Patellar instability is a condition where the patella dislocates 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

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.

### Comparing Model Performances

<table>
<thead>
<tr>
<th>Model</th>
<th>CDI error</th>
<th>Key Point Distance</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>0.42</td>
<td>8.68</td>
<td>-05</td>
</tr>
<tr>
<td>U-Net</td>
<td>0.25</td>
<td>4.5</td>
<td>-16</td>
</tr>
<tr>
<td>ResNet</td>
<td>0.21</td>
<td>3.14</td>
<td>0.22</td>
</tr>
<tr>
<td>AlexNet</td>
<td>0.22</td>
<td>4.77</td>
<td>0.26</td>
</tr>
<tr>
<td>VGG16</td>
<td>0.19</td>
<td>2.05</td>
<td>0.33</td>
</tr>
<tr>
<td>ModelY</td>
<td>0.06</td>
<td>0.06</td>
<td>0.91</td>
</tr>
</tbody>
</table>

### Success vs Failure Examples

- Baseline Model vs VGG16 Model
- VGG16 predictions (red) vs Ground Truth (green)
- VGG16 feature map visualization pre-trained (left) vs not pre-trained (right)
- Saliency Map visualization for selected X-Ray Samples
- Train and Validation loss plots comparison between Baseline (left) and VGG16 (right)

### 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
