Part 5. 생성

2.α 생성 모델링 그리고 미래

2014 2015 2016 2017 2018 2019

- VAE
- 인코더-디코더
- GAN
- 어텐션

- U-Net
- Pointer Network
- 뉴럴 스타일
- ResNet

- pix2pix
- WGAN
- WGAN GP
- CycleGAN
- 트랜스포머
- MuseGAN
- Progressice GAN

- World Models
- Self-Attention GAN
- BigGAN
- BERT
- StyleGAN

- GPT-2
- MuseNet

GAN에서의 트랜스포머는 순환 신경망이나
합성곱 신경망을 완전히 제거한 어텐션 기반의 모듈

트랜스포머는 BERT, GPT-2, MuseNet과 같이
최고의 성능을 내는 시퀀스 모델의 기반

※ 디코더에서 출력 단어를 예측하는 매 시점마다,
인코더에서의 전체 입력 문장을 다시 한 번 참고.
단, 전체 입력 문장을 전부 다 동일한 비율로 참고하는 것이 아닌,
해당 시점에서 예측해야할 단어와 연관이 있는 입력 단어 부분을
좀 더 집중해서 보게 되는 구조.

2.α 생성 모델링 그리고 미래

2014 2015 2016 2017 2018 2019

- VAE
- 인코더-디코더
- GAN
- 어텐션

- U-Net
- Pointer Network
- 뉴럴 스타일
- ResNet

- pix2pix

- WGAN
- WGAN GP
- CycleGAN

- 트랜스포머

Attention Is All You Need

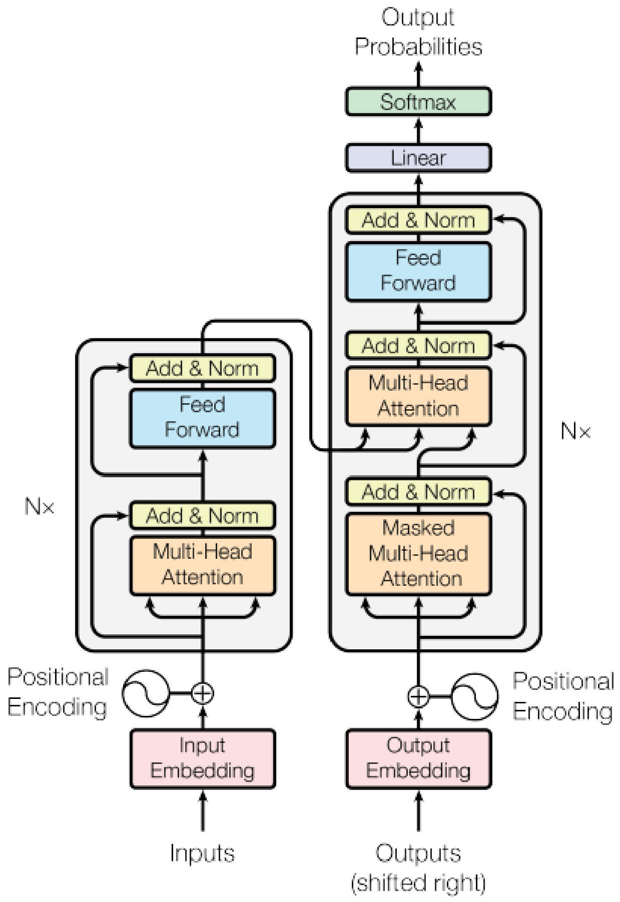


Figure 1: The Transformer - model architecture.

✓ 복잡한 순환 구조, 합성곱 구조를 사용하지 않고 Sequential modeling 위한 강력한 신경망을 만들 수 있는 방법을 소개. 그 대신 오로지 어텐션 메커니즘만 사용

✓ 인코더와 디코더에 lstm과 같은 순환 층을 사용하지 않고 어텐션 층을 쌓아 사용

3.1 Encoder and Decoder Stacks

Encoder: The encoder is composed of a stack of $N = 6$ identical layers. Each layer has two sub-layers. The first is a multi-head self-attention mechanism, and the second is a simple, position-wise fully connected feed-forward network. We employ a residual connection [11] around each of the two sub-layers, followed by layer normalization [1]. That is, the output of each sub-layer is $\mathrm{LayerNorm}(x + \mathrm{Sublayer}(x))$, where $\mathrm{Sublayer}(x)$ is the function implemented by the sub-layer itself. To facilitate these residual connections, all sub-layers in the model, as well as the embedding layers, produce outputs of dimension $d_{\text{model}} = 512$.

Decoder: The decoder is also composed of a stack of $N = 6$ identical layers. In addition to the two sub-layers in each encoder layer, the decoder inserts a third sub-layer, which performs multi-head attention over the output of the encoder stack. Similar to the encoder, we employ residual connections around each of the sub-layers, followed by layer normalization. We also modify the self-attention sub-layer in the decoder stack to prevent positions from attending to subsequent positions. This masking, combined with fact that the output embeddings are offset by one position, ensures that the predictions for position i can depend only on the known outputs at positions less than i.

https://colab.research.google.com/github/tensorflow/tensor2tensor/blob/master/tensor2tensor/notebooks/hello_t2t.ipynb#scrollTo=OJKU36QAfqOC

Part 5. 생성

2.α 생성 모델링 그리고 미래

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

Jacob Devlin Ming-Wei Chang Kenton Lee Kristina Toutanova
Google AI Language
{jacobdevlin,mingweichang,kentonl,kristout}@google.com

2014 2015 2016 2017 2018 2019

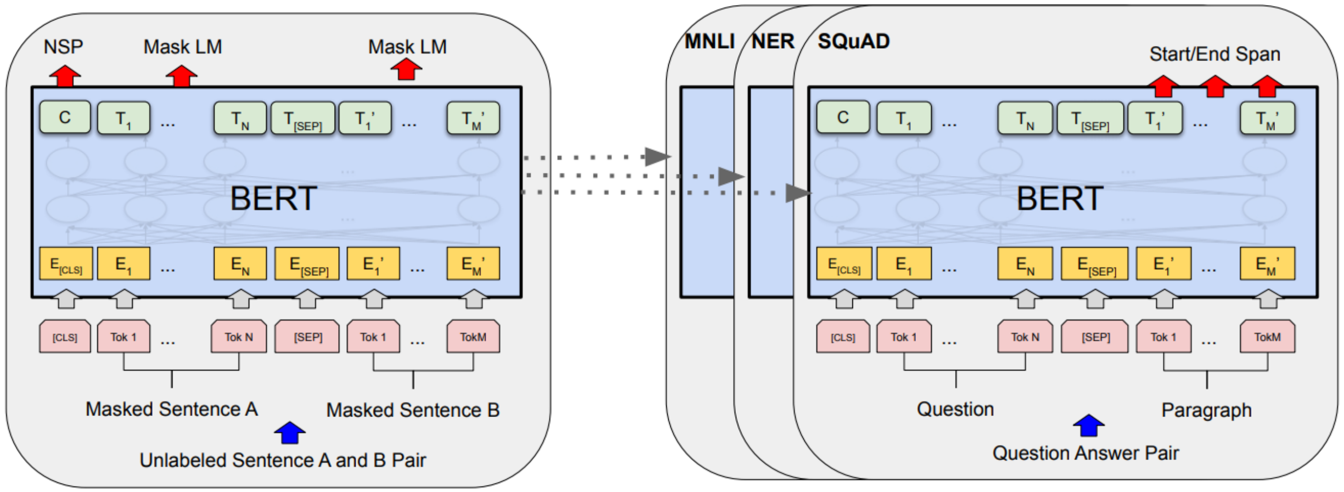
2014
- VAE
- 인코더-디코더
- GAN
- 어텐션

2015
- U-Net
- Pointer Network
- 뉴럴 스타일
- ResNet

2016
- pix2pix

2017
- WGAN
- WGAN GP
- CycleGAN
- 트랜스포머
- MuseGAN
- Progressice GAN

2018
- World Models
- Self-Attention GAN
- BigGAN
- **BERT**



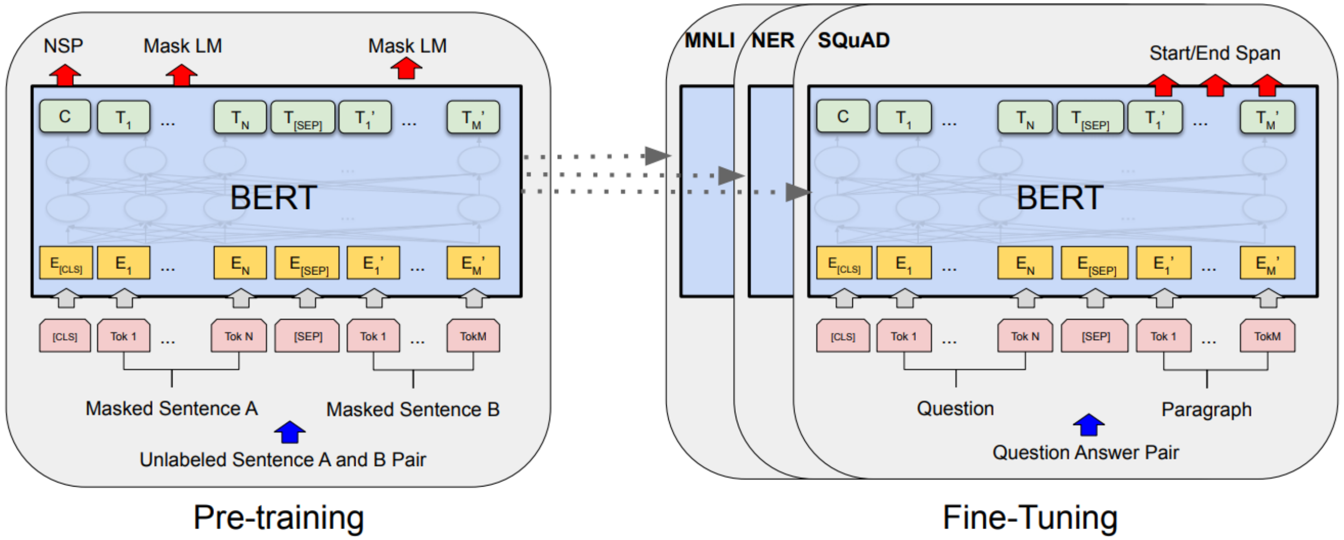
Pre-training Fine-Tuning

✓ 모든 층에 누락된 단어, 전후 맥락을 주면 문장에서 누락된 단어를 예측

✓ MLM 이라고 하는 Masked Language Model 사용

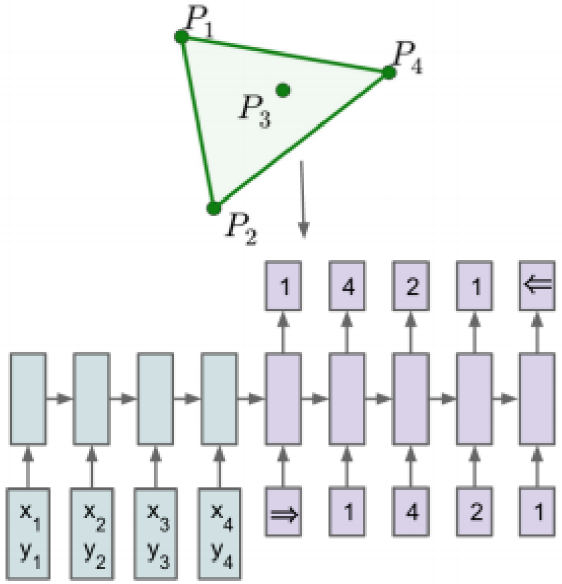
- 훈련하는 동안 15%의 단어를 랜덤하게 마스킹 처리
- 모델은 마스킹된 입력에서 원본문장을 재생성
- 마스킹된 토큰중 10%는 <MASK> 토근 대신 다른 단어로 변환.
- 이 모델은 <MASK> 토큰을 실제 단어로 바꾸는 방법을 학습뿐 아니라 입력 문장에서 다른 것으로 바뀐 적합하지 않은 단어 또한 찾아냄.

- Rather than *always* replacing the chosen words with [MASK], the data generator will do the following:
- 80% of the time: Replace the word with the [MASK] token, e.g., `my dog is hairy` → `my dog is [MASK]`
- 10% of the time: Replace the word with a random word, e.g., `my dog is hairy` → `my dog is apple`
- 10% of the time: Keep the word unchanged, e.g., `my dog is hairy` → `my dog is hairy`. The purpose of this is to bias the representation towards the actual observed word.



✚ *Downstream task Output layer*

ex) 감성 분석 같은 분류작업에는 언어 모델 출력 값에 긍정,부정처럼 분류층을 추가
마찬가지로『질문 – 대답』작업에서 ***pointer network***를 BERT의 *Outputlayer*로 사용하여 입력 시퀀스에 대한 대답을 표시



(a) Sequence-to-Sequence

GloVe ***'water'***

문맥에 상관없이 표현을 ***water***에 할당

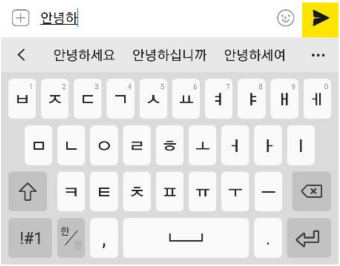
BERT ***'water'***

단어 주위 정보를 사용하여 동사 혹은 명사 중에 문맥에 맞는 적합한 단어표현을 생성

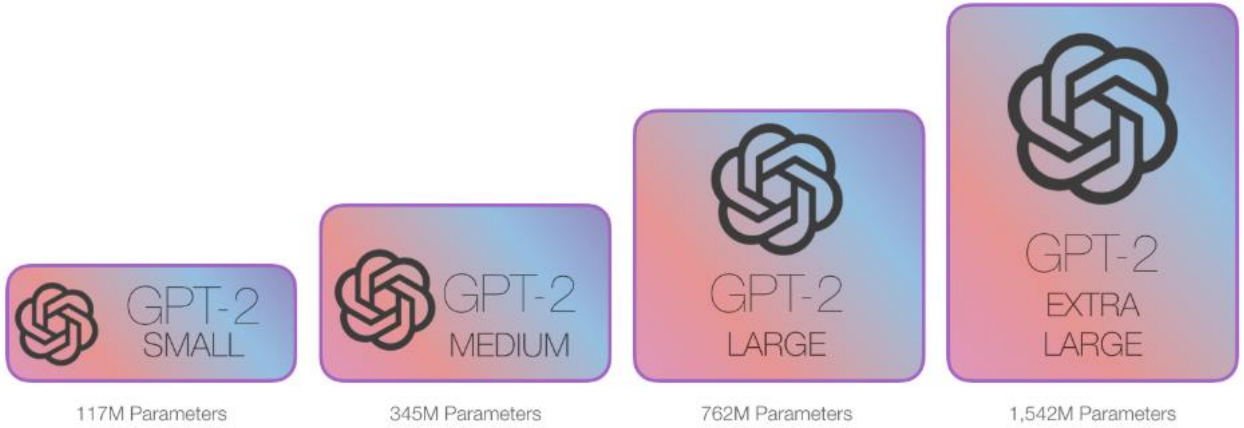
어텐션 메커니즘을 가진 seq2seq 구조로, 입력의 인덱스를 출력.
출력 단어가 시퀀스 길이에 따라 달라지게 돼서, 다양한 크기의 입력을 다룰 수 있다.

2.α 생성 모델링 그리고 미래

- ✓ Open AI에서 개발한 텍스트 문장에서 다음 단어를 예측하도록 훈련된 모델.
- ✓ GPT-2는 단방향



- ✓ 기본적인 문장 생성 모델은 문장의 일부를 보고 다음 단어를 예측할 수 있는것이 일반적



- ✓ WebText라는 40GB 데이터 세트로 훈련
- ✓ 훈련된 GPT-2의 가장 작은 변형은 모든 매개 변수를 저장하기 위해 500MB의 용량을 차지
- ✓ 가장 큰 GPT-2는 13배 크기를 가진 6.5GB 이상의 저장 공간

2.α 생성 모델링 그리고 미래

AI2 Allen Institute for AI

AllenNLP

- Answer a question
  - Reading Comprehension
- Annotate a sentence
  - Named Entity Recognition
  - Open Information Extraction
  - Sentiment Analysis
  - Dependency Parsing
  - Constituency Parsing
  - Semantic Role Labeling

Language Modeling

Language modeling is the task of determining the probability of a given sequence of words occurring in a sentence.

This demonstration uses the public 345M parameter OpenAI GPT-2 language model to generate sentences.

Provide some initial text, and the model will generate a list of the most-likely next words. You can click on one of those candidate words to choose it and continue, or you can keep typing. Click the left arrow at the bottom to undo your last choice.

Sentence:

AllenNLP is

Predictions:

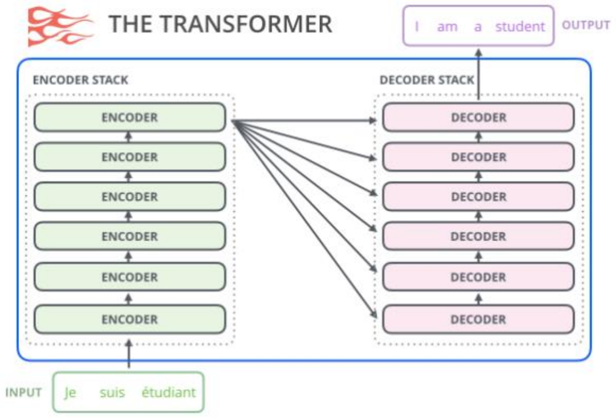
21.5% a
5.7% the
5.3% an
4.0% not
1.5% one
← Undo

Sentence:

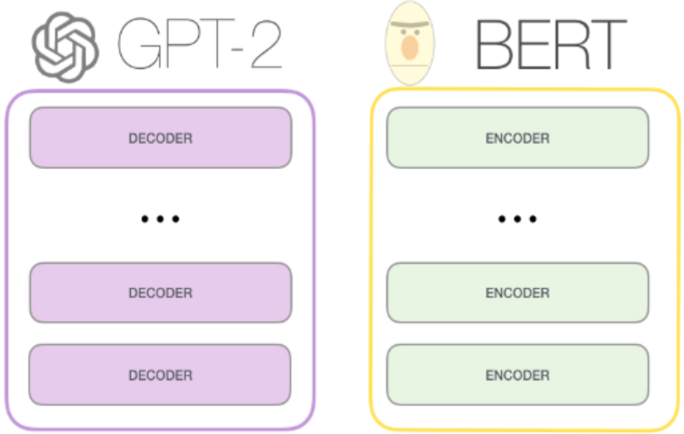
AllenNLP is a member of the National Association of Realtors.

https://demo.allennlp.org/next-token-lm?text=AllenNLP%20is%20a%20member%20of%20the%20National%20Association%20of%20Realtors.

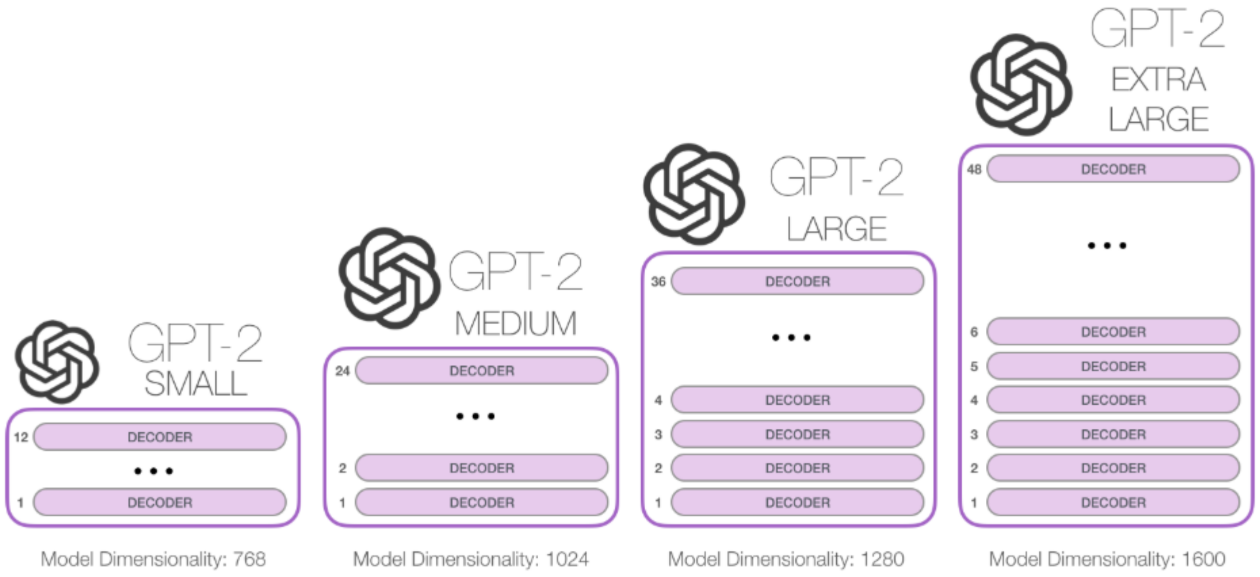
2.α 생성 모델링 그리고 미래



- ✓ 트랜스포머 모델은 인코더와 디코더로 구성되어



- ✓ GPT-2는 트랜스포머 디코더 블록을 사용하여 제작
  반면 BERT는 트랜스포머의 인코더 블록을 사용
- ✓ 둘 사이의 중요한 차이점은 GPT2가 전통적인 언어 모델처럼
  한 번에 하나의 토큰을 출력.



- ✓ 후속 연구에서는 트렌스포머 블록을 한 타입만 사용
- ✓ 실제로 가능한 한 높게 쌓고 방대한 양의 학습 텍스트를 제공

https://openai.com/blog/better-language-models/#sample1

2.α 생성 모델링 그리고 미래

GPT-2

SYSTEM PROMPT (HUMAN-WRITTEN)

In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

MODEL COMPLETION (MACHINE-WRITTEN, 10 TRIES)

The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science.

Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved.

Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow.

Pérez and the others then ventured further into the valley. "By the time we reached the top of one peak, the water looked blue, with some crystals on top," said Pérez.

Pérez and his friends were astonished to see the unicorn herd. These creatures could be seen from the air without having to move too much to see them – they were so close they could touch their horns.

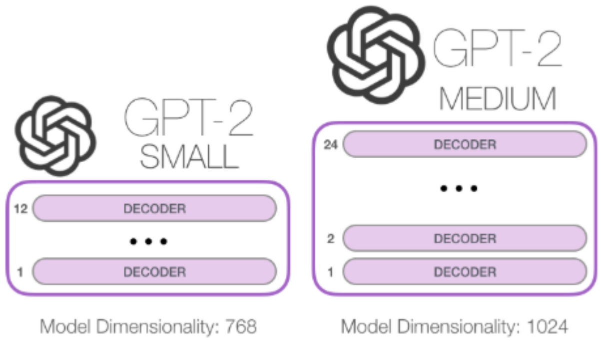
While examining these bizarre creatures the scientists discovered that the creatures also spoke some fairly regular English. Pérez stated, "We can see, for example, that they have a common 'language,' something like a dialect or dialectic."

Dr. Pérez believes that the unicorns may have originated in Argentina, where the animals were believed to be descendants of a lost race of people who lived there before the arrival of humans in those parts of South America.

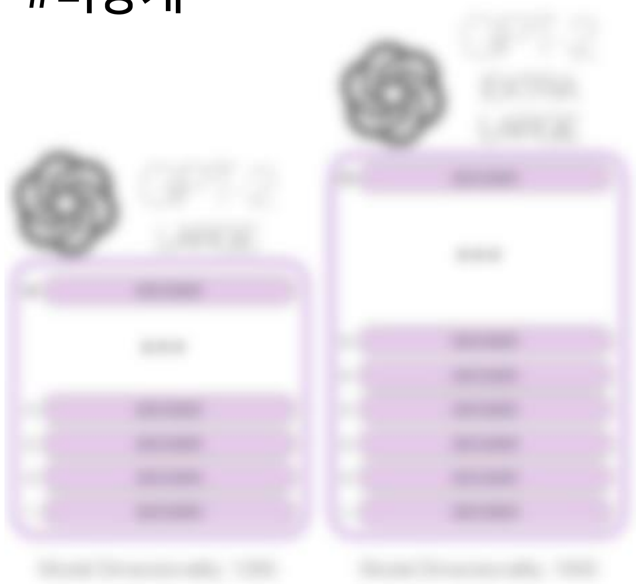
While their origins are still unclear, some believe that perhaps the creatures were created when a human and a unicorn met each other in a time before human civilization. According to Pérez, "In South America, such incidents seem to be quite common."

However, Pérez also pointed out that it is likely that the only way of knowing for sure if unicorns are indeed the descendants of a lost alien race is through DNA. "But they seem to be able to communicate in English quite well, which I believe is a sign of evolution, or at least a change in social organization," said the scientist.

#공개



#비공개



2.α 생성 모델링 그리고 미래

PROGRESSIVE GROWING OF GANS FOR IMPROVED QUALITY, STABILITY, AND VARIATION

Tero Karras NVIDIA **Timo Aila** NVIDIA **Samuli Laine** NVIDIA **Jaakko Lehtinen** NVIDIA and Aalto University
{tkarras,taila,slaine,jlehtinen}@nvidia.com

- ➢ GAN 훈련의 속도와 안정성을 향상시키기 위해 개발한 새로운 기술

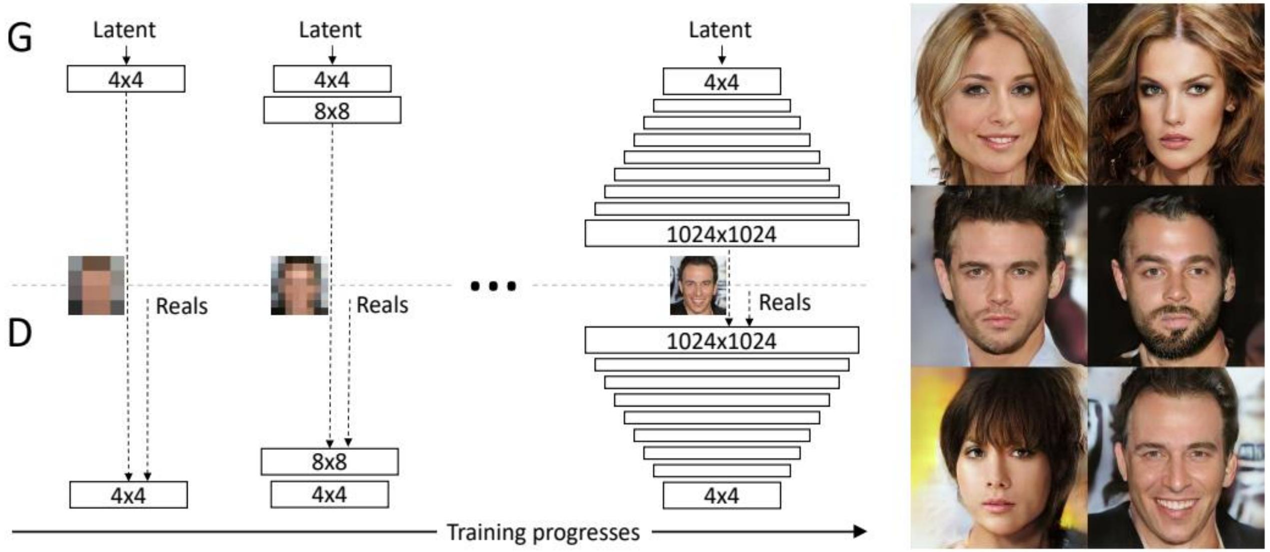


Figure 1: Our training starts with both the generator (G) and discriminator (D) having a low spatial resolution of 4×4 pixels. As the training advances, we incrementally add layers to G and D, thus increasing the spatial resolution of the generated images. All existing layers remain trainable throughout the process. Here $\boxed{N \times N}$ refers to convolutional layers operating on $N \times N$ spatial resolution. This allows stable synthesis in high resolutions and also speeds up training considerably. One the right we show six example images generated using progressive growing at 1024 × 1024.

- ✓ 고해상도 이미지에 GAN을 직접 훈련하는 대신
  4X4 픽셀의 초저해상도 이미지에서 생성자와 판별자를 훈련
- ✓ 점차 훈련을 진행하면서 층을 추가하고 해상도를 높여 간다.
- ✓ 추가된 층이 훈련 과정에서 동결되지 않고 계속 전체가 훈련.

Figure 7: Selection of 256 × 256 images generated from different LSUN categories.

https://arxiv.org/pdf/1710.10196.pdf

2.α 생성 모델링 그리고 미래

Self-Attention Generative Adversarial Networks

Han Zhang [1,2] Ian Goodfellow [2] Dimitris Metaxas [1] Augustus Odena [2]

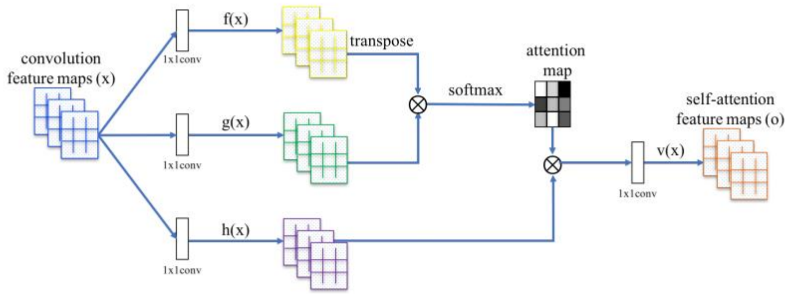
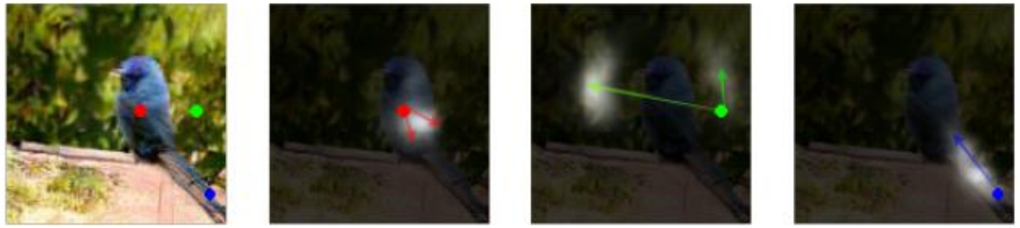


Figure 2. The proposed self-attention module for the SAGAN. The ⊗ denotes matrix multiplication. The softmax operation is performed on each row.



- ✓ 트랜스포머와 같이 순차 모델에 사용되는 어텐션 메커니즘을 이미지 생성을 위한 gan 기반 모델에 적용할 수 있는지 보여주는 GAN 분야의 주요 발전을 보여줌.
- ✓ SAGAN의 셀프 어텐션 메커니즘은 앞서 살펴본 트랜스포머 구조의 어텐션헤드 구조와 비슷.
- ✓ 이 방식의 효과는 어텐션 맵을 통해 확인.

http://proceedings.mlr.press/v97/zhang19d/zhang19d.pdf

2.α 생성 모델링 그리고 미래

AI 아트

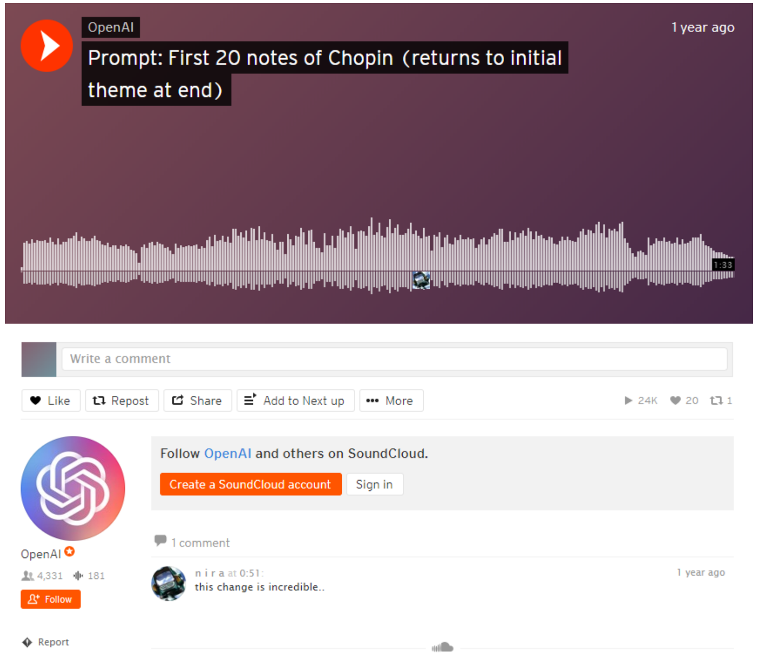
pix2pix

<Form, Figures and BigGAN>

AI 아트 큐레이터 Luba Elliott 주최
Scott Eaton 발표
Andrew Brock 발표
(BigGAN 창시자)

https://www.youtube.com/watch?v=TN7Ydx9ygPo&feature=youtu.be

MuseNet



<OpenAI – 쇼팽 그리고 모차르트 >

https://soundcloud.com/openai_audio/chopin-return

https://soundcloud.com/openai_audio/sonatina