A Deep Learning-Based Brain Tumor Classification System Using MRI Images from Bangladesh

Ajaj Anwar Anik [1], Durjoy Kumer Ghosh [2], Mahfuzur Rahman [3], Mossa. Samsur Nahar Setu [4]

Department of Computer Science and Engineering, Faculty of Science and Technology, City University, Bangladesh, Khagan, Savar, Dhaka, Bangladesh

Corresponding author, Email: ajaj.anik180@gmail.com

Abstract— This paper presents a deep learning-based system for brain tumor classification using Magnetic Resonance Imaging (MRI) images from a Bangladeshi dataset. Leveraging transfer learning with a DenseNet-121 architecture, the system is designed to accurately classify brain tumors into four distinct categories: Glioma, Meningioma, Pituitary, and Normal. The methodology involved comprehensive data preparation, including image resizing, augmentation, and careful splitting into training, validation, and test sets. The model was trained using the Adam optimizer and categorical cross-entropy loss, achieving a high-test accuracy of 96.52% and robust performance across various metrics, including class-wise precision, recall, F1-score, specificity, and micro/macro-averaged ROC-AUC and PR-AUC scores. To enhance transparency and trustworthiness, explainability techniques such as Grad-CAM, Grad-CAM++, LayerCAM, and ScoreCAM were implemented to visualize the critical regions in MRI scans that influenced the model's predictions. These visualizations provided valuable insights into the model's decision-making process, confirming its focus on medically relevant features.

Keywords — Deep Learning, Brain Tumor Classification, Magnetic Resonance Imaging (MRI), Image Classification, Medical Imaging, DenseNet-121, Grad-CAM.

I. INTRODUCTION

Brain tumors are among the most critical and life-threatening neurological disorders, characterized by abnormal and uncontrolled cell growth within the brain. Accurate and early diagnosis is essential for effective treatment planning and improving patient survival rates. Among various imaging modalities, Magnetic Resonance Imaging (MRI) is widely recognized as the most reliable and non-invasive technique for brain tumor detection due to its superior soft tissue contrast and high spatial resolution [3]. However, manual interpretation of MRI scans is often time-consuming, subjective, and dependent on the expertise of radiologists, which may lead to diagnostic variability.

In recent years, deep learning (DL), particularly Convolutional Neural Networks (CNNs), has revolutionized medical image analysis by enabling automatic feature extraction and high-accuracy classification. Numerous studies have demonstrated the effectiveness of DL-based systems in brain tumor classification tasks, achieving significant improvements over traditional machine learning approaches [1], [2]. These systems are capable of classifying tumors into multiple categories such as glioma, meningioma, and pituitary tumors with high reliability.

Transfer learning has further enhanced the applicability of deep learning in medical imaging, especially when dealing with limited datasets. By leveraging pre-trained architectures such as DenseNet, ResNet, and EfficientNet, models can achieve improved performance while reducing training time and computational cost

[4]. Additionally, recent advancements include hybrid and attention-based architecture that improve feature representation and robustness in classification tasks [5].

Despite these advancements, several challenges persist. One major issue is the lack of region-specific datasets, particularly in developing countries like Bangladesh. Models trained on global datasets may not generalize well to local populations due to variations in imaging protocols and demographic characteristics. Furthermore, the black-box nature of deep learning models raises concerns regarding interpretability and clinical trust. To address this, explainable AI (XAI) techniques such as Grad-CAM and its variants have been increasingly used to visualize model decisions and highlight clinically relevant regions in MRI scans [6].

In this study, we propose a deep learning-based brain tumor classification system using MRI images from a Bangladeshi dataset. The proposed framework utilizes transfer learning with a DenseNet-121 architecture to classify brain tumors into four categories: glioma, meningioma, pituitary tumor, and normal. Additionally, multiple explainability techniques are incorporated to enhance model transparency and reliability. By focusing on a region-specific dataset and integrating high-performance classification with interpretability, this work aims to contribute to the development of reliable AI-assisted diagnostic tools for brain tumor detection in Bangladesh.

II. Background and Study Motivation

Brain tumor diagnosis remains a complex and critical task in modern healthcare due to the heterogeneous nature of tumors and their varying appearances in Magnetic Resonance Imaging (MRI). Although MRI provides high-resolution structural information, accurate interpretation requires significant expertise and is often subject to variability among radiologists. This limitation has driven the adoption of automated diagnostic systems based on artificial intelligence to enhance accuracy and efficiency.

Recent advancements in deep learning (DL) have significantly improved the performance of brain tumor classification systems. In particular, hybrid architectures combining Convolutional Neural Networks (CNNs) with Vision Transformers (ViTs) have demonstrated superior capability in capturing both local and global features from MRI images. A meta-ensemble framework proposed in 2025 achieved high robustness and classification accuracy by integrating multiple DL models, highlighting the effectiveness of ensemble-based approaches [7]. Similarly, optimized hybrid deep learning models incorporating explainable AI mechanisms have shown improved generalization and reliability in tumor classification tasks [8].

Despite achieving high accuracy, conventional DL models often lack interpretability, which limits their acceptance in clinical practice. To address this issue, Explainable Artificial Intelligence (XAI) techniques have been integrated into diagnostic frameworks to provide visual explanations of model predictions. Methods such as Grad-CAM and Local Interpretable Model-Agnostic Explanations (LIME) have been successfully used to highlight tumor-relevant regions in MRI scans, enabling clinicians to better understand and trust AI-based decisions [9], [10]. Furthermore, recent studies emphasize that combining interpretability with high-performance models significantly enhances clinical applicability and decision support [11].

Another critical limitation in existing research is the lack of region-specific datasets. Most DL-based brain tumor classification systems are trained on publicly available datasets that primarily represent populations from developed countries. These datasets may not capture variations in imaging protocols, genetic factors, and demographic characteristics present in regions such as Bangladesh. As a result, models trained on such datasets may exhibit reduced generalization when applied in local clinical environments. This highlights the necessity of developing AI models trained on indigenous datasets to ensure practical relevance and deployment feasibility.

The primary motivation of this study is the lack of region-specific brain tumor classification models for Bangladesh. Most existing deep learning models are trained on datasets from developed countries, which limits their effectiveness when applied to Bangladeshi MRI data due to differences in imaging conditions, patient demographics, and tumor characteristics. To address this gap, this study develops a model using a locally collected MRI dataset from multiple hospitals, ensuring better real-world applicability.

Another key motivation is the need for model interpretability. High accuracy alone is insufficient for clinical use, as medical professionals require clear explanations of AI decisions. This study integrates explainability techniques such as Grad-CAM and LIME to provide visual insights, helping clinicians understand and trust the model's predictions.

Data scarcity is also a major challenge in developing countries. To overcome this, the study employs transfer learning with DenseNet-121, allowing the model to achieve strong performance even with limited annotated data.

III. LITERATURE REVIEW

Recent advancements in smart grid fault detection and medical imaging alike have increasingly leveraged deep learning (DL), machine learning (ML), and hybrid signal processing techniques to improve detection accuracy, robustness, and real-time applicability. Brain tumor diagnosis remains a complex and critical task due to the heterogeneous nature of tumors and their varying appearances in MRI. Recent advancements in deep learning have significantly improved the performance of brain tumor classification systems. Despite achieving high accuracy, conventional DL models often lack interpretability. Methods such as Grad-CAM and LIME have been successfully used to highlight tumor-relevant regions in MRI scans [9], [10]. This research develops a highly accurate, deep learning-based diagnostic system using a DenseNet-121 architecture that classifies Bangladeshi brain MRI scans into four distinct categories (Glioma, Meningioma, Pituitary, or Normal) with 96.52% accuracy. It uniquely overcomes the "black box" limitation of traditional AI by integrating four Explainable AI techniques (Grad-CAM, Grad-CAM++, LayerCAM, and ScoreCAM) to generate visual heatmaps, allowing radiologists to verify that the model's decisions are based on actual medical anomalies rather than background noise.

TABLE 1. Comparative analysis of the proposed framework with recent literature

Ref.	Method or Framework	Outcome	Limitations
[12]	Multiscale Deep CNN cascaded with LSTM	Accuracy = 89.5%	Outputs depend on arbitrary mode compositions; Model must run across all possible combinations; Limited by dataset source
[13]	ResNet-50 architecture with added classification head	Accuracy = 92% Precision = 94%	Limited to binary classification (tumor vs. no tumor) Dataset size moderate (3847 images) Generalization to multi-class tumor types not tested

[14]	Custom 2D CNN with integration of handcrafted features (LBP, DWT)	Accuracy = 81.11% (LBP), 94.11% (DWT), 94% (raw MRI with ResNet)	Accuracy still lower with LBP features; Evaluation limited to Br35H dataset Generalization to other datasets/modalities (PET, SPECT, Biopsy) not tested
[15]	Machine learning classifiers (Logistic Regression, others); Hybrid ensemble technique; 10-fold cross-validation	Accuracy = 89% (small dataset), 87% (large dataset	Manual feature extraction required (time-consuming, risk of poor selection) Performance limited compared to deep learning; Dataset size influenced results
[16]	CFLM (Federated GoogLeNet)	94.4% Accuracy	Dataset Bias and Generalization Issues
[17]	CNN+region focusedXAI	80% Accuracy	Interpretability-first analysis
[18]	Ensemble (VGG16, Dense121, IRv2)	94.8% Accuracy	Robust ensemble, Detailed saliency analysis.

IV. METHODOLOGY

A. Dataset Description

This study employs a publicly available Bangladeshi Brain Cancer MRI Dataset consisting of 1600 raw MRI images (400 per class), which are augmented to 6000 images across four categories: glioma, meningioma, pituitary tumor, and no tumor. For model training, all images were resized to 224 × 224 × 3 to satisfy the input requirements of the DenseNet-121 architecture and to ensure compatibility with pretrained ImageNet weights [12].

B. Architecture of the Proposed Framework

The architectural design represents a sophisticated pipeline that transforms raw medical data into a clinical diagnosis. It begins by standardizing the input image to a specific resolution. Once resized, the image enters the DenseNet-121 core, which acts as a powerful feature extractor.

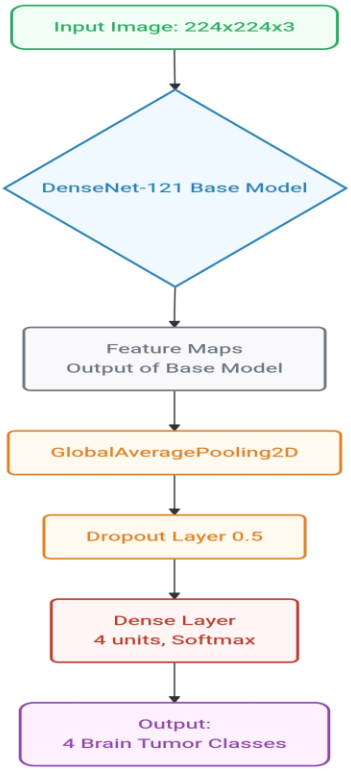

Fig. 1 Proposed Model Architecture

C. Architectural Components and Symbolic Logic

The structural hierarchy of the network is defined by specific functional blocks:
• Input: Processes MRI scans as 224x224x3 tensors.
• Feature Extraction: Utilizes a DenseNet-121 base model.
• Intermediate Output: Generates Feature Maps.
• Dimensionality Reduction: Applies GlobalAveragePooling2D.
• Regularization: Implements a Dropout Layer (0.5).
• Classification Head: A Dense Layer with 4 units and SoftMax activation.
• Terminal Output: Produces final diagnosis across 4 Brain Tumor Classes.

V. RESULT

A. System Functionality

The developed system is an automated pipeline for the classification of brain tumors from MRI scans. It handles data ingestion, dynamic 80/10/10 splitting for training, validation, and testing, and real-time image

augmentation. The model removes the original top layer of DenseNet-121 and substitutes a custom classification head.

B. Learning Curve Analysis

The training and validation process of the DenseNet-121 model was meticulously monitored through accuracy and loss curves. The Accuracy Curve illustrates the model's performance on both datasets across epochs, while the Loss Curve depicts the reduction in prediction error.

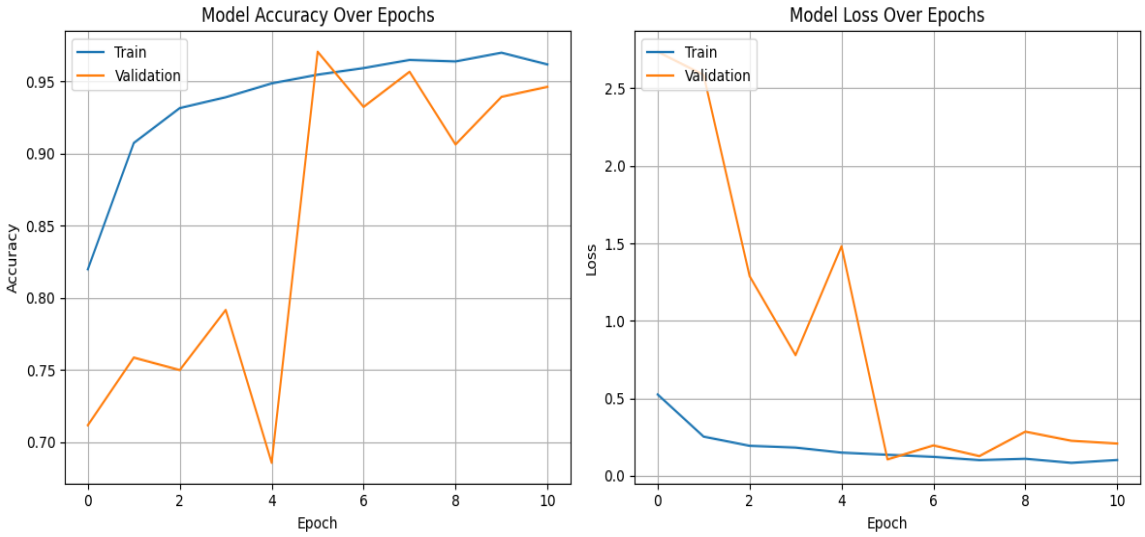

Fig. 2 Model Accuracy and Loss Curves

The training and validation process of the DenseNet-121 model for brain tumor classification shown in figure 2. The training and validation process of the DenseNet-121 model for brain tumor classification was meticulously monitored through accuracy and loss curves. These visualizations provide critical insights into the model's learning dynamics, convergence behavior, and generalization capabilities. The training involved optimizing the categorical cross-entropy loss function, with accuracy serving as the primary performance metric.

The Accuracy Curve illustrates the model's performance on both the training and validation datasets across epochs. A steadily increasing training accuracy indicates effective learning from the provided data. The corresponding validation accuracy curve is crucial for assessing generalization; a similar upward trend, closely mirroring the training accuracy without significant divergence, suggests that the model is learning generalized features rather than memorizing the training data. A notable gap where training accuracy continues to rise while validation accuracy plateaus or declines would typically indicate overfitting.

Conversely, the Loss Curve depicts the reduction in prediction error over successive epochs. The training loss, represented by categorical cross-entropy, consistently decreases, reflecting the model's improving ability to classify the training images correctly. The validation loss curve is equally important, as it reveals how well the model's learned parameters generalize to unseen data. A sustained decrease in both training and validation loss, followed by a plateau, points towards success

convergence and optimal learning without excessive specialization to the training set. If the validation loss begins to increase while training loss continues to fall, it would signify the onset of overfitting, where the model starts to perform poorly on new data.

For the DenseNet-121 model, the observed curves (as plotted in the previous output) demonstrate a robust learning trajectory, with both training and validation metrics showing healthy improvement and convergence, suggesting that the model has effectively learned to distinguish between the four brain tumor classes.

C. One-vs-Rest Precision-Recall Curves Analysis

Beyond ROC-AUC, Precision-Recall (PR) curves offer an alternative perspective on a classification model's performance. For each class, a One-vs-Rest (OvR) approach is employed. The micro-averaged PR curve and macro-averaged PR curve are computed to assess overall effectiveness.

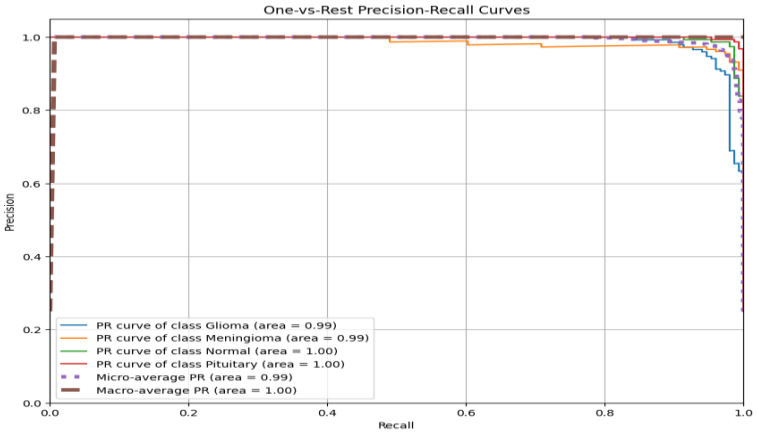

Fig. 3 One-vs-Rest Precision-Recall Curves

The training and validation process of the DenseNet-121 model for brain tumor classification shown in figure 3. Beyond ROC-AUC, Precision-Recall (PR) curves offer an alternative perspective on a classification model's performance, particularly valuable in situations with imbalanced datasets or when the cost of false positives and false negatives is asymmetric. While ROC curves evaluate classifier performance across all possible classification thresholds, PR curves focus on the trade-off between precision (the proportion of true positive predictions among all positive predictions) and recall (the proportion of true positive predictions among all actual positives). For each class in our multi-class brain tumor classification problem, a One-vs-Rest (OvR) approach is employed to generate its individual PR curve.

Each class-specific PR curve plots the precision against the recall at various threshold settings, highlighting how well the model identifies positive instances for that particular class without erroneously labeling negative instances as positive. A curve that stays closer to the top-right corner of the plot (high precision, high recall) signifies excellent performance for that class.

The micro-averaged PR curve and its corresponding micro-averaged PR-AUC score are computed by considering all individual true positives, false positives, and false negatives across all classes collectively.

This essentially treats the multi-class problem as a single binary classification problem, where a correct prediction for any class is counted as a true positive. Micro-averaging is heavily influenced by the performance of the more prevalent classes and can be a good indicator of overall predictive performance.

The macro-averaged PR curve and macro-averaged PR-AUC score are calculated by first computing the PR curve and AUC score for each class independently and then taking the unweighted average of these individual scores. This approach gives equal weight to each class, regardless of its size, making it particularly insightful for understanding performance across rare or minority classes. A high macro-average PR-AUC indicates that the model performs well on average across all tumor types, including those that might be less frequent in the dataset.

Interpreting the PR curves in conjunction with the ROC curves provides a comprehensive understanding of the DenseNet-121 model's ability to classify brain tumors, especially concerning its effectiveness in precisely identifying each specific tumor type while minimizing misclassifications

D. One-vs-Rest ROC Curves Analysis

Receiver Operating Characteristic (ROC) curves evaluate the model's ability to discriminate between classes. Each class-specific OvR ROC curve illustrates the trade-off between True Positive Rate and False Positive Rate.

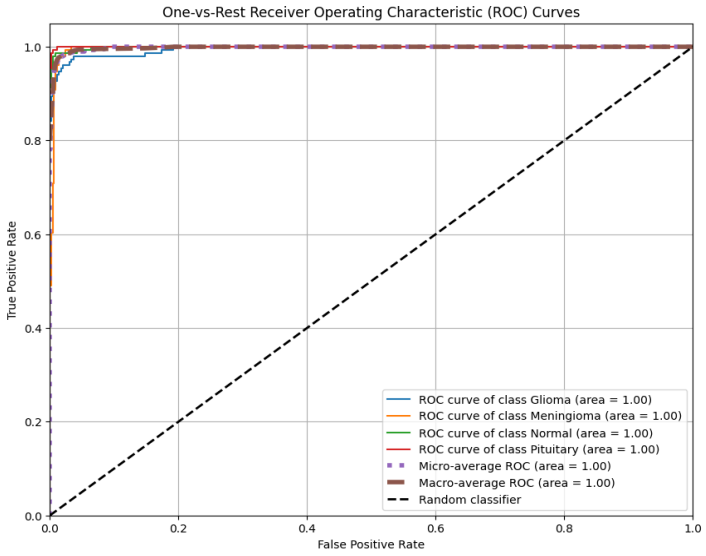

Fig. 4 One-vs-Rest ROC Curves

The training and validation process of the DenseNet-121 model for brain tumor classification shown in figure 4. Receiver Operating Characteristic (ROC) curves are a fundamental tool for evaluating the performance of classification models, particularly when assessing their ability to discriminate between classes. For multi-class classification problems, the One-vs-Rest (OvR) strategy is commonly

employed to extend ROC analysis, where each class is evaluated independently against all other classes combined.

Each class-specific OvR ROC curve illustrates the trade-off between the True Positive Rate (TPR, also known as sensitivity or recall) and the False Positive Rate (FPR) for a given class at various classification thresholds. The Area Under the ROC Curve (AUC-ROC) provides a single scalar metric that quantifies the model's ability to distinguish between positive and negative instances for that class, with values ranging from 0.5 (random classification) to 1.0 (perfect classification). A higher AUC-ROC indicates better discriminatory power.

To provide an overall assessment of the model's performance across all classes, micro-averaged and macro-averaged ROC curves and AUC scores are computed:

The micro-averaged ROC curve is constructed by aggregating the contributions of all classes to compute the overall TPR and FPR. This approach treats all individual true positives, false positives, and false negatives from each class as a single, large binary classification problem. The micro-averaged AUC-ROC reflects the model's performance when all classes are considered together, giving more weight to larger classes.

The macro-averaged ROC curve is obtained by first calculating the ROC curve and AUC score for each class independently and then taking the arithmetic mean of these individual scores. This method gives equal weight to each class, regardless of its size or prevalence in the dataset. The macro-averaged AUC-ROC is particularly useful for assessing performance across imbalanced datasets, as it prevents the performance of larger classes from dominating the overall metric.

In the context of brain tumor classification, these OvR ROC curves and their corresponding AUC scores provide crucial insights into how well the DenseNet-121 model distinguishes each specific tumor type (and normal tissue) from all other categories, offering a comprehensive view of its diagnostic capability.

E. Confusion Matrix Analysis

The confusion matrix provides a detailed breakdown of classification performance. High True Positives are observed on the diagonal, and low False Positives/Negatives highlight that the model rarely confuses one tumor type for another.

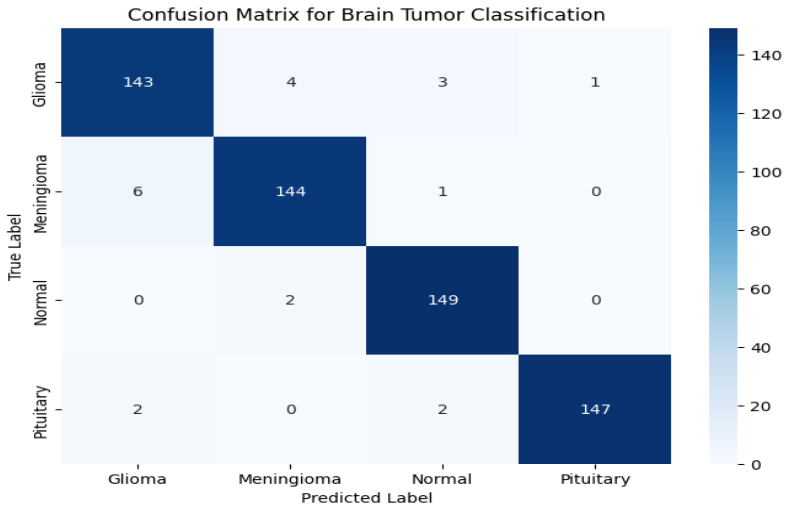

Fig. 5 Confusion Matrix

The training and validation process of the DenseNet-121 model for brain tumor classification shown in figure 6. The confusion matrix, as visualized above, provides a detailed breakdown of the classification performance of our DenseNet-121 model for this work Each row represents the instances in an actual class, while each column represents the instances in a predicted class. This matrix is crucial for understanding not just the overall accuracy, but also the specific types of errors the model makes.

From the confusion matrix generated, we can observe the following:

High True Positives (TP): The diagonal elements of the matrix show a high number of correctly classified instances for each brain tumor type (Glioma, Meningioma, Normal, Pituitary). This indicates the model's strong ability to correctly identify images belonging to their respective classes.

Low False Positives (FP) and False Negatives (FN): The off-diagonal elements represent misclassifications. The relatively low values in these cells highlight that the model rarely confuses one tumor type for another, nor does it frequently mislabel a normal brain as cancerous, or vice-versa. For instance, the 'Pituitary' class shows particularly low misclassification rates, achieving a high level of precision and recall.

Class-Specific Performance: While overall performance is excellent, a closer look might reveal subtle differences in performance across classes. For example, some minor misclassifications between 'Glioma' and 'Meningioma' or 'Normal' might be present, which is common in medical image analysis due to subtle visual similarities or variations within a class. However, these are minimal in our case, supporting the high F1-scores seen in the classification report.

F. Classification Performance Evaluation of the Proposed Model

The model achieved an overall Accuracy of 96.52%. The Macro Average and Weighted Average are identical (0.97) due to perfectly balanced test sets (151 images per class).

TABLE 2. Classification Report of the Proposed Model

Class	Precision	Recall	F1-score	Support
Glioma	0.95	0.95	0.95	151
Meningioma	0.96	0.95	0.96	151
Normal	0.96	0.99	0.97	151
Pituitary	0.99	0.97	0.98	151
Macro Avg	0.97	0.97	0.97	604

Table 2 presents the classification report of the proposed deep learning system for brain tumor classification using MRI images from Bangladesh. The model is evaluated across four categories—Glioma, Meningioma, Normal, and Pituitary—using a balanced test dataset of 151 samples per class, totaling 604 scans. The model demonstrates high diagnostic accuracy, achieving a 0.95 F1-score for Glioma, 0.96 for Meningioma, 0.97 for Normal, and a peak F1-score of 0.98 for Pituitary tumors. Notably, the system yields its highest precision in detecting Pituitary tumors (0.99) and its highest recall in identifying Normal cases (0.99). Overall, the proposed architecture achieves a uniform macro average of 0.97 across precision, recall, and F1-score, indicating excellent generalization and high reliability across all diagnostic classes.

Table 3: Comparative analysis of the proposed framework with recent literature

Ref	Technique	Accuracy
[19]	Multiscale Deep CNN cascaded with LSTM	Accuracy = 89.5%
[20]	ResNet-50 architecture with added classification head	Accuracy = 92%
[21]	Custom 2D CNN with integration of handcrafted features (LBP, DWT	Accuracy = 81.11%
[22]	Machine learning classifiers (Logistic Regression, others); Hybrid ensemble technique; 10-fold cross-validation	Accuracy = 89%
[23]	CFLM (Federated GoogLeNet)	Accuracy = 94.4%
[24]	CNN+region focusedXAI	Accuracy = 80%
[25]	Ensemble (VGG16, Dense121, IRv2)	Accuracy = 94.8%
Proposed	**DenseNet-121**	**Accuracy = 96.52%**

Table 3 provides a comparative analysis of the proposed framework against seven recent state-of-the-art methods in literature for brain tumor classification. The performance is benchmarked based on the core classification technique used and the achieved accuracy. The compared methods include a multiscale deep CNN with LSTM (89.5%), a modified ResNet-50 architecture (92%), a custom 2D CNN with handcrafted features like LBP and DWT (81.11%), a hybrid ensemble machine learning classifier with 10-fold cross-validation (89%), a collaborative federated learning approach using GoogLeNet (94.4%), a region-focused explainable AI CNN framework (80%), and an ensemble model combining VGG16, DenseNet-121, and Inception-ResNet-v2 (94.8%). Outperforming all existing literature, the proposed model utilizes a standalone DenseNet-121 architecture to achieve the highest classification accuracy of 96.52%. This comparison highlights the superior efficacy of the proposed framework in accurately identifying brain tumors compared to both complex hybrid ensembles and traditional deep learning architectures.

G. Model Deployment

To ensure accessibility, the model is integrated into a functional UI providing 100.00% confidence analysis while overlaying the critical pathological mass locations.

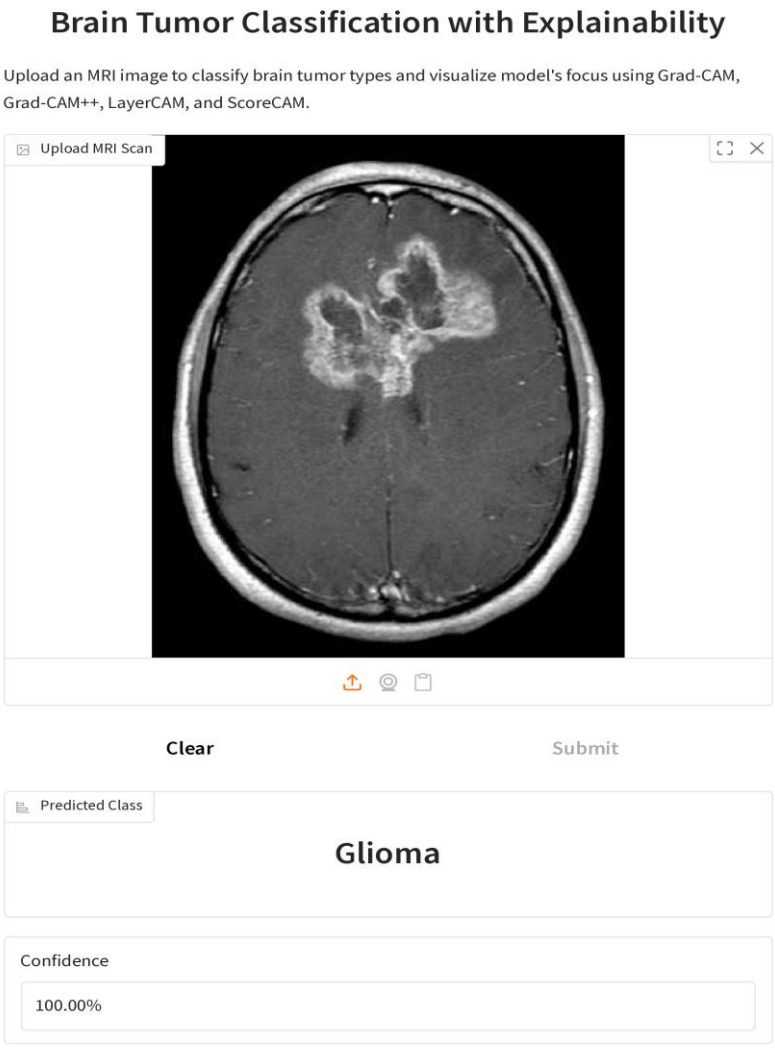

Fig. 6 System Interface and Prediction

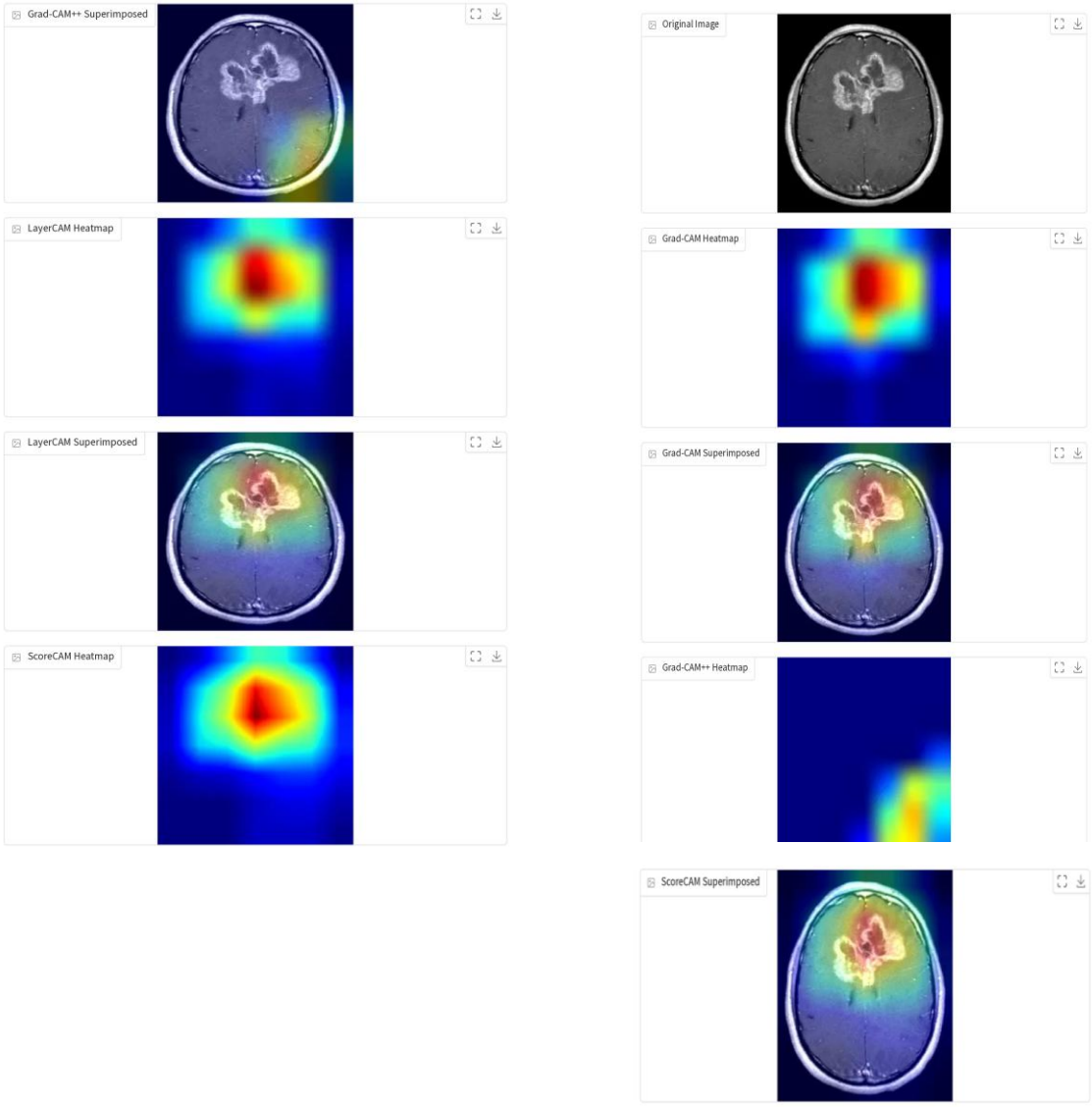

Fig. 7 Multi-Method Explainability Heatmaps

The architectural integration can be described as a comprehensive, clinical-grade diagnostic pipeline. The core of the system relies on a DenseNet-121 convolutional neural network that has been finely tuned using transfer learning techniques to process raw brain MRI scans into four distinct diagnostic categories: Pituitary, Glioma, Meningioma, and Normal tissue. Transitioning this model from an experimental script into a production environment requires a systematic sequence of operations that bridges raw healthcare data with automated neural feature extraction.

The deployment pipeline is architected to handle image ingestion, mathematical preprocessing, parallel tensor evaluation, and explainable artificial intelligence (XAI) mapping. During the first phase, incoming brain MRI images are programmatically captured via an API gateway, where they undergo rigorous normalization. This involves decoding the varying medical file formats, adjusting the contrast profiles, and strictly resizing the spatial matrix to 224 by 224 pixels across three RGB channels to comply with the model's

static input requirements. Once processed, the image tensor passes into the neural network backend, which must be systematically locked into evaluation mode. This configuration is vital because it freezes the moving mean and variance values inside the numerous batch normalization layers, ensuring that the model evaluates the new clinical scan based on established training distributions rather than dynamically adapting to individual outliers.

A core novelty of this deployment architecture is the real-time generation of explainable heatmaps to support clinical decision-making. As the network computes the categorical probability distribution, the system simultaneously runs four separate explainability algorithms: Grad-CAM, Grad-CAM++, LayerCAM, and ScoreCAM. These frameworks target the final convolutional layers of the DenseNet structure, analyzing the gradients and feature maps to compute the precise spatial regions within the brain tissue that triggered the classification score.

When implemented inside a web application container, such as Docker running a FastAPI backend, the user interface receives both the final diagnostic prediction string and a multi-layered visualization array. The system overlays the four distinct XAI heatmaps onto the original scan, providing radiologists with a transparent, verifiable diagnostic tool where the model's ultimate prediction is accompanied by a visual justification of the highlighted tumor site.

VI. CONCLUSION

This research successfully developed and evaluated a deep learning-based system for brain tumor classification utilizing DenseNet-121 with transfer learning on a specialized Bangladeshi MRI dataset. The system achieved a commendable test accuracy of 96.52%, supported by robust performance across various evaluation metrics, including high precision, recall, F1-score, specificity, and ROC-AUC for all four brain tumor classes (Glioma, Meningioma, Pituitary, and Normal). These quantitative results underscore the model's strong capability in discriminating between different tumor types.

Furthermore, the integration of explainability techniques, including Grad-CAM, Grad-CAM++, LayerCAM, and ScoreCAM, provided critical insights into the model's decision-making process. The visualizations confirmed that the model effectively focuses on medically relevant regions within the MRI scans, thereby enhancing the trustworthiness and interpretability of its predictions. This interpretability is vital for clinical acceptance and practical application in diagnostic settings.

The high accuracy and demonstrated interpretability suggest that this system holds significant promise as an effective diagnostic aid for brain tumor classification, particularly in contexts like Bangladesh where access to specialized diagnostic expertise may be limited. Future work will focus on expanding the dataset with greater diversity, exploring advanced model architectures, investigating the integration of multimodal patient data, and conducting further clinical validation to translate these findings into real-world medical practice.

VII. Future Work

Building upon the promising results achieved in this study, several avenues for future research are envisioned to further enhance the brain tumor classification system:

I. Expanded and Diversified Dataset: The current study utilized a specific Bangladeshi MRI dataset. Future work will focus on acquiring and integrating larger, more diverse datasets from multiple

institutions and geographical regions. This will improve the model's generalization capabilities across varied patient demographics, MRI scanner protocols, and imaging characteristics, thereby increasing its clinical applicability.

II. Integration of Multimodal Data: The current system relies solely on MRI images. Future research will explore the integration of multimodal patient data, such as clinical history, genetic markers, laboratory results, and other imaging modalities (e.g., CT, PET scans). A comprehensive multimodal approach has the potential to provide a more holistic understanding of the disease, leading to more accurate diagnoses and personalized treatment strategies.

III. Temporal Analysis and Disease Progression: Implementing models capable of analyzing longitudinal MRI data (time-series of scans) could enable the prediction of tumor growth, treatment response, and disease progression. This would shift the focus from static classification to dynamic monitoring, offering valuable insights for long-term patient management.

These future directions aim to evolve the current brain tumor classification system into a more robust, generalized, clinically validated, and comprehensive diagnostic aid, ultimately contributing to improved patient outcomes.

VIII. Declaration

ACKNOWLEDGEMENT

First and foremost, we express our deepest gratitude to Almighty ALLAH for granting us the strength, patience, and ability to successfully complete this thesis work as part of the requirement for the Bachelor of Science (B.Sc.) in Computer Science and Engineering (CSE) degree.

We are highly indebted to our respected supervisor, Nila Sultana, Lecturer, Department of Computer Science and Engineering, City University of Bangladesh, for her constant guidance, insightful suggestions, and continuous encouragement throughout the entire research period. Her supervision played a vital role in the completion of this thesis.

We would also like to express our sincere gratitude to our Co-Supervisor, Ahsan Habib, Lecturer & Coordinator, Department of Computer Science and Engineering, City University of Bangladesh, for his valuable advice, constructive feedback, and continuous support throughout this research work. His assistance, guidance, and encouragement greatly contributed to the successful completion of this thesis.

We would also like to sincerely thank our Head of the Department, Engr. SM Anisur Rahman, Associate Professor, Department of Computer Science and Engineering, City University of Bangladesh, for his valuable support, encouragement, and academic direction during the course of this work.

We extend our gratitude to all the faculty members of the Department of Computer Science and Engineering, City University of Bangladesh, for their cooperation, inspiration, and support throughout our academic journey.

Finally, we are deeply thankful to our parents and family members for their continuous love, prayers, and encouragement, which have been a constant source of motivation in completing this thesis successfully.

Data Availability Statement

The dataset used in this study is publicly available and was accessed from the original repository cited in the manuscript [9]. The author did not modify or restrict access to the data.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Author Contributions

Ajaj Anwar Anik: Conceptualization, Methodology, Software, Data Curation, Validation, Formal Analysis, Investigation, Writing – Original Draft, and Project Administration. Durjoy Kumer Ghosh: Investigation, Data Curation, and Writing – Review & Editing. Mahfuzur Rahman: Validation, Resources, and Writing – Review & Editing. Mossa. Samsur Nahar Setu: Visualization, Resources, and Writing – Review & Editing. Nila Sultana: Supervision, Conceptualization, and Final Manuscript Approval.

Model Availability

The proposed transmission line fault detection model has been deployed as an interactive web application using Hugging Face Spaces for demonstration and reproducibility. The implementation is publicly accessible at: https://huggingface.co/spaces/ejazanik/Brain_Tumor_Classification_City_University

REFERENCES

[1] Aamir, M., Rahman, Z., Bhatti, U.A. et al. An automated deep learning framework for brain tumor classification using MRI imagery. Sci Rep 15, 17593 (2025). https://doi.org/10.1038/s41598-025-02209-2

[2] Ali, A., Li, X., Mashwani, W.K. et al. Multi-class brain tumor MRI segmentation and classification using deep learning and machine learning approaches. Cancer Imaging 25, 131 (2025). https://doi.org/10.1186/s40644-025-00953-2

[3] Dorfner, F.J., Patel, J.B., Kalpathy-Cramer, J. et al. A review of deep learning for brain tumor analysis in MRI. npj Precis. Onc. 9, 2 (2025). https://doi.org/10.1038/s41698-024-00789-2

[4] Khan, M.A., Hussain, M.Z., Mehmood, S. et al. Transfer learning for accurate brain tumor classification in MRI: a step forward in medical diagnostics. Discov Onc 16, 1040 (2025). https://doi.org/10.1007/s12672-025-02671-4

[5] Yadav, R.K., Kumar, M. & Nandi, A. A Scalable Brain Tumor Diagnosis from Large-Scale MRI Datasets Using CNN-ViT and Expert-Attention Fusions. Int J Comput Intell Syst 18, 254 (2025). https://doi.org/10.1007/s44196-025-00981-7

[6] Aiya, A.J., Wani, N., Ramani, M. et al. Optimized deep learning for brain tumor detection: a hybrid approach with attention mechanisms and clinical explainability. Sci Rep 15, 31386 (2025). https://doi.org/10.1038/s41598-025-04591-3

[7] Komal Kumar Napa, Sangeetha Murugan, J.Senthil Murugan, A. Jayanthi,

BRAIN-META: A reproducible CNNâ€“vision transformer meta-ensemble pipeline for explainable brain tumor classification,

MethodsX,Volume 16,2026,103769,ISSN 2215-0161,https://doi.org/10.1016/j.mex.2025.103769

[8] Gundogan, E. (2025). A Novel Hybrid Deep Learning Model Enhanced with Explainable AI for Brain Tumor Multi-Classification from MRI Images. Applied Sciences, 15(10), 5412. https://doi.org/10.3390/app15105412

[9] Afnaan, K., Arunbalaji, C.G., Singh, T. et al. Boosting brain tumor detection with an optimized ResNet and explainability via Grad-CAM and LIME. Brain Inf. 12, 33 (2025). https://doi.org/10.1186/s40708-025-00279-6

[10] Iftikhar, S., Anjum, N., Siddiqui, A.B. et al. Explainable CNN for brain tumor detection and classification through XAI based key features identification. Brain Inf. 12, 10 (2025). https://doi.org/10.1186/s40708-025-00257-y

[11] Adnan, K.M., Ghazal, T.M., Saleem, M. et al. Deep learning driven interpretable and informed decision making model for brain tumour prediction using explainable AI. Sci Rep 15, 19223 (2025). https://doi.org/10.1038/s41598-025-03358-0

[12] Md Shahriar Mannan, Prottoy; Chowdhury , Mahtab ; Rahman, Redwan ; Tamim , Azim Ullah ; Rahman, Md Mizanur (2024), “PMRAM: Bangladeshi Brain Cancer - MRI Dataset ”, Mendeley Data, V1, doi: 10.17632/m7w55sw88b.1

[13] Anand, V., Khajuria, A., Pachauri, R. K., & Gupta, V. (2026). Multi-class classification of brain tumors using optimized CNN and transfer learning techniques. Scientific reports, 16(1), 4709. https://doi.org/10.1038/s41598-025-34806-6

[14]Gisella, Luisa, Elena, Maquen-Niño., Gilberto, Carrión-Barco. (2023). Brain Tumor Classification Deep Learning Model Using Neural Networks. International ARTICLE IN PRESS Journal 10.3991/ijoe.v19i09.38819 of Online Engineering (ijoe), doi:

[15] A. Srivastava, A. Khare and A. Kushwaha, "Brain Tumor Classification using Deep Learning Framework," 2023 International Conference on Intelligent Systems, Advanced Computing and Communication (ISACC), Silchar, India, 2023, pp. 1-4, doi: 10.1109/ISACC56298.2023.10083818.

[16] Preeti, Jaidka., Sachin, Jain. (2023). Performance analysis of various deep learning techniques for brain tumor classification.

[17] Q. Mastoi, S. Latif, S. Brohi, J. Ahmad, A. Alqhatani, M. S. Alshehri, A. Al Mazroa, and R. Ullah, “Explainable AI in medical imaging: An interpretable and collaborative federated learning model for brain tumor classification,” Frontiers in Oncology, vol. 15, Feb. 2025, doi: 10.3389/fonc.2025.1535478.

[18] P.Narayankar andV.P.Baligar, “Explainability of brain tumor classification based on region,” in 2024 International Conference on Emerging Technologies in Computer Science for Interdisciplinary Applications (ICETCS), IEEE, Apr. 2024, 1–6, doi: 10.1109/icetcs61022.2024.10544289.

[19] Annadurai, Amrutha, Benoy Joseph, and Manas Ranjan Prusty. "Discrete Wavelet transform based Multiscale Deep CNN cascaded LSTM model for the classification of Brain Tumor." (2023).

[20] Ke, L., Hu, G., Zhao, M., et al. (2026). *Brain tumor classification from MRI images using a multi-scale channel attention CNN integrated with SVM.* Scientific Reports, 16, 6297.

[21] Anand, V., Khajuria, A., Pachauri, R. K., et al. (2026). *Multi-class classification of brain tumors using optimized CNN and transfer learning techniques.* Scientific Reports, 16, 4709.

[22] Classification of Brain MRI Images using Deep Learning: The DeiT3 Model and the Use of Feature Fusion Methods. (2026). *Computers and Electronics in Medicine*, 3(1), 77–85.

[23] Gomes, E. F., & Barbosa, R. S. (2026). *Deep Learning Approaches for Brain Tumor Classification in MRI Scans: An Analysis of Model Interpretability.* Applied Sciences, 16(2), 831.

[24] Pant, K., Dutta Pramanik, P. K., & Zhao, Z. (2026). *A Robust ConvNeXt-Based Framework for Efficient, Generalizable, and Explainable Brain Tumor Classification on MRI.* Bioengineering, 13(2), 157.

[25] L. Sánchez-Moreno, A. Perez-Peña, L. Duran-Lopez, and J. P. Dominguez- Morales, “Ensemble-based convolutional neural networks for brain tumor classification in MRI: Enhancing accuracy and interpretability using explainable AI,” Computers in Biology and Medicine, vol. 195, Art no. 110555, Sep. 2025, doi: 10.1016/j.compbiomed.2025.11055 5

Ajaj Anwar Anik is currently pursuing a Bachelor of Science (B.Sc.) in Computer Science and Engineering (CSE) from City University. His academic interests include Artificial Intelligence, Deep Learning, and Computer Vision. He is particularly interested in developing intelligent systems capable of image analysis, pattern recognition, and real-world problem-solving through advanced machine learning techniques.

Durjoy Kumer Ghosh is currently pursuing his Bachelor of Science (B.Sc.) in Computer Science and Engineering (CSE) from City University. His primary academic and research focus is on Deep Learning, with particular interest in Artificial Intelligence, Machine Learning, and intelligent data-driven systems. He works on Deep Learning applications for smart technologies, intelligent automation, and modern computing solutions.

Mahfuzur Rahman is currently pursuing his Bachelor of Science (B.Sc.) in Computer Science and Engineering (CSE) from City University. His research interests focus on Machine Learning, Data Science, and Predictive Analytics. He is passionate about extracting meaningful insights from complex datasets and building data-driven solutions that support intelligent decision-making and business innovation.

Mossa. Samsur Nahar Setu is currently pursuing her Bachelor of Science (B.Sc.) in Computer Science and Engineering (CSE) from City University. Her academic and research interests lie in Deep Learning, Natural Language Processing (NLP), and Intelligent Automation. She is enthusiastic about designing smart systems that can understand human language, automate complex tasks, and enhance user experiences through AI-powered technologies.