XAI-empowered MRI Analysis for Consumer Electronic Health

Al Amin, Kamrul Hasan, *Member, IEEE*, and M. Shamim Hossain, *Senior Member, IEEE*

***Abstract*—The intelligent use of artificial intelligence-generated content (AIGC) in magnetic resonance imaging (MRI) analysis is a significant step towards the rapidly advancing field of consumer electronics (CE) used in healthcare. It greatly improves the accuracy and usefulness of diagnostic processes. This paper introduces a novel MRI analysis approach through the lens of AIGC, leveraging physics-informed deep learning (PIDL) models. This integration pioneers a new paradigm in consumer healthcare diagnostics and embeds physics principles into deep learning (DL) models, thus improving interpretability and adherence to physical constraints. Additionally, this method improves the visibility of the healthcare community by integrating explainable AI (XAI) techniques, including gradient-weighted class activation mapping (Grad-CAM) and Local Interpretable Model-Agnostic Explanations (LIME). This approach demonstrates a reasonable precision rate of 96% when applied to brain tumor MRI images. Therefore, this research introduces a new para diagram of applying AIGC in medical imaging analysis within Consumer Healthcare Electronics (CHE).**

***Index Terms*—Explainable AI (XAI), Physics-Informed Deep Learning (PIDL), AIGC, Feature extraction method, Consumer Health Electronics.**

I. Introduction

In Consumer Health Electronics (CHE), magnetic resonance imaging (MRI) provides high-resolution images for accurate diagnosis without intrusive intervention [1], [2]. Despite its importance, MRI data interpretation is difficult and often requires advanced computational methods. These strategies require more transparency, making clinical acceptance and confidence harder [3]–[5]. Physics-informed deep learning (PIDL) models, a cornerstone of this approach, redefine Artificial Intelligence-Generated Content (AIGC) in medical imaging by integrating domain-specific knowledge, thereby enhancing the reliability of AI systems for complex tasks like brain tumor identification within Consumer Electronics/Technologies (CE/CT). The analysis of MRI images has been considerably improved by recent advancements in Artificial Intelligence (AI), mainly through deep learning (DL) [6]. Physics-informed deep learning (PIDL) is a novel method incorporating domain-specific physics knowledge into neural network architectures. This implies that the diagnostic outputs are enhanced regarding interpretability and reliability as the neural network acquires knowledge from the data and the underlying physics principles that govern it.

Al Amin, Kamrul Hasan are with the Electrical and Computer Engineering Department, Tennessee State University, Nashville, TN, USA. (emails: aamin2@tnstate.edu, mhasan1@tnstate.edu)

M. Shamim Hossain is with the Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 12372, Saudi Arabia (email: mshossain@ksu.edu.sa)

Manuscript received Oct 11, 2021; revised .....

We aim to establish a new standard for the interpretability and clinical usefulness of AI-driven magnetic resonance imaging (MRI) diagnostics, with the potential to completely transform how medical imaging is carried out in consumer health technology. Integrating deep learning (DL), particularly convolutional neural networks (CNNs), has marked a significant milestone in automating magnetic resonance images (MRI) analysis, a critical component of Consumer Healthcare Electronics (CHE) [7], [8], yielding substantial improvements in detecting and classifying pathological abnormalities. However, the deep learning (DL) ”black box” nature poses a critical barrier to its clinical deployment, as it provides limited insights into its inferential processes [9]. Although recent developments in explainable artificial intelligence (XAI) have endeavored to mitigate this issue by introducing interpretability in AI systems [10], these approaches often need to incorporate the domain-specific knowledge intrinsic to MRI physics. This gap underlines the need for creative ways that combine deep learning (DL) prediction capacity with MRI physics' fundamental principles to improve diagnostic and consumer technology accuracy and interpretability.

This research introduces a novel approach to address the critical requirements of Consumer Electronics/Technologies in healthcare diagnostics. It is the first to introduce a distinctive system architecture within the Artificial Intelligence-Generated Content (AIGC) framework, which integrates physics-informed deep learning (PIDL) with explainable artificial intelligence (XAI). The PIDL model meets the strict demands of Consumer Health Electronics (CHE) by combining the deep learning (DL) model with physics-based constraints. This makes MRI diagnostics much more accurate and easier to understand. The methodology, validated through extensive experimentation, underscores the potential of this novel approach in revolutionizing medical imaging analysis for consumer health applications.

Figure 1 illustrates a practicable and innovative framework for Physics-Informed and Explainable AI-enhanced AI-generated content (PIE-AIGC) to improve MRI diagnostics in Consumer Healthcare Electronics (CHE). The procedure begins with acquiring data from CE-certified MRI machines, which is transferred securely and seamlessly to Consumer Electronics-compliant servers. The data, representing a diverse range of imaging, is strategically divided, with 80% allocated for training and 20% for testing, laying a robust foundation for the AI's learning phase. The data is meticulously preprocessed in the subsequent stage, optimizing it for the deep learning (DL) model. This stage is critical for the precise extraction of features, as it includes essential quality enhancements and normalization.

Integration of XAI and PIDL for Enhanced Diagnostics: A dualistic methodology is employed during extraction, integrating physics-based constraints with convolutional neural network (CNN) driven pattern recognition. This approach enhances the clinical findings' interpretability and relevance. The synthesized information feeds into the Consumer Electronics/Technologies optimized physics-informed deep learning (PIDL) model, which integrates these insights with a fine-tuned loss function to minimize false diagnostics. The resultant model is not merely a tool for identification but a beacon of reliability in clinical diagnostics. Completing the architecture, XAI techniques—Local Interpretable Model-Agnostic Explanation (LIME) and gradient-weighted class activation mapping (Grad-CAM)—provide a window into the AI's inferential mechanics, offering healthcare practitioners transparent and understandable diagnostic results. This comprehensive system streamlines the diagnostic workflow and assures precision and clarity in interpreting complex medical images, solidifying the commitment to enhancing CHE through intelligent innovation.

Transparency and Future Expansion of the PIE-AIGC Framework: The Physics-Informed and Explainable AI-Enhanced AIGC Framework (PIE-AIGC) integrates advanced XAI techniques, notably LIME and Grad-CAM, to make the AI's decision-making process transparent. In experiments, Grad-CAM visually highlights the crucial regions of an MRI scan in diagnosing a brain tumor, providing clinicians with intuitive visual evidence of the AI's focus areas. This enhances trust in the AI's capabilities and enables medical professionals to understand and verify the underlying reasoning for its diagnoses. This transparency is precious for educational purposes, as it aids healthcare providers in understanding and adopting AI technology. The aim is to explore additional XAI techniques that could offer even deeper insights into the model's predictions and adherence to physical principles. This expansion of the scope of the Physics-Informed and Explainable AI-enhanced AI-generated content (PIE-AIGC) Framework is intended to encompass a broader range of medical applications.

In summary, the main contributions of this work are:

- A novel Physics-Informed and XAI-Enhanced Artificial Intelligence-Generated Content (AIGC) Framework for Consumer Health System Architecture (PIE-AIGC) leverages a PIDL hybrid approach tailored for advanced brain tumor MR image analysis. This model elevates feature extraction efficiency and minimizes loss, marking a significant leap forward in medical image analysis.
- Custom Explainable AI (XAI) techniques, specifically Local Interpretable Model-Agnostic Explanation (LIME) and gradient-weighted class activation mapping (Grad-CAM) are introduced to demystify the intricate decision-making processes within deep learning (DL) black-box architectures. This initiative ensures the diagnostic procedures are transparent and understandable, aligning with the essential requisites for medical interpretability and reinforcing PIDL's role in AIGC content.

The remainder of this paper is structured as follows: Section II explores the current state of research, specifically focusing on combining advanced deep learning techniques with artificial intelligence (XAI) in medical diagnostics for consumer healthcare electronics. Section III introduces an innovative methodology, the Physics-Informed and Explainable AI-Enhanced AI-Generated Content (PIE-AIGC) framework for Consumer Electronic (CE). Section IV discusses the dataset utilized and provides a detailed analysis of the experimental results, showcasing the approach's efficacy. Section V summarizes the paper by discussing fundamental discoveries and their significant impact on consumer healthcare diagnostics, and it also outlines future research directions that could further refine AI applications in medical imaging.

II. Related Works

Globally, scholars, industry experts, healthcare practitioners, and regulatory agencies are collaboratively emphasizing the extraction of significant features from MR images by employing mathematical model [11], [12] statistical models [13], [14], principles of physics [15], [16], advanced DL algorithms [6], [17], and recently, AIGC content has been popular for medical imaging [18], [19]. AIGC, leveraging generative AI in Consumer Health Electronics (CHE), particularly for feature extraction in MRI imaging, presents challenges related to patient data privacy and diagnostic accuracy due to its potential to disclose sensitive information inadvertently [20].

Li et al. [21] present an AIGC comprehensive case study on a rare ectopic pituitary neuroendocrine tumor (PitNET), showcasing how ChatGPT advances medical literature on neuroendocrine tumors. It thoroughly details the patient's diagnosis, treatment, and prognosis, illustrating the benefits of AI in managing intricate medical cases. However, a potential limitation of this research arises from its emphasis on feature extraction techniques without a corresponding focus on the underlying physics-based laws that govern biological processes and tumor growth dynamics. This gap suggests that AI can significantly enhance data analysis and presentation. However, a deeper integration of physics-based models could further improve the understanding of disease mechanisms and the efficacy of treatment approaches, particularly in complex and rare conditions like ectopic PitNET.

Yang et al. [22], represent a novel feature extraction technique leveraging multi-level features from Inception-v3 and DenseNet201 pre-trained models, diverging from standard practices focusing on initial layers for feature extraction. This methodology, characterized by extracting and concatenating features from varying network depths, significantly enhances brain tumor classification capabilities. The study reports 99.34% and 99.51% testing accuracies for Inception-v3 and DenseNet201, respectively. Despite its high accuracy, the model's complexity and lack of explainability are a notable limitation. This opacity challenges its trustworthiness and practical application, as the "black box" nature of the fusion model may hinder its real-life deployment [3].

Lin et al. [23] identified significant features in brain MR images and established that CNNs and FreeSurfer are crucial for extracting features from MRI data. CNN-based methods leverage DL to transform MR images into 2.5D patches,

This article has been accepted for publication in IEEE Transactions on Consumer Electronics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCE.2024.3443203

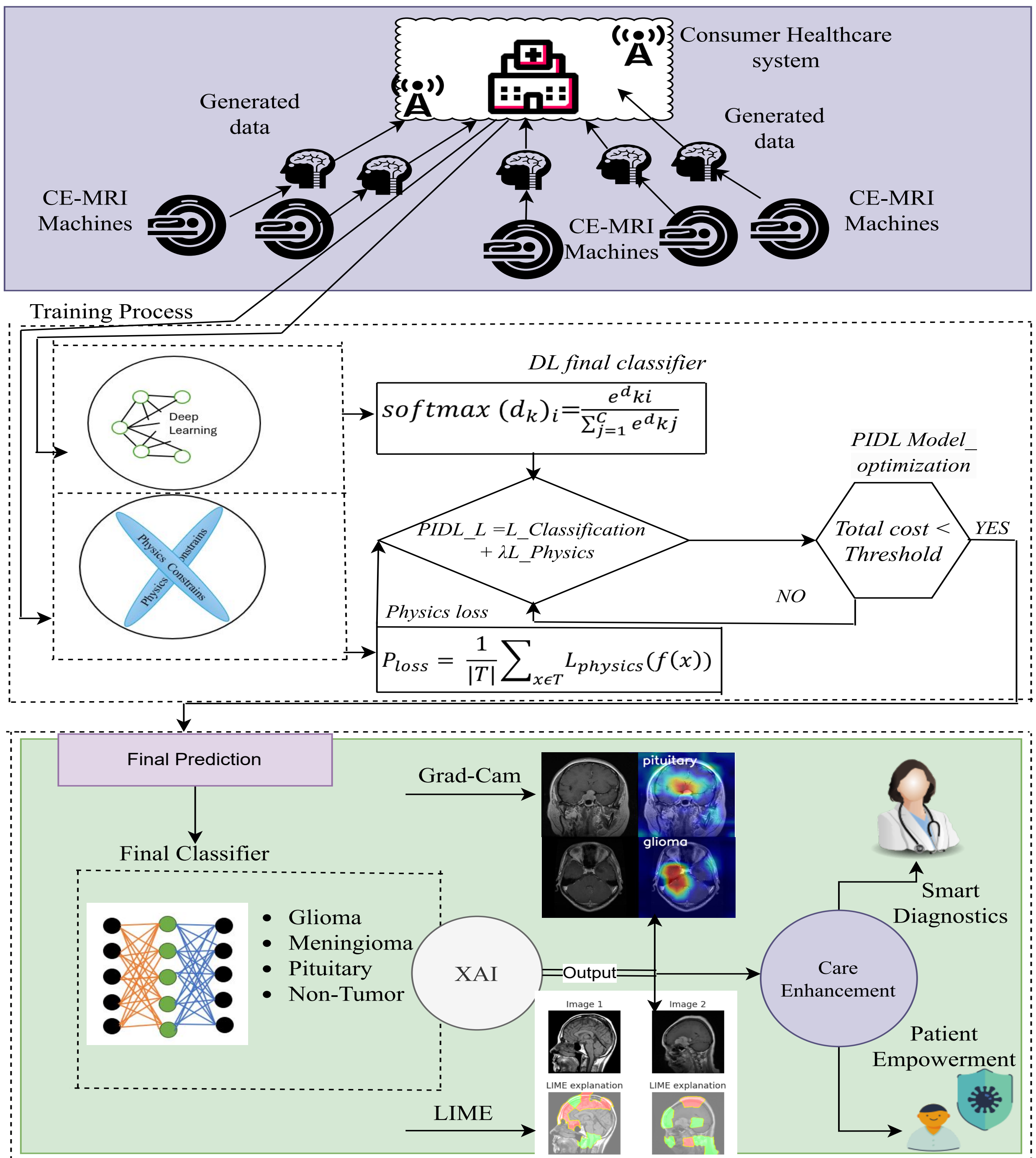

Fig. 1: Physics-Informed and Explainable AI-Enhanced AI-Generated Content (PIE-AIGC) Framework for Consumer Electronic Health System

capturing essential neuroimaging biomarkers. Concurrently, FreeSurfer provides critical volumetric and morphometric data, delineating structural brain variations. Integrating CNN-derived features with FreeSurfer metrics fosters a comprehensive, multimodal neuroimaging framework, significantly enhancing the model's predictive accuracy of 79.9%. The primary drawback of non-linear decision-making is the opacity of DL "black-box" processes.

Pattanaik et al.'s [24] research introduces a robust feature extraction fusion method that combines Stacked BiLSTM with Resnet50 and the Adaswarm optimizer. This method is applied to consumer device-collected datasets such as COVID-19, Pneumonia, Malaria, lung cancer, and brain tumors, achieving a satisfactory classification accuracy of 99%. However, there is a research gap in the explainability of this model decision, and DL could combine with biological law-like physics constraints for robust feature extraction techniques.

Feng et al. [25] consortiumed a Deep Boltzmann Machine (DBM) model for extracting features from multiple groups of medical images in the training datasets, achieving alignment and registration processing in advance. While the results indicate that the proposed method is effective in image fusion and somewhat better than Super-Resolution (SR) in tests of CT and MR images, the research faces limitations regarding Relative Entropy (REL) and Standard Deviation. This suggests that the fusion process can cause a loss of information. Consequently, significant, meaningful features have disappeared [26].

Zuo et al. [27], Another study introduced the Prior-Guided Adversarial Learning with Hypergraph (PALH) model for predicting Alzheimer's Disease (AD) progression, effectively identifying brain connectivity changes and offering insights into compensatory mechanisms and deterioration across AD stages. The PALH model, utilizing anatomical knowledge and hypergraph techniques, demonstrated superior performance with accuracies of 96.47% for AD vs. NC, 92.20% for LMCI vs. NC, and 87.50% for EMCI vs. NC predictions. Despite its success, the model's reliance on limited ROIs and neuroimaging data points to potential enhancements through incorporating genetic data, like SNPs, in future iterations.

Table 1 compares the AIGC framework with existing methodologies. It highlights the advancements in feature extraction, precision, and interpretability by integrating PIDL and XAI techniques in MRI-based brain tumor analysis.

TABLE I: Comprehensive comparison of the AIGC framework with existing methodologies

Research Title	Model	Validation criteria	Validation Accuracy
brain tumor classification using MRI data [28]	Fully convolutional neural networks (FCNs); Optimizers: Adam and Nesterov momentum, and RMSprop.	Precision, Recall, and F1	95%
Physics-Informed Discretization and Feature Extraction Using Quantitative MRI	Physics-Informed Discretization (PID) method [29]	Correlation Coefficient (ICC)	ICC >0.9
Classification of brain tumors using MRI [30]	DL architectures such as VGG, ResNet, EfficientNet, and ConvNeXt.	Precision, Recall, and F1	95.30%
Super-resolution of brain tumor utilizing MRI [31]	MRBT-SR-GAN to super-resolution	dice ratio	From 0.724 to 0.786
Our research: XAI-empowered MRI analysis for consumer health	PIDL With XAI	Precision, Recall, and F1	96.00%

The research introduces a groundbreaking AIGC framework that integrates PIDL with XAI techniques, marking a significant advancement in the analysis of brain tumors. This innovative approach, blending PIDL's precise modeling with the clarity offered by XAI, creates a powerful tool for MRI-based diagnostics. The framework excels in extracting critical features, employing DL and physics-informed principles to accurately pinpoint disease-specific changes in brain anatomy. Incorporating XAI technologies like LIME and Grad-CAM enhances the model's interpretability. This analyzes complex medical images transparent and grounded in scientific rigor. This approach sets a new standard for AI applications in healthcare diagnostics.

III. Methodology

A. Overview

The PIDL model refines the precision and dependability of AI for brain tumor detection in CHE by merging physical laws with domain knowledge. This integration aligns with biophysical standards, significantly decreasing imaging errors. The PIE-AIGC framework, a novel contribution to CHE, capitalizes on this fusion, incorporating XAI techniques like LIME and Grad-CAM to decode complex DL learning decisions. Thus, it increases the transparency and credibility of AI diagnostics in consumer healthcare.

B. Fusion of DL and Physics for Advanced Brain Tumor Detection in CHE MRI Images

In the CHE domain, a transformative framework for brain tumor identification is introduced through a PIDL model applied to CHE-compliant MRI datasets [32]. The PIDL model of MRI images is denoted as

$$X = \{x_1, x_2, ..., x_n\} \tag{1}$$

And analyzing it through a custom CNN function $f(x;\theta)$. This model is uniquely enhanced by integrating physics-informed constraints encapsulated in a sophisticated loss function.

$$L = L_{\text{traditional}}(f(x;\theta), y) + \lambda L_{\text{physics}}(f(x;\theta)), \tag{2}$$

Where y represents the true label and λ is a tuning parameter pivotal in CHE applications. The optimization process involves updating the model parameters θ, adhering to the rule

$$\theta_{\text{new}} = \text{Optimizer}(\theta_{\text{old}}, \nabla_\theta L). \tag{3}$$

The update formula governs training across epochs and batches

$$\theta = \theta - \eta \nabla_\theta L_B, \tag{4}$$

with η as the learning rate. For model interpretability, CHE-tailored XAI methods are implemented, applying functions $I_{\text{LIME}} = G_{\text{LIME}}(f, x)$ and $I_{\text{GradCAM}} = G_{\text{GradCAM}}(f, x)$ to produce comprehensible visualizations. The model's effectiveness is assessed on a test set T, calculating accuracy.

$$\text{Acc} = \frac{1}{|T|} \sum_{x \in T} 1[\arg\max(f(x)) = y] \tag{5}$$

and physics-informed loss

$$\text{PLoss} = \frac{1}{|T|} \sum_{x \in T} L_{\text{physics}}(f(x)). \tag{6}$$

The comprehensive and CHE-centric analytical algorithm culminates in the visualization and interpretation phase, rendering predictions as.

$$\hat{y} = \arg\max(f(x)). \tag{7}$$

Algorithm 1 represents the comprehensive approach that accurately identifies brain tumors and offers profound insights into the model's decision-making processes.

Algorithm 1: PIDL Model for Brain Tumor Identification

```
Data: MRI image dataset
Result: final prediction, model interpretability
Function Main(MRI image dataset):
    Step 1: Input - MRI Images
    Load MRI images from dataset
    preprocessed_images ← preprocess_images(dataset)
    Step 2: Custom CNN
    model ← define_cnn_architecture()
    Add layers (conv, activation, pool, drop, flatten)
    Step 3: Physics-Informed Constraints
    Define physics_informed_loss_function()
    Apply physics constraints to model
    Step 4: Optimization and Learning Rate
    Set up optimizer (e.g., Adam)
    Define learning rate strategy
    Step 5: Training the Model
    for epoch in epochs do
        for batch in train_generator do
            Train model on batch
        Validate model on validation_generator
        if early_stopping_criteria_met then
            break
    Step 6: Combined Loss Function
    During training, use combined loss function
    Step 7: Explainable AI (XAI) Techniques
    Implement LIME for local interpretability
    Apply Grad-CAM for visualization
    Step 8: Evaluation
    Evaluate on the test dataset
    Calculate accuracy, physics loss, etc.
    Step 9: Visualization and Interpretation
    Apply LIME and Grad-CAM on test images
    final_prediction ← model.predict(test_images)
    print("Final prediction:", final_prediction)
end
```

C. Mathematical Foundations of CNN in Physics-Informed MRI Feature Extraction for CHE

The proposed custom CNN model for CHE applications excels in extracting features from MRI images, starting with convolutional layers l, where the input undergoes a series of filter operations. Mathematically, this is represented as:

$$f_l^{ij} = \sigma\left(\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} W_l^{mn} \cdot x^{(i+m)(j+n)} + b_l\right) \tag{8}$$

where f_l^{ij} is the feature map output, σ denotes the ReLU activation function, W_l^{mn} are the filter weights, $x^{(i+m)(j+n)}$ is the input, and b_l is the bias. Subsequent pooling layers reduce spatial dimensions, formulated as

$$p_l^{ij} = \max_{k,l \in R_{ij}} f_{l-1}^{kl} \tag{9}$$

with p_l^{ij} being the pooling output and R_{ij} the pooling region. The flattened layer converts 2D feature maps into a 1D feature vector

$$v = \text{flatten}(p_L^{ij}) \tag{10}$$

The fully connected layers then process this vector, where each neuron's output d_k is given by

$$d_k = \sigma\left(\sum_{n=1}^{N} V_{kn} \cdot v_n + c_k\right) \tag{11}$$

with V_{kn} as weights and c_k as bias. The final classification is performed using a softmax function in the output layer:

$$\text{softmax}(d_k)_i = \frac{e^{d_{ki}}}{\sum_{j=1}^{C} e^{d_{kj}}} \tag{12}$$

This paper presents a probabilistic approach to the multiclass classification vital in CHE MRI diagnostics. This mathematical framework underpins the CHE-optimized CNN model, ensuring precision in feature extraction critical for accurate brain tumor detection.

D. Integrating PIDL for Enhanced MRI Feature Extraction

PIDL for MRI brain tumor analysis uniquely integrates the power of CNNs with physics-based constraints to enhance feature extraction, as shown in Figure 2. In this approach, CNN's ability to learn intricate, data-driven features from MR images is augmented by embedding physics-informed elements. The combined loss function, pivotal in guiding this integration, is mathematically represented as

$$\mathcal{L}_{\text{PIDL}} = \mathcal{L}_{\text{DL}}(f_{\text{CNN}}(x), y_{\text{class}}) + \lambda\mathcal{L}_{\text{Physics}}(f_{\text{CNN}}(x), y_{\text{physics}}) \tag{13}$$

This formulation entails a dual-focus loss function, where $\mathcal{L}_{\text{DL}}$ is the traditional DL loss for classification accuracy, and $\mathcal{L}_{\text{Physics}}$ introduces the physics-informed constraints (ϕ) into the learning process. The parameter λ strategically balances these two aspects. By balancing data-driven learning with physics-informed constraints, PIDL enhances the reliability and clinical relevance of MRI analyses for brain tumor detection within the CHE framework.

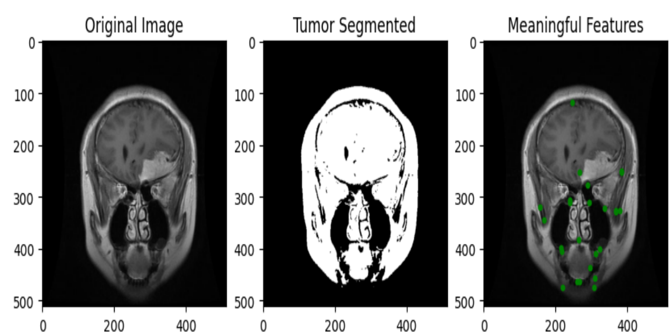

Fig. 2: PIDL-based meaningful feature extraction

E. Hybrid PIDL Loss Function in CHE: Optimizing Classification Accuracy and Physics-Based Insights

In the CHE-focused approach, the PIDL model's hybrid loss function adeptly manages the dual objectives of tumor classification and physics-informed analysis. It's articulated as:

$$\mathcal{L} = \mathcal{L}_{\text{Classification}} + \lambda\mathcal{L}_{\text{Physics}} \tag{14}$$

This article has been accepted for publication in IEEE Transactions on Consumer Electronics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCE.2024.3443203

Here, $\mathcal{L}$ represents the total loss function, $\mathcal{L}_{\text{Classification}}$ is the classification loss, such as categorical cross-entropy, used for tumor classification, and $\mathcal{L}_{\text{Physics}}$ is the loss associated with the physics-informed task, like mean squared error for a regression-based physical output. The hyperparameter λ balances these two components of the loss function. This can be further detailed as

$$\mathcal{L} = -\sum_{c=1}^{M} y_{o,c} \log(p_{o,c}) + \lambda \frac{1}{N} \sum_{i=1}^{N} (y_i^{(\text{Physics})} - \hat{y}_i^{(\text{Physics})})^2 \tag{15}$$

where M is the number of classes in the classification task, $y_{o,c}$ indicates if class label c is the correct classification for observation o, $p_{o,c}$ is the predicted probability that observation o is of class c, N is the number of samples in the physics-informed task, $y_i^{(\text{Physics})}$ is the actual value for the physics-informed output, and $\hat{y}_i^{(\text{Physics})}$ is the predicted value for the physics-informed output. This sophisticated loss framework ensures the model's accuracy of tumor classification and adherence to physics principles, which is crucial for advancing diagnostic precision in CHE.

F. Enhancing MRI Brain Tumor Analysis Interpretability with PIDL and LIME in CHE

The research developed a sophisticated PIDL model complemented by LIME to enhance the interpretability of AI-driven diagnostic processes significantly. This integration is crucial in making the AI's decision-making transparent and comprehensible to healthcare professionals, as depicted in Figure 3. The dual-outputs are mathematically defined as:

$$M(x) = \{M_{\text{class}}(x), M_{\text{phys}}(x)\} \tag{16}$$

where $M_{\text{class}}(x)$ utilizes a CNN for feature extraction and classification:

$$M_{\text{class}}(x) = \text{softmax}(W_{\text{class}} \cdot f_{\text{CNN}}(x) + b_{\text{class}}) \tag{17}$$

and $M_{\text{phys}}(x)$ applies physics-informed constraints to ensure the model adheres to established physical laws:

$$M_{\text{phys}}(x) = W_{\text{phys}} \cdot f_{\text{CNN}}(x) + b_{\text{phys}} \tag{18}$$

The integrated loss function $\mathcal{L}$ combines the classification loss $\mathcal{L}_{\text{class}}$ and the physics-based loss $\mathcal{L}_{\text{phys}}$, with a balancing parameter λ:

$$\mathcal{L} = \mathcal{L}_{\text{class}}(M_{\text{class}}(x), y_{\text{class}}) + \lambda \mathcal{L}_{\text{phys}}(M_{\text{phys}}(x), y_{\text{phys}}) \tag{19}$$

To enhance model transparency and support educational use in clinical settings, LIME is employed to elucidate the AI's inferential process:

$$E(x) = \arg\min_{g \in G} \left(\sum_{z \in Z} \pi_x(z)(M(z) - g(z))^2 + \Omega(g) \right) \tag{20}$$

This formulation ensures that g, a simpler model, approximates the complex behavior of M within a localized vicinity around the input x. The approximation focuses on balancing local fidelity—where $g(z)$ closely mirrors $M(z)$ for perturbations z within a neighborhood defined by $\pi_x(z)$—and the simplicity of g, controlled by $\Omega(g)$. This approach not only provides insights into the priorities of the model during classification but also makes these insights accessible and understandable to clinicians, effectively bridging the gap between advanced AI techniques and practical clinical applications. This rigorous methodological approach ensures that AI-driven diagnostics are effective but also transparent and accountable, fostering trust among healthcare providers and facilitating deeper integration into clinical workflows.

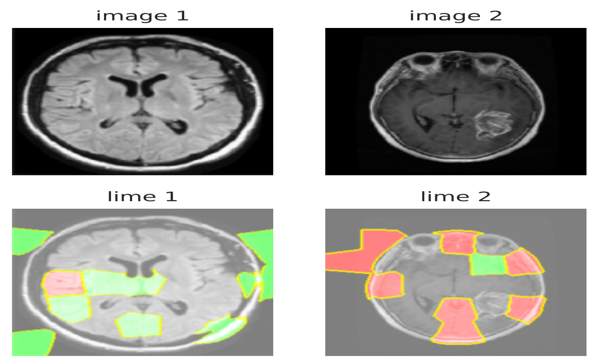

Fig. 3: LIME Visualization for Interpreting Brain Tumor

G. Leveraging Grad-CAM for Interpretability in CHE MRI Brain Tumor Detection

Figure 4 illustrates the application of Grad-CAM to visualize regions of interest in MRI images, which play a critical role in the predictions made by the DL model. Grad-CAM works by first performing a forward pass through the CNN to produce feature maps $f(x)$ for a given input image x. The class score S_c for the target class c is then computed as follows:

$$S_c = \sum_{i=1}^{C} w_{c,i} \cdot f_i(x) \tag{21}$$

Subsequently, the gradient of S_c is computed concerning $f(x)$, and global average pooling across the feature map dimensions to calculate neuron importance weights α_c^k:

$$\alpha_c^k = \frac{1}{Z} \sum_i \sum_j \frac{\partial S_c}{\partial f_k(x)_{ij}} \tag{22}$$

Where Z is the total number of elements in the feature map. The Grad-CAM heatmap H_c, which highlights the influential regions for the decision of the class c, is obtained using the ReLU activation function to focus only on the features with a positive influence:

$$H_c = \text{ReLU}\left(\sum_{k=1}^{C} \alpha_c^k \cdot f^k(x) \right) \tag{23}$$

This heatmap overlaps the original diagnostic image, enhancing interpretability by visually representing areas significantly impacting the model's prediction. Integrating Grad-CAM into the diagnostic process allows clinicians and healthcare professionals to verify the AI's focus areas visually, making the model's complex decision-making process accessible and interpretable. This technique not only aids in building trust in AI-powered diagnostic tools but also facilitates a deeper understanding of model behavior, which is crucial for clinical acceptance and real-world application. It represents a significant advance towards transparent and interpretable AI-driven diagnostics in Consumer Health Electronics, ensuring that the methodologies are scientifically robust and practically valuable in medical settings.

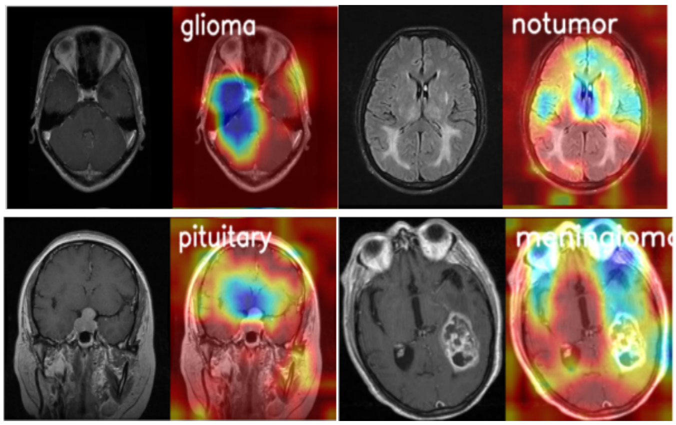

Fig. 4: Grade-Cam Visualization for Interpreting Brain Tumor

IV. Dataset and Experimental Result Analysis

A. Dataset Description

The research used a dataset of 6,771 MRI brain scans sourced from the Kaggle data repository for brain tumor classification [32]. This dataset encompasses four categories: glioma (1,621 images), meningioma (1,645 images), pituitary tumors (1,775 images), and non-tumorous scans (2,000 images). The dataset was strategically divided for model training and validation, adhering to an 80-20 split. Consequently, 5,360 samples were allocated for training, forming the core dataset for the DL model. The remaining 1,311 samples were designated for model validation, enabling the assessment of the model's performance and accuracy in brain tumor classification.

B. Experimental Setup

These investigations were performed on a High-Performance Computing (HPC) system configured as an edge server to accommodate the rigorous computational requirements of data processing and deep learning. This system, configured with 31 GB of DDR4 RAM, could effectively manage computationally intensive machine-learning tasks and large-scale datasets. It was operating on the robust Linux Ubuntu distribution. The NVIDIA GeForce GTX 3070 GPU, which possessed 8 GB of DDR4 memory, was the critical element of the computational configuration. The Jupyter Notebook integrated development environment was utilized to conduct model development and training meticulously. This was executed with the uttermost care. The data pipeline was optimized to optimize throughput and minimize input/output constraints, ensuring seamless data flow during training. Accuracy, precision, recall, and loss were meticulously monitored during the exercise. Hyperparameters, including the learning rate, were dynamically adjusted using these metrics to optimize model performance, showcasing the model's adaptability and ability to respond to changing conditions. The PIDL model's effectiveness and reliability in diagnosing brain tumors were demonstrated by the model's attainment of the most optimal results on the validation dataset, which was facilitated by this meticulous and adaptive approach.

C. Graphical Representations of Model Performance in Brain Tumor Detection

Figure 5 illustrates the relationship between physics constraints loss and accuracy over training epochs. On the x-axis, the epochs are represented, while the y-axis simultaneously depicts loss and accuracy. Observing the physics loss graph, a clear trend is noticeable where the loss value decreases progressively as the number of epochs increases. Conversely, in the physics accuracy graph, a steady increase in model accuracy correlates positively with the increasing number of epochs. This demonstrates the model's improving performance over time during the training phase. However, Figure 6 shows the PIDL training and validation classification loss and accuracy where the model performs smoothly over epochs, such as increasing accuracy and decreasing loss.

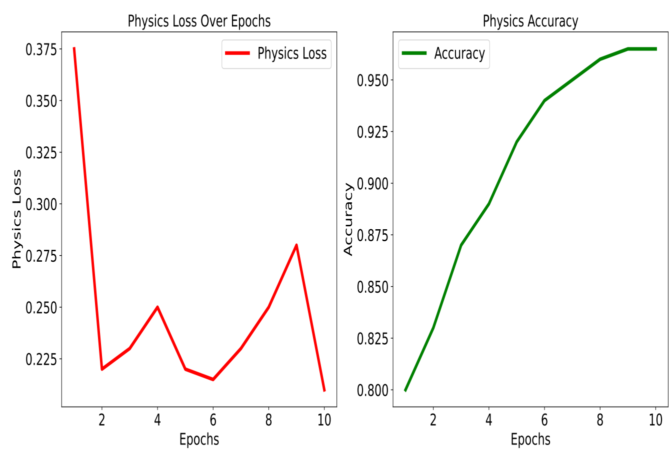

Fig. 5: Physics Constrains loss and accuracy

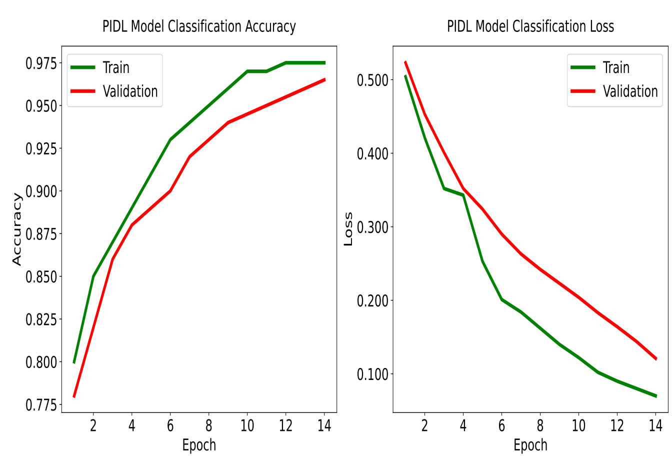

Fig. 6: PIDL model classification loss and accuracy

Figure 7 showcases a multi-class confusion matrix for the PIDL model employed in classifying brain tumors, including gliomas, meningiomas, pituitary tumors, and non-tumorous cases. The confusion matrix serves as an essential instrument for appraising the efficacy of classification models. It provides a visual depiction of the comparison between the actual and predicted classifications, facilitating a comprehensive evaluation of the model's predictive accuracy.

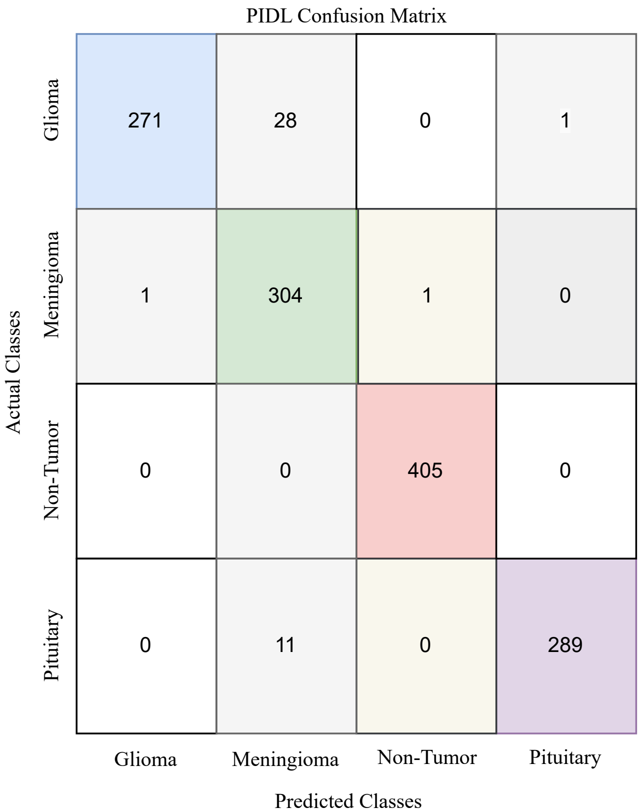

Fig. 7: PIDL model Confusion matrix

TABLE II: Summary of AIGC, leveraging physics-informed deep learning (PIDL) model performance

Categories of Tumors	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1 Score (%)**
Glioma	96	98	97	95
Meningioma	98	97	96	96
Pituitary	95	97	95	95
Non-tumorous	95	96	98	97

V. Conclusion & Future work

collaborative study has proven that merging PIDL and XAI approaches dramatically advances MRI-based medical diagnoses, as evidenced by AI-generated content. PIDL, a novel approach that integrates physical laws and constraints into the learning process of deep neural networks, has significantly enhanced the accuracy of the models. In the same way, explainability tools like LIME and Grad-CAM have improved the capability to understand and trust AI-generated diagnostics. The proposed PIE-AIGC framework has some limitations because this experiment utilized a single MRI dataset from a public repository, Kaggle. To generalize this work, future scholars should emphasize diverse medical datasets like CT scans and X-rays to extract some meaningful features. However, the existing PIE-AIGC framework outperforms MRI diagnosis in the healthcare industry. It focuses on explainability, accuracy, and integrating the physics principle in medical imaging.

Acknowledgment

This work is supported partially by the NSF project number 2205773 at Tennessee State University, TN, USA, and partially by the Researchers Supporting Project number (RSP2024R32), King Saud University, Riyadh, Saudi Arabia.

References

[1] C. K. Wu, C.-T. Cheng, Y. Uwate, G. Chen, S. Mumtaz, and K. F. Tsang, "State-of-the-art and research opportunities for next-generation consumer electronics," *IEEE Transactions on Consumer Electronics*, vol. 69, no. 4, pp. 937–948, 2023.

[2] S. Qiao, S. Pang, P. Xie, W. Yin, S. Yu, H. Gui, M. Wang, and Z. Lyu, "Hcmmnet: Hierarchical conv-mlp-mixed network for medical image segmentation in metaverse for consumer health," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 2078–2089, 2024.

[3] K. Doulani, A. Rajput, A. Hazra, M. Adhikari, and A. K. Singh, "Explainable ai for communicable disease prediction and sustainable living: Implications for consumer electronics," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 2460–2467, 2024.

[4] S. He, C. Du, and M. S. Hossain, "6g-enabled consumer electronics device intrusion detection with federated meta-learning and digital twins in a meta-verse environment," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 3111–3119, 2024.

[5] S. Ali, B. Sahoo, A. Zelikovsky, P.-Y. Chen, and M. Patterson, "Benchmarking machine learning robustness in covid-19 genome sequence classification," *Scientific Reports*, vol. 13, no. 1, p. 4154, 2023.

[6] A. Ashraf, Z. Qingjie, W. H. K. Bangyal, and M. Iqbal, "Analysis of brain imaging data for the detection of early age autism spectrum disorder using transfer learning approaches for internet of things," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 4478–4489, 2024.

[7] S. Hossain, A. Chakrabarty, T. R. Gadekallu, M. Alazab, and M. J. Piran, "Vision transformers, ensemble model, and transfer learning leveraging explainable ai for brain tumor detection and classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 3, pp. 1261–1272, 2024.

[8] P. Udayakumar and R. Subhashini, "Diffusion mri preprocessing for deep learning analysis of brain structural connectivity," in *2023 12th International Conference on Advanced Computing (ICoAC)*, 2023, pp. 1–5.

[9] W. Jin, X. Li, M. Fatehi, and G. Hamarneh, "Guidelines and evaluation of clinical explainable ai in medical image analysis," *Medical Image Analysis*, vol. 84, p. 102684, 2023.

[10] R. Kumar, D. Javeed, A. Aljuhani, A. Jolfaei, P. Kumar, and A. K. M. N. Islam, "Blockchain-based authentication and explainable ai for securing consumer iot applications," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 1145–1154, 2024.

[11] M. Doneva, "Mathematical models for magnetic resonance imaging reconstruction: An overview of the approaches, problems, and future research areas," *IEEE Signal Processing Magazine*, vol. 37, no. 1, pp. 24–32, 2020.

[12] M. Bhatia, S. Bhatia, M. Hooda, S. Namasudra, and D. Taniar, "Analyzing and classifying mri images using robust mathematical modeling," *Multimedia Tools and Applications*, vol. 81, no. 26, pp. 37 519–37 540, 2022.

[13] S. Larivière, Ş. Bayrak, R. V. de Wael, O. Benkarim, P. Herholz, R. Rodriguez-Cruces, C. Paquola, S.-J. Hong, B. Misic, A. C. Evans *et al.*, "Brainstat: A toolbox for brain-wide statistics and multimodal feature associations," *NeuroImage*, vol. 266, p. 119807, 2023.

[14] Y.-H. Kim and J. Lee, "Image feature and noise detection based on statistical hypothesis tests and their applications in noise reduction," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 4, pp. 1367–1378, 2005.

[15] A. Shamaei, J. Starcukova, and Z. Starcuk Jr, “Physics-informed deep learning approach to quantification of human brain metabolites from magnetic resonance spectroscopy data,” *Computers in Biology and Medicine*, vol. 158, p. 106837, 2023.

[16] M. Sarabian, H. Babaee, and K. Laksari, “Physics-informed neural networks for brain hemodynamic predictions using medical imaging,” *IEEE transactions on medical imaging*, vol. 41, no. 9, pp. 2285–2303, 2022.

[17] C. Sun, X. Fan, and D. Zhao, “Jpeg decoding with nonlinear inverse transform network and progressive recurrent residual network,” *IEEE Transactions on Consumer Electronics*, vol. 69, no. 3, pp. 499–509, 2023.

[18] Y. Wang, Y. Pan, M. Yan, Z. Su, and T. H. Luan, “A survey on chatgpt: Ai–generated contents, challenges, and solutions,” *IEEE Open Journal of the Computer Society*, vol. 4, pp. 280–302, 2023.

[19] X. Liu, G. Yang, Y. Zhou, Y. Yang, W. Huang, D. Ding, and J. Wu, “Fine-grained multi-modal fundus image generation based on diffusion models for glaucoma classification,” in *International Conference on Multimedia Modeling*. Springer, 2024, pp. 58–70.

[20] T. Zhou, X. Chen, Y. Shen, M. Nieuwoudt, C.-M. Pun, and S. Wang, “Generative ai enables eeg data augmentation for alzheimer’s disease detection via diffusion model,” in *2023 IEEE International Symposium on Product Compliance Engineering-Asia (ISPCE-ASIA)*. IEEE, 2023, pp. 1–6.

[21] J. Li, J. Zhong, Z. Li, Y. Xiao, S. Wang, and J. LI, “Ectopic pituitary neuroendocrine tumor: A case report written with the help of chatgpt,” *Cureus*, vol. 15, no. 10, 2023.

[22] G. Yang, Q. Ye, and J. Xia, “Unbox the black-box for the medical explainable ai via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond,” *Information Fusion*, vol. 77, pp. 29–52, 2022.

[23] W. Lin, T. Tong, Q. Gao, D. Guo, X. Du, Y. Yang, G. Guo, M. Xiao, M. Du, X. Qu *et al.*, “Convolutional neural networks-based mri image analysis for the alzheimer’s disease prediction from mild cognitive impairment,” *Frontiers in neuroscience*, vol. 12, p. 777, 2018.

[24] S. Pattanaik, C. Chakraborty, S. Behera, S. K. Majhi, and S. K. Pani, “An miot framework of consumer technology for medical diseases prediction,” *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 3754–3761, 2024.

[25] S. Feng, H. Zhao, F. Shi, X. Cheng, M. Wang, Y. Ma, D. Xiang, W. Zhu, and X. Chen, “Cpfnet: Context pyramid fusion network for medical image segmentation,” *IEEE transactions on medical imaging*, vol. 39, no. 10, pp. 3008–3018, 2020.

[26] Y. Li, J. Zhao, Z. Lv, and J. Li, “Medical image fusion method by deep learning,” *International Journal of Cognitive Computing in Engineering*, vol. 2, pp. 21–29, 2021.

[27] Q. Zuo, H. Wu, C. L. P. Chen, B. Lei, and S. Wang, “Prior-guided adversarial learning with hypergraph for predicting abnormal connections in alzheimer’s disease,” *IEEE Transactions on Cybernetics*, vol. 54, no. 6, pp. 3652–3665, 2024.

[28] A. M. D. Simo, A. T. Kouanou, V. Monthe, M. K. Nana, and B. M. Lonla, “Introducing a deep learning method for brain tumor classification using mri data towards better performance,” *Informatics in Medicine Unlocked*, vol. 44, p. 101423, 2024.

[29] W. Zhao, Z. Hu, A. F. Kazerooni, G. Körzdörfer, M. Nittka, C. Davatzikos, S. E. Viswanath, X. Wang, C. Badve, and D. Ma, “Physics-informed discretization for reproducible and robust radiomic feature extraction using quantitative mri,” *Investigative Radiology*, pp. 10–1097, 2023.

[30] D. Reyes and J. Sánchez, “Performance of convolutional neural networks for the classification of brain tumors using magnetic resonance imaging,” *Heliyon*, vol. 10, no. 3, 2024.

[31] Z. Zhou, A. Ma, Q. Feng, R. Wang, L. Cheng, X. Chen, X. Yang, K. Liao, Y. Miao, and Y. Qiu, “Super-resolution of brain tumor mri images based on deep learning,” *Journal of Applied Clinical Medical Physics*, vol. 23, no. 11, p. e13758, 2022.

[32] M. NICKPARVAR, “Brain tumor mri dataset,” 2021, accessed: 2023-04-08. [Online]. Available: https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset

Al Amin (Graduate Student Member, IEEE) is a Ph.D. student in Electrical and Computer Engineering at Tennessee State University, USA. His research focuses on applying AI, Machine Learning, and Deep Learning to healthcare, especially in medical image processing. Al Amin earned his M.Sc. in Electronic and Electrical Engineering from the University of Portsmouth, UK, in 2022.

He has developed physics-informed deep learning models and explainable AI frameworks for medical imaging and contributed to AI-powered safety systems in railroad crossings.

Kamrul Hasan (Member, IEEE) is an Assistant Professor in the Department of Electrical and Computer Engineering (ECE) at Tennessee State University (TSU), Nashville, TN, USA. He earned his Ph.D. in Computational Modeling and Simulation Engineering from Old Dominion University, Norfolk, VA, in 2020.

His research interests encompass a broad range of topics within the realm of Cyber-Physical Systems (CPS), including CPS Security, Computer Networks and Security, and Software-Enabled Network Security. In addition to these areas, Dr. Hasan’s work is significantly centered on the integration of Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) within cybersecurity and automation frameworks. A key focus of his research lies in the healthcare domain, where he has pioneered the application of Explainable AI (XAI) in medical image processing, enhancing the interpretability and transparency of AI-driven diagnostic tools.

He has also made significant contributions by developing physics-based mathematical models that integrate seamlessly with deep learning techniques, providing robust and accurate solutions for analyzing and interpreting medical images. His expertise in this area has been instrumental in advancing the application of AI in healthcare, particularly in enhancing diagnostic precision and improving patient outcomes.

Since 2020, he has served as Principal Investigator (PI) or co-PI on four research projects funded by prestigious organizations such as the National Science Foundation (NSF) and the United States Department of Transportation (USDOT). His scholarly output includes more than 30 peer-reviewed journal and conference papers, a book chapter, and a U.S. patent, reflecting his active contribution to advancing knowledge in his field.

M. Shamim Hossain (Senior Member, IEEE) received a Ph.D. in electrical and computer engineering from the University of Ottawa, Ottawa, ON, Canada, in 2009. He is currently a Professor at the Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia. He is also an adjunct professor at the School of Electrical Engineering and Computer Science, University of Ottawa, ON, Canada. His research interests include cloud networking, smart environment (smart city, smart health), AI, deep learning, edge computing, Internet of Things (IoT), multimedia for health care, and multimedia big data. He has authored and co-authored more than 380 publications.

He was the recipient of several awards, including the Best Conference Paper Award, The 2024 IEEE Communications Society Outstanding Paper Award, the 2016 ACM Transactions on Multimedia Computing, Communications and Applications (TOMM) Nicolas D. Georganas Best Paper Award, the 2019 King Saud University Scientific Excellence Award (Research Quality), and the Research in Excellence Award from the College of Computer and Information Sciences (CCIS), King Saud University (3 times in a row). He is the chair of the IEEE Special Interest Group on Artificial Intelligence (AI) for Health with the IEEE ComSoc eHealth Technical Committee. He is the Symposium Chair of Selected Areas in Communications (E-Health) with IEEE GLOBECOM 2024. He is the Track Chair of the IEEE International Conference on Consumer Electronics (ICCE 2024). He is the Technical Program Co-Chair of ACM Multimedia 2023. He is on the editorial board of the IEEE Transactions on Instrumentation and Measurement (TIM), IEEE Transactions on Multimedia (TMM), ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), IEEE Multimedia, IEEE Network, IEEE Wireless Communications, Journal of Network and Computer Applications (Elsevier). He is a senior member of the IEEE and a Distinguished Member of the ACM. He is an IEEE Distinguished Lecturer (DL). He is a highly Cited Researcher in Computer Science (Web of Science™).