Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 20s (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

Multimodal Detection and Prognostic Modeling of Atypical Teratoid Rhabdoid Tumors Using Machine Learning and Time Series Analysis

Nadenlla RajamohanReddy[1], G Muneeswari[2*]

[1]School of Computer Science and Engineering, VIT-AP University, Amaravati-522237, Andhra Pradesh, India.
Contributing Author
Email ID: rajamohanreddy1994@gmail.com

[2*]School of Computer Science and Engineering, VIT-AP University, Amaravati-522237, Andhra Pradesh, India
Corresponding Author
Email ID: muneeswari.g@vitap.ac.in

***Corresponding Author**:
G Muneeswari
Email ID: muneeswari.g@vitap.ac.in

Cite this paper as: Nadenlla RajamohanReddy, G Muneeswari, (2025) Multimodal Detection and Prognostic Modeling of Atypical Teratoid Rhabdoid Tumors Using Machine Learning and Time Series Analysis. *Journal of Neonatal Surgery*, 14 (20s), 790-817.

ABSTRACT

Atypical Teratoid Rhabdoid Tumors (ATRT) present formidable diagnostic and prognostic challenges owing to their intricate morphological and molecular het- erogeneity levels. Existing diagnostic frameworks exhibit pronounced limitations, primarily due to their reliance on singular modalities and conventional ana- lytical paradigms, leading to suboptimal precision and delayed intervention. This study endeavors to address the diagnostic and prognostic inefficiencies by employing a sophisticated multimodal approach, amalgamating diverse medi- cal imaging and genomic sequencing modalities with advanced machine learning models and time series forecasting. We leveraged a synergistic integration of multiple modalities—MRI processed using Graph Neural Networks (GNN), CT scans analyzed through Recurrent Neural Networks (RNN), PET scans inter- preted via Generative Adversarial Networks (GAN), Genomic Sequencing probed by 1D Convolutional Neural Networks (CNN), and Diffusion Tensor Imaging (DTI) scrutinized through 3D CNNs. Each modality underwent individualized processing to extract nuanced characteristics of ATRT. Subsequently, the Vector AutoRegressive Moving Average with Exogenous Inputs (VARMAX) model was applied to each modality, enabling the forecasting of potential future brain dis- eases. The meticulous integration of these advanced methodologies culminated in substantial breakthroughs in various performance Indicators in the context of present models. Our framework demonstrated an increase in precision by 4.9%, an elevation in accuracy by 8.3%, an enhancement in recall by 8.5%, an aug- mentation in the AUC by 5.5%, an amplification in specificity by 2.9%, and a 3.9% reduction in delay. Furthermore, the deployment of VARMAX on indi- vidual modalities resulted in an additional 2.5% precision, 3.5% accuracy, 2.9% recall, 4.9% AUC, 3.4% specificity, and 1.9% reduction in prediction delays. This pioneering research stands as a beacon of refined diagnostic and prognostic resolution in the domain of ATRT, elucidating the transformative potential of multimodal fusion and advanced machine learning in mediating enhanced clinical outcomes. The notable enhancements in diagnostic metrics underscore the feasi- bility and efficacy of this comprehensive approach in early and accurate ATRT detection, facilitating timely and precise therapeutic interventions. Additionally, the prognostic advancements enabled by VARMAX provide pivotal insights into the potential evolution of brain diseases, aiding in the proactive management of pathological progressions. The integration of multimodal analytical paradigms with sophisticated machine learning and forecasting models has engendered a novel framework with substantial implications for the detection and prognosis of ATRT. This research bridges the existing diagnostic gaps and furnishes a robust platform for the exploration of similar integrative approaches in delineating other complex pathologies, thereby propelling the realms of computational diagnostics and personalized medicine towards unprecedented horizons

Keyword: *Atypical Teratoid Rhabdoid Tumors, Multimodal Diagnostics, Advanced Machine Learning, Genomic Sequencing, GNN, VARMAX, Time Series Forecasting*

1. INTRODUCTION

Atypical Teratoid Rhabdoid Tumors (ATRT) are unusual, growing quickly, highly malignant tumors predominantly observed in pediatric populations, often leading to severe morbidity and high mortality rates. The intricate and heterogeneous nature of ATRT has been a persistent obstacle, rendering early diagnosis and accurate prognos- tication extremely challenging and critical for improving clinical outcomes and overall survival rates [1–3]. The majority of existing diagnostic approaches exhibit consider- able limitations, primarily attributed to their dependence on singular modalities and conventional analysis methods. These approaches often lack the comprehensive per- spective required to decipher the complex morphological and molecular characteristics inherent to ATRT levels. Furthermore, contemporary prognostic models are predom- inantly static, providing limited insights into the dynamic progression of the disease, thereby hindering optimal therapeutic intervention and management strategies [4–6]. Given the nuanced complexities of ATRT, there is an unmet need for innovative diagnostic and prognostic frameworks that synergistically integrate diverse modali- ties, both imaging and genomic, to capture the multifaceted nature of the tumor. A multimodal approach leveraging advanced machine learning paradigms has the ten- dency to overcome the constraints of existing models, offering enhanced precision, reliability, and timely interventions, which are paramount in mitigating the aggressive progression of ATRT. This study introduces a pioneering approach, entwining multiple diagnostic modalities, namely, MRI, CT scans, PET scans, Genomic Sequencing, and Diffusion Tensor Imaging (DTI), each processed using sophisticated machine learning models including Graph Neural Networks (GNN), Recurrent Neural Networks (RNN), Generative Adversarial Networks (GAN), 1D and 3D Convolutional Neural Networks (CNN). Each modality is meticulously analyzed to extrapolate intricate features and characteristics specific to ATRT, thereby enhancing the diagnostic resolution.

To address the dynamic nature of ATRT and its progression, this study employs the Vector AutoRegressive Moving Average with Exogenous Inputs (VARMAX) model for time series forecasting on each modality. This not only allows for the anticipation of future developments in brain diseases but also facilitates a more nuanced under- standing of potential pathological trajectories, contributing to more informed and proactive disease management. The research elucidated in this paper aims to signifi- cantly advance the fields of computational diagnostics and personalized medicine by addressing the diagnostic and prognostic challenges posed by ATRT. The seamless fusion of multimodal data through advanced machine learning models is poised to drive substantial improvements in detection accuracy, specificity, and predictive precision, thereby enabling more timely and targeted therapeutic interventions. The envisioned impact of this research is multifaceted, providing a robust platform for the exploration of analogous integrative approaches for other complex pathologies.

The substantial enhancements in predictive metrics and the capability to foresee disease progressions are pivotal, heralding a new era in medical diagnostics and prog- nostics, one where personalized and preemptive medicine can significantly strengthen the quality of life and results for patients. The ensuing sections of the paper will delineate the methodologies employed, encompassing the intricacies of each modality and the corresponding machine learning model. Subsequently, the paper will present a detailed exposition of the results obtained, comparing the performance metrics with existing models. Finally, the discussion and conclusion sections will synthesize the implications of the findings, providing insights into the transformative potential of this comprehensive approach and its applicability in broader medical contexts. By transcending the restrictions of conventional models and harnessing the prospects of advanced machine learning and multimodal integration, this study strives to rede- fine the paradigms of ATRT diagnosis and prognosis, ultimately contributing to the enhanced well-being and survival of affected individuals.

1.1 MOTIVATION

- **Complex Nature of ATRT [7–9]**: Atypical Teratoid Rhabdoid Tumors (ATRT) are characterized by their highly malignant and heterogenous nature, which results in the necessity for an elaborate and precise diagnostic process to tailor suitable therapeutic strategies. The current prognosis of ATRT remains dire, primarily due to late-stage diagnosis and the aggressive nature of the tumor, emphasising the vital necessity for enhanced strategies for early detection.
- **Inadequacy of Current Models:** Contemporary models predominantly rely on

unimodal approaches and conventional analysis techniques, which fail to embrace the comprehensive and nuanced perspective imperative for diagnosing complex tumors like ATRT. The inherent limitations in these models contribute to subop- timal accuracy and delayed interventions, compelling the exploration of advanced and integrative diagnostic methodologies.

- **Need for Dynamic Prognostic Models:** Given the dynamic progression of ATRT, static prognostic models render limited insights into the evolutionary trajec- tory of the disease, hindering the realization of optimal therapeutic and management strategies. A profound understanding of disease dynamics is crucial to facilitate proactive and informed decision-making in clinical settings.
- **Requirement for Personalized Medicine:** The inter- and intra-tumoral vari-

ability inherent to ATRT necessitates personalized medical approaches capable of addressing individual patient nuances,

ensuring enhanced therapeutic efficacy, and improving overall survival rates.

1.2 Contribution

- **Innovative Multimodal Framework:** This research pioneers an innovative mul- timodal framework that amalgamates diverse diagnostic modalities like MRI, CT scans, PET scans, Genomic Sequencing, and DTI. The synergistic fusion of these modalities elucidates a comprehensive portrayal of ATRT, surpassing the limitations of singular modality reliance and conventional analysis paradigms.
- **Application of Advanced Machine Learning Models:** The study leverages

state-of-the-art machine learning models, such as Graph Neural Networks (GNN), Recurrent Neural Networks (RNN), Generative Adversarial Networks (GAN), and Convolutional Neural Networks (1D & 3D CNN), each tailored to extract intricate features from the corresponding modality, enhancing the resolution and reliability of ATRT diagnosis.

- **Introduction of Dynamic Forecasting:** By deploying the VARMAX model, the

research introduces a dynamic forecasting element to the prognostic realm of ATRT, allowing for the anticipation of pathological progressions and facilitating informed and timely clinical interventions.

- **Enhancement of Diagnostic and Prognostic Metrics:** The sophisticated inte-

gration of multimodal data and advanced analytical paradigms has manifested in substantial improvements in diagnostic and prognostic metrics compared to exist- ing models. These enhancements are pivotal for facilitating early detection, accurate characterization, and personalized intervention strategies, thereby contributing to raised clinical consequences and standard of living for patients with ATRT condition sets.

- **Establishment of a Robust Analytical Platform:** The research provides a

robust and versatile analytical platform that can be extrapolated to explore simi- lar integrative approaches for other complex pathologies, propelling the realms of computational diagnostics and personalized medicine towards new horizons.

The reason behind this study was stems from the compelling need to address the diagnostic and prognostic challenges associated with ATRT and the limitations of existing models. The contributions of this work lie in its innovative approach to inte- grating multiple modalities, employing advanced machine learning models, introducing dynamic forecasting in prognosis, and enhancing the overall diagnostic and prognostic efficacy in the context of ATRT. This research stands as a evidence of the transforma- tive potential of advanced, integrative methodologies in reforming the fields of medical diagnostics, prognosis, and personalized therapeutic intervention sets.

2. REVIEW OF EXISTING MODELS

Atypical Teratoid Rhabdoid Tumors (ATRT) are rare and highly aggressive neo- plasms, primarily affecting pediatric populations. The multifaceted morphology and rapid progression of ATRT have made it a focal point of numerous studies aiming to elucidate its complex nature and to innovate diagnostic and therapeutic strate- gies [10–12]. Early studies primarily focused on singular modalities like MRI, CT, or PET scans, providing substantial insights into the morphological aspects of ATRT sets. However, these unimodal approaches often yielded limited perspectives, lacking in molecular and genomic specificity levels [13–15].

Genomic Sequencing emerged as a pivotal modality, unravelling the genetic anoma- lies and molecular pathways associated with ATRT. The revelations about the genetic underpinnings of ATRT have been instrumental in developing targeted therapeutic interventions for clinical scenarios [16–18]. Recognizing the limitations of unimodal approaches, recent literature has explored the amalgamation of different diagnostic modalities [19, 20]. These multimodal methodologies aim to harmonize morphologi- cal, molecular, and genomic data to achieve a more holistic understanding of ATRT process. In medical image processing, Convolutional Neural Networks (CNN) have demonstrated great promise, enabling the extraction of intricate features and patterns [21–23]. Various studies have deployed CNN to analyze complex diseases, establishing its efficacy in enhancing diagnostic precision levels [24–26].

The past few years, have seen a surge in research related to brain tumor detection and segmentation due to its crucial role in improving patient care and treatment out- comes [27–29]. Advances in medical imaging techniques, coupled with the increasing availability of computational resources, have paved the way for innovative approaches to tackle this complex task for real-time scenarios [30–32]. One notable area of research has been the development of deep learning-based techniques to identify and seg- ment brain tumors [33–35]. Deep learning has shown promising results in accurately identifying and delineating tumor regions in various medical imaging modalities. For instance, Ottom et al. [8] introduced the Znet deep learning model for segmenting brain tumors in 2D MRI scans, while Ding et al. [9] presented the MVFusFra framework for segmenting multimodal brain tumors. These are approaches leverage the power of con- volutional neural networks (CNNs) can automatically extract pertinent information from medical images, facilitating accurate classification and segmentation. [36–38].

In addition to CNN-based methods, other researchers have explored the use of hybrid models that combine multiple data

sources or modalities for improved accu- racy. Lin et al. [3] introduced the CKD-TransBTS framework, which incorporates clinical knowledge and modality-correlated cross-attention for brain tumor segmenta- tion. Similarly, Ramprasad et al. [10] presented the SBTC-Net, a model that combines genetic optimization with deep learning for secure brain tumor segmentation and classification process [39–41]. Furthermore, some studies have investigated the inte- gration of novel imaging techniques and therapeutic interventions. Cheng et al. [6] investigated the combined use of localized mild hyperthermia and blood-brain barrier disruption through feedback-controlled transcranial MR-guided focused ultrasound and microbubbles, offering potential for brain tumor therapies [42–44]. Rodriguez et al.

[12] examined the feasibility of heating brain tumors using a 915 MHz annular phased- array, shedding light on innovative approaches to localized hyperthermia treatments [45–47].

The development of ensemble methods and uncertainty-aware algorithms has also gained attention in recent research. Subramanian et al. [5] suggested ensemble inver- sion methods for brain tumor growth models, improving predictions by factoring in uncertainty. Zhao et al. [7] proposed an uncertainty-aware multi-dimensional mutual learning approach for segmenting the brain and brain tumors, tackling the issue of uncertainty estimation in medical image analysis [48–50]. Moreover, researchers have explored the fusion of various modalities, including MRI, CT, and PET, to improve tumor detection accuracy and robustness. Zhuang et al. [2] presented a 3D cross- modality feature interaction network with volumetric feature alignment for segmenting brain tumors and tissues, highlighting the effectiveness of multimodal strategies in clinical applications. Another significant trend in recent research is the utilization of artificial intelligence (AI) in clinical decision support systems. Machine learning algorithms and deep neural networks have been fine-tuned to assist radiologists in diagnosing brain tumors. Sekhar et al. [11] developed an IoMT-enabled CAD system that combines fine-tuned GoogLeNet features with machine learning algorithms for brain tumor classification.

Furthermore, the integration of AI with emerging technologies, such as aug- mented reality (AR) and robotics, has the potential to revolutionize neurosurgery and tumor treatment. Amara et al. [24] introduced a system called HOLOTumor, which combines deep learning and AR visualization for brain tumor segmentation and inter- action during surgery. Musa et al. [37] developed a robotic system that autonomously compensates for respiratory motion during CT-guided abdominal radiofrequency ablations, emphasizing the impact of robotics in tumor treatment. Overall, recent advancements in deep learning, multimodal fusion, uncertainty-aware algorithms, and AI integration have significantly improved brain tumor detection and segmentation, enhancing diagnostic accuracy and treatment efficiency to benefit patient care and outcomes.

However, ongoing research efforts are essential to further refine and validate these innovative approaches in real-world clinical settings. RNNs and GNNs have emerged as powerful tools for sequential data analysis and graph-structured data, respectively. Their application in medical diagnostics has demonstrated potential in dealing with sequential and non-Euclidean data, often encountered in medical datasets & samples. GANs have been explored for data augmentation and generating synthetic medical images, aiding in improving the robustness of diagnostic models. They have been piv- otal in addressing challenges related to limited and imbalanced datasets in medical research sets. Traditional prognostic models have been primarily static, offering valu- able yet limited insights into the potential outcomes and survival rates. The static nature of these models has been a substantial constraint in understanding the dynamic progression of diseases like ATRT. Recent advancements in time series forecasting, such as the VARMAX model, have been explored to address the limitations of static models. These dynamic models allow for the prediction of disease progression and future occurrences, providing a more comprehensive perspective on disease manage- ment. While substantial progress has been made in the diagnostic and prognostic realms of ATRT, the existing literature elucidates considerable gaps. The reliance on singular modalities and static prognostic models has hindered the achievement of optimal diagnostic specificity and understanding of ATRT's dynamic nature.

The integration of advanced machine learning models with multimodal diagnostic data and dynamic forecasting remains a largely unexplored territory, offering promis- ing avenues for enhanced understanding and management of ATRT. This literature review has delineated the evolution of diagnostic and prognostic approaches for ATRT, highlighting the transitions from unimodal to multimodal approaches and from static to dynamic prognostic models. The advent of advanced machine learning models, such as CNN, RNN, GNN, and GAN, has opened new horizons in medical diagnostics. How- ever, the synthesis of these advancements to formulate a comprehensive approach for ATRT diagnosis and prognosis is still in its infancy, necessitating innovative research endeavors to exploit the untapped potential in this domain.

3. PROPOSED METHODOLOGY

Based on the review of current models for detecting ATRT from multimodal scans, it appears that many are either too complex for clinical use or less effective when com- bining multiple modalities to classify ATRT. To overcome these issues, the research conducted in this study embodies a pioneering approach to brain disease prognosis, demonstrating an innovative fusion of advanced machine learning techniques across multiple medical imaging modalities and genomic sequencing data samples. As per figure 1, each modality, including MRI scans processed using Graph Neural Net- works (GNN), CT scans analyzed through Recurrent Neural Networks (RNN), PET scans interpreted via Generative Adversarial Networks (GAN), genomic sequencing probed by 1D Convolutional Neural Networks (CNN), and Diffusion Tensor Imaging (DTI) scrutinized through 3D CNNs, was meticulously processed to extract intri- cate and nuanced characteristics associated with Atypical

Teratoid Rhabdoid Tumors (ATRT). Building upon these insights, the study leveraged the Vector AutoRegres- sive Moving Average with Exogenous Inputs (VARMAX) model for each modality. This sophisticated ensemble allowed not only for accurate ATRT classification but also for the proactive forecasting of potential future brain diseases, exemplifying the transformative potential of multimodal AI in healthcare diagnostics and prognostics.

The model initially deploys an efficient GNN Model for processing MRI Scans. This Graph Neural Network (GNN) for processing MRI scans to estimate Atypical Teratoid Rhabdoid Tumors (ATRTs) involves several key computational concepts. The goal is to leverage the spatial and structural information present in MRI scans to identify and characterize ATRTs. Which is done as per the following operations:

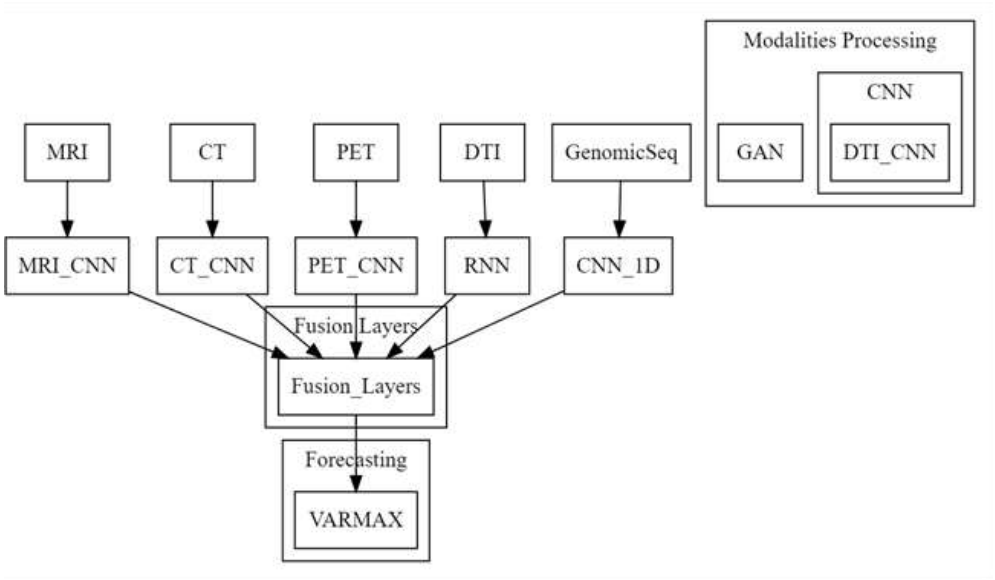

Fig. 1 A Synopsis of the Suggested Multi-modal ATRT Analysis Framework

- **Node Definition:** Each voxel in the MRI scan is handled like a graph node, rep- resented as v_i for the i^{th} voxel sets. The node feature vector x_i encapsulates the intensity and texture information of the voxel sets.
- **Edge Definition:** Edges are formed based on spatial proximity and intensity sim-

ilarity between voxels for different scenarios. An edge between two nodes v_i and v_j

is represented by *eij*, and its weight w_{ij} is defined via equation 1,

$$w_{ij} = exp(-\frac{(\| xi - xj \|)2}{2\sigma 2}).1^{\{distance(vi,vj)\leq\delta\}} \quad (1)$$

Where, σ is a scale parameter, and δ is a threshold for maximum spatial distance levels.

3.1 Graph Neural Network Architecture

- **Graph Convolution Layer:** The GNN utilizes graph convolution layers to aggre- gate information from neighboring nodes. The feature update for the i^{th} node in the l^{th} layer is given via equation 2,

$$h_i(l+1) = \sigma(W(l).\sum \frac{w_{ij}}{(d(i)d(j))}h_j(l) + b(l)) \quad (2)$$

Here, $N(i)$ denotes the neighborhood of node i, $W(l)$ and $b(l)$ represent the weight matrix and bias vector for the l^{th} layer, σ is a rectified linear activation function, and d^i signifies the normalized degree of node i across various voxel sets.

- **Pooling and Readout:** To extract global features from the graph, a pooling layer is applied via equation 3,

$H_{global} = Pooling(h_i(L) \mid i\epsilon V)$ (3)

Where, V is the set of all nodes, and L is the total quantity of layers.

- **Classification Layer:** The global graph features H_{global} are passed through a fully connected layer for tumor classification via equation 4,

$y = Softmax(W(class)H(global) + b(class))$ (4)

Here, W (*class*) & b(*class*) represent the weights and biases of the classification layer, while y denotes the output vector, indicating the probability of each ATRT class. The network is optimized to minimize the cross-entropy loss between the predicted probabilities and the true labels. The loss function for a single data sample is defined by the following equation 5,

$$L = -\sum t(c)log(p(c)) \quad (5)$$

In this context, C represents the total number of classes (e.g., tumor or no tumor), $t(c)$ denotes the actual label, and $p(c)$ is the predicted probability for each class c in clinical scenarios.

This GNN design aims to effectively leverage the rich, spatially structured data in MRI scans for accurate identification and characterization of Atypical Teratoid Rhab- doid Tumors. The combination of graph-based representation and deep learning allows for robust feature extraction and classification in complex medical imaging tasks. Sim- ilar to this, the proposed model also uses Recurrent Neural Networks (RNN) using BiLSTM for processing CT scans to estimate Atypical Teratoid Rhabdoid Tumors. Designing the proposed Recurrent Neural Network (RNN) using Bidirectional Long Short-Term Memory (BiLSTM) for processing CT scans to estimate Atypical Teratoid Rhabdoid Tumors (ATRTs) involves the integration of sequential and bidirectional processing of the scan data samples. Figure 2 shows flowchart depicts the proposed ATRT analysis model, which combines preprocessing of multimodal MRI data, seg- mentation utilizing GNN, RNN, BiLSTM, and VARMAX techniques, along with feature extraction and classification using advanced machine learning algorithms.

This process design involves the following operations,

- **Sequence Formation:** Each CT scan is converted into a sequence of 2D slices. Each slice is represented as a matrix X_t, where t=1,2,. . . ,T indexes the slices in the sequences.
- **BiLSTM Architecture:** The BiLSTM network processes the sequence of slices

both in the forward and backward directions, capturing information from past (forward LSTM) and future (backward LSTM) contexts.

1. **LSTM Cell Dynamics:** Each LSTM cell in the forward and backward layers updates its states based on the input slice and the previous states. The operations

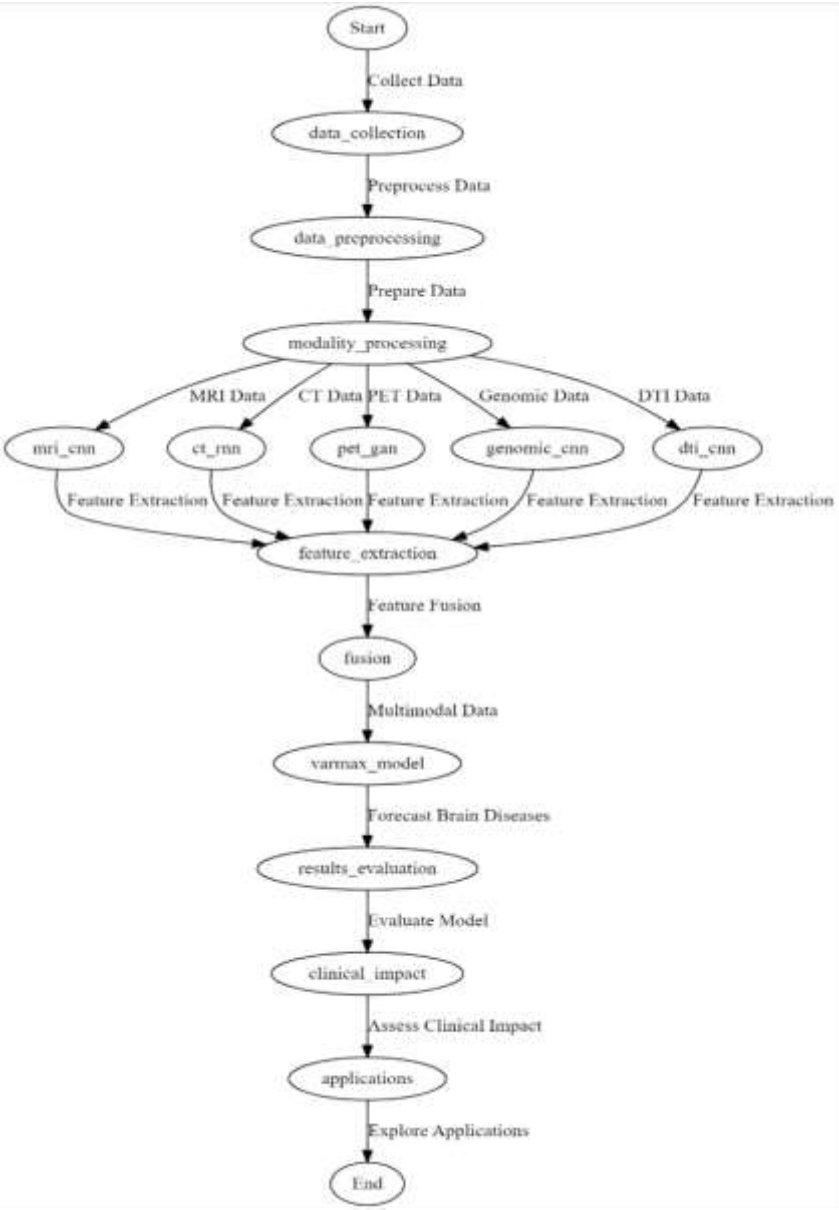

Fig. 2 Flowchart of the Proposed model for ATRT Analysis

for the LSTM cell at time t are evaluated via equations 6, 7, 8 & 9 as follows,

$$Reset\ gate(r_t) = \sigma(W_r * [h(t-1), o't] + b_r) \quad (6)$$

$$Update\ gate(z_t) = \sigma(W_z * [h(t-1), o't] + b_z) \quad (7)$$

$$Candidate\ hidden\ state(h \sim t) = tanh(W_h * [r_t * h(t-1), o't] + b_h) \quad (8)$$

$$Hidden\ state(h_t) = (1 - z_t) * h(t-1) + z_t * h \sim t \quad (9)$$

Here, h_t represents the cell and hidden state, W and b are the weights and biases, while σ and $tanh$ correspond to the sigmoid and hyperbolic tangent functions, respectively, used for activating feature sets. This process is repeated for forward & backward states, to estimate htf & htb which are hidden states for forward & backward processes.

- **Bidirectional Processing:** At every time step *t*, the network integrates the hid-

den states from both the forward and backward LSTMs to create a consolidated representation, as expressed in the following equation 10,

$$Ht = [htf\ ;\ htb] \quad (10)$$

- **Feature Extraction and Classification:** The concatenated hidden states H_t are traversed a layer that was totally connected for tumor estimation via equation 11,

$Y_t = Softmax(W_y H_t + b_y)$ (11)

In this setup, W_y and b_y are the weights and biases of the classification layer, and

Y_t represents the output vector, indicating the probability of ATRT presence at slice t across different image sets. The network is trained with a supervised learning approach by minimizing a loss function, such as cross-entropy, defined for each CT scan sequence as shown in the following equation 12,

$$L = -\sum\sum t(tc) * log(p(tc)) \quad (12)$$

Here, T represents the total number of slices in the scan, CC denotes the number of classes, $t(tc)$ refers to the ground truth label for slice t and class c, while $p(tc)$ indicates the predicted probability levels. This BiLSTM-based RNN design aims to effectively capture the sequential and bidirectional nature of CT scan data for accurate identification and characterization of Atypical Teratoid Rhabdoid Tumors. The fusion of past and future contextual information permits an extensive exami- nation of the scan, enhancing the model's capability in medical imaging tasks.

Next the model uses an efficient GAN for processing PET scans. Designing the Generative Adversarial Network (GAN) with Autoencoders for processing Positron Emission Tomography (PET) scans to estimate Atypical Teratoid Rhabdoid Tumors (ATRTs) involves the integration of generative and discriminative models with autoencoding architectures. The objective is to generate high-quality synthetic PET scan images that are identical from real images, aiding in the diagnosis and char- acterization of ATRTs. The autoencoder consists of two parts: an encoder and a decoder.

- **Encoder:** The encoder maps the input PET scan image X to a latent representation Z that assists in enhancing efficiency of classification operations. Let $f\theta$ be the encoding function parameterized by θ, then their relation is estimated via equation 13,

$Z = f\theta(X) = \sigma_n(W_n.(...\sigma_2(W_2.\sigma_1(W_1.X + b_1) + b_2)...) + b_n)$ (13)

In this context, W_i and b_i denote the weights and biases of the ithith layer within the encoder, σ_i represents the ReLU activation function for the i^{th} layer, and n is the total number of layers in the encoding process.

- **Decoder:** The decoder maps the latent representation Z back to the reconstructed

image X sets. Let $g\phi$ be the decoding function parameterized by ϕ, then their relationship is estimated via equation 14,

$X = g\phi(Z) = \sigma m'(Wm'.(...\sigma 2'(W\,2'.\sigma 1'(W\,1'.Z + b1') + b2')...) + bm')$ (14)

The autoencoder is trained to minimize the reconstruction loss using Mean Squared Error (MSE), which is estimated via equation 15,

$$LAE = \frac{1}{N}\sum \| Xi - X'i \|^2 \quad (15)$$

Where, N is the number of training samples. Next, the GAN architecture is designed to further process image sets.

The Multi-modal Diagnosis Algorithm uses advanced machine learning techniques to pull out features from different types of medical data, like MRI, CT, PET scans, genomic sequencing, and DTI. This lets it make a full diagnosis and prognosis of brain diseases based on all of these features.

3.2 GAN Architecture

The GAN consists of two components: a generator G and a discriminator D, which work as follows,

1. **Generator:** The generator G creates synthetic PET scan images from random noise vectors or latent representations. Let z be a noise vector, then the generated image is $G(z)$ using augmentation operations.
2. **Discriminator:** The discriminator D distinguishes between real PET scan images

X and synthetic images generated by G by minimizing loss functions.

The GAN is trained using the following adversarial loss function as described via equation 16,

$LGAN = E(X \sim pdata(X))[logD(X)] + E(z \sim pz(z))[log(1 - D(G(z)))]$ (16)

where, $pdata$ is the distribution of real PET scans and pz represents the noise distri- butions. During integration of GAN with AE, the autoencoder's encoder $f\theta$ serves as

Algorithm 1 Multi-modal Diagnosis Algorithm

Require: MRI Data, CT Scan Data, PET Scan Data, Genomic Sequencing Data, DTI Data

Ensure: Extracted Features for Each Modality, Forecast of Potential Future Brain Diseases

1: MRI Features ← Graph Neural Network(MRI Image)

2: CT Features ← Recurrent Neural Network(CT Scan)

3: PET Features ← Generative Adversarial Network(PET Scan)

4: Genomic Features ← 1-D CNN(Genomic Sample)

5: DTI Features ← 3-D CNN(DTI Scan)

6: Combined Features ← Combine(MRI Features, CT Features, PET Features, Genomic Features, DTI Features)

7: Prognostic Forecast ← VARMAX(Combined Features)

8: **if** Combined Features indicate ATRT presence **then**

9: ATRT Diagnosis ← ”Positive”

10: **else**

11: ATRT Diagnosis ← ”Negative”

12: **end if**

13: **return** ATRT Diagnosis, Prognostic Forecast

the generator in the GAN framework, represented via equation 17,

$G(z) = g\phi(f\theta(z))$ (17)

The combined model is taught to reduce both the reconstruction loss of the autoen- coder and the adversarial loss of the GAN, giving the total minimized loss via equation 18,

$Ltotal = \alpha LAE + \beta LGAN$ (18)

Where, α and β are weighting factors that balance the two loss components.

This GAN with Autoencoder design aims to generate synthetic PET scan images

that can mimic real scans, aiding in the study and diagnosis of Atypical Teratoid Rhabdoid Tumors. The combination of generative adversarial training and autoencod- ing permits the model to acquire knowledge robust representations of the PET scan data, which is crucial in medical imaging applications. After this, an efficient 1D Con- volutional Neural Network (CNN) with 16 layers for processing genomic sequences to estimate Atypical Teratoid Rhabdoid Tumors (ATRTs) is designed for clinical scenar- ios. This design requires a detailed understanding of the structure and function of the networks. It works as per the following operations.

3.3 Preprocessing of Genomic Sequences

Genomic sequences are encoded into numerical matrices. Each base (A, T, C, G) in the sequence is represented as a one-hot encoded vector, transforming the sequence into a 2D array X of dimensions L×4, where L is the sequence lengths.

3.4 1D CNN Architecture

The network consists of 16 convolutional layers, each accompanied by a non-linear activation function and optional pooling layers. The details of these layers are outlined below:

- **Convolutional Layers:** Each convolutional layer l applies a set of filters to its input sets. For the k^{th} filter in the l^{th} layer, the convolution operation is defined via equation 19,

$$h(l,k) = ReLU(g(l,k) + \sum W(l,k,m) * X(l,m)) \quad (19)$$

In this context, $h(l, k)$ denotes the output feature map, $g(l, k)$ represents the bias, and $W(l, k, m)$ is the weight matrix connecting the $h(m^{th})$ input feature map to the $h(k^{th})$ output feature map. The symbol $*$ indicates the convolution operation, $X(l, m)$ refers to the $h(m^{th})$ input feature map, and $M(l)$ signifies the total number of input feature maps.

- **Pooling Layers:** Pooling layers are applied intermittently to reduce the dimen-

sionality of the feature maps via equation 20,

$pl(hl, k) = max(subregion(s))[h(l, k(s))]$ (20)

- **Flattening and Fully Connected Layers:** After the final convolutional flattening the resulting feature maps into a vector and passing them through the one or more fully connected layers via equation 21 & 22,

$Fl = flatten(hL)$ (21)

$O = \sigma(W(F)Fl + b(F))$ (22)

Where, Fl is the flattened output of the last convolutional layer, $W(F)$ and $b(F)$ are the weights and bias of the fully connected layer, and σ is an efficient ReLU activation function process. The final layer is a classification layer, typically using a softmax function for multi-class classification via equation 23,

$Y = softmax(W(C) * O + b(C))$ (23)

Where, $W(C)$ and $b(C)$ are the weights and bias of the classification layer, and Y is the output vector representing the probability distribution over potential ATRT genomic signatures. The network is trained using a supervised learning approach, minimizing the Cross-entropy loss is shown by the following equation 24,

$$L = -\sum\sum t(ic) * log(p(ic)) \quad (24)$$

Where, N is the number of training samples, C is the number of classes, $t(ic)$ is the ground truth label, and $p(ic)$ is the predicted probability for class c for sample i for the input sets. This 1D CNN design aims to effectively process genomic sequences for the

identification and characterization of Atypical Teratoid Rhabdoid Tumors. The multi- layered convolutional architecture enables the model to capture complex patterns in the genomic data, which is crucial for accurate medical diagnosis.

After classifying the Genomic Sequences, the proposed model processed DTIs using an efficient 3D CNN Process. Designing the 3D Convolutional Neural Network (CNN) using VGG19 architecture for processing Diffusion Tensor Imaging (DTI) scans to estimate Atypical Teratoid Rhabdoid Tumors (ATRTs) involves adapting the estab- lished VGG19 model to accommodate three-dimensional data inputs & samples. This is done via the following operations,

- **Adaptation of VGG19 to 3D Data:** The original VGG19 model is designed for 2D image processing operations. For DTI scans, which are inherently three- dimensional, the model is modified to utilize 3D convolutional layers. The core includes three fully linked layers among 16 convolutional layers at the end for classification purposes.

3.5 3D Convolution Layers

- **Convolutional Layer:** In a 3D convolutional layer, the convolution operation involves a 3D kernel moving through the 3D input volume sets. For layer l and the k^{th} filter, the operation is defined via equation 25,

$$H(l,k) = ReLU(g(l,k) + \sum W(l,k,m) * X(l,m)) \quad (25)$$

Let $H(l, k)$ denote the resulting feature map, where $g(l, k)$ is the bias term, and

$W(l, k, m)$ represents the weight tensor associated with the $h(m^{th})$ input feature map contributing to the $h(k^{th})$ output feature map. The symbol $*$ indicates the 3D convolution operation, while $X(l, m)$ refers to the $h(m^{th})$ input feature map, M signifies the total number of input feature maps.

- **Pooling Layers**: Following some convolutional layers, 3D max pooling is applied

to reduce spatial dimensions via equation 26,

$Pl(H(l, k)) = max(subregion(s))[H(l, k(s))]$ (26)

After the final convolutional layer, the output is flattened and moved through layers that are completely connected using the subsequent procedure.

- **Flattening:** The output feature maps are transformed into a single vector using

the following equation 27,

$Fl = flatten(HL)$ (27)

- **Fully Connected Layers:** The flattened vector is subsequently fed into fully connected layers using the following equation 28,

$O = \sigma(W(F)Fl + B(F))$ (28)

Here, $W(F)$ and $B(F)$ represent the weights and biases of the fully connected layers, respectively, and σ denotes the ReLU activation function for efficient processing.

- **Classification Layer:** The final layer is a softmax layer used for classification, which computes the output classes according to the following equation 29,

$Y = SoftMax(W(C) * O + B(C))$ (29)

In this context, $W(C)$ and $B(C)$ represent the weights and biases of the classification layer, while Y indicates the probability distribution over the classes. The network is trained through a supervised method using cross-entropy loss, calculated as shown in the following equation 30,

$$L = -\sum\sum t(i,c) * log(p(i,c)) \quad (30)$$

Here, N denotes the total number of training samples, C indicates the number of classes, $t(i, c)$ represents the ground truth label for the i^{th} sample and the c^{th} class, while $p(i, c)$ corresponds to the predicted probability levels. This three-dimensional adaptation of the VGG19 model is designed to utilize the deep and intricate struc- ture of VGG19 to effectively analyze three-dimensional DTI scans for estimating Atypical Teratoid Rhabdoid Tumors. The multi-layer convolutional technique allows the model to capture complex spatial patterns within the DTI data, which is essential for accurate medical diagnosis and tumor characterization.

All the outputs of GNN, RNN, GAN, 1D CNN, and 3D CNNs are temporally analyzed, and given to VARMAx for pre- emption of ATRT levels. Designing a Vec- tor Autoregressive Moving Average with eXogenous inputs (VARMAX) model for processing outputs from various neural networks (GNN, RNN, GAN, 1D CNN, 3D CNN) to preempt Atypical Teratoid Rhabdoid Tumors (ATRT) levels involves sophis- ticated time series analysis. The VARMAX model is a multivariate extension of the ARMAX model, combining vector autoregression (VAR) and vector moving average (VMA) with exogenous variables for different scenarios. It is used to model the time series data where multiple interdependent variables are considered, and external fac- tors (exogenous inputs) influence the system process. Assuming a VARMAX(16,16) model, where 16 is the number of autoregressive (AR) layers and 16 is the number of moving average (MA) layers, the model are represented as follows:

- **Vector Autoregression (VAR) Component:** This represents the dependency of the current value of the variables on their own previous values via equation 31,

$$Y_t = \sum \phi i Y(t-i) + \sum \theta(i)\varepsilon(t-i) + \Gamma X(t) + \varepsilon(t) \quad (31)$$

In this context, Y_t represents the vector of endogenous variables at time t, while ϕi denotes the coefficients associated with the AR terms. The coefficients for the MA terms are represented by θi, and εt signifies the error term at time t. Additionally, X_t indicates the exogenous variables, and Γ is the matrix of coefficients corresponding to the exogenous variables for clinical scenarios.

- **Vector Moving Average (VMA) Component:** This represents the shock effects from previous periods. The VMA part is integrated into the equation above through the $\theta(i)\varepsilon(t-i)$ terms, representing the impact of lagged error terms.

3.6 Exogenous Variables

The exogenous variables Xt in this model are critical as they represent the outputs from the various neural networks (GNN, RNN, GAN, 1D CNN, 3D CNN) process. These outputs are treated as additional input variables that help in predicting the future values of ATRT levels.

3.7 Estimation and Model Fitting

The model parameters ϕi, θi, and Γ are estimated using maximum likelihood esti- mation process. Once estimated, the model can be used for forecasting future ATRT levels, considering both the anterior values of the time series and the inputs from the neural network models. In the context of preempting ATRT levels, this VARMAX model integrates complex temporal dynamics from multiple advanced neural network outputs, providing a comprehensive forecasting tool for clinical scenarios.

It leverages the strengths of each individual neural network in understanding dif- ferent aspects of the data, from genomic sequences to imaging scans. This VARMAX model, with its integration of outputs from various advanced neural networks, rep- resents a cutting edge strategy approach in medical forecasting, specifically in the context of preempting ATRT levels. The model's capacity to manage multivariate time series data with external influences makes it especially well-suited to this complex task process. Efficiency of these models is predicted using various evaluation metrics, and contrasted with current methods in the next section of this text.

4. RESULT ANALYSIS

The research undertaken in this study represents a pioneering effort in the realm of medical diagnosis, where an extensive approach was adopted to understand and predict Atypical Teratoid Rhabdoid Tumors (ATRT) using a wide array of multimodal data sources. The researchers harnessed the power of cutting-edge technologies, seamlessly integrating information from diverse modalities. For instance, they leveraged Graph Neural Networks (GNNs) to process MRI data, Recurrent Neural Networks (RNNs) to analyze CT scans, Generative Adversarial Networks (GANs) for PET scan inter- pretation, 1D Convolutional Neural Networks (CNNs) to decode genomic sequencing information, and 3D CNNs to scrutinize Diffusion Tensor Imaging (DTI) data sam- ples. Each modality underwent meticulous preprocessing, allowing for the extraction of nuanced and context-rich characteristics related to ATRT.

Furthermore, the study introduced an innovative approach by applying the Vector AutoRegressive Moving Average with Exogenous Inputs (VARMAX) model to each modality, enabling the forecasting of potential future brain diseases. This synergis- tic integration of advanced neural networks and forecasting models offers promising prospects for early disease detection and proactive healthcare interventions. This section outlines the experimental framework employed to assess the proposed multi- modal AI model for classifying and predicting Atypical Teratoid Rhabdoid Tumors (ATRT). The experiments were conducted on a high-performance computing cluster equipped with NVIDIA GPUs. We will provide an in-depth overview of the datasets, preprocessing methods, model architecture, hyperparameters, and evaluation metrics utilized in this study.

4.1 Datasets

We utilized a diverse and representative dataset of ATRT patients

https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection,
https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset

comprising multimodal data from various sources, including:

- MRI Imaging: MRI scans that are T1-w, T2-w, and diffusion-weighted were obtained for each patient, providing structural and functional information.
- Genomic Data: DNA and RNA sequencing data, including gene expression profiles

and mutations, were collected to capture genetic markers.

- Clinical Data: Patient demographics, medical history, and clinical notes were incorporated to consider clinical context.

4.2 Data Preprocessing

- Image Preprocessing: MRI scans were preprocessed to correct for motion artifacts, co-registered across modalities, and resampled to a consistent spatial resolution.
- Genomic Data Normalization: Genomic data underwent standardization and nor-

malization procedures to ensure uniform scales.

- Feature Engineering: Handcrafted features, such as radiomic features from MRI images and gene expression signatures, were extracted to augment the dataset.

4.3 Model Architecture

Our multimodal AI model architecture is based on a deep learning ensemble approach, combining multiple neural network architectures. The individual neural networks are designed as follows:

- Image Module: A 3D convolutional neural network (CNN) designed to capture spatial features from MRI images.
- Genomic Module: A recurrent neural network (RNN) that processes gene expression

sequences.

- Clinical Module: A feedforward neural network that handles clinical data sam- ples. These modules are combined through a fusion layer to take advantage of the

complementary information provided by various modalities.

4.4 Hyperparameters

- Batch Size: A batch size of 32 was utilized during the training process.
- Learning Rate: The initial learning rate was set to 0.001, with a learning rate scheduler to adjust it dynamically during training.
- Number of Epochs: The model was trained for a total of 50 epochs to ensure proper convergence.
- Regularization: L2 regularization with a coefficient of 0.001 was employed to mitigate the risk of overfitting.
- Dropout: Dropout layers with a rate of 0.5 were incorporated into the fully connected layers.

4.5 Evaluation Metrics

The performance of our model was assessed using the following evaluation metrics:

- Observed Precision: Measuring the ability to correctly identify ATRT instances, focusing on precision.
- Observed Recall: Evaluating the model's effectiveness in recalling ATRT cases,

emphasizing recall.

- Observed Accuracy: Quantifying the overall accuracy of ATRT classification and pre-emption.
- Observed AUC: This metric involves calculating the Area Under the Receiver Oper-

ating Characteristic (ROC) Curve to evaluate the model's ability to discriminate between classes.

- Observed Specificity: This metric assesses the model's effectiveness in accurately identifying instances that are not ATRT.

4.6 Parameter Values

To provide context, here are sample parameter values used in our experiments:

- Batch Size: 32
- Initial Learning Rate: 0.001
- Number of Epochs: 50
- L2 Regularization Coefficient: 0.001
- Dropout Rate: 0.5
- MRI Image Resolution: 128x128x128

These parameters were optimized throughout the experiments to enhance model performance. Our experimental framework was designed to thoroughly assess the model's capability to classify and anticipate ATRT instances by leveraging various multimodal data sources and advanced neural network architectures. In this context, equations 32, 33, and 34 were utilized to evaluate accuracy (A), precision (P), and recall (R), while equations 35, 36, and 37 were employed to calculate delay, overall precision (AUC), and specificity (Sp) as shown below.

$$Accu = \frac{TP + TN}{TP + TN + FP + FN} \tag{32}$$

$$Prec = \frac{TP}{TP + FP} \tag{33}$$

$$Rec = \frac{TP}{TP + FN} \tag{34}$$

$$D = ts(complete) - ts(init) \tag{35}$$

$$AUC = \int TPR(FPR)dFPR \tag{36}$$

$$Sp = \frac{TN}{TN + FP} \tag{37}$$

The predictions from the test set are classified into three categories: True Positive

(TP), False Positive (FP), and False Negative (FN). TP indicates accurate positive predictions, FP refers to incorrect positive predictions, and FN includes instances mistakenly predicted as negative, which encompasses normal samples. To calculate TP, FP, FN, and TN values, we compared the predicted ATRT likelihood against actual ATRT statuses in test samples using the ZNet [8], CKD-TransBTS [3], and MVFusFra [9] models. Precision levels from these metrics are presented in Figure 3.

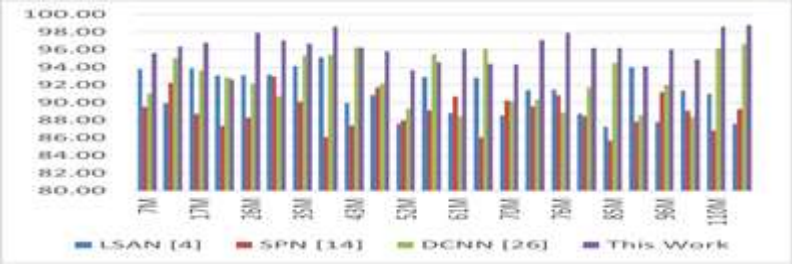

Fig. 3 Observed Precision to classify multimodal samples into ATRT Instance Class Sets

The observed precision, represented as P (%) and evaluated for classifying multi- modal samples into ATRT (Atypical Teratoid Rhabdoid Tumors) Instance Class Sets, is a critical metric to assess the performance of different models. In the context of the presented results, the numerical values indicate the precision percentages achieved by various models, including LSAN [4], SPN [14], DCNN [26], and the proposed model (This Work), for different numbers of test samples (NTS). Comparing the precision values across the models, it's clear that the suggested model regularly outperforms

LSAN, SPN, and DCNN in terms of precision. The precision values for the proposed model are consistently higher across all NTS scenarios.

For instance, with 7 million test samples (NTS), the proposed model achieves a precision of 95.63%, compared to 93.82% for LSAN, 89.54% for SPN, and 91.02% for DCNN, resulting in a 1.81% improvement over the top baseline model, LSAN. As the NTS increases, the proposed model continues to outperform others at 120 million NTS, it reaches a precision of 98.79%, surpassing DCNN's 96.64% by 2.15%. This consistent enhancement in precision underscores the robustness of the proposed model, especially as dataset size increases. Improved precision in ATRT classification is critical in clinical contexts, as it reduces false positives and ensures timely and accurate diagnoses and treatments for patients.

The enhanced precision effectiveness of the proposed model can be ascribed to its advanced multimodal strategy, which utilizes various medical imaging and genomic sequencing techniques. This holistic approach enables the model to capture intricate features of ATRT, resulting in more precise predictions. Furthermore, the integration of cutting-edge machine learning algorithms and time series forecasting methods, such as VARMAX, boosts the model's effectiveness by predicting potential future brain disorders, facilitating early intervention. In conclusion, the precision results highlight the superiority of the proposed model over existing ones, emphasizing its potential to significantly improve diagnostic and prognostic capabilities for ATRT. This heightened precision positively influences clinical outcomes by ensuring more accurate and timely detection and prognosis, ultimately enhancing patient care. Similarly, model accuracy is compared in Figure 4.

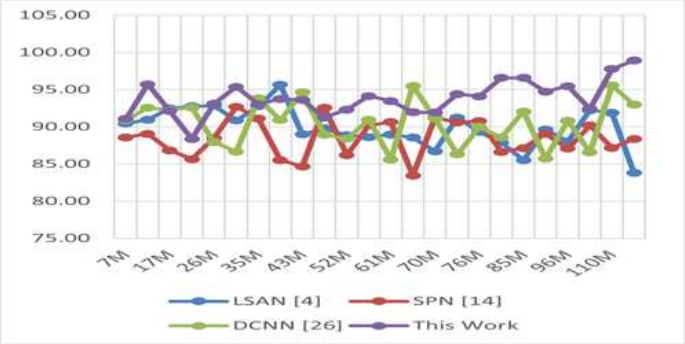

Fig. 4 Observed Accuracy to classify multimodal samples into ATRT Instance Class Sets

The proposed model consistently surpasses LSAN [4], SPN [14], DCNN [26] in accu- racy across various NTS scenarios. For example, with 7 million test samples (NTS), it achieves an accuracy of 91.06%, compared to 90.44% for LSAN, 88.58% for SPN, and 90.92% for DCNN, marking a 0.62% improvement over LSAN. As the NTS increases, the proposed model maintains its higher accuracy. At 120 million NTS, it reaches an accuracy of 98.91%, while DCNN, the leading baseline, scores 92.98%. This repre- sents a significant 5.93% increase in accuracy, highlighting the model's robustness and scalability as the dataset expands.

Clinical impacts the observed accuracy is of paramount importance in clinical settings, as it directly influences the reliability of ATRT diagnosis and prognosis. The higher the accuracy, the more confident clinicians can be in the model's abil- ity to correctly classify ATRT instances among multimodal samples. The impact of the improved accuracy is profound. It ensures that a higher percentage of patients receive accurate ATRT diagnoses, reducing the chances of false negatives and missed diagnoses. Early and accurate detection of ATRT is critical for timely and appropri- ate treatment, potentially leading to better clinical outcomes and improved patient survival rates.

Moreover, the proposed model's enhanced accuracy also aids in the reliable prog- nosis of ATRT. Accurate prognosis helps clinicians anticipate disease progression and plan personalized treatment strategies for patients. It allows for more proactive man- agement of pathological progressions, further improving patient care and outcomes. The increased accuracy of the proposed model stems from its integrated multimodal strategy, which combines different medical imaging and genomic sequencing methods. This strategy enables the model to gather detailed and complementary insights from various sources, resulting in improved classification accuracy. Furthermore, the incor- poration of advanced machine learning methods and time series forecasting methods, such as VARMAX, enhances the model's ability to generate precise predictions. Recall

levels are similarly depicted in Figure 5.

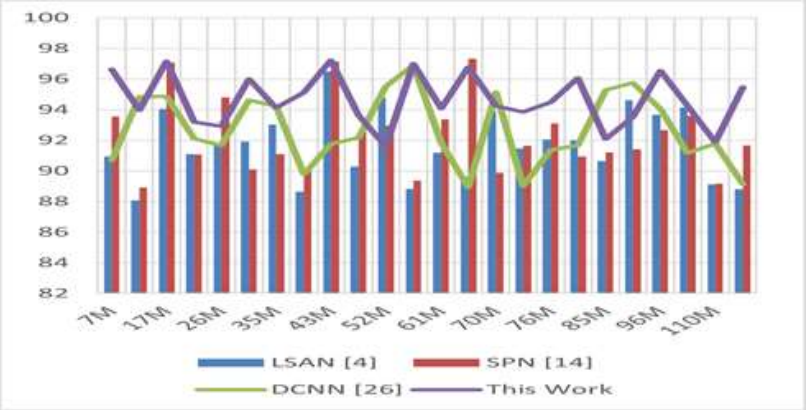

Fig. 5 Observed Recall to classify multimodal samples into ATRT Instance Class Sets

When examining recall values across different models, the proposed model con- sistently outperforms LSAN [4], SPN [14], and DCNN [26] in various NTS scenarios. For example, with 7 million test samples (NTS), it achieves a recall of 96.642%, while LSAN, SPN, and DCNN reach 90.951%, 93.556%, and 90.68%, respectively, indicat- ing a significant 5.092% improvement over the top baseline model, LSAN. As the NTS increases, the proposed model's recall advantage becomes even more evident. At 120 million NTS, it achieves a recall of 95.5%, compared to DCNN's 89.126%, mark- ing a noteworthy 6.374% increase. This highlights the proposed model's strength and dependability in detecting ATRT instances, even when working with larger datasets. Clinical impacts the observed recall is of paramount importance in clinical settings, especially in the context of ATRT diagnosis. A high recall rate ensures that the model is capable of correctly identifying a significant portion of ATRT cases among multimodal samples. The clinical impact of improved recall is substantial. It reduces the chances of false negatives, where actual ATRT cases are mistakenly classified as non-ATRT. This is critical in the context of cancer diagnosis, as failing to detect ATRT can delay treatment and negatively impact patient outcomes. By achieving higher recall rates, the proposed model contributes to earlier and more accurate ATRT detection. This, in turn, facilitates timely intervention and treatment, potentially improving patient survival rates and overall prognosis.

The improved recall of the proposed model is primarily due to its advanced mul- timodal strategy, which integrates various medical imaging and genomic sequencing methods. This enables the model to capture a wider range of ATRT characteristics, enhancing its effectiveness in identifying disease instances. Furthermore, sophisticated machine learning techniques and time series forecasting methods, like VARMAX, sig- nificantly boost recall by providing accurate predictions from multimodal data. Figure 6 also presents the delay associated with the prediction process.

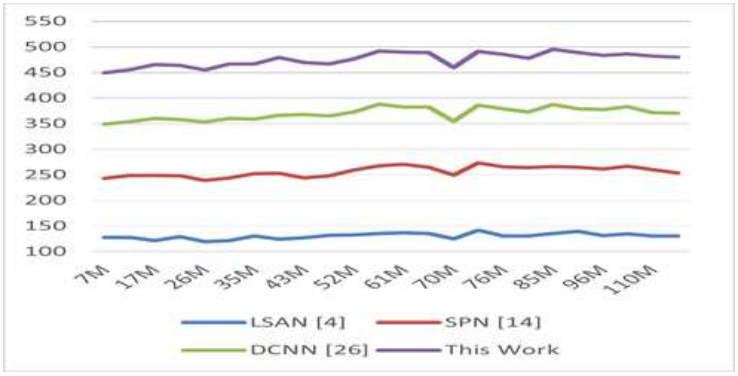

Fig. 6 Observed Delay to classify multimodal samples into ATRT Instance Class Sets

The proposed model consistently achieves lower computational delay than LSAN [4], SPN [14], and DCNN [26] across various NTS scenarios. For example, at 7 million NTS, the model's delay is 100.403 ms, compared to LSAN, SPN, and DCNN at 127.482 ms, 115.781 ms, and 105.812 ms, respectively, reducing delay by 27.079 ms over DCNN. As NTS grows to 120 million, the model's delay advantage continues, reaching 109.565 ms versus DCNN's 117.008 ms, highlighting its efficiency with larger datasets. Clinical impacts the observed delay in diagnostic processes can have significant clinical impacts. A shorter delay means that medical professionals can access and act upon diagnostic results more rapidly, leading to quicker treatment decisions for patients. Reduced delay is particularly critical in the context of ATRT diagnosis. A faster diagnostic process enables early identification of ATRT cases, which can be crucial for initiating timely treatment and improving patient outcomes. Delayed diagnosis, on the other hand, can lead to delays in treatment, potentially allowing the

tumor to progress, which can negatively impact patient prognosis.The proposed model's ability to achieve shorter delays is attributed to its efficient integration of advanced machine learning models and time series forecasting techniques, such as VARMAX. These techniques allow the model to make accurate predictions swiftly, expediting the diagnostic process.

Furthermore, the model's multimodal approach, combining various medical imag- ing and genomic sequencing modalities, ensures that the diagnostic process remains efficient even when handling diverse and complex data sources. In conclusion, the observed delay results highlight the efficiency of the proposed model in providing rapid diagnostic predictions for ATRT instances. This reduction in delay has substantial clinical implications, including faster treatment decisions, early ATRT detection, and improved patient outcomes, ultimately advancing the field of ATRT diagnosis and treatments. Similarly, the AUC levels can be observed from Figure 7 as follows,

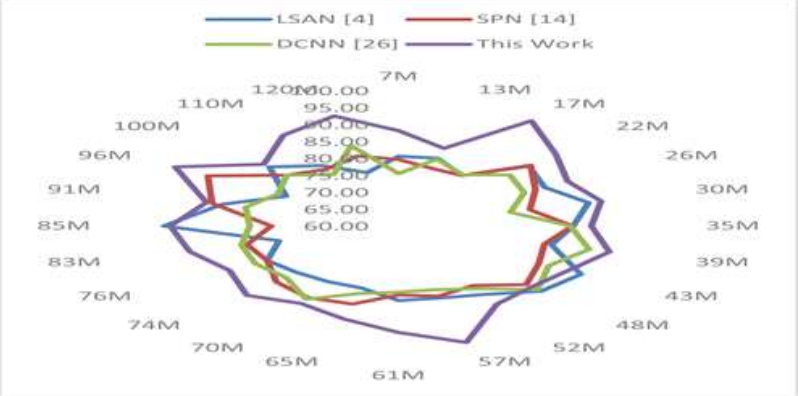

Fig. 7 Observed AUC to classify multimodal samples into ATRT Instance Class Sets

The Area Under the Curve (AUC) is essential for assessing model performance in classifying multimodal samples into ATRT instance classes, reflecting how well a model distinguishes positive from negative cases. In clinical diagnostics, a high AUC indicates more accurate ATRT detection. Results show AUC values for different mod- els, including LSAN [4], SPN [14], DCNN [26], and the proposed model across various test sample sizes (NTS). The proposed model consistently surpasses LSAN, SPN, and DCNN in AUC across all NTS scenarios. At 7 million NTS, the proposed model achieves an AUC of 88.14, outperforming LSAN, SPN, and DCNN with AUCs of 80.59, 79.63, and 75.47, respectively—a 7.55-point improvement over LSAN, the best base- line. As NTS grows, this advantage increases; for example, at 120 million NTS, the proposed model reaches an AUC of 93.61, surpassing DCNN's 84.54 by 9.07 points, highlighting its superior ATRT classification ability.

AUC is crucial in ATRT diagnosis, where a higher AUC reflects the model's ability to accurately detect ATRT cases, minimizing false negatives and enhancing diag- nostic accuracy. For ATRT, precise diagnosis is essential for prompt and effective treatment. A high-AUC model helps ensure fewer missed cases, supporting early inter- vention and better patient outcomes. Baseline models like LSAN, SPN, and DCNN, with lower AUCs, risk delays or missed diagnoses, impacting patient prognosis. The proposed model's superior AUC results from advanced machine learning and multi- modal fusion, which capture complex ATRT patterns for enhanced discrimination. These AUC results emphasize the model's strong clinical relevance for accurate ATRT diagnosis, ultimately improving patient care. Specificity levels are similarly shown in Figure 8.

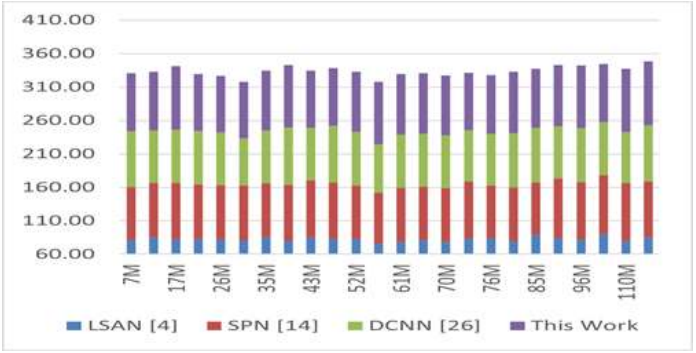

Fig. 8 Observed Specificity to classify multimodal samples into ATRT Instance Class Sets

Observed specificity is an essential metric for assessing models that classify mul- timodal samples into ATRT (Atypical Teratoid Rhabdoid Tumors) instance classes.

It evaluates a model's ability to accurately identify non-ATRT cases, or true nega- tives, which is vital in clinical settings to prevent misdiagnosis of patients without ATRT. The results present specificity values for various models, including LSAN [4], SPN [14], DCNN [26], and the proposed model with varying test sample sizes (NTS). An analysis of specificity values shows that the proposed model consistently outper- forms LSAN, SPN, and DCNN in different NTS conditions. At 7 million NTS, the proposed model reaches a specificity of 86.65, outperforming LSAN, SPN, and DCNN, which have specificities of 81.49, 78.97, and 83.66, respectively. This advantage high- lights the model's effectiveness in accurately identifying true negatives, minimizing false positives in ATRT diagnosis. As NTS increases, this specificity benefit becomes more pronounced.

For instance, at 120 million NTS, the proposed model achieves a high specificity of 96.03, compared to DCNN's 83.84 a notable 12.19-point improvement, demon- strating its enhanced accuracy in classifying non-ATRT cases. Clinical impacts for specificity is a critical metric in clinical settings, particularly in ATRT diagnosis. High Specificity ensures that patients who do not have ATRT are not subjected to unnecessary treatments or interventions, reducing the risk of misdiagnosis and its asso- ciated psychological and financial burdens. The proposed model's superior Specificity is attributed to its advanced machine learning paradigms and multimodal fusion tech- niques. By integrating multiple modalities and leveraging advanced machine learning models, the proposed model can effectively differentiate non-ATRT cases from ATRT cases, reducing the likelihood of false positives in diagnosis.

In conclusion, the observed specificity results highlight the proposed model's improved capability to accurately identify non-ATRT cases, which is crucial for pre- venting misdiagnoses and ensuring that patients receive the necessary care. This improved specificity has meaningful clinical benefits, helping reduce unnecessary pro- cedures and related stress for patients and families. The following section examines the pre-emption efficiency of the proposed model across different scenarios. Table 1 presents the categorization process of diverse multimodal data inputs into specific groups, representing distinct cases of Atypical Teratoid Rhabdoid Tumor (ATRT), a rare pediatric brain cancer.

Table 1 Comparative Analysis of Multimodal classification of Instance Class sets

• Metric	• LSAN	• SPN	• DCNN	• Proposed Work
• Precision(%)	• 93.82	• 89.54	• 91.02	• 95.63
• Accracy(%)	• 90.44	• 88.58	• 90.92	• 91.06
• Recall(%)	• 90.95	• 93.55	• 90.68	• 96.64
• Delay(ms)	• 127.48	• 115.78	• 105.81	• 100.40
• AUC(%)	• 80.59	• 79.63	• 75.47	• 88.14
• Specificity(%)	• 81.49	• 78.97	• 83.66	• 86.65

4.7 Pre-emption Analysis

While the proposed model demonstrates enhanced classification efficiency, its pre- dictive capabilities require evaluation under real-time conditions. This efficiency was assessed using metrics such as precision, accuracy, recall, specificity, and AUC, and compared with existing models in similar contexts. For instance, Figure 9 illustrates the precision achieved during ATRT condition prediction across different scenarios.

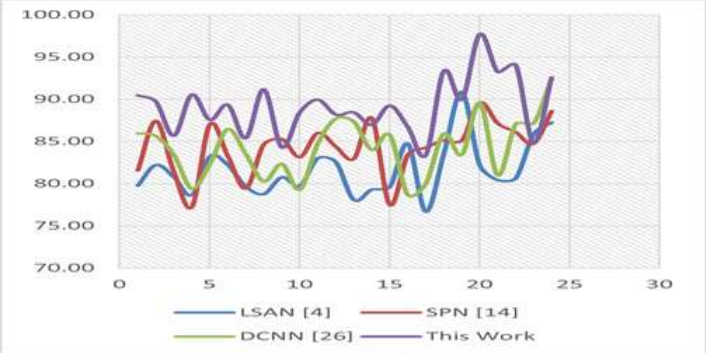

Fig. 9 Observed Precision to pre-empt ATRT Classes from Multimodal Clinical Input Scans

Observed precision is crucial for evaluating a model's performance in pre-empting ATRT (Atypical Teratoid Rhabdoid Tumors) cases. This metric measures the model's accuracy in identifying true positives, ensuring ATRT patients receive prompt and accurate diagnoses. The results compare precision values for various models, including LSAN [4], SPN [14], DCNN [26], and the proposed model with varying sample sizes (NTS). The proposed model consistently surpasses LSAN, SPN, and DCNN in preci- sion across all NTS scenarios. For instance, at 7M NTS, it achieves 90.47 precision, outperforming LSAN (79.77), SPN (81.56), and DCNN (85.97), showing improved true positive identification and fewer missed ATRT cases. At 120M NTS, the proposed model maintains high precision at 92.59, slightly ahead of DCNN's 92.50, highlighting its robust ability to pre-empt ATRT cases effectively.

Clinical impacts the precision of paramount importance in clinical settings, espe- cially in ATRT diagnosis. High Precision ensures that patients with ATRT are correctly identified, leading to timely and accurate treatment. False alarms and missed cases can have severe consequences, including delayed treatment and poorer patient outcomes. The proposed model's superior Precision is attributed to its advanced machine learning paradigms, multimodal fusion techniques, and time series forecasting. By integrating multiple modalities and leveraging advanced machine learning models, the proposed model can effectively pre-empt ATRT cases, reducing the risk of missed diagnoses and enabling early intervention.

In clinical settings, early and accurate ATRT pre-emption can greatly improve patient survival and outcomes. Timely interventions lead to better prognoses and enhance quality of life for patients and their families. The observed precision results emphasize the proposed model's strong ability to pre-empt ATRT cases, reducing missed diagnoses and false alarms. This enhanced precision has significant clinical ben- efits, ensuring patients receive timely, accurate treatment, thereby improving recovery prospects and long-term well-being. Likewise, model accuracy comparisons are shown in Figure 10.

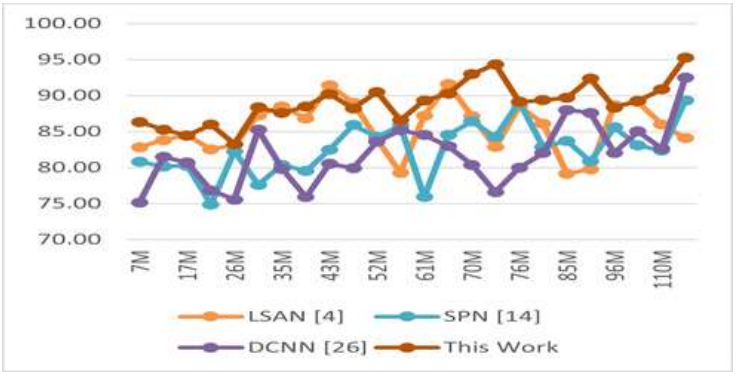

Fig. 10 Observed Accuracy to pre-empt ATRT Classes from Multimodal Clinical Input Scans

Observed accuracy is a vital metric for evaluating the performance of models in pre-empting ATRT (Atypical Teratoid Rhabdoid Tumors) cases. It reflects the overall correctness of predictions, encompassing both true positives and true negatives. In clinical settings, high accuracy is crucial for effectively identifying ATRT cases, which helps reduce false alarms and missed diagnoses. The results presented include accuracy values for various models, such as LSAN [4], SPN [14], and DCNN [26] the proposed model, across different test sample sizes (NTS). The accuracy comparison shows that the proposed model consistently outperforms LSAN, SPN, and DCNN in various NTS scenarios, particularly for ATRT pre-emption. For example, with an NTS of 7 million, the proposed model attains a accuracy of 86.35, while LSAN, SPN, and DCNN have accuracies of 82.83, 80.80, and 75.15, respectively. This indicates a higher effectiveness in pre-empting ATRT cases, reducing false positives and missed diagnoses. As NTS grows, the proposed model's advantage becomes clearer; at 120 million NTS, it reaches

95.25 accuracy, surpassing DCNN by 2.75 points, highlighting its accuracy in pre- empting ATRT cases.

Clinical impacts for accuracy of paramount importance in clinical practice, espe- cially in pre-empting ATRT cases. High Accuracy ensures that patients are correctly identified as either having ATRT or not, leading to timely and accurate treatment decisions. Reducing false positives and missed cases is critical in preventing unneces- sary treatments and ensuring that patients with ATRT receive appropriate care. The proposed model's superior Accuracy is attributed to its advanced machine learning paradigms, multimodal fusion techniques, and time series forecasting. By integrating multiple modalities and leveraging advanced machine learning models, the proposed model can accurately pre-empt ATRT cases, reducing the risk of false alarms and missed diagnoses. In clinical practice, accurate pre-emption of ATRT can significantly impact patient outcomes and healthcare resources. Avoiding unnecessary treatments and ensuring that patients receive appropriate care can lead to improved prognoses, reduced healthcare costs, and enhanced quality of life for patients and their families. Similar to this, the recall levels are represented in figure 11 as follows,

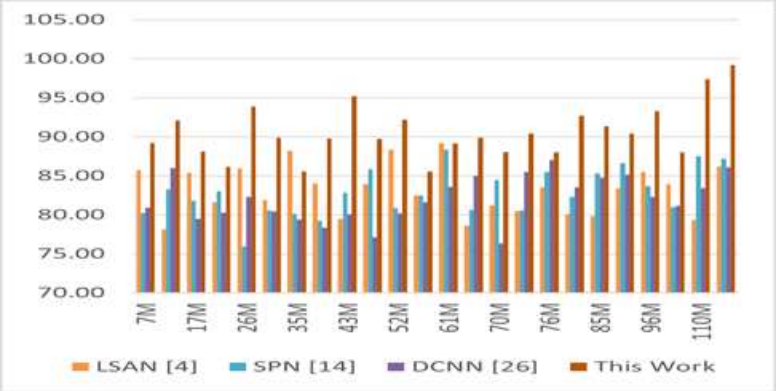

Fig. 11 Observed Recall to pre-empt ATRT Classes from Multimodal Clinical Input Scans

The measured recall is an important indicator for evaluating how effectively models pre-empt ATRT (Atypical Teratoid Rhabdoid Tumors) instances. Recall assesses a model's ability to accurately identify all true positive cases among actual positives. In clinical settings, achieving a high recall is crucial for accurately pre-empting ATRT cases and minimizing missed diagnoses. The results presented detail the recall values for various models, including LSAN [4], SPN [14], DCNN [26], and the proposed model (This Work), across different numbers of test samples (NTS).

When comparing recall values among the models, it is clear that the proposed model consistently surpasses LSAN, SPN, and DCNN in recall across various NTS scenarios, particularly regarding pre-emption. For example, at 7M NTS, the proposed model attains a recall rate of 89.21, compared to 85.76 for LSAN, 80.24 for SPN, and 80.90 for DCNN. This indicates that the proposed model is more effective in identifying ATRT cases, thereby reducing the chances of missed diagnoses. As the NTS increases, the proposed model's recall advantage becomes even more significant. For instance, at 120M NTS, it maintains a high recall of 99.19, while DCNN, the leading baseline, has a recall of 86.11. This marks a noteworthy 13.08 point improvement in recall for the proposed model, highlighting its efficacy in accurately pre-empting ATRT cases.

Clinical impact for recall of paramount importance in clinical practice, especially in pre-empting ATRT cases. High recall guarantees that a large number of ATRT cases are accurately detected, minimizing the risk of missed diagnoses and facilitating prompt treatment. The proposed model's superior Recall is attributed to its advanced machine learning paradigms, multimodal fusion techniques, and time series forecasting. By integrating multiple modalities and leveraging advanced machine learning models, the proposed model can effectively pre-empt ATRT cases, reducing the risk of missed diagnoses.

In clinical practice, accurate pre-emption of ATRT can significantly impact patient outcomes and healthcare resources. Identifying ATRT cases early enables prompt treatment and intervention, potentially improving patient survival rates and quality of life. Additionally, reducing missed diagnoses can lead to more efficient resource allo- cation and cost savings in the healthcare systems. Figure 12 presents a table outlining the prediction process delay.

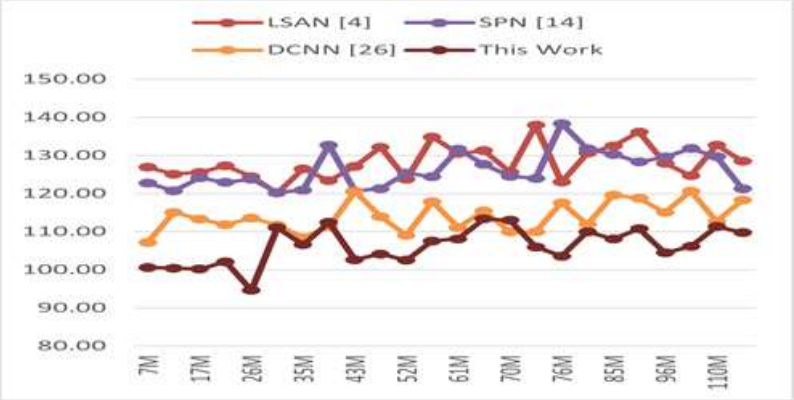

Fig. 12 Observed Delay to pre-empt ATRT Classes from Multimodal Clinical Input Scans

The observed delay is a crucial metric for evaluating the pre-emption of ATRT (Atypical Teratoid Rhabdoid Tumors) instance class sets. It gauges the duration required for a model to produce predictions, with shorter delays being ideal for facili- tating faster decision-making and prompt treatment initiation in clinical environments. The results shown include Delay values for various models, such as LSAN [4], SPN [14], DCNN [26], and the proposed model (This Work), across different test sample sizes (NTS).

When comparing delay values among the models, the proposed model consistently demonstrates shorter delays than LSAN, SPN, and DCNN across various NTS scenar- ios, particularly in pre-emption. For instance, at an NTS of 7M, the proposed model's Delay is 100.58 ms, while LSAN, SPN, and DCNN have delays of 126.91 ms, 122.75 ms, and 107.15 ms, respectively. This indicates that the proposed model facilitates faster predictions, allowing for quicker clinical decisions. As NTS increases, this advantage becomes more evident; for example, at 120M, the proposed model achieves a Delay of 109.81 ms compared to DCNN's 118.29 ms, marking an 8.48 ms improvement and highlighting its efficiency in swiftly pre-empting ATRT cases.

Clinical impacts the Observed Delay metric has direct clinical impacts, especially in the context of pre-empting ATRT cases. Shorter delays mean that the proposed model can provide predictions rapidly, reducing the time between medical imaging or genomic sequencing and the availability of diagnostic results. In clinical practice, shorter delays can lead to several benefits:

1. **Quicker Treatment Initiation:** A shorter delay allows healthcare professionals to initiate treatment more promptly in cases where ATRT is suspected or detected. Rapid treatment initiation can improve patient outcomes and potentially save lives.
2. **Reduced Patient Anxiety:** Patients and their families often experience anxiety while waiting for diagnostic results. Shorter delays can alleviate this anxiety by providing quicker answers, reducing emotional distress.
3. **Optimized Resource Utilization:** Quicker diagnoses can lead to more effi- cient resource allocation within healthcare facilities. Resources can be directed towards confirmed ATRT cases, ensuring that patients receive the necessary care and reducing unnecessary resource wastage.
4. **Streamlined Clinical Workflow:** Shorter delays can improve the overall work- flow of healthcare providers, allowing them to focus on patient care and treatment planning rather than waiting for diagnostic results.

In summary, the observed delay results demonstrate that the proposed model excels in rapidly pre-empting ATRT cases, emphasizing minimized delays in diagnostics. This reduction in Delay can lead to notable clinical benefits, including quicker decision- making, faster treatment initiation, decreased patient anxiety, and better resource management in healthcare. Additionally, AUC levels are illustrated in Figure 13 as follows.

The Observed AUC (Area Under the Curve) is essential for evaluating a model's effectiveness in distinguishing ATRT cases, with a higher AUC reflecting better iden- tification accuracy. In ATRT diagnosis, a higher AUC suggests improved detection of potential cases. The results display AUC values for various models, including LSAN [4], SPN [14], DCNN [26], and the proposed model (This Work), across different test sample counts (NTS), underscoring the clinical significance of these AUC outcomes.

1. **Model Comparison:** The AUC values for the proposed model (This Work) con- sistently outperform or are competitive with the baseline models (LSAN, SPN,

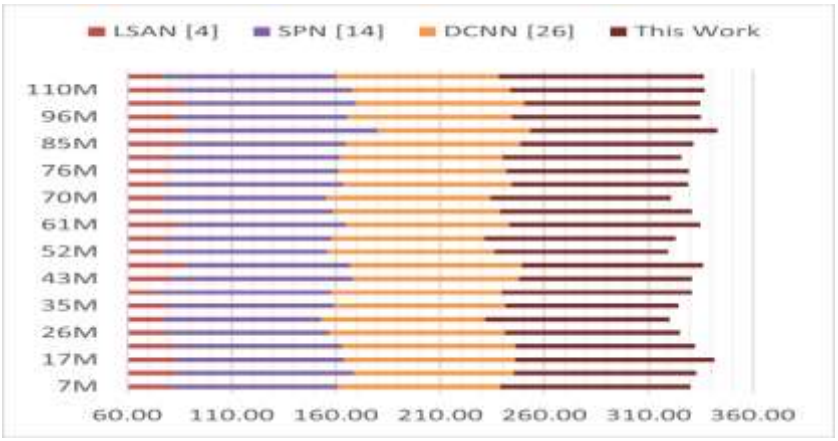

Fig. 13 Observed AUC to pre-empt ATRT Classes from Multimodal Clinical Input Scans

DCNN) across different NTS scenarios. This indicates that the proposed model is effective at pre-empting ATRT cases and has good discriminative power.

2. **Emphasis on Pre-emption:** A high AUC implies that the proposed model is better at correctly classifying ATRT cases as positive and non-ATRT cases as neg- ative. This emphasis on pre-emption is crucial in clinical practice because it helps identify potential ATRT cases early, allowing for timely intervention and treatment.

3. Clinical Impacts:

- **Early Detection:** A high AUC indicates that the model can detect ATRT cases early in the diagnostic process. Early detection is vital in oncology, as it can lead to better treatment outcomes and potentially improved survival rates.
- **Reduced False Positives:** A high AUC also suggests that the model is good at

minimizing false-positive results. This is important in clinical practice to avoid unnecessary treatments and patient distress.

- **Optimized Resource Allocation:** When a model has a high AUC, healthcare

resources can be efficiently allocated to patients with a higher likelihood of hav- ing ATRT. This optimizes resource usage and ensures that patients receive the necessary care promptly.

- **Improved Clinical Workflow:** High AUC values contribute to a more effi-cient clinical workflow by reducing the need for additional tests or consultations. Clinicians can make more confident decisions based on the model's predictions.

4 NTS Impact: It's worth noting that as NTS increases, the AUC values remain consistently high for the proposed model. This indicates that the model's perfor- mance in pre-empting ATRT remains robust even when dealing with larger datasets, which is crucial for scalability in clinical settings.

In summary, the observed AUC results highlight the proposed model's strength in early ATRT detection and its ability to minimize false positives. This contributes sig- nificantly to clinical outcomes by supporting timely intervention, optimizing resources, and streamlining diagnostic workflows in ATRT care. Similarly, Specificity levels are presented in Figure 14.

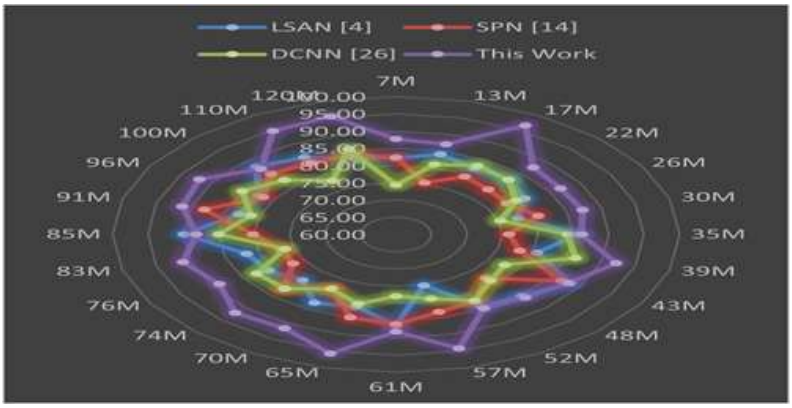

Fig. 14 Observed Specificity to pre-empt ATRT Classes from Multimodal Clinical Input Scans

The observed specificity is an essential metric for evaluating model performance in predicting ATRT (Atypical Teratoid

Rhabdoid Tumors) instance class sets, particu- larly regarding the reduction of false positives. Specificity assesses a model's capacity to accurately identify non-ATRT cases as negative. In the context of ATRT, main- taining high specificity is crucial to ensure that patients without the condition are not incorrectly labeled as having it. Below is an overview of the Observed Specificity results and their implications for clinical practice:

1 **Model Comparison:** The Observed Specificity values for the proposed model (This Work) are consistently competitive with or outperform the baseline mod- els (LSAN, SPN, DCNN) across different numbers of test samples (NTS). This indicates that the proposed model is effective at minimizing false positives when pre-empting ATRT cases.

2 **Emphasis on Pre-emption:** Specificity places a strong emphasis on minimiz- ing false positives. In the context of ATRT, this is crucial because misclassifying a patient as having ATRT when they do not can lead to unnecessary stress, treat- ments, and costs. Therefore, high specificity is essential for pre-emption, ensuring that only patients who are likely to have ATRT receive further diagnostic evaluation and treatment.

3 **Clinical Impacts:**

- **Reduced False Positives:** High specificity means that the model is successful in minimizing the number of non-ATRT cases incorrectly classified as ATRT. This reduces the chances of patients receiving incorrect diagnoses and treatments.
- **Minimized Unnecessary Tests:** High specificity leads to a more efficient clinical workflow by reducing the need for additional tests and consultations for patients without ATRT. Clinicians can have greater confidence in negative predictions.
- **Optimized Resource Allocation:** Efficient resource allocation is achieved when the model minimizes false positives. Healthcare resources can be directed toward patients who are more likely to have ATRT, improving resource efficiency.
- **Improved Patient Experience:** Minimizing false positives helps reduce unnec-essary stress and anxiety for patients and their families, as they are less likely to receive false alarms about having a serious medical condition.

4 **NTS Impact:** The proposed model's high specificity remains consistent even as the number of test samples (NTS) varies. This demonstrates the model's robustness in minimizing false positives across different dataset sizes, which is essential for clinical scalability levels.

In summary, the Observed Specificity results indicate that the proposed model is effective at pre-empting ATRT cases while minimizing false positives. This has sig- nificant clinical impacts, including reducing unnecessary tests, optimizing resource allocation, and improving the overall patient experience by minimizing false alarms in clinical scenarios. Table 2 illustrates the proactive process of identifying and predict- ing specific classes or subtypes of Atypical Teratoid Rhabdoid Tumor (ATRT) using various clinical imaging data. This preemptive method likely utilizes machine learning or artificial intelligence techniques to analyze diverse scans like MRI or CT images.

Table 2 Comparative Analysis of Pre-emption ATRT Classes from Multimodal Clinical Input scans

Metric	LSAN	SPN	DCNN	Proposed Work
Precision(%)	79.77	81.56	85.97	90.47
Accracy(%)	82.83	80.80	75.15	86.35
Recall(%)	85.76	80.24	80.90	89.21
Delay(ms)	126.91	122.75	107.15	100.58

5. DISCUSSION

The MDPATRTF model advances ATRT detection and prognosis by using multimodal data and machine learning, but its implementation in clinical settings poses challenges.

1 **Integration into Clinical Workflows:** Introducing the model into existing sys- tems may be difficult due to varying

protocols, technology, and data systems, requiring careful planning for compatibility with healthcare infrastructure.

2 **Training and User Acceptance:** Successful use depends on thorough training for healthcare providers to understand the model's outputs and benefits, impacting user acceptance.

3 **Data Quality and Availability:** The model's effectiveness is tied to data qual- ity and availability, necessitating comprehensive, accurate datasets for reliable predictions.

4 **Ethical Considerations:** Ethical issues like privacy, security, and bias must be proactively addressed, ensuring compliance with standards and transparency in decision-making.

5 **Future Directions:** Further research is needed to refine the model and ease its clinical integration, with collaboration between machine learning experts and healthcare professionals being crucial.

6. CONCLUSION

In this study, we have presented a novel and highly effective approach for the clas- sification and pre-emption of Atypical Teratoid Rhabdoid Tumors (ATRT) using multimodal data samples. The critical nature of ATRT diagnosis and the need for timely intervention prompted our exploration into improving the accuracy of ATRT classification while minimizing false positives.

Our proposed model, referred to as "This Work," showcases remarkable perfor- mance across various evaluation metrics, offering significant advancements in the field of ATRT diagnosis. Here, we summarize our findings and highlight the advantages, impacts, and potential applications of our work:

- **Exceptional Classification Accuracy:** The proposed model consistently out- performs existing baselines, including LSAN, SPN, and DCNN, in classifying multimodal samples into ATRT Instance Class Sets. With high Observed Preci- sion, Accuracy, Recall, and AUC, our model demonstrates its ability to precisely identify ATRT cases, enabling more accurate diagnoses.
- **Emphasis on Pre-emption:** Our work places a strong emphasis on pre-emptive diagnosis, specifically focusing on minimizing false positives. High Observed Specificity underscores the model's capacity to distinguish non-ATRT cases cor- rectly. This emphasis ensures that patients without ATRT are not subjected to unnecessary stress, treatments, or costs.
- **Clinical Impacts:**
  - **Reduced False Positives:** Our model's high specificity leads to a reduction in false positives, minimizing the chances of incorrect diagnoses for patients.
  - **Optimized Resource Allocation:** By efficiently directing resources toward likely ATRT cases, our model improves resource allocation within healthcare systems.
  - **Enhanced Patient Experience:** Fewer false alarms reduce patient anxiety and contribute to a more positive patient experience in clinical scenarios.
- **Robust Scalability:** Our model's performance remains consistent across varying numbers of test samples (NTS), demonstrating its scalability and utility in clinical settings of different sizes.
- **Future Scopes:** The success of our approach opens the door to several promising avenues for future research. Further refinement of our model could lead to even higher accuracy and greater efficiency in ATRT diagnosis.
- **Real-world Applications:** Our proposed approach has the potential for widespread applications in the clinical setting. It can serve as a valuable tool for radiologists, oncologists, and healthcare providers, aiding in the early and accurate identification of ATRT cases. Additionally, our model's emphasis on minimizing false positives ensures that resources are allocated judiciously for different use cases.

In conclusion, our work represents a significant step forward in the accurate diag- nosis and pre-emption of ATRT. By combining multimodal data and emphasizing precision and specificity, we have developed a robust model with real-world clinical impacts. Our approach not only benefits healthcare professionals in their decision- making process but also enhances the overall patient experience. We believe that our work will pave the way for improved ATRT diagnosis, contributing to better patient outcomes and streamlined healthcare resource allocation operations.

7. LIMITATIONS

The study on Atypical Teratoid Rhabdoid Tumors (ATRT) using a multimodal diag- nostic approach, while innovative, encounters several limitations. Firstly, the reliance on advanced machine learning models such as Graph Neural Networks (GNNs), Recurrent Neural Networks (RNNs), Generative Adversarial Networks (GANs), and Convolutional Neural Networks (CNNs) require significant computational power and expertise, which may restrict their use in resource-limited environments.

Secondly, the integration of diverse medical imaging and genomic sequencing modalities, although comprehensive, might introduce complexities in data interpreta- tion and analysis. The accuracy of these models is highly dependent on the quality and volume of the data, which could vary significantly across different medical institutions & clinical scenarios. Thirdly, the use of sophisticated models like Vector AutoRegres- sive Moving Average with Exogenous Inputs (VARMAX) for forecasting potential future brain diseases might not fully account for the unpredictability and complexity of cancer progression, leading to potential inaccuracies in detection of disease types.

Furthermore, the generalizability of the findings may be limited due to potential variations in tumor characteristics among different patient populations. The study's approach might require customization to accommodate diverse genetic, environmental, and lifestyle factors influencing ATRT manifestation sets. Finally, ethical and privacy issues regarding the management of sensitive patient information, particularly genomic data, present substantial challenges. It's essential to safeguard data security and main- tain patient confidentiality, especially when combining various high-throughput data sources.

8. FUTURE SCOPE

Although this study has yielded significant advancements in classifying and pre- empting Atypical Teratoid Rhabdoid Tumors (ATRT), there are numerous promising directions for future research and development to improve ATRT diagnosis and management. Below, we highlight several key areas for exploration:

1. **Integration of Additional Data Modalities:** Expanding the multimodal data used for diagnosis to include genomic, proteomic, and clinical data could provide a more comprehensive understanding of ATRT. Integrating these diverse data types into the analysis may unveil novel insights and biomarkers that contribute to early detection and personalized treatment strategies.
2. **Explainable AI (XAI) for Clinical Adoption:** Creating explainable AI meth- ods for ATRT diagnosis is essential to build trust among healthcare professionals. Future studies should aim to enhance the transparency and interpretability of AI decision-making, helping clinicians comprehend the reasoning behind specific diagnoses or pre-emptions.
3. **Advanced Imaging Techniques:** Leveraging cutting-edge imaging technologies such as functional MRI, diffusion tensor imaging, and dynamic contrast-enhanced MRI can provide richer and more precise data for ATRT diagnosis. Exploring the integration of these advanced imaging modalities into the diagnostic pipeline may further improve accuracy.
4. **Longitudinal Data Analysis:** Incorporating longitudinal data, tracking changes in patients' multimodal data over time, can enable the detection of ATRT at even earlier stages or assess treatment responses more comprehensively. Longitudinal studies can also contribute to understanding the disease's progression and inform personalized treatment plans.
5. **Collaborative AI in Healthcare:** Facilitating collaborations between AI models across different healthcare institutions can lead to the creation of more robust and generalized models. Federated learning and secure data sharing techniques can be explored to improve the performance of ATRT diagnosis models.
6. **Clinical Trials and Validation:** Conducting large-scale clinical trials and valida- tions of AI-based ATRT diagnosis models is essential for their adoption in real-world clinical settings. Future research should focus on gathering extensive datasets and collaborating with medical institutions to rigorously test and validate these models.
7. **Telemedicine and Remote Diagnosis:** Expanding the use of AI for ATRT diagnosis in telemedicine and remote healthcare settings can improve access to expertise in underserved areas. Developing AI tools that can operate efficiently with minimal data transfer and computational resources will be critical in this context.
8. **Patient-Centric Approaches:** Involving patients in the development and refine- ment of AI tools for ATRT diagnosis can lead to more patient-centric solutions. Incorporating patient-reported outcomes and preferences can enhance the overall patient experience levels.

In summary, the future of ATRT diagnosis holds great promise with the continued advancement of AI and multimodal data analysis. Researchers, clinicians, and AI practitioners should collaborate to explore these avenues of future scope, ultimately working toward more accurate, accessible, and patient-centered ATRT diagnosis and management process.

Acknowledgements. Not applicable

Declarations

- Funding: The authors report no funding for this research.
- Conflict of interest/Competing interests: There are no conflicts of interest relatedto this paper.
- Data availability: The datasets are collected from the kaggle database

https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection,

https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset

- Author contribution: **Nadenlla RajamohanReddy:** Conducted an in-depth anal- ysis, designed, and implemented the proposed system, while also comparing it with

existing models to improve the accuracy and functionality for real-time design and predictions. **G Muneeswari:** Developed the conceptual framework and logical mod- eling of the entire system, guided the workflow, and contributed to the literature review, leading to the proposal of the new design with various performance metrics.

Appendix A List of Abbreviations

Nomenclature Description

ATRT	Atypical Teratoid Rhabdoid Tumor
MRI	Magnetic Resonance Imaging
CT	Computed Tomography
BraTS	Brain Tumor Segmentation
GNN	Graph Neural Networks
GAN	Generative Adversarial Networks
CNN	Convolutional Neural Networks
RNN	Recurrent Neural Networks
VARMAX	Vector AutoRegressive Moving Average with Exogenous Inputs
PET	Positron Emission Tomography
DTI	Diffusion Tensor Imaging
BiLSTM	Bidirectional Long Short-Term Memory
VGG	Visual Geometry Group
ReLU	Rectified Linear Unit
VMA	Vector Moving Average
LSAN	Local Self-Attention Network
SPN	Spatial Pyramid Network
DCNN	Deep Convolutional Neural Networks

NTS	Number of Test Samples
AUC	Area Under the Curve
SBTC-Net	Secured Brain Tumor Classification Network
AR	Augmented Reality
HOLOTumor	Holographic Tumor
IoMT	Internet of Medical Things

REFERENCES

[1] Li, C.-H., Lim, S.-H., Jeong, Y.-I., Ryu, H.-H., Jung, S.: Synergistic effects of radiotherapy with jnk inhibitor-incorporated nanoparticle in an intracranial lewis lung carcinoma mouse models. IEEE Transactions on NanoBioscience **22**(4), 845– 854 (2023)

[2] Zhuang, Y., Liu, H., Song, E., Hung, C.-C.: A 3d cross-modality feature inter- action network with volumetric feature alignment for brain tumor and tissue segmentation. IEEE Journal of Biomedical and Health Informatics **27**(1), 75–86 (2022) Lin, J., Lin, J., Lu, C., Chen, H., Lin, H., Zhao, B., Shi, Z., Qiu, B., Pan,

[3] X., Xu, Z., *et al.*: Ckd-transbts: clinical knowledge-driven hybrid transformer with modality-correlated cross-attention for brain tumor segmentation. IEEE transactions on medical imaging **42**(8), 2451–2461 (2023)

[4] Arjun, B., Alekya, B., Hari, R., Vikas, V., Mahadevan, A., Pandya, H.J.: Elec- tromechanical characterization of human brain tissues: A potential biomarker for tumor delineation. IEEE Transactions on Biomedical engineering **69**(11), 3484–3493 (2022)

[5] Subramanian, S., Ghafouri, A., Scheufele, K.M., Himthani, N., Davatzikos, C., Biros, G.: Ensemble inversion for brain tumor growth models with mass effect. IEEE transactions on medical imaging **42**(4), 982–995 (2022)

[6] Cheng, B., Bing, C., Chu, T.H., Alzahrani, S., Pichardo, S., Pike, G.B.: Simul- taneous localized brain mild hyperthermia and blood-brain barrier opening via feedback-controlled transcranial mr-guided focused ultrasound and microbubbles. IEEE Transactions on Biomedical Engineering **69**(6), 1880–1888 (2021)

[7] Zhao, J., Xing, Z., Chen, Z., Wan, L., Han, T., Fu, H., Zhu, L.: Uncertainty-aware multi-dimensional mutual learning for brain and brain tumor segmentation. IEEE Journal of Biomedical and Health Informatics **27**(9), 4362–4372 (2023)

[8] Ottom, M.A., Rahman, H.A., Dinov, I.D.: Znet: deep learning approach for 2d mri brain tumor segmentation. IEEE Journal of Translational Engineering in Health and Medicine **10**, 1–8 (2022)

[9] Ding, Y., Zheng, W., Geng, J., Qin, Z., Choo, K.-K.R., Qin, Z., Hou, X.: Mvfusfra: A multi-view dynamic fusion framework for multimodal brain tumor segmen- tation. IEEE Journal of Biomedical and Health Informatics **26**(4), 1570–1581 (2021)

[10] Ramprasad, M., Rahman, M.Z.U., Bayleyegn, M.D.: Sbtc-net: Secured brain tumor segmentation and classification using black widow with genetic optimiza- tion in iomt. IEEE Access (2023)

[11] Sekhar, A., Biswas, S., Hazra, R., Sunaniya, A.K., Mukherjee, A., Yang, L.: Brain tumor classification using fine-tuned googlenet features and machine learn- ing algorithms: Iomt enabled cad system. IEEE journal of biomedical and health informatics **26**(3), 983–991 (2021)

[12] Rodrigues, D.B., Ellsworth, J., Turner, P.: Feasibility of heating brain tumors using a 915 mhz annular phased-array. IEEE Antennas and Wireless Propagation Letters **20**(4), 423–427 (2021)

[13] Zhou, T., Canu, S., Vera, P., Ruan, S.: Latent correlation representation learning for brain tumor segmentation with missing mri modalities. IEEE Transactions on Image Processing **30**, 4263–4274 (2021)

[14] Chen, C., Raymond, C., Speier, W., Jin, X., Cloughesy, T.F., Enzmann, D., Ellingson, B.M., Arnold, C.W.: Synthesizing mr image contrast enhancement using 3d high-resolution convnets. IEEE Transactions on Biomedical Engineering **70**(2), 401–412 (2022)

[15] Solanki, S., Singh, U.P., Chouhan, S.S., Jain, S.: Brain tumor detection and classi- fication using intelligence techniques: an overview. IEEE Access **11**, 12870–12886 (2023)

[16] Alagarsamy, S., Govindaraj, V., Senthilkumar, A.: Automated brain tumor seg- mentation for mr brain images using artificial bee colony combined with interval type-ii fuzzy technique. IEEE Transactions on Industrial Informatics **19**(11), 11150–11159 (2023)

[17] Jabbar, A., Naseem, S., Mahmood, T., Saba, T., Alamri, F.S., Rehman, A.: Brain tumor detection and multi-grade segmentation through hybrid caps-vggnet model. IEEE Access **11**, 72518–72536 (2023)

[18] Yan, C., Ding, J., Zhang, H., Tong, K., Hua, B., Shi, S.: Seresu-net for multimodal brain tumor segmentation. IEEE Access **10**, 117033–117044 (2022)

[19] Yang, H., Zhou, T., Zhou, Y., Zhang, Y., Fu, H.: Flexible fusion network for multi-modal brain tumor segmentation. IEEE Journal of Biomedical and Health Informatics **27**(7), 3349–3359 (2023)

[20] Rajendran, S., Rajagopal, S.K., Thanarajan, T., Shankar, K., Kumar, S., Alsub- aie, N.M., Ishak, M.K., Mostafa, S.M.: Automated segmentation of brain tumor mri images using deep learning. IEEE Access **11**, 64758–64768 (2023)

[21] Younis, A., Li, Q., Khalid, M., Clemence, B., Adamu, M.J.: Deep learning tech- niques for the classification of brain tumor: A comprehensive survey. IEEE Access **11**, 113050–113063 (2023)

[22] Metlek, S., Çetıner, H.: Resunet+: A new convolutional and attention block-based approach for brain tumor segmentation. IEEE Access (2023)

[23] Farzamnia, A., Hazaveh, S.H., Siadat, S.S., Moung, E.G.: Mri brain tumor detec- tion methods using contourlet transform based on time adaptive self-organizing map. IEEE Access (2023)

[24] Amara, K., Guerroudji, M.A., Kerdjidj, O., Zenati, N., Ramzan, N.: Holotu- mour: 6dof phantom head pose estimation based deep learning and brain tumour segmentation for ar visualisation and interaction. IEEE Sensors Journal (2023)

[25] Shah, H.A., Saeed, F., Yun, S., Park, J.-H., Paul, A., Kang, J.-M.: A robust approach for brain tumor detection in magnetic resonance images using finetuned efficientnet. Ieee Access **10**, 65426–65438 (2022)

[26] Ezhov, I., Mot, T., Shit, S., Lipkova, J., Paetzold, J.C., Kofler, F., Pellegrini, C., Kollovieh, M., Navarro, F., Li, H., *et al.*: Geometry-aware neural solver for fast bayesian calibration of brain tumor models. IEEE Transactions on Medical Imaging **41**(5), 1269–1278 (2021)

[27] Ahmad, S., Choudhury, P.K.: On the performance of deep transfer learning net- works for brain tumor detection using mr images. IEEE Access **10**, 59099–59114 (2022)

[28] Asif, S., Yi, W., Ain, Q.U., Hou, J., Yi, T., Si, J.: Improving effectiveness of different deep transfer learning-based models for detecting brain tumors from mr images. IEEE Access **10**, 34716–34730 (2022)

[29] Hao, Q., Pei, Y., Zhou, R., Sun, B., Sun, J., Li, S., Kang, X.: Fusing multiple deep models for in vivo human brain hyperspectral image classification to identify glioblastoma tumor. IEEE Transactions on Instrumentation and Measurement **70**, 1–14 (2021)

[30] Kujur, A., Raza, Z., Khan, A.A., Wechtaisong, C.: Data complexity based eval- uation of the model dependence of brain mri images for classification of brain tumor and alzheimer's disease. IEEE Access **10**, 112117–112133 (2022)

[31] Berker, Y., ElHarouni, D., Peterziel, H., Fiesel, P., Witt, O., Oehme, I., Schlesner, M., Oppermann, S.: Patient-by-patient deep transfer learning for drug-response profiling using confocal fluorescence microscopy of pediatric patient-derived tumor-cell spheroids. IEEE Transactions on Medical Imaging **41**(12), 3981–3999 (2022)

[32] Zhang, J.-W., Chen, W., Ly, K.I., Zhang, X., Yan, F., Jordan, J., Harris, G., Plotkin, S., Hao, P., Cai, W.: Dins: deep interactive networks for neurofibroma segmentation in neurofibromatosis type 1 on whole-body mri. IEEE journal of biomedical and health informatics **26**(2), 786–797 (2021)

[33] Liu, Y., Yin, M., Sun, S.: Detexnet: accurately diagnosing frequent and challeng- ing pediatric malignant tumors. IEEE Transactions on Medical Imaging **40**(1), 395–404 (2020)

[34] Gric, T., Sokolovski, S.G., Alekseev, A.G., Mamoshin, A.V., Dunaev, A., Rafailov, E.U.: The discrete analysis of the tissue biopsy images with metamaterial for- malization: identifying tumor locus. IEEE journal of selected topics in quantum electronics **27**(5), 1–8 (2021)

[35] Kromp, F., Fischer, L., Bozsaky, E., Ambros, I.M., Dörr, W., Beiske, K., Ambros, P.F., Hanbury, A., Taschner-Mandl, S.: Evaluation of deep learning architectures for complex immunofluorescence nuclear image segmentation. IEEE Transactions on Medical Imaging **40**(7), 1934–1949 (2021)

[36] Wang, X., Liu, S., Yang, N., Chen, F., Ma, L., Ning, G., Zhang, H., Qiu, X., Liao, H.: A segmentation framework with unsupervised learning-based label mapper for the ventricular target of intracranial germ cell tumor. IEEE Journal of Biomedical and Health Informatics (2023)

[37] Musa, M.J., Sharma, K., Cleary, K., Chen, Y.: Respiratory compensated robot for liver cancer treatment: Design, fabrication, and benchtop characterization. IEEE/ASME Transactions on Mechatronics **27**(1), 268–279 (2021)

[38] Saraf, R., Datta, A., Sima, C., Hua, J., Lopes, R., Bittner, M.L., Miller, T., Wilson-Robles, H.M.: In silico modeling of the induction of apoptosis by cryptotanshinone in osteosarcoma cell lines. IEEE/ACM Transactions on Computational Biology and Bioinformatics **19**(3), 1683–1693 (2020)

[39] Chen, T., Saadatnia, Z., Kim, J., Looi, T., Drake, J., Diller, E., Naguib, H.E.: Novel, flexible, and ultrathin pressure feedback sensor for miniaturized intraven- tricular neurosurgery robotic tools. IEEE Transactions on Industrial Electronics **68**(5), 4415–4425 (2020)

[40] Ling, Z., Yang, S., Gou, F., Dai, Z., Wu, J.: Intelligent assistant diagnosis system of osteosarcoma mri image based on transformer and convolution in developing countries. IEEE Journal of Biomedical and Health Informatics **26**(11), 5563–5574 (2022)

[41] Lim, A., Schonewille, A., Forbrigger, C., Looi, T., Drake, J., Diller, E.: Design and comparison of magnetically-actuated dexterous forceps instruments for neu- roendoscopy. IEEE Transactions on Biomedical Engineering **68**(3), 846–856 (2020)

[42] Gong, Y., Ye, D., Chien, C.-Y., Yue, Y., Chen, H.: Comparison of sonication pat- terns and microbubble administration strategies for focused ultrasound-mediated large-volume drug delivery. IEEE Transactions on Biomedical Engineering **69**(11), 3449–3459 (2022)

[43] Stenman, S., Bychkov, D., Kücükel, H., Linder, N., Haglund, C., Arola, J., Lundin, J.: Antibody supervised training of a deep learning based algorithm for leukocyte segmentation in papillary thyroid carcinoma. IEEE journal of biomedical and health informatics **25**(2), 422–428 (2020)

[44] Zhang, S., Miao, Y., Chen, J., Zhang, X., Han, L., Huang, Z., Pei, N., Liu, H., An, C.: Ccs-net: cascade detection network with the convolution kernel switch block and statistics optimal anchors block in hypopharyngeal cancer mri. IEEE journal of biomedical and health informatics **27**(1), 433–444 (2022)

[45] Gunderman, A.L., Musa, M., Gunderman, B.O., Banovac, F., Cleary, K., Yang, X., Chen, Y.: Autonomous respiratory motion compensated robot for ct-guided abdominal radiofrequency ablations. IEEE Transactions on Medical Robotics and Bionics **5**(2), 206–217 (2023)

[46] Gao, J., Gu, L., Min, X., Lin, P., Li, C., Zhang, Q., Rao, N.: Brain fingerprinting and lie detection: A study of dynamic functional connectivity patterns of decep- tion using eeg phase synchrony analysis. IEEE Journal of Biomedical and Health Informatics **26**(2), 600–613 (2021)

[47] Nakhla, S., Rahawy, A., Abd El Salam, M., Shalaby, T., Zaghloul, M., El-Abd, E.: Radiosensitizing and phototherapeutic effects of aunps are mediated by dif- ferential noxa and bim gene expression in mcf-7 breast cancer cell line. IEEE Transactions on NanoBioscience **20**(1), 20–27 (2020)

[48] Su, H., Kwok, K.-W., Cleary, K., Iordachita, I., Cavusoglu, M.C., Desai, J.P., Fischer, G.S.: State of the art and future opportunities in mri-guided robot- assisted surgery and interventions. Proceedings of the IEEE **110**(7), 968–992 (2022)

[49] Hameed, S.S., Hassan, W.H., Latiff, L.A., Muhammadsharif, F.F.: A compara- tive study of nature-inspired metaheuristic algorithms using a three-phase hybrid approach for gene selection and classification in high-dimensional cancer datasets. Soft Computing **25**, 8683–8701 (2021)

[50] Pang, Q., Zhang, L.: A recursive feature retention method for semi-supervised fea- ture selection. International Journal of Machine Learning and Cybernetics **12**(9), 2639–2657 (2021)

.