Luka Tom Eerens

Team D: Deeply Emotional

Teammates: Keerthana P G, Ritwik Das

ILR05

November 22, 2017

Individual Progress

My progress since the last ILR has mainly revolved around debugging further issues with the dataset and pre-processing, as well as mechanical design and assembly of the robot. I was also able to obtain the CREMA-D dataset from one of its researchers who I emailed a while back when we were doing the literature review. We will not our training solely to it, but rather tap into the diversity of facial expressions from a large number of people in order to improve the networks performance for the visual modality.

I have started returning the parts that we do not need so that we can get some reimbursements as we spent quite a fortune on expensive items like the Kinect that we did not need. Since the FVE is close, I have also placed an order for a large supply of contingency parts so that if a calamity where to befall our assembled robot, that we could promptly reassemble it.

On the topic of mechanical design and assembly of the robot: I redesigned the robot to be 3D printable in its entirety. The result of this was something that was lighter and that we had a better control off in terms of geometry of parts that mate with the camera. Through trial and error I was able to design a robot that really tightly secures the webcam in place and prevents it from sliding all over the place. We are even able to hold the robot upside down and the camera will not slip off. This 3D printed body was printed in 2 components, one, which was the main body which resembled an I-beam column and the other which was mounted at the top of this body to mate with the camera.

Later this design was changed and is currently in the process of being manufactured (None of the 3D printers in the MRSD lab are working). The changed design reflects the outcome of an integrated test of the robot body with the stepper. It turns out that the vibrations caused by the stepper motor reverberated pretty strongly through the 3D printed body of the previous version and created a horrible buzzing noise at the contact point between the 2 3D printed pieces. This noise is unacceptable because of how it with would affect the acoustic sampling from the microphone for the SVE robot.

Next, on to the topic of debugging further issues, we have found that yet again, there were confusing outputs that did not necessarily break the code, but certainly affected the ability for the network to learn. The reasons for this yet again are not due to my code but rather due to the dataset that contained incomplete chunks of data sampling or completely ludicrous and impossible data. An explanation of this section will be provided in the challenges section of this report.

I discussed ways around this with Ritwik and decided that we would simply ignore any of the data that was rogue after evaluation whether there was any merit based on the time we had left to develop code that handles these frustrating rogue data samples.

Challenges

What do I mean by rogue? An example would be a video that is 3 minutes long being rated for only 6 seconds in one of the dimensions of emotion. This is of course unhelpful

to us because thought our strategy for dealing with varying duration of rating would be to learn for as long as the length of the shortest rating. This was originally chosen because you cannot add artificial labels to label files that are not long as the others, and so trimming the longer ones to the length of the shorter ones was the solution.

We now have a case however where all label files are around 3 minutes long and one is 6 seconds long, this would make the 6 second long label file the longest and so every other label file would be trimmed down to 6 seconds in length. This of course reduces the quantity of data going into these networks.

Ritwik was able to narrow down the search space to around a dozen subfolders that had this problem, and within each of these we had to manually scan through every txt file to see how long it was and record that. This tedious task took a while but we found around 5 folders that contained rogue data. One of them, folder 55 of the SEMAINE dataset contains labelling by all the raters that lasts almost twice as long as the entire video.

It is possible that the correct intended labels of the video in folder 55 lies within this longer than expected label file but we would not be able to detect where it starts or ends. We just decided to discard file 55 altogether. For the other files we were lucky enough to generally have 2 folders by the same rater for the same emotion dimension, one for example called A5R22TUCPrDE.txt and the other A5R22TUCPrDE2.txt with the latter being the full length one. We could thus only pay attention to the latter and ignore the former. What is annoying is that the rest of the correct files in SEMAINE are named after the former and so building an automated data pre-processing pipeline was ultimately limited by the lack of organisation of the dataset.

As far as the actual design of the robot and its assembly, the main challenges were technical but now seem to related to the 3D printers not working. I have tried using the 3D printers from the Microsoft Space but they are too small to make the part we are looking for. I did not anticipate the vibration from the stepper motors to resonate so much and if I did I would have certainly designed the 3D printed part as one piece even though it would have been heavier than if printed separately.

Teamwork

Ritwik:

Trained the neural network, and suggested avenues we should follow based on the results of the training. He found ways to get around a loss saturation problem by unfreezing the layers of our Resnet-18 network so that it could learn to recognize emotions as opposed to just recognizing objects. He has also begun training our network on facial expressions, namely the CREMA-D dataset.

Luxing:

Is no longer in our team.

Keerthana:

Integrated the face detection system with the serial communication to Arduino and the Arduino stepper control. She therefore tested the ability for the system to track people

in 1D. Aside from this she has been helping the two of us with our current challenges. She has done the bulk of the tasks that we would have otherwise assigned to Luxing.

Plans

The FVE is looming and we will need to take full advantage of the thanksgiving holidays in order to troubleshoot issues related to the accuracy of the emotion recognition system.

The plans are to have reached all of our FVE deliverables namely getting the 1D human tracking robot to work with the performance requirements, and the other being the emotion recognition performing with the required accuracy on a subset of the SEMAINE dataset.

We will also need to rehearse the human tracking for the demo so that we can showcase exactly what we intend to, nothing more and nothing less. We will hopefully solve all the main issues by then, and are slightly relieved by the good progress Keerthana has made with regards to the Stepper Motor based human tracker.