Luka Tom Eerens

Team D: Deeply Emotional

Teammates: Luxing Jiang, Keerthana P G, Ritwik Das

ILR03

October 27, 2017

Individual Progress

Since our last progress review, we were notified that we needed to have the hardware complete by the FVE. As a result we quickly re-steered the course of the ship so as to quickly design the final robot and order parts. I was solely responsible for doing this, and this involved conducting trade studies on components. Given the small size of our robot, which has very few components, I had a lot of freedom from a financial point of view to choose components that worked well given their performance requirements.

I conducted the trade study as I built the skeleton of the robot, choosing components as I added links and joints to the SolidWorks model. We ended up deciding to use the Microsoft Kinect camera for visual input into the neural network, and a cheap 48 kHz microphone that can be found on eBay or Amazon for the audio input into the neural network.

Based on the weight of the metal frame, as well as the Microsoft Kinect and the microphone, Stepper Motors were used instead of servos. The NEMA-34 Steppers will surely provide enough power to rotate the camera about the Y-axis which is perpendicular to the table-top, and about the X-axis so as to tilt it upwards or downwards.

Version 1 of the CAD rendering is presented below. Notice how it omits the microphone. This has been currently left out as we are still debating whether to have the human hold the microphone or place it on the robot. Microphones placed on the robot will pick up noise made by the steppers quite well and the result could corrupt the purity of the audio waveform that is fed into the system.

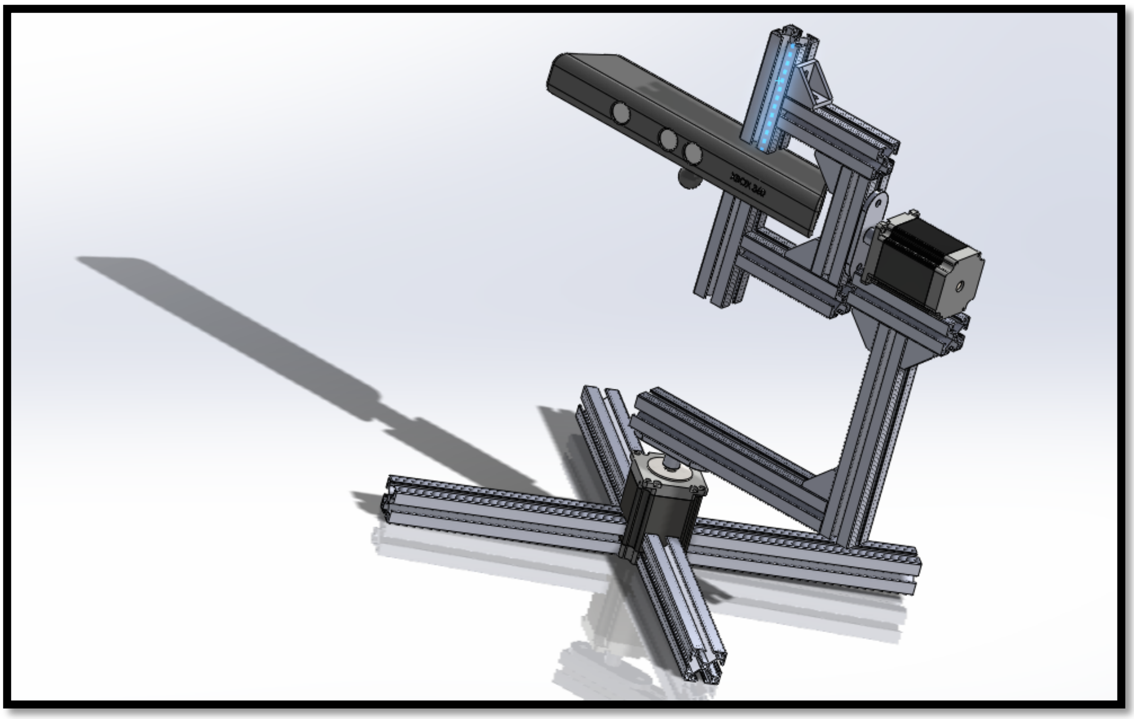

Figure 1: Version 1 of Team D Robot

It is also called version 1 because one of the parts has not been found on the market, and will need to be 3D printed. This is a piece that mates with the shaft of the stepper and mates with the metal frame that the stepper is suppose to rotate. After we conduct loading simulations on the part, and confirm that it can perform within the expected stress regime and survive sufficient loading cycles, we will set the design in stone and proceed to ordering the parts. Until then, this design will be the called version1.

Update—
We actually conducted load simulations of this piece and have found that given its current design, it holds up well with the expected loading cycles.

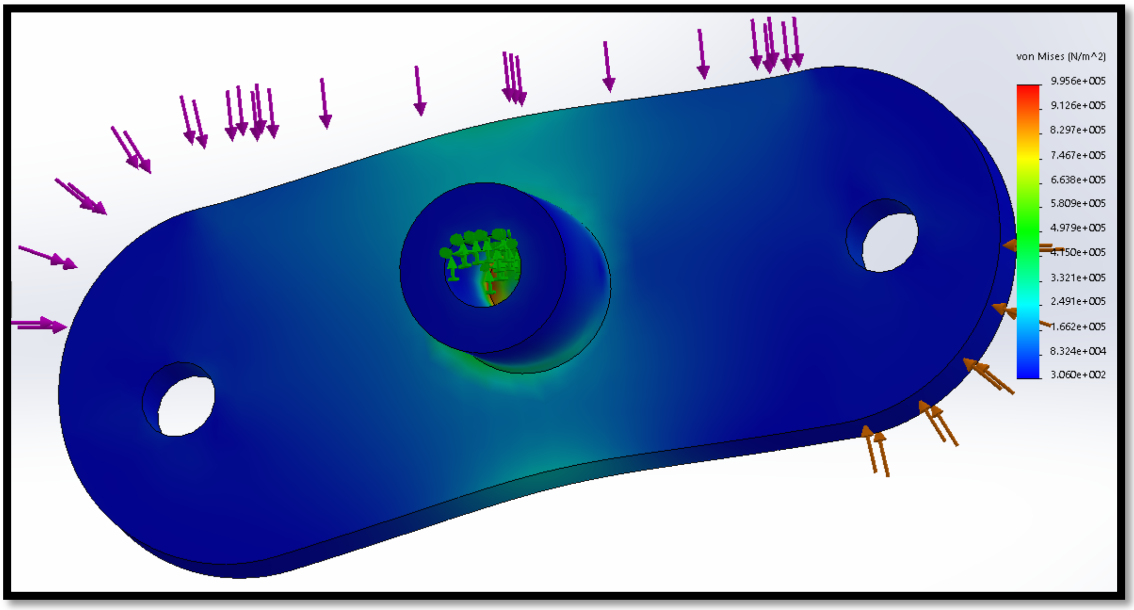

Figure 2: Von Mises Stress (Yield Stress of ABS: 2x10^9GPa, Highest Von Mises Stress: 1x10^6)

We will however modify it slightly by making it more athletic and thicker, just to be cautious and for it to be able to handle sudden loads as we may choose to put our hands on the robot.

An example of this updated version is:

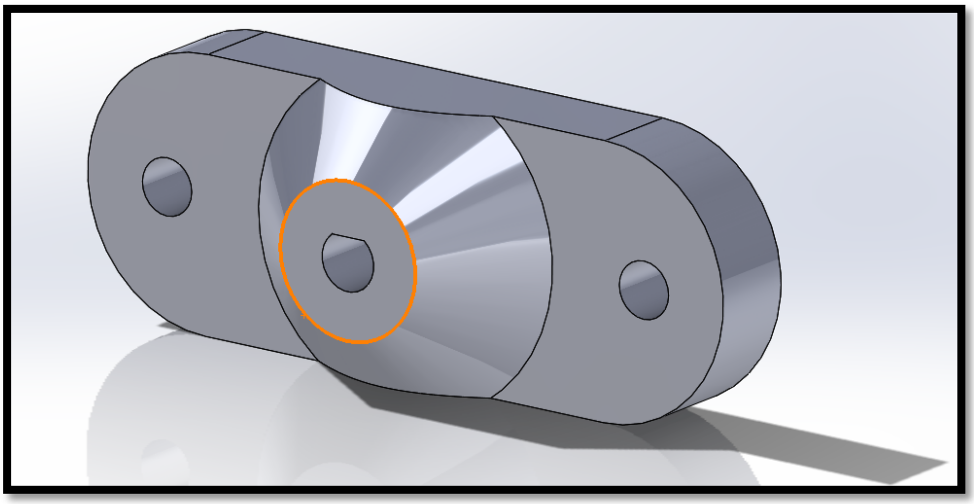

Figure 3: Version 2 of Stepper Motor Shaft Connector

Additional progress that I have made was on the software side. I developed a script that automates the parsing through waveforms and lexical dyadic script files for every folder in the Semaine dataset. This is essentially an automated data pre-processing pipeline.

A reminder of what this does… The computer parses through the text script of the interaction between a host and the test humans. It finds the time interval where the test human is speaking and then nulls the waveform where the test person is not speaking.

This is done because the waveforms for vocal input are not partitioned based on speaker, but contain the joint superimposition of waveforms from both speakers together as they respond to each other or make noise over each other. The operation I described earlier only isolates the speakers if they do not speak over each other, which is the case for about 90% of the time.

Thankfully Semaine contains waveforms of a Microphone that is placed closer to the test human and so the acoustic impurity of that waveform is not as bad. The network will hopefully be able to handle this as the loudness of the responses from the other person is slight notch higher than background noise.

<u>Challenges</u>

There was no challenge with developing the automated data pre-processing pipeline.

However, there was a challenge in designing the robot on the spot. We did not anticipate ever having to build hardware for our project, and as a matter of fact picked this project because we all were mechanical engineers/electrical engineers who wanted to gain experience in deep learning and so avoided mechanical projects like the plague.

So it was clear that there was no enthusiasm among the troops in regards to designing this robot. Since I had industrial design experience I decided to quickly get something done.

The main challenge since last week was having to be agile based on sudden requirements while still weathering the ship in the right directions with respect to our other deliverables. Deep learning related tasks I had to do were transferred over to others as I had to commit to designing the robot.

Teamwork

Ritwik:

Developed the LSTM and Speech Net further. He connected the ResNet18, Speech Net, and LSTM together and integrated them together into a single network.

Luxing:

Helped design the schematic for PCB for our project. He also helped Ritwik with concatenating the Speech Net and LSTMs together. He also tested the entire network with some random input.

Keerthana:

Keerthana worked a lot on the attention models for the speech analysis. She developed the attention model for the text input of the network. She also made the word embedding file that converts words into 300 dimensional vectors for each word. She also built a Bi-modal LSTM to convert the vectors into the encodings.

Plans

We hope to have ordered the parts to build our robot and ideally have some of them delivered and ready for assembly.

We also hope to train our current model.

During the period of training, we also hope to move onto the next version of our system that will include a better ability to digest tri-modal data.