

Applications of Machine Learning in Schizophrenia Prediction

Swapnali Kulkarni¹, Pravin Kumawat², Dipesh Nikam³, Manisha Mali⁴

^{1,2,3,4}

Department of Computer Engineering, Vishwakarma Institute of Information Technology, Pune, India
Email: ¹swapnali.22211572@viit.ac.in, ²pravin.22210292@viit.ac.in, ³dipesh.22210027@viit.ac.in,

⁴

manisha.mali@viit.ac.in

Abstract - A comprehensive review of recent challenges faced by schizophrenic patients and how machine learning can be applied to overcome these challenges is being reviewed in this paper. In this paper, we explore various ML methodologies and techniques that contribute to development of personalized care and treatment systems. Showing adverse effects on behavioral patterns of an individual, such as disorganized speech and delusions, Schizophrenia is a critical disorder. Schizophrenic persons' may feel like they have lost touch with reality, bothering their family and friends. Symptoms for schizophrenia vary from person to person, however, three common categories are: psychotic, negative, and cognitive. Recognizing symptoms and seeking help at the earliest is very important in the case of this disorder. While there's no cure for schizophrenia, the study is leading to innovative and safer treatments. The current methods are not effective as the disease takes an average of 10 years to become diagnosed. To address this need, ML (machine learning) techniques are applied. The purpose of this study is to summarize data on the use of ML techniques in the prediction of schizophrenia, thereby aiding in the earliest and timeliest diagnosis of the disorder.

Keywords: Schizophrenia, Machine learning, EEG(Electroencephalogram), SVM (Support Vector Machine), CNN (Convolutional Neural Networks)

I. INTRODUCTION

Seeming to appear from disturbances with brain growth brought by inheritable or environmental factors, or both, schizophrenia is a complex, miscellaneous behavioral and cognitive disorder. Schizophrenia affects around 1% of the global population.[4] It is imperative to treat the disorder as soon as possible after its onset. A delay in treatment can lead to brain volume depletion, which has adverse effects on long-term care for patients. Diagnosing schizophrenia is chiefly done using psychiatric inquiry forms and clinical investigations. The various tools that help psychiatrists diagnose the disorder are very important for the reasons discussed above. One of the most important executive functions associated

with schizophrenia is response inhibition, which underlies behavioural flexibility, allowing the suspension of highly automated but contextually inappropriate actions.[1]

The formal thought problem of schizophrenia patients may typically be noticed directly in terms of thought, language, and communication (TLC) impairments. Positive and negative disorders are additional classifications for these types of conditions. The term "poor speech" or "content poverty" primarily describes negative mental disorders. Patients with positive thought disorder, on the other hand, can speak incoherently, too quickly, or with loosely connected ideas, which makes it hard for other people to grasp what they're saying. Schizophrenia patients exhibit typical positive (hallucinations, delusions, etc.) and negative (alogia, antisociality, anhedonia, etc.) symptoms in addition to formal thought dysfunction.

Several scales can be used to assess schizophrenia symptoms objectively, including the Scale for the Assessment of Thought, Language, and Communication (TLC), Positive and Negative Syndrome Scale (PANSS), Negative Symptom Assessment, and so on. These examinations typically involve a lengthy interview between mental health specialists and patients. As a result, these scales are not commonly utilized in clinical practice. Schizophrenia usually begins in youth or early adulthood. Patients who receive treatment



early on have a better likelihood of achieving clinical remission and even functional recovery.[3]

In clinical settings, well-trained clinicians manually evaluate schizophrenia patients' speeches based on semantic content, grammatical coherence, dialogue rapport, and so on. [2] Schizophrenia disorder has high mortality rates, often due to suicide, and significantly impacts patients' lives and society. ML offers the potential to enhance early diagnosis, reducing the substantial economic burden on patients and society. Various ML algorithms, including random forests, support vector machines, and gradient boosting, have been employed to predict these disorders by analyzing speech signals, imaging data, and other features.[6]

II. LITERATURE REVIEW

Schizophrenia is a complex behavioral disorder that affects millions of people worldwide. Early detection and accurate prognosis of schizophrenia are crucial for timely intervention and better outcomes. In recent years, the application of machine learning (ML) techniques to the prediction of schizophrenia has become important. This literature review explores the various studies and approaches employed in utilizing ML algorithms for predicting schizophrenia.

Historically, diagnosing schizophrenia relied heavily on clinical assessments and expert opinions. However, the limitations of traditional methods have led researchers to explore machine-learning techniques.

In the following study, data from 132 healthy people and 68 patients diagnosed with schizophrenia were collected.

EEG(Electroencephalogram) data were recorded using a 19-channel system and processed to

remove artifacts. The experimental setup involved a modified visual cued Go/NoGo paradigm with stimuli categories (animals, plants, and people with novel sounds) presented in pairs. Participants responded to specific stimulus pairs by pressing a button. Event-related potentials (ERPs) were analyzed for various conditions, including 'Plus1', 'Plus2', 'NoGo', 'Go', and 'P-H', each capturing different aspects of neural responses. Behavioral data, including response accuracy and reaction times, were recorded and analyzed. The data processing involved splitting ERP signals into overlapping time windows, calculating minimum, maximum, and average values, as well as global extremes and their corresponding window numbers. Feature selection methods, such as random feature addition, sequential feature selection, and truncated SVD, were applied to prevent overfitting.[1]

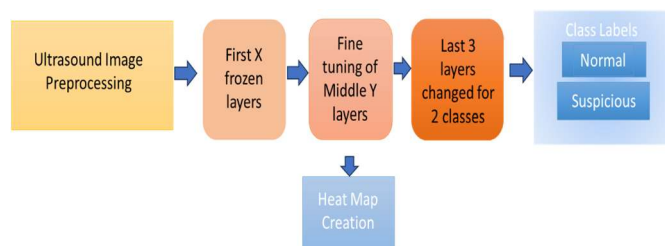
Researchers have employed various linguistic and non-verbal speech features to assess schizophrenia symptoms. Semantic features, including speech graphs capturing word co-occurrence patterns and topic modeling, were utilized to distinguish schizophrenia from other conditions. Syntactic analysis, incorporating part-of-speech tagging and sentence structure consideration, provided insights into patients' speech syntax abnormalities. Additionally, acoustic analysis, focusing on speech speed, pitch changes, and pause duration, proved valuable in identifying schizophrenia patients. Non-verbal cues, such as mutual silence and response time, were also effective markers. Deep learning techniques, including Word2Vec and CNN (Convolutional Neural Networks), were employed for semantic feature extraction. These diverse approaches highlight the multifaceted methods used in automatic assessments of schizophrenia symptoms.[2]

The study conducted a comprehensive review of Machine Learning (ML) applications in schizophrenia and bipolar disorder (BD). The research highlighted the rapid growth in ML applications in these domains, primarily focusing on MRI and fMRI data. The most widely utilized algorithms today are SVM (Support Vector Machine) and RF (Random Forest), because of their capacity to handle big data sets and identify non-linear relationships. A technique for supervised machine learning called SVM is utilized for applications involving regression or classification. To classify the data points into different classes, SVM searches an N-dimensional space for a hyperplane, where N is the number of features. SVM model achieved a precision of 100 percent.[5] Regression and classification tasks are both handled by the ensemble learning technique known as RF. To get the class mode (classification) or average prediction (regression) of each tree, it builds several decision trees during training. When large, labelled datasets are available, RF is a good option for forecasting schizophrenia outcomes due to its effectiveness in managing complex data and producing results that are easy to understand.[6]

III. PROPOSED METHODOLOGY

Creating a CNN model for predicting schizophrenia involves several steps: In the first step of the ultrasound schizophrenic brain image dataset preparation, the Canny edge detection algorithm was applied to the original dataset, creating a new edge-based image dataset. Subsequently, a Convolutional Neural Network (CNN) was directly applied to both the original and edge-based datasets to calculate their respective time and space complexities. The performance results were compared between the two

approaches. The edge-based dataset with Canny



preprocessing potentially improved CNN's efficiency by reducing computational load and enhancing feature extraction. However, the trade-off between accuracy and efficiency needs to be carefully considered, as the edge-based approach may sacrifice some level of detail in favour of speed.

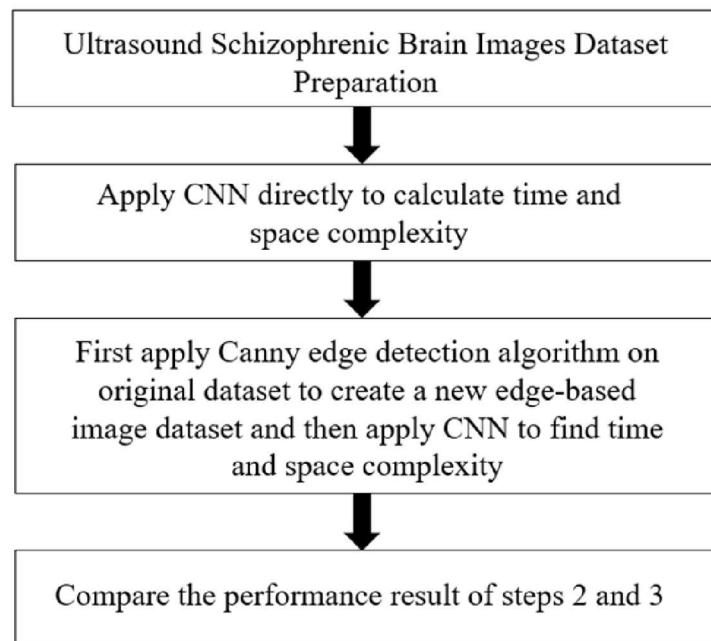


Fig 1: Procedure for CNN model

Fig 2: CNN System Architecture

IV. TECHNIQUE INVOLVED IN PROPOSED METHODOLOGY

A. Data Preprocessing

Here are some preprocessing steps:

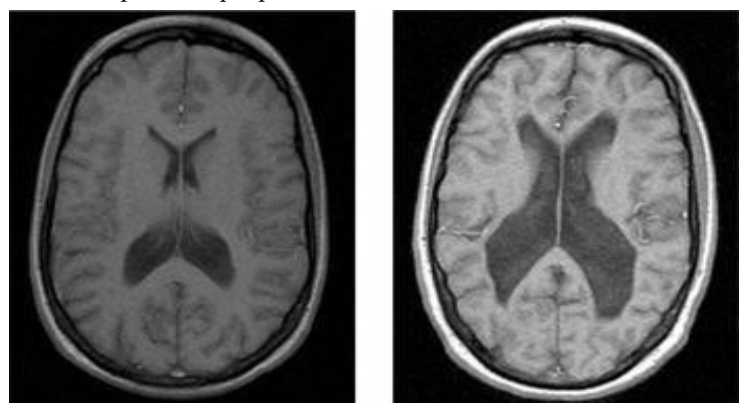
- 1) *Collection of Data*: Gather a dataset of schizophrenia-related data, which includes medical images, and other relevant information.
- 2) *Cleaning of Data*: Remove irrelevant data.
- 3) *Data Augmentation*: Increase the size of your dataset by applying data augmentation techniques to images, such as rotation, flipping, or resizing.
- 4) *Normalization*: Normalize the input data to have zero mean and unit variance. For images, this typically involves scaling pixel values to a range like $[0, 1]$ or $[-1, 1]$.
- 5) *Image Preprocessing*: If working with medical images, you may need to perform additional preprocessing like resizing images to a consistent size, cropping, zooming, etc.
- 6) *Label Encoding*: If your labels are in text or categorical form, encode them into numerical values. For binary classification (schizophrenia vs. non-schizophrenia), you can use 0 and 1.
- 7) *Splitting the Data*: Split your dataset into training, validation and test set for efficient model evaluation and performance. A general split is 70-15-15% or 80-10-10% for training, validation and testing.
- 8) *Data Balancing*: Ensure that your dataset is balanced, especially if you have imbalanced classes (i.e., more samples of one class than the other). Techniques like oversampling, under sampling, or using class weights can be helpful.
- 9) *Data Generators*: Use data generators in your deep learning framework to efficiently load and preprocess data in batches, especially if you have a large dataset that can't fit into memory.
- 10) *Data Loading and Batching*: Implement an efficient data loader that loads and batches your pre-processed data. Libraries like TensorFlow and PyTorch provide tools for this purpose.
- 11) *Pretrained Models*: Consider using pre-trained CNN models (e.g., VGG, ResNet, or Inception) as a starting point. You can fine-tune these models on the basis of your schizophrenic-patient dataset, saving both training time and potentially improving performance.
- 12) *Data Privacy*: Ensure that you handle sensitive medical data with the utmost care, following privacy

regulations and best practices for data security and anonymization.

B. CNN Model

The primary applications of convolutional neural networks (CNNs) in computer vision are object identification, image segmentation, and image classification. Figure 1 and 2 gives us an overall idea of CNN operation. Although CNNs can be applied to a variety of medical image analysis tasks, including tasks related to psychiatry, they are not commonly used as a stand-alone algorithm for schizophrenia monitoring or schizophrenia-related disease diagnosis. Instead, they can be a component of larger systems or tools used in the field of psychiatry. Here are some ways in which CNNs can be applied in the context of schizophrenia:

- 1) *Ultrasound Image Analysis*: CNNs can be used to analyse ultrasound images of schizophrenic brains and classify regions associated with schizophrenia. This can aid in early diagnosis and understanding of the affected brain regions.
- 2) *Brain Segmentation & Monitoring*: CNNs can be employed for segmenting brain regions in MRI scans. Identifying specific regions affected by schizophrenia is crucial for studying the disorder's progression and effects on different parts of the brain.
- 3) *Predictive Analysis*: Utilize CNNs for predicting schizophrenia based on brain imaging data. Once trained, these models can assist in diagnosing new patients or individuals at risk by analysing their brain images. For example, Figure 3 shows T2-weighted MRI scan showing ventriculomegaly in discordant monozygotic twins. Twins without schizophrenia (right) and with schizophrenia (left) are contrasted, while Figure 4 depicts differences in brains of schizophrenic and non-schizophrenic people shown under PET scan. It's



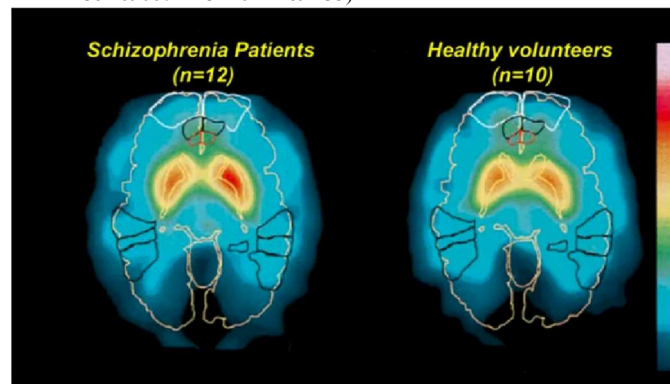
important to note that the application of CNNs in psychiatry-related tasks typically involves collaboration between medical professionals and machine learning experts. Moreover, the use of CNNs in medical contexts must adhere to strict ethical and regulatory standards, as patient privacy and safety are of utmost importance. Overall, while CNNs are a valuable tool in medical image analysis, they are just one component of a larger system used in neuro-mental monitoring and healthcare. Schizophrenic patients should always seek care from qualified healthcare professionals for proper monitoring and diagnosis of schizophrenia-related mental health conditions.

Figure 3: The left twin is healthy, but the right twin has a schizophrenic brain.

prediction of schizophrenia as a valuable tool in modern psychiatry. By harnessing the power of CNN algorithms for the analysis of ultrasonic brain images, we can improve the accuracy of neuro-mental condition diagnosis and provide patients with tailored guidance for better treatment. The application of machine learning in predicting schizophrenia marks an important moment in the history of mental healthcare. Its ability to facilitate early diagnosis, personalized treatment plans, and advance scientific research is transforming how we look up to and address this complex disorder.

The results obtained in our study are based on the use of the Convolutional Neural Network (CNN) algorithm to predict schizophrenia by analysing ultrasound images. The experiments were conducted on a dataset comprising a diverse range of ultrasound images obtained from schizophrenic individuals. One of the primary objectives of our research was to predict schizophrenia based on the detected brain conditions. This personalization led

(<https://www.cambridge.org/core/journals/advances-in-psychiatric-treatment/article/neuroimaging-in-schizophrenia-what-does-it-tell-the-clinician/B58F2DC2B35B69067F223D2924E89A33>)



Our paper studies the potential of ML in the

Figure 4: PET scan displaying differences in brain activity in people with and without schizophrenia. Science History Images / Alamy Stock (Photo: <https://www.healthline.com/health/schizophrenia/schizophrenia-brain-scan>)

V. CONCLUSION

to enhanced patient satisfaction and potentially improved prediction and timely treatment. The CNN model demonstrated promising results in the classification of mental(schizophrenic) conditions.

ACKNOWLEDGMENT

We are deeply grateful to our mentor for their important direction and assistance in completing this research paper on schizophrenia prediction using Machine Learning approaches.

REFERENCES

- [1] Ali Afraz, Kambiz Bahaadinbeigy, Mahdiah Montazeri, Mitra Montazeri and Mohadeseh Montazeri. "Application of machine learning methods in predicting schizophrenia and bipolar disorders: a systematic review". Health Sci Rep. 2022; 6:e962. doi:10.1002/hsr2.962 Application of machine learning methods in predicting schizophrenia and bipolar disorders: A systematic review - Montazeri - 2023 - Health Science Reports - Wiley Online Library
- [2] Andreas Müller, Juri Kropotov, Marina Pronina, Mikhail Lipkovich, Nadezhda, Shanarova and Valery Ponomarev. 2023. "Application of Machine Learning to Diagnostics of Schizophrenia Patients Based on Event-Related Potentials" Diagnostics 13, no. 3:

509. <https://doi.org/10.3390/diagnostics13030509><https://www.mdpi.com/2075-4418/13/3/509>

- [3] A. Mathur, R. K. Dwivedi and R. Rastogi, "*A Survey of Machine Learning Based Approaches for Neurological Disorder Predictions*," 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, India, 2022, pp. 586-590, doi: 10.1109/SMART55829.2022.10046944.<https://ieeexplore.ieee.org/document/10046944>
- [4] A. V. Nimkar and D. R. Kubal, "*Optimization of Schizophrenia Diagnosis Prediction using Machine Learning Techniques*," 2018 4th International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, Malaysia, 2018, pp. 1-6, doi: 10.1109/ICCOINS.2018.8510599. <https://ieeexplore.ieee.org/document/8510599>
- [5] R. Sharma, S. Tripathi and K. K. Sekhon, "*Detection of Schizophrenia using Machine Learning*," 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2023, pp. 505-512, doi:10.1109/ICACITE57410.2023.10183069. <https://ieeexplore.ieee.org/document/10183069>
- [6] Y. -J. Huang et al., "*Assessing Schizophrenia Patients Through Linguistic and Acoustic Features Using Deep Learning Techniques*," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, pp. 947-956, 2022, doi: 10.1109/TNSRE.2022.3163777. Assessing Schizophrenia Patients Through Linguistic and Acoustic Features Using Deep Learning Techniques | IEEE Journals & Magazine | IEEE Xplore



This is an open source journal