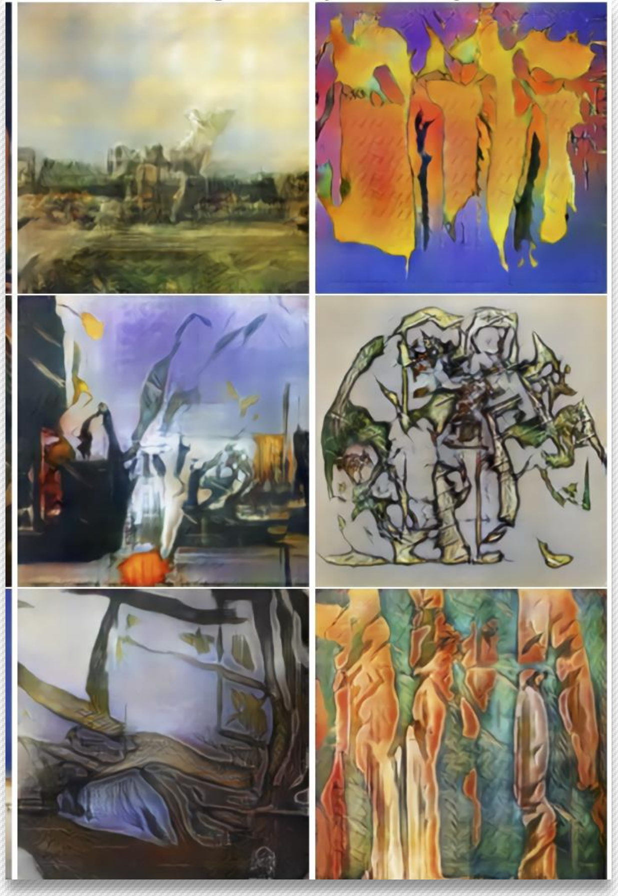


Creative GANs for art generation

Exploring the creative space to learn about the history of art and its creation process

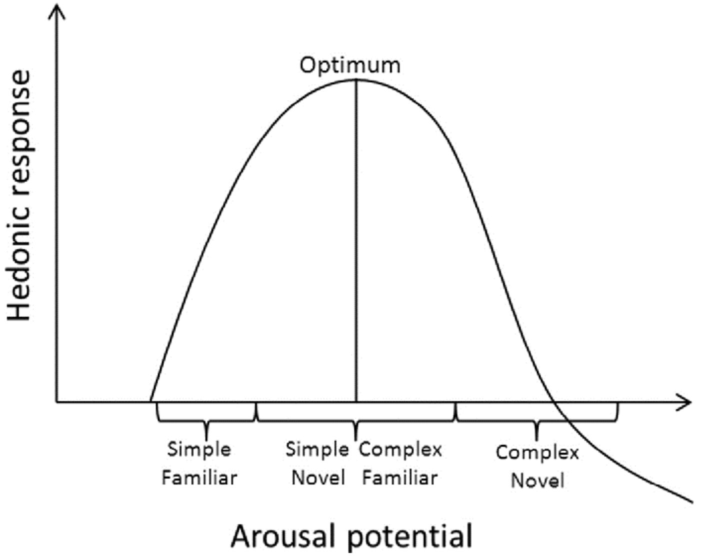
Davide Montagno Bozzone

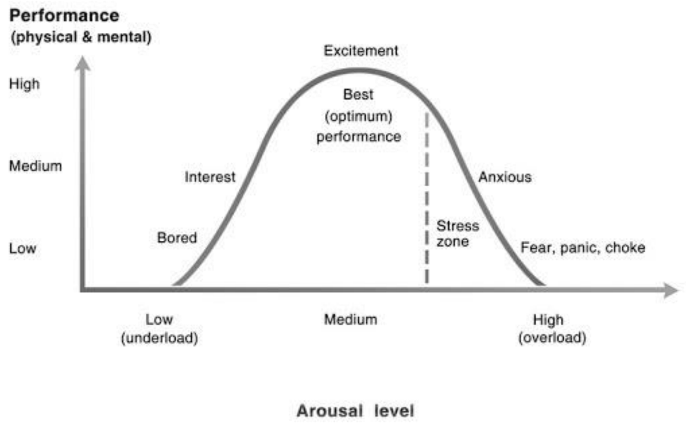
–

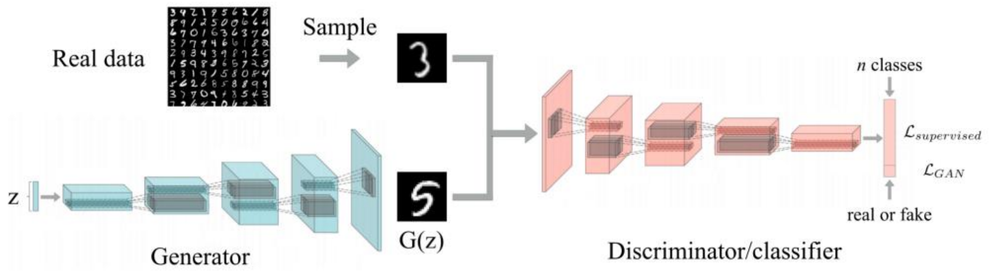
535910 (Project 4)

Human Brain and cognitive emotions – Create a new *creative Agent*

- Arousal potential. How people are affected by stimuli
    - Wundt Curve
    - Colin Martindale hypothesis for abituation
    - Berlyne theory used for experiments
- History of Creative Agents (with human feedbacks)
    - Emulative Agent
- Machine generated art recognized as hallucination-like (need ambiguity)

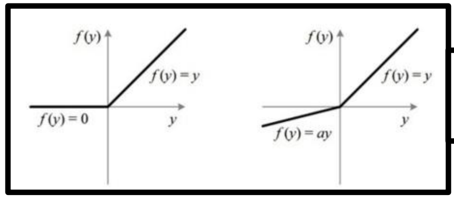




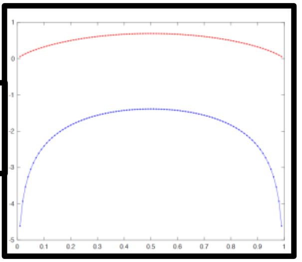


$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{data}}[\log D(x)] + \mathbb{E}_{z\sim p_z}[\log(1 - D(G(z)))]$$

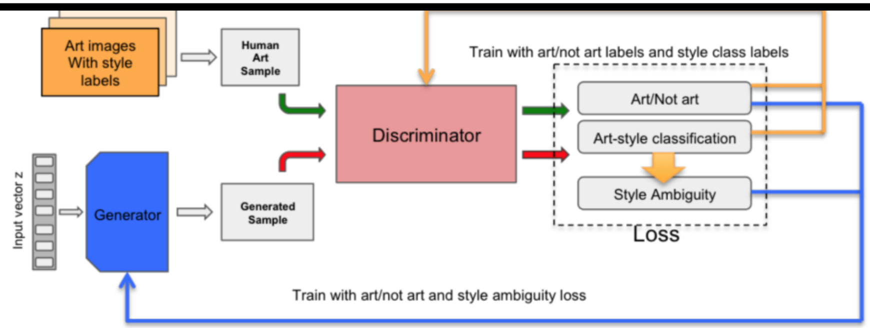
GAN



Activation and Loss



CAN



$$\begin{aligned}\min_G \max_D V(D,G) = \\ &\mathbb{E}_{x,\hat{c}\sim p_{data}}[\log D_r(x) + \log D_c(c = \hat{c}|x)] + \\ &\mathbb{E}_{z\sim p_z}[\log(1 - D_r(G(z))) - \sum_{k=1}^{K}(\frac{1}{K} log(D_c(c_k|G(z)) + \\ &(1 - \frac{1}{K}) log(1 - D_c(c_k|G(z)))]\end{aligned}$$

Generator

starting from $z \in \mathbb{R}^{100} \to 4 \times 4 \times 1024 \to 8 \times 8 \times 1024 \to 16 \times 16 \times 512 \to 32 \times 32 \times 256 \to 64 \times 64 \times 128 \to 128 \times 128 \times 64 \to 256 \times 256 \times 3$ (the generated image size).

Discriminator (stride 2 and 1 pixel padding)

- 6 convolution layer: 4 × 4 x 32 → 4 × 4 x 64 → 4 × 4 x 128 → 4 × 4 x 256 → 4 × 4 x 512 → 4 × 4x 512
- Final Feature map: 4 × 4 × 512
- Real/fake Dr head: dense layer to produce **Dr(c|x) - image comes from real distribution.**
- The multi-label head: 3 dense layers {1024, 512, K} (K is the number of style classes to produce **Dc(c_k|x) - multilabel probabilities**

GAN VS CAN

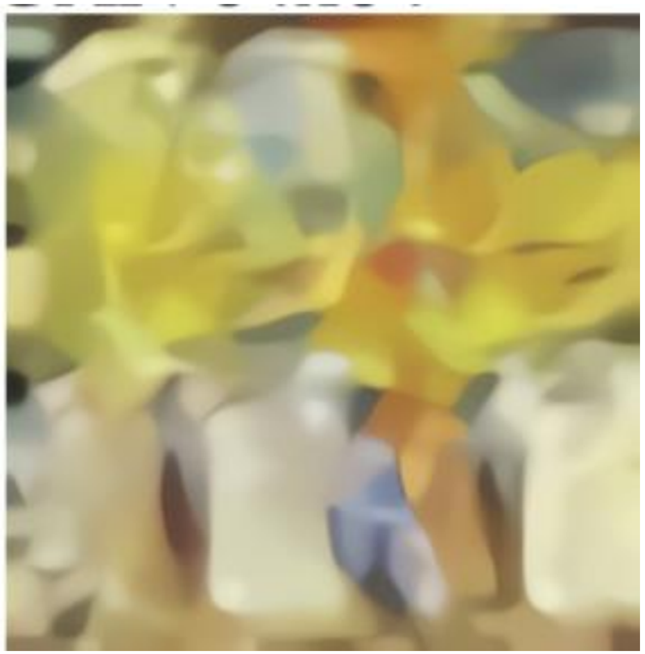
Transform an Emulative Generative Adversarial Network into a new *creative* one

Training CAN hyperparameters:

Weights zero-centered Normal distribution σ = 0.02. Mini-batch = 128 and mini-batch stochastic gradient descent (SGD) for training with eta = 0.0001. Activation function = LeakyReLU with slope = 0.2 in all model. Finally, they used data augumentation 5 crops within for each image

Qualitative Test

- Compared four models to assess novelty and creativeness
  - DCGAN (no emulation)
  - DCGAN + 2 more layers to resize images (improvement on colors and structures)
  - Style-Classification CAN (sc-CAN) (recognizible figures , religious subjects)
  - CAN (no properties of previous model but creates aestetically appealing images)



Quantitative Test

WIKIART
VISUAL ART ENCYCLOPEDIA

- 82k images from 1200 artist with 25 arstic styles used to train, generate and evaluate the images. Users (18 Mturk and History and art students judged comparisons between *actual real art* and the generated one).
  - Impressionist Dataset: 25 images
  - Art Basel: 25 images
  - DCGAN:
    - 100 images with resolution 64x64
    - 76 images with resolution 256x256
  - Sc-CAN: 100 images
  - CAN: 125
- t-test with nulll hypothesis

All images are up-sampled to 512x512

Means and standard deviations of responses of Experiment I

Painting set	Q1 (std)	Q2 (std)
CAN	53% (18%)†	3.2 (1.5) ‡
DCGAN [18] (64x64)	35% (15%) †	2.8 (0.54) ‡
Abstract Expressionist	85% (16%)	3.3 (0.43)
Art Basel 2016	41% (29%)	2.8 (0.68)
Artist sets combined	62% (32%)	3.1 (0.63)

All images are resized to 512x512 resolution
† Q1 t-test (CAN vs. DCGAN) p-value = $1.9932e-15$
‡ Q2 t-test (CAN vs. DCGAN) p-value = $9.3634e-06$

Means and standard deviations of responses of Experiment II

Image set	Q1 (std) Likeness	Q2 (std) Novelty	Q3 (std) Surprising	Q4 (std) Ambiguity	Q5 (std) Complexity	Q6 (std) human/computer
DCGAN [18] (256x256)	3.23 (0.53)	3.08 (0.50)	3.21 (0.59)	3.37 (0.48)	3.18 (0.63)	0.65 (0.17)
CAN	3.30 (0.43)	3.27 (0.44)	3.13 (0.46)	3.54 (0.45)	3.34 (0.50)	0.75 (0.14)
Abstract Expresionist	3.38 (0.43)	3.03 (0.38)	2.95 (0.50)	3.17 (0.35)	2.90 (0.35)	0.85 (0.11)
Art Basel 2016	2.95 (0.70)	2.69 (0.59)	2.36 (0.66)	2.79 (0.59)	2.46 (0.68)	0.48 (0.23)

Berlyne theory

CAN is creative!

Means and standard deviations of the responses of Experiment III

Painting set	Q1 (std) Intentionality	Q2 (std) Visual Structure	Q3 (std) Communication	Q4 (std) Inspiration
CAN	3.3 (0.47)	3.2 (0.47)	2.7 (0.46)	2.5 (0.41)
Abstract Expressionist	2.8 (0.43)	2.6 (0.35)	2.4 (0.41)	2.3 (0.27)
Art Basel 2016	2.5 (0.72)	2.4 (0.64)	2.1 (0.59)	1.9(0.54)
Artist sets combined	2.7 (0.6)	2.5 (0.52)	2.2 (0.54)	2.1 (0.45)

(Experiment 4) - Sc-CAN vs CAN for novel and aestetically appealing
59.47% subjects selected CAN images as more novel, and 60% the found CAN images more aestetically appealing ⇒ Style ambiguity loss has real effect on new generated images

Final Consideration

Is this machine really creative?

- Colton theory related to CAN
- Tests were good, but not enough people to assess creativeness
- What means creative for us?
- Creative for economics purposes

- Example of CAN (Phillip Kravtsov and Phillip Kuznetsov) on https://github.com/mlberkeley/Creative-Adversarial-Networks
  Downgrade tensorflow version to 1.X. If not, you can not use *Contrib lib* from *tensorflow*

DCGAN image sold for $423.520

256x256

128x128

CAN - Github