ART AND MACHINE LEARNING
CMU 2019 SPRING
PROJECT 2

Medusozoa

Yixiao Fu, Yi-Chin Li, Jiyuan Li, Zixin Yao

DESCRIPTION

Concept

Hindra, a gawky and self-loathing yet wickedly talented and whimsical goddess of creation depicted by Kristen Wiig on Saturday Night Live, is a beneficiary of favoritism from Jupiter for her erratic eruption of captivating novelties. Hindra's portfolio flaunts a constellation of curios ranging from flamingo, clam and peacock to jellyfish, volcano and music. While her more conservative and evolutionist colleagues relying on morphology and adaptation based varieties, such as bears in distinct biomes or an elongated version of a horse for recreational purposes, Hindra excels through singularities that defy a continuous narrative of natural history. Hindra's repertoire documents our own enchantment with the discoveries of biology. The glamour comes with a blow within the lacuna between established phylums and genuses. It both resembles a chimera of arbitrary contingents and harnesses tokens of humanoid faculties. It is nature's work of contraptions, delivered through the lense introspection.

Jellyfishes, despite flamboyant in HDR representations that embody human beings' triumph of the deep sea, even in their sexually mature phase as Medusas, are not intended to be appreciated visually. Their existence is ethereal, almost an ephemeral substance that is homogenous with its surroundings. Their crimson guts are the solutions of efficiency, an utterance not meant to be heard, for red doesn't reach the deep and black is a tautology. The luring iridescence is an elaborate fabrication, since bioluminescence is predominantly a defense mechanism and it takes great toll on the 'light artist' given the scarcity of enzyme. Presentations of discoveries misinterpret the scintilla of deep horror and survival instinct as the signature of extraterrestrial paradise. Convolution Neural Nets greedily sample those fantasies as essences and populates the reconstructions with those enhanced projections relentlessly. Just as veiled human biases could be exposed in an alarmingly stark manner with the brute force of machine learning, our creative whims could also be machine learnt and laid bare with an under-trained GANs model, as a romanticist trope in art criticism loves to tout that deskilling and naivety follows sophistication and maturity.

Technique

Pix2Pix

Pix2pix[1] learns a mapping from input image to target image, and in our project the input image is edge/contour of an object (jellyfish), and the target image is the actual object. Pix2pix is a variant of conditional GAN, in which the generator (which has an encoder-decoder architecture) learns the mapping, and discriminator tried to distinguish between the output image from the generator and from the original data set.

Edge Extraction
The edge detection/extraction algorithm we used is called Holistically-Nested Edge Detection[2]. It uses essentially fully convolutional neural nets and by “nested” they mean multiple receptive field sizes (and thus multiple layers). The method is supervised so we used their pretrained model.

Photoshop
generate bitmap and mosaic as semantic labels in training data

Process, Results and Reflections

1. Conditional GANs: A Potential Interface for Participatory Art Making

Among the gangsters of GANs, conditional GANs allow for greater control/out of control over the final output from the generator. Semantic labels are embedded in the noise vector input to the generator and the image input to the discriminator. This additional mechanism allows the misuse of correlation learnt in trained models for novel imageries. We resort to Pix2Pix for post-training participation where imputend test inputs could be applied to the trained model as confounders. This workflow anticipates a visual synthesis where dollops of learnt semantics ramming into each other following nonsensical commands creating unruly compositions of maimed familiarities, a reductionist model of representational impressionist art making, despite one with a lopsided distribution of impulse and articulation.

2. Ideologies of Datasets

We initially set out to explore three distinct visual datasets, each harbours a contended theme in iconology. Could GANs help flog those schemes out in revealing visual forms? What does it mean to learn a semantic model of visuo-cultural concepts?

Jellyfish:
Human centered representation of objecthood

Visual Anthropology(Vivian Maier Collection):
Posthumous discovery of amateur representation

Style Junk(Japaneseque Vignettes):
Amicable kitsch distributed through networks of cultural ordinaries.

Considering feasibility we proceeded to work with the Jellyfish datasets. We gathered around 500 professional photographs of jellyfishes. The size is viable according to the author of Pix2Pix.

3. Implementation and Workflow Improvement

Edge Extraction

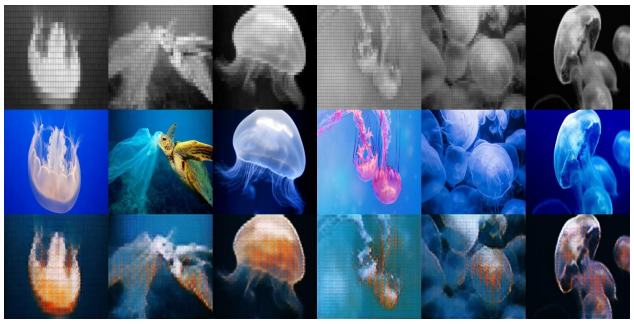
The edge extraction was provided by the pix2pix authors' source code [1, 2] and we did not do any modification on that. However, we wrote a script so that we can generate datasets from original image to resized {edge; original image} in one command line. There are four steps: resize original image, generate edges images in .mat format, convert edges images from .mat to .jpg, and finally combine the two images in one image.

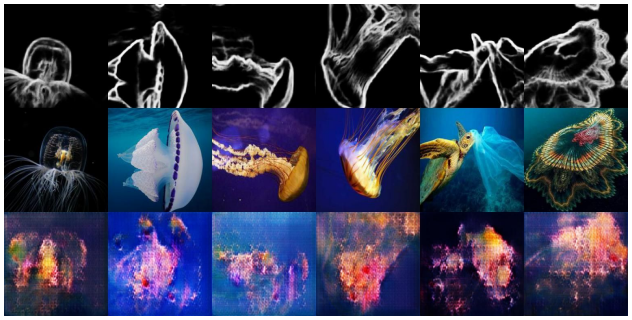
Pix2pix

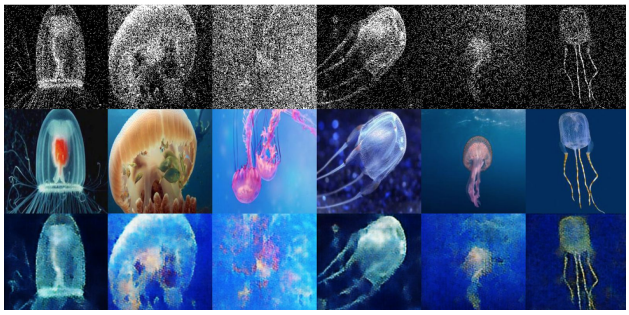
Since pix2pix takes longer to train, we imported the given notebook code into python files, which could facilitate the training and testing process. We just need to run train.py and test.py to train or generate. The parameters are adjustable in the command line argument, and also we do not need to let screen on all the time to monitor by using tmux. Importantly, we added support to save and load models so that we could continue training.

4. An Assortment of Semantic Labels

We tried three different semantic labels for the jellyfish dataset, mosaics, edges and bitmaps. Since loss functions are also learnt with Pix2Pix, the mosaics and bitmaps groups had provided overwhelming visual cues that caused the model's atrophy into merely colorization learner.
For a semantically richer model, we decided to proceed with the edge group after inspecting the early training results of the three.


Using 'mosaics' as semantic labels

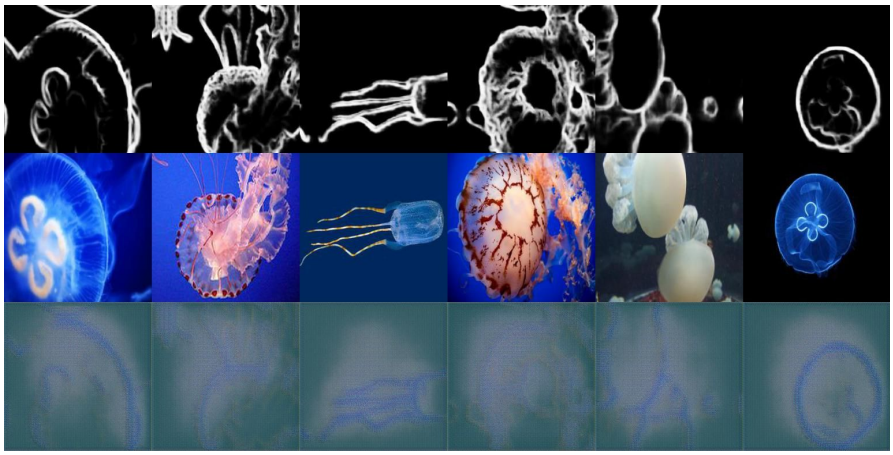

Using edges as semantic labels


Using bitmap as semantic labels

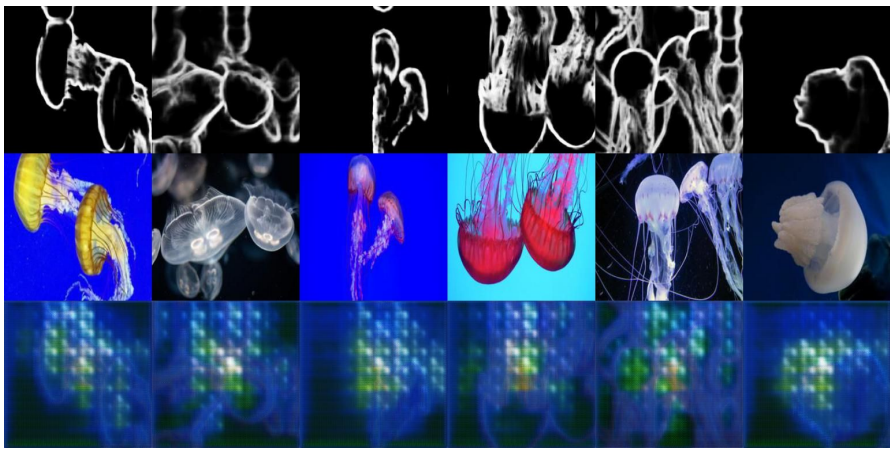
5. A Journey Through 200 Epochs Pix2Pix w/Edges as Semantic Labels

We handpicked representative moments of our training process, some of which bear visual traits whose frameworks of appreciation are well established within our visual culture.

Epoch 10



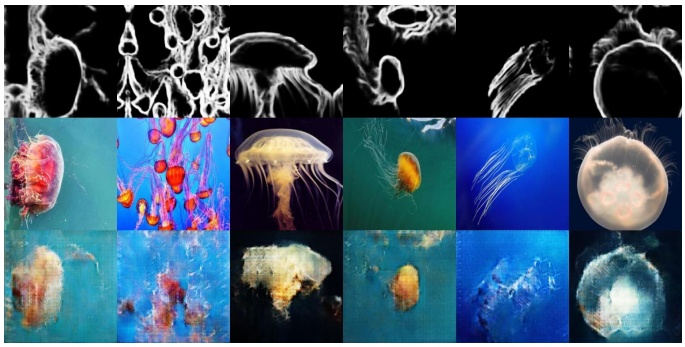
Epoch 30 (Retro Voxel Stage)



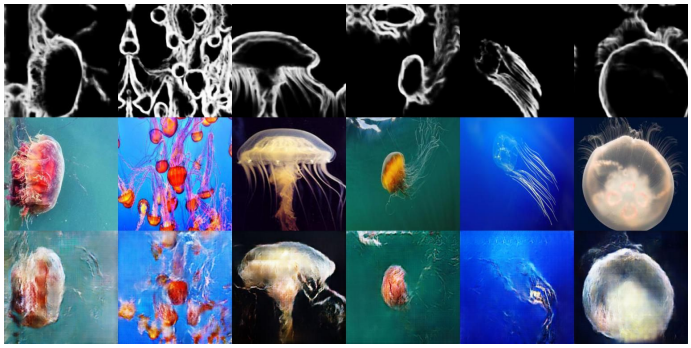
Epoch 50 (Monet's Water Lilies Stage)

Epoch 60

Epoch 100



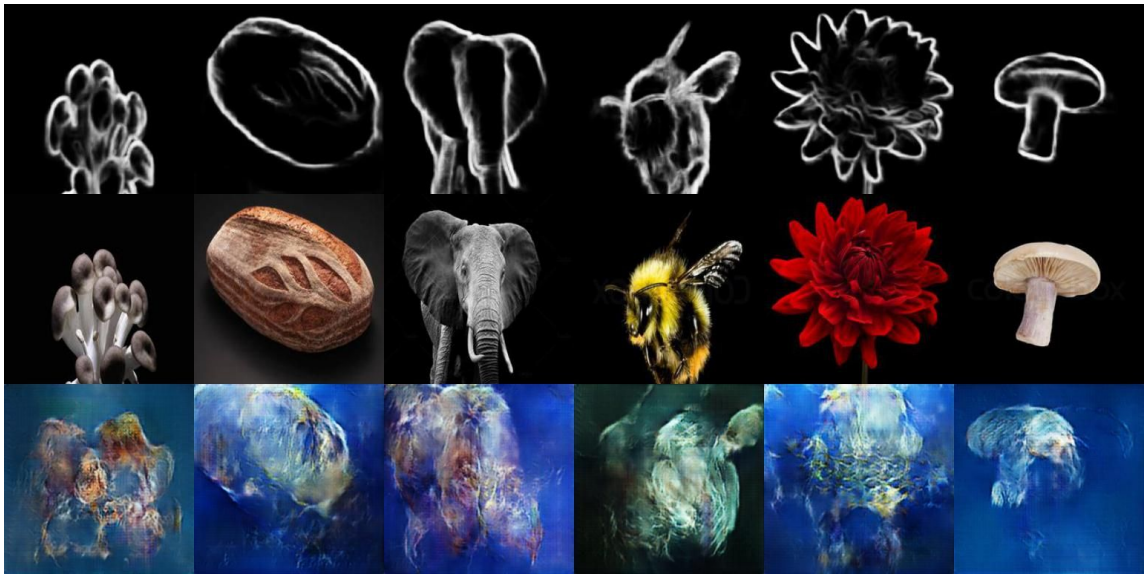
Epoch 200



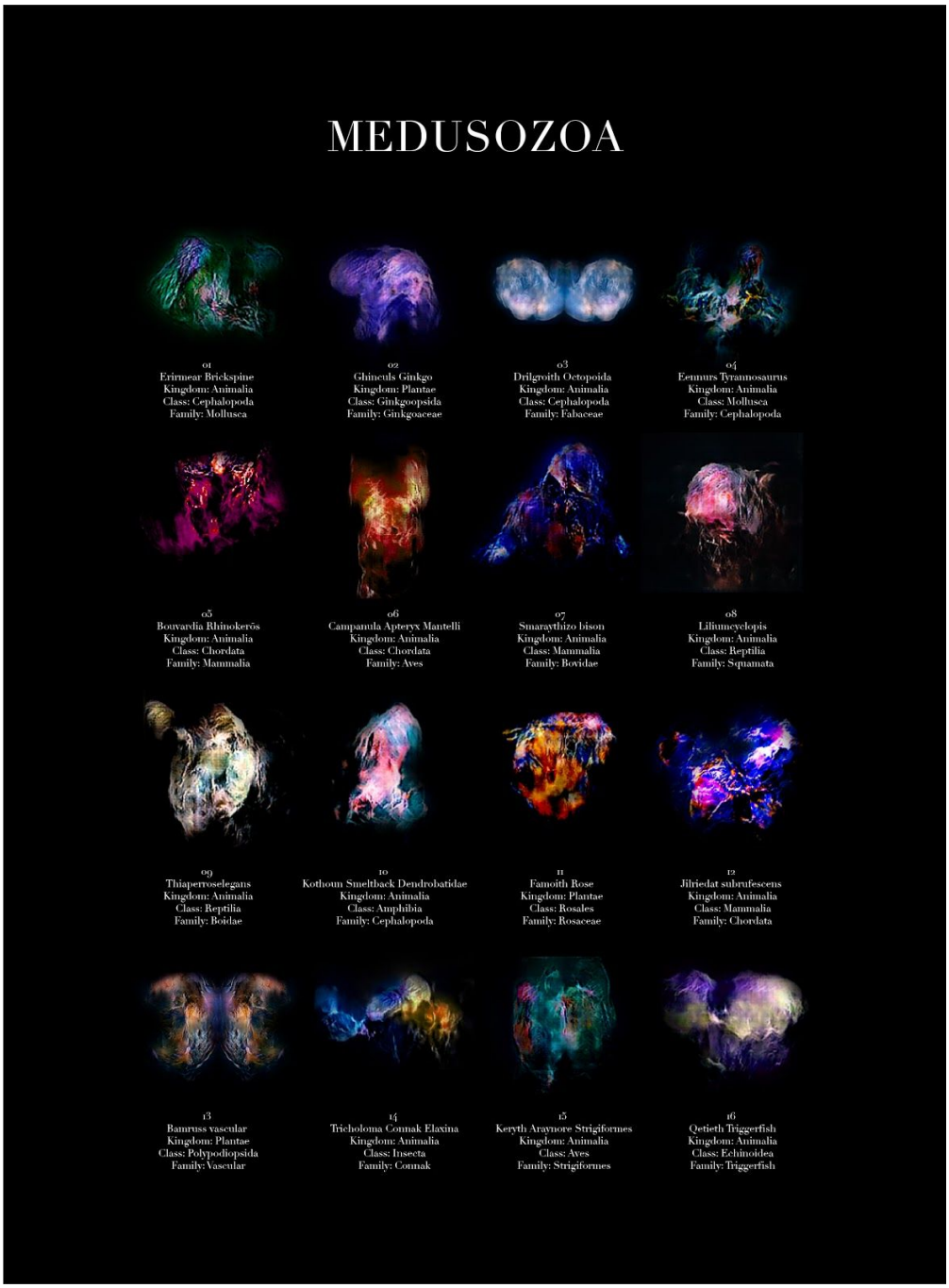
6. Populate an Otherworldly Waterfront

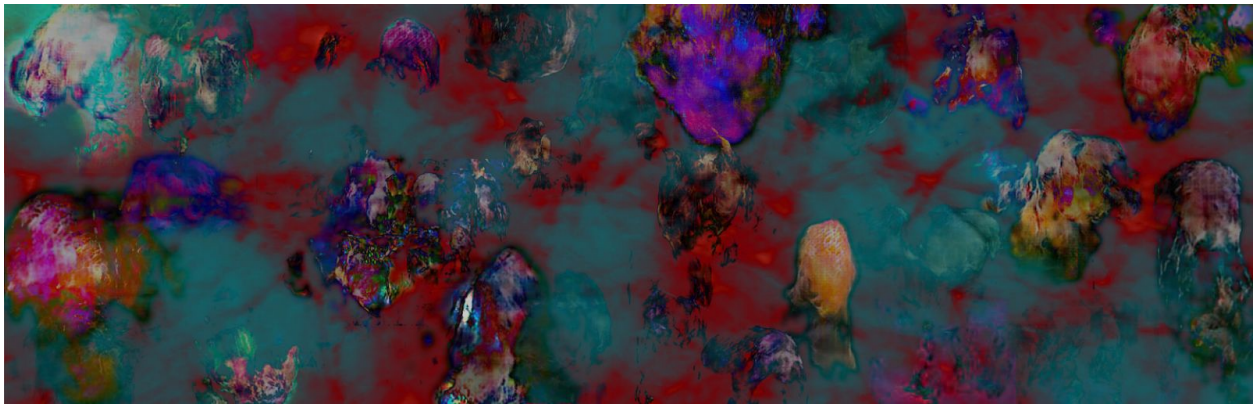
After more than 100 hours of training, at 200 epochs, our model start to deliver decent results on the training data. That's where we terminated the training process and started experimenting with test inputs.

Our training set consists exclusively of jellyfish images. For the test set, we employed edge information of a wide range of species and objects. The pulsating bells and long trailing tentacles of the jellyfishes recede into the dynamics of deep sea ebbies and currents. This convolution releases the interplay of light and gel from the jellyfish gestaut, leaving only the mesmerizing swirl of scintillating biosignals in those adaptations, which lend themselves to exterrestrial interpretations.



We selected the creature that has intact shape into a collection of extraterrestrial beings . Since most of the output images have different background color, we used photoshop to adjust the image through color latitude, color saturation and Channel Mixer. After processed all the selected image, we named the creature according to the object in the test data and mix with the name generated online (https://www.fantasynamegenerators.com) For example: If the object was tested by octopus, the name will be selected fantasy name of sea creature + scientific name of octopus = Gothagoctopoda





In the final visual output, we contextualized our extraterrestrial exemplars in both an idyllic waterfront and a battle royale post apocalypse. The versatility of our learnt semantics got its DNA from our original motif, the picturesque medusas, which embodies the dubious dyads of exploration and annihilation, fanfare and intoxication, survival and sacrifice, geniality and menace.

CODE: https://github.com/zxyao/jellyfish

RESULT: https://github.com/zxyao/jellyfish

Reference:
[1] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-toimage translation with conditional adversarial networks. In CVPR, 2017.
[2] S. Xie and Z. Tu. Holistically-nested edge detection. In ICCV, 2015.