Canadian Association of Radiologists' Journal
2021, Vol. 72(1) 45-59

DOI: 10.1177/0846537120947148
journals.sagepub.com/home/caj
SAGE

Applications of Artificial Intelligence in Musculoskeletal Imaging: From the Request to the Report

Natalia Gorelik, MD[1], and Soterios Gyftopoulos, MD, MSc[2,3]

Abstract

Artificial intelligence (AI) will transform every step in the imaging value chain, including interpretive and noninterpretive components. Radiologists should familiarize themselves with AI developments to become leaders in their clinical implementation. This article explores the impact of AI through the entire imaging cycle of musculoskeletal radiology, from the placement of the requisition to the generation of the report, with an added Canadian perspective. Noninterpretive tasks which may be assisted by AI include the ordering of appropriate imaging tests, automatic exam protocoling, optimized scheduling, shorter magnetic resonance imaging acquisition time, computed tomography imaging with reduced artifact and radiation dose, and new methods of generation and utilization of radiology reports. Applications of AI for image interpretation consist of the determination of bone age, body composition measurements, screening for osteoporosis, identification of fractures, evaluation of segmental spine pathology, detection and temporal monitoring of osseous metastases, diagnosis of primary bone and soft tissue tumors, and grading of osteoarthritis.

Résumé

L'intelligence artificielle (IA) promet de transformer toutes les étapes de la chaine de valeur de l'imagerie médicale, notamment les volets interprétatifs et non interprétatifs. Les radiologistes doivent se familiariser avec les évolutions de l'IA afin de devenir des leaders de leur application clinique. Cet article s'intéresse à l'incidence de l'IA à travers le cycle complet de l'imagerie en radiologie musculosquelettique, à partir de la demande jusqu'à la génération du rapport final, en tenant compte du contexte canadien. Les missions non interprétatives pour lesquelles l'IA peut être utilisée comprennent la commande des examens d'imagerie appropriés, un protocole d'imagerie automatisé, une prise de rendez-vous optimisée, un temps d'acquisition des images par résonance magnétique réduit, une imagerie par tomographie assistée avec des artefacts et des doses de radiations réduites, et de nouvelles méthodes de génération et d'utilisation des rapports de radiologie. Les applications de l'IA pour l'interprétation des images consistent en la détermination de l'âge osseux, la mesure de la composition corporelle, le dépistage de l'ostéoporose, l'identification des fractures, l'évaluation des pathologies de la colonne vertébrale, la détection et le suivi temporel des métastases osseuses, le diagnostic des tumeurs primaires des os et des tissus mous, et la classification de l'ostéoarthrite.

Keywords

musculoskeletal imaging, artificial intelligence, deep learning, machine learning, neural networks, appropriate utilization, image interpretation, image acquisition

Artificial intelligence (AI) has the potential to redefine every step of the imaging value chain, including both interpretive and noninterpretive components. It is important for radiologists to acquire expertise in AI developments in order to become leaders in their upcoming clinical implementation. Several excellent review articles have explored applications of AI in musculoskeletal (MSK) imaging,[1-5] including a recent issue of Seminars in Musculoskeletal Radiology dedicated to AI.[6] The present article provides an overview of the impact of AI through the entire imaging cycle of MSK radiology, from the placement of the requisition to the generation of the report, with a Canadian perspective.

[1] Department of Diagnostic Radiology, McGill University Health Center, Montreal, Quebec, Canada
[2] Department of Radiology, NYU Langone Medical Center/NYU Langone Orthopedic Center, New York, NY, USA
[3] Department of Orthopedic Surgery, NYU Langone Medical Center/NYU Langone Orthopedic Center, New York, NY, USA

Corresponding Author:
Natalia Gorelik, MD, Department of Diagnostic Radiology, McGill University Health Center, 1001 Decarie Blvd, Montreal, Quebec, Canada H4A 3J1.
Email: natalia.gorelik@mail.mcgill.ca

Ordering of Imaging Tests

Canada, like many other countries, has witnessed an increase in utilization of medical imaging over the past years. There is evidence of some inappropriate use; for example, in one study by Manta et al on the appropriateness of referrals for outpatient magnetic resonance imaging (MRI) of the hip, 32.1% of the requests were deemed inappropriate.[7] Imaging of the spine for uncomplicated low back pain and MRI of joints in older patients with osteoarthritis are important drivers of overuse in MSK radiology. Overutilization of imaging not only leads to unnecessary health care expenditures but may also have some potential harms, such as radiation exposure in the setting of computed tomography (CT) and potential risks associated with gadolinium-based contrast administration for MRI. In the setting of limited resources in the Canadian public health care system with lengthy waiting times for MRI, responsible stewardship of imaging resources is essential.

Multiple factors contribute to excess use of imaging, including rapidly evolving technology and the exponential growth of the medical knowledge base, making it challenging for clinicians to keep track of current evidence-based imaging guidelines. Efforts to reduce overutilization have been implemented. For example, the Choosing Wisely Canada campaign discourages the use of spine imaging for low back pain without red flags or the use of ankle X-ray series in adults with minor trauma.[8] There was also an initiative by the Ontario government to withdraw public insurance coverage for imaging of uncomplicated lower back pain.[9] Another attempted approach, which is more laborious, is the manual triage of imaging requisitions for appropriateness for MRI and CT arthrography of the knee, hip, and shoulder based on the presence of osteoarthritis on radiographs.[10] Clinical decision support (CDS) systems could also help reduce utilization rates, as noted in some American studies.[11] The Canadian Association of Radiologists (CAR) has collaborated with Medicalis Inc, a Canadian health care technology company based in Waterloo, Ontario, to integrate the CAR guidelines into an electronic CDS software.[12]

There has been a long-standing interest in using AI techniques to assist clinicians with medical decisions, including for radiology order entry, with early works by Kahn and Swett dating over 3 decades ago.[13-15] A CDS created with machine learning algorithms could automatically evaluate the clinical query considering a holistic clinical picture using extracted information from the electronic medical records, such as symptoms, physical examination findings, laboratory results, pathology reports, and previous imaging data to determine which imaging examination is most appropriate based on local guidelines.[3] The CDS would be seamlessly integrated to the computerized order entry system and would provide immediate guidance to clinicians, without disrupting the normal workflow.

Automatic Protocoling

Protocoling is a crucial step in the radiology workflow to ensure that the optimal test is performed to allow proper diagnosis. Alas, it is also a very time-consuming task, often performed by radiologists or radiology trainees. Radiologist time and abilities could be put to better use by having a machine protocol requisitions, with radiologists remaining available to verify that the adequate protocol was selected when technologists have doubts.

Trivedi et al evaluated a deep learning–based natural language classifier from IBM Watson which assessed free-text clinical indications to determine automatically whether to protocol the MRI examination with or without intravenous contrast.[16] For cases which were protocolled identically by the original protocol and the second evaluator, the system demonstrated up to 90% agreement.

Lee assessed a convolutional neural networks (CNNs) classifier which utilized short-text classification to evaluate whether MSK MRI studies should be performed according to a routine or tumor protocol.[17] The κ agreement for protocol assignment by the CNN and radiologists was 0.88. The system demonstrated a sensitivity of 92.10%, a specificity of 95.76%, an area under curve (AUC) of 0.977, and an overall accuracy of 94.2%.

An AI-based protocoling system could screen the patient's records for contrast allergy, renal dysfunction, pregnancy, noncompatible implantable devices, and presence of metallic foreign bodies which could impact on the type of imaging which the patient can safely undergo.

Artificial intelligence could also help optimize MR protocols. Until now, the selection of imaging planes and pulse sequences for a protocol was based on an assessment by human reviewers. However, this task can be performed using machine learning. For example, Richardson demonstrated the ability of a CNN to evaluate the value of MR sequences for the diagnosis of anterior cruciate ligament tears.[18]

Scheduling

Given the long waiting times for MRI in Canada, AI tools to automatically prioritize more urgent exams would be helpful to avoid delays in diagnosis and treatment of time-sensitive conditions, such as sarcoma. The radiology appointments could be automatically coordinated with clinical follow-up appointments or dialysis sessions, when applicable. They should also be planned for the adequate scanner, such as dual-energy CT or a 1.5T strength magnet to reduce metal artifact, if there is metallic hardware at the area of interest, with the AI system screening for the presence of hardware in the clinical and past radiological data.[4]

Moreover, AI could help in maximizing patient throughput to optimize the use of the limited MRI resources in the Canadian setting. For example, Nelson et al presented complex, nonlinear, high-dimensional models using machine learning to predict missed MRI appointments.[19] A "no-show" is a lost opportunity for another patient to be scanned or to undergo an image-guided procedure and contributes to waiting times; therefore such predictive models could help adapt booking strategies. The work by Muelly et al suggests that MR scanner utilization could be increased by improving scheduling

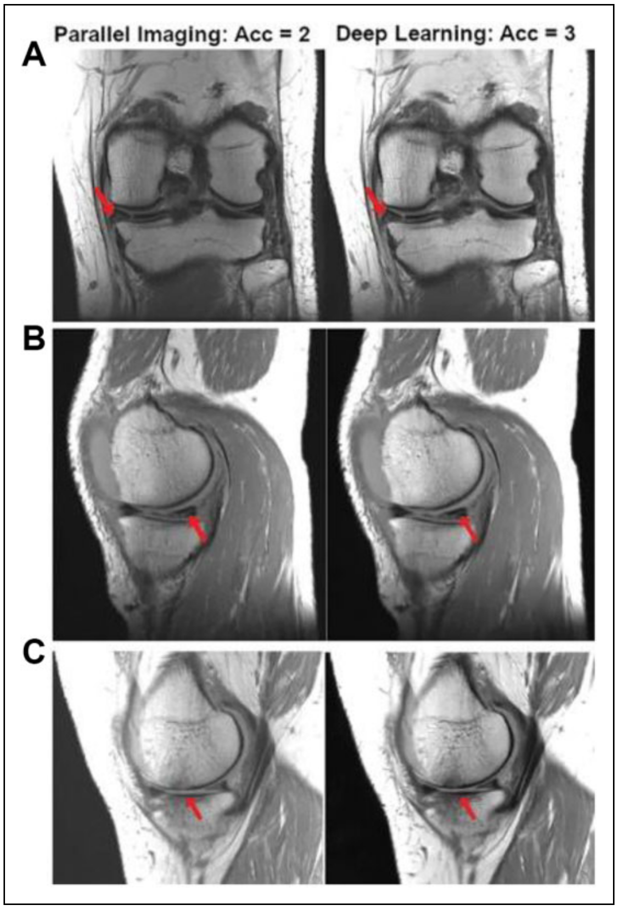

Figure 1. Examples of magnetic resonance images reconstructed with parallel imaging (left) and variational network deep learning (right), both adequately depicting (A, B) medial meniscal tears and (C) cartilage thinning with subchondral marrow signal alterations in 3 different patients. Reprinted Figure 10 in Johnson et al.[21]

efficiency using dynamic exam slot lengths determined by a feed-forward neural network.[20]

Magnetic Resonance Imaging Image Acquisition

The use of limited MRI resources can also be optimized by reducing scan times, thereby increasing patient throughput. Artificial intelligence tools can help accelerate MRI examinations, such as with undersampling and super-resolution.[21] Such techniques have permitted the acquisition of excellent quality images, without compromising diagnostic accuracy (Figure 1). To foster developments in image reconstruction for accelerated MRI, in a collaborative effort, Facebook AI Research and NYU Langone Health released the fastMRI data set.[22] It is the first publicly available large deidentified imaging data set, comprised of MRI k-space data as well as Digital Imaging and Communications in Medicine images from knee MRI examinations. Artificial intelligence systems may also provide automated quality control, thereby reducing the need to recall patients for repeat examinations.[23]

Another exciting innovation in the production of MR images is the creation of synthetic MR images from CT images. Lee et al studied the use of generative adversarial networks to transform spine CT images into axial T2-weighted MR images (Figure 2).[24] When 2 experienced MSK radiologists evaluated the similarities between the synthetic and the real MR images (Figure 3) based on the disc signal, degree of disc protrusion, muscle, fat tissue, facet joint signal, degree of stenosis, thecal sac, bone, and overall appearance, the average similarity was 80.2%. When 2 radiologists, 2 spine surgeons, and 2 residents blindly classified real and synthetic MR images, the failure rate ranged from 0% to 40%. On quantitative analysis, the mean absolute error value of synthetic MR images was 13.75 to 34.24 pixels (average 21.19 pixels; Figure 4) and the peak signal to noise ratio of 61.96 to 68.16 dB (mean 64.92 dB). Generating synthetic MR images from CT spine images may be particularly useful for patients who are unable to undergo MRI.

Computed Tomography Image Acquisition

Developments in AI may also help improve image quality in computed tomography. For example, they may assist in decreasing artifacts related to orthopedic hardware. Zhang and Yu described the use of a CNN for metal artifact reduction which merges original and corrected images data for artifact suppression.[25] Artificial intelligence algorithms have also shown promise in reducing CT radiation dose while still ensuring a high quality of images.[26]

Image Interpretation

Although AI can assist the imaging value chain in many of its components, it is AI's capacity to detect findings and suggest diagnoses that has received the most attention in recent years. The following section outlines some of AI's achievements in image interpretation for MSK radiology.

Bone Age

Leaps in AI developments have been facilitated by competitions. Most notably, in the field of computer vision, the 2012 ImageNet Large Scale Visual Recognition Challenge played a major role in promoting the advancement of CNNs. The first prize of that competition was won by the Canadian team of Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, who presented a CNN-based algorithm now known as AlexNet, which demonstrated an impressive top-5 test error rate of 15.3%, better than the second-best model by 10.9%.[27] Their work illustrated the advantages of CNNs and played a highly influential role in the field of computer vision.

Similarly, in the field of Radiology, the Radiological Society of North America Pediatric Bone Age Machine Learning Challenge promoted collaborative efforts in furthering AI developments in medical imaging through competitions.[28] This challenge aimed at determining the best machine learning–based approaches for most accurately determining bone age.

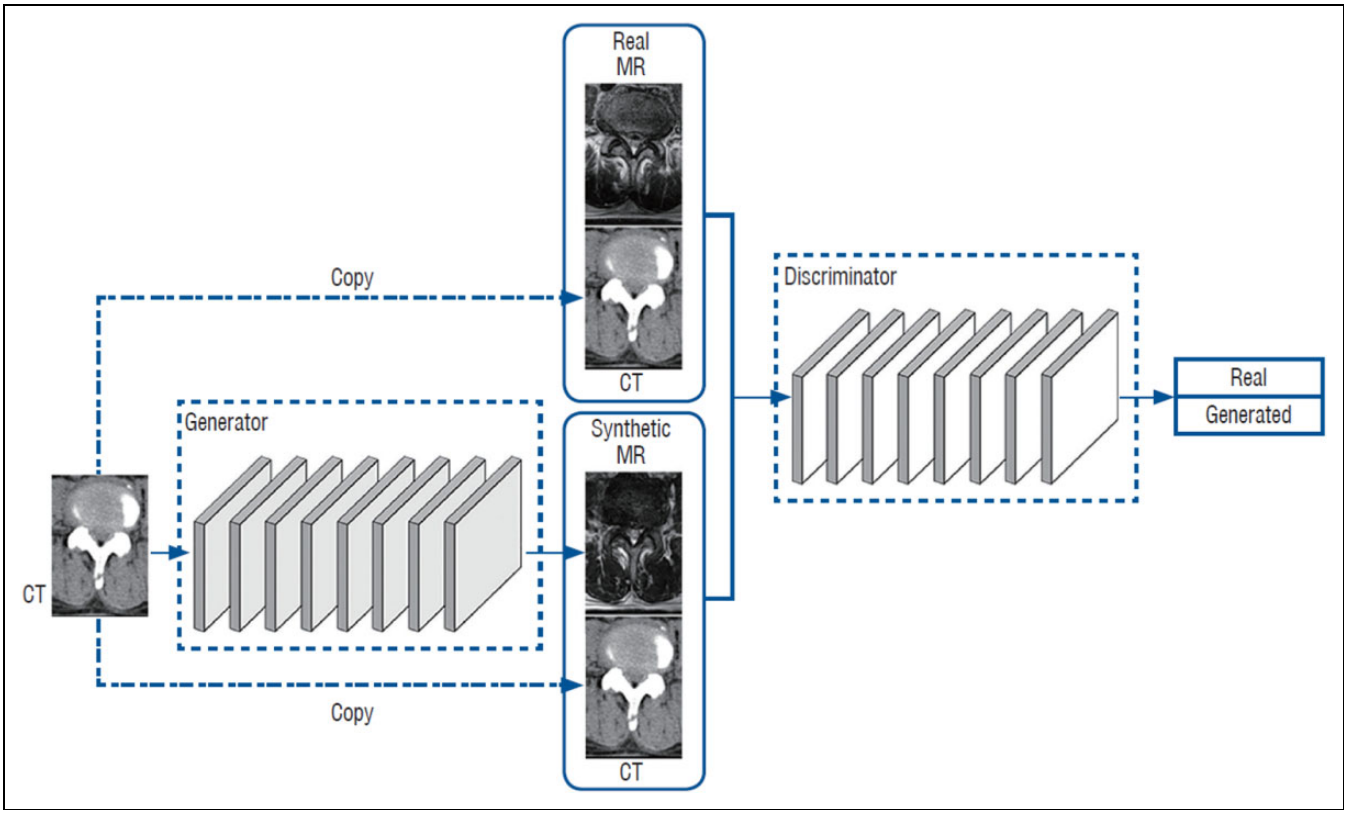

Figure 2. Flow diagram illustrating the synthesis of magnetic resonance (MR) images from computed tomography (CT) images by generative adversarial networks (GANs). The generator learns to produce synthetic MR images which cannot be differentiated from real MR images by an adversarially trained discriminate network. Reprinted Figure 1 from Lee et al.[24]

There has been a long-standing interest in automating radiographic evaluation of skeletal maturity as the traditional ways, such as using the Greulich and Pyle atlas or the Tanner-Whitehouse method, are time-consuming, have a limited precision, and are prone to substantial inter- and intraobserver variation. The model developed by Drs Alexander Bilbily and Mark Cicero from University of Toronto won the first prize.[28] Their algorithm used the Inception V3 architecture for pixel information along with concatenation for sex information, together with data augmentation.

Although many AI developments for image interpretation remain at the research stage, including several published techniques for skeletal age determination, BoneXpert is a commercially available machine learning tool for automated bone age evaluation.[29] This method uses conventional machine learning techniques, automatically segmenting 15 bones, then assessing bone age based on 13 bones using handcrafted features of shape, intensity, and texture (Figure 5). It outputs the skeletal maturity according to the Greulich and Pyle or Tanner-Whitehouse standards. BoneXpert has a precision of 0.17 years with the Greulich and Pyle method, nearly 3 times better than humans.

Body Composition

Differences in body composition may have important health outcomes implications. For example, increased visceral fat is associated with impaired glucose and lipid metabolism with increased cardiometabolic risk and all-cause mortality, decreased bone mineral density, nonalcoholic fatty liver disease, and increased risk for neoplasm.[30] Decreased muscle mass or sarcopenia is linked to poorer prognosis in intensive care unit patients, soft tissue sarcoma recurrence, as well as increased mortality post liver transplant, after colorectal surgery, and in hepatocellular carcinoma.[30] Various methods exist to measure the different tissue types in the human body. Conventional approaches include anthropomorphic measures, bioelectrical impedance analysis, and dual-energy X-ray absorptiometry.[30] The quantitative assessment of body composition may be obtained from segmentation of cross-sectional images, including computed tomography and MRI examinations. Manual segmentation is tedious and impractical in the clinical setting. Such a task is however amenable to automation by machine learning techniques. Weston et al developed a completely automated deep learning algorithm for the segmentation of abdominal CT examinations aimed for body composition analysis (Figure 6).[31] Their CNN model utilized a U-Net architecture trained on 2430 CT examinations. It demonstrated Dice scores with mean (standard deviation) of 0.98 (0.03), 0.96 (0.02), and 0.97 (0.01) in the test set of 270 CT examinations and 0.94 (0.05), 0.92 (0.04), and 0.98 (0.02) in a data set of 2369 patients with hepatocellular carcinoma, for the subcutaneous, muscle, and visceral adipose tissue compartments,

Figure 3. Axial images from the 15 computed tomography (CT) examinations (CT01-CT15) (left) from which synthetic T2-weighted magnetic resonance (MR) images (middle) were generated. The right-sided images are the corresponding real MR images. Reprinted Figure 6 from Lee et al.[24]

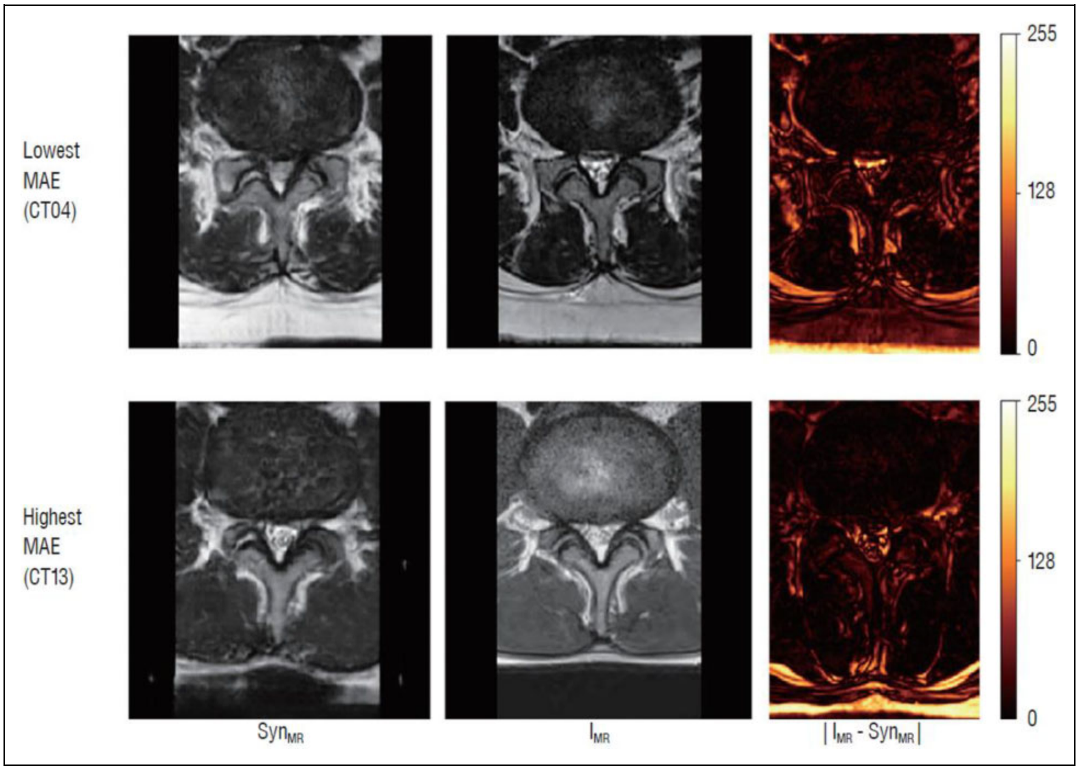

Figure 4. Cases with the lowest (top) and highest (bottom) mean absolute error (MAE) between the real (IMR) and synthesized (SynMR) magnetic resonance images. Reprinted Figure 7 from Lee et al.[24]

respectively. Its results were comparable to or better than those of expert manual segmentation.

Automated body composition measurements may serve for opportunistic screening in radiology. That is, information other than the primary indication for the examination may be gathered from the study without need for additional radiation or exam time, thereby generating added value.[1] Conceivably, such an algorithm could evaluate all the abdominal CTs performed at an institution and provide body composition information in the report, which could be used as a clinical indicator or for epidemiological research.

Bone Fragility

Another chance for opportunistic radiologic screening is the assessment for bone fragility. Early detection of decreased bone mineral density offers an opportunity for prompt treatment aimed at decreasing the risk of fracture and improving quality of life and survival.[32] Several computer-aided diagnosis systems have been studied for the evaluation of bone quality on dental panoramic radiographs.[33] Such a tool gives the chance to incidentally screen for osteoporosis on visits aimed at dental care. Kathirvelu et al presented such an semiautomated measure of mandibular cortical thickness on a dental panoramic radiograph to identify patients at risk for low bone mineral density.[32] Their approach consisted first of selecting a region of interest inferior to the mental foramen with median filtering and intensity normalization for image enhancement. Then,

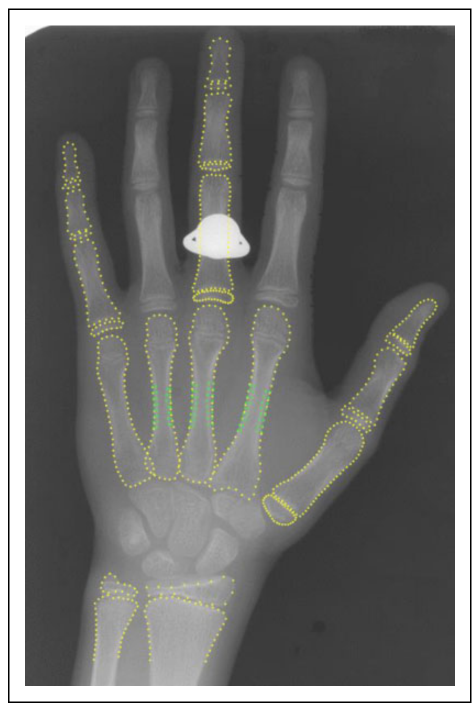

Figure 5. Fifteen bones were automatically reconstructed by the active appearance model, to subsequently compute the age of individual bones using shape, intensity, and texture scores. Of note, the ring on the third proximal phalanx does not impede reconstruction, as the "blind spot method" was utilized by the active appearance model. Reprinted Figure 2 from Thodberg et al.[29]

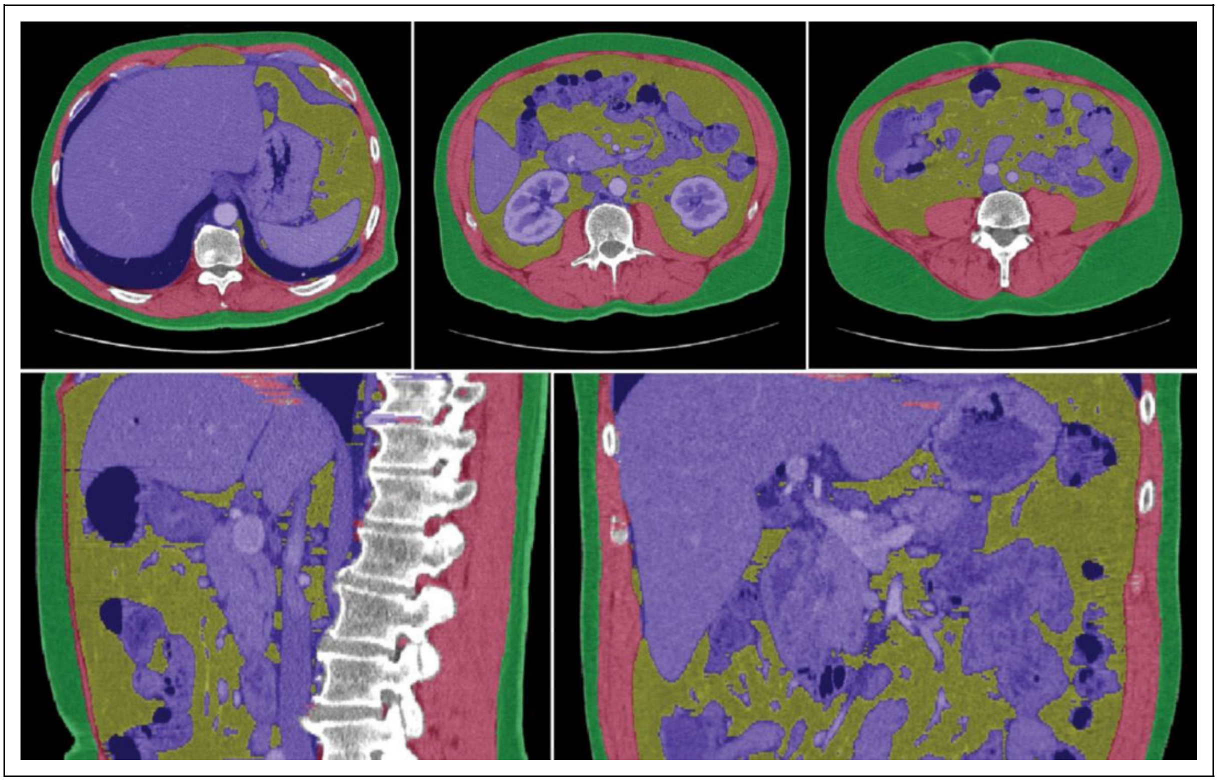

Figure 6. Automated 3-dimensional segmentation of subcutaneous adipose tissue, muscle, visceral adipose and fat-free tissue, and bone compartments on abdominal computed tomography for the quantification of body composition. Reprinted Figure 5 from Weston et al.[31]

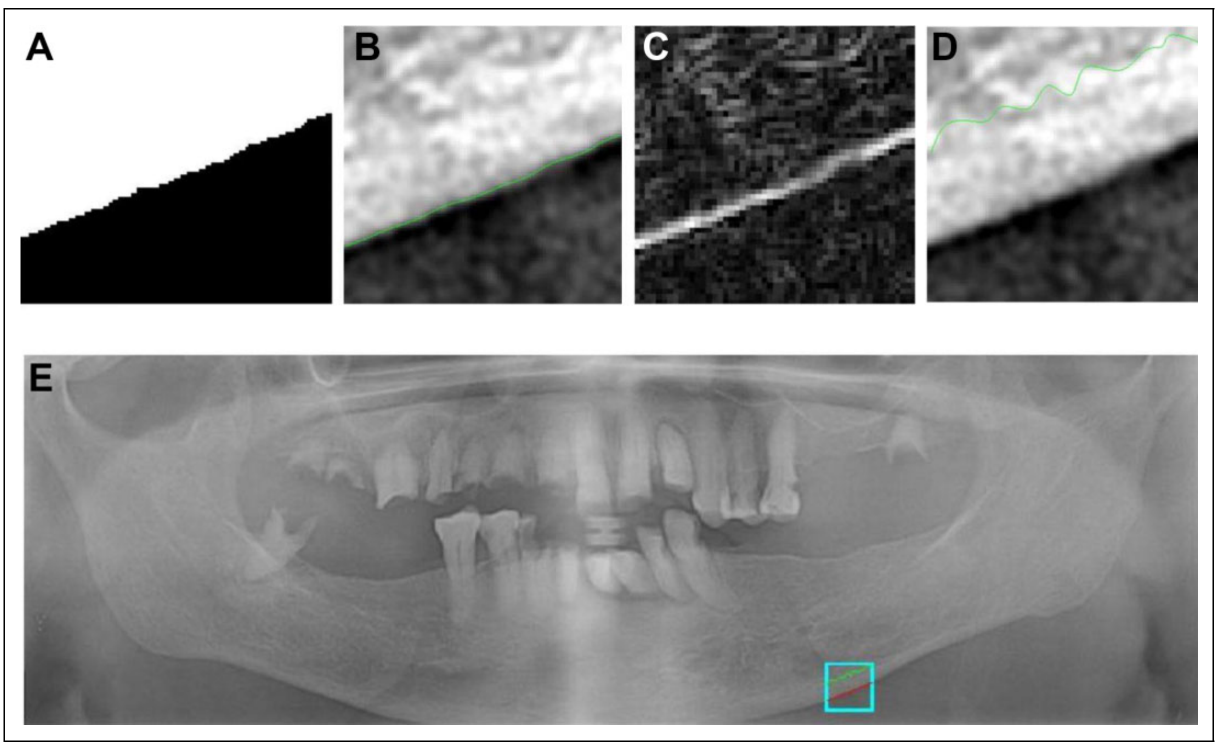

Figure 7. Performance of mandibular cortical thickness measurement depicting (A) the segmented foreground, (B) the identified lower boundary, (C) the Haar wavelet magnitude, (D) the identified upper boundary, and (E) the identified upper and lower boundaries on the dental panoramic radiograph. Reprinted Figure 2 from Kathirvelu et al.[32]

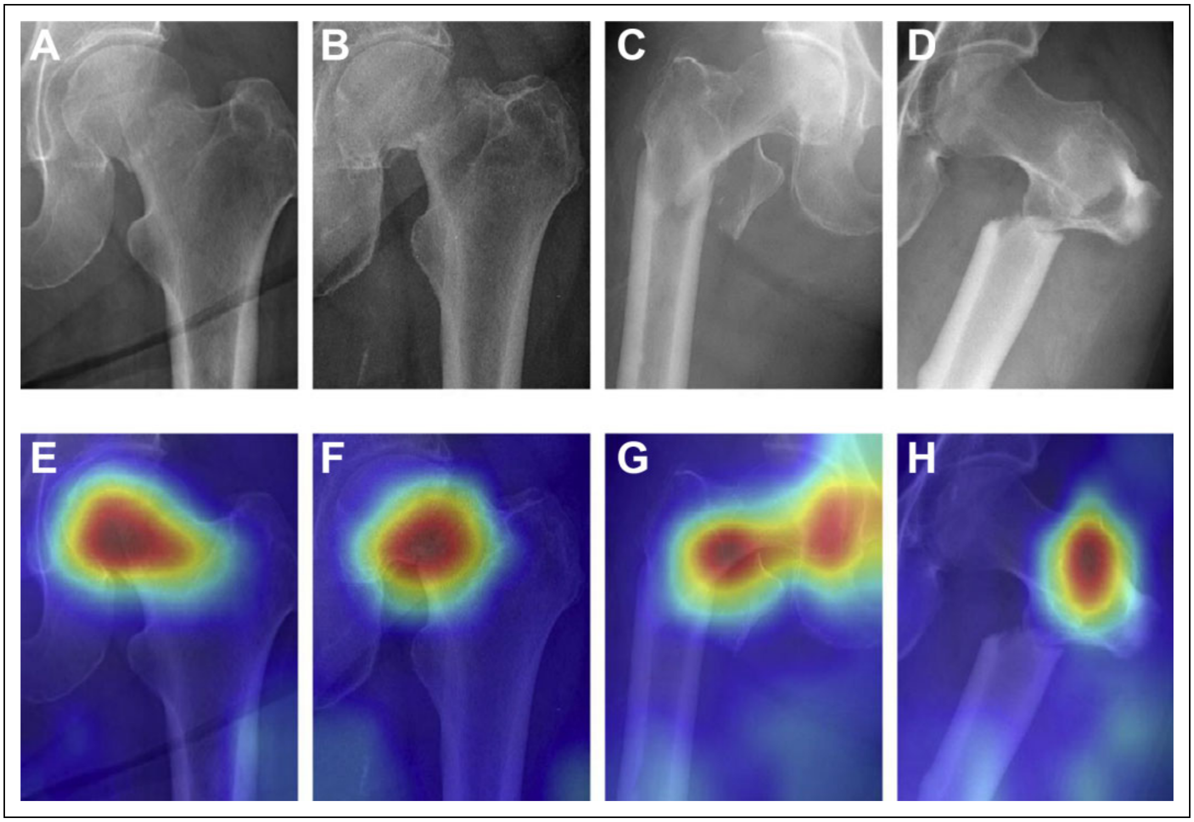

Figure 8. Saliency maps in 4 representative examples of automatically detected hip fractures, including a nondisplaced transcervical (A, B), angulated transcervical (C, D), displaced and minimally angulated intertrochanteric (E, F), and displaced and angulated subtrochanteric (G, H) fractures. Areas which most contributed to image classification were the fracture site, except for subtrochanteric fractures, where it was the trochanteric region. Reprinted Figure 7 from Yu et al.[39]

Otsu segmentation was used for lower boundary detection with smothering using Snakes algorithm. The upper boundary was detected with a Haar wavelet operation. The mandibular cortical thickness was determined by the distance between the upper and lower boundary pixels (Figure 7). The semiautomated mandibular cortical thickness measure correlated with the manual measure, with a Pearson $r = 0.96$ ($P < .01$). With a mandibular cortical thickness of 2.3 mm or less, the sensitivity, specificity, and accuracy for detecting low bone mineral density as detected at the right femur were of 91%, 70%, and 79%. The use of deep CNN-based computer-assisted diagnosis systems developed for the identification of osteoporosis on panoramic dental radiographs.[34] Deep CNN-based computer-assisted diagnosis systems have also been developed for the identification of osteoporosis on panoramic dental radiographs.[34]

Opportunistic osteoporosis screening can also be performed on computed tomography. Pan et al described a deep learning–based system for bone mineral density measurement on low-dose chest CTs performed for lung cancer screening.[35] In this system, segmentation was performed using a 3D CNN model with U-net architecture and dense connections. Conventional image processing algorithms were used for vertebral body labeling. For bone mineral density measurement, mean CT numbers of trabecular area in cylinder volumes of interest at target vertebral bodies were acquired, using segmentation mask with geometric operations. These numbers were mapped to bone mineral density values with a one-degree linear function. The Dice coefficient was 86.6% for vertebral body segmentation and 97.5% for vertebral body labeling. Good agreement was demonstrated between the predicted bone mineral density by the model and the ground truth obtained by quantitative computed tomography, with correlation coefficients of 0.964 to 0.968 (mean errors of 2.2-4.0 mg/cm^3). The AUC was 0.927 for osteoporosis and 0.942 for low bone mineral density.

There has also been interest in using machine learning algorithms to predict osteoporotic fractures from MRI data.[36] Convolutional neural networks have also been applied for automatic segmentation of the proximal femur, which would be useful for measuring bone quality on MRI.[37]

Fractures

In many Canadian hospitals, there is no routine after-hours coverage by radiologists for the interpretation of radiographs. In the evening and overnight, emergency department physicians therefore make decisions based on their own assessment of radiographs, which subsequently get interpreted by the radiologist during the daytime. This may lead to discrepancies

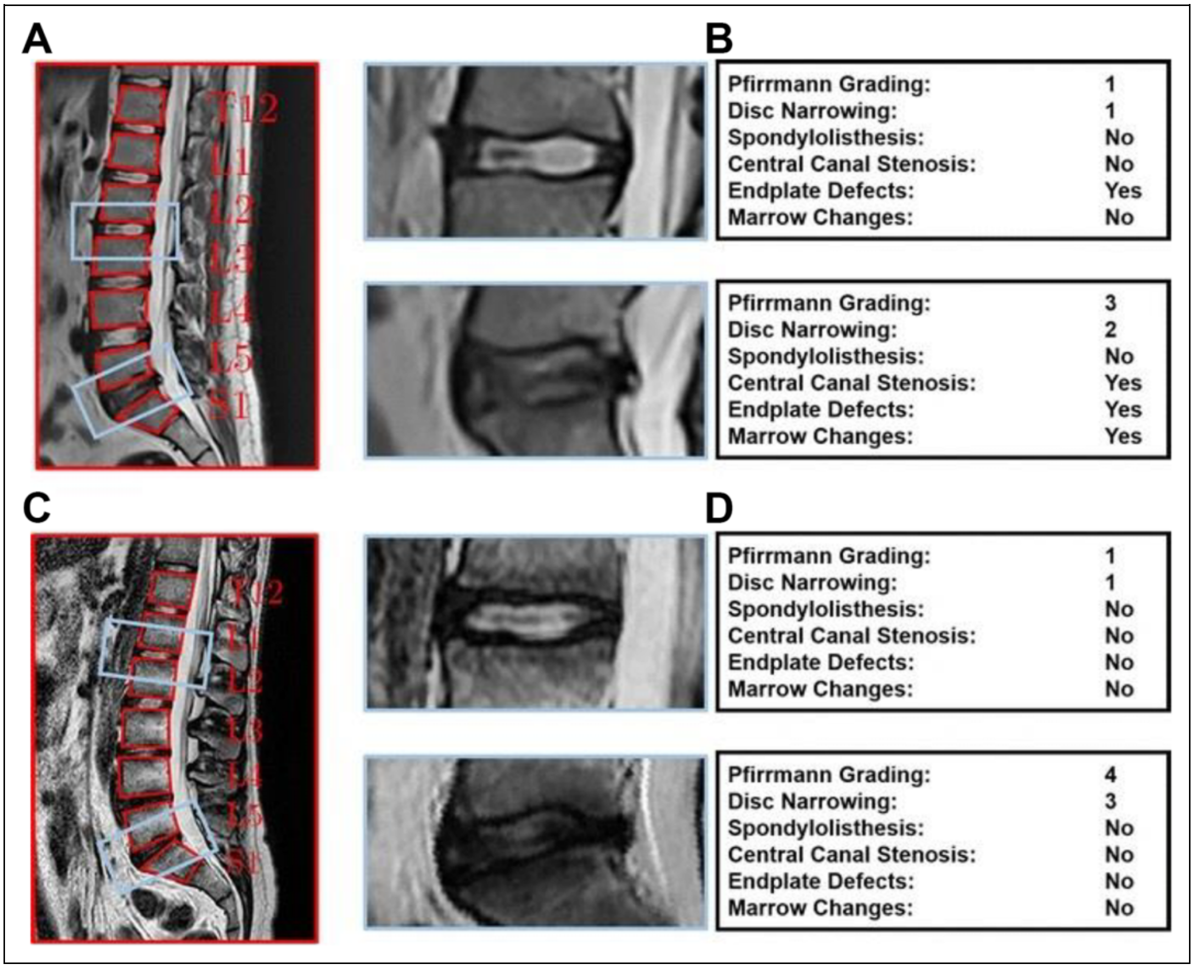

Figure 9. Midsagittal magnetic resonance images of the spine with detected vertebrae (red boxes) and examples of extracted disc regions (blue boxes) (A, C), with the associated automated evaluation of the Pfirrmann grading, disc narrowing, spondylolisthesis, central canal stenosis, end plate defects, and Modic changes (B, D). Reprinted Figure 5 from Jamaludin et al.[44] Coloured version of this figure is available online.

which may potentially impact patients' outcomes, such as in the case of hip fractures where a delayed diagnosis and surgery may portend a poorer prognosis.[38] Artificial intelligence could aid with automatic fracture detection.

Multiple studies have assessed the use of AI for the detection (and classification) of fractures both in the axial and appendicular skeleton on radiographs and computed tomography and have shown promise (Figure 8).[2,33,39,40] In a systematic review which included 10 studies (8 for fracture detection at the ankle, hand, hip, spine, wrist, and ulna; 1 for classification of femoral diaphyseal fractures; and 1 for both detection and classification of proximal humeral fractures), the area under the receiving operating characteristic curve (AUC) achieved in 5 studies was 0.95 to 1.0, and the accuracy for 7 studies was 83% to 98% for fracture detection.[40] The AUC was 0.94 in 1 study and the accuracy was 77% to 90% in 2 studies for fracture classification. The performance of AI was higher than that of human readers in 2 studies for the detection and classification of hip and humeral fractures and was similar to that of human readers in 1 study for the detection of wrist, hand, and ankle fractures. Authors noted that fractures in the studied areas are frequently displaced therefore easier to identify. Artificial intelligence models could possibly be less accurate for less evident fractures, such as nondisplaced femoral neck or scaphoid fractures.[40] A limitation of CNN-based models is that they must be trained for separate body parts, unlike humans who translate their knowledge of fractures to any site.[2]

Spine Imaging

The interpretation of multilevel degenerative spine disease can be a tedious task, prone to high interobserver variability. This task is amenable to AI, which could improve its consistency. Several groups of researchers have presented AI algorithms for the detection and labeling of vertebral bodies and intervertebral discs on CT and MRI.[41-44] Certain models can also use this information to grade segmental pathology.[44] Jamaludin et al presented a method to identify and label vertebrae and discs on MRI with an accuracy of 95.6%. Subsequently, a CNN was used to evaluate the Pfirrmann grading, disc narrowing,

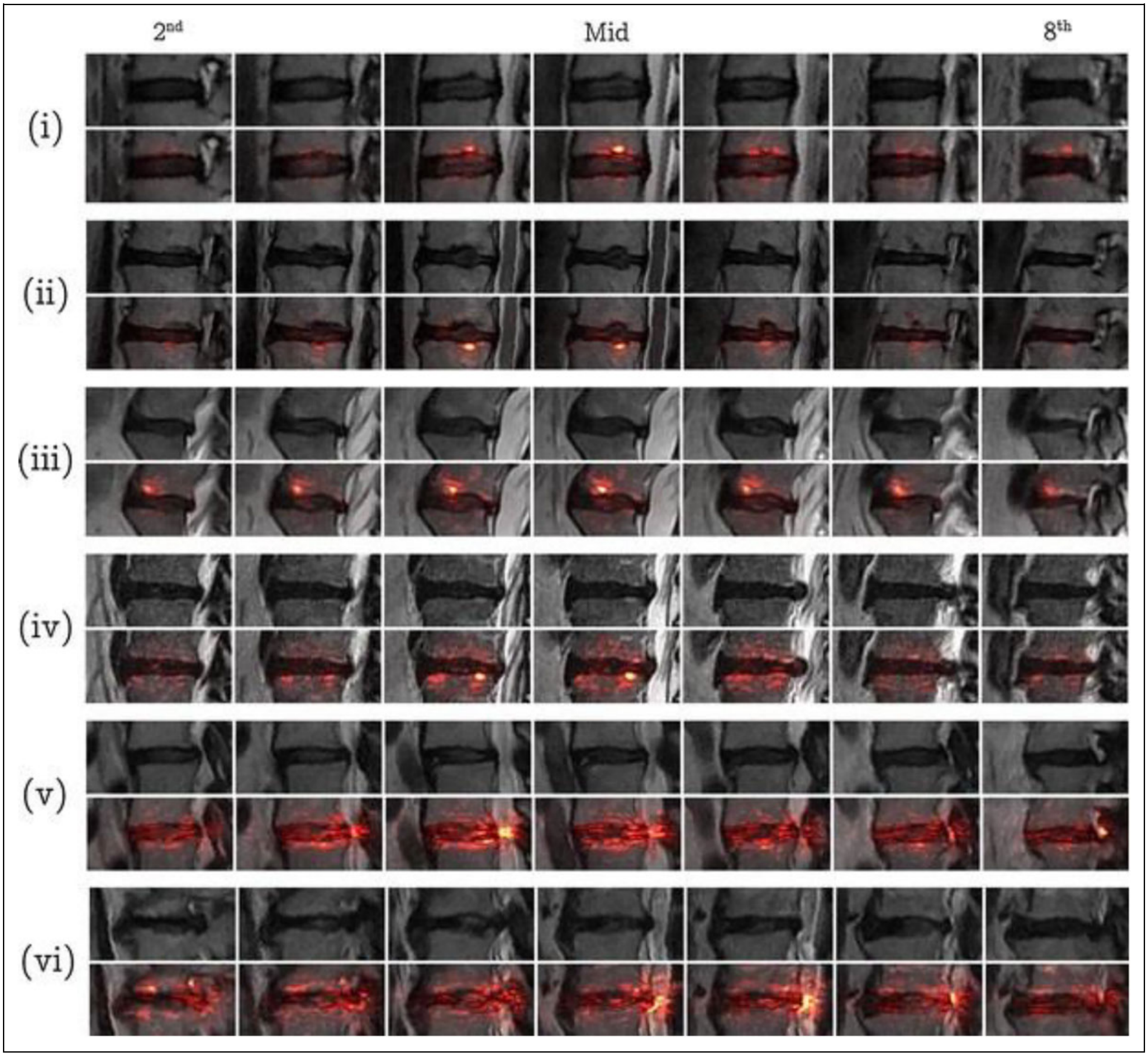

Figure 10. Examples of pairs of disc volumes (upper) and associated evidence hotspots (lower) on 7 of 9 slices (second slice on left and eighth slice on the right) depicting voxels that were most responsible for the classification. Images show (i) upper end plate defects, (ii) lower end plate defects, (iii) upper marrow change, (iv) lower marrow change, (v) spondylolisthesis, and (vi) central canal stenosis. Reprinted Figure 7 from Jamaludin et al.[44]

spondylolisthesis, central canal stenosis, end plate defects, and marrow signal variations, with a performance comparable to a radiologist (Figure 9).[44] Figure 10 illustrates the voxels which played the greatest importance in determining the grade by their automated technique. Publicly available data sets, such as SpineWeb[45] and the MICCAI 2018 Challenge on Automatic Intervertebral Disc Localization and Segmentation[46] data set, help promote the development of this line of research. Other applications of AI in spine imaging include the use of CNNs to distinguish tuberculous from pyogenic spondylitis[47] and radiographic measurements of spinal alignment.[48,49]

Musculoskeletal Oncology

The detection of metastatic bone lesions[50,51] and the monitoring of their temporal evolution on serial scans can be facilitated by AI methods.[52] Machine learning can also be useful in assessing treatment response for osseous metastases. In a study by Acar et al, sclerotic bone lesions in patients with prostate cancer could be classified as metastases (with 68Ga-prostate-specific membrane antigen expression on positron emission tomography/computed tomography) versus sclerotic lesions with complete response (without 68Ga-prostate-specific membrane antigen expression) using CT texture analysis and

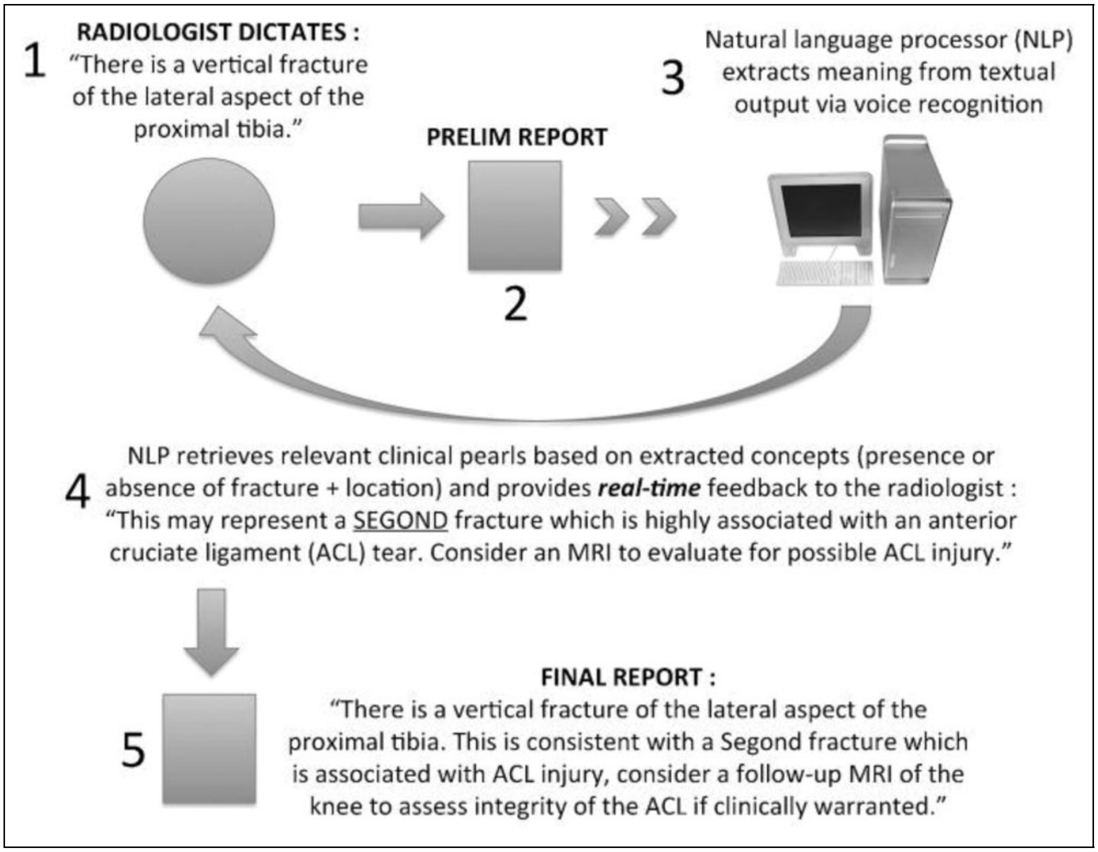

Figure 11. Flow diagram illustrating the function of a decision support system based on speech recognition and natural language processing. When the radiologist dictates the fracture description, the natural language processor identifies the disease (fracture) and anatomy (tibia) concepts from the unstructured text and proceeds to retrieve pertinent fracture knowledge (eg, classification, diagnostic criteria, disease probability). Reprinted Figure 1 from Do et al.[78]

machine learning, with the best performance obtained using a weighted k-nearest neighbor technique, which demonstrated an AUC of 0.76.[53] Artificial intelligence can also help diagnose the origin of metastases, such as in a study by Lang et al, in which radiomics and deep learning methods could differentiate spinal metastases from lung and other cancers on dynamic contrast-enhanced MRI with an accuracy of up to 0.81.[54]

The diagnosis of primary bone tumors can also be assisted by AI, with works published on this topic by Lodwick et al as early as the 1960s.[55] Prediction of tumor recurrence can also be achieved with AI, as illustrated by the work of He et al who used an Inception V3 CNN to predict local recurrence of giant cell tumor of bone based on MR and clinical data.[56]

Similarly, AI may help with the diagnosis of soft tissue tumors.[57,58] For example, Malinauskaite et al presented a machine-learning classifier which could distinguish between lipoma and liposarcoma on MRI using radiomic features with an AUC of 0.926, with a performance superior to that of 3 MSK radiologists.[58] Moreover, using radiomic features and machine learning classification techniques, the histopathological grade of soft tissue sarcomas can be predicted with an AUC of up to 0.92 (accuracy of 0.88).[59] Machine learning methods may also assist in monitoring posttreatment changes of soft tissue sarcoma on MRI.[60]

Osteoarthritis and Cartilage Imaging

The aging Canadian population entails an increased prevalence of osteoarthritis with an associated rise in demands on the health care system, including on medical imaging.[61] The radiographic grading of osteoarthritis may be automated by AI, which would expedite and standardize its interpretation. Thomas et al presented an automated deep learning model to stage knee osteoarthritis according to the Kellgren-Lawrence system using the Osteoarthritis Initiative radiographs, which reached an average F1 score of 0.64 and accuracy of 0.66 compared to the best individual radiologist's F1 score of 0.60 and accuracy of 0.60, with no manual image preprocessing required.[62] The qualitative and quantitative assessment of cartilage on MRI can also be augmented with AI methods, which would make the assessment of osteoarthritis more precise and consistent. Several deep learning systems for the automated detection of cartilage lesions have already been developed.[63,64]

Miscellaneous

Multiple other applications of AI for MSK imaging interpretation have been published, including detection, classification, and segmentation tasks. Some examples include the identification of anterior cruciate ligament[65] and meniscal[66] tears on MRI of the knee, detection of epidural masses on CT,[67] evaluation of the joint space,[68] and presence of erosions[69] in rheumatoid arthritis, sex estimation from sacrum and coccyx CT[70] or from hand and wrist radiographs,[71] assessment of Achilles tendon healing,[72] and automated segmentation of peripheral nerves of the thigh,[73] supraspinatus muscle,[74] proximal femur,[37] and bones of the wrist[75] on MRI. Additional repetitive and time-consuming tasks suitable to automation will likely be tackled in coming years.

Results Reporting

Artificial intelligence could transform the production and utilization of radiology reports. Already, speech recognition has revolutionized how reports are created. Speech recognition could be further optimized with deep learning methods.[76] Artificial intelligence could translate radiology reports into other languages, such as from English to French and vice versa, which would be helpful in Canadian regions where patients and providers use French and English languages to various extent. Artificial intelligence can also automatically generate structured reports, such as in the model published by Lee et al for bone age.[77] Do et al presented a natural language processing system able to identify fracture and anatomy data from text obtained with a speech recognition software and synchronously extract fracture knowledge.[78] Such a system could suggest management recommendations to the radiologist during the dictation of a report. For example, when the radiologist reports a Segond fracture, the AI system could suggest further evaluation with an MRI to assess for a potential anterior cruciate ligament injury (Figure 11). Tan et al described a natural language processing system able to extract lumbar spine findings related to low back pain on MRI and X-ray radiology reports.[79] As the authors suggested, such a tool could be beneficial in research, to help elucidate the link between imaging findings, symptoms, and prognosis. Artificial intelligence can also be used to extract follow-up recommendations from radiology reports, which could help ensure that adequate management for actionable findings is obtained.[80]

In conclusion, a myriad of AI applications has been developed, including tools to improve appropriate ordering, study protocoling, scheduling, imaging acquisition, workflow optimization, hanging protocols, image interpretation, reporting of results, and billing. These developments hold promise to enhance radiologists' job satisfaction and patients' outcomes. However, few of these products are currently commercially available. There remain some challenges in adopting AI in the clinical workflow, such as the limited generalizability of certain algorithms, the "blackbox" phenomenon, and uncertainties surrounding medicolegal aspects of AI.[2,4] As academia and industry collaborate in solving these obstacles, it is likely that more AI applications will become integrated in the clinical practice in the near future.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Dr. Natalia Gorelik's research work is supported by an AUR GE Radiology Research Academic Fellowship Award.

References

1. Burns JE, Yao J, Summers RM. Artificial intelligence in musculoskeletal imaging: a paradigm shift. *J Bone Miner Res*. 2020;35(1):28-35. doi:10.1002/jbmr.3849
2. Chea P, Mandell JC. Current applications and future directions of deep learning in musculoskeletal radiology. *Skeletal Radiol*. 2020;49(2):183-197. doi:10.1007/s00256-019-03284-z
3. Gyftopoulos S, Lin D, Knoll F, Doshi AM, Rodrigues TC, Recht MP. Artificial intelligence in musculoskeletal imaging: current status and future directions. *AJR Am J Roentgenol*. 2019;213(3):506-513. doi:10.2214/ajr.19.21117
4. 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
5. Syed AB, Zoga AC. Artificial intelligence in radiology: current technology and future directions. *Semin Musculoskelet Radiol*. 2018;22(5):540-545. doi:10.1055/s-0038-1673383
6. Gyftopoulos S, Subhas N. Musculoskeletal Imaging applications of artificial intelligence. *Semin Musculoskelet Radiol*. 2020;24(1):1-2. doi:10.1055/s-0039-3400511
7. Manta A, O'Grady J, Bleakney R, et al. Determining the appropriateness of requests for outpatient magnetic resonance imaging of the hip. *Can J Surg*. 2019;62:224-226. doi:10.1503/cjs.003718
8. Choosing Wisely Canada. *Five Things Physicians and Patients Should Question by Canadian Association of Radiologists*. Published April 2014. Updated February 2020. Accessed April 2, 2020. https://choosingwiselycanada.org/radiology
9. Fine B, Schultz SE, White L, et al. Impact of restricting diagnostic imaging reimbursement for uncomplicated low back pain in Ontario: a population-based interrupted time series analysis. *CMAJ Open*. 2017;5(4):E760-E767. doi:10.9778/cmajo.20160151
10. Kandiah JW, Chan VWY, Luo J, et al. Reducing the volume of low-value outpatient MRI joint examinations in patients $\geq$55 years of age. *Can Assoc Radiol J*. 2020;71(1):83-91. doi:10.1177/0846537119885686
11. Canadian Agency for Drugs and Technologies in Health. Clinical decision support systems for diagnostic imaging. *Issues in Emerging Health Technologies*; 2012.
12. Canadian Association of Radiologists. *Clinical Decision Support [Internet]*. 2017. Accessed April 2, 2020. https://car.ca/patient-care/referral-guidelines/clinical-decision-support
13. Kahn CE, Kovatsis PG, Messersmith RN, Lehr JL. Automated entry of radiology requisition information with artificial-

intelligence techniques. *AJR Am J Roentgenol*. 1989;153(5): 1085-1088. doi:10.2214/ajr.153.5.1085
14. Kahn CE. Artificial intelligence in radiology: decision support systems. *Radiographics*. 1994;14(4):849-861. doi:10.1148/radiographics.14.4.7938772
15. Swett HA, Rothschild M, Weltin GG, Fisher PR, Miller PL. Optimizing radiologic workup: an artificial intelligence approach. *J Digit Imaging*. 1989;2(1):15-20. doi:10.1007/bf03168010
16. Trivedi H, Mesterhazy J, Laguna B, Vu T, Sohn JH. Automatic determination of the need for intravenous contrast in musculoskeletal MRI examinations using IBM Watson's natural language processing algorithm. *J Digit Imaging*. 2018;31(2):245-251. doi: 10.1007/s10278-017-0021-3
17. Lee YH.Efficiency improvement in a busy radiology practice: determination of musculoskeletal magnetic resonance imaging protocol using deep-learning convolutional neural networks. *J Digit Imaging*. 2018;31(5):604-610. doi:10.1007/s10278-018-0066-y
18. Richardson ML. MR protocol optimization with deep learning: a proof of concept. *Curr Probl Diagn Radiol*. 2019. doi:10.1067/j.cpradiol.2019.10.004
19. Nelson A, Herron D, Rees G, et al. Predicting scheduled hospital attendance with artificial intelligence. *NPJ Digit Med*. 2019;2(1): 1-7. doi:10.1038/s41746-019-0103-3
20. Muelly MC, Stoddard PB, Vasanwala SS. Using machine learning with dynamic exam block lengths to decrease patient wait time and optimize MRI schedule fill rate. *International Society for Magnetic Resonance in Medicine*. Honolulu, HI:2017.
21. Johnson PM, Recht MP, Knoll F. Improving the speed of MRI with artificial intelligence. *Semin Musculoskelet Radiol*. 2020; 24(1):12-20. doi:10.1055/s-0039-3400265
22. Knoll F, Zbontar J, Sriram A, et al. FastMRI: a publicly available raw k-space and DICOM dataset of knee images for accelerated mr image reconstruction using machine learning. *Radiol Artif Intell*. 2020;2(1):e190007. doi:10.1148/ryai.2020190007
23. Küstner T, Gatidis S, Liebgott A, et al. A machine-learning framework for automatic reference-free quality assessment in MRI. *Magn Reson Imaging*. 2018;53:134-147. doi:10.1016/j.mri.2018.07.003
24. Lee JH, Han IH, Kim DH, et al. Spine computed tomography to magnetic resonance image synthesis using generative adversarial networks: a preliminary study. *J Korean Neurosurg Soc*. 2020; 63(3):386. doi:10.3340/jkns.2019.0084
25. Zhang Y, Yu H. Convolutional neural network based metal artifact reduction in x-ray computed tomography. *IEEE Trans Med Imaging* 2018;37(6):1370-1381. doi:10.1109/TMI.2018.2823083
26. Kambadakone A. Artificial intelligence and CT image reconstruction: potential of a new era in radiation dose reduction. *J Am Coll Radiol*. 2020;17(5):649-651. doi:10.1016/j.jacr.2019.12.025
27. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems*; Lake Tahoe, Nevada, USA, December 2012, pp. 1097-1105.
28. Halabi SS, Prevedello LM, Cramer JK, et al. The RSNA pediatric bone age machine learning challenge. *Radiology*. 2019;290(2): 498-503. doi:10.1148/radiol.2018180736
29. Thodberg HH, Kreiborg S, Juul A, Pedersen KD. The BoneXpert method for automated determination of skeletal maturity. *IEEE Trans Med Imaging* 2009;28(1):52-66. doi:10.1109/TMI.2008.926067
30. Wang B, Torriani M. Artificial intelligence in the evaluation of body composition. *Semin Musculoskelet Radiol*. 2020;24(1): 30-37. doi:10.1055/s-0039-3400267
31. Weston AD, Korfiatis P, Kline TL, et al. Automated abdominal segmentation of CT scans for body composition analysis using deep learning. *Radiology*. 2019;290(3):669-679. doi:10.1148/radiol.2018181432
32. Kathirvelu D, Vinupritha P, Kalpana V. A computer aided diagnosis system for measurement of mandibular cortical thickness on dental panoramic radiographs in prediction of women with low bone mineral density. *J Med Syst*. 2019;43(6):148. doi:10.1007/s10916-019-1268-7
33. Gundry M, Knapp K, Meertens R, Meakin JR. Computer-aided detection in musculoskeletal projection radiography: a systematic review. *Radiography*. 2018;24(2):165-174. doi:10.1016/j.radi.2017.11.002
34. Lee JS, Adhikari S, Liu L, Jeong HG, Kim H, Yoon SJ. Osteoporosis detection in panoramic radiographs using a deep convolutional neural network-based computer-assisted diagnosis system: a preliminary study. *Dentomaxillofac Radiol*. 2018;48(1): 20170344. doi:10.1259/dmfr.20170344
35. Pan Y, Shi D, Wang H, et al. Automatic opportunistic osteoporosis screening using low-dose chest computed tomography scans obtained for lung cancer screening. *Eur Radiol*. 2020;30(7): 4107-4116. doi:10.1007/s00330-020-06679-y
36. Ferizi U, Besser H, Hysi P, et al. Artificial intelligence applied to osteoporosis: a performance comparison of machine learning algorithms in predicting fragility fractures from MRI Data. *J Magn Reson Imaging*. 2019;49(4):1029-1038. doi:10.1002/jmri.26280
37. Deniz CM, Xiang S, Hallyburton RS, et al. Segmentation of the proximal femur from MR images using deep convolutional neural networks. *Sci Rep*. 2018;8(1):16485. doi:10.1038/s41598-018-34817-6
38. Weller I, Wai EK, Jaglal S. The effect of hospital type and surgical delay on mortality after surgery for hip fracture. *J Bone Joint Surg Br*. 2005;87(3):361-366. doi:10.1302/0301-620x.87b3.15300
39. Yu JS, Yu SM, Erdal BS, et al. Detection and localisation of hip fractures on anteroposterior radiographs with artificial intelligence: proof of concept. *Clin Radiol* 2020;75(3):e231-237.e239. doi:10.1016/j.crad.2019.10.022
40. Langerhuizen DWG, Janssen SJ, Mallee WH, et al. What are the applications and limitations of artificial intelligence for fracture detection and classification in orthopaedic trauma imaging? a systematic review. *Clin Orthop Relat Res*. 2019;477(11): 2482-2491. doi:10.1097/CORR.0000000000000848
41. Pastor AJ, Alberich-Bayarri A, Guarinos BF, et al. Automated vertebrae localization and identification by decision forests and image-based refinement on real-world CT data. *Radiol Med*. 2020;125(1):48-56. doi:10.1007/s11547-019-01079-9

42. Lessmann N, van Ginneken B, de Jong PA, Isgum I. Iterative fully convolutional neural networks for automatic vertebra segmentation and identification. *Med Image Anal*. 2019;53:142-155. doi:10.1016/j.media.2019.02.005
43. Wimmer M, Major D, Novikov AA, et al. Fully automatic cross-modality localization and labeling of vertebral bodies and intervertebral discs in 3D spinal images. *Int J Comput Assist Radiol Surg*. 2018;13(10):1591-1603. doi:10.1007/s11548-018-1818-3
44. Jamaludin A, Lootus M, Kadir T, et al. ISSLS PRIZE IN BIOENGINEERING SCIENCE 2017: Automation of reading of radiological features from magnetic resonance images (MRIs) of the lumbar spine without human intervention is comparable with an expert radiologist. *Eur Spine J*. 2017;26(5):1374-1383. doi:10.1007/s00586-017-4956-3
45. SpineWeb. Accessed April 18, 2020. http://spineweb.digitalimaginggroup.ca
46. IVDM3Seg. *MICCAI 2018 Challenge Automatic Intervertebral Disc Localization and Segmentation from 3D Multi-modality MR (M3) Images*. Published 2018. Accessed April 18, 2020. https://ivdm3seg.weebly.com/
47. Kim K, Kim S, Lee YH, et al. Performance of the deep convolutional neural network based magnetic resonance image scoring algorithm for differentiating between tuberculous and pyogenic spondylitis. *Sci Rep*. 2018;8(1):13124. doi:10.1038/s41598-018-31486-3
48. Pan Y, Chen Q, Chen T, et al. Evaluation of a computer-aided method for measuring the Cobb angle on chest X-rays. *Eur Spine J*. 2019;28(12):3035-3043. doi:10.1007/s00586-019-06115-w
49. Weng CH, Wang CL, Huang YJ, et al. Artificial intelligence for automatic measurement of sagittal vertical axis using ResUNet framework. *J Clin Med*. 2019;8(11):1826. doi:10.3390/jcm8111826
50. Hammon M, Dankerl P, Tsymbal A, et al. Automatic detection of lytic and blastic thoracolumbar spine metastases on computed tomography. *Eur Radiol*. 2013;23(7):1862-1870. doi:10.1007/s00330-013-2774-5
51. Yao J, Burns JE, Sanoria V, et al. Mixed spine metastasis detection through positron emission tomography/computed tomography synthesis and multiclassifier. *J Med Imaging (Bellingham)*. 2017;4(2):024504. doi:10.1117/1.JMI.4.2.024504
52. Ueno M, Aoki T, Murakami S, et al. CT temporal subtraction method for detection of sclerotic bone metastasis in the thoracolumbar spine. *Eur J Radiol*. 2018;107:54-59. doi:10.1016/j.ejrad.2018.07.017
53. Acar E, Leblebici A, Ellidokuz BE, Başbınar Y, Çapa KG. Machine learning for differentiating metastatic and completely responded sclerotic bone lesion in prostate cancer: a retrospective radiomics study. *Br J Radiol* 2019;92(1101):20190286. doi:10.1259/bjr.20190286
54. Lang N, Zhang Y, Zhang E, et al. Differentiation of spinal metastases originated from lung and other cancers using radiomics and deep learning based on DCE-MRI. *Magn Reson Imaging*. 2019;64:4-12. doi:10.1016/j.mri.2019.02.013
55. Lodwick GS, Haun CL, Smith WE, Keller RF, Robertson ED. Computer diagnosis of primary bone tumors: a preliminary report. *Radiology* 1963;80(2):273-275.
56. He Y, Guo J, Ding X, et al. Convolutional neural network to predict the local recurrence of giant cell tumor of bone after curettage based on pre-surgery magnetic resonance images. *Eur Radiol*. 2019;29(10):5441-5451. doi:10.1007/s00330-019-06082-2
57. Chen CY, Chiou HJ, Chou SY, et al. Computer-aided diagnosis of soft-tissue tumors using sonographic morphologic and texture features. *Acad Radiol*. 2009;16(12):1531-1538. doi:10.1016/j.acra.2009.07.024
58. Malinauskaite I, Hofmeister J, Burgermeister S, et al. Radiomics and machine learning differentiate soft-tissue lipoma and liposarcoma better than musculoskeletal radiologists. *Sarcoma*. 2020;2020:7163453. doi:10.1155/2020/7163453
59. Zhang Y, Zhu Y, Shi X, et al. Soft tissue sarcomas: preoperative predictive histopathological grading based on radiomics of MRI. *Acad Radiol*. 2019; 26(9):1262-1268. doi:10.1016/j.acra.2018.09.025
60. Blackledge MD, Winfield JM, Miah A, et al. Supervised machine-learning enables segmentation and evaluation of heterogeneous post-treatment changes in multi-parametric MRI of soft-tissue sarcoma. *Front Oncol*. 2019;9:941. doi:10.3389/fonc.2019.00941
61. Birtwhistle R, Morkem R, Peat G, et al. Prevalence and management of osteoarthritis in primary care: an epidemiologic cohort study from the canadian primary care sentinel surveillance network. *CMAJ Open*. 2015;3(3):E270-275. doi:10.9778/cmajo.20150018
62. Thomas KA, Kidziński Ł, Halilaj E, et al. Automated classification of radiographic knee osteoarthritis severity using deep neural networks. *Radiol Artif Intell*. 2020;2(2):e190065. doi:10.1148/ryai.2020190065
63. Pedoia V, Norman B, Mehany SN, et al. 3D convolutional neural networks for detection and severity staging of meniscus and PFJ cartilage morphological degenerative changes in osteoarthritis and anterior cruciate ligament subjects. *J Magn Reson Imaging*. 2019;49(2):400-410. doi:10.1002/jmri.26246
64. Liu F, Zhou Z, Samsonov A, et al. Deep learning approach for evaluating knee mr images: achieving high diagnostic performance for cartilage lesion detection. *Radiology*. 2018;289(1):160-169. doi:10.1148/radiol.2018172986
65. Liu F, Guan B, Zhou Z, et al. Fully automated diagnosis of anterior cruciate ligament tears on knee MR images by using deep learning. *Radiol Artif Intell*. 2019;1(3):180091. doi:10.1148/ryai.2019180091
66. Fritz B, Marbach G, Civardi F, Fucentese SF, Pfirrmann WA. Deep convolutional neural network-based detection of meniscus tears: comparison with radiologists and surgery as standard of reference. *Skeletal Radiol*. 2020;49(8):1207-1217. doi:10.1007/s00256-020-03410-2
67. Liu J, Pattanaik S, Yao J, et al. Computer aided detection of epidural masses on computed tomography scans. *Comput Med Imaging Graph*. 2014;38(7):606-612. doi:10.1016/j.compmedimag.2014.04.007
68. Pfeil A, Renz DM, Hansch A, et al. The usefulness of computer-aided joint space analysis in the assessment of rheumatoid arthritis. *Joint Bone Spine*. 2013;80(4):380-385. doi:10.1016/j.jbspin.2012.10.022

69. Langs G, Peloschek P, Bischof H, Kainberger F. Model-based erosion spotting and visualization in rheumatoid arthritis. *Acad Radiol*. 2007;14(10):1179-1188. doi:10.1016/j.acra.2007.06.013
70. Etli Y, Asirdizer M, Hekimoglu Y, Keskin S, Yavuz A. Sex estimation from sacrum and coccyx with discriminant analyses and neural networks in an equally distributed population by age and sex. *Forensic Sci Int*. 2019;303:109955. doi:10.1016/j.forsciint.2019.109955
71. Yune S, Lee H, Kim M, Tajmir SH, Gee MS, Do S. Beyond human perception: sexual dimorphism in hand and wrist radiographs is discernible by a deep learning model. *J Digit Imaging*. 2019;32(4):665-671. doi:10.1007/s10278-018-0148-x
72. Kapiński N, Zieliński J, Borucki BA, et al. Monitoring of the Achilles tendon healing process: can artificial intelligence be helpful? *Acta Bioeng Biomech*. 2019;21(1):103-111.
73. Balsiger F, Steindel C, Arn M, et al. Segmentation of peripheral nerves from magnetic resonance neurography: a fully-automatic, deep learning-based approach. *Front Neurol*. 2018;9:777. doi:10.3389/fneur.2018.00777
74. Kim S, Lee D, Park S, Oh KS, Chung SW, Kim Y. Automatic segmentation of supraspinatus from MRI by internal shape fitting and autocorrection. *Comput Methods Programs Biomed*. 2017;140:165-174. doi:10.1016/j.cmpb.2016.12.008
75. Włodarczyk J, Czaplicka K, Tabor Z, Wojciechowski W, Urbanik A. Segmentation of bones in magnetic resonance images of the wrist. *Int J Comput Assist Radiol Surg*. 2015;10(4):419-431. doi:10.1007/s11548-014-1105-x
76. Hannun A, Case C, Casper J, et al. Deep speech: Scaling up end-to-end speech recognition. *arXiv preprint arXiv:14125567*. 2014.
77. Lee H, Tajmir S, Lee J, et al. Fully automated deep learning system for bone age assessment. *J Digit Imaging*. 2017;30(4):427-441. doi:10.1007/s10278-017-9955-8
78. Do BH, Wu AS, Maley J, et al. Automatic retrieval of bone fracture knowledge using natural language processing. *J Digit Imaging*. 2013;26(4):709-713. doi:10.1007/s10278-012-9531 -1
79. Tan WK, Hassanpour S, Heagerty PJ, et al. Comparison of natural language processing rules-based and machine-learning systems to identify lumbar spine imaging findings related to low back pain. *Acad Radiol*. 2018;25(11):1422-1432. doi:10.1016/j.acra.2018.03.008
80. Carrodeguas E, Lacson R, Swanson W, Khorasani R. Use of machine learning to identify follow-up recommendations in radiology reports. *J Am Coll Radiol*. 2019;16(3):336-343. doi:10.1016/j.jacr.2018.10.020