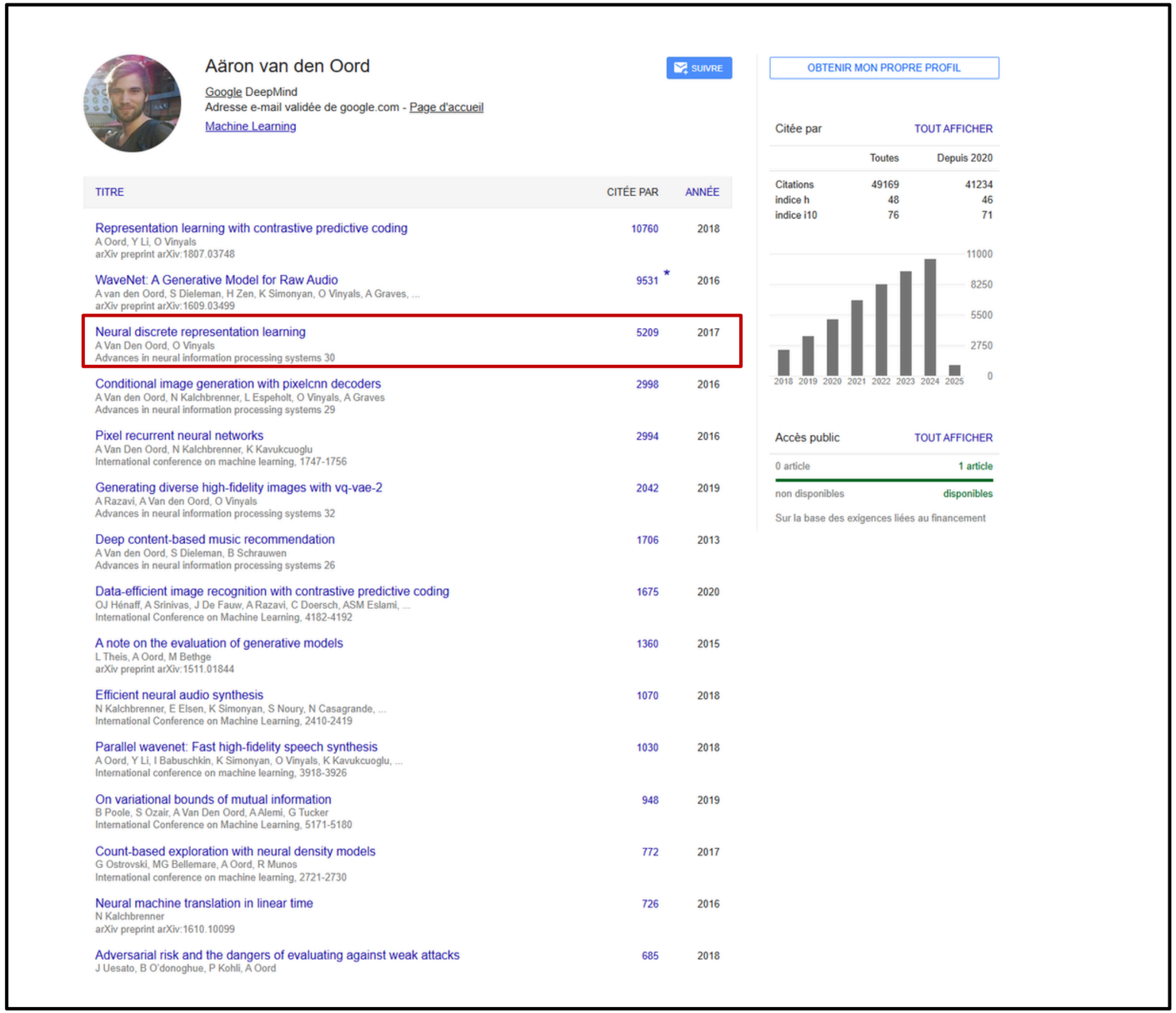
BrAIn Lab.
Brain and Artificial Intelligence Lab.

Neural Discrete Representation Learning

(NeurIPS-2017)

0.

BrAIn Lab.
Brain and Artificial Intelligence Lab.



Aäron van den Oord

SUIVRE

Google DeepMind
Adresse e-mail validée de google.com - Page d'accueil
Machine Learning

OBTENIR MON PROPRE PROFIL

TITRE	CITÉE PAR	ANNÉE
Representation learning with contrastive predictive coding A Oord, Y Li, O Vinyals arXiv preprint arXiv:1807.03748	10760	2018
WaveNet: A Generative Model for Raw Audio A van den Oord, S Dieleman, H Zen, K Simonyan, O Vinyals, A Graves, ... arXiv preprint arXiv:1609.03499	9531 *	2016
Neural discrete representation learning A Van Den Oord, O Vinyals Advances in neural information processing systems 30	5209	2017
Conditional image generation with pixelcnn decoders A Van den Oord, N Kalchbrenner, L Espeholt, O Vinyals, A Graves Advances in neural information processing systems 29	2998	2016
Pixel recurrent neural networks A Van Den Oord, N Kalchbrenner, K Kavukcuoglu International conference on machine learning, 1747-1756	2994	2016
Generating diverse high-fidelity images with vq-vae-2 A Razavi, A Van den Oord, O Vinyals Advances in neural information processing systems 32	2042	2019
Deep content-based music recommendation A Van den Oord, S Dieleman, B Schrauwen Advances in neural information processing systems 26	1706	2013
Data-efficient image recognition with contrastive predictive coding OJ Hénaff, A Srinivas, J De Fauw, A Razavi, C Doersch, ASM Eslami, ... International Conference on Machine Learning, 4182-4192	1675	2020
A note on the evaluation of generative models L Theis, A Oord, M Bethge arXiv preprint arXiv:1511.01844	1360	2015
Efficient neural audio synthesis N Kalchbrenner, E Elsen, K Simonyan, S Noury, N Casagrande, ... International Conference on Machine Learning, 2410-2419	1070	2018
Parallel wavenet: Fast high-fidelity speech synthesis A Oord, Y Li, I Babuschkin, K Simonyan, O Vinyals, K Kavukcuoglu, ... International conference on machine learning, 3918-3926	1030	2018
On variational bounds of mutual information B Poole, S Ozair, A Van Den Oord, A Alemi, G Tucker International Conference on Machine Learning, 5171-5180	948	2019
Count-based exploration with neural density models G Ostrovski, MG Bellemare, A Oord, R Munos International conference on machine learning, 2721-2730	772	2017
Neural machine translation in linear time N Kalchbrenner arXiv preprint arXiv:1610.10099	726	2016
Adversarial risk and the dangers of evaluating against weak attacks J Uesato, B O'donoghue, P Kohli, A Oord	685	2018

Citée par — TOUT AFFICHER

	Toutes	Depuis 2020
Citations	49169	41234
indice h	48	46
indice i10	76	71

Accès public — TOUT AFFICHER

0 article	1 article
non disponibles	disponibles

Sur la base des exigences liées au financement

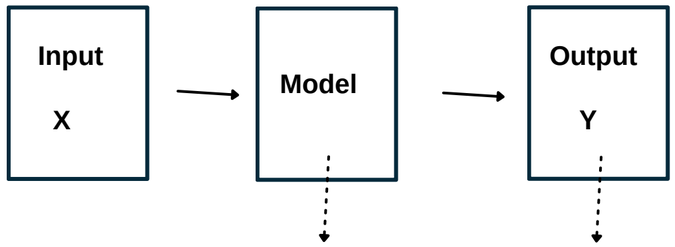
0. Deep Learning

BrAIn Lab.
Brain and Artificial Intelligence Lab.

Deep Learing ?

-> Probabilistic Inference Machine

Why?



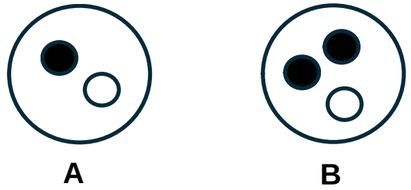
서로 다른 정도

Loss
- Cross Entropy(CE)
- MSE
- Etc…

MLE(Maximum Likelihood Estimation)

- Likelihood

“결과를 바탕으로, 그 결과를 만드는 조건을 찾음”



MLE : **결과를 최대화 하는 조건을 선택하는 방법**

MAP(Maximum A Posterior)

“사후 분포 p(z|x) 를 최대화 하는 x값을 추정 "

- Cross entropy

두 확률분포 P,Q의 차이를 계산

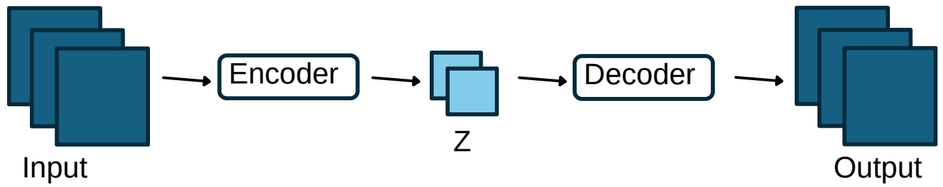
- Kullback-Leibler divergence(KDL)

- Jenson-Shannon divergence(JSD)

1. AutoEncoder

BrAIn Lab.
Brain and Artificial Intelligence Lab.

AutoEncoder



Where is it used?

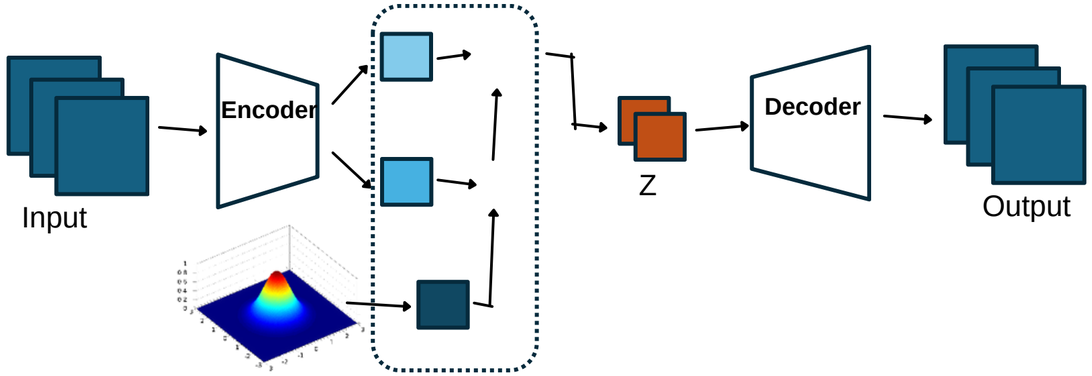
- Image generation
- **Denoise**
- **Anommaly Detection**
- **Deimension Reduction**

-> **목적에 따라서 중점을 두는 부분이 다르다**

Loss

-> Reconstruction Error + Etc..

VAE(Variational AutoEncoder)



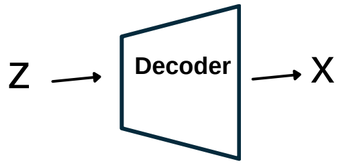
Loss

Loss =

“Reconstruction Error”

“Regularzation”

VAE

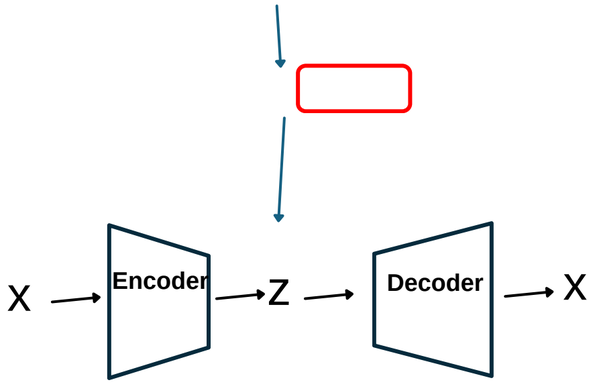


"Intractible"

Posterior 로 계산? ->

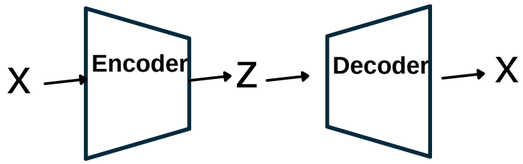
"알수 없다."

- **Decoder 학습을 위하여 Encoder를 도입**



2. VQ-VAE

VAE-Loss



대한 (log)likelihood 적용

=

||)

“Variational lower bound (ELBO)”

“Intractible” ,

“Reconstruction Error”

“Regularzation”

2. VQ-VAE

VQ-VAE

- **AutoEncoder 기반의 생성 모델**
- **Encoder 의 output이 discrete 함**
- **VAE의 “Posterior collapse” 문제를 해결**

What is Posterior Collapse?

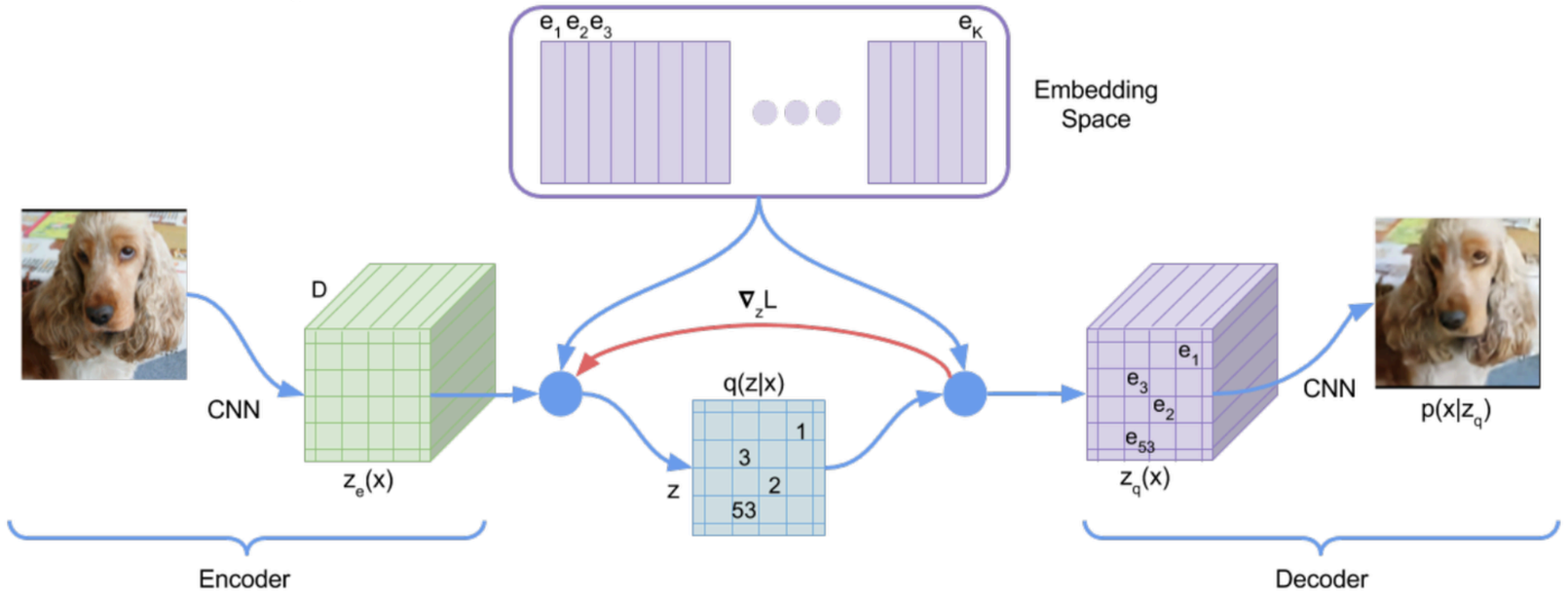
- **Decoder의 영향이 너무 커서 학습중에 KL항을 과도하게 줄이는 방향으로 업데이트 되는 것**

→

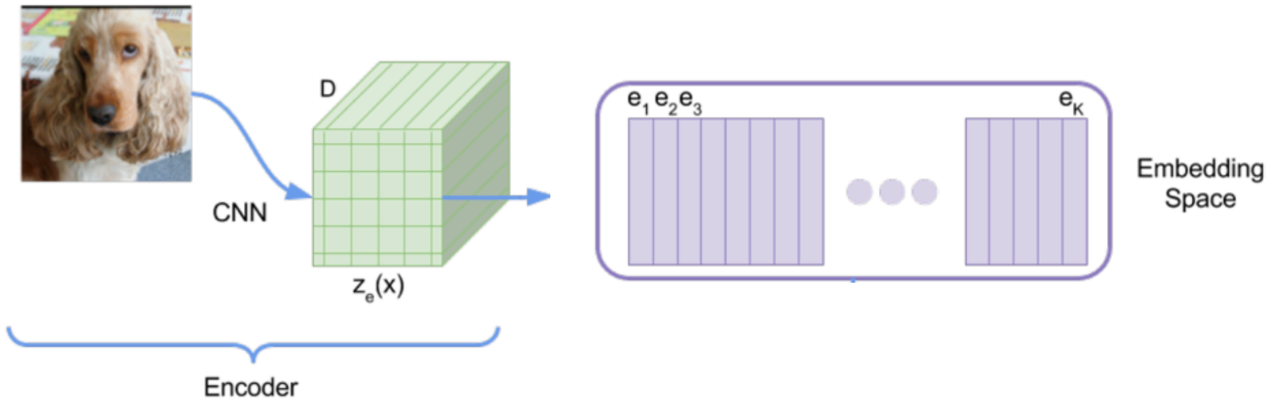
- **입력 x에 무관하게 사전분포 p(z) 에 붕괴(collapse) 된다.**
- **결론적으로 인코더는 데이터 x를 반영하지 않게되고 z는 거의 동일해진다.**

2. VQ-VAE

Overview of VQ-VAE

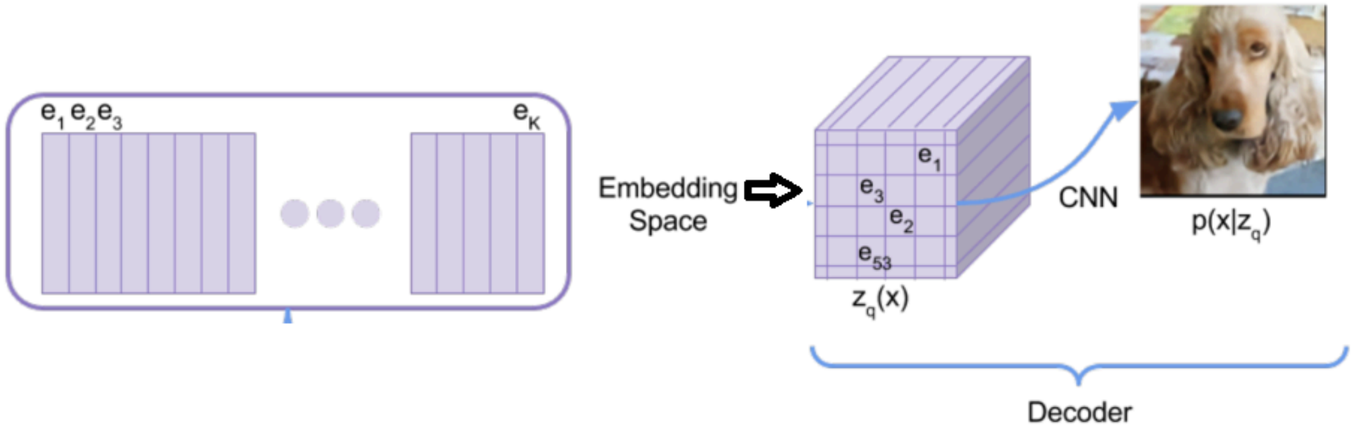


Encoder of VQ-VAE



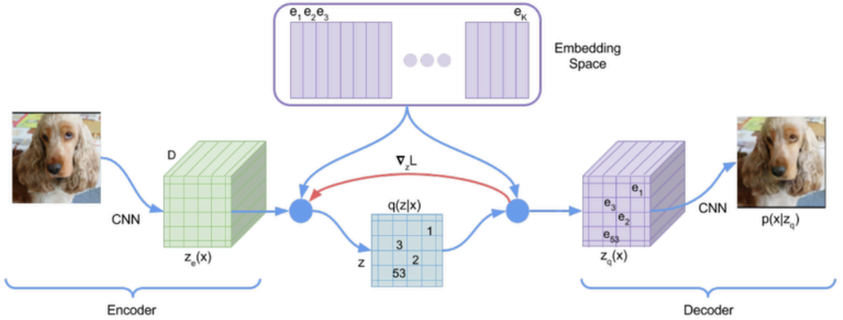
1. **Input x 가 Encoder를 통과하여 성**
2. **: 코드북 벡터와 인코더의 결과를 유클리드 거리 계산**
3. **가장 가까운 벡터를 코드북에서 선택 :**
4. **포는 one-hot 분포**
5. **(KL항은 상수)**

Decoder of VQ-VAE



1. **Quntization vector 가 decoder의 입력으로 들어감**

Loss of VQ-VAE



"Reconstruction" "Code book" "Commitment"

1. Reconstruction

- 원본을 재구성
- Straight through-Gradient 로 gradient 복사
- 따라서 코드북은 디코더의 gradient 를 받지 않는다.

2. Code book

- Code book을 업데이트
- 임베딩 벡터를 l2 error를 이용하여 업데이트 한다 (SG: Stop-Gradient)
- 인코더의 출력과 embedding vector를 가까워지게 한다.

3. Commitment

- Encoder를 업데이트
- 임베딩 l2 error를 이용하여 업데이트 한다 (SG: Stop-Gradient)
- Embedding vector와 인코더의 출력을 가까워지게 한다.

→ 한쪽으로 치우치는게 아닌 서로 당기는 방향으로 업데이트

Result

1. Image reconstruction

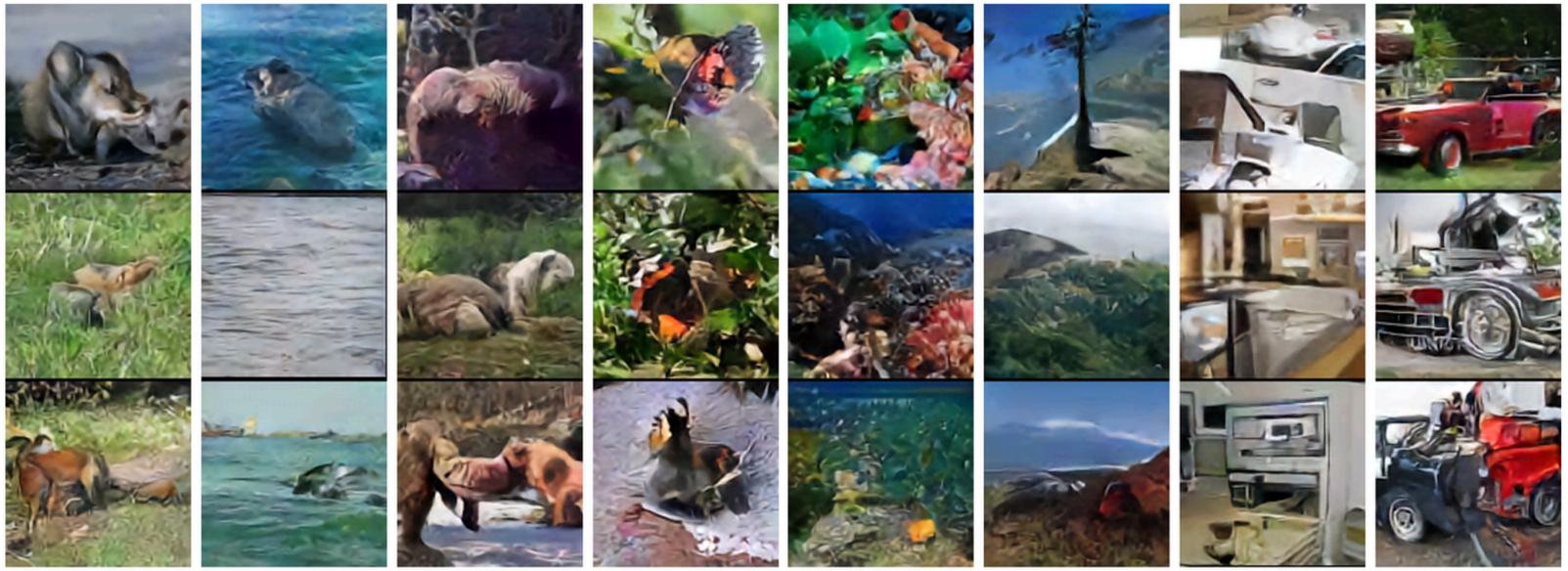
Imagenet images

Imagenet image reconstruction

-

Result

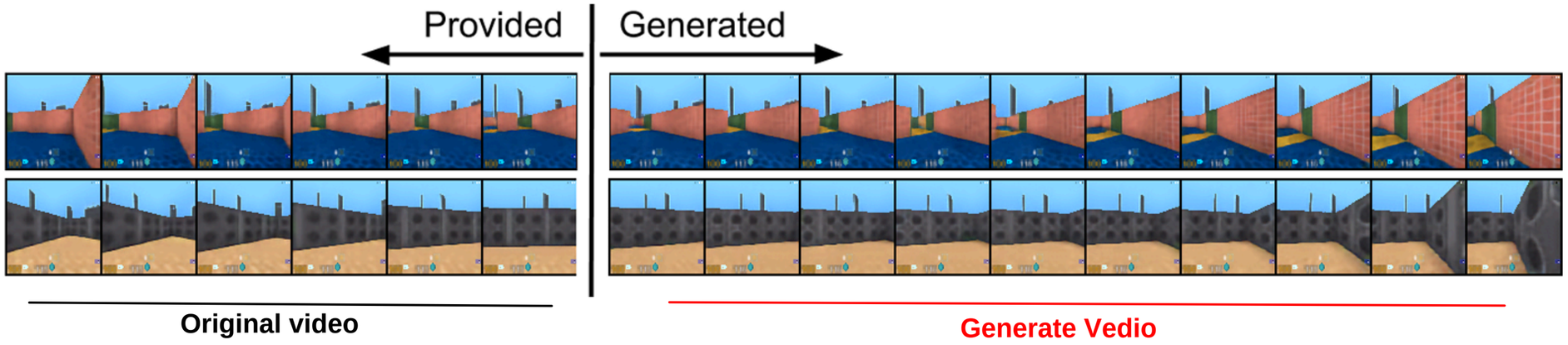
2. Image generation



fox	Gray whale	Brown bear	butterfly	Coral reef	alp	micro wave	picup

Result

3. Video generation



- **Top : “move forward”**
- **Bottom : “move right”**