- Aim was to use infinite amt of unlabeled imgs and vids to learn good feature reps. GANs are a good alternative to maximum likelihood and their implicit cost func is good for representation learning. Their intermediate layer reps can be used as feature extractors.
- GAN architecture was constrained for stable training.
    - Spatial pooling was replaced with strided conv to allow model to learn slowly. It was used in generator also to make it learn spatial upsampling as fraction-stride.
    - FCs after conv features are removed. It was found global avg pooling increased stability but hurts convergence. So highest conv features were directly connected to input of gen and output of discriminator.
    - BatchNorm was removed after output of generator and input of gen to decrease instability.
    - ReLU is used in gen but Leaky ReLU in discriminator. Tanh bounds output of gen which helps to learn quickly to saturate and cover color space of train distribution.
- Only preprocessing was scaling imgs to [-1,1] range. Batch size = 128, Weight initialisation = N(0,0.02), Adam with LR = 0.0002, and $\beta = 0.5$ for stability.
- It was recently analysed that learning speed and generalisation performance are directly linked. So generated imgs after 1 epoch and after convergence were visualized to show that the model did not memorize train data as imgs improve in quality slowly. Duplicates were removed to decrease chance of over-fitting.
- Using feature reps from DCGAN outperformed K-means baselines but was inferior to exemplar CNNs (unsupervised) which trains model to classify surrogate classes using heavily augmented centre crops. DCGAN is robust as it was not trained on CIFAR.
- Visualisation : Nearest neighbour is not used as it is easily fooled by small changes. Log likelihood was also not used as it is also poor. Walking the latent space by interpolation shows semantic changes in imgs => model learnt good reps.
- Window filters/feature maps were removed which resulted in generator using other objects in place of windows.
- Vector arithmetic to manipulate generated imgs was applied with avg of 3 imgs for more stability. It is an alternative to conditional GANs.