Denoising: Supervised, Self-Supervised and Unsupervised (LEAPS Innov WP7)

https://events.hifis.net/event/758/ (https://events.hifis.net/event/758/)

Please use this pad to collect notes, questions or thoughts on the presentations given. Please use positive language and be mindful of your peers.

Notes & Questions

Questions

- Is denoising is an ill-posed problem?
    - AK: yes in the sense that there is no one true solution to this for a given noisy image; simply don't know what the original true image was
    - AK: the less light you have, the more noise you get -> what would you say about the original image then? -> best shot
    - Majid: Problems that are not well-posed in the sense of Hadamard are termed ill-posed. Inverse problems are often ill-posed. For example, the inverse heat equation, deducing a previous distribution of temperature from final data, is not well-posed in that the solution is highly sensitive to changes in the final data.
- What is the difference between denoising, deconvolution, reconstruction & super resolution as they all are inverse problem?
    - PS:
- As most of these methods use SSIM, PSNR, Resolution etc as an evaluation metrics can we compare the results of denoising with a result of SuperResolution?
    - AK: PSNR = quadratic error to clean image (in a logarithm)
    - AK: denoising uses the quadratic error for training -> mission accomplished
    - AK: downside -> compromise at the end: good PSNR doesn't mean a plausible image
    - AK: note, denoising changes the intensity of the image -> be careful if you are interested in this
        - currently in the works to tackle this
- N2V: is output then an image where true values and predicted ones are combined? /

How are predicted pixel values “linked”? (I.e. how is full image created?)

- AK: use standard Unet -> mask input images -> during training, model predicts specific output pixels (the ones masked before)
- AK: get one output predicted per masked pixel

- Naive one: The bayesian inverse problems theory is mainly based on the accuracy of the prior probability. Is there a possibility to benchmark all the reasonable prior probas? We can even imagine to use MonteCarlo methods on top to use the best prior proba?
  - AK: back in 1990/2000 -> model $p(s)$ directly by markov random field, experts than engineered these MRFs for smoothness etc
  - AK: supervised learning approach, never engineer $p(s)$ … inherently learned by model from data
  - AK: VAE based approaches provides distribution over $p(s)$, can sample from this and generate clean images
- Do all of the presented methods assume a zero-centered noise? What to do with non-zero-centered noise?
  - AK: likely all of the denoising methods make that assumption
  - AK: what if it is not zero-centered, device not linear how it operates, i.e. pixels are systematically made too bright
  - Example - overflow effects might distort noise behaviour
- Do you need to train the model in N2V for each single image? For instance in tomography we have thousands projections or thousand reconstructed slices. Do you need to train the model for each single projection or reconstruction slice or assuming that there is a similarity you can train on a few and apply the model to the entire set of images?
  - AK: In general, if you can assume all images come from same distribution, could train either on all or a subset. But if experimental conditions different, train separately. Potential problem is that if training data has different prior, could get structure that’s not there
  - AB: @CT-Issue: I think a “Denoising Model” should be trained with real the raw data. As you have to apply the Beer-Lamberts Law (log(Ix/I0)) before reconstruction, the “nature” of the noise is totally distorted.
- Chat: What of the case where the noise is not uncorrelated, i.e., where bright spots produce more noise in their surroundings? How do the methods you present respond to that?

    - AK - In paper with structured noise, assumed noise was not dependent on signal itself. Looking at this topic now - both not independent, and conditioned on signal.

- Chat: For applications on XRF data, did you apply those methods on the result of fitting or you've tried it on the complete raw spectrum which is the original source/container of noise?
    - AK: go with raw data
    - GK: one spectrum per image in the raw data -> try to denoise the original spectrum (per pixel)
    - AK: depends on type of noise, images are not special -> offline
    - PS: some folks from the hyperspectral community are looking into this problem: https://github.com/Helmholtz-AI-Energy/HyDe (https://github.com/Helmholtz-AI-Energy/HyDe)

Use the section below to collect notes that might be relevant for the audience. Please also note down your questions here.

- Initial problem - fluorescence microscopy,
    - use more light and get clean images but destroy sample OR
    - use less light, get noisy images and retain good sample
- Noise sources
    - Poissonian photon statistics
    - readout noise
    - amplification
    - digitization
- Empirically measure noise behaviour $p(x_i|s_i)$ - sample independent
- Traditional supervised learning with CNN
    - need pairs of clean and noisy images
    - CNN - loss function results in returning expected value of si for given xi
    - Downside - often impractical to gather pairs of clean and noisy data
- Noise2Noise, https://arxiv.org/abs/1803.04189 (https://arxiv.org/abs/1803.04189)
    - input data: pair of same signal, different noise
    - CNN returning expected value of xi (noise)
- Noise2Void
    - Blind spot network: Predict pixel based on values of neighbouring pixels
    - Predict signal (assume noise is independent in pixels, given signal)
    - Does not adress structure in noise; Assumes same chance for noise in every

pixel

- DivNoising and HDN https://arxiv.org/abs/2006.06072 (https://arxiv.org/abs/2006.06072)
    - Variational Autoencoder (VAE)
    - input data: noisy images
    - noise model