# State of art of generative models in robotics applications

Facoltà di Ingegneria dell'informazione, informatica e statistica

SAPIENZA
UNIVERSITÀ DI ROMA

Master thesis in engineering in computer science

Supervisor: **Prof. Daniele Nardi**

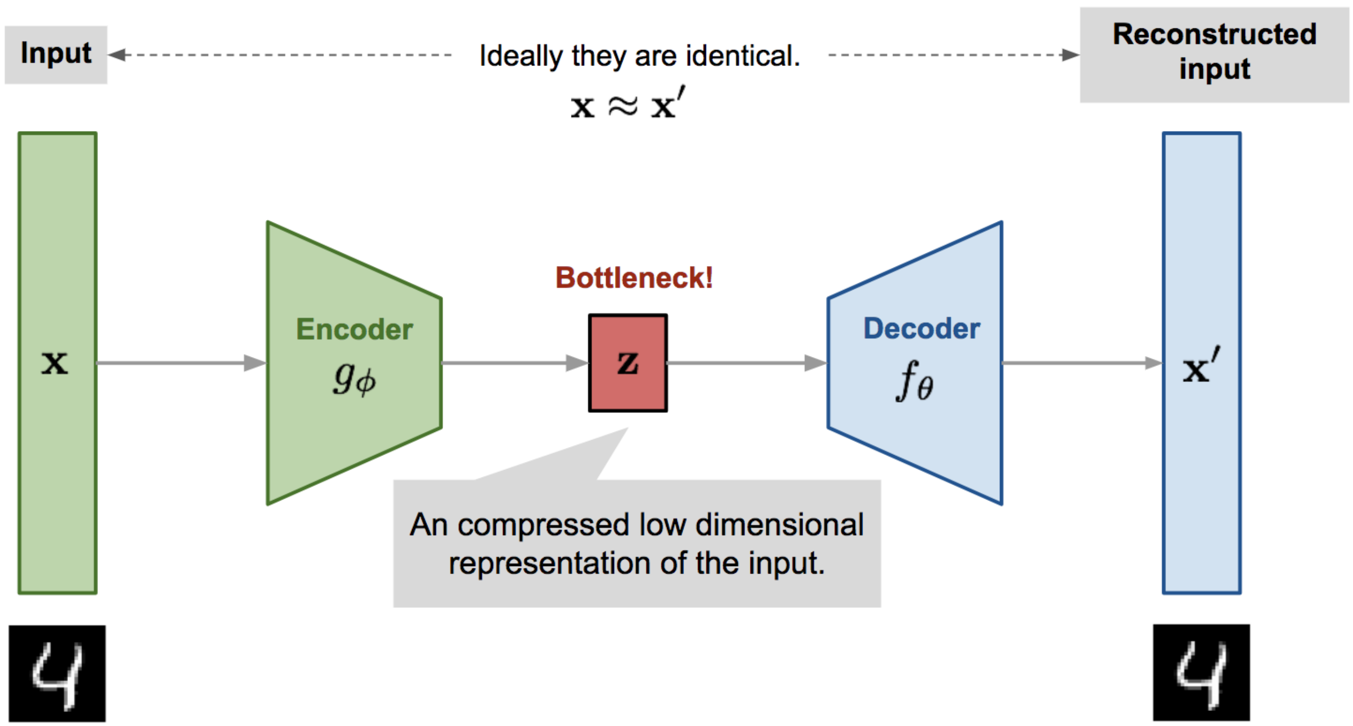
Correlator: **Ing. Francesco Riccio**

Candidate:

**Milad Kiwan**

Academic year 2019-2020

# Variational Autoencoders VAEs



# Variational Autoencoder VAE

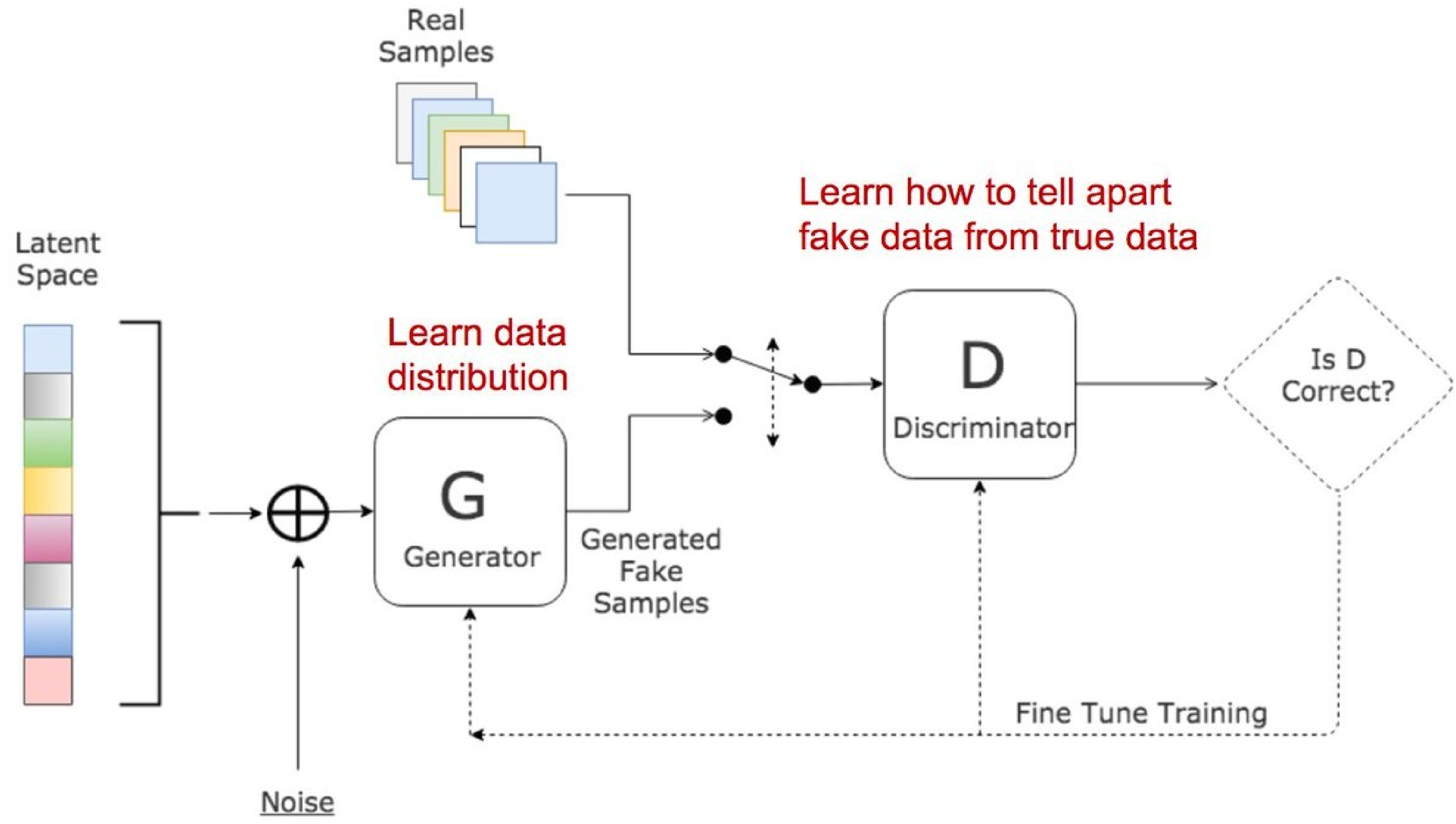
***Encoder:***

- Find latent space.
- Low dimensional representation of the original data.

***Decoder:***

- Given the latent or the data representation returned  from the encoder, generates new data.
- Reconstruct the original data.

# Generative Adversarial networks



# Generative Adversarial Networks GANs

- Distinguish between authentic and fake data.
- Discriminator loss function:

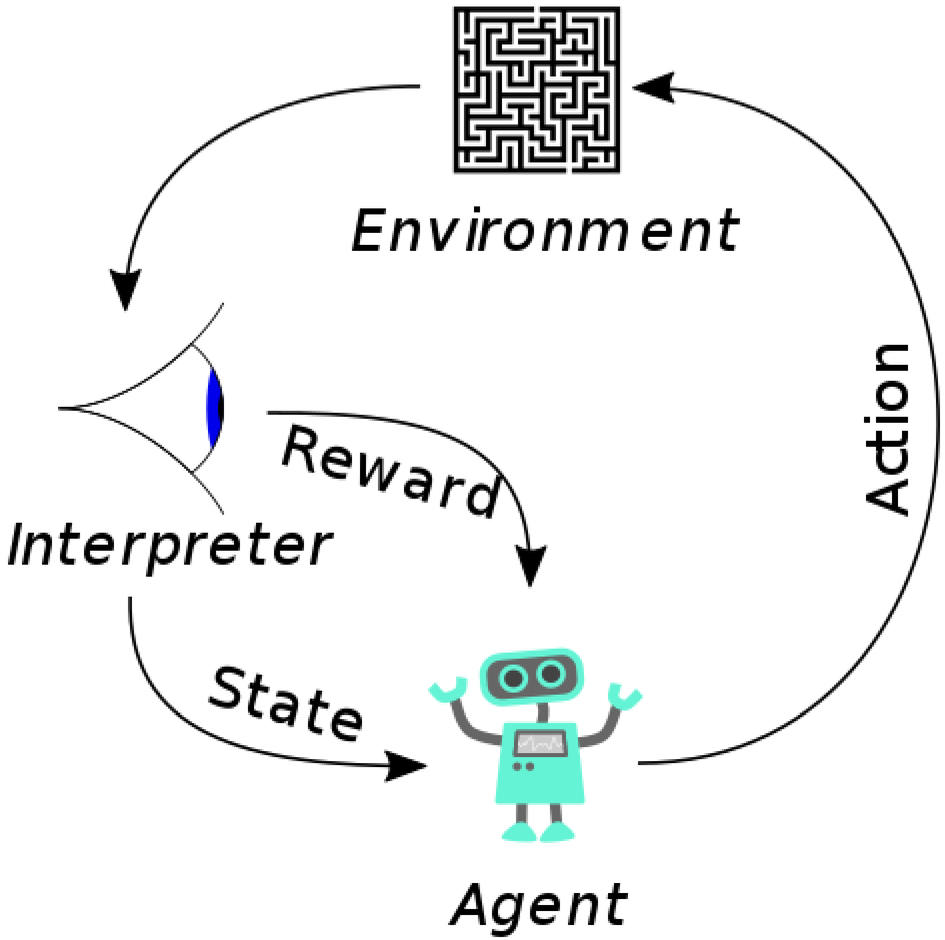
$$E_x[\log(D(x))] + E_z[\log(1 - D(G(z)))]$$

- Generate new data.
- Generator loss function:

$$E_z[\log(1 - D(G(z)))]$$

- Find saddle point

# Reinforcement Learning



# Reinforcement Learning RL

Basically is modeled as a Markov decision process:

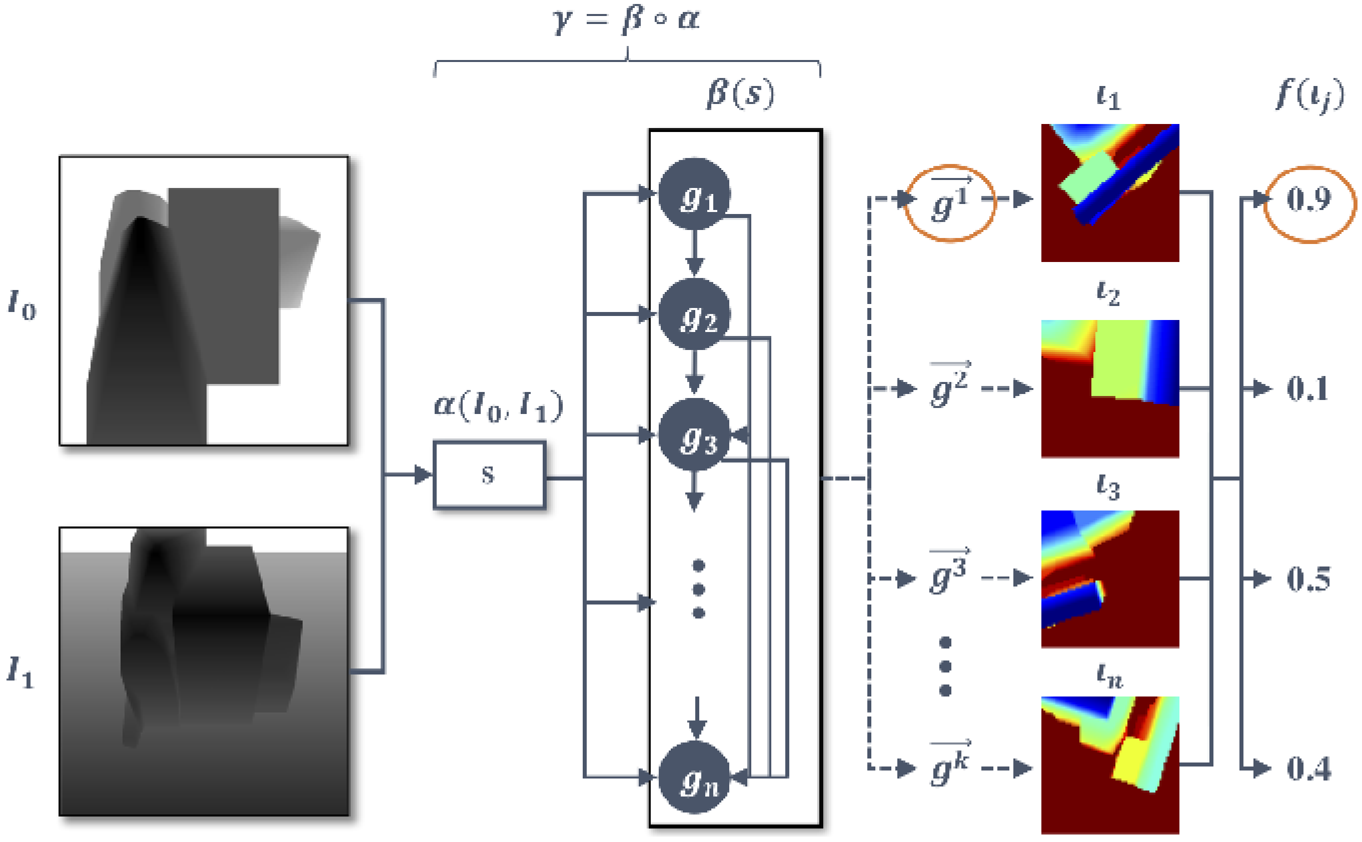
$$(S, A, P, R, \gamma)$$

- S is a set of environment a agent states.
- A is a set of actions the agent can perform.
- P is the transition probability between the states.
- R is the  immediate reward after transition from s to s' with action a.
- $\gamma$ is a discount factor, to discount the future reward.

# Domain randomization and generative models for robotic grasping

- Perform the grasp planning for randomized objects domain.
- Objects generation was achieved by placing random meshes sequentially so that each mesh intersects with at least one of the preceding meshes.
- The first part is the grasp planning module that outputs is a probability distribution over possible grasps.
- The second part is grasp evaluation model that outputs a value corresponding to the likelihood of success of that grasp.

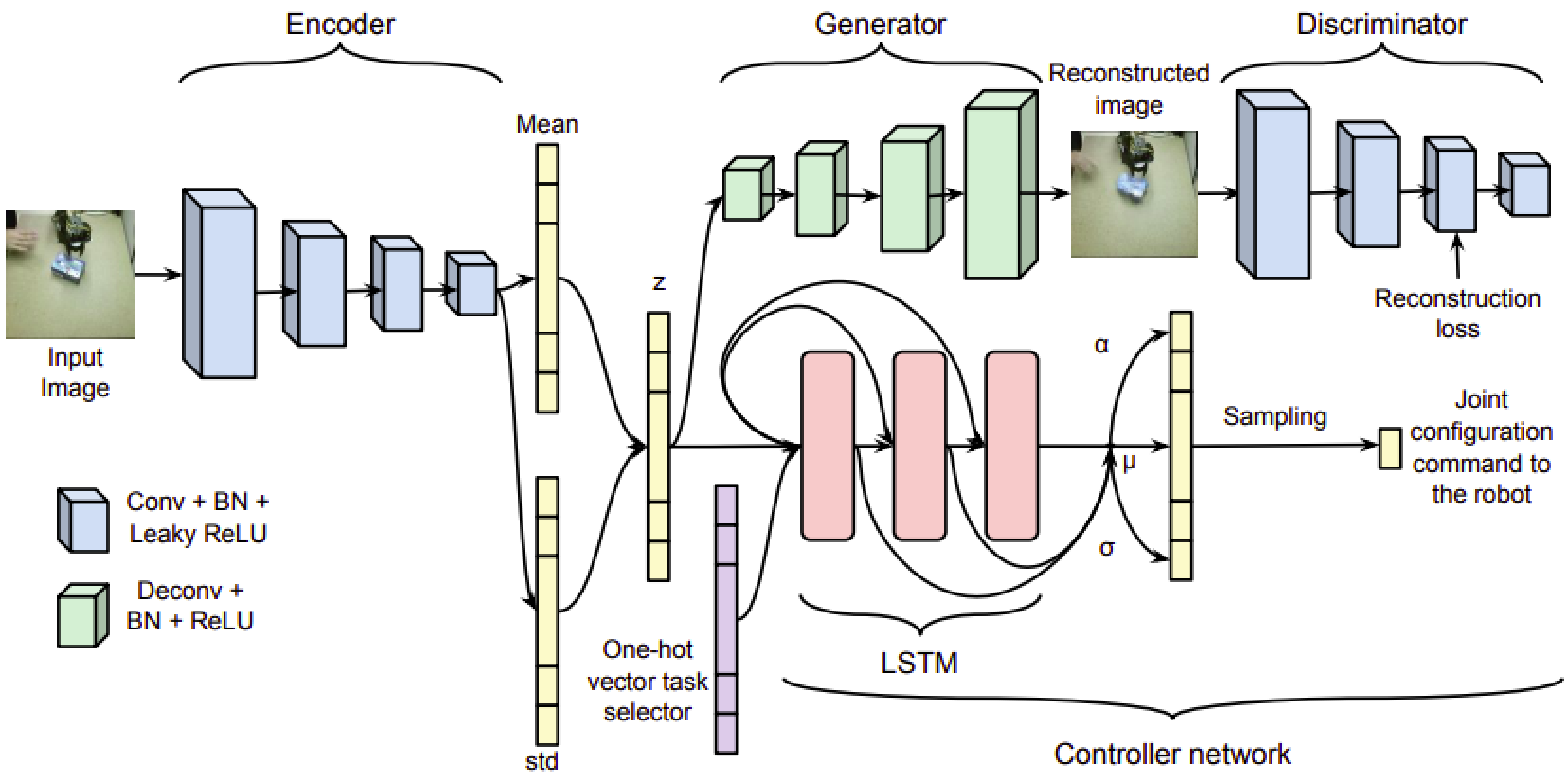
# Domain randomization and generative models for robotic grasping



## Vision-Based Multi-Task Manipulation for Inexpensive Robots Using End-To-End Learning from Demonstration

- The encoder finds the distribution of the data and then sample a latent representation.
- The decoder is a GAN.
- The policy controller that interact with the environment is an LSTM.
  - The output of the LSTM is a GMM.
- The approach does not require a reward signal as RL.

# Vision-Based Multi-Task Manipulation for Inexpensive Robots Using End-To-End Learning from Demonstration



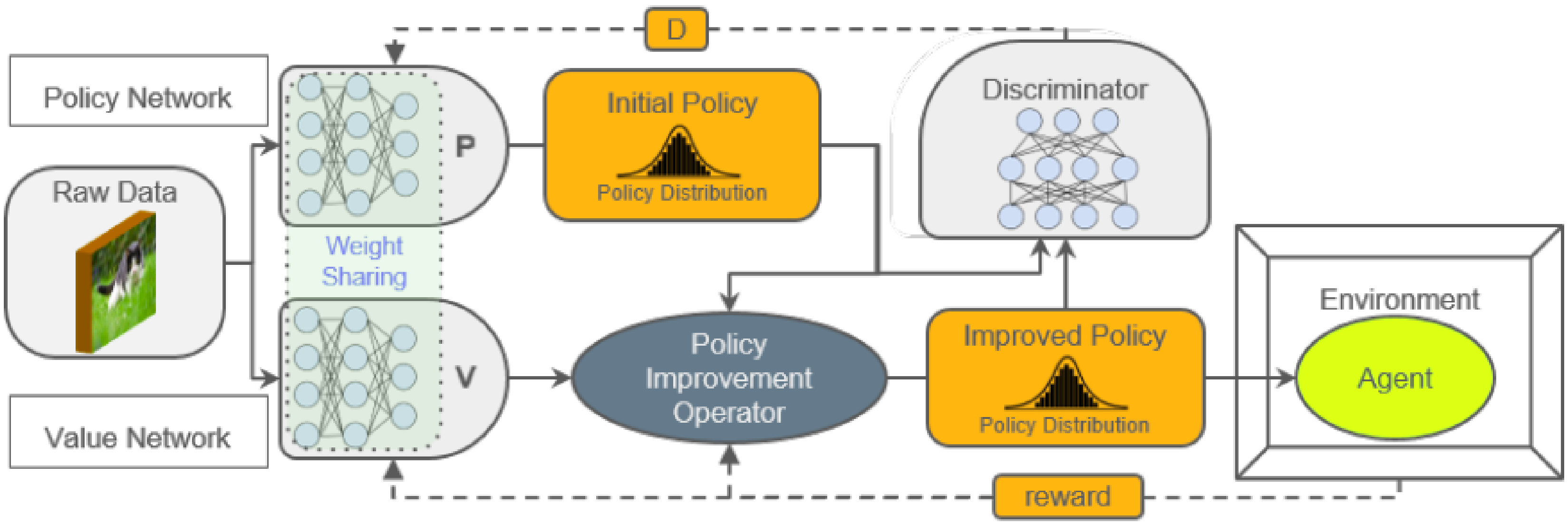
# Generative Adversarial Imitation Learning (GAIL)

- GAIL is a combination between GAN and RL.
- The generator is the policy to be learned, and follows TRPO algorithm.
- discriminator distinguishes between the expert trajectories and those generated by the generator.
- Again there is no need to design a reward function.

## Self-Improving Generative Adversarial Reinforcement Learning

- A network to generate initial policies.
- A value function network to approximate how good is the current state.
- Policy improvement black box (TRPO).
- The improved policy interact with the environment.
- A discriminator represents a classifier and assigns a low loss to the improved policies and a high loss to the initial ones.

# Self-Improving Generative Adversarial Reinforcement Learning



# Discussion

- **Cons**
- Generative models are hard to train.
- Generally requires a large amount of data.
- VAEs explore places that does not always make sense.
- **Pros**
- Once the saddle point is found, GAN achieve a very good results.
- VAEs are the best option to move to latent space.

# Conclusion

- Work in latent space using VAEs.
- The introduction of GANs in RL transformed the actor-critic approach of reinforcement learning to generative adversarial network.
- Avoiding reward function, led to achieve multi-task frameworks for robotics.
- Private robots able to learn and perform multi-tasks.


learning
policy
data
used
model
can
models
training
space
robotics
GAN
VAE
GANs
latent
function
generate
Generative
framework
network
probability
tasks
learn
state
like
one
goal
environment
reward
networks
algorithms
adversarial
applications
VAEs
input
objects
object
expert
different
parameters
structure
architecture
demonstrations
generated
dataset
Robotics
achieve
GMM
encoder
task
make
process
neural
images
train
point
using
well
robot
problem


# Thank you!