We thank the reviewers for their thorough and constructive reviews. We first address a concern raised by all reviewers and then respond to each reviewer individually.

Application to neural recordings: Although all reviewers found the work interesting, and the simulations realistic and compelling, a shared concern was that we have not yet tested gaudy images with real neural data collected from macaque V4. We are eager to try these experiments, but we respecfully submit that requiring an application to neural data is too high a bar for the following reasons:

1. There is strong precedent in NeurIPS and the computational neuroscience field to perform realistic simulations on new methods/approaches that then inspire new neuroscientific experiments (see [12, 14, 48-55]). Almost all NeurIPS studies rely on either simulations or previously-collected data to validate their approaches/methods. Our paper is no different, and we use state-of-the-art models and realistic paradigms. Our work goes beyond simulations, as we propose a modeling framework (i.e., a readout network) and provide intuition about DNN models that can be used for future experiments and models (e.g., DNN responses are strongly driven by high-contrast edges).
2. The reviews highlight a chicken-and-egg problem: We wrote this paper to convince experimentalists to run these costly experiments, but the reviewers suggest we should have already collected this data to convince experimentalists. We believe NeurIPS is an appropriate place for our work, as we propose a proof-of-concept, backed by thorough and realistic simulations, which will inspire experimenters and machine learning researchers to test the efficacy of gaudy images in their own work.
3. If we had already collected neural data, we almost certainly would not submit this work to NeurIPS, which is not an appropriate venue for new neuroscientific findings. This is because NeurIPS does not have ethical requirements for the treatment of animals nor statistical requirements for number of subjects, etc. In addition, due to the 8-page limit of NeurIPS, we would need to include experimental details—critical for assessing neuroscientific claims—as Supplemental Material, which is optional and thus not guaranteed to be peer-reviewed for NeurIPS.

For these reasons, we believe that our work advances the field and should not be penalized for not performing new neuroscientific experiments. **We respectfully ask reviewers to increase their score, if they agree.**

We will update the text with all of the reviewer's comments. We respond to a few of those comments here:

R1: We agree about linear mappings and will provide broader scope about them. We are unaware of any successful end-to-end DNNs to predict V4/IT responses (although we cite two for V1 responses [8,9]) and request R1 provides citations for them to improve our paper. Although R1 finds that a linear model outperforms the 3-layer relu network, it does not outperform the sigmoid or Resnet-like networks. This suggests gaudy images are more useful for models that are good fits to the ground truth responses.

R3: re: results from simulations would generalize to real neural data. We base this assumption on many previous studies that find DNNs are appriopriate models for neural responses [3-10,13-16,23] and that our results generalize across multiple DNN models. Thus, we have provided strong evidence to convince experimenters to collect real data. We feel this is the epitome of a NeurIPS paper: Propose a computational approach/method and demonstrate its effectiveness in order to inspire new experiments. re: "gaudy versions and the original images share the same neural responses at least to a significant extent". We disagree with R2 that we make this assumption. This assumption is more in line with data augmentation, which transforms images but assumes the labels remain the same. Here, gaudy images drive *different* responses than normal images (Fig. 3a). We believe this will hold for neural data and is one of the reasons why gaudy images train DNNs more efficiently.

R4: re: "no clear applications outside...visual cortex". We provide evidence that gaudy images may be useful for data augmentation for object recognition (Supp. Fig. 6). Please see our comments to all reviewers. re: other models of visual cortex to assess gaudy images. To our knowledge, for our simulations we have used models that achieve the highest predictive performance for neural data [7]. We are unaware of any other models that achieve such high performance and request R4 to provide citations to these other models to improve our paper. re: active learning of the linear model. We note that $\Sigma = \mathbf{X}\mathbf{X}^T$ is the covariance matrix of images $\mathbf{X}$ and does not depend on responses y. Thus, all images can be chosen before recording any responses. re: PCA on normal vs. gaudy images. We have applied PCA to the pixels (Supp. Fig. 1a-c) and find gaudy images do increase the variance for the lower PCs. re: other adaptive stimulus techniques. These methods are not applicable to our setting. Some methods do not consider training any model [13-16,50] and some methods consider models with $<$ 1k params [11,12,51,52] (we have 1.5 million params).

R5: re: "drop in final prediction accuracy". Because AL algorithms explicitly mismatch the training distr. from the test distr., this drop is expected from any AL algorithm when sample size grows large (when the prior of the AL algo becomes too strong). We propose a hybrid approach for large-data regimes (Supp. Fig. 3a caption). re: mean vs. median vs. contrast. We have found that all of these yield similar performance (including binarizing based on each channel separately), so we chose the simplest (i.e., thresholding on the mean). We will include these results. We note that one can scale gaudy images just as one would scale contrast.