# Predicting Bone Age from Hand Radiographs Using Deep Convolutional Neural Networks

**Caroline Kimmel, Amin Ojjeh, and Samir Safwan**
Stanford University

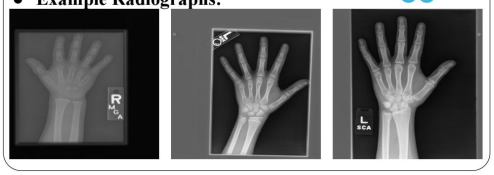
## Project Objective

- Predict skeletal bone age from hand radiographs.
- Use deep CNNs to capture important features used in making skeletal bone age prediction.
- Use visualization techniques to understand what our CNN is learning.

## RSNA Challenge Dataset

- **Training and Test Sets:**
  - The training set was provided on Kaggle and included 12,612 unique arbitrary-sized left and right hand radiographs and a CSV file of identifier, bone age, and gender.
  - The test set included 200 similarly labeled radiographs from RSNA on which we evaluated our models.
- **Processing:**
  - Randomized 80:20 split of RSNA Challenge data for training and validation set, respectively.
  - Resized images to 299 by 299 and converted to RGB (3 channels).
- **Example Radiographs:**

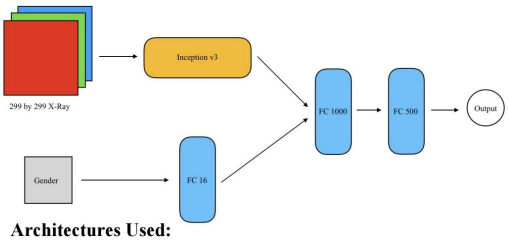
kaggle



## Methodology

- **Loss Function:** Mean Absolute Deviance (MAD)
- No pretrained weights were used in our models.
- Incorporated random horizontal and vertical flips.
- Trained using AWS GPUs and Stanford ICME GPUs.
- Focused on training two architectures.

## CNN Architectures




- **Architectures Used:**
  - Deep CNN with Inception v3 module for image input.
  - Gender input goes through fully connected ReLU layer, concatenates with Inception v3 flattened output and goes through two fully connected layers with ReLU activation and a final linear output layer.
  - Also tried replacing Inception v3 with ResNet 101.
- **Hyperparameters After Tuning:**

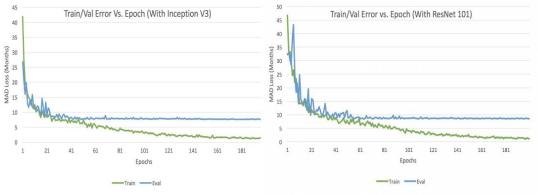
| Hyperparameters | Value |
|---|---|
| Learning Rate | 0.001 |
| LR Decay | On Plateau |
| Epochs | 200 |
| Batch-size | 16 |
| Optimizer | Adam |

## Performance

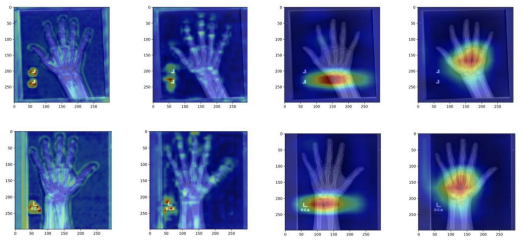
- We selected the model with Inception v3 as our final model, which had MAD error of 7.279 months on the test set, while model with ResNet 101 had MAD error of 8.766 on test set.
- Below are examples of radiographs in our test set:
  - Left: Male, 73.26 months, Right: Female, 192.59
  - Predicted 60.6 months and 189.9 months, respectively.

## Model Analysis

- **Train/Val Error vs. Training Duration (Epochs):**




- **Activation Maps of Various Layers of Our Final Model:**



## Conclusions

- Our model learns detailed hand features such as epiphyses in the metacarpals earlier in the model and emphasizes larger areas like the wrist or multiple fingers later on.
- The performance of our model is very competitive with the winners of the RSNA challenge (MAD 5.99 months), considering the smaller size of our input radiographs.

## Future Work

- Train longer and on larger X-rays.
- Try more data augmentation techniques.
- Balance train and validation sets more.
- Ensemble various models.