Week 1

Paper Survey

- **Grasp LDM**
  - **Introducing Diffusion model in the latent space of VAE improved performance.**
  - **Input: Mesh file**
  - **Output: Can visualize Grasps**
- **Economic Grasp**
  - **For input it demands many parameters like multiple RGB images at different camera angles, camera intrinsic and extrinsic properties....**

So, we decided to stick with Grasp LDM.

Week2,3

Training Data Generation

- **Modifications of DNJPs script were made several times so that it behaves similar to Acronym Dataset (Data used to train Grasp LDM)**
- **A few major changes that are made involve:**
  - **Addition of both good and bad grasps. Which is crucial for a model to predict the type of the grasp so that it won't learn to tell all grasps as good but instead takes help of that label and get trained.**
  - **Meshes are also generated along with point clouds. Since the input format is a mesh, instead of converting a point cloud into a convex hull/mesh**

where a few data points can be erased, from script itself we now generate meshes.

Week 4,5

Training of the model

- **Training of the model consists of two major steps**
  - **Training VAE**
  - **Training LDM**

  Code is implemented in such that we get different loss plots in a tensorBorad platform called 'Weights and Biases'

- **Training is done from scratch so that we can observe the notable changes in our plots.**
- **Evaluation part of code was not explicitly part of the paper's implementation. But with a few in-built classes functions they can be achieved.**
  - **generate_grasp.py generates a grasp using our new trained model, predicting the label and comparing with the ground truth gives accuracy.**
  - **A good accuracy (when using LDM) has been achieved (Around 85-90%)**

Now working on visualization of the results such that they look like expected. Because visualization of Grasp LDM grasps is different from ours. Our plots have different colors (gives green to good, grey to bad grasps, blue for the robot hand) whereas the Grasp LDM plots have anchor shape like structures to indicate the grasps.