CNN for spatial alias removal.

Abe Vos (11779810) **Nikita Bortych** (12342076)[1]

Abstract

1. Introduction

- Talk about process for which superresolution is necessary.
- Talk about limitations of naive approach (bilinear interpolation causes artifacts in f-k domain) - Talk about how deep learning SR can solve this problem.

2. Related works

Since this project is multidisciplinary, the related works section is split in to two parts. The first part will discuss the work related to the general problem of super resolution in the field of Deep Learning (DL). The second part will focus on the ways researchers in the geophysics have approached the problem of super resolution.

2.1. Super resolution

The first attempt at the super resolution problem using DL was Super Resolution CNN (SRCNN) model (Dong et al., 2015). This model is relatively shallow, with only 3 convolutional layers with kernel sizes of 9x9 $\rightarrow$ 1x1 $\rightarrow$ 5x5, with the first later having 64 kernels and the second one having 32. The model has beaten all the benchmarks at the time of its creation, although it was pretty shallow and simple. The significance of this model is that it opened research into the problem of single image super resolution for the field of DL.

Theoretical DL, as well as practical successes have shown that there is a benefit to increasing the depth of the network in order to achieve better hierarchical representations. Very Deep Super Resolution (VDSR) model (Kim et al., 2015) applies this principal by creating a VGG-style network of depth 20 with 3x3 kernels. The paper, aside from novelty of depth, has made two contributions: using different scales for low-resolution data generation, which is not relevant to our case and using residual approach: adding the image to the output of the network, so the network learns only the residual function, instead of the full transformation. The concept of residual learning is even further advanced in the Enhanced Deep Residual Networks for Single Image Super-Resolution paper (Lim et al., 2017) which introduced the EDSR model, that is still the state of the art in Super Resolution task (Yang et al., 2019). Due to the successes of

2.2. Geophysics domain

Within the field of geophysics, a lot of research has been done the mitigate the problem of limited resolution data using a diverse set of approaches (Valenciano et al., 2015; Lecomte, 2008; Yu et al., 2006). However, more recently a trend can be seen toward applying deep learning techniques in order to approach these problems from a data-centered perspective. For example, Halpert 2018 uses a GAN to generate high resolution images from a synthetic dataset.

Additionally, researchers in the field have been using deep learning models - other than superresolution models - to tackle a multitude of geophysical problems like modelling the Hyperbolic Radon Transform (Kaur et al., 2019) or applying domain transfer to emulate wave inversion methods (Mosser et al., 2018). The main goal of these applications is to reduce computational cost and alleviate the strict assumptions that are implied by the more 'classical' approaches.

3. Method

3.1. Baselines

The models discussed in this section are compared to results obtained by plain bilinear interpolation and the SRCNN architecture. As shown in 1, bilinear interpolation - although fast to compute - causes unwanted artifacts to appear when the image is transformed to the f-k domain. In order to compare the performance of different model architectures, we also use a SRCNN model (Dong et al., 2015) as baseline. SRCNN was chosen because it is one of the seminal models to be used for image superresolution. Furthermore, it has a simple implementation and trains relatively quickly.

[1]Supervisor: Dr. A. Garg.
Abe Vos <abe_vos@msn.com>
Nikita Bortych <nikita.bortych@gmail.com>.

Project AI 2019, MSc Artificial Intelligence, Univ. of Amsterdam

3.2. Statistical testing

3.3. Model architectures

3.4. Different loss terms

Aside from trying out different model architectures from literature, we also experimented with a set of different loss functions. Below is a short overview of the loss terms that we experimented with:

- L1 and L2 loss: these two loss functions are natural choices since they are computationally cheap and allow for stable training. However, since they compute the average of a pixel-wise loss, these two terms can fail to capture qualitative visual features in the generated image, causing the results to appear blurred (Isola et al., 2017).

- GAN discriminator loss: Since GANs (Goodfellow et al., 2014) learn a custom loss function at the level of visual features instead of individual pixels, the generated image does not suffer the same problems as L1 and L2 loss (Ledig et al., 2017). However, due to their adversarial nature, GANs are notoriously difficult to train, harbouring a number of instabilities that require a number of heuristics to solve (Arjovsky & Bottou, 2017) and (Salimans et al., 2016). The discriminator architecture used across experiments is based on the DCGAN architecture (Radford et al., 2015) with a large fully connected layer at the end to account for the larger image size (the original paper used 64x64 images).

- Transformation to f-k domain: the overall objective of this research is to generate high resolution images that - when transformed to the f-k domain - displays the same features as the ground-truth high resolution image. For this reason introduce a specialized loss function for this particular problem: After applying the model, both the super- and high resolution images are transformed to the f-k domain. From these transformed image, the reconstruction loss L_{f-k} is computed (using any of the methods described above). This way, when optimizing the model is forced to reduce visual artifacts like the ones displayed by the bilinearly interpolated image.

In section 4 we try out different combinations of these building blocks in order to find a good model.

4. Experiments/model search

5. Results

5.1. Influence of data on results

Since a real seismic data are often proprietary or difficult to come by, it is useful to be able to train the superresolution model well without having to acquire vast amounts of samples. In order to quantify the influence the amount of data has on the model's performance, we trained the model from scratch for a number of times. Each time the model was trained, a random split was made in the data such that a fraction t was assigned to be training data and the remainder was used for validation. Values of t where uniformly selected in the range $(0.1, 0.9)$ to ensure that there always where some data available to train and/or validate the model.

Figure 1 shows the PSNR and SSIM values for the validation set at the end of training for each split. It is clearly visible that more data have a positive effect on model performance. However, even when the model is trained on just 10% of the data (which amounts to 40 samples) the model still performs reasonably well at a PSNR of around 38. Surprisingly, although the descend of the structural similarity seems to be rapid as t approaches 0, the change in values is actually very small.

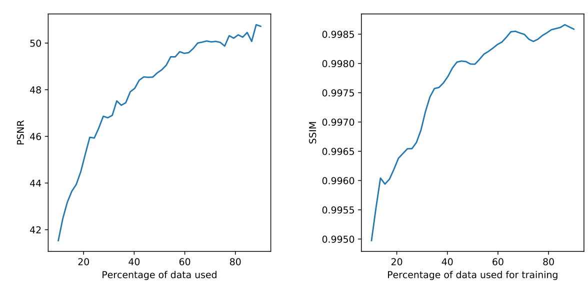

Figure 1. The validation PSNR and SSIM values of models trained on different amounts of data.

6. Issues

is it challenges related to the project? maybe that we have no domain knowledge ? Probably challenges encountered with our method/model, like training difficulties for the GAN. (ethical, legal, software/code quality, security),

7. Conclusion

Figure 2. Detailed caption of the figure goes here.

7.1. Citations and References

7.2. Software, Data, Demo and Dissemination

Our git?

References

Arjovsky, M. and Bottou, L. Towards principled methods for training generative adversarial networks. arxiv. 2017.

Dong, C., Loy, C. C., He, K., and Tang, X. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2):295–307, 2015.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Generative adversarial nets. In *Advances in neural information processing systems*, pp. 2672–2680, 2014.

Halpert, A. D. Deep learning-enabled seismic image enhancement. In *SEG Technical Program Expanded Abstracts 2018*, pp. 2081–2085. Society of Exploration Geophysicists, 2018.

Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125–1134, 2017.

Kaur, H., Pham, N., and Fomel, S. A fast hyperbolic radon transform algorithm using deep neural network. *EAGE COnference Exhibition*, 2019.

Kim, J., Lee, J. K., and Lee, K. M. Accurate image super-resolution using very deep convolutional networks. *CoRR*, abs/1511.04587, 2015. URL http://arxiv.org/abs/1511.04587.

Lecomte, I. Resolution and illumination analyses in psdm: A ray-based approach. *The Leading Edge*, 27(5):650–663, 2008.

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4681–4690, 2017.

Lim, B., Son, S., Kim, H., Nah, S., and Lee, K. M. Enhanced deep residual networks for single image super-resolution. *CoRR*, abs/1707.02921, 2017. URL http://arxiv.org/abs/1707.02921.

Mosser, L., Kimman, W., Dramsch, J., Purves, S., De la Fuente Briceño, A., and Ganssle, G. Rapid seismic domain transfer: Seismic velocity inversion and modeling using deep generative neural networks. In *80th EAGE Conference and Exhibition 2018*, 2018.

Radford, A., Metz, L., and Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., and Chen, X. Improved techniques for training gans. In *Advances in neural information processing systems*, pp. 2234–2242, 2016.

Valenciano, A., Lu, S., Chemingui, N., and Yang, J. High resolution imaging by wave equation reflectivity inversion: 77th annual international conference and exhibition, eage. *Extended Abstracts, We N*, 103:15, 2015.

Yang, W., Zhang, X., Tian, Y., Wang, W., and Xue, J. Deep learning for single image super-resolution: A brief review. *CoRR*, abs/1808.03344, 2019. URL http://arxiv.org/abs/1808.03344.

Yu, J., Hu, J., Schuster, G. T., and Estill, R. Prestack migration deconvolution. *Geophysics*, 71(2):S53–S62, 2006.

A. Appendix