

## **Cohort 12 3<sup>rd</sup> Meeting Minutes**

**Group:** 15, **Chair:** Rael Ware, **Minutes:** Gia Croock

**Mohammed Haffejee and Thabo Tshabalala- App to scan pool water test strips:** Developed an application (the camera part). Struggling with detecting the edges. The algorithm used was giving gaps. Managed to use a system called dilation which fills the gaps. Main aim to use unsupervised learning technique to classify colours. Trying to find other algorithms for colours. Using the K mean clustering. Need to consider colour spaces. Supervised requires a lot of time to create data set therefore decided to use unsupervised. Idea to consider from Dr Bekker is to make the white next to the test strip a feature so that the algorithm uses this as a feature. The supervised learning does actually do this, it looks at the strip as a whole (the block and the bit next to it). The algorithm extracts all the colours. Questions to consider: What do you use as your baseline? Are you weighting your white? Can you also use the white to measure the lighting conditions and if you know its not white then maybe you need to adjust all the colours.

**Benjamin Palay and Raphi Druion- Facial recognition:** Developing a static web app in angular to get a user interface. Managed to deploy. Managed to filter faces based on ethnicity. Carrying on with the facial recognition. Managed to get algorithm that reads in everything about colour. Grouped 800 colours into 5 different colour names. An issue they are facing is that they are getting eyes that are meant to be blue to be silver. Don't know if they going to use prebuild models. Next week they going to try train their own model for the limited scope they created. The problem is the pretrained models are too specific. They thought of a continuous scale, but the problem is they must classify that scale. Suggestion: just let the user choose a colour visually. Only want the most common eye colours so that all images can be filtered. Questions to consider: Can you use the reduction already made to group eye colour and then refine search. There are lots of manual classification. Potentially can run averages.

**Ahmed Ibrahim and Kevin Naidoo - The lost cities of South Africa:** Preparing the pipeline for the model. The satellite images for training and testing are missing. Goal for next week is to obtain these images and prepare it to see what kind of results they are getting. Contacted the computer science student to get data.

**Johann Gouws and James Allsop - Mapping the Rainbow (twitter):** Not a very productive week. Managed to generate data sets needed for the graphs. Managed to visualise the graphs on HTML. Next week they are moving on to cleaning up the data.

**Joseph Baggott and Hraklis Papageorgiou- The world's protest capital:** Worked on trying to link the process data. They have their two data sets and are on the verge of linking them. Once this is done, they can be automated. Had problems with the locations but they solved it. Twitter doesn't mention how they get the location. Content sentiment analyses as a general concept will be worked on next week. Mainly considering English and then may expand.

**Devlan Mckenzie and Muchaveleli Manjat- Network firewall management and visualisation:** They worked on visualising the firewalls. Saved the firewall rules into a text file. Trying to make sense of the different tags. Managed to quantify the firewall rules. Looking into how to run the app on the cluster. If they classify their app as containers, they can be classified. They are using C++ which is not compatible with Doca. Manually identifying the tags. Made a simpler set of rules that read from file and assign IP data. Goal for next week is to get the visualisation working with some basic data.

**Nkosingiphile Ndabandaba and Ishmael Sithole- Fast genotype calling:** Started by programming a code that will extract data. Instead of specifying the path that the data is extracted from the user should specify the path. Biggest challenge was getting logged on to the cluster, but they managed to get that working. Looking at how to copy the local work to the cluster. Next step is to compile the code. Look at using git to copy to the cluster. Use slurm to run your stuff you can see the progress. Goal is to get the cluster working so that they can test to see if what they have created is able to extract the correct stuff.

**Robin Jonker and Tristan Lilford- Location aware scientific workflows:** Last week the system was working statically. This week's challenge was to get it working dynamically. It is working but not as they want it to. Started getting the Execution reports done. Getting comparable testing environments is not ideal because the cluster is being used by many groups. Using slurm scheduler for running the jobs. Will look to use AWS if they have time.

**Rael Ware and Gia Croock- Cheap Air Quality Reporting Station:** The device needs to be fully completed by the end of this week as the device is going to be evaluated and tested using the air quality reference monitors at SBIMB next week. Still waiting on the components necessary to finish implementing the GSM module. Spent the week preparing the device for testing this included 3d printing the devices housing, soldering the components onto the Vero board, and adjusting the code to ensure the results obtained from the device are comparable to the results obtained from the reference monitor. Spent time refining the evaluation and testing procedure. Reducing the power consumption of the device is still providing a significant challenge. While the device is being tested next week, they will develop our own API at the moment they are using things speak.

#### **General notes:**

It is important to have a GitHub repo and make sure you are doing things in a proper engineering way and not just doing it at the last minute.

Next week Mohammed and Thabo are chairing the meeting and taking minutes.

Next week there will be a projector for demonstration.