GROUP-6

Video Prediction and Forecasting on the KITTI Dataset

**1.Selected Problem and Reason for Selection:**

We have chosen the problem of predicting future video frames using the KITTI dataset. This problem is motivated by the potential real-world applications in autonomous driving.

**2.Database/Dataset Used:**

We will use the KITTI dataset, a comprehensive computer vision dataset that includes over 100,000 video frames along with synchronized data from various sensors like RGB cameras

**3.** **What deep network will you use? Will it be a standard form of the network, or will you have to customize it?**

Our approach will utilize a convolutional encoder-decoder architecture, incorporating 3D convolutional layers to capture the spatial and temporal dynamics of the video. Additionally, we will explore integrating recurrent neural network layers, such as ConvLSTMs

**4. What framework will you use to implement the network? Why?**

Pytorch- PyTorch is particularly well-suited for rapid prototyping and experimenting with novel deep learning architectures, which aligns with our project needs.

**5. What reference materials will you use to obtain sufficient background on applying the chosen network to the specific problem that you selected?**

We are going to use research papers on this topic like- Video Prediction Using Improved Transition Dynamics, PredNet: A Predictive Coding Network for Video Prediction and Unsupervised Learning

**6. How will you judge the performance of the network? What metrics will you use?**

Mean squared error (MSE) and the structural similarity index (SSIM), along with other video-specific metrics.

**7. Provide a rough schedule for completing the project.**

* Data Preprocessing and Exploration (3 days)
* Model Design and Implementation (2 week)
* Training and Optimization (1week)
* Evaluation and Analysis (1 week)
* Documentation and Presentation (1 week)