



Published on *SpineUniverse* (<https://www.spineuniverse.com>)

Using Artificial Intelligence Algorithms to Detect Vertebral Fractures

Artificial intelligence can augment existing diagnostic capabilities for spine specialists. Learn how a Japanese team tested AI diagnosis against experienced spine surgeons, and hear from an expert on the tech's potential.

Written by [Kristen Lovell](#) [1]; Reviewed by [Jun S. Kim, MD](#) [2]

Peer Reviewed

- [Understanding AI](#) [3]
- [Glossary of AI Terms](#) [4]
- [CNNs and Diagnosis](#) [5]
- [Better Outcomes?](#) [6]
- [CNN Weaknesses](#) [7]
- [Case Study](#) [8]
- [The Future](#) [9]

Use of image processing technology employing artificial intelligence (AI) has increased in recent years across an array of medical imaging disciplines. Particularly, deep learning technologies are receiving much of the attention due to their rapid technological improvement and their potential to improve delivery methods of the healthcare industry.



[Click to enlarge](#)

[10]

How does artificial intelligence stack up against experienced clinicians in the diagnosis of compression fractures?

This article will use a recent study published in *The Spine Journal* as a case study for use of AI in diagnosing vertebral fractures. Jun S. Kim, MD of Mount Sinai Hospital is an expert in the field of AI and using computer science to aid in the use of spinal surgery and orthopedics. He provides an in-depth assessment and explanation on the benefits and drawbacks of using this technology.



A Framework for Understanding AI

Available open source AI options can be classified as low-level or high-level deep learning frameworks. Though this is not official industry-recognized terminology, this classification provides an easier and more intuitive understanding of the frameworks. *Low-level frameworks* provide for a basic block of abstraction, flexibility and room for customization. *High-level frameworks* are used to further aggregate the abstraction.

The purpose of this is to decrease human work. High-level frameworks limit the scope of customization and flexibility. High-level frameworks also use a low-level framework as a backend, often working by converting the data source into the chosen, customizable low-level framework for execution of the final model.¹

Popular open source frameworks for clinicians and researchers include the low-level frameworks:

- [TensorFlow](#) [11] by Google (currently most popular; used in the *Case Study* below)
- [MXNet](#) [12] by Apache
- [PyTorch](#) [13] Geometric (developed initially by Facebook; easier learning curve than TensorFlow)

...and high-level frameworks:

- [Keras](#) [14] (uses TensorFlow as a backend; used in the *Case Study* below)
- [Gluon](#) [15] (uses MXNet as a backend)

Terminology

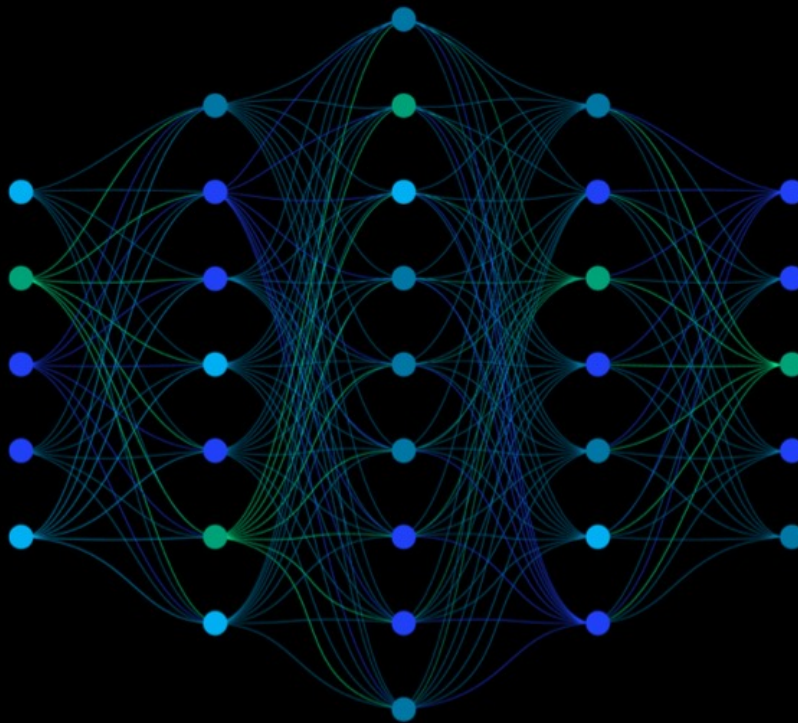
Artificial intelligence and machine learning comes with its own vocabulary. If you're new to AI and ML, here are some of the most pertinent terms for this discussion:

- *Deep learning* (DL) is an AI machine learning method dealing with algorithms based on artificial neural networks that mimic the biological structure and functioning of a human brain combined with representation learning.
 - In *shallow learning*, input data is readily understandable and patterns are learned from the model. With *deep learning*, the model must learn meaningful representations of the data it's been given, as well as learn parameters, all with the ability to generate classifications based on its learned representations.
 - The most common uses of deep learning are datasets involving images, text, or sound, because these contain data not readily interpretable by a computer – in effect, non-numeric.
- *Representation learning* is a set of techniques the machine uses allowing it to automatically find the representations required for a feature's detection or classification. Besides replacing manual feature engineering, representation learning allows machines to learn the features and apply them to perform a specific task.
- *Convolutional Neural Networks* (CNNs) are deep neural networks that use a special linear operation called convolution, which is an important and distinctive element of CNNs. A CNN has a multilayer structure of a neural network, and is a DL algorithm developed based on animals' visual functioning.
- *Convolution* is a mathematical operation fundamental to common image processing. Convolution works by multiplying two arrays of numbers (usually of different sizes, but of the same dimensionality) in order to produce a third array of numbers of the same dimensionality.
- A *deep neural network* (DNN) is an *artificial neural network* (ANN) with multiple layers "stacked" between the input and output layers. Regardless of the type, all neural networks consist of the same components:
 - Neurons
 - Synapses
 - Weights
 - Biases
 - Functions

"Medicine is complex and it doesn't always lend itself to simple patterns, and thus can't be modeled through shallow learning," per Dr. Kim.



Artificial Neural Network



Click to enlarge

[16]

A neural network representation

Use of CNN in Detecting Osteoporotic Vertebral Fractures and Other Spinal Conditions

AI carries great potential for early detection of osteoporotic vertebral fractures (OVF). Standard OVF detection relies on computed tomography (CT) images. One argument for using CNN for disease and fracture detection is the potential to maximize diagnostic capability and minimize the human factors of subjectivity and errors occurring because of distraction and fatigue. However, this benefit has not been proven thus far, with available research pointing to a fairly equal diagnostic ability between humans and CNNs.

According to Dr. Kim, CNNs “can be used to rapidly diagnose abnormal pathology; or, at the very least can change the priority at which they are examined by a physician. There are many different potential utilities of CNNs as it relates to spine.”

Potential uses of CNN in diagnosis of spinal conditions, scoliosis classification and identification of:

- Implants
- Tumors
- Infections
- Fractures
- Stenosis

Using CNNs in Spine Surgery

Dr. Kim uses a combination of his own programming and tool others have built. “When I was a resident in training, I took an online course in machine learning as a means of pursuing and applying deep learning to orthopedics. There aren’t many orthopedic spine surgeons with domain expertise in machine learning, and most spine companies are making implants rather than software. More recently this has changed with the application of increasing technology in the operating room and in the perioperative period.”

Using AI mainly for research, Dr. Kim posits, “More recently spine companies have pivoted and begun providing preoperative planning tools that use AI algorithms/neural networks to predict the correction a deformity or scoliosis patient will get after surgery. It can help you plan your fusion levels, your osteotomies, rod contour, etc.” Besides using CNN in diagnostic imaging, Dr. Kim also uses it for *natural language processing* (NLP), patient complication and prognostication (deep neural networks), and *generative adversarial networks* (GANs).

Can Using CNNs in Spine Imaging Translate into Better Patient Outcomes?



Dr. Kim believes these models will lead to better outcomes in the future. Currently though, “these models will require validation in the clinical setting before they become commonplace. I think where we can use CNNs is still being explored. At this point in time, I believe they may be safer to use as a tool to prioritize certain X-rays, CT, [and] MRI studies rather than classifying pathologies outright.” says Dr. Kim.

Since the beginning of his career, use of AI has changed Dr Kim’s approach to his practice of medicine. “It can change my surgical plan. There are preoperative planning tools that allow me to plan bony cuts and the contour of metal rods that I place during surgery to correct scoliosis or spinal deformity patients.”

Accounting for CNN Weaknesses

CNN can be excellent in extracting features, such as edges, corners, etc. However, CNNs can be “brittle” – a term used by Dr. Kim to describe the propensity of a CNN to be unable to bend itself like a human in terms of interpretation.

CNNs can only learn and adapt from the information they are given. “In CNNs, high-level details are done by high-level neurons. They check whether features are present or absent. They lose information about the composition or the position of the components. Humans have a much easier time with objects:

1. Under different angles
2. Under different backgrounds
3. Under several different lighting conditions...

...because we have an ability to find signs and other pieces of information to infer what we are seeing. Thus, CNNs are brittle because they have great performance when the images they are classifying are very similar to the dataset on which they are trained,” but not outside those data sets.

For example, if one hospital uses another hospital’s CNN dataset to try and detect fresh OVFs, the output will be skewed. The data set used is from another location using different imaging systems, different computers and different sets of patients with different demographics and conditions.

Other weaknesses include examples such as adversarial attacks, where adding noise to an image confuses the CNN. There are potential solutions to some of these issues with newer networks, like capsule networks and transformer visual networks.

Case Study: Promising Results from New Study Using CNN to Diagnose Fresh OVF

CNN is primarily used for detecting and classifying objects. A [2021 study](#) [17] used CNN to create a diagnostic support tool for magnetic resonance images (MRI). The authors then compared the performance of two spine surgeons (20 and 7 years of experience) to the results of the tool.

The method employed was a retrospective analysis of patient data from a clinical trial of patients that suffered fresh OVF. Patients were required to be over 65 and have a fresh OVF (defined as a fracture <3 months from OVF injury). The OVF could not be a result of disease-related fracture or a high-energy injury. The data was collected from January 2005 to September 2016.

The spine surgeons each independently evaluated 100 vertebrae (42 fresh; 58 old) randomly extracted from the test data for classification as either old and fresh OVFs.

Dataset

Several months after fracture onset, MRI was used with either 1.5 T or 3.0 T MR systems. There were institutional variations on MRI acquisition protocol and the timing of the MRIs. Midsagittal images output in jpeg format and used for analysis. Sixteen hundred twenty-four slices for T1 weighted images (T1WI) were obtained from 814 patients with OVF.

Algorithms Used

The training data was imbalanced (see *Limitations* below) and represented a relatively small dataset. To offset, data augmentation was used to improve the performance of the deep learning system. Also, a binary classification model of new and old fractures was created by using algorithm training data of 5,785 vertebrae. This binary classification was performed by using a combination of nine

CNNs: VGG16², VGG19², DenseNet121³, DenseNet169³, DenseNet201³, InceptionResNetV2⁴, InceptionV3⁵, ResNet50⁶, and Xception⁷.

All models were then normalized as one-hot expressions, which corresponded to old and fresh fractures using binary cross entropy as the loss function. These models were applied to the test data. Using 511 combinations, the AUC with the highest average value of the loss function of each model was selected as the optimal combination model.

To construct the CNN architecture, the authors used the Keras framework on top of Google’s open source deep learning framework TensorFlow, version 1.12.0. They performed image processing using the OpenCV library. Training and validation of CNN was performed using a computer with a GeForce GTX 1080 Ti (NVIDIA) graphics processing unit.

Study Limitations

- MR images were used, but due to the limitations caused by cost and equipment, MRI was not able to be used with all patients.
- Different MRI systems were used to gather data dependent upon the institution the data came from.
- 2D slices were used only from the 3D MRI modality and the authors then converted the DICOM files to JPEG before inputting them into the CNN. Some information may have been lost in these steps.
- Distinguishing OVF from other spinal pathologies can be difficult, especially metastatic spinal tumors, pyogenic spondylitis, and tuberculous spondylitis. Any system built to diagnose OVFs in the future would require a system to identify these differential diagnoses.
- The two spine surgeons used to compare accuracy with the CNN did not have access to clinical information, such as symptoms and days of injury. They only had access to the MR images. The authors believe if clinical information had been provided, their accuracy rate would also have been higher.

One workaround to some limitations included only using images of patients with OVF for deep learning, as well as eliminating all images from normal individuals or patients with pathological fractures. This solution caused the CNN system to only be able to diagnose an OVF – it could not diagnose anything

else.

Study Summary

Deep learning using T1WI was successful in the creation of an accurate system with an area under the curve (AUC) of 0.949 for the detection of fresh OVF. This custom AI system was comparable to two spine surgeons. The authors believe the system, “can contribute to the daily care of OVF by helping to properly diagnose the OVF.” An AUC of 1.0 is a perfect classifier, says Dr. Kim, while one of 0.0 is a classifier that gets everything wrong.

The Future of Machine Learning in Medicine

Dr. Kim believes machine learning is the key to taking better advantage of unstructured data such as those found in electronic medical records.

“With appropriate preprocessing, this data can be analyzed by machine learning algorithms,” says Dr. Kim. “So broadly speaking, not just CNNs, but machine learning in general, may allow us to perform research that better captures the population rather than from smaller clinical studies. This may lead to better patient outcomes from our ability to leverage all this data from EMRs.”

At present, the use of AI in medicine needs to have a checks and balances system. Data scientists must work closely with physicians to avoid datasets and other AI output misleading physicians. Currently, a human must still read the data. A radiologist needs to view diagnostic images. While developing or using an existing CNN should not be used for purely diagnostic purposes, it can be very useful for pushing more emergent patient cases to the front of the line in terms of their images being examined. This would be especially useful for fracture and infections.

Footnotes:

1. Moolayil JJ. DEEP LEARNING FOUNDATIONS: [A Layman’s Guide to Deep Neural Networks](#). [18] Towards Data Science; Jul 24, 2019.
2. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556 2014.
3. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017.
4. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, inception-resnet and the impact of residual connections on learning. Thirty-first AAAI conference on artificial intelligence; 2017.
5. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Angulov D, et al. Going deeper with convolutions. In: Szegedy, Christian, editor. Proceedings of the IEEE conference on computer vision and pattern recognition; 2015. p. 1–9.
6. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016.
7. Chollet F. Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017.

Ankile LL, Heggland MF, Krangle K: Deep Convolutional Neural Networks: [A survey of the foundations, selected improvements, and some current applications](#). [19] Computer Science > Computer Vision and Pattern Recognition. [Submitted on 25 Nov 2020] Cornell University.

Moolayil JJ. DEEP LEARNING FOUNDATIONS: [A Layman’s Guide to Deep Neural Networks](#). [18] Towards Data Science; Jul 24, 2019.

Murata, K., Endo, K., Aihara, T. *et al.* [Artificial intelligence for the detection of vertebral fractures on plain spinal radiography](#). [20] *Sci Rep* **10**, 20031 (2020). <https://doi.org/10.1038/s41598-020-76866-w> [21]

Schwartz JT, Cho BH, Tang P, Schefflein J, Arvind V, Kim JS, Doshi AH, Cho SK. [Deep Learning Automates Measurement of Spinopelvic Parameters on Lateral Lumbar Radiographs](#). [22] *Spine (Phila Pa 1976)*. 2020 Dec 1. doi: 10.1097/BRS.0000000000003830. Epub ahead of print. PMID: 33273436.

Yabu A, Hoshino M, Tabuchi M, et al. [Using artificial intelligence to diagnose fresh osteoporotic vertebral fractures on magnetic resonance images](#). [23] *The Spine Journal*. March 12, 2021. doi:<https://doi.org/10.1016/j.spinee.2021.03.006>

– [View Sources](#)

Updated on: 04/27/21

[Continue Reading](#)

[Heads Up! Docs Perform First Augmented Reality-Guided Spinal Fusion](#)

[24]

Updated on: 04/27/21

- [Remedy Health Media Sites](#)
- [About SpineUniverse](#)
- [Editorial Board](#)
- [Contact Us](#)
- [Advertising Opportunities](#)
- [Advertising Policy](#)
- [Partners and Sponsors](#)
- [Editorial Guidelines](#)
- [Case Submission Guidelines](#)
- [Privacy Policy](#)
- [Do Not Sell My Info](#)
- [Cookie Preferences](#)



- [Terms of Use](#)
- [Premium Practice Listing](#)

© 2021 Remedy Health Media, LLC ALL RIGHTS RESERVED

This information is not designed to replace a physician's independent judgment about the appropriateness or risks of a procedure for a given patient. Always consult your doctor about your medical conditions or back problem. SpineUniverse does not provide medical advice, diagnosis or treatment. Use of the SpineUniverse.com site is conditional upon your acceptance of our [User Agreement](#)



```
if (isOsanoBlockingScripts() && cloudfrontCountry !== 'US') { document.getElementById('cookiepreference').style.display = 'list-item' }
window.DD_LOGS && DD_LOGS.init({ clientToken: 'pub8e7a014aeb329eadda732052682ea942', site: 'datadoghq.com', service: 'SpineUniverse', env:
'prod', forwardErrorsToLogs: true, sampleRate: 100 }); div.osano-cm-window > div.osano-cm-info-dialog-wrapper > div > ul > li:nth-child(5){ display: none;
} .osano-cm-widget { display: none; } .osano-cm-info-dialog-close { background-color: transparent; background-image: none; }
```

Source URL: <https://www.spineuniverse.com/professional/news/AI-detect-vertebral-fractures>

Links:

- [1] <https://www.spineuniverse.com/author/51393/lovell>
- [2] <https://www.spineuniverse.com/author/51522/kim>
- [3] <https://www.spineuniverse.com/print/professional/news/AI-detect-vertebral-fractures#understanding>
- [4] <https://www.spineuniverse.com/print/professional/news/AI-detect-vertebral-fractures#terms>
- [5] <https://www.spineuniverse.com/print/professional/news/AI-detect-vertebral-fractures#diagnosis>
- [6] <https://www.spineuniverse.com/print/professional/news/AI-detect-vertebral-fractures#outcomes>
- [7] <https://www.spineuniverse.com/print/professional/news/AI-detect-vertebral-fractures#weaknesses>
- [8] <https://www.spineuniverse.com/print/professional/news/AI-detect-vertebral-fractures#case>
- [9] <https://www.spineuniverse.com/print/professional/news/AI-detect-vertebral-fractures#future>
- [10] https://www.spineuniverse.com/sites/default/files/imagecache/gallery-large/wysiwyg_imageupload/49571/2021/04/27/AI%20compression%20fracture%20diagnosis.jpg
- [11] <https://www.tensorflow.org/>
- [12] <https://mxnet.apache.org/versions/1.8.0/>
- [13] <https://pytorch.org/>
- [14] <https://keras.io/>
- [15] <https://mxnet.apache.org/versions/1.4.1/gluon/index.html>
- [16] https://www.spineuniverse.com/sites/default/files/imagecache/gallery-large/wysiwyg_imageupload/49571/2021/04/27/Ai%20compression%20fracture%20diagnosis%20neural%20network.jpg
- [17] [https://www.thespinejournalonline.com/article/S1529-9430\(21\)00117-0/fulltext](https://www.thespinejournalonline.com/article/S1529-9430(21)00117-0/fulltext)
- [18] <https://towardsdatascience.com/a-laymans-guide-to-deep-neural-networks-ddcea24847fb>
- [19] <https://arxiv.org/abs/2011.12960v1>
- [20] <https://www.nature.com/articles/s41598-020-76866-w>
- [21] <https://doi.org/10.1038/s41598-020-76866-w>
- [22] <https://pubmed.ncbi.nlm.nih.gov/33273436/>
- [23] [https://www.thespinejournalonline.com/article/S1529-9430\(21\)00117-0/fulltext#articleInformation](https://www.thespinejournalonline.com/article/S1529-9430(21)00117-0/fulltext#articleInformation)
- [24] <https://www.spineuniverse.com/professional/news/first-augmented-reality-guided-spinal-fusion>

© 1999-2021 Vertical Health, LLC

This information is not designed to replace a physician's independent judgment about the appropriateness or risks of a procedure for a given patient. Always consult your doctor about your medical conditions or back problem. SpineUniverse does not provide medical advice, diagnosis or treatment. Use of the SpineUniverse.com site is conditional upon your acceptance of our [User Agreement](#)

Vertical Health, LLC

