Response Letter for Paper ID 733

On The Classification-Distortion-Perception Tradeoff

We are grateful to the three reviewers for the time and effort in reviewing this paper, and for the recognition of the originality, quality, clarity, and significance of this paper. We will improve the paper according to the comments.

Reviewer 1

Plots in Figure 3. Thank you for the suggestion. We try a new plot style, as shown in Figure I (bottom left). We will provide new plots in this style in the revised paper or in the supplementary.

Using different classifiers. In Figure 3, we present the results of two classifiers: a CNN-based classifier and a logistic regression classifier. Different classifiers indeed lead to different tradeoff boundaries (please note the values of error rate). To address your comment, we plan to use a third classifier and redo the experiment. Due to limited time of rebuttal, that experiment cannot be finished now. We will add the experimental results into the revised paper.

Reviewer 2

Results on general natural images. We agree with you that more experimental results can make the conclusion more solid. We plan to test on the CIFAR-10 dataset. Due to limited time of rebuttal, that experiment cannot be finished now. We will add the experimental results into the revised paper.

Results under other degradations. Thank you for the suggestion. We conduct a new experiment about image super-resolution (SR) on the MNIST dataset. Original images are down-sampled by bicubic with a factor of 6. We use the structure of the well-known SRCNN [Dong *et al.*, ECCV 2014] for training SR networks. Other settings are the same as for the denoising experiment (using the same pretrained CNN-based classifier). The results are shown in Figure I (top row). They can confirm the conclusion drawn in the paper.

Reviewer 3

Clarification about the statement in Section 4 (line 247-251). Thank you for the suggestion. In line 247-251 we discuss the correlation between classification error rate (which is evaluated by a classifier) and human recognition (which is evaluated by ourselves). The human recognition is different from the visual quality: human recognition means whether the class can be correctly recognized by human, visual quality (perceptual naturalness as defined in this paper) means whether the image looks like a natural image. We will revise the wording to avoid confusion.

Visual comparisons for distortion and classification. We will include more visual results in the revised paper or in the supplementary, such as CIFAR-10 results and SR results. As an example, please check Figure I (bottom right).
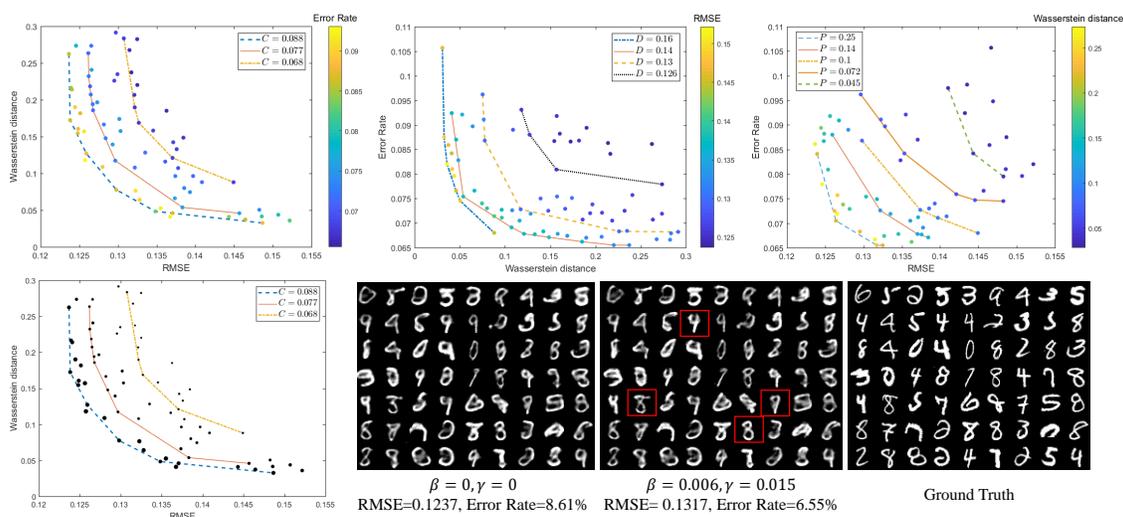

Figure I: Top: profiles of the CDP function for MNIST and SR. Bottom left: using a new style for the top left plot, where the size of each point indicates the corresponding error rate (quantized). Bottom right: some visual results to display the C-D tradeoff, where red boxes indicate examples that appear more recognizable by human.