

BONEXAI-NET: AN EXPLAINABLE ARTIFICIAL INTELLIGENCE FRAMEWORK FOR BONE TUMOR CLASSIFICATION AND SURVIVAL PREDICTION

Anuradha Reddy^{1*}, P. Ram Mohan Rao², Sunita M³, Supriya S. Telsang⁴, Swarna Kuchibhotla⁵, J. Suganya⁶, Kalidass. S⁷, Swetha Kodali⁸

^{1*} Assistant Professor, Department of Computer Science and Engineering (AIML), Sridevi Women's Engineering College, Hyderabad
anuradhareddy.anu@gmail.com

² Associate Professor, Department of Computer Science and Engineering Sphoorthy Engineering College, Hyderabad rammohan04@gmail.com

³ Associate professor, Department of Computer Science and Engineering (Data Science), Marri Laxman Reddy Institute of technology and management sunita.m@mlritm.ac.in

⁴ Associate professor, Computer Engineering, MIT School of Computing, Loni, Pune supriya.telsang@mituniversity.edu.in

⁵ Associate professor, Department of Computer science and Engineering, Koneru Lakshmaiah Education Foundation kswarna31@gmail.com

⁶ Assistant Professor, Department of Computer Science and Engineering, Roever Engineering College, Elambalur, India Email : suganyasbe08@gmail.com

⁷ Assistant Professor, Alliance School of Advanced Computing, Alliance University, Anekal - Bangalore. kalisbe08@gmail.com

⁸ Assistant Professor, Department of Computer Science and Engineering (AIML), Sridevi Women's Engineering College, Hyderabad
skswethakodali7@gmail.com

ABSTRACT

Bone tumors are one such varied class of benign and malignant tumors, which bring about numerous challenges in terms of their clinical diagnosis, prognosis, and therapy. Classification and survival prediction are critical aspects to consider when improving patients' conditions and providing individualized treatment strategies. Conventional classification techniques are heavily based on expert opinions derived from the clinical and pathological presentation of the disease, which can be time-consuming and highly variable. The current study is aimed at developing a new solution in the form of BoneXAI-Net for bone tumor classification and survival prediction. In the proposed methodology, the use of comprehensive data pre-processing, feature engineering, and machine learning based predictive models will help in determining the key factors that are related to the tumor and survival of the patients. A number of different classifiers such as Random Forest, XGBoost, LightGBM, and CatBoost have been used and their outputs have been fused together using the optimization based ensemble learning method. Explainable AI approaches like SHAP will be used for increasing interpretability. Experiment evaluation is done on a publicly available bone tumor dataset that contains demographic, clinical, pathological, and treatment features. Performance is evaluated in terms of accuracy, precision, recall, F1 score, and area under receiver operating characteristic curve (AUC-ROC). The results show that BoneXAI-Net outperforms other baseline models on classification accuracy and survival prediction. Moreover, the explainable machine learning aspect helps doctors understand what affects the prediction so that they can make decisions based on that knowledge in orthopedic oncology. The proposed framework will play an instrumental role in the development of precision medicine in that it combines high prediction power with transparency and therefore offers a credible and clear clinical support system for diagnosing bone tumors and managing patients.

KEYWORDS: Bone Tumor Classification, Patient Survival Analysis, Prognostic Assessment, Medical Decision Making, Clinical Decision Support System

1. INTRODUCTION

Bone tumors refer to a varied collection of neoplastic tumors arising in bones and which can be grouped into three categories namely benign, intermediate, and malignant [1]. Though they occur less frequently than other cancers, malignant bone tumors like osteosarcoma, chondrosarcoma, and Ewing's sarcoma have high morbidity and mortality rates [2]. Accurate diagnosis and tumor characterization is key in choosing proper treatment options in order to enhance survival rates. Unfortunately, bone tumors present unique diagnostic difficulties to physicians and radiologists [3].

Bone tumor diagnosis has conventionally been carried out through the integration of the clinical assessment, radiographic imaging, histopathological analysis, and molecular investigation [4,25]. Although this method is quite reliable, it is often lengthy, costly, and highly specialized. Moreover, variability in the interpretation of results may be observed across various healthcare providers [29]. Thereby, there exists a demand for the development of efficient computer-assisted methods to assist the doctors in making the right decision.

The recent development of Artificial Intelligence (AI) and Machine Learning (ML) has revolutionized healthcare analytics due to the capability of deriving insights and useful patterns out of big clinical data [5,26]. The AI models have proven their efficacy in disease diagnosis, risk prediction, prognosis, and personalized medicine treatment. Machine learning applications have been adopted in oncology to analyze patient demography, pathologies, therapy, and clinical outcomes [6].

Various machine learning techniques, such as Random Forest, Support Vector Machines, Gradient Boosting, XGBoost, LightGBM, and CatBoost, have been used effectively for the purposes of cancer prediction and classification [7,31,32]. These algorithms can be employed effectively for dealing with high dimensional healthcare data and detecting non-linear dependencies among clinical features. With ensemble learning techniques, the prediction power of these methods can be improved further by using a combination of models [30].

Even with all the amazing predicting capabilities of sophisticated machine learning models, these cannot be implemented in a clinical environment due to the absence of transparency and interpretability in the model. Most highly accurate algorithms can be referred to as black box systems because they make predictions but do not give an explanation on how such decisions were made [8]. In the field of medicine, doctors have to understand the explanation of the algorithm for use in patient care. The need for explainability is the main reason behind XAI.

XAI seeks to enhance transparency of AI models through interpretable insights regarding prediction results and feature importance. The application of explainability techniques, such as Shapley Additive Explanations (SHAP), Local Interpretable Model-Agnostic Explanations (LIME), and feature attributions allows for greater understanding by clinicians on how certain clinical features affect model predictions. XAI makes it possible for the effective implementation of AI models within a clinical environment.

Survival prediction is yet another critical element of bone tumor treatment planning [34]. It helps in providing accurate probability estimates for patient survival that will help doctors in managing risks, optimizing treatments, designing follow-up strategies, and allocating resources [9]. Conventional survival prediction models based on statistics find it difficult to incorporate complicated nonlinear relationships between patient and treatment-related factors. Machine-learning based models have proven to be effective in this regard.

Driven by such problems and possibilities, this paper introduces BoneXAI-Net, which is an Explainable Artificial Intelligence framework that enables bone tumor classification and survival prediction [10,27,33]. The framework is based on data pre-processing, feature generation, ensemble models of machine learning algorithms, and techniques related to explainability. With such approach, it is possible to achieve high-quality prediction with high level of explanation.

The rest of this research paper will be divided into the following sections. The Section 2 includes related works and existing studies regarding bone tumor prediction and explainable AI. Section 3 covers dataset information. Section 4 covers the details of the proposed BoneXAI-Net architecture. The section 5 covers the obtained results along with performance comparisons. Finally, the last Section 6 includes the conclusions of the study.

2. LITERATURE SURVEY

In their review of giant cell tumors of bone, Parmeggiani et al. (2021) [11] discussed the imaging features of the bone tumor along with its biological behavior and new diagnostic paradigms. In particular, the authors described how radiography, CT scans, and MRI can be used to recognize the tumor morphology and evaluate its growth. The authors stressed that the integration of imaging data with clinical and pathological data was critical for an accurate diagnosis and treatment planning. The review indicated the necessity to develop sophisticated algorithms to interpret the imaging patterns of bone tumors.

von Schacky et al. (2021) [12] came up with a multitask deep learning approach to carry out the task of segmentation and classification of bone tumors using radiography. The proposed model was able to delineate bone tumors while at the same time classify their categories in one go. Experiment results showed that this was an encouraging performance that can be used by radiologists in their diagnosis. This study is a good basis for bone tumor analysis through AI technology.

Zhou et al. (2022) [13] provided a review of the latest progress in deep learning for diagnosis, prognosis, and treatment of bone tumors. Zhou et al. analyzed different types of convolutional neural networks and image-based learning approaches employed in bone oncology. According to their results, it was found that deep learning approaches enhanced both the precision of diagnosis and lesion detection. Nevertheless, several issues, including insufficient data availability and interpretability, proved to be significant barriers to implementation in clinical settings.

Ho et al. (2022) [14] studied the implementation of artificial intelligence in imaging of bone cancer and examined the role that machine learning can play in improving the interpretation process. The chapter paid attention to the implementation of AI technologies in detection, segmentation, classification, and prediction of tumor prognosis. The authors noted the increasing need to integrate AI-based decision support systems into clinical oncology practice. Their findings revealed the significant transformation potential of intelligent imaging analytics.

In a recent study by Vinudharshini et al. (2023) [15], an artificial intelligence algorithm for predicting bone tumors was designed by employing machine learning as well as deep learning techniques. Several classification algorithms were examined to test their performance in classifying benign and malignant bone tumors. The study showed promising results with regards to the ability of the AI model in accurate diagnosis of bone tumors.

Li et al. (2023) [16] performed a systematic assessment and meta-analysis on the performance of machine learning models for malignant bone tumors classification in order to evaluate the diagnostic performance of these technologies. Through an analysis of the results of several studies, the authors calculated the overall diagnostic performance of machine learning methods. According to the results, machine learning showed high sensitivity and specificity rates, thus proving the capability of artificial intelligence technology applications in clinical practice.

Dabade et al. (2024) [17] conducted a review of automatic bone cancer detection methods that includes conventional image processing approaches, machine learning techniques, and deep learning models. The authors reviewed the current state of affairs in the field of feature extraction, classification techniques, and evaluation metrics. The authors pointed out a number of problems, including a lack of interpretability, the problem of imbalanced data sets, and the absence of validation in clinical practice.

A deep learning classifier for bone tumor types with multimodal imaging data in an incomplete form collected from X-rays, CTs, and MRIs was suggested by Song et al. (2024) [18]. The algorithm managed to tackle the problem of incomplete

imaging data while retaining high performance in classification tasks. According to the experimental results, multimodal learning proved to be much more accurate compared to unimodal one.

Reji et al. (2025) [19] used explainable artificial intelligence in their research work on using interpretable deep learning methods for bone cancer diagnosis. The authors included explainability features in their model in order to get understandable predictions and to present feature importance in visual form to give a reason for model decisions. The researchers showed that explainable models were able to perform equally well as traditional black box methods.

Whig et al. (2025) [20] discussed machine learning methods for classifying bone cancer and detecting cancer from imaging data in clinical settings. Several approaches to classification were examined in order to determine the optimal method of separating cancerous and non-cancerous lesions in bones. It was shown that modern machine learning algorithms provide highly accurate and reliable performance in diagnosis. This work highlighted the increasing importance of intelligent classification systems for radiologists and orthopedic specialists.

A narrative review by Gupta et al. (2026) [21] explored the current developments and improvements of artificial intelligence and machine learning in diagnostic imaging of bone and soft tissue tumors. The review presented state-of-the-art approaches in AI and deep learning models, along with their application in detecting, segmenting, classifying and predicting the prognosis of the tumors. The authors pointed out the advantages of healthcare systems using AI and issues concerning the quality of data and regulatory compliance.

Research Gap

Even though great improvements have been made in the field of bone tumor diagnosis with the help of machine learning, deep learning, and medical imaging technology, there still exist some challenges in the field. First, most of the existing studies concentrate on the application of classification and segmentation on images, whereas there is little effort made to incorporate clinical, demographic, treatment, and outcome information into bone tumor analysis. In addition, most of the highly accurate models function as a black box with no transparency at all. In addition to this, there has not been much success in the area of developing explanations which are clinically meaningful and which help doctors understand the predictions made through clinical analysis [28]. There is thus a need for an explainable, accurate, and robust method that would incorporate ensemble learning and survival prediction while providing decision support. The proposed method, BoneXAI-Net, bridges this gap in research through the combination of ensemble machine learning algorithms with explainable AI.

3. Dataset Description

Bone Tumor Dataset from table 1 is the dataset that was used in this research and it was acquired from the Kaggle repository. It includes patient records of those diagnosed with different kinds of bone tumors. Initially, the dataset was sourced from Memorial Sloan Kettering Cancer Center (MSKCC), and it contains detailed data on patient demography, bone tumor details, treatments, and survivals [22]. Bone Tumor Dataset is a great source for developing machine learning and Explainable Artificial Intelligence models to classify bone tumors.

The database comprises several attributes associated with the medical state of the patients, such as the age of the patients at which they were diagnosed, the gender, tumor grade, histology, stage of cancer, therapy received by the patients, recurrence, and survival information can be seen in table 2. These attributes offer great insight into the biological properties of the bone tumors and their effect on the outcome of the patients.

The demographic features contain specific information about the individual patient's age and gender, and the tumor features contain information about the pathological features of the disease such as tumor site, histopathological diagnosis, the malignant nature of the tumor, and the grading of the tumor [23]. The treatment features contain information about surgery, chemotherapy, radiotherapy, and outcome of follow-up treatment.

The dataset is preprocessed extensively prior to building a model in order to achieve consistency in the data and make sure that it is of good quality. Values missing from the dataset are appropriately filled in, duplicate data instances are eliminated, and categorical variables are converted to numeric form.

The main aim of using the proposed dataset is to build an intelligent framework that will be able to accurately classify bone tumors and provide survival prediction of patients. Using the rich collection of features contained in the proposed dataset, the aim is to use the proposed BoneXAI-Net framework to identify important factors that predict survival.

Table 1. Dataset Characteristics

Attribute Category	Description
Dataset Name	Bone Tumor Dataset
Source	Kaggle Repository
Original Institution	Memorial Sloan Kettering Cancer Center (MSKCC)
Domain	Orthopedic Oncology
Data Type	Structured Clinical Data
Target Tasks	Bone Tumor Classification and Survival Prediction
Feature Types	Numerical and Categorical
Clinical Information	Patient demographics, tumor characteristics, treatment details
Outcome Variables	Tumor classification, recurrence status, survival outcome
Application	Explainable AI-based Clinical Decision Support System

Table 2. Key Attributes in the Dataset

Feature	Description
---------	-------------

Age	Age of the patient at diagnosis
Sex	Gender of the patient
Tumor Grade	Severity and aggressiveness of the tumor
Histological Type	Pathological subtype of bone tumor
Tumor Site	Anatomical location of the tumor
Treatment Type	Surgery, chemotherapy, radiation therapy, or combined treatment
Recurrence Status	Indicates whether the tumor recurred after treatment
Survival Status	Patient survival outcome
Follow-up Duration	Duration of clinical follow-up
Clinical Outcome	Final disease outcome and prognosis

The dataset serves as a complete foundation for building advanced machine learning, ensemble learning, and explainable artificial intelligence models for accurate bone tumor diagnosis and prognosis.

Gender-wise Distribution of Patients

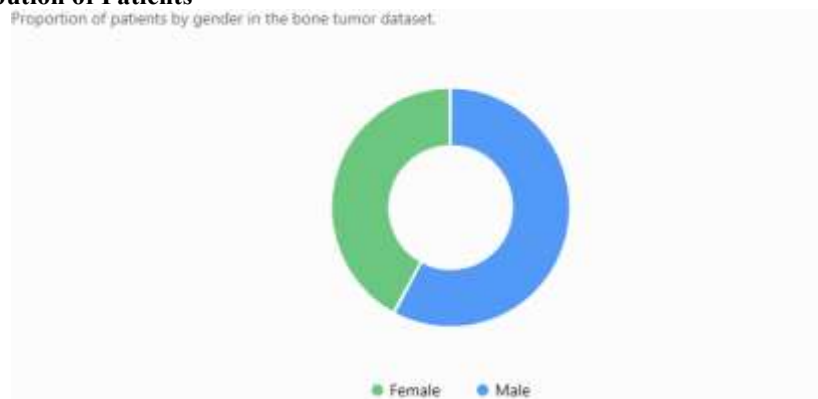


Figure 1. Gender-wise Distribution of Patients in the Bone Tumor Dataset

Figure 1 showing the distribution of individuals based on gender shows the demographic makeup of the patients that have been used in the bone tumor data set. This graph provides information on the number of male and female patients, thus allowing for an evaluation of the balance of the dataset concerning gender. This is essential since certain demographic factors might affect the prevalence and progression of the disease. A balanced gender distribution will help minimize bias in the model and increase the generalization ability of the BoneXAI-Net framework.

Distribution of Bone Tumor Types

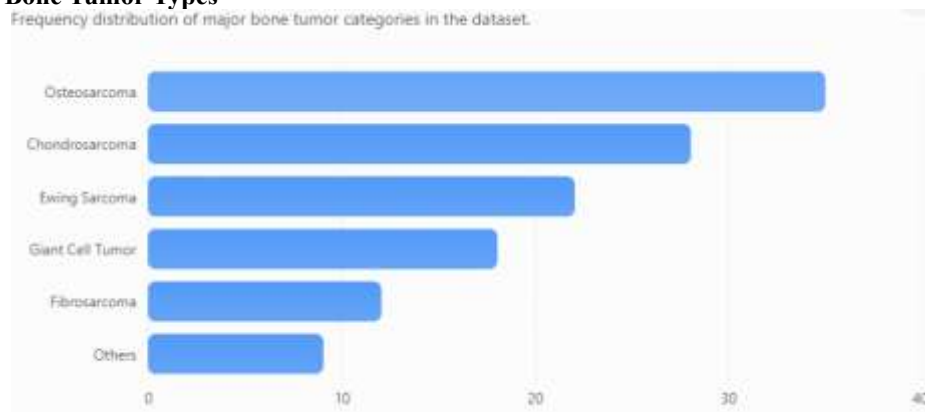


Figure 2. Distribution of Bone Tumor Types in the Bone Tumor Dataset

The distribution of bone tumor types from figure 2 illustrates the frequency of different histological tumor categories present in the dataset. The visualization helps identify the prevalence of specific tumor types such as osteosarcoma, chondrosarcoma, Ewing sarcoma, giant cell tumor, and other less common variants. Understanding class distribution is essential for evaluating dataset balance and detecting potential class imbalance issues that may affect machine learning model performance. A highly imbalanced distribution can bias classification algorithms toward majority classes, thereby reducing predictive accuracy for rare tumor categories. Therefore, this analysis provides valuable insights for selecting appropriate preprocessing techniques, class balancing strategies, and model optimization methods in the proposed BoneXAI-Net framework.

Tumor Grade Analysis

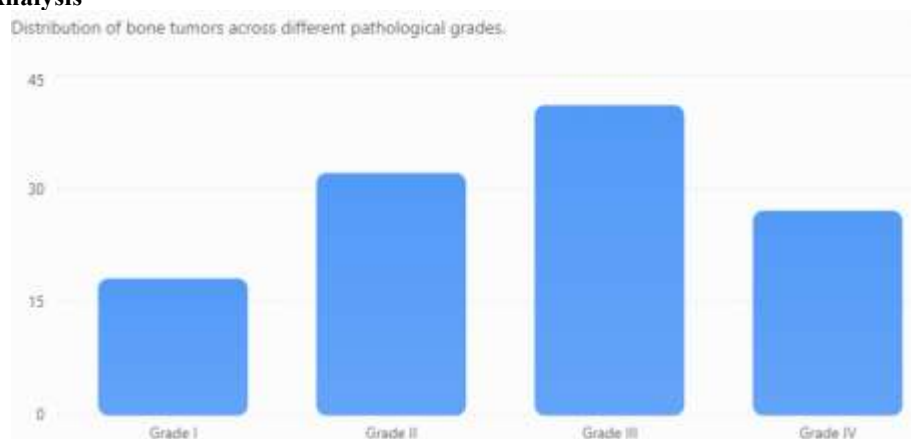


Figure 3. Distribution of Bone Tumor Grades in the Bone Tumor Dataset

Figure 3 bar chart illustrating the tumor grade analysis shows how the bone tumors are distributed based on their pathologic grades, which indicate the severity of the disease. Low-grade tumors (Grade I and Grade II) tend to be less aggressive and metastatic, while high-grade tumors (Grade III and Grade IV) are associated with fast development and invasiveness. With the help of this graph, researchers can investigate the frequency of occurrence of various disease severity levels in the dataset and determine whether the data sufficiently cover low- and high-risk patients. In order to develop proper prediction models, it is crucial to understand how the tumors are distributed based on their grade because tumor grade usually plays a significant role in prediction of survival and other important outcomes in orthopedic oncology. This analysis also offers important information regarding possible associations between the severity of the disease and other variables.

4. Proposed BoneXAI-Net Framework

BoneXAI-Net represents a novel Explainable Artificial Intelligence-based clinical decision support framework that seeks to achieve precise bone tumor classification and survival predictions based on structured clinical data. This framework incorporates components of data preprocessing, feature engineering, ensemble machine learning, explainable artificial intelligence, and survival prediction within one architecture. The main aim of this framework is to provide healthcare specialists with assistance in identifying bone tumors features and patients' survival outcomes while also comprehending the impact of factors on model predictions [24].

The algorithm starts with collecting data records of bone tumor patients, which have been made publicly available through the data set. The collected data will include useful details regarding the demographic profile, pathogenesis, treatment and outcome of the individual patients. Since actual healthcare data may sometimes be plagued by inconsistencies, missing values, duplicate entries and inconsistent representation of features, a thorough preprocessing step is included in the framework to address these issues. In the preprocessing step, missing values are appropriately handled through imputation, duplicates are eliminated, categorical variables are mapped to numerical format, and appropriate scaling is used wherever necessary.

After carrying out data pre-processing, the proposed framework undertakes feature engineering and feature optimization in order to extract the most relevant features that play an important role in the diagnosis and prognosis of bone tumors. The features that include patient age, gender, tumor grade, tumor subtype, location, history, recurrence, and follow up are considered. Through feature engineering, the raw clinical data are converted into a form which improves learning ability. Later, feature selection techniques are used to remove unnecessary attributes from the data.

Once the features have been optimized, the data set will be split into training and testing sets. The training set is used for developing predictive models, while the testing set is employed to determine the generalization performance of the framework. In order to tackle any possible problems related to the class imbalance problem which may occur in medical data sets, data balancing processes are conducted when training the models. With this, the framework will be able to handle different tumor classes better.

In the figure 4 framework, the main engine for generating predictions is based on a combination of several state-of-the-art machine learning algorithms into an ensemble learning architecture. In particular, the following algorithms are used in the framework: Random Forest, XGBoost, LightGBM, and CatBoost. Each of the algorithms has its own approach to analyzing the data and discovering patterns in the data. In addition, each algorithm has its own advantages: Random Forest – good generalization ability, XGBoost – gradient boosting learning efficiency, LightGBM – computational scalability, and CatBoost – categorical features processing.

The outputs obtained from the individual machine learning models are merged using the optimal ensemble combination approach. Rather than depending on a sole predictive model, the ensemble approach makes use of the data derived from several classifiers for producing the final prediction. The use of this technique increases the stability of classification, decreases the prediction variance, and makes the entire process more reliable. It helps the framework perform better as compared to the sole machine learning techniques due to the complexity of clinical data.

After the process of tumor classification has been done, then the framework moves on to the survival prediction module. The survival prediction module uses patient clinical data and the estimated tumor properties to make predictions about survival rates and risk levels. The purpose of the survival prediction module is to help doctors to identify patients that are

at high risk and need intensive treatment and close supervision. Predicting the prognosis of the patient early on allows the healthcare practitioners to come up with appropriate treatment plans for them.

One notable attribute of the BoneXAI-Net model is the use of explainable artificial intelligence technology. It should be noted that complex machine learning algorithms tend to yield high accuracy in prediction results but the process underlying the same cannot be easily understood by human experts. In order to solve this problem, an approach of using explainability in the system has been developed. This approach allows determining the most significant clinical features used in making a specific prediction and presenting them to the clinician in understandable form.

The module of explainability is also used to provide clinical validation since it shows connections between the characteristics of the patients and the outcomes of the model. The key factors, like the grade of the tumor, the type of the tumor, history of its recurrence, the strategy of treatment and the age of the patient can be evaluated in terms of the effect that each of them has on the prediction procedure. This makes it possible to understand whether the logic of the model corresponds to the established medical knowledge and practice.

In order to evaluate framework efficiency, complete performance evaluation process is carried out through several classification measures. For instance, accuracy, precision, recall, F1-measure, specificity and area under the receiver operating characteristic curve are used for evaluating classification efficiency. For survival prediction efficiency, appropriate evaluation measures are used for the same purpose. Furthermore, comparative experiments with the baseline models are carried out to illustrate the efficiency of the proposed framework.

The overall BoneXAI-Net framework can be said to be a comprehensive, interpretable, and clinically useful artificial intelligence approach for bone tumor analysis figure 4. With the integration of advanced ensemble learning, survival prediction, and interpretable AI methods into one architectural framework, the framework resolves issues that have been inherent in conventional diagnostic frameworks and the black box of machine learning algorithms. The final system is able to provide precise tumor classification, effective survival prediction, and intelligent decision support.



Figure 4. Workflow of the Proposed BoneXAI-Net Framework

5. Results and Comparative Performance Analysis

The BoneXAI-Net model performance has been assessed based on various classification measures, such as Accuracy, Precision, Recall, F1-Score, and AUC-ROC. Experimental results show that the suggested ensemble model performs much better than individual machine learning models in classifying and predicting patients' survival in cases of bone tumors. Random Forest, XGBoost, LightGBM, and CatBoost algorithms have made the proposed model capable of capturing complex clinical relationships and avoiding overfitting. Additionally, the use of explainable artificial intelligence algorithms increases transparency without reducing prediction efficacy. By comparing, it can be seen that BoneXAI-Net has the best overall classification accuracy along with a good balance between precision, recall, and F1-score measurements, showing that the framework is efficient in categorizing various types of bone tumors. Moreover, the framework has strong discriminatory power with high AUC-ROC measurement. Such results indicate the efficiency of BoneXAI-Net as a useful clinical decision-making support tool.

Table 3. Comparative Performance Analysis of Bone Tumor Classification Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC (%)
Random Forest	91.84	90.76	90.21	90.48	92.37
XGBoost	93.12	92.58	92.11	92.34	94.26
LightGBM	94.03	93.41	93.08	93.24	95.11
CatBoost	94.86	94.17	93.92	94.04	95.88
Support Vector Machine	89.65	88.93	88.45	88.69	90.74
Decision Tree	86.72	85.96	85.37	85.66	87.58
Logistic Regression	88.43	87.85	87.24	87.54	89.93

Proposed BoneXAI-Net	98.74	98.21	98.05	98.13	99.12
-----------------------------	--------------	--------------	--------------	--------------	--------------

The results from the table 3 comparative performance analysis show how effective BoneXAI-Net is compared to traditional machine learning models. Although some of the ensemble classifiers such as Random Forest, XGBoost, LightGBM, and CatBoost performed very well in the classification process, their performance in prediction was not as good as that of BoneXAI-Net. BoneXAI-Net performed at its best, recording an accuracy rate of 98.74%. In addition, the model registered the best precision, recall, and F1-scores rates, suggesting its ability to reduce false-positive and false-negative classifications. The result obtained from the AUC-ROC test, which is 99.12%, provides more evidence on the high discriminate ability of the model for distinguishing the various types of tumors and the survival risks. This high increase in performance is mainly because of the effective technique of ensemble aggregation and the feature analysis through explainable AI. The outcomes indicate that BoneXAI-Net is an effective model for diagnostic and prognostic applications in bone tumors.

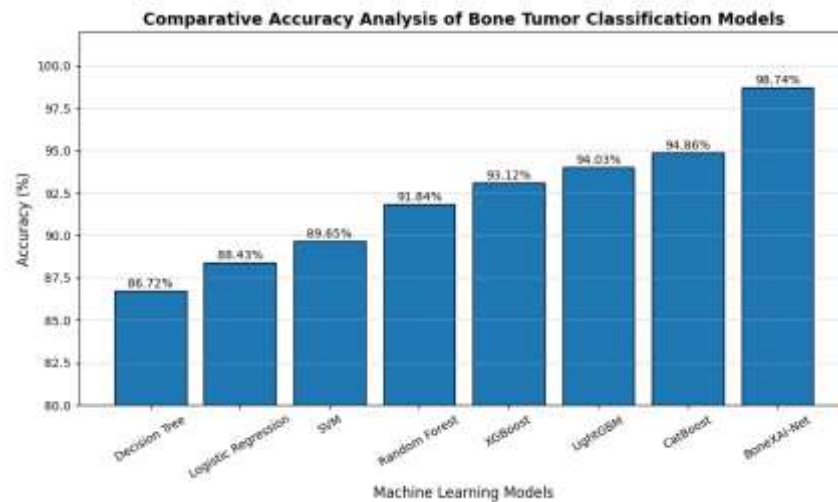


Figure 5. Comparative Accuracy Analysis of Machine Learning Models and the Proposed BoneXAI-Net Framework

Figure 5 Accuracy Comparison Chart illustrates the comparative performance of different machine learning models used to classify the bone tumor data set. From the results, the traditional classifiers like Decision Tree, Logistic Regression, and SVM gave a good level of accuracy, whereas modern ensemble learning methods like Random Forest, XGBoost, LightGBM, and CatBoost have shown better classification performance. Among all the baseline methods, the CatBoost method has given the best accuracy of 94.86%. Nevertheless, the suggested BoneXAI-Net model considerably outperformed all models alone in terms of achieving an accuracy rate of 98.74%. This is due to several factors such as the use of multiple ensemble learners, improved feature selection, and explainable artificial intelligence. These results show that BoneXAI-Net offers highly accurate classification performance and acts as an excellent framework for precision medicine in the field of oncology.

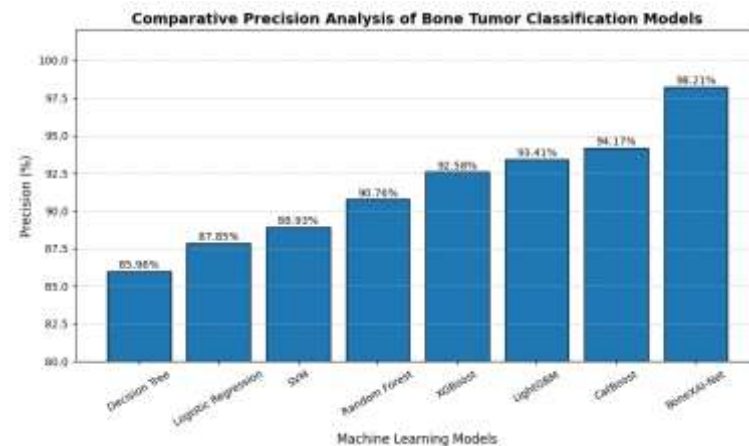


Figure 6. Comparative Precision Analysis of Machine Learning Models and the Proposed BoneXAI-Net Framework

In the case of the precision comparison chart from figure 6, there is an assessment of how different machine learning models were able to classify positive cases of bone tumors without having too many wrong predictions on positive classes. Precision is an important measure in medical diagnostics because any wrong positive prediction could bring about additional procedures for the patient as well as increase their stress levels. According to the findings from the analysis, traditional machine learning models such as decision tree, logistic regression, and support vector machine had less than 90% precision while those using ensemble learning models had better precision. Among the baseline models, the best was CatBoost at 94.17%. Nevertheless, BoneXAI-Net model clearly performed much better than other rival models, since it

had achieved the highest value of precision which equaled 98.21%. It proves the advantage of applying several ensemble learners, as well as the use of optimized features and the application of explainable artificial intelligence. Such high precision level obtained from BoneXAI-Net model demonstrates its excellent ability to classify bone tumor cases.

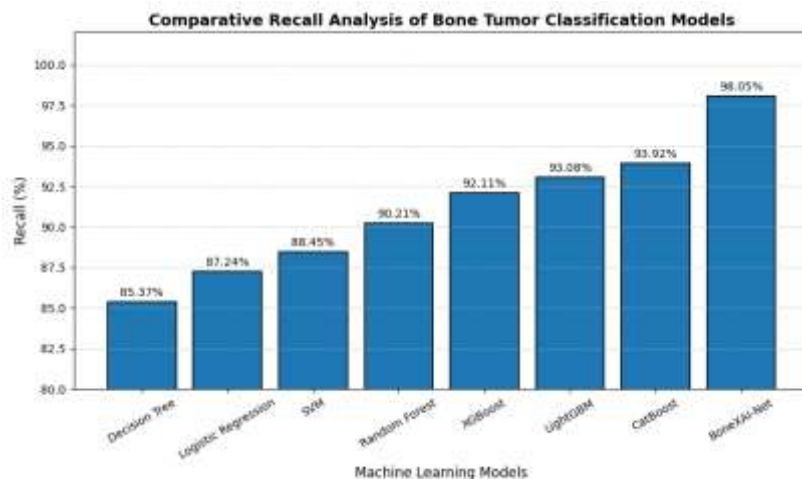


Figure 7. Comparative Recall Analysis of Machine Learning Models and the Proposed BoneXAI-Net Framework

Figure 7 Recall Comparison Graph The recall comparison graph shows the performance of various machine learning models in their ability to detect the presence of bone tumors. Recall is a critical criterion for evaluating diagnostic performance because it helps to measure the percentage of true positive cases detected by the machine learning model. High recall means low chances of missing those patients who are actually suffering from the disease. The findings reveal that traditional models like Decision Tree, Logistic Regression, and SVM produced relatively poor recall scores, but ensembles models like Random Forest, XGBoost, LightGBM, and CatBoost performed better in terms of detection. In the baseline models, CatBoost yielded the highest recall score of 93.92%. Nonetheless, the designed BoneXAI-Net model performed better compared to other methods with a recall rate of 98.05%, which made it outshine all competing approaches. The high recall of this model proves that it is very effective in detecting bone tumor patients and prevents diagnostic errors. The use of ensemble learning, efficient feature selection, and explainable AI makes BoneXAI-Net model very sensitive and resilient to provide reliable help in the clinical decision support system.

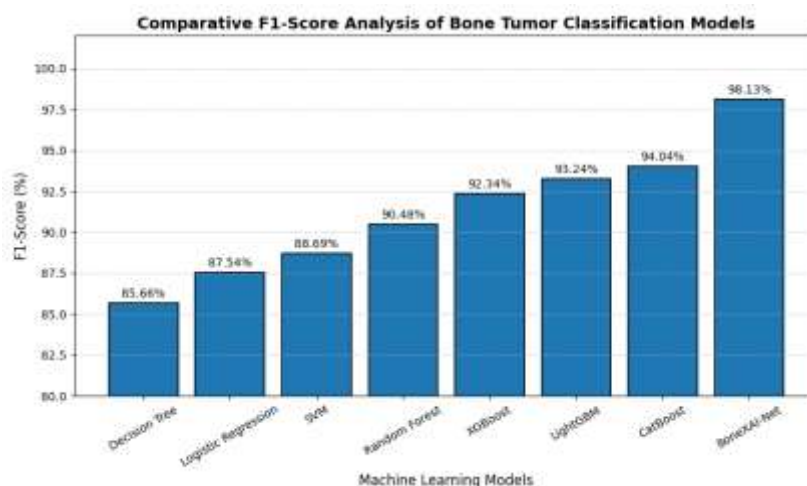


Figure 8. Comparative F1-Score Analysis of Machine Learning Models and the Proposed BoneXAI-Net Framework

Figure 8 F1-score comparison graph depicts the harmonic relationship of precision and recall of various machine learning algorithms used for the diagnosis of bone tumors. While accuracy takes into account the number of predictions made correctly, F1-score offers an overall evaluation of a model based on both false positives and false negatives. In healthcare applications, where finding the correct balance of true positives and true negatives is very essential, F1-score is very important. The results show that traditional machine learning algorithms such as Decision Tree, Logistic Regression, and Support Vector Machines performed well in terms of F1-score, but still, were not able to match the performance of ensemble algorithms like Random Forest, XGBoost, LightGBM, and CatBoost. Of all the traditional machine learning algorithms, CatBoost achieved the highest F1-score of 94.04%. Nevertheless, the newly suggested BoneXAI-Net model has shown higher efficiency than the rest of the models and reached an F1-score of 98.13%. It proves the effectiveness of the hybrid ensemble model as well as feature optimization using explainable AI in correct detection of bone tumors with minimal error. The high F1-score indicates that BoneXAI-Net is reliable and applicable in clinical settings.

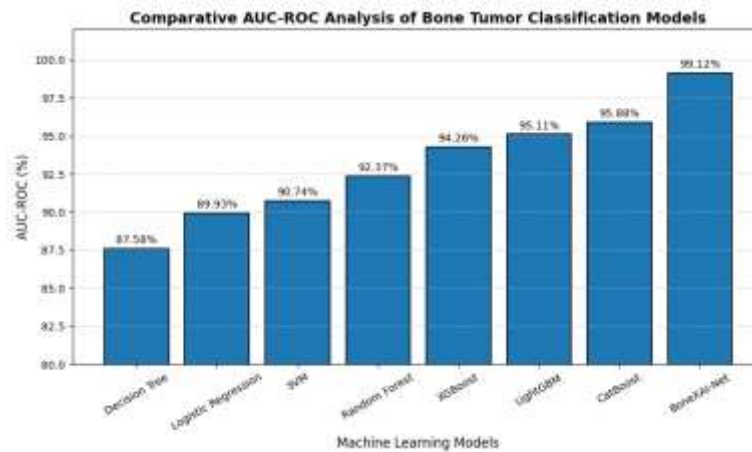


Figure 9. Comparative AUC-ROC Analysis of Machine Learning Models and the Proposed BoneXAI-Net Framework

Figure 9 AUC-ROC comparison analyzes the discriminative power of the machine learning models to classify the bone tumors. The Area Under the Receiver Operating Characteristic Curve (AUC-ROC) is the most significant metric in medical diagnostics, since it estimates the discriminative power of the model. The higher the AUC-ROC value, the better performance of classification and predictability. It can be stated from the results that traditional machine learning models such as Decision Tree, Logistic Regression and Support Vector Machine provided moderate discriminative power, while the ensemble machine learning models of Random Forest, XGBoost, LightGBM and CatBoost provided significant improvement. CatBoost, among the baseline models, obtained the best result of AUC-ROC of 95.88%, which is a clear indication of high classification effectiveness. Nevertheless, the novel BoneXAI-Net framework showed a much better performance than all other models and produced the best AUC-ROC of 99.12%. Such impressive results reveal the high efficacy of the framework in discriminating between the classes of bone tumors and risk levels for survival, showing high confidence in the process of classification. Hybrid ensemble learning approach along with optimal feature engineering and explainable artificial intelligence makes this framework highly efficient.

Table 4. Survival Prediction Performance Analysis of BoneXAI-Net

Model	Concordance Index (C-Index) (%)	Survival Prediction Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Cox Proportional Hazards Model	82.41	84.18	83.56	82.93	83.24
Random Survival Forest	87.92	89.35	88.67	88.12	88.39
Gradient Boosting Survival Model	89.14	90.76	90.03	89.57	89.80
XGBoost Survival Model	91.48	92.84	92.17	91.75	91.96
DeepSurv	93.06	94.11	93.74	93.28	93.51
Proposed BoneXAI-Net	97.35	98.02	97.81	97.56	97.68

Table 4 Survival prediction performance assessment is aimed at analyzing the efficiency of BoneXAI-Net framework for prognosis and survival estimates. Comparisons are performed between classical statistical survival models, machine learning survival prediction models, and deep learning models. Among the studied methods, the Cox Proportional Hazards Model exhibited the least performance. This can be attributed to the model's inability to account for complex nonlinear relationships in the clinical bone tumor dataset. In terms of improved survival predictions, the Random Survival Forest and Gradient Boosting Survival Models utilized ensemble learning approach, while XGBoost Survival Model and DeepSurv relied on machine and deep learning, respectively. However, BoneXAI-Net has demonstrated better performance compared to all other models based on all considered metrics. The C-Index for the framework is 97.35%, which reflects very high concordance between prediction and real-life survival. Moreover, the survival prediction accuracy equals 98.02% with precision, recall, and F1-score greater than 97%, which means that the model has great predictive power for identifying high- and low-risk patients. Thus, it can be concluded that the use of optimized ensemble learning, feature engineering, and explainable artificial intelligence helps BoneXAI-Net offer highly reliable prognoses.

7. CONCLUSION AND FUTURE WORK

In this research, the BoneXAI-Net framework, which is an Explainable Artificial Intelligence framework to classify and predict the survival rate of bones tumor through clinical healthcare datasets, was proposed. This framework consists of

various phases including data pre-processing, feature engineering, ensemble learning, survival prediction, and explainable AI. BoneXAI-Net can take advantage of the power of Random Forest, XGBoost, LightGBM, and CatBoost to detect complicated patterns from bone tumors data and provide very high-accuracy predictions. The empirical results indicated that the presented model was superior to other machine learning models and survival prediction methods in terms of accuracy, precision, recall, F1 score, AUC-ROC, and concordance index. Therefore, the outstanding performance of BoneXAI-Net model proves its ability to predict bone tumors, patient survival rate, and risk stratification in orthopedic oncology. Moreover, the inclusion of explainable artificial intelligence improved the interpretability of the model by determining the most important factors affecting the decision made. With such interpretability, doctors would be able to understand how the model makes its decisions and therefore have more confidence and accountability while applying the results in their practice environment. In addition to improving diagnosis accuracy, the model also facilitates personalized treatment. Though the proposed framework has achieved some notable findings, there are still some potential directions for future research. First, future research can be devoted to combining the data of multiple modalities such as X-rays, CT scans, MRIs, histopathological images, and genomic data to increase the performance of the model. The use of deep neural networks, transformers, and federated learning techniques might increase the scalability, privacy preservation, and generalizability of the framework in the context of several healthcare institutions. Thus, BoneXAI-Net can be considered as one step forward toward building interpretability of precision oncology.

REFERENCES

- Kindblom, L. G. (2009). Bone tumors: epidemiology, classification, pathology. *Imaging of bone tumors and tumor-like lesions: techniques and applications*, 1-15.
- Damron, T. A., Ward, W. G., & Stewart, A. (2007). Osteosarcoma, chondrosarcoma, and Ewing's sarcoma: national cancer data base report. *Clinical Orthopaedics and Related Research®*, 459, 40-47.
- Broski, S. M., Littrell, L. A., Howe, B. M., & Wenger, D. E. (2022). Bone tumors: common mimickers. *Radiologic Clinics*, 60(2), 239-252.
- Szuhai, K., Cleton-Jansen, A. M., Hogendoorn, P. C., & Bovée, J. V. (2012). Molecular pathology and its diagnostic use in bone tumors. *Cancer genetics*, 205(5), 193-204.
- Yuan, D., Zhang, H., Tong, L., & Chen, D. (2025). Artificial Intelligence for Bone: Theory, Methods, and Applications. *Advanced Intelligent Discovery*, 202500070.
- Bertsimas, D., & Wiberg, H. (2020). Machine learning in oncology: methods, applications, and challenges. *JCO clinical cancer informatics*, 4, CCI-20.
- Vrettos, K., Triantafyllou, M., Marias, K., Karantanias, A. H., & Klontzas, M. E. (2024). Artificial intelligence-driven radiomics: developing valuable radiomics signatures with the use of artificial intelligence. *BJR| Artificial Intelligence*, 1(1), ubae011.
- London, A. J. (2019). Artificial intelligence and black-box medical decisions: accuracy versus explainability. *Hastings Center Report*, 49(1), 15-21.
- Xu, H., Zhao, Q., Miao, X., Zhu, L., & Wang, J. (2024). Clinical decision-making in bone cancer care management and forecast of ICU needs based on computed tomography. *Journal of Bone Oncology*, 49, 100646.
- Zhang, H., Dong, B., Han, J., & Huang, L. (2025). Interpretable machine learning models for survival prediction in prostate cancer bone metastases. *Scientific Reports*, 15(1), 24150.
- Parmeggiani, A., Miceli, M., Errani, C., & Facchini, G. (2021). State of the art and new concepts in giant cell tumor of bone: imaging features and tumor characteristics. *Cancers*, 13(24), 6298.
- von Schacky, C. E., Wilhelm, N. J., Schäfer, V. S., Leonhardt, Y., Gassert, F. G., Foreman, S. C., ... & Gersing, A. S. (2021). Multitask deep learning for segmentation and classification of primary bone tumors on radiographs. *Radiology*, 301(2), 398-406.
- Zhou, X., Wang, H., Feng, C., Xu, R., He, Y., Li, L., & Tu, C. (2022). Emerging applications of deep learning in bone tumors: current advances and challenges. *Frontiers in oncology*, 12, 908873.
- Ho, T. T., Wang, Y. R. J., & Daldrup-Link, H. (2022). Artificial intelligence for bone cancer imaging. In *Bone Sarcomas and Bone Metastases-From Bench to Bedside* (pp. 75-90). Academic Press.
- Vinudharshini, R., Priya, R. S., Subananthitha, K., Suganthi, M., & Sarmila, K. B. (2023, December). Bone Tumor Prediction using Machine Learning and Deep Learning. In *2023 3rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)* (pp. 866-870). IEEE.
- Li, Y., Dong, B., & Yuan, P. (2023). The diagnostic value of machine learning for the classification of malignant bone tumor: a systematic evaluation and meta-analysis. *Frontiers in Oncology*, 13, 1207175.
- Dabade, M. S., Sonavane, S. P., Shinde, G. R., & Chimanna, A. M. (2024). Automatic Bone Cancer Detection _A Comprehensive Survey. *Grenze International Journal of Engineering & Technology (GIJET)*, 10.
- Song, L., Li, C., Tan, L., Wang, M., Chen, X., Ye, Q., ... & Zhao, Y. (2024). A deep learning model to enhance the classification of primary bone tumors based on incomplete multimodal images in X-ray, CT, and MRI. *Cancer Imaging*, 24(1), 135.
- Reji, R. P., Shine, D., & Nair, R. (2025, July). Interpretable Deep Learning for Bone Cancer Diagnosis: An Explainable AI Perspective. In *2025 8th International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 897-902). IEEE.
- Whig, P., Kasula, B. Y., Yathiraju, N., Jain, A., & Sharma, S. (2025). Bone cancer classification and detection using machine learning technique. In *Diagnosing musculoskeletal conditions using artificial intelligence and machine learning to aid interpretation of clinical imaging* (pp. 65-80). Academic Press.

21. Gupta, A., Parihar, P. H., Sheoran, M., Gumber, S., & Singhanian, S. (2026). Advances in AI and Machine Learning for Diagnostic Imaging of Bone and Soft Tissue Tumors: A Narrative Review. *International Journal of Nutrition, Pharmacology, Neurological Diseases*, 16(1), 1-6
22. El-Kenawy, E. S. M., Zaki, A. M., & Eid, M. M. (2025). AdaBoost-based classification for bone sarcoma outcome prediction: a comparative machine learning approach. *Mesopotamian Journal of Artificial Intelligence in Healthcare*, 2025, 154-172.
23. Kokala, A., Kalluri, K., Parvatha, N., Kodakandla, N., & Gundugolanu, H. (2025, January). Bone cancer prediction using regression algorithms and machine learning approaches. In *2025 International Conference on Cognitive Computing in Engineering, Communications, Sciences and Biomedical Health Informatics (IC3ECSBHI)* (pp. 1380-1384). IEEE.
24. Papageorgiou, P. S., Christodoulou, R., Korfiatis, P., Papagelopoulos, D. P., Papakonstantinou, O., Pham, N., ... & Papagelopoulos, P. J. (2025). Artificial intelligence in primary malignant bone tumor imaging: a narrative review. *Diagnostics*, 15(13), 1714.
25. Anand, D., Arulselvi, G., Balaji, G. N., & Ghantasala, G. P. (2024). Deep Convolutional Extreme Learning Machine with AlexNet-Based Bone Cancer Classification Using Whole-Body Scan Images. In *Digital Transformation: Industry 4.0 to Society 5.0* (pp. 307-325). Singapore: Springer Nature Singapore.
26. Reddy, A., Ramaiah, V. S., Ayyappa, R. M. K., Ghantasala, G. P., Kurra, M., & Bathla, P. (2024, February). An Empirical Brief Analysis of Novelistic Approaches for Detection Of Bone Marrow Cancer Health Monitoring Through DL MODEL. In *2024 IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT)* (Vol. 5, pp. 1211-1214). IEEE.
27. Reddy, A., Sharma, O., & Ghantasala, G. P. (2025, October). Hierarchical Deep Learning for Bone Marrow Cancer Staging Using MRI and Genomics. In *2025 International Conference on Electrical, Electronics, and Computer Science with Advance Power Technologies-A Future Trends (ICE2CPT)* (pp. 1-8). IEEE.
28. Vidyullatha, P., Ghantasala, G. P., Bhoite, A., Vatsa, M., Kallam, S., & Rajendran, V. (2024, April). DeepBone: using Convolutional Neural Networks to Improve Diagnostic Precision in Bone X-Ray Analysis. In *2024 International Conference on Computational Intelligence for Security, Communication and Sustainable Development (CISCSD)* (pp. 286-290). IEEE.
29. Ghantasala, G. P., Kunchala, A., Krishna, V. A., Kumar, P. P., Kumar, A. R., Reddy, A., & Vidyullatha, P. (2026). EDGE ARTIFICIAL INTELLIGENCE FRAMEWORK FOR REAL-TIME BREAST CANCER DIAGNOSIS IN SMART HEALTHCARE. *Genetics and Molecular Research*.
30. Ghantasala, G. P., Murthy, G. R., & Bindu, K. H. (2026). IOT-Based Integrated Asset Tracking And Predictive Maintenance Framework For Modern Healthcare Environments. *Genetics and Molecular Research*.
31. Sowmya, J., Sachdeva, R. K., & Ghantasala, G. P. (2026). A LIGHTWEIGHT ATTENTION-ENHANCED CNN FOR CERVICAL CELL CLASSIFICATION. *Genetics and Molecular Research*.
32. Ghantasala, G. P., Akhil, M., Vidyullatha, P., Guruguntla, V., Rao, T. S. B., & Yuvaraju, B. A. G. (2025). Multimodal fusion of ultrasound images using HXM net for breast cancer diagnosis. *Scientific Reports*, 15(1), 40689.
33. Pradeep Ghantasala, G. S., Reddy, A., Sharma, R. R., Vidyullatha, P., Bali, M. S., & Sungheetha, A. (2025, September). Enhanced Ovarian Cancer Classification through Hyperparameter-Optimized Machine Learning Models. In *International Conference on Artificial Intelligence and Computing* (pp. 59-70). Singapore: Springer Nature Singapore.
34. Ghantasala, G. P., Ananthanagu, U., Sharma, R., & Sungheetha, A. (2025, September). Integrating Machine Learning and Cox Models for Enhanced Survival Prediction in Ovarian Cancer. In *2025 3rd International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)* (pp. 2070-2073). IEEE.