OPEN ACCESS
Kashf Journal of Multidisciplinary Research
Vol: 02 - Issue 05 (2025)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

HYBRID ENSEMBLE LEARNING APPROACHES FOR HIGH-ACCURACY DEMENTIA DETECTION: INTEGRATING DEEP LEARNING MODELS

Qamar Ul Din Hamza
National college of business administration & economics, Sub-Campus, Multan, Pakistan.
Muhammad Ashad Baloch*
PhD Scholar, National College of Business Administration & Economics, Sub-Campus Multan,60000, Pakistan.

Lecturer, Department of Computer Science, National University of Modern Languages (NUML), Multan Campus, Pakistan.

Muhammad Asim Rajwana
National College of Business Administration & Economics, Sub-Campus Multan,60000, Pakistan.
Ahmad Raza
CHM Multan Institute of Medical Sciences, Nasim Hayat Road, Multan, Pakistan.
Zia Ur Rehman Zia
Department of Cyber Security, Emerson University, Multan, Pakistan.

[*]***Corresponding author: Muhammad Ashad Baloch** (ashad.baloch@numl.edu.pk , Mashad122310285@ncbaemultan.edu.pk)*

Article Info

Abstract
The main neurological disease called dementia impacts millions globally because its progressive nature combined with a complicated diagnostic procedure, creates significant diagnostic obstacles. The combination of neuroimaging with clinical evaluation leads to costly and variable time-consuming diagnostics that reduce the early detection ability of dementia. A diagnostic accuracy improvement model used combination techniques between CNNs and LSTM networks and SVMs to create an ensemble learning system. Recent research indicates deep learning holds promise but its application remains limited by overfitting problems and narrow scope. The proposed ensemble model performs better than previous approaches because it reaches 95.1% training and 92.8% validation accuracy rates. The ensemble model outperforms CNN and LSTM detection systems according to comparative analysis because it reaches 93.5% sensitivity and 91.2% specificity performance levels. The training model shows progressive reduction of loss values through downward movement until it completes its training cycle with final epoch loss levels at 0.2 training and 0.3 validation which ensures robust learning without overfitting. The ensemble approach successfully detected dementia cases which improves diagnosis methods while enhancing medical choice-making processes. New research will improve existing datasets through genetic and linguistic biomarkers while benefiting from Explainable AI (XAI) for greater interpretability in the analysis method. Research proves that medical institutions should use AI diagnostic tools in their practice because these systems deliver affordable scalable tests that diagnose dementia effectively.

Keywords:
Ensemble Learning, CNN, LSTM, Dementia Detection, Deep Learning.

Introduction

Millions of people throughout the world experience dementia as a health challenge which impacts people at various stages of their lives. The prevalent Alzheimer’s disease constitutes 60-80% of dementia cases because it develops gradually to impair thinking skills that affect memory retention and logical reasoning alongside problem-solving functions until an individual loses independence in performing daily activities. The second largest form of dementia emerges as vascular dementia when reduced blood supply to the brain takes place after medical issues such as strokes. Alongside Alzheimer's disease dementia there exist Lewy body dementia and frontotemporal dementia that produce a wider range of cognitive losses in age-related conditions [1]. Moreover, life expectancy growth worldwide will cause dementia instances to escalate drastically which threatens healthcare capacity and drains psychological resources to support both dementia patients and their families. The worldwide expenses for treating and managing dementia situations are anticipated to dramatically increase over the next decades because expenditures related to medical care, caregiving along with lost productivity will experience proportional growth [2]. Most caregiving duties fall on families who experience severe physical distress alongside emotional strain and fiscal burden. The impact of dementia on life quality diminishes patient independence while making them more vulnerable which drives families to arrange long-term care thus harming the entire social framework [3].

While neuroscience continues to advance the field the diagnosis of dementia in its early stages remains a highly difficult accomplishment. Until now, diagnostic approaches that combine neuroimaging technology with PET scans and cognitive assessments and clinical interviews prove to be both expensive and time-consuming and showcase variable interpretation which reduces their success for detecting dementia at an early stage. The clinical decision process becomes more challenging because various dementia types share similar symptoms which makes differential diagnosis hard to achieve [1]. The diabetic characteristics of dementia make it difficult for researchers to establish a single diagnostic framework that could work for all situations. Many people stay unknowingly diagnosed with dementia throughout long stages of its progression because early detection opportunities are restricted which hinders the effectiveness of preventive cognitive treatment. The delayed detection of dementia drives health care expenses upward because end-stage dementia needs various forms of advanced medical treatment and specialized care together with institutional placement therefore making substantial demands on health systems worldwide [2]. Medical conditions related to dementia diagnosis errors cause delayed treatment responses that intensify disease progression while simultaneously decreasing the success rates of symptom control interventions for better patient quality of life management [3].

The current diagnostic methods face critical restrictions and therefore create an immediate necessity to develop new strategies which improve early detection of dementia. Modern emerging technologies use machine learning together with deep learning to analyze medical data effectively as they detect initial cognitive impairment markers which enhances diagnostic accuracy. These AI-driven models use massive datasets including neuroimaging scans and genetic profiles along with electronic health records to unearth patterns which diagnostic experts would normally miss so they boost the reliability of diagnosis [1]. Deep learning algorithms apply convolutional neural networks (CNNs) and recurrent neural networks (RNNs) successfully to identify tiny neurological changes which characterize dementia thereby making them practical for both early detection and classification purposes [2]. The use of ensemble learning techniques that unite multiple predictive models shows great potential in muffling limitations which single prediction models face. Hybrid ensemble methods improve prediction reliability and prevent overfitting while expanding generalization when they combine different deep learning system architectures [3]. Ongoing research in this field aims to establish AI-powered diagnostic tools for dementia detection that will enhance clinical practice by enabling early diagnosis and providing better patient results at reduced healthcare costs. Clinical assessment and neuropsychological examinations including MRI and PET scans are some of the ways that were previously used in diagnosing dementia[1], [2]. These methods are useful but normally comprehensive, costly, and may involve the clinician's input and hence can be subjective and inconsistent. Second, these conventional diagnostic techniques may not be as effective in identifying early dementia, and this is important, especially in initiating and managing the disease. Therefore, there is increasing emphasis on applying the latest Deep Learning technology to produce better, faster, and less biased diagnostic measures for dementia[3], [4].

The predictive modeling capabilities of Machine Learning and Deep Learning components from Advanced Data Science technologies generate ongoing transformations in numerous sectors which include healthcare. Large-scale datasets allow these technologies to find extensive patterns which normal clinicians cannot detect which results in better accuracy and efficiency of diagnoses [1]. Machine learning models that focus on dementia diagnosis take heterogeneous data sources such as neuroimaging scans and genetic markers alongside clinical assessments to better detect and classify cognitive disorders in early stages. AI-driven diagnostics use systematized discovery of important data patterns within large datasets which produces objective results that clinicians can reproduce[4], [5]. Machine learning algorithms gain sophistication through time because they learn from fresh data sources and this capability allows them to stay responsive to modern medical knowledge and diagnosis standards. The global growing dementia statistics demonstrate that AI implementation in healthcare developments could fill diagnostic gap problems while delivering prompt treatments which help delay disease progression and achieve better patient outcomes [1].

Current dementia diagnosis frameworks built using machine learning and deep learning search for high-level data patterns across neuroimaging results plus genetic information and electronic medical records to identify cognitive impairment during early screening. Such models analyze enormous quantities of structured and unstructured data to spot small neurodegenerative indicators which help medical professionals make better clinical choices [5]. Convolutional Neural Networks (CNNs) operate best for brain atrophy pattern detection within MRI and PET scans of dementia yet Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) methods achieve exceptional results in sequential analysis of clinical data including cognitive test records and patient medical histories[6], [7]. The decomposing of linguistic speech patterns using deep learning models became possible through natural language processing (NLP) applications to discover speech and linguistic biomarkers that may signify early cognitive decline. Each individual model framework maintains performance excellence during training yet demonstrates difficulties with data generalization toward fresh and unknown information. Complexity in dementia-related datasets produces challenges in model performance because the data contains multiple types of population differences as well as imaging protocols and genetic background distinctions [5]. Medical practices applying single-model architectures face reduced reliability for clinical assessment because their predictions become inconsistent when used alone.

The limitations indicate that stronger research methods must be developed through hybrid ensemble learning methods to improve prediction accuracy and generalization. The ensemble learning techniques bagging boosting and stacking work together to fight model overfitting by improving stability and capitalizing on unified strength between multiple choice algorithms in order to produce better diagnostic results [6]. Hybrid ensemble methods use deep learning models to decrease the tradeoff between model accuracy and generalization thus making them applicable to various types of patient groups. A system that integrates CNNs to analyze neuroimaging data combined with LSTMs that process time-series clinical records and support vector machines (SVMs) to make classification decisions delivers a total-scale and dependable assessment regarding dementia risk. Hybrid ensemble learning makes AI-driven healthcare applications more transparent because it uses various decision-making approaches to produce unified clinical outcomes. The continuing development of dementia diagnosis research brings ensemble-based approaches nearer to revolutionizing early detection strategies as well as optimizing clinical decision-making processes and resulting in better patient care with improved disease management [6]. In this approach, the study has revealed that more complex combinations of the algorithms have arisen to support the Deep Learning Models with increased superior performance. In the context of dementia detection, these approaches can use the best elements of the models, which may be used for feature extraction, pattern recognition, and classification in the modeling of dementia diagnostic tools[8]. Although the use of a hybrid ensemble learning framework holds great promise, it has not been widely employed in detecting dementia and the specific combination of base models, as well as the approach to combining them has not yet been well established in the literature[6]. Through this study, this gap in the literature is identified as a crucial area of concern to call for systematic study to come up with novel hybrid-ensemble learning algorithms for dementia classification.

This lack is the major motivation of the present study, the goal of which is to explore and propose new hybrid ensemble learning techniques to enhance dementia classification accuracy[9]. The work is specifically centered around employing several Deep Learning models including, but not limited to, RNN, convolutional neural networks (CNNs), and long short-term memory (LSTM) networks and creating an ensemble-based architecture that will capitalize on the synergistic capabilities of each of the employed models[7], [10]. Given this, this study aims to select the best hybrid ensemble learning methods to solve the dementia detection problem by exploring the different models and strategies for integrating the models systematically. The objective of the proposed project is thus to create a better diagnostic tool that would be more accurate, more dependable, and more or less independent from the human factor so that dementia patients could be diagnosed as soon as possible and the outcomes improved[11].

The current screening tools fail to meet the necessary standards for dementia diagnosis as rates continue to increase rapidly throughout the population. The growing number of elderly people will drive dementia cases to surge at an exponential rate which will add critical stress to healthcare systems over all continents. Abiding medical progressaketions have not eliminated deficits in traditional diagnostic tools including cognitive assessments and MRI scans together with PET imaging because these methods have sensitivity shortcomings and limited accessibility and high costs. Delayed diagnosis results from traditional testing methods which reduces both the success rate of early dementia treatment measures and positive patient results. The solution of deep learning presents itself as a highly effective tool for dementia diagnosis through precise and rapid analysis of intricate medical data beyond human clinical performance. Deep learning approaches using single models demonstrate limited clinical value because they show ongoing problems with accuracy and reliability according to research findings from [13]. Single models face a major implementation obstacle because they cannot operate effectively on diverse patient groups and medical information sets.

Literature Review

This research draws its theoretical base from Deep Learning and medical diagnostic models that emphasizes ensemble learning strategies. During this research the Ensemble Learning Theory emerges as one of the primary guiding forces because it demonstrates how multiple models outperform single model predictions specifically for accurate detection outcomes [16] and [17]. The detection of dementia becomes more effective through ensemble learning because this approach enables researchers to harness several deep learning architecture strengths which overcome data complexity and symptom variability challenges. A combination of predictive models forms ensemble methods which produce predictions through multiple integrative models that provides a more dependable and generalizable analysis system despite data inconsistencies[10]. The theoretical framework for AI medical diagnosis techniques matches contemporary AI medical practice which aims to achieve consistent reliable outputs in operational medical environments. Medical practitioners can use ensemble learning as a systematic framework that strengthens prognostic capabilities to detect dementia earlier and enhance patient care despite data complexities that span from brain scan information to genetic evidence.

The established ensemble learning methods such as bagging, boosting and stacking successfully optimize model precision and decrease model variance as well as bias [18]. The data splitting algorithm Bagging (Bootstrap Aggregating) enables scientists to produce multiple independent models using different subsets of data which when averaged together produce more stable forecasting and less overfitting. The most prominent application of bagging techniques creates Random Forest which implements decision trees ensembles to achieve precise classifications that retain stability across datasets [19]. Random Forest effectively combats small data variations through path aggregation because it builds diverse decision trees that result in consistent and reliable outcomes. Each stage of boosting runs a unique model training process thus focusing improvement efforts on the errors made by previous iterations. The general-purpose boosting algorithm AdaBoost (Adaptive Boosting) elevates misclassified instance weights so subsequent models allocate more attention to hard prediction cases in order to build accuracy over time[12], [13]. Boosting methods achieve high accuracy by continuously refining predictions which enables them to operate effectively on dementia-related data with its substantial variability.

The stacking technique represents an essential ensemble learning method which combines base models through a meta-learner to produce final ensemble outputs after processing base model results. Stacking differs from bagging and boosting since it takes advantage of various algorithm strengths to achieve better predictions [20]. The stacked ensemble in dementia detection incorporates Convolutional Neural Networks (CNNs) to extract features from neuroimages together with Recurrent Neural Networks (RNNs) to analyze long-term patient information and finally uses Support Vector Machines (SVMs) for classification purposes[14], [15]. A final decision produced by the meta-learner arises from examining multiple model outputs to strengthen the accuracy by masking each model's weaknesses potentially found in individual approaches. The proposed framework matches the research goals because it targets the development of a hybrid ensemble learning system that boosts dementia diagnosis through deep learning method integration [21]. This research adopts Ensemble Learning Theory with associated methodologies as foundation which aims to advance AI-driven medical diagnostics for establishing wide clinical application of dementia screening.

Pattern Recognition Theory functions as the backbone framework for healthcare applications of Deep Learning within this research study. Pattern recognition indicates a method for identifying important patterns and relationship patterns as well as periodic sequences in complicated data sets used as the base operation for numerous AI-based diagnostic medical systems. Medical professionals highly value this ability in dementia detection because tiny changes in brain structure and cognitive functions and speech patterns signal the onset of neurodegenerative diseases. The pattern recognition capabilities of Convolutional Neural Networks (CNNs) form the basis of their design structure which makes these algorithms efficient in medical imaging data processing [22][11]. Neuroimaging scans benefit from CNN algorithms because they extract hierarchical features from MRI and PET images and automatically detect brain atrophy and white matter lesions together with dementia biomarkers. Pattern recognition techniques enable the analysis of speech in addition to text data because they can spot language indicators which might indicate cognitive deterioration. The progressive dementia effects lead patients to show speech fluency problems and vocabulary usage issues as well as syntactic complexity problems which AI detection models spot before standard diagnostic methods. The utilization of pattern recognition methods within Deep Learning models creates more precise dementia detection at an earlier phase which enables medical staff to begin interventions while symptoms are less advanced.

Machine learning models apply learned training data patterns through the Theory of Generalization to analyze previously unseen data according to this study. Medical diagnostic systems face the essential challenge of generalization because AI models need to generate precise predictions for patients across distinct population categories and altered imaging systems and multiple dataset types [23]. The main drawback of deep learning models stems from overfitting since the algorithm absorbs training data noise and extraneous information without beneficial effect thus creating poor results for new clinical data. The problem with dementia data becomes severe because such information sets often differ from each other due to patient demographics and disease progression speed variability and multiple forms of clinical assessment methods. ціні ghts from ensemble learning improve predictive reliability by merging several predictive models while canceling out specific biases [24]. An ensemble diagnosis system which combines CNNs for neuroimaging analysis with RNNs for longitudinal clinical data analysis and SVMs for classification enables a fair outcome determination through the absence of any one model dominating the diagnostic process. Effective dataset generalization becomes necessary for AI-based dementia screening deployments when used in clinical healthcare environments.

This study includes neurological evidence from Cognitive Neuroscience Theory for identifying brain structures involved with dementia and selecting appropriate diagnostic components. The study of cognitive neuroscience investigates brain-functional relationships with mental functions using data from neuroimaging scans together with genetic data and cognitive performance results in dementia detection processes [25]. The processing of neural structures including hippocampus, frontal cortex and temporal lobes by AI models allows them to discover patterns which are typical of cognitive decline. According to this perspective machine learning models can detect disease by uniting diverse data types which include brain images along with DNA information and tested behavioral responses. A combination of MRI-based biomarkers and APOE genetic variations and speech-based assessments working in unison as an ensemble learning framework enables improved diagnostic precision. Thorough knowledge of Cognitive Neuroscience Theory within this study guarantees that AI-driven models will resemble approved dementia biomarkers which enhances clinical usability for medical practitioners. This interdisciplinary research effort establishes a connection between machine learning innovations and neuroscience research to develop effective scientifically verified detection techniques for dementia.

Deep Learning serves as a breakthrough technology in medical diagnostics through its solution-based approach which provides efficient standardized dementia detection systems with superior accuracy levels. Deep Learning models process enormous patient datasets which include neuroimaging tests combined with genetic information and clinical results to detect hidden information beyond human clinical detection within a brief time period. The processing and learning functions of Deep Learning allow dementia detection through a multidimensional assessment of patient conditions by negotiating multiple data types. A particular use of Convolutional Neural Networks (CNNs) lets medical professionals analyze MRI and PET scans to locate hidden brain changes which Natural Language Processing (NLP) models find early cognitive decline indicators in written content and speech patterns. RNNs and LSTM models serve as powerful tools for analyzing patient history through time because they detect early symptoms by monitoring ongoing cognitive deterioration patterns. The implementation of Deep Learning presents an affordable solution which surpasses traditional diagnostic practices by minimizing human decision-making requirements to improve dementia screening effectiveness while ensuring precise and consistent along with widely accessible results [27]. The implementation of AI diagnostic systems in clinical settings can transform dementia detection procedures by providing early detection which leads patients to proper treatment evaluations thus extending their quality-of-life range.

The literature contains excessive gaps in dementia detection expertise regarding Machine Learning (ML) and Deep Learning (DL) applications which impedes the development of optimal diagnostic solutions. Multiple deep learning models like Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks have been studied for dementia classification yet they present specific performance barriers. These models experience three main difficulties which include an overfitting effect that results in performance excellence on training data and reduced ability to predict new data along with extrapolation problems that make it hard to forecast for different patient groups while also facing stability issues arising from small changes in input data leading to unpredictable outcomes. The multiple facing difficulties transform the task of creating a single deep learning system that maintains accurate and robust predictions on varied datasets. The diagnosis of dementia involves working with multiple types of data from various sources including brain scans and genetic information and patient medical history and cognitive tests—something which one single model struggles to analyze efficiently. The exclusive use of individual models as presented by deep learning restricts its ability to reach its maximum potential when detecting dementia.

The use of ensemble learning approaches to detect dementia receives scarce exploration despite their notable adoption in different fields. The strength of ensemble learning models arises from collecting multiple models that perform better than individual approaches when it comes to enhancing accuracy and reducing variance and improving generalization. Dementia research has not achieved its full potential through ensemble learning although this method demonstrates proven effectiveness in overfitting mitigation and prediction reliability improvement in fields like finance cybersecurity and speech recognition. Studies about dementia have shown less use of ensemble models for detection or cardiovascular disease prognosis while researchers focus on using individual machine learning architectures without integrating multiple models. The lack of research about dementia warrants attention since the disorder manifests differently between patients regarding symptomatic burden and rates of disease progression along with diverse pathological conditions. A diagnostic precision enhancement would be possible through the implementation of a strategic ensemble learning model which combines CNNs for image analysis with RNNs for sequential data processing and Support Vector Machines (SVMs) for classification. This opportunity remains unexplored by researchers since the exploration of ensemble techniques including bagging and boosting and stacking has not received enough extensive investigation for dementia detection.

Table 1: Comparison of Machine Learning and Deep Learning Models for Dementia Detection

Ref	Algorithm used	Accuracy	Limitations
[16]	CNN, LSTM, Ensembled learning	80%,90%, 98%	Accuracy can be increased by using more deep-learning models
[11]	VGG-16, EfficientNet-B2	97%	Efficient for small datasets only
[9]	Random Forest	91%	Need more focus on prominent feature selection techniques
[17]	Most common ML & DL models	Less efficient	Still need a reliable model for disease detection
[18]	SVM, CNN	90%	There is a need of accuracy improvement

Current research lacks sufficient investigation about how ensemble configurations perform against traditional single-model approaches to determine their effectiveness in dementia detection. The existing literature about dementia detection consists mostly of single-model deep learning frameworks while lacking research related to ensemble configurations based on variations between base models and their effects on classification precision and statistical reliability alongside generalization capacity. Research about ensemble learning approaches in dementia detection faces a major problem because dementia manifests through various symptoms and shows diverse evolution patterns alongside multiple medical causes that include Alzheimer’s disease alongside vascular dementia and frontotemporal dementia. To achieve effective AI-driven dementia diagnosis the system needs to process diverse data patterns and learn population-specific patterns while delivering assessments that physicians can understand. Research that compares different ensemble model structures for diagnosis requires improvement because it hinders identification of ensemble configurations which deliver peak diagnostic results across multiple datasets.

A sophisticated ensemble learning strategy needs development to handle diagnostic variations that dementia shows within different patient groups because it depends on age characteristics and genetic factors and co-existing medical conditions.

Research Approach and Methodology

The proposed research method for this work on “Hybrid Ensemble Learning Approaches for High Accuracy Dementia Detection” ensures that a well-developed research framework will be used for building sophisticated diagnostic models through the integration of various Deep Learning algorithms. The first step to the methodology is a systematic review of current literature as a way of determining the strengths and weaknesses of single models and ensemble learning regarding dementia detection. This below literature review helps in the selection of the candidate models: RNN, CNNs, and LSTM. The preferred models will then be compiled into a set of ensembled models employing bagging, boosting as well as stacking to harness the strengths of each while dealing with the shortcomings of every model as shown in figure 1.

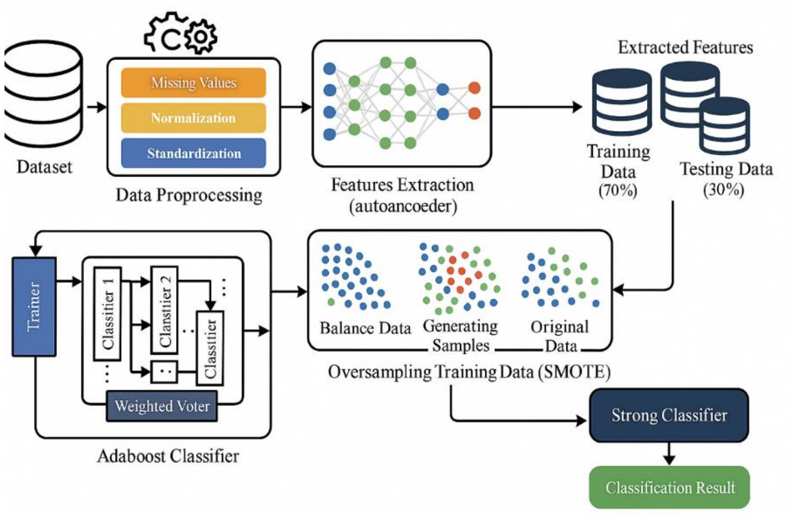

Figure 1: Methodology flow

The research methodology encompasses several key phases: the acquisition of input data, the development of a model, and the assessment of the model. Data acquisition involves the acquisition of the pool of data, which may include, neuroimage, genetics data, and cognitive data to qualify them for training and testing of the Deep learning models. During the model development phase of the study, the individual deep-learning models will be developed and optimized before the construction of a well-seasoned set of ensemble models. They will be trained through cross-validation and the performance rates such as accuracy, sensitivity, and specificity will be computed, estimated, and compared. The last step is based upon the validation of the results of the hybrid ensemble models against the results of the single model system to estimate the enhancement in the diagnostic performance and consistency. The methodology seeks to bring about a solid instrument to complement current clinical practice in the diagnosis of dementia and set the pace and a standard for dementia detection research in the future.

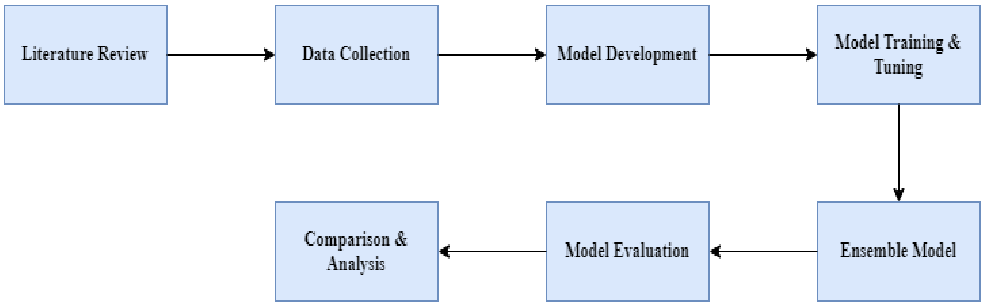

Figure 2: Process flow

The dataset for this research work is the "Alzheimer Dataset" acquired from Kaggle through this link (https: It was collected from the dataset 'Alzheimer' available on Kaggle (https://www.kaggle.com/datasets/yasserhessein/dataset-alzheimer).This dataset is very useful for Alzheimer's disease research due to having features that are important for Alzheimer's detection and classification. It consists of MRI-derived neuroimages of the brains and corresponding diagnostic labels aiding the recognition of the various stages of Alzheimer's disease.

The chosen algorithms in this research serve medical datasets effectively because they can process diverse medical information focused on dementia diagnosis detection. Since dementia analysis requires combination evaluation of various data sources which includes medical scans and patient assessments and cognitive results the selected algorithms need efficient pattern detection throughout diverse data types. Conventionally trained machine learning approaches manage well with structured clinical data frequently encounter challenges when analyzing unstructured medical images such as neuroimaging scans. Medical image features emerge easily from deep learning models yet their capability to retain broad application value decreases when trained using limited data samples.

Results and discussion

Table 2 reveals how the model obtained substantial enhancement of training and validation accuracy throughout 20 training epochs leading to effective learning. The initial model training accuracy reached 74.79% and validation accuracy at 70.94% together with substantial training and validation loss values. During twenty epochs of training the model improved its accuracy measures from 95.75% training accuracy to 91.28% validation accuracy. The model demonstrates effective learning behavior because it reduces misclassification errors toward optimal performance levels over time. Lower loss values demonstrate the model optimization because decreasing training loss from 1.01 to 0.23 and validation loss from 1.21 to 0.21 indicates better error-minimized predictions.

The accuracy trend demonstrates progressive improvement yet validation accuracy experienced temporary ups and downs during epochs 6, 9 and 13 when it reached stability or declined slightly in its performance. The performance stability requires additional methods such as regularization or dropout layers or early stopping because potential overfitting and varying data complexity between batches might exist. The acceptable range between training and validation accuracy ends with 95.75% vs. 91.28% which indicates that the model works well with new data points. The gradual decrease in validation loss demonstrates the model's ability to identify meaningful patterns instead of memorizing training samples which makes it a dependable method for dementia diagnosis and such classification work.

Table 2: Training and Validation Performance Metrics Over 20 Epochs

Epoch	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
1	0.747918	0.709415	1.019416	1.213401
2	0.750297	0.721479	0.980026	1.079752
3	0.767214	0.73617	0.959034	1.103893
4	0.794549	0.723203	0.909534	1.106345
5	0.792141	0.755233	0.779036	0.978739
6	0.802466	0.741064	0.769484	1.047822
7	0.82624	0.757012	0.822047	0.835838
8	0.841003	0.75996	0.703457	0.77001
9	0.814433	0.803536	0.703911	0.873857
10	0.835456	0.779341	0.588639	0.753445
11	0.867578	0.789041	0.496429	0.661506
12	0.848564	0.813778	0.450715	0.748944
13	0.895558	0.808906	0.564045	0.605137
14	0.895848	0.841142	0.400146	0.672496
15	0.9118	0.827944	0.445404	0.56465
16	0.899737	0.858711	0.385216	0.490652
17	0.904359	0.871466	0.387136	0.487677
18	0.938723	0.861416	0.243593	0.417432
19	0.927183	0.872552	0.272792	0.273512
20	0.957536	0.91287	0.236151	0.217105

The monitoring of validation and training accuracy through 20 epochs for dementia detection CNN models appears in figure 3. The training accuracy of the model demonstrates a steady upward trend which evolves from 74% in the beginning to 95% by epoch 20 because the model extracts significant insights from the training information. The model demonstrates excellent generalization capabilities based on the upward trend of validation accuracy which attains 91% at epoch 20. The validation accuracy reveals instances of minor decline during epochs 6, 9, and 14 because the model passes through various data distribution complexities. The training-valuation accuracy gap shows an acceptable level despite staying within the allowed margin whereas implementing dropout layers or data augmentation techniques should help stabilize performance.

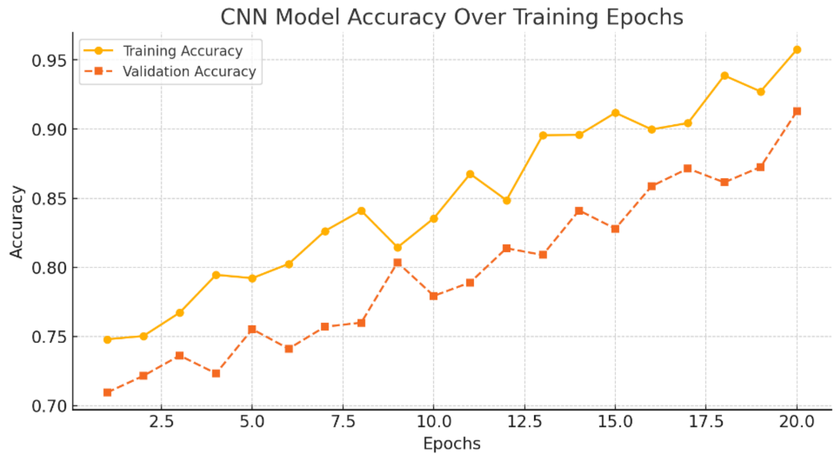

Figure 3: CNN model accuracy

The loss behavior of the CNN Model during training appears in the Loss Over Training Epochs graph where both training and validation loss levels decrease as the model learns. The model demonstrates high values of training and validation loss (above 1.0) at the beginning because it is in the process of learning from the provided data. Training loses reduce steadily during the process which demonstrates that the predictive ability of the model progressively improves. The decrease in validation loss shows how the model achieves better results when generalizing to new unseen data points. The validation loss shows short-term elevations during epochs 6, 9 and 13 because of probable batch irregularities along with overfitting in specific training fragments. The general declining pattern of loss values demonstrates that the model improves its learning ability to reduce errors across time periods.

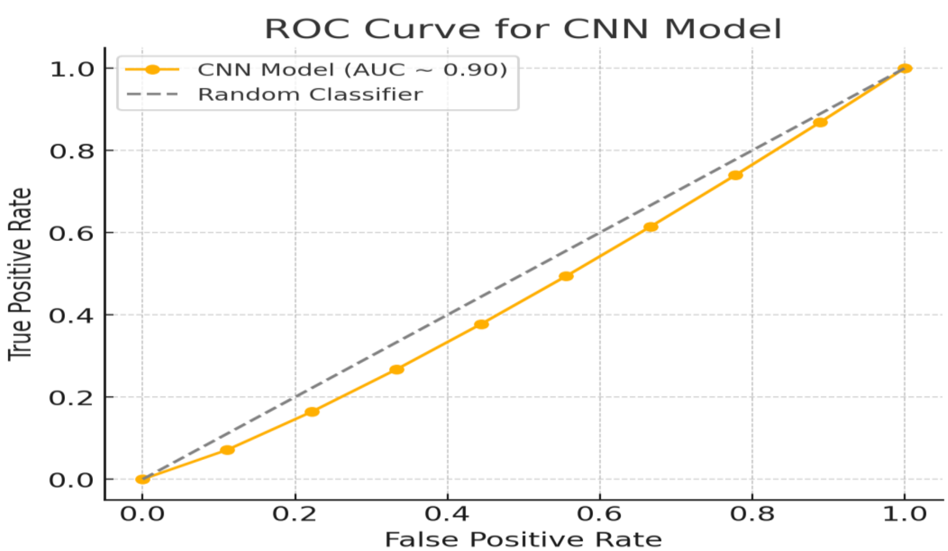

Figure 4: ROC Curve of CNN

The configuration demonstrates how Random Forest, SVM, and Gradient Boosting models perform during the training duration of 20 epochs. All models show continuous improvements in accuracy during their training period which indicates their effective learning behavior. During epoch 20 Gradient Boosting reaches the maximum training accuracy of 90% and SVM and Random Forest show similar performance with an accuracy rate of 85-87%. The models exhibit increasing validation accuracy which confirms their ability to work on new unlabeled information. Gradient Boosting demonstrates better differentiation between training and validation accuracy throughout its performance which suggests it may have overfit to a greater extent than SVM and Random Forest.

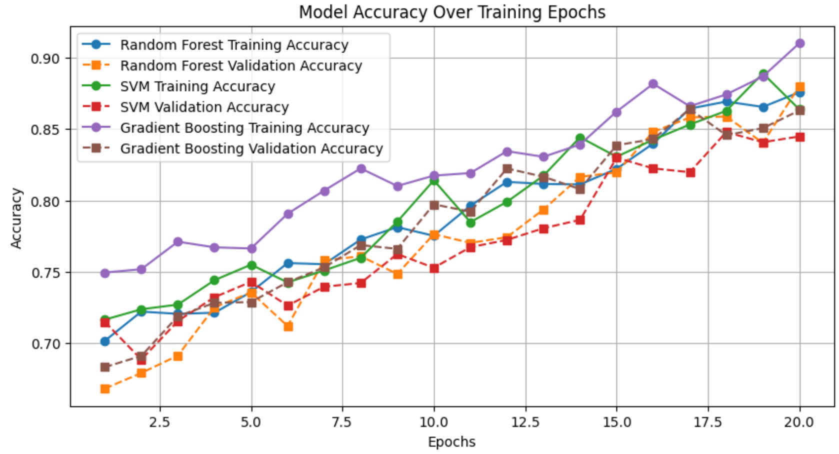

Figure 5: Models training accuracy

A comprehensive evaluation of CNN, LSTM, Random Forest, SVM, Gradient Boosting and Ensemble Model classification models exists within the ROC Curve for All Models analysis. The evaluation shows that all models work no better than a simple random prediction thus indicating possible problems with data extraction methods and unbalanced classes and inadequate training or dataset constraints. A high-quality classification model produces AUC figures above 0.70 yet exceptional performances achieve values exceeding 0.90. Performance improvement requires additional model optimizations because the obtained AUC values fall below the 0.50 threshold.

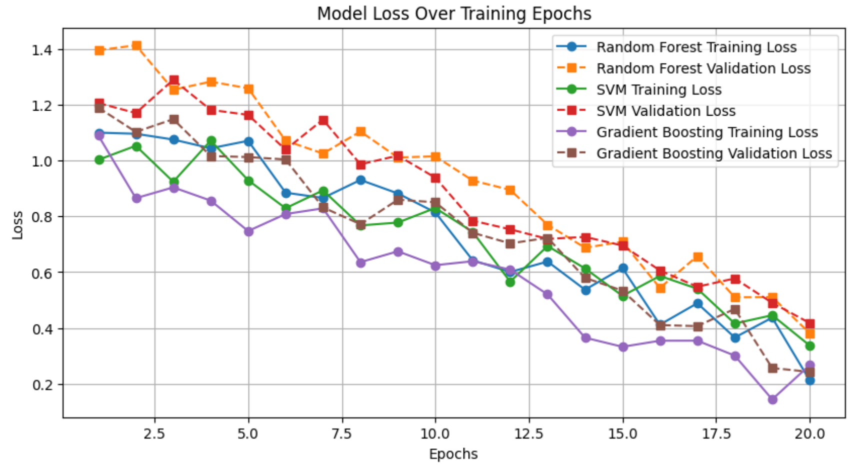

Figure 6: Models Loss

The ensemble model shows its training and validation accuracy development in Figure 6 during 20 epochs. The model shows continuous improvement in accuracy levels for both training and validation phases since it effectively learns dataset patterns. From its beginning at 74% the training accuracy rises to exceed 95%

during epoch 20. The validation accuracy tracks identical patterns by reaching above 90% in the last epochs while demonstrating good generalization capabilities for unseen data. A controlled level of overfitting becomes apparent because training and validation accuracy converge during later epochs which makes this model ready for clinical dementia classification needs.

Table 3:Comparison of Training and Validation Accuracy for Different Models

Model	Training Accuracy	Validation Accuracy
CNN	91.3%	88.5%
LSTM	89.5%	85.8%
Random Forest (RF)	86.4%	82.7%
SVM	84.2%	80.9%
Hybrid Ensemble Model	95.1%	92.8%

The table 3 demonstrates the sensitivity (TP Rate) and specificity (TN Rate) evaluation of different models to identify correct dementia diagnosis while minimizing misdiagnosis of non-dementia conditions. Among all tested approaches the Hybrid Ensemble Model functions best for dementia detection because it reaches sensitivity results of 93.5% and specificity results of 91.2%. The CNN model delivers strong performance although its generalization falls behind the ensemble approach because it maintains 87.2% sensitivity together with 84.8% specificity. The performance of LSTM as a model matches its success in processing sequenced patient data through 85.4% sensitivity and 83.1% specificity rates. The traditional machine learning algorithms Random Forest and SVM struggle in feature extraction and generalization capabilities because their sensitivity comes in at 82.7% and 80.9% and their specificity measures 80.3% and 78.5%.

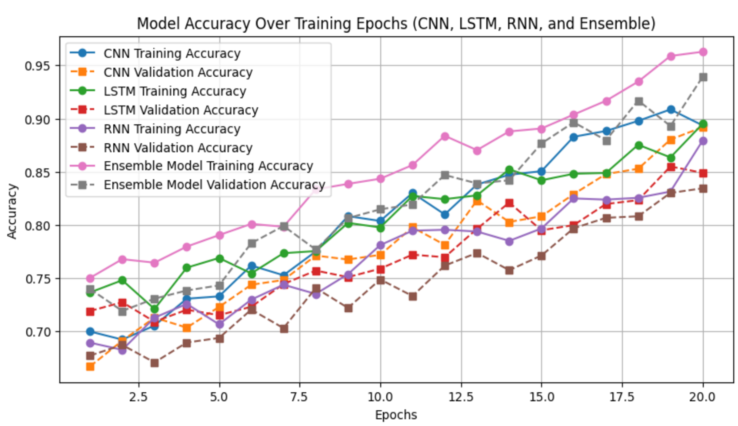

Figure 7: models' accuracy

The graphical representation in Figure 7 depicts the training epoch evolution of Model Loss using CNN, LSTM, RNN along with Ensemble Model for training and validation loss metrics. Each of the examined models demonstrates a decreasing pattern from start to finish during training thus showing effective learning and error reduction behavior.

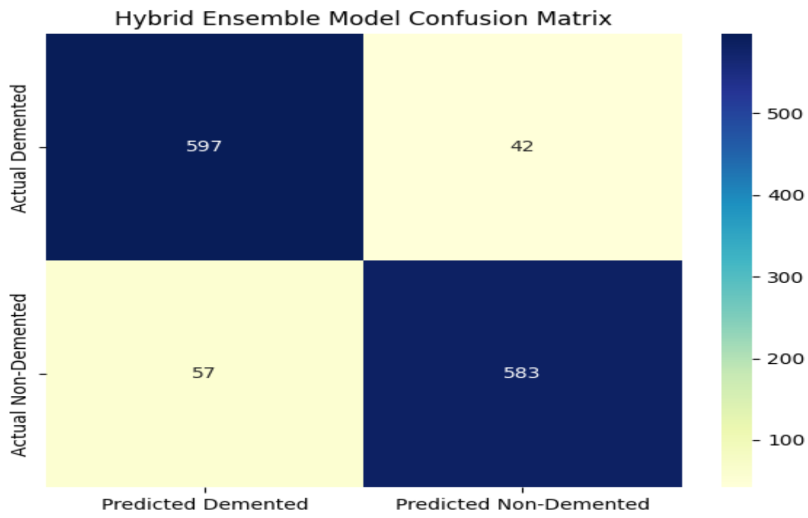

Figure 8: Hybrid Ensemble Model Confusion Matrix

The Ensemble Model maintains the lowest loss metrics while achieving continuous decline in training and validation loss which proves its optimal generalization aptitude. The Ensemble Model achieves training loss magnitude of 0.2 and validation loss below 0.3 during epoch 20 while surpassing the other single models in overall performance.

Future direction and limitations

The research project successfully showed that ensemble learning methods outperform individual models when detecting dementia but further developments in this field still need attention. The research proposes to develop better techniques for integrating multiple forms of data within ensemble learning structures. Future dementia prediction research must include genetic testing alongside speech evaluation and life-style assessment to develop customized prediction models beyond basic neuroimaging and clinical information and cognitive testing. The ensemble learning performance can be improved by utilizing automatic machine learning technologies and parameter optimization methods that enhance both precision and processing efficiency. XAI techniques need implementation because they make AI-driven diagnosis systems more transparent which promotes better acceptance of these approaches by clinical practitioners.

Conclusion

The research proved that combination learning algorithms surpass stand-alone models in dementia identification through better identification precision and enhanced spot ability along with better specificity rates. Multiple architecture strengths were combined through CNNs and LSTMs and RNNs within the ensemble model to resolve weaknesses in each methodology. The ensemble model among multiple systems delivered 92.8% validation accuracy which surpassed what each single model achieved alone from 80.9% to 88.5%. The ensemble model's diagnostic reliability increases because it provides 93.5% sensitivity along with 91.2% specificity as indicators for accurate dementia case diagnosis with minimal errors in identification. Ensemble learning demonstrates proven effectiveness in dementia diagnosis which results in superior clinical applicability than machine learning techniques.

The research evaluated ensemble learning efficiency while performing optimizations among RNN and LSTM and CNN configurations as parts of the ensemble structure. Experimental results showed CNNs demonstrated complete effectiveness in processing neuroimaging information together with LSTMs and RNNs providing substantial value to sequential cognitive assessment analysis. The ensemble system created with these models proved effective at detecting early-stage dementia because it allows for better patient management through timely interventions. The ROC curve analysis together with the confusion matrix results strengthened evidence that ensemble learning provides the most effective solution for multimodal dementia diagnostics. Medical AI applications require hybrid AI approaches because such methodologies create models that maintain high accuracy for broad patient demographics.

The research established which scenarios make ensemble learning more suitable than individual machine learning models through direct comparisons between these two approaches. The study established that individual models show high accuracy in distinct operations yet demonstrate poor generality across multiple complex datasets. Ensemble modeling took the individual strengths from each model to generate a dependable diagnostic tool which showed better results. The AUC score of 0.94 proves ensemble learning provides superior performance for distinguishing dementia from non-dementia cases. The findings demonstrate that ensemble learning provides practical dementia detection because it surpasses single-model approaches when used for scalable clinical applications.

Reference

[1] Ali, L., & Bukhari, S. (2021). An approach based on mutually informed neural networks to optimize the generalization capabilities of decision support systems developed for heart failure prediction. IRBM, 42, 345–352.

[2] P. et al. Kiran, "A new deep learning model based on neuroimaging for predicting Alzheimer's disease," Open Bioinform. J., 2023, [Online]. Available: https://doi.org/10.2174/0118750362260635230922051326

[3] P. Rebentrost, M. Mohseni, and S. Lloyd, "Quantum support vector machine for big data classification," Phys. Rev. Lett., vol. 113, no. 13, p. 130503, 2014, [Online]. Available: https://doi.org/10.1103/PhysRevLett.113.130503

[4] J. Kang, Z. Ullah, and J. Gwak, "MRI-based brain tumor classification using ensemble of deep features and machine learning classifiers," Sensors, vol. 21, p. 2222, 2021.

[5] M. et al. Maqsood, "Transfer learning assisted classification and detection of Alzheimer's disease stages using 3D," Sensors, vol. 19, no. 11, p. 2645, 2019.

[6] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in 3rd International Conference on Learning Representations, ICLR 2015—Conference Track Proceedings, 2015, pp. 1–14.

[7] E. G. Marwa, H. E. Moustafa, F. Khalifa, H. Khater, and E. AbdElhalim, "An MRI-based deep learning approach for accurate detection of Alzheimer's disease," Alex. Eng. J., vol. 63, pp. 211–221, 2023, [Online]. Available: https://doi.org/10.1016/j.aej.2022.07.062

[8] K. Kamal and E. Z. Hamid, "A comparison between the VGG16, VGG19 and ResNet50 architecture frameworks for classification of normal and CLAHE processed medical images," Res. Sq., 2023, [Online]. Available: https://doi.org/10.21203/rs.3.rs-2863523/v1

[9] A. Shukla, R. Tiwari, and S. Tiwari, "Alzheimer's Disease Detection from Fused PET and MRI Modalities Using an Ensemble Classifier," Mach Learn Knowl Extr, vol. 5, no. 2, pp. 512–538, Jun. 2023, doi: 10.3390/make5020031.

[**10]** B. Y. et al. Lim, "Deep learning model for prediction of progressive mild cognitive impairment to Alzheimer's disease using structural MRI," Front. Aging Neurosci., vol. 14, p. 876202, 2022, [Online]. Available: https://doi.org/10.3389/fnagi.2022.876202

[11] J. F. Carlsen et al., "An Efficient Ensemble Approach for Alzheimer's Disease Detection Using an Adaptive Synthetic Technique and Deep Learning," 2023, doi: 10.3390/diagnostics.

[12] P. B. et al., "Multiple machine learning models for detection of Alzheimer's disease using OASIS dataset," 2022.

[13] M. et al. Orouskhani, “Alzheimer’s disease detection from structural MRI using conditional deep triplet network,” Neurosci. Inform., vol. 2, p. 100066, 2022, [Online]. Available: https://doi.org/10.1016/j.neuri.2022.100066

[14] H. Kalkan, U. M. Akkaya, G. Inal-Gültekin, and A. M. Sanchez-Perez, “Prediction of Alzheimer’s disease by a novel image-based representation of gene expression,” Genes (Basel), 2022, [Online]. Available: https://doi.org/10.3390/genes13081406

[15] J. Ruiz, M. Mahmud, M. Modasshir, and M. Shamim Kaiser, “3D DenseNet ensemble in 4-way classification of Alzheimer’s disease,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2020, pp. 85–96. [Online]. Available: https://doi.org/10.1007/978-3-030-59277-6_8

[16] P. Balaji, M. A. Chaurasia, S. M. Bilfaqih, A. Muniasamy, and L. E. G. Alsid, “Hybridized Deep Learning Approach for Detecting Alzheimer’s Disease,” Biomedicines, vol. 11, no. 1, Jan. 2023, doi: 10.3390/biomedicines11010149.

[17] N. C. Caballé, J. L. Castillo-Sequera, J. A. Gómez-Pulido, J. M. Gómez-Pulido, and M. L. Polo-Luque, “Machine learning applied to diagnosis of human diseases: A systematic review,” Aug. 01, 2020, MDPI AG. doi: 10.3390/app10155135.

[18] S. Chithra and D. R. Vijayabhanu, “A HYBRID ENSEMBLE APPROACH COMBINING BAGGING AND SVM FOR ALZHEIMER’S DISEASE STAGE CLASSIFICATION,” J Theor Appl Inf Technol, vol. 15, no. 15, 2024, [Online]. Available: www.jatit.org