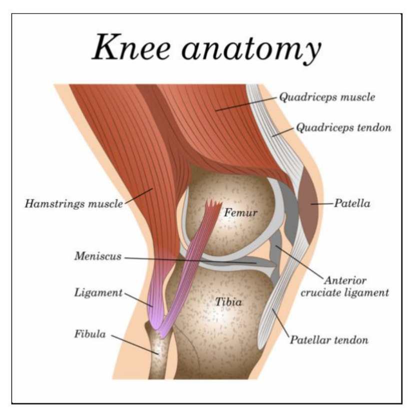
Measuring Patellar Instability using Deep Learning Predicted Caton-Deschamps Index

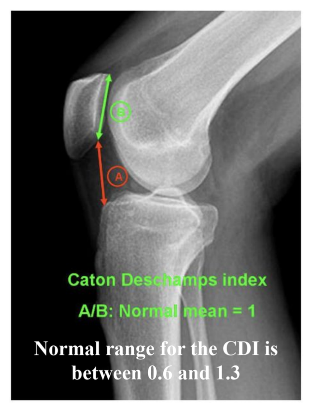
Micael Tchapmi, Jacob Azoulay, Daniel Lee

Stanford University

Introduction

- Patellar instability is a condition where the patella dislodges partially (subluxation) or completely (dislocation) from the groove at the end of the femur resulting in an unstable kneecap.
- Caton-Deschamps Index (CDI) is measured from lateral x-rays and can be used to determine severity of the knee injury and inform treatment decisions.
- Determining patellar height through CDI from x-rays manually is laborious and time consuming.
- Deep learning, relatively nascent in medical imaging, can automate the task.



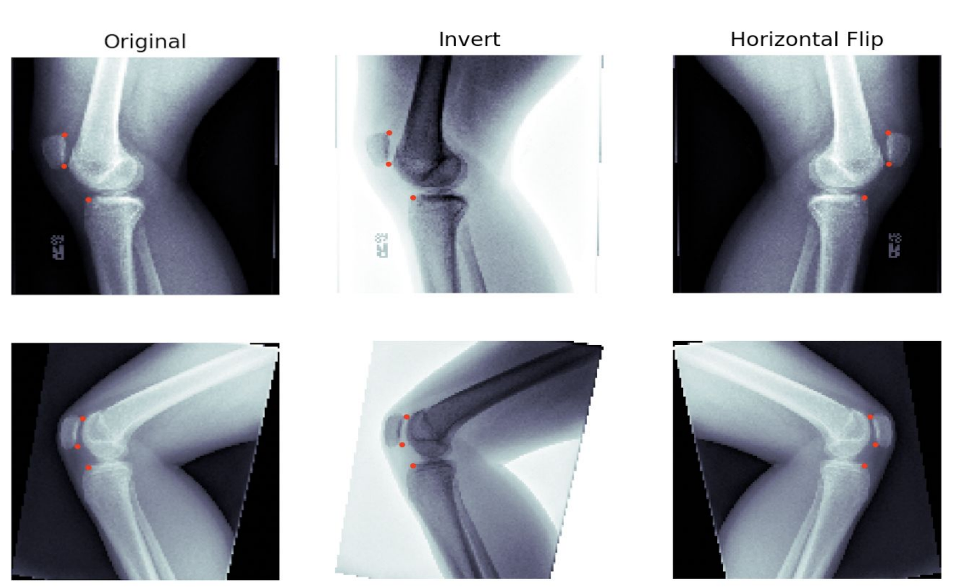


Problem Statement

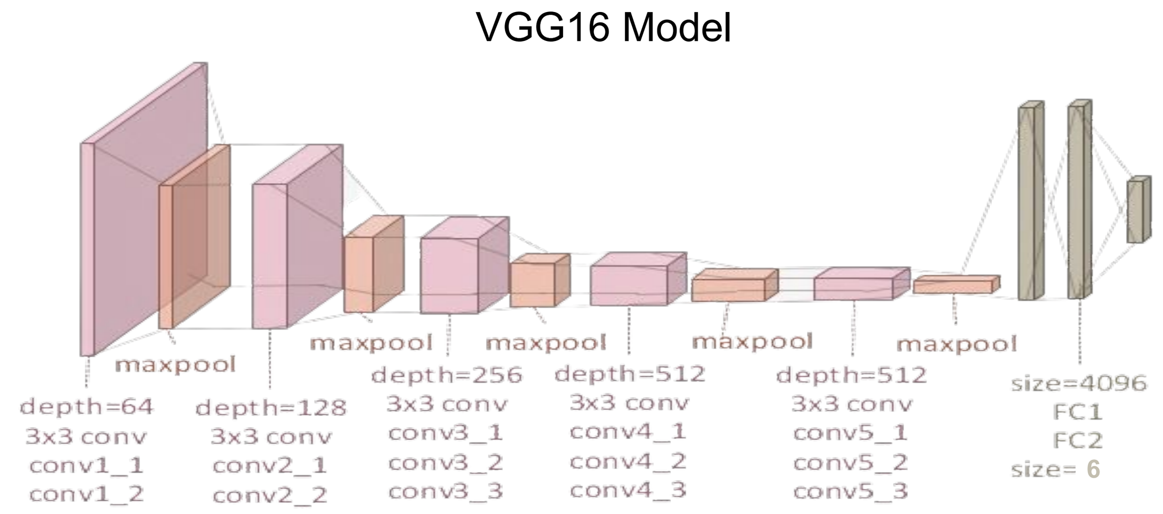
- Generate a convolutional neural network (CNN) model that takes a lateral x-ray as an input and outputs the three key points that can be used to compute CDI.
- Results are measured using CDI error, key point distance, and intra-class correlation coefficient.

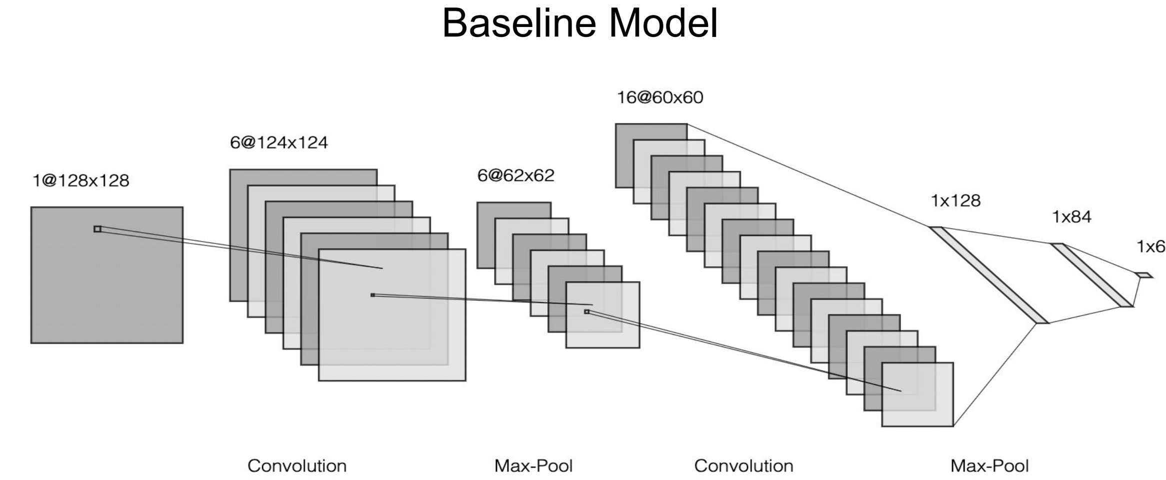
Data Processing

- Data consists of 304 x-ray images of the knees of patients aged 10-25 who have sustained patellar dislocation or subluxation.
- Key points are labeled marking the superior patella, inferior patella, and the tibial plateau.
- Augmentation methods such as inversion, horizontal flip, rotation, brightness shift, and rescaling have been implemented.



Model Architecture

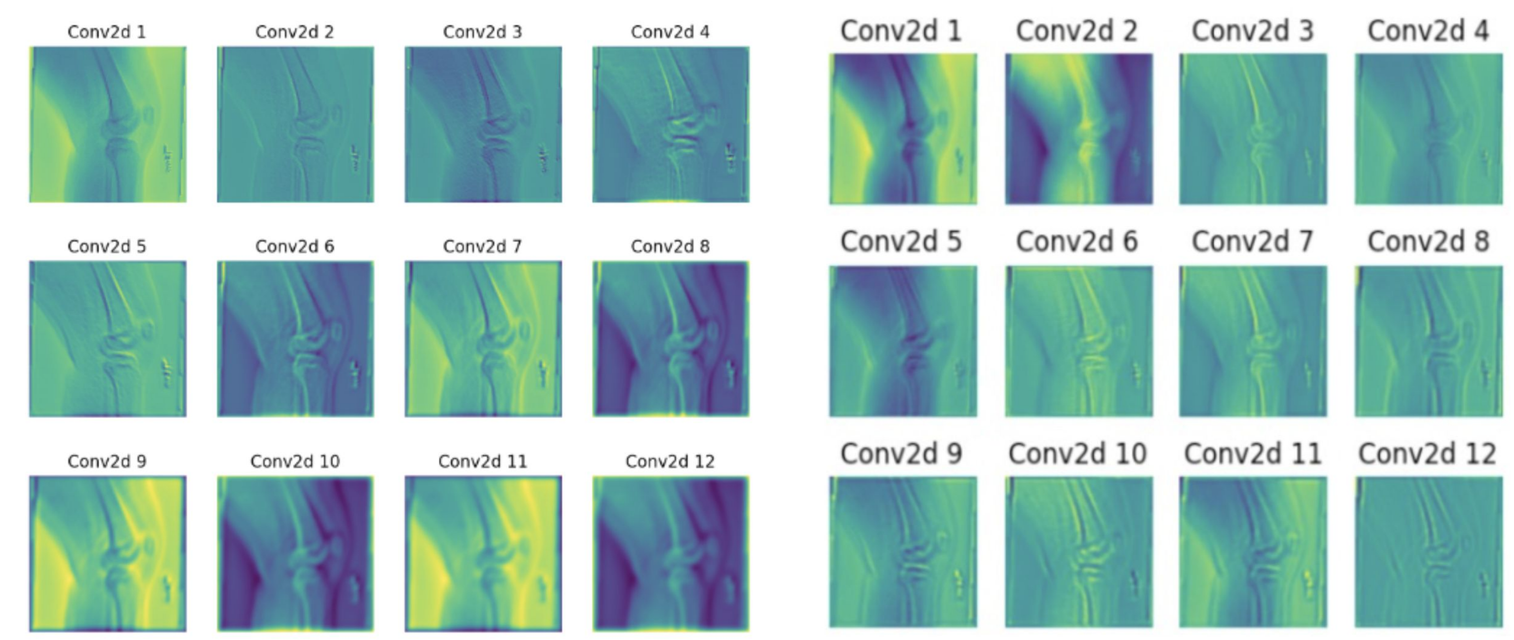




Experiments and Analysis

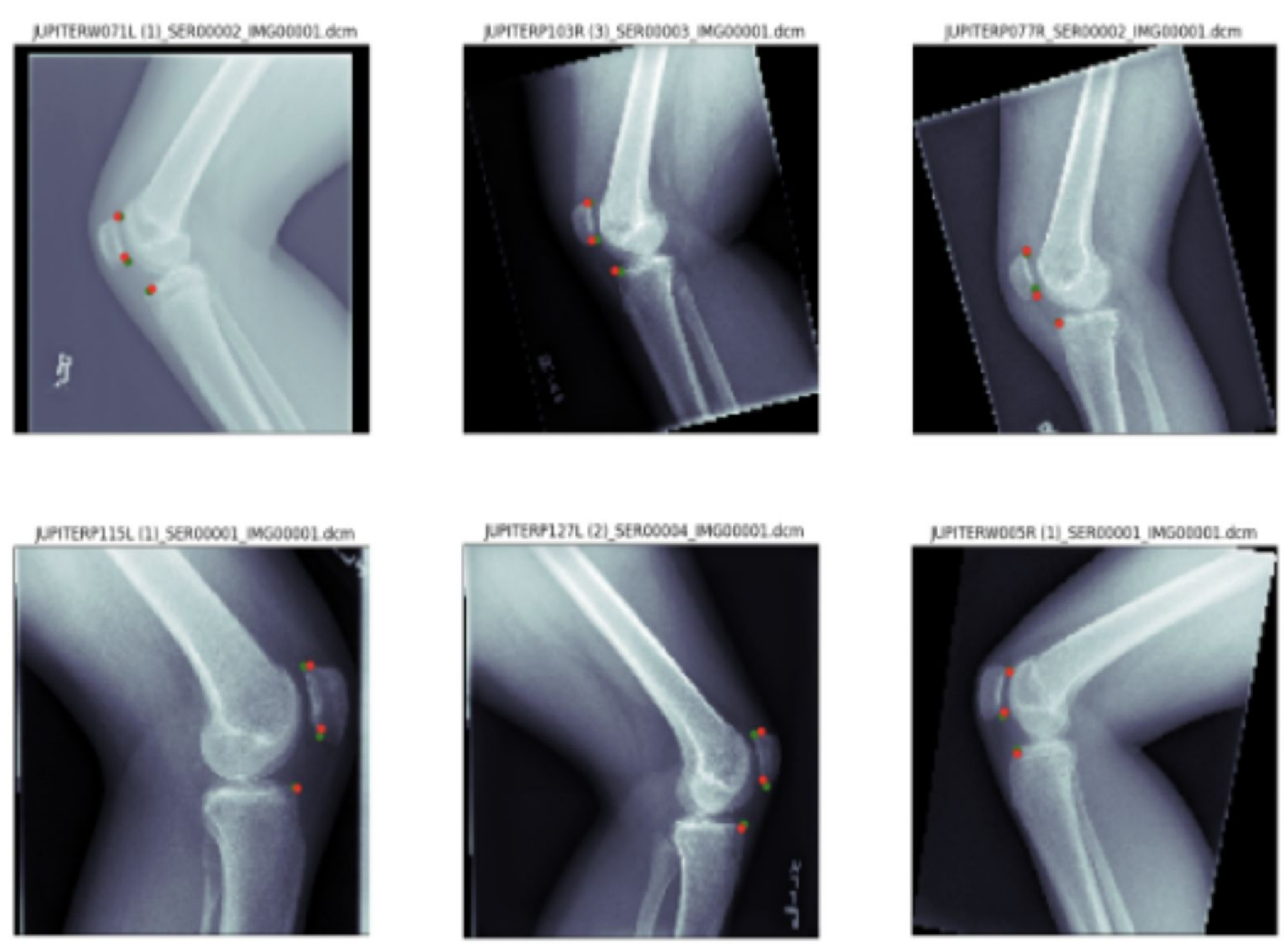
Model	CDI error	Key Point Distance	ICC
Baseline	0.42	8.68	-0.05
U-Net	0.25	4.5	-0.16
ResNet	0.21	3.14	0.22
AlexNet	0.22	4.77	0.26
VGG16	**0.19**	**2.05**	**0.33**
ModelY	0.06	0.06	0.91

Comparing Model Performances

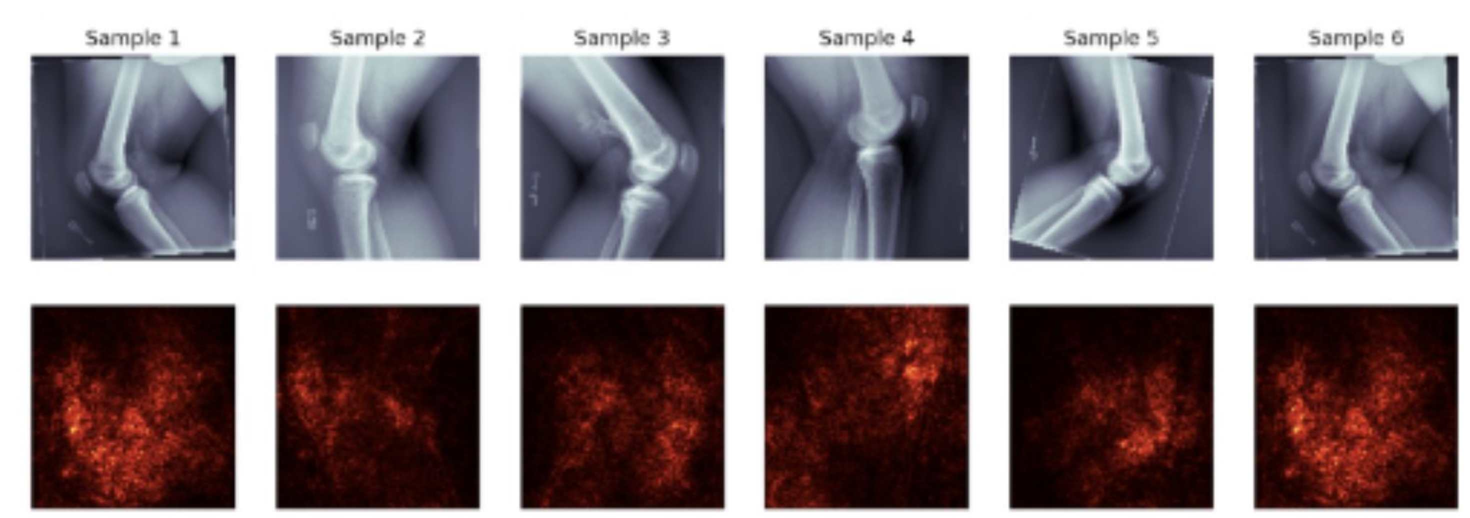


Key Point Dist: 1.4 (Train), 2.1 (Test) Key Point Dist: 1.6 (Train), 3.5 (Test)

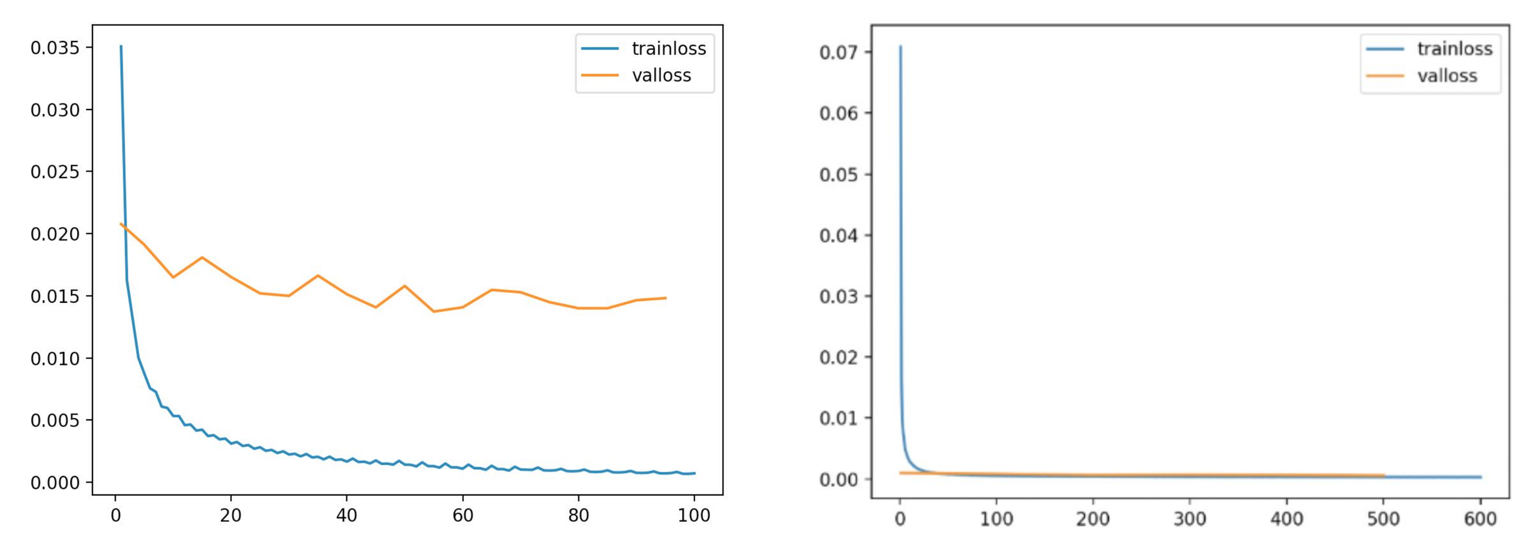
VGG16 feature map visualization pre-trained (left) vs not pre-trained (right)



VGG16 predictions (red) vs Ground Truth (green)

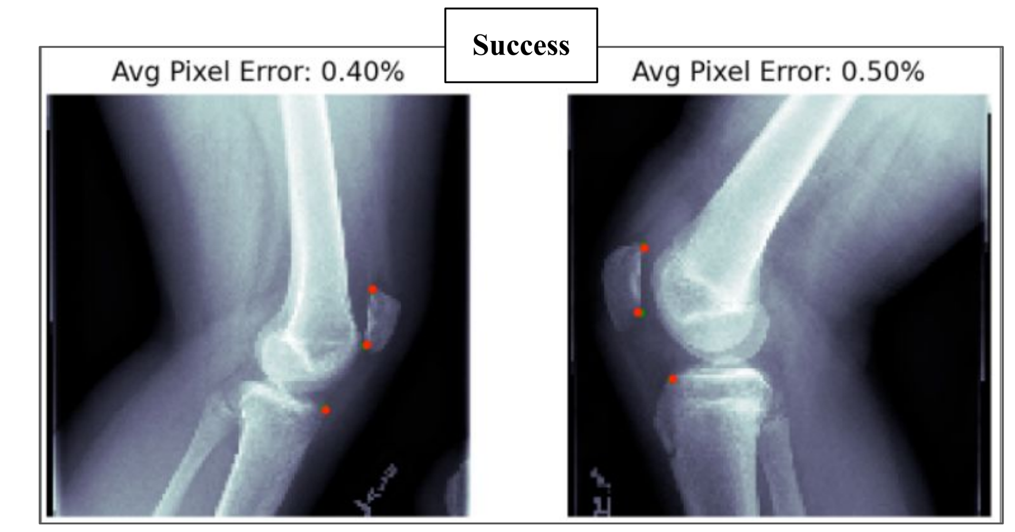


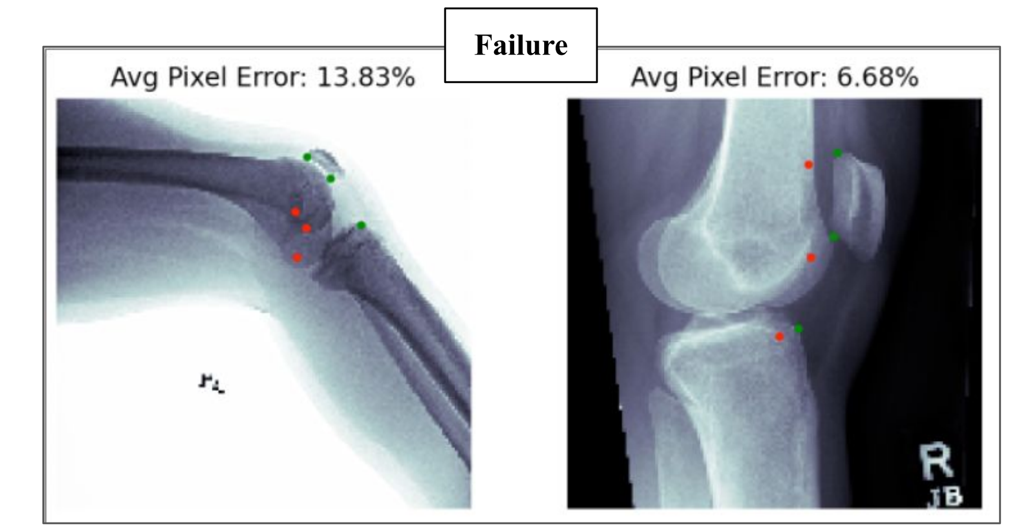
Saliency Map visualization for selected X-Ray Samples



Train and Validation loss plots comparison between Baseline (left) and VGG16 (right)

Success vs Failure Examples





Conclusion

- Outperforming all other models, VGG16 deep CNN can identify the 6 key points required on a knee x-ray to calculate the CDI even when trained on a small dataset.
- This superior performance is obtained by leveraging data augmentations and starting with a pre-trained model on the ImageNet dataset.
- Pre-trained VGG16 outperforms a VGG16 trained from scratch.
- Detecting the patella plays a significant role in the model's performance while the location of the tibia and femur also contribute but to a lesser extent.

Future Work

- Additional hyperparameter tuning related to augmentation and optimization methods.
- Experimenting with additional CNN architectures such as a pre-trained U-Net or deeper networks such as VGG19.
- Collecting more x-ray images to train on and exploring x-ray image quality improvement which may help minimize overfitting.

References

[1] Qin Ye, Qiang Shen, Wei Yang, Shuai Huang, Zhiqiang Jiang, Linyang He, and Xiangyang Gong. Development of automatic measurement for patellar height based on deep learning and knee radiographs. European Radiology, 30(9):4974–4984, 2020.

[2] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

Stanford University

Contact Information

Micael Tchapmi (mtchapmi@stanford.edu) | Jacob Azoulay (jazoulay@stanford.edu) | Daniel Lee (dklee@stanford.edu)

Stanford ENGINEERING