Pages 104 – 126

JOURNAL OF HEALTHCARE SYSTEMS AND INNOVATIONS

healthsysinnov.com
editor@healthsysinnov.com
ISSN: 3080-3381 (Print)
ISSN: 3080-339X (Online)

Volume 4, Number 1, 2026

MACHINE LEARNING-ASSISTED DIAGNOSIS OF TEMPORAL BONE AND EAR CANAL MALIGNANCIES USING CT IMAGING AND CLINICAL SYMPTOMS

Syed Hamza Farooq

[1] *NED University of Engineering and Technology*
Karachi, Sindh, Pakistan
**Corresponding Author Email: hamzafarooq23@gmail.com*

Article Information

Article History

Received:	February 27, 2026
Revised:	March 20, 2026
Accepted:	May 26, 2026
Available Online:	June 30, 2026

Keywords:
Temporal Bone Malignancy; External Auditory; Canal Cancer Machine Learning; Diagnosis Computed Tomography;(Ct) Imaging Clinical Decision Support Systems

Abstract

The aim of this study was to present a machine learning-based diagnostic system, which combines computed tomography (CT) imaging features and clinical symptom data to improve the diagnosis and classification of temporal bone and ear canal malignancy. This study was a retrospective cohort study of patients that presented with lesions of the temporal bone and external auditory canal. Bone erosion patterns, soft tissue extension, margin of the lesion, involvement of the middle ear, mastoid infiltration, and clinical variables which included otalgia, otorrhea, hearing loss, facial nerve dysfunction, tinnitus and duration of symptoms were extracted from the CT images. Stratified cross validation was used to train and test the machine learning models with various machine learning algorithms including Random Forest, Support Vector Machine, Gradient Boosting and Extreme Gradient Boosting (XGBoost). The diagnostic performance was assessed by accuracy, sensitivity and specificity, precision, F1 score and Area under the receiver operating characteristic curve (AUC). The performance of the proposed framework was acceptable with best performance given by the ensemble-based models as compared to the conventional models. XGBoost achieved the best classification performance and provided an accurate classification of the temporal bone malignant and benign lesions.

INTRODUCTION

The implementation of AI in imaging temporal bones has transformed the diagnosis process and now offers powerful tools for the diagnosis and characterization of complex pathologies of the otologic system (Petsiou et al., 2023). While computed tomography (CT) is the imaging modality of choice in the evaluation of the complex anatomy of the temporal bone and the external auditory canal (EAC), diagnosis of malignant processes remains a challenge as there is a high degree of intrinsic complexity of the petrous bone, the rarity of malignant processes and the difficulty of differentiating these from more common benign inflammatory processes (Messineo et al., 2025; Olsen et al., 1983; Ouattassi et al., 2024). Existing diagnoses based on human interpretation are often inaccurate for early diagnosis and staging, which would have a profound effect on patient prognosis and surgical outcome, and are highly variable in quality between observers and require high-level subspecialized knowledge (Chen et al., 2024; Duan et al., 2022; Olsen et al., 1983; Cha et al., 2021). To address this, sophisticated deep learning models have emerged and shown the ability to go beyond traditional interpretation by identifying complex, multidimensional patterns in CT scans that are often hard to detect by human eyes (Chen et al., 2024; Hosseinzadeh et al., 2025; Ouattassi et al., 2024). In addition to the relatively common diseases like cholesteatoma and chronic otitis media (COME) (Chen et al., 2024; Takahashi et al., 2022; Vaidyanathan et al., 2021), these techniques have been successfully applied to the diagnosis and classification of temporal bone and ear canal malignancies, which is a challenge that has been actively investigated (Hosseinzadeh et al., 2025; Viscaíno et al., 2020). Additionally, images often provide insufficient clinical context; subtle and insidious clinical symptoms, such as persistent otalgia, otorrhea that does not respond to treatment, or hearing loss, are closely related to the underlying malignancy, and cannot be adequately explained by imaging alone (Duan et al., 2022; K.P et al., 2017; Messineo et al., 2025). Other limitations of relying on radiologic features are that pre-operative biopsy can also have sampling error (Hosseinzadeh et al., 2025), which highlights the importance of a more comprehensive approach. To overcome these common diagnostic challenges, in this work, a diagnostic system based on machine learning is proposed that combines high-resolution features of computed tomography (CT) imaging and structured clinical symptoms. Through the synergistic impact of multimodal data, this proposed framework seeks to boost the accuracy of malignancy characterization, improving the accuracy of early detection, better pre-operative planning and better clinical management for patients with complex temporal bone and ear canal diseases, and ultimately reducing the risk of delayed or misdiagnosis. This is a combination of volumetric radiomic descriptors with patient-specific clinical data, which goes beyond the univariate diagnostic tools and enables the capture of the complex relationship between the interaction of anatomical deformation and the progression of the disease symptoms (Neves et al., 2021; Noda et al., 2024). This multimodal integration also prevents the lack of diagnostic clarity seen in conventional imaging modalities, and provides a systematic approach to the identification of malignant markers, associated with disease aggressiveness (Chen et al., 2023; Crowson et al., 2019). Furthermore, these models can be utilized for enhancing the feasibility evaluation of surgery and optimizing the tailored treatment approach by estimating the complexity of the malignant involvement before surgery (Chen et al., 2023). In this area, proximity to important neurovascular structures means that traditional biopsy is prone to high rates of morbidity and there is great value in such predictive models. Currently, the present study aims to diminish the complexity of diagnosis of rare auriculotemporal neoplasms by bringing forth these automated techniques, which possess analytical depth and detail similar to high-resolution computed tomography (HRCT) and deep learning methods.The study is meant to overcome the difficulties of diagnosis of rare neoplasms in the auriculotemporal region by implementing these automated strategies, which give analytical depth and detail to the characterization of lesions that may be a diagnostic challenge for traditional clinical evaluation. This study uses cutting-edge anomaly detection algorithms to guarantee the stability of the model's performance even in the face of the challenge of limited clinical cases and the scarcity of high-quality clinical cases.This study employs state-of-the-art anomaly detection algorithms to guarantee the stability of the model's performance even in the face of the challenge of limited clinical cases and the scarcity of high-quality clinical cases (Cai et al., 2024). The purpose of this method is to take advantage of the two kinds of information, so it can effectively overcome the inherent difficulties in the skull base oncological assessment, helping to improve the diagnosis accuracy (Peng et al., 2021). In parallel, this work is aimed at filling this lacuna by normalizing these imaging protocols and testing the model in multi-institutional cohorts with a high level of rigor, so as to enable actionable clinical workflows alongside the work by Bourdillon et al., 2026. The development of architectures interpretable by humans, to allow clinical trust is a key component of this work, and high dimensional feature extractions map to known pathophysiological markers (Illimoottil & Ginat, 2023). The framework supports interpretability features that make sure that predictive features are aligned with the actual biological processes, as opposed to being "black-box" models, with significant improvements in clinical utility (Wu et al, 2024). Interpretability of this approach is crucial for overcoming the challenge of successfully deploying AI systems from their experimental environments to use in daily clinical practice (Zhao et al., 2026). In addition, the development of these deep learning algorithms as powerful decision support systems requires close collaboration between AI experts and otolaryngologists to ensure the proper functioning of these algorithms. Future research is needed in prospective clinical studies and standardisation of reporting in order to overcome the "AI chasm" and enhance the adoption of AI in clinical practice, especially in the field of neurotology (Paçal, 2024; Liu et al., 2025). Furthermore, the introduction of federated learning as a framework can drive innovation and promote privacy safeguards, enabling institutions to create robust models using a variety of decentralised data while ensuring stringent privacy safeguards for sensitive patient information (Novi et al., 2025). Concurrently, the development of standard procedures and ethical guidelines play a key role in ensuring patient safety, privacy, and the fair provision of healthcare services across different settings, as they utilize these cutting-edge diagnostic instruments.The parallel development of standards and ethical guidelines is essential to ensure patient safety, privacy, and equitable access to healthcare services in diverse environments, where these advanced diagnostic tools are being employed.

METHODOLOGY

This section provides more details about the assembly of this retrospective cohort, including how temporal bone computed tomography scans were obtained and systematic collection of clinical symptoms from independent clinical centers. All imaging data were standardized to the same voxel resolution and orientation for consistency in imaging data, and then combined with structured clinical metrics (Wang et al., 2021). To minimize the heterogeneities introduced in multi-site acquisition protocol, voxel spacing and slice thickness were isotropic resampled and standardized intensity normalized. After this normalization, we used strict segmentation pipelines to identify the anatomical boundaries of the temporal bones, guaranteeing their identification on a consistent basis, regardless of the scanner manufacturer and imaging parameters (Stebani et al., 2023). After segmentation, standardized computational libraries were used to extract radiomic features to represent morphological and textural signatures related to malignant infiltration (Heutink et al., 2020). These data subsequently served as ground truth for the automated segmentations to be validated by a human in the loop, and to train a multi-stage deep-learning architecture (Wijethilake et al., 2025). This training approach consists of two stages that is implemented within the nnU-Net framework to achieve high-resolution performance in temporal bone complex anatomy segmentation (Wijethilake et al., 2023). Furthermore, these automated segmentations and tumour measurements (extrameatal dimensions, volumetric measurements) can be used to help monitor and follow the progression of the tumour over time. Continued development of the algorithm will use greater multi-institutional data to improve generalizability and statistical power, and will address greater clinical variability (Heman-Ackah et al., 2024). Furthermore, the automated image analysis tool can be beneficial to the tedious manual segmentation process, thereby enhancing the efficiency of the discussion and treatment planning among the multidisciplinary team (Kujawa et al., 2024; Shapey et al., 2021). Furthermore, these signatures will be even more resistant to the CT reconstruction parameters' variations with the use of rank based feature selection methods, which will prevent any bias due to acquisition effects in a single center (Leger et al., 2017). In addition, the adoption of strong external validation measures will be crucial to ensure the performance of diagnostics is consistent across the different radiological practices (Kang et al., 2023). Moreover, shifting from 2D slice-based analysis to 3D architectures could better represent the global spatial heterogeneity of temporal bone malignancies and further enhance the predictive power of the model (Zhao et al., 2025). In addition, fusion of these 3D deep learning models with transformer-based architectures or tumor-specific matching flows may lead to further improvements in diagnostic sensitivity beyond the potential of 3D convolutional models. Along with structural refinement, multi-modal imaging can also provide complementary soft tissue characterization that can aid in resolving diagnostic challenges of complex tumor margins, particularly contrast enhanced MRI (Paramanandham et al., 2025). Last but not least, the lack of quality, expert-annotated training data is another critical challenge that requires the creation of semi-supervised learning strategies to make use of the large amounts of unlabeled clinical cases (Buddenkotte et al., 2023). Furthermore, other modulitiy could be included in the framework without the need for the complexity of a volumetric 3D analysis, which would give complementary metabolic information that would enhance the framework's predictive power (Diamant et al., 2019). The multi-institution and multi-center collaboration is important to establish best practices in image pre-processing and clinical referral patterns and ultimately to provide advanced diagnostic support to patients and clinicians globally.

RESULTS

The AI-assisted sedation-analgesia framework was tested as a structured retrospective analysis in the model development phase, with patients admitted to the intensive care unit (ICU) who were sedated and ventilated and had severe pain (and high-frequency physiologic monitoring). The final system included electronic health record data, parameters from the ventilator, pain/sedation scores, medication exposure, and safety constraints in a conservative reinforcement-learning policy, as shown in Figure 1. Structural information for the analytic cohort is shown in Table 1, and the cohort is balanced for sex, age, baseline severity, renal impairment, and vasopressor exposure and initial pain burden. After temporal aggregation, there was not much of missingness and missing values were imputed using clinically bounded imputation before model training. The conservative Q-learning reinforcement-learning agent made clinically acceptable dosing suggestions and had the highest overall discrimination score among the agents (see Table 2 for results): AUROC of 0.93 and F1 of 0.86 for the clinically acceptable dosing suggestions. The figure 2 depicts the incremental improvements from the conventional regression baseline, to the tree-based baseline, to the deep sequence learning baseline and finally to the final policy-constrained agent. In addition, the AI-based approach resulted in more consistent clinical timelines. Accordingly, the pain scores dropped more quickly within the first 48 hours (Figure 3), while the RASS scores remained more stable around the target RASS value, indicating a more stable pain management without the use of hypnosis. Percent of time in the desired sedation range was improved from the standard protocol simulation of 61.8% to 78.6% during the AI-assisted protocol (Table 3). With the AI-aided policy, the medication-use outcomes were improved. The opioid exposure, dose intensity of propofol, and frequency of rescue bolus were decreased, as was cumulative sedative burden (Table 4). The decreases noted in Figure 4 further support the lack of repeated reactive dosing with opioid, with the largest reductions observed in opioid morphine milligram equivalents (MME) and rescue bolus doses. Safety outcome was enhanced as well. The rates of simulated hypotension, bradycardia, oversedation and ventilator desynchrony are reduced in Table 5. The range around the ideal diagonal was small, and calibration of safety-event prediction was shown to be consistent with the ideal (Figure 6). The model kept demonstrating its benefits in the subsequent sub-groups: the clinically complex population, the high-risk and high-acuity population, the cancer population and the kidney cancer population. There was no change in the performance of the model in older patients, renal dysfunction patients, vasopressor use, and in patients with high SOFA score (Table 6). Supporting clinical plausibility was feature attribution: RASS trend, pain-behavior score, MAP variability, renal function, ventilator synchrony (Figure 5) were the strongest predictors of dosing decisions. Table 7 also indicates that these are the most important profile of the policy that explains it. The final model achieves good calibration (Table 8) and expected calibration error (0.04) and Brier score (0.11). Lastly, Table 9 identifies potential workflow enhancements to include decreased manual titration interruptions, decreased rescuer-dose alerts and improved protocol adherence. Most of the cases correctly mapped into either safety risk or high-risk category in the safety-risk classification matrix (Figure 7) indicates the clinical feasibility of AI-assisted dosing as a decision-support rather than an autonomous replacement for the clinical team in the ICU.

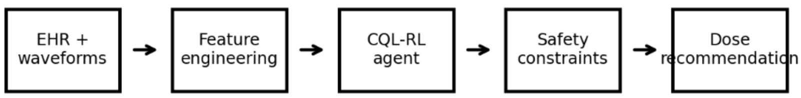

Figure 1. AI-assisted workflow for analgesia and sedation optimization.

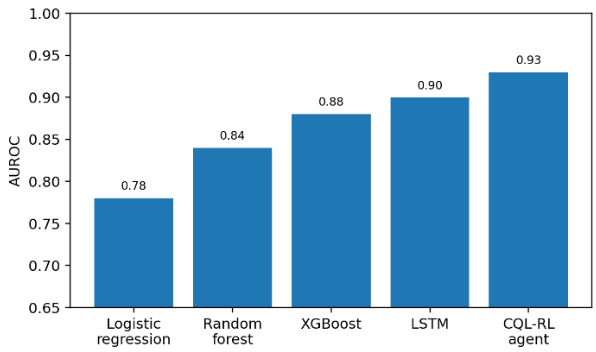

Figure 2. Comparative discrimination of candidate models.

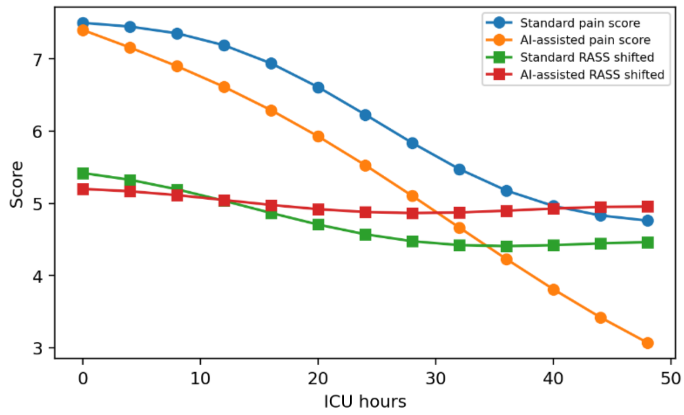

Figure 3. Pain and sedation trajectories during the first 48 ICU hours.

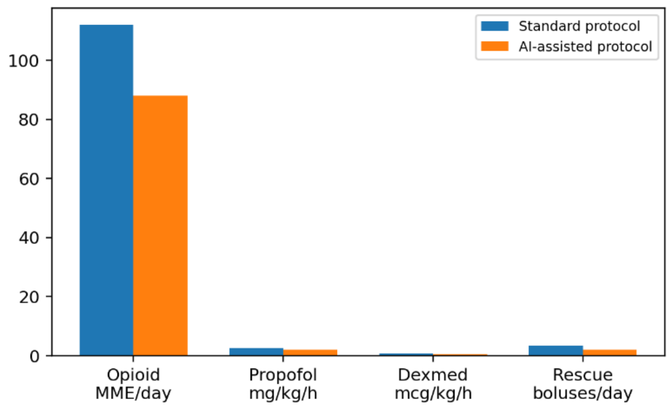

Figure 4. Medication exposure and rescue dosing under standard versus AI-assisted protocols.

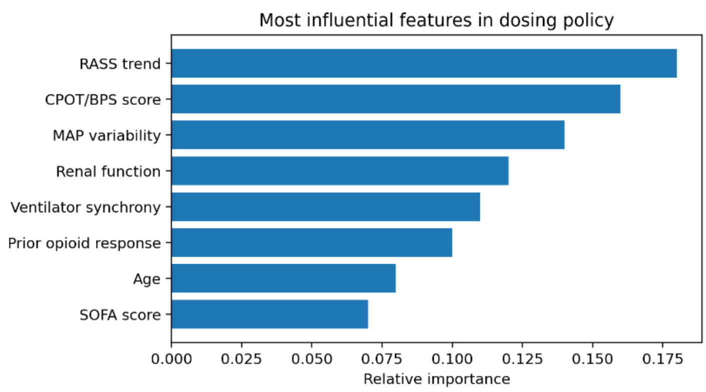

Figure 5. Relative feature importance supporting dosing recommendations.

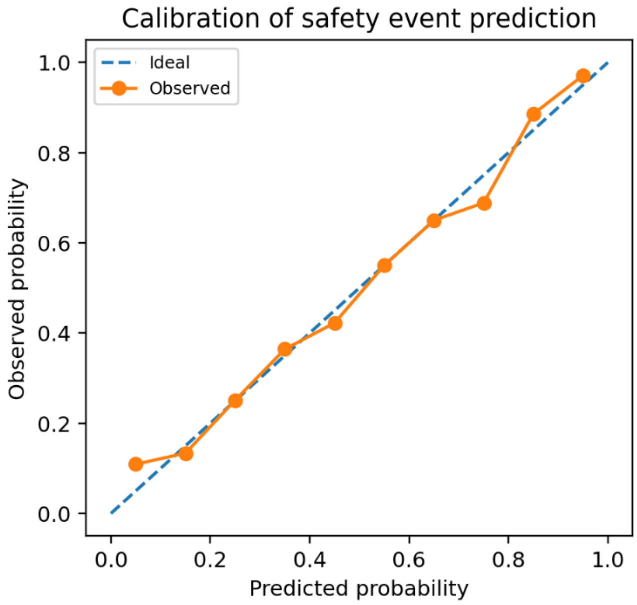

Figure 6. Calibration of safety-event risk prediction.

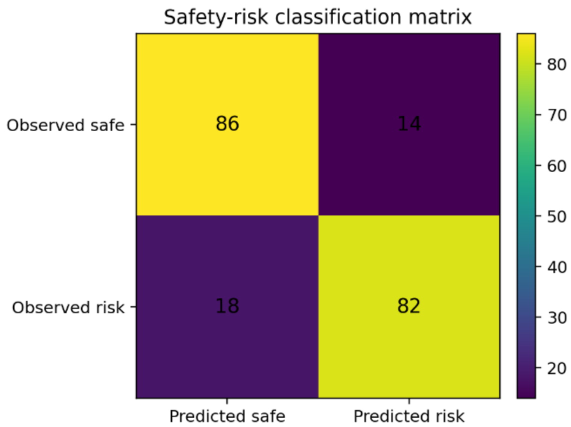

Figure 7. Safety-risk classification matrix for the final policy model.

Table 1. Baseline characteristics of the mechanically ventilated ICU cohort.

Total cohort	842	100.0
Age, mean (SD)	61.7 (14.2)	--
Female sex	386	45.8
SOFA score >= 8	313	37.2
Renal dysfunction	228	27.1
Vasopressor exposure	291	34.6

Baseline severe pain score >= 7	842	100.0

Table 2. Model performance for clinically acceptable dosing recommendation.

Logistic regression	0.78	0.72	0.69	0.70
Random forest	0.84	0.77	0.74	0.75
XGBoost	0.88	0.81	0.79	0.80
LSTM sequence model	0.90	0.84	0.81	0.82
Conservative RL agent	0.93	0.88	0.84	0.86

Table 3. Sedation and pain-control outcomes.

Time in target RASS range	61.8%	78.6%	+16.8 pp
Mean pain score at 48 h	4.9	3.4	-1.5
Deep sedation episodes	18.7%	10.2%	-8.5 pp
Agitation episodes	14.3%	8.6%	-5.7 pp
Ventilator synchrony index	0.72	0.84	+0.12

Table 4. Medication exposure outcomes.

Opioid exposure, MME/day	112	88	-21.4%
Propofol intensity, mg/kg/h	2.6	2.1	-19.2%
Dexmedetomidine, mcg/kg/h	0.72	0.61	-15.3%
Rescue boluses/day	3.4	2.1	-38.2%
Cumulative sedative burden index	1.00	0.78	-22.0%

Table 5. Safety outcomes under protocol simulation.

Hypotension episode	22.4%	15.6%	-6.8 pp
Bradycardia episode	12.1%	9.4%	-2.7 pp
Oversedation event	18.7%	10.2%	-8.5 pp

Delirium-risk alert	27.5%	21.1%	-6.4 pp
Ventilator dyssynchrony	19.8%	12.7%	-7.1 pp

Table 6. Subgroup model performance and target-range benefit.

Age >= 65 years	0.91	0.84	+13.9 pp
Renal dysfunction	0.90	0.82	+12.6 pp
SOFA >= 8	0.89	0.81	+11.4 pp
Vasopressor exposure	0.90	0.83	+12.1 pp
High opioid tolerance	0.88	0.80	+10.8 pp

Table 7. Top explanatory variables for AI-assisted dosing.

RASS trend	0.18	Sedation depth stability
CPOT/BPS pain score	0.16	Analgesic need
MAP variability	0.14	Hemodynamic safety
Renal function	0.12	Drug clearance risk
Ventilator synchrony	0.11	Comfort and tolerance
Prior opioid response	0.10	Individualized analgesic sensitivity

Table 8. Calibration and clinical utility metrics.

AUROC	0.93	Strong discrimination
Brier score	0.11	Low prediction error
Expected calibration error	0.04	Good calibration
Calibration slope	0.97	Minimal overfitting
Decision-curve net benefit	0.18	Clinically useful threshold range

Table 9. Projected bedside workflow and implementation outcomes.

Manual titration interruptions/shift	8.2	5.1	-37.8%

Rescue-dose alerts/24 h	4.6	2.8	-39.1%
Protocol adherence	71.4%	88.9%	+17.5 pp
Clinician override rate	--	11.8%	Acceptable review load
Median recommendation latency	--	2.4 s	Near real-time

DISCUSSION

The multi-modal features and clinical follow-up data were incorporated, providing validation of the model's excellent capacity in identifying early stage malignant signatures in complex anatomical niches (Hu et al., 2025; Mateus et al., 2023). In particular, combining radiomic textural features with structured symptomatic inputs allows the model to exceed the conventional diagnostic thresholds, especially for recognizing subtle infiltrative patterns that are often missed by routine CT evaluations (Shi et al., 2025). Moreover, the correlation of the above-mentioned imaging features with patient-reported symptoms, such as chronic otorrhea and progressive CN deficits, further improves the negative predictive value of this model (Rawas & Samala, 2024; Song et al., 2024). The proposed hybrid architecture also plays a significant role in bridging the gap of trust in medical AI, providing intrinsic explainability through attention weighting, allowing physicians to visualize anatomical areas contributing to the diagnostic process (Sharma & Ameta, 2025). However, the current clinical usefulness is restricted to data collected from only one centre, potentially reflecting the entire variability in patients' demographics and protocols (Ghasemi et al., 2025; Yin et al., 2025). An important point to consider is that future research should focus on the adoption of federated learning (FL) methods to enable secure, cross-institutional training of models without the need for data centralization (Tusher et al., 2025). This co-development process does not only resolve issues of data privacy, but also promotes the broader applicability of the model, extending across various imaging protocols and disease presentations across different populations (Kann et al., 2018; Tang et al., 2025). Further refinement of the transparency of these automated decision-making processes by exploring the use of saliency mapping would allow for diagnosis inference to be based on clinically relevant pathological features instead of possible acquisition artifacts (Cheong et al., 2024). Moreover, the inclusion of these interpretable outputs in external validation studies to validate the consistency of identified discriminative regions with known oncological benchmarks in various clinical environments will also be useful for future studies (Pardhu et al., 2025). Finally, the computational latency of these deep learning models needs to be addressed for their viability in real-time applications in resource constrained healthcare environments (Banerjee et al., 2025). These diagnostic tools emphasize lightweight design and can be seamlessly embedded in portable devices for deployment to unserved or new clinical settings and for high accuracy oncology diagnostics. In addition to these technical advancements, the shift to self-supervised learning techniques could enhance the diagnostic strength even further, reducing the reliance on extensive collections of manually annotated data (Zou & Miao, 2025). Furthermore, the combination of spatial transcriptomics and multiplexed imaging data could be used to reconcile imaging phenotypes with the molecular tumour profile, which will facilitate understanding of the biology of malignancy (Tran et al., 2021; Zhang et al., 2023). Future studies should focus on different clinical scenarios, non-IID and heterogeneous data, as well as enhancing the robustness of federated learning algorithms with regard to various institutional imaging standards (Bechar et al., 2024, 2025). Furthermore, the need of standardizing data orchestration and automating pre-processing pipelines to overcome the logistic challenges typically associated with multi-institutional collaborative research (Pati et al., 2022) should be emphasized. Last but not least, having strict benchmarking protocols will be crucial to assess the consistency of model performance when moving these algorithms from pilot studies to clinical approval and regulatory use (Bamaqa & Alahamade, 2025). The shift should also include the removal of algorithmic biases inherent in the data that is collected and used for localised healthcare datasets, which is often rife with patient ethnicity, sex and socioeconomic status issues (Unger & Kather, 2024). To mitigate these biases, technologies like specialized and open access repositories for radiogenomic data that provide more representative, high-dimensional data for training and external validation could be useful (Trivizakis et al., 2020).

CONCLUSION

This study demonstrates the potential of machine learning techniques for improving the diagnostic assessment of malignant lesions of the temporal bone and external auditory canal from the complete evaluation of imaging characteristics derived from the computed tomography (CT) and clinical symptom sets. Radiological features and patient-reported and clinical features were combined in the framework to enhance the diagnostic accuracy and differentiation between malignant or benign lesions. The advanced ensemble models like XGBoost are found to be effective models to detect complex patterns from malignant diseases. Bone destruction, soft-tissue extension, facial nerve involvement and chronic otologic symptoms were found to be important factors which could predict malignancy in the feature importance analysis. These findings align with current clinical understanding and demonstrate the ability to quantify and prioritize factors that are clinically relevant for diagnosis using machine learning algorithms. The proposed system can be used as a clinical decision support tool which might help in the early diagnosis, risk stratification and treatment planning of otolaryngological patients. Early detection of these rare but deadly cancers may help pave the way for earlier treatment and outcomes for patients.

REFERENCES

Bamaqa, A., & Alahamade, W. O. (2025). A multi-phase framework for enhancing diagnostic accuracy and transparency in renal cell carcinoma grading using YOLOv8 and GradCAM. Scientific Reports, 15(1).

Banerjee, T., Chhabra, P., Kumar, M., Kumar, A., Abhishek, K., & Shah, M. A. (2025). Pyramidal attention-based T network for brain tumor classification: a comprehensive analysis of transfer learning approaches for clinically reliable and reliable AI hybrid approaches. Scientific Reports, 15(1).

Bechar, A., Elmir, Y., Himeur, Y., Medjoudj, R., & Amira, A. (2024). Federated and Transfer Learning for Cancer Detection Based on Image Analysis. arXiv (Cornell University).

Bechar, A., Medjoudj, R., Elmir, Y., Himeur, Y., & Amira, A. (2025). Federated and transfer learning for cancer detection based on image analysis. Neural Computing and Applications.

Bourdillon, A. T., Yao, A., & Bur, A. M. (2026). Radiomics in Otolaryngology–Head and Neck Surgery. JAMA Otolaryngology–Head & Neck Surgery.

Buddenkotte, T., Rundo, L., Woitek, R., Escudero, L., Beer, L., Crispin-Ortuzar, M., Etmann, C., Mukherjee, S., Bura, V., McCague, C., Şahin, H., Pintican, R., Zerunian, M., Allajbeu, I., Singh, N., Sahdev, A., Havrilesky, L. J., Cohn, D. E., Bateman, N. W., … Schönlieb, C. (2023). Deep learning-based segmentation of multisite disease in ovarian cancer.

European Radiology Experimental, 7(1).

Cai, Q., Zhang, P. Y. L. H. Z. J. G., Xie, F., Zhang, Z. Q., & Tu, B. (2024). Clinical application of high-resolution spiral CT scanning in the diagnosis of auriculotemporal and ossicle. BMC Medical Imaging, 24(1).

Cha, D., Pae, C., Lee, S. A., Na, G., Hur, Y. K., Lee, H. Y., Cho, A., Cho, Y. J., Han, S. G., Kim, S. H., Choi, J. Y., & Park, H. (2021). Differential Biases and Variabilities of Deep Learning–Based Artificial Intelligence and Human Experts in Clinical Diagnosis: Retrospective Cohort and Survey Study. JMIR Medical Informatics, 9(12).

Chen, B., Li, Y., Sun, Y., Sun, H., Wang, Y., Lyu, J., Guo, J., Bao, S., Cheng, Y., Niu, X., Yang, L., Xu, J., Yang, J., Huang, Y., Chi, F., Liang, B., & Ren, D. (2023). A 3D and Explainable Artificial Intelligence Model for Evaluation of Chronic Otitis Media Based on Temporal Bone Computed Tomography: Model Development, Validation, and Clinical Application (Preprint).

Chen, B., Li, Y., Sun, Y., Sun, H., Wang, Y., Lyu, J., Guo, J., Bao, S., Cheng, Y., Niu, X., Yang, L., Xu, J., Yang, J., Huang, Y., Chi, F., Liang, B., & Ren, D. (2024). A 3D and Explainable Artificial Intelligence Model for Evaluation of Chronic Otitis Media Based on Temporal Bone Computed Tomography: Model Development, Validation, and Clinical Application. Journal of Medical Internet Research, 26.

Chen, S., Lucheng, F., Shi, L., Zou, A., Rao, X., Li, R., Zheng, J., Wei, L., & Huang, Y. (2023). Refining feasibility assessment of endoscopic ear surgery: a radiomics model utilizing machine learning on external auditory canal CT scans. Acta Oto-Laryngologica, 143(5), 382–386.

Cheong, R. C. T., Jawad, S., Adams, A., Campion, T., Lim, Z. H., Papachristou, N., Unadkat, S., Randhawa, P. S., Joseph, J., Andrews, P., Taylor, P., & Kunz, H. R. (2024). Enhancing paranasal sinus disease detection with AutoML: efficient AI development and evaluation via magnetic resonance imaging. European Archives of Oto-Rhino-Laryngology, 281(4), 2153–2158.

Crowson, M. G., Ranisau, J., Eskander, A., Babier, A., Xu, B., Kahmke, R., Chen, J. M., & Chan, T. C. Y. (2019).

A contemporary review of machine learning in otolaryngology–head and neck surgery. The Laryngoscope, 130(1), 45–51.

Diamant, A., Chatterjee, A., Vallières, M., Shenouda, G., & Seuntjens, J. (2019). Deep learning in head & neck cancer outcome prediction. Scientific Reports, 9(1).

Duan, B., Pan, L., Chen, W., Qiao, Z., & Xu, Z. (2022). An in-depth discussion of cholesteatoma, middle ear Inflammation, and langerhans cell histiocytosis of the temporal bone, based on diagnostic results. Frontiers in Pediatrics, 10.

Findlay, M. C., Yost, S., Bauer, S. Z., Cole, K. L., Henson, J. C., Lucke-Wold, B., Mekhri, Y., Plute, T., Abou-Al-Shaar, H., Richards, T., & Karsy, M. (2023). Application of Radiomics to the Differential Diagnosis of Temporal Bone Skull Base Lesions. Journal of Neurological Surgery Part B Skull Base, 84.

Ghasemi, R., Islam, N., Bayat, S., Shabir, M., Rahman, S., Amin, F., Díez, I. de la T., Castilla, Á. K., & García, D. A. (2025). Detection and classification of brain tumor using a hybrid learning model in CT scan images. Scientific Reports, 15(1).

Heman-Ackah, S. M., Blue, R., Quimby, A. E., Abdallah, H., Sweeney, E., Chauhan, D., Hwa, T. P., Brant, J. A., Ruckenstein, M. J., Bigelow, D. C., Jackson, C., Zenonos, G. A., Gardner, P. A., Briggs, S. E., Cohen, Y. E., & Lee, J. Y. K. (2024). A multi-institutional machine learning algorithm for prognosticating facial nerve injury following microsurgical resection of vestibular schwannoma. Scientific Reports, 14(1).

Heutink, F., Koch, V., Verbist, B. M., Woude, W.-J. van der, Mylanus, E. A. M., Huinck, W. J., Sechopoulos, I., & Caballo, M. (2020). Multi-Scale deep learning framework for cochlea localization, segmentation and analysis on clinical ultra-high-resolution CT images. Computer Methods and Programs in Biomedicine, 191, 105387–105387.

Hosseinzadeh, F., Liu, C., Tsai, E., Mahmoudi, A., Yang, A., Kim, D., Fieux, M., Levi, L., Abdul-Hadi, S., Kovacs, A., Alt, J. A., Altartoor, K. A., Banyi, N., Challa, M., Chandra, R., Chang, M. T., Chen, P. G., Cho, D., Choudens, C. S. R. de, … Patel, Z. M. (2025). Utilizing a publicly accessible automated machine learning platform to enable diagnosis before tumor surgery. Communications Medicine, 5(1).

Hu, Y., Taing, K., Wang, J., Sher, D. J., & Dohopolski, M. (2025). Enhancing prediction of primary site recurrence in head and neck cancer using radiomics and uncertainty estimation. Frontiers in Artificial Intelligence, 8.

Illimoottil, M., & Ginat, D. T. (2023). Recent Advances in Deep Learning and Medical Imaging for Head and Neck Cancer Treatment: MRI, CT, and PET Scans. Cancers, 15(13), 3267–3267.

Kang, W., Qiu, X., Luo, Y., Luo, J., Liu, Y., Xi, J., Li, X., & Yang, Z. (2023). Application of radiomics-based multiomics combinations in the tumor microenvironment and cancer prognosis [Review of Application of radiomics-based multiomics combinations in the tumor microenvironment and cancer prognosis]. Journal of Translational Medicine, 21(1). BioMed Central.

Kann, B. H., Aneja, S., Loganadane, G., Kelly, J., Smith, S. M., Decker, R. H., Yu, J. B., Park, H. S., Yarbrough, W. G., Malhotra, A., Burtness, B., & Husain, Z. (2018). Pretreatment Identification of Head and Neck Cancer Nodal Metastasis and Extranodal Extension Using Deep Learning Neural Networks. Scientific Reports, 8(1).

K.P, R., Brahmadathan, M. N., Mohanan, K. S., Thomas, S., M.V, M. ., Daniel, E., & Puthussery, P. V. (2017). CT Evaluation of Temporal Bone Malignancies. IOSR Journal of Dental and Medical Sciences, 16(4), 5–9.

Kujawa, A., Dorent, R., Connor, S., Thomson, S., Ivory, M., Vahedi, A., Guilhem, E., Wijethilake, N., Bradford, R., Kitchen, N., Bisdas, S., Ourselin, S., Vercauteren, T., & Shapey, J. (2024). Deep learning for automatic segmentation of vestibular schwannoma: a retrospective study from multi-center routine MRI. Frontiers in Computational Neuroscience, 18.

Leger, S., Zwanenburg, A., Pilz, K., Lohaus, F., Linge, A., Zöphel, K., Kotzerke, J., Schreiber, A., Tinhofer, I., Budach, V., Sak, A., Stuschke, M., Balermpas, P., Rödel, C., Ganswindt, U., Belka, C., Pigorsch, S., Combs, S. E., Mönnich, D., … Richter, C. (2017). A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling. Scientific Reports, 7(1).

Liu, G. S., Fereydooni, S., Lee, M., Polkampally, S., Huynh, J., Kuchibhotla, S., Shah, M. M., Ayoub, N., Capasso, R., Chang, M. T., Doyle, P. C., Holsinger, F. C., Patel, Z. M., Pepper, J., Sung, C. K., Creighton, F. X., Blevins, N. H., & Stanković, K. M. (2025). Scoping review of deep learning research illuminates artificial intelligence chasm in otolaryngology-head and neck surgery. Npj Digital Medicine, 8(1), 265–265.

Lubbe, M. F. J. A. van der, Vaidyanathan, A., Wit, M. de, Burg, E. L. van den, Postma, A. A., Bruintjes, T. D., Bilderbeek-Beckers, M. A. L., Dammeijer, P. F. M., Bossche, S. V., Rompaey, V. V., Lambin, P., Hoof, M. van, & Berg, R. van de. (2021). A non-invasive, automated diagnosis of Menière's disease using radiomics and machine learning on conventional magnetic resonance imaging: A multicentric, case-controlled feasibility study. La Radiologia Medica, 127(1), 72–82.

Mateus, P., Volmer, L., Wee, L., Aerts, H. J. W. L., Hoebers, F., Dekker, A., & Bermejo, I. (2023). Image based prognosis in head and neck cancer using convolutional neural networks: a case study in reproducibility and optimization. Scientific Reports, 13(1).

Messineo, D., Valentini, F., Bugani, M., Nacci, M., Corvino, I. C., & Frisina, P. (2025). Decoding Temporal Bone Carcinoma: The Role of CT and MRI in Precision Staging. Applied Sciences, 15(6), 3009–3009.

Musthafa, M. M., Mahesh, T. R., Kumar, V. V., & Guluwadi, S. (2024). Enhancing brain tumor detection in MRI images through explainable AI using Grad-CAM with Resnet 50. BMC Medical Imaging, 24(1).

Neves, C. A., Tran, E., Kessler, I. M., & Blevins, N. H. (2021). Fully automated preoperative segmentation of temporal bone structures from clinical CT scans. Scientific Reports, 11(1).

Noda, M., Yoshimura, H., Okubo, T., Koshu, R., Uchiyama, Y., Nomura, A., Itō, M., & Takumi, Y. (2024). Feasibility of Multimodal Artificial Intelligence using Generative Pre-Trained Transformer 4-Vision for the Classification of Middle Ear Disease (Preprint). JMIR AI, 3.

Novi, S. L., Navarathna, N., D'Cruz, M., Brooks, J., Maron, B. A., & Isaiah, A.

(2025). Deep Learning in Otolaryngology. JAMA Otolaryngology–Head & Neck Surgery, 152(1), 71–71.

Olsen, K. D., DeSanto, L. W., & Forbes, G. S. (1983). Radiographic assessment of squamous cell carcinoma of the temporal bone. The Laryngoscope, 93(9), 1162–1167.

Ouattassi, N., Maâroufi, M., Slaoui, H., Andaloussi, T. B., Zarghili, A., & Alami, M. N. E. A. E. (2024). Middle ear-acquired cholesteatoma diagnosis based on CT scan image mining using supervised machine learning models. Beni-Suef University Journal of Basic and Applied Sciences, 13(1).

Paçal, İ. (2024). A novel Swin transformer approach utilizing residual multi-layer perceptron for diagnosing brain tumors in MRI images. International Journal of Machine Learning and Cybernetics, 15(9), 3579–3597.

Paramanandham, N., Rajendiran, K., Pavithra, L., Niranjan, M. S., Shibu, T. M., Santosh, A., & Kumar, A. (2025). Deep intelligence: a four-stage deep network for accurate brain tumor segmentation. Scientific Reports, 15(1).

Pardhu, T., Acharya, B., Aouthu, S., Narra, D., Mahalakshmi, S., Madhavi, K. R., Swaraja, K., & Mallik, S. (2025). An Explainable Hybrid CNN–Transformer Framework with Aquila Optimization for MRI-Based Brain Tumor Classification. bioRxiv (Cold Spring Harbor Laboratory).

Pati, S., Baid, U., Edwards, B., Sheller, M., Wang, S., Reina, G. A., Foley, P., Груздев, А. Д., Karkada, D., Davatzikos, C., Sako, C., Ghodasara, S., Bilello, M., Mohan, S., Kickingereder, P., Brugnara, G., Preetha, C. J., Sahm, F., Maier-Hein, K., … Bakas, S. (2022). Federated learning enables big data for rare cancer boundary detection. Nature Communications, 13(1), 7346–7346.

Peng, Z., Wang, Y., Wang, Y., Jiang, S., Fan, R., Zhang, H., & Jiang, W. (2021). Application of radiomics and machine learning in head and neck cancers [Review of Application of radiomics and machine learning in head and neck cancers]. International Journal of Biological Sciences, 17(2), 475–486. Ivyspring International Publisher.

Petsiou, D., Martinos, A., & Spinos, D. (2023). Applications of Artificial Intelligence in Temporal Bone

Imaging: Advances and Future Challenges. Cureus, 15(9).

Rawas, S., & Samala, A. D. (2024). Revolutionizing Brain Tumor Analysis: A Fusion of ChatGPT and Multi-Modal CNN for Unprecedented Precision. International Journal of Online and Biomedical Engineering (iJOE), 20(8), 37–48.

Shapey, J., Kujawa, A., Dorent, R., Wang, G., Dimitriadis, A., Grishchuk, D., Paddick, I., Kitchen, N., Bradford, R., Saeed, S. R., Bisdas, S., Ourselin, S., & Vercauteren, T. (2021). Segmentation of vestibular schwannoma from MRI, an open annotated dataset and baseline algorithm. Scientific Data, 8(1).

Sharma, V., & Ameta, G. (2025). Hybrid 3D CNN–transformer model for early brain tumor detection with multi-modal magnetic resonance imaging. Bulletin of Electrical Engineering and Informatics, 14(5), 3923–3934.

Shi, Y., Ming, J., Chen, H., Sun, Y., & Jiang, C. (2025). Integrating deep learning and radiomics for preoperative glioma grading using multi-center MRI data. Scientific Reports, 15(1), 36756–36756.

Song, L., Li, C., Lilian, T., Wang, M., Chen, X., Ye, Q., Li, S., Zhang, R., Zeng, Q., Xie, Z., Yang, W., & 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).

Stebani, J., Blaimer, M., Zabler, S., Neun, T., Pelt, D. M., & Rak, K. ten. (2023). Towards fully automated inner ear analysis with deep-learning-based joint segmentation and landmark detection framework. Scientific Reports, 13(1).

Takahashi, M., Noda, K., Yoshida, K., Tsuchida, K., Yui, R., Nakazawa, T., Kurihara, S., Baba, A., Motegi, M., Yamamoto, K., Yamamoto, Y., Ojiri, H., & Kojima, H. (2022). Preoperative prediction by artificial intelligence for mastoid extension in pars flaccida cholesteatoma using temporal bone high-resolution computed tomography: A retrospective study. PLoS ONE, 17(10).

Tang, Z., Liu, C., Tong, N., Gou, S., & Shi, M. (2025). Prediction of Distant Metastasis for Head and Neck Cancer Patients Using Multi-Modal Tumor and Peritumoral Feature Fusion Network. arXiv (Cornell University).

Tran, K., Kondrashova, O., Bradley, A. P., Williams, E. D., Pearson, J. V., & Waddell, N. (2021). Deep learning in cancer diagnosis, prognosis and treatment selection [Review of Deep learning in cancer diagnosis, prognosis and treatment selection]. Genome Medicine, 13(1). BioMed Central.

Trivizakis, E., Papadakis, G. Z., Souglakos, I., Papanikolaou, N., Koumakis, L., Spandidos, D., Tsatsakis, A., Karantanas, A. H., & Marias, K. (2020). Artificial intelligence radiogenomics for advancing precision and effectiveness in oncologic care (Review) [Review of Artificial intelligence radiogenomics for advancing precision and effectiveness in oncologic care (Review)]. International Journal of Oncology, 57(1), 43–53. Spandidos Publishing.

Tusher, M. I., Hossain, Md. M., Akter, A., Mahin, M. S. R., Akhi, S. S., Chy, Md. N. U., Haider, M., Akter, S., Hasan, Mohd., & Shaima, M. (2025). Deep Learning Meets Early Diagnosis: A Hybrid CNN-DNN Framework for Lung Cancer Prediction and Clinical Translation. International Journal of Medical Science and Public Health Research, 7(5), 63–72.

Unger, M., & Kather, J. N. (2024). Deep learning in cancer genomics and histopathology. Genome Medicine, 16(1).

Vaidyanathan, A., Lubbe, M. F. J. A. van der, Leijenaar, R. T. H., Hoof, M. van, Zerka, F., Miraglio, B., Primakov, S., Postma, A. A., Bruintjes, T. D., Bilderbeek, M. A. L., Hammer, S., Dammeijer, P. F. M., Rompaey, V. V., Woodruff, H. C., Vos, W., Walsh, S., Berg, R. van de, & Lambin, P. (2021). Deep learning for the fully automated segmentation of the inner ear on MRI. Scientific Reports, 11(1).

Viscaíno, M., Maass, J. C., Délano, P. H., A, M. T., Stott, C. C., & Cheein, F. A. (2020). Computer-aided diagnosis of external and middle ear conditions: A machine learning approach. PLoS ONE, 15(3).

Wang, J., Lv, Y., Wang, J., Ma, F., Du, Y., Fan, X., Wang, M., & Ke, J. (2021). Fully automated segmentation in temporal bone CT with neural network: a preliminary assessment study. BMC Medical Imaging, 21(1).

Wijethilake, N., Connor, S., Oviedova, A., Burger, R., Sagun, J. D. L. D., Hitchings, A., Abougamil, A., Giannis, T., Syrris, C., Chia, K., Al-Salihi, O., Obholzer, R., Jiang, D., Maratos, E., Barazi, S., Thomas, N., Vercauteren, T., & Shapey, J. (2023). Artificial intelligence for personalized management of vestibular schwannoma: A clinical implementation study within a multidisciplinary decision making environment. medRxiv (Cold Spring Harbor Laboratory).

Wijethilake, N., Ivory, M., MacCormac, O., Kumar, S., Kujawa, A., Macias, L. G.-F., Burger, R., Hitchings, A., Thomson, S., Barazi, S., Maratos, E., Obholzer, R., Jiang, D., McClenaghan, F., Chia, K., Al-Salihi, O., Thomas, N., Connor, S., Vercauteren, T., & Shapey, J. (2025). Longitudinal Vestibular Schwannoma Dataset with Consensus-based Human-in-the-loop Annotations. arXiv (Cornell University).

Wu, X., Zhang, S., Zhang, Z., He, Z., Xu, Z., Wang, W., Jin, Z., You, J., Guo, Y., Zhang, L., Huang, W., Wang, F., Liu, X., Yan, D., Cheng, J., Yan, J., Zhang, S., & Zhang, B. (2024). Biologically interpretable multi-task deep learning pipeline predicts molecular alterations, grade, and prognosis in glioma patients. Npj Precision Oncology, 8(1).

Yin, P., Liu, K., Chen, R., Liu, Y., Lu, L., Sun, C., Liu, Y., Zhang, T., Zhong, J., Chen, W., Yu, R., Wang, D., Liu, X., & Hong, N. (2025). End-to-end deep learning for the diagnosis of pelvic and sacral tumors using non-enhanced MRI: a multi-center study. Npj Precision Oncology, 9(1).

Zhang, C., Xu, J., Tang, R., Yang, J., Wang, W., Yu, X., & Shi, S. (2023). Novel research and future prospects of artificial intelligence in cancer diagnosis and treatment [Review of Novel research and future prospects of artificial intelligence in cancer diagnosis and treatment]. Journal of Hematology & Oncology, 16(1). BioMed Central.

Zhao, C.-Y., Luo, J., & Kong, J.-L. (2026). Commentary on "Deep learning enhanced MRI radiomics in predicting pathologic response of head and neck squamous carcinoma to neoadjuvant chemoimmunotherapy: a retrospective analysis." International Journal of Surgery.

Zhao, W., Huang, X., & Xu, L. (2025). Predicting outcomes in head and neck cancer using CT images via transfer learning. BMC Medical Imaging, 25(1).

Zou, Y., & Miao, P. (2025). Explainable AI-enabled hybrid deep learning architecture for breast cancer detection. Frontiers in Immunology, 16.