3D images in .nii format. They are matched to the corresponding class and split into 2D slices, giving the same label to all slices from the same subject. The most informative slices (70% from the middle of the stack) are preprocessed / augmented, standardized and fed into an RBF-kernel SVM, which classifies the slices as schizophrenic or non-schizophrenic.

3.4 Model Interpretability

To identify key brain regions associated with schizophrenia prediction, we utilized 3D Grad-CAM++ visualization on the trained ResNet-18 3D model. The model was trained MRI scans preprocessed and augmented with the determined optimal pipeline (resampling + normalization + brain extraction + translation + enhancement) to classify scans as healthy or schizophrenic. After model training, we generated Grad-CAM++ heatmaps for individual MRI scans. Heatmaps were analyzed in three main anatomical planes: Sagittal (side-to-side, ear-to-ear slices), Coronal (front-to-back, face-to-back head slices) and Axial (top-to-bottom, bird's eye view slices).

3.5 Model Deployment

We deployed the final model on a web application built with Streamlit. The application takes in 3D MRI scans, performs classification and gives an output of the prediction- schizophrenic or non-schizophrenic. Test the application here.

4.0 Results

4.1 Preprocessing and Augmentation

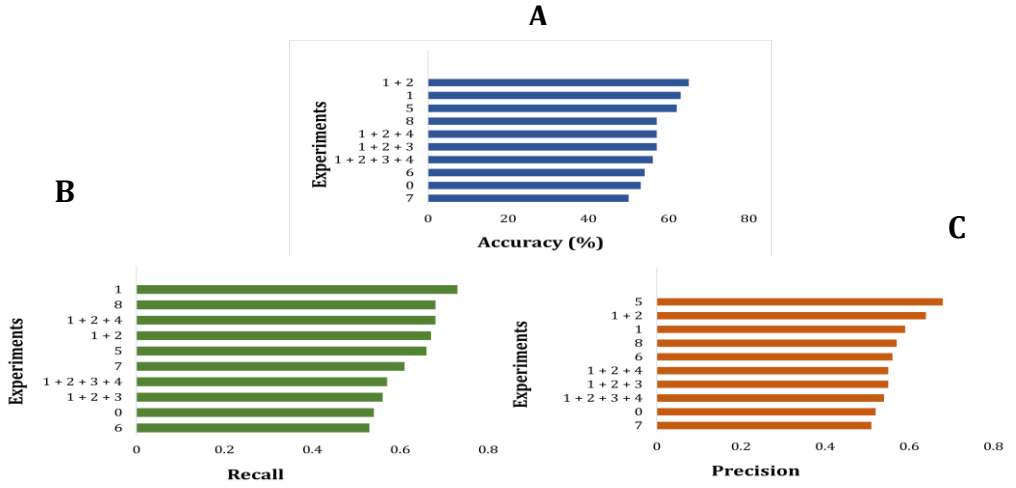

Figure 5: Summary of the metrics across different experiments. A) comparison of accuracy scores across the experiments. The highest accuracy was recorded in the third experiment combining resampling, normalization and brain extraction. B) comparison of recall scores. The second experiment combining resampling and normalization yielded the highest recall score. C) comparison of precision scores. The translation experiment produced the highest score. The numbers along the y-axis represent preprocessing and augmentation scenarios: 0 — raw images, 1 — resampling and normalization, 2 — brain extraction, 3 — cropping, 4 — smoothing, 5 — translation, 6 — rotation, 7 — shearing, 8— contrast enhancement.

4.2 Model Interpretability

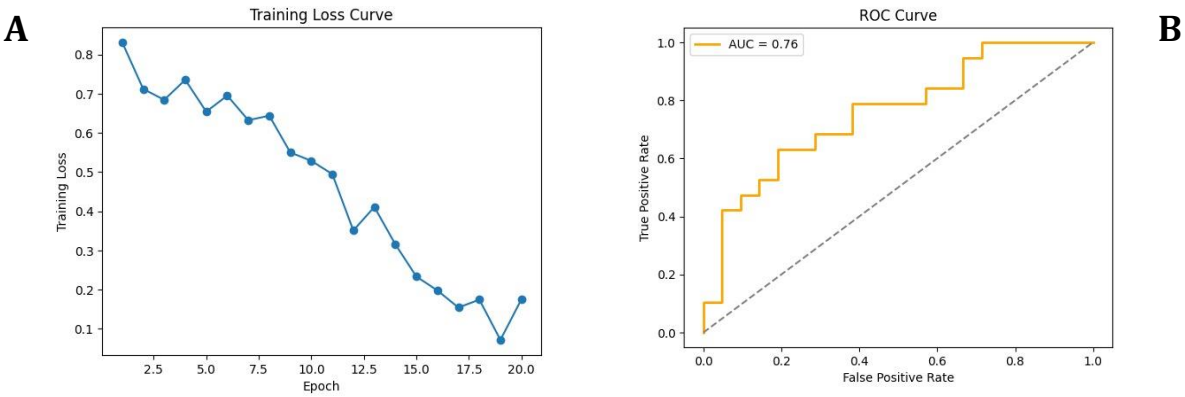

Figure 6: Results from final training for model interpretability with an accuracy of 83%. A) Training loss curve B) ROC Curve

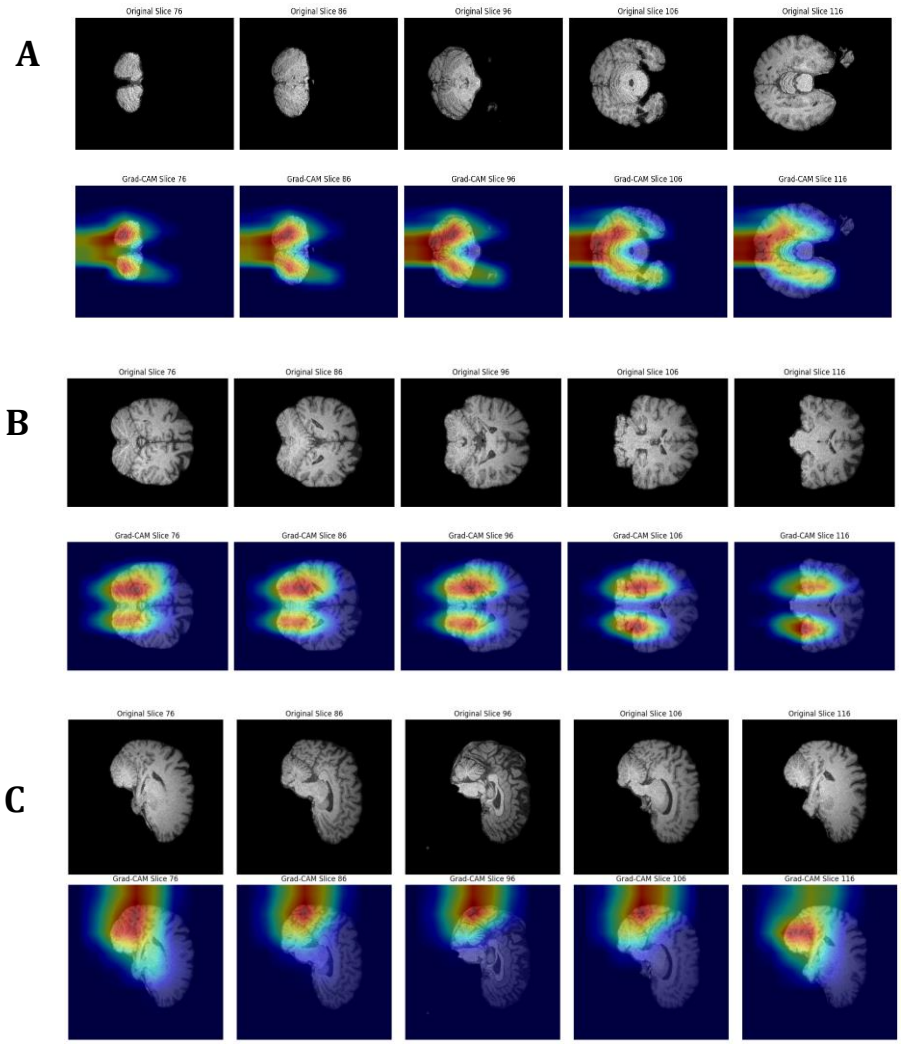

Figure 7: Brain regions with the greatest effect on the model's predictions. A) Axial plane B) Coronal plane C) Sagittal plane.

5.0 Discussion

Among the preprocessing experiments, resampling and normalization yielded the highest performance boost, increasing accuracy to 63% (a 18.87% improvement over baseline). This likely stems from balancing feature contributions, enabling the model to make more reliable predictions. Interestingly, this result contrasts with the conventional view that normalization may obscure clinically important intensity variations in medical images. The success of normalization in this context underscores the need to carefully tailor preprocessing to the specific imaging task.

Brain extraction (skull stripping) improved the accuracy obtained from combining resampling, normalization and brain extraction, but slightly lowered the recall, which is a major focus in solving medical problems. This suggests that removing non-brain tissues heightens the model's sensitivity to peripheral features that contribute to accurate classification, but may reduce the model's sensitivity to schizophrenia cases.

Cropping and smoothing generally reduced the model's accuracy and recall scores, which can be attributed to cutting out significant brain regions and blurring fine anatomical details respectively.

Although some preprocessing techniques improve model metrics individually, combining multiple preprocessing steps often degraded performance, particularly when all techniques were applied together. Some preprocessing methods may interact adversely, introducing distortions that obscure key features instead of clarifying them. This highlights the importance of carefully evaluating preprocessing combinations rather than assuming additive benefits.

Regarding data augmentation, all the methods increased the model's accuracy, consistent with prior studies showing that geometric transformations help models generalize better by diversifying the training data. However, the magnitude of the effects of increasing the data were lower than expected and could point to the inherent difficulty of identifying informative features for schizophrenia detection from structural MRIs, because the changes in brain structure are often subtle and may not manifest until years after the disease onset.

Considering the metrics from the experiments, the combination of resampling, normalization, brain extraction, translation, and contrast enhancement is the optimal combination for this task. The final model, a 3D ResNet-18, was trained with this combination and achieved the best overall performance: Accuracy: 83%, Recall: 0.85, F1-Score: 0.86

These results demonstrate that a carefully optimized combination of preprocessing and augmentation can substantially enhance classification performance. Importantly, the high recall value is crucial for medical applications, where correctly identifying positive cases is a top priority.

Additionally, the study reinforces that preprocessing techniques like skull stripping, cropping, and smoothing, although beneficial in some contexts, do not universally improve model performance. Their effectiveness heavily depends on dataset characteristics and the overall pipeline design.

In conclusion, the findings highlight the need for standardized yet adaptable preprocessing frameworks in medical imaging, particularly for challenging tasks like schizophrenia detection. Caution is warranted when stacking multiple preprocessing steps, as poorly chosen combinations may hurt rather than improve performance.

The Grad-CAM++ analysis suggests that the posterior brain regions, particularly the cerebellum, temporal lobes, and possibly parts of the occipital lobe, play an important role in the model's decision-making when predicting schizophrenia. These findings align with existing neuroimaging literature: structural and functional abnormalities in the cerebellum have been increasingly implicated in schizophrenia, particularly in theories of "cognitive dysmetria" (Andreasen et al., 1998). Alterations in the temporal lobes, including the hippocampus, have been consistently reported, especially regarding auditory hallucinations and memory disturbances. (Heckers, S., & Konradi, C. et al., 2015). Occipital lobe involvement has been less emphasized historically, but some studies have reported visual processing abnormalities in schizophrenia patients (Butler, P. D., Silverstein, S. M., & Dakin, S. C., 2008). The observed bilateral symmetry in coronal slices matches the clinical understanding that schizophrenia affects brain structures relatively symmetrically in many cases. Overall, Grad-CAM++ successfully highlighted biologically plausible regions, supporting both the model's internal logic and known pathophysiology of schizophrenia.

Resources

Supplementary materials GitHub repository Web application

References

Andreasen, N. C., Paradiso, S., & O'Leary, D. S. (1998). "Cognitive dysmetria" as an integrative theory of schizophrenia: a dysfunction in cortical-subcortical-cerebellar circuitry? Schizophrenia bulletin, 24(2), 203-218.

Butler, P. D., Silverstein, S. M., & Dakin, S. C. (2008). Visual perception and its impairment in schizophrenia. Biological psychiatry, 64(1), 40-47.

Charlson, F. J., Ferrari, A. J., Santomauro, D. F., Diminic, S., Stockings, E., Scott, J. G., ... & Whiteford, H. A. (2018). Global epidemiology and burden of schizophrenia: findings from the global burden of disease study 2016. *Schizophrenia bulletin*, *44*(6), 1195-1203.

Fuchs, T. (2010). Subjectivity and intersubjectivity in psychiatric diagnosis. *Psychopathology*, *43*(4), 268-274.

Gengeç Benli, Ş., & Andaç, M. (2023). Constructing the schizophrenia recognition method employing GLCM features from multiple brain regions and machine learning techniques. *Diagnostics*, *13*(13), 2140.

Goldman, A. L., Pezawas, L., Mattay, V. S., Fischl, B., Verchinski, B. A., Chen, Q., ... & Meyer-Lindenberg, A. (2009). Widespread reductions of cortical thickness in schizophrenia and spectrum disorders and evidence of heritability. *Archives of general psychiatry*, *66*(5), 467-477.

Heckers, S., & Konradi, C. (2015). GABAergic mechanisms of hippocampal hyperactivity in schizophrenia. Schizophrenia research, 167(1-3), 4-11.

Keshavan, M. S., Collin, G., Guimond, S., Kelly, S., Prasad, K. M., & Lizano, P. (2019). Neuroimaging in schizophrenia. *Neuroimaging Clinics of North America*, *30*(1), 73.

Lee, D. K., Lee, H., Park, K., Joh, E., Kim, C. E., & Ryu, S. (2020). Common gray and white matter abnormalities in schizophrenia and bipolar disorder. *PLoS One*, *15*(5), e0232826.

Madre, M., Canales-Rodríguez, E. J., Fuentes-Claramonte, P., Alonso-Lana, S., Salgado-Pineda, P., Guerrero-Pedraza, A., ... & Pomarol-Clotet, E. (2020). Structural abnormality in schizophrenia versus bipolar disorder: a whole brain cortical thickness, surface area, volume and gyrification analyses. *NeuroImage: Clinical*, *25*, 102131.

Nimkar, A. V., & Kubal, D. R. (2018, August). Optimization of schizophrenia diagnosis prediction using machine learning techniques. In 2018 4th international conference on computer and information sciences (ICCOINS) (pp. 1-6). IEEE.

Oh, J., Oh, B. L., Lee, K. U., Chae, J. H., & Yun, K. (2020). Identifying schizophrenia using structural MRI with a deep learning algorithm. Frontiers in psychiatry, 11, 16.

Oh, K., Kim, W., Shen, G., Piao, Y., Kang, N. I., Oh, I. S., & Chung, Y. C. (2019). Classification of schizophrenia and normal controls using 3D convolutional neural network and outcome visualization. Schizophrenia research, 212, 186-195.

Rustam, Z., & Saragih, G. S. (2020). Prediction schizophrenia using random forest. TELKOMNIKA (Telecommunication Computing Electronics and Control), 18(3), 1433-1438.

Sadeghi, D., Shoeibi, A., Ghassemi, N., Moridian, P., Khadem, A., Alizadehsani, R., ... & Acharya, U. R. (2022). An overview of artificial intelligence techniques for diagnosis of Schizophrenia based on magnetic resonance imaging modalities: Methods, challenges, and future works. *Computers in Biology and Medicine*, *146*, 105554.

Svancer, P., & Spaniel, F. (2021). Brain ventricular volume changes in schizophrenia. A narrative review. *Neuroscience Letters*, *759*, 136065.

Verma, S., Goel, T., Tanveer, M., Ding, W., Sharma, R., & Murugan, R. (2023). Machine learning techniques for the Schizophrenia diagnosis: A comprehensive review and future research directions. *Journal of Ambient Intelligence and Humanized Computing*, *14*(5), 4795-4807.

Xiao, Y., Sun, H., Shi, S., Jiang, D., Tao, B., Zhao, Y., ... & Lui, S. (2018). White matter abnormalities in never-treated patients with long-term schizophrenia. *American Journal of Psychiatry*, *175*(11), 1129-1136.

Zhang, J., Rao, V. M., Tian, Y., Yang, Y., Acosta, N., Wan, Z., ... & Guo, J. (2023). Detecting schizophrenia with 3D structural brain MRI using deep learning. Scientific reports, 13(1), 14433.

Zhang, W., Deng, W., Yao, L., Xiao, Y., Li, F., Liu, J., ... & Gong, Q. (2015). Brain structural abnormalities in a group of never-medicated patients with long-term schizophrenia. *American Journal of Psychiatry*, *172*(10), 995-1003.