**Thank you for your concerns. We hope this response can address all your questions. Best wish :)**

For Reviewer #1: Thank you for your great approval. For the first question, Figure 8(b) in the original paper shows that the proposed method outperforms others. We have also checked more randomly selected samples. Honestly, some samples have similar results visually, especially the ones with fewer details. Like Figure 9(q) and 9(s) in the paper, the difference of error maps is not significant while the PSNR is 0.21db better.The former has lower intensity visually in the top-right area but it leaves more aliased artifact in the top-middle area. However, both areas are not important for clinical diagnosis and our methods has more distinct edge in the zoom-in area. As for lesion images, our methods can restore details with enough sampling rate (like 15%). While with aggressive sampling rate, the details might be hard to reconstruct because of the lack of high frequecy phasing lines as shown in the Figure 8 of the supplementary material. Anyway, the low frequency part are de-ailiased well. Up to now, we haven't noticed any poor result in our dataset yet. The checking process will be continued, and the trained weights will be released with our codes if the paper is accepted.

For Reviewer #2: Thank you for your suggestions. We take 11-fold cross validation experiments on our dataset. The result of PSNR is shown in Figure (a) and it demonstrates the robustness. As MRI reconstruction serves clinic diagnosis, radiologists' interpretation is indeedly a more reasonable standard. A formal surver for radiogists should be taken in the future. Acutally, we showed our result to some clincians and won their recognition. Our result have rather good quality and keep consistency with groud truth in distinguishing the edge of tissue is distinct or not, which is a major evidence for diagnose of certain diseases. We hope to conduct further experiments on the FastMRI dataset but it will take a long time for training. Because of our insufficiency of GPU devices and the limitaion of time, what we can show is the result trained with FastMRI eval dataset. The loss curves are shown at Figure (b), we believe further training and fine-tuning on hyper-parameters will promote the convergence. An qualitative result is shown in Figure (c) and (d). It can be seen that our method can reconstructed accuracy details for FastMRI dataset. The training code for FastMRI will be available as well if the paper is accepted and the training process will be continued.

(a) 11-fold Cross Validation (b) Result on Reduced FastMRI (c) GT (esc) (d) Rec
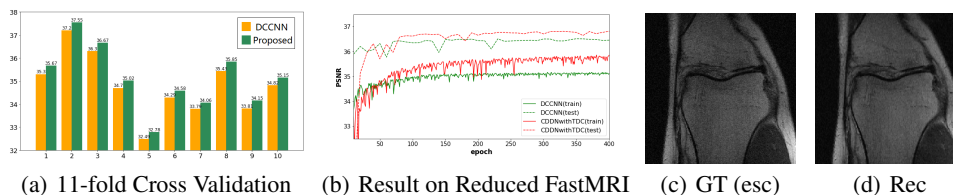
For Reviewer #3: I'm sorry for your confusion. A revision will be taken on the paper and we're trying to provide more details and emphasize key points here. In this paper, we propose a novel DNN architecture for MRI reconstruction. Our main contributions are (1) Cascaded dense blocks, (2) combination of dilation convolution and dense connection, (3) novel Two-Step Data Consistency layer. Cascading technology is proved to be effective in our experiments and other related works as it alleviates the difficulty of residual learning. Discussion of the benefits can be found in the Section 6 of the Supplementary Material. As mentionded in Section 3.4 of the paper, dilation convolution can expand receptive field for the demand of de-aliasing task. With dense connection, the feature maps from different receptive field sizes can be fused together for reconstruction. Experiment in Section 4.3 shows the necessity. All the designs of De-Aliase Module are proposed base on above analysis and are proved to be effective by extensive experiments in the paper.

As for data consistency term, the first thing to confirm is that the raw MRI data is k-space (frequency-domain) data. Unlike traditional image restoration task, the *corrupted* data is exactly true in the sampled location. During reconstruction, we have to ensure the invariance of the *true* part (i.e. the first term in Eq.2 of the paper). Direct replacement can meet the requirement, while it brokes the *self-consistency* of frequency information. It means the hybird result are unnatural in image-domain. In other word, direct replacement only corrects specific(sampled location) k-space data while leaving others in *outdated* state. As stated in Section 3.2 of the paper, we use complex-value structure to preserve information. The complex part of intermedia result can benefits the method with its latent information. However, only real-valued image is accepted in Human Visual System. In our data consistency unit, we apply modulus calculation on the firstly corrected data to imporve image-domain reasonablity. Finally, another direct replacement is taken for ensuring invariance as we have no reason to abandon *exactly true* data. Empirical experiments prove the necessity. Section 3.5 states our TDC module, Figure 7 illustrates the benefits and Section 2 of the Supplementary Material takes further discussion. We hope these explanations can help you follow the exact details.

Our advantage is reasonable in the present state of MRI reconstruction research. As shown in the Table 2 of the paper, our method(Proposed-C10) outperforms the secondary method by about 0.7db in PSNR. The minimun is 0.68db while the maximin is 0.80db. Accroding to recently published articles of MRI reconstruction, even if the advantage is hardly "very significant", we thought it can be "pretty good". Figure 9(t) also shows our advantage in the detail restoration by comparing with 9(p) and 9(q). More validation experiments in **"For Reviewer #2"** show the robustness and superiority.