Volume (1) No (10), 2025
Journal of Global Research in Multidisciplinary Studies (JGRMS)
Review Paper/Research Paper
Available online at https://saanvipublications.com/journals/index.php/jgrms/index

Deep Learning for Integrative Healthcare Analytics: A Focus on Multimodal Brain Imaging Data

Vivek Sharma
Designation: System Analyst
MITS Deemed University
Viveks@Mitsgwalior.in

***Abstract*—Medical imaging faces several significant challenges, including image segmentation, cross-modal translation, and real-value prediction. Link CT and MRI scans using a well-liked method. Image quality and diagnostic efficacy can be enhanced with this technique. One new area of data science is brain imaging genetics, which aims to better understand the brain's normal and abnormal phenotypic, molecular, and genetic features and how they influence its function and behaviour. Enhancing diagnostic accuracy in neurological diseases, this work proposes a multimodal approach to analyzing brain tumor imaging data for healthcare analytics. It leverages modern predictive models. Multiple models were tested using extensive data, such as CNN, Report Guided Net, MLP, and ResNet18. With an F1-score of 99.12%, recall of 99.67%, precision of 99%, and accuracy of 99.28%, the Convolutional Neural Network (CNN) model stood out from the contest. These results prove that deep learning techniques, such as CNNs, can understand complex multimodal images and extract valuable information from them. The results provide credence to the idea that these models can be useful in automating the identification of brain disorders by incorporating them into clinical procedures.**

Keywords—Smart Healthcare, Artificial Intelligence, Neuroimaging, Magnetic Resonance Imaging, Computed Tomography.

I. INTRODUCTION

Diagnosing diseases is a major obstacle in contemporary healthcare since prompt treatment interventions are possible with early and correct detection of pathological conditions, which in turn minimizes the risks of illness progression and sequelae. Considering the prevalence and complexity of conditions such as Alzheimer's, breast cancer, depression, cardiovascular disease, and epilepsy, this becomes even more crucial [1]. A brain tumor is a benign, abnormal growth found in or around the brain. A tumor is an abnormal mass of tissues that might be solid or filled with fluid. The tumor is formally called a neoplasm. Worldwide, there are about 18,000,000 new cases of cancer per year, with about 20,000 of those being brain tumors [2][3]. Most cases (102,260 or 34.4% of the total) and deaths (77,815 or 32.3% of the total) occurred in areas with very high human development index (HDI). Surgical navigation, clinical diagnosis, and radiation surgery are just a few examples of the numerous medical applications made possible by medical pictures, which have led to their extensive use in healthcare systems. This is where MRI and computed tomography (CT) come in as two of the most important medical images [4][5]. While magnetic resonance imaging (MRI) is superior at depicting the finer details of soft tissues, such as blood vessels, computed tomography (CT) scans reveal the exact locations of solid structures, including skeletal tissues. But to get a whole picture of a patient's condition, it's usually not enough to use just one diagnostic tool. One potential remedy to this issue is the fusion of CT and MRI images, which combines the best features of both imaging modalities [6]. The fusion technique allows for the efficient merging of CT and MRI images, which enhances the information and completeness of the merged image [7]. Clinicians can now possess a powerful tool that helps them attain greater diagnostic precision, assists them in real-world treatment planning and gauging, and in utilizing the synergistic combination of structural and soft tissue data.

Neuro-imaging has become a revolutionary tool of learning the structure and functioning of the brain and thus equally provides valuable insight into neurological and psychiatric disorders [8][9]. sMRI can be used to examine cortical thickness, the volume of gray matter, and the presence of anatomical abnormalities, whereas fMRI can be used to analyze dynamic brain activity and brain connections, thus allowing the exploration of functional and anatomical changes in a number of diseases [10][11][12]. Imaging genomics is the field that combines neuroimaging with genomic data to study the genetic, molecular, and phenotypic basis of brain function and dysfunction [13]. Data analytics methods that can glean useful insights from high-dimensional datasets are essential for dealing with the enormous computational and statistical difficulties posed by multimodal brain imaging data due to its complexity and size.

Deep learning, machine learning, and AI in general have emerged as effective tools for tackling these issues in the last several years [14][15]. Classification, detection, segmentation, and registration are just a few of the many medical imaging tasks for which DL models have shown outstanding performance, and the literature is full of examples [16][17]. Most frequent are classification, detection and segmentation that would allow the characterization of pathological regions, quantification of structural or functional abnormalities, and clinical decision support [18][19][20]. Use of deep learning with multimodal brain imaging data can combine complementary brain information that sMRI, fMRI and other brain imaging modalities provide, to increase diagnostic accuracy, prognosis and clarity of understanding complex brain conditions[21][22].

A. Motivation and Contribution

The increasing demand for automated diagnostic tools that are accurate, efficient, and well-suited to neurological illnesses like dementia is driving this work. Manually analysing brain tumour imaging data is a time-consuming and, frequently, inconsistently accurate method of diagnosis. With the increasing availability of Brain tumor multimodal image Datasets, there is a significant opportunity to apply advanced ML and DL techniques to extract meaningful patterns and improve diagnostic outcomes. This research is driven by the potential to leverage these technologies, especially the CNN model, to enhance early detection, support clinical decision-making, and ultimately contribute to better patient care and management. The study below makes a number of important contributions to the field:

- Developed a robust predictive framework using Brain tumor multimodal image Datasets for healthcare analytics.
- Implemented a complete data pipeline including image encryption-decryption, preprocessing, normalization, and feature extraction.
- Addressed class imbalance issues through data visualization and analysis of distribution patterns.
- Demonstrated the practical applicability of DL models in improving diagnostic accuracy in brain imaging analysis.
- Conceived of a CNN model specifically for the purpose of using MRI scans of the brain to identify different stages of dementia.
- Metrics like F1-score, memory, accuracy, and precision were used to carefully evaluate the model's performance.

B. *Justification and Novelty*

Utilizing brain tumor multimodal image data to aid in early dementia detection is crucial, and this study is being conducted because manual interpretation is time-consuming and not always accurate. A key innovation of the suggested method involves creating and testing a custom Convolutional Neural Network (CNN) model that excels with multimodal brain MRI data, outperforming standard models. Additionally, the integration of an image encryption-decryption mechanism ensures data security without compromising diagnostic quality a crucial aspect in medical imaging. This combination of high-accuracy classification and secure data handling represents a novel contribution toward building practical, AI-driven solutions for healthcare analytics.

C. *Structure of the paper*

The following is the paper's outline: It examines previous research and identifies areas where further study is needed in Section II. Section III details the datasets, preprocessing methods, and models that were employed. In Section IV, the results and a critical review are given. Section V presents our decision and suggestions for further study.

II. LITERATURE REVIEW

A comprehensive literature analysis was conducted to inform and enhance the creation of this work, which focusses on healthcare analytics using multimodal brain imaging data.

Wu *et al.* (2025) Despite CLIP's practical success in the real world, its medicinal applications remain underexplored. In order to tackle these obstacles, we explored three possible directions: 1) present a new CLIP variant that classifies brain and skin cancers using four convolutional neural networks (CNNs) and eight vein-invasive tumors (VITs) as image encoders. 2) To prevent data privacy breaches, integrate 12 deep models with two federated learning techniques. 3) To enhance the deep models' ability to generalize to unseen domain data, employ traditional machine learning (ML) methods. In the HAM10000 dataset, maxvit exhibits the greatest averaged (AVG) test metrics (AVG = 87.03%) with multimodal learning, while convnext_l shows remarkable test performance with an F1-score of 83.98% compared to swin_b's 81.33% in the FL model [23].

Chen *et al.* (2025) investigate and contrast single-modal and multimodal breast cancer prediction models using medical imaging modalities. Using 790 patients' medical imaging data, including 2,235 mammography images and 1,348 ultrasound images, an ideal model for constructing the multimodal classification model was identified. Examining and contrasting six distinct DL classifiers was the subsequent step. Several metrics, such as AUC, sensitivity, specificity, precision, and accuracy, were used to assess the performance of the multi- and single-modal classification models. According to the experimental data, the multimodal classification model achieves higher specificity (96.41% vs. 93.78% vs. 76.27% vs. 91.05% overall) than the single-modal models [24].

Asish *et al.* (2024) applied various ML algorithms, including kNN, RF, 1D-CNN-LSTM, and 2D-CNN, to classify a group of subjects based on their multi-modal features and the aspects of their grouping tests, including cross-subject, cross-session, and gender tests. They found that the RF classifier achieves the highest accuracy over 83% in the cross-subject test, around 68% to 78% in the cross-session test, and around 90% in the gender-based grouping test compared to other models. The extracted features and their SHAP analysis showed higher scores of the occipital, prefrontal areas of the brain, gaze angle, gaze origin, and head rotation features of the eye tracking [25].

Odusami *et al. (*2024) The goal is to find out how well machine learning can correctly group the different stages of Alzheimer's disease using a mix of neuroimaging data. We also use the Wilcoxon signed-rank test to statistically assess the accuracy ratings of the present models. The combined sensitivity in discriminating between NC and MCI was 83.77% (95% CI: 78.8, 87.7%), between AD and NC 94.60% (90.8%, 96.9%), and between pMCI and sMCI 80.41% (74.7%, 85.1%). The sensitivity to differentiate NC and mild cognitive impairment (MCI) was 83.77% with a 95% confidence range of 78.87% to 87.71%. A total of 94.60 sensitivity was obtained to differentiate between NC and Alzheimer disease (AD). Similarly, 86.41% could distinguish the progressive (pMCI) and stable (sMCI), and 86.63 could distinguish between NC and early moderate cognitive impairment (EMCI) (82.4332 89.95) [26].

Jansi *et al.* (2023) The greatest hindrance to the detection of brain tumors is its variability in case of location, structure, and arrangement of the tumor size. Using CNNs, data augmentation, and picture preparation, this study lays forth a comprehensive method for detecting brain tumors. This research makes use of preprocessed, enlarged, and erosion-enhanced multimodal MRI data from the Brats dataset, which includes brain tumors. By improving visibility of the tumor regions, dilation and erosion allow for more precise detection. The next step is to train a CNN model, prioritizing data shuffling for greater performance. The TensorFlow and Keras libraries played a crucial role in constructing the suggested system. Using the Brain Tumor Brats dataset, the suggested framework achieved a respectable detection accuracy of 98.2% in brain tumor identification [27].

Ghosh *et al.* (2023) GCNN is trained to distinguish between normal and schizophrenia by analyzing multimodal human brain connectomes. Specifically, construct structural connectivity graphs using diffusion tensor imaging data and functional connectivity graphs using functional magnetic resonance imaging data to train and evaluate a network. GCNN method is compared to various popular classification benchmarks, including one that is based on support vector machines. results demonstrate that the suggested graph convolution outperforms the alternatives in terms of F1 scores (0.75 for schizophrenia classification). A multimodal approach to diagnosing and predicting the course of mental illness may be possible using this approach [28].

The table I provides a summary of a few recent studies on multimodal brain imaging in healthcare analytics, with an emphasis on models, data used by the studies, the main results, and reported challenges by the research

TABLE I. OVERVIEW OF RECENT STUDIES ON PREDICTIVE MODELING OF MULTIMODAL BRAIN IMAGING DATA FOR HEALTHCARE ANALYTICS USING DEEP LEARNING

Author	Proposed Work	Dataset	Key Findings	Challenges/recommendations
Wu et al., (2025)	Introduced a novel CLIP variant with CNNs & ViTs for brain and skin cancer; combined 12 deep models with FL; integrated traditional ML for generalization.	HAM10000 dataset	MaxViT achieved highest AVG = 87.03% (multimodal learning); ConvNeXt_1 had best F1-score (83.98%) vs Swin_b (81.33%) in FL.	Limited exploration of CLIP in medical field; need stronger domain generalization & privacy-preserving methods.
Chen et al., (2025)	Models for predicting breast cancer using single-modal imaging versus multi-modal imaging (mammography + ultrasound).	790 individuals with a total of 2,235 mammography and 1,348 ultrasound scans	Specificity (96.41%), accuracy (93.78%), precision (83.66%), and area under the curve (0.968) were all better for the multimodal model than for the single-modal.	Sensitivity was better in single-modal models; multimodal models require bigger datasets and modal balancing.
Asish et al., (2024)	Applied ML methods (kNN, RF, 1D-CNN-LSTM, 2D-CNN) to classify distraction states from multimodal features.	Multimodal features incl. brain (EEG), eye tracking	RF achieved highest accuracy: >83% cross-subject, ~68–78% cross-session, ~90% gender-based grouping. SHAP showed brain (occipital, prefrontal) & gaze features were key.	Performance dropped in cross-session tests, indicating a need for robustness across sessions.
Odusami et al., (2024)	Statistical validation with the Wilcoxon test; a systematic review on ML for Alzheimer's disease staging utilizing multimodal neuroimaging.	Multimodal neuroimaging datasets (literature review)	The pooled sensitivity for MCI compared to NC is 83.77%, for AD it is 94.60%, for pMCI it is 80.41%, and for EMCI it is 86.63%. Joint specificity: AD against NC = 93.49 percent, etc.	Variability across studies; requires harmonization of datasets & standardized benchmarks.
Jansi et al., (2023)	Brain tumor detection via preprocessing, augmentation & CNN on multimodal MRI.	Brain Tumor Brats dataset	Achieved 98.2% accuracy; preprocessing (dilation, erosion) improved tumor region visibility.	Performance dependent on preprocessing quality; real-time clinical validation required.
Ghosh et al., (2023)	Used GCNN to classify schizophrenia from multimodal brain connectomes (DTI + fMRI).	Multimodal connectome data (DTI, fMRI)	GCNN outperformed SVM and benchmarks, achieving the best F1-score of 0.75.	Limited F1-score leaves room for improvement; larger sample sizes & robust models recommended.

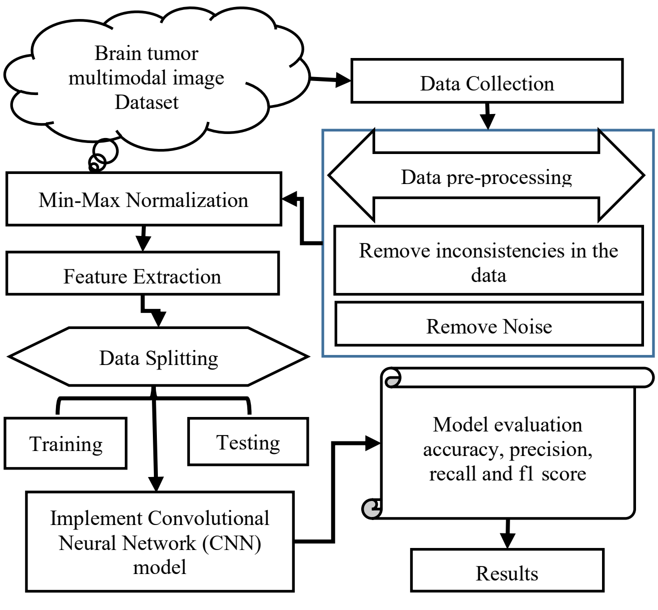

Fig. 1. Proposed flowchart for Multimodal Brain Imaging Data for Healthcare Analytics

III. RESEARCH METHODOLOGY

The proposed methodology of Focus on Multimodal Brain Imaging starts with the gathering of the Brain tumor multimodal image collection. The preprocessing steps covered data cleansing, elimination of inconsistency and noise, scaling and min-max normalization and feature extraction. The data were then partitioned into training and testing in order to effectively learn and assess the performance of the training. The last step was to train the proposed CNN model on the cleaned data. This model excels at managing time-dependent dependencies and patterns, and it successfully classified the stages of dementia. The next step was to evaluate the model's performance using industry-standard metrics including F1-score, recall, accuracy, and precision. This would ensure that healthcare analytics produced correct predictions and classifications. The entire steps involved are presented in Figure 1.

Each step of the proposed flowchart for using multimodal brain imaging in healthcare is explained in detail below.

A. Data collection

The dataset includes a variety of medical pictures, including MRI and CT scans, which are used for the detection and study of brain tumors. Brain tumors cause a wide variety of structural and functional alterations, and the combined pictures from the two modalities reveal a great deal about these changes. Each image in the collection is anatomically labelled with the type of tumor (e.g., glioma, meningioma, etc.) and its location in the brain. The dataset consists of high-resolution CT and MRI scans taken from various individuals. For the experiment, they manipulated 50 sets of greyscale images and 256×256 scaled images from CT and MRI scans. One example is the test set shown in Figure 2, which consists of 50 image pairs. Data visualizations such as bar plots were used to examine data distribution, feature correlations etc., are given below:

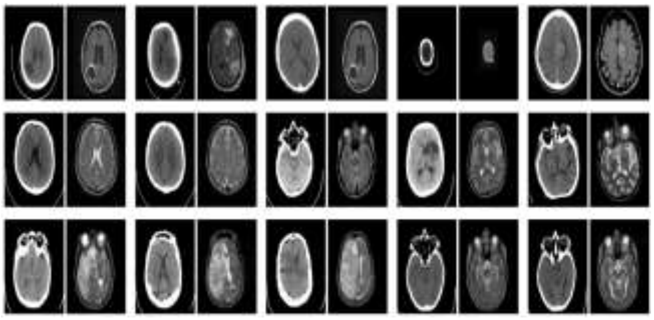

Fig. 2. Fifteen pairs of CT-MRI images for evaluation. In each pair, the left is the CT image, and the right is the MRI image

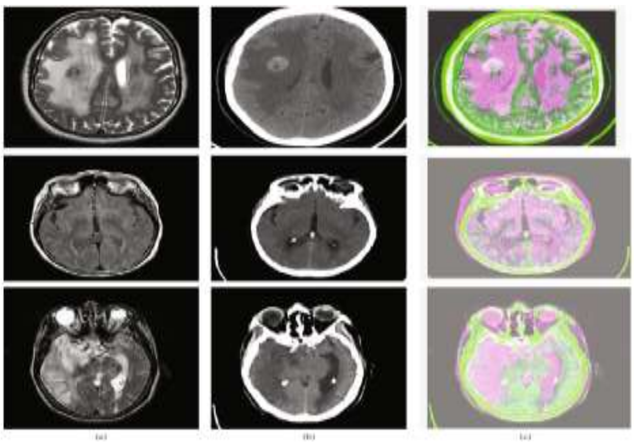

Fig. 3. MRI and CT Brain of size 256 × 256 pixels and its fused images. (a) Input-MRI image. (b) Input-CT image. (c) Fused MRI-CT image

Figure 3 illustrates brain imaging using MRI and CT modalities along with their fused counterparts, each resized to 256 × 256 pixels. Column (a) shows the input MRI images, which provide detailed soft tissue contrast useful for visualizing brain structures and abnormalities. Column (b) presents the corresponding CT images, which highlight bone structures and calcifications with high spatial resolution. Column (c) depicts the fused MRI-CT images, where information from both modalities is integrated, enhancing diagnostic interpretation by combining the anatomical clarity of CT with the soft tissue detail of MRI. This fusion facilitates comprehensive analysis for improved medical decision-making.

Figure 4 the image sequence demonstrates the process of image encryption and decryption applied to a brain CT and MRI scans. The first panel shows the original image, clearly depicting the structural details of the brain. The middle panel presents the encrypted image, where the original features are completely obscured, ensuring data confidentiality and protection against unauthorized access. The final panel displays the decrypted image, which accurately reconstructs the original scan, preserving the anatomical integrity. This figure illustrates the effective use of the encryption-decryption method to secure important medical imaging information without degrading image quality.

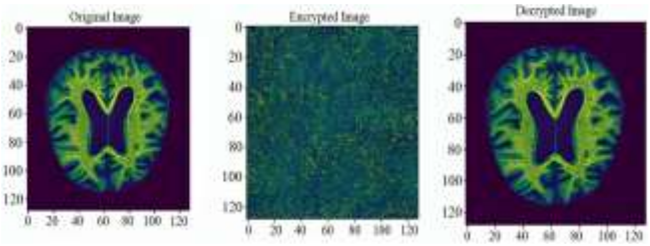

Fig. 4. Encryption and decryption of images

B. Data pre-processing

Data preparation included the organization of brain tumor multimodal images, combining their data and cleansing to maintain similarity. Relevant characterics were then discovered that were to facilitate handy analysis. Any form of data was pre-processed to eliminate noise and inconsistency to ensure quality of data. Model training was based on the standardization of the data through normalization and data transformation in this phase. The key features of the pre-processing step are listed below:

- **Data exploration.** The dataset sample size, the distribution of features, and analysis add some insights into the inner nature of the dataset and establish a good basis to choose preprocessing techniques.
- **Remove inconsistencies in the data**: To avoid the lack of consistency, focus on correcting errors, standardizing the formats, and enforcing data governance. This involves data validation, data cleaning, data transformation and so on.
- **Remove Noise:** The Gaussian blur filter was used to remove noise in photos and to enhance the output of a better quality. This kind of filtering actually lightens up the picture and dims down notable detail. Images adjusted with high-pass filter to sharpen image and complex features are retrieved. This filter makes edges and smaller details sharper and enables easier discernment of the significant details in an image.

C. Min-Max Normalization

The data were put in normalized form by applying the min-max method of keeping values within a range between 0 and 1. The aim of such an action was to get the classifiers to perform better and to reduce the influence of outliers. In order to standardise, the following formula was used (1):

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

X indicates the initial feature value, X' stands for the normalized value, X_{min} denotes the feature's minimum, and X_{max} denotes its maximum.

D. Feature Extraction

The feature extraction technique was binary thresholding. In this approach, images in grey scale can be represented in binary form (with the foreground pixels assigned to white and the background pixels in black) given a hackneyed size set by a user. This aids the model in differentiating between tumor kinds and healthy parts by drawing attention to the object's outline, in this case the tumor regions [29]. Employed contour detection algorithms to determine the largest contour in each image- the tumor region- through thresholding, and then cropped the image to highlight that location within the region of interest (ROI) [30]. Due to this, there is the model capable of concentrating on the essential sections of the picture without the need to incorporate excessive background. In the case of input size homogeneity, generated photos were trimmed and reduced to a normalized size of 256 X 256 pixels.

E. Data Splitting

Training and testing sets were split so that we could assess the performance of the model. 70 % of the data was used to train the model and calculate model parameters. The remaining 30% was not dedicated to training purposes but reserved to test the model performance and estimation of parameters.

F. Proposed Recurrent Neural Network (CNN) model

Deep learning has greatly changed the rate at which the doctors will be able to diagnose diseases in an accurate and efficient manner by improving the processing of images using CNNs. CNNs have helped tremendously in medical image processing and diagnosis [31]. The several tests that have been done on CNNs include but are not limited to object segmentation, object detection, and image categorizing which have been reported that CNNs are better than conventional CAD systems. A major advantage compared to more traditional machine learning techniques is that CNNs can learn complex picture features without the need to engineer feature extraction, which requires the assistance of a human being. CNNs have been applied in several diagnostic imaging modalities in the field of medicine such as in MRI and CT. With the adaptive use of CNNs, radiologists and doctors can make more accurate and expedited diagnoses since the accuracy levels of these variables are impressive when used to interpret medical images. In order to learn features, convolutional neural networks (CNNs) use a multi-layered architecture that incorporates non-linear transformations. The input data is displayed at the visible layer as a tensor, which is a multidimensional data array. Examples of this type of topology include time-series data, which is essentially a 1D grid sampling at regular intervals, 2D picture pixels, 3D video structure, etc. The next step involves extracting multiple abstract characteristics through a series of hidden layers. For example, using equation (2), a two-dimensional kernel h can calculate the 2D convolution given an input x that is two dimensions in size:

$$(x * h)i,j = x[i,j] * h[i,j] = \sum n \sum m x[n,m] \cdot h[i-n][j-m] \quad (2)$$

their individual weights multiplied with regard to a small input area to which they are linked.

A feature map is created at the filter's output by including a bias term and applying a point-wise nonlinearity g following the convolution. The filters are defined by the input x, the bias bl, and the coefficients or weights, given a convolutional layer and a l-th feature map h^l. Equation (3) can be used to obtain the feature map h^l from W^l.:

$$h_{i,j}^l = g(W^l * x)_{ij} + b_l \quad (3)$$

where $g(\cdot)$ is the activation function and $*$ is the 2D convolution defined by Equation (3).

Deep neural networks often use the rectifier activation function, which is defined as

$$g(x) = x^+ = max(0,x) \quad (4)$$

G. Evaluation metrics

TP, FP, TN, and FN are the four classifications used in AI diagnostic tasks. It is common practice to compute performance evaluation measures for binary classification tasks using the confusion matrix. Positive cases, or the positive class that was accurately classified as positive, are represented by TP [32]. A negative instance is represented by TN if and only if the negative class was accurately identified as negative. The term "false positive" (FP) describes cases where the negative class was mistakenly classified as positive. When members of the positive class are mistakenly classified as negative, this is known as a false negative (FN). Important evaluation measures such as recall, accuracy, precision, and F1-score were calculated using these data, as displayed in the matrix below:

Accuracy: A metric that compares the trained model's output predictions to those of the entire dataset (input samples). 5 is the provided value-

$$Accuracy = \frac{\text{TP+TN}}{\text{TP+Fp+TN+FN}} \quad (5)$$

Precision: Precision measures how well a model predicts positive occurrences relative to all positive occurrences. Precision indicates. How good the classifier is in predicting the positive classes is expressed as (6)-

$$Precision = \frac{\text{TP}}{\text{TP+FP}} \quad (6)$$

Recall: This metric measures the accuracy of positive event predictions as a percentage of the total number of instances that ought to have been positive. The formula for it in mathematics is (7)-

$$Recall = \frac{\text{TP}}{TP+FN} \quad (7)$$

F1 score: In other words, it aids in maintaining a healthy equilibrium between recall and precision by combining the two. Its range is [0, 1]. Mathematically, it is given as (8)-

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (8)$$

IV. Results And Discussion

This section provides an overview of the experimental setup and presents the results of the suggested model's training and testing. Here is the experimental setup used in this article: The specifications include Windows 10 OS, NVIDIA RTX 2080 graphics processing unit, PyTorch machine learning framework, Python 3.10, and an experimental environment created in PyCharm. Results from experiments using the suggested CNN model for healthcare analytics, including brain tumour multimodal images, show exceptional performance on all assessment criteria (Table II). Impressively, the model was able to minimise false positives, as evidenced by its 99% precision and 99.28% accuracy. Recalling genuine cases with a sensitivity of 99.67% and an F1-score of 99.12% respectively, it shows a good trade-off between the two. The CNN model's reliability and resilience in analysing brain tumour multimodal image datasets for healthcare applications are demonstrated by these results.

TABLE II. EXPERIMENT RESULTS OF CNN FOR BRAIN TUMOR MULTIMODAL IMAGE DATASET

Performance matrix	CNN
Accuracy	99.28
Precision	99
Recall	99.67
F1-score	99.12

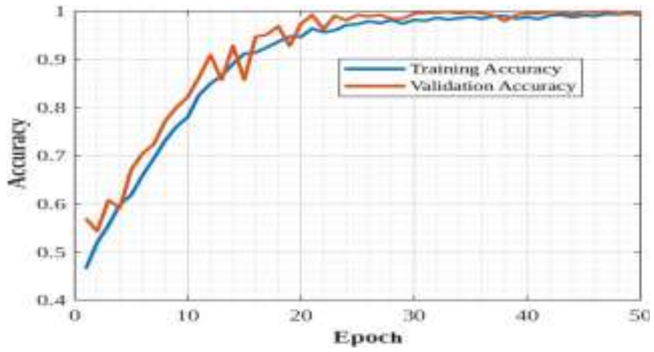

Fig. 5. Accuracy curves for the CNN Model

Figure 5 shows tendencies in the model's validation and training accuracy throughout 50 epochs. The number of epochs is one variable, while the accuracy—which can range from 0.4 to 1.0—is another. The blue line represents validation accuracy, which begins at around 0.55 and has early variability; the orange line represents training accuracy, which begins at about 0.48 and grows with each epoch. Both curves show rapid improvement in the first 10–15 epochs. After about epoch 20, both training and validation accuracies converge and begin to plateau, reaching near-perfect accuracy (0.99) by epoch 30 and maintaining stability through epoch 50.

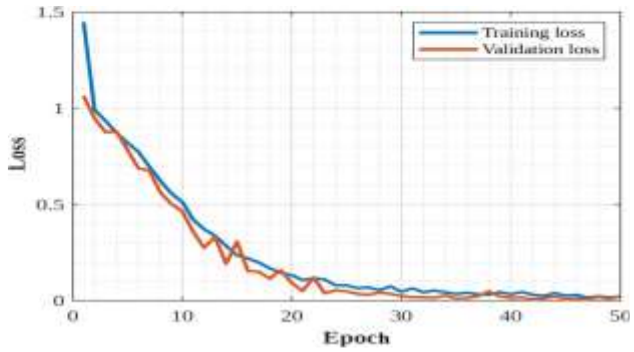

Fig. 6. Loss curves for the CNN Model

Figure 6 displays the 50-epoch training and validation loss curves. On one side, we have the number of epochs, and on the other, we have the amount of loss, which might range from zero to one and a half. Initially, both the training loss (blue line) and the validation loss (orange line) start at levels slightly below 1.4. Both curves exhibit a sharp decline during the first 10 epochs, indicating rapid model learning. The losses continue to decrease steadily and begin to plateau after epoch 20, approaching near-zero values around epoch 40. This close alignment, combined with the consistently decreasing trend, suggests effective model convergence with minimal overfitting or underfitting. The low final loss values indicate high model accuracy and stability.

A. Comparative analysis

A comparative accuracy examination was carried out against other existing models to illustrate the efficacy of the suggested CNN model. Table III presents the experimental evaluation of the models, demonstrating that traditional deep learning architectures and advanced neural networks exhibit varying performance levels. ReportGuidedNet achieved an accuracy of 70.3%, while the MLP model slightly improved the results with 72.69%. A significant jump in performance was observed with ResNet18, which reached 96.0% accuracy, highlighting the effectiveness of residual connections in extracting discriminative features. Among all, the CNN model outperformed the others with the highest accuracy of 99.28%, indicating its strong capability in handling sequential dependencies and delivering superior classification performance.

TABLE III. ACCURACY COMPARISON OF DIFFERENT PREDICTIVE MODELS, FOR BRAIN IMAGING IN HEALTHCARE

Models	Accuracy
ReportGuidedNet [33]	70.3%
MLP[34]	72.69
ResNet18[35]	96.00
CNN	99.28

The proposed CNN model demonstrates a clear advantage by achieving an exceptional accuracy of 99.28%, highlighting its superior capability in classifying dementia stages from brain tumor imaging data. The extremely high accuracy is an indication of the efficiency of this model in understanding complex patterns and temporal dependencies, and therefore, would be highly dependable in healthcare analytics and clinical practice in the real world.

V. CONCLUSION AND FUTURE STUDY

Multimodal medical imaging is central in the realm of clinical diagnosis and research, since it integrates data of multiple imaging modalities into a single, more detailed picture of the pathology. Multimodal fusion techniques based on the deep learning concept have recently emerged as effective ways to enhance medical image classification. This review provides an analytical overview of advances in the field of multimodal fusion using deep learning approaches to medical classification problems. The performance of DL in Brain tumour multimodal image Datasets as a healthcare analytics tool is evaluated. The CNN model, was found to give the best result with a 99.28% accuracy. The conventional models, such as MLP and ReportGuidedNet, demonstrated relatively smaller accuracies, implying their limited capability to learn high-dimensional information of medical images. The results overall show that CNN-based models are highly effective in predictive healthcare applications involving brain imaging, and hold promising potential as to accurately and early predict neurological conditions. Future research may investigate the transferability of these computational models to other medical data sets and assess their distribution in clinical conditions. An additional potential direction may be investigating encrypted medical images as a further step to enforcing privacy and security provisions.

REFERENCES

[1] S. Anto, "Supervised Machine Learning Approaches for Medical Data Set Classification - A Review," *Int. J. Comput. Sci. Technol.*, vol. 2, no. 4, pp. 234–240, 2011, [Online]. Available: http://www.ijcst.com/vol24/2/santo.pdf

[2] M. A. Mostafiz, "Machine Learning for Early Cancer Detection and Classification : AI- Based Medical Imaging Analysis in Healthcare," *Int. J. Curr. Eng. Technol.*, vol. 15, no. 3, pp. 251–260, 2025, doi: https://doi.org/10.14741/ijcet/v.15.3.7.

[3] R. P. Mahajan, "Transfer Learning for MRI image reconstruction: Enhancing model performance with pretrained networks," *Int. J.*

Sci. Res. Arch., vol. 15, no. 1, pp. 298–309, Apr. 2025, doi: 10.30574/ijsra.2025.15.1.0939.

[4] Q. Mao, W. Zhai, X. Lei, Z. Wang, and Y. Liang, "CT and MRI Image Fusion via Coupled Feature-Learning GAN," *Electronics*, vol. 13, no. 17, p. 3491, Sep. 2024, doi: 10.3390/electronics13173491.

[5] S. S. S. Neeli, "Key Challenges and Strategies in Managing Databases for Data Science and Machine Learning," *Int. J. Lead. Res. Publ.*, vol. 2, no. 3, pp. 1–9, 2021, doi: https://doi.org/10.5281/zenodo.14672937.

[6] S. R. Sagili, S. B, P. RVS, and A. P, "Sentiment Classification for Depression Detection: Integrating Capsule Networks with CNNs on Review Data," in *2025 Emerging Technologies for Intelligent Systems (ETIS)*, 2025, pp. 1–7. doi: 10.1109/ETIS64005.2025.10961685.

[7] R. Dattangire, D. Biradar, and A. Joon, "AI-Enhanced U-Net for Accurate Low-Grade Glioma Segmentation in Brain MRI: Transforming Healthcare Imaging," in *2024 3rd International Conference on Electrical, Electronics, Information and Communication Technologies, ICEEICT 2024*, 2024, pp. 1–6. doi: 10.1109/ICEEICT61591.2024.10718440.

[8] K. M. Kuntzelman, J. M. Williams, P. C. Lim, A. Samal, P. K. Rao, and M. R. Johnson, "Deep-Learning-Based Multivariate Pattern Analysis (dMVPA): A Tutorial and a Toolbox," *Front. Hum. Neurosci.*, vol. 15, pp. 1–18, Mar. 2021, doi: 10.3389/fnhum.2021.638052.

[9] S. Pandya, "Integrating Smart IoT and AI-Enhanced Systems for Predictive Diagnostics Disease in Healthcare," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 10, no. 6, pp. 2093–2105, Dec. 2023, doi: 10.32628/CSEIT2410612406.

[10] M. W. Weiner *et al.*, "Recent publications from the Alzheimer's Disease Neuroimaging Initiative: Reviewing progress toward improved AD clinical trials," 2017. doi: 10.1016/j.jalz.2016.11.007.

[11] P. M. Thompson *et al.*, "ENIGMA and global neuroscience: A decade of large-scale studies of the brain in health and disease across more than 40 countries," 2020. doi: 10.1038/s41398-020-0705-1.

[12] R. P. Mahajan, "Optimizing Pneumonia Identification in Chest X-Rays Using Deep Learning Pre-Trained Architecture for Image Reconstruction in Medical Imaging," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 5, no. 1, pp. 52–63, Apr. 2025, doi: 10.48175/ijarsct-24808.

[13] N. Patel, "Quantum Cryptography In Healthcare Information Systems: Enhancing Security In Medical Data Storage And Communication," *J. Emerg. Technol. Innov. Res.*, vol. 9, no. 8, pp. g193–g202, 2022.

[14] L. Shen and P. M. Thompson, "Brain Imaging Genomics: Integrated Analysis and Machine Learning," *Proc. IEEE*, vol. 108, no. 1, pp. 125–162, 2020, doi: 10.1109/JPROC.2019.2947272.

[15] R. Q. Majumder, "Machine Learning for Predictive Analytics: Trends and Future Directions," *Int. J. Innov. Sci. Res. Technol.*, vol. 10, no. 04, pp. 3557–3564, 2025.

[16] S. R. Sagili, S. Chidambaranathan, N. Nallametti, H. M. Bodele, L. Raja, and P. G. Gayathri, "NeuroPCA: Enhancing Alzheimer's disorder Disease Detection through Optimized Feature Reduction and Machine Learning," in *2024 Third International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*, IEEE, Jul. 2024, pp. 1–9. doi: 10.1109/ICEEICT61591.2024.10718628.

[17] R. Dattangire, R. Vaidya, D. Biradar, and A. Joon, "Exploring the Tangible Impact of Artificial Intelligence and Machine Learning: Bridging the Gap between Hype and Reality," in *2024 1st International Conference on Advanced Computing and Emerging Technologies (ACET)*, IEEE, Aug. 2024, pp. 1–6. doi: 10.1109/ACET61898.2024.10730334.

[18] H. T. Shen *et al.*, "Heterogeneous data fusion for predicting mild cognitive impairment conversion," *Inf. Fusion*, vol. 66, pp. 54–63, 2021, doi: 10.1016/j.inffus.2020.08.023.

[19] Y. Zhu, M. Kim, X. Zhu, D. Kaufer, and G. Wu, "Long range early diagnosis of Alzheimer's disease using longitudinal MR imaging data," *Med. Image Anal.*, vol. 67, pp. 1–12, 2021, doi: 10.1016/j.media.2020.101825.

[20] S. P. Kalava, "The Role of AI in Reinventing Hospitality Safety Measures After COVID-19," *Sci. Res. Community*, no. 2754–6659, p. 3, 2024.

[21] X. Zhu *et al.*, "Joint prediction and time estimation of COVID-19 developing severe symptoms using chest CT scan," *Med. Image Anal.*, 2021, doi: 10.1016/j.media.2020.101824.

[22] E. Sugawara and H. Nikaido, "Properties of AdeABC and AdeIJK efflux systems of Acinetobacter baumannii compared with those of the AcrAB-TolC system of Escherichia coli," *Antimicrob. Agents Chemother.*, vol. 58, no. 12, pp. 7250–7257, 2014, doi: 10.1128/AAC.03728-14.

[23] Y. Wu, M. Owais, R. Kateb, and A. Chaddad, "Deep Modeling and Optimization of Medical Image Classification," in *2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI)*, IEEE, Apr. 2025, pp. 1–4. doi: 10.1109/ISBI60581.2025.10981184.

[24] J. Chen *et al.*, "A deep learning-based multimodal medical imaging model for breast cancer screening," *Sci. Rep.*, vol. 15, no. 1, pp. 1–13, 2025, doi: 10.1038/s41598-025-99535-2.

[25] S. M. Asish, A. K. Kulshreshth, C. W. Borst, and S. Sutradhar, "Classification of Internal and External Distractions in an Educational VR Environment Using Multimodal Features," *IEEE Trans. Vis. Comput. Graph.*, vol. 30, no. 11, pp. 7332–7342, 2024, doi: 10.1109/TVCG.2024.3456207.

[26] M. Odusami, R. Maskeliūnas, R. Damaševičius, and S. Misra, "Machine learning with multimodal neuroimaging data to classify stages of Alzheimer's disease: a systematic review and meta-analysis," 2024. doi: 10.1007/s11571-023-09993-5.

[27] R. Jansi, S. Kowsalya, S. Seetha, and A. Yogadharshini, "A Deep Learning based Brain Tumour Detection using Multimodal MRI Images," in *2nd International Conference on Automation, Computing and Renewable Systems, ICACRS 2023 - Proceedings*, 2023, pp. 582–587. doi: 10.1109/ICACRS58579.2023.10404952.

[28] S. Ghosh, E. Bhargava, C.-T. Lin, and S. S. Nagarajan, "Graph Convolutional Learning of Multimodal Brain Connectome Data for Schizophrenia Classification," in *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, IEEE, Apr. 2023, pp. 1–5. doi: 10.1109/ISBI53787.2023.10230441.

[29] R. Disci, F. Gurcan, and A. Soylu, "Advanced Brain Tumor Classification in MR Images Using Transfer Learning and Pre-Trained Deep CNN Models," *Cancers (Basel).*, vol. 17, no. 1, p. 121, Jan. 2025, doi: 10.3390/cancers17010121.

[30] S. Pandya, "Predictive Modeling for Cancer Detection Based on Machine Learning Algorithms and AI in the Healthcare Sector," *Int. Res. J.*, vol. 11, no. 12, pp. 549–555, 2024.

[31] A. M. Al-Zoghby, A. Ismail Ebada, A. S. Saleh, M. Abdelhay, and W. A. Awad, "A Comprehensive Review of Multimodal Deep Learning for Enhanced Medical Diagnostics," *Comput. Mater. Contin.*, vol. 84, no. 3, pp. 4155–4193, 2025, doi: 10.32604/cmc.2025.065571.

[32] X. Xu *et al.*, "A Comprehensive Review on Synergy of Multi-Modal Data and AI Technologies in Medical Diagnosis," *Bioengineering*, vol. 11, no. 3, pp. 1–51, 2024, doi: 10.3390/bioengineering11030219.

[33] L. Dai *et al.*, "Boosting Deep Learning for Interpretable Brain MRI Lesion Detection through the Integration of Radiology Report Information," *Radiol. Artif. Intell.*, vol. 6, no. 6, 2024, doi: 10.1148/ryai.230520.

[34] L. Cai *et al.*, "MM-GTUNets: Unified Multi-Modal Graph Deep Learning for Brain Disorders Prediction," *IEEE Trans. Med. Imaging*, vol. 44, no. 9, pp. 3705–3716, 2025, doi: 10.1109/TMI.2025.3556420.

[35] C. Macfadyen, A. Duraiswamy, and D. Harris-Birtill, "Classification of hyper-scale multimodal imaging datasets," *PLOS Digit. Heal.*, vol. 2, no. 12, pp. 1–15, 2023, doi: 10.1371/journal.pdig.0000191.

