Received: 5 May 2024 | Revised: 27 June 2024 | Accepted: 3 July 2024

DOI: 10.1002/jor.25938

RESEARCH ARTICLE

Journal of Orthopaedic Research®

Enhanced enchondroma detection from x-ray images using deep learning: A step towards accurate and cost-effective diagnosis

Şafak Aydin Şimşek[1] | Ayhan Aydin[2] | Ferhat Say[1] | Tolgahan Cengiz[3] | Caner Özcan[4] | Mesut Öztürk[5] | Erhan Okay[6] | Korhan Özkan[7]

[1]Department of Orthopedics and Traumatology, Faculty of Medicine, Ondokuz Mayis University, Samsun, Turkey

[2]Department of Computer Engineering, Karabuk University, Karabük, Turkey

[3]Clinic of Orthopedics and Traumatology, Inebolu State Hospital, Kastamonu, Turkey

[4]Department of Software Engineering, Karabuk University, Karabük, Turkey

[5]Department of Radiology, Faculty of Medicine, Samsun University, Samsun, Turkey

[6]Department of Orthopedics and Traumatology, Istanbul Medeniyet University Goztepe Education and Research Hospital, İstanbul, Turkey

[7]Department of Orthopedics and Traumatology, Acıbadem Atasehir Hospital, Istanbul, Turkey

Correspondence
Ferhat Say, Department of Orthopaedics and Traumatology, Ondokuz Mayis University, Samsun 55200, Turkey.
Email: ferhatsay@gmail.com

Funding information
The Scientific and Technological Research Council of Turkey (TUBITAK), Grant/Award Number: 1002 - Priority Support - A with the project code; Scientific and Technological Research Council of Turkey (TUBITAK)

Abstract

This study investigates the automated detection of enchondromas, benign cartilage tumors, from x-ray images using deep learning techniques. Enchondromas pose diagnostic challenges due to their potential for malignant transformation and overlapping radiographic features with other conditions. Leveraging a data set comprising 1645 x-ray images from 1173 patients, a deep-learning model implemented with Detectron2 achieved an accuracy of 0.9899 in detecting enchondromas. The study employed rigorous validation processes and compared its findings with the existing literature, highlighting the superior performance of the deep learning approach. Results indicate the potential of machine learning in improving diagnostic accuracy and reducing healthcare costs associated with advanced imaging modalities. The study underscores the significance of early and accurate detection of enchondromas for effective patient management and suggests avenues for further research in musculoskeletal tumor detection.

KEYWORDS

deep learning, Detectron2, enchondromas, machine learning, x-ray

STATEMENT OF CLINICAL SIGNIFICANCE

1. Diagnostic Challenges of Enchondromas: Enchondromas, benign cartilage tumors commonly found in bones, present diagnostic complexities due to their potential for malignant transformation and overlapping radiographic features with other conditions.
2. Importance of Accurate Detection: Accurate detection of enchondromas is crucial for effective patient management, as misdiagnosis or delayed diagnosis can lead to adverse outcomes, including malignant transformation and unnecessary interventions.
3. Role of Deep Learning in Enhancing Diagnosis: This study demonstrates the utility of deep-learning techniques, specifically Detectron2, in automating the detection of enchondromas from x-ray images. Such advancements in imaging technology hold promise for improving diagnostic accuracy and efficiency.

4. Clinical Implications of Study Findings: The high accuracy achieved by the deep-learning model underscores its potential as a valuable tool for radiologists and orthopedic surgeons in diagnosing enchondromas. Early and accurate detection can facilitate timely intervention and appropriate patient management strategies.
5. Potential for Cost Reduction in Healthcare: By reducing the reliance on advanced imaging modalities and streamlining the diagnostic process, deep learning-based approaches have the potential to mitigate healthcare costs associated with unnecessary imaging studies and interventions.
6. Future Directions in Musculoskeletal Tumor Detection: This study sets the stage for further research in leveraging machine learning techniques for musculoskeletal tumor detection. Future studies could explore the integration of deep-learning algorithms into clinical practice for real-time diagnosis and decision support.

1 | INTRODUCTION

Enchondromas frequently occur in benign cartilage tumors found within the medullary cavity, displaying imaging characteristics that indicate their benign nature. Histologically, they exhibit similarities to low-grade chondrosarcoma and are occasionally categorized as part of the broader spectrum of low-grade chondral series tumors.[1] Enchondromas are commonly found in the long tubular bones, especially in the metaphysis of the humerus and femur, typically situated at the center of the bone's medullary cavity.[2] Enchondroma is the second most common primary bone tumor following osteochondroma, comprising approximately 3%–10% of all bone tumors and 20%–27% of all benign bone tumors with a reported incidence of 1:200,000–1:500,000.[3] Individuals usually discover asymptomatic lesions incidentally on x-rays between the ages of 20 and 50.[4] While most cases of enchondroma are incidentally discovered, the most common signs and symptoms are nonspecific, including pain, swelling, and deformities. In many cases, around 40%–60% of patients exhibit pathological fractures during their initial presentation in most series.[5] Additionally, enchondromas were identified in 3% of routine knee magnetic resonance imaging (MRI) and 2% of routine shoulder MRI studies.[6]

Enchondroma can be diagnosed based on typical radiological appearance and clinical evaluation.[7] The classic calcification pattern on x-ray known as rings and arcs is typical for enchondroma. However, in long bones, these calcifications may be challenging to differentiate from the dystrophic calcifications seen in bone infarction. Based solely on radiographic findings, the enchondroma and low-grade chondrosarcoma cannot be reliably distinguished. Further investigations are warranted when suspicion arises.[8] MRI can distinguish enchondromas from bone infarcts in long bones.

The true rate of malignant transformation in solitary enchondromas is unknown because most enchondromas are asymptomatic and undetected. However, the lifetime risk of secondary chondrosarcoma arising from solitary enchondroma was reported to be as high as 4.2%.[9] Enchondromatosis patients face an estimated 25%–30% risk of malignant transformation, with recent studies suggesting rates as high as 40%.[9,10] Although the risk of malignant transformation is low in cases of solitary enchondroma, the risk of malignant transformation from enchondromas to chondrosarcomas has been reported to be up to 50% in Ollier disease and Maffucci syndromes.[11] Differentiating between enchondroma and low-grade chondrosarcoma presents a significant diagnostic challenge for orthopedic surgeon, radiologists, and pathologists. Therefore, accurate diagnosis and diligent follow-up of these patients are crucial.

The identification and categorization of tumor formations within organ tissues have been subject to thorough investigation through the application of image-processing techniques and deep-learning methodologies.[12] Recent studies have demonstrated the potential of deep learning in the detection, segmentation, and classification of bone tumors. Zhou et al.[13] and Furuo et al.[14] both highlight the use of deep-learning models in these areas, with Zhou discussing its application in detection, segmentation, and prognosis prediction, and Furuo specifically focusing on the automatic estimation of benign and malignant bone tumors. Santhanalakshmi et al.[15] further emphasizes the significance of early and accurate detection, particularly using MRI. These studies collectively underscore the potential of deep learning in improving the accuracy and efficiency of bone tumor diagnosis and classification.

Enchondromas, despite being one of the common primary bone tumors, pose a significant challenge in diagnosis due to the risk of malignant transformation and the potential for confusion with low-grade chondrosarcoma. Therefore, diagnosing patients and closely monitoring those diagnosed with enchondroma is crucial. Our study aims to identify enchondroma lesions in bone tissue using x-ray data by deep learning and ensuring that these lesions are not overlooked.

2 | MATERIALS AND METHODS

2.1 | Power analysis

Statistical power analysis was performed to determine the sample size required to study enchondroma detection with deep-learning approaches. G*Power software was used to perform the power analysis. Results were determined to be within a 0.95 confidence interval ($p < 0.05$). In addition, the paired t-test performed in conjunction with the research was based on a Cohen's d low effect size value.[16] Power values corresponding to the effect size of $d = 0.200$ were calculated for different sample sizes.

Based on these values, a power of approximately 0.994 will be obtained if the research uses 500 observations. The calculated power value is more significant than 0.80, which is statistically sufficient. Therefore, as determined by the power analysis, the research can be conducted with 500 observations.

2.2 | Data set

The images used in the study belong to the archive at three tertiary centers. Patients were identified by scanning the keyword "enchondroma" in the records of three tertiary centers, MR and computerized tomography (CT) reports, and pathology reports. In addition, patients diagnosed with enchondroma by clinical and direct radiography were included in the study. Clinical and radiological enchondroma diagnosis was made by an orthopedist and radiologist working in this field. The images were collected by archive search, randomly selected, and anonymized. Ondokuz Mayıs University Non-Interventional Clinical Research Ethics Committee approved this study with reference number 2022/120.

Before generating the data set, a power analysis was performed to determine the minimum number of samples required for optimal power. According to the analysis results, data were collected well above the minimum data threshold.

The data set consisted of 1645 unique radiographs from 1173 different patients. The area containing the enchondroma on the image was marked polygonally by two orthopedic specialists. An expert then checked this. A number of 1316 images in the data set were randomly selected as training images. To train the model. The training data set was validated using k-fold cross-validation ($k = 5$), 1316 training images, and 329 validation images at each fold.

2.3 | Detectron2

Detectron2 is a new generation library of state-of-the-art detection and segmentation algorithms from Facebook AI Research (FAIR). It is the successor to the Detectron and Mask RCNN benchmarks and can be used in various computer vision research projects and production applications, including those at Facebook. It implements several object recognition algorithms, including Mask R-CNN, RetinaNet, Faster R-CNN, RPN, Fast R-CNN, and TensorMask. The Detectron2 framework shown in Figure 1 is analyzed in several segments. The Data Input Module is designed to load large amounts of data from disk using optimization techniques such as caching and multiworker systems. Users can also quickly implement data augmentation techniques into the module's data loader. In addition, the module is flexible, allowing users to customize and register their own data sets.

The backbone module extracts feature from the supplied images. This is achieved using advanced convolutional neural networks, including ResNet or ResNeXt.[18] The module can be customized to use any standard convolutional neural network adequate for a particular image classification task. It is noteworthy that the module has extensive insights into transfer learning. In this context, we could use pretrained models to deploy a state-of-the-art convolutional neural network that works effectively on large image data sets such as ImageNet. If not, we could opt for simpler networks in this module to improve our efficiency (regarding training and prediction time) without compromising accuracy. The features extracted from the backbone are processed by region proposal, which proposes image regions with location information and scores. The scores indicate whether the regions contain objects with objectness scores. The objectness score of the proposed region can be either 0 or 1. The objectness score is specifically concerned with whether the region contains an object or background rather than the likelihood of it being a class of interest.[19]

3 | RESULTS

3.1 | Evaluation metrics

Among the various annotated data sets used for object detection, the most common metric used by challenges and the scientific community to measure the accuracy of detections is average precision (AP). Within AP there are several variants. First, we will explain the concepts that makeup AP;

True positive: Accurate identification of the ground truth bounding box was achieved.

False positive: A nonexistent object is incorrectly detected.

False negative: An undetected object that should have been included in the ground truth bounding box.[20]

It is important to note that true negative (TN) does not apply to object detection because images contain countless bounding boxes that need not be present. It is necessary to define "true detection" and "false detection" for the definitions presented here. In object detection, the intersection over union (IOU) calculates the overlap between the predicted bounding box (B_p) and the ground truth bounding box (B_{gt}). This is then divided by the area of their union.

$$J(Bp,\ Bgt) = \text{IOU} = \frac{\text{area}(Bp \cap Bgt)}{\text{area}(Bp \cup Bgt)}. \tag{1}$$

Correct and incorrect detections can be determined by comparing their IOU with a predetermined threshold, t. The detection is correct if the IOU is greater than or equal to t. If the IOU is less than t, the detection is incorrect. It is worth noting that TNs are not used in object detection frameworks. Therefore, metrics such as TPR, FPR, and ROC curves based on TNs should not be relied upon.

The evaluation of object detection techniques relies heavily on the principles of precision P and recall R, which are defined as follows:

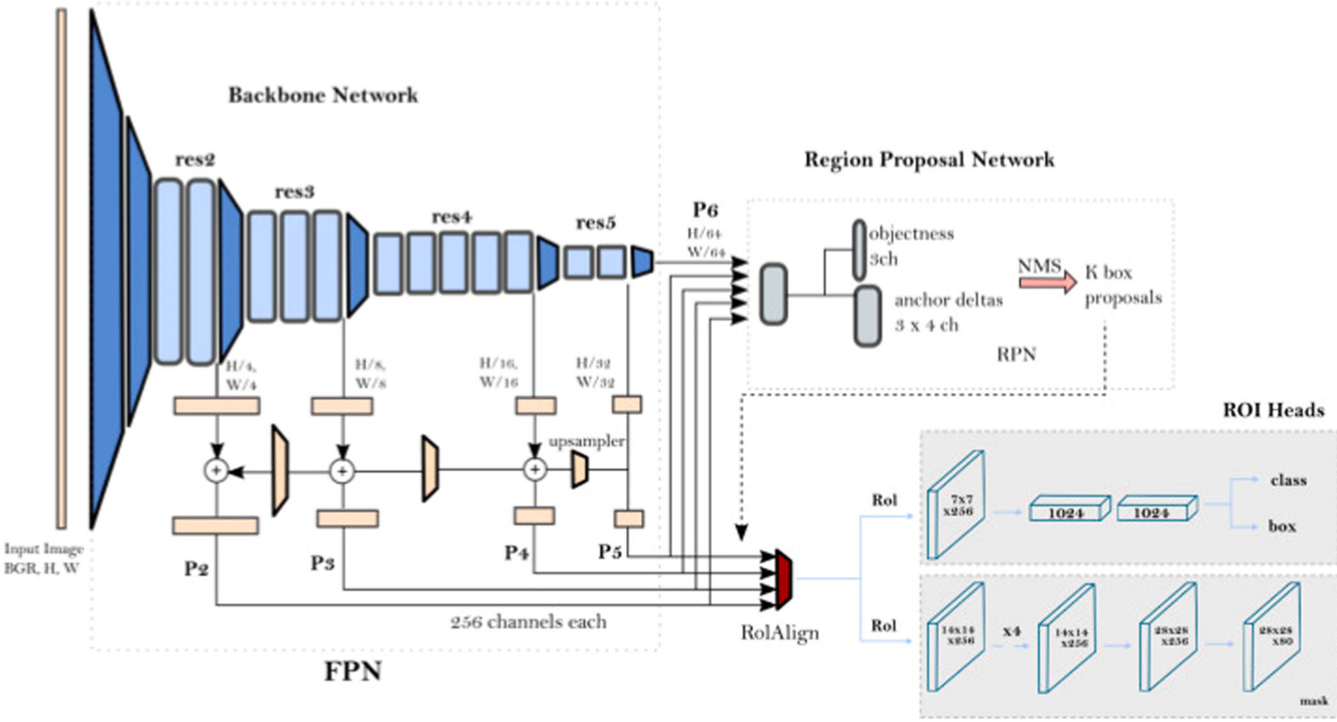

FIGURE 1 Diagram of the Detectron2 architecture. The backbone network provides feature maps (P1–P5) to the region proposal network (RPN). The ROI module identifies (bounding box) and separates (mask) the objects together with their respective classification.[17] ROI, region of interest.

$$P = \frac{TP}{TP + FP} = \frac{TP}{\text{all detections}}, \tag{2}$$

$$R = \frac{TP}{TP + FN} = \frac{TP}{\text{all ground truths}}. \tag{3}$$

Precision is the ability of a model to detect only relevant objects, measured by the proportion of correct positive predictions. Recall is the ability of a model to detect all applicable instances, encompassing all ground truth bounding boxes, measured by the proportion of correct positive predictions among all ground truth cases. The precision–recall trade-off curve shows different levels of confidence in the bounding boxes generated by a detector in a graph where precision and recall are plotted.[21]

In the definition of *AP*, instead of using the precision $P(R)$ observed at recall level R for each different object identification, *AP* is obtained by considering the maximum precision whose recall value is greater than R. And interpolation can be done over these points.

$$mAP = \frac{1}{N}\sum_{i=1}^{N} APi. \tag{4}$$

3.2 | Model results

Data from 1173 patients were converted from DCM to PNG format. The images were then resized to a standard size of 1024 × 1024 pixels. Table 1 shows the Detectron2 (DETR2) parameters used for training.

TABLE 1 Train parameters.

Maximum iteration	2000
Evaluation period	200
Base learning rate	0.001
ROI heads batch size per image	64
Data loader worker	2
Training example per iteration	2

Abbreviation: ROI, region of inerest.

Detectron2 takes 88 min to learn on a Tesla K80 16 GB GPU using the parameters in Table 1.

Figure 2 presents the performance of Fast R-CNN and Mask R-CNN, alongside accuracy curves illustrating the accuracy values attained using Detectron2. Additionally, Figure 3 illustrates the prediction samples of Detectron2. As seen in Figure 3, metal objects and orthopedic implants do not hinder the detection of the model's prediction results.

4 | DISCUSSION

In our study, we found that our deep-learning approach showed good diagnostic performance for enchondroma detection on x-ray. We analyzed cases of enchondromas, including those in long bones,

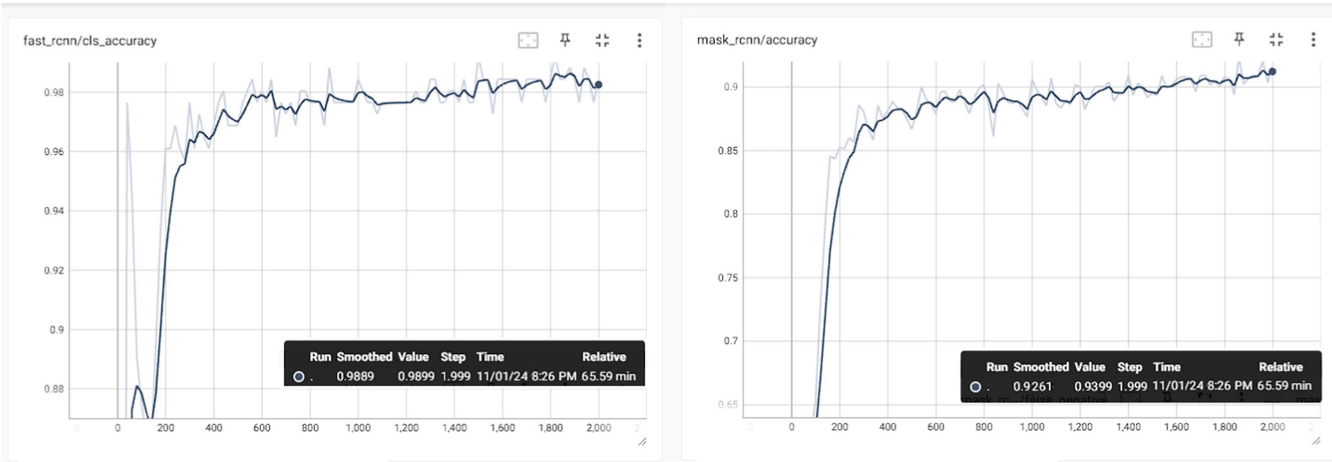

FIGURE 2 DETR2 accuracy graphs. DETR2, Detectron2.

without distinction. The data set comprised x-ray images from 1173 unique patients, totaling 1645 images and our deep-learning model achieved an accuracy of 0.9899.

Radiographic findings of enchondromas are usually diagnostic. The classic appearance of an enchondroma on x-ray is a well-defined lytic lesion containing stippled calcifications. MRI is not always necessary for diagnosis, but it is used to determine the intramedullary extension and soft tissue involvement of suspected enchondromas, distinguishing them from chondrosarcomas and bone infarctions.[22]

Akoh et al.[23] conducted a retrospective study on 55 cases, which they followed for at least 1 year. They compared the costs of plain radiographs versus advanced imaging surveillance during the final follow-up. Their findings revealed a significant cost associated with prereferral imaging, with 42% of the costs attributed to MRI scans. Wilson et al.,[24] in their retrospective review of 121 patients (105 enchondromas and 19 chondrosarcomas), employed decision analysis to determine the number of unnecessary images. They discovered that 85% of enchondromas had at least one unnecessary advanced image, and 58% had two unnecessary images. Similarly, Akoh et al.[23] found that 54.8% of prereferral and 29.2% of postreferral imaging were advanced images. In their series, enchondroma growth was infrequent and typically occurred at 2 years of follow-up. Radiographically, distinguishing between enchondromas and low-grade chondrosarcomas is often challenging. However, certain features such as a size greater than 5 cm, location in the axial skeleton, and ill-defined margins may indicate low-grade chondrosarcomas.[22]

In this context, reducing the need for advanced radiological examinations can effectively decrease healthcare expenses. Specifically, monitoring stable lesions using x-ray imaging and detecting any changes, including minimal size increases, can lead to notable cost reductions.

In our study, we aim to identify enchondroma lesions in bone tissue using x-ray data and detect differences in tumoral lesions using deep learning. The x-ray is the first step of the radiological evaluation of musculoskeletal tumors. It is a cheap test, available in many hospitals, and provides quick and detailed results.

The increase in demand for radiology services has placed a significant burden on the workforce, exacerbated by the creation of one of the largest image data sets ever. In the UK, a staggering 41 million imaging studies were conducted in 2016 alone.[25] In the Netherlands, over 1.7 million CT examinations were conducted across all hospitals in 2016.[26] Moreover, the analysis of medical images can frequently prove challenging and time-consuming. Machine learning appears to be a potential tool to solve this problem. The machine learning approach on musculoskeletal tumors will give a good way for screening. Also, machine learning approach can detect any changes during the follow-up of the bone tumor patients.

Machine learning is a branch of artificial intelligence that uses mathematical algorithms to identify patterns in data and make predictions. Machine learning algorithms are categorized as classical and deep, where in classical machine learning, features of interest are defined by the practitioner, whereas in deep machine learning, after machine learning with input data, new features are automatically computed by the algorithm.[27] Deep learning, a subset of machine learning, employs multilayered neural network algorithms inspired by the structure of the brain to make predictions.[28]

In Guan et al., dilated convolutional feature pyramid network (FPN) was trained for fracture detection with 3842 femur fracture x-ray images, 82.1% AP which is 3.9% higher than that of state-of-the-art FPN values were obtained.[29] He showed us that precision in results can increase with advanced methods in deep learning. Using the entire data set, Ma and Luo conducted fracture detection employing Faster R-CNN, and subsequently applied the proposed CrackNet model for fracture/nonfracture classification, achieving an accuracy rate of 90.11%.[30] Studies on fracture detection have indicated the potential for similar research to be conducted on bone lesions as well.

He et al.[31] demonstrated that by employing convolutional neural networks, a sizable research team achieved the discrimination of three classes (benign, malignant, and intermediate) of bone tumors on radiographs based on histopathological categories, with an accuracy of 0.73. However, the study categorized the data based on various age

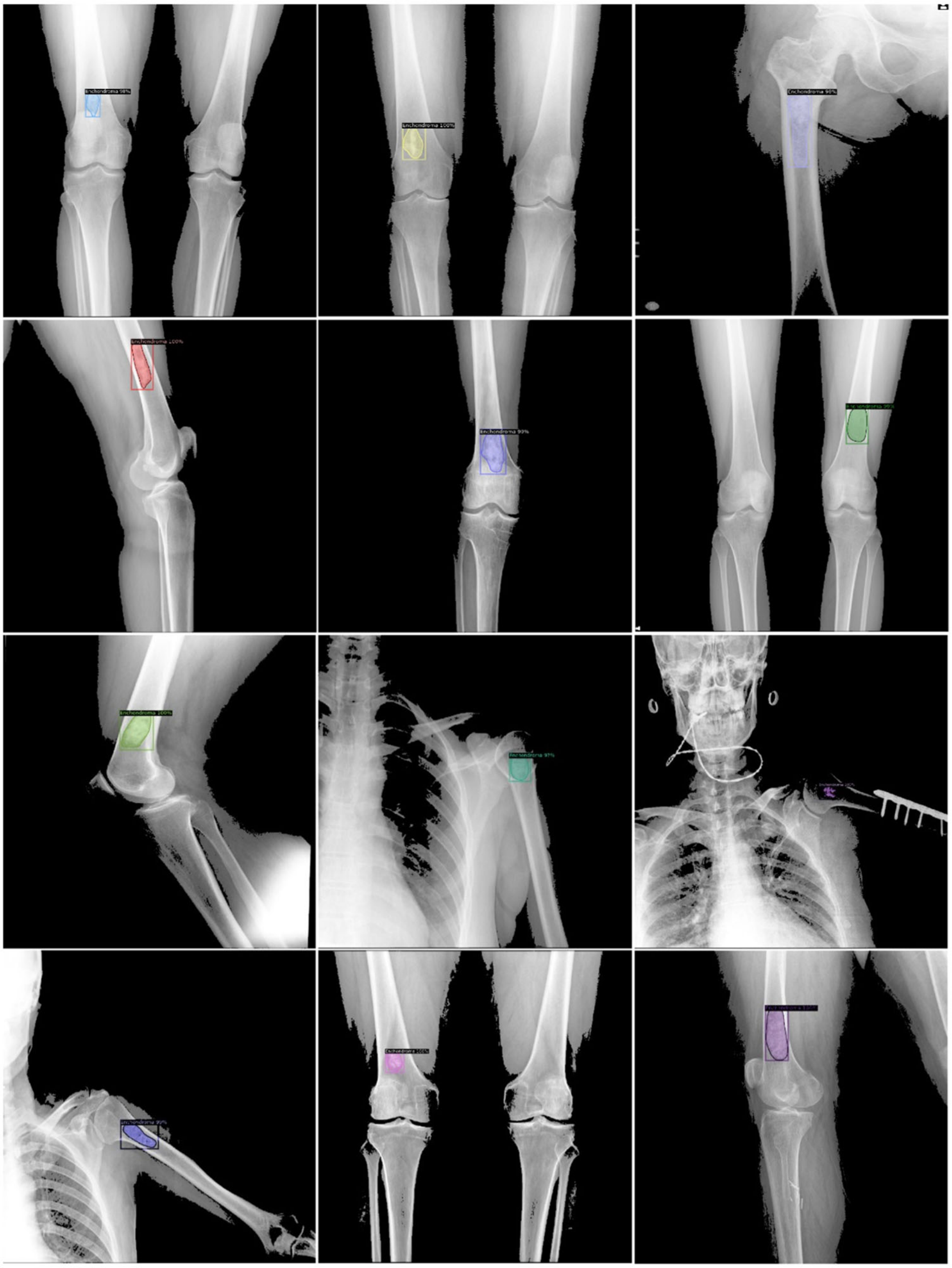

FIGURE 3 Detectron2 prediction samples. DETR2, Detectron2.

groups, and the success of the detection process involved querying 291 data points for different classes. In a study led by Chianca et al.,[32] they analyzed data from 146 patients to classify primary bone tumors of the spine and spinal bone lesions into three categories (benign, malignant, and primary malignant) using radiomic data mining on the Weka platform. Deep artificial neural networks and radiologists' diagnoses were compared, yielding success rates of 0.71 and 0.86, respectively. In this study, low success rates were obtained in both group.[32]

Vezakis et al.[33] utilized publicly accessible osteosarcoma cross-sectional images to assess and contrast the effectiveness of cutting-edge deep neural networks in histopathological assessment of osteosarcomas. Following training with fivefold cross-validation, the MobileNetV2 network demonstrated an overall accuracy of 91%.[33] However, the study is limited by the images from only four patients selected by pathologists, and since different tumor types and stages may result in different images, it does not represent a broad population.

Gitto et al.[34] reported that using a machine learning approach developed based on MR images of 58 patients diagnosed with histologically low-grade and high-grade chondrosarcoma, they achieved good diagnostic performance. However, there was no significant difference in performance between the radiologist and the machine-learning classifier.[34]

In another study by Gitto et al.,[34] 158 patients who underwent surgical treatment and were histologically diagnosed with cartilaginous tumors were evaluated using MRI radiomics-based machine learning. After tuning on the training cohort, the machine-learning classifier demonstrated 92% accuracy in identifying the lesions in the external test cohort. Its accuracies in correctly classifying atypical cartilaginous tumor and grade II chondrosarcoma were 98% and 80%, respectively. The radiologist achieved 98% accuracy with no significant difference compared to the classifier.[34]

Sharma et al.[35] employed x-ray image data consisting of 65 images with cancer and 45 healthy images, totaling 105 images. Initially, they delineated the bone boundaries using edge detectors and subsequently examined cancer structure detection using HOG and support vector machines. Their research achieved an accuracy of 0.92 and precision of 0.93. However, their study diverges from ours due to the limited amount of visual data utilized and the absence of numerous deep-learning methods.[35]

In their study, Anttila et al.[36] aimed to diagnose enchondroma from hand radiographs using a developed deep-learning algorithm. They trained the model with 414 enchondroma radiographs and evaluated its performance using a separate test set of 131 radiographs, of which 47% had an enchondroma. The deep-learning model successfully detected 56 out of 62 enchondromas from the radiographs, achieving an area under the receiver operator curve of 0.95. The F1 score for area statistical overlapping was 69.5%. In this study, detailed information regarding the methods and model used has not been provided, and the analysis was conducted on a limited data set of the hand region. Table 2 provides studies in the literature that detect differences in bone tissue using various methods and their success rates.

When an overview of the detection studies on medical images is made, there are areas for improvement in some points that require method-dependent and specialized expertise. Detection architectures obtained with deep-learning methods achieve much superior success compared to conventional image-processing methods. Moreover, deep-learning models obtained with state-of-the-art architectures and backbone structures are about to reach 100% detection success.[42] Detectron2 object detection platform is one of the most widely used open source projects of FAIR. The platform is implemented in PyTorch. With a new, more modular design, Detectron2 is flexible and extensible and can provide fast training on single or multiple GPU servers. The *k*-Fold cross-validation method mentioned in the method section is a resampling procedure used to evaluate machine learning models. It allows the evaluation of the designed model more than once by using different clusters in the total data set. In this way, the consistency of the model test results is proved.[43]

This study has some limitations. X-ray images of patients diagnosed with enchondroma radiologically or pathologically were used in our research. When the data of patients with chondroid matrix tumors at different stages and patients showing progression in long-term follow-ups are compared, information about differential diagnosis and follow-up through x-ray can be obtained using deep-learning methods.

However, the strengths of this study are that it achieved high accuracy compared to the literature with the method we used and the large amount of data. In the data set we used, the lesions are not located in a specific anatomical location but in different parts of the skeletal system, such as the skull, costa, and long bones. Additionally, Detectron2 deep-learning architecture and the *k*-fold cross-validation method have not been previously used in object detection on medical imaging. The use of the *k*-fold cross-validation method further enhances the reliability of the system. Moreover, a multidisciplinary study was conducted with the involvement of domain experts in the medical field and deep-learning architecture coding.

TABLE 2 Recent research on the detection of objects in radiological images.

References	Year	Aim	Data set	Architecture	Evaluation
Guan et al.[29]	2019	Femur fracture detection	3842	Dilated convolutional feature pyramid network	0.821 precision
He et al.[31]	2020	Bone tumor classification	2899	EfficientNet	0.746 accuracy
Ma and Luo[30]	2021	Bone detection and classification	1052	CrackNet	90.11% accuracy
Sharma et al.[35]	2021	Bone tumors classification	105	Edge detection+SVM	0.92 accuracy
Chianca et al.[32]	2021	Bone tumors classification	146	Weka platform	0.80 sensitivity and 0.75 specificity
Cheng et al.[37]	2021	Bone metastasis detection	576	Yolo4	0.90 accuracy, 0.92 precision
von Schacky et al.[38]	2021	Primary bone tumors classification and segmentation	934	Multitask DL Model	0.80 accuracy, 0.62 sensitivity, 0.88 specificity
Li et al.[39]	2023	Primary bone detection and classification	1430	Yolo Model	0.95 accuracy
Gawade et al.[40]	2023	Bone tumor detection	1144	ResNet101	0.90 accuracy, 0.89 precision
Breden et al.[41]	2023	Pediatric bone tumor detection	366	Grad-CAM	89.1% accuracy, 82.2% sensitivity, 93.2% specificity
Anttila et al.[36]	2023	Bone enchondroma	500	Not specified	0.93 accuracy

Abbreviations: CAM, computer-aided manufacturing; DL, deep learning; SVM, support vector machine.

In conclusion the deep-learning model achieved an accuracy of 0.9899, surpassing the training speed of various anomaly detection methods reported in the literature. These attributes set our study apart from recent research initiatives, demonstrating superior performance. Deep-learning techniques could be employed to assess the risk of recurrence and malignant transformation in lesions demonstrating growth. Restricting MRI imaging for stable lesions may lead to significant cost reductions.

AUTHOR CONTRIBUTIONS

Preprocessing on visual data, data set creation, artificial intelligence coding, architecture optimization, experimental applications, and preparation of publications: Ayan Aydin. *Visual data processing, evaluation and consultancy of deep-learning models, analysis and improvement of model results, and preparation of publications*: Caner Özcan. *Data management, labeling check, comparison of model inferences with confirmed diagnoses, and preparation of publications*: Ferhat Say. *Data collection, anonymization, preliminary image processing, data labeling*: Mesut Öztürk. *Data collection, anonymization, data labeling, and comparison of model inferences with confirmed diagnoses*: Şafak Aydın Şimşek. *Data collection*: Erhan Okay. *Data collection*: Korhan Özkan. *Preparation of publications*: Tolgahan Cengiz. All authors have read and approved the final version of submitted manuscript.

ACKNOWLEDGMENTS

The Scientific and Technological Research Council of Turkey (TUBITAK) supports the study within the scope of 1002–Priority Support–A with the project code 122E636. The executive, researcher, and scholarship holder are the paper's authors.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data used in the study were collected and processed by researchers approved by the Ethics Committee. By the Regulation on the Protection of Personal Health Data, the researchers are prohibited from disclosing the data to third parties. It can be used for scientific studies with permission from relevant institutions. The authors will share the codes produced in the study following the completion of the project. It can be requested through the authors' GitHub pages or e-mail.

ETHICS STATEMENT

As stated in the data set Section 2.2, the study on real patient x-ray images was authorized by Ondokuz Mayıs University Non-Interventional Studies Ethics Committee with reference number 2022/120.

ORCID

Şafak Aydin Şimşek https://orcid.org/0000-0003-2250-8043
Ayhan Aydin http://orcid.org/0000-0001-9127-0951
Ferhat Say http://orcid.org/0000-0002-8021-0942
Tolgahan Cengiz http://orcid.org/0000-0003-2363-0198
Caner Özcan http://orcid.org/0000-0002-2854-4005
Mesut Öztürk http://orcid.org/0000-0003-4059-2656
Erhan Okay http://orcid.org/0000-0003-2443-2505
Korhan Özkan http://orcid.org/0000-0002-3755-1813

REFERENCES

1. Mulligan ME. How to diagnose enchondroma, bone infarct, and chondrosarcoma. *Curr Probl Diagn Radiol*. 2019;48(3):262-273. doi:10.1067/j.cpradiol.2018.04.002
2. Engel H, Herget GW, Füllgraf H, et al. Chondrogenic bone tumors: the importance of imaging characteristics. *RöFo*. 2021;193(3):262-275. doi:10.1055/a-1288-1209
3. Dierselhuis EF, Goulding KA, Stevens M, Jutte PC. Intralesional treatment versus wide resection for central low-grade chondrosarcoma of the long bones. *Cochrane Database Syst Rev*. 2019;3(3):010778. doi:10.1002/14651858.CD010778.pub2
4. Trăistaru MR, Kamal D, Ciurea RN, et al. Hand enchondroma: complete evaluation and rehabilitation. *Rom J Morphol Embryol*. 2018;59(4):1083-1089.
5. Tang C, Chan M, Fok M, Fung B. Current management of hand enchondroma: a review. *J Hand Surg Asian Pac Vol*. 2015;20(1):191-195. doi:10.1142/s0218810415300028
6. Afonso P, Isaac A, Villagrán J. Chondroid tumors as incidental findings and differential diagnosis between enchondromas and low-grade chondrosarcomas. *Semin Musculoskelet Radiol*. 2019;23(1):003-018. doi:10.1055/s-0038-1675550
7. Lubahn JD, Bachoura A. Enchondroma of the hand: evaluation and management. *J Am Acad Orthop Surg*. 2016;24(9):625-633. doi:10.5435/jaaos-d-15-00452
8. Douis H, Parry M, Vaiyapuri S, Davies AM. What are the differentiating clinical and MRI-features of enchondromas from low-grade chondrosarcomas? *Eur Radiol*. 2018;28(1):398-409. doi:10.1007/s00330-017-4947-0
9. Herget GW, Strohm P, Rottenburger C, et al. Insights into enchondroma, enchondromatosis and the risk of secondary chondrosarcoma: review of the literature with an emphasis on the clinical behaviour, radiology, malignant transformation and the follow up. *Neoplasma*. 2014;61(4):365-378. doi:10.4149/neo_2014_046
10. Verdegaal SH, Bovée JV, Pansuriya TC, et al. Incidence, predictive factors, and prognosis of chondrosarcoma in patients with Ollier disease and Maffucci syndrome: an international multicenter study of 161 patients. *Oncologist*. 2011;16(12):1771-1779. doi:10.1634/theoncologist.2011-0200
11. Wu J, Huang Y, Yu C, et al. The key gene expression patterns and prognostic factors in malignant transformation from enchondroma to chondrosarcoma. *Front Oncol*. 2021;11:693034. doi:10.3389/fonc.2021.693034
12. Maurya S, Tiwari S, Mothukuri MC, Tangeda CM, Nandigam RNS, Addagiri DC. A review on recent developments in cancer detection using machine learning and deep learning models. *Biomed Signal Process Control*. 2023;80:104398. doi:10.1016/j.bspc.2022.104398
13. Zhou X, Wang H, Feng C, et al. Emerging applications of deep learning in bone tumors: current advances and challenges. *Front Oncol*. 2022;12:908873.
14. Furuo K, Morita K, Hagi T, Nakamura T, Wakabayashi T, eds. Automatic benign and malignant estimation of bone tumors using deep learning. *2021 5th IEEE International Conference on Cybernetics (CYBCONF)*. IEEE; 2021.
15. Santhanalakshmi S, Abinaya T, Affinasel T, Dimple P. Deep learning based bone tumor detection with real time datasets. *Int Res J Eng Technol*. 2020;7(3):2391-2394.

16. Faul F, Erdfelder E, Lang A-G, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39(2):175-191.
17. Yuxin Wu AK, Massa F, Lo W-Y, Girshick R. Detectron2. 2019.
18. Khan A, Sohail A, Zahoora U, Qureshi AS. A survey of the recent architectures of deep convolutional neural networks. *Artif Intell Rev*. 2020;53:5455-5516.
19. Shangguan Z, Rostami M. Improved region proposal network for enhanced few-shot object detection. *arXiv Preprint*. Published online August 15, 2023.
20. Padilla R, Netto SL, Da Silva EA, eds. A survey on performance metrics for object-detection algorithms In: *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*. IEEE; 2020.
21. Lin T-Y, Goyal P, Girshick R, He K, Dollár P, eds. Focal loss for dense object detection In: *Proceedings of the IEEE International Conference on Computer Vision*. IEEE; 2017.
22. Lubahn JD, Bachoura A. Enchondroma of the hand: evaluation and management. *J Am Acad Orthop Surg*. 2016;24(9):625-633. doi:10.5435/jaaos-d-15-00452
23. Akoh CC, Craig E, Troester AM, Miller BJ. Radiographic enchondroma surveillance: assessing clinical outcomes and costs effectiveness. *Iowa Orthop J*. 2019;39(1):185-193.
24. Wilson RJ, Zumsteg JW, Hartley KA, et al. Overutilization and cost of advanced imaging for long-bone cartilaginous lesions. *Ann Surg Oncol*. 2015;22(11):3466-3473. doi:10.1245/s10434-014-4325-y
25. Kim DH, MacKinnon T. Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. *Clin Radiol*. 2018;73(5):439-445. doi:10.1016/j.crad.2017.11.015
26. Kalmet PHS, Sanduleanu S, Primakov S, et al. Deep learning in fracture detection: a narrative review. *Acta Orthop*. 2020;91(2):215-220. doi:10.1080/17453674.2019.1711323
27. Hirschmann A, Cyriac J, Stieltjes B, Kober T, Richiardi J, Omoumi P. Artificial intelligence in musculoskeletal imaging: review of current literature, challenges, and trends. *Semin Musculoskelet Radiol*. 2019;23(3):304-311. doi:10.1055/s-0039-1684024
28. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444. doi:10.1038/nature14539
29. Guan B, Yao J, Zhang G, Wang X. Thigh fracture detection using deep learning method based on new dilated convolutional feature pyramid network. *Pattern Recognit Lett*. 2019;125(C):521-526. doi:10.1016/j.patrec.2019.06.015
30. Ma Y, Luo Y. Bone fracture detection through the two-stage system of crack-sensitive convolutional neural network. *Int J Med Inform*. 2021;22:100452. doi:10.1016/j.imu.2020.100452
31. He Y, Pan I, Bao B, et al. Deep learning-based classification of primary bone tumors on radiographs: a preliminary study. *EBioMedicine*. 2020;62:103121. doi:10.1016/j.ebiom.2020.103121
32. Chianca V, Cuocolo R, Gitto S, et al. Radiomic machine learning classifiers in spine bone tumors: a multi-software, multi-scanner study. *Eur J Radiol*. 2021;137:109586. doi:10.1016/j.ejrad.2021.109586
33. Vezakis IA, Lambrou GI, Matsopoulos GK. Deep learning approaches to osteosarcoma diagnosis and classification: a comparative methodological approach. *Cancers*. 2023;15(8):2290. doi:10.3390/cancers15082290
34. Gitto S, Cuocolo R, Albano D, et al. MRI radiomics-based machine-learning classification of bone chondrosarcoma. *Eur J Radiol*. 2020;128:109043. doi:10.1016/j.ejrad.2020.109043
35. Sharma A, Yadav DP, Garg H, Kumar M, Sharma B, Koundal D. Bone cancer detection using feature extraction based machine learning model. *Comput Math Methods Med*. 2021;2021:1-13. doi:10.1155/2021/7433186
36. Anttila TT, Aspinen S, Pierides G, Haapamäki V, Laitinen MK, Ryhänen J. Enchondroma detection from hand radiographs with an interactive deep learning segmentation tool: a feasibility study. *J Clin Med*. 2023;12(22):7129. doi:10.3390/jcm12227129
37. Cheng DC, Hsieh TC, Yen KY, Kao CH. Lesion-based bone metastasis detection in chest bone scintigraphy images of prostate cancer patients using pre-train, negative mining, and deep learning. *Diagnostics*. 2021;11(3):518. doi:10.3390/diagnostics11030518
38. von Schacky CE, Wilhelm NJ, Schäfer VS, et al. Multitask deep learning for segmentation and classification of primary bone tumors on radiographs. *Radiology*. 2021;301(2):398-406. doi:10.1148/radiol.2021204531
39. Li J, Li S, Li X, et al. Primary bone tumor detection and classification in full-field bone radiographs via YOLO deep learning model. *Eur Radiol*. 2023;33(6):4237-4248. doi:10.1007/s00330-022-09289-y
40. Gawade S, Bhansali A, Patil K, Shaikh D. Application of the convolutional neural networks and supervised deep-learning methods for osteosarcoma bone cancer detection. *Healthcare Anal*. 2023;3:100153. doi:10.1016/j.health.2023.100153
41. Breden S, Hinterwimmer F, Consalvo S, et al. Deep learning-based detection of bone tumors around the knee in x-rays of children. *J Clin Med*. 2023;12(18):5960. doi:10.3390/jcm12185960
42. Merz G, Liu Y, Burke CJ, et al. Detection, instance segmentation, and classification for astronomical surveys with deep learning (DEEPDISC): DETECTRON2 implementation and demonstration with Hyper Suprime-Cam data. *Mon Not R Astron Soc*. 2023;526(1):1122-1137.
43. Rodriguez JD, Perez A, Lozano JA. Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE Trans Pattern Anal Mach Intell*. 2010;32(3):569-575.

How to cite this article: Aydin Şimşek Ş, Aydin A, Say F, et al. Enhanced enchondroma detection from x-ray images using deep learning: a step towards accurate and cost-effective diagnosis. *J Orthop Res*. 2024;42:2826-2834. doi:10.1002/jor.25938