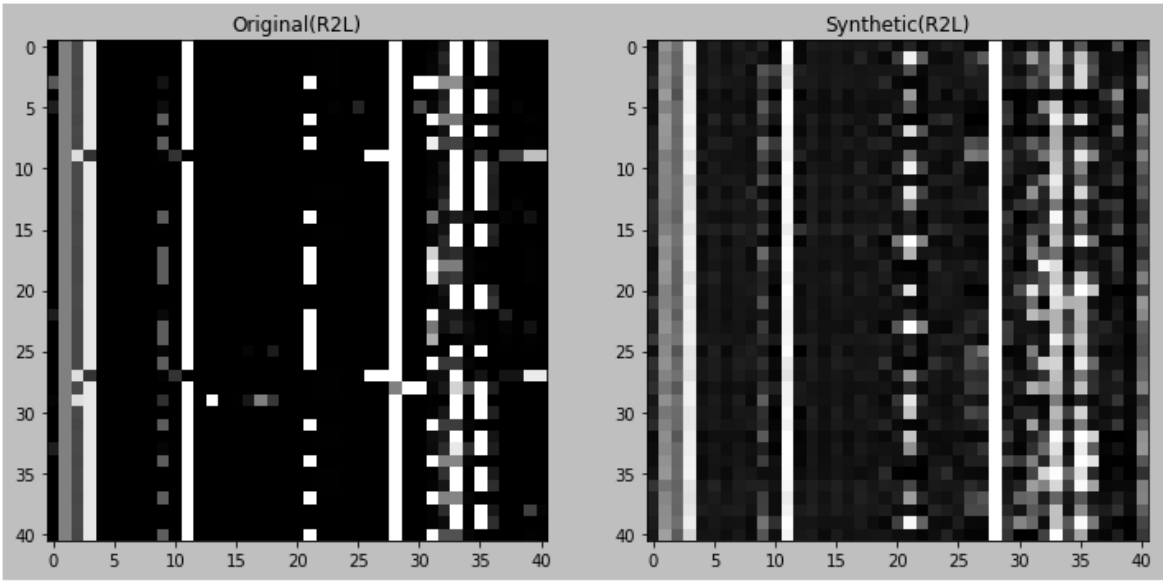
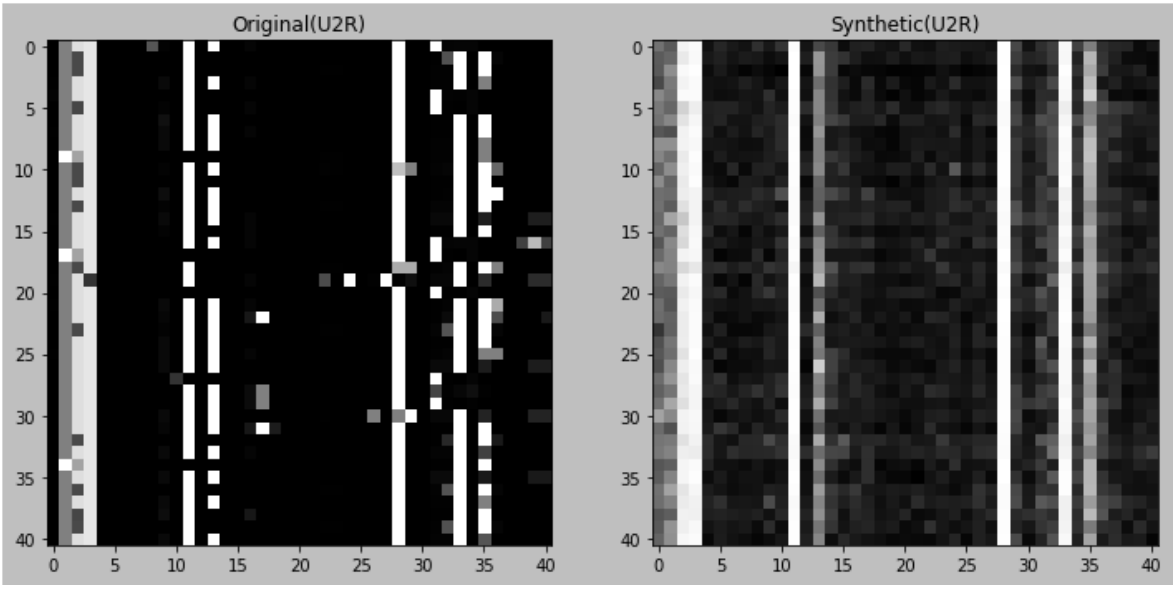
Task:
Use GAN to generate samples for minority attack classes
Output:
Visualized 41 instances from Original and Synthesized data
X-axis: Features (41)
Y-axis: Instances (41)

Following are the traffic classes along with percentage in the original Dataset

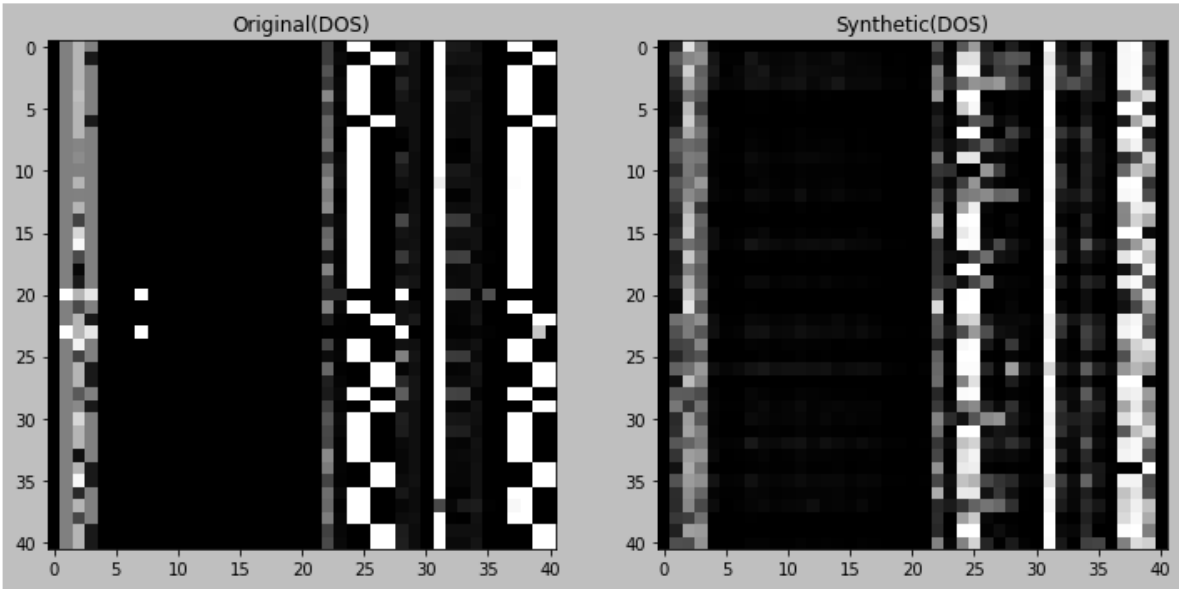
R2L Attacks (0.79%)



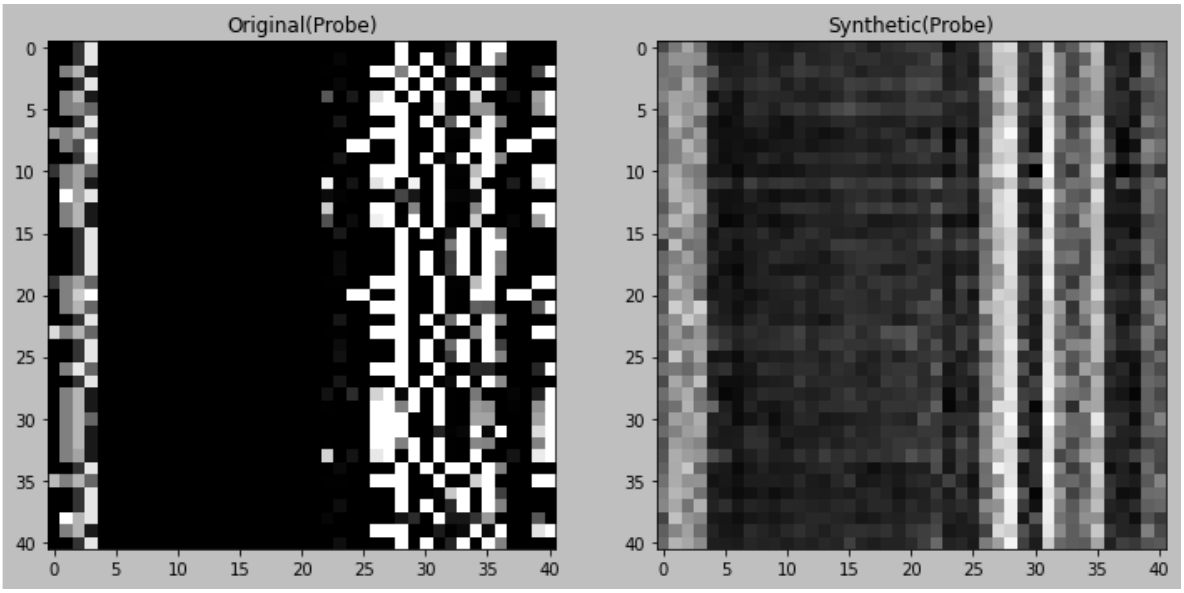
U2R Attacks (0.04%)



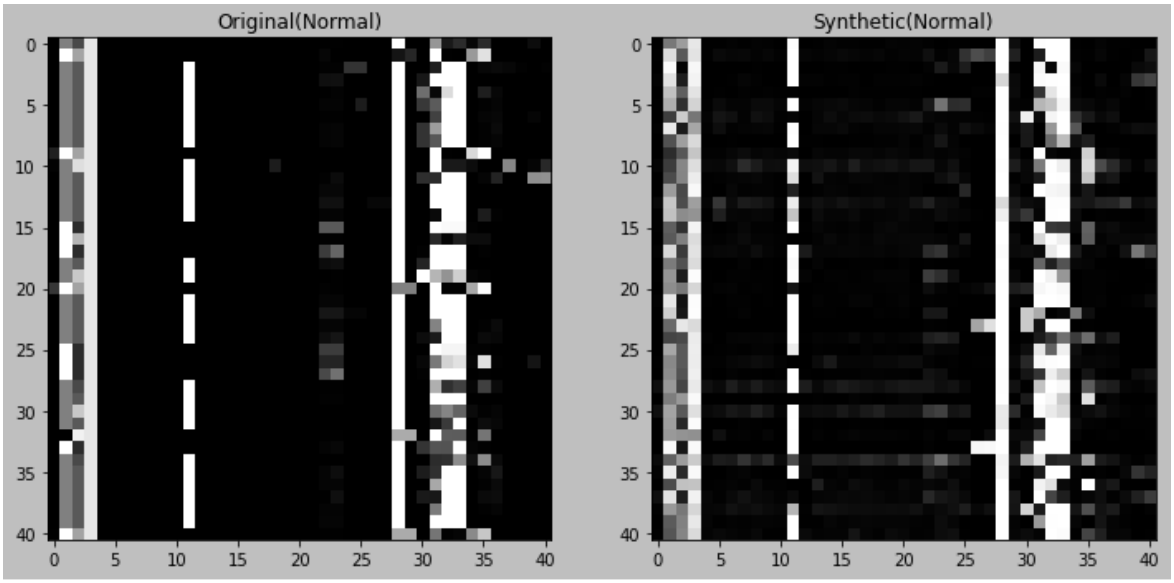
DOS Attacks (36.45%)



Probe Attacks (9.25%)



Normal Traffic (53.46%)



Conclusion:

1. Used DCGAN with Wasserstein Loss to learn the distributions.
2. This technique was able to catch some of the patterns as observed from the visualizations.
3. Added a few thousand R2L attacks to the original dataset and used a deep CNN for the detection task.
4. Adding a few thousand R2L attacks to the original dataset slightly improved the accuracy.
5. There's still room for improvement in Generators, especially for the **U2R** and **Probe** attacks.
6. Different variants of GAN can be used to improve the Generators, such as WGAN-GP.

Next:

1. Improve the GAN architecture (WGAN-GP).
2. Tune GAN parameters for each attack class.
3. After improving U2R and Probe Generators, generate an equal number of samples and add them to the original dataset.
4. Evaluate a deep CNN on the new balanced dataset.