Robotics Research Center, International Institute of Information Technology, Hyderabad

Project Report - FRAME INTERPOLATION USING GENERATIVE ADVERSARIAL NETWORK Sep 2017

Autonomous navigation using visual systems has been a significant subject for researchers. One of the challenges faced during navigation of mobile robots is to overcome the unexpected obstacles. Such situations demand strong rotation of the vehicle which leads to significant loss in data from the camera feed resulting in inaccurate localization of the vehicle. This error in localization could be reduced by increasing the frame rate.

GANs have been successful in synthesizing high quality images. Hence, Frame Interpolation with Generative Adversarial networks (GANs) has been a novel way to this. We have shown significant improvement from FinniGAN [1] in terms of sharpness in the images and reduction of architecture.

Convolutions account for short range dependencies between pixels limited by the size of the filters and pooling layers for long range dependencies. But for frame interpolation the input and output should be of same size, as a result pooling would lead to loss of data. Modified Skip connections in the form of UNets, served as generator network. To improve the generated output, additional losses L1 (to capture lower frequencies) and Image gradient loss (Sharpness enhancement) [2] were incorporated into the output, which was then given to the discriminator along with the ground truth. We have considered training with WGAN-gp to see if we can overcome the convergence problem with GANs. Also experimented with conditional GANs [3], Pix2pix [4] and Cyclic GANs [5] during this process.

ARCHITECURE

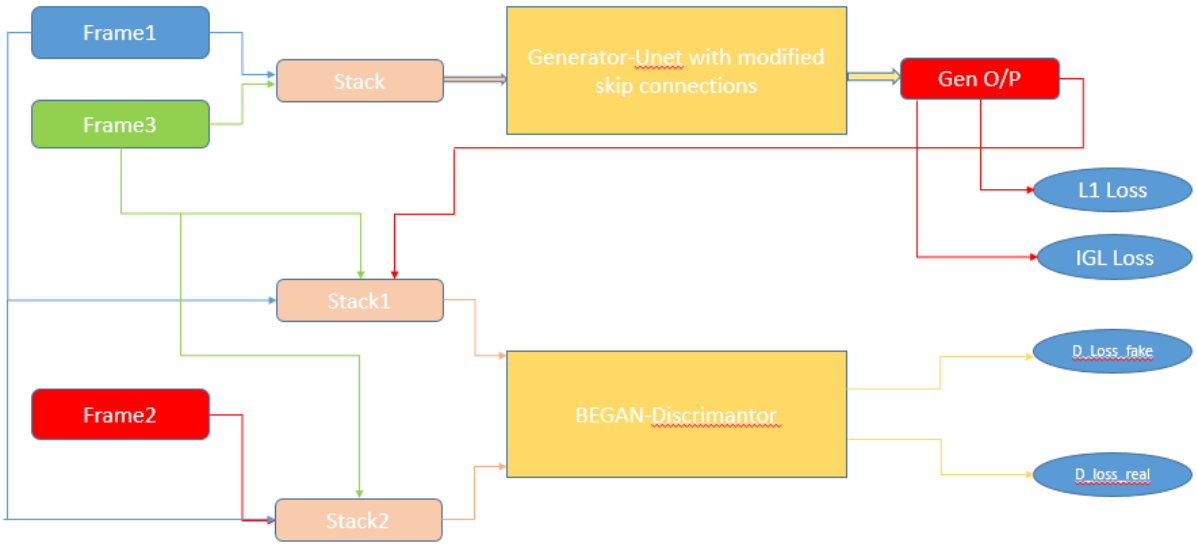

Fig: GAN Architecture for Frame Interpolation

RESULTS

Fig: Left- OUTPUT FRAME, Right – TARGET FRAME

The left frame is the output frame generated from the network and right frame is the ground truth to be attained. The output is well reconstructed in the parts where there is less movement within the frames stacked together. On rapidly changing frames there is still scope for improvement. Exploring different loss functions and fine tune for optimal weights will enhance the texture further.

References

[1] http://cs231n.stanford.edu/reports/2017/pdfs/317.pdf

[2] https://arxiv.org/pdf/1511.05440.pdf

[3] https://arxiv.org/pdf/1411.1784.pdf

[4] https://arxiv.org/pdf/1611.07004.pdf

[5] https://arxiv.org/pdf/1703.10593.pdf

Project Team-Krishna Sumanth, Badrinath **Professor- K Madhava Krishna**