

Lean Comic GAN (LC-GAN): a Light-Weight GAN Architecture Leveraging Factorized Convolution and Teacher Forcing Distillation Style Loss Aimed to Capture Two Dimensional Animated Filtered Still Shots Using Mobile Phone Camera and Edge Devices

Authors : Kaustav Mukherjee

Abstract : In this paper we propose a Neural Style Transfer solution whereby we have created a Lightweight Separable Convolution Kernel Based GAN Architecture (SC-GAN) which will very useful for designing filter for Mobile Phone Cameras and also Edge Devices which will convert any image to its 2D ANIMATED COMIC STYLE Movies like HEMAN, SUPERMAN, JUNGLE-BOOK. This will help the 2D animation artist by relieving to create new characters from real life person's images without having to go for endless hours of manual labour drawing each and every pose of a cartoon. It can even be used to create scenes from real life images. This will reduce a huge amount of turn around time to make 2D animated movies and decrease cost in terms of manpower and time. In addition to that being extreme light-weight it can be used as camera filters capable of taking Comic Style Shots using mobile phone camera or edge device cameras like Raspberry Pi 4, NVIDIA Jetson NANO etc. Existing Methods like CartoonGAN with the model size close to 170 MB is too heavy weight for mobile phones and edge devices due to their scarcity in resources. Compared to the current state of the art our proposed method which has a total model size of 31 MB which clearly makes it ideal and ultra-efficient for designing of camera filters on low resource devices like mobile phones, tablets and edge devices running OS or RTOS. Owing to use of high resolution input and usage of bigger convolution kernel size it produces richer resolution Comic-Style Pictures implementation with 6 times lesser number of parameters and with just 25 extra epoch trained on a dataset of less than 1000 which breaks the myth that all GAN need mammoth amount of data. Our network reduces the density of the Gan architecture by using Depthwise Separable Convolution which does the convolution operation on each of the RGB channels separately then we use a Point-Wise Convolution to bring back the network into required channel number using 1 by 1 kernel. This reduces the number of parameters substantially and makes it extreme light-weight and suitable for mobile phones and edge devices. The architecture mentioned in the present paper make use of Parameterised Batch Normalization Goodfellow etc al. (Deep Learning OPTIMIZATION FOR TRAINING DEEP MODELS page 320) which makes the network to use the advantage of Batch Norm for easier training while maintaining the non-linear feature capture by inducing the learnable parameters

Keywords : comic stylisation from camera image using GAN, creating 2D animated movie style custom stickers from images, depth-wise separable convolutional neural network for light-weight GAN architecture for EDGE devices, GAN architecture for 2D animated cartoonizing neural style, neural style transfer for edge, model distillation, perceptual loss

Conference Title : ICCV 2021 : International Conference on Computer Vision

Conference Location : Istanbul, Türkiye

Conference Dates : January 28-29, 2021