Mini-project 1: Using the standard PyTorch framework

Stephen Monnet, Thibaut Stoltz, Arina Lozhkina
EPFL
Lausanne, Switzerland
stephen.monnet@epfl.ch, thibaut.stoltz@epfl.ch, arina.lozhkina@epfl.ch

May 27, 2022

Abstract

The field of noise reduction in images is becoming increasingly relevant [1]. Its application is especially important in the areas of processing medical [2], SAR images [3] etc. The aim of our work is to apply deep neural networks to solve the problem of noise to noise translation in low-resolution images. As methods, a modification of the Noise2Noise [4] architecture is used. Others methods are implemented for comparison. In addition, experiments were carried out for training the model, setting its parameters and configurations. Peak signal-to-noise ratio (PSNR) and mean squared error (MSE) were used as metrics for assessing the quality of the solution to the problem. As a result, the PSNR of 25.56 dB is achieved.

1 Introduction

Few methods are currently known for converting an image characterized by one noise into an image with another noise. The main one is the article proposing the Noise2Noise[4] model, as well as its modification [5]. However, the application of neural networks for image reconstruction from a noisy image has started to develop alone, starting with the use of the MLP [8] and continuing with the development of convolutional neural networks [6, 7]. The idea of deep neural networks led to a change in optimization methods for learning [9, 10]. Further works applied more sophisticated methods: DnCNN with batch normalization [11], residual encoding-decoding [12]. All these models are mostly based on the use of neural networks, which we will try to implement in our work.

2 Methodology

The proposed solution is based on the implementation of a neural network architecture based on the PyTorch. The bases for the architectures were taken: Noise2Noise [4], DnCNN [11], Convolutional Neural Network, Linear network.

When training the network, the following instruments were used: Adam [13] as optimizer, ExponentialLR as scheduler, Mean Squared Error as criterion to optimize.

2.1 Dataset

The training sample consists of 50,000 pairs of $3 \times H \times W$ images - the original noisy image and the target image. The validation sample is similarly composed and contains 1000 pairs of images.

To expand the data sample and increase its representativeness, the data were augmented by flip along the x-axis, y-axis with probabilities of 0.3 each. Resizing, rotating with image enlargement, cropping, and color palette transformations were not applied due to the fact that pixel intensity changes could lead to added or altered noise.

2.2 Models description

For comparison, several models have been implemented to translate noise from the image, as described in the Introduction. These models are:

- Linear model; In applying this model, each image was flattened into a vector, causing the structure of the image and the relationship between the values of neighboring pixels to be lost. The model consists of 4 linear layers with a ReLU activation function.
- Simple convolution model; This model contains 4 convolution layers, 4 Transposed convolution layers with padding "same" and Dropout. For all convolution layers the kernel size is (3, 3).
- DnCNN; This is the deepest network we used because of training time constraints. It consists of a convolutional layer with ReLU activation and a sequence of 15 blocks of the form: Convolutional layer -> BatchNorm -> ReLU; and a final convolutional layer.
- Noise2Noise. This model was created to transform one noisy image into another. The idea of its implementation is to analyze the signal transformation based on deep learning. The architecture is an encoder and decoder consisting of convolution layers, MaxPooling, Upsampling and concatinating with the results of previous layers.

2.3 Experiments

When training the model, we tried to apply different parameters:

- Number of training epochs: 3, 7, 20, 50.
- Size of batch: 16, 32, 64, 512.
- Learning rate: 5e-3, 1e-3, 1e-4.
- Betas: (0.9, 0.99) for Adam.

We also tried stochastic gradient descent. In addition, we tried learning with and without Scheduler with a gamma parameter of 0.9.

After the experiments, due to the low dimensionality of the input images, the parameters turned out to be optimal for this problem: Number of epochs of training: 7, Size of batch: 16, Learning rate: 1e-4, Optimizer Adam with parameters beta = (0.9, 0.99), Exponential Scheduler with gamma parameter 0.9.

3 Results

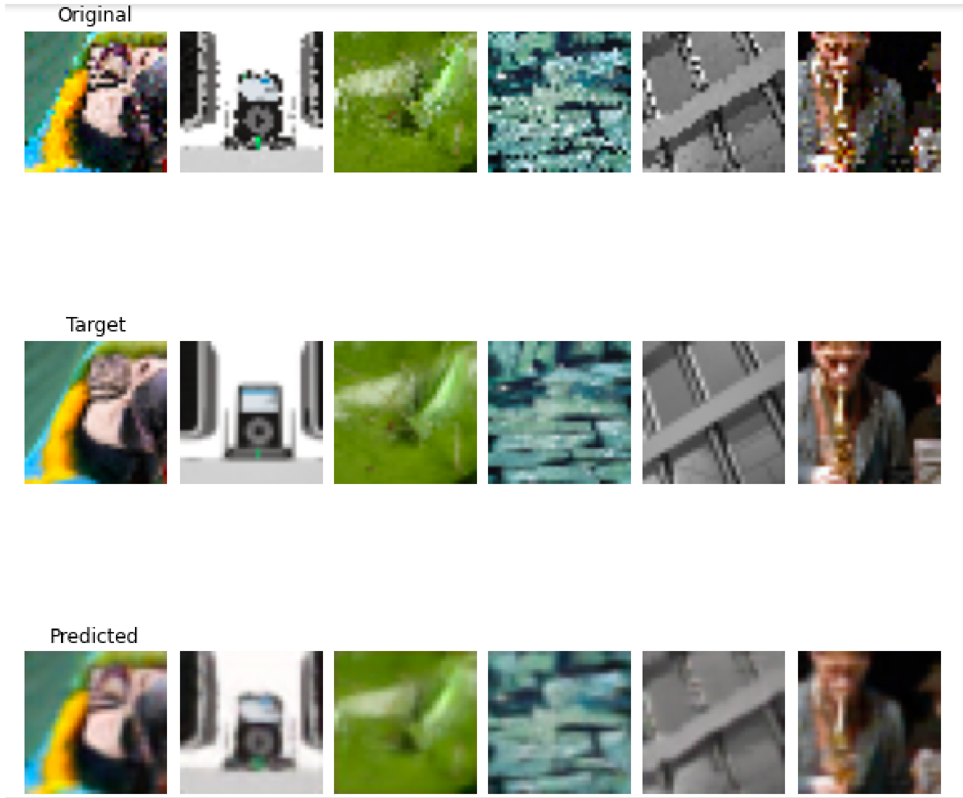
As a result, the Noise2Noise model with PSNR on the validation sample equal to 25.56 showed the best results.

The following results were obtained when comparing the training time for training with equal parameters for 7 epochs using Google Compute Engine ((GPU) server gas pedal) and comparing PSNR values for the validation sample :

Results for different models			
Model	PSNR	Mean Loss	Minutes to train
Linear	22.97	1208	01:40
CNN	25.01	1010	01:48
DnCNN	24.50	1542	06:07
Noise2Noise	25.56	937	03:51

Finally, the results of the validation images for the best model are presented in the Figure 1. The first line shows the images that come to the input of the neural network, the next line the target images, the last line the images obtained by applying the algorithm. You can see that the results are more blurred, this is due to the application of convolution on low resolution images.

Figure 1: Results of Noise2Noise



4 Discussion

Despite attempts to consider a variety of approaches, apply parameters over a wide range, and explore possible signal processing and noise conversion materials, the results obtained in the paper for different architectures do not differ significantly.

Due to the high risk of overfitting, models with few hidden layers were chosen. Nevertheless, only the first few epochs turned out to be the most effective in training all models. Such learning peculiarities were explained for us by the fact that the images are characterized by low resolution. However, even in our best solution, the blurring by convolutional layers leads to a mismatch of some pixels and, as a result, to a low value of accuracy metrics. At the same time, in the absence of convolutional layers (in the linear model) there is pepper and salt noise. Perhaps wavelets or GANs should be used to get the best one.

5 References

- Fan L., Zhang F., Fan H. and Zhang C., Brief review of image denoising techniques, Visual Computing for Industry, Biomedicine, and Art, 2019, 2:7, https://doi.org/10.1186/s42492-019-0016-7

- Sameera V. Mohd Sagheer, Sudhish N. George, A review on medical image denoising algorithms, Biomedical Signal Processing and Control, Volume 61, 2020, 102036, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2020.102036.

- Balnarsaiah B., G Rajitha, Denoising and Optical and SAR Image Classifications Based on Feature Extraction and Sparse Representation, Image and Video Processing (eess.IV); Computer Vision and Pattern Recognition (cs.CV), arXiv:2106.01896 [eess.IV], https://doi.org/10.48550/arXiv.2106.01896

- Lehtinen J., Munkberg J., Hasselgren J., Laine S., Karras T., Aittala M., Aila T., Noise2Noise: Learning Image Restoration without Clean Data, Computer Vision and Pattern Recognition (cs.CV); Machine Learning (cs.LG); Machine Learning (stat.ML), arXiv:1803.04189

- Calvarons A.F., Improved Noise2Noise Denoising with Limited Data, 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2021, pp. 796-805, doi: 10.1109/CVPRW53098.2021.00089.
- V. Jain and S. Seung. Natural image denoising with convolutional networks. In Advances in Neural Information Processing Systems, pages 769–776, 2009
- J. Xie, L. Xu, and E. Chen. Image denoising and inpainting with deep neural networks. In NIPS, 2012.
- C. J. Schuler, H. Christopher Burger, S. Harmeling, and B. Scholkopf. A machine learning approach for non-blind image deconvolution. In CVPR, 2013
- U. Schmidt and S. Roth. Shrinkage fields for effective image restoration. In IEEE Conference on Computer Vision and Pattern Recognition, 2014
- Y. Chen and T. Pock. Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. IEEE transactions on pattern analysis and machine intelligence, 2017
- K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. IEEE Transactions on Image Processing, 2017.
- X. Mao, C. Shen, and Y.-B. Yang. Image restoration using very deep convolutional encoderdecoder networks with symmetric skip connections. In Advances in neural information processing systems, pages 2802–2810, 2016
- Kingma D.P., Ba J., Adam: A Method for Stochastic Optimization, Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015, Machine Learning (cs.LG), arXiv:1412.6980