Investigation into the use of Convolutional Neural Networks for the purpose of image exposure correction

u1921415

12th June 2022

Contents

1 Abstract

This report examines the implementation of a Convolutional Neural Network for image exposure correction. A Generative Adversarial Network (GAN) style network architecture is used and it's loss functions analysed.

2 Introduction

Image exposure correction is a typical activity for photographers to adjust the light levels present in their photos. The exposure is set by the quantity of light let in during image capture. Adjustments are often made using image processing software such as adobe lightroom after the images are taken by the photographer. The image processing is a manual process that the photographer must perform for every image that they capture. This report explores whether a Generative Adversarial Network (GAN) could be used in order to generate well exposed images from a given under or over exposed image.

3 Network Architecture

3.1 Overview

To achieve image translation from an image of poor exposure to one of more professionally exposed quality, a Generative Adversarial Network archtitecture was used. A GAN is a type of Convolutional Neural Network architecture with a unique construction. One of the ways in which this implementation varies from others is that the network recieves an initial image and a target image. The aim of the network is to correctly hone its weights to consistently generate images or perform image transforms based on training cycles of those input/target paired images.

The 'generative' part of the GAN's structure is a generator model, whose role is to produce an output image from a given input image. On the other side of the GAN structure is the discriminator model which is trained to determine whether an image is real or fake. Here real can be considered to be straight from the dataset of target images, and fake being anything that has too high a loss value post evaluation. This architecture allows the discriminator to critique the generated images. The criticism is used to influence the generator model which reacts by updating it's weights to attempt to create more lifelike images after every epoch.

3.2 Technical Implementation

To create new images from an input image, the generator takes advantage of a U-net style, encoder-decoder design. U-net is an architecture that effectivly downsamples and then upsamples an image. It is a fully connected, convoultional neural net that takes advantage of stride convolution and max-pooling. It also includes skip connections that make localized image information glabally available in the symmetric model, increasing the predictive ability of the generator (Gayathri and Aarthy, 2021). The downscaling or encoding section attempts to gather context from the input whilst the decoder works to identify areas of interest during its upscaling (Ali et al., 2022).

In this implementation, an image of size 256x256 is shrunk down to a 1x1 feature map before being blown up again back into original size. This is performed by six layers down and six up which convolve the image. Stride is set at 2 in order to downscale and upscale the image. During downscaling, a leaky relu function is used

For the discriminator, a PatchGAN is used. A PatchGAN is a comparably small array of values that represent whether a corresponding area or patch of an image is real or fake (Demir and Ünal, 2018). This relationship between patch and image can be visalized better in figure 1. By utilising this technique, arbitrarily large images can be processed. Other benefits such as fewer parameters and faster runtime can be realised using this method (Isola et al., 2016).
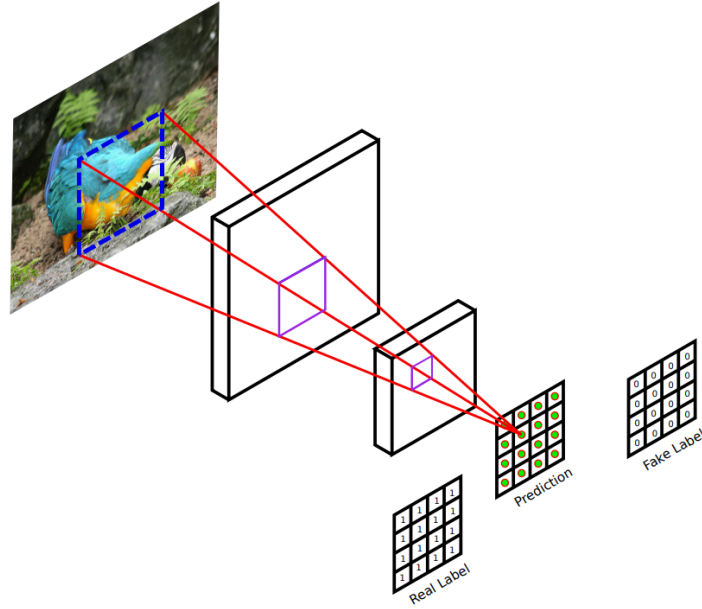
Figure 1: Demonstration of PatchGAN generation (Demir and Ünal, 2018)

4 Input Manipulations

To increase the variety of training data, augmentations were applied to the images. During training, a horizontal flip was applied at a probability of 0.5. Another transformation was the addition of colour jitter at a probability of 0.2 to increase the range of colours the model experiences. To aid in training, a normalisation transform was applied giving every training image a level of blur. The ground truth images however only recieved the same normalised blur but neither the colours nor the orientation augmentations since they are the target for the model to work towards.

5 Loss function/s used

6 Training and testing process followed

7 Model evaluation metric/s used

8 Analysis and conclusions

9 Description of possible alternative approaches

References

Ali, O., Ali, H., Shah, S.A.A., Shahzad, A., 2022. Implementation of a modified u-net for medical image segmentation on edge devices. IEEE Transactions on Circuits and Systems II: Express Briefs 1–1. https://doi.org/10.1109/tcsii.2022.3181132

Demir, U., Ünal, G.B., 2018. Patch-based image inpainting with generative adversarial networks. CoRR abs/1803.07422.

Gayathri, S., Aarthy, D.K., 2021. Computer-aided detection of malignant mass in mammogram using u-net architecture, in: Advances in Intelligent Systems and Computing. Springer Singapore, pp. 179–185. https://doi.org/10.1007/978-981-16-2674-6_14

Isola, P., Zhu, J., Zhou, T., Efros, A.A., 2016. Image-to-image translation with conditional adversarial networks. CoRR abs/1611.07004.