**Brief research for Visual Question Answering**

The topic I have chosen is visual question answering, and the topic is about using a picture as input, analyze it using machine learning. After the process, we can achieve the function of a user asks a question using plain text, and we use natural language processing tools to analyze what the user is asking and feedback the answer.

I chose this topic because if someday, we can improve its efficiency to the point that quick analysis and feedback are possible. It can help visual-impaired people by providing things like hazard warnings or helping them locate buildings, which can be very beneficial for society. Right now, there aren’t many real-life applications I can’t find, and this might be because the processing time is taking very long, limiting its usage. In most situations, we are expecting real-time feedbacks, and in most cases, if this feedback is not in real-time, it is not helpful for us.

After doing some research, I found an open-source framework for visual question answering called MMF. It is from Facebook AI Research and is powered by PyTorch using the Python language. Its last GitHub release was back in Aug 2019, but it is still maintained by more than 20 people. The latest update was within this month, so I think it’s safe to say it is still an ongoing project. The license it uses is a 3-cause BSD license, which means that we can use it to redistribute and modify it any way we want as long as we keep the 3-cause BSD license and cite the original author.

Right now, the most common way to achieve visual question answering is by extract information from using methods like CNN, the language is processed separately using language processing methods, things as natural language parser, and then these two data are going through a classifier, and give the feedback from a set of predefined answers.

Overall, I think visual question answering topic still has a long way to go before we can use it in our daily life, but I do think when that day comes, it can help a lot of people in many ways include education, picture classification/search in social media and helping visual-impaired people.