Critical Analysis

Deep learning CNN for super-resolution (Dong et al PAMI'15)

Main Trends[1]:

- GAN-based approaches generally deliver visually pleasing outputs while the reconstruction error-based methods more accurately preserve spatial details in an image,
- for the case of high magnification rates (8x or above), the existing models generally deliver sub-optimal results,
- the top-performing methods generally have a higher computational complexity and are deeper than their counterparts,
- residual learning has been major contributing factor for performance improvement due to its signal decomposition that makes the learning task easier.

Overall, we note that the super-resolution performance has been greatly enhanced in recent years with a corresponding increase in the network complexity.

Key Ideas since the publication:

- [2]VDSR: Very Deep Super Resolution. Has 20 weight layers which is much deeper compared with SRCNN which only got 3 layers. The network is able to improve the given image for PSNR, SSIM and MTF values. This proved that the network is able to increase the spatial resolution of given images, which is useful for when images have been distorted by effects such as diffraction. when analysing a high frequency component of the resolution chart (black and white lines), the network was able to increase the visual contrast of the image.
- [3]DnSRGAN : Denoising Super-resolution Generative Adversarial Network. Method for high-quality super-resolution reconstruction of noisy cardiac magnetic resonance (CMR) images. high-quality reconstruction of noisy cardiac images.
- [4]IrCNN: Inception recurrent convolutional neural network for object recognition
- [5] FSRCNN: Fast-Super-Resolution Convolutional Neural Network. A more efficient network structure to achieve high running speed without the loss of restoration quality, thus making it possible to realize real-time video SR.

Main Problems:

- The choice of the network scale trade-off between performance and speed. Larger filter size and network width significantly improve the performance. If a fast restoration speed is desired, a small network width is preferred. The deployment speed will also decrease with a larger filter size.
- Slow convergence. Given enough training time, the deeper networks will finally catch up and converge to the three-layer ones. Author specified that the CNN network contains no pooling layer or full-connected layer; thus, it is sensitive to the initialization parameters and learning rate. When gone deeper (e.g., 4 or 5 layers), it's hard to set appropriate learning rates that guarantee convergence. Even it converges, the network may fall into a bad local minimum, and the learned filters are of less diversity even given enough training time.
- Only Single scale is handled by the network.
- The per-pixel loss function (MSE) focusses on the pixel wise intensity differences alone. Better reconstruction can be obtained by using the perceptual loss functions.

Remaining Problems:

There remain several open research questions:

- [1]Incorporation of Priors: If information about the sensor, imaged object/scene and acquisition conditions is known, useful priors can be designed to obtain high-resolution images. Recent works focusing on this direction have proposed both deep network and sparse coding-based priors for better super-resolution.
- [1]Objective Functions and Metrics: Mainly uses pixel-level error measures the resulting images do not always provide perceptually sound results. There is no universal perceptual metric that optimally works in all conditions and perfectly quantifies the image quality. Therefore, the development of new objective functions is an open research problem.
- [1]Need for Unified Solutions: Two or more degradations often happen simultaneously in real life situations. An important consideration in such cases is how to jointly recover images with higher resolution, low noise, and enhanced details. Current models developed for SR are generally restricted to only one case and suffer in the presence of other degradations It is a challenge to design unified models that perform well for several low-level vision tasks, simultaneously.
- [1]Unsupervised Image SR: Explore how SR can be performed for cases where corresponding HR images are not available.
- [1]Higher SR rates: Tackle extreme super-resolution which can be useful for cases such as super-resolving faces in crowd scenes. Challenging to preserve accurate local details in the image and high perceptual quality in these super-resolved images.
- [1]Arbitrary SR rates: Not known which upsampling factor is the optimal one for a given input as the downsampling factor is not known for all the images in the dataset.
- [1]Real vs Artificial Degradation: SR networks trained on artificially created degradations do not generalize well to actual LR images in practical scenarios.

Most interesting unsolved problem:

- It will be interesting to explore the implementation and application of SR on live streaming videos. The use of unsupervised method as opposed to the example-based strategy (which either exploit internal similarities of the same image or learn mapping functions from external low- and high-resolution exemplar pairs) might help to explore wider image ranges without restrictions.

References:

[1] S. Anwar, S. Khan, and N. Barnes, "A Deep Journey into Super-resolution: A survey," pp. 1–20, 2020 arXiv:1904.07523v3. [Online]. Available: https://arxiv.org/abs/1904.07523v3

[2]D. Vint, G. Di Caterina, J. J. Soraghan, R. A. Lamb and D. Humphreys, "Evaluation of Performance of VDSR Super Resolution on Real and Synthetic Images," 2019 Sensor Signal Processing for Defence Conference (SSPD), Brighton, United Kingdom, 2019, pp. 1-5, doi: 10.1109/SSPD.2019.8751651.

[3] A Generative Adversarial Network technique for high-quality super-resolution reconstruction of cardiac magnetic resonance images. Ming Zhao, Yang Wei, and Kelvin K.L. Wong

[4] Alom, M.Z., Hasan, M., Yakopcic, C. *et al.* Inception recurrent convolutional neural network for object recognition. *Machine Vision and Applications* **32,** 28 (2021).

[5] Improved Super-Resolution Convolution Neural Network for Large Images by Rong Song and Jason Wang.