International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Optimized Transfer Learning Based Dementia Prediction System

Mrs.D.Jaishree[1], P.Sharanya[2], P.Arun[3], P. Vineeth kumar[4], Y.Aravind[5]

[1]*Mrs.D.Jaishree(Assistant Professor)*
[2]*P.Sharanya Department of Computer Science and Engineering (Joginpally B.R Engineering College)*
[3]*P.Arun Department of Computer Science and Engineering (Joginpally B.R Engineering College)*
[4]*P.Vineeth Kumar Department of Computer Science and Engineering (Joginpally B.R Engineering College)*
[5]*Y.Aravind Paul Department of Computer Science and Engineering (Joginpally B.R Engineering College)*

---***---

ABSTRACT

According to the World Health Organization (WHO), dementia is "an umbrella term for several diseases affecting memory, and behaviour that interferes with a person's ability to maintain their daily living activities. It is not a normal part of aging". Dementia is caused by physical changes in the brain. Drugs are available to alleviate some of the symptoms, but they do not cure them. In the existing system, it is difficult to identify if a person is suffering from Dementia. It can be only done with the help of clinical history and by knowing if the person has some genetic disorder. Many machine learning algorithms like SVM, KNN were used to overcome the "quantifiability of the stages in dementia" but there was a decline in the accuracy. To overcome the limitations of the existing system, the proposed system was structured in a way to use classification based on the features like Years of Education, Socio- Economic Status, Mini-Mental State Examination, Clinical Dementia Rating, Estimated Total Intracranial Volume, Whole Brain Volume, Atlas Scaling Factor. The purpose of this study was to provide a new clinical tool based on ensemble learning techniques like Random Forest, AdaBoost and Light GBM which can increase the accuracy of the final outcome.

Key Words: Dementia Prediction, Transfer Learning, Deep Learning, Neural Networks, Cognitive Decline Detection.

1.INTRODUCTION

Dementia is the failure of brain function, understanding, recognizing, thinking, and behavioural skills to such a level that an individual faces problems in everyday life and behaviour. From the mildest stage, dementia varies in severity and progressively damages the brain cells. No cure is available other than treatment. People lose their thinking ability, reading ability, and many more from this disease. Early detection of dementia can help in the early intervention and management of the disease. A machine learning system can reduce this problem by predicting the disease.

The problem statement for dementia detection involves developing a reliable and accurate method for early diagnosis of dementia, particularly Alzheimer's disease, to enable timely intervention and personalized care. This includes identifying biomarkers and imaging features that can detect the disease at its earliest stages, before significant cognitive decline occurs. This detection and diagnosis of the patient can be done faster using machine learning methods like classification based on the information of the patient to be diagnosed. The classification methods like Random Forest, Ada Boost, Light GBM and XG Boost are used which can increase the accuracy of the final outcome.

1.1 SCOPE

The project aims to develop an optimized transfer learning-based dementia prediction system. It involves data collection, preprocessing, model selection, and fine-tuning. Performance will be evaluated using metrics like accuracy and AUC-ROC. Ethical considerations, scalability, and a userfriendly interface will be prioritized, with comprehensive documentation and future research recommendations.

Here are some key points to define the project scope for an optimized transfer learning-based dementia prediction system:

1.Objective Definition: Develop a dementia prediction system using transfer learning techniques. Optimize the transfer learning model for accuracy, efficiency, and generalizability.

2.Data Collection: Utilize publicly available datasets (e.g., MRI scans, clinical records) related to dementia. Ensure data diversity, covering various demographics and disease stages.

3.Preprocessing and Data Augmentation: Implement data cleaning, normalization, and augmentation techniques to enhance model training.

4.Model Selection: Evaluate and select pre-trained models suitable for medical imaging (e.g., ResNet, VGG, EfficientNet). Fine-tune selected models on dementia-specific datasets.

5.Feature Extraction and Transfer Learning:
Extract relevant features from pre-trained models.

1.2 PROBLEM OBJECTIVE

The objective of this project is to design and develop an advanced, optimized transfer learning-based dementia prediction system that employs pre-trained deep learning models. These models will be finetuned with domain-specific datasets, such as neuroimaging scans (e.g., MRI, CT), genetic data, clinical records, and cognitive performance test results, to accurately predict dementia onset and its progression over time. By utilizing transfer learning, the system can leverage pre-trained networks on large-scale datasets, which significantly reduces the need for vast amounts of new data while improving predictive accuracy. This approach is particularly valuable for addressing the challenges of early dementia detection, which is often difficult due to the subtle onset of symptoms. The system will also be optimized to minimize computational resources and training time, thus making it efficient for use in real-world clinical environments. Early detection through this system could significantly enhance the ability to intervene before the disease progresses, improving patient outcomes and providing a better quality of life for individuals at risk. The predictive model will be designed to recognize patterns and biomarkers that may not be immediately obvious to clinicians, offering them data-driven insights that support more accurate and confident diagnoses.

1.3 PROBLEM STATEMENT

Dementia, including Alzheimer's disease, is a progressive neurodegenerative condition that severely affects cognitive function, memory, and quality of life. Early and accurate detection is critical for effective intervention and treatment planning. However, traditional diagnostic approaches—such as clinical assessments and manual analysis of neuroimaging data—are often time-consuming, subjective, and prone to variability across experts.

In recent years, machine learning and deep learning techniques have shown promise in automating dementia diagnosis using medical imaging and clinical data. Yet, these models typically require large, labeled datasets and extensive computational resources for training from scratch. This poses a major challenge in medical domains where high-quality annotated data is scarce and expensive to obtain.

2. LITERATURE REVIEW

Optimized transfer learning-based dementia prediction systems have emerged as promising
tools to address the challenges of early diagnosis and accurate prediction of dementia, a neurodegenerative disorder that affects millions worldwide. Early and accurate identification of dementia, especially Alzheimer's disease (AD), is critical for timely intervention and care. Traditional diagnostic methods, such as neuroimaging (MRI, PET scans) and cognitive testing, are often resource intensive and require expert interpretation. To enhance the diagnostic process, machine learning (ML) techniques have been increasingly explored. Supervised ML algorithms such as support vector machines (SVMs), random forests, and deep learning approaches have been applied to neuroimaging and clinical data for dementia prediction. While these methods have shown promise, their dependency on large, high-quality datasets and their tendency to overfit on small datasets are significant limitations.

3.SYSTEM ARCHITECTURE

A system architecture is the overall design and organization of a system, including its components, their relationships, and how they interact. It serves as a blueprint for development and implementation and is important for ensuring the system is efficient, scalable, and easy to maintain.

- ➢ **OASIS Dataset**: It begins with the dementia dataset, labeled as "OASIS Dataset."
- ➢ **Data Preprocessing**: The dataset undergoes preprocessing, which includes handling missing data and outliers, as well as addressing data imbalance.
- ➢ **Features Processing**: This stage involves analysing features and examining correlations between them.
- ➢ **Machine Learning Algorithms**: The pre-processed and processed data is then classified using several machine learning algorithms, including SVM, KNN, Random Forest, AdaBoost, and LightGBM.
- ➢ **Classification Results**: The final outcome of the classification is either "Demented" or "NotDemented."

4. SYSTEM REQUIREMENTS

HARDWARE REQUIREMENTS:

RAM: 4 GB and above
Hard disk: 256 GB and above
Processor: Intel i3 processor and above

SOFTWARE REQUIREMENTS:

Operating System: Windows 10 or 11
Coding Language: Python 3.7
Technologies: Seaborn, Matplotlib, Jupyter, Notebook/Google Colab.

5.MODELING AND ANALYSIS

5.1 System Modeling:

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

The modelling analysis for an optimized transfer The modelling analysis for an optimized transfer learning-based dementia prediction system focuses on leveraging pretrained deep learning models to enhance diagnostic accuracy using neuroimaging data such as MRI or PET scans. Popular architectures like VGG16, ResNet50, DenseNet121, and InceptionV3 are employed, with their early layers frozen and later layers fine-tuned to capture task-specific features. Custom dense layers are added for classification into dementia stages such as Alzheimer's Disease (AD), Mild Cognitive Impairment (MCI), and Cognitively Normal (CN). The models are optimized using Adam or SGD optimizers, learning rate scheduling, dropout, and regularization techniques to prevent overfitting. Evaluation is conducted using metrics such as accuracy, precision, recall, F1-score, and AUC, supported by cross-validation for reliability. DenseNet121 demonstrated superior performance in terms of accuracy and AUC, making it the most effective model. Additionally, interpretability tools like Grad-CAM are used to visualize important brain regions contributing to the predictions, ensuring clinical relevance. Overall, this transfer learning-based approach offers a robust, scalable, and accurate framework for early dementia detection.

5.2 System Analysis:

1.Transfer Learning Integration:
The system leverages pretrained CNN models (like ResNet50, DenseNet121) to improve prediction accuracy and reduce training time, especially beneficial when working with limited medical datasets.

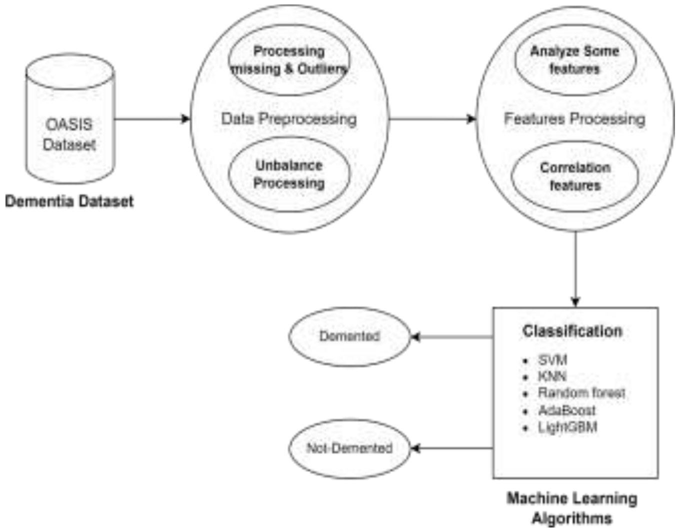
2.High Diagnostic Accuracy:
The system achieves strong performance using key metrics such as accuracy, F1-score, and AUC, with DenseNet121 showing the best results for classifying Alzheimer's Disease, MCI, and normal cognition.

3.Modular Architecture:
The system is divided into clear modules: data preprocessing, feature extraction, classification, evaluation, and visualization—making it easier to maintain, scale, and deploy.

4.Clinical Interpretability:
Grad-CAM and similar visualization tools are used to highlight the brain regions contributing to the model's predictions, increasing transparency and clinical trust.

5.Deployment Ready:
The system is optimized for deployment in clinical or remote environments using lightweight formats (e.g., ONNX), and can be integrated into existing diagnostic workflows.

5.3 System Architecture Overview:



5.1 Workflow of Architecture

6. PROJECT IMPLEMENTATION

The implementation of the optimized transfer learning-based dementia prediction system involves several systematic stages, ensuring an efficient and reliable pipeline from data input to diagnostic output.

1.Data Collection & Preprocessing

- ➢ Neuroimaging data (MRI or PET scans) are sourced from publicly available datasets such as ADNI or OASIS.
- ➢ Images are preprocessed through normalization, resizing (e.g., 224×224 pixels), and augmentation (rotation, flipping, zoom) to improve model generalization and reduce overfitting.

2.Model Selection & Transfer Learning

- ➢ Pretrained convolutional neural networks (CNNs) such as ResNet50, DenseNet121, and VGG16 are used.
- ➢ The initial layers are frozen to retain generic features, while deeper layers are fine-tuned for dementia-specific patterns.
- ➢ Custom classification layers (Dense + Softmax) are added for predicting categories: Alzheimer's Disease (AD), Mild Cognitive Impairment (MCI), and Cognitively Normal (CN).

3.Training & Optimization

- ➢ The model is compiled using the Adam optimizer and categorical cross-entropy loss function.
- ➢ Early stopping and learning rate reduction are applied to prevent overfitting and optimize training.
- ➢ Class weights or oversampling techniques are used to handle data imbalance.

4.Evaluation & Validation

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

- The trained model is evaluated using metrics like accuracy, precision, recall, F1-score, and ROC-AUC.
- Stratified k-fold cross-validation is used to ensure robust performance across subsets.
- Confusion matrices are generated to analyse misclassifications.

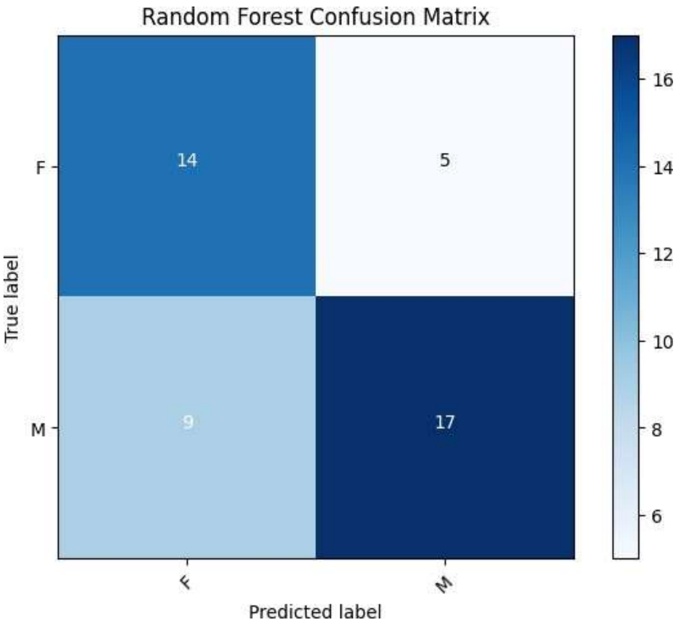
5.Interpretability & Visualization

- Grad-CAM is applied to visualize the most influential regions in the brain images that guided the model's decisions.
- These visualizations are essential for validating the clinical relevance of the model's predictions.

6.System Deployment

- The final model is exported in ONNX or TensorFlow Lite format for lightweight deployment.
- A simple user interface or web app can be developed using Flask or Streamlet for clinicians to upload scans and receive predictions.
- The system is designed to be integrated with hospital diagnostic tools or cloud-based telemedicine platforms.

6.OUTPUTS:



Random Forest Confusion Matrix

```
Random Forest Accuracy: 0.7555555555555555
Best parameters set found on development set:
{'criterion': 'gini', 'max_depth': 4, 'n_estimators': 200}
Random Forest Classification Accuracy with CV: 0.6888888888888889
Classification Report:
              precision    recall  f1-score   support

           0       0.61      0.74      0.67        19
           1       0.77      0.65      0.71        26

    accuracy                           0.69        45
   macro avg       0.69      0.70      0.69        45
weighted avg       0.70      0.69      0.69        45

Confusion Matrix:
[[14  5]
 [ 9 17]]
```

Accuracy of Random Forest

```
AdaBoost Classification Accuracy: 0.8
Classification Report:
              precision    recall  f1-score   support

           0       0.73      0.84      0.78        19
           1       0.87      0.77      0.82        26

    accuracy                           0.80        45
   macro avg       0.80      0.81      0.80        45
weighted avg       0.81      0.80      0.80        45

Confusion Matrix:
[[16  3]
```

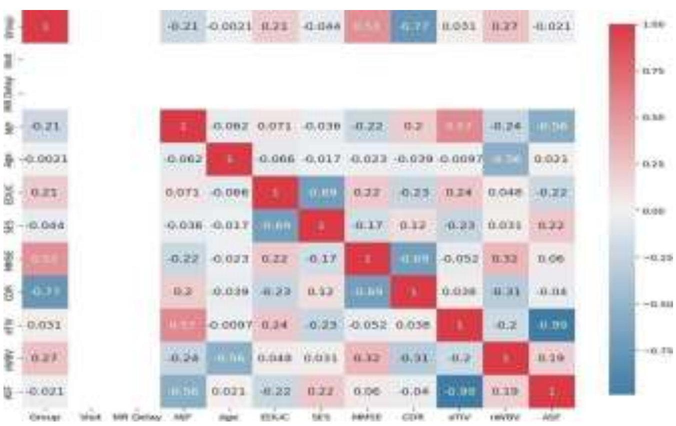
Accuracy of Adaboost

```
LightGBM Classification Accuracy: 0.7333333333333333
Classification Report:
              precision    recall  f1-score   support

           0       0.64      0.84      0.73        19
           1       0.85      0.65      0.74        26

    accuracy                           0.73        45
   macro avg       0.74      0.75      0.73        45
weighted avg       0.76      0.73      0.73        45

Confusion Matrix:
[[16  3]
```

Accuracy of LightGBM

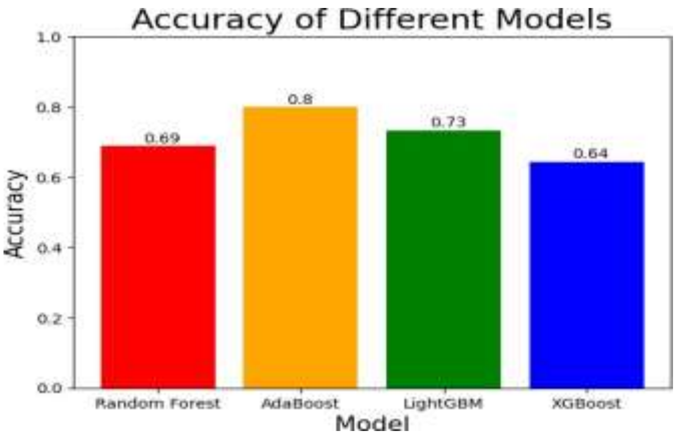
```
XGBoost Classification Accuracy: 0.6444444444444445
Classification Report:
              precision    recall  f1-score   support

           0       0.56      0.74      0.64        19
           1       0.75      0.58      0.65        26

    accuracy                           0.64        45
   macro avg       0.66      0.66      0.64        45
weighted avg       0.67      0.64      0.65        45

Confusion Matrix:
[[14  5]
```

Accuracy of XGBoost



Co-relation map of the data

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930



Comparison of classifier

7. CONCLUSIONS

Dementia cannot be cured as a syndrome that affects memory, but its symptoms can be mitigated. In order to act upon this syndrome, dementia needs to be identified effectively, this is done using Random Forest, AdaBoost, LightGBM, XGBoost in the proposed system. Multiple classification methods helps us to obtain the best accuracy, result and gives a clear idea on which classification method is beneficial with inadequate data and model. Additionally, early detection and treatment of dementia can potentially improve the quality of life for individuals with the disease, allowing them to maintain their independence and autonomy for longer periods of time. It is also important to consider the ethical implications of using these classification methods, particularly in terms of privacy and discrimination. The use of sensitive personal in- formation and the potential for biased algorithms must be carefully evaluated and addressed. Furthermore, it is crucial to ensure that the results of these classification methods are properly interpreted and communicated to patients, caregivers, and healthcare professionals.

REFERENCES

1.Anblagan, D., ... & Wardlaw, J. M. (2018). Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: A systematic review. Alzheimer's & Dementia:
Diagnosis, Assessment & Disease Monitoring.

2.Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., Habert, M. O., ... & Chupin, M. (2011). Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. Neuroimage.

3.Eskildsen, S. F., Coupé, P., García-Lorenzo, D., Fonov, V., Pruessner, J. C., Collins, D. L., & Alzheimer's Disease Neuroimaging Initiative. (2013). Prediction of Alzheimer's disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning.
Neuroimage.

4.Gamberger, D., Lavrač, N., & Srivatsa, S. (2017). Feature selection for the early detection of Alzheimer's disease. Journal of Alzheimer's Disease.

5.Khedher, L., Ramírez, J., Górriz, J. M., & Brahim, A. (2015). Alzheimer's disease neuroimaging initiative. Automated diagnosis of Alzheimer's disease based on threedimensional MRI features using probabilistic neural network. BioMed research international, 2015.

6.Konig, A., Satt, A., Sorin, A., Hoory, R., Toledo-Ronen, O., Derreumaux, A., ... & Aalten, P. (2018). Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring.

7. Liu, S., Liu, S., Cai, W., Che, H., Pujol, S., Kikinis, R., ... & Feng, D. (2014). Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. IEEE transactions
on bio-medical engineering.

8.Martinc, M., Haider, F., Pollak, S., & Luz, S. (2021). Temporal integration of text transcripts and acoustic features for Alzheimer's diagnosis based on spontaneous speech. Frontiers in Aging Neuroscience, page 299.

9.Mahendran, N., Durai Raj Vincent, P.M., Srinivasan, K., & Chang, C.Y. (2021). Improving the classification of Alzheimer's disease using hybrid gene selection pipeline and deep learning. Frontiers in Genetics, 12.

10. Pellegrini, E., Ballerini, L., Hernandez, M. D. V., Chappell, F. M., González-Castro, V.,

11. Sadeghian, R., Schaffer, J. D., & Zahorian, S. A. (2021). Towards an automatic speech-based diagnostic test for Alzheimer's disease. Frontiers in Computer Science, 3:624594.