SCIENTIFIC ARTICLE

Automated detection and segmentation of sclerotic spinal lesions on body CTs using a deep convolutional neural network

Connie Y. Chang[1] · Colleen Buckless[1] · Kaitlyn J. Yeh[1] · Martin Torriani[1]

Abstract

Purpose To develop a deep convolutional neural network capable of detecting spinal sclerotic metastases on body CTs.

Materials and methods Our study was IRB-approved and HIPAA-compliant. Cases of confirmed sclerotic bone metastases in chest, abdomen, and pelvis CTs were identified. Images were manually segmented for 3 classes: background, normal bone, and sclerotic lesion(s). If multiple lesions were present on a slice, all lesions were segmented. A total of 600 images were obtained, with a 90/10 training/testing split. Images were stored as 128×128 pixel grayscale and the training dataset underwent a processing pipeline of histogram equalization and data augmentation. We trained our model from scratch on Keras/TensorFlow using an 80/20 training/validation split and a U-Net architecture (64 batch size, 100 epochs, dropout 0.25, initial learning rate 0.0001, sigmoid activation). We also tested our model's true negative and false positive rate with 1104 non-pathologic images. Global sensitivity measured model detection of any lesion on a single image, local sensitivity and positive predictive value (PPV) measured model detection of each lesion on a given image, and local specificity measured the false positive rate in non-pathologic bone.

Results Dice scores were 0.83 for lesion, 0.96 for non-pathologic bone, and 0.99 for background. Global sensitivity was 95% (57/60), local sensitivity was 92% (89/97), local PPV was 97% (89/92), and local specificity was 87% (958/1104).

Conclusion A deep convolutional neural network has the potential to assist in detecting sclerotic spinal metastases.

Keywords Deep convolutional neural network · Artificial intelligence · Bone lesions · Sclerotic

Introduction

Osseous metastases are common, affecting approximately 400,000 adults in the USA, often affecting patients with primary breast, prostate, or lung cancer, and are a source of morbidity and pain [1–3]. Detection of metastatic disease is also important for staging, outcome prediction, and treatment planning [4–6]. The majority of these metastases occur in the axial skeleton, which is most commonly evaluated on chest, abdomen, and pelvic CTs, where findings can be subtle [3, 7]. In addition, these CTs are primarily interpreted by general radiologists or body radiologists, rather than musculoskeletal radiologists.

It has been widely recognized that automated lesion detection could help improve radiologist sensitivity for osseous metastases, and multiple computer-aided detection (CAD) software systems have been developed [3, 8–12]. Although CNNs have been used to detect focal lesions in the brain, liver, teeth, lung, ocular fundus, skin, and breast, use of deep convolutional neural networks (CNNs) and other newer techniques has not been fully explored for this particular task [13–21]. We believe that CNNs can be more accurate than the previously developed models in detecting osseous metastases. Our purpose is to develop a deep CNN to accurately detect sclerotic spinal metastases.

Materials and methods

This study was IRB-approved and was Health Insurance and Portability and Accountability Act (HIPAA)-compliant, with exemption status for individual informed consent.

✉ Connie Y. Chang
cychang@mgh.harvard.edu

1   Division of Musculoskeletal Imaging and Intervention, Department of Radiology, Massachusetts General Hospital, Harvard Medical School, 55 Fruit Street, YAW 6, Boston, MA 02114, USA

Dataset

A retrospective search of an internal report search engine developed at our institution was performed for CT reports with "sclerotic lesion" in the report impression between December 2000 and December 2019. Only chest, abdomen, and pelvis CTs in patients with a history of malignancy were included. If a patient had more than one CT scan, then the timepoint where the lesions were the most visible was selected. If a lesion was visible on more than one slice, then only one slice per lesion was selected, in order to avoid over-fitting the model to our dataset. For a similar reason, if there were more than one lesion in a vertebral level, only one slice was chosen per vertebral level. These lesions were confirmed across multiple contiguous slices, prior and/or follow-up imaging, and/or MRI. A total of 242 CT scans were used for the study, from which 600 unique images were collected. The images were divided into 90% training ($N = 540$) and 10% test ($N = 60$) datasets. The 540 training images were augmented to 20,000 training images.

The CTs were performed using multi-detector CT scanners (General Electric, Waukesha, WI, USA; Siemens Healthcare, Erlangen, Germany) at our institution. Studies with and without contrast enhancement were included. Patients were scanned supine, headfirst, and with arms overhead to decrease streak artifact.

Ground truth labeling (manual segmentation)

All studies were reviewed by two investigators with 1 and 6 years of medical image analysis experience, supervised by a musculoskeletal radiologist with 10 years of experience to identify sclerotic osseous lesions—presumably metastases. Lesions were considered sclerotic if denser than adjacent unaffected marrow. Only lesions that were 100% sclerotic were included in the study; mixed lytic-sclerotic lesions were excluded. For each vertebra with a sclerotic lesion, a single axial image at the largest cross-sectional diameter was used for segmentation.

Manual and semi-automated segmentation was performed using the Osirix DICOM viewer (version 6.5.2, www.osirix-viewer.com/index.html). Each image was segmented into 3 classes:

(1) lesion (any number of well-defined abnormal sclerotic foci; green),
(2) bone (non-pathologic marrow and vertebral cortex, including osteophytes; magenta),
(3) background (all remaining image pixels; black).

Initial segmentation of the entire vertebral body was performed using an extension of a previously established model [22]. Then, this segmentation was checked and fine-tuned by the two investigators with 1 (200 images) and 6 years (400 images) of medical image analysis experience. Only the spine (vertebral body and posterior elements) was included in the "bone" class. All other bones (ribs, pelvis) were considered "background." If multiple lesions were present on a slice, they were all segmented separately, but all labeled as the same class (lesion) (Fig. 1). The segmented images were stored as Tag Image File Format (TIFF), with masks being 8-bit RGB and corresponding CT images were 8-bit single-channel grayscale. Finally, the images were manually cropped centered on the vertebrae with final dimensions of 128×128 pixels.

Training

The dataset was randomly divided into 90% training and 10% test datasets, ensuring fully segregated datasets with no overlap. Contrast-limited adaptive histogram equalization (CLAHE) was performed on all grayscale images, which were then saved as JPEG files. Image augmentation was performed on grayscale and ground truth pairs to enlarge the training dataset by applying random rotation, horizontal
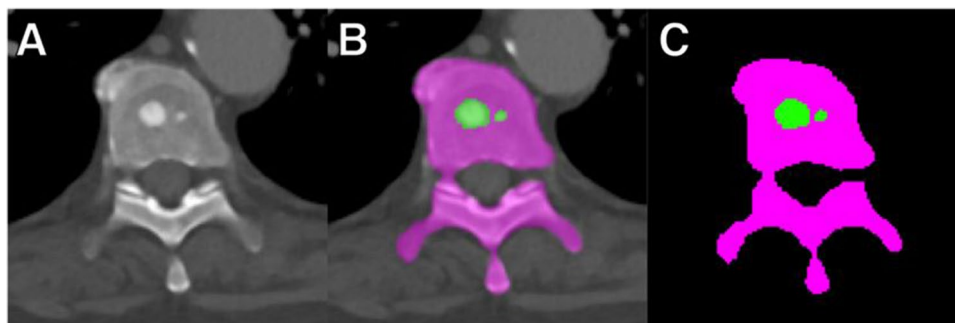

Fig. 1 Example of manual ground truth segmentation. The axial CT image of a vertebra (**A**) is manually segmented (**B**) resulting in a mask (**C**) with three classes: lesion (any number of well-defined abnormal sclerotic foci; *green*), bone (non-pathologic marrow and vertebral cortex, including osteophytes; *magenta*), and background (all remaining image pixels; *black*)

flipping, cropping, and scaling ($N = 20,000$). To further increase variability and improve generalizability, we applied Poisson noise to 50% of randomly selected augmented gray-scale images. All images were normalized to the training dataset mean and standard deviation.

We trained our model from scratch on Keras/Tensor-Flow using an 80/20 training/validation split and a U-Net architecture (64 batch size, 100 epochs, dropout 0.25, initial learning rate 0.0001, sigmoid activation) [23]. Briefly, images were input in our U-Net pipeline that consisted of five layers with four down-sampling steps followed by four up-sampling steps. Each step consisted of two successive 3×3 padded convolutions, and in the down-sizing steps, a dropout of 0.25 was applied. Next, a rectified linear unit (ReLU) activation function and a max-pooling operation with a 2×2 pixel kernel size. The up-sampling operations were performed using a 2×2 transposed convolution followed by a 3×3 filter size convolution, after which the output concatenates with the corresponding decoding step. The final layer consisted of a 1×1 convolution followed by a sigmoid function, resulting in an output pixelwise prediction score for each class.

Our model was written and trained in Python 3.7 (Python Software Foundation, Beaverton, OR) using the Keras library (v2.2.4, https://keras.io) with TensorFlow 1.13.1 (Google, Mountain View, CA, USA) [24]. During training, the 3 classes were adjusted for imbalances by weighting prevalence, penalizing predictions of classes with highest pixel count (e.g., background). Batch size was 64 and we used the Adadelta optimizer (initial learning rate, 0.0001). The model trained for 100 epochs with early stopping enabled. Multi-class Dice loss was used as cost function. Training was performed using a Linux workstation (Ubuntu 14.04) with 4 NVIDIA Titan Xp Graphic Processing Units.

Testing and further validation

The Dice (F1) score was used to assess similarity between the manual segmentations and the CNN-predicted segmentations [25]. A Dice score of 1.00 is a perfect similarity. To further examine model performance, we tested 1104 images of vertebrae from CTs of the chest, abdomen, and pelvis without bone pathology based on report and confirmed to be non-pathologic bone by one of the authors (hereafter referred to as the "non-pathologic dataset"). These images did not contain sclerotic foci suggestive of metastases.

Statistical analysis

Sensitivity was defined as TP / (TP + FN), PPV was defined as TP / (TP + FP), and specificity was defined as TN / (FP + TN).

In the **lesion test dataset**, each image and lesion were evaluated as follows:

- *global sensitivity*: a measure to identify one or more lesions on an image

    - True positive (TP) if the model correctly captured any lesion in that image.
    - False negative (FN) if the model failed to identify any lesion within an image.

    local sensitivity and positive predictive value (PPV): measures to identify each lesion separately in a given image

    - True positive if each lesion was captured regardless of segmentation accuracy.
    - False positive (FP) if the model assigned "lesion" class to non-pathologic bone.
    - False negative if the model assigned "bone" class to a lesion.

In the **non-pathologic test dataset**, each image was evaluated as follows:

- *local specificity*: a measure to identify the false positive rate in non-pathologic bone.

    - True negative if non-pathologic bone was correctly identified as such.
    - False positive if the model assigned the "lesion" class to non-pathologic bone.

Results

Among the 600 images, there were 13 (2%) cervical, 384 (64%) thoracic, and 203 (34%) lumbar vertebral images. Lesions predominantly involved the vertebral body (497, 83%), followed by posterior elements (48, 8%), and images with lesions in both vertebral body and posterior elements (55, 9%).

Lesion test dataset

Dice scores were 0.83 for lesion, 0.96 for non-pathologic bone, and 0.99 for background (Fig. 2). Global sensitivity was 95% (57/60), with a false negative rate of 5% (3/60 images, each containing one missed lesion). Local sensitivity was 92% (89/97) with 3 false positive lesions: a pronounced anterior osteophyte (Fig. 3A), volume averaging from the pedicle, and an area of heterogeneous bone marrow (Fig. 3B). Local PPV was 97% (89/92). We observed

Fig. 2 Examples of accurate lesion segmentation using our deep CNN. Each row shows distinct test images with varied lesion sizes and count (**A**, **B**, **C**). Manual, manual tracing; Deep CNN, model prediction
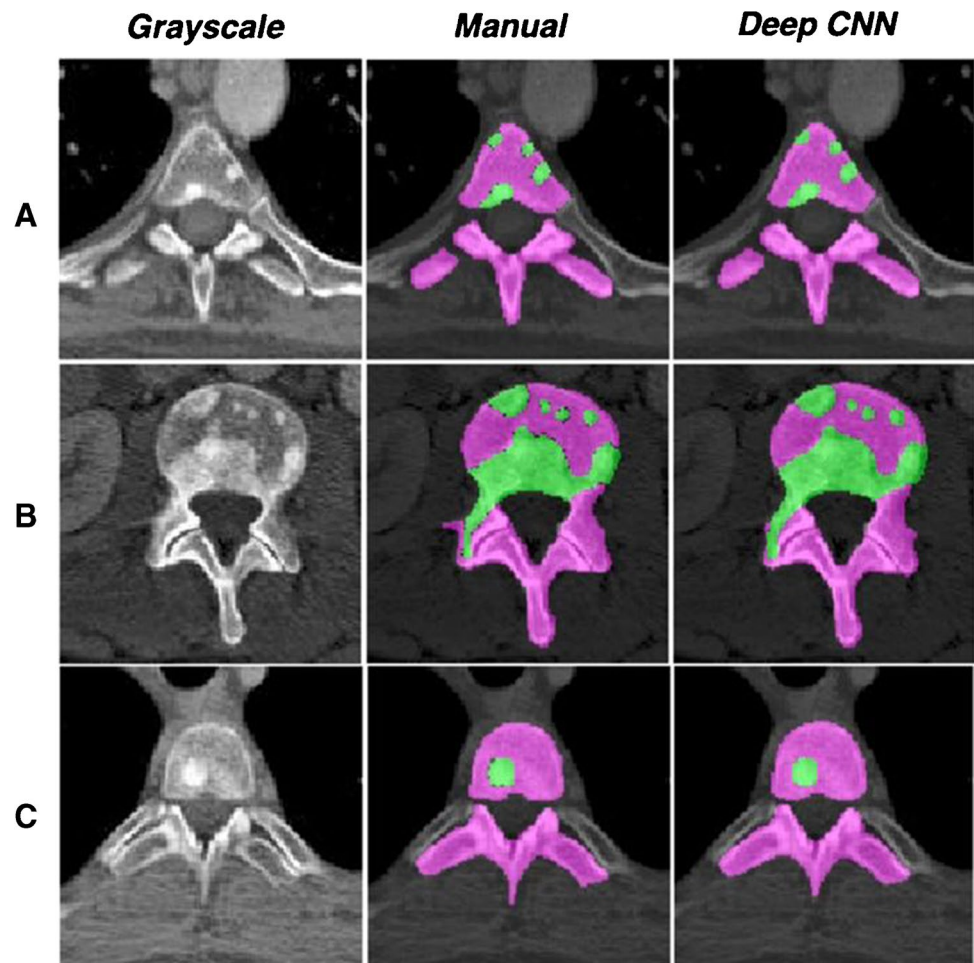

Fig. 3 Prediction errors on 2 distinct test images (one per row) from the lesion dataset, showing 3 false positive lesions. (**A**) Pronounced anterior osteophyte (arrowhead) was mislabeled as a lesion. (**B**) Volume averaging from the right pedicle (arrow) and an area of heterogeneous bone marrow (arrowhead) were mislabeled as lesions. Manual, manual tracing; Deep CNN, model prediction
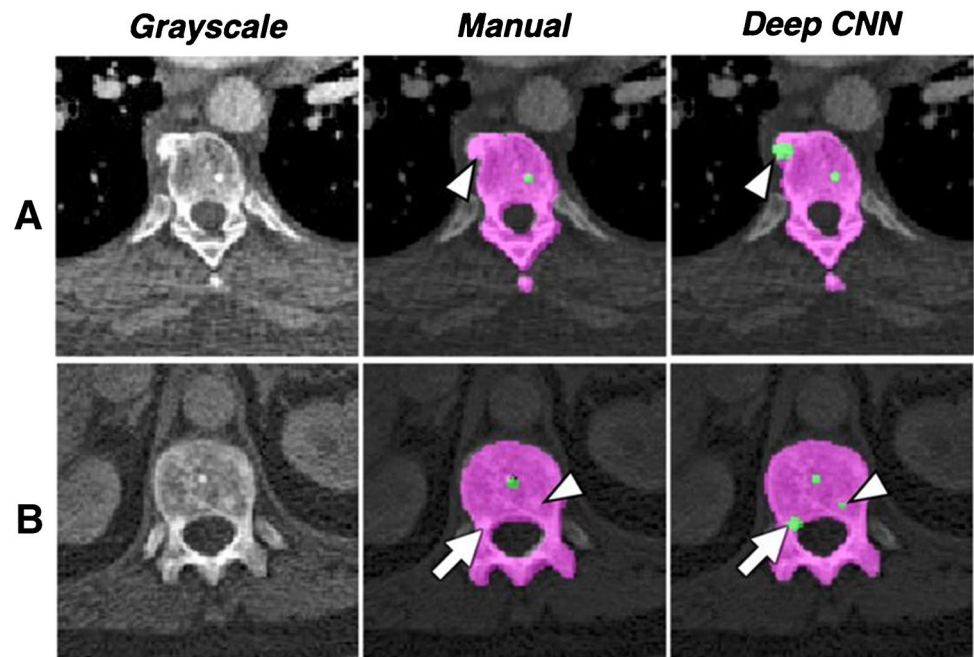
Fig. 4 Prediction errors on 5 distinct test images (one per row) showing 8 false negative lesions. (**A**) Single lesion in bone marrow (arrowhead). (**B**) Two lesions (arrowhead and arrow) in bone marrow. (**C**) Single lesion adjacent to endplate (arrowhead). (**D**) Single lesion adjacent to left costovertebral junction (arrowhead). (**E**) Three lesions, one in bone marrow (black arrow), one in right transverse process (white arrow), and one in left transverse process (arrowhead). Manual, manual tracing; Deep CNN, model prediction
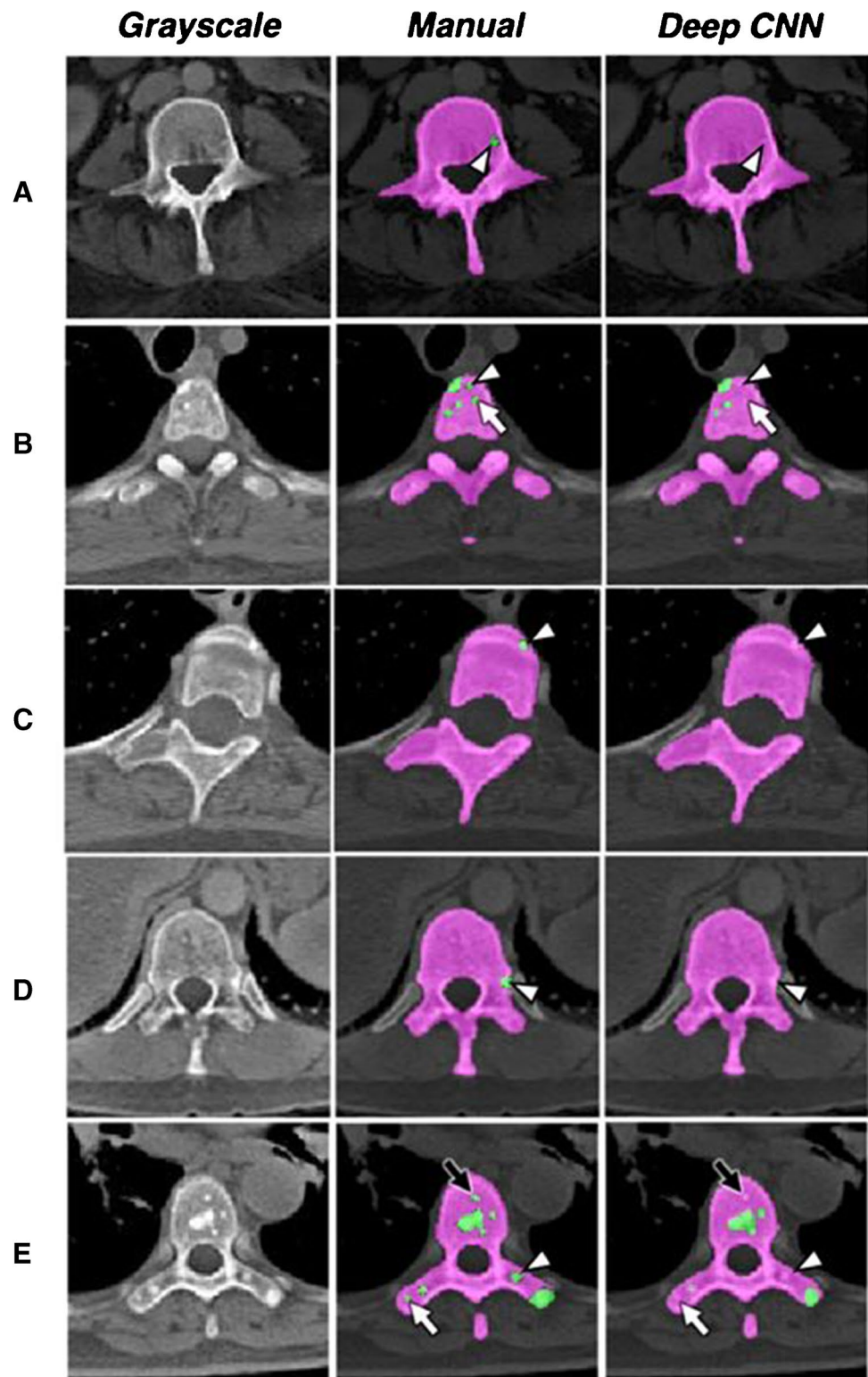
8/97 (8%) local false negative rate comprising 5 images, which are shown in Fig. 4. In 2 of the 5 images, lesions were missed in the context of multiple lesions, most of which were identified. In 3 of the 5 images, missed lesions were near cortical bone.

Non-pathologic test dataset

Additional testing performed on 1104 images without lesions covered the following anatomic regions: 21/1104 (2%) cervical, 787/1104 (69%) thoracic, and 318 (29%)

lumbar spine. Local specificity was 87% (958/1104). At least one false positive lesion was noted in 146/1104 (13%) images, of which 47/146 (32%) were 4 pixels or fewer, considered to be below detection threshold (Fig. 5A). A single pixel for a given class was not reliably visualized upon qualitative inspection of images and could represent a spurious prediction. Therefore, we used the next step larger in scale that had a square aspect ratio, which is comprised of 4 pixels. The most common location for false positive lesions was the vertebral body (131/1104, 12%), most commonly at the endplate or in regions of increased sclerosis secondary to endplate degenerative change (73/131, 56%) (Fig. 5B), as well as osteophytes (25/131, 19%) (Fig. 5C). The second most common location for false positive lesions was the posterior elements (36/1104, 3%), with volume averaging in the pedicle (18/36, 50%) (Fig. 5D) and lamina (8/36, 22%) being more frequent. Other likely reasons for false positive lesions included disc calcification, facet degenerative change, and intravenous contrast in adjacent blood vessels (Fig. 5E).

Discussion

This study aimed to develop a deep CNN capable of detecting spinal sclerotic metastases on body CTs. We found our proposed technique yielded high Dice scores for lesion detection and high global and local sensitivity. Further, examination of a non-pathologic dataset showed a high local specificity. Altogether, our study showed that a deep CNN has the potential to assist in detecting sclerotic spinal metastases.

The earliest applications of artificial intelligence in detection of focal lesions on imaging were in the fields of breast imaging (computer-assisted diagnosis or CAD) and CT colonography [14, 18]. CAD software packages have also been developed to detect focal bone lesions. Burns et al. used a "watershed algorithm" originally designed for detection of lytic lesions to evaluate the entire vertebral body for lesion candidates [3]. Additional algorithms using graph cuts and threshold levels were used to refine the segmentations. This CAD system detected 439 out of 532 (83%) lesions in 49 patients, with a testing set sensitivity of 79% (95% CI: 74–84%) [3]. Huang and Chiang described a CAD system for PET/CT able to detect both lytic and sclerotic lesions with a sensitivity, specificity, and accuracy of 85%, 92%, and 90%, respectively [26]. Hammon et al. described another CAD system able to detect sclerotic and lytic lesions on CT with a sensitivity by case of 83% for sclerotic lesions [8]. In comparison, our neural network had a global sensitivity of 95% and local sensitivity of 92%. The model accurately predicted the borders of the lesions, which was reflected in the high Dice scores (Fig. 2). The 8 missed lesions in the 5 images were in locations of higher difficulty to detect. In the first instance, the lesion was slightly denser than the background marrow and was difficult to identify on a single axial slice (Fig. 4A). In the second instance, there were 5 lesions on the image, and the model identified 3 out of 5 lesions. The two that were not identified were also only slightly denser than the surrounding bone marrow (Fig. 4B). In the third instance, the lesion was adjacent to the endplate, and there was also disc and the adjacent vertebral body on the same image (Fig. 4C). In the fourth instance, the model

Fig. 5 Examples of model false positive lesions in the non-pathologic test dataset. (**A**) Single pixel in the left transverse process (arrowhead) was considered below detection threshold (< 4 pixels). (**B**) Vertebral body endplate change (arrowheads) was the most common location for false positive lesions (arrowheads). (**C**) Osteophytes (arrowhead) were the second most common, followed by (**D**) volume averaging in the posterior elements (arrowhead), and (**E**) contrast in adjacent vessels (arrowhead). Deep CNN, model prediction
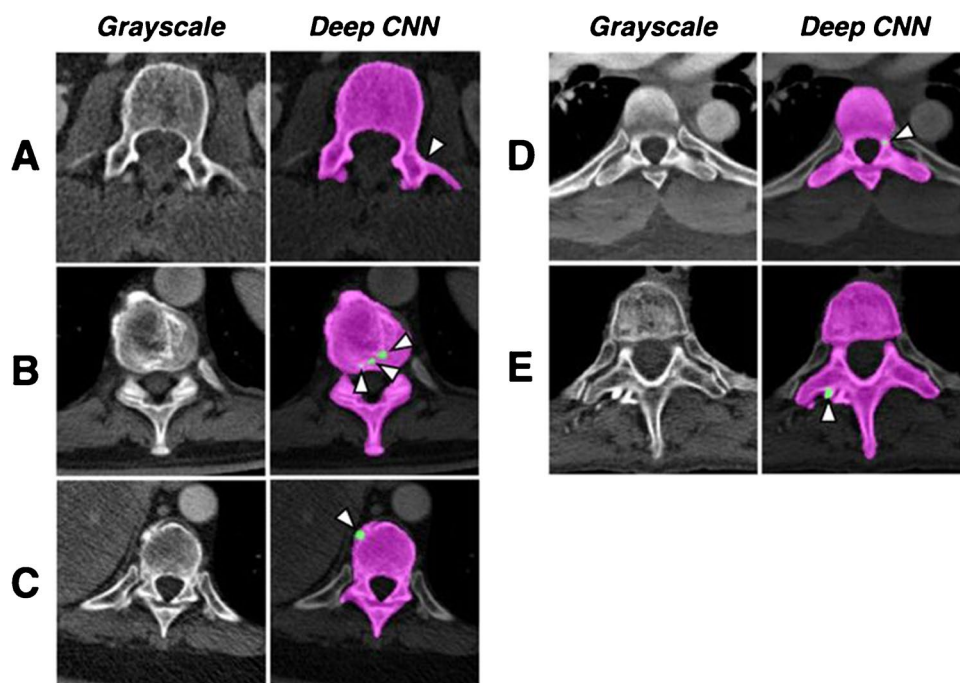
missed a small lesion adjacent to the left costovertebral junction (Fig. 4D). In the fifth instance, the model identified 4 out of 7 lesions. Two of the missed lesions were in the posterior elements, which is a less common location for lesions than the vertebral body (Fig. 4E). Additional training images focused on these challenging areas may be helpful to improve detection in these areas.

To our knowledge, there is no previous study using a deep CNN to help detect vertebral sclerotic lesions. Roth et al. used a CNN and random view aggregation to improve the performance of a CAD system [10]. However, specific lesion segmentation was not performed. There has been a single study by Klein et al. that looked at CNN segmentation of normal bone, and they achieved a Dice score of 0.95, similar to our study (0.96) [27]. Although segmentation of normal bone and background was not the goal of our study, this task could also have useful applications in trauma or osteoporosis screening [28, 29].

Using a CNN to detect focal lesions has been explored in other organ systems, including focal bladder lesions in CT urography, prostate cancer, and breast masses [30, 31]. CNNs have also been shown to be able to assist in the detection of non-neoplastic focal lesions such as strokes, multiple sclerosis plaques, focal cortical dysplasia, liver cirrhosis, and fatty liver disease [13, 32–41]. Other non-radiology medical fields using deep learning methods to detect focal lesions include focal esophageal lesions on endoscopy, segmentation of cervical cancer cells in pathology, and diabetic retinopathy in ophthalmology [19, 42, 43]. These studies found that their deep learning algorithms can be helpful to physicians in improving lesion detection rate. For example, Cai et al. sound that their model sensitivity values for the detection of early esophageal squamous cell carcinoma were 98%, 85%, 91%, 86%, and 98%, compared to diagnostic accuracy for 89% of a senior group of endoscopists, and 77% of a junior group of endoscopists [40]. The sensitivity of our model is similar to the results of these studies.

In this study, we observed 3 false positive lesions: a pronounced anterior osteophyte, volume averaging from the pedicle, and an area of heterogeneous bone marrow (Fig. 3). This finding is similar to findings on prior CAD studies. In Burns et al., the most common causes for false positive lesions were endplate changes and vertebral endplate cortex. The most common cases for false negative findings were endplate proximity (and therefore volume averaging with the intervertebral disc), lower attenuation, and small size [3]. CAD errors were also most often near the endplates in Huang and Chiang [26]. Hammon et al. found that CAD errors were most commonly secondary to endplate changes, osteophytes, and small size [8].

The limitations of this study include the relatively small dataset that may impact generalizability, and the application of the model only to sclerotic lesions, not evaluating its performance on lytic or mixed density lesions. Also, the model was trained and tested on single slices, rather than having to scan through a stack of contiguous spine images. Each scan was checked at multiple adjacent levels by an experienced musculoskeletal radiologist prior to image selection for lesion identification. However, it is possible volume averaging artifacts may still be present in the selected images. The model inputs were also cropped to isolate vertebra, so that the sclerotic lesion represents a higher percentage of pixels on the image; in reality, the sclerotic lesion is a very small portion of the image on a full chest, abdomen, or pelvic CT slice. Further, the model is limited to lesions in the spine and not in other structures, such as the pelvic bones and ribs. Multiple lesions were used from the same patient, which has a risk of overfitting the data; however, this risk was minimized by using only one image per lesion and vertebral level. Finally, external validation has not been performed on this model. As presented, this model is not sufficient to deploy in routine clinical practice. These limitations are hurdles that must be overcome in order to achieve a clinically useful application.

In conclusion, a deep CNN has the potential to assist in detecting sclerotic spinal metastases and could become an important component of radiology workflow in the interpretation of body CTs. This model is highly accurate in defining lesions, non-pathologic bone, and background. Additional model training may help overcome areas prone to false positives, especially endplate degenerative changes and osteophytes.

Declarations

References

1. Coleman RE. Clinical features of metastatic bone disease and risk of skeletal morbidity. Clin Cancer Res Off J Am Assoc Cancer Res. 2006;12(20 Pt 2):6243s-s6249.
2. Li S, Peng Y, Weinhandl ED, Blaes AH, Cetin K, Chia VM, et al. Estimated number of prevalent cases of metastatic bone disease in the US adult population. Clin Epidemiol. 2012;4:87–93.

3. Burns JE, Yao J, Wiese TS, Muñoz HE, Jones EC, Summers RM. Automated detection of sclerotic metastases in the thoracolumbar spine at CT. Radiology. 2013;268(1):69–78.

4. Söderlund V. Radiological diagnosis of skeletal metastases. Eur Radiol. 1996;6(5):587–95.

5. Clemons M, Gelmon KA, Pritchard KI, Paterson AHG. Bone-targeted agents and skeletal-related events in breast cancer patients with bone metastases: the state of the art. Curr Oncol Tor Ont. 2012;19(5):259–68.

6. Hamaoka T, Madewell JE, Podoloff DA, Hortobagyi GN, Ueno NT. Bone imaging in metastatic breast cancer. J Clin Oncol Off J Am Soc Clin Oncol. 2004;22(14):2942–53.

7. Chang CY, Gill CM, Joseph Simeone F, Taneja AK, Huang AJ, Torriani M, et al. Comparison of the diagnostic accuracy of 99 m-Tc-MDP bone scintigraphy and 18 F-FDG PET/CT for the detection of skeletal metastases. Acta Radiol Stockh Swed 1987. 2016;57(1):58–65.

8. Hammon M, Dankerl P, Tsymbal A, Wels M, Kelm M, May M, et al. Automatic detection of lytic and blastic thoracolumbar spine metastases on computed tomography. Eur Radiol. 2013;23(7):1862–70.

9. Gorelik N, Chong J, Lin DJ. Pattern recognition in musculoskeletal imaging using artificial intelligence. Semin Musculoskelet Radiol. 2020;24(1):38–49.

10. Roth HR, Lu L, Liu J, Yao J, Seff A, Cherry K, et al. Improving computer-aided detection using convolutional neural networks and random view aggregation. IEEE Trans Med Imaging. 2016;35(5):1170–81.

11. O'Connor SD, Yao J, Summers RM. Lytic metastases in thoracolumbar spine: computer-aided detection at CT–preliminary study. Radiology. 2007;242(3):811–6.

12. Yao J, Burns JE, Sanoria V, Summers RM. Mixed spine metastasis detection through positron emission tomography/computed tomography synthesis and multiclassifier. J Med Imaging Bellingham Wash. 2017;4(2):024504.

13. Chen L, Bentley P, Rueckert D. Fully automatic acute ischemic lesion segmentation in DWI using convolutional neural networks. NeuroImage Clin. 2017;15:633–43.

14. Le EPV, Wang Y, Huang Y, Hickman S, Gilbert FJ. Artificial intelligence in breast imaging. Clin Radiol. 2019;74(5):357–66.

15. Zhou L-Q, Wang J-Y, Yu S-Y, Wu G-G, Wei Q, Deng Y-B, et al. Artificial intelligence in medical imaging of the liver. World J Gastroenterol. 2019;25(6):672–82.

16. Orhan K, Bayrakdar IS, Ezhov M, Kravtsov A, Özyürek T. Evaluation of artificial intelligence for detecting periapical pathosis on cone-beam computed tomography scans. Int Endod J. 2020;53(5):680–9.

17. Chen X-Z, Liu C-G, Chen Y, Wang L-Q, Zhu Q-Z, Lin P. Arthroscopy-assisted surgery for tibial plateau fractures. Arthrosc J Arthrosc Relat Surg Off Publ Arthrosc Assoc N Am Int Arthrosc Assoc. 2015;31(1):143–53.

18. Helbren E, Fanshawe TR, Phillips P, Mallett S, Boone D, Gale A, et al. The effect of computer-aided detection markers on visual search and reader performance during concurrent reading of CT colonography. Eur Radiol. 2015;25(6):1570–8.

19. Zhang S, Han F, Liang Z, Tan J, Cao W, Gao Y, et al. An investigation of CNN models for differentiating malignant from benign lesions using small pathologically proven datasets. Comput Med Imaging Graph Off J Comput Med Imaging Soc. 2019;77:101645.

20. Pesce E, Joseph Withey S, Ypsilantis P-P, Bakewell R, Goh V, Montana G. Learning to detect chest radiographs containing pulmonary lesions using visual attention networks. Med Image Anal. 2019;53:26–38.

21. Jaworek-Korjakowska J. A deep learning approach to vascular structure segmentation in dermoscopy colour images. BioMed Res Int. 2018;2018:5049390.

22. Hemke R, Buckless CG, Tsao A, Wang B, Torriani M. Deep learning for automated segmentation of pelvic muscles, fat, and bone from CT studies for body composition assessment. Skeletal Radiol. 2020;49(3):387–95.

23. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention. Springer; 2015. p. 234–41.

24. Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J. TensorFlow: a system for large-scale machine learning. OSDI; 2016. 265–83 p. Available from: usenix.org

25. Dice L. Measure of the amount of ecologic assocation between species. Ecology. 1945;26(3):297–302.

26. Huang S, Chang K. Automatic detection of bone metastases in vertebrae by using CT images. Paper presented at: Proceedings of the World Congress on Engineering; July 4–6, 2012; London, UK.

27. Klein A, Warszawski J, Hillengaß J, Maier-Hein KH. Automatic bone segmentation in whole-body CT images. Int J Comput Assist Radiol Surg. 2019;14(1):21–9.

28. Löffler MT, Sekuboyina A, Jacob A, Grau A-L, Scharr A, El Husseini M, et al. A vertebral segmentation dataset with fracture grading. Radiol Artif Intell. 2020;2(4):e190138.

29. Jang S, Graffy PM, Ziemlewicz TJ, Lee SJ, Summers RM, Pickhardt PJ. Opportunistic osteoporosis screening at routine abdominal and thoracic CT: normative L1 trabecular attenuation values in more than 20 000 adults. Radiology. 2019;291(2):360–7.

30. Cha KH, Hadjiiski L, Samala RK, Chan H-P, Caoili EM, Cohan RH. Urinary bladder segmentation in CT urography using deep-learning convolutional neural network and level sets. Med Phys. 2016;43(4):1882.

31. Cuocolo R, Cipullo MB, Stanzione A, Ugga L, Romeo V, Radice L, et al. Machine learning applications in prostate cancer magnetic resonance imaging. Eur Radiol Exp. 2019;3(1):35.

32. Chen Y, Villanueva-Meyer JE, Morrison MA, Lupo JM. Toward automatic detection of radiation-induced cerebral microbleeds using a 3D deep residual network. J Digit Imaging. 2019;32(5):766–72.

33. Nazari-Farsani S, Nyman M, Karjalainen T, Bucci M, Isojärvi J, Nummenmaa L. Automated segmentation of acute stroke lesions using a data-driven anomaly detection on diffusion weighted MRI. J Neurosci Methods. 2020;333:108575.

34. Hsieh Y-Z, Luo Y-C, Pan C, Su M-C, Chen C-J, Hsieh KL-C. Cerebral small vessel disease biomarkers detection on MRI-sensor-based image and deep learning. Sensors. 2019;19(11).

35. Kamnitsas K, Ledig C, Newcombe VFJ, Simpson JP, Kane AD, Menon DK, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med Image Anal. 2017;36:61–78.

36. Hashemi L, Webster BS, Clancy EA. Trends in disability duration and cost of workers' compensation low back pain claims (1988–1996). J Occup Environ Med. 1998;40(12):1110–9.

37. Yoo Y, Tang LYW, Brosch T, Li DKB, Kolind S, Vavasour I, et al. Deep learning of joint myelin and T1w MRI features in normal-appearing brain tissue to distinguish between multiple sclerosis patients and healthy controls. NeuroImage Clin. 2018;17:169–78.

38. Wang H, Ahmed SN, Mandal M. Automated detection of focal cortical dysplasia using a deep convolutional neural network. Comput Med Imaging Graph Off J Comput Med Imaging Soc. 2020;79:101662.

39. Nair T, Precup D, Arnold DL, Arbel T. Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation. Med Image Anal. 2020;59:101557.

40. Alex V, Vaidhya K, Thirunavukkarasu S, Kesavadas C, Krishnamurthi G. Semisupervised learning using denoising autoencoders

for brain lesion detection and segmentation. J Med Imaging Bellingham Wash. 2017;4(4):041311.

41. Sundaresan V, Zamboni G, Le Heron C, Rothwell PM, Husain M, Battaglini M, et al. Automated lesion segmentation with BIANCA: impact of population-level features, classification algorithm and locally adaptive thresholding. NeuroImage. 2019;202:116056.

42. Cai S-L, Li B, Tan W-M, Niu X-J, Yu H-H, Yao L-Q, et al. Using a deep learning system in endoscopy for screening of early esophageal squamous cell carcinoma (with video). Gastrointest Endosc. 2019;90(5):745-753.e2.

43. Zago GT, Andreão RV, Dorizzi B, TeatiniSalles EO. Diabetic retinopathy detection using red lesion localization and convolutional neural networks. Comput Biol Med. 2020;116:103537.