



The latest news from Google AI

## Introducing NIMA: Neural Image Assessment

Monday, December 18, 2017

Posted by Hossein Talebi, Software Engineer and Peyman Milanfar Research Scientist, Machine Perception

Quantification of image quality and aesthetics has been a long-standing problem in image processing and computer vision. While technical quality assessment deals with measuring pixel-level degradations such as noise, blur, compression artifacts, etc., aesthetic assessment captures semantic level characteristics associated with emotions and beauty in images. Recently, deep convolutional neural networks (CNNs) trained with human-labelled data have been used to address the subjective nature of image quality for specific classes of images, such as landscapes. However, these approaches can be limited in their scope, as they typically categorize images to two classes of low and high quality. Our proposed method predicts the distribution of ratings. This leads to a more accurate quality prediction with higher correlation to the ground truth ratings, and is applicable to general images.

In “[NIMA: Neural Image Assessment](#)” we introduce a deep CNN that is trained to predict which images a typical user would rate as looking good (technically) or attractive (aesthetically). NIMA relies on the success of state-of-the-art deep [object recognition](#) networks, building on their ability to understand general categories of objects despite many variations. Our proposed network can be used to not only score images reliably and with high correlation to human perception, but also it is useful for a variety of labor intensive and subjective tasks such as intelligent photo editing, optimizing visual quality for increased user engagement, or minimizing perceived visual errors in an imaging pipeline.

### Background

In general, image quality assessment can be categorized into full-reference and no-reference approaches. If a reference “ideal” image is available, image quality metrics such as [PSNR](#), [SSIM](#), etc. have been developed. When a reference image is not available, “blind” (or no-reference) approaches rely on statistical models to predict image quality. The main goal of both approaches is to predict a quality score that correlates well with human perception. In a deep CNN approach to image quality assessment, weights are initialized by training on object classification related

datasets (e.g. [ImageNet](#)), and then fine-tuned on annotated data for perceptual quality assessment tasks.

## NIMA

Typical aesthetic prediction methods categorize images as low/high quality. This is despite the fact that each image in the training data is associated to a histogram of human ratings, rather than a single binary score. A histogram of ratings is an indicator of overall quality of an image, as well as agreements among raters. In our approach, instead of classifying images a low/high score or regressing to the mean score, the NIMA model produces a distribution of ratings for any given image – on a scale of 1 to 10, NIMA assigns likelihoods to each of the possible scores. This is more directly in line with how training data is typically captured, and it turns out to be a better predictor of human preferences when measured against other approaches (more details are available in our [paper](#)).

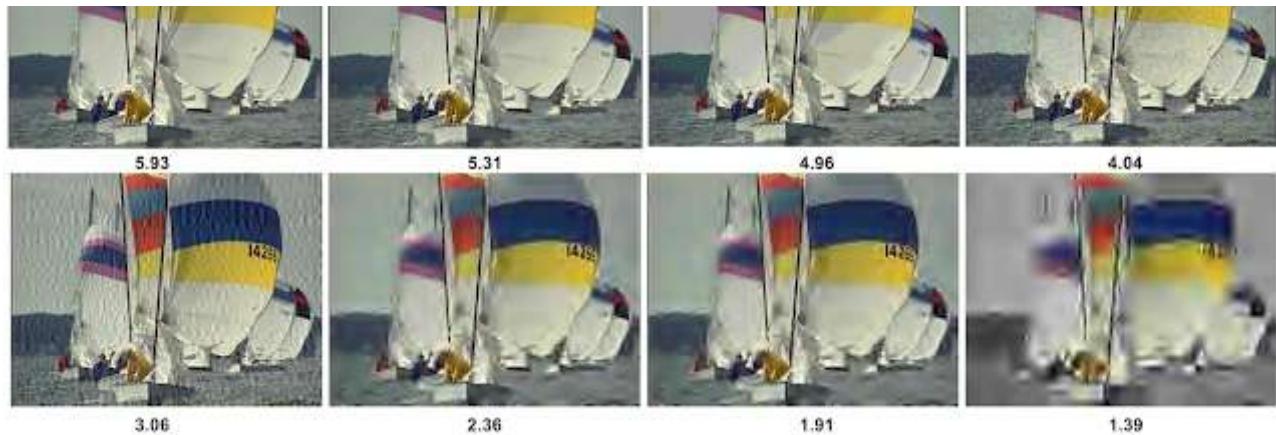
Various functions of the NIMA vector score (such as the mean) can then be used to rank photos aesthetically. Some test photos from the large-scale database for Aesthetic Visual Analysis ([AVA](#)) dataset, as ranked by NIMA, are shown below. Each AVA photo is scored by an average of 200 people in response to [photography contests](#). After training, the aesthetic ranking of these photos by NIMA closely matches the mean scores given by human raters. We find that NIMA performs equally well on other datasets, with predicted quality scores close to human ratings.



Ranking some examples labelled with the “landscape” tag from [AVA](#) dataset using NIMA. Predicted NIMA (and ground truth) scores are shown below each image.

NIMA scores can also be used to compare the quality of images of the same subject which may have been distorted in various ways. Images shown in the following example are part of the [TID2013](#) test set, which contain various types and levels of distortions.





Ranking some examples from [TID2013](#) dataset using NIMA. Predicted NIMA scores are shown below each image.

### Perceptual Image Enhancement

As we've shown in another recent [paper](#), quality and aesthetic scores can also be used to perceptually tune image enhancement operators. In other words, maximizing NIMA score as part of a loss function can increase the likelihood of enhancing perceptual quality of an image. The following example shows that NIMA can be used as a training loss to tune a tone enhancement algorithm. We observed that the baseline aesthetic ratings can be improved by contrast adjustments directed by the NIMA score. Consequently, our model is able to guide a deep CNN filter to find aesthetically near-optimal settings of its parameters, such as brightness, highlights and shadows.





NIMA can be used as a training loss to enhance images. In this example, local tone and contrast of images is enhanced by training a deep CNN with NIMA as its loss. Test images are obtained from the [MIT-Adobe FiveK dataset](#).

## Looking Ahead

Our work on NIMA suggests that quality assessment models based on machine learning may be capable of a wide range of useful functions. For instance, we may enable users to easily find the best pictures among many; or to even enable improved picture-taking with real-time feedback to the user. On the post-processing side, these models may be used to guide enhancement operators to produce perceptually superior results. In a direct sense, the NIMA network (and others like it) can act as reasonable, though imperfect, proxies for human taste in photos and possibly videos. We're excited to share these results, though we know that the quest to do better in understanding what quality and aesthetics mean is an ongoing challenge — one that will involve continuing retraining and testing of our models.



[Google](#) · [Privacy](#) · [Terms](#)