

Representation learning

발표자 권선우



**SOGANG
UNIVERSITY**

Index

- Representation learning
- Representation example
 - Greedy layer-wise unsupervised pretraining
 - Transfer learning
 - Domain adaptive
 - DCGAN

Representation learning

Feature learning

From Wikipedia, the free encyclopedia

In [machine learning](#), **feature learning** or **representation learning**^[1] is a set of techniques that allows a system to automatically discover the representations needed for [feature detection](#) or [classification](#) from raw data. This replaces manual [feature engineering](#) and allows a machine to both learn the features and use them to perform a specific task.

Feature learning is motivated by the fact that machine learning tasks such as [classification](#) often require input that is mathematically and computationally convenient to process. However, real-world data such as images, video, and sensor data has not yielded to attempts to algorithmically define specific features. An alternative is to discover such features or representations through examination, without relying on explicit algorithms.

Feature learning can be either supervised or unsupervised.

- In supervised feature learning, features are learned using labeled input data. Examples include supervised neural networks, multilayer perceptron and (supervised) dictionary learning.
- In unsupervised feature learning, features are learned with unlabeled input data. Examples include dictionary learning, independent component analysis, autoencoders, matrix factorization^[2] and various forms of clustering.^{[3][4][5]}

Contents [hide]

- 1 Supervised
 - 1.1 Supervised dictionary learning
 - 1.2 Neural networks
- 2 Unsupervised
 - 2.1 *K*-means clustering
 - 2.2 Principal component analysis
 - 2.3 Local linear embedding
 - 2.4 Independent component analysis
 - 2.5 Unsupervised dictionary learning
- 3 Multilayer/deep architectures
 - 3.1 Restricted Boltzmann machine
 - 3.2 Autoencoder
- 4 See also

- Representation learning은 시스템이 필요로 하는 표현을 자동으로 발견하도록 하는 기법
- 시스템은 수학적으로 편리한 입력을 요구하는 것에 동기가 있다.

Representation learning

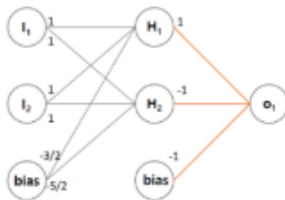
- 간단한 예시
 - 로마숫자 CCL을 VI로 나누기
 - > 250을 6으로 나누기

'나누기' 시스템을 동작 시킬 때 장제법을 적용하기
편한 아라비아 숫자가 '나누기' 시스템에 있어서 좋은
표현이다.

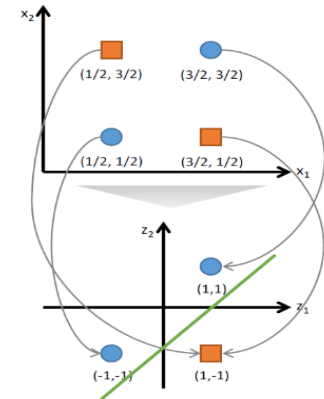
6 $\overline{) 250}$

- XOR 문제를 갖는 data로 linear classification
 - > XOR 문제를 representation learning 으로 제거한 data로 linear classification

$$o_1 = z_1 - z_2 - 1, \quad z = g(o_1)$$



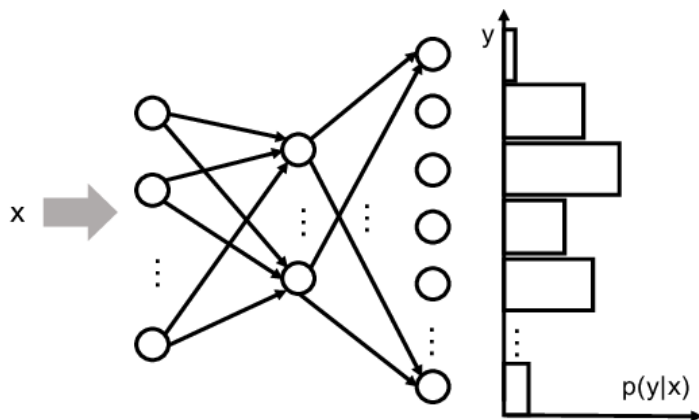
	x_1	x_2	h_1	z_1	h_2	z_2	o_1	z
●	1/2	1/2	-1/2	-1	-3/2	-1	-1	-1
■	3/2	1/2	1/2	1	-1/2	-1	1	1
■	1/2	3/2	1/2	1	-1/2	-1	1	1
●	3/2	3/2	3/2	1	1/2	1	-1	-1



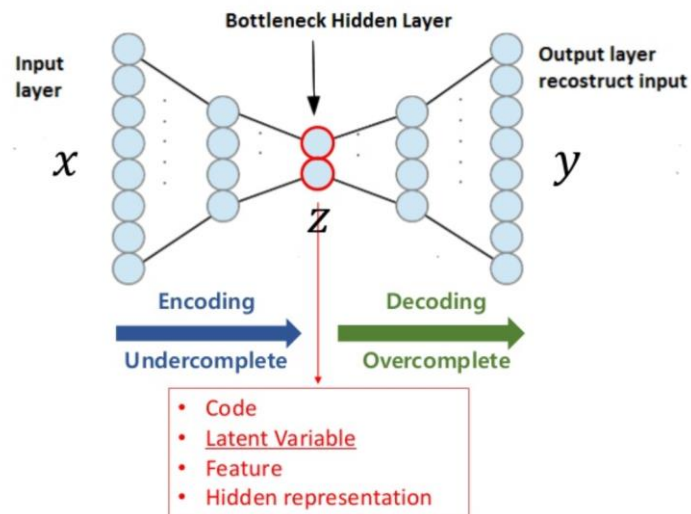
$(x_1, x_2) \rightarrow (z_1, z_2)$ domain 변경

Representation learning

- 간단한 예시



Softmax regression network



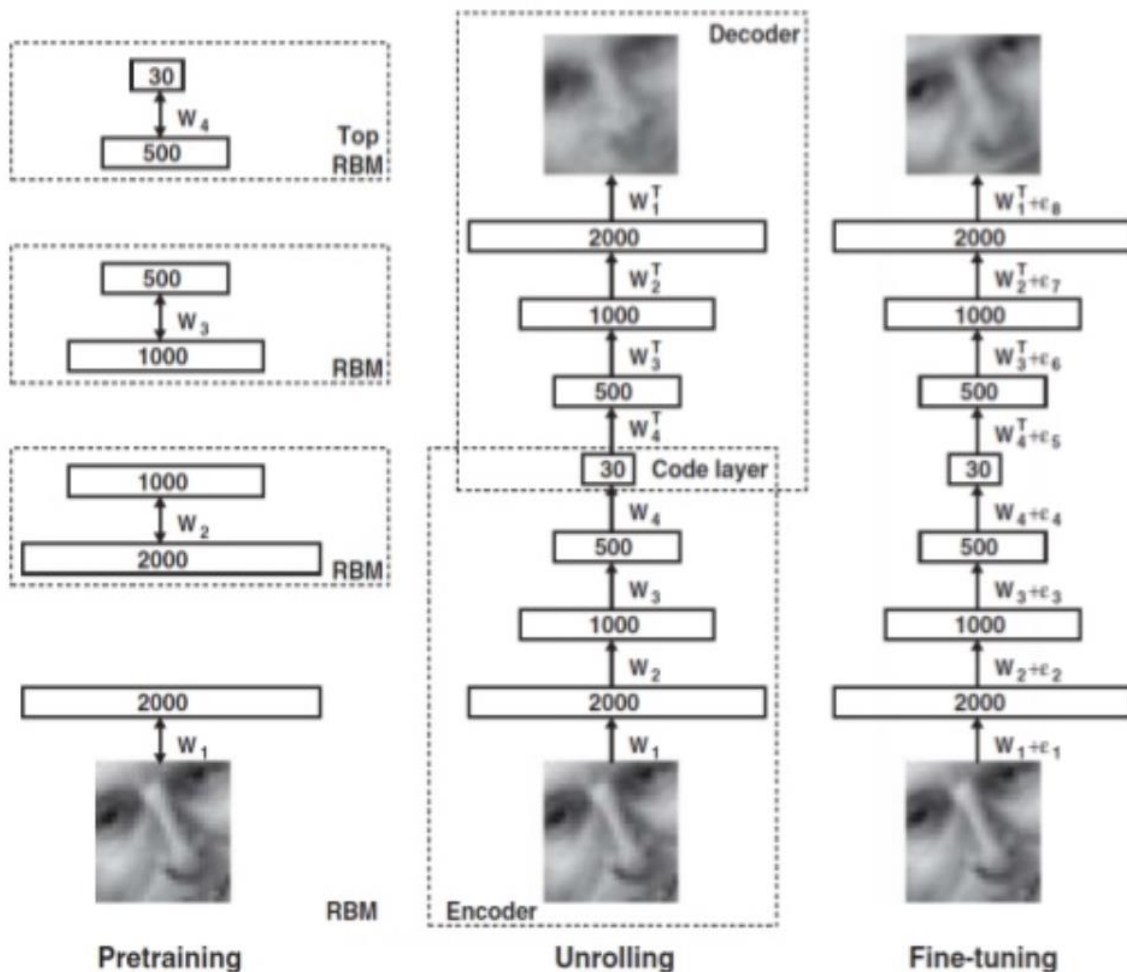
Autoencoder

- Representation learning은 이어지는 시스템에 적합한 입력을 생성

Greedy layer-wise unsupervised pretraining

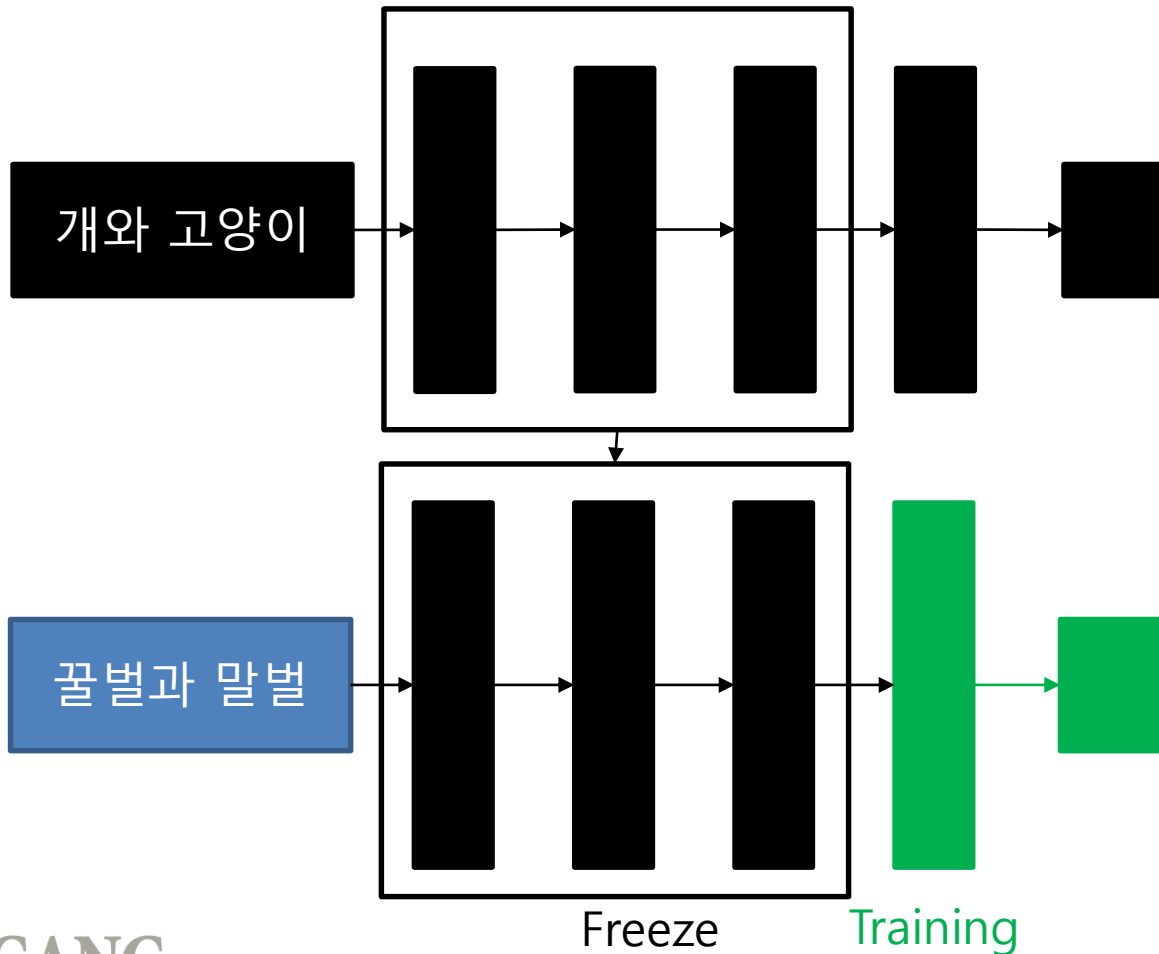
- 목적
 - Parameter 초기화
- 의도
 - 유의미한 weight
- 장점
 - 적합한 초기치
 - Unsupervised learning
- 쓰이는 분야
 - 언어

Reducing the Dimensionality of Data with Neural Networks



Transfer learning

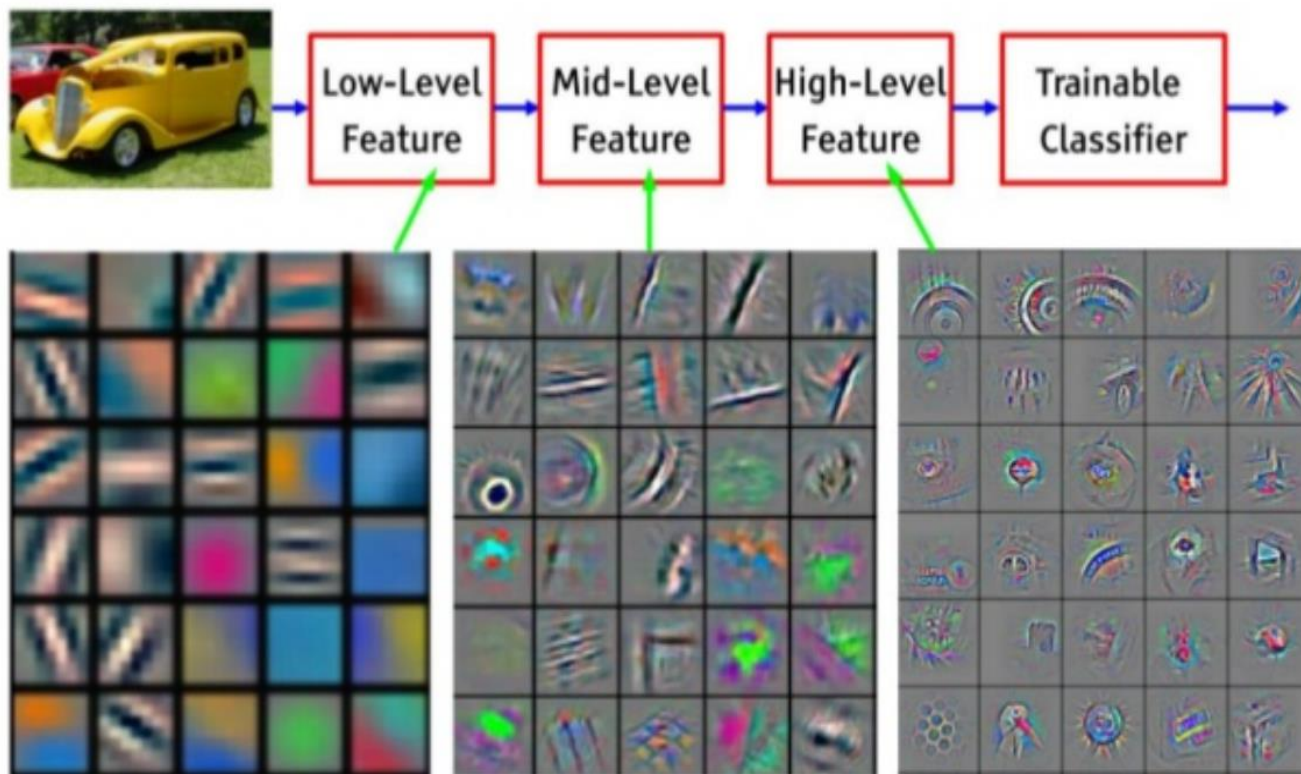
- 목적 : data가 부족한 label을 학습시킬 때 기존의 유사한 유형의 학습 된 모델을 사용하여 data를 학습하는 기법



아주 적은 데이터로
도 빠르게 일반화
되는데 도움

Transfer learning

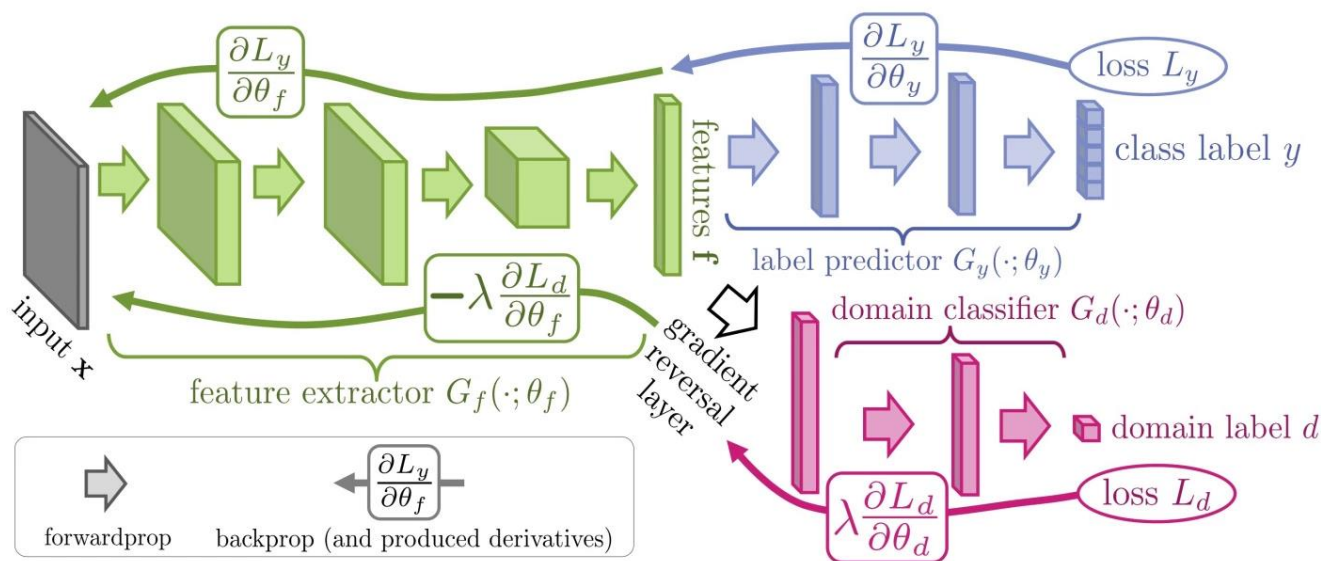
- Transfer learning의 또다른 장점
 - 복잡한 모델일 수록 학습시키기 어렵다.
 - Layer의 개수, activation hyper parameters 등등 고려해야 할 사항들이 많으며, 처음부터 학습 시키려면 많은 시도가 필요하다.



Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

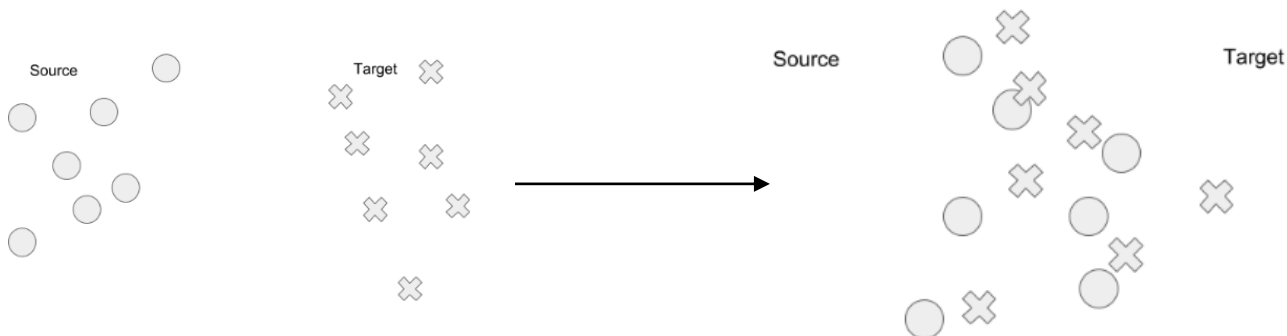
Domain adaptation

- 목적 : Label을 얻기 힘든 데이터 환경에서 synthetic 환경에서 얻은 데이터로 실제 환경에 적용하기 원하는 경우
- 예시 : 영화에 대한 리뷰가 긍정적인지 부정적인지를 판별하는 것을 학습시킨 모델에서 책에 대한 리뷰를 판별하도록 하는 sentiment classification



Domain adaptation

- Label predictor
 - Source Domain data로 label 예측하는 일반적인 classifier
- Domain classifier
 - Source domain 과 Target domain을 구분하는 classifier(**gradient reversal layer**)
- Feature extraction
 - Domain classifier 에게 구분 당하지 않기 위해 두 domain의 동일 한 feature를 추출 하면서 label을 예측하기에 좋은 feature를 뽑는 layer



Domain adaptation

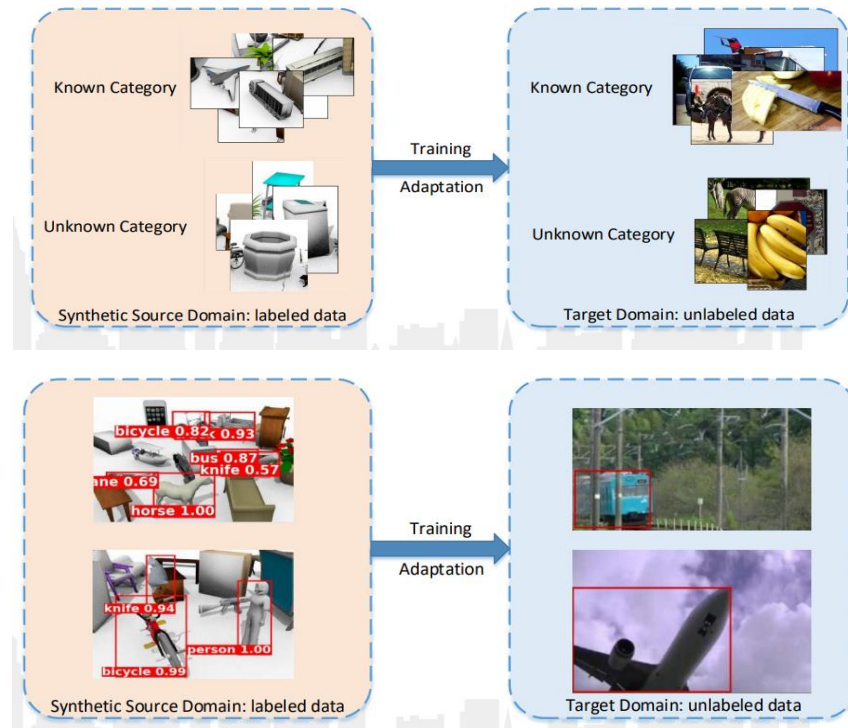
- Visual Domain Adaptation Challenge(2018)

Openset Classification [full leaderboard]

#	Team Name	Affiliation	Score
1	VARMS	JD AI Research, CV Lab	92.3 [slides]
2	Diggers	University of Electronic Science and Technology of China	69.0 [slides]
3	THUML	Tsinghua University	68.3 [slides]

Detection [full leaderboard]

#	Team Name	Affiliation	Score
1	VARMS	JD AI Research, CV Lab	48.6 [slides]
2	GF_ColourLab_UEA	University of East Anglia, Colour Lab	13.5 [slides]
3	UQ_SAS	University of Queensland, Australia	12.1 [slides]



Unsupervised representation learning with DCGAN(2016)

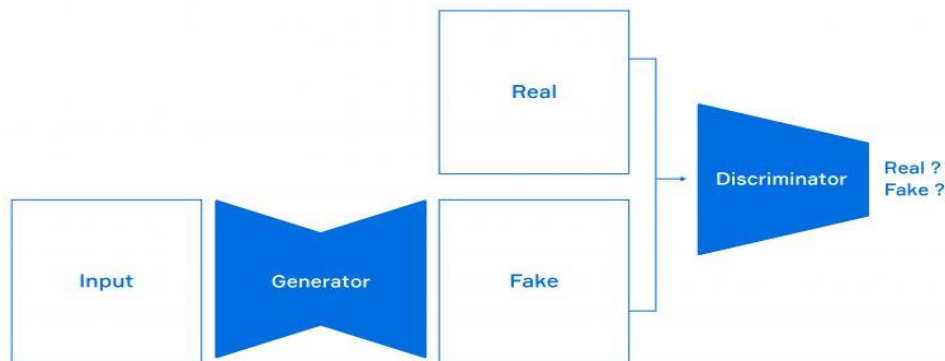
- GAN

- Unsupervised learning

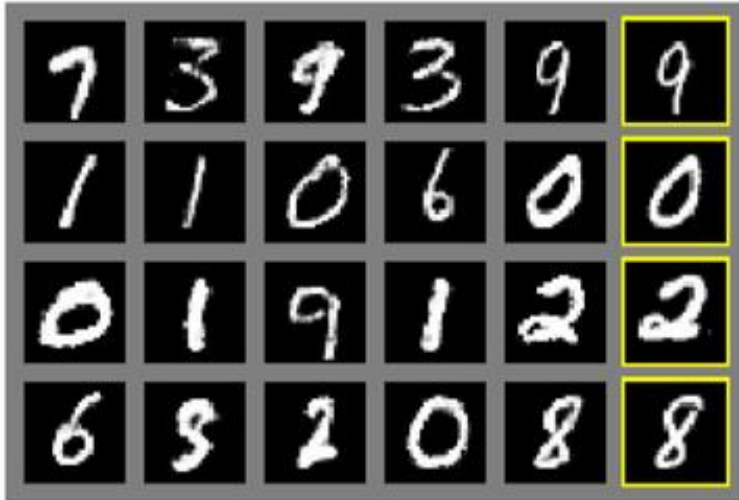
- Generator : Discriminator를 속일 데이터 생성
 - Discriminator : Real image 와 Fake image 구분
 - Input : noise

- 단점

- 복잡한 영상에 대해 결과가 나쁨.
 - 학습 시 안정성이 떨어지는 문제



Unsupervised representation learning with DCGAN(2016)



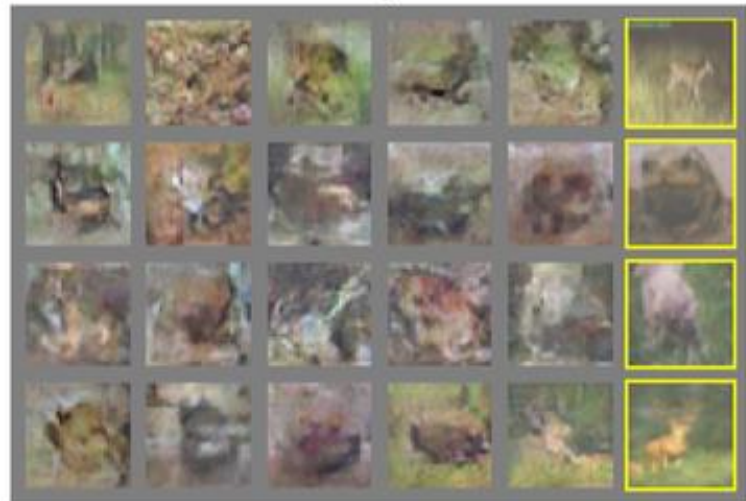
a)



b)



c)

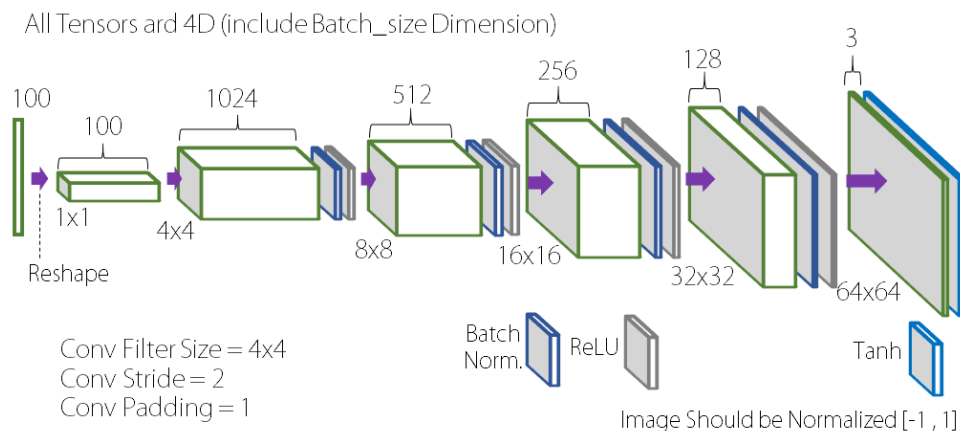


d)

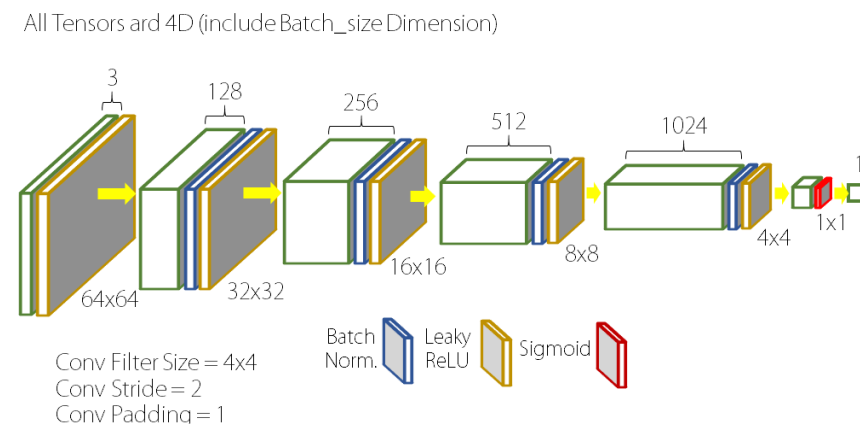
Unsupervised representation learning with DCGAN(2016)

- **DCGAN**

- CNN을 적용시킨 GAN
- 구조적 특징
- Max-pooling layer를 없애고, strided convolution이나 fractional-strided convolution을 사용하여 feature-map의 크기를 조절.
- Batch normalization을 적용하기.
- Fully connected hidden layer를 제거하기.
- Generator의 출력단의 활성화함수로 Tanh를 사용하고, 나머지 layer는 ReLU를 사용.
- Discriminator의 활성화함수로 LeakyReLU를 사용.



Generator



Discriminator

Unsupervised representation learning with DCGAN(2016)

- DCGAN

- 입력 잡음에 대한 의미

- Latent space walking

- Input 값 latent space 를 조금씩 바꿔가면서 생성되는 output 이미지가 어떻게 바뀌는지



LSUN dataset

Unsupervised representation learning with DCGAN(2016)

- DCGAN

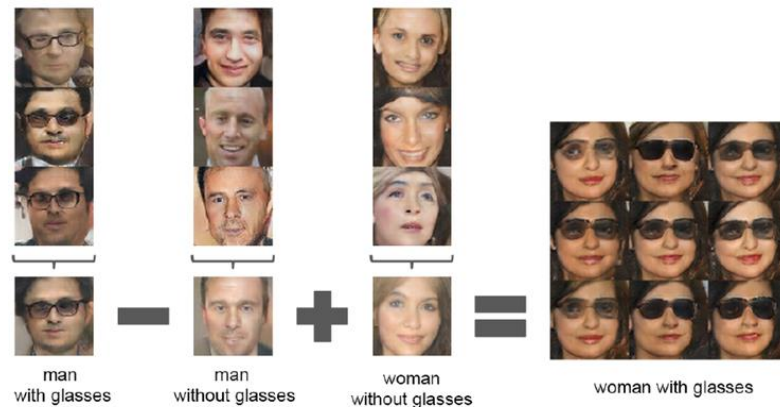
- Vector arithmetic

- [1]논문에서 단어들에 대한 representation을 벡터 공간에 매핑을 시키면 재미있는 linear arithmetic이 가능함을 보였다.

- Ex) $\text{vec}(\text{'Madrid'}) - \text{vec}(\text{'spain'}) + \text{vec}(\text{'France'}) = \text{vec}(\text{'paris'})$

- DCGAN에서도 유사한 실험

- 학습을 통해 얻은 latent space(z)값이 의미 없는 값이 아니라 각각의 영상이 갖고 있는 representation을 제대로 나타낸다는 것



[1] Mikolov, Tomas, et al. "Distributed representations of words and phrases and their compositionality." *Advances in neural information processing systems*. 2013.

Unsupervised representation learning with DCGAN(2016)

- DCGAN



왼쪽을 보고 있는 얼굴의 벡터와 오른쪽을 보고 있는 얼굴에 대한 벡터에 대한 interpolation

결론

- Representation learning
 - Feature extraction의 용도로 자주 사용
 - unsupervised representation learning의 장점 및 예시들
 - Data에 자유로움
 - Unsupervised representation learning은 언어 쪽에서 자주 사용
 - 초기 표현이 좋지 않아서 $P(x)$ 데이터로 입력 x 를 다시 복원하기 힘든 경우 자주 사용
 - 이전 발표인 manifold learning의 manifold 영역을 시스템이 사용하기 편하도록 변환하는 과정
- 다음 발표