Making AI Models Interpretable and Explainable for Medical Image Analysis

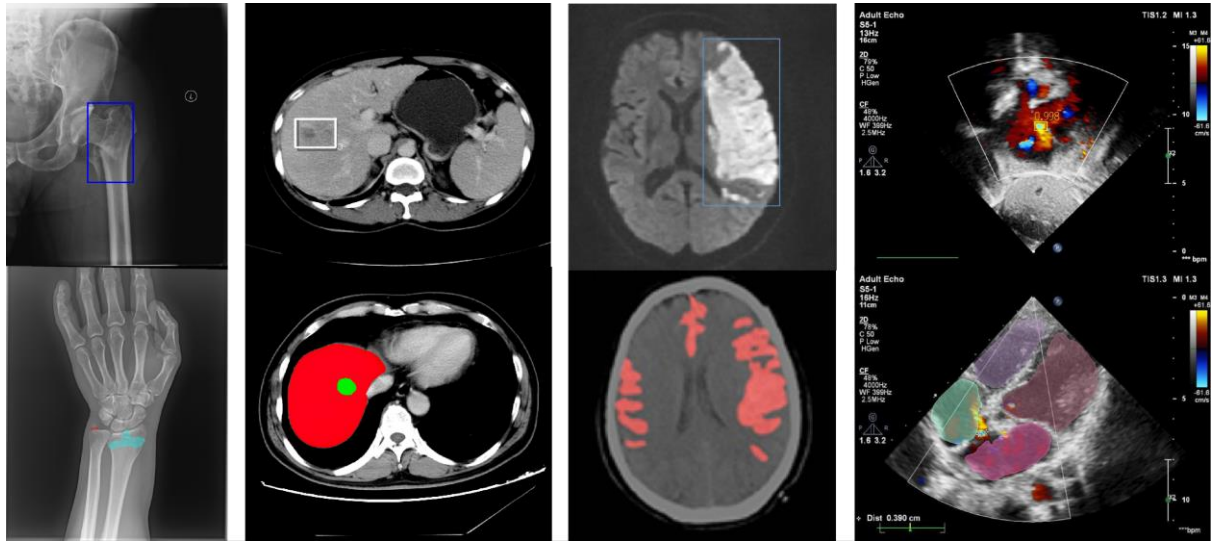
Ahmad P. Tafti, PhD
Assistant Professor of Computer Science
Director, HexAI Research Laboratory
University of Southern Maine

Outline

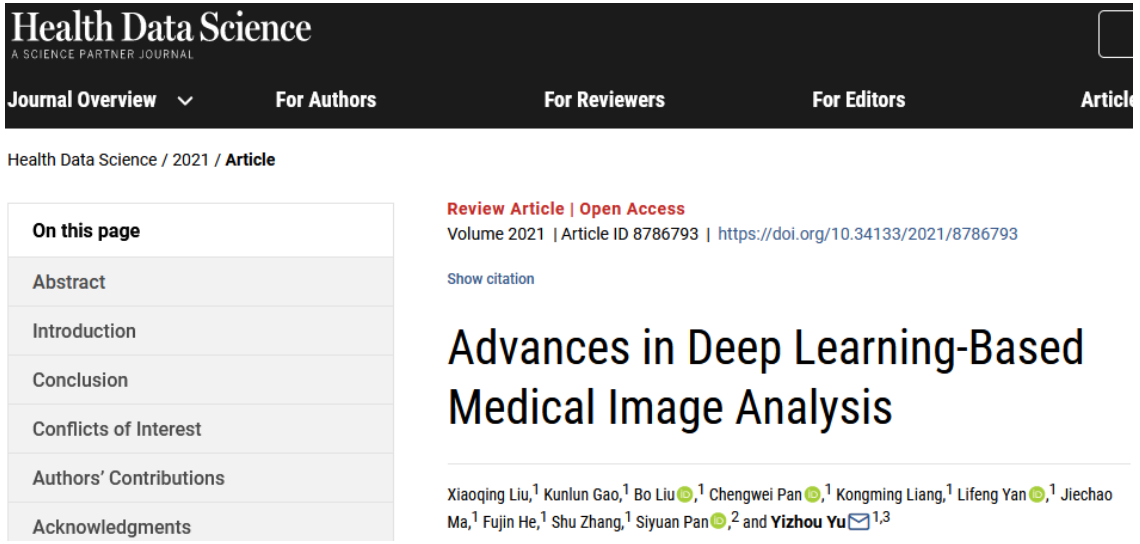
- Recent Advances in AI-Powered Medical Image Analysis
- Two Main Message in AI-Powered Medical Image Analysis
- Explainability in Computer Vision
- Making AI Models Interpretable and Explainable for Medical Image Analysis

AI-Powered Medical Image Analysis; Recent Advances

- Computer Vision, and particularly Deep Learning Computer Vision has already demonstrated successful applications in a variety of medical image analysis problems, including image registration, image segmentation, anomaly detection, object localization, and classification.



Citation: Liu, X., Gao, K., Liu, B., Pan, C., Liang, K., Yan, L., Ma, J., He, F., Zhang, S., Pan, S. and Yu, Y., 2021. Advances in deep learning-based medical image analysis. Health Data Science, 2021.



Overall, according to the best available evidence, deep learning models performed well in medical image analysis.

AI-Powered Medical Image Analysis; Recent Advances

MEDICAL PHYSICS
The International Journal of Medical Physics Research and Practice

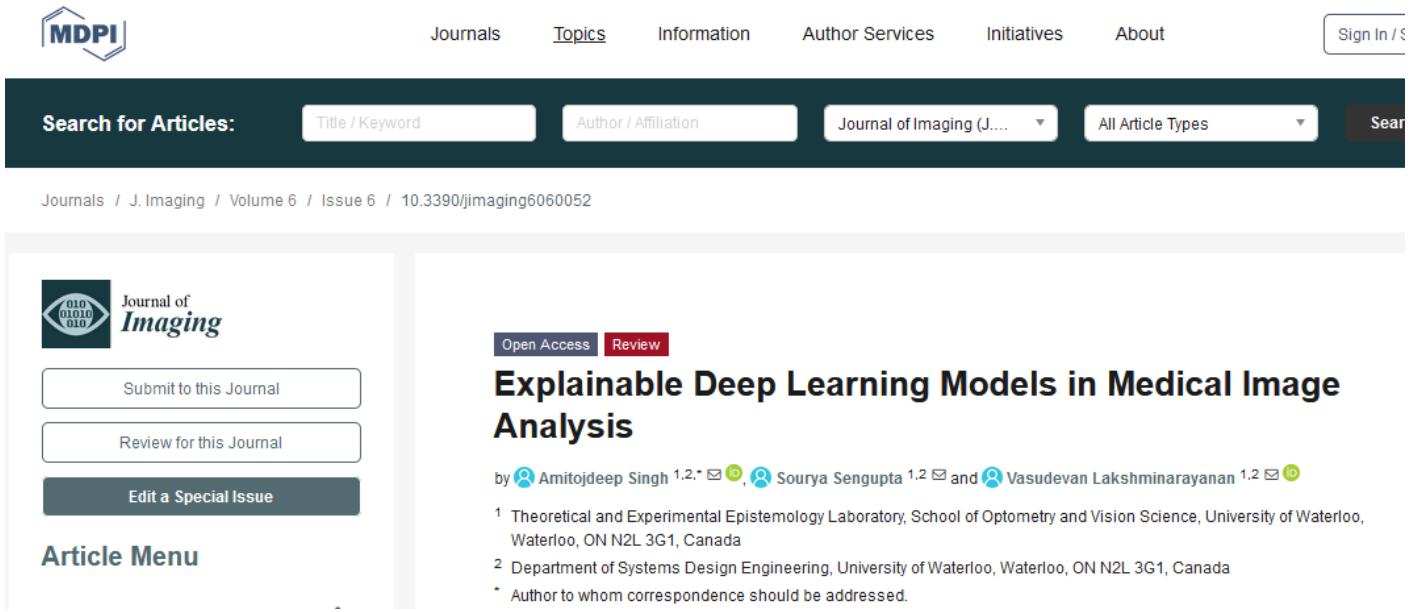
Research Article

Comparing the performance of a deep convolutional neural network with orthopedic surgeons on the identification of total hip prosthesis design from plain radiographs

Alireza Borjali, Antonia F. Chen, Hany S. Bedair, Christopher M. Melnic, Orhun K. Muratoglu, Mohammad A. Morid, Kartik M. Varadarajan

The CNN achieved the same or higher performance than at least one of the surgeons in identifying eight of nine THR implant designs and underperformed all of the surgeons in identifying one THR implant design (Anthology)

AI-Powered Medical Image Analysis; Recent Advances



Deep learning methods have been very effective for a variety of medical diagnostic tasks and have even outperformed human experts on some of those. However, the black-box nature of the algorithms has restricted their clinical use.

AI-Powered Medical Image Analysis; Two Main Messages

Message #1

It is **AND**. It is not **OR**.

Deep Learning Computer Vision **AND** Surgeons
Deep Learning Computer Vision **AND** Radiologists
Deep Learning Computer Vision **AND** Physicians

Message #2

While scientific progress in deep learning computer vision and pattern recognition has led to advanced modeling strategies with almost human-like performance, they are limited in their explainability and interpretability.

(1) What is the rationale behind AI-Powered decision making? **(Explainability)**
(2) What is the meaning of this AI-Powered decision making? **(Interpretability)**

Rule-based vs. Learning-based Computer Vision

Step

Door

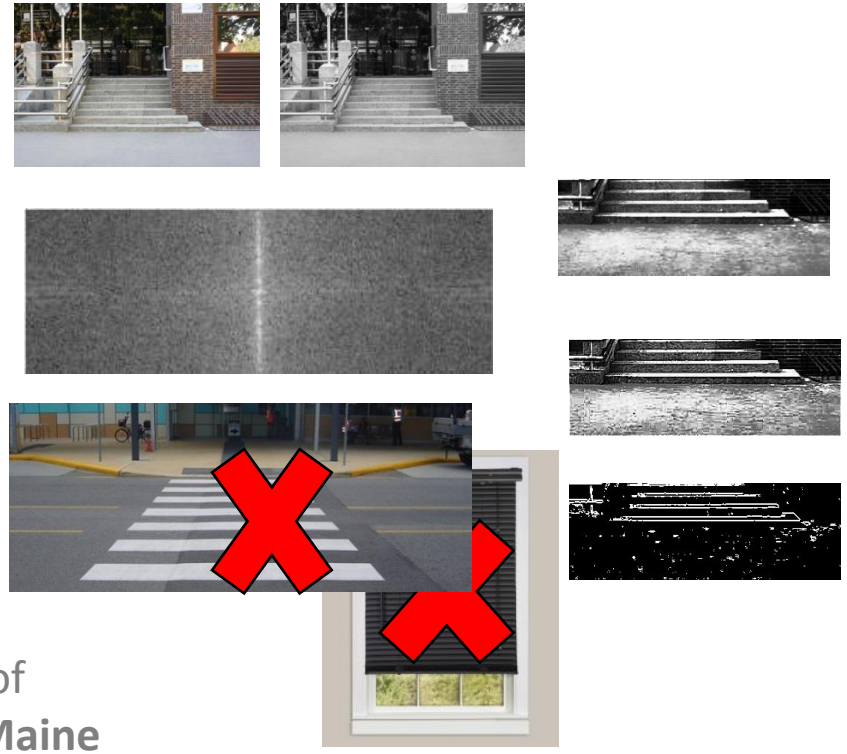
How to implement **singularities**???

Rule-based algorithms

Learning-based algorithms

Deep and machine learning strategies vs. **Traditional** (rule-based) methods

Traditional (there is no any learning technique)

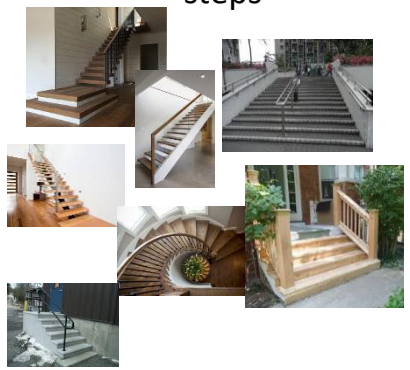


Deep and machine learning techniques

We train computers at recognizing doors from steps by showing them a **large amount** of:

(object_type, picture) pairs → (door, image1.jpg)
(step, image10.jpg)

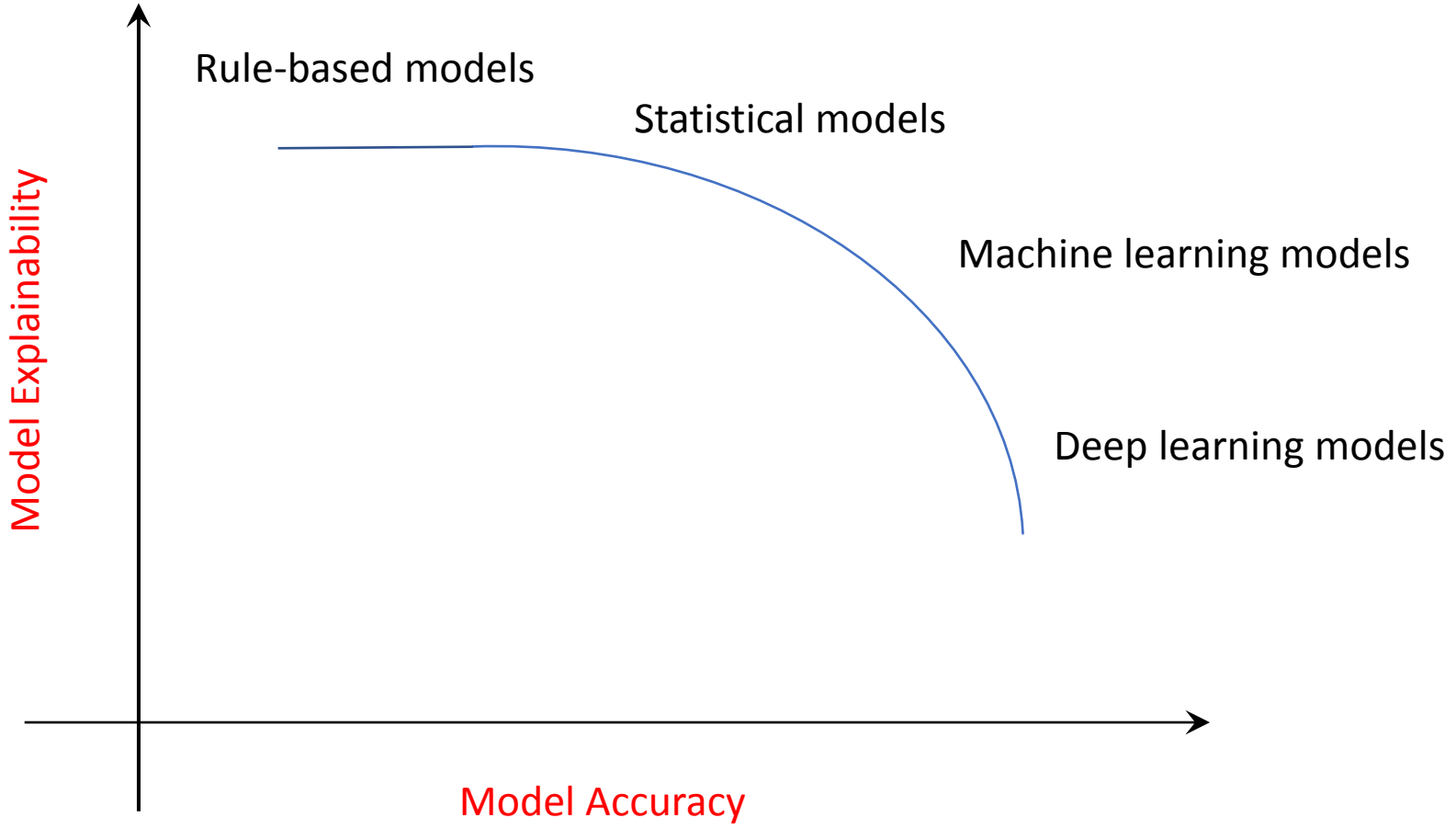
steps



doors

University of Southern Maine

Computer Vision Models; Explainability



Computer Vision Models; Explainability

- **As part of computer vision model explainability, we need:**
  - **Model interpretability** to ensure that the AI-Powered decision making is natural, and we can't see biasness in the prediction.
  - **Model transparency** to make sure that we are differentiating the false causality from true causality.
  - Enable real-world decision makers (domain experts) to **trust** the AI-Powered models

Computer Vision Models; Explainability Objectives

- **The AI-Powered models explainability aims to achieve:**
  - **Trust:** Convey confidence in the model to our end users.
  - **Correlations:** Capture correct correlations and associations among various features.
  - **Identity:** Preserve privacy and identity management
  - **Fairness:** Be fair and ethically compliant.
  - **Reliability:** Confidence in the AI-Powered models

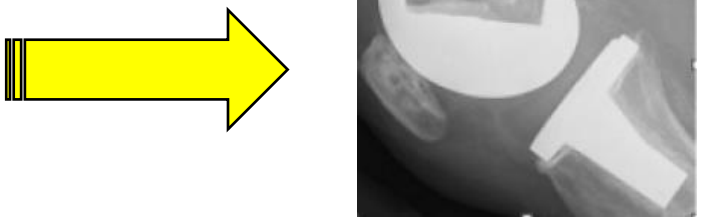
Explainability in Computer Vision Models; How

Example-based

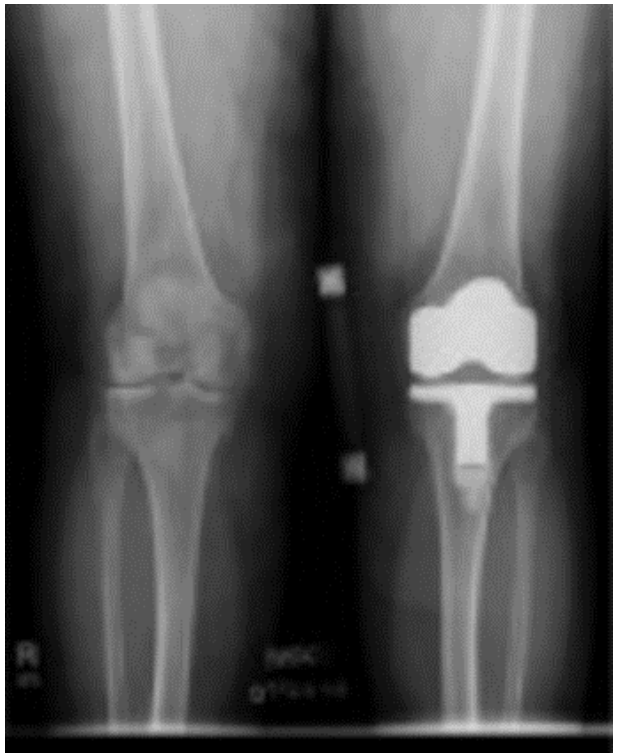
Textual-based

Visual-based

What X-ray machines generate

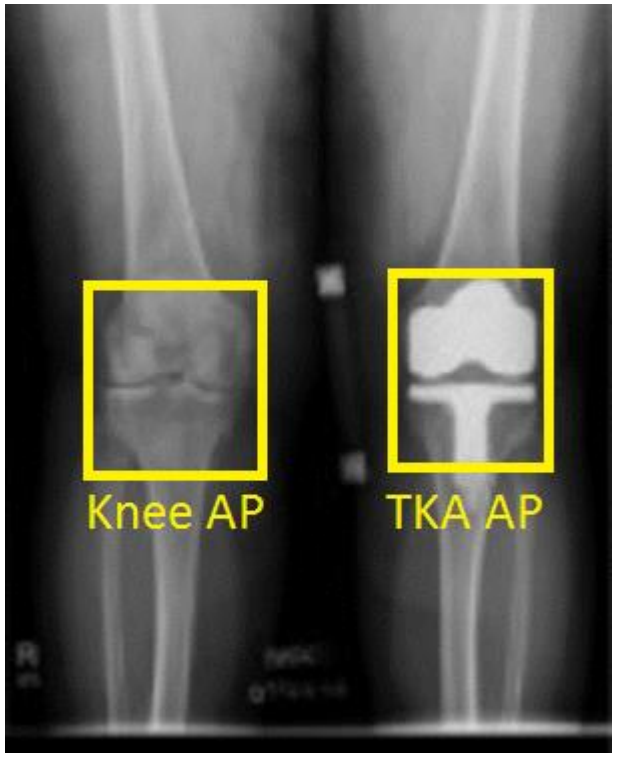


What our practitioners want



What X-ray machines generate

What our practitioners want



Example-based

Citation: Yan, S., Ramazanian, T., Sagheb, E., Kremers, W.K., Chaudhary, V., Taunton, M., Kremers, H.M. and **Tafti, AP**. *Give me a knee radiograph, I will tell you where the knee joint area is: a deep convolutional neural network adventure*. arXiv preprint arXiv:2202.05382.

Clinical Significance:
Pre/Post Operative Knee Surgery

(a) TKA AP view
left:1.00 right:1.00

(b) knee AP view 0.99
TKA AP view 1.00

(c) TKA AP view 1.00
knee AP view 0.99

(g) knee lateral view 0.97

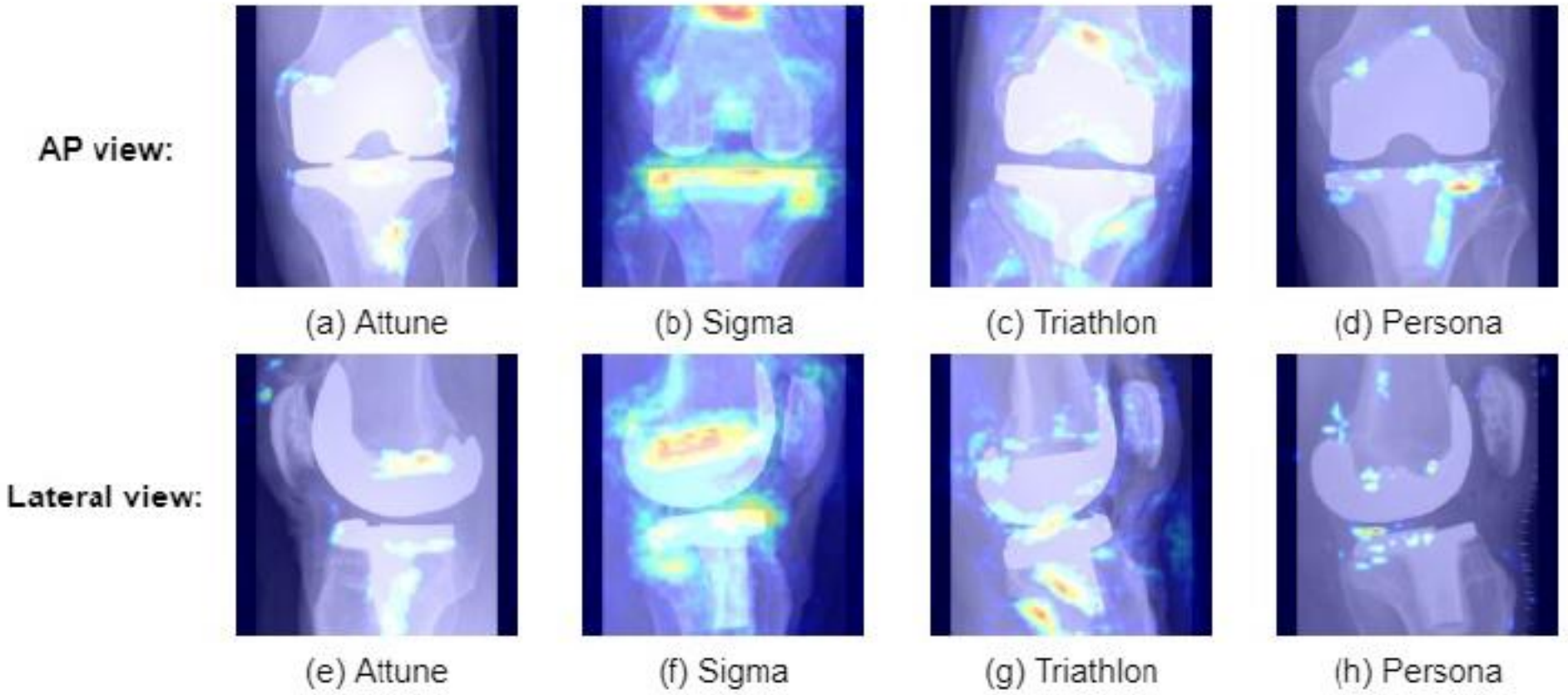
(h) TKA lateral view 0.98

(i) knee lateral view 1.00

University of Southern Maine

Visual-based

Citation: Yan S, Ramazanian T, Sagheb E, Fu S, Sohn S, Lewallen DG, Liu H, Kremers WK, Chaudhary V, Taunton M, Kremers HM, **Tafti AP**. *DeepTKAClassifier: Brand Classification of Total Knee Arthroplasty Implants Using Explainable Deep Convolutional Neural Networks*. In International Symposium on Visual Computing 2020 Oct 5 (pp. 154-165). Springer, Cham.



Clinical Significance:
Knee revision surgery

Thank you!