Network-to-Network Translation with Conditional Invertible Neural Networks

Author Response

We thank all the reviewers for their positive feedback and for valuing the importance of the problem and the novelty of our approach, and for acknowledging its potential benefits for a variety of research communities as demonstrated by a comprehensive set of experiments.

We will incorporate valuable suggestions regarding related work into the final version and will go into more detail about multi-task learning and INNs in general. Thanks also for the comments on writing style and title, which we will gladly take into account. In particular, as suggested by **Reviewer 1**, we will replace the abbreviation "cINN" in the title by the full name of the method. We furthermore agree with **Reviewer 3** about the importance of the comparison of our cINN with an MLP and will move parts of this section from the supplementary material into the main text.

Reviewer 2 suggests that the quantitative comparison of our method with the given other works for the *text-to-image translation* task will benefit from additional metrics beside the Inception Score (see Tab. 1 in the main paper). We generally agree with this statement and therefore provide an additional comparison in terms of FID scores:

	our	DM-GAN	AttnGAN	Mirror GAN	SD-GAN
FID ↓	**30.63**	32.64	35.49	no pretrained model	no code available

To obtain these scores, we used the authors' official implementation available at `https://github.com/MinfengZhu/DM-GAN`. Note, however, that our method utilizes a pretrained expert generator which was trained on the *ImageNet* dataset (i.e. BigGAN, see l.173). Thus, we evaluate FID scores w.r.t. the validation split of the *ImageNet* dataset.

We furthermore agree with **Reviewer 4** that a detailed study of dual-way translation is an interesting avenue for future works.