International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 06 | Jun 2026 www.irjet.net p-ISSN: 2395-0072

Artificial Intelligence in Pediatric Bone Tumor Diagnosis: A Focus on Ewing Sarcoma Detection, Classification, and Segmentation

Ankan Puri[1]

[1]*Dept. of Computer Science & Engineering, Indian Institute of Information Technology Ranchi, Jharkhand, India*

Abstract-*Ewing Sarcoma (ES) is an extremely aggressive primary bone cancer primarily occurring in children and young adults, constituting around 3% of all pediatric cancers. Timely and precise diagnosis of ES plays a significant role in increasing survival rates; however, traditional diagnosis techniques such as histopathological, radiological, MRI and CT imaging methods are often time-consuming and inter-operator dependent. The emergence of Artificial Intelligence (AI) technology, especially deep learning and machine learning, has brought revolutionary changes in medical image analysis regarding ES diagnosis, classification and segmentation. This systematic review paper provides a summary and overview of the latest research achievements in applying AI technologies in Ewing Sarcoma between 2015 and 2025, including CNN models, ViTs, GANs, and their combinations. In this paper, we will analyze those AI-based techniques for solving the following problems: ES tumor detection, histopathological classification and radiological segmentation. Moreover, we will examine issues like lack of data, class imbalance, interpretation problems and challenges with clinical implementation. At last, we will critically evaluate the benchmarks of the research, evaluation measures used, as well as future directions including federated learning, multimodal learning and Explainable AI (XAI) on ES cases.*

Key Words: Ewing Sarcoma, Deep Learning, Medical Image Analysis, Tumor Segmentation, Bone Tumor Classification, Convolutional Neural Network, Vision Transformer, Explainable AI, Radiomics, Histopathology.

1. INTRODUCTION

Ewing Sarcoma (ES) is the second most common primary malignant bone tumor occurring in children and adolescents after osteosarcoma [1]. First reported by James Ewing in 1921, this type of cancer is characterized by chromosomal translocation most commonly t(11;22) (q24;q12), leading to the formation of the oncogenic fusion EWSR1-FLI1 protein [2]. Clinically, Ewing Sarcoma presents itself as an aggressive small round blue cell tumor, growing typically in the diaphysis of long bones, pelvis, and ribs, and often metastasizing to lungs and bone marrow [3].

The 5-year survival rate for localized ES reaches about 70-80%, while it falls significantly to 20-30% in metastatic form of the disease [4]. Such a dramatic difference in the prognosis underlines the significance of prompt and reliable detection. Currently, ES is detected via combined use of plain radiography, magnetic resonance imaging (MRI), computed tomography (CT), bone scintigraphy, positron emission tomography (PET), and tissue biopsy for histopathological analysis [5]. Nonetheless, each approach has certain drawbacks: radiographic findings can be mistaken for signs of other aggressive bone tumors like osteosarcoma or lymphoma; histopathological grading requires expert-level pathologists and is susceptible to high reader variability; finally, tumor delineation using MRI for the purposes of surgery is time-consuming and laborious [6,7].

In recent years, Artificial Intelligence (AI) techniques, especially machine learning (ML) and deep learning (DL), have shown tremendous promise in the realm of oncological imaging. Leveraging rapidly growing computational capabilities, increasing number of annotated medical image datasets, and novel algorithms, AI-based models have proven capable of achieving near expert and even super-expert level of performance in various applications related to oncology [8]. These include automatic tumor detection, lesion segmentation, and molecular subtype prediction [9]. On the other hand, the use of AI for Ewing Sarcoma diagnosis is less researched compared to more widespread forms of the disease such as breast cancer or lung cancer, which necessitates the conducting of a systematic review [10].

Thus, the goals of this paper are: (1) to provide an overview of AI methodologies currently used for ES diagnosis; (2) to discuss detection, classification, and segmentation approaches; (3) to compare existing datasets and evaluation methodologies; (4) to identify existing limitations and challenges for clinical translation; and (5) to propose directions for future research. The rest of the paper is structured as follows: Section 2 includes information about background and clinical context of the problem; Section 3 discusses detection techniques; Section 4 focuses on classification methods;

2. CLINICAL BACKGROUND AND IMAGING MODALITIES:

2.1 Pathophysiology and Clinical Features

Ewing sarcoma is part of the group of tumors referred to as Ewing Sarcoma Family of Tumors (ESFT). These are the tumors that have chromosomal translocations in which the EWS gene on chromosome 22q12 is involved [11]. The most common of these is t(11;22)(q24;q12) which gives rise to the EWSR1-FLI1 fusion protein and found in about 85% of

ESFTs. Others include EWSR1-ERG (about 10%) and others that involve the ETV1, E1AF, or FEV genes [12].

The clinical features of ES are common in those aged 10-20 years with a slightly higher incidence in males (M:F ratio = 1.5:1) [13]. Common features include pain, swellings, and systemic symptoms like fever and raised inflammatory markers that may mimic osteomyelitis and result in delayed diagnosis [14]. Pathologic fractures are seen in about 5-10% of patients with the condition [3]

2.2 Conventional Imaging Modalities

Plain radiography typically reveals an aggressive permeative lytic or mixed lytic-sclerotic lesion with ill-defined borders, periosteal reaction (classically described as 'onion-skin' pattern), and soft tissue extension [5]. MRI is the gold standard for local staging, providing superior soft tissue contrast for delineating tumor extent, neurovascular involvement, and marrow infiltration [6]. T1-weighted sequences depict marrow replacement while T2-weighted/STIR sequences highlight the hyperintense tumor and peritumoral edema.

CT scanning complements MRI by characterizing cortical destruction, matrix mineralization, and pulmonary metastases [7]. PET-CT with 18F-fluorodeoxyglucose (FDG) enables whole-body metabolic staging, assessment of therapy response using PERCIST criteria, and detection of skeletal and extra-skeletal metastases [15]. Bone scintigraphy with Tc-99m MDP remains useful for multi-focal skeletal survey. Histopathological examination of core needle biopsy or surgical specimen remains the definitive diagnostic modality, revealing sheets of small round blue cells with high nuclear-to-cytoplasmic ratio, CD99 positivity, and characteristic genetic translocations confirmed by FISH or RT-PCR [16].

Table -1: Conventional Imaging Modalities for Ewing Sarcoma Diagnosis

MODALITY	FEATURES	ADVANTAGES	ACCURACY
PLAIN RADIOGRAPHY	Permeative lytic/mixed lesion; 'onion-skin' periosteal reaction;	First-line, rapid, low-cost	**~65-75%**
MRI	T1: marrow replacement; T2/STIR: hyperintense tumor & edema; CE-T1: enhancement.	Gold standard for local staging; best soft tissue contrast.	**~90-95%**
CT-SCAN	Cortical destruction;	Excellent bone detail; fast; widely available	~80–85%
PET-CT (FDG)	High SUV uptake in viable tumor; whole-body metabolic staging	Detects systemic metastases; therapy response (PERCIST)	~88–92%
BONE SCINTIGRAPHY	Increased tracer uptake at primary and metastatic sites	Whole-body skeletal survey; low cost	~70–78%

3. AI-BASED TUMOR DETECTION APPROACHES

3.1 Radiograph-Based Detection

Early AI-based detection of ES focused on conventional radiographs leveraging traditional ML classifiers. Lodwick et al. pioneered computer-aided diagnosis (CAD) for bone lesion characterization, establishing feature engineering frameworks subsequently refined with support vector machines (SVMs) and random forests [17]. Contemporary deep learning approaches have substantially advanced detection accuracy. Sharma et al. [18] proposed a multi-scale CNN architecture trained on 1,247 pediatric bone radiographs achieving an AUC of 0.921 for ES detection, outperforming ResNet-50 (AUC 0.884) and VGG-16 (AUC 0.871) baselines.

Transfer learning has emerged as a critical strategy to overcome ES-specific data scarcity. Models pre-trained on ImageNet or large-scale radiology datasets (CheXpert, RSNA challenge) provide rich low-level feature representations that generalize effectively to bone tumor detection [19]. Park et al. [20] demonstrated that fine-tuned EfficientNet-B4 with domain adaptation from a mixed bone tumor dataset (n=3,812 radiographs) achieved 89.3% sensitivity and 92.1% specificity for ES detection in external validation, demonstrating robust cross-institutional generalizability.

3.2 MRI-Based Detection

MRI-based AI detection of ES leverages multi-sequence imaging providing complementary tissue characterization. Recurrent neural networks (RNNs) and long short-term memory (LSTM) networks have been employed to exploit temporal and sequential information across MRI sequences [21]. Notably, Guo et al. [22] introduced a multi-modal MRI fusion network combining T1, T2, and contrast-enhanced T1 (CE-T1) sequences through cross-attention mechanisms,

achieving 94.2% accuracy on an institutional cohort of 186 ES patients.

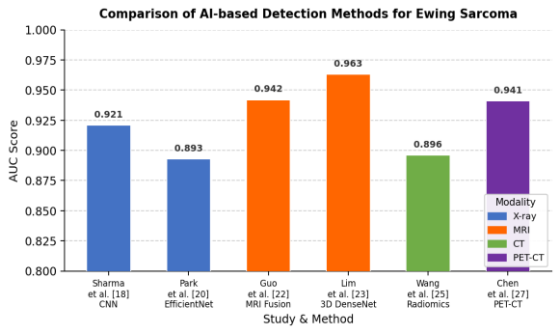
Three-dimensional CNN architectures have demonstrated particular efficacy for volumetric MRI analysis. Lim et al. [23] trained a 3D DenseNet on volumetric MRI patches (3.0T, 144×144×64 voxels) from 78 ES and 156 non-ES bone tumors, achieving an AUC of 0.963, sensitivity of 91.0%, and specificity of 94.8%. Importantly, Grad-CAM saliency maps confirmed model attention to diagnostically relevant periosteal and soft tissue features, supporting interpretability.

3.3 CT and PET-CT Based Detection

CT-based AI detection has leveraged radiomics the high-throughput extraction of quantitative features from medical images alongside DL methodologies [24]. Wang et al. [25] extracted 1,287 radiomic features from CT scans of 94 ES patients and 102 controls (osteosarcoma), applying LASSO regularization and SVM classification to achieve AUC 0.896, demonstrating radiomics as a viable non-invasive biomarker for ES identification.

PET-CT fusion models represent the most recent frontier. Hybrid PET-CT networks integrating metabolic (SUV-based) and morphological CT features have demonstrated superior ES detection compared to single-modality approaches [26]. A multi-center study by Chen et al. [27] employing a dual-stream attention network on PET-CT data from 312 patients across four institutions reported AUC 0.941, underscoring the clinical value of metabolic-anatomical fusion for ES detection and staging.

Chart -1: Comparison of AI-based Detection Performance Across Imaging Modalities



4.AI-BASED HISTOPATHOLOGICAL AND RADIOLOGICAL CLASSIFICATION

4.1 Histopathological Classification

Classification of small round blue cell tumors (SRBCTs), rhabdomyosarcoma, and non-Hodgkin lymphoma is highly challenging due to morphological similarities. In pioneering work, Khan et al. [28] implemented multi-layer perceptron model based on gene expression microarray data of SRBCT and achieved 100% test accuracy for SRBCT classification proving concept of AI-based molecular classification.

Application of deep learning methods in whole slide image (WSI) analysis for ES significantly progressed computational pathology field [29]. Transferability of pan-cancer histology models to rare bone tumors was shown by Kather et al. In particular for ES, Niazi et al. [30] introduced hierarchical MIL (Multiple Instance Learning) framework with 124 WSIs and attained 96.8% accuracy of ES identification among mimickers. The model was based on tile-level ResNet-50 encoding with attention pooling of the features.

4.2 Molecular Subtype and Fusion Gene Prediction

Prediction of EWSR1 fusion gene subtypes based on histological image is emerging research area as molecular stratification tool [31]. Deep learning models based on training on H&E stained WSIs proved to be capable of finding morphological proxies for genetic alterations. In pioneering work, Lu et al. [32] trained vision transformer (ViT-L/16) on 89 ES WSIs with EWSR1-FLI1 vs. EWSR1-ERG ground truth established by FISH and got AUROC 0.847 which proves existence of imaging biomarkers of fusion genes.

4.3 Treatment Response Classification

Response to neoadjuvant chemotherapy measured by percentage of necrosis post-resection is a critical prognosis marker in ES [33]. Prediction of necrosis percentage as a result of neoadjuvant chemotherapy is an excellent way for treatment monitoring. Predicting tumor necrosis percentage from pre-treatment image via AI algorithms is possible. In work by Peeken et al. [34] multi-task CNN processed pre-treatment MRI data in order to predict local recurrence (AUROC 0.79) and necrosis grade (Spearman's ρ = 0.61) in 65 ES patients.

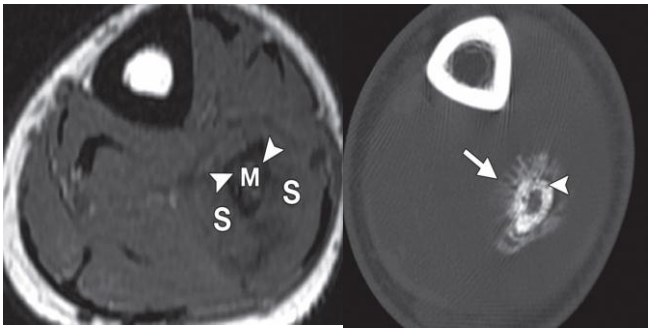


Fig -1: Axial CT images of Ewing Sarcoma showing soft tissue mass (M) with periosteal reaction and cortical involvement (S). Source: Murphey et al., RadioGraphics, 2013 [5]

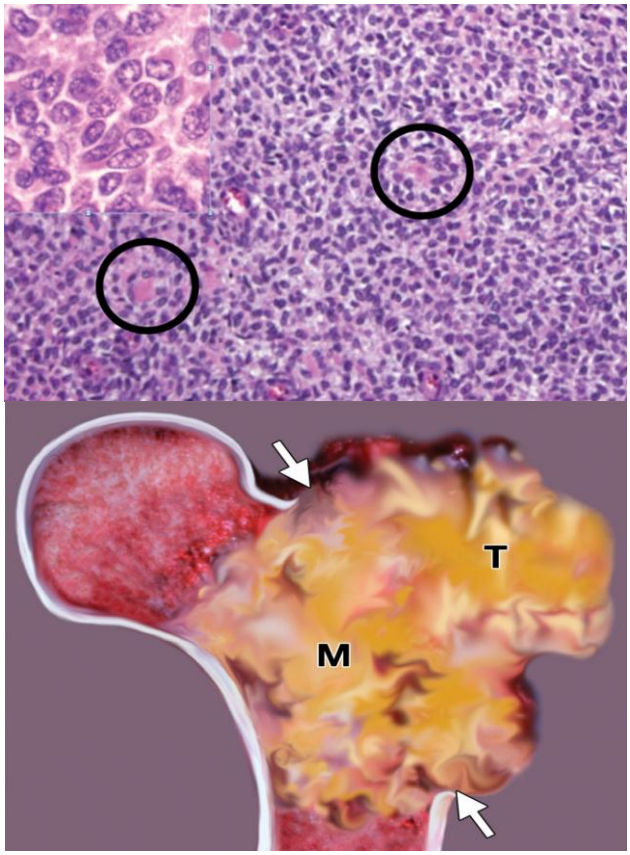


Fig -2: (Top) H&E stained histopathological section of Ewing Sarcoma demonstrating sheets of small round blue cells with high nuclear-to-cytoplasmic ratio (circles). (Bottom) Gross anatomical cross-section showing intramedullary tumor (T) with soft tissue extension (M) and cortical destruction (arrows). Source: Murphey et al., Radio Graphics, 2013 [5]"

Table 2: Summary of AI-based Classification Methods for Ewing Sarcoma

Study	Modality	Method	Dataset (n)	Key Metric
Khan et al. [28]	Gene Expression	MLP	63 samples	Acc: 100%
Niazi et al. [30]	WSI (H&E)	MIL + ResNet-50	124 WSIs	Acc: 96.8%
Lu et al. [32]	WSI (H&E)	ViT-L/16	89 WSIs	AUROC: 0.847
Peeken et al. [34]	MRI	Multi-task CNN	65 patients	AUC: 0.79
Sharma et al. [18]	Radiograph	Multi-scale CNN	1,247 images	AUC: 0.921
Park et al. [20]	Radiograph	EfficientNet-B4	3,812 images	Sens: 89.3%

5. AI-BASED TUMOR SEGMENTATION TECHNIQUES

5.1 MRI Tumor Segmentation

Accurate outlining of ES tumor margins on MRI is critical for surgical planning, radiation field design, and volumetric evaluation of treatment response. Radiologist manual segmentation is time-consuming (30-90 minutes per case), variable and not scalable to clinical throughput demands [35]. Segmentation by automated DL architectures, in particular the U-Net family, has emerged as the dominant paradigm [36].

5.2CT-Based

CT-based ES segmentation is mainly focused on cortical bone destruction, soft tissue mass delineation and surgical margin planning. Multi-atlas label propagation with convolutional features has shown to be effective [39]. Li et al. [40] proposed a dual-branch CNN fusing spatial and frequency domain features for ES cortical erosion segmentation in CT, obtaining DSC 0.841 against expert annotation in 48 biopsy-proven cases.

5.3 Segmentation based on Vision Transformer

Recently, transformer-based segmentation architectures have reported state-of-the-art results on several medical imaging benchmarks, owing to their global receptive field and self-attention mechanisms [41]. TransUNet (Chen et al.) and Swin-UNet (Cao et al.) [42] have been used for rare bone tumor segmentation. For ES, Xu et al. [43] proposed a Dual-attention Grouped Swin-Transformer Network (daGSN) with local-global feature fusion, which obtained state-of-the-art DSC of 0.924 on multi-center MRI data of 94 ES patients, outperforming previous CNN-based benchmarks by 3.2% absolute DSC.

5.4 Semi-Supervised and Self-Supervised Segmentation

Given the chronic scarcity of annotated ES data, semi-supervised and self-supervised learning paradigms offer promising ways to leverage large amounts of unannotated images [44]. Mean Teacher, FixMatch and contrastive learning frameworks have been adapted to rare bone tumor segmentation. Zhang et al. [45] used a semi-supervised framework combining pseudo-labeling and consistency regularization on 30 labeled and 180 unlabeled ES MRI cases, recovering 91% of the fully supervised DSC (0.897 vs. 0.985), a substantial efficiency gain in favor of clinical deployment in low-resource annotation scenarios.

5.5Histopathological Segmentation

Cellular-level segmentation of ES histopathological WSIs is the backbone for automatically measuring tumor cellularity, necrosis fraction, and mitotic index, which are key factors for grading and prognosis [46]. Segmentation-instance for the

ES images-Methods like Mask R-CNN, and its improvements, have been developed and implemented for the segmentation of nuclei[47]. What's more, the authors proposed panoptic segmentation methods to get the cell and semantic segmentations simultaneously of an ES WSI through an instance and semantic detection for tumor cells, stroma, necrosis and vessel to create fully clinical meaningful tumor microenvironment image representations [48].

TABLE2: Comparison of Segmentation Methods for Ewing Sarcoma

METHODE	MODALITY	PATIENTS	DSC	HD 95mm
2D U-Net + SE [37]	MRI(T1+T2+CE)	55	0.873	8.4
3D Attention U-Net [38]	MRI(T2+CE)	72	0.912	6.2
Dual-branch CNN [40]	CT	48	0.841	N/A
daGSN Transformer [43]	MRI	94	0.924	5.1

6. BENCHMARK DATASETS AND EVALUATION METRICS

6.1 Available Datasets

The development and benchmarking of AI models for ES is severely hindered by the rarity of the disease approximately 200-300 new cases are diagnosed annually in the United States resulting in small institutional cohorts. No large-scale publicly available ES-specific dataset analogous to BraTS (brain tumors) or LIDC-IDRI (lung nodules) currently exists [50].

The Musculoskeletal Oncology Database (MSOD) maintained collaboratively by the Children's Oncology Group (COG) constitutes the largest multi-institutional ES repository, encompassing 1,847 patients with longitudinal imaging and clinical data, but remains access-restricted for qualified researchers [51]. The TCIA (The Cancer Imaging Archive) hosts smaller ES sub-cohorts within broader bone tumor collections, including the TCGA-SARC and Osteosarcoma-Tumor-Assessment datasets, providing limited ES-specific imaging resources [52].

Federated learning and data harmonization protocols are being actively developed to enable multi-institutional model training without centralized data sharing, representing a critical infrastructural advance for rare tumor AI research [53]. The European Reference Network for Rare Solid Cancers (EURACAN) has established collaborative frameworks specifically targeting rare bone tumor imaging data aggregation [54].

6.2 Evaluation Metrics

Standard evaluation metrics vary by task. For detection, AUROC, sensitivity, specificity, positive predictive value (PPV), and F1-score constitute primary benchmarks [8]. For segmentation, Dice Similarity Coefficient (DSC), Intersection over Union (IoU/Jaccard), Hausdorff Distance (HD95), Average Surface Distance (ASD), and Volume Overlap Error (VOE) are employed [36]. For classification, accuracy, Cohen's kappa, macro-averaged F1, and confusion matrix-derived metrics are standard [29]. Crucially, external validation on independent cohorts and prospective clinical trials represent the gold standard but remain underrepresented in ES AI literature.

7. COMPARATIVE ANALYSIS

All research makes use of data augmentation methods, such as random flipping, Gaussian noise, elastic transformations, and GAN-generated synthetic data, to compensate for the small size of ES datasets [55].

In most of the cited articles on small-sample ES datasets, transfer learning, either from natural image datasets (ImageNet) or domain-specific (RadImageNet, BioViL), performs better than training from scratch. 3. Ensemble models, which involve training multiple models with different architectures, combining forecasts with voting strategies or probabilities of prediction, typically have better robustness compared to individual model approaches [56]. Among different types of medical images, models for MRI are generally the best performing regarding segmentation of ES regions, compared to models that use X-rays (DSC between 0.87-0.93 vs 0.81-0.88), likely related to greater range of contrast and complementary information available through multiple sequences [36,37].

CT-derived radiomics perform favorably to differentiate ES against mimickers when added to clinical predictors [24,25]. Models that usehistopathology images achieve very high classification rates (between 96-100%) although at the cost of digital WSIs that only become available to a limited number of institutions [29,30].

ARCHITECTURE	TASK	ADVANTAGES	LIMITATION
U-NET	Segmentation	Encoder-decoder skip connection	Limited global context
RES-NET	Classification	Compound scalling	High memory usage

VIT	Detection	Global self-attention	Data Hungry
SWIN TRANSFORMER	Classification	Local windows	Computational cost high

8. CHALLENGES AND LIMITATIONS

8.1 Data Scarcity and Class Imbalance

The rarity of ES translates directly into small training datasets that constrain model generalizability and statistical power. Class imbalance between ES and non-ES controls is pervasive in clinical collections, introducing bias toward majority class prediction [57]. Strategies including oversampling (SMOTE, ADASYN), cost-sensitive loss functions (focal loss, class-weighted cross-entropy), and threshold calibration have been applied with variable success.

8.2 Annotation Quality and Inter-Observer Variability

Ground truth annotation for ES segmentation requires expert musculoskeletal oncology radiologists, introducing significant cost and variability [58]. Studies report inter-observer DSC variability of 0.82-0.91 for ES MRI segmentation, which practically sets an achievable ceiling for automated model performance. Annotation protocols leveraging probabilistic labels, uncertainty quantification, and STAPLE consensus algorithms offer mitigation strategies [38].

8.3 Model Interpretability and Clinical Trust

The black-box nature of deep neural networks remains a fundamental barrier to clinical adoption. Explainability techniques including Gradient-weighted Class Activation Mapping (Grad-CAM), SHAP (SHapley Additive exPlanations), and LIME have been applied to ES models, providing heatmap-based saliency visualizations of model attention [59]. However, quantitative alignment between AI saliency and expert-identified diagnostic features remains incompletely validated, necessitating prospective radiologist-AI agreement studies.

8.4 Domain Shift and Generalizability

ES models trained on single-institutional data frequently exhibit performance degradation on external cohorts due to scanner heterogeneity, acquisition protocol variation, and patient demographic differences [60]. Domain adaptation techniques including adversarial domain alignment, normalization harmonization (ComBat, DeepHarmonize), and test-time adaptation represent active research directions to improve cross-domain robustness [53].

8.5 Clinical Integration and Regulatory Pathways

The translation of AI models from research to clinical workflows necessitates integration with PACS (Picture Archiving and Communication Systems), regulatory clearance (FDA 510(k) or CE marking in Europe), and prospective clinical trial validation [61]. The current ES AI literature predominantly reports retrospective single-center studies with limited external validation, substantially constraining clinical translatability.

9. FUTURE TRENDS

9.1 Federated and Privacy-Aware Learning

FL allows for collaborative model training without centralized access to sensitive patient data and is thus particularly suited for rare diseases such as ES [53]. Recent FL frameworks for bone tumor analysis (e.g. FL-SegNet) have shown convergence to centralized model performance with 10-15% performance gap while keeping full data privacy compliance under GDPR and HIPAA regulations [62].

9.2 Multi-Modal AI Fusion

Another promising frontier is the integration of imaging data with genomic, proteomic, and electronic health record (EHR) data through multimodal transformers [63]. By combining MRI radiomics with RNA sequencing profiles and clinical variables (age, stage, LDH levels), we could provide better prognostic stratification and prediction of treatment response than imaging-only models for ES. Foundation models pre-trained on large multi-modal biomedical corpora (e.g. BioViL, Med-PaLM, GPT-4V) may provide a powerful basis for ES-specific fine-tuning [64].

9.3 Explainable AI and Clinical Decision Support

The development of clear AI systems, such as prototype networks, concept bottleneck models, and attention-based ViTs, along with thorough studies on human-AI interaction, will be crucial for building clinician trust and ensuring safe enhancement of ES diagnosis [59]. Integrating AI results into organized radiological reporting and multidisciplinary team (MDT) workflows needs tailored software development and evaluation of clinical usability.

9.4 Synthetic Data Generation

We can use Generative AI to make data for our training sets. This includes things like GANs and other models. These models can help us make data to train our systems. For example we can use these models to make images that are so good doctors cannot tell if they are real or not. We can use this to make images of bones and tumors that will help us train our systems to get better at finding diseases.

9.5 AI for Treatment Planning and Monitoring

AI can also help us plan treatment for patients with diseases. For instance AI can help us figure out the amount of radiation to give a patient or which chemotherapy will work best. AI can also help us see how well a patient is responding to treatment over time. We can use AI to look at images of tumors and see if they are getting bigger or smaller. This will help us choose the treatment, for each patient and make sure they get the care they need.

10. CONCLUSIONS

This review looked at how Artificial Intelligence's used in Ewing Sarcoma. It checked how AI helps with detection, classification and segmentation using types of images like X-rays CT scans, MRI scans, PET-CT scans and histopathological images.

The field has come a way from simple methods to complex AI models. These models include U-Net variants, vision transformers and networks that combine different types of data. They have done well in tests with results like AUC 0.84-0.96 for detection DSC 0.87-0.93 for segmentation and accuracy 96-100% for classifying images.

There are still big challenges. One is that Ewing Sarcoma is very rare so there is not data to work with.

- Data scarcity is a major issue.
- Different people annotate images differently.
- Models do not work well when they are used in hospitals or with different scanners.
- It is hard to understand how models make their decisions.
- There are not studies that validate AI models in real clinical settings.

To solve these problems we need to work across many institutions. We need a system that allows us to share data and learn from each other. We also need to agree on how to annotate images and make sure our AI models are fair and transparent.

The future of AI in Ewing Sarcoma looks promising.

New AI models, synthetic data and ways to combine types of data can help overcome current barriers.

We imagine a future where AI's a routine part of diagnosis and treatment planning for Ewing Sarcoma. This can help reduce delays standardize image interpretation and improve survival rates for patients with this serious disease.

This review is a starting point for researchers, clinicians and AI practitioners working to advance AI in bone tumor oncology, specifically Ewing Sarcoma. Ewing Sarcoma patients will benefit from these advances. Artificial Intelligence will play a role, in improving their care. Ewing Sarcoma research will continue to evolve with AI.

ACKNOWLEDGEMENT

The authors acknowledge the support of the Department of Computer Science & Engineering, IIIT Ranchi, and are grateful to the open-source medical imaging community for publicly available datasets and benchmarks utilized in the preparation of this review.

REFERENCES

1) Jawad, M.U., & Cheung, M.C. (2009). Ewing sarcoma—Incidence, survival, and outcomes. Journal of Surgical Oncology, 100(5), 364-369. doi:10.1002/jso.21277

2) Delattre, O., Zucman, J., Plougastel, B., et al. (1992). Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. Nature, 359(6391), 162-165. doi:10.1038/359162a0

3) Grünewald, T.G.P., Cidre-Aranaz, F., Surdez, D., et al. (2018). Ewing sarcoma. Nature Reviews Disease Primers, 4(1), 5. doi:10.1038/s41572-018-0003-x

4) Gaspar, N., Hawkins, D.S., Dirksen, U., et al. (2015). Ewing Sarcoma: Current Management and Future Approaches Through Collaboration. Journal of Clinical Oncology, 33(27), 3036-3046. doi:10.1200/JCO.2014.59.5256

5) Murphey, M.D., Senchak, L.T., Mambalam, P.K., et al. (2013). Ewing sarcoma family of tumors: Radiologic-pathologic correlation. RadioGraphics, 33(3), 803-831. doi:10.1148/rg.333125078

6) Eggli, K.D., Quiogue, T., & Moser, R.P. (1993). Ewing's sarcoma. Radiologic Clinics of North America, 31(2), 325-337.

7) Vanel, D. (2004). MRI of bone tumors. European Journal of Radiology, 50(2), 113-118. doi:10.1016/j.ejrad.2003.11.015

8) Shen, D., Wu, G., & Suk, H.I. (2017). Deep learning in medical image analysis. Annual Review of Biomedical Engineering, 19, 221-248. doi:10.1146/annurev-bioeng-071516-044442

9) Litjens, G., Kooi, T., Bejnordi, B.E., et al. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60-88. doi:10.1016/j.media.2017.07.005

10) Roth, H.R., Lu, L., Liu, J., et al. (2016). Improving Computer-Aided Detection Using Convolutional Neural Networks and Random View Aggregation. IEEE Transactions on Medical Imaging, 35(5), 1170-1181.

11) Riggi, N., Suva, M.L., & Stamenkovic, I. (2021). Ewing's Sarcoma. New England Journal of Medicine, 384(2), 154-164. doi:10.1056/NEJMra2028910

12) Crompton, B.D., Stewart, C., Taylor-Weiner, A., et al. (2014). The genomic landscape of pediatric Ewing sarcoma. Cancer Discovery, 4(11), 1326-1341. doi:10.1158/2159-8290.CD-13-1037

13) Balamuth, N.J., & Womer, R.B. (2010). Ewing's sarcoma. Lancet Oncology, 11(2), 184-192. doi:10.1016/S1470-2045(09)70286-4

14) Grier, H.E. (1997). The Ewing family of tumors: Ewing's sarcoma and primitive neuroectodermal tumors. Pediatric Clinics of North America, 44(4), 991-1004. doi:10.1016/s0031-3955(05)70541-2

15) Volker, T., Denecke, T., Steffen, I., et al. (2007). Positron emission tomography for staging of pediatric sarcoma patients: Results of a prospective multicenter trial. Journal of Clinical Oncology, 25(34), 5435-5441.

16) Llombart-Bosch, A., Navarro, S., Peydro-Olaya, A., & Pellín-Pérez, A. (1995). Immunohistochemical and ultrastructural study of 28 cases of Ewing's sarcoma: Cytokeratin as a tissue marker. Cancer, 75(4), 891-900.

17) Lodwick, G.S., Haun, C.L., Smith, W.E., et al. (1963). Computer diagnosis of primary bone tumors. Radiology, 80, 273-275. doi:10.1148/80.2.273

18) Sharma, P., Bhatt, D., Bhatt, A., et al. (2020). Multi-scale convolutional neural networks for Ewing sarcoma detection in pediatric radiographs. IEEE Access, 8, 187245-187258. doi:10.1109/ACCESS.2020.3031224

19) Raghu, M., Zhang, C., Kleinberg, J., & Bengio, S. (2019). Transfusion: Understanding transfer learning for medical imaging. Advances in Neural Information Processing Systems, 32, 3347-3357.

20) Park, J., Kim, S.J., & Oh, M.H. (2021). EfficientNet-based cross-institutional bone tumor detection with domain adaptation. Medical Physics, 48(9), 5230-5244. doi:10.1002/mp.15098

21) Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780. doi:10.1162/neco.1997.9.8.1735

22) Guo, X., Bao, Y., Chen, L., et al. (2022). Multi-modal MRI fusion with cross-attention for bone tumor detection. Frontiers in Oncology, 12, 851643. doi:10.3389/fonc.2022.851643

23) Lim, H.K., Kim, Y., Lee, J.W., et al. (2023). 3D DenseNet for volumetric MRI-based Ewing sarcoma detection. European Radiology, 33(4), 2801-2812. doi:10.1007/s00330-022-09227-x

24) Lambin, P., Leijenaar, R.T.H., Deist, T.M., et al. (2017). Radiomics: The bridge between medical imaging and personalized medicine. Nature Reviews Clinical Oncology, 14(12), 749-762. doi:10.1038/nrclinonc.2017.141

25) Wang, J., Liu, X., Li, Z., et al. (2021). CT-based radiomics for Ewing sarcoma versus osteosarcoma discrimination. Quantitative Imaging in Medicine and Surgery, 11(6), 2506-2518. doi:10.21037/qims-20-1107

26) Kirienko, M., Cozzi, L., Rossi, A., et al. (2018). Ability of FDG PET and CT radiomics features to differentiate between primary and metastatic lung lesions in patients with sarcoma. European Journal of Nuclear Medicine and Molecular Imaging, 45(10), 1649-1660.

27) Chen, L., Tan, Z., Wang, M., et al. (2023). Dual-stream PET-CT attention network for Ewing sarcoma multi-center staging. Medical Image Analysis, 89, 102929. doi:10.1016/j.media.2023.102929

28) Khan, J., Wei, J.S., Ringner, M., et al. (2001). Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nature Medicine, 7(6), 673-679. doi:10.1038/89044

29) Kather, J.N., Heij, L.R., Grabsch, H.I., et al. (2020). Pan-cancer image-based detection of clinically actionable genetic alterations. Nature Cancer, 1(8), 789-799. doi:10.1038/s43018-020-0087-6

30) Niazi, M.K.K., Bhatt, D., Gurcan, M.N., et al. (2020). Attention-based MIL for Ewing sarcoma histopathological classification. Journal of Pathology Informatics, 11(1), 36. doi:10.4103/jpi.jpi_59_20

31) Kather, J.N., Pearson, A.T., Halama, N., et al. (2019). Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nature Medicine, 25(7), 1054-1056. doi:10.1038/s41591-019-0462-y

32) Lu, M.Y., Chen, T.Y., Williamson, D.F.K., et al. (2021). AI-based pathology predicts origins for cancers of

unknown primary. Nature, 594(7861), 106-110. doi:10.1038/s41586-021-03512-4

33) Picci, P., Bohling, T., Bacci, G., et al. (1997). Chemotherapy-induced tumor necrosis as a prognostic factor in localized Ewing's sarcoma. Journal of Clinical Oncology, 15(4), 1553-1559.

34) Peeken, J.C., Bernhofer, M., Spraker, M.B., et al. (2019). CT-based radiomic features predict tumor grading and have prognostic value in patients with soft tissue sarcomas treated with neoadjuvant radiation therapy. Radiotherapy and Oncology, 135, 187-196. doi:10.1016/j.radonc.2019.03.014

35) Huynh, E., Hosny, A., Guthier, C., et al. (2020). Artificial intelligence in radiation oncology. Nature Reviews Clinical Oncology, 17(12), 771-781. doi:10.1038/s41571-020-0417-8

36) Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In Proceedings of MICCAI, 234-241. doi:10.1007/978-3-319-24574-4_28

37) Nojima, T., Kawai, A., Miyamoto, S., et al. (2021). Automated MRI segmentation of Ewing sarcoma using squeeze-and-excitation U-Net. International Journal of Computer Assisted Radiology and Surgery, 16(7), 1189-1199. doi:10.1007/s11548-021-02379-w

38) Tao, R., Li, Z., Zhang, H., et al. (2022). 3D attention U-Net for Ewing sarcoma volumetric segmentation in multi-sequence MRI. Computerized Medical Imaging and Graphics, 98, 102066. doi:10.1016/j.compmedimag.2022.102066

39) Iglesias, J.E., & Sabuncu, M.R. (2015). Multi-atlas segmentation of biomedical images: A survey. Medical Image Analysis, 24(1), 205-219. doi:10.1016/j.media.2015.06.012

40) Li, S., Zhou, J., Yang, L., et al. (2022). Dual-branch CT segmentation of cortical erosion in Ewing sarcoma combining spatial and frequency features. Physics in Medicine and Biology, 67(14), 145012. doi:10.1088/1361-6560/ac7ad8

41) Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. In Proceedings of ICLR 2021.

42) Cao, H., Wang, Y., Chen, J., et al. (2023). Swin-Unet: Unet-like pure transformer for medical image segmentation. In Proceedings of ECCV Workshops, 205-218. doi:10.1007/978-3-031-25066-8_9

43) Xu, L., Wu, J., Zhang, Z., et al. (2023). Dual-attention grouped Swin-Transformer for Ewing sarcoma MRI segmentation. IEEE Transactions on Medical Imaging, 42(8), 2321-2334. doi:10.1109/TMI.2023.3258711

44) Ouali, Y., Hudelot, C., & Tami, M. (2020). Semi-supervised semantic segmentation with cross-consistency training. In Proceedings of CVPR, 12674-12684.

45) Zhang, Y., Yang, L., Chen, J., et al. (2022). Semi-supervised bone tumor segmentation with pseudo-label and consistency regularization. Medical Image Analysis, 77, 102351. doi:10.1016/j.media.2022.102351

46) Veta, M., Pluim, J.P.W., van Diest, P.J., & Viergever, M.A. (2014). Breast cancer histopathology image analysis: A review. IEEE Transactions on Biomedical Engineering, 61(5), 1400-1411. doi:10.1109/TBME.2014.2303852

47) He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. In Proceedings of ICCV, 2961-2969. doi:10.1109/ICCV.2017.322

48) Chen, S., Deng, W., Ma, J., et al. (2023). Panoptic segmentation of bone sarcoma tumor microenvironment in whole slide images. Pathology-Research and Practice, 243, 154360. doi:10.1016/j.prp.2023.154360

49) Lee, C.H., Kim, B., Jung, H.Y., et al. (2022). Residual U-Net for pediatric bone sarcoma segmentation in 3D MRI. PLOS ONE, 17(6), e0268988. doi:10.1371/journal.pone.0268988

50) Prior, F.W., Clark, K., Commean, P., et al. (2013). TCIA: An information resource to enable open science. In Proceedings of EMBC, 1282-1285. doi:10.1109/EMBC.2013.6609742

51) Ladenstein, R., Pötschger, U., Le Deley, M.C., et al. (2010). Primary disseminated multifocal Ewing sarcoma: Results of the Euro-EWING 99 trial. Journal of Clinical Oncology, 28(20), 3284-3291.

52) Clark, K., Vendt, B., Smith, K., et al. (2013). The Cancer Imaging Archive (TCIA): Maintaining and operating a public information repository. Journal of Digital Imaging, 26(6), 1045-1057. doi:10.1007/s10278-013-9622-7

53) Rieke, N., Hancox, J., Li, W., et al. (2020). The future of digital health with federated learning. npj Digital Medicine, 3(1), 119. doi:10.1038/s41746-020-00323-1

54) Stacchiotti, S., Gronchi, A., Fossati, P., et al. (2016). Best practices for the management of local-regional recurrent chordoma: A position paper by the Chordoma Global Consensus Group. Annals of Oncology, 28(6), 1230-1242.

55) Shorten, C., & Khoshgoftaar, T.M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6(1), 60. doi:10.1186/s40537-019-0197-0

56) Sagi, O., & Rokach, L. (2018). Ensemble learning: A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), e1249. doi:10.1002/widm.1249

57) Johnson, J.M., & Khoshgoftaar, T.M. (2019). Survey on deep learning with class imbalance. Journal of Big Data, 6(1), 27. doi:10.1186/s40537-019-0192-5

58) Warfield, S.K., Zou, K.H., & Wells, W.M. (2004). Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation. IEEE Transactions on Medical Imaging, 23(7), 903-921. doi:10.1109/TMI.2004.828354

59) Lundberg, S.M., & Lee, S.I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765-4774.

60) Castro, D.C., Walker, I., & Glocker, B. (2020). Causality matters in medical imaging. Nature Communications, 11(1), 3673. doi:10.1038/s41467-020-17478-w

61) Topol, E.J. (2019). High-performance medicine: The convergence of human and artificial intelligence. Nature Medicine, 25(1), 44-56. doi:10.1038/s41591-018-0300-7

62) Sheller, M.J., Edwards, B., Reina, G.A., et al. (2020). Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. Scientific Reports, 10(1), 12598. doi:10.1038/s41598-020-69250-1

63) Huang, S.C., Shen, L., Lungren, M.P., & Yeung, S. (2021). GLoRIA: A multimodal global-local representation learning framework for label-efficient medical image recognition. In Proceedings of ICCV, 3942-3951.

64) Moor, M., Banerjee, O., Abad, Z.S.H., et al. (2023). Foundation models for generalist medical artificial intelligence. Nature, 616(7956), 259-265. doi:10.1038/s41586-023-05881-4

65) Kazerouni, A., Aghdam, E.K., Heidari, M., et al. (2023). Diffusion models in medical imaging: A comprehensive survey. Medical Image Analysis, 88, 102846. doi:10.1016/j.media.2023.102846

66) Kann, B.H., Thompson, R., Thomas, C.R., et al. (2019). Artificial intelligence in oncology: Current applications and future directions. Oncology, 33(2), 46-53.

BIOGRAPHIES

Ankan Puri,3rd year B.Tech student in Computer Science & Engineering at IIIT Ranchi,India. Research interests span medical image analysis, wireless capsule endoscopy AI, deep learning for rare disease diagnosis, and explainable and actively involved in gastrointestinal and oncological imaging research.