DL4CV - Final Report

Dan Bar
Weizmann Institute
dan.bar@weizmann.ac.il

Neta Shaul
Weizmann Institute
neta.shaul@weizmann.ac.il

Yeari Vigder
Weizmann Institute
yeari.vigder@weizmann.ac.il

1 Abstract

Supervised deep learning models that claim to solve the 'Super Resolution' (SR) problem preform poorly on images that where not taken from a specific data distribution. That motivated an image-specific network models. The models learn the internal patch distribution of the image and then use it to generate the SR solution. It was shown that such models give state of the art (SOTA) results, see ZSSR[1],KernelGAN[2] and SinGAN[3]. We offer a model based on the previous works, that use adversarial training both to find the downscaling kernel and to find the SR solution. We put emphasis on training time saving that can be costly when training a GAN.

2 Problem statement/Motivation

The project considers the task of 'Super Resolution' (SR). It focused on Blind-SR, that is SR task when the downscaling kernel (SR-kernel) is unknown. Our model is based on a work done in ZSSR[1] and kernelGAN[2], where a GAN was used to learn the SR-kernel. SR is a problem with no unique solution, that motivates us to introduce a GAN also for generating the SR solution. Such an approach was proved useful in SinGAN[3], but required long running times. We aim to shorten the running times while improving previous results.

3 Related work

First to offer an image-specific CNN to solve SR problem is ZSSR[1]. It has shown that SR with deep learning can be done only by learning the image internal statistics. Later work such as kernelGAN[2], used an image-specific GAN to learn the SR-kernel and showed that training ZSSR with the learned kernel improve result for Blind-SR. Additional work in SinGAN[3], used a more complex architecture of image-specific GANs to generate the actual SR solution.

4 Method

Our approach uses image-specific GAN to learn the SR-kernel. It then utilize its discriminator combined with a reconstruction loss (L1 loss) to train an image-specific CNN that generates the SR solution. We offer a multiple options for training such network.

The first is without adversarial training for the generating network. That is, the discriminator is trained only on the Kernel network, it is then stays fixed and used as an additional loss for the generating network. See Figure 1 for visualization.

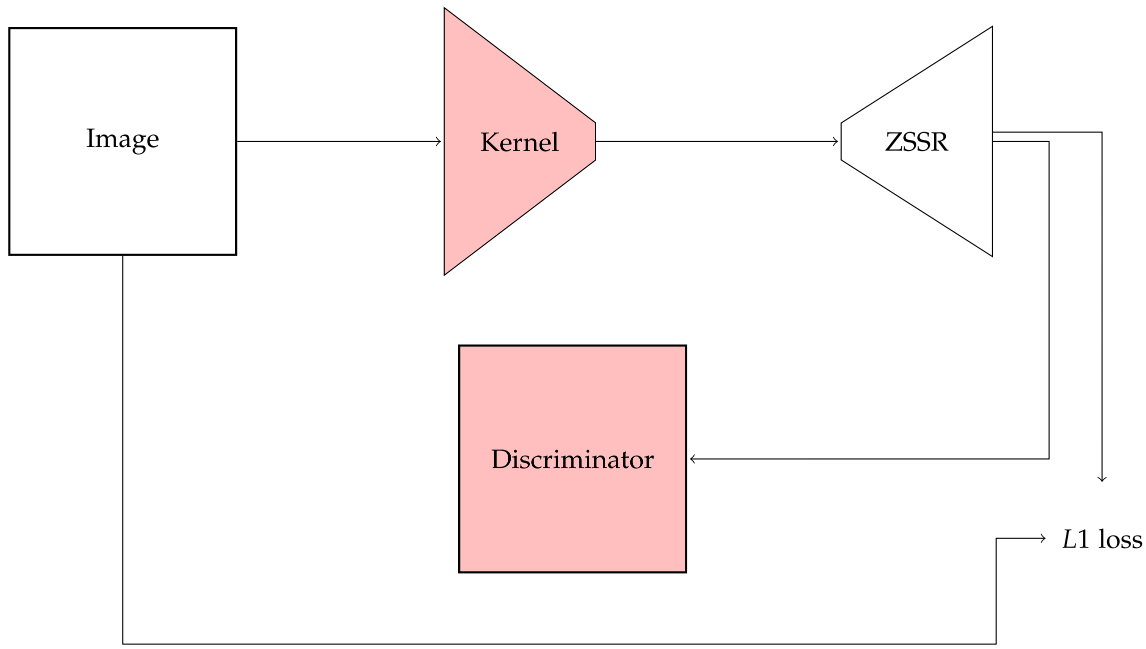
The Second includes adversarial training for the generating network, but in a serial order. That is, the discriminator is first trained on the Kernel network and then goes through additional adversarial training with the generating network.

The third also includes adversarial training for both networks, but in parallel. It trains the two generators and the discriminator simultaneously. At each epoch, all network components are trained but one at a time. That is, when one of the three component is trained the other two are fixed, so no end-to-end training. This ensures that the first generator still learns the SR-Kernel. See Figure 2 for visualization.

5 Data

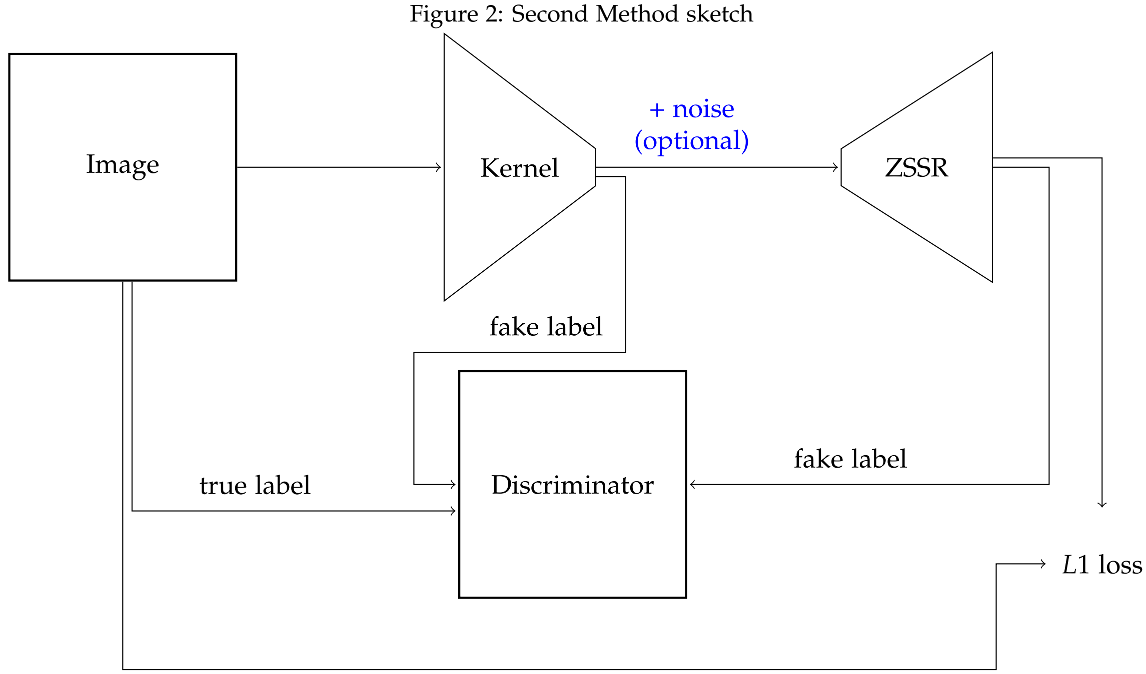
Considering the nature of this problem, we do not need a specific data collection. But to be consistent with previous work, we tested our method on the data that was used by [1, 2]. That will give us a good

Figure 1: Basic Method sketch



While colored nodes are pre-trained and fixed in the second training.

Figure 2: Second Method sketch



measure of the results. We also included additional data. The data can be found here.

6 Experiments and Analysis

As described at the method section, our experiment included three architectures: fix - the first type, serial - the second type and semi E2E - the third type. Our main focus was on the first and the second architectures, the fixed and the serial, that manage to produce superior results over the semi E2E architecture. In fig. 3 and fig. 4 we can see the comparison between our method with the serial architecture and the results of the two existing methods, ZSSR and kernelGAN. The two comparisons were tested on two photos with unknown downscaling kernel. Visually, we can see that our network emphasize the outline of the objects in the images. In fig. 5, the experiment subject was an image with known kernel, the Bicubic kernel. We were able to produce results with slightly better score on the PSNR/SSIM benchmarks compared to the kernelGAN method on average, and we lost to the ZSSR method - see table 1. This result is understandable, ZSSR uses the Bicubic kernel which is the true kernel of the image, while kernelGAN and our method try to approximate it. In addition to the visual results, we managed to make our network learn the two loss functions, the L1 and the GAN loss, see fig. 6. More results can be found in our webpage.

Figure 3: cars comparison

Figure 4: Lincoln comparison

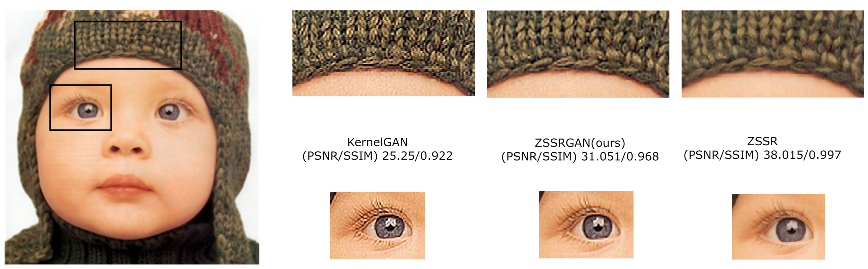


Figure 5: baby comparison

7 Conclusions

We were able to add the discriminator trained by KernelGAN to the objective function of ZSSR training and optimize it while still optimizing the reconstruction loss. In addition, we implemented GAN training in the architecture that improved our result. We saw that there is Sharpness-Smoothness trade-off, and our model seems to give sharper results but looses on smoothness.

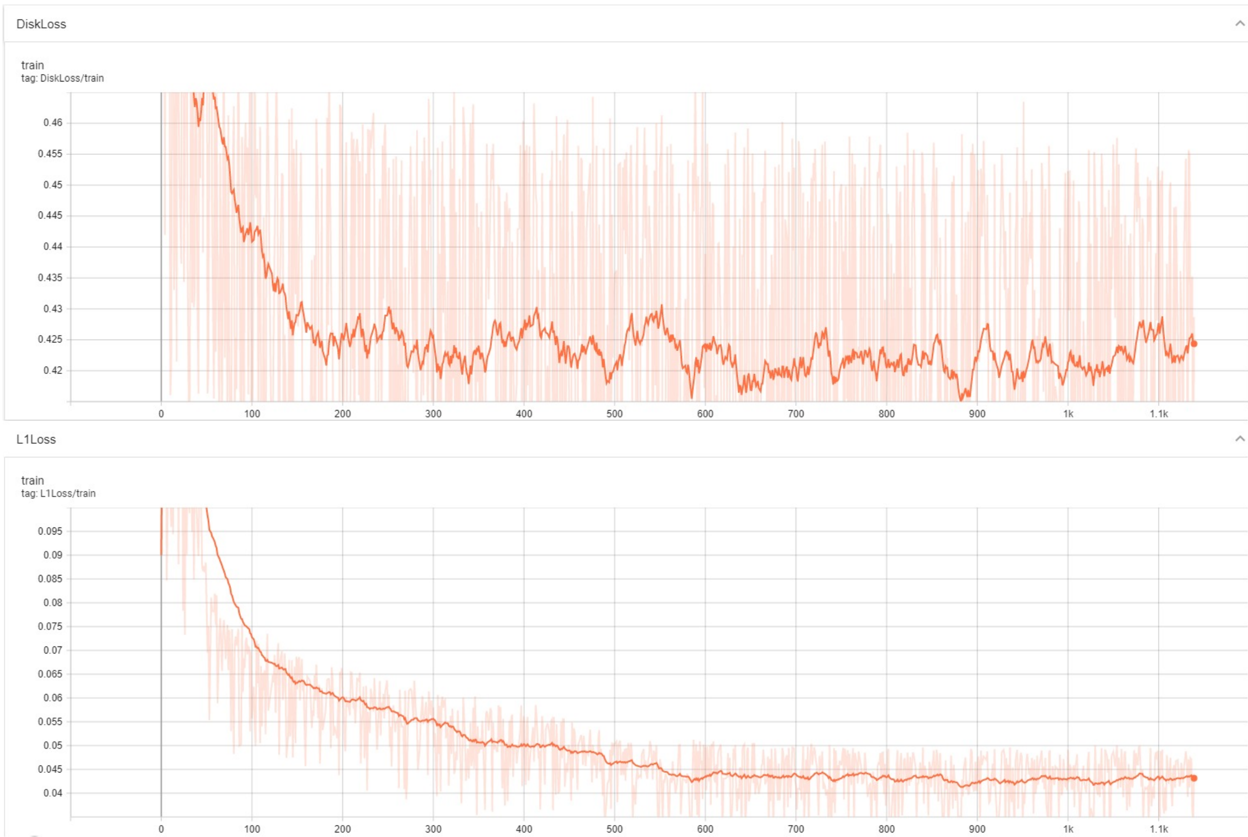


Figure 6: Both losses trained on Lincoln image in SERIAL training process

Model	PSNR/SSIM
ZSSR	**33.554/0.992**
ZSSRGAN	23.367/0.914
KernelGAN	22.592/0.906

Table 1: PSNR/SSIM score of the our method, ZSSR and kerGAN on average

References

[1] Michal Irani Assaf Shocher, Nadav Cohen. ”zero-shot” super-resolution using deep internal learning. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[2] Sefi Bell-Kligler, Assaf Shocher, and Michal Irani. Blind super-resolution kernel estimation using an internal-gan. In *NeurIPS*, 2019.

[3] Tamar Rott Shaham, Tali Dekel, and Tomer Michaeli. Singan: Learning a generative model from a single natural image. *CoRR*, abs/1905.01164, 2019.