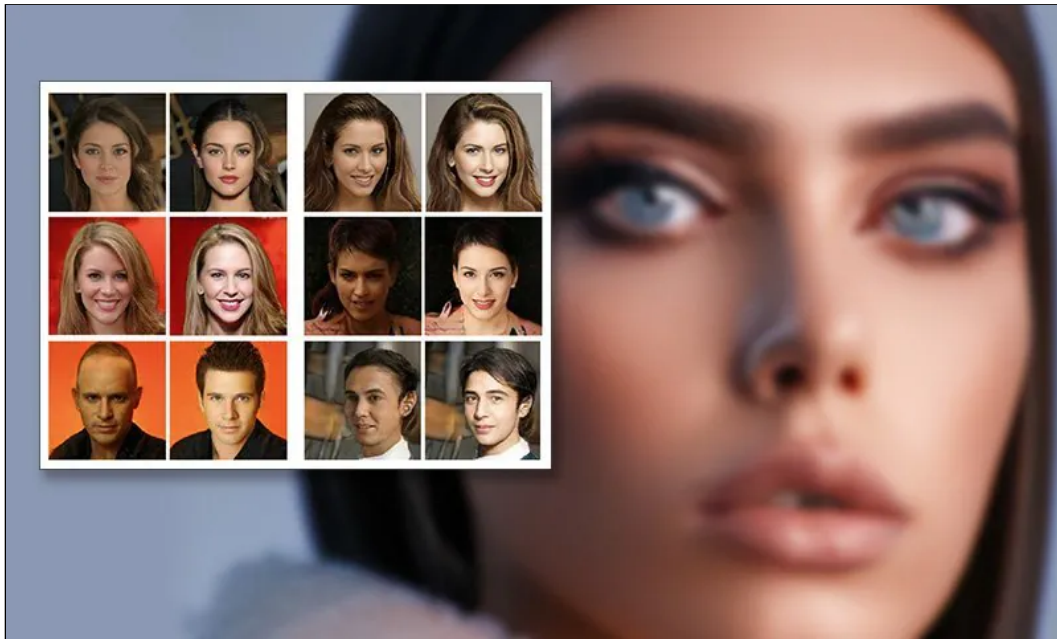


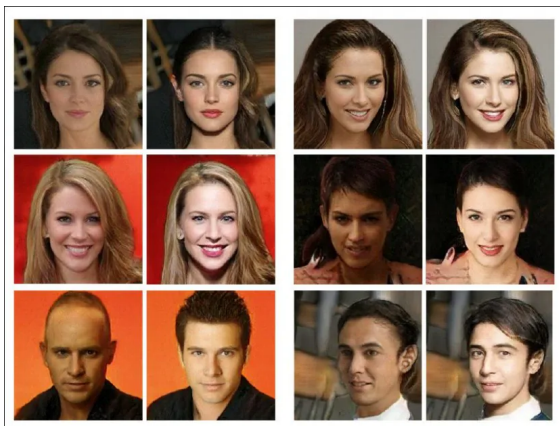
An AI System That Can Make Images of People More 'Beautiful'

Originally published at <https://www.unite.ai/an-ai-system-that-can-make-images-of-people-more-beautiful/>



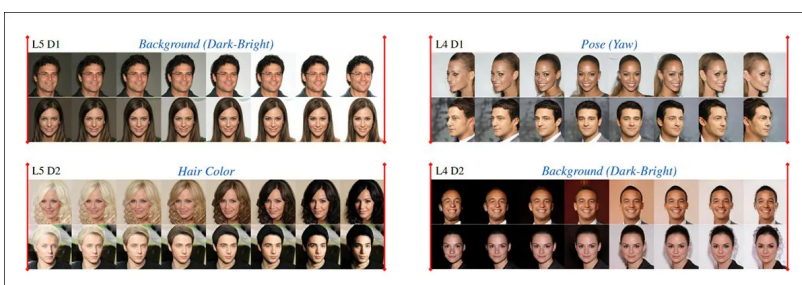
Background image: DALL-E 2 'Award-winning 8K photo of the most beautiful Caucasian catwalk model in the world' - <https://labs.openai.com/s/kRXusxOR5GcYyb6pqZjNH2AA>

Researchers from China have developed a new AI-based image enhancement system that's capable of making images of a person more 'beautiful', based on a novel approach to reinforcement learning.



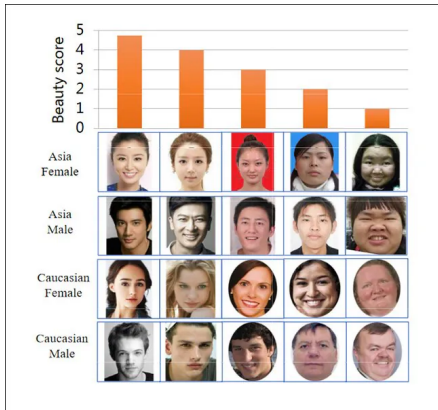
The new approach uses a 'facial beauty prediction network' to iterate through variations on an image based on a number of factors, among which 'lighting' and eye poses may be critical factors. Here the original sources (on the left of each column) are from the EigenGAN system, with the new results to the right of these. Source: <https://arxiv.org/pdf/2208.04517.pdf>

The technique draws on innovations discovered for the [EigenGAN generator](#), another Chinese project, from 2021, that made notable strides in identifying and gaining some control over the diverse [semantic attributes](#) within the latent space of Generative Adversarial Networks (GANs).



The 2021 EigenGAN generator was able to individuate high-level concepts such as 'hair color' within the latent space of a generative adversarial network. The new instrumentality to deliver a system that can 'beautify' source images, but without changing the recognizable identity – a problem in previous approaches. Source: <https://arxiv.org/pdf/2104.12476.pdf>

The system makes use of an 'aesthetics score network' derived from [SCUT-FBP5500](#) (SCUT), a 2018 benchmark dataset for facial beauty prediction, from the South China University of Technology at Guangzhou.



From the 2018 paper 'SCUT-FBP5500: A Diverse Benchmark Dataset for Multi-Paradigm Facial Beauty Prediction', which proffered a 'Facial beauty prediction' (FBP) network capable of ranking faces in terms of perceived attractiveness, but which could not actually transform or 'upgrade' faces. Source: <https://arxiv.org/pdf/1801.06345.pdf>

Unlike the new work, the 2018 project cannot actually execute transformations, but contains algorithmic value judgements for 5,500 faces, supplied by 60 mixed gender labelers (a 50/50 split). These have been incorporated into the new system as an effective [discriminator](#), to inform transformations that are likely to enhance the 'attractiveness' of an image.

Interestingly, the [new paper](#) is titled *Attribute Controllable Beautiful Caucasian Face Generation by Aesthetics Driven Reinforcement Learning*. The reason that all races except Caucasian are excluded from the system (consider also that the researchers themselves are Chinese) is because the source data for SCUT skews notably to Asian sources (4000 evenly-divided Asian females/males, 1500 evenly-divided Caucasian females/males), making the 'average person' in that dataset brown-haired and brown-eyed.

Therefore, in order to accommodate coloring variation at least within one race, it was necessary to exclude the Asian component from the original data, or else go to the considerable expense of reconstituting the data to develop a method that might not have panned out. Additionally, [variation in cultural perceptions of beauty](#) inevitably mean that such systems will need some degree of geographical configurability in regard to what constitutes 'attractiveness'.

Pertinent Attributes

To determine the primary contributing factors to an 'attractive' photo of a person, the researchers also tested the effect of various changes to images, in terms of how well such augmentations boosted the algorithmic perception of 'beauty'. They found that at least one of the facets is more central to good photography than good genetics:

D	Attribute	Spearman's ρ	p -value
$z_{2,5}$	Body Pose	0.418	0.0
$z_{3,4}$	Bangs	-0.857	0.0
$z_{3,5}$	Hair Style	-0.125	2.875×10^{-3}
$z_{3,6}$	Body Side	0.188	0.0
$z_{4,1}$	Pose	0.328	0.0
$z_{4,2}$	Background	-0.978	0.0
$z_{4,3}$	Pose	0.027	0.1551
$z_{4,4}$	Lighting	-0.561	0.0
$z_{4,5}$	Smiling	0.786	0.0
$z_{4,6}$	Face Shape	-0.720	0.0
$z_{5,1}$	Background	-0.198	0.0
$z_{5,2}$	Hair Color	-0.164	0.0
$z_{5,3}$	Hair Color	0.685	0.0
$z_{5,4}$	Lighting	-0.061	0.1484
$z_{5,5}$	Gaze	0.585	0.0
$z_{5,6}$	Lipstick Color	0.471	0.0

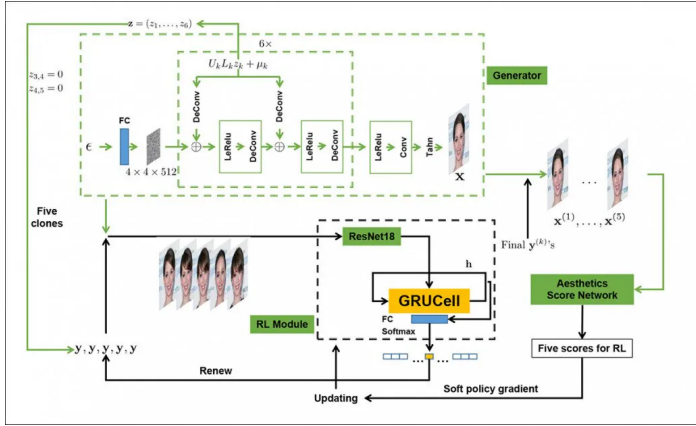
Besides lighting, the aspects that had the biggest impact on beauty score were bangs (which, in the case of men, can often be equivalent to having a full head of hair at all), body pose, and eye disposition (where engagement with the camera viewpoint is a fillip to attractiveness).

(Regarding 'lipstick color', the new system, which can work effectively on both male and female presentations of gender, does not individuate gender appearance, but rather relies on the novel discriminator system as a 'filter' in this respect)

Method

The reward function in the reinforcement learning mechanism in the new system is powered by a straightforward regression over the SCUT data, which outputs facial beauty predictions.

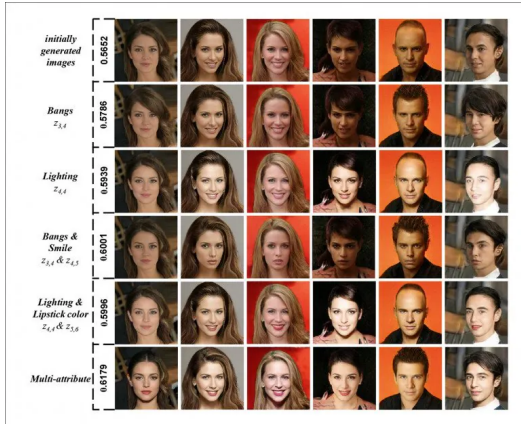
The training system iterates over the data input images (bottom left in the schematic below). Initially a pretrained [ResNet18](#) model (trained on [ImageNet](#)) extracts features from the five identical ('y') images. Next, a potential transformative action is derived from the hidden state of a [fully connected layer](#) ([GRUCell](#), in image below), and the transformations applied, leading to five altered images which are fed into the aesthetics score network, whose rankings, Darwin-style, will determine which variations will be developed and which discarded.



An illustration of the workflow for the new system.

The aesthetics score network uses an Efficient Channel Attention ([ECA](#)) module, while an adaptation of a pre-trained instance of [EfficientNet-B4](#) is tasked with extracting 1,792 features from each image.

After normalization through a [ReLU activation function](#), a 4-dimensional vector is obtained back from the ECA module, which is then flattened to a one-dimensional vector following activation and [adaptive average pooling](#). Finally, the results are fed into the [regression network](#), which retrieves an aesthetics score.



A qualitative comparison of output from the system. In the bottom row, we see the aggregated sum of all the individuated facets that have been identified by the EigenGAN method and subsequently enhanced. Averaged FID scores for the images are to the left of the image rows (higher is better).

Tests and User Study

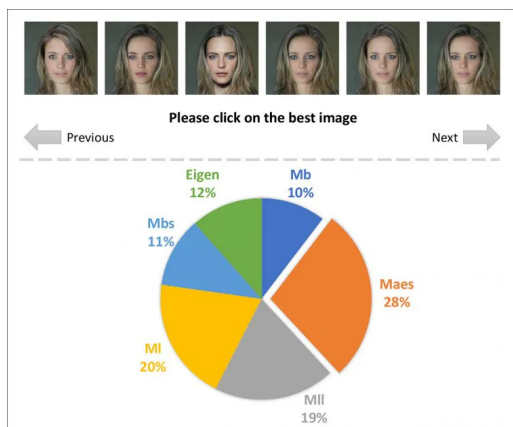
Five variants of the proposed method were evaluated algorithmically (see image above), with Fréchet inception distance (FID, [controversial in some quarters](#)) scores assigned to a total of 1000 images put through the system.

The researchers note that improving the lighting achieved a better attractiveness score for the subjects in the photos than several other more 'obvious' possible changes (i.e. to the actual appearance of the person depicted).

To a certain extent, testing the system in this way is limited by the eccentricities of the SCUT data, which does not have many 'bright smiles', and the authors argue that this could excessively over-rank the more typical 'enigmatic' look in the data, in comparison to the likely preferences of potential target end users (presumably, in this case, a western market).

However, since the entire system hangs on the mean average opinions of just 60 people (in the EigenGAN paper), and since the quality being studied is far from empirical, it could be argued that the procedure is more sound than the dataset.

Though it is dealt with very briefly in the paper, images from EigenGAN and the system's own five variants were also shown in a limited user study (eight participants), who were asked to select the 'best image' (the word 'attractive' was avoided).



Above, the GUI presented to the small study group; below, the results.

The results indicate that the new system's output achieved the highest selection rate among the participants ('MAES' in the image above).

The (Aimless?) Pursuit of Beauty

The utility of such a system is difficult to establish, despite what appears to be a [notable locus of effort in China](#) towards these goals. None is outlined in the new publication.

The previous EigenGAN paper suggests* that a beauty-recognition system could be used in facial [make-up synthesis recommendation systems](#), aesthetic surgery, [face beautification](#), or content-based image retrieval.

Presumably such an approach could also be used in dating sites, by end-users, to 'enhance' their own profile photos into a guaranteed 'lucky shot', as an alternative to using outdated photos, or photos [of other people](#).

Likewise, dating sites themselves could also 'score' their clients to create ratings and even [restricted-access tiers](#), though this would presumably only work via a liveness authentication capture, rather than submitted photos (which could likewise be 'enhanced' by the clients, if the approach were to become popular).

In advertising, an algorithmic method to assess beauty (a technology predicted by the late science-fiction author Michael Crichton in his 1982 cinematic outing [Looker](#)) could be used to select the non-enhanced creative output most likely to engage a target audience, while the capacity to actually maximize the aesthetic impact of face images, without actually overwriting them in the style of deepfakes, could boost already-effective images intended to garner public interest.

The new work is supported by the National Natural Science Foundation of China, the Open Fund Project of the State Key Laboratory of Complex System Management and Control, and the Project of Philosophy and Social Science Research from China's ministry of education, among other supporters.

** Many of the EigenGAN paper's recommendations point towards a commercially available 2016 book titled 'Computer Models for Facial Beauty Analysis', rather than academic resources.*

First published 11th August 2022.

Writer on machine learning, domain specialist in human image synthesis. Former head of research content at Metaphysic.ai, until its dissolution into DNEG's Brahma.ai.

Portfolio site: martinanderson.ai

Contact: martin@martinanderson.ai



[Discover More](#)