European Journal of Therapeutics
pISSN: 2564-7784
eISSN: 2564-7040

Eur J Ther. 2025;31(1):44-50.
https://doi.org/10.58600/eurjther2612

Original Research

Magnetic Resonance Imaging-Based Differential Diagnosis of Parotid Gland Tumors Using Deep Learning

Esat Kaba[1], Kubilay Muhammed Sunnetci[2,3], Ahmet Alkan[3], Fatma Beyazal Celiker[1]

[1] Department of Radiology, Recep Tayyip Erdogan University, Rize,Türkiye
[2] Department of Electrical and Electronics Engineering, Osmaniye Korkut Ata University, Osmaniye,Türkiye
[3] Department of Electrical and Electronics Engineering, Kahramanmaraş Sütçü İmam University, Kahramanmaraş,Türkiye

Received: 2025-01-27 **Accepted:** 2025-02-22 **Published Online:** 2025-02-28

Corresponding Author
Esat Kaba, MD
Address: Department of Radiology, Recep Tayyip Erdogan University, 53100, Türkiye
E-mail: esatkaba04@gmail.com

ABSTRACT

Objectives: It is aimed at magnetic resonance imaging (MRI)-based differential diagnosis of parotid gland tumors (PGTs) using deep learning.

Methods: This study included 117 PGTs obtained from 113 patients. T2-w, T1-w, contrast-enhanced T1-w, Diffusion Weighted Imaging-b0, Diffusion Weighted Imaging-b2000 (DWI-2000), and apparent diffusion coefficient sequences of these patients were used in the study. We implemented four different classification models, and we categorized the images as benign-malignant, pleomorphic adenoma (PA)-Warthin, Warthin-malignant, and all classes (mucoepidermoid carcinoma-other benign-other malignant-PA-Warthin). We constructed classification for each sequence separately using the ResNet18 architecture, with the dataset split into 80% for training and 20% for validation.

Results: The most successful model in this study, achieving an accuracy of 95.37% and an F1-score of 94.74% in classifying malignant-Warthin images in T1-w sequences, also demonstrated the highest accuracy among all models evaluated. For the classification of benign-malignant and the differentiation across all classes, the highest accuracies were achieved with the T2-w sequence at 93.75% and 86.67%, respectively. In the differentiation of PA-Warthin, T1-w and DWI-b0 sequences demonstrated the highest performance, both with an accuracy of 90.36%.

Conclusion: The deep networks proposed in the study supported MRI-based differential diagnosis of PGTs with high accuracy, and the user-friendly software classified images with high accuracy in about 10 seconds.

Keywords: classification, deep learning, magnetic resonance imaging, parotid gland tumors

INTRODUCTION

Parotid gland tumors (PGTs) are the most common salivary gland tumors, constituting 2-6% of head and neck malignancies. The most common benign tumors are pleomorphic adenoma (PA) and Warthin tumors, while mucoepidermoid carcinoma (MC) is the most common malignant type [1,2]. Most tumors are in the superficial lobe of the parotid gland. Treatment varies by subtype; enucleation or partial parotidectomy is used for benign tumors, while total parotidectomy or aggressive surgery is needed for malignant tumors [3]. Preoperative differential diagnosis is crucial for determining the treatment algorithm. Fine needle aspiration biopsy aids in diagnosis but has limitations, including insufficient material, technical failure in deep lobe lesions, local infection, and tumor seeding [4].

On the other hand, imaging modalities are frequently used for preoperative differential diagnosis and the detection of tumor locations. Although ultrasonography and computed tomography (CT) are also used for this purpose, magnetic resonance imaging (MRI) is the most sensitive imaging modality due to its high soft tissue resolution [5,6]. In addition to determining the nature of the tumor, MRI successfully reveals facial nerve involvement and its relationship with adjacent structures [7, 8]. Conventional MRI sequences like T1-w and T2-w evaluate tumor size and its relationship with adjacent structures. Diffusion-weighted imaging (DWI) and apparent diffusion coefficient (ADC) values help distinguish malignancy, though exceptions like Warthin tumors exist due to their unique tissue composition. Contrast-enhanced sequences aid in detecting perineural spread [6-9].

Recently, methods such as dynamic contrast-enhanced MRI, Susceptibility-Weighted Imaging, and MR spectroscopy have also been used in differential diagnosis [10]. However, despite all these sequences, the diagnostic challenge of MRI remains [11,12]. For this reason, decision support systems have been created with artificial intelligence (AI) techniques developed in recent years. In this context, we aimed to perform an MRI-based classification of different PGT subtypes using deep learning (DL) models.

Main Points

- The differential diagnosis of parotid gland tumors is crucial for determining treatment options and this remains a challenge.
- The model proposed in this study differentiates parotid gland tumors with an accuracy rate of 95.37%.
- Deep networks can successfully and rapidly classify parotid gland tumors and improve the efficiency of MRI.

MATERIAL AND METHODS

This retrospective study was approved by the institutional ethics committee (2021/175). Informed consent was obtained from all individual participants included in the study.

Patient Selection

Patients who underwent surgery for PGTs in the hospital between 2017 and 2023 and had pre-operative contrast-enhanced neck MRIs were included in the study. Our exclusion criteria were patients whose images could not be accessed, patients with poor image quality, and patients with histopathologic results of non-parotid pathologies. A total of 148 patients were initially analyzed and 117 PGTs from 113 patients were included in the study following exclusion criteria. Forty-two patients (43 tumors) were PA, forty-three patients (45 tumors) were Warthin tumors, nine patients (10 tumors) were other benign tumors (hemangioma, basal cell adenoma, lymph node), nine patients were MC, and ten patients were other malignant tumors (adenoid cystic carcinoma, polymorphous adenocarcinoma, metastasis).

Dataset

Non-contrast T1 (T1-w), T2 (T2-w), DWI-b0 (DWI-0), DWI-b2000 (DWI-2000), ADC, and contrast-enhanced T1 (T1c-w) sequences of all patients were included in this study. All images were reviewed and axial slices showing the maximum tumor dimension were selected and saved for further analysis. A radiology resident with 5 years of experience and a head and neck radiologist with 13 years of experience jointly evaluated all slices. All images were exported as Digital Imaging and Communications in Medicine (DICOM) files. These images were cropped into PNG format with the tumor in the center and the parotid gland around it and used as the final data set for the study. To understand the dataset used in the study, we provided different sequence images obtained from a patient in Figure 1. The sizes of these images were usually 224*224*3 after cropping.

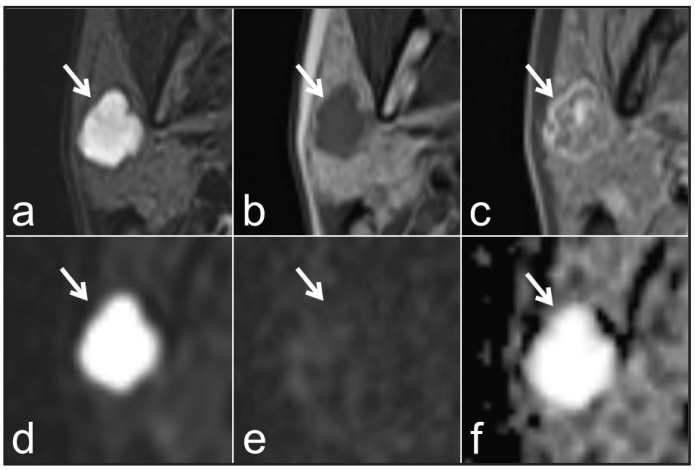

Figure 1. Sample images of the six MRI sequences from the same patient a) T2-w b) T1-w c) T1c-w d) DWI-0 e) DWI-2000 f) ADC

Methods

We prepared the dataset using T2-w, T1-w, T1c-w, DWI-0, DWI-2000, and ADC sequences. In this dataset (117 PGTs and 6 different sequences), we aimed to differentiate between benign-malignant, PA-Warthin, Warthin-malignant, and all groups. We trained the models separately for each sequence and we aimed to develop a total of 24 models for the differential diagnosis of benign-malignant, PA-Warthin, Warthin-malignant, and all classes for six different sequences. For this purpose, we set the input size of these images to be equal to 224*224*3, and here they were 24-bit images. We determined the train and validation set ratios of each of the 24 DL models trained as 0.8 and 0.2, respectively.

Preprocessing, Classification Models, and Training Parameters

We augmented related images for the dataset with the help of the rotate command, and we rotated them using different degree angles [13]. ResNet18 is used in this study [14]. This architecture is a convolutional neural network with 18 layers deep. This model, which can classify up to 1000 object categories, has an image input size of 224*224*3. The number of layers and connections for this architecture is equal to 71 and 78, respectively. The features extracted from the fc1000 and the probe layers immediately after it is a column matrix of 1*1*1000, and it is stated in Matlab that the total learnable value is equal to 11.6M. We set Momentum, Learn Rate Drop Factor, Learn Rate Drop Period, and L2Regularization to 0.9, 0.1, 10, and 1e-4, respectively, and Initial Learn Rate and Max Epoch to 1e-3 and 30. Validation Patience was set to Inf, with Mini Batch Size and Validation Frequency at 128 and 10 for ResNet18. The study utilized Matlab R2022a on a computer with an 11th Gen Intel Core i5-1135G7, 2.40GHz, and 8GB RAM.

GUI Design

We also designed a user-friendly Graphical User Interface (GUI), which allows users to upload three images to the application to determine which of the benign-malignant, PA-Warthin, and all classes (MC-other benign-other malignant-PA-Warthin) these images belong to. Before starting the design of this software, it is worth noting that we completed the training of all the models used in the study and then saved these trained networks in a suitable folder. Then, we first created the figure file of this GUI and added three axes and three user-controlled buttons and text sections that can print the prediction result on the screen.

RESULTS

The mean age of 47 female and 66 male patients included in this study was 56.2±14.9 years. Table 1 presents the performance metrics obtained using the ResNet18 architecture for the validation set, comprising 20% of the dataset. According to table 1, the architecture categorizing benign-malignant for the T2-w sequence had an accuracy of 93.75% and an F1-score of 94.66%. In the T1-w sequence, the model was most successful in classifying malignant-Warthin images, with an accuracy value equal to 95.37%. This value is also the highest accuracy achieved in the entire study. In T1c-w, DWI-2000, and ADC, the most successful model categorizing benign-malignant images, with their accuracy values equaling 91.07%, 91.84%, and 90.91%, respectively. In addition, for DWI-0, malignant-Warthin images were classified with an accuracy and F1 score of 93.52%, and 92.47%, respectively.

Figure 2 shows the user-friendly GUI, where three images can be easily uploaded, alongside a screenshot of the results. The software displayed the addresses of the uploaded images and successfully predicted 'Benign' when a benign tumor from the T1-w sequence was input. A pleomorphic adenoma from the T2-w sequence was applied to the middle axis, yielding the correct label 'Pleomorphic Adenoma.' Lastly, a mucoepidermoid carcinoma from the ADC sequence was uploaded, and the GUI accurately detected it as 'Mucoepidermoid Carcinoma.' These results demonstrate the proposed GUI application's effectiveness in correctly identifying image labels.

Table 1. Performance metrics obtained for the validation set using ResNet18 architecture

	Train (80%), Validation (20%)	Accuracy %	Sensitivity %	Specificity %	Precision %	F1 score %
T2-w	Benign-Malignant	93.75	100	86	89.86	94.66
	PA-Warthin	87.95	87.50	88.37	87.50	87.50
	Malignant-Warthin	93.46	88	98.25	97.78	92.63
	All classes	86.67	85.52	96.51	88.33	86.47
T1-w	Benign-Malignant	91.07	98.39	82	87.14	92.42
	PA-Warthin	90.36	87.50	93.02	92.11	89.74
	Malignant-Warthin	95.37	90	100	100	94.74
	All classes	81.21	79.95	95.06	84.65	81.68
T1c-w	Benign-Malignant	91.07	100	80	86.11	92.54
	PA-Warthin	86.75	85	88.37	87.18	86.08
	Malignant-Warthin	90.74	82	98.28	97.62	89.13
	All classes	83.64	84.02	95.76	85.59	84.39
DWI-0	Benign-Malignant	90.18	100	78	84.93	91.85
	PA-Warthin	90.36	92.50	88.37	88.10	90.24
	Malignant-Warthin	93.52	86	100	100	92.47
	All classes	84.85	83.90	96.01	87.65	85.20
DWI-2000	Benign-Malignant	91.84	100	78.95	88.24	93.75
	PA-Warthin	83.95	97.37	72.09	75.51	85.06
	Malignant-Warthin	91.58	78.95	100	100	88.24
	All classes	80.95	78.29	94.79	86.61	81.21
ADC	Benign-Malignant	90.91	100	79.17	86.11	92.54
	PA-Warthin	86.75	87.50	86.05	85.37	86.42
	Malignant-Warthin	85.85	70.83	98.28	97.14	81.93
	All classes	84.05	82.45	95.81	86.91	83.99

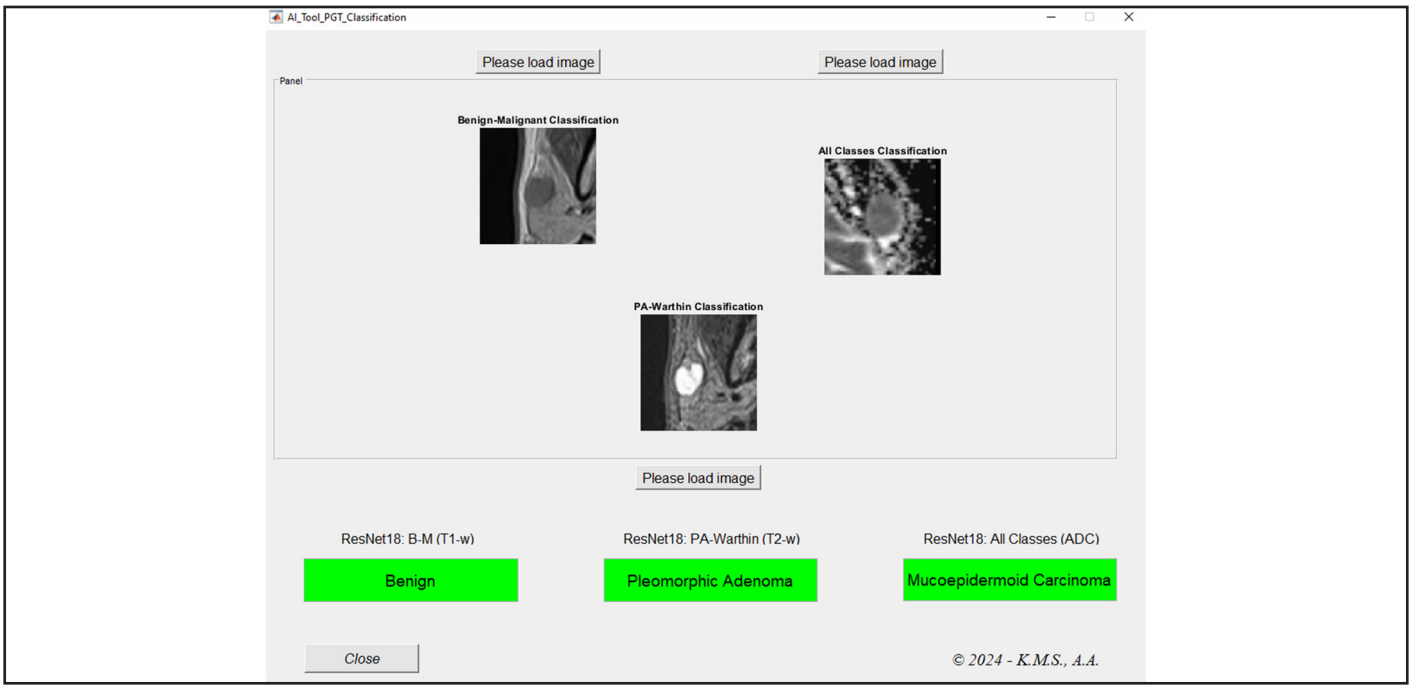

Figure 2. Screenshot of the GUI application designed for this study

DISCUSSION

Parotid gland tumors are the most common salivary gland tumors and their treatment depends on whether they are benign or malignant, less aggressive approaches are used for benign tumors, while more aggressive approaches are preferred for malignant tumors [9]. Therefore, preoperative differential diagnosis of PGTs is critical. The most used and most sensitive technique for differential diagnosis is MRI. However, despite current MRI sequences, the success of differential diagnosis is not at the desired level [9,12].

Artificial intelligence-based architecture is now observed to be usable in almost every field. It is significant to note that optimization and AI-based models are utilized in various subjects, from education to health optimization to renewable energy sources [15,16]. In the scope of this study, we aim to successfully classify PGTs.

In this study, we focused on four classifications. First, we aimed at benign-malignant differentiation, the most significant and initial step in the differential diagnosis process, testing the success of different sequences. Then, we differentiated PA from Warthin tumors, both benign but with malignant potential if untreated, evaluating the performance of all sequences. We then assessed the success of all sequences in differentiating Warthin and malignant tumors, which are challenging due to similar ADC values, and finally in distinguishing all five groups.

When evaluating the results of our study, it is evident that satisfactory success values were achieved, with the highest accuracy of 93.75% in benign-malignant differentiation. A notable finding is that the pre-contrast T1-w sequence outperformed the contrast-enhanced T1-w sequence in PA-Warthin and Warthin-malignant differentiation. Based on these results, we anticipate that DL could reduce the need for contrast media in future tumor differentiation, potentially lowering costs and side effects, though further studies are required.

Regarding the models used, the ResNet18 architecture proved advantageous in terms of training time and performance metrics. Additionally, the user-friendly GUI application developed in the study may aid users in diagnosis and treatment planning. Considering this information, we describe the studies in literature in the following paragraph.

In [17], an accuracy of 0.81 was achieved in pixel-wise tumor classification with U-net using five different sequences. In [11], the highest success rate for benign-malignant discrimination was 0.739 when the sequences were used separately, and 0.822 when all sequences were combined. In another study, classification was performed with five different sequences and the classification success between the control group, PA, Warthin, and malignant tumors was reported as 92.86% [12]. X. Liu et al. [18] published a classification accuracy of 0.84 in benign-malignant discrimination with DWI sequence. Matsuo et al. [19] presented an AUC value of 0.86 with the VGG16 model. When these results are evaluated with the results of our study, it is seen that our results are competitive and successful according to the studies in the literature. The GUI model we developed constitutes the innovative aspect of our study, which is different from the studies in the literature. In addition, not only MRI was used in the literature; CT was used in [20], ultrasound imaging and clinical data in [21]. Moreover, we have conducted some segmentation and fusion-based classification studies [22-24]. We believe that these studies will contribute to differential diagnosis and preoperative preparation in the future.

This study has some limitations. Firstly, the number of patients was limited and different in benign and malignant groups. We tried to address this with augmentation techniques. Secondly, only ResNet18 architecture was used in this study. In the future, comparative studies with other popular models for image classification such as DenseNet or EfficientNet can be conducted. Thirdly, we performed differential diagnosis with deep networks through single slices with two-dimensional images. In the future, with a larger number of images, studies can be performed with three-dimensional images that include the entire tumor.

CONCLUSION

This study demonstrated the effectiveness of DL models in the MRI-based differential diagnosis of PGTs. The ResNet18 architecture achieved high accuracy across various MRI sequences, with the best performance observed in malignant-Warthin classification using T1-w. The T2-w sequence was most effective for benign-malignant classification, while T1-w and DWI-b0 showed the highest performance in PA-Warthin differentiation. The proposed DL approach enables automated and highly accurate classification of PGTs. The developed software quickly and efficiently classifies images, highlighting

its potential for clinical implementation. Future prospective studies involving larger datasets and advanced architecture could further increase the generalizability of the software.

Competing Interests: The authors have no relevant financial or non-financial interests to disclose.

Informed Consent: Informed consent was obtained from all individual participants included in the study.

Funding: This research did not receive a specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ethics Approval: This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of Recep Tayyip Erdogan University, Faculty of Medicine (2021/175).

Author Contributions: Research conception and design: EK, KMS, AA, FBÇ. Data acquisition: EK, KMS, AA, FBÇ. Software: EK, KMS, AA. Data analysis and interpretation: EK, KMS, AA, FBÇ. Drafting of the manuscript: EK, KMS, AA, FBÇ. Critical revision of the manuscript: EK, KMS, AA, FBÇ. Approval of the final manuscript: EK, KMS, AA, FBÇ.

REFERENCES

[1] Coudert H, Mirafzal S, Dissard A, Boyer L, Montoriol PF (2021) Multiparametric magnetic resonance imaging of parotid tumors: A systematic review. Diagn Interv Imaging. 102(3):121-130. https://doi.org/10.1016/j.diii.2020.08.002

[2] de Oliveira FA, Duarte EC, Taveira CT, Máximo AA, de Aquino EC, Alencar RdeC, Vencio EF (2009) Salivary gland tumor: a review of 599 cases in a Brazilian population. Head Neck Pathol. 3(4):271-275. https://doi.org/10.1007/s12105-009-0139-9

[3] Guzzo M, Locati LD, Prott FJ, Gatta G, McGurk M, Licitra L (2010) Major and minor salivary gland tumors. Crit Rev Oncol Hematol. 74(2):134-148. https://doi.org/10.1016/j.critrevonc.2009.10.004

[4] Piludu F, Marzi S, Ravanelli M, Pellini R, Covello R, Terrenato I, Farina D, Campora R, Ferrazzoli V, Vidiri A (2021) MRI-Based Radiomics to Differentiate between Benign and Malignant Parotid Tumors With External Validation. Front Oncol. 11:656918. https://doi.org/10.3389/fonc.2021.656918

[5] Cantisani V, David E, Sidhu PS, Sacconi B, Greco A, Pandolfi F, Tombolini M, Lo Mele L, Calliada F, Brunese L, Catalano C, De Vincentiis M, Di Leo N, Ascenti G, D'Ambrosio F (2016) Parotid Gland Lesions: Multiparametric Ultrasound and MRI Features. Ultraschall Med. 37(5):454-471. https://doi.org/10.1055/s-0042-109171

[6] Kim SY, Borner U, Lee JH, Wagner F, Tshering Vogel DW (2022) Magnetic resonance imaging of parotid gland tumors: a pictorial essay. BMC Med Imaging. 22(1):191. https://doi.org/10.1186/s12880-022-00924-0

[7] Thoeny HC (2007) Imaging of salivary gland tumours. Cancer Imaging. 7(1):52-62. https://doi.org/10.1102/1470-7330.2007.0008

[8] Christe A, Waldherr C, Hallett R, Zbaeren P, Thoeny H (2011) MR imaging of parotid tumors: typical lesion characteristics in MR imaging improve discrimination between benign and malignant disease. AJNR Am J Neuroradiol. 32(7):1202-1207. https://doi.org/10.3174/ajnr.A2520

[9] Liang YY, Xu F, Guo Y, Wang J (2018) Diagnostic accuracy of magnetic resonance imaging techniques for parotid tumors, a systematic review and meta-analysis. Clin Imaging. 52:36-43. https://doi.org/10.1016/j.clinimag.2018.05.026

[10] Xu Z, Chen M, Zheng S, Chen S, Xiao J, Hu Z, Lu L, Yang Z, Lin D (2022) Differential diagnosis of parotid gland tumours: Application of SWI combined with DWI and DCE-MRI. Eur J Radiol. 146:110094. https://doi.org/10.1016/j.ejrad.2021.110094

[11] Xia X, Feng B, Wang J, Hua Q, Yang Y, Sheng L, Mou Y, Hu W (2021) Deep learning for differentiating benign from malignant parotid lesions on MR images. Front Oncol. 11:632104. https://doi.org/10.3389/fonc.2021.632104

[12] Gunduz E, Alçin OF, Kizilay A, Yildirim IO (2022) Deep learning model developed by multiparametric MRI in differential diagnosis of parotid gland tumors. Eur Arch Otorhinolaryngol. 279(11):5389-5399. https://doi.org/10.1007/s00405-022-07455-y

[13] Chlap P, Min H, Vandenberg N, Dowling J, Holloway L, Haworth A (2021) A review of medical image data augmentation techniques for deep learning applications. J Med Imaging Radiat Oncol. 65(5):545-563. https://doi.org/10.1111/1754-9485.13261

[14] "Resnet18 @ Www.Mathworks.Com." [Online]. Available: https://www.mathworks.com/help/deeplearning/ref/resnet18.html

[15] Xiang H, Zou Q, Nawaz MA, Huang X, Zhang F, Yu H (2023) Deep learning for image inpainting: A survey. Pattern Recognit. 134:109046. https://doi.org/10.1016/j.patcog.2022.109046

[16] Guher AB, Tasdemir S, Yaniktepe B (2020) Effective estimation of hourly global solar radiation using machine learning algorithms. Int J Photoenergy. 2020:8843620. https://doi.org/10.1155/2020/8843620

[17] Chang YJ, Huang TY, Liu YJ, Chung HW, Juan CJ (2021) Classification of parotid gland tumors by using multimodal MRI and deep learning. NMR Biomed. 34(1):e4408. https://doi.org/10.1002/nbm.4408

[18] Liu X, Pan Y, Zhang X, Sha Y, Wang S, Li H, Liu J (2023) A deep learning model for classification of parotid neoplasms based on multimodal magnetic resonance image sequences. Laryngoscope. 133(2):327-335. https://doi.org/10.1002/lary.30154

[19] Matsuo H, Nishio M, Kanda T, Kojita Y, Kono AK, Hori M, Teshima M, Otsuki N, Nibu KI, Murakami T (2020) Diagnostic accuracy of deep-learning with anomaly detection for a small amount of imbalanced data: discriminating malignant parotid tumors in MRI. Sci Rep. 10(1):19388. https://doi.org/10.1038/s41598-020-76389-4

[20] Shen XM, Mao L, Yang ZY, Chai ZK, Sun TG, Xu Y, Sun ZJ (2023) Deep learning-assisted diagnosis of parotid gland tumors by using contrast-enhanced CT imaging. Oral Dis. 29(8):3325-3336. https://doi.org/10.1111/odi.14474

[21] Zhang G, Zhu L, Huang R, Xu Y, Lu X, Chen Y, Li C, Lei Y, Luo X, Li Z, Yi S, He J, Zheng C (2023) A deep learning model for the differential diagnosis of benign and malignant salivary gland tumors based on ultrasound imaging and clinical data. Quant Imaging Med Surg. 13(5):2989-3000. https://doi.org/10.21037/qims-22-950

[22] Sunnetci KM, Kaba E, Celiker FB, Alkan A (2024) MR Image Fusion-Based Parotid Gland Tumor Detection. J Imaging Inform Med. https://doi.org/10.1007/s10278-024-01137-3

[23] Sunnetci KM, Kaba E, Beyazal Çeliker F, Alkan A (2023) Comparative parotid gland segmentation by using ResNet-18 and MobileNetV2-based DeepLab v3+ architectures from magnetic resonance images. Concurr Comput Pract Exp. 35(1):1-14. https://doi.org/10.1002/cpe.7405

[24] Sunnetci KM, Kaba E, Celiker FB, Alkan A (2024) Deep Network-Based Comprehensive Parotid Gland Tumor Detection. Acad Radiol. 31(1):157-167. https://doi.org/10.1016/j.acra.2023.04.028

How to Cite;

Kaba E, Sunnetci KM, Alkan A, Beyazal Celiker F (2025) Magnetic Resonance Imaging-Based Differential Diagnosis of Parotid Gland Tumors Using Deep Learning. Eur J Ther. 31(1):44-50. https://doi.org/10.58600/eurjther2612