[1]Venkata Nagaraju Thatha

[2]Sikhi Yalavarthi

[3]Dr. P. A. Abdul Saleem

[4]P. Ramani

[5]Kotha Chandana

[6]Kasturi Sai Sandeep

[7]Dr. Maithili K

Advanced Recognizing and Categorizing System for Alzheimer's and Parkinson's Syndromes using CNN Algorithm and Clinical Data

Abstract: - Parkinson's Disease (PD) and Alzheimer's Disease (AD) are advancing neurodegenerative diseases that impact a large number of individuals globally. This study provides a novel method to identify and categorize individuals in the early stages of PD and AD by analyzing online handwriting structures, leveraging the current progress in Machine Learning (ML) in the medical domain. The research explores multiple data augmentation strategies to address the limited training information available for classifying neuro-degenerative disorders based on behavioral information. The study used a Convolutional Neural Network (CNN) classification as a foundation model by identifying essential radiomics characteristics from the T1-weighted Magnetic Resonance Imaging (MRI) data. The research explores data augmentation strategies designed explicitly for time-series information, such as online handwriting patterns, instead of standard methods used for hardware-based detection. This approach, with more validation, might assist in distinguishing between difficult instances of AD and PD that have a comparable mild motor and non-motor symptoms.

Keywords: Neurodegenerative diseases, Convolutional Neural Network (CNN), Alzheimer's Disease (AD), Parkinson's Disease (PD), Structural Magnetic Resonance Imaging (MRI), Deep Learning (DL) in neuroimaging, Data augmentation strategies, Online handwriting patterns, Machine Learning (ML) in medical domain, Classification model for AD and PD.

1.Introduction alzheimer's and parkinson's disease

[1] 1Associate Professor, Department of Information Technology, MLR Institute of Technology,

Hyderabad.

Email ID: nagaraju.thatha@gmail.com

2Software Engineer II Lexisnexis risk solutions

Email ID: ysikhi@gmail.com

3Professor, Department of Computer Science and Engineering (Data Science), CVR College of Engineering, Ibrahimpatanam, R R Dist.

Email ID: drsaleemprincipal@gmail.com

4Assistant Professor (SG), Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai-600088, India.

Email ID: ramanip@srmist.edu.in

5Assistant Professor, R.V.R & J.C College of Engineering, Chowdavaram, Guntur-522019.

Email ID: Chandanakotha@gmail.com

6Assistant Professor, K L University, Vijayawada.

Email ID: kssandeep12@gmail.com

7Associate Professor, Department of CSE (AI & ML), KG Reddy College of Engineering and Technology, Hyderabad, Telangana, INDIA-500086.

Email ID: drkmaithili@gmail.com

Alzheimer's Disease (AD) [1] and Parkinson's Disease (PD) [2] are the predominant neurodegenerative disorders globally. According to the World Health Organization (WHO), 50 million

individuals globally have dementia, with AD being the probable cause in 65-75% of these instances. A survey found that PD impacted 8.2 million individuals worldwide. Structural Magnetic Resonance Imaging (MRI) is used to identify biomarkers indicative of PD and AD in people [3]. Dimensional T1-weighted MRI brain is a common and readily available imaging technique that does not expose the patient to ionizing radiation [4].

A multi-center research was done by the PD task force of the Enhancing Brain Imaging Genomics via Meta-Analysis (ENIGMA) collaboration [5]. The study included 2000 PD patients and 1180 normal controls, utilizing MRI images from 19 worldwide locations. The scientists examined the cortical thickness, local dimension, and percentages of subcortical brain areas, uncovering a widespread pattern of atrophy that worsens as the illness develops. Research on AD has been conducted based on extensive cohort studies like the AD Neuroimaging Initiative (ADNI), which has released several research showing significant alterations in brain measurements in the condition [6]. Although substantial, case-control studies merely highlight unique characteristics across groups and do not provide a method to categorize illness status in new people. More than 80% of PD patients have cognitive impairments, which are often seen in the early stages of the disease and in newly diagnosed individuals. Thus, distinguishing between early AD and early PD is a crucial clinical concern.

Deep Learning (DL) and Machine Learning (ML) algorithms have been created for various neuroimaging tasks, such as picture reconstruction, augmentation, area segmentation, illness detection, and subtyping. When building models for these tasks, it is essential to consider data availability, the scale and effectiveness of the neural network construction, and how it performs on other datasets apart from those used for training. Several research has shown the potential of using Convolutional Neural Networks (CNNs) on brain MRI images for regression analysis, including predicting brain age, a frequently employed benchmarking test to estimate a person's age based on their brain scans [7].

This paper proposes the Alzheimer's and Parkinson's Syndromes Classification Model (APSCM), a CNN that extracts features from MRI scans and classifies subject-level PD and AD. The critical contribution of this work includes:

- A CNN architecture for classification of PD and AD versus healthy control subjects using MRI.
- To test the proposed model independently, the research evaluates the models on multiple datasets, including external datasets from different centers.

The rest of the paper is listed as follows: section 2 indicates the background and literature survey of the PD and AD. The proposed Alzheimer's and Parkinson's Syndromes Classification Model is designed and discussed in section 3. Section 4 shows the simulation analysis and its findings. Section 5 discusses the conclusion and future research area about PD and AD.

2.Background and related works

Several research studies have made remarkable advances in the field of dementia diagnosis. Many focus on diagnosing a specific kind of dementia via computer-aided diagnostics. They used many forms of clinical data, including history, physical examinations, cognitive tests, laboratory research, and imaging. Instead of visualizing the complete range of medical tests, it identifies problems as normal or pathological. The most effective method for classifying various kinds of memory using one or a combination of techniques is imaging in medicine. The suggested research aims to identify several types of dementia using a computer program. It is challenging to discern between various kinds of dementia due to the similarity of their symptoms. The most effective method for identifying diseases is studying the changes in the human brain under different dementia conditions.

Various ML approaches used in the same field are examined and elucidated below: Fouladi et al. proposed a neural network-based tool for diagnosing and monitoring AD advancement [8]. The individual gathered ADNI data from 1700 patients and trained the data using the "All-Pairs" approach. When dealing with missing data, preprocessing the data to highlight missing items is a crucial benefit in this study. Belyaev et al. used model-

based ML approaches for diagnosing and classifying PD [9]. They utilized clinical, socioeconomic, and neuroimaging databases in their research. Li et al. suggested a study to enhance dementia screening tests via Artificial Intelligence (AI) [10]. The obtained data comprises history data, physical examination, cognitive testing, laboratory research, and radiology. They categorized the dementia state using a Naïve Bayes classification. El-Sappagh et al. researched identifying dementia using ML [11]. A two-layer model was created for the early detection of AD using the brief mental state assessment. They compared the suggested method with various ML techniques.

Other studies have suggested techniques for categorizing neurodegenerative disorders. Sethi et al. developed a 2D model using the Inception-V4 framework for AD categorization [12]. Rezaee et al. presented a CNN model based on the AlexNet structure for PD categorization [13]. Previous techniques often make forecasts in a segmented fashion. Veetil et al. discuss data leakage concerns in previous neuroimaging studies, some of which included making projections based on slices of data [14]. Many 2D approaches used before only consider the spatial connections among some of the segments in the MRI region. Hu et al. demonstrate that CNNs perform better than other MLs in specific neuro-imaging categorization tasks [15].

3.Alzheimer's and parkinson's syndromes classification model

The process flow of the suggested system is shown in the architectural diagram in Figure 1.

3.1Pre-processing

The MRI dimensions underwent preprocessing to align the images with standard template pictures. The research conducted a skull-stripping brain separation utilizing the architecture. The research then obtained masks for gray and white material. The study performed a nonparametric brightness adjustment. The research then ran linear registrations to a UK Biobank with minimal distortion using 6 degrees of variation. The research resized the photos to 2 mm utilizing the Resampled Image instrument, resulting in pre-processed pictures measuring 92 x 110 x 93. The MRI clinical data were normalized to a range of 0 to 1 by min-max scaled for the CNN algorithm.

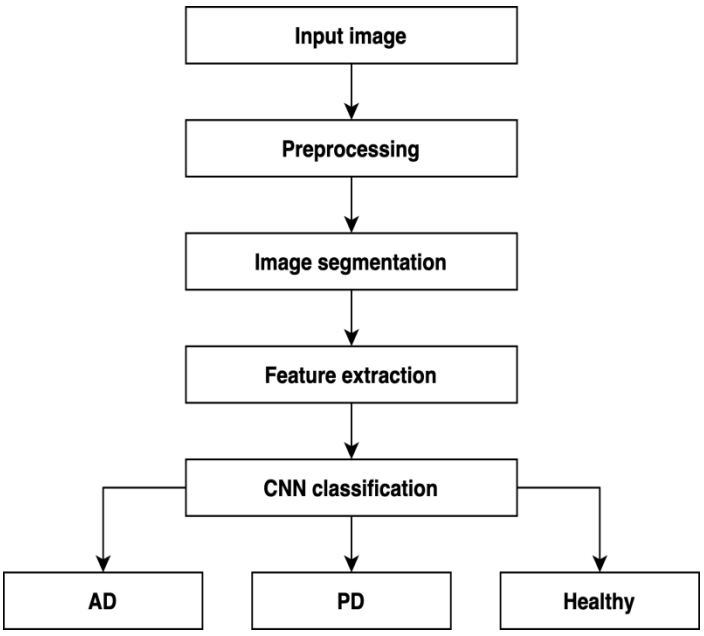

Fig. 1. Workflow of the proposed APSCM system

3.2. Clinical data preparation

The preparation method follows a series of stages to achieve consistency in the databases, which includes nonparametric intensity standardization and orthogonal cubic resampling to 2 mm. The research generated distinct data divisions to train, validate, and test the suggested model. The research matched both subsets of the UPenn database (PD) and OASIS database (AD) by selecting an equal number of instances and controls for testing independently since there were more PD patients in UPenn than normal controls and more controllers in OASIS than AD instances. Test data sets for both cohorts have been balanced by merging the smaller group with randomized, unique groups of identical proportions from the more prominent categories.

3.3. Segmentation

Picture segmentation is the process of dividing a digital picture into several parts. The segmentation approach used in the suggested study is an improved K-means. The K-means technique is employed in multiple fields, such as segmenting pictures, mining data, and bioinformatics. The standard K-means method is best suited for a small amount of data. The approach utilizes an altered version of the K-means method. The improved K-means clustering algorithm utilizes the Euclidean distance to calculate centroids. The local optimum solution is circumvented, decreasing the use of the cluster error criteria.

3.4. Feature extraction

The picture features are retrieved from the segmented layer utilizing the Gray-Level Co-Occurrence Matrix (GLCM), Speeded-Up Robust Features (SURF), and Gradient Magnitude Histogram (GHM) techniques and then saved in the database.

SURF is a trademarked method for detecting and describing local image features. It primarily uses picture categorization, object identification, and reconstruction. The SURF method has three components: identifying interest locations, defining the local neighborhood, and performing matching. The SURF technique is used on the segmentation input picture in the suggested work, and the resulting data sets are saved in the repository.

GLCM analyzes the texture of the fragmented image and provides information about the spatial arrangement of pixels. Several GLCM characteristics are accessible, including Autocorrelation, Contrasting, Entropy, Energies, Highest probability, and Average estimation, which are used in this study.

GHM quantifies the frequency of gradient directions in certain picture regions. The gradient estimates are computed from the spectrum and then saved in the database.

3.5.Classification

A CNN was traine and categorized using a repository including photos of AD, PD, and healthy individuals. The network used values obtained from the feature extraction layer. A neural network is a unique collection of inputs and output units, where each unit has its weighting and biases. Training a neural network classification involves modifying the weights to anticipate the proper class. CNN structure is seen in Figure 2.

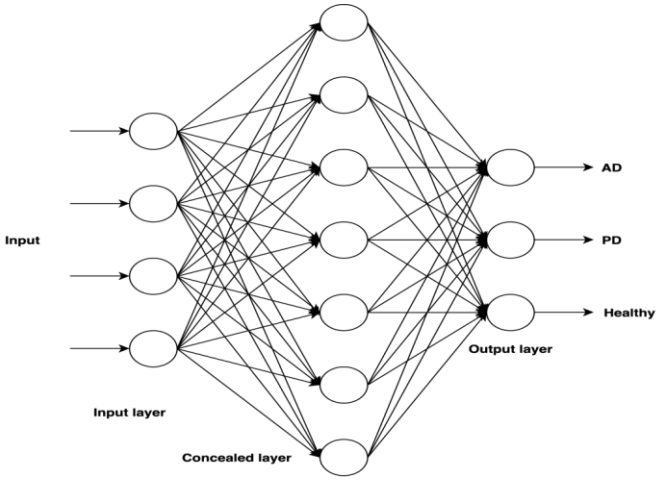

Fig. 2. Architectural design of the CNN model

Three kinds of parameters usually characterize a CNN:

1. The connectivity pattern links the various levels together.
2. The learning method involves updating the related weights as they progress from one layer to another.
3. The activation function extracts the output from the provided input sample.

This article utilizes a CNN that combines feedforward and backpropagation techniques. The input layer computes and transmits the output values to the concealed layer. Each hidden layer receives information from all input layers via individual weights and bias functions. The output layer receives these data to generate the output. Activation is employed to process and generate the output based on the inputs. Every input symbol is compared to an upper limit to determine whether it is more than or less than the value set for the threshold. The suggested APSCM divides the categorization procedure into training and evaluation. During the training phase, photos are picked for training, and their characteristics are retrieved while already knowing the label of each image. The information is then placed in a trained database. A picture is chosen from the test library during the test stage, and the illness classification is determined.

3.6. CNN training and hyperparameter optimization

The CNN's tiers were taught using the information for PD categorization and the ADNI information for AD categorization. The CNN's hyperparameters were fine-tuned using a random searching method. The primary hyperparameters that were adjusted were the optimization, learning percentage, rate of learning, planning, dropout percentage, biased initializer for imbalanced classes, early halting, number of training sessions, and batch length. The hyperparameters for classifying AD were 200 iterations, a dropout rate of 0.2, a batch length of 4, Adam optimization with a learning speed of $2e^{-4}$, and a weight decay of 1e-4. The categorization of PD was done using hyperparameters: trained for 200 iterations with early halting if there was no change in verification loss for 30 intervals, dropout percentage of 0.2, batch dimension of 4, Adam optimization with an acquiring rate of $2e^{-4}$, and a weight decay of $1e^{-4}$. The research set a bias in the result level to the logarithmic of the ratio between the amount of specimens in class 0 and the number of specimens in class 1. This was done to correct the discrepancy in the database. The algorithm that demonstrated the highest validation effectiveness was applied to evaluate the two categories on the test datasets.

3.7. Model testing

The research assessed the models using conventional metrics for performance, such as accuracy, precision, specificity, and sensitivity. The performance measures were computed by averaging all the balancing subgroups of the UPenn and OASIS separate testing sets.

4.Simulation outcomes

MRI pictures were obtained from the Laboratory of Neuroimaging (LONI) in DICOM format, including images related to AD, PD, and healthy brains [16]. Out of a database containing 1000 photos, 700 were allocated for learning and 300 for validation. A CNN processed both the training data and test data. The tests are implemented using Python, and a framework is constructed and trained using the Pytorch and Sklearn tools. The research uses the assets Google Colab provided to utilize a Graphical Processing Unit (GPU) and enhance training velocity.

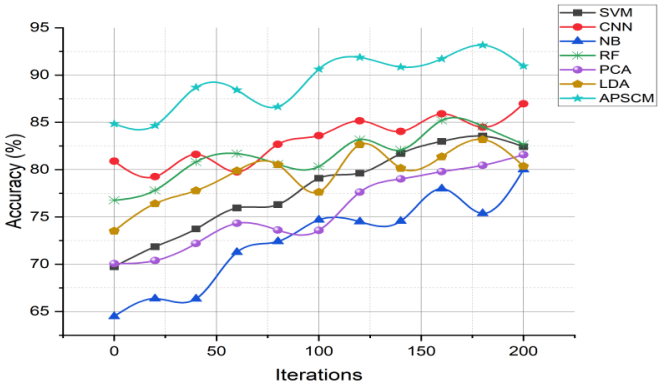

Fig. 3. Accuracy evaluation of MRI images

Figure 3 displays how accuracy changes throughout iterations from 0 to 200, with increments of 20, demonstrating the influence of iterative modifications on the performance of Support Vector Machine (SVM), CNN, Naïve Bayes (NB), Random Forest (RF), Principle Component Analysis (PCA), Linear Discriminant Analysis (LDA), and the proposed APSCM algorithms. Accuracy is calculated by dividing the number of accurately predicted cases by the total number of instances and expressing it as a percentage. By the 200th iteration, the average accuracy for each approach is SVM (82.42%), CNN (86.96%), NB (80.00%), RF (82.69%), PCA (81.58%), LDA (80.37%), and APSCM (90.96%). APSCM's excellent performance is due to its specialized processing of time-series data, showing promise for improved categorization of neurodegenerative diseases compared to conventional approaches.

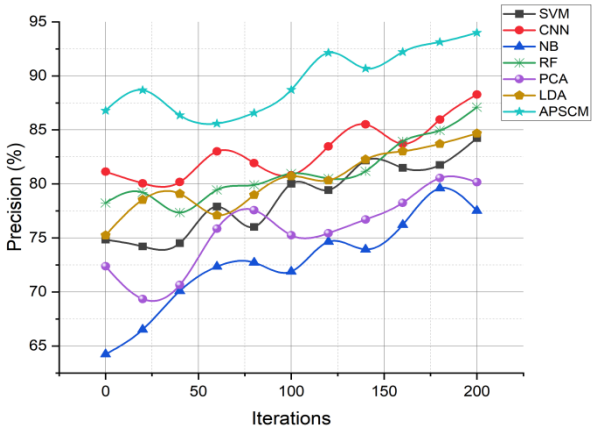

Fig. 4. Precision evaluation of MRI images

Figure 4 displays accuracy variations throughout iterations (0 to 200 with a step size of 20) for SVM, CNN, NB, RF, PCA, LDA, and APSCM algorithms. Precision is determined by dividing the number of actual positive cases by the total of virtual positive and false positive instances and is given as a percentage. The average accuracy values for each approach are SVM (84.25%), CNN (88.28%), NB (77.53%), RF (87.10%), PCA (80.16%), LDA (84.68%), and APSCM (94.00%). The APSCM technique regularly shows greater accuracy in formulating correct optimistic predictions, perhaps because of its specialized handling of time-series information,

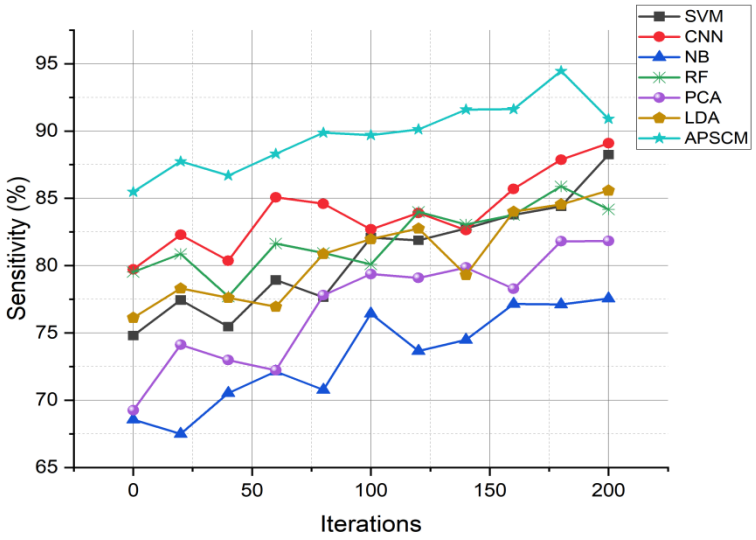

Fig. 5. Sensitivity evaluation of MRI images

Figure 5 displays the sensitivity trends for SVM, CNN, NB, RF, PCA, LDA, and APSCM algorithms across iterations ranging from 0 to 200 with a step size of 20. Sensitivity is measured as the ratio of true positive cases to the total of actual positive and false negative instances represented as a percentage. The average sensitivity for each technique at the last iteration (200) is as follows: SVM (88.25%), CNN (89.09%), NB (77.56%), RF (84.17%), PCA (81.83%), LDA (85.58%), and APSCM (90.91%). The suggested APSCM regularly shows increased sensitivity, highlighting its efficacy in precisely detecting positive cases, perhaps because of its specialized method of managing time-series data.

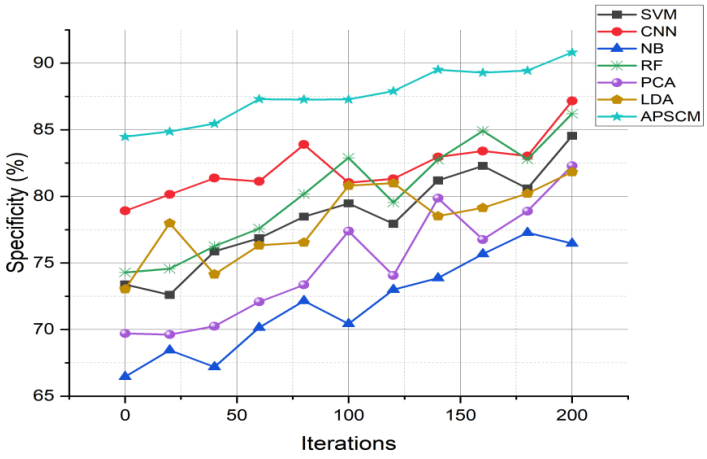

Fig. 6. Specificity evaluation of MRI images

Figure 6 displays the specificity patterns for SVM, CNN, NB, RF, PCA, LDA, and APSCM algorithms throughout iterations ranging from 0 to 200 with a step size of 20. Specificity is the proportion of actual negative cases to the total of true negative and false positive instances, represented as a percentage. The average specificity for each approach at the last iteration (200) is as follows: SVM (84.55%), CNN (87.16%), NB (76.47%), RF (86.21%), PCA (82.30%), LDA (81.83%), and APSCM (90.81%). The suggested APSCM shows superior specificity, indicating its effectiveness in reliably recognizing negative occurrences, likely because of its unique method of dealing with time-series information.

5.Conclusion and future scope

The studies and findings demonstrate the capability of the CNN model to categorize neurodegenerative disorders using MRI. The CNN scored well when trained on a partial dataset and surpassed the starting point random forest

classification. The research demonstrated variations in the training and testing performance of the suggested approach in tasks related to classifying PD and AD. The proposed pipeline might be improved by training it on more data and doing more validation. This could assist in screening patients for PD and AD non-invasively using structural MRI information. Classification algorithms with modest accuracy are valuable screening tools before invasive tests, which are utilized to define biological AD and diagnose PD. The research explores multiple ML and optimization methods, such as swarm optimizations and genetic algorithms, for analyzing AD and PD at different degrees in the future.

References

1. A.A. Tahami Monfared, M.J. Byrnes, L.A. White, Q. Zhang, "Alzheimer's disease: epidemiology and clinical progression. Neurol", Ther., vol.11, no.2, pp.553-569, 2022.
2. G.M. Lawrence, C.L. Holley, K. Schroder, "Parkinson's disease: connecting mitochondria to inflammasomes," *Trends Immunol*, vol.43, no.11, pp.877-885, 2022.
3. T.T. Pham, B. Whelan, B.M. Oborn, G.P. Delaney, S. Vinod, C. Brighi, and P. Keall, "Magnetic resonance imaging (MRI) guided proton therapy: A review of the clinical challenges, potential benefits, and pathway to implementation," *Radiother Oncol*, vol.**170**, pp.37-47, 2022.
4. V. Kaplanoğlu, H. Kaplanoğlu, A. Turan, and A. Dilli, "Evaluation of Arachnoid Granulations in Cranial Dural Sinuses with Contrast-Enhanced 3-Dimensional T1-Weighted Magnetic Resonance Imaging," *Eurasian j. med*, vol.**55**, no.2, pp.95-99, 2023.
5. P.M. Thompson, N. Jahanshad, L. Schmaal, J.A. Turner, A.M. Winkler, S.I. Thomopoulos, and P. Kochunov, "The enhancing NeuroImaging genetics through meta-analysis consortium: 10 Years of global collaborations in human brain mapping", *Hum. Brain Mapp*, vol.**43**, no.1, pp.15-22, 2022.
6. D.P. Veitch, M.W. Weiner, P.S. Aisen, L.A. Beckett, C. DeCarli, and R.C. Green, "Alzheimer's Disease Neuroimaging Initiative", Using the Alzheimer's Disease Neuroimaging Initiative to improve early detection, diagnosis, and treatment of Alzheimer's disease, *Alzheimer's & Dementia*, vol.**18, no.**4, pp.824-857, 2022.
7. A.S. Musallam, A.S. Sherif, and M.K. Hussein. "A new convolutional neural network architecture for automatic detection of brain tumors in magnetic resonance imaging images", *IEEE Access*, vol.**10**, pp.2775-2782, 2022.
8. S. Fouladi, A.A. Safaei, N.I. Arshad, M.J. Ebadi, and A. Ahmadian, "The use of artificial neural networks to diagnose Alzheimer's disease from brain images", *Multimed. Tools Appl*, vol.**81**, no.26, pp.37681-37721, 2022.
9. M. Belyaev, M. Murugappan, A. Velichko and D. Korzun, "Entropy-Based Machine Learning Model for Fast Diagnosis and Monitoring of Parkinson's Disease," *Sensors*, vol.**23**, no.20, pp.1-20, 2023.
10. R. Li, X. Wang, K. Lawler, S. Garg, Q. Bai, and J. Alty, "Applications of artificial intelligence to aid early detection of dementia: a scoping review on current capabilities and future directions", *J. Biomed. Inform*, vol.**127**, 2022.
11. Mohan, A., Prabha, G. and V., A. 2023. Multi Sensor System and Automatic Shutters for Bridge- An Approach. International Journal of Intelligent Systems and Applications in Engineering. 11, 4s (Feb. 2023), 278–281.
12. Prabha , G. , Mohan, A. , Kumar, R.D. and Velrajkumar, G. 2023. Computational Analogies of Polyvinyl Alcohol Fibres Processed Intellgent Systems with Ferrocement Slabs. International Journal of Intelligent Systems and Applications in Engineering. 11, 4s (Feb. 2023), 313–321.
13. Study On Structural Behaviour Of Ductile High-Performance Concrete Under Impact And Penetration Loads, Lavanayaprabha, S. Mohan, A. Velrajkumar, G., Mohammedharoonzubair, A. Journal of Environmental Protection and Ecology., 2022, 23(6), pp. 2380–2388.
14. Mohan, A., & K, S. . (2023). Computational Technologies in Geopolymer Concrete by Partial Replacement of C&D Waste. International Journal of Intelligent Systems and Applications in Engineering, 11(4s), 282–292.
15. Mohan, A., Dinesh Kumar, R. and J., S. 2023. Simulation for Modified Bitumen Incorporated with Crumb Rubber Waste for Flexible Pavement. International Journal of Intelligent Systems and Applications in Engineering. 11, 4s (Feb. 2023), 56–60.
16. R.Gopalakrishnan, Mohan, "Characterisation on Toughness Property of Self-Compacting Fibre Reinforced Concrete", Journal of Environmental Protection and Ecology 21, No 6, 2153–2163 (2020)
17. S. El-Sappagh, H. Saleh, F. Ali, E. Amer, and T. Abuhmed, "Two-stage deep learning model for Alzheimer's disease detection and prediction of the mild cognitive impairment time", *Neural Comput. Appl*, vol.**34, no.**17, pp.14487-14509, 2022.
18. M. Sethi, S. Ahuja, S. Rani, D. Koundal, A. Zaguia, and W. Enbeyle, "An Exploration: Alzheimer's disease classification based on convolutional neural network", *Biomed Res. Int*, **2022**.
19. K. Rezaee, S. Savarkar, X. Yu, and J. Zhang, "A hybrid deep transfer learning-based approach for Parkinson's disease classification in surface electromyography signals. Biomed," *Signal Process. Control*, vol.**71**, 2022.
20. I.K. Veetil, D.E. Chowdary, P.N. Chowdary, V. Sowmya, and E.A. Gopalakrishnan, "An analysis of data leakage and generalizability in MRI-based classification of Parkinson's Disease using Explainable 2D Convolutional Neural Networks," *Digit. Signal Process*, 2024.
21. M. Hu, X. Qian, S. Liu, A.J. Koh, K. Sim, X. Jiang, and J.H. Zhou. "Structural and diffusion MRI-based schizophrenia classification using 2D pre-trained and 3D naive Convolutional Neural Networks," *Schizophr. Res*, vol.**243**, pp.330-341, 2022.
22. https://adni.loni.usc.edu/data-samples/data-types/mri/