Two-Phase Deep Learning Algorithm for Detection and Differentiation of Ewing Sarcoma and Acute Osteomyelitis in Paediatric Radiographs

SARAH CONSALVO[1], FLORIAN HINTERWIMMER[1,2], JAN NEUMANN[3],
MARC STEINBORN[4], MAYA SALZMANN[1], FRITZ SEIDL[5], ULRICH LENZE[1],
CAROLIN KNEBEL[1], DANIEL RUECKERT[2] and RAINER H.H. BURGKART[1]

[1]Department of Orthopaedics and Sports Orthopaedic, Klinikum Rechts der Isar,
Technical University of Munich, Munich, Germany;
[2]Institute for AI and Informatics in Medicine, Technical University of Munich, Munich, Germany;
[3]Department of Diagnostic and Interventional Radiology, Klinikum Rechts der Isar,
Technical University of Munich, Munich, Germany;
[4]Institute for Diagnostic and Interventional Radiology and Paediatric Radiology,
Klinikum Schwabing, Munich, Germany;
[5]Department of Trauma Surgery, Klinikum Rechts der Isar, Technical University of Munich, Munich, Germany

Abstract. *Background/Aim: Ewing sarcoma is a highly malignant tumour predominantly found in children. The radiological signs of this malignancy can be mistaken for acute osteomyelitis. These entities require profoundly different treatments and result in completely different prognoses. The purpose of this study was to develop an artificial intelligence algorithm, which can determine imaging features in a common radiograph to distinguish osteomyelitis from Ewing sarcoma. Materials and Methods: A total of 182 radiographs from our Sarcoma Centre (118 healthy, 44 Ewing, 20 osteomyelitis) from 58 different paediatric (≤18 years) patients were collected. All localisations were taken into consideration. Cases of acute, acute on chronic osteomyelitis and intraosseous Ewing sarcoma were included. Chronic osteomyelitis, extra-skeletal Ewing sarcoma, malignant small cell tumour and soft tissue-based primitive neuroectodermal tumours were excluded. The algorithm development was split into two phases and two different classifiers were built and* combined with a Transfer Learning approach to cope with the very limited amount of data. In phase 1, pathological findings were differentiated from healthy findings. In phase 2, osteomyelitis was distinguished from Ewing sarcoma. Data augmentation and median frequency balancing were implemented. A data split of 70%, 15%, 15% for training, validation and hold-out testing was applied, respectively. Results: The algorithm achieved an accuracy of 94.4% on validation and 90.6% on test data in phase 1. In phase 2, an accuracy of 90.3% on validation and 86.7% on test data was achieved. Grad-CAM results revealed regions, which were significant for the algorithms decision making. Conclusion: Our AI algorithm can become a valuable support for any physician involved in treating musculoskeletal lesions to support the diagnostic process of detection and differentiation of osteomyelitis from Ewing sarcoma. Through a Transfer Learning approach, the algorithm was able to cope with very limited data. However, a systematic and structured data acquisition is necessary to further develop the algorithm and increase results to clinical relevance.*

Correspondence to: Sarah Consalvo, Ismaninger Straße 22, 81675 Munich, Germany. Tel: +49 8941408012, e-mail: sarah.consalvo@mri.tum.de

Ewing sarcomas (ES) represent 7-10% of all bone malignancies and have the second highest incidence after osteosarcomas (1). The main differential diagnoses of Ewing sarcoma are acute osteomyelitis (OM) and Langerhans Histiocytosis. Acute osteomyelitis is a severe bone infection which most often has a haematogenous origin (2). Other causes can be trauma, surgery, or contiguously infected soft tissue. It occurs in 8 out of 100,000 children per year in high-income countries, yet it is extremely common in developing countries as well. Male

children are affected twice as often as female children (3). Clinical and laboratory exams might be normal. Blood cultures and biopsy samples are positive for bacteria in only 32-62% and 40-60%, respectively. *Staphylococcus aureus*, *β-haemolytic Streptococcus*, *Streptococcus pneumomiae*, *Escherichia coli* and *Pseudomonas aeruginosa* are the most common bacteria involved in this acute bone infection (4). The symptoms include pain, ROM (Range of Motion) limitations and fever (5). After all, with proper treatment, the outcome for OM is usually good. Conservative treatment with antibiotics is effective in 90% of the early diagnosed paediatric cases (5, 6).

However, Ewing sarcoma is a highly malignant blue round cell tumour, 90% of whose cases occur in patients between age 5 to 25. Worldwide, 2.9 out of 1,000,000 children per year are affected by this malignancy, with a slightly higher incidence in male patients (1.5 male: 1 female) (7). Children usually present with load-independent local pain and ROM limitation without a history of trauma, lasting for at least four to six weeks. Ewing sarcoma treatment begins and ends with chemotherapy. Surgery to remove the cancer is normally performed after neoadjuvant chemotherapy.

Taking into consideration the completely contrasting course of these two diseases, early diagnosis and referral to a specialised centre is crucial for a successful treatment. However differential diagnosis is extremely difficult.

Radiographs and MR images have a relatively low diagnostic value in this crucial differential diagnosis (8, 9), if not interpreted by a trained and experienced musculoskeletal radiologist.

In brief summary, the symptoms, blood screening, as well as the localisation (10) are extremely similar in both diseases. The first radiological exam to conduct a differential diagnosis apart from an ultrasound will be an X-ray. Even with this imaging modality, the diagnosis will not be clear. Although methods of nuclear medicine such as PET and SPECT are currently the most accurate techniques, they are too elaborate to be used in the phase of differential diagnosis and they are usually not available for outpatient clinics (11, 12).

In radiographs, both entities can present bone destruction and periosteal reaction. The typical periosteal reaction associated with Ewing sarcoma – lamellated, "onion skin" – or "Codman's triangle" can also be present in acute osteomyelitis due to a subperiosteal abscess (4). Instead, MR T1-weighted images in comparison with short tau inversion recovery (STIR) showing sharp margins are one of the most significant signs of Ewing sarcoma for the differentiation from osteomyelitis (13). Hence, MRI, PET and SPECT are complex techniques that are indicated when a solid suspicion is provided or when the diagnosis is to be validated. The resemblance of the radiological features as well as the clinical course makes it demanding to distinguish these two entities.

According to Bacci *et al*. (14), the overall delay between initial symptoms and biopsy for Ewing sarcoma is approximately four months. If we consider that the estimated five-year survival for Ewing sarcoma patients shifts from 50-70% in early diagnosed localised cases to 18-30% in metastatic cases (15) and that unfortunately, 25% of all Ewing sarcoma patients have a metastatic disease at the time of diagnosis (16), four months "until" or "since" the first diagnosis make a huge difference in the prognosis of these young patients. To shorten the delay of referral to a specialised centre, it is crucial to improve the ability of outpatient clinics to address a suspicious case. In this process, radiographs represent the first obligatory step. In order to prevent delays and limitation of the prognosis, it is decisive to develop a new form of assistance which can support precision and accuracy of the diagnostic process.

Image interpretation as a part of precision medicine will play an increasingly important role in the future of orthopaedic oncology and novel, more comprehensive and specific analysis tools are urgently needed, especially for outpatient clinics with limited experience and resources for detection and interpretation of rare bone malignancies. Deep learning (DL) represents a subset of Machine Learning and a distinct application of artificial intelligence (AI), which evolved from pattern recognition and learning theory. While complex data analysis of cancerous tissue by AI models and imaging data is already widely applied in some medical specialties (*e.g.* lung and breast cancer), the application of these methods in orthopaedic oncology is still very limited (17). The fact that globally no far-reaching structures for systematic data acquisition have yet been established and that sarcomas are very rare and heterogeneous entities makes modern AI applications, for which a sufficient and qualitative amount of data is crucial, considerably more difficult. While this is a common obstacle – particularly in medicine – several techniques to cope with limited data have emerged. One popular technique is called data augmentation (18), in which new data is created artificially by applying minor transformations to initial data. Another even more powerful method is Transfer Learning (19), where a model is developed for a source task and then reused as a starting point for the target task (Figure 1).

The focus of this study was to develop a real-time support tool for the detection and distinction between Ewing sarcoma and acute osteomyelitis using a two-phase DL algorithm.

Materials and Methods

Data and ethics approval. The local institutional review and ethics board (Klinikum rechts der Isar, Technical University of Munich) approved this retrospective study (N°48/20S). The study was performed in accordance with national and international guidelines. The study is a purely retrospective study in which all data are already available and are collected in pseudonymised form with the help of the musculoskeletal tumour database or by studying files. To increase the quality of the presented observational study and its prediction model, reporting was derived from the Transparent
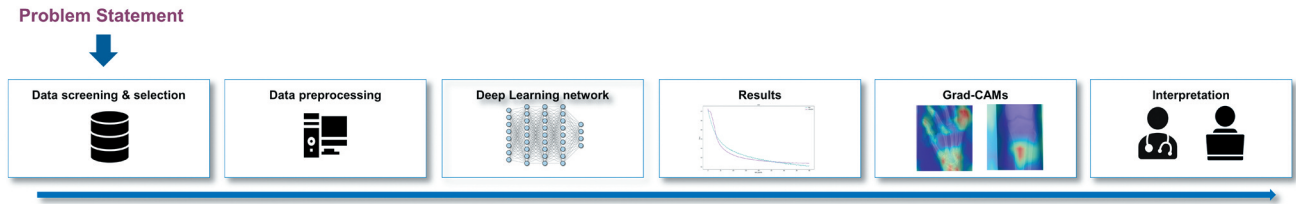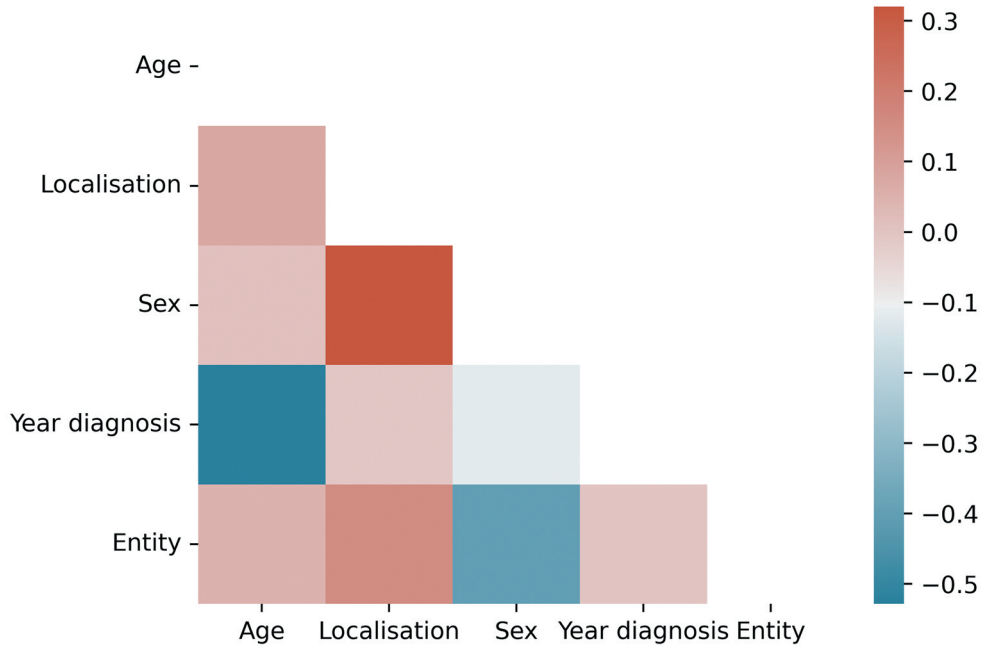
Figure 1. *Study work flow.*

Figure 2. *Correlation matrix: a matrix describing the correlation between age, localization, sex, year of diagnosis and entity.*

Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) guidelines (20) and the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement (21).

Eligibility criteria. All patients from our database with the according ICD-10 code for OM and ES were selected. For all patients, the diagnoses were validated through a histopathological examination as reference standard. The data was retrieved from our hospital information system (HIS) and the picture archiving and communications system (PACS).

The following inclusion criteria were applied:
- patients younger than or equal to 18 years;
- intraosseous Ewing sarcoma;
- histopathologically confirmed cases of acute osteomyelitis or acute on chronic osteomyelitis;
- images prior to treatment.

Patients older than 18 years, chronic osteomyelitis, extraosseous Ewing sarcoma, malignant small cell tumour, soft tissue-based primitive neuroectodermal tumours (PNET) cases were excluded.

Statistical analysis. For statistical analysis and evaluation, accuracy, sensitivity, and specificity were computed for each phase, cross-validated and interpreted by an orthopaedic surgeon (S.C.) and a computer scientist (F.H.). The metrics were implemented using the scikit-learn library (https://scikit-learn.org/stable/modules/model_evaluation.html).

Considering that the control group was selected, only the patients with acute osteomyelitis and Ewing sarcoma were included in the statistical analysis. Nevertheless, a control group was needed to develop an algorithm for detection of pathological cases in the first place.

Except for the 'Localisation', none of the patient meta data is normally distributed according to normality test by D'Agostino-Pearson. Figure 2 shows a correlation matrix according to values of Spearman's rank-order correlation coefficient, which is a measure for linear correlation between two datasets and does not assume that both datasets are normally distributed. Only 'Age'/'Year' of diagnosis' and 'Sex'/'Entity' show a slight indirect correlation ($|\varrho|$ >0.4). It is to be expected with small datasets that no high and stable correlations can be found.
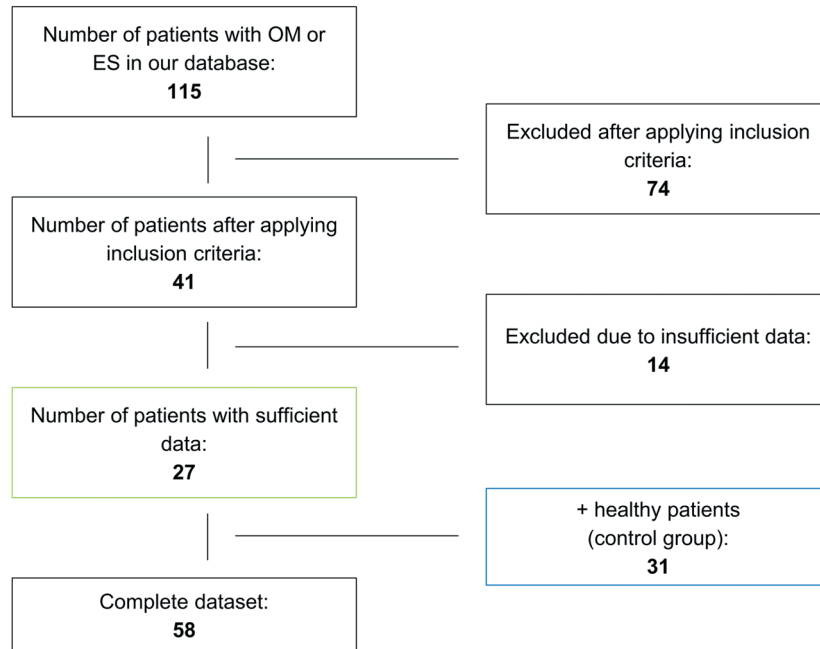
Figure 3. *Flow chart: description of patient selection according to eligibility criteria.*

Model training. Model training and inference was conducted on a DGX Station A100 with four 80GB graphical processing units (Nvidia Corporation, Santa Clara, CA, USA), 64 2.25 GHz cores and 512 GB DDR4 system memory running on a Linux/Ubuntu 20.04 distribution (Canonical, London, England). Preprocessing and model implementation were performed in Python 3.9.6 (https://www.python.org/) using PyTorch 1.9.0 and cuda toolkit 11.1 (https://pytorch.org/).

The source code for this study is provided on GitHub (https://github.com/FlorianH3000/ewing).

A supervised DL algorithm for image classification in two phases was developed: phase 1 for detection of pathological cases and phase 2 for differentiation of ES and OM cases. For both phases, a ResNet (22) was selected. Beforehand, the model was pretrained on 42.608 sarcoma related X-ray images for Transfer learning. For phase 1 and 2 a ResNet18 architecture was chosen. To tackle the overall limited amount of data and integrate regularization, extensive data augmentation was implemented to artificially create more input data during training. In order to manage the class imbalances in both phases, median frequency balancing was utilized to weight the loss of classes accordingly and support the robustness of the algorithm (23). A data split of 70%, 15%, 15% was applied for training, validation, and testing, respectively. Since up to four images from single patients were collected, the data split was applied to patients in order to avoid cross-contamination and therefore provide a higher statistical significance. An additional 6-fold cross validation supported this task, while random chosen hold-out test data for final evaluation remained untouched.

Plausibility. To provide plausibility and more insight into the AI model, Grad-CAMs were implemented in the final inference step (24). Grad-CAMs utilize the gradient information from the last convolutional layer of a deep learning network to understand specific neurons and their impact for decision-making. The result is a coloured heat map, which is co-registered to the original input image and indicates where the algorithm found relevant information for the task at hand. This technique was applied to get a better understanding where the algorithm detects relevant information. To provide a higher statistical significance, the Grad-CAM results were averaged from the 6-fold cross validation.

Results

Dataset. A total of 115 patients treated in our institution for OM or ES between 2000 and 2021 were retrospectively reviewed. After applying the inclusion criteria, 74 cases were excluded, and 41 cases remained. After screening the data, another 14 cases were excluded due to insufficient or invalid data. Ultimately, 27 cases, 9 with acute osteomyelitis and 18 with Ewing sarcoma were collected.

Additionally, 31 healthy cases were included in order to balance the dataset and create a "control group". These patients were treated in our emergency room with a history of acute trauma of a joint. The performed X-ray could exclude any kind of fracture or bone anomalies so that these cases were diagnosed as bruises or contusions. Consequently, a "healthy group" without exposing children to X-ray radiation for our study was obtained. The control group was chosen with similar localisation to our "pathological group". Overall, 182 radiographs (healthy 118, 44 Ewing, 20

Table I. *Distribution of Ewing sarcoma (ES) and osteomyelitis (OM) dataset according to patient characteristics (sex and localisation).*

	Entity			Sex			Localisation	
		#	%		#	%		#
Pathological cases (ES & OM)	Acute osteomyelitis	9	33.3%	Female	5	18.5%	Upper extr.	2
				Male	4	14.8%	Lower extr.	6
							Other	1
	Ewing sarcoma	18	66.7%	Female	3	11.1%	Upper extr.	6
				Male	15	55.6%	Lower extr.	9
							Other	3
	Total	27	100.0%	Female	8	29.6%	Upper extr.	8
				Male	19	70.4%	Lower extr.	15
							Other	4
Control group (healthy)	Total	31	100.0%	Female	15	48.4%	Upper extr.	11
				Male	16	51.6%	Lower extr.	20
							Other	0
Complete dataset	Pathological cases (relative to whole dataset)	27	46.6%	Female	8	13.8%	Upper extr.	8
				Male	19	32.8%	Lower extr.	15
							Other	4
	Control group (relative to whole dataset)	31	53.4%	Female	15	25.9%	Upper extr.	11
				Male	16	27.6%	Lower extr.	20
							Other	0
	Total	58	100.0%	Female	23	39.7%	Upper extr.	19
				Male	35	60.3%	Lower extr.	35
							Other	4

Table II. *Age distribution of involved patients classified in Ewing sarcoma (ES) group and osteomyelitis (OM) group: pathological cases, control group and complete dataset.*

	Entity			Age		
		#	%	Average	Variance	Standard deviation
Pathological cases (ES & OM)	Acute osteomyelitis	9	33.3%	13.6	5.6	2.4
	Ewing sarcoma	18	66.7%	12.8	20.5	4.5
	Total	27	100.0%	13.0	15.7	4.0
Control group (healthy)	Total	31	100.0%	6.4	16.6	4.1
Complete dataset	Pathological cases	27	46.6%	13.0	15.7	4.0
	Control group	31	53.4%	6.4	16.6	4.1
	Total	58	100.0%	9.5	27.1	5.2

osteomyelitis) from 58 patients were collected including data from external imaging data (Figure 3).

Patient characteristics. The dataset including the healthy control group, Ewing sarcoma and acute osteomyelitis consists of 23 females (39.7%) and 35 males (60.3%). While 19 (32.8%) of the patients were affected at their upper extremities, 35 (60.3%) were affected at their lower extremities and 4 (6.9%)

at other localisations. The average age of patients at the time of the initial diagnosis resulted in 9.5 years with a variance of 27.6 and a standard deviation of 5.2 (Table I and Table II).

Model performance in phase 1. All results were cross-validated. The first two-entity classification of the healthy control group and the pathological group resulted in an accuracy of 94.4%/90.6%, sensitivity of 90.0%/89.4% and
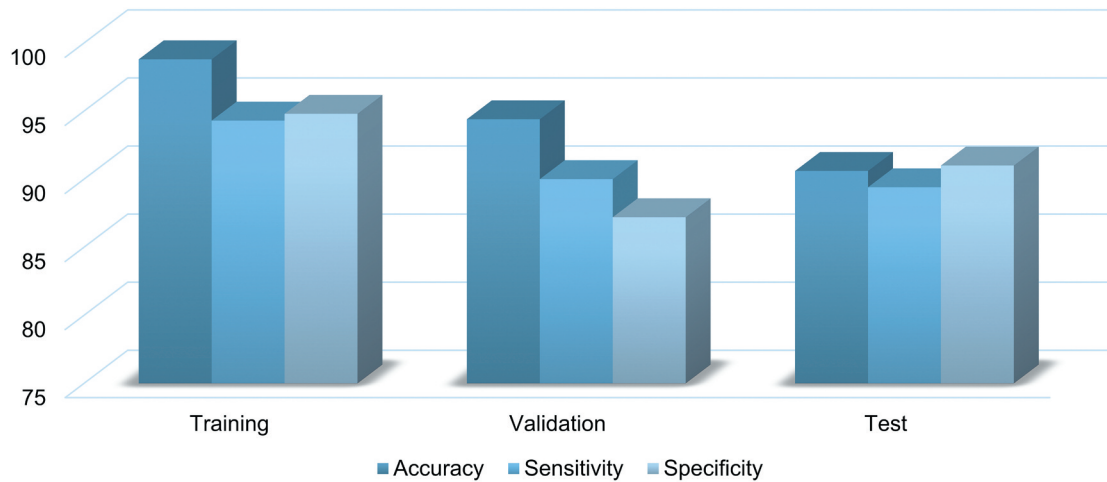
Performance phase 1

Figure 4. *Prediction of performance in Phase 1.*
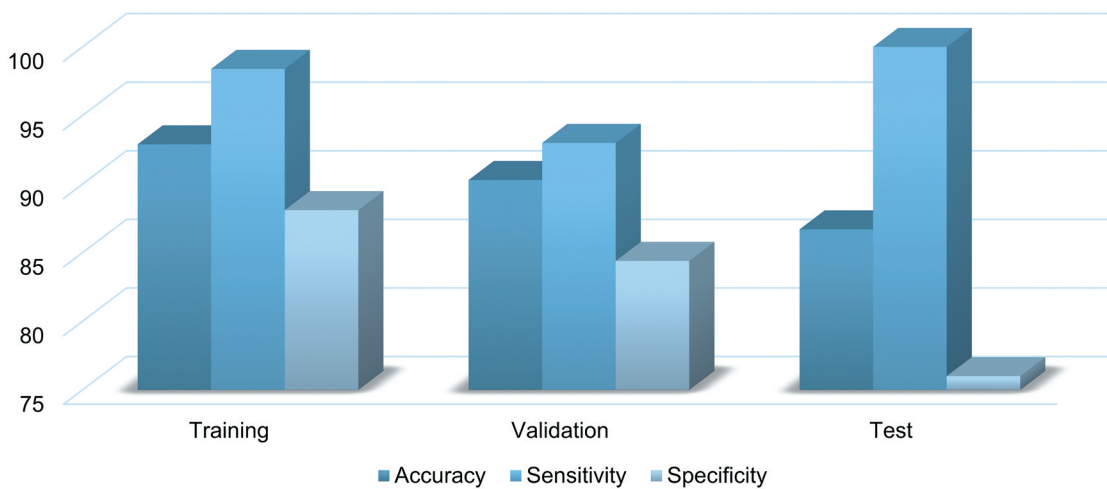
Performance phase 2

Figure 5. *Prediction of performance in Phase 2.*

specificity of 87.2%/91.0% for the validation and test split, respectively (Figure 4).

Model performance in phase 2. All results were cross-validated. The second two-entity classification of OM and ES cases resulted in an accuracy of 90.3%/86.7%, sensitivity of 93.0%/100.0% and specificity of 84.4%/76.0% for the validation and test dataset, respectively (Figure 5).

Grad-CAM results. Figure 6 and Figure 7 display the results of Grad-CAM visualizations from the test dataset of each

entity. The displayed Figures show that the algorithm did in fact find relevant information in very similar areas where a trained radiologist or an orthopaedic surgeon would look at when diagnosing a patient based on a radiograph.

Discussion

The most important finding of this study is that even with a very limited amount of data, good results in detecting and distinguishing Ewing sarcoma from acute osteomyelitis can be achieved through data augmentation and particularly
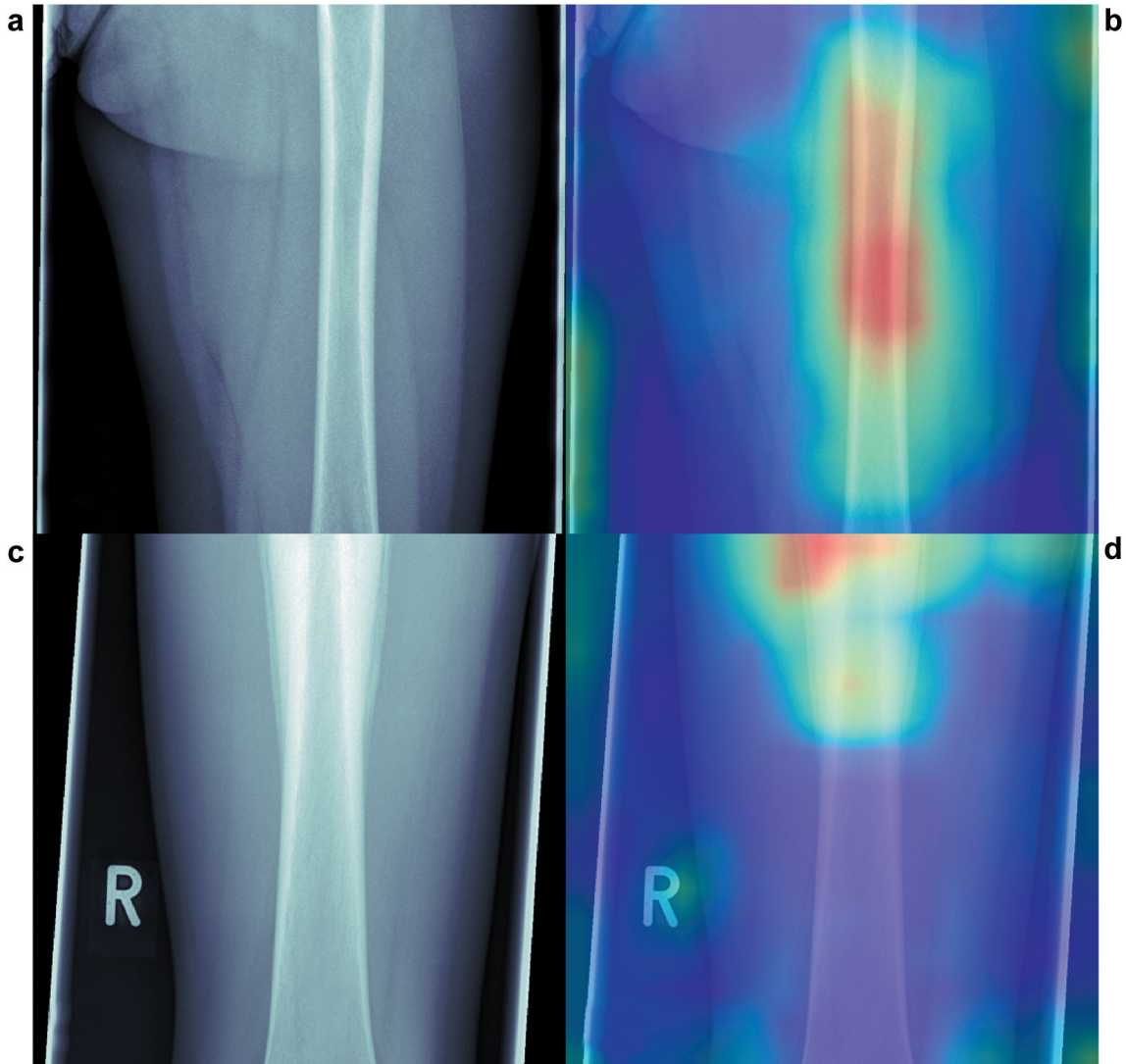
Figure 6. *Grad-CAM of healthy (a/b) and pathological cases (c/d) in phase 1: Grad-CAM results displaying that the algorithm focused on pixels similar to the areas a radiologist or orthopaedic surgeon would look at.*

Transfer Learning. Nevertheless, to further increase the results, a systematic and structured data acquisition is necessary to gather sufficient data and improve the overall accuracy.

Limitations. The main limitation of studying these entities is the extreme rarity of Ewing sarcoma. This makes it very challenging to acquire sufficient imaging data that could enhance the accuracy and stability of the algorithm. Additionally, in most centres data infrastructures are not yet fully adapted to the needs of modern AI applications. Current HIS and PACS systems were often initially set up years ago and were not designed to retrieve data for AI research. Thus,

a considerable amount of data was lost over the years (14 patients excluded due to insufficient data).

While several precautions to provide statistical significance were applied – such as cross validation, loss weighting or incorporating Transfer Learning via pretrained networks - limited amount of data for final validation and testing might still bias the accuracy of the algorithm compared to real-world scenarios. However, this issue can most likely be addressed with further establishment of collaboration of specialised centres, the according data infrastructure, and therefore more sufficient datasets.

Another limitation of this study is that the DL model did not use demographics or other important patient characteristics as
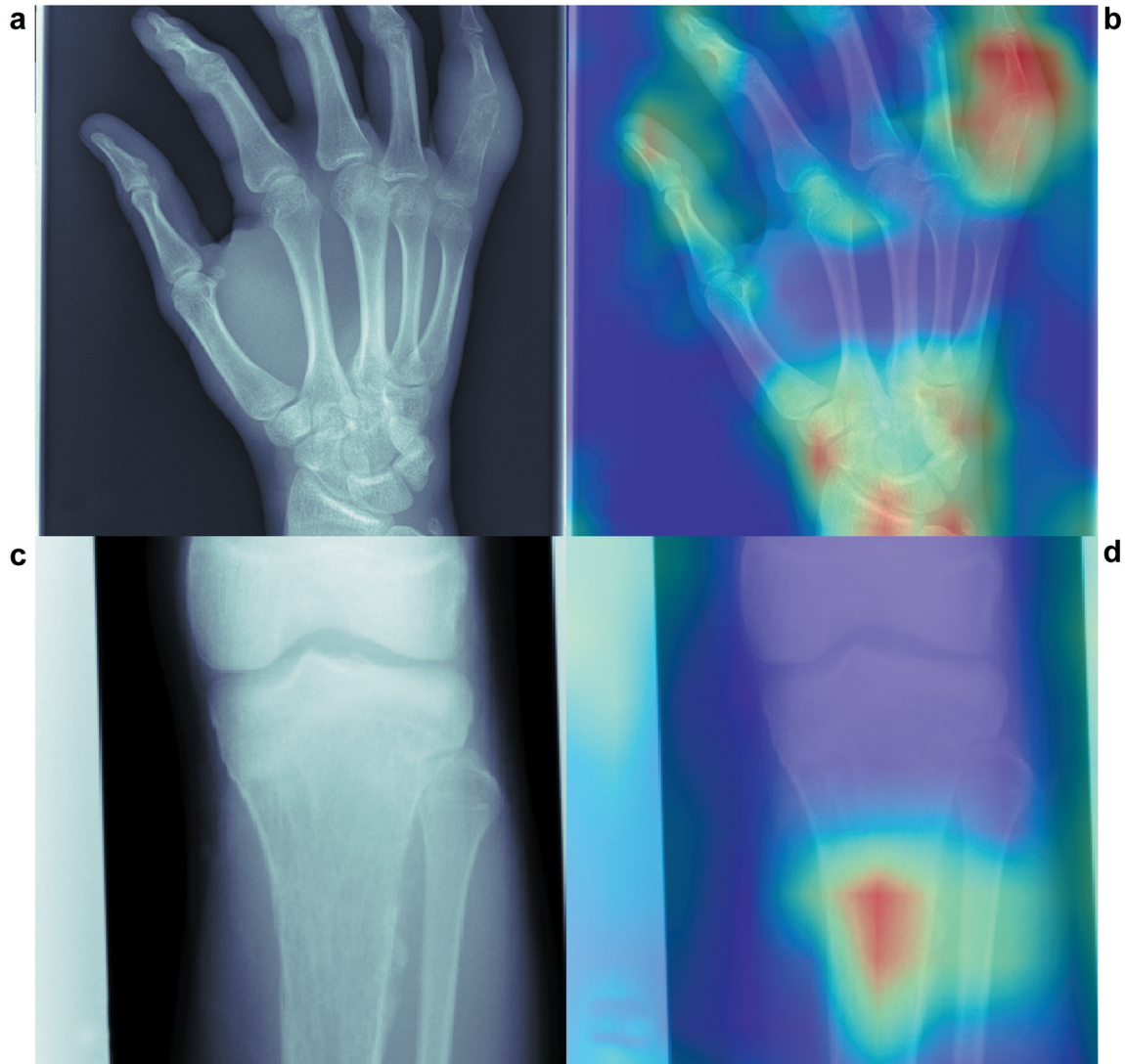
Figure 7. *Grad-CAMs of an OM patient (a/b) and an ES patient (c/d) in phase 2: Grad-CAM results showing that for ES the algorithm did also focus on similar areas a radiologist or orthopaedic surgeon would look at, but in the acute osteomyelitis case several areas were in focus of the algorithm.*

input. This study is supposed to be a feasibility study for radiographs, ES and OM. Nevertheless, integrating meta information into the algorithm is one of the next steps.

Interpretation of results. From a medical as well as a computer science point of view, the performances are very promising considering the complexity of the radiological manifestation of the diseases and the very limited amount of available data. Not only the overall accuracy, but the sensitivity and specificity (also incorporating true positive rate and the true negative rate), concluded in considerably high results.

The model accuracy obtained in the study of von Schacky *et al.* (25) involving all primary bone tumours was comparable with a musculoskeletal fellowship-trained radiologist (71.2% and 64.9%, respectively) and even higher than the one obtained by radiologic residents (83.8% and 82.9%; respectively). Therefore, we can hypothesize that deep learning algorithms, such as the one presented in this study, can potentially become a significant support - particularly for outpatient clinic doctors who do not have access to expert orthopaedic tumour radiologists. The algorithm could help to reduce the delay of referral to a specialised centre and improve the overall survival of young patients.

While this study demonstrates the feasibility of interpreting X-ray images with ES and OM through DL and most likely also surpasses the accuracy of outpatient clinics (no literature

was found to underline this statement), however, statistical robustness must be further investigated before a decision support tool can be integrated into a clinical environment.

Interpretation of Grad-CAMs. While the significance and validity of Grad-CAMs is for some tasks also controversially discussed in the field of computer science, we still believe that it is worth analysing and interpreting specific Grad-CAM results. For example, Figure 7 (Grad-CAM c) shows that the suspect region around the middle phalange of the 4th finger was detected by the algorithm, but additionally several other spots in the wrist area affected the algorithm's decision. Such findings can help to unravel the "black box" behind state-of-the-art DL algorithms, might indicate new ways to evaluate radiographs (and also other imaging modalities) and on the long run assist the process of making precise and fast diagnosis.

Future application. The primary application of the developed algorithm is focused on outpatient clinics. While specialised centres usually have several sarcoma experts as well as more sophisticated imaging modalities, an outpatient clinic doctor has to rely on his/her expertise and radiographic diagnostics to conclude a first diagnosis and potentially refer a patient to a specialised centre, while having seen only about three musculoskeletal malignancies in his/her professional life (26). In such a case, a support tool to highlight suspect cases and even identify ES or OM could have a significant impact.

Conclusion

Radiography is a common and largely available imaging technique that is often used for first clinical assessment. Although radiographs only consist of two-dimensional greyscale information, the high resolution and the considerably standardised technique still make it a very suitable input for modern algorithms. We believe that AI algorithms can become a valuable real-time support for any outpatient clinic involved in the crucial processes of detecting and differentiating a case of acute osteomyelitis from a possible case of an Ewing sarcoma. This allows for a minimal loss of time between diagnosis and specific treatment, which is crucial for patients with Ewing sarcoma. While our algorithm was developed for a specific dataset, it can function as a template for other entities with minor adjustments, where a radiograph can be utilised for early and precise detection for various diseases.

Conflicts of Interest

The Authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authors' Contributions

FH - Project administration, Software, Writing - Original Draft, Visualization. SC - Project administration, Conceptualization, Writing - Original Draft, Data Curation. JN - Writing - Review & Editing, Formal analysis. MS - Investigation, Data Curation. AS - Investigation, Data Curation. MS - Investigation, Data Curation. FS - Writing - Review & Editing. UL - Writing - Review & Editing. CK - Writing - Review & Editing. DR - Methodology, Supervision, Validation. RE - Supervision, Validation. RB - Supervision, Validation

Acknowledgements

References

1 Weber MA, Papakonstantinou O, Nikodinovska VV and Vanhoenacker FM: Ewing's sarcoma and primary osseous lymphoma: Spectrum of imaging appearances. Semin Musculoskelet Radiol *23(1)*: 36-57, 2019. PMID: 30699452. DOI: 10.1055/s-0038-1676125

2 Peltola H and Pääkkönen M: Acute osteomyelitis in children. N Engl J Med *370(4)*: 352-360, 2014. PMID: 24450893. DOI: 10.1056/NEJMra1213956

3 Chiappini E, Mastrangelo G and Lazzeri S: A case of acute osteomyelitis: an update on diagnosis and treatment. Int J Environ Res Public Health *13(6)*: 539, 2016. PMID: 27240392. DOI: 10.3390/ijerph13060539

4 Pineda C, Vargas A and Rodríguez AV: Imaging of osteomyelitis: current concepts. Infect Dis Clin North Am *20(4)*: 789-825, 2006. PMID: 17118291. DOI: 10.1016/j.idc.2006.09.009

5 McNeil JC: Acute hematogenous osteomyelitis in children: Clinical presentation and management. Infect Drug Resist *13*: 4459-4473, 2020. PMID: 33364793. DOI: 10.2147/IDR.S257517

6 Peltola H, Pääkkönen M, Kallio P, Kallio MJ and Osteomyelitis-Septic Arthritis Study Group: Short- versus long-term antimicrobial treatment for acute hematogenous osteomyelitis of childhood: prospective, randomized trial on 131 culture-positive cases. Pediatr Infect Dis J *29(12)*: 1123-1128, 2010. PMID: 20842069. DOI: 10.1097/INF.0b013e3181f55a89

7 Picci P, Manfrini M, Fabbri N, Gambarotti M and Vanel D: Atlas of musculoskeletal tumors and tumorlike lesions: The rizzoli case archive. Springer International Publishing, 2016.

8 McCarville MB, Chen JY, Coleman JL, Li Y, Li X, Adderson EE, Neel MD, Gold RE and Kaufman RA: Distinguishing osteomyelitis from Ewing sarcoma on radiography and MRI. AJR Am J Roentgenol *205(3)*: 640-50; quiz 651, 2015. PMID: 26295653. DOI: 10.2214/AJR.15.14341

9 Kuleta-Bosak E, Kluczewska E, Machnik-Broncel J, Madziara W, Ciupińska-Kajor M, Sojka D, Rogala W, Juszczyk J and Wilk R: Suitability of imaging methods (X-ray, CT, MRI) in the diagnostics of Ewing's sarcoma in children - analysis of own material. Pol J Radiol *75(1)*: 18-28, 2010. PMID: 22802757.

10 Mar WA, Taljanovic MS, Bagatell R, Graham AR, Speer DP, Hunter TB and Rogers LF: Update on imaging and treatment of Ewing sarcoma family tumors: what the radiologist needs to know. J Comput Assist Tomogr *32(1)*: 108-118, 2008. PMID: 18303298. DOI: 10.1097/RCT.0b013e31805c030f

11 Aggarwal H, D'souza M, Panwar P, Jyotsna N, Alvi T, Solanki Y, Kumar T and Sharma R: Role of fluoroethyl tyrosine positron emission tomography-computed tomography scan in differentiating ewing's sarcoma from osteomyelitis. World J Nucl Med *18(1)*: 77-80, 2019. PMID: 30774555. DOI: 10.4103/wjnm.WJNM_23_18

12 Santiago Restrepo C, Giménez CR and McCarthy K: Imaging of osteomyelitis and musculoskeletal soft tissue infections: current concepts. Rheum Dis Clin North Am *29(1)*: 89-109, 2003. PMID: 12635502. DOI: 10.1016/s0889-857x(02)00078-9

13 Henninger B, Glodny B, Rudisch A, Trieb T, Loizides A, Putzer D, Judmaier W and Schocke MF: Ewing sarcoma versus osteomyelitis: differential diagnosis with magnetic resonance imaging. Skeletal Radiol *42(8)*: 1097-1104, 2013. PMID: 23685708. DOI: 10.1007/s00256-013-1632-5

14 Bacci G, Di Fiore M, Rimondini S and Baldini N: Delayed diagnosis and tumor stage in Ewing's sarcoma. Oncol Rep *6(2)*: 465-466, 1999. PMID: 10023023.

15 Bellan DG, Filho RJ, Garcia JG, de Toledo Petrilli M, Maia Viola DC, Schoedl MF and Petrilli AS: Ewing's sarcoma: Epidemiology and prognosis for patients treated at the pediatric oncology institute, IOP-GRAACC-UNIFESP. Rev Bras Ortop *47(4)*: 446-450, 2015. PMID: 27047848. DOI: 10.1016/S2255-4971(15)30126-9

16 Esiashvili N, Goodman M and Marcus RB Jr: Changes in incidence and survival of Ewing sarcoma patients over the past 3 decades: Surveillance Epidemiology and End Results data. J Pediatr Hematol Oncol *30(6)*: 425-430, 2008. PMID: 18525458. DOI: 10.1097/MPH.0b013e31816e22f3

17 Vogrin M, Trojner T and Kelc R: Artificial intelligence in musculoskeletal oncological radiology. Radiol Oncol *55(1)*: 1-6, 2020. PMID: 33885240. DOI: 10.2478/raon-2020-0068

18 Moradi M, Madani A, Karargyris A and Syeda-mahmood T: Chest x-ray generation and data augmentation for cardiovascular abnormality classification. Medical Imaging 2018: Image Processing, 2018. DOI: 10.1117/12.2293971

19 Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, Xiong H and He Q: A comprehensive survey on transfer learning. Proceedings of the IEEE *109(1)*: 43-76, 2022. DOI: 10.1109/JPROC.2020.3004555

20 Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF and Collins GS: Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med *162(1)*: W1-73, 2015. PMID: 25560730. DOI: 10.7326/M14-0698

21 von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP and STROBE Initiative: The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. Int J Surg *12(12)*: 1495-1499, 2014. PMID: 25046131. DOI: 10.1016/j.ijsu.2014.07.013

22 He K, Zhang X, Ren S and Sun J: Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019. DOI: 10.1109/CVPR.2016.90

23 Eigen D and Fergus R: Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture. 2015 IEEE International Conference on Computer Vision (ICCV), 2017. DOI: 10.1109/ICCV.2015.304

24 Selvaraju R, Cogswell M, Das A, Vedantam R, Parikh D and Batra D: Grad-CAM: Visual explanations from deep networks via gradient-based localization. 2017 IEEE International Conference on Computer Vision (ICCV), 2022. DOI: 10.1109/ICCV.2017.74

25 von Schacky CE, Wilhelm NJ, Schäfer VS, Leonhardt Y, Gassert FG, Foreman SC, Gassert FT, Jung M, Jungmann PM, Russe MF, Mogler C, Knebel C, von Eisenhart-Rothe R, Makowski MR, Woertler K, Burgkart R and Gersing AS: Multitask deep learning for segmentation and classification of primary bone tumors on radiographs. Radiology *301(2)*: 398-406, 2021. PMID: 34491126. DOI: 10.1148/radiol.2021204531

26 Clark MA and Thomas JM: Delay in referral to a specialist soft-tissue sarcoma unit. Eur J Surg Oncol *31(4)*: 443-448, 2005. PMID: 15837054. DOI: 10.1016/j.ejso.2004.11.016

27 Biswas B, Rastogi S, Khan SA, Mohanti BK, Sharma DN, Sharma MC, Mridha AR and Bakhshi S: Outcomes and prognostic factors for Ewing-family tumors of the extremities. J Bone Joint Surg Am *96(10)*: 841-849, 2014. PMID: 24875025. DOI: 10.2106/JBJS.M.00411