딥 러닝으로 무엇을 하고 싶은가?

얼굴 이미지 편집하는 GAN만들기

조영주 | ETRI

ETRI 한국전자통신연구원
Electronics and Telecommunications Research Institute

저는

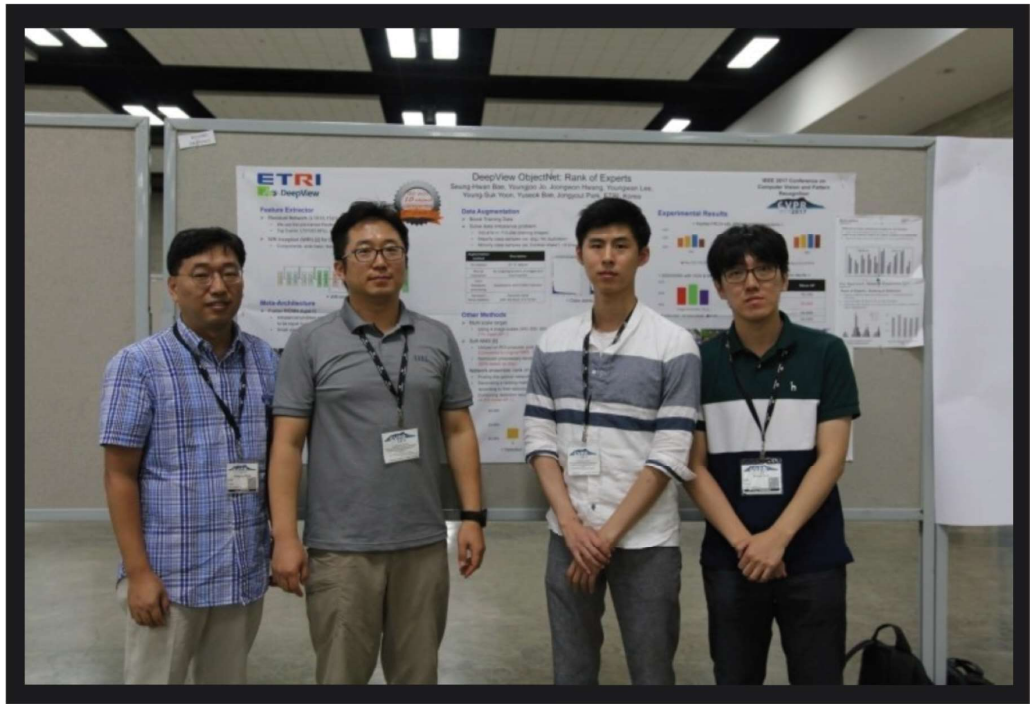
ETRI 병특

KAIST 학부/석사 졸업

Won 2nd place on detection task in 2017 ImageNet challenge

Make SC-FEGAN(★2300+)

2017 ImageNet challenge



<국제영상인식대회 세계 2,3위 쾌거>
<국제저전력영상인식대회 세계 2위 달성>

ETRI, '이미지넷 2017' 서 기술력 입증

- 시각 지능 및 콘텐츠 비주얼 검색 분야 핵심 원천 기술 확보
- CCTV, SNS, 블랙박스 심층분석, 영상 내 장소, 상품 정보검색
- 향후 공공서비스 및 지능형 콘텐츠 분야 활용

국내 연구진이 세계적 이슈인 인공지능과 관련된 『국제 영상인식 대회』 사물 검출 분야 및 『국제 저전력 영상인식 대회』에서 기술력을 입증하는 성적을 거두었다.

ETRI(한국전자통신연구원)는 27일, 미국 하와이 컨벤션센터에서 개최된 국제 영상인식대회(ILSVRC[1], 이미지넷) 사물검출 분야에서 전세계 기업, 대학 연합팀들과 겨루어 사물 종류별 검출 성능 기준 2위, 평균 검출 정확도 기준 3위 성적을 달성했다고 밝혔다.

출처: ETRI 보도자료

2019 SC-FEGAN

JoYoungjoo / SC-FEGAN

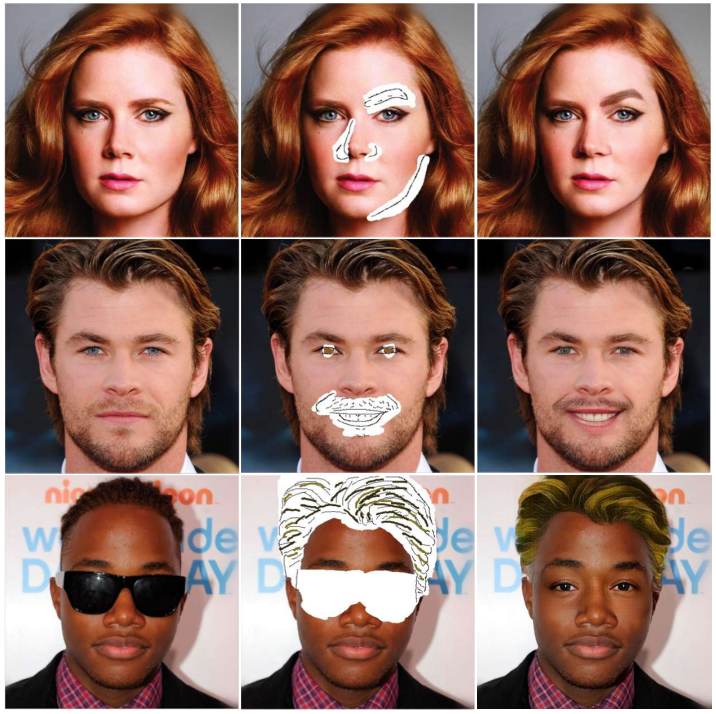
Watch 75 | Unstar 2,334 | Fork 325

Code | Issues 12 | Pull requests 2 | Projects 0 | Wiki | Insights | Settings

SC-FEGAN : Face Editing Generative Adversarial Network with User's Sketch and Color

Edit

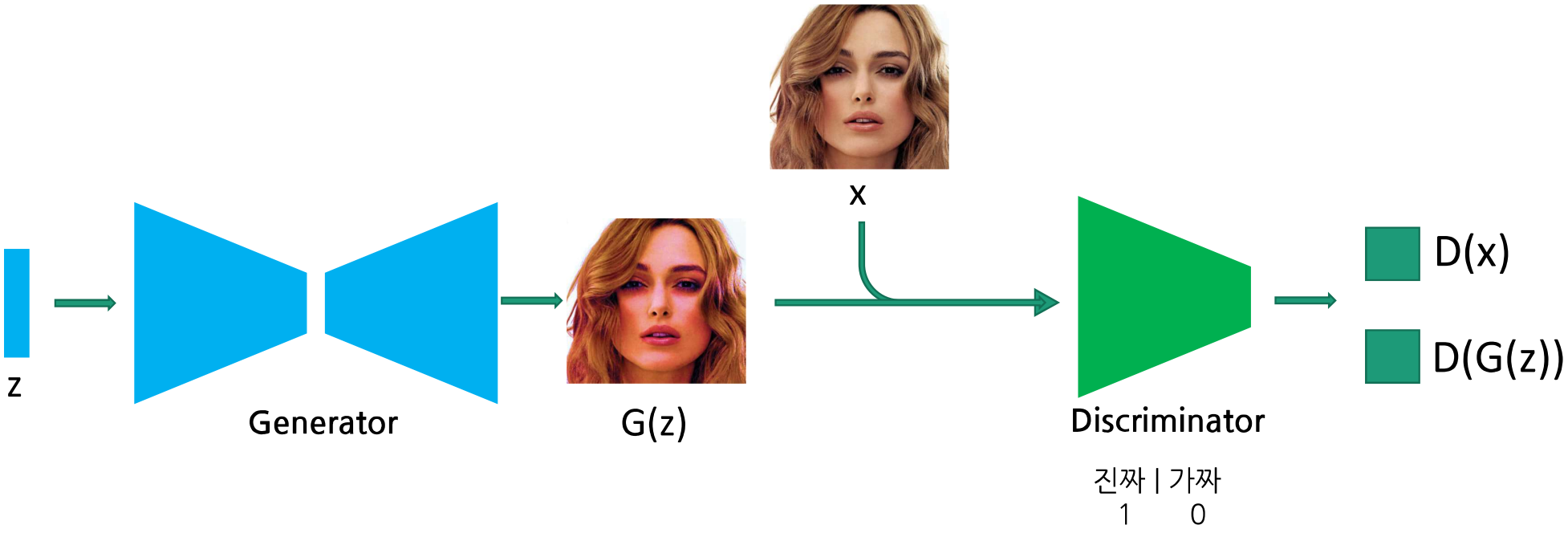
Manage topics



왜 GAN을 연구하기 시작하였는가?

그전에 잠깐 간략하게 GAN이란?

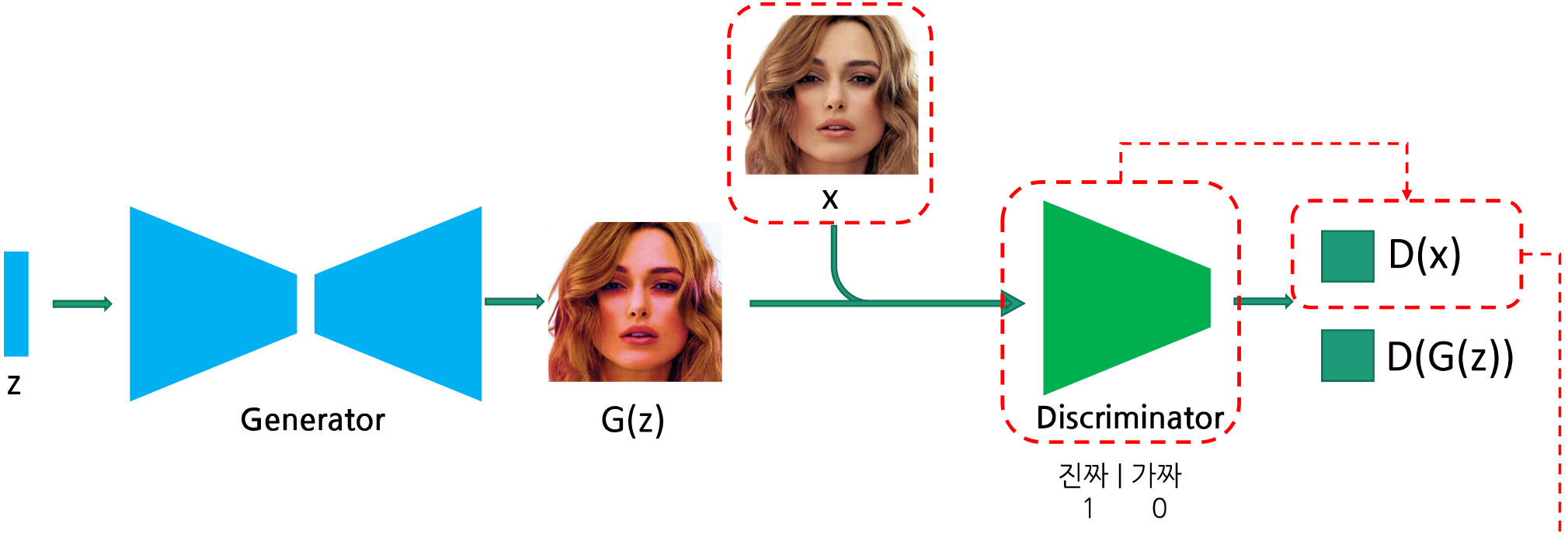
기원 - Ian Goodfellow의 NIPS 2014 paper



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

기원 - Ian Goodfellow의 NIPS 2014 paper

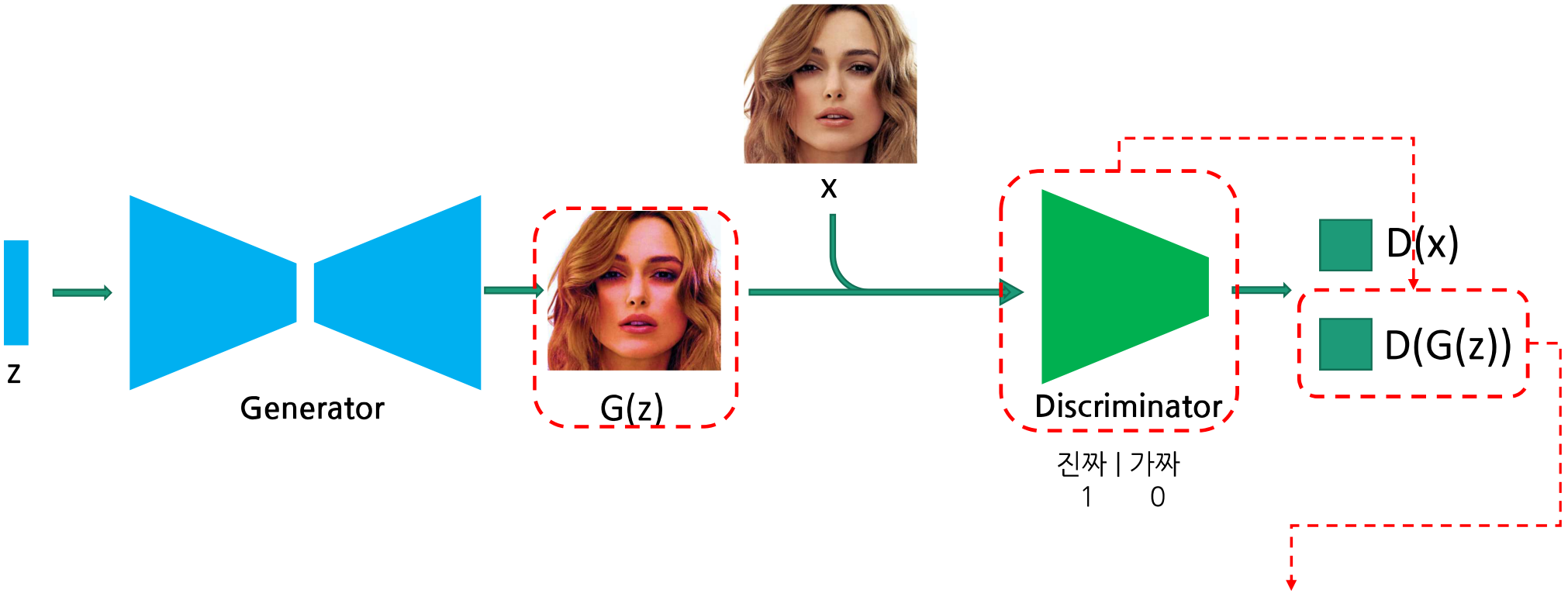
D입장



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

기원 - Ian Goodfellow의 NIPS 2014 paper

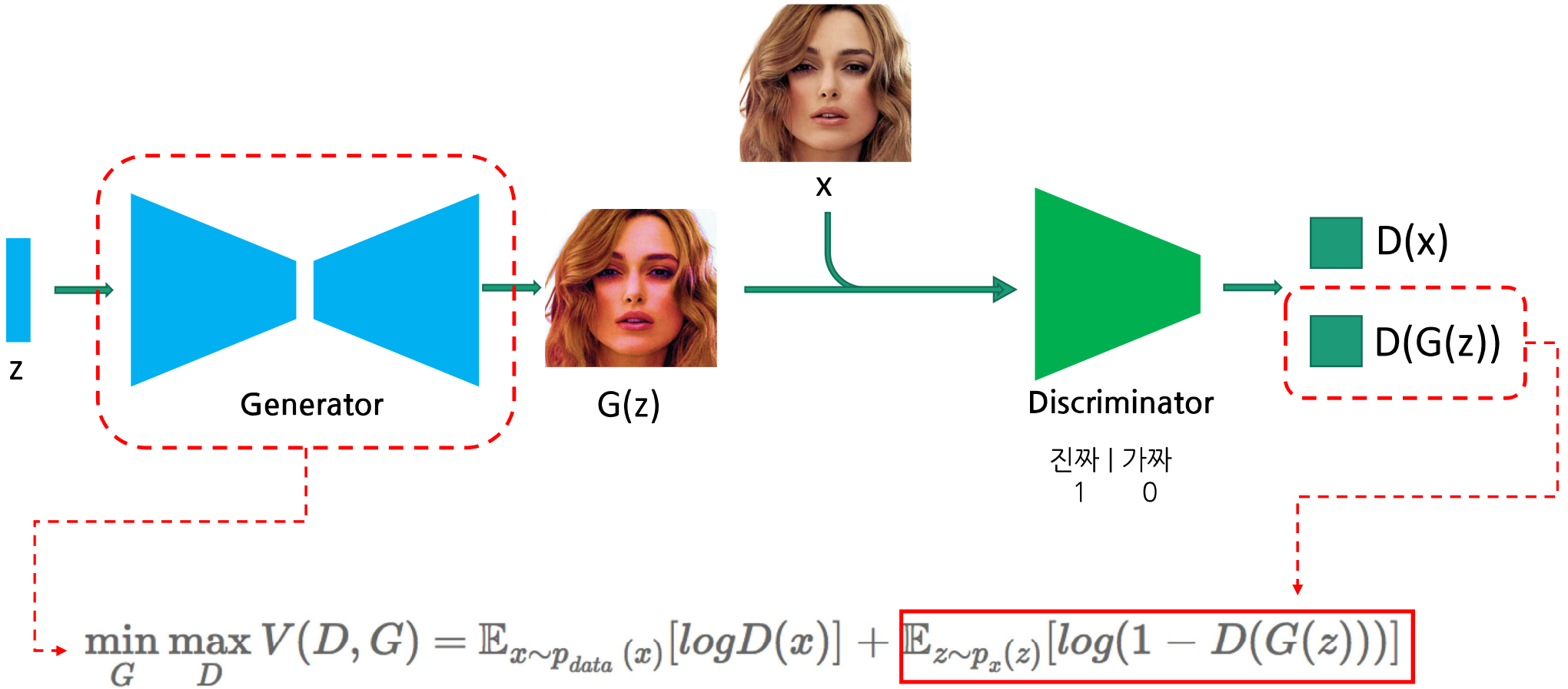
D입장



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

기원 - Ian Goodfellow의 NIPS 2014 paper

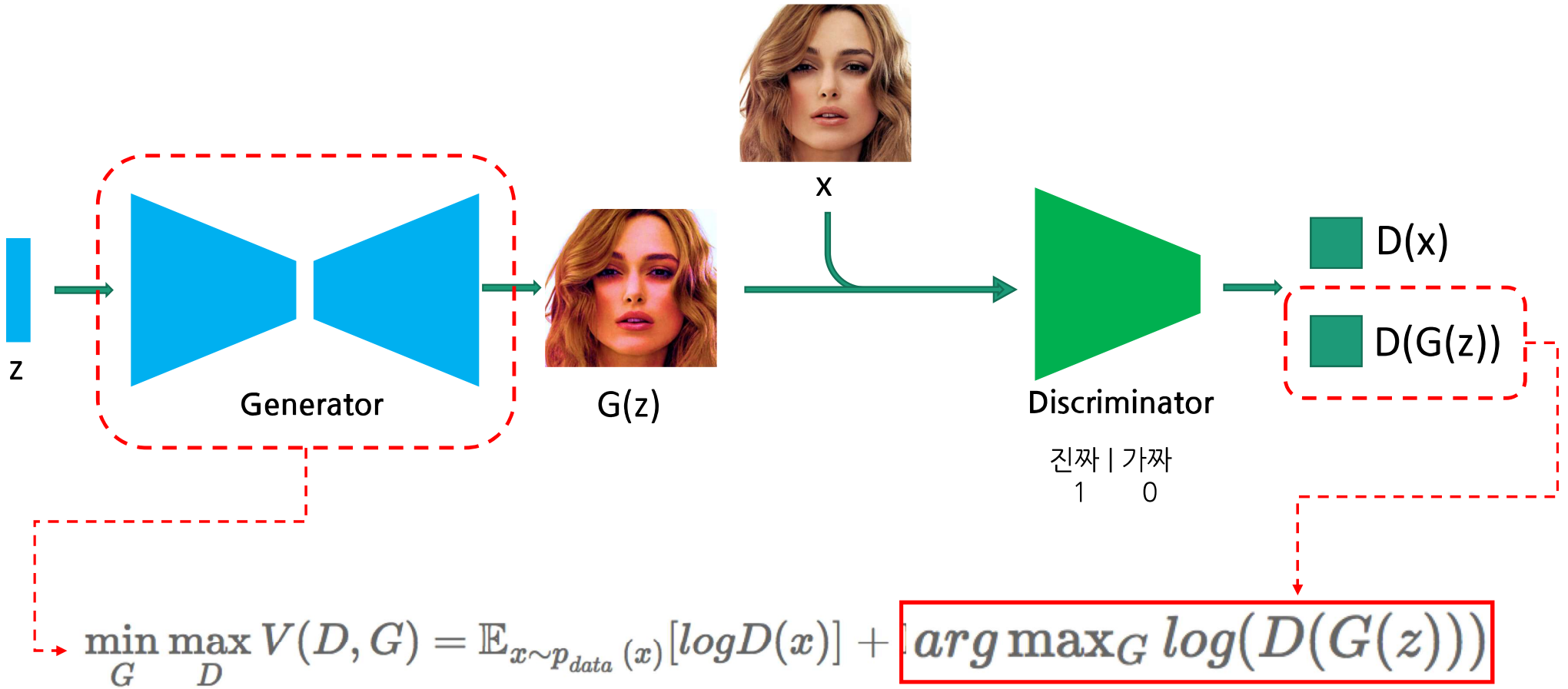
G입장



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

기원 - Ian Goodfellow의 NIPS 2014 paper

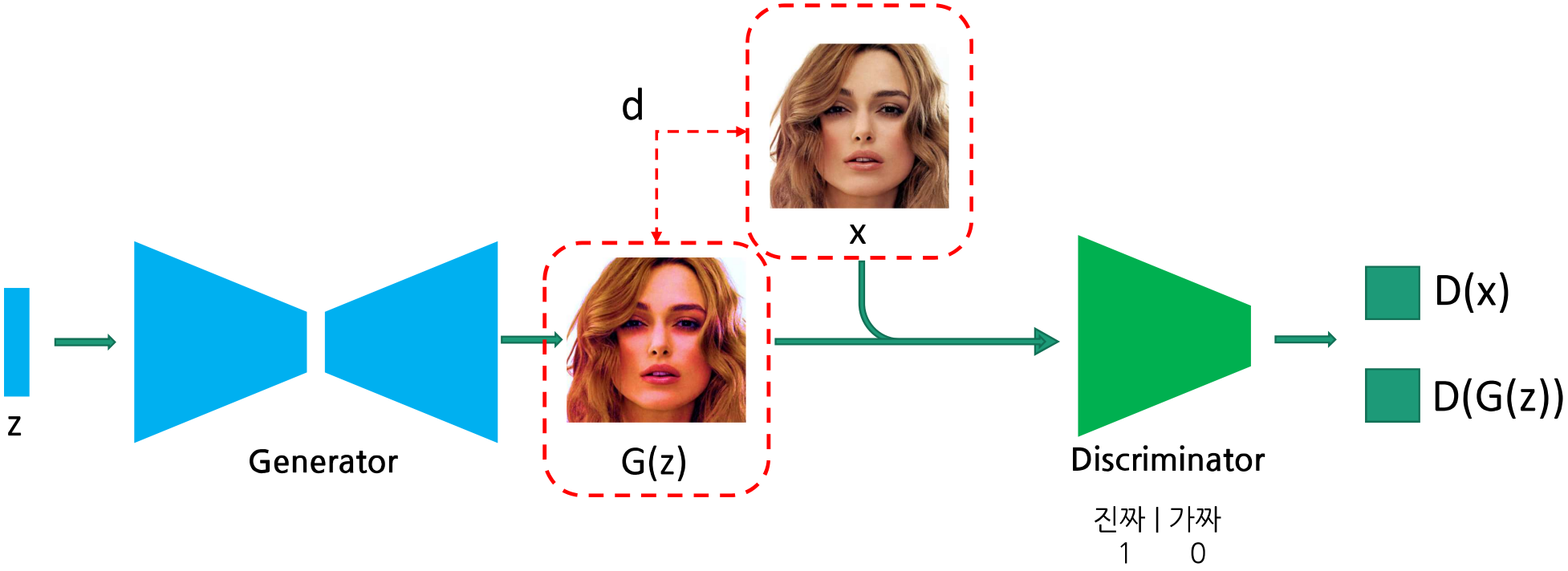
G입장



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

기원 - Ian Goodfellow의 NIPS 2014 paper

G입장



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

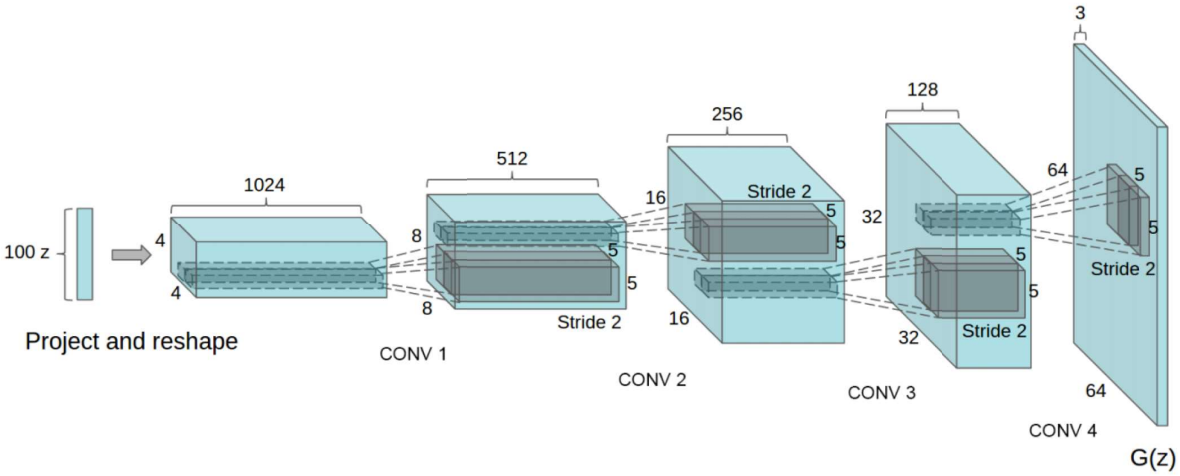
Generator는 Discriminator를 속이는 방향으로!

Discriminator는 Generator가 만든 건지 아닌지 판단하도록!

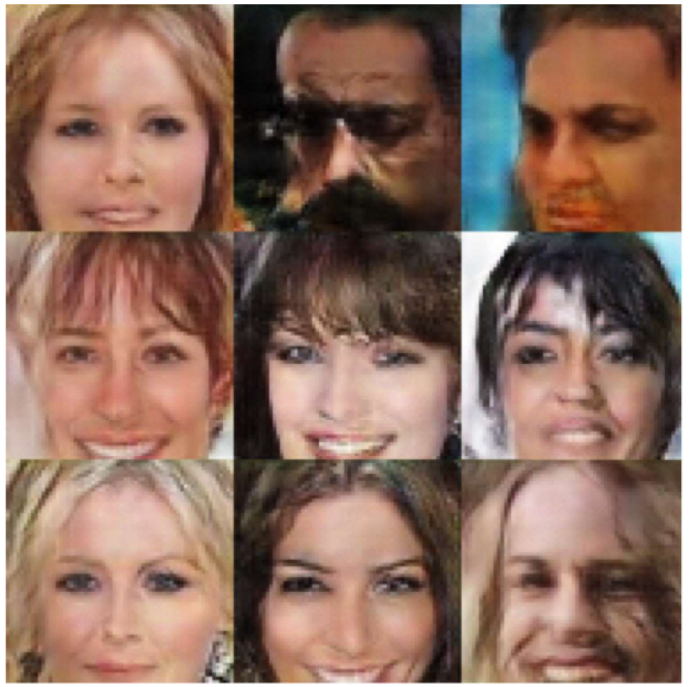
Ian Goodfellow가라사대… 위조지폐범과 경찰을 두고 위조지폐범은 경찰을 속일 수 있게 노력하고 경찰은 위조지폐를 잘 구분할 수 있도록 노력하면 위조지폐를 잘 만드는 범인과 잘 잡는 경찰을 모두 얻을 수 있다…

시작은 신기하기만 했으나…

- **Noise to Image(첫 시작)**
  - DCGAN, BEGAN, WGAN

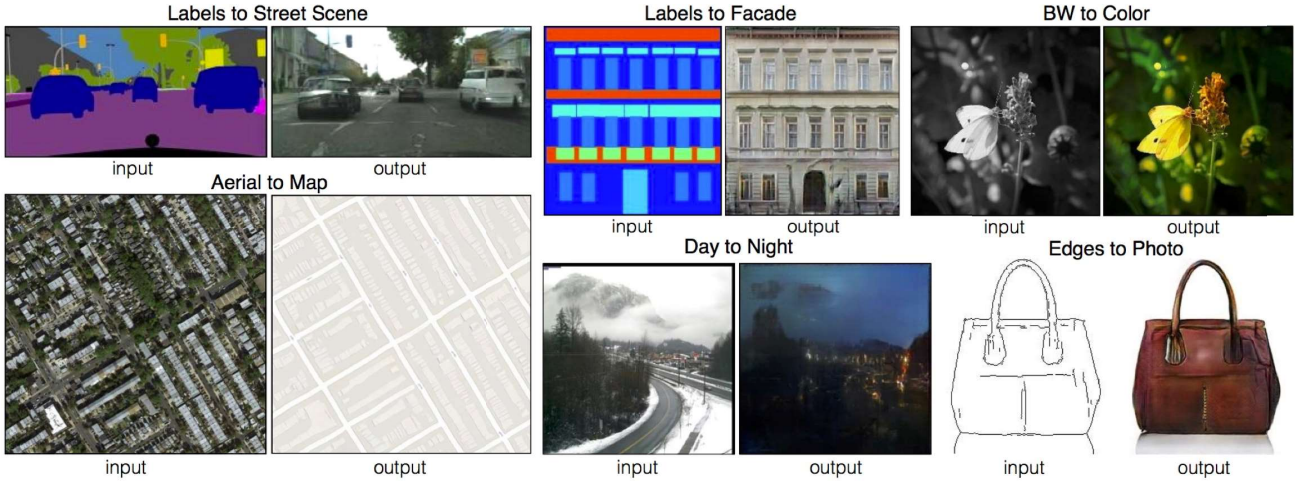


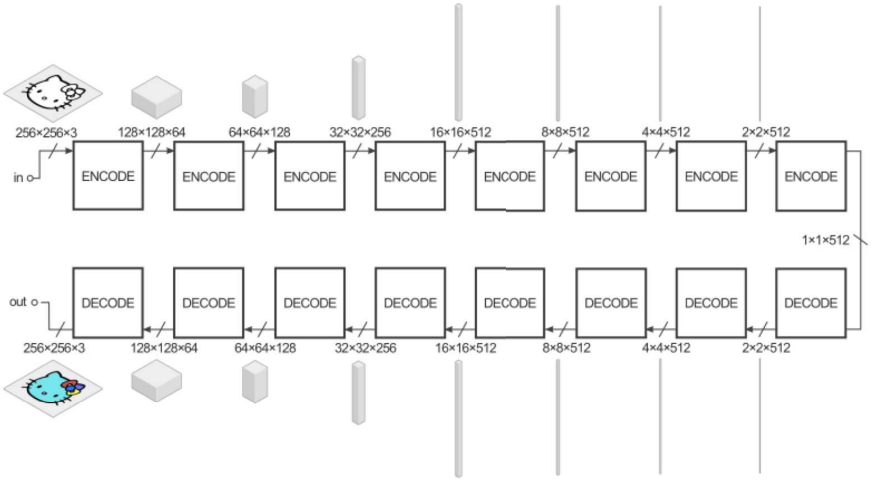
DCGAN



지금은 활용도도 무궁무진…

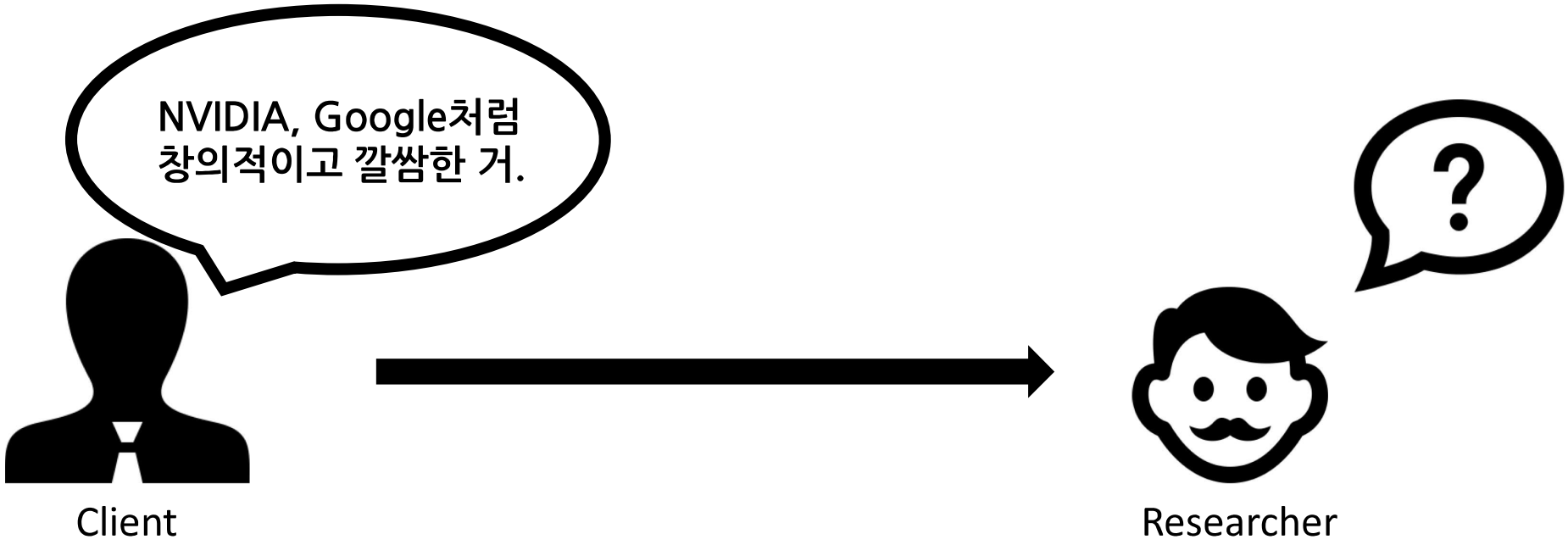
- **Image to Image (pix2pix)**
  - 최근에 나오는 거의 모든 GAN
  - 이미지 변환에 해당하는 거의 모든 분야에 활용





저는 왜 GAN을 공부 했을까요?

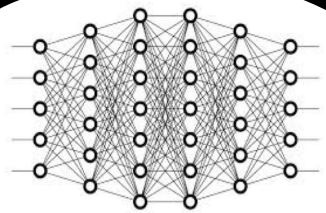
Deep Learning 연구자의 현실?


NVIDIA, Google처럼
창의적이고 깔쌈한 거.
?
Client
Researcher


Deep Learning은 만능!
Client

저걸 만들려면…

Researcher



NLP & Voice

Google

Baidu 百度

Microsoft

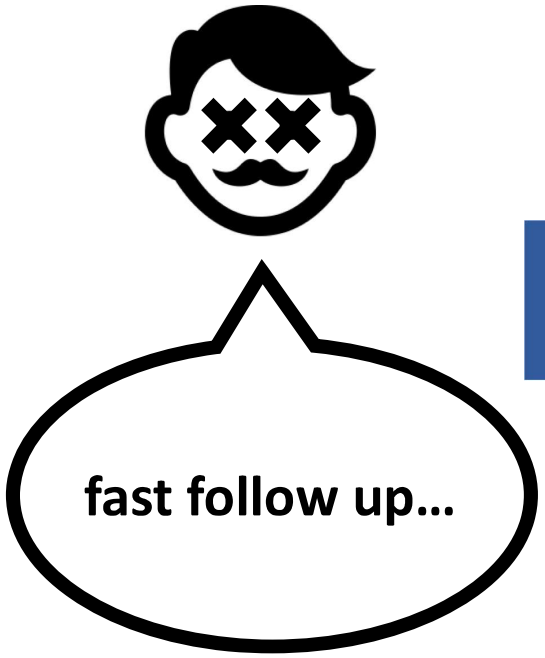
amazon

kakao

SK telecom

NAVER

SAMSUNG



Vision

Google

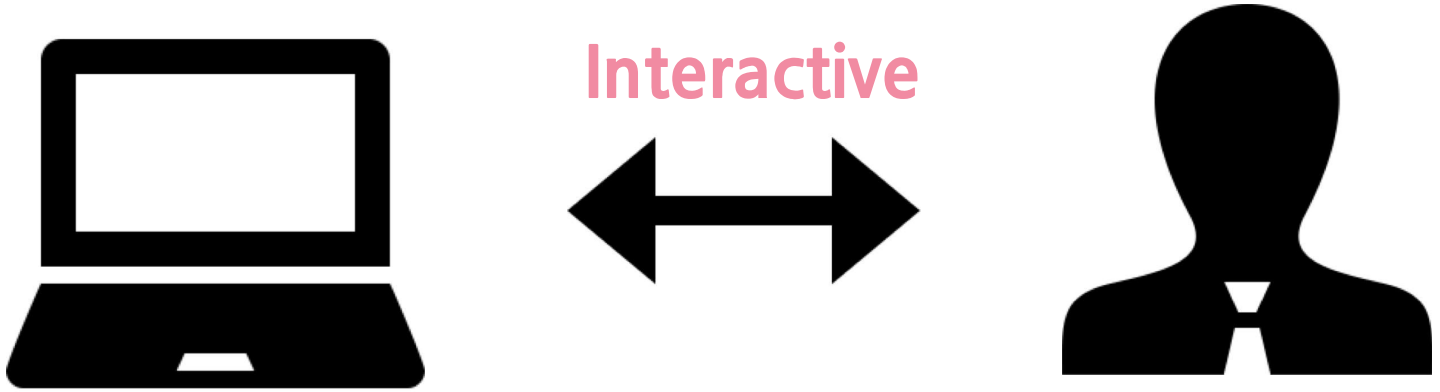
Face++ 旷视

facebook

SAMSUNG


완성도 있는 창의적
서비스 만들기는 힘들고...
뭘 해야 할까?
보조적인 역할에
집중하는 AI를 만들자!

ㅌㄴㅅ처럼 튕기면 다되는 것이 아니라..


Interactive

머가 Interactive한 걸까?

내가 생각하는 interactive

1. 결과물이 어색해도 괜찮은 것! → 성능 평가 하기 어려운 것!
2. 사용자가 의도를 입력할 수 있는 것! → 입력을 정의하기 쉬운 것!

→ Design & Editor

결과물이 어색해도 사용자가 다시 쉽게 수정해서

입력으로 넣으면서 상호작용 할 수 있는 걸 만들어보자!

검색 검색 검색…

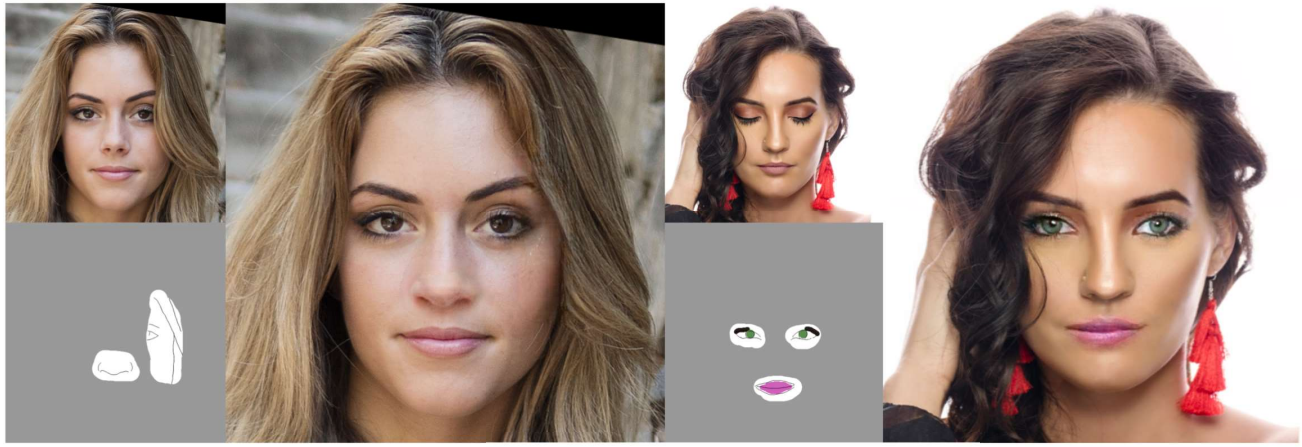
공부 공부…

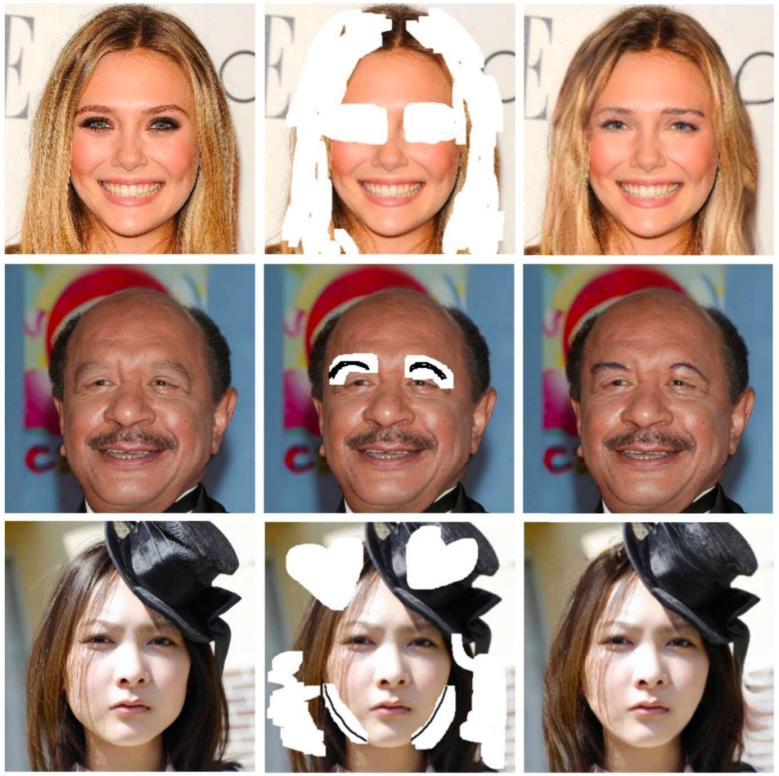
관련 Code implementation…

Key reference papers

Faceshop: Deep sketch-based face image editing.
T. Portenier, Q. Hu, A. Szabo, S. Bigdeli, P. Favaro, and M. Zwicker.

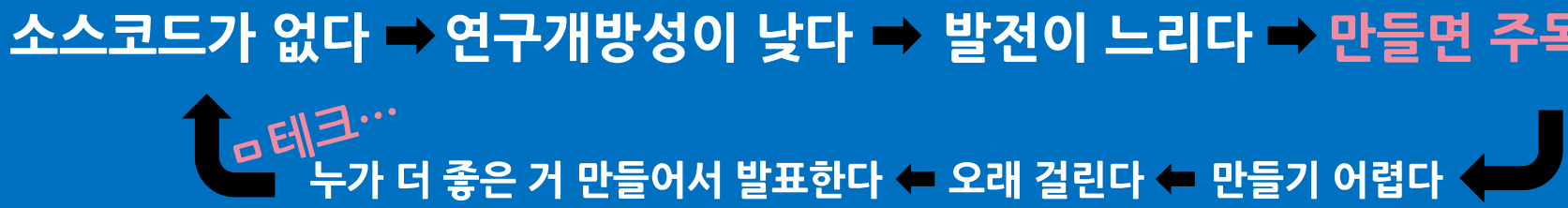
Free-form image inpainting with gated convolution. (Deepfillv2)
J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang.





진리 : 내가 생각한 건 누가 이미 만들어 놨다...

어? 근데 공개되어있는 소스코드가 아무것도 없네?


소스코드가 없다
연구개방성이 낮다
발전이 느리다
만들면 주목 받는다
ㅁ테크…
누가 더 좋은 거 만들어서 발표한다
오래 걸린다
만들기 어렵다

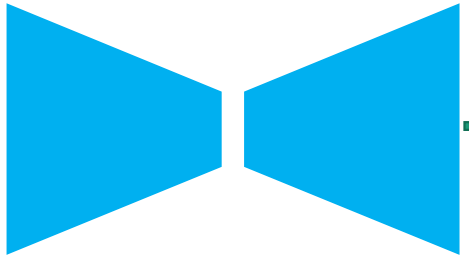
어차피 성능 지표도 없는 거 한번 해보자!

이미지 수정을 위한 GAN

학습을 위한 GT가 필요

입력이
그냥 이미지라면?

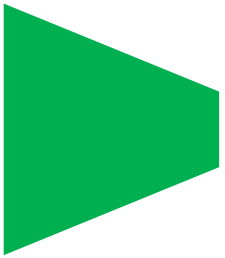
z



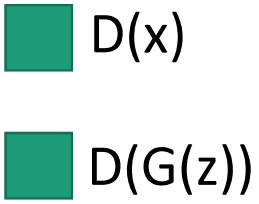
Generator

G(z)

x



Discriminator



진짜 | 가짜
1 0

머하러 변형하지?
그냥 그대로 내보내면 GT인데?

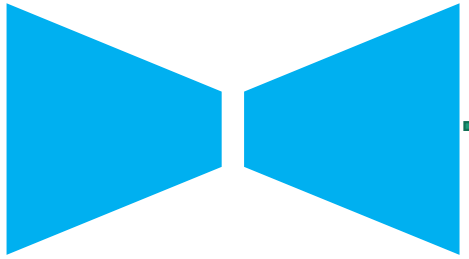
이미지 수정을 위한 GAN

학습을 위한 GT가 필요

x

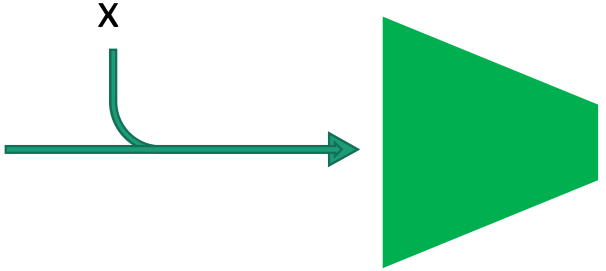
입력이
그냥 이미지라면?

z

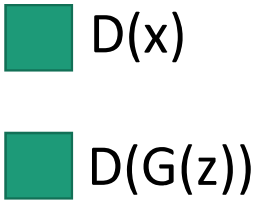


Generator

G(z)



Discriminator



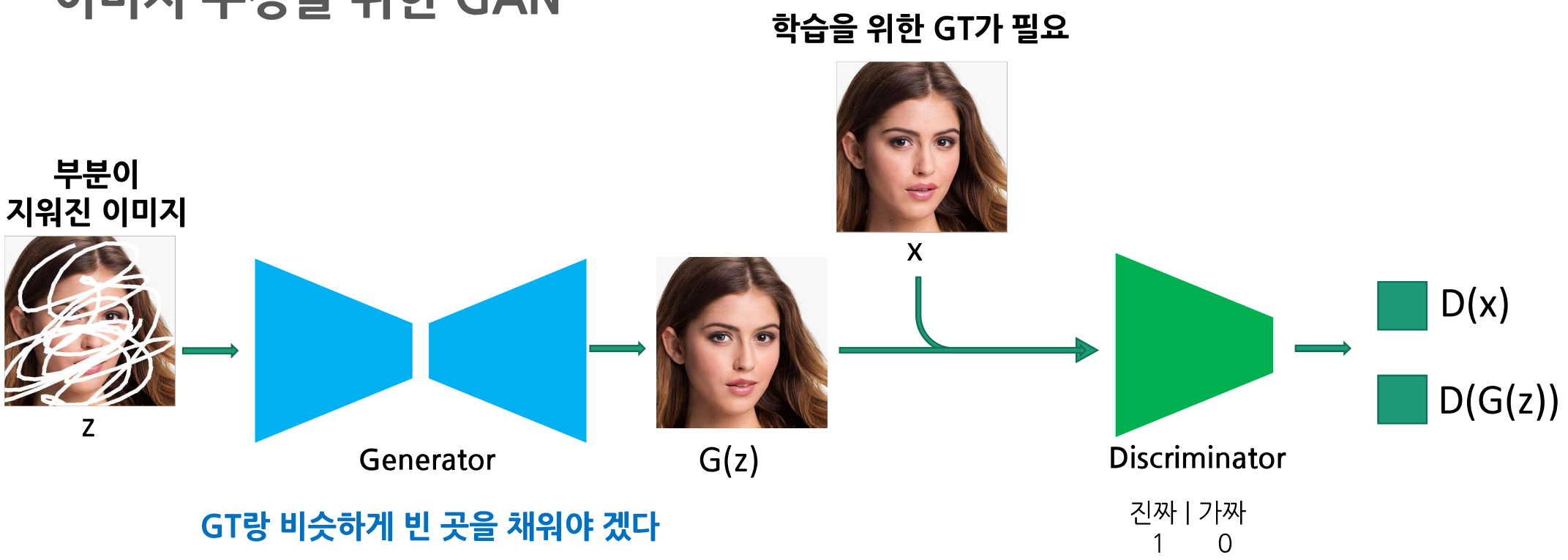
D(x)

D(G(z))

뭐하러 변형하지?
그냥 그대로 내보내면 GT인데?

진짜 | 가짜
1 0

이미지 수정을 위한 GAN



GT랑 비슷하게 빈 곳을 채워야 겠다

Image inpainting

이미지 수정을 위한 GAN

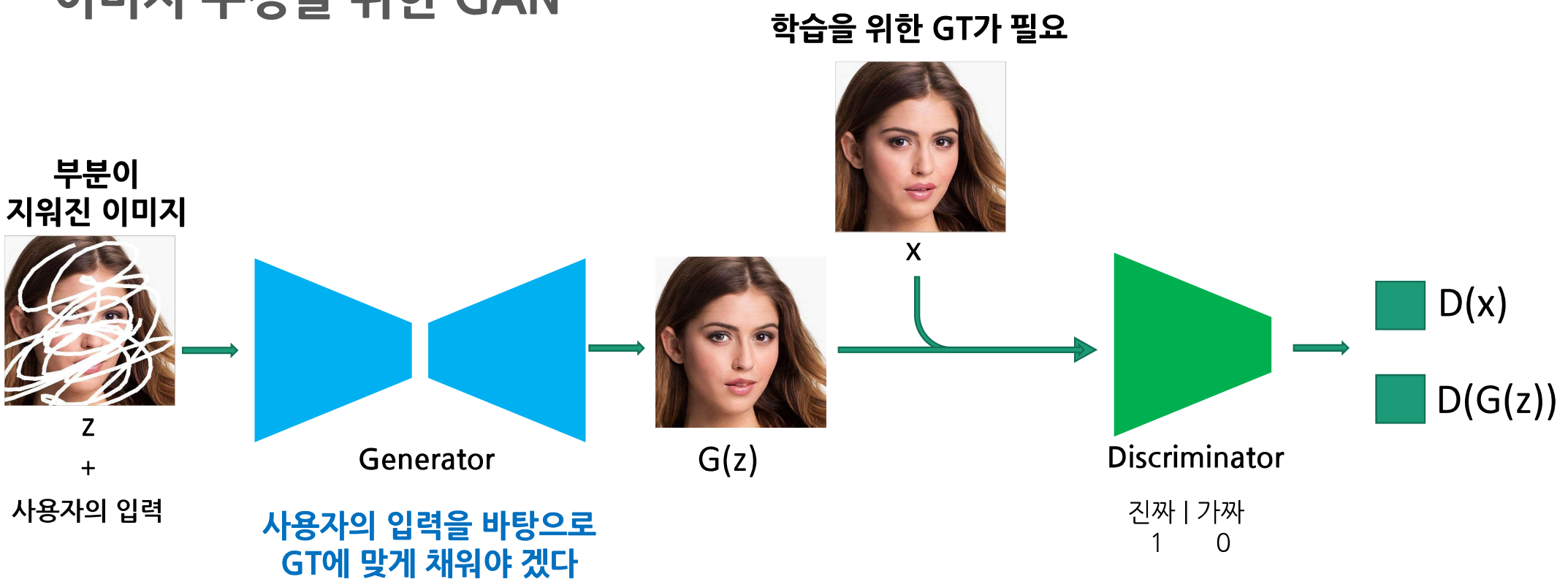


Image inpainting을 활용한 이미지 수정

Target : Faceshop

1. 데이터 준비

   CelebA data download, Align, Crop, Sketch domain, Color domain, Random rotable mask…

2. Network 구조, Loss 설정

   U-net, Coarse-Refined net, global-local discriminator, L2 loss, GAN loss…

3. Training


제발...이미지 답게만...

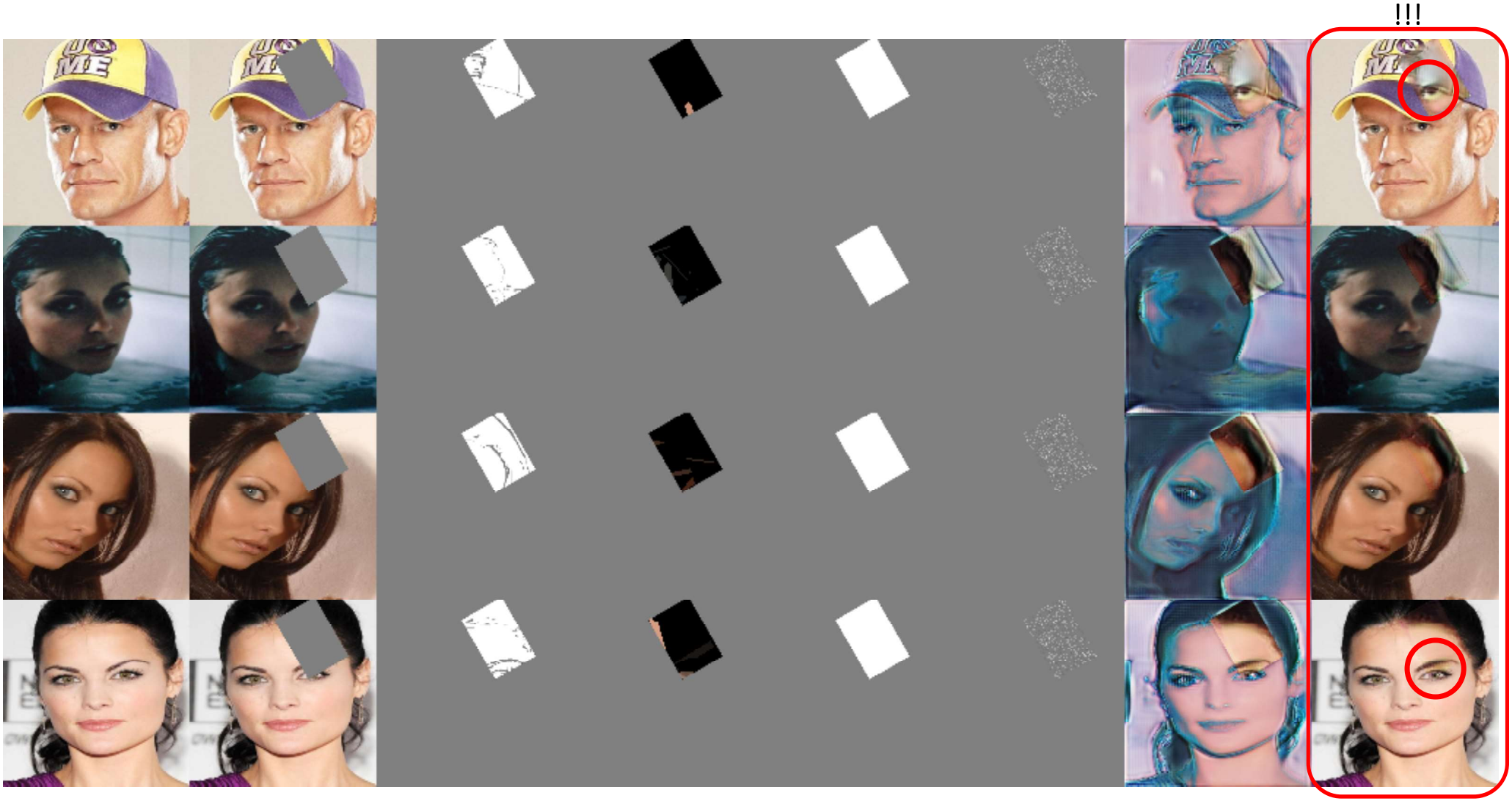

아... ㄱ망했...

Target : Faceshop

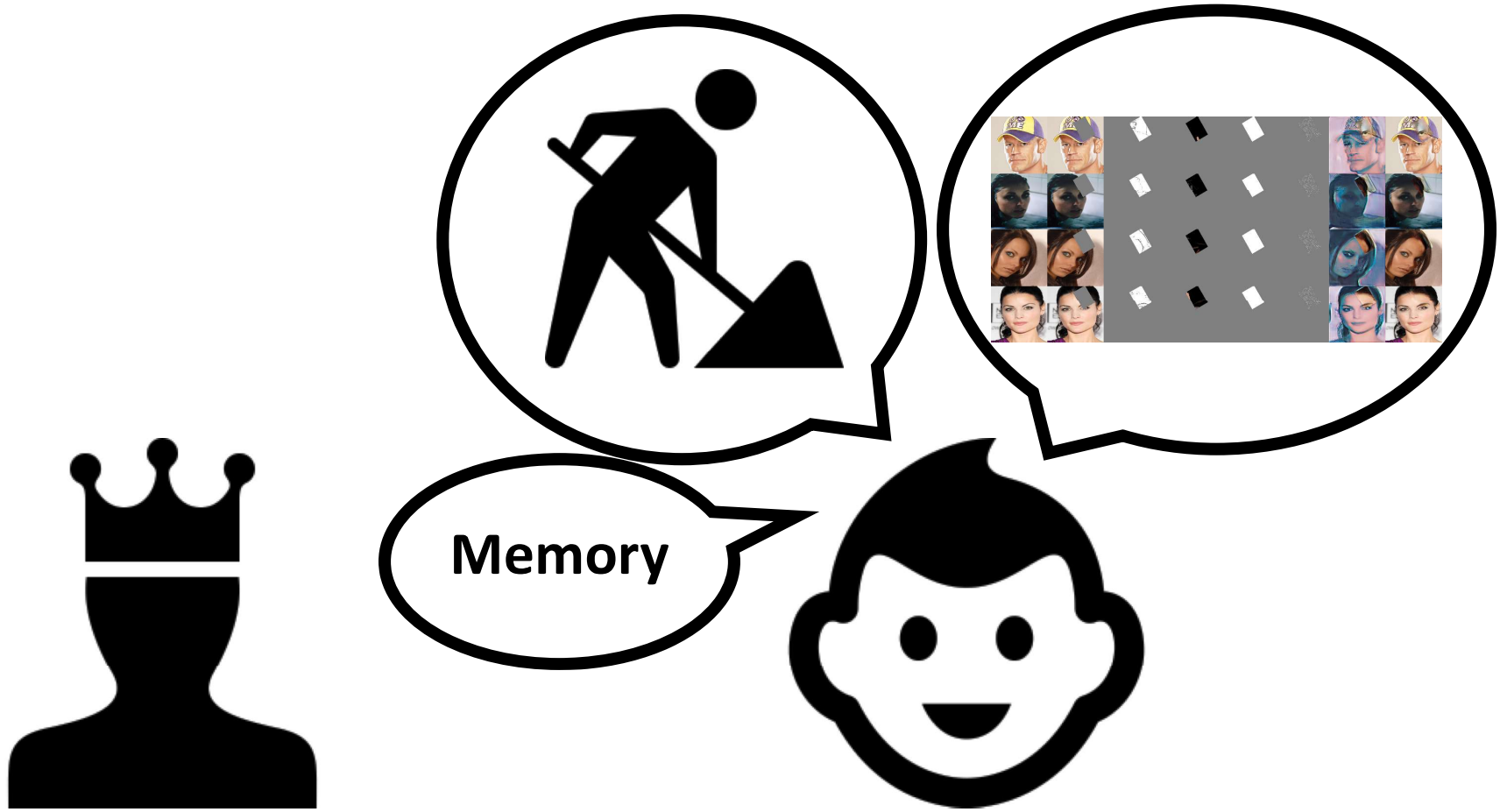
1. Loss 개선, 구조 개선
   SN-discriminator, perceptual loss, L2 loss, GAN loss…

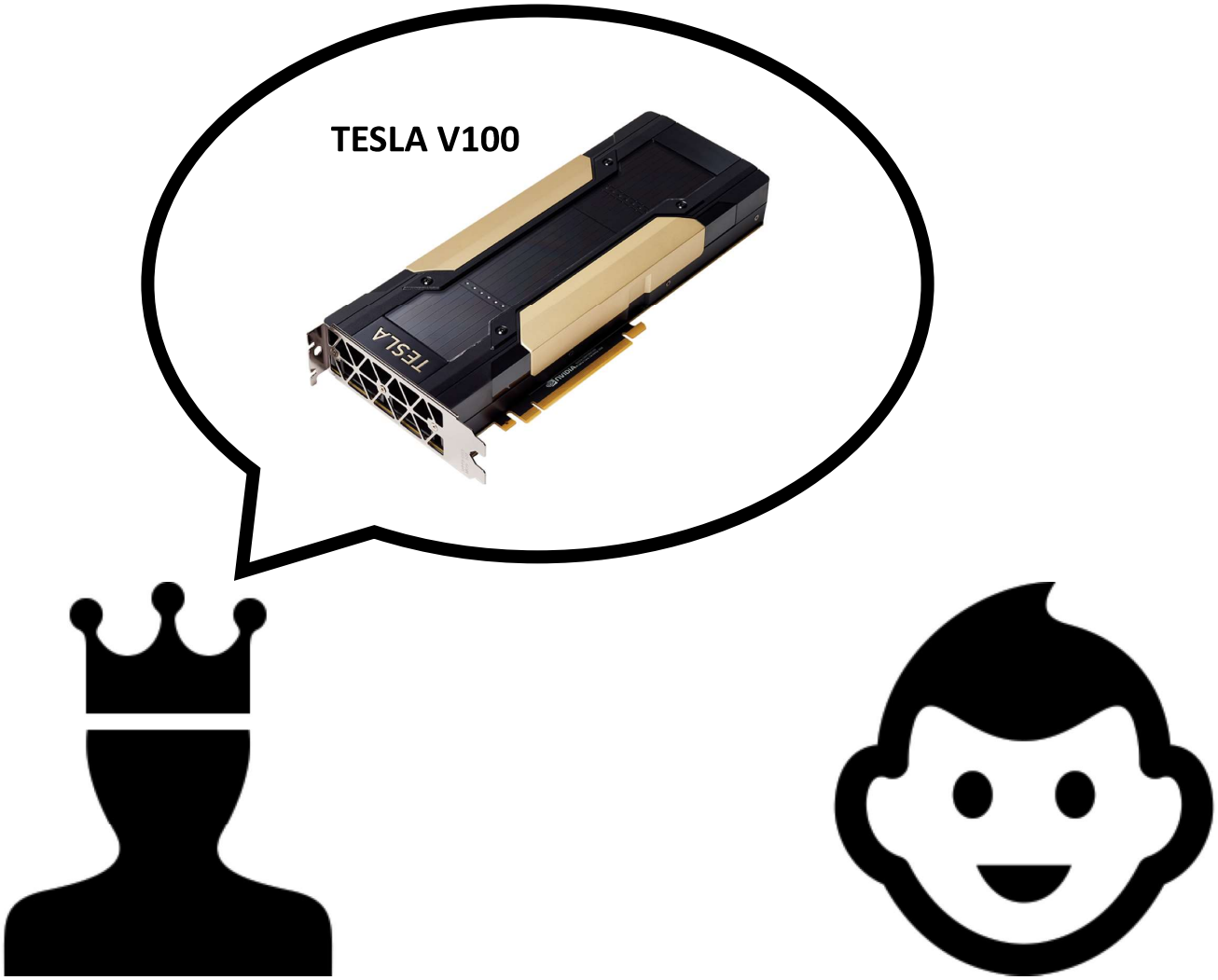
2. Training


제발...한번만...

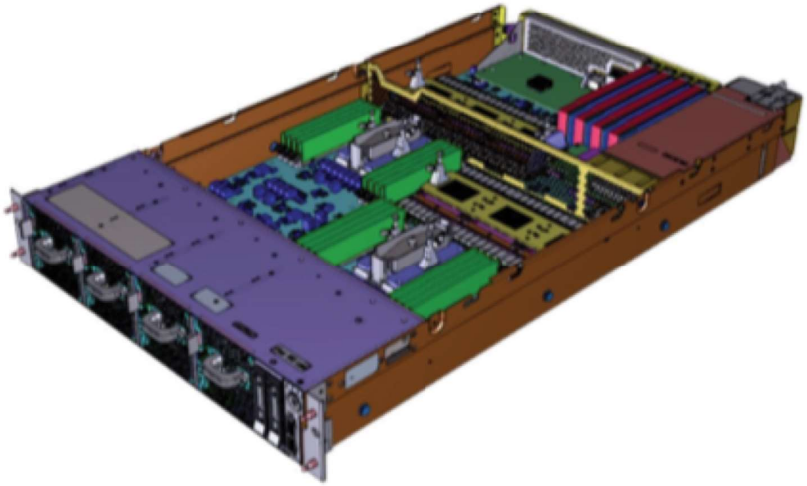


오! 먼가 됐어! 정리해서 Upgrade 고고!


Memory


TESLA V100
TESLA

Power Systems AC922



19" rack compatible
2U rack-mounted form factor
3 year, 9x5 warranty

- CPU: POWER9 2.6GHz 16 cores x 2 or
  POWER9 2.0GHz 20 cores x 2
- Memory: DDR4 128GB ~ 2048GB
  8 DIMM slots / CPU
- Storage Bay: 2.5" bay(HDD or SSD) x 2
- GPU: Nvidia V100(SXM2 type, HBM2 16GB)
  2 or 4 or 6(Water-cooling only)
- I/O Slots: PCIe Gen4 x 4
- CPU-GPU & inter-GPU interface: NVLink 2.0
- Power Supply Unit: 2,200W x 2
- OS: RHEL, Ubuntu
- S/W: PowerAI, Power HPC Stack

1500만원 GPU 4개(32GB)

POWER 9 CPU

서버 가격 = 약 1억

돈의 힘이란…

Loss, Network, Layer 개선

SN-discriminator, perceptual loss,
L2 loss, GAN loss, Gated conv

입력 방법 개선

Free-from mask, color and sketch

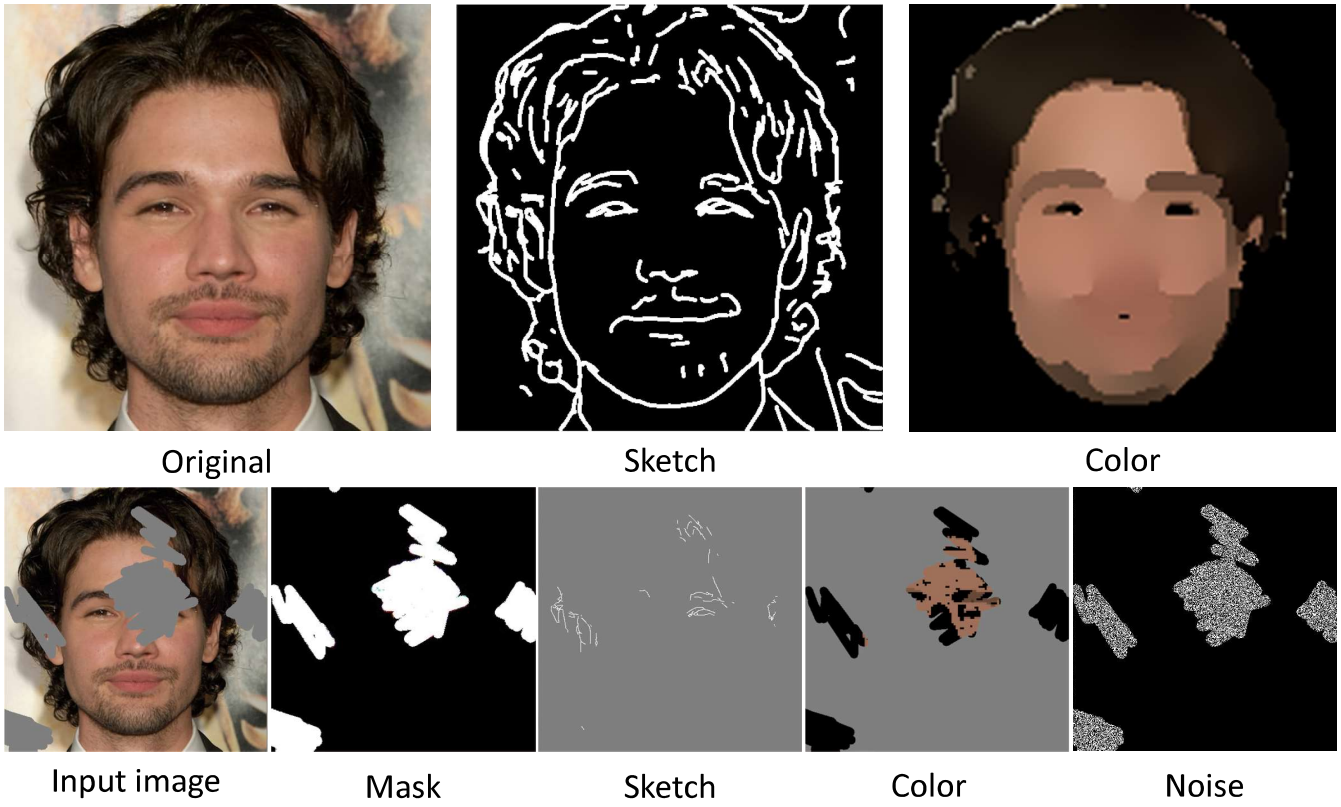
Data 정제 작업

HED edge detector, smoothing,
GFC map, median color, CelebA-HQ data

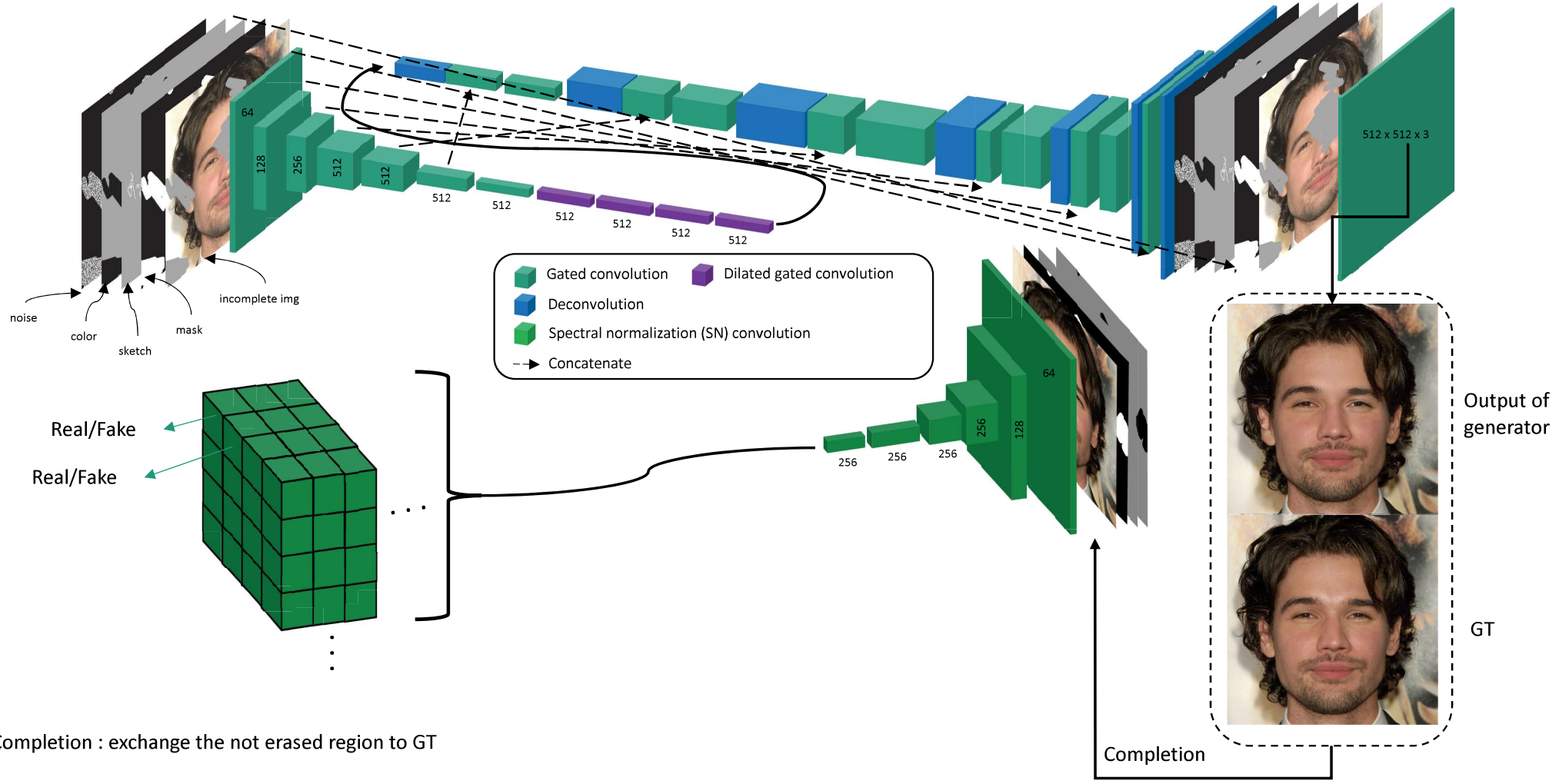
사용자 편의 개선

GUI 작업

- **Data**
  - **CelebA-HQ**
  - **Mask – Free-from mask**
  - **Sketch – HED edge detector**
  - **Color – GFC & median color**



- **Network**



Completion : exchange the not erased region to GT

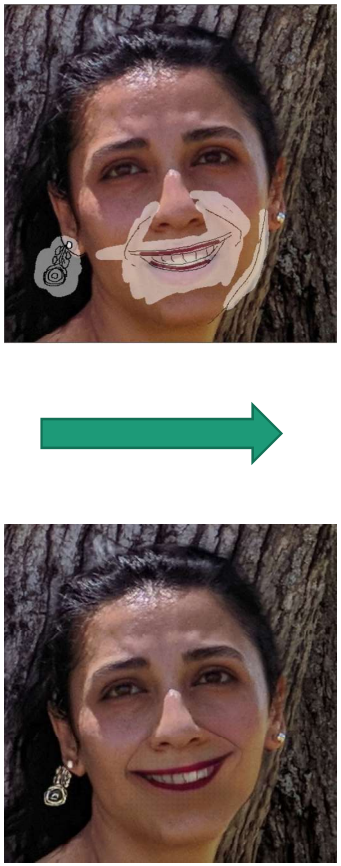
결과는?

2019 SC-FEGAN

보지 않은 종류의 데이터 셋에 적용 가능할까?

- **Results**
  - Editing HDR photo

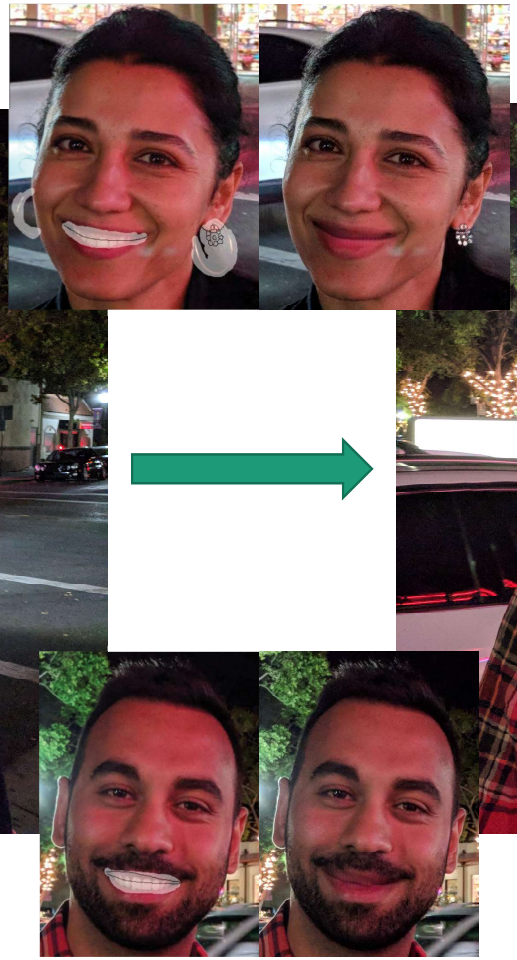
Original



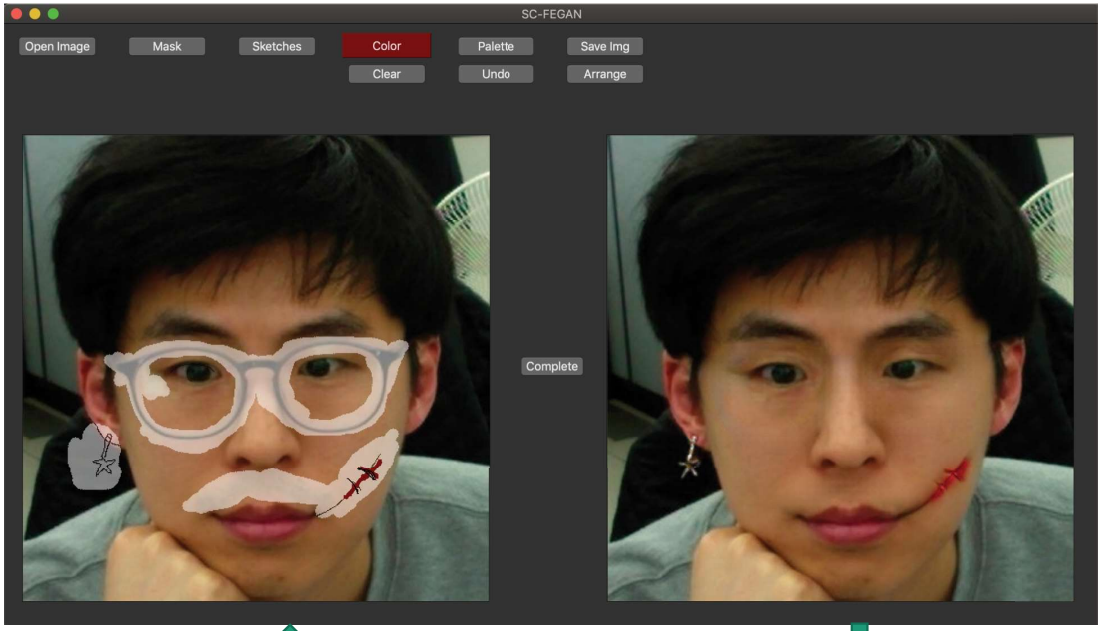
Editing

- **Results**
  - Editing HDR photo

Original



Editing


SC-FEGAN
Open Image
Mask
Sketches
Color
Clear
Palette
Undo
Save Img
Arrange
Complete

네 어느정도 됩니다.

제일 중요하다고 느낀 것들은?

1. Input & output 정의
2. Data 정제
3. Loss의 중요성

1. Input & output 정의
   - Free-form mask
   - Color & Sketch method
   2주
2. Data 정제
3. Loss의 중요성

1. Input & output 정의

2. Data 정제

- CelebA-HQ save
- Make domain
- Free-form algorithm
- Training batch

3달

3. Loss의 중요성

1. Input & output 정의

2. Data 정제

3. **Loss의 중요성** - **Perceptual loss**의 중요성! **2달**

구조적으로 중요한 부분은?

Generator

vs

Discriminator

Generator

vs

Discriminator

GAN의 최근 연구 방향은…

1. Loss
2. Normalization
3. SR

개발 과정은 여기까지…

2017 Imagenet challenge 2nd place

지금부터는 왜 Detection을 하다가

GAN을 하게 되었는지에 대해 이야기 해보겠습니다.

2017 ImageNet challenge

IMAGENET Large Scale Visual Recognition Challenge 2017 (ILSVRC2017)

DET LOC VID **Team information**

Legend:
Yellow background = winner in this task according to this metric; authors are willing to reveal the method
White background = authors are willing to reveal the method
Grey background = authors chose not to reveal the method
Italics = authors requested entry not participate in competition

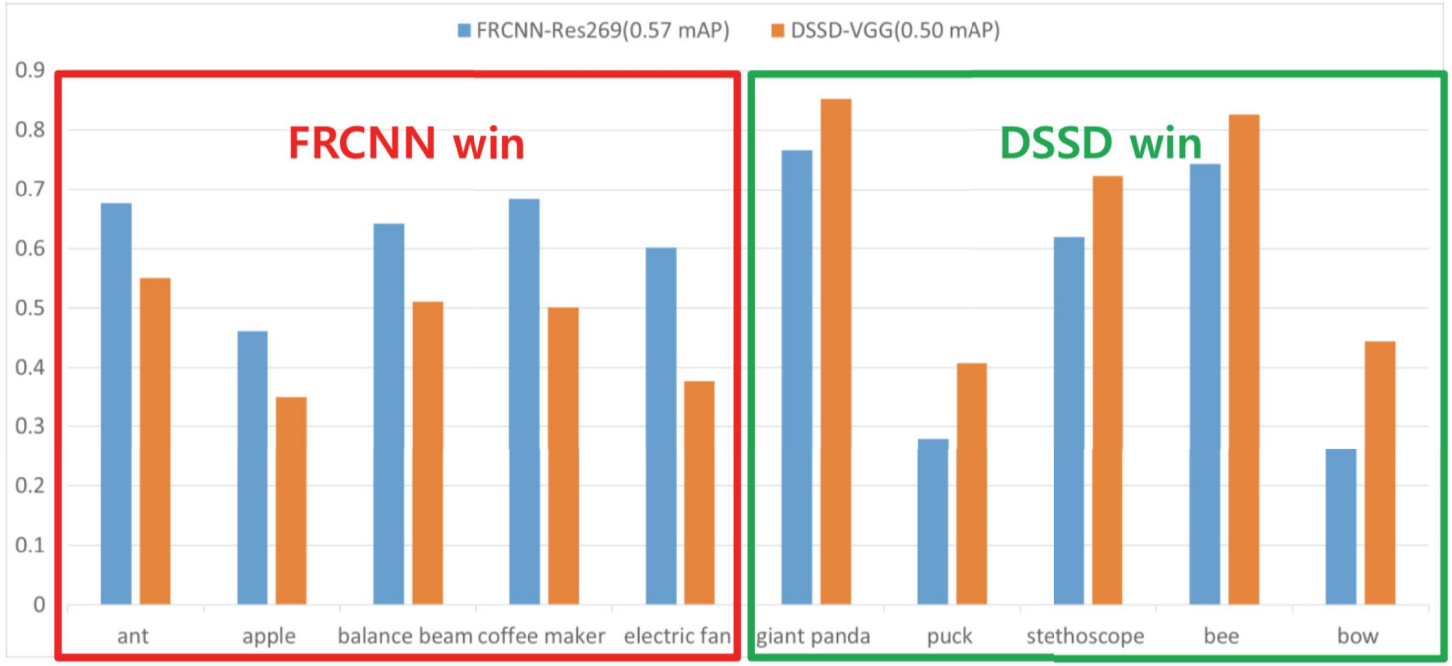
Object detection (DET)[top]

Task 1a: Object detection with provided training data

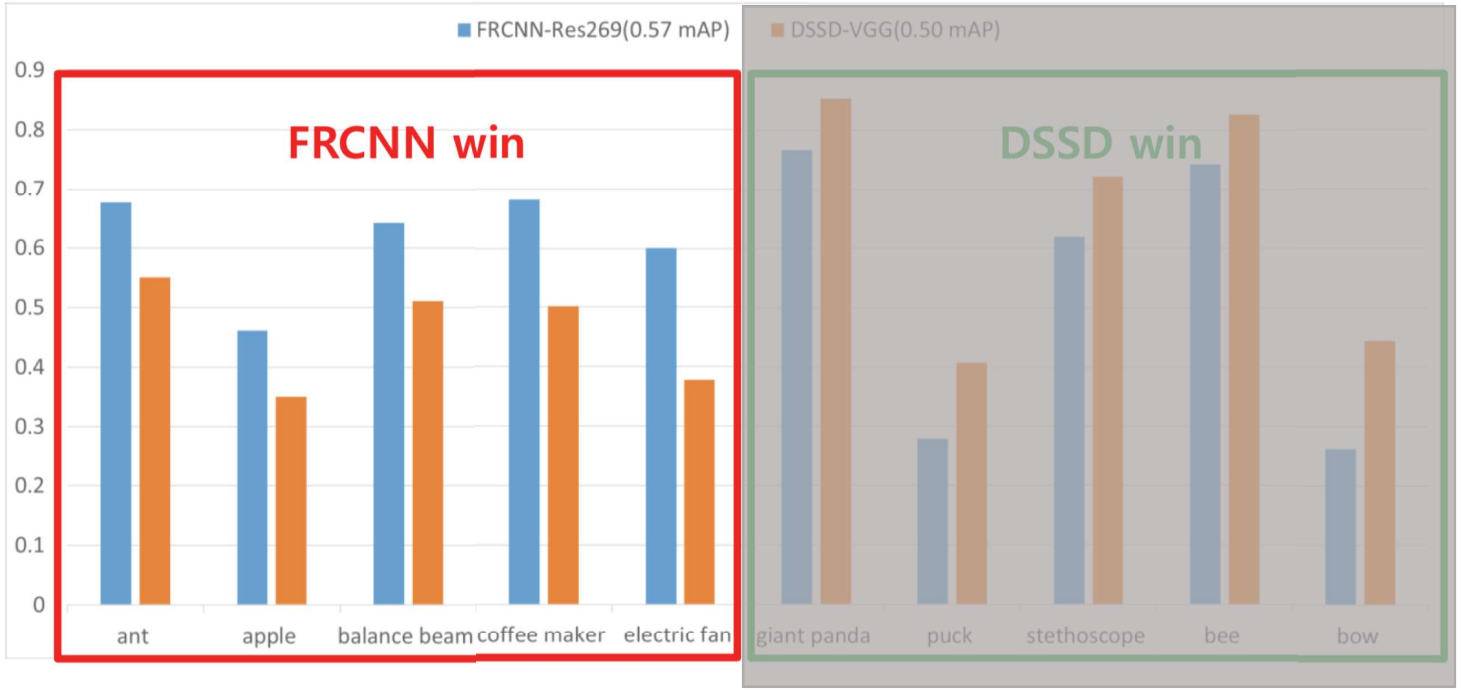
Ordered by number of categories won

Team name	Entry description	**Number of object categories won**	mean AP
BDAT	submission4	85	0.731392
BDAT	submission3	65	0.732227
BDAT	submission2	30	0.723712
DeepView(ETRI)	Ensemble_A	10	0.593084
NUS-Qihoo_DPNs (DET)	Ensemble of DPN models	9	0.656932

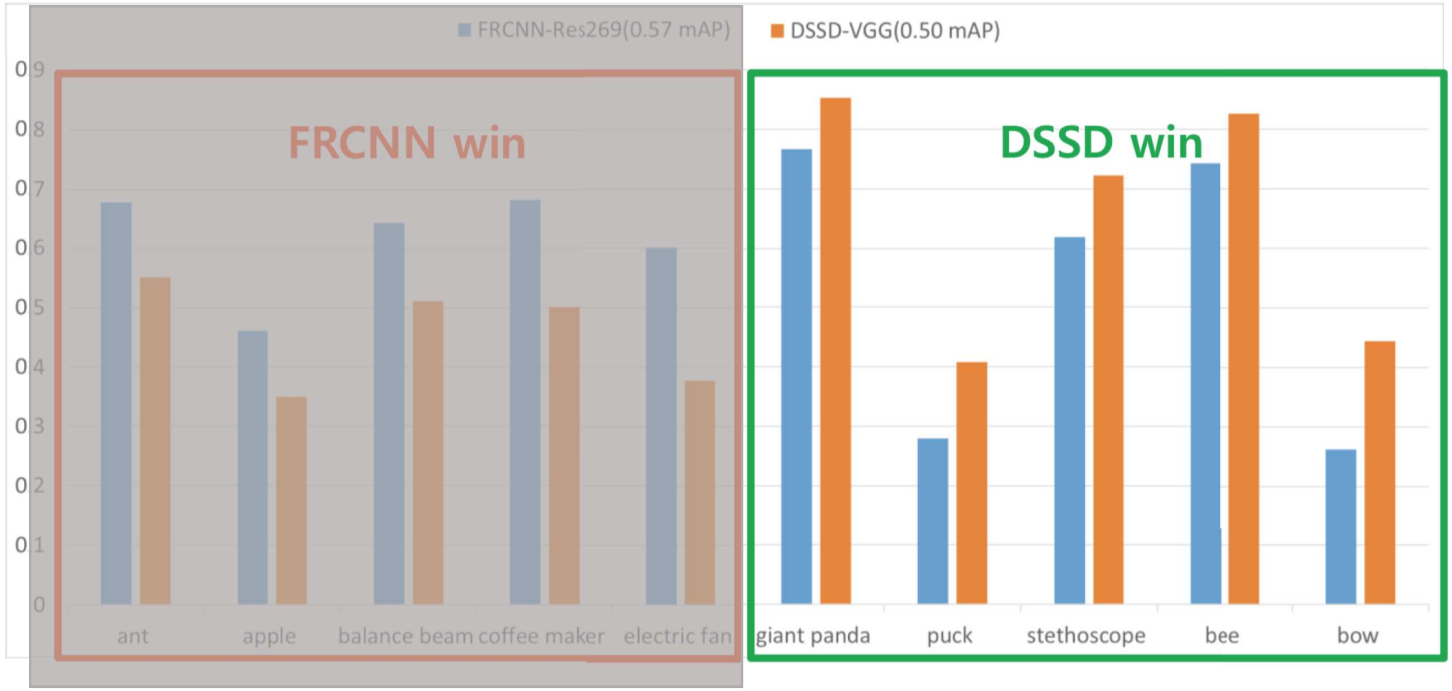
Rank of expert란?



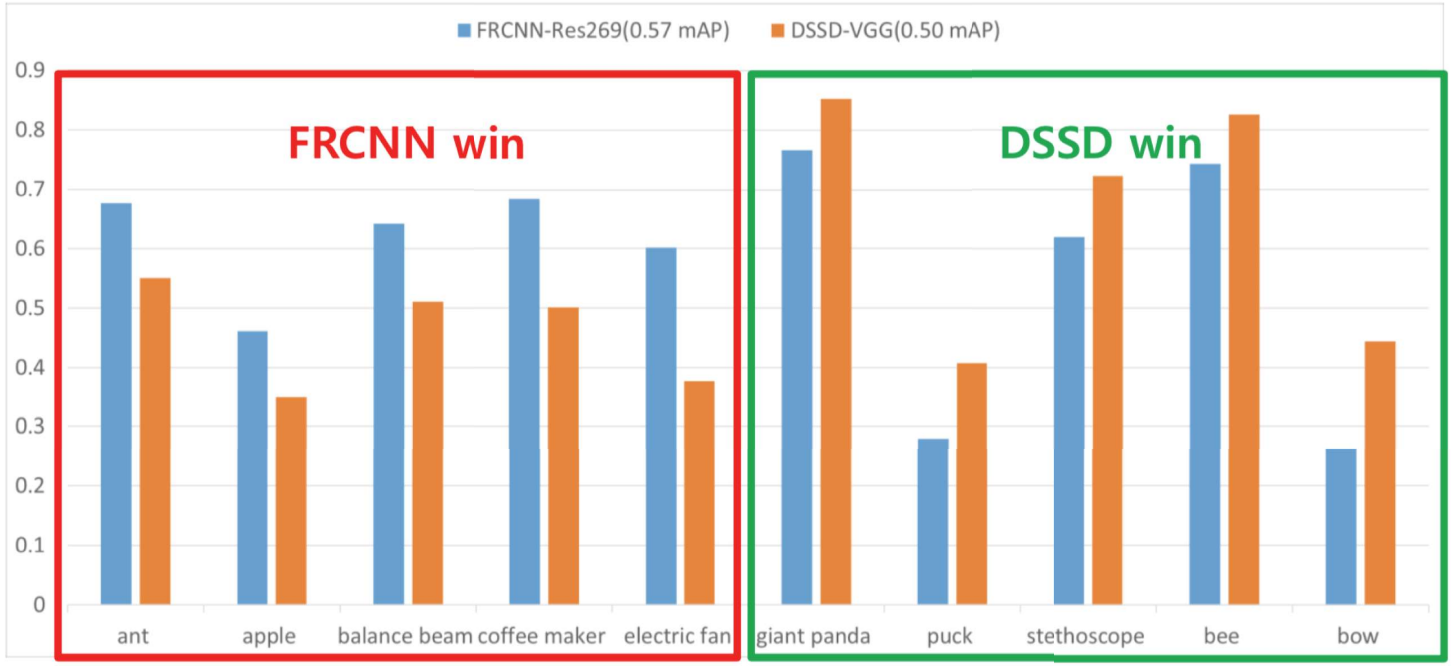
Rank of expert란?



Rank of expert란?

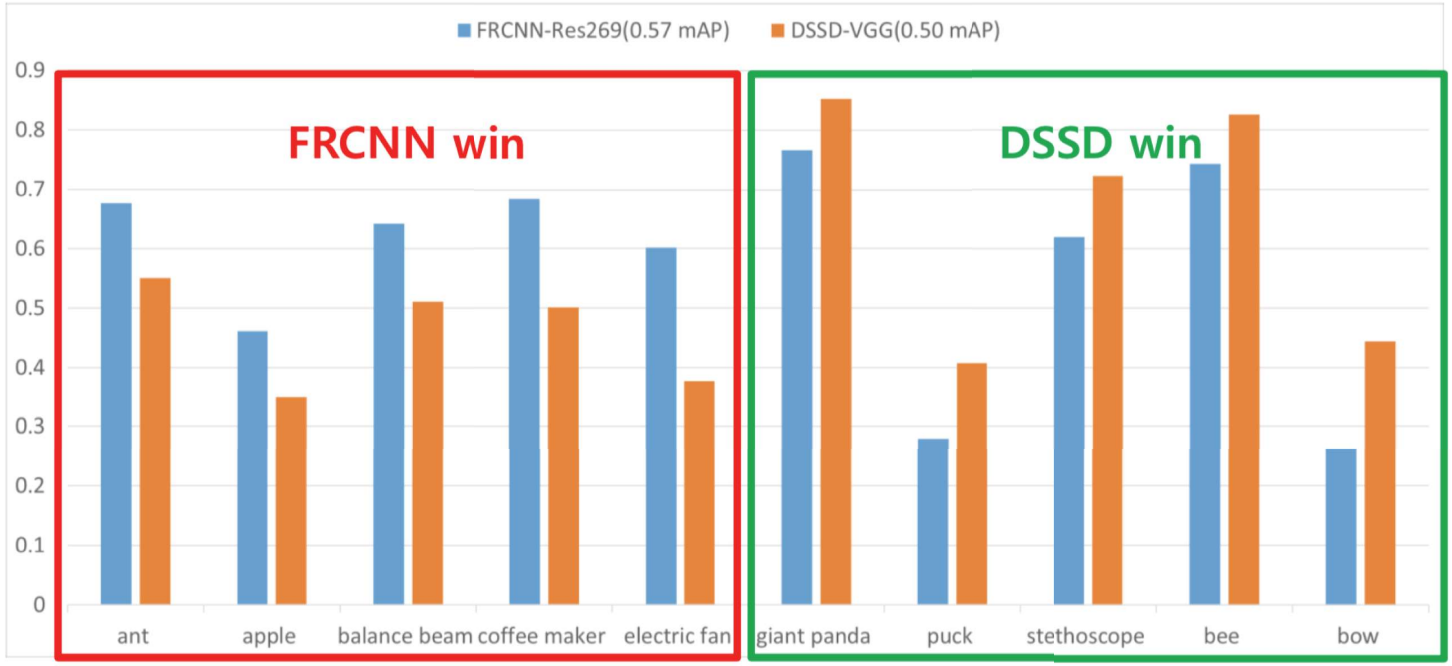


Rank of expert란?



Category 별로 성능이 높은 네트워크의 결과를 따르자!

Rank of expert란?



Category 별로 성능이 높은 네트워크의 결과를 따르자!

좋은 걸로 다 짬뽕하자!

19개 네트워크

Network	Feature extractor	Meta architecture	RoI warping	Training Dataset	mAP (# of selection) per image resolution				
					400	512	600	800	900
D1	Res101	FRCN-Type1	Pooling	train	50.07% (0)	-	53.13% (13)	53.25% (16)	52.08% (10)
D2	Res101	FRCN-Type1	Pooling	train+val1	49.79% (2)	-	53.57% (12)	53.32% (7)	51.96% (5)
D3	Res152	FRCN-Type1	Pooling	train+val1	52.16% (11)	-	55.77% (32)	54.94% (20)	53.59% (19)
D4	Res152	FRCN-Type1	Pooling	train	52.59% (13)	-	55.71% (23)	55.27% (26)	-
D5	Res152	FRCN-Type1	Pooling	train+val1	51.41% (6)	-	55.35% (32)	54.21% (19)	52.91% (11)
D6	Res152	FRCN-Type1	Pooling	train	51.92% (5)	-	56.00% (28)	55.41% (18)	54.38% (24)
D7	Res152	FRCN-Type1	Pooling	train+val1	52.41% (8)	-	56.19% (32)	55.54% (25)	54.34% (22)
D8	Res269	FRCN-Type1	Pooling	train+val1	-	-	56.92% (49)	-	-
D9	Res269	FRCN-Type1	Pooling	train+val1	54.09% **(24)**	-	57.65% (69)	56.29% (38)	54.94% (19)
D10	Res269	FRCN-Type1	Pooling	trainval+val1+aug	53.21% (17)	-	56.76% (43)	55.84% (34)	54.49% (20)
D11	Res269	FRCN-Type1	Pooling	trainval1	**53.98%** (23)	-	**57.72%** **(76)**	**56.64%** **(45)**	55.41% (26)
D12	Res269	FRCN-Type1	Pooling	train	53.59% (17)	-	57.34% (63)	56.56% (33)	**55.53%** (29)
D13	Res152	FRCN-Type2	Alingment	train+val1	48.91% (6)	-	54.65% (46)	54.53% (30)	54.49% **(42)**
D14	Res152	FRCN-Type2	Alingment	trainval+val1+aug	-	-	54.57% (40)	53.92% (29)	-
D15	Res152	FRCN-Type2	Alignment	train+val1	51.95% (8)	-	56.15% (35)	55.41% (25)	54.73% (20)
D16	Res152	FRCN-Type2	Alignment	trainval+val1+aug	43.42% (1)	-	51.39% (10)	49.00% (5)	47.40% (2)
D17	VGG	SSD	-	trainval+val1+aug	-	50.48% (14)	-	-	-
D18	VGG	DSSD	-	trainval+val1+aug	-	49.98% (15)	-	-	-
D19	WRI	SSD	-	trainval+val1+aug	-	49.21% (8)	-	-	-
Rank of Experts (for 19 detectors)							**62.54%**		

19개 네트워크

4개/1개의 image pyramid

Network	Feature extractor	Meta architecture	RoI warping	Training Dataset	mAP (# of selection) per image resolution				
					400	512	600	800	900
D1	Res101	FRCN-Type1	Pooling	train	50.07% (0)	-	53.13% (13)	53.25% (16)	52.08% (10)
D2	Res101	FRCN-Type1	Pooling	train+val1	49.79% (2)	-	53.57% (12)	53.32% (7)	51.96% (5)
D3	Res152	FRCN-Type1	Pooling	train+val1	52.16% (11)	-	55.77% (32)	54.94% (20)	53.59% (19)
D4	Res152	FRCN-Type1	Pooling	train	52.59% (13)	-	55.71% (23)	55.27% (26)	-
D5	Res152	FRCN-Type1	Pooling	train+val1	51.41% (6)	-	55.35% (32)	54.21% (19)	52.91% (11)
D6	Res152	FRCN-Type1	Pooling	train	51.92% (5)	-	56.00% (28)	55.41% (18)	54.38% (24)
D7	Res152	FRCN-Type1	Pooling	train+val1	52.41% (8)	-	56.19% (32)	55.54% (25)	54.34% (22)
D8	Res269	FRCN-Type1	Pooling	train+val1	-	-	56.92% (49)	-	-
D9	Res269	FRCN-Type1	Pooling	train+val1	54.09% **(24)**	-	57.65% (69)	56.29% (38)	54.94% (19)
D10	Res269	FRCN-Type1	Pooling	trainval+val1+aug	53.21% (17)	-	56.76% (43)	55.84% (34)	54.49% (20)
D11	Res269	FRCN-Type1	Pooling	trainval1	**53.98%** (23)	-	**57.72%** **(76)**	**56.64%** **(45)**	55.41% (26)
D12	Res269	FRCN-Type1	Pooling	train	53.59% (17)	-	57.34% (63)	56.56% (33)	**55.53%** (29)
D13	Res152	FRCN-Type2	Alingment	train+val1	48.91% (6)	-	54.65% (46)	54.53% (30)	54.49% **(42)**
D14	Res152	FRCN-Type2	Alingment	trainval+val1+aug	-	-	54.57% (40)	53.92% (29)	-
D15	Res152	FRCN-Type2	Alignment	train+val1	51.95% (8)	-	56.15% (35)	55.41% (25)	54.73% (20)
D16	Res152	FRCN-Type2	Alignment	trainval+val1+aug	43.42% (1)	-	51.39% (10)	49.00% (5)	47.40% (2)
D17	VGG	SSD	-	trainval+val1+aug	-	50.48% (14)	-	-	-
D18	VGG	DSSD	-	trainval+val1+aug	-	49.98% (15)	-	-	-
D19	WRI	SSD	-	trainval+val1+aug	-	49.21% (8)	-	-	-
Rank of Experts (for 19 detectors)							**62.54%**		

19개 네트워크 → 16x4+3 =67번 실험 데이터 ← 4개/1개의 image pyramid

Network	Feature extractor	Meta architecture	RoI warping	Training Dataset	mAP (# of selection) per image resolution 400	512	600	800	900
D1	Res101	FRCN-Type1	Pooling	train	50.07% (0)	-	53.13% (13)	53.25% (16)	52.08% (10)
D2	Res101	FRCN-Type1	Pooling	train+val1	49.79% (2)	-	53.57% (12)	53.32% (7)	51.96% (5)
D3	Res152	FRCN-Type1	Pooling	train+val1	52.16% (11)	-	55.77% (32)	54.94% (20)	53.59% (19)
D4	Res152	FRCN-Type1	Pooling	train	52.59% (13)	-	55.71% (23)	55.27% (26)	-
D5	Res152	FRCN-Type1	Pooling	train+val1	51.41% (6)	-	55.35% (32)	54.21% (19)	52.91% (11)
D6	Res152	FRCN-Type1	Pooling	train	51.92% (5)	-	56.00% (28)	55.41% (18)	54.38% (24)
D7	Res152	FRCN-Type1	Pooling	train+val1	52.41% (8)	-	56.19% (32)	55.54% (25)	54.34% (22)
D8	Res269	FRCN-Type1	Pooling	train+val1	-	-	56.92% (49)	-	-
D9	Res269	FRCN-Type1	Pooling	train+val1	54.09% **(24)**	-	57.65% (69)	56.29% (38)	54.94% (19)
D10	Res269	FRCN-Type1	Pooling	trainval+val1+aug	53.21% (17)	-	56.76% (43)	55.84% (34)	54.49% (20)
D11	Res269	FRCN-Type1	Pooling	trainval1	**53.98%** (23)	-	**57.72%** **(76)**	**56.64%** **(45)**	55.41% (26)
D12	Res269	FRCN-Type1	Pooling	train	53.59% (17)	-	57.34% (63)	56.56% (33)	**55.53%** (29)
D13	Res152	FRCN-Type2	Alingment	train+val1	48.91% (6)	-	54.65% (46)	54.53% (30)	54.49% **(42)**
D14	Res152	FRCN-Type2	Alingment	trainval+val1+aug	-	-	54.57% (40)	53.92% (29)	-
D15	Res152	FRCN-Type2	Alignment	train+val1	51.95% (8)	-	56.15% (35)	55.41% (25)	54.73% (20)
D16	Res152	FRCN-Type2	Alignment	trainval+val1+aug	43.42% (1)	-	51.39% (10)	49.00% (5)	47.40% (2)
D17	VGG	SSD	-	trainval+val1+aug	-	50.48% (14)	-	-	-
D18	VGG	DSSD	-	trainval+val1+aug	-	49.98% (15)	-	-	-
D19	WRI	SSD	-	trainval+val1+aug	-	49.21% (8)	-	-	-
Rank of Experts (for 19 detectors)							**62.54%**		

약 120만장의 이미지 데이터

67번의 실험 데이터

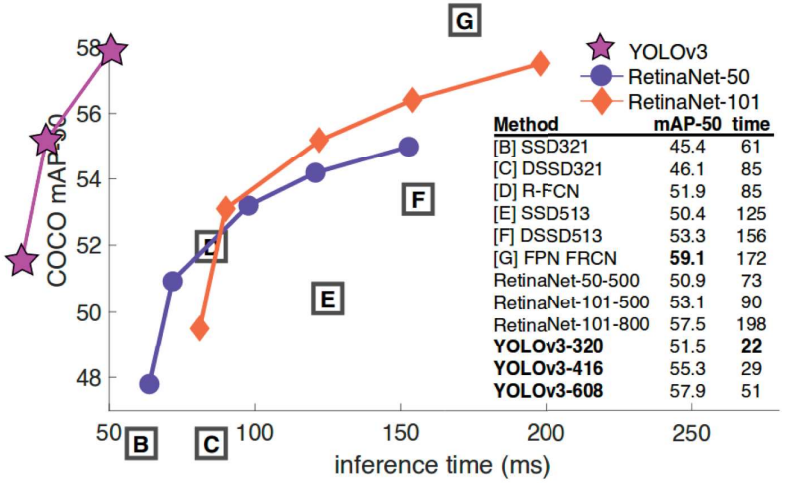
6명 x 12시간 x 6달

고생은 고생대로 시간은 시간대로…

상용화 가능한 기술인가?

그 후…

- 2017
  - Mask RCNN (facebook)
  - Yolov2
  - DenseNet
- 2018
  - Yolov3
  - RetinaNet (facebook)
- 2019
  - YOLACT



Method	mAP-50	time
[B] SSD321	45.4	61
[C] DSSD321	46.1	85
[D] R-FCN	51.9	85
[E] SSD513	50.4	125
[F] DSSD513	53.3	156
[G] FPN FRCN	**59.1**	172
RetinaNet-50-500	50.9	73
RetinaNet-101-500	53.1	90
RetinaNet-101-800	57.5	198
YOLOv3-320	51.5	**22**
YOLOv3-416	55.3	29
YOLOv3-608	57.9	51

Accurate, Large Minibatch SGD: Training ImageNet in 1 Hour

Priya Goyal Piotr Dollár Ross Girshick Pieter Noordhuis
Lukasz Wesolowski Aapo Kyrola Andrew Tulloch Yangqing Jia Kaiming He

Facebook

Abstract

Deep learning thrives with large neural networks and large datasets. However, larger networks and larger datasets result in longer training times that impede research and development progress. Distributed synchronous SGD offers a potential solution to this problem by dividing SGD minibatches over a pool of parallel workers. Yet to make this scheme efficient, the per-worker workload must be large, which implies nontrivial growth in the SGD minibatch size. In this paper, we empirically show that on the ImageNet dataset large minibatches cause optimization difficulties, but when these are addressed the trained networks exhibit good generalization. Specifically, we show no loss of accuracy when training with large minibatch sizes up to 8192 images. To achieve this result, we adopt a hyper-parameter-free linear scaling rule for adjusting learning rates as a function of minibatch size and develop a new warmup scheme that overcomes optimization challenges early in training. With these simple techniques, our Caffe2-based system trains ResNet-50 with a minibatch size of 8192 on 256 GPUs in one hour, while matching small minibatch accuracy. Using commodity hardware, our implementation achieves ~90% scaling efficiency when moving from 8 to 256 GPUs. Our findings enable training visual recognition models on internet-scale data with high efficiency.

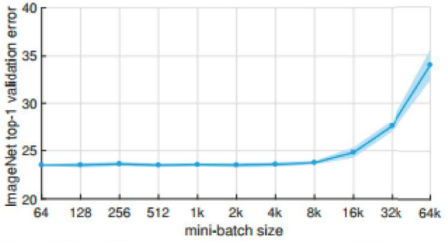
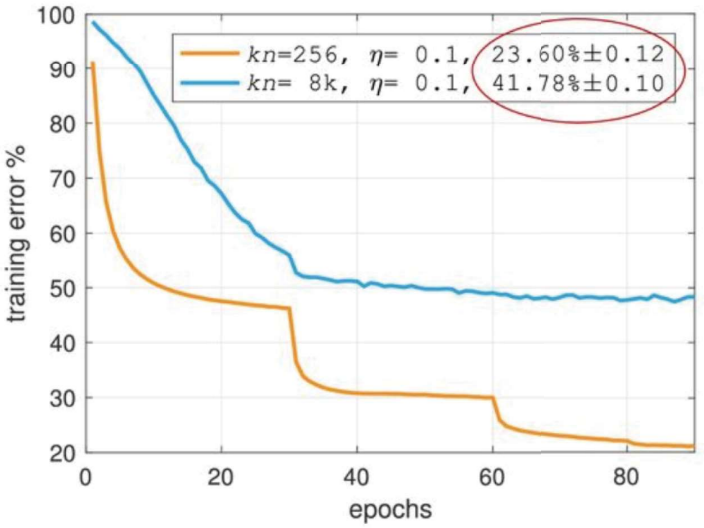


Figure 1. **ImageNet top-1 validation error *vs.* minibatch size.** Error range of plus/minus *two* standard deviations is shown. We present a simple and general technique for scaling distributed synchronous SGD to minibatches of up to 8k images *while maintaining the top-1 error of small minibatch training*. For all minibatch sizes we set the learning rate as a *linear* function of the minibatch size and apply a simple warmup phase for the first few epochs of training. All other hyper-parameters are kept fixed. Using this simple approach, accuracy of our models is invariant to minibatch size (up to an 8k minibatch size). Our techniques enable a linear reduction in training time with ~90% efficiency as we scale to large minibatch sizes, allowing us to train an accurate 8k minibatch ResNet-50 model in 1 hour on 256 GPUs.

tation [8, 10, 28]. Moreover, this pattern generalizes: larger datasets and neural network architectures consistently yield improved accuracy across all tasks that benefit from pre-

Naïve Scaling



8192 = 256 × 32
#gpus
per gpu batch

facebook research

이런 선두 그룹이 엄청난 머신 파워도 지니고 있는데…

공개도 다하는데…

6명 x 12시간 x 6달…??? 해야 되나???

네 그래서 안하기로 했습니다.^^

그 후의 고민

CV 분야에서

현재 가장 많이 알려진 활용 방법

CCTV

자율 자동차

얼굴인식

영상 분석

CCTV	감시감독원을 대체
자율 자동차	운전자를 대체
얼굴인식	감시감독원을 **보조**
영상 분석	분석가(의사 등)를 **보조** L 유니크한 데이터를 활용 가능하다면 굿!

사람을 대체하는 것이 아닌 사람이 사용하는 Tool의 역할에 집중

인공지능 스피커/비서가 가장 빠르게 상용화된 이유라고 생각

Tool의 역할 : 보조

완성도가 완벽하면 좋겠지만 떨어져도, 사용자가 다시 입력하기 편해야…

CV전공자로서 그래서 디자인 쪽을 타겟팅 해보았습니다.

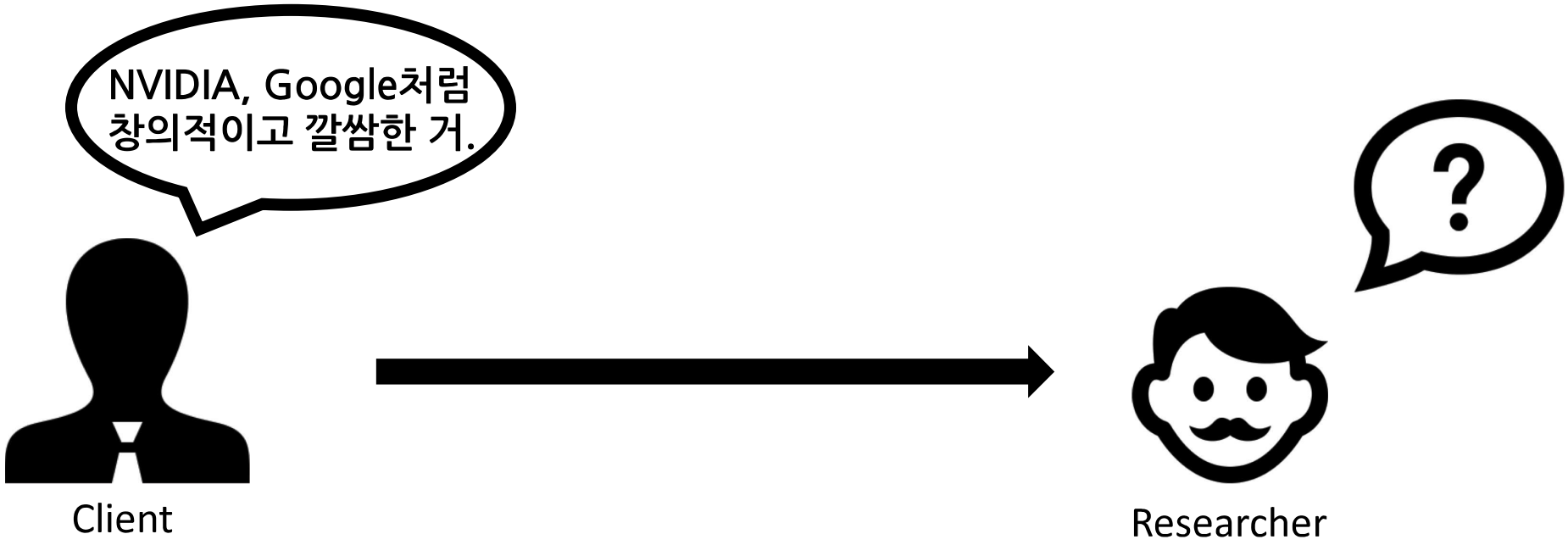
그래서 ㄱㄱㅅ 하면서 GUI를 만들어서 같이 공개를…

누구에게 도움이 되는지

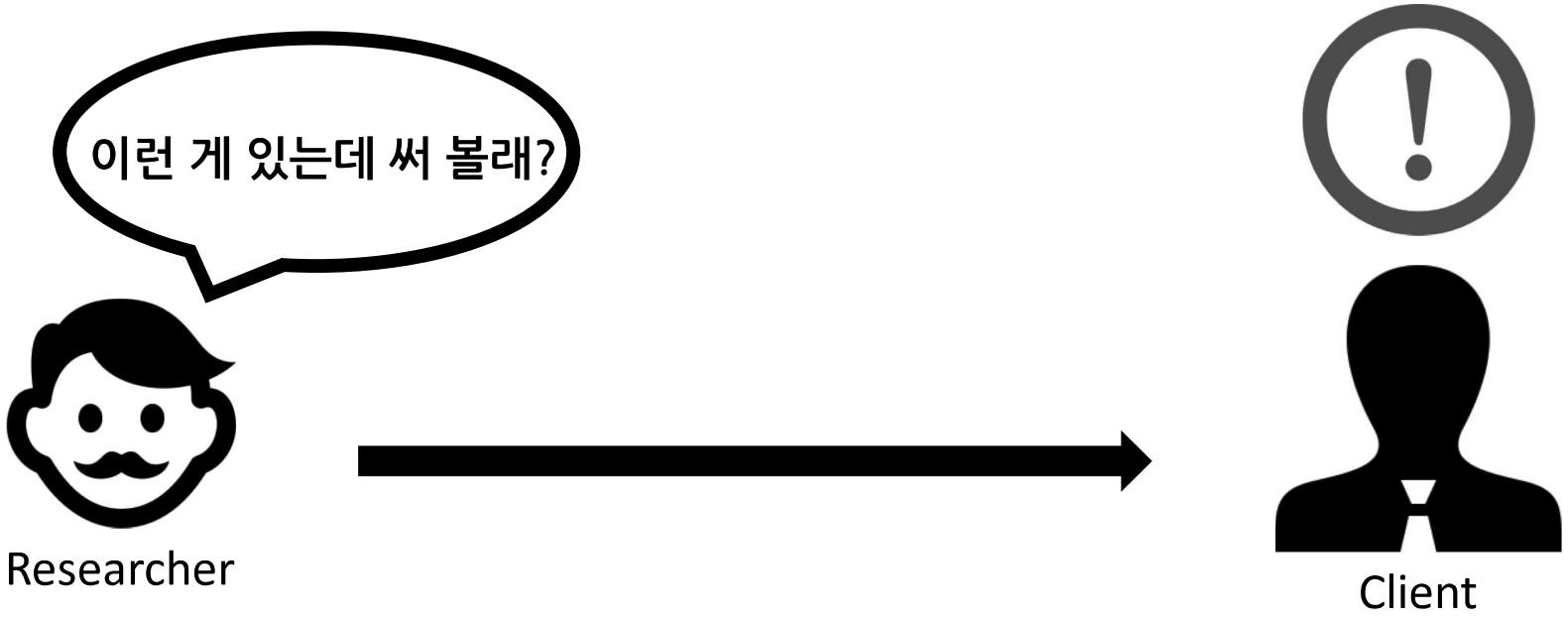
어떻게 도움이 되는지

필요한 공부는 무엇인지

고민하면서 연구하면


NVIDIA, Google처럼
창의적이고 깔쌈한 거.
?
Client
Researcher

이런 날에서


이런 게 있는데 써 볼래?
Researcher
Client

이런 날이 오길 희망해 봅니다

run.youngjoo