We spent more than four months investigating related literature and it really helps us fully understand the project logic and detailed information. Instead of traditional methods like IMU-based tracking or color-based tracking, researchers are more focused on performing markless shape estimation in recent studies. Also, it's popular to use machine learning based methods for robotics shape estimation and use images captured from different types of sensors as raw dataset. The previous research has applied several types of methods: Supervised learning, Semi-Supervised learning, Unsupervised learning, Transfer learning, Reinforcement learning. The input and output of models are vary based on individual tasks.

For our project, we utilized transfer learning which can be understood as Storing knowledge gained solving one problem and applying it to a different but related problem. We mainly focus on the pre-trained model approach in transfer learning and this type of transfer learning is common in the field of deep learning. These four pre-trained models are the major models we used and all four models are pre-trained on ImageNet. The ResNet model is a strong model and 50 represents layers deep. The MobileNet V2 is a lightweight model since it uses depthwise separable convolutions. Different numbers stand for different width_multiplier. The EfficientNet is a powerful model and normally has faster training speed and better parameter efficiency. Starting from B0 to B6, with the number increase, the ability of the model increases.

DeepLabCut is the main tool we used in this project. It is a toolbox for markerless pose estimation of animals performing various tasks. It performs frame-by-frame prediction therefore it can be also used for intermittent occlusions. It also helps to solve the detecting body parts problem in dynamic environments. The latest version improves the model's robustness.

To better present the project to you, I will first illustrate the overall technical logic of the project here. We started with the **Data Curation** which includes Frame Extraction, Frame labeling and training set creation. In our **Methodology** part, we will explain how we train the neural network to extract image features and how we perform the bending angle shape estimation. During the Risk identification and Alternative approach, we will present the project risks and the methods we use to address those risks. After that, we will demonstrate our model results and elevation metrics in the model performance part.
Now, Let's welcome Steven to illustrate our project data curation workflow.

Before we present our result, it is pretty important to introduce some model evaluation metrics. It will help us to evaluate the model's performance precisely. We applied our manually labeled coordinates as our ground truth data since the Deeplabcut predicted all image frames of the video. We first evaluated the initial model performance by computing the mean average Euclidean error, the MAE (which is proportional to the average root mean square error between the manual labels and the ones predicted by DeepLabCut). So if the MAE of the initial model doesn't look good, we utilized the likelihood function from tensorflow and visualized the overall label likelihood to set up the optimal p-cutoff value for fine-tuning. Normally, the frames with a likelihood lower than 0.8 will be extracted as outliers. After label fine-tuning, we utilized our own average distance function to check the new model shape estimate performance (RMSE) and then created a video with labels and skeleton. With angle detection, we adopted Bending angle accuracy based on the RMSE value to present the application of shape estimation.

All the results we display today are the videos after refinement. We started our model implementation with the Pink Actuator video. We assessed different pre-train model performance with this same label position - Pink sealing. The resolution of the Pink actuator video is 1080*1920 which results in 2073699 total pixels. From these videos, it is clear to see the model pre-trained by ResNet_50 is more stable than others with the smallest test error (3.5px)

To determine which kind of label position has the most suitable tracking movement, we experimented on different label positions with different pre-trained model comparisons. ResNet_50 tends to have the best performance among all these label positions.We only display the result videos of Resnet_50 due to the time limitation. From these videos, we can easily observe the sealing label is more suitable for performing shape estimation of this actuator

We also adopted this pipeline on other color actuators. Since the pixel sizes of these raw videos are different from the Pink one. We can only compare these four colors at this phase. These videos have the same resolution 640 *480 with 307200 total pixels. The videos we saw now are pre-trained by Resnet_50. The orange one resulted in a good performance with 1.35 pixel test error in average distance.

We utilized angle detection as an shape estimation application in rehabilitation soft-robotics, which means the accuracy of the angle estimation is curial. We created a real-time bending angle estimation animation as it tracks the robot. On the right you can see again the model that is tracking the pink sealing of the actuator. On the left, you can see the results of our predicted angle shape estimation, which calculates the bending angle at each time stamp. Each angle is calculated by looking at the vector between the first two points and the vector between the last two points in the skeleton. The vectors of any two points in the skeleton could potentially be used to infer other information about the shape of the robot. We only had the bumping graph at first. After we utilized the moving average convolution with the numpy library, we obtained a smoother graph. The angle plot you see here is the result of smoothing with a moving average convolution of 3 time steps. This Smooth graph actually improves the angle estimation performance.

Here are the relative references we used in this project.

Now we are heading up to our discussion time~ Any question about the project or anything you want to discuss with us regarding this project?

Before closing our project section , we would like to thank CUSP for giving us this valuable opportunity to participate in an industry-related project, which allowed us to learn to communicate with our clients and understand their intentions and accomplish their needs while we were still in school. We would also like to thank MERIIT Lab for providing us with this special health related project that allowed us to practice machine learning and deep learning in a practical way. A special thanks to Dr. Jacquline Libby for her advice and guidance during this project over 8 months. The whole team chose this project for the same reason: we are passionate to utilize high technology and artificial intelligence to improve the quality of life for the whole society and better help with human medical development. Do our best to reduce unnecessary deaths and disabilities.

.