Conditional GANs for Single Image Dehazing

Archit Sharma and Kanishk Gandhi
Electrical Engineering, IIT Kanpur

Problem Statement

Haze is a common atmospheric phenomenon caused by dust, smoke and other dry particles. Haze causes a visual loss in contrast due to scattering of light by haze particles. Single image dehazing aims to restore the image without scatter.

Introduction

Single image dehazing is a complex and ill-posed task. Hazing is attributed to dust, fog and other environmental factors which severely degrades the images. Image hazing is a function of depth, where the visual contrast reduces rapidly as the depth of objects in the images increases. From a perspective of vision, this severely impacts feature retrieval based tasks. Degraded photos often lack visual appeal and offer poor visibility of scene contents. Thus, dehazing images forms an important task in consumer photography and a crucial preprocessing step in vision tasks.

Initial classical dehazing algorithms required additional annotated data such as depth maps, 3D geographic models of the scene and so on. Other approaches used multiple images of the same scene, to capture the differences between images, in retrieval of haze free scene. Classical work in single image dehazing assumes a prior for hazed images. Dark channel prior formed the state-of-the-art in this domain. Other works include maximizing the contrast in haze free images and image-category specific methods.

Image dehazing has only recently been explored in the purview of deep learning. Various deep architectures have been proposed for the purpose of single image dehazing, with the aim of estimating haze depth as human brains do. The general architectural approach involves CNNs to estimate medium transmission maps and learn features from hazy features. Another work proposes multi-scale CNNs to dehaze images.

GANs have been shown to be extremely effective in synthesizing images. To the best of our knowledge, this is the first work which uses GANs for dehazing.

Approach

We use the D-Hazy dataset for our problem [1]. The dataset uses the NYU Depth dataset to artificially haze 1400+ images. The haze is modelled using $I(x) = J(x)t(x) + \alpha(1 - t(x))$, where I and J respectively are the hazed and real scenes, t is the medium transmission and α is the global atmospheric light.

Generative adversarial networks (GANs) is an architecture in which two competing frameworks learn simultaneously. The generators learns to synthesize images, conditioned on the hazy images and the discriminator learns to discriminate fake images from the real ones. The discriminator network can be thought to learn an appropriate loss function for the generator network to optimize. The learning process runs iteratively, where both are optimized separately in batches.

Although GANs synthesize realistic images, for tasks such as dehazing, GANs often fail to construct images fail to learn desired mappings to ground truth because of lack of penalty on the distance from ground truth. To this effect, a L1 penalty is introduced, which forces the generator to produce images which are not only realistic, but close to the ground truth.

Despite the introduction of penalty, some artefacts are still observed in images. Image objects at large depths are particularly difficult to restore. As shown in [2], we try stacking two generators to ease the task of generation. However, training such a stacked GAN is not an easy task. So, we simplify the task by training the generators separately. We also, experiment with L1 loss-Adversarial loss ratio in the two stacked layers.

For the general GAN implementation, we follow the pix2pix implementation [3] which has been shown to be succesful on many tasks.

Results

Figure 1: Input - Adversarial Loss - Adversarial + L1 loss - Stacked Approach - Ground Truth

Observations

In general, GANs perform fairly well in image restoration. Training with a pure adversarial loss, produces realistic images, but, generates random sections when restoring sections of images at depth. Adding L1 loss, in general reduces this problem. Restoration of objects seems to be a tricky issue, which was not helped by stacking two generators on top of each other.

Evaluation Criteria

We plan to evaluate results of dehazing using human evaluation. After training on a few images, we will ask the subjects to judge images as fake or real. Images generated via different GANs will be mixed with the ground truth images, and subjects will be asked to judge a subset of those images.

Further Work

The results contain substantial artifacts causing them to appear fake compared to the ground truth images. It is proposed that these can be improved upon by stacking another GAN upon the existing network and training the model end to end. The first stage would generate a preliminary image with a large number of artifacts while the second stage, conditioned upon the original input and the output of the first generator would act as a refining network.

References

[1] Christophe De Vleeschouwer Cosmin Ancuti, Codruta O. Ancuti. D-hazy: A dataset to evaluate quantitatively dehazing algorithms. In *IEEE International Conference on Image Processing (ICIP)*, ICIP'16, 2016.

[2] Han Zhang, Tao Xu, Hongsheng Li, Shaoting Zhang, Xiaolei Huang, Xiaogang Wang, and Dimitris Metaxas. Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks. *arXiv:1612.03242*, 2016.

[3] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and A. Efros. Image-to-image translation with conditional adversarial networks. *arxiv*, 2016.