Pre-Screening for Spine Red Flags in the Emergency Room

A Radiologist's Assistant

Josh Cyphers, Heaven Klair, Fengshou Liang, Joe Schwab, Brian Tung

Fewer than 14% of EDs achieve 90% Threshold of seeing patients triaged for immediate Care (< 1 Hr)

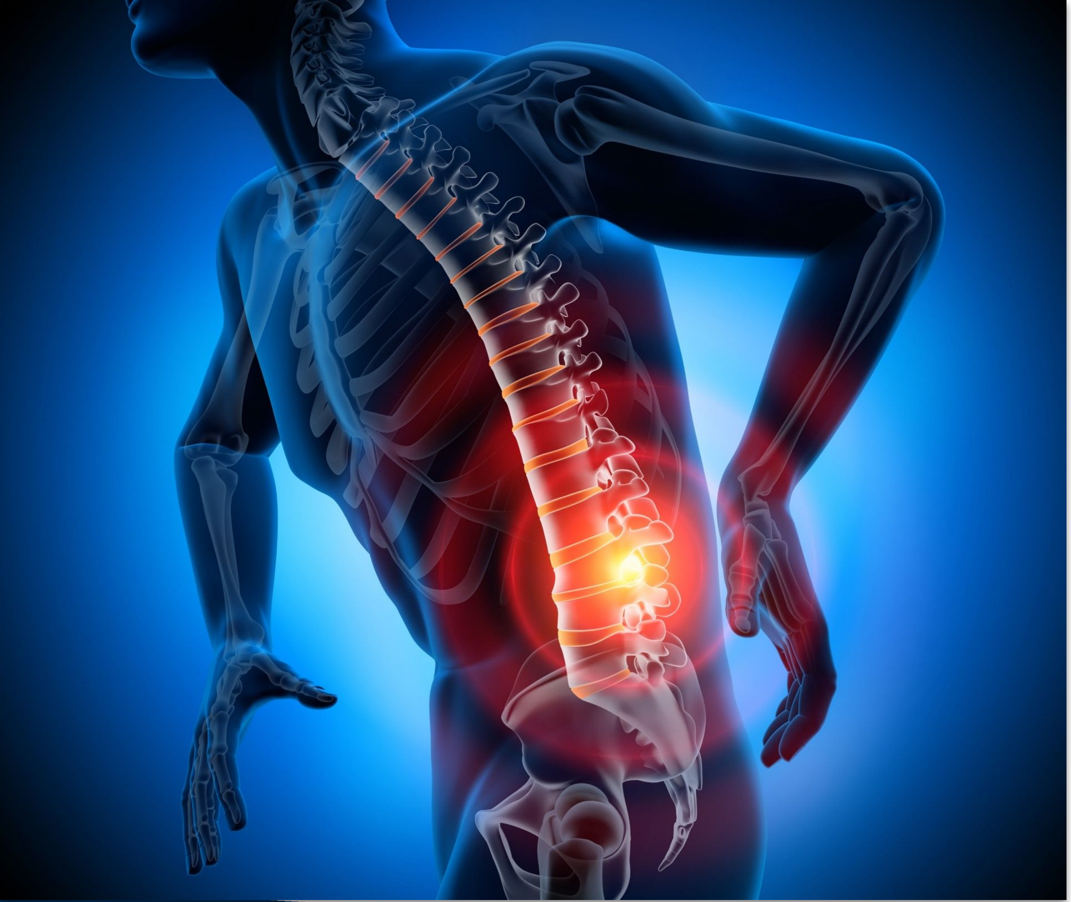
Horwitz LI et al, US emergency department performance on wait time and length of visit. Ann Emerg Med. 2010 Feb;55(2):133-41

Berkeley
UNIVERSITY OF CALIFORNIA

Overcrowding in the ER leads to Adverse Outcomes and Increased Overall Mortality

Sprivulis PC et al The association between hospital overcrowding and mortality among patients admitted via Western Australian emergency departments. *Med J Aust.* 2006 Mar 6;184(5):208–212

Berkeley
UNIVERSITY OF CALIFORNIA



Back Pain

2.6 million Emergency Room visits per year (3% of all ER visits)

Second most common reason for missing work (after common cold)

Pitts SR et al, National Hospital Ambulatory Medical Care Survey: 2006 emergency department summary.*Natl Health Stat Report.* 2008:1–38


Yet, Back Pain is Almost Always Self Limiting and Does not Require ER or Hospital Care.
Unless…..

Spine Red Flags

Infection

Fracture

Tumor

Paralysis, Death, Disability Unless Treated

Spine Red Flags must be confirmed via an MRI

However, Radiologists often have little effective means to prioritize their workload in the Emergency Department

All cases are emergencies by definition

Goals

Assist the Overworked Radiologist

Pre-Review the Spinal MRIs ordered in the ER

Screen for "Red Flags" & Place "Red Flag" Cases at the top of the Queue

Decrease Wait Times in EDs efficiency

Problem Statement

This project focuses on the critical "Red Flag" symptoms in lumbar spine MRI images: Tumors, Infections, and Fractures. With emergency departments facing overcrowding issues, our goal is to develop and implement a Convolutional Neural Network (CNN) capable of detecting these "Red Flag" indicators.

By incorporating this model into clinical workflows, we aim to enable radiologists to receive immediate alerts for cases requiring urgent attention. We seeks to streamline this diagnosis process, ensuring timely intervention for patients with serious health conditions.

Data Collection

- **Dataset Overview:** Comprises 1372 hand-picked MRI images from 88 different patients categorized into 4 labels: Tumor, Fracture, Infection, and No Red Flag.
    - Total Images: 1372
    - Patients Represented: 88 unique patients
    - Categorization: Tumor, Fracture, Infection, and No Red Flag

- **Data Source**
    - [https://radiopaedia.org/?lang=us]: "**Radiopaedia.org** is a rapidly growing, peer-reviewed open-edit radiology resource, compiled by radiologists and other health professionals from across the globe"

- **Data Quality**
    - Expert Validation: Radiologists/Medical professionals reviewed the images to confirm the accuracy of the categorization and the clarity of the images.

```
Total cases: 88
Training files: 924, Cases: 61
Validation files: 335, Cases: 13
Testing files: 100, Cases: 14
```

Exploratory Data Analysis

- **Varying Image Dimensions:** Dataset contains MRI images of different sizes, impacting uniformity and model training efficiency.
- **Standardization Approach:**
  - **Dimension Standardization:** To achieve uniformity, applied black padding to resize all images to 255x255 pixels without distorting aspect ratios.
  - **Grayscale Normalization:** Standardized image intensity by converting all images to grayscale and normalizing pixel values for improved model consistency.
- **Label Distribution:**
  - [Insert the counts for each label]

Data and ML Pipelines

Modeling Phase 1:
RadImageNet Dataset
25,570 Images

Orientation
Classification Task
(Binary)

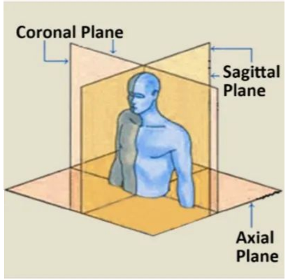
Modeling Phase 2:
Red Flags
1372 Flags Images

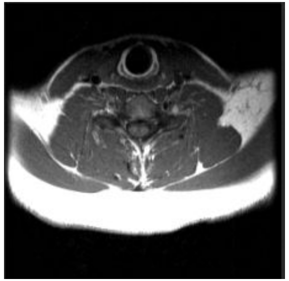
Red Flags
Classification Task
(Binary and Multiclass)

Spine Image Binary Classification

```
Distribution of Axial vs Sagittal Images:
Axial: 572
Sagittal: 585
```

```
26/26 [==============================] - 133s 5s/step - loss: 0.1363 - accuracy: 0.9543 - val_loss: 0.0035 - val_accuracy: 1.0000
Epoch 2/10
26/26 [==============================] - 129s 5s/step - loss: 0.0523 - accuracy: 0.9928 - val_loss: 0.0997 - val_accuracy: 0.9844
Epoch 3/10
26/26 [==============================] - 126s 5s/step - loss: 0.0580 - accuracy: 0.9904 - val_loss: 0.1611 - val_accuracy: 0.9844
Epoch 4/10
26/26 [==============================] - 130s 5s/step - loss: 0.0322 - accuracy: 0.9940 - val_loss: 0.0071 - val_accuracy: 1.0000
Epoch 5/10
26/26 [==============================] - 129s 5s/step - loss: 0.0182 - accuracy: 0.9952 - val_loss: 0.1184 - val_accuracy: 0.9844
Epoch 6/10
26/26 [==============================] - 129s 5s/step - loss: 0.0058 - accuracy: 0.9976 - val_loss: 0.1188 - val_accuracy: 0.9844
Epoch 7/10
26/26 [==============================] - 129s 5s/step - loss: 0.0041 - accuracy: 0.9988 - val_loss: 0.1279 - val_accuracy: 0.9844
Epoch 8/10
26/26 [==============================] - 129s 5s/step - loss: 0.0037 - accuracy: 0.9988 - val_loss: 0.2075 - val_accuracy: 0.9844
Epoch 9/10
26/26 [==============================] - 127s 5s/step - loss: 9.0466e-04 - accuracy: 1.0000 - val_loss: 0.1128 - val_accuracy: 0.9844
Epoch 10/10
26/26 [==============================] - 126s 5s/step - loss: 7.5549e-05 - accuracy: 1.0000 - val_loss: 0.1607 - val_accuracy: 0.9844
7/7 [==============================] - 10s 1s/step - loss: 0.0375 - accuracy: 0.9911
Test accuracy: 0.9910714030265808
```

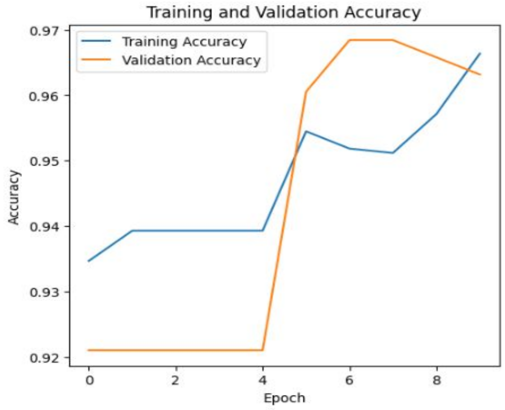


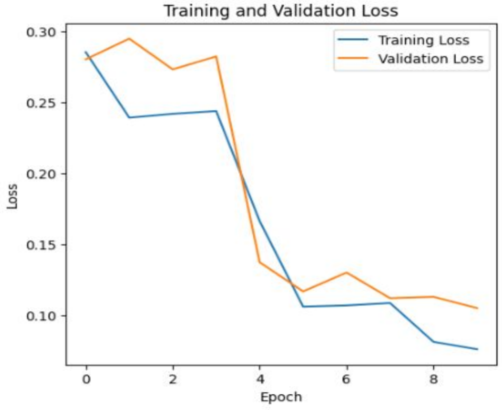


Pipeline Results

```
Epoch 1/10
48/48 [==============================] - 8s 93ms/step - loss: 0.2855 - accuracy: 0.9347 - val_loss: 0.2804 - val_accuracy: 0.9211
Epoch 2/10
48/48 [==============================] - 3s 58ms/step - loss: 0.2393 - accuracy: 0.9393 - val_loss: 0.2950 - val_accuracy: 0.9211
Epoch 3/10
48/48 [==============================] - 3s 58ms/step - loss: 0.2420 - accuracy: 0.9393 - val_loss: 0.2733 - val_accuracy: 0.9211
Epoch 4/10
48/48 [==============================] - 3s 58ms/step - loss: 0.2440 - accuracy: 0.9393 - val_loss: 0.2825 - val_accuracy: 0.9211
Epoch 5/10
48/48 [==============================] - 3s 57ms/step - loss: 0.1665 - accuracy: 0.9393 - val_loss: 0.1375 - val_accuracy: 0.9211
Epoch 6/10
48/48 [==============================] - 3s 57ms/step - loss: 0.1062 - accuracy: 0.9545 - val_loss: 0.1169 - val_accuracy: 0.9605
Epoch 7/10
48/48 [==============================] - 3s 57ms/step - loss: 0.1071 - accuracy: 0.9518 - val_loss: 0.1302 - val_accuracy: 0.9684
Epoch 8/10
48/48 [==============================] - 3s 57ms/step - loss: 0.1088 - accuracy: 0.9512 - val_loss: 0.1122 - val_accuracy: 0.9684
Epoch 9/10
48/48 [==============================] - 3s 58ms/step - loss: 0.0814 - accuracy: 0.9571 - val_loss: 0.1131 - val_accuracy: 0.9658
Epoch 10/10
48/48 [==============================] - 3s 57ms/step - loss: 0.0762 - accuracy: 0.9664 - val_loss: 0.1052 - val_accuracy: 0.9632
15/15 [==============================] - 1s 46ms/step - loss: 0.0749 - accuracy: 0.9768
Test accuracy: 97.68%
```

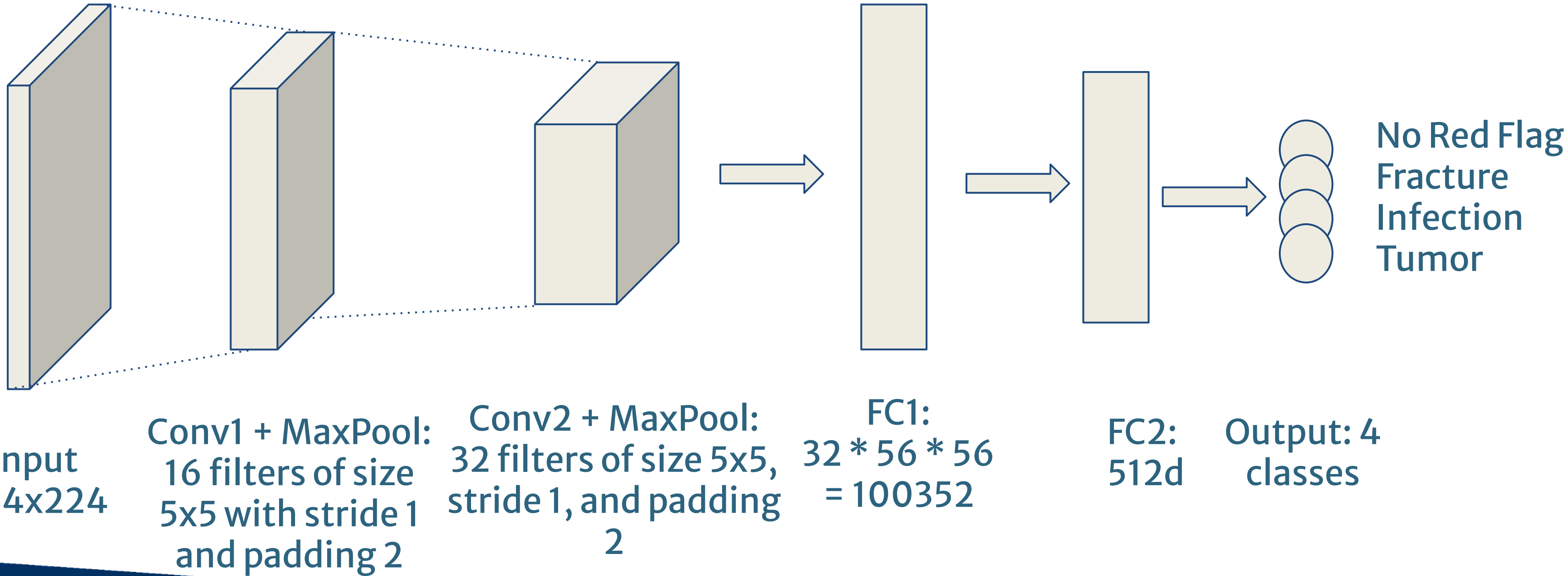
```
There are 721 files in the RIN_MRI_SPINE_CORD_Pathology folder.
There are 8320 files in the RIN_MRI_SPINE_Disc_Pathology folder.
There are 1580 files in the RIN_MRI_SPINE_Foraminal_Pathology folder.
There are 10847 files in the RIN_MRI_SPINE_NORMAL/normal folder.
There are 2291 files in the RIN_MRI_SPINE_OSSEOUS_ABN folder.
There are 1991 files in the RIN_MRI_SPINE_SCOLIOSIS folder.
```



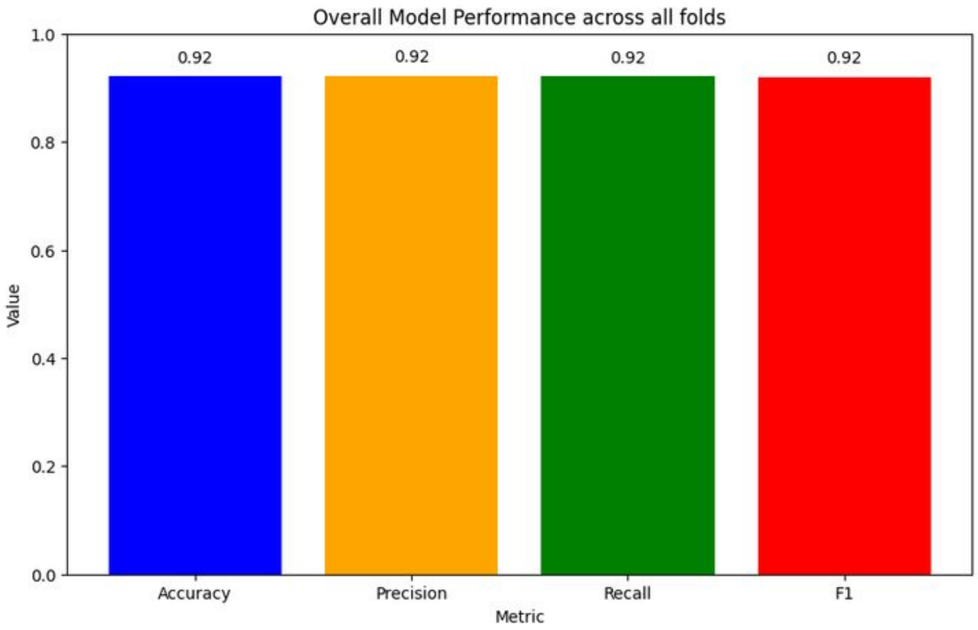


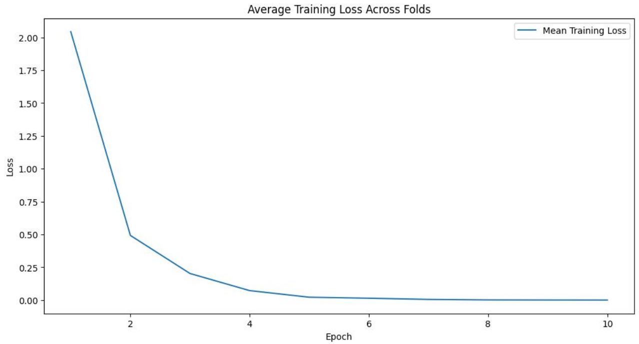
Berkeley
UNIVERSITY OF CALIFORNIA

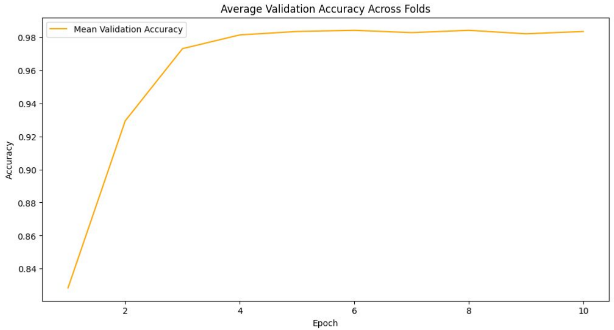
Multiclass Classification Setup



Multiclass Classification Model Results

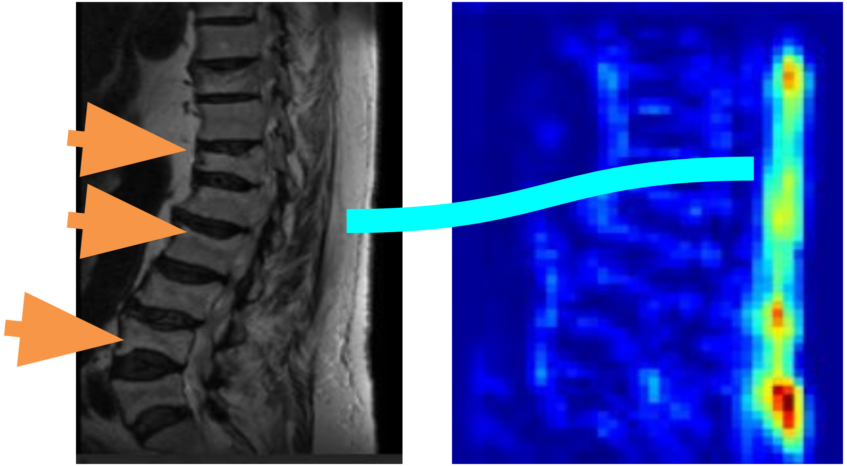


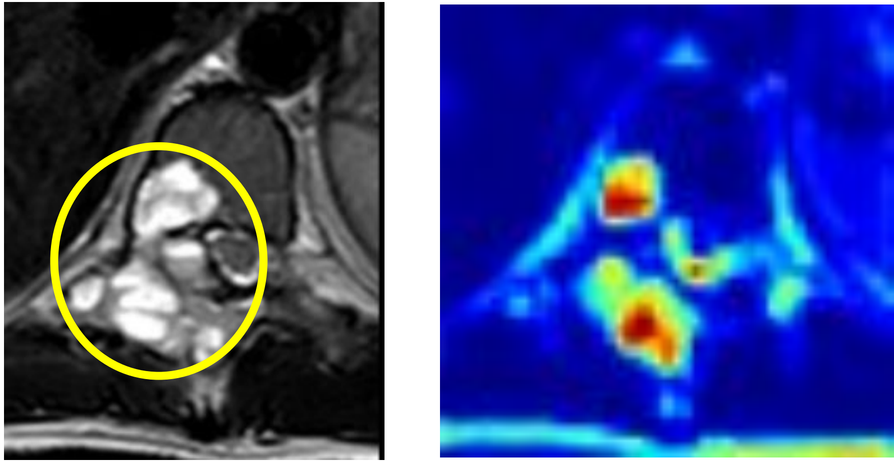




Berkeley
UNIVERSITY OF CALIFORNIA

CAM Activation and ROI Labeling





Exploring Alternatives: Model Experimentation

Model Type	Accuracy	Precision	Recall	F1-Score	Remarks
SimpleCNN	0.89	0.92	0.92	0.92	Best Model
DenseNet	0.75	0.76	0.75	0.74	Overfitting
SimpleCNN + ROI	0.78	0.88	0.78	0.81	
Dual PathWay SimpleCNN + ROI	0.84	0.87	0.84	0.84	

Berkeley
UNIVERSITY OF CALIFORNIA

All Models evaluated via K-Fold cross validation

Minimum Viable Product (MVP)

- Desktop/Mobile application: a classifier that identifies MRI images which present a red flag vs those that do not
- Values:
    - Less wait time for the patient obtaining the MRI
    - Less wait time for the patients waiting to enter the ED but cannot due to the bed being occupied by the patient waiting for the MRI
    - Improved overall throughput for the ED which translates to more money for the institution which can be reinvested into the ED providing better care

Client Application

- Client application developed as a fast and efficient way to utilize the algorithm for inference.
- Two usage modes:
  - Personal: A user can upload an image and observe an inferred pathology (or lack thereof) from their uploaded image
  - Professional: Radiologists can view MRIs in their queue prioritized by pathology type
- Designed for simplicity
- Mobile application in the near future

Application Demo

Ethics and Privacy

- As a “successor” to RadImageNet, we follow the same ethics guidance as that original work. For that study, the institutional review boards waived the requirement for written informed consent for this retrospective, Health Insurance Portability and Accountability Act–compliant study.
- All images de-identified before classification and processing.
- No link between the patients, data provider, and data receiver(s) was made available
- Privacy concerns regarding the client application will need to be analyzed in a different light – radiologists may need to access patient data.

Conclusion:

- Our project uses ML to detect 'red flag' indicators in lumbar spine MRIs.
- Achieving over 90% accuracy, it could speed up emergency care.
- Next steps involve refining, integrating, and ensuring ethical use.
- Ultimately, our goal is to improve care by alerting radiologists promptly.

Berkeley
UNIVERSITY OF CALIFORNIA