## 7/13 - 7/19 Progress Report

1. **Reading List**
    a. Rough Reading
        i. Ross, Andrew Slavin, and Finale Doshi-Velez. "Improving the adversarial robustness and interpretability of deep neural networks by regularizing their input gradients." *arXiv preprint arXiv:1711.09404* (2017).
        ii. Smilkov, Daniel, et al. "Smoothgrad: removing noise by adding noise." *arXiv preprint arXiv:1706.03825* (2017).
    b. Intensive Reading
        i. Goodfellow, Ian J., Jonathon Shlens, and Christian Szegedy. "Explaining and harnessing adversarial examples (2014)." *arXiv preprint arXiv:1412.6572*.
2. **Ideas and Deliverables**
    a. Comparing Saliency map of Original and FGSM images (started a week ago)
        i. The goal is to get the saliency maps of any input image as well as the saliency map of the corresponding FGSM image. This way we can compare how the model evaluates the pictures differently
    b. Integrating Inception V3 (started a few days ago)
        i. The goal is to use the same model on both the Adversarial FGSM code and SmoothGrad. I have chosen to use the Inception V3 pretrained model so I can use images of 1000 classes.
3. **Deliverables**
    a. RIght now I am able to generate the adversarial image of pictures in the classes of the Inception V3 model. I will soon be able to apply SmoothGrad on both images to see how the saliency map changes.
4. **Plan**
    a. The next step is to figure out what to do now I have the entire baseline set up. Depending on the results I get after applying SmoothGrad from it, there are different directions to go. This is something we can discuss on our meeting Tuesday morning 7/2