**Paper review of:   
Multi-Contrast Super-Resolution MRI Through a Progressive Network**

Application domain.

A summary of the application domain of this paper would be the use of a one-level non-progressive neural network and a two-level progressive neural network to sample multi-contrast super-resolution (SR) MRI images. The non-progressive network is used for low up-sampling and the two-level progressive network is used for high up-sampling.

The paper has three main contributions which will be described to give more detail:

1) The Wasserstein generative adversarial network with gradient penalty (WGAN-GP) architecture, which is used in the two-level progressive neural that can obtain excellent MCSR results with the use of a high up-sampling factor.

2) When combining multi-contrast information in a high-level feature space leads to a significantly improved results over the combination in the low level pixel space

3) The contribution of a composite loss function including the mean-squared-error (MSE), perceptual loss and a texture matching loss to ensure that the generated images are able to recover texture details and are faithful to the ground truth.

Methods.

The methodology is divided in 5 sections.  
In section 1 the **overall super-resolution process** is described. A super resolution process is finding the inverse solution to a down-sampling process. It is impossible to find the exact inverse, however close estimations are possible using a deep learning-based multi-contrast super resolution method for SR T2 weighted imaging by incorporating high-resolution PD or T1 weighted images as reference images.

In section 2 **down-sampling and zero-filling** is described to produce low resolution (LR) images. By down-sampling and zero-filling in Fourier space the image size is unchanged but the image quality is degraded.

In section 3 the **one-level non-progressive network** is described. The network is also based on the WGAN-GP, which includes a generator and discriminator. The generator consists of an encoder which has eight sequential convolutional layers, each of which is followed by a rectified linear unit (ReLu). The generator also consists of a reference feature extraction network.

In section 4 the **objective function** is described. The function includes four parts:

1. The **adversarial loss** in the generative adversarial network framework is used to train the generator.
2. The **Mean-Squared Error** evaluates the difference between the output of the generator and the corresponding ground truth at pixel-wise level. It can greatly improve the signal-to-noise ratio of generated images.
3. The **perceptual loss** overcomes the problem that some details may be lost due to over-smoothed SR results. The perceptual loss recovers more details by measuring image similarity in a high-level feature space.
4. The **texture matching loss** contributes to generate an image with great similarity between the output of the generator and the ground truth by statistically matching extracted features.

Section 5 is about the **two-level progressive network**. The proposed network sequentially up-samples the image in small steps, resulting in a large up-sampling factor.

The image quality **evaluation metrics** that are used in this paper are structural similarity, peak signal-to-noise ratio and information fidelity criterion.

Discussion.

One common problem in training a Generative Adversarial Network is that it is highly unstable. This is inherently the case since often the generator and the discriminator are trained simultaneously while competing against each other. In the aforementioned paper the discriminator was trained four times before the generator was trained once. This might help stabilize the model.

Another strong point of the methodology is the use of multiple loss functions. Each loss function has its own (dis)advantages. By using the four loss functions, the power of each one is combined and results in a generally better applicable loss function for this application.

Using peak signal-to-noise ratio (PSNR) as a metric is sometimes problematic, since a higher PSNR usually indicates a reconstruction of higher quality, but this is not guaranteed. PSNR is also proven to be outperformed by most other popular evaluation metrics, so the use of PSNR does not seem necessary in this paper. But on the other hand it is beneficial to use multiple evaluation metrics to ensure that the model is working optimally and correctly.

Recommendations.

* Although MSE leads to a high signal-to-noise ratio in reference to the ground truth, it tends to produce over-smoothed SR results. Therefore they also use perceptual loss. **Is there a way to combine these two to avoid one problem being created that has to be solved with another loss function.**
* Also two distorted images with the same MSE may have very different types of errors, some of which are much more visible than others. One error might be preferred over the other, which is now not addressed with the use of MSE.
* PSNR is no longer regarded as a reliable indicator of image quality degradation [1]. Although a higher PSNR generally indicates that the reconstruction is of higher quality, in some cases it may not. Bear in mind that none of these objective measures are particularly good at predicting human visual response to image quality. Sometimes PSNRs vary wildly between two almost indistinguishable images. SSIM is recommended. So since this is already used we propose just skipping PSNR.
* The Cognitive Interaction Problem. It is widely known that cognitive understanding and interactive visual processing (e.g., eye movements) influence the perceived quality of images. For example, a human observer will give different quality scores to the same image if given different instructions [1].

[1] <https://ece.uwaterloo.ca/~z70wang/publications/ssim.pdf>