Gitto et al. European Radiology Experimental (2025)9:91
https://doi.org/10.1186/s41747-025-00633-7

European Radiology EXPERIMENTAL

ORIGINAL ARTICLE Open Access

Deep learning for automated segmentation of central cartilage tumors on MRI

Salvatore Gitto[1,2], Anna Corti[3], Kirsten van Langevelde[4], Ana Navas Cañete[4], Antonino Cincotta[2], Carmelo Messina[1,5], Domenico Albano[2,6], Carlotta Vignaga[3], Laura Ferrari[3], Luca Mainardi[3], Valentina D. A. Corino[3,7] and Luca Maria Sconfienza[1,2*]

Abstract

Background Automated segmentation methods may potentially increase the reliability and applicability of radiomics in skeletal oncology. Our aim was to propose a deep learning-based method for automated segmentation of atypical cartilaginous tumor (ACT) and grade II chondrosarcoma (CS2) of long bones on magnetic resonance imaging (MRI).

Materials and methods This institutional review board-approved retrospective study included 164 patients with surgically treated and histology-proven cartilaginous tumors at two tertiary bone tumor centers. The first cohort consisted of 99 MRI scans from center 1 (79 ACT, 20 CS2). The second cohort consisted of 65 MRI scans from center 2 (45 ACT, 20 CS2). Supervised Edge-Attention Guidance segmentation Network (SEAGNET) architecture was employed for automated image segmentation on T1-weighted images, using manual segmentations drawn by musculoskeletal radiologists as the ground truth. In the first cohort, a total of 1,037 slices containing the tumor out of 99 patients were split into 70% training, 15% validation, and 15% internal test sets, respectively, and used for model tuning. The second cohort was used for independent external testing.

Results In the first cohort, Dice Score (DS) and Intersection over Union (IoU) per patient were 0.782 ± 0.148 and 0.663 ± 0.175, and 0.748 ± 0.191 and 0.630 ± 0.210 in the validation and internal test sets, respectively. DS and IoU per slice were 0.742 ± 0.273 and 0.646 ± 0.266, and 0.752 ± 0.256 and 0.656 ± 0.261 in the validation and internal test sets, respectively. In the independent external test dataset, the model achieved DS of 0.828 ± 0.175 and IoU of 0.706 ± 0.180.

Conclusion Deep learning proved excellent for automated segmentation of central cartilage tumors on MRI.

Relevance statement A deep learning model based on SEAGNET architecture achieved excellent performance for automated segmentation of cartilage tumors of long bones on MRI and may be beneficial, given the increasing detection rate of these lesions in clinical practice.

Key Points

- Automated segmentation may potentially increase the reliability and applicability of radiomics-based models.
- A deep learning architecture was proposed for automated segmentation of appendicular cartilage tumors on MRI.
- Deep learning proved excellent with a mean Dice Score of 0.828 in the external test cohort.

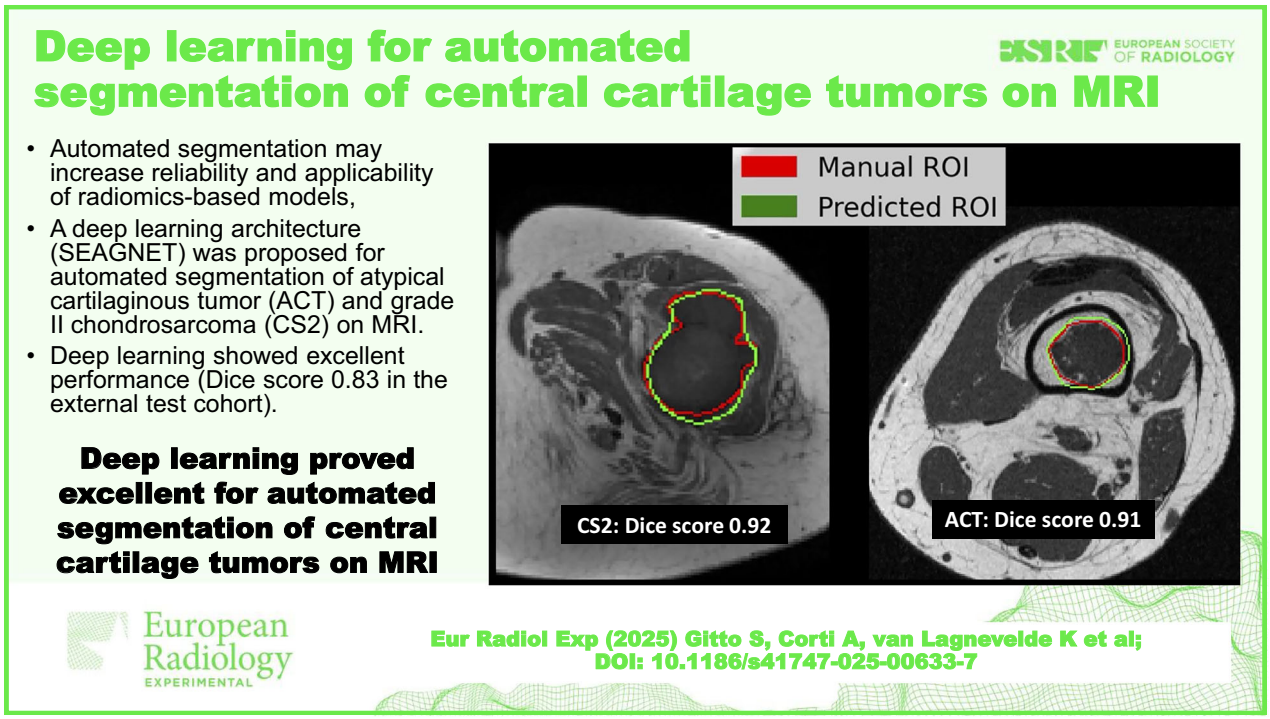
Salvatore Gitto and Anna Corti contributed equally to this work.

Valentina D.A. Corino and Luca Maria Sconfienza jointly supervised this work.

*Correspondence:
Luca Maria Sconfienza
io@lucasconfienza.it
Full list of author information is available at the end of the article

Keywords Chondrosarcoma, Deep learning, Machine learning, Magnetic resonance imaging, Radiomics

Graphical Abstract



Background

Over the last three decades, the reported incidence of central cartilage tumors has markedly increased, especially in the category of atypical cartilaginous tumors (ACTs), due to an increase in incidental findings on imaging studies performed for unrelated reasons [1, 2]. According to the 2020 edition of the World Health Organization classification of bone tumors, the term ACT denotes an intermediate (locally aggressive) cartilage lesion located in long bones, which is low grade and shows indolent behavior with unlikelihood to metastasize [3]. Cartilage tumors with the same histology as ACT, but located in the axial skeleton, are classified as grade I chondrosarcoma [3]. ACTs are entirely different lesions from grade I axial and high-grade (grades II and higher, regardless of the axial or appendicular location) chondrosarcomas, which are malignant tumors with metastatic potential and require wide resection with free margins [3]. On the other hand, ACTs can be treated with intralesional curettage for sufficient local control [4]. Additionally, watchful waiting has progressively become an increasingly favored option for asymptomatic ACTs, based on new insights on their natural course with low growth rates and no reported transformation into high-grade chondrosarcoma [5–8]. Therefore, in long bones, the difference in treatment strategies between ACT and enchondroma—its benign counterpart—is progressively disappearing, and the radiological focus has shifted towards the clinically most relevant distinction between ACT and high-grade chondrosarcoma, particularly grade II chondrosarcoma (CS2) [9].

Several studies have recently attempted to diagnose and grade central cartilage tumors using radiomics, either alone [10] or combined with machine learning algorithms [11–13]. Particularly, MRI radiomics achieved excellent accuracy for the classification of ACT and CS2 of long bones [12]. In most of these studies, manual segmentation was employed to draw regions of interest (ROIs) along the lesion borders, which were subsequently embedded in the radiomics workflow [14]. One study employed a semi-automated segmentation process, which was based on manual segmentation of the upper and lower portions of the lesion and automated computation of the lesion volume on the slices in between [15]. Deep learning (DL)-based fully automated segmentation methods could ideally achieve higher reliability than both manual and

semiautomated approaches, but their applications have been limited in musculoskeletal oncology [16–19].

In this study, we propose a DL-based method for automated segmentation of central cartilage bone tumors on MRI. In detail, we focused on ACT and CS2 of long bones, as the issue of differentiating these two entities is clinically most relevant and automated segmentation may be combined with radiomics tools recently described for their accurate classification and grading [12].

Materials and methods

Datasets

Institutional Review Board approved this retrospective study and waived the need for informed consent (study protocol: "RADIO-BOSTT," approved on 13 July 2022—Università degli Studi di Milano, Milan, Italy). Consecutive series of patients with ACT or CS2 of long bones and MRI available at one of the two participating institutions (center 1: IRCCS Orthopedic Institute Galeazzi, Milan, Italy; center 2: Leiden University Medical Center, Leiden, Netherlands) were considered for inclusion. Information was collected through the medical records from the orthopedic surgery, pathology and radiology departments. Inclusion criteria were: (1) central ACT or CS2 of long bones; (2) definitive diagnosis based on post-surgical pathology available after resection or curettage; and (3) MRI examination performed within 3 months before surgery and including a T1-weighted sequence in at least one plane. Metacarpal, metatarsal or phalangeal cartilaginous tumors, recurrent lesions and those associated with pathological fractures were excluded.

Overall, 164 patients were retrospectively included at the two participating institutions. The first cohort (from center 1) comprised 99 patients (79 ACT and 20 CS2, located in the femur ($n = 43$), fibula ($n = 10$), humerus ($n = 39$), radius ($n = 1$), and tibia ($n = 6$)) and was employed for training, validation and internal testing. The second cohort (from center 2) comprised 65 patients (45 ACT and 20 CS2, located in the femur ($n = 46$), humerus ($n = 10$) and tibia ($n = 9$)) and was employed for independent external testing. In the first cohort, MRI scans were performed on a 1.5-T unit (Magnetom Avanto or Magnetom Espree, Siemens Healthineers). In the second cohort, MRI scans were performed on a 3-T (Ingenia or Intera, Philips Medical System) or 1.5-T (Ingenia, Philips Medical System) unit. In addition, externally acquired MRI scans of patients referred to both participating institutions were included if they were performed on at least 1.5-T MRI units and T1-weighted sequence was available in the study protocol. MRI specifications for both centers are reported in the Supplementary material. The same cohorts of patients were used in a previous study focused on radiomics-based machine learning classification of cartilaginous bone tumors, where manual segmentation was performed only on the slice showing the largest tumor diameter to obtain the ROIs subsequently embedded in the radiomics workflow [12].

Ground truth segmentation

Noncontrast T1-weighted sequence was chosen for image segmentation in the present study, as MRI radiomics-based machine learning already showed excellent accuracy in differentiating ACT from CS2 using T1-weighted sequence only [12] and could ideally be combined with automated segmentation tools to facilitate the workflow. A musculoskeletal radiologist per center/cohort performed manual contour-focused segmentation (A.C. in center 1 and S.G. in center 2, with 3 and 5 years of work experience in a tertiary bone sarcoma center, respectively). In the first cohort, an ROI including the whole tumor volume was drawn along the tumor borders using the "polygon mode," slice by slice, on the open-source software ITK-SNAP (version 4.2.2) [20]. Cartilage nodules separated by interspersed fatty marrow were also included in the ROI. In the second cohort, for each patient, an ROI was drawn on the slice showing the largest tumor diameter using the same segmentation method and software. Axial T1-weighted sequences were used as the first choice, and coronal or sagittal sequences were used as the second choice for image segmentation, respectively. These manually drawn ROIs were used as the ground truth when assessing DL performance for automated segmentation.

Image preprocessing

Image preprocessing involved different steps. First, MRI images and their corresponding ROIs were resized to 256 × 256 pixels. Second, MRI images were normalized to rescale intensities to a range of 0 to 1. Specifically, pixel values were scaled by subtracting the minimum pixel value and dividing by the maximum-minimum difference. For each slice, the minimum and maximum pixel values were computed independently, and the pixel intensities were adjusted accordingly. This per-slice normalization accounts for variability of pixel intensities across different slices and ensures that each slice is rescaled individually.

DL-based automated segmentation

We adopted the Supervised Edge-Attention Guidance segmentation Network (SEAGNET) DL architecture for image segmentation described by Zhan et al for malignant bone tumor segmentation [21]. The SEAGNET architecture was implemented using a ResNet50 backbone for feature extraction, augmented by a Feature Pyramid Network and a decoder incorporating mixed attention layers and dilation convolutions. The model output was a single channel with sigmoid activation for binary

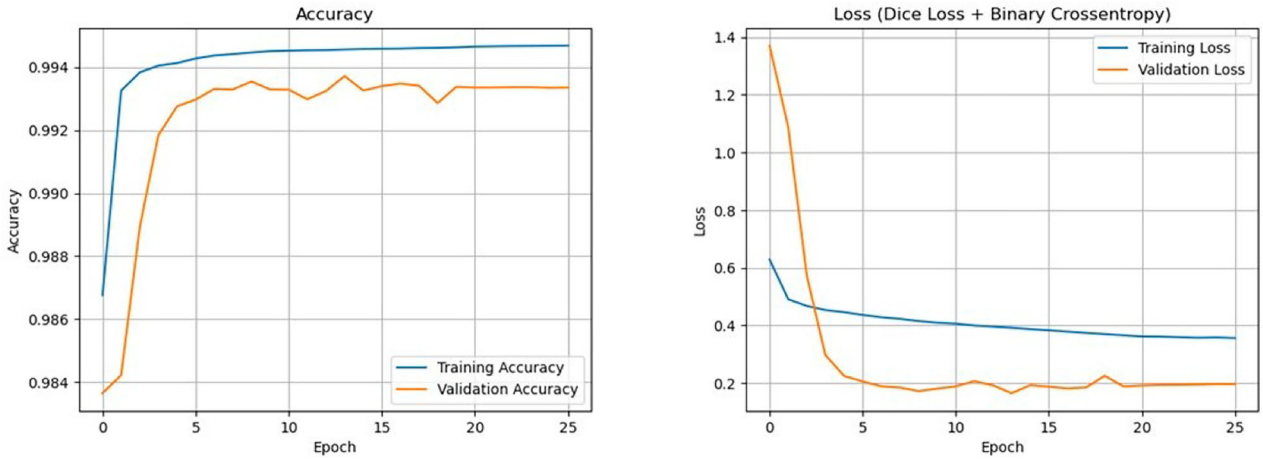


Fig. 1 Accuracy and loss of the deep learning model (SEAGNET) during the training and validation phases

Table 1 Dice score (DS) and intersection over union (IoU) computed per patient and per slice in the validation and internal testing sets from the first cohort (center 1)

Per-patient metrics	Validation	Internal test
DS	0.782 ± 0.148	0.748 ± 0.191
IoU	0.663 ± 0.175	0.630 ± 0.210
Per-slice metrics	**Validation**	**Internal test**
DS	0.742 ± 0.273	0.752 ± 0.256
IoU	0.646 ± 0.266	0.656 ± 0.261

Data are given as mean ± standard deviation

segmentation. A detailed description of the SEAGNET architecture is provided in the Supplementary material. In the Supplementary material, the performance of another DL architecture (nnU-Net: https://github.com/MIC-DKFZ/nnUNet) commonly used for segmentation tasks is also assessed for comparison. Of note, the terms "DL architecture" and "DL model" refer to SEAGNET in the following sections of this manuscript unless otherwise specified.

The first cohort (center 1) was used for training, validation and internal testing. Specifically, only the slices containing the tumor were included, resulting in a total of 1,037 slices from 99 patients. This dataset was split into 70% training, 15% validation and 15% internal test sets to ensure robust model assessment. Specifically, 718 slices from 68 patients constituted the training set, 160 slices from 17 patients constituted the validation set and 159 slices from 14 patients constituted the internal test set, respectively. Data augmentation was applied to the training set to improve generalizability and address the imbalance in tumor slices in the first cohort (center 1), thus increasing the number of slices with tumor from 718 to 3,590. The augmentation techniques included random image flipping (horizontal and vertical), rotation (angle range: 10° to 70°), width and height translations (pixel range: -20 to 20), and zoom-in and zoom-out (zoom range: 0.8 to 1.2). DL modeling was performed using two different approaches: (1) by computing the metrics for each patient and then averaging them (not taking into account the number of slices per patient); and (2) by computing the metrics for each slice and then averaging them, ensuring that each patient gave a weighted contribution to the final metrics (proportional to the number of slices with tumor per patient). The performance metrics included accuracy, Dice Score (DS) and the Intersection over Union (IoU). DS measures the similarity between the predicted ROI and the ground truth segmentation (range 0–1, where 1 indicates the highest similarity). IoU evaluates the ratio of the intersection area between the predicted ROI and the ground truth segmentation to their union area (range 0–1, where 1 means perfect overlap). The model was trained using the Adam optimizer with an initial learning rate of 1×10^{-4}. The used loss function was a combination of binary cross-entropy and Dice loss, which balances pixel-wise accuracy with segmentation overlap. The model was trained for a maximum of 100 epochs with a batch size of 16. Early stopping was employed with a patience of 12 epochs, monitoring the validation loss to prevent overfitting. A learning rate reduction strategy (ReduceLROnPlateau) was applied with a patience of 5 epochs, reducing the learning rate by a factor of 10 when the validation loss plateaued. The minimum learning rate threshold was 1×10^{-6}. The validation loss, accuracy, and IoU were monitored throughout the training process to ensure consistency and detect potential overfitting. Finally, the DL model was used for internal (15% of the cohort from center 1) and independent external (entire cohort from

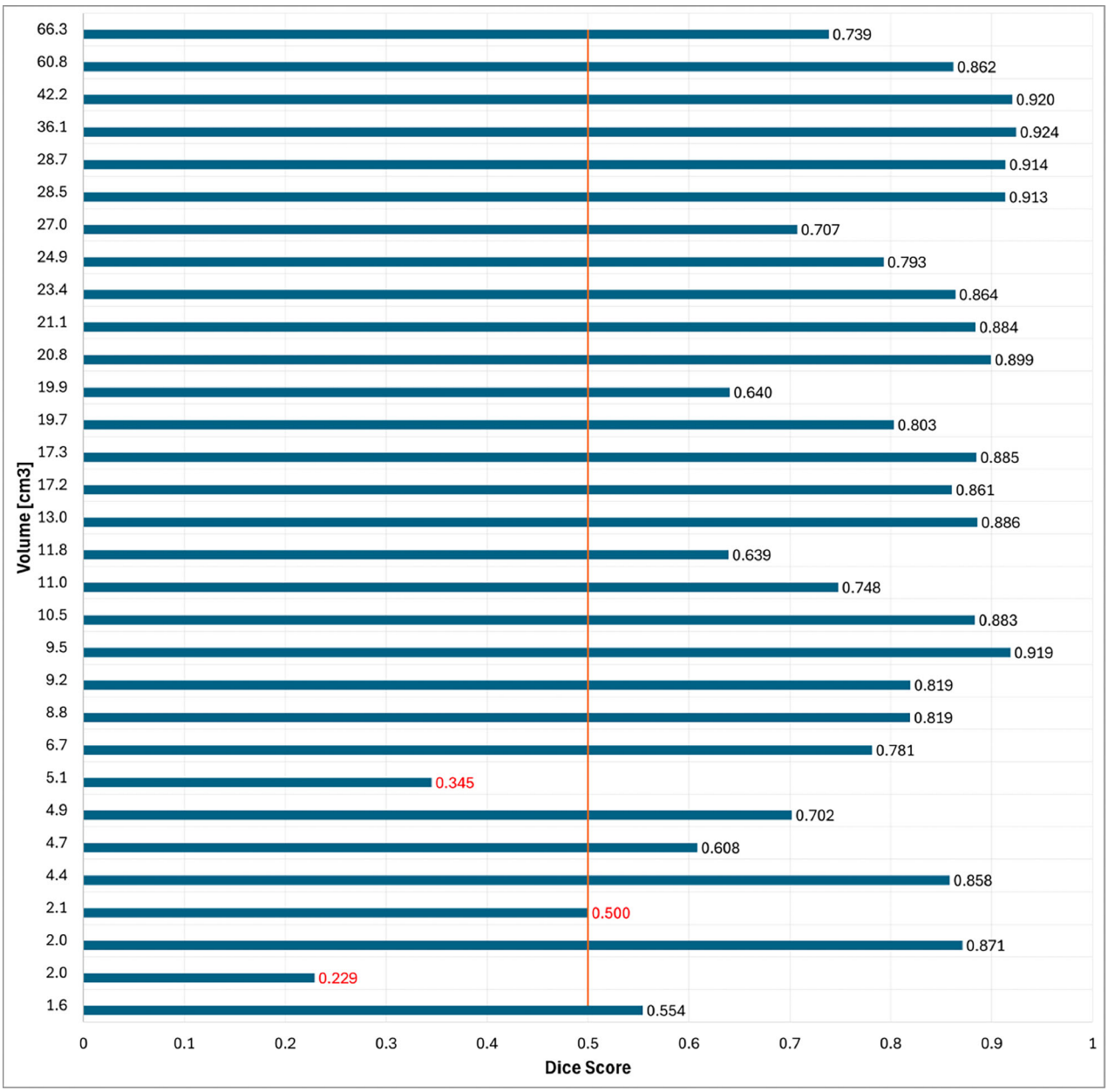


Fig. 2 Dice score per lesion volume in all patients from the validation and internal test sets

center 2, including only the slice showing the largest tumor diameter per patient and therefore 65 slices overall) testing, respectively. Metrics were expressed as mean ± standard deviation.

Results

Accuracy and loss of the DL model (SEAGNET) during the training and validation phases are shown in Fig. 1. DS and IoU computed per patient and per slice in the validation and internal test sets from the first cohort (center 1) are reported in Table 1. In detail, in the internal test set, the DL model showed DS per patient of 0.748 ± 0.191 overall, 0.740 ± 0.213 for ACT and 0.800 ± 0.087 for CS2, respectively. In the internal test set, it showed DS per slice of 0.752 ± 0.256 overall, 0.746 ± 0.268 for ACT and 0.771 ± 0.220 for CS2, respectively. DS showed lower values than the threshold of 0.5 in 3 patients from the first cohort, particularly in 1 patient from the validation set and 2 patients from the internal test set, respectively. In these three cases, tumor volume was 5.1 cm^3 or smaller, as shown in Fig. 2. In the independent external test set (center 2), where one slice with tumor per patient was

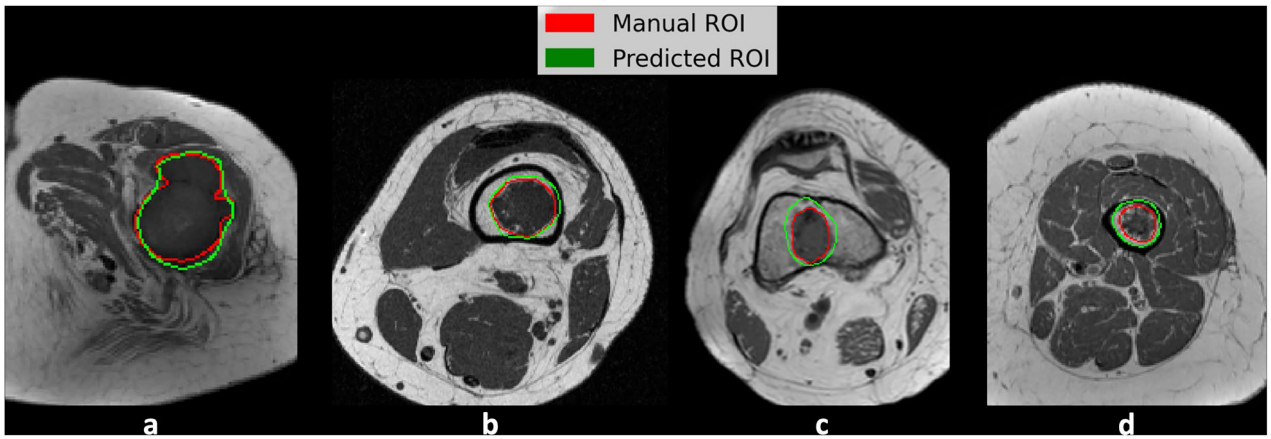


Fig. 3 Examples of manual and automated (predicted) segmentations from the external test dataset (center 2). Highly-to-moderately accurate automated segmentations are shown in a CS2 (**a**) and three ACTs (**b**–**d**) of the femur in different patients, with DS = 0.92 and IoU = 0.86 (**a**), DS = 0.91 and IoU = 0.83 (**b**), DS = 0.76 and IoU = 0.61 (**c**), and DS = 0.74 and IoU = 0.59 (**d**), respectively. ACT, Atypical cartilaginous tumor; CS2, Grade II chondrosarcoma; DS, Dice score; IoU, Intersection over union

included, the DL model showed DS of 0.828 ± 0.175 and IoU of 0.706 ± 0.180, respectively. Figure 3 shows examples of highly-to-moderately accurate DL-based automated segmentations from the external test dataset, respectively. In the Supplementary material, the performance of another DL architecture (nnU-Net) commonly used for segmentation tasks is reported for comparison, with lower results than SEAGNET in our population of study.

Discussion

In the present study, we developed and externally validated a DL-based method to perform automated segmentation of ACT and CS2 of long bones on T1-weighted MRI sequence, which achieved excellent performance with DS of 0.828 ± 0.175 and IoU of 0.706 ± 0.180 in the independent external test dataset, respectively. In addition, our method showed lower DS than the threshold of 0.5 in small lesions (volume ≤ 5.1 cm^3) from the validation and internal test datasets and, therefore, seemed to perform better with larger tumors.

Nowadays, radiologists have to deal with the increasing detection rate of cartilaginous tumors as incidental findings on imaging studies, especially ACTs [1, 2]. MRI is the imaging modality of choice for diagnosis, although a certain degree of interobserver variability in lesion grading is reported even among experts [22]. As an emerging quantitative technique, MRI radiomics holds the potential to address the issue of interobserver variability and shows promise for the classification of cartilage tumors [12]. However, in orthopedic oncology, most radiomics models to date have been built upon manual ROI annotations around the lesion margins, thus resulting in low time efficiency and the need for a preliminary reliability analysis aimed at evaluating segmentation variability [23–26]. This study proposed a DL method for automated segmentation of cartilaginous tumors on MRI, which may eventually be combined with radiomics models for lesion classification and grading. Existing segmentation methods based on different DL architectures were developed using mixed datasets of bone and soft-tissue lesions [27, 28] or were tailored to specific entities such as osteosarcoma [29–32] and vertebral metastases [33]. In our study, we used the SEAGNET architecture, which previously showed good results for automated segmentation of osteosarcoma [21] and now proved to be effective also for automated segmentation of cartilage tumors on MRI. Specifically, in a previous study on osteosarcoma [21], the authors demonstrated the superiority of SEAGNET architecture compared to classical segmentation methods, such as U-Net [34], Fully Convolutional Network [35] and DeepLab V3+ [36], which were commonly employed in other organs [37–39]. Similarly, in our study, SEAGNET showed better performance than nnU-Net (Supplementary material). This superiority can be due to the more sophisticated attention mechanisms of the SEAGNET architecture, as these layers are mixed attention layers that focus on both spatial and channel attention, allowing the model to selectively emphasize critical features within the images. Our results in terms of DS and IoU were in line with previously published studies focusing on different bone lesions [28].

Some limitations of this study should be addressed. First, ACT was over-represented compared to CS2 in both cohorts of our study. Nonetheless, this reflects the incidence of ACT and CS2 in clinical practice [1], and data

augmentation was performed to overcome the imbalance in the training set. Second, the ROIs manually drawn by two radiologists were used as ground truth segmentations, and interobserver reliability was not evaluated as part of this study. However, manual contour-focused segmentation has been demonstrated to be a reproducible technique both on CT and MRI in a recent study focusing on central cartilage tumors [40]. Third, multiple slices per patient (*i.e.*, all slices containing the tumor) were used for model training, validation and internal testing in the first cohort, resulting in a potential risk of overfitting. Nonetheless, the model achieved excellent performance when externally tested against independent data from the second cohort, where only one slice per patient was employed, therefore proving to be robust and effective.

In conclusion, a DL model based on SEAGNET architecture achieved excellent performance for automated segmentation of ACT and CS2 of long bones on T1-weighted MRI sequences. This accurate method may be beneficial given the increasing incidence (or detection rate) of these lesions, especially ACTs. Future studies should focus on combining automated segmentation with radiomics-based tools for the classification and grading of cartilaginous tumors, thus aiding clinicians and radiologists in decision-making.

Abbreviations

ACT	Atypical cartilaginous tumor
CS2	Grade II chondrosarcoma
DL	Deep learning
DS	Dice score
IoU	Intersection over union
ROI	Region of interest
SEAGNET	Supervised edge-attention guidance segmentation network

Supplementary information

The online version contains supplementary material available at https://doi.org/10.1186/s41747-025-00633-7.

ELECTRONIC SUPPLEMENTARY MATERIAL

Acknowledgements

No large language model was used for the preparation of this manuscript.

Author contributions

SG and LMS contributed to the study conception and design. Material preparation and data collection were performed by SG, KvL, ANC, ACi, CM, and DA. Data analysis was performed by ACo, CV, LF, LM, and VDAC. The first draft of the manuscript was written by SG and ACo. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

This research was supported by the Investigator Grant awarded by Fondazione AIRC per la Ricerca sul Cancro for the project "RADIOmics-based machine-learning classification of BOne and Soft Tissue Tumors (RADIO-BOSTT)" (LMS). The funding source provided financial support without any influence on the study design; on the collection, analysis, and interpretation of data; and on the writing of the report. The first and last authors had the final responsibility for the decision to submit the paper for publication.

Data availability

The employed deep learning model is available online in the GitHub repository (https://github.com/annacorti/DL-segmentation-of-central-cartilage-tumors).

Declarations

Ethics approval and consent to participate

Institutional Review Board approved this retrospective study and waived the need for informed consent (study protocol: "RADIO-BOSTT," approved on 13 July 2022—Università degli Studi di Milano, Milan, Italy).

Consent for publication

Institutional Review Board approved this retrospective study and waived the need for informed consent.

Competing interests

SG and LMS are members of the Scientific Editorial Board of *European Radiology Experimental* and did not take part in the review or selection processes of this article. The remaining authors declare no conflicts of interest related to this manuscript.

Author details

[1]Dipartimento di Scienze Biomediche per la Salute, Università degli Studi di Milano, Milan, Italy. [2]IRCCS Istituto Ortopedico Galeazzi, Milan, Italy. [3]Department of Electronics, Information and Bioengineering (DEIB), Politecnico Di Milano, Milan, Italy. [4]Department of Radiology, Leiden University Medical Center (LUMC), Leiden, Netherlands. [5]UOC Radiodiagnostica, ASST Centro Specialistico Ortopedico Traumatologico Gaetano Pini-CTO, Milan, Italy. [6]Dipartimento di Scienze Biomediche, Chirurgiche ed Odontoiatriche, Università Degli Studi di Milano, Milan, Italy. [7]Cardiotech Lab, Centro Cardiologico Monzino IRCCS, Milan, Italy.

Received: 13 April 2025 Accepted: 25 August 2025
Published online: 12 September 2025

References

1. van Praag Veroniek VM, Rueten-Budde AJ, Ho V et al (2018) Incidence, outcomes and prognostic factors during 25 years of treatment of chondrosarcomas. Surg Oncol 27:402–408. https://doi.org/10.1016/j.suronc.2018.05.009
2. Davies AM, Patel A, Botchu R, Azzopardi C, James S, Jeys L (2021) The changing face of central chondrosarcoma of bone. One UK-based orthopaedic oncology unit's experience of 33 years referrals. J Clin Orthop Trauma 17:106–111. https://doi.org/10.1016/j.jcot.2021.02.017
3. WHO Classification of Tumours Editorial Board (2020) WHO classification of tumours: soft tissue and bone tumours. International Agency for Research on Cancer Press, Lyon
4. Strauss SJ, Frezza AM, Abecassis N et al (2021) Bone sarcomas: ESMO–EURACAN–GENTURIS–ERN PaedCan clinical practice guideline for diagnosis, treatment and follow-up. Ann Oncol 32:1520–1536. https://doi.org/10.1016/j.annonc.2021.08.1995
5. Deckers C, Schreuder BHW, Hannink G, de Rooy JWJ, van der Geest ICM (2016) Radiologic follow-up of untreated enchondroma and atypical cartilaginous tumors in the long bones. J Surg Oncol 114:987–991. https://doi.org/10.1002/jso.24465
6. Deckers C, de Rooy JWJ, Flucke U, Schreuder HWB, Dierselhuis EF, van der Geest ICM (2021) Midterm MRI follow-up of untreated enchondroma and atypical cartilaginous tumors in the long bones. Cancers (Basel) 13:4093. https://doi.org/10.3390/cancers13164093
7. Omlor GW, Lohnherr V, Lange J et al (2019) Outcome of conservative and surgical treatment of enchondromas and atypical cartilaginous tumors of

the long bones: retrospective analysis of 228 patients. BMC Musculoskelet Disord 20:134. https://doi.org/10.1186/s12891-019-2502-7
8. Scholte CHJ, Dorleijn DMJ, Krijvenaar DT, van de Sande MAJ, van Langevelde K (2024) Wait-and-scan: an alternative for curettage in atypical cartilaginous tumours of the long bones. Bone Joint J 106-B:86–92. https://doi.org/10.1302/0301-620X.106B1.BJJ-2023-0467.R1
9. van de Sande MAJ, van der Wal RJP, Navas Cañete A et al (2019) Radiologic differentiation of enchondromas, atypical cartilaginous tumors, and high-grade chondrosarcomas—improving tumor-specific treatment: a paradigm in transit? Cancer 125:3288–3291. https://doi.org/10.1002/cncr.32404
10. Fritz B, Müller DA, Sutter R et al (2018) Magnetic resonance imaging–based grading of cartilaginous bone tumors. Invest Radiol 53:663–672. https://doi.org/10.1097/RLI.0000000000000486
11. Gitto S, Cuocolo R, Annovazzi A et al (2021) CT radiomics-based machine learning classification of atypical cartilaginous tumours and appendicular chondrosarcomas. EBioMedicine 68:103407. https://doi.org/10.1016/j.ebiom.2021.103407
12. Gitto S, Cuocolo R, van Langevelde K et al (2022) MRI radiomics-based machine learning classification of atypical cartilaginous tumour and grade II chondrosarcoma of long bones. EBioMedicine 75:103757. https://doi.org/10.1016/j.ebiom.2021.103757
13. Gitto S, Annovazzi A, Nulle K et al (2024) X-rays radiomics-based machine learning classification of atypical cartilaginous tumour and high-grade chondrosarcoma of long bones. EBioMedicine 101:105018. https://doi.org/10.1016/j.ebiom.2024.105018
14. Zhong J, Hu Y, Ge X et al (2022) A systematic review of radiomics in chondrosarcoma: assessment of study quality and clinical value needs handy tools. Eur Radiol 33:1433–1444. https://doi.org/10.1007/s00330-022-09060-3
15. Lisson CS, Lisson CG, Flosdorf K et al (2018) Diagnostic value of MRI-based 3D texture analysis for tissue characterisation and discrimination of low-grade chondrosarcoma from enchondroma: a pilot study. Eur Radiol 28:468–477. https://doi.org/10.1007/s00330-017-5014-6
16. Erdem F, Gitto S, Fusco S et al (2024) Automated detection of bone lesions using CT and MRI: a systematic review. Radiol Med 129:1898–1905. https://doi.org/10.1007/s11547-024-01913-9
17. Gitto S, Serpi F, Albano D et al (2024) AI applications in musculoskeletal imaging: a narrative review. Eur Radiol Exp 8:22. https://doi.org/10.1186/s41747-024-00422-8
18. Gitto S, Cuocolo R, Huisman M et al (2024) CT and MRI radiomics of bone and soft-tissue sarcomas: an updated systematic review of reproducibility and validation strategies. Insights Imaging 15:54. https://doi.org/10.1186/s13244-024-01614-x
19. Gitto S, Cuocolo R, Albano D et al (2021) CT and MRI radiomics of bone and soft-tissue sarcomas: a systematic review of reproducibility and validation strategies. Insights Imaging 12:68. https://doi.org/10.1186/s13244-021-01008-3
20. Yushkevich PA, Piven J, Hazlett HC et al (2006) User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 31:1116–1128. https://doi.org/10.1016/j.neuroimage.2006.01.015
21. Zhan X, Liu J, Long H et al (2023) An intelligent auxiliary framework for bone malignant tumor lesion segmentation in medical image analysis. Diagnostics (Basel) 13:223. https://doi.org/10.3390/diagnostics13020223
22. Zamora T, Urrutia J, Schweitzer D, Amenabar PP, Botello E (2017) Do orthopaedic oncologists agree on the diagnosis and treatment of cartilage tumors of the appendicular skeleton? Clin Orthop Relat Res 475:2176–2186. https://doi.org/10.1007/s11999-017-5276-y
23. Gitto S, Corino VDA, Annovazzi A et al (2022) 3D vs. 2D MRI radiomics in skeletal Ewing sarcoma: feature reproducibility and preliminary machine learning analysis on neoadjuvant chemotherapy response prediction. Front Oncol 12:1016123. https://doi.org/10.3389/fonc.2022.1016123
24. Gitto S, Bologna M, Corino VDA et al (2022) Diffusion-weighted MRI radiomics of spine bone tumors: feature stability and machine learning-based classification performance. Radiol Med 127:518–525. https://doi.org/10.1007/s11547-022-01468-7
25. Gitto S, Cuocolo R, Giannetta V et al (2024) Effects of interobserver segmentation variability and intensity discretization on MRI-based radiomic feature reproducibility of lipoma and atypical lipomatous tumor. J Imaging Inform Med 37:1187–1200. https://doi.org/10.1007/s10278-024-00999-x
26. Gitto S, Cuocolo R, Klontzas ME, Albano D, Messina C, Sconfienza LM (2025) Quality appraisal of radiomics-based studies on chondrosarcoma using METhodological RadiomICs Score (METRICS) and Radiomics Quality Score (RQS). Insights Imaging 16:129. https://doi.org/10.1186/s13244-025-02016-3
27. Qu Y, Li X, Yan Z et al (2021) Surgical planning of pelvic tumor using multi-view CNN with relation-context representation learning. Med Image Anal 69:101954. https://doi.org/10.1016/j.media.2020.101954
28. Wang S, Sun M, Sun J et al (2024) Advancing musculoskeletal tumor diagnosis: automated segmentation and predictive classification using deep learning and radiomics. Comput Biol Med 175:108502. https://doi.org/10.1016/j.compbiomed.2024.108502
29. Liu F, Zhu J, Lv B et al (2022) Auxiliary segmentation method of osteosarcoma MRI image based on Transformer and U-Net. Comput Intell Neurosci 2022:9990092. https://doi.org/10.1155/2022/9990092
30. Wu J, Xiao P, Huang H, Gou F, Zhou Z, Dai Z (2022) An artificial intelligence multiprocessing scheme for the diagnosis of osteosarcoma MRI images. IEEE J Biomed Health Inform 26:4656–4667. https://doi.org/10.1109/JBHI.2022.3184930
31. Wu J, Yang S, Gou F et al (2022) Intelligent segmentation medical assistance system for MRI images of osteosarcoma in developing countries. Comput Math Methods Med 2022:7703583. https://doi.org/10.1155/2022/7703583
32. Zhong J, Zhang C, Hu Y et al (2022) Automated prediction of the neoadjuvant chemotherapy response in osteosarcoma with deep learning and an MRI-based radiomics nomogram. Eur Radiol 32:6196–6206. https://doi.org/10.1007/s00330-022-08735-1
33. Kim DH, Seo J, Lee JH et al (2024) Automated detection and segmentation of bone metastases on spine MRI using U-Net: a multicenter study. Korean J Radiol 25:363. https://doi.org/10.3348/kjr.2023.0671
34. Ronneberger O, Fischer P, Brox T (2015) U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, Frangi A (eds) Medical image computing and computer-assisted intervention (MICCAI) 2015. Lecture notes in computer science. Springer, Cham, pp 234–241
35. Shelhamer E, Long J, Darrell T (2017) Fully convolutional networks for semantic segmentation. IEEE Trans Pattern Anal Mach Intell 39:640–651. https://doi.org/10.1109/TPAMI.2016.2572683
36. Chen L-C, Zhu Y, Papandreou G, Schroff F, Adam H (2018) Encoder-Decoder with atrous separable convolution for semantic image segmentation. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y (eds) Computer vision (ECCV) 2018. Lecture notes in computer science. Springer, Cham, pp 833–851
37. Sore R, Cathier P, Vlachomitrou AS et al (2024) Deep learning-based segmentation of kidneys and renal cysts on T2-weighted MRI from patients with autosomal dominant polycystic kidney disease. Eur Radiol Exp 8:122. https://doi.org/10.1186/s41747-024-00520-7
38. Rinneburger M, Carolus H, Iuga A-I et al (2023) Automated localization and segmentation of cervical lymph nodes on contrast-enhanced CT using a 3D foveal fully convolutional neural network. Eur Radiol Exp 7:45. https://doi.org/10.1186/s41747-023-00360-x
39. Assadi H, Alabed S, Li R et al (2024) Development and validation of AI-derived segmentation of four-chamber cine cardiac magnetic resonance. Eur Radiol Exp 8:77. https://doi.org/10.1186/s41747-024-00477-7
40. Gitto S, Cuocolo R, Emili I et al (2021) Effects of interobserver variability on 2D and 3D CT- and MRI-based texture feature reproducibility of cartilaginous bone tumors. J Digit Imaging 34:820–832. https://doi.org/10.1007/s10278-021-00498-3

Publisher's Note