**Main idea:**

the performance of Noise2Noise drops when the amount of training data is reduced, limiting its capability in practical scenarios.

using diverse network architectures and loss functions, that the duplicity of information in the noisy pairs can be exploited to reach increased denoising performance of Noise2Noise. Additionally, the issue of overfitting in Noise2Noise is analyzed, given its relevance when training with limited data, and an interpretable early termination criterion is proposed.

The core of this paper is the proposal of a simple modification of the Noise2Noise training method that consistently enhances its denoising performance, getting closer to and often surpassing equivalent traditional learning methods, even when training on relatively small datasets.

**Summary:**

the performance of Noise2Noise drops when the amount of training data is reduced, limiting its capability in practical scenarios. Commonly used datasets for training denoising networks are designed for synthetic noise, which can be added on-the-fly, and are usually made up of several hundred to several thousand images. However, the smaller training datasets get, the less one can theoretically expect Noise2Noise to remain competitive.

The assumption of unlimited noisy samples per scene is not realistic in many practical scenarios (e.g. if a noise model is not known). Let us assume only two noisy samples (,) per scene are available.

Assuming noisy samples are drawn from the same distribution, one obvious idea is that they can both be used as input and target respectively(e.g. the pair and can be used in the network){refered as “AltN2N”}

Furthermore, changes in only a few pixels in the input or the target can yield virtually unseen samples of a scene. In a Noise2Noise setting, the pixels (or pixel regions) in and are interchangeable as long as the images are well-aligned, and the noise is not correlated (or correlation is not destroyed in the process). Under this assumption, one can swap one or more single pixels (or pixel regions) between and , such that two new unseen yet plausible images and are generated[Figure 1]. {refered as “surrogate N2N” or “SN2N}

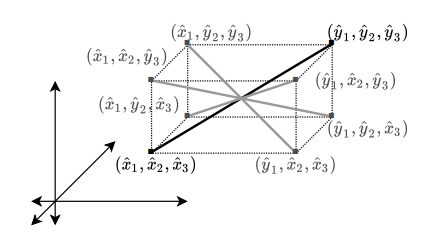


Figure - Illustration of the noise surrogates strategy, for a Noise2Noise image pair of 3 pixels(or pixel regions) each

Noise2Noise training might be more susceptible to overfitting than equivalent Noise2Clean training, especially with high levels of noise.

If the case of AN2N(also applies to SN2N), if the network is truly efficient at denoising, the difference between and should be minimal since the base image is the same in both inputs, the only difference between them is the noise.

this cannot be a loss function (if the network sends all pixels to 0, the minimum will be reached). However, when computed on the training set, besides the actual loss function, this measure would be an indicator of when the noise in the targets is attempted to be reproduced, and would, upon divergence, signal the need for training to be terminated, yielding an Early Termination rule.

For and to be available at every iteration, they both need to be computed, and could equally contribute to the regression problem:

the termination criterion is relatively orthogonal to the loss function, and therefore, the proposed interpretable measure derived from the training set alone can be indicative of overfitting.