MASSACHUSETTS GENERAL HOSPITAL
IMAGING

RADIOLOGY ROUNDS
A newsletter for referring physicians

September 2017 - Volume 15, Issue 9 *Massachusetts General Hospital Department of Radiology*

Gary Boas, Author

Raul N. Uppot, M.D., Editor

Bone Age Assessment with Artificial Intelligence

- Bone age assessment is the evaluation of skeletal maturity. For decades, it was determined by examining the X-ray of a patient's hand and wrist and matching it using an atlas of 200 images.
- Artificial intelligence (AI) is manifest when machines develop an ability to not only recognize patterns but also apply those patterns in meaningful ways.
- Use of AI for semi-automated evaluation of bone age is the first deep learning medical image application to be deployed clinically at Mass General.

The artificial intelligence approach known as deep learning is often associated with uses such as autonomous navigation and face recognition, but it can do more than just tag your photos on social media sites. In recent years, researchers and clinicians have begun developing deep learning for a range of possible uses in radiology, where it can aid in the diagnosis of disease by discerning clinically relevant patterns in medical images. They are also preparing to introduce the approach in the clinical environment, with the goal of advancing an application long in need of an update: bone age assessment.

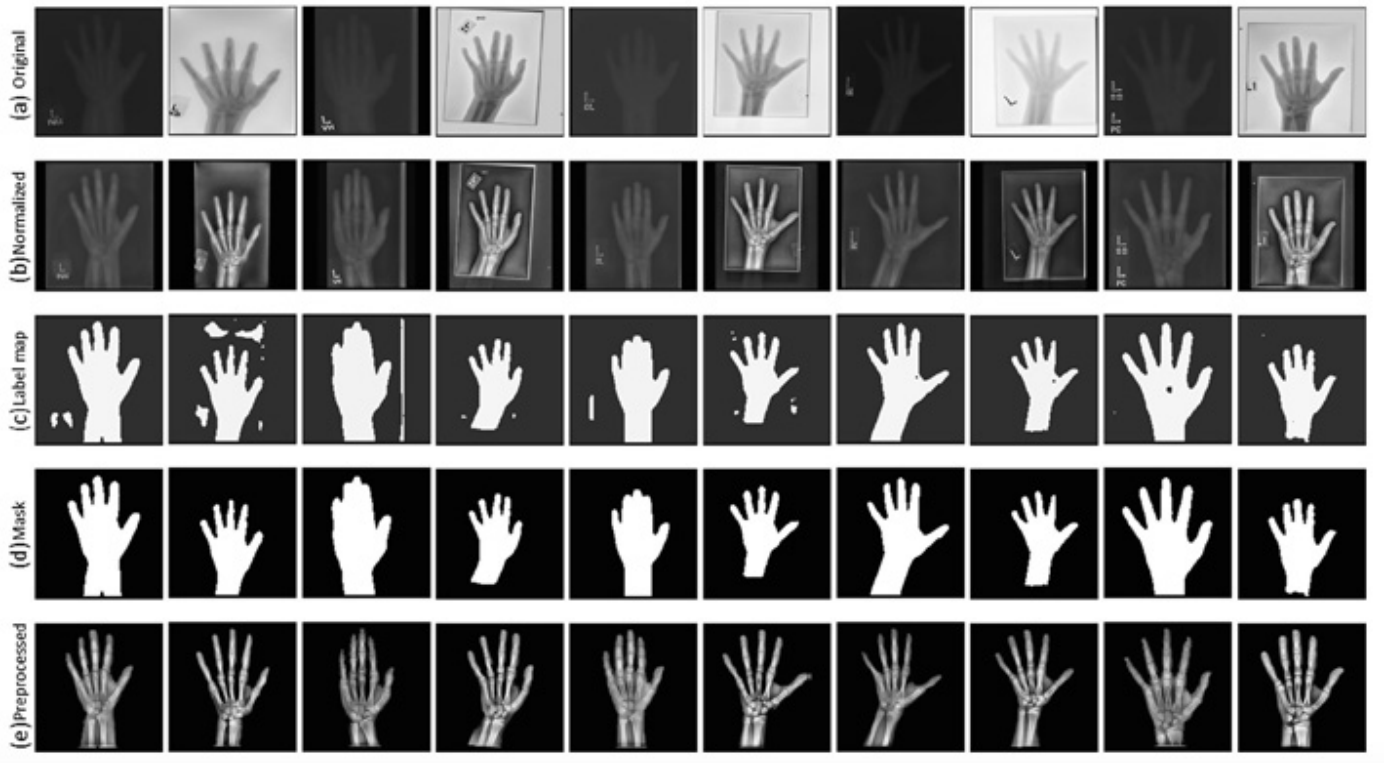

Figure 1. Researchers at Massachusetts General Hospital have reported a deep learning system for fully automated bone age assessment and are now preparing to implement the system in the clinical environment. Shown here are ten examples of hand X-rays at each stage of preprocessing for bone age assessment. Image courtesy of Journal of Digital Imaging (Lee, H., Tajmir, S., Lee, J. et al. J Digit Imaging (2017) 30: 427). (Click here to enlarge image)

Intelligence exhibited by something other than the human or animal brain, artificial intelligence (AI) is manifest when machines develop an ability to not only recognize patterns but also apply those patterns in meaningful ways. Initially, machine learning algorithms required structured input, *i.e.,* labeled data with clearly defined parameters to

facilitate the learning process. Over time, the algorithms have become more sophisticated and robust, eliminating the need for structured input. This newer approach to machine learning, using neural networks designed to mimic the decision-making processes in the human brain, is referred to as deep learning.

Deep learning already permeates much of our daily lives. In addition to autonomous navigation and face recognition, it underlies speech recognition software, online movie and purchasing recommendations, and augmented reality selfie filters for video chat platforms. Medicine is a more recent focus. Over the past several years, researchers have been exploring the potential of deep learning for applications including breast cancer detection and diagnosis with mammography, brain tumor segmentation with MRI, and diagnosis of Alzheimer's and other neurological disorders with functional MRI.

Deep Learning Advances Bone Age Assessment

Use of AI for semi-automated evaluation of bone age is the first deep learning medical image application to be deployed clinically at Mass General. The evaluation of skeletal maturity, bone age assessment has a range of applications, for example to predict adult height or assess cases of genetic under- or overgrowth syndromes. Its use dates to 1950, when anatomy professor William Walter Greulich and research associate S. Idell Pyle published their atlas of skeletal development of the hand and wrist. Since then, the procedure has essentially not changed. Today, to determine bone age, the radiologist will examine an X-ray of the patient's hand and wrist and compare it to 200 sample images to find the best match. This approach can take a considerable amount of time—as much as four minutes per image—and involves significant inter-radiologist variability, which can complicate decisions about therapy.

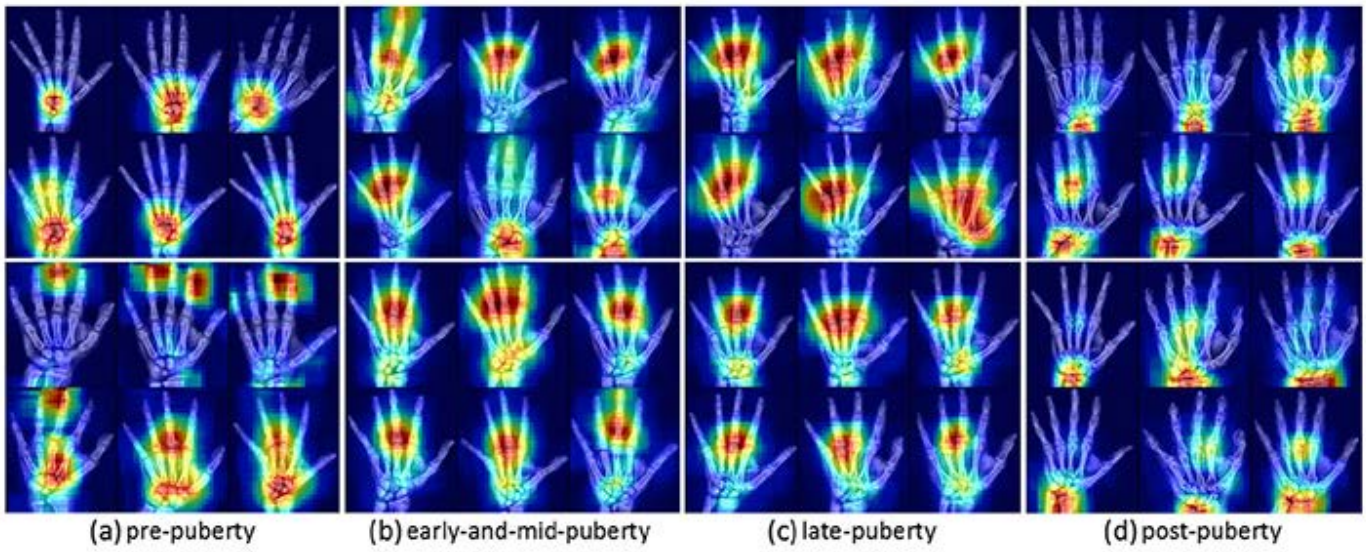

Figure 2. This image includes examples of attention maps for females (upper rows) and males (lower rows) during the four major skeletal maturity stages: pre-puberty, early and mid-puberty, late puberty, and post-puberty. The attention maps highlight regions of the hand and wrist that are key to understanding skeletal maturity at each of these stages. Image courtesy of Journal of Digital Imaging. (Click here to enlarge image)

Researchers have spent years trying to develop a fully automated approach to bone age assessment, with some success. For example, a software package has been approved for use in Europe. But typical solutions rely on hand-crafted features, *i.e.,* the structured input that drives conventional machine learning algorithms.

Clinical Implementation of Deep Learning at Mass General

In a *Journal of Digital Imaging* paper published online in March 2017, researchers at Massachusetts General Hospital described a deep learning system for bone age assessment that addresses this limitation and provides a fully automated approach for clinical implementation. Here, AI infers the bone age from an X-ray image with no other patient information, trained by a convolutional neural network (CNN) using a data set including age and associated ideal X-rays. A CNN is a class of deep learning networks that mimics the neural connectivity patterns found in the animal visual cortex. It can provide confidence level and predicted age for any X-ray within seconds.

The researchers tested the new deep learning system by applying it to more than 10,000 radiographs obtained at Mass General between 2005 and 2015. The retrospective study confirmed that it offers improved classification accuracy and dramatically reduced read times compared to the manually performed bone age assessment. It demonstrated an accuracy of greater than 90 percent when assigning a bone age within one year and greater than 98 percent when assigning a bone age within two years—all with read times of less than two seconds.

With an eye toward implementation in the clinic, the researchers developed an application that automatically scans the reference images and, for each patient, outputs the three or four that are most likely to match. The radiologist then selects from among these. Further streamlining the process, the application incorporates a reporting system the researchers had previously developed that allows the radiologist to generate a standardized report with the click of a button instead of dictating findings.

The researchers are now preparing the system for clinical use at Mass General; it will be the first deep learning medical image application to be deployed clinically at the hospital. They anticipate introducing deep learning for bone age assessment within the next few months, after additional development of the system to better account for different races and to further improve inter-radiologist variability.

Further Information

We thank Synho Do, PhD, and Tarik Alkasab, MD, PhD, Department of Radiology, Massachusetts General Hospital, for their advice and assistance in preparing this article. Drs. Do and Alkasab and the Center for Clinical Data Science at Massachusetts General Hospital and Brigham & Women's Hospital are currently exploring other possible applications of deep learning in radiology. They welcome any feedback about areas that could benefit.

Gary Boas, Author
Raul N. Uppot, M.D., Editor