



Moreno La Quatra
PhD Student @ DAUIN



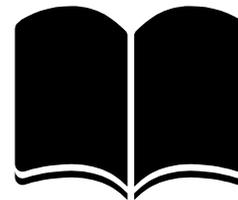
Generative Adversarial Networks

Beyond discriminative models



Afternoon walk through the alleys of Ortygia in Siracusa Sicily.
Joanne Hastie original 2019 - CycleGAN
<https://joannehastie.com/product/sicilian-alleyway/>

Outline



- Discriminate or generate
- Short story: Gen & Dis
- Introduction to GANs
- Tasks in Computer Vision
- Demo
- Conclusions

Generative

VS

Discriminative

1.

Generative models

A generative model describes how a dataset is generated, in terms of a probabilistic model.

By sampling from this model, we are able to generate new data.

Generative deep learning ([David Foster](#) - O'Reilly)



<https://thispersondoesnotexist.com>

Francis Picabia Paintings

Discriminative

We could train a discriminative model to predict if a given painting was painted by Francis Picabia.



Generative

We could train a *generative* model to produce paintings that seem to be drawn by Francis Picabia himself.



What can we do?

Discriminative

$$p(y|X = x)$$



What would you like to know?

Generative

$$p(x|Y = y)$$





Generative Adversarial Networks is the most interesting idea in the last ten years in machine learning.

- Yann LeCun, Director, Facebook AI (Turing Prize)

2.

Mr. Gen

&

Prof. Dis

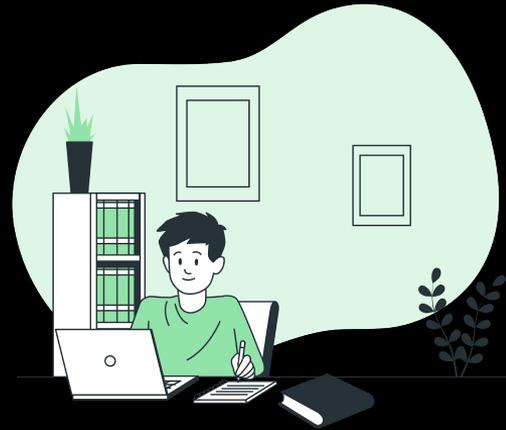
Studying for DSL Exam

DSL Exam

This is Gen, a PoliTO student



Gen studies for his DSL exam.



DSL Exam

Time for exam arrived. **Gen** answers to questions creating his output for the exam.



Prof. **Dis** examines the output produced by Gen and assign it a score.



DSL Exam

Is this good enough?



Unfortunately not. **Gen** need to train harder to get a higher score.



DSL Exam

Second chance, now **Gen** knows what's wrong with his exam!



Again, Prof. **Dis** analyzes the exam and assign it a score.

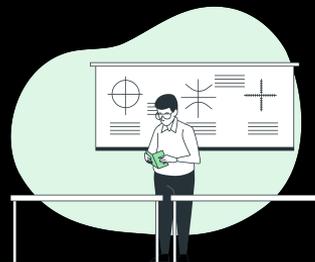


Studying = Training

Gen is a generative model.
It is trained to produce more
and more accurate results.



Prof. **Dis** is a discriminative model.
It evaluates samples.



[Training Phase]

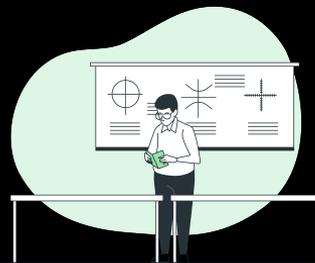
Depending on the results of the exam, Gen adjust the
competencies to have better grades.

Studying = Training

Gen is a generative model.
It is trained to produce more
and more accurate results.



Prof. **Dis** is a discriminative model.
It evaluates samples.



[Evaluation Phase]

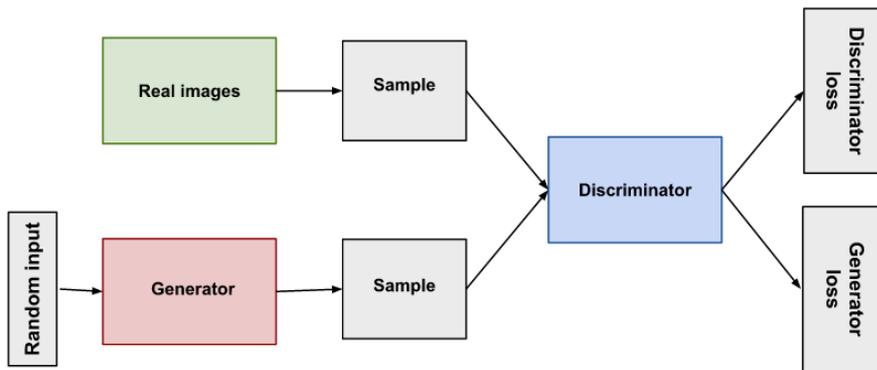
Analyzing the exam solution and using his previous
experience, Prof. Dis can evaluate the exam.

3.

**GAN model
Architecture**

GAN Architecture

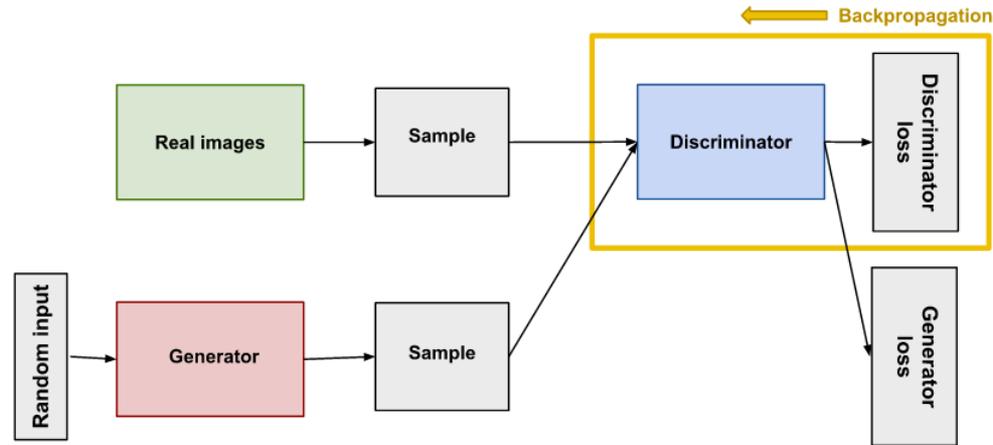
Similarly to the characters of our story, GANs have two main components.



Generator: learns to create data by incorporating feedback from the discriminator.

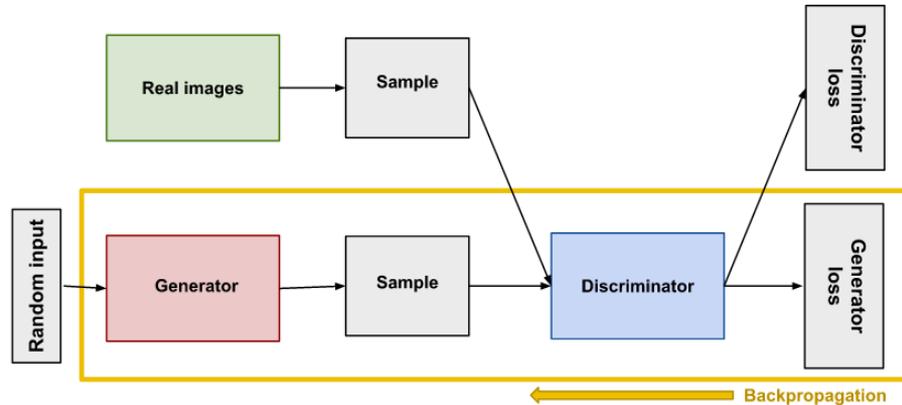
Discriminator: tries to distinguish real data from the data created by the generator.

Discriminator (Prof. Dis)



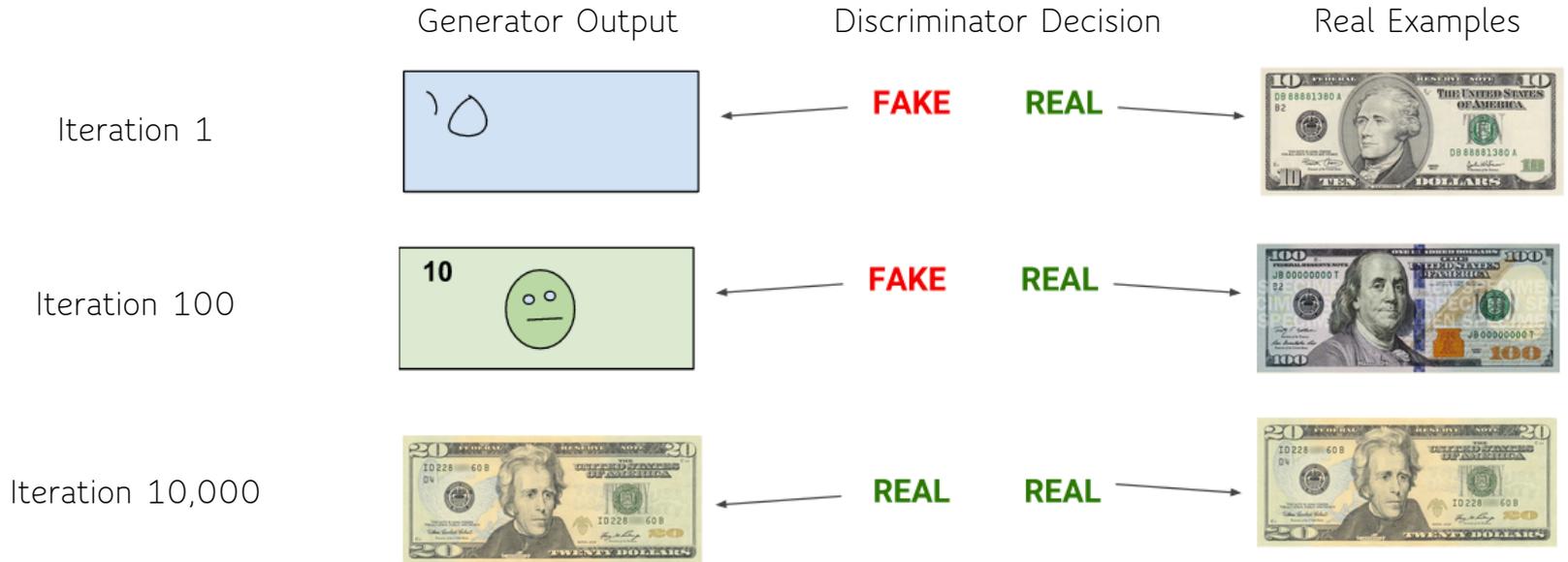
- It is trained to predict **False** on fake examples and **True** for real ones.
- It uses both real and fake examples and classify them using knowledge.
- Better the discriminator, harder the task for the generator.

Generator (Mr. Gen)



- Aims at fooling the discriminator (let it classify generated images as real).
- Uses random noise as input.
- The discriminator give it **feedbacks** by using the loss function.

Training Process

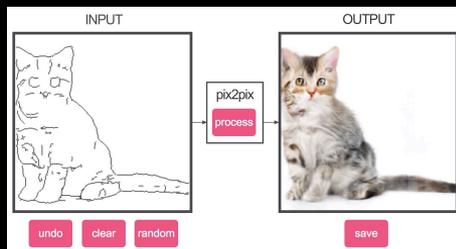


4.

**GAN Tasks in
Computer Vision**

Computer vision tasks

Image-to-image
translation



Super resolution



Semantic Image
Synthesis



Image-to-image translation

GANs take an image as input and map it to a generated output image with different properties.



<https://affinelayer.com/pixsrv/>

Super-resolution

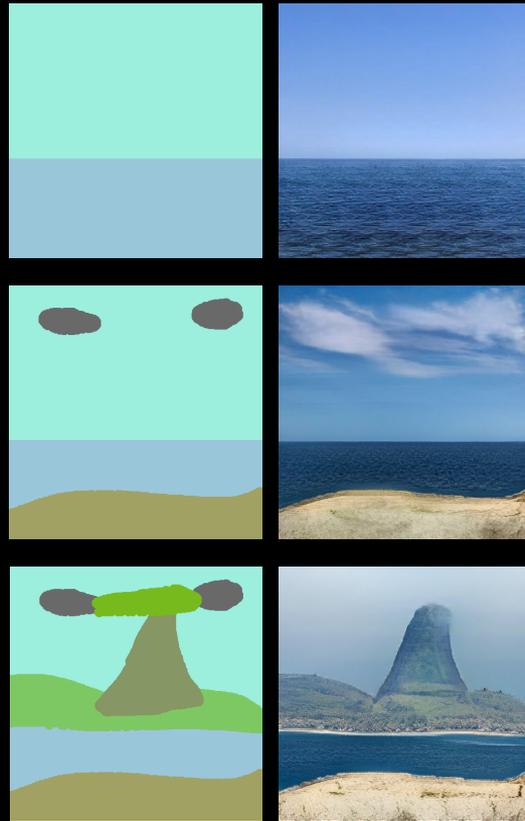
GANs increase the resolution of images, adding detail where necessary to fill in blurry areas.



<https://deepai.org/machine-learning-model/torch-srgan>

Semantic Image Synthesis

GANs take an image as input and map it to a generated output image with different properties.



<http://nvidia-research-mingyu-liu.com/gaugan/>

Demo Time

How to generate people faces sampling from the **human** distribution.



https://colab.research.google.com/drive/1b0qTrP8_jBkay8u6oE610IAneCa15_2f?usp=sharing

Final Remarks

GANs can be used in **other fields** (NLP, Financial data, ...)

GANs are very un-optimized on **data usage**.

In the next future, it will be harder for humans to **distinguish** fake or real data.

GANs can be used for good but also for **illegal/evil** purposes.

It is relatively simple to find **blind spots** in the generator.

References

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems* (pp. 2672-2680).

Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1125-1134).

Park, T., Liu, M. Y., Wang, T. C., & Zhu, J. Y. (2019). Semantic image synthesis with spatially-adaptive normalization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2337-2346).

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4681-4690).

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Agarwal, S. (2020). Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*.

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and improving the image quality of stylegan. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8110-8119).

Thanks!



Any question? I'll try to generate an answer for them!

Moreno La Quatra

PhD Student @ DAUIN

moreno.laquatra@polito.it