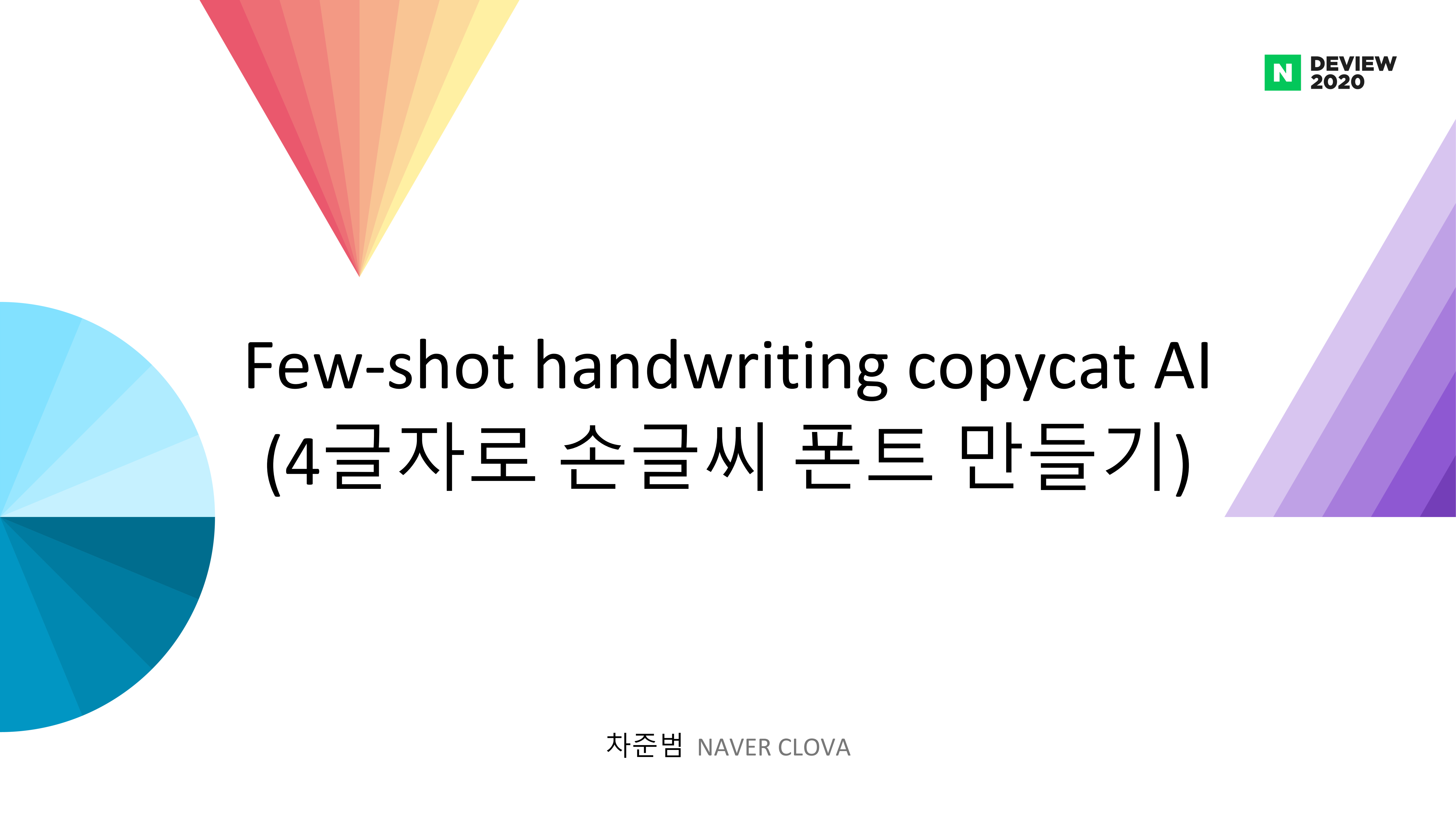




Few-shot handwriting copycat AI (4글자로 손글씨 폰트 만들기)

차준범 NAVER CLOVA

CONTENTS

2019년 한글날 손글씨 공모전

사람의 꿈을 꾸는 손글씨 모델 Few-shot Font Generation

28글자로 손글씨 폰트 만들기 DM-Font

더 적은 글자로, 한국을 넘어 세계로! LF-Font

2019년 한글날 손글씨 공모전

한글 폰트 만들기의 어려움

영어 문자: 알파벳 26자 + 26자 = **52**자

한글 문자: 초성 (19) x 중성 (21) x 종성 (27+1) = **11,172** 자

30분에 1글자씩 만들면 700일 동안 만들어야 하는 작업량
폰트 하나에 3년...? 🤔

하루 8시간

연간 250 근무일

한글 폰트 만들기의 어려움

영어 문자: 알파벳 26자 + 26자 = **52**자

한글 문자: 초성 (19) x 중성 (21) x 종성 (27+1) = **11,172** 자

30분에 1글자씩 만들면 700일 동안 만들어야 하는 작업량
폰트 하나에 3년...? 🤔

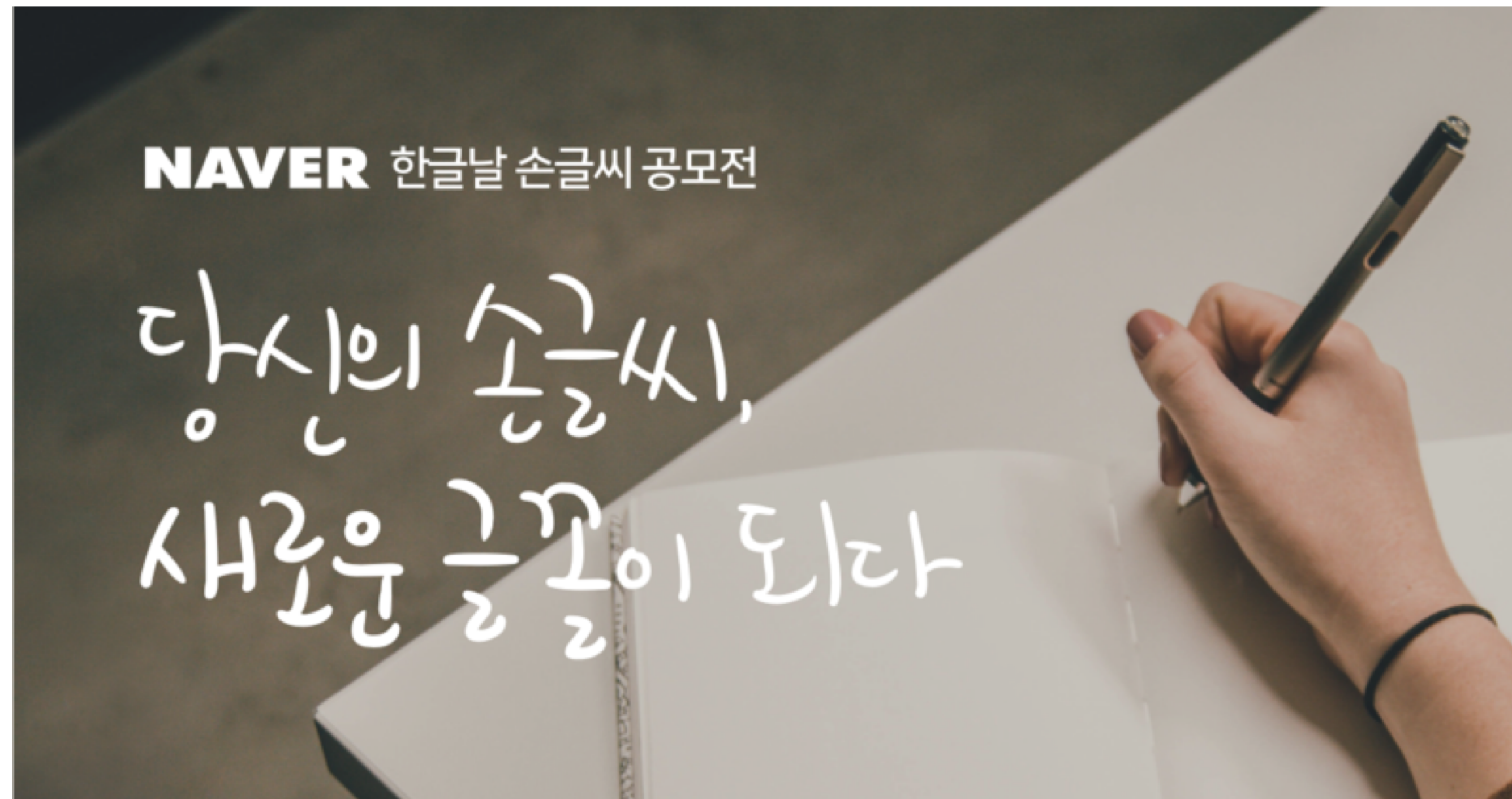
하루 8시간

연간 250 근무일

7일 경기도에 따르면 도는 지난해 5월부터 10개월간 폰트디자인 전문업체 (주)타이포랩과 경기도 전용서체인 '경기천년체'를 제작했다. 한글 1만1천172자, 영문 94자로 구성된 경기천년체 개발에는 총 1억4천300여만 원이 투입됐다.

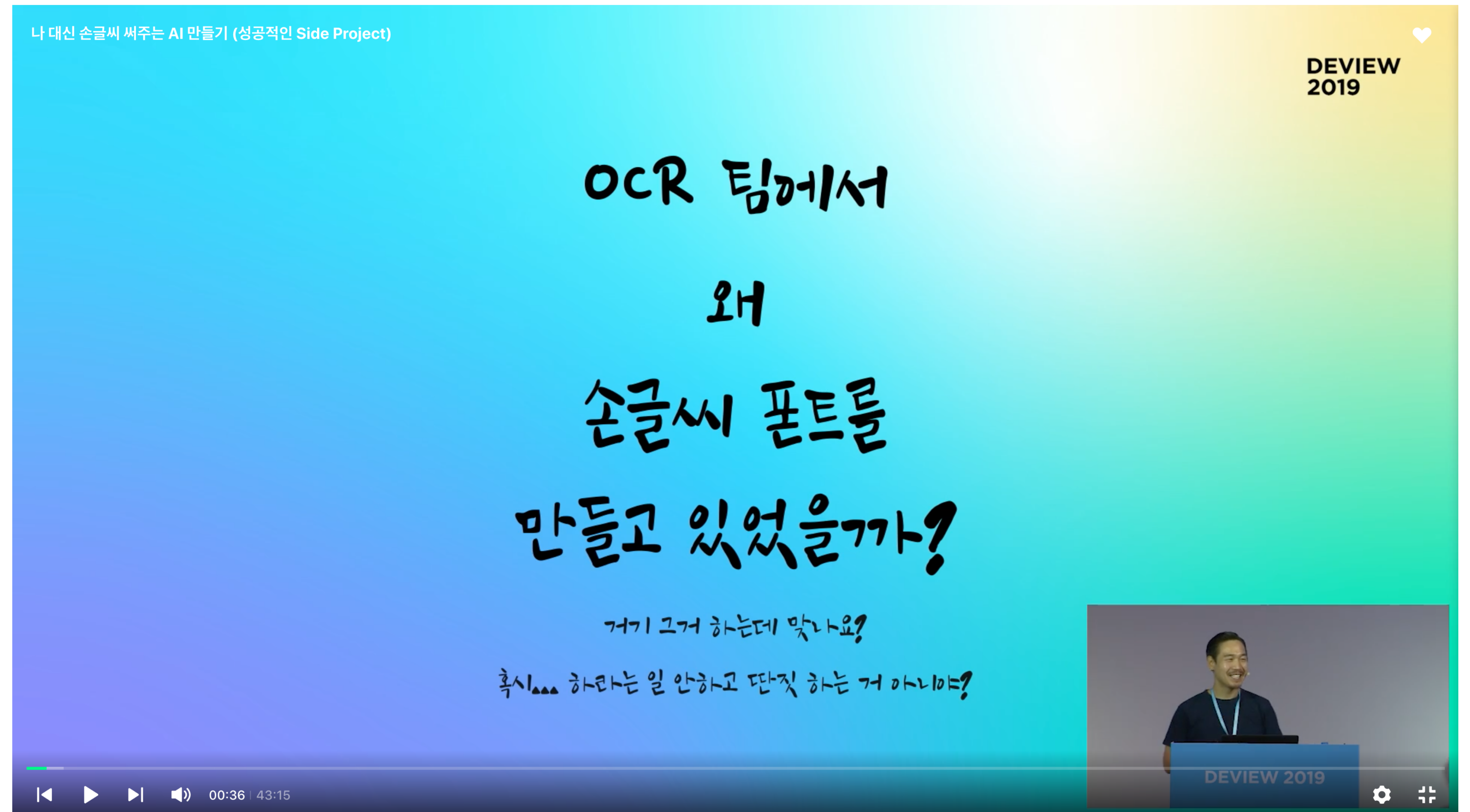
2019년 한글날 손글씨 공모전

AI 로 한글 폰트를 만들면 어떨까?
256자만 있으면, 30분에 폰트 하나가 똑딱 🌟😄



자세한 내용은

DEVIEW 2019!



나 대신 손글씨 써주는 AI 만들기, 이바도

폰트 생성 과정

폰트 생성 파이프라인에는 이것저것 많지만,
오늘은 모델에 집중할 거예요

Data → Preprocess → **Model (finetune & generate)** → Postprocess → TTF

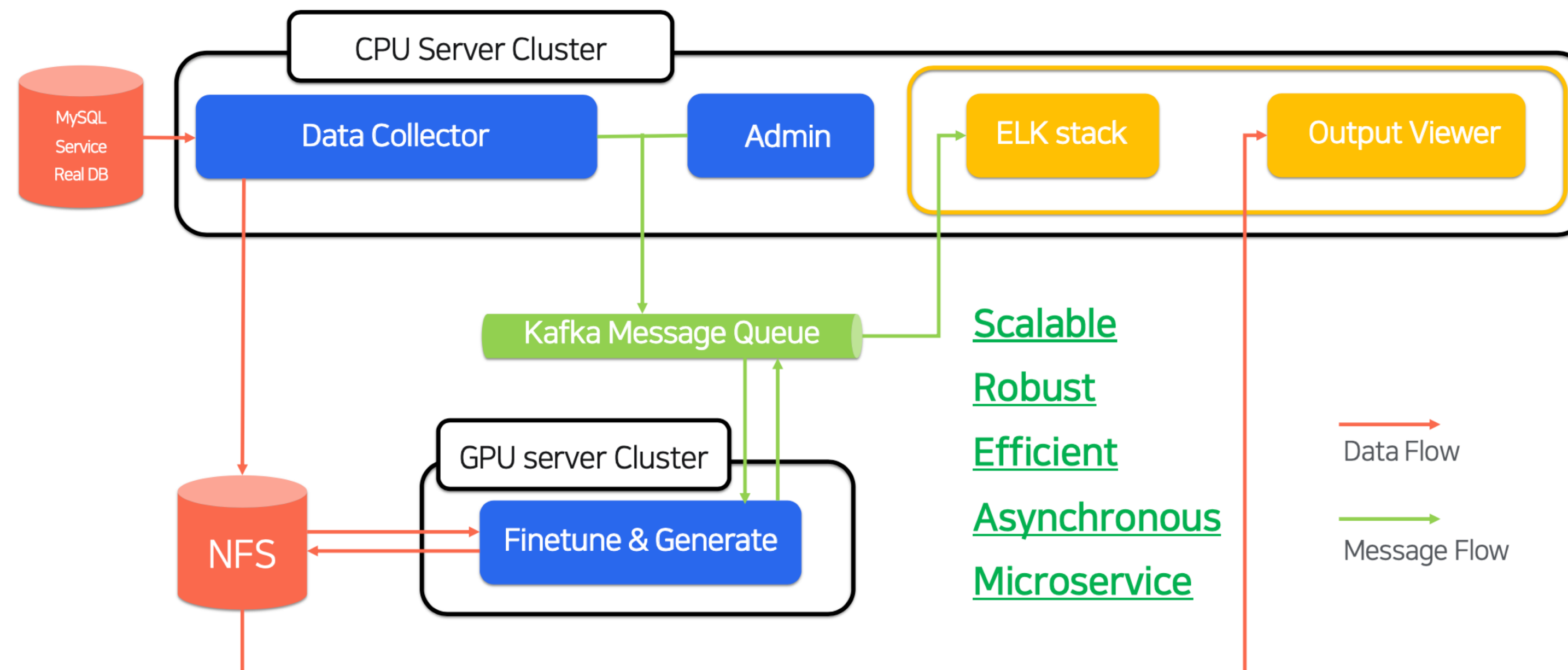
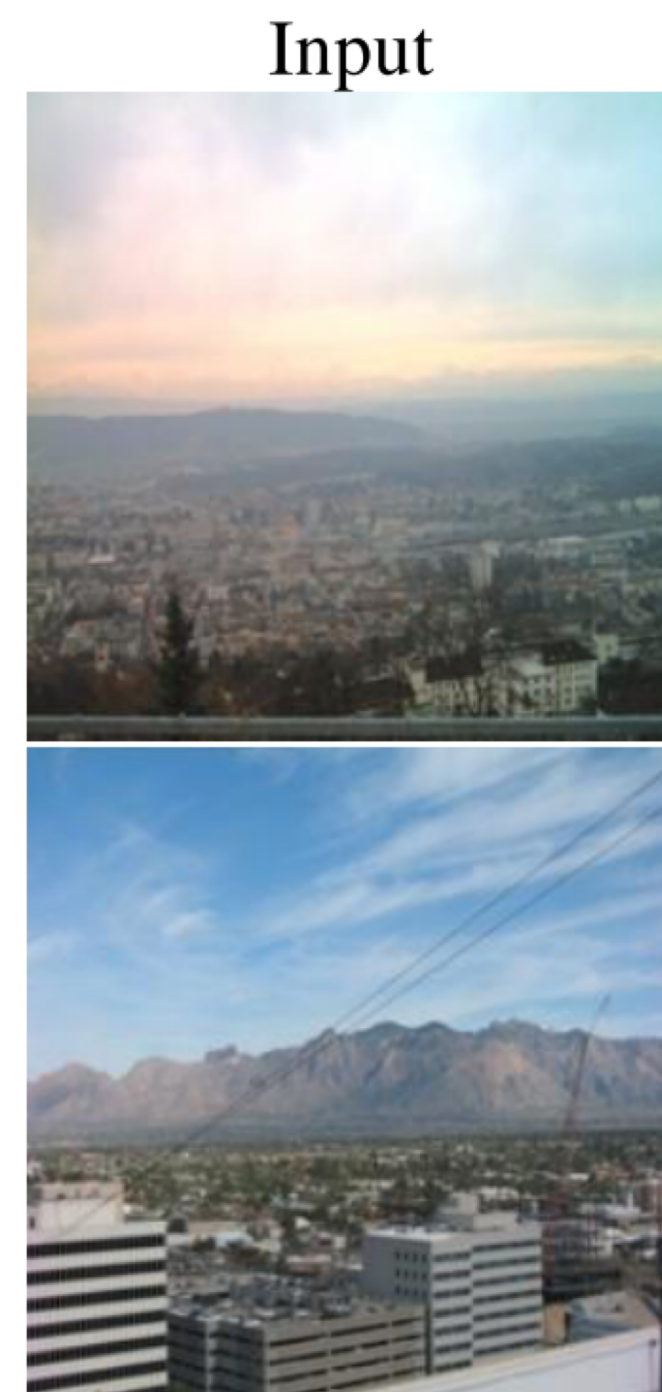


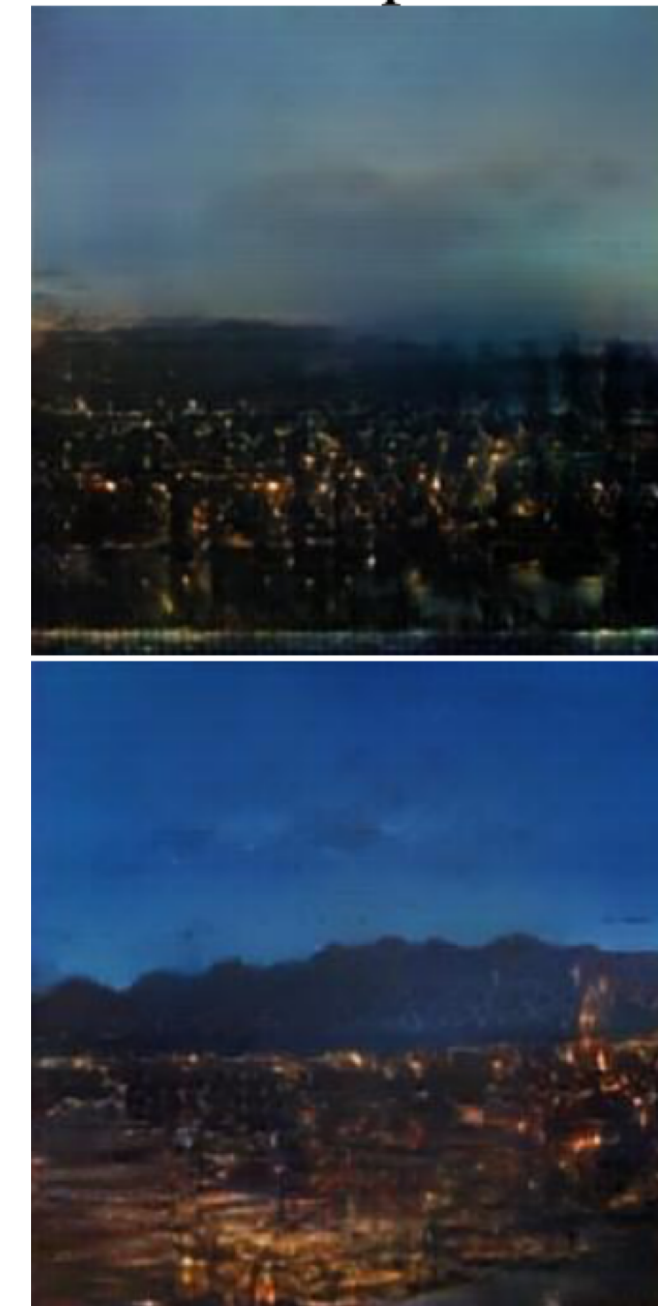
Image-to-Image translation

이미지의 내용 (content) 을 유지하면서 다른 스타일로 바꾸는 모델

낮 사진을



Output



밤 사진으로!

pix2pix on font generation

폰트 생성

글자 이미지의 내용을 유지하면서 글자 스타일을 바꾸는 문제

가



가

“나눔스퀘어” 를

나

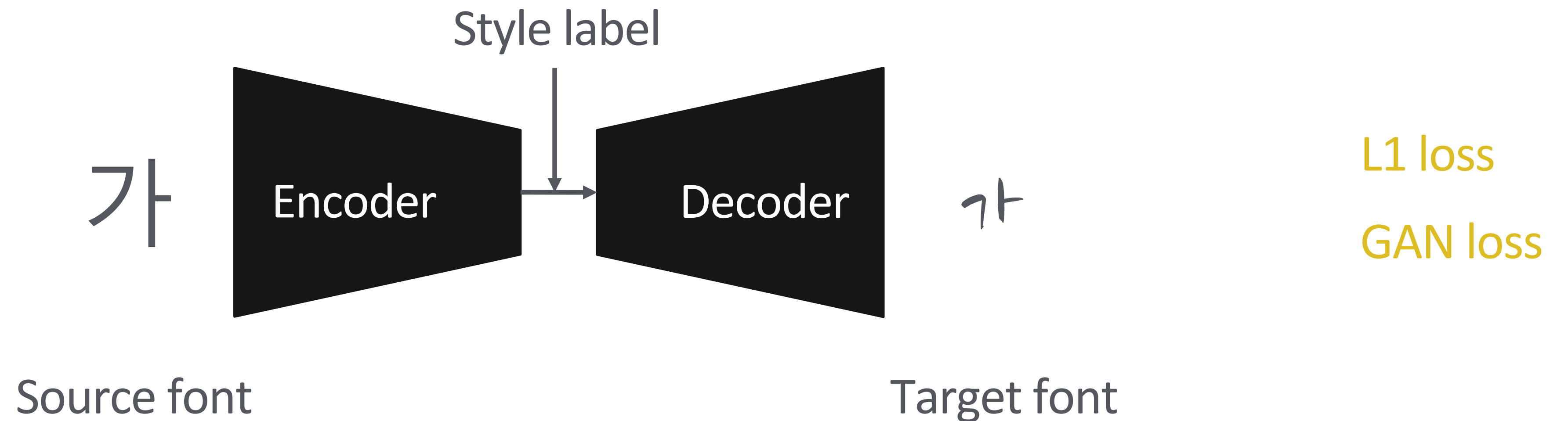


나

“나눔손글씨 붓” 으로!

한글날 손글씨 생성 모델

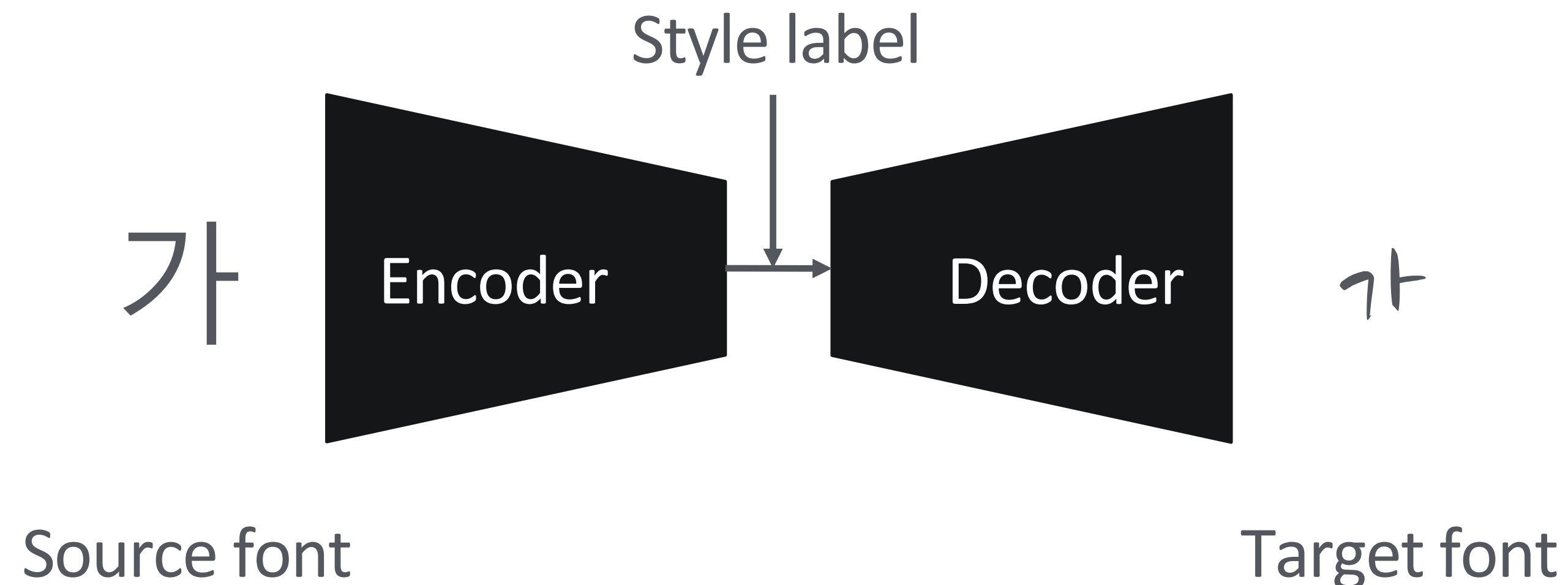
pix2pix 기반 실제론 훨씬 복잡하지만 큰 구조는 pix2pix 와 비슷해요
새로운 폰트를 생성하기 위해서는 finetuning 필요



한글날 손글씨 생성 모델

1. 다양한 폰트를 이용한 모델 pretraining
2. 256개의 입력 레퍼런스를 이용한 finetuning
3. 생성!

약 30분 소요

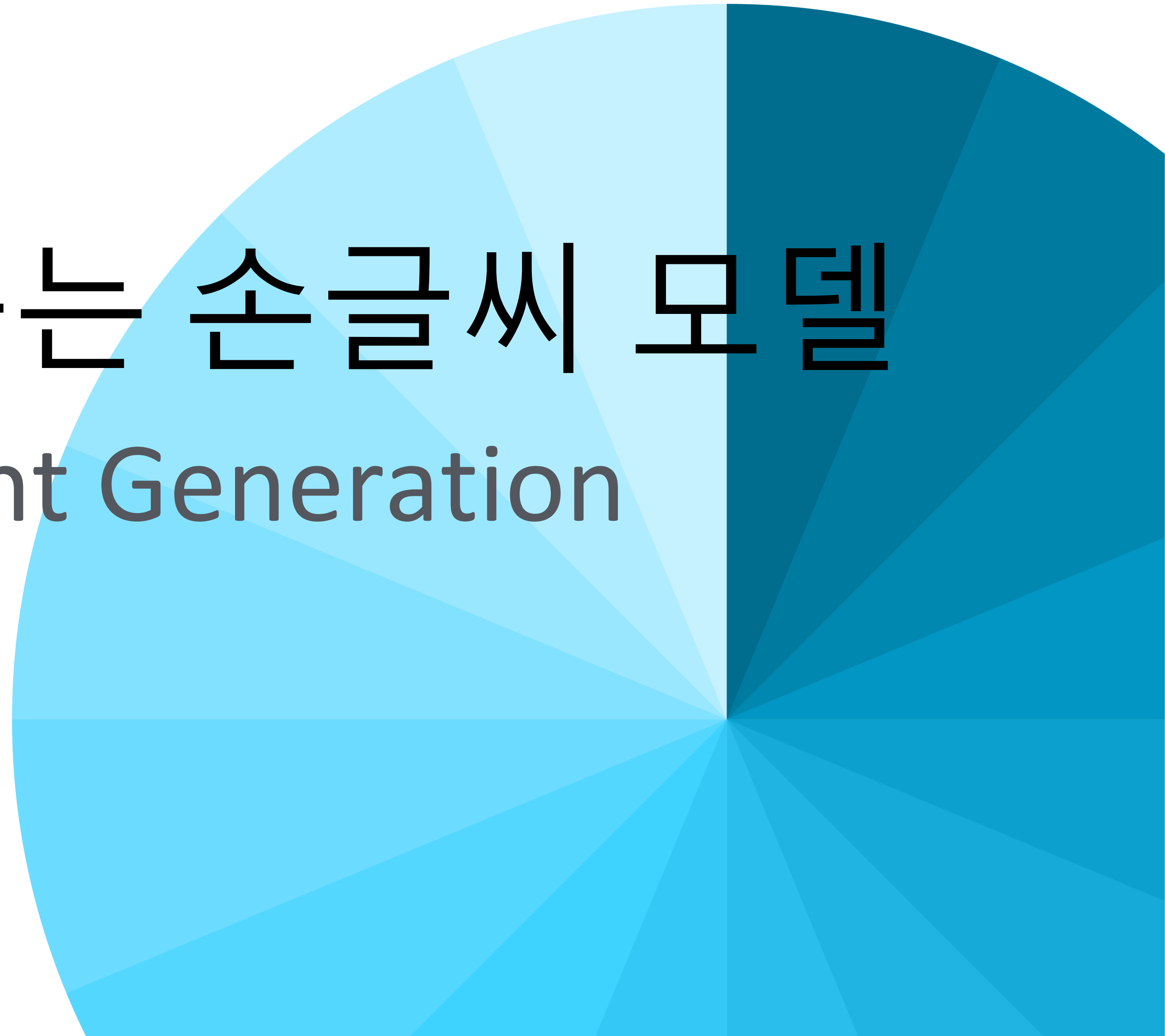


L1 loss

GAN loss

사람의 꿈을 꾸는 손글씨 모델

Few-shot Font Generation



Few-shot font generation

모든 모델의 꿈

사람처럼 적은 데이터로도 잘 하는 것

자원 관점에서

더 적은 글자로, 더 적은 리소스로, 더 빨리 생성하는 것

Few-shot vs. Many-shot

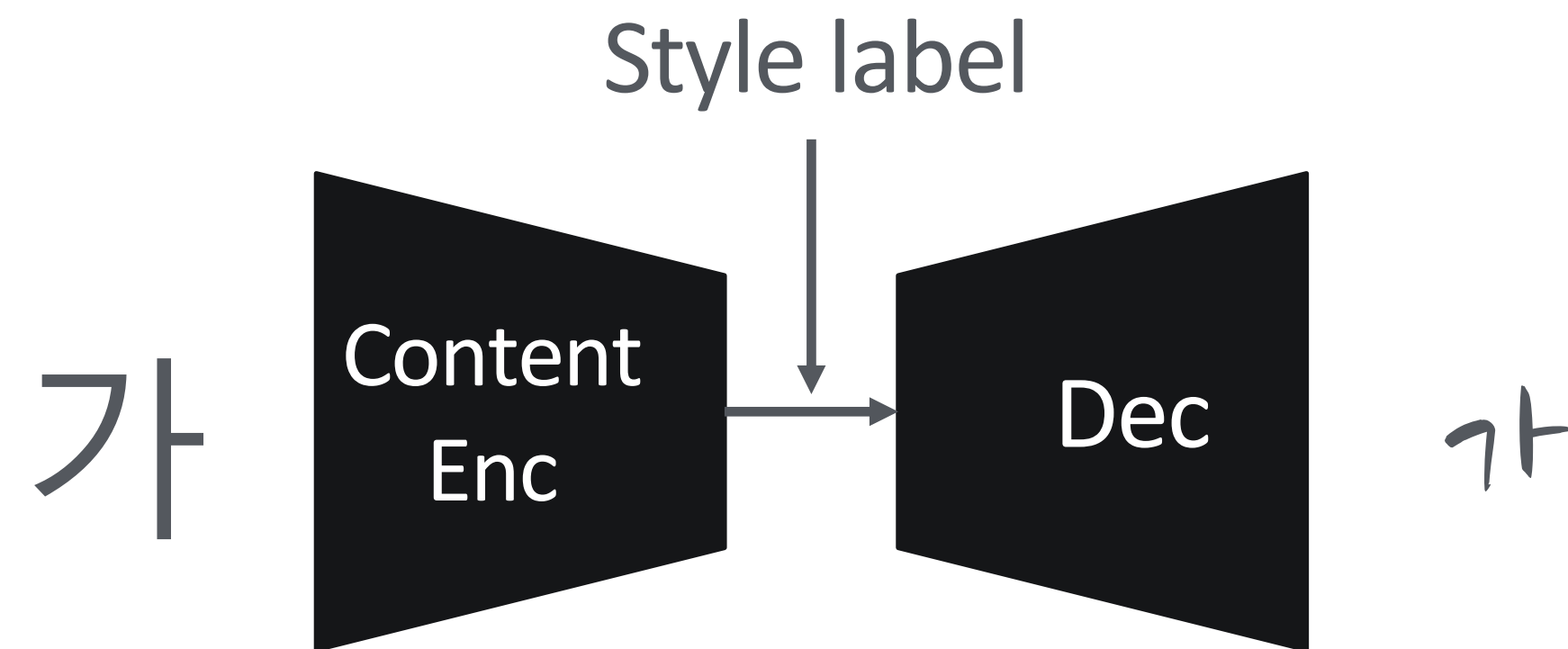
	Few-shot	Many-shot
레퍼런스 글자 수	1 ~ 수십	수백 ~ 수천
생성 시 학습 필요	X	O
생성 소요시간	1분 미만	30분 이상

더 적은 글자로, 더 적은 리소스로, 더 빠르게!

구조적 차이

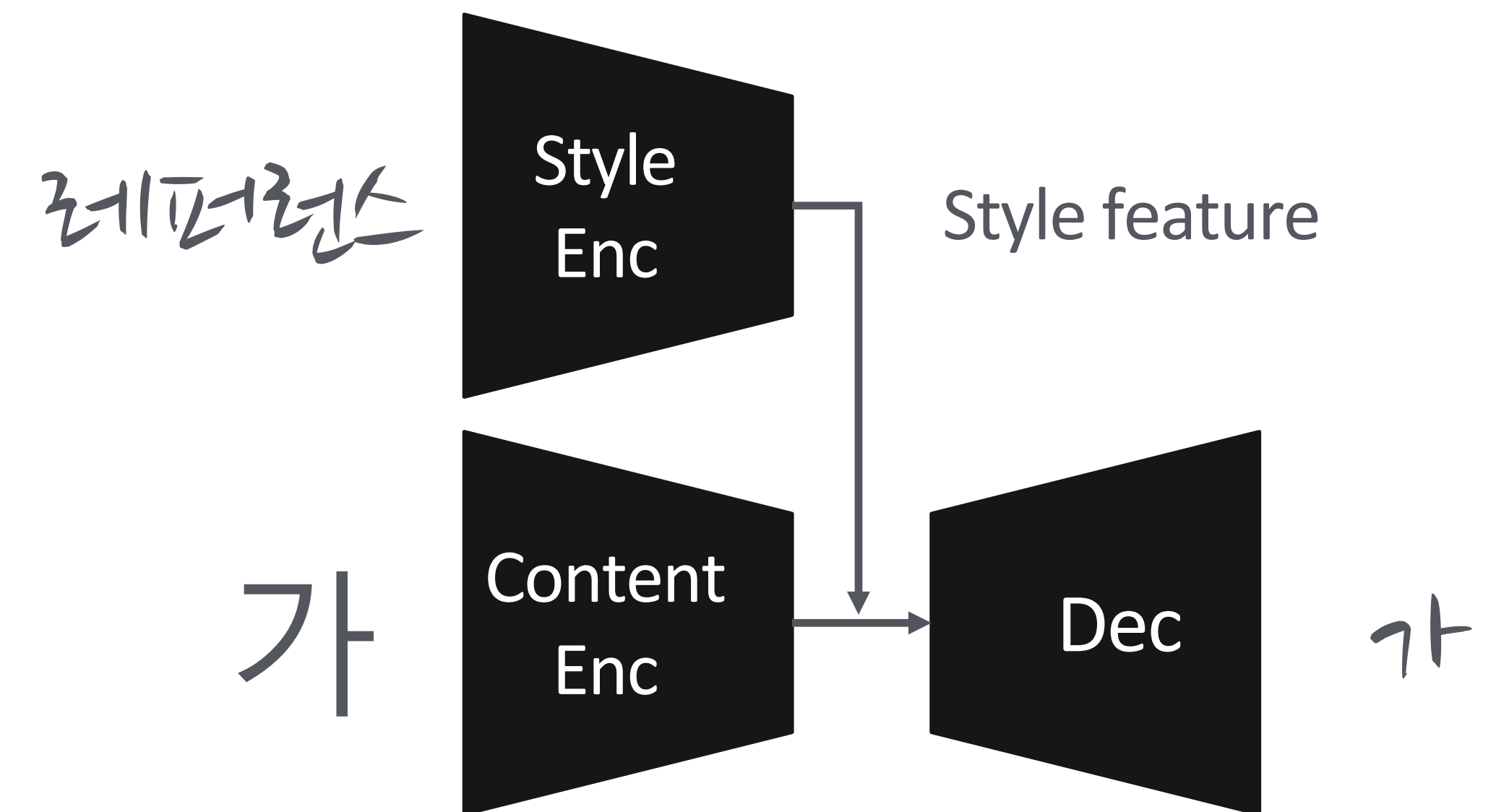
Many-shot model

$f(\text{content image}, \text{style label})$



Few-shot model

$f(\text{content image}, \text{style image})$

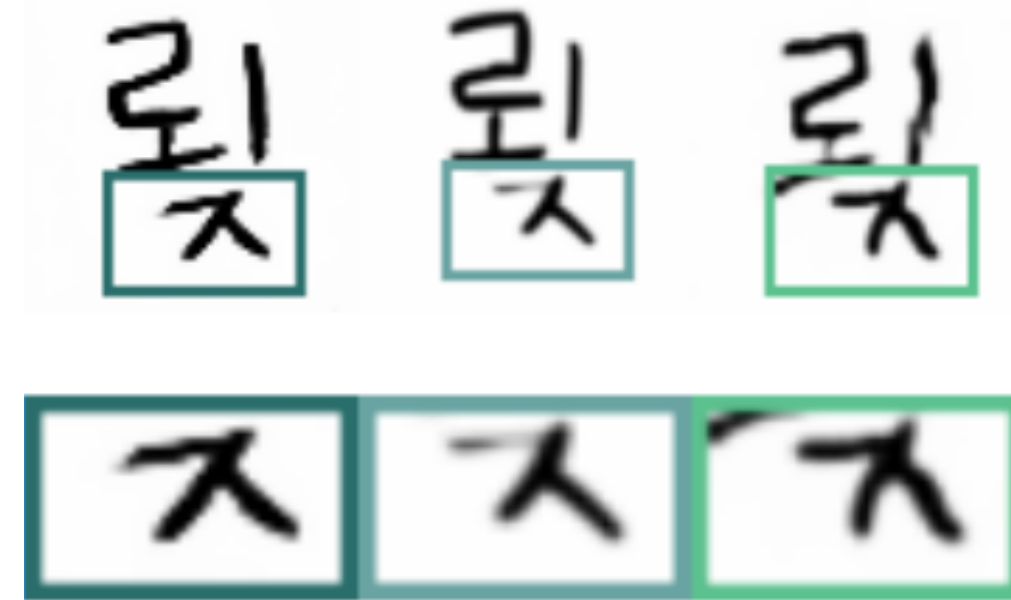


기존 연구들

더 많지만 좋은 곳에 실린 애들만 골라 왔어요



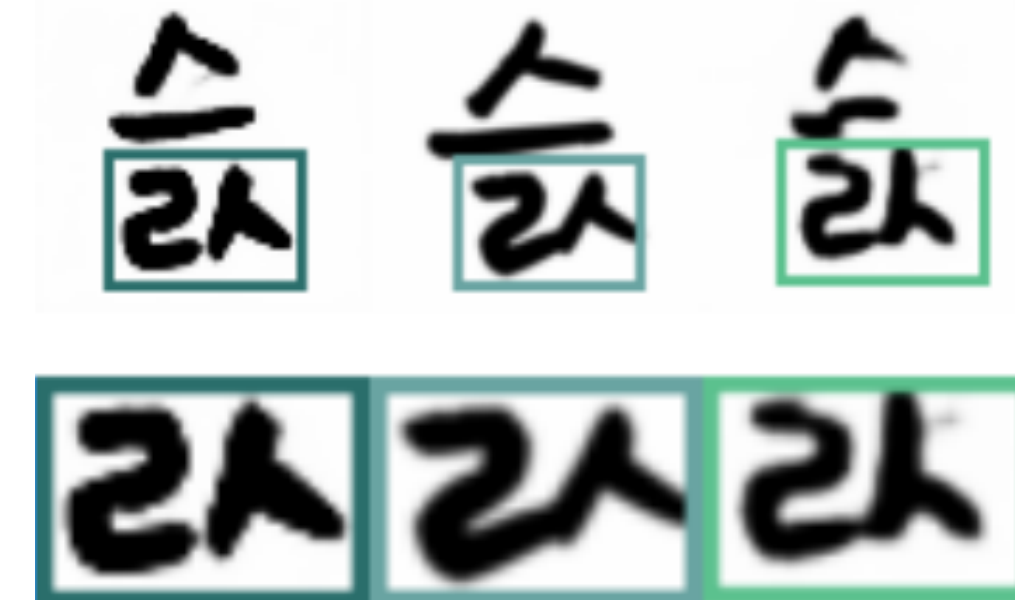
GT



AGIS FUNIT EMD



GT



AGIS FUNIT EMD

DM-Font, ECCV'20

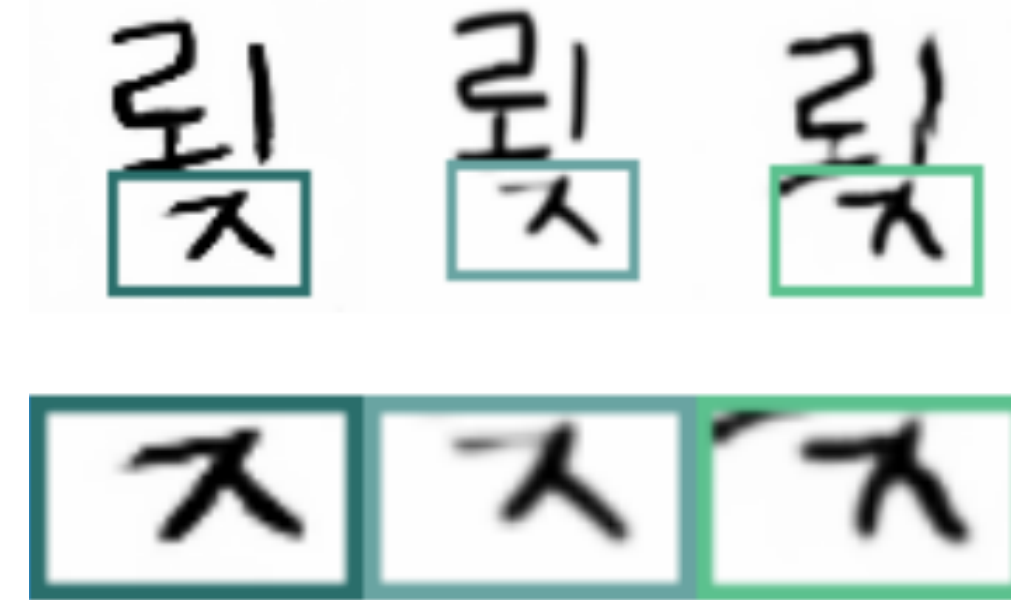
- [AGIS] Gao, et al. "Artistic glyph image synthesis via one-stage few-shot learning." *Transactions on Graphics (TOG)*, 2019.
- [FUNIT] Liu, et al. "Few-shot unsupervised image-to-image translation." *International Conference on Computer Vision (ICCV)*, 2019.
- [EMD] Zhang, et al. "Separating style and content for generalized style transfer." *Computer Vision and Pattern Recognition (CVPR)*, 2018.

기존 연구들의 한계점

대충 보면 비슷해 보이지만...



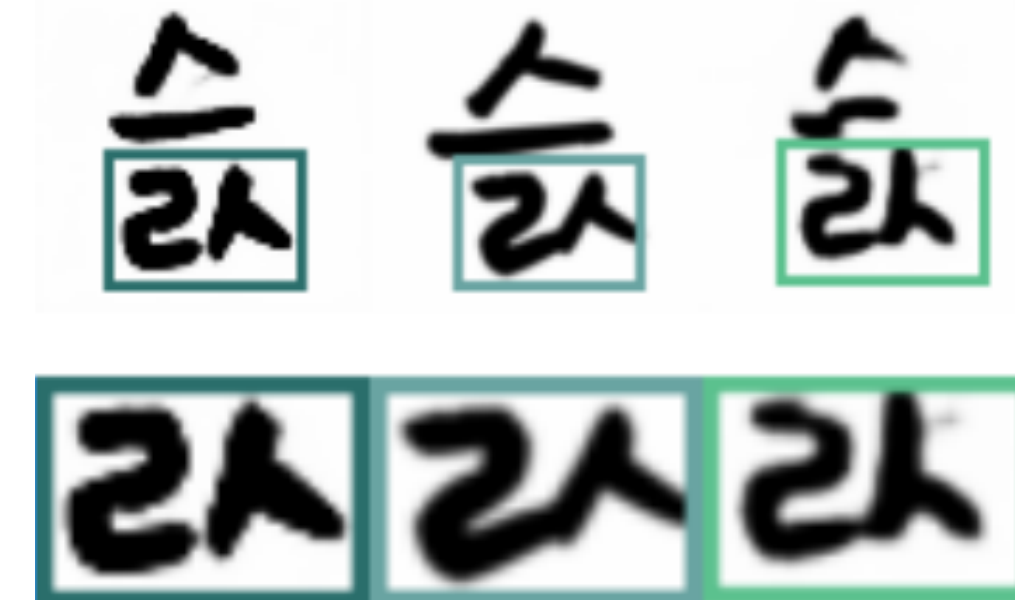
GT



AGIS FUNIT EMD



GT



AGIS FUNIT EMD

DM-Font, ECCV'20

- [AGIS] Gao, et al. "Artistic glyph image synthesis via one-stage few-shot learning." *Transactions on Graphics (TOG)*, 2019.
- [FUNIT] Liu, et al. "Few-shot unsupervised image-to-image translation." *International Conference on Computer Vision (ICCV)*, 2019.
- [EMD] Zhang, et al. "Separating style and content for generalized style transfer." *Computer Vision and Pattern Recognition (CVPR)*, 2018.

기존 모델들은 왜 못할까?

기존 모델들이 풀고 있는 문제

Ref. 가 나 $\xrightarrow{?}$ 최 최 최 최 Generated

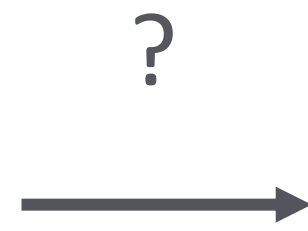
손글씨 “가,나” 를 참조해서 새로운 글자 “최” 를 만든다면,
이 글자는 어떻게 생겨야 할까요?

기존 모델들은 왜 못할까?

원하는 글자의 초중종성을 미리 관측할 수 있다면

Ref.

노 잡

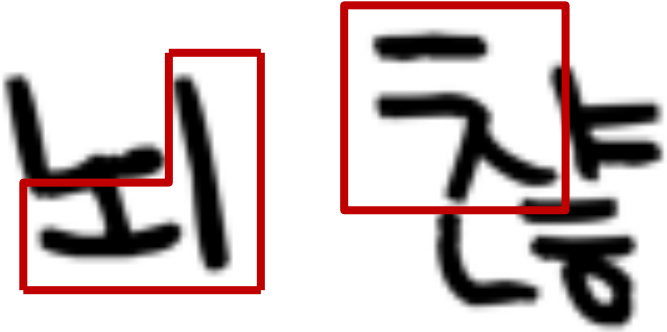
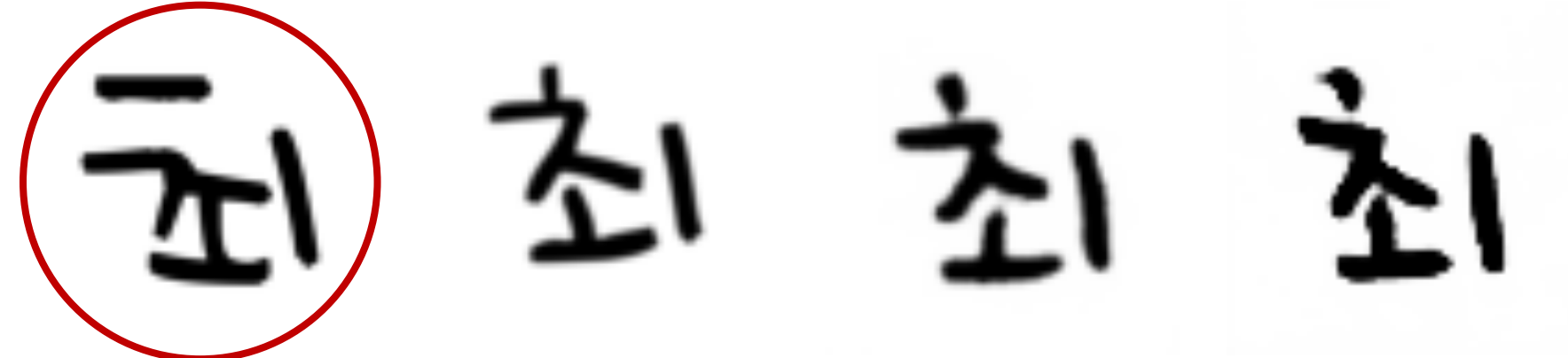


최 최 최 최

Generated

기존 모델들은 왜 못할까?

원하는 글자의 초중종성을 미리 관측할 수 있다면

Ref.  $\xrightarrow{!}$  Generated

예측이 가능해진다!

28글자로 한글 폰트 만들기

DM-Font

Paper: <https://arxiv.org/abs/2005.10510>

Code: <https://github.com/clovaai/dmfont>

차준범, 전상혁, 이가영, 이바도, 김성현, 이활석

직관을 모델에 녹여보자

우리 모델에게 필요한 것

1. 필요한 모든 컴포넌트(초중종성) 보기
2. 한글의 조합성 (compositionality) 이해

Ref.

노인

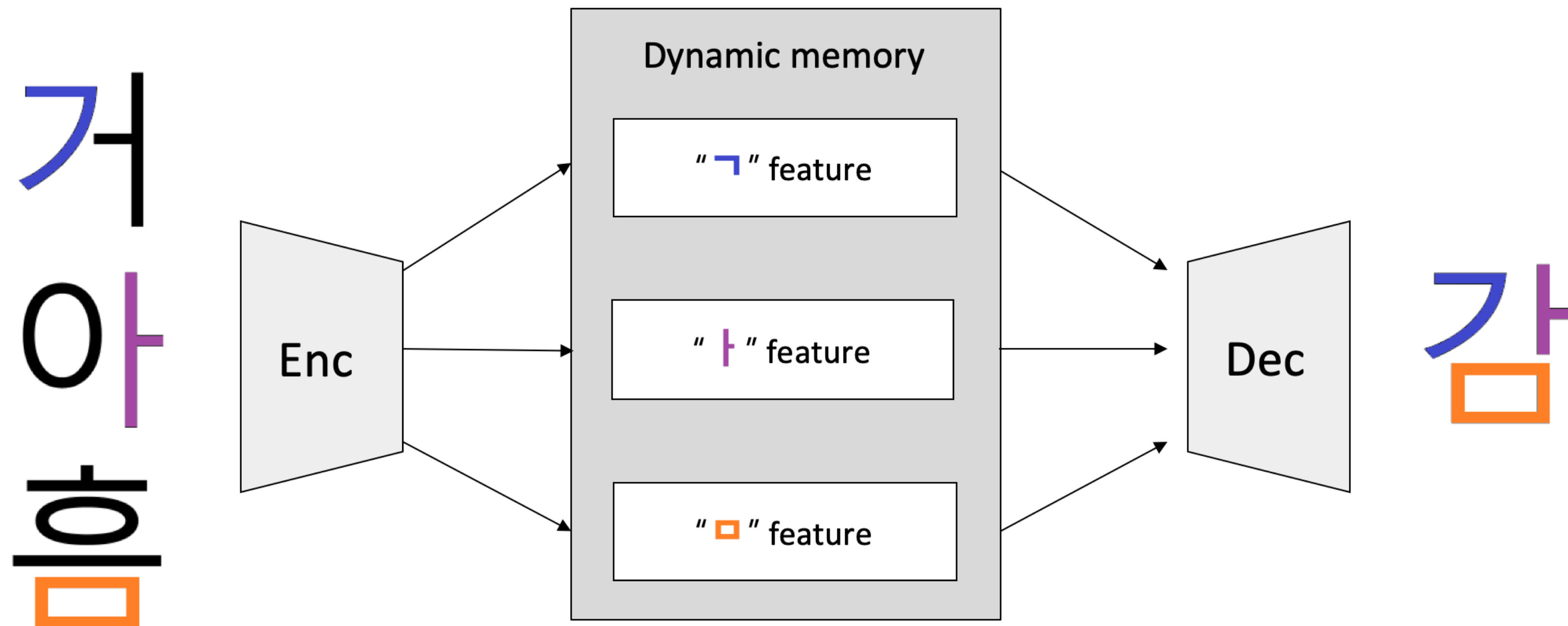


최 최 최 최

Generated

외우지 않고도 풀 수 있도록

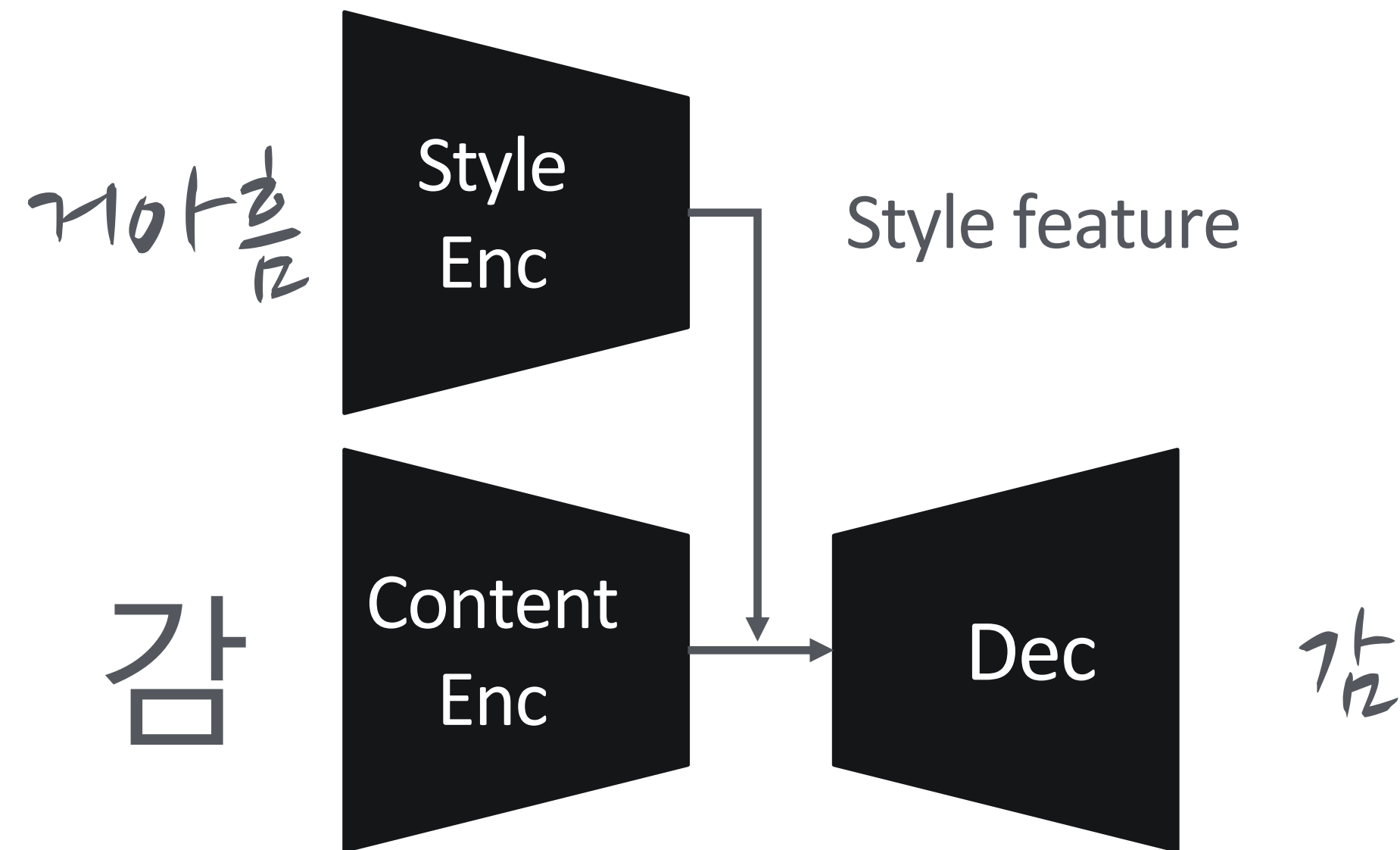
타겟 글자를 만들기 위해 필요한 모든 컴포넌트를 보여주자



조합성을 알아야만 풀 수 있도록

Content encoder 제거

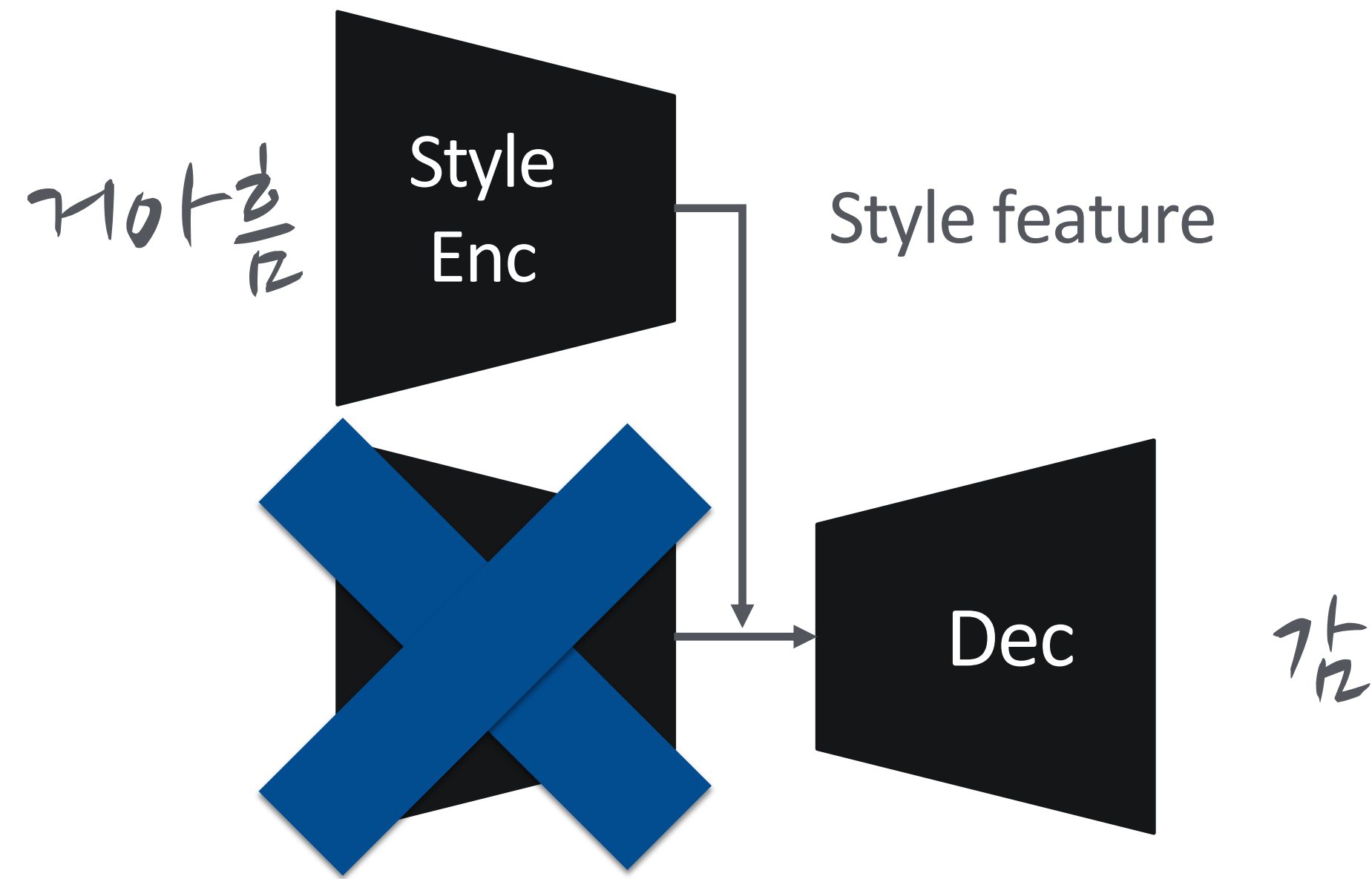
조합성을 이해하지 않고서는 원하는 글자를 만들 수 없음



조합성을 알아야만 풀 수 있도록

Content encoder 제거

조합성을 이해하지 않고서는 원하는 글자를 만들 수 없음



폰트를 생성해보자

초성 (19) x 중성 (21) x 종성 (27+1) = 11,172

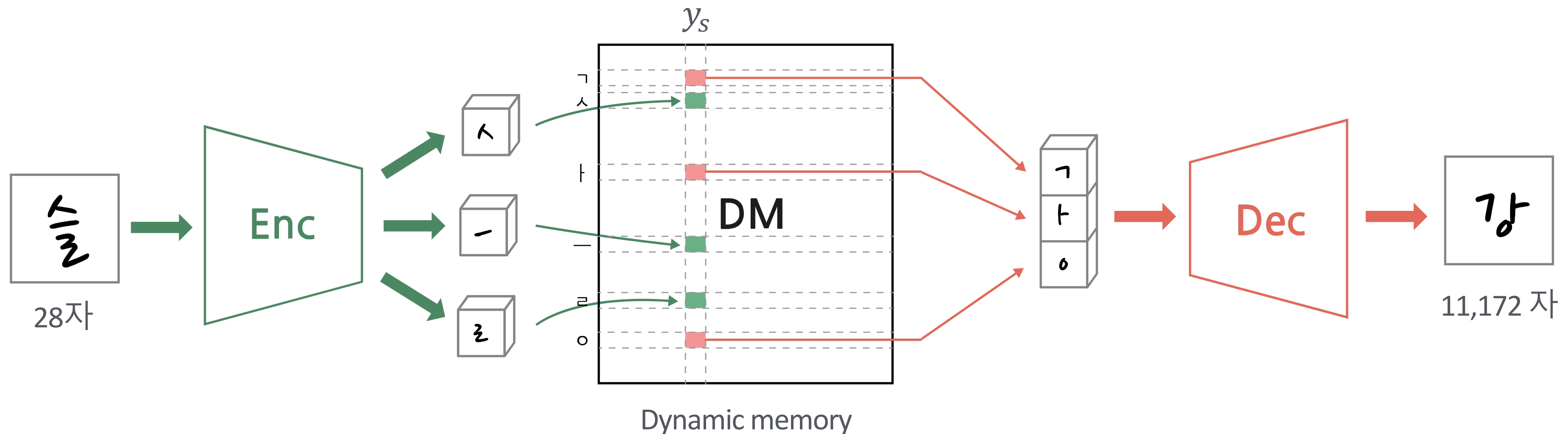
한 글자에 초중종성 하나씩 볼 수 있으니,
한글의 모든 컴포넌트는 28자면 전부 볼 수 있다!

Example)

가객낙뎡뎡렷명벤뽕쓱쓱웁죇죇죇퀘틔푼흙긱긱낭댓뎡랠맏뽕뽕

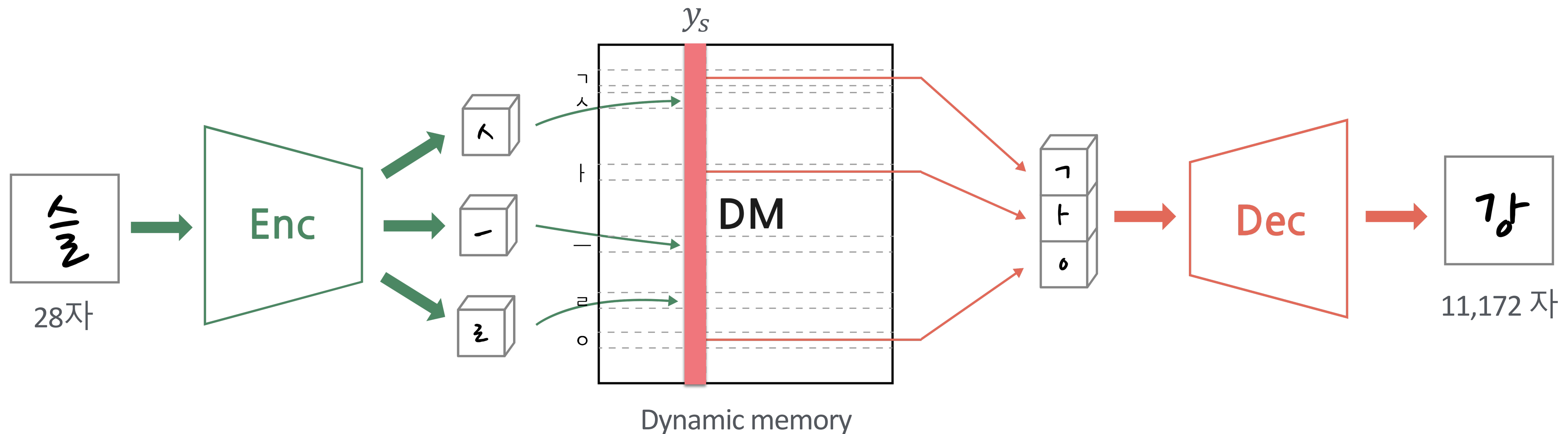
Inference: 폰트 생성

1. 28개의 레퍼런스 글자를 인코딩해서 메모리를 꽉 채우고



Inference: 폰트 생성

1. 28개의 레퍼런스 글자를 인코딩해서 메모리를 꽉 채우고
2. 디코더가 필요한 피쳐들을 가져와서 타겟 글자 생성

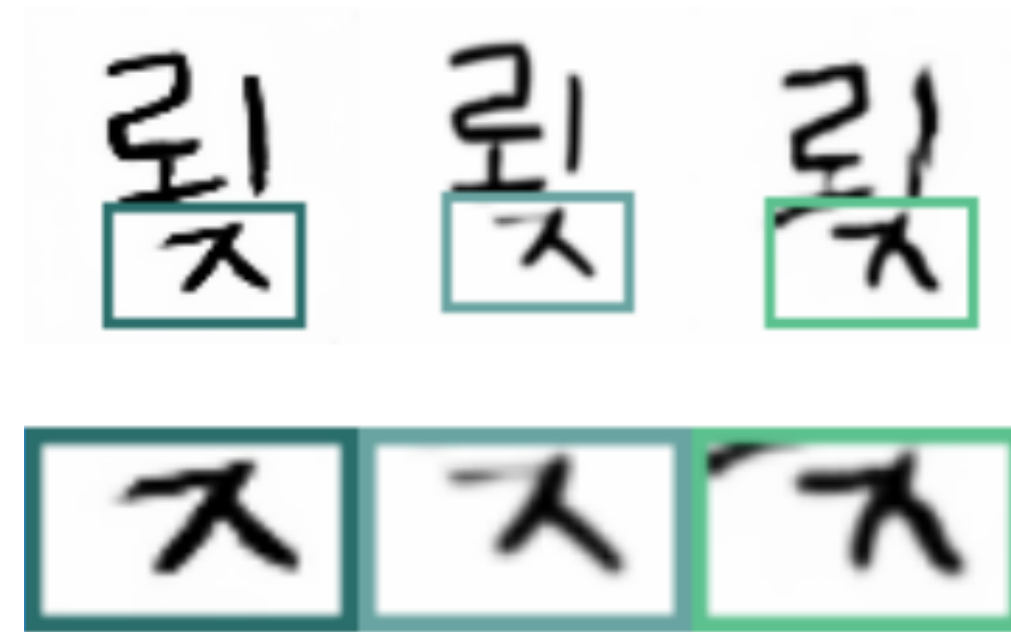


결과

앞서 보았던 문제의 이 그림



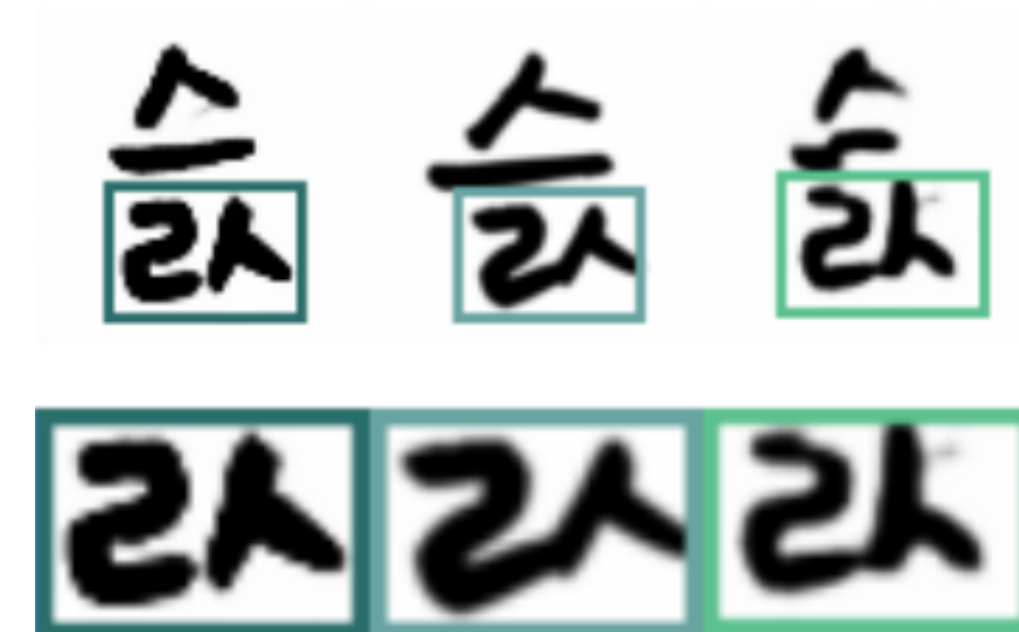
GT



AGIS FUNIT EMD



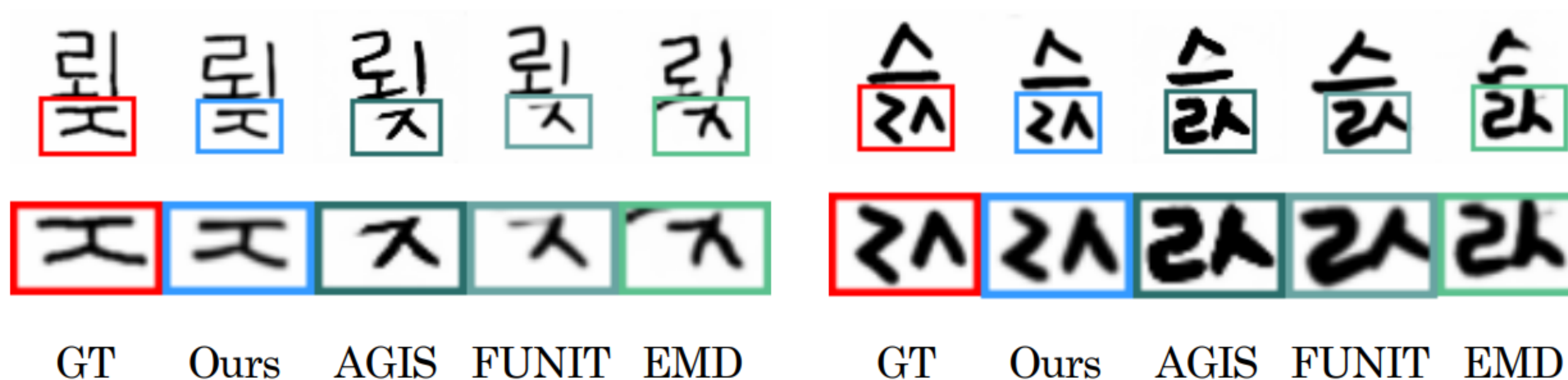
GT



AGIS FUNIT EMD

결과

우리는 잘 한다!



더 많은 결과

우리는 항상 잘 한다!

EMD	혜	밀	여	영	푸	웁	갸	테	빈	관	털	섯	췌	눅
FUNIT	혜	밀	여	영	푸	웁	갸	테	빈	관	털	섯	췌	눅
AGIS-Net	혜	밀	여	영	푸	웁	갸	테	빈	관	털	섯	췌	눅
Ours	혜	밀	여	영	푸	웁	갸	테	빈	관	털	섯	췌	눅
GT	혜	밀	여	영	푸	웁	갸	테	빈	관	털	섯	췌	눅

혜	밀	여	영	푸	웁	갸	테	빈	관	털	섯	췌	눅
---	---	---	---	---	---	---	---	---	---	---	---	---	---

다른 언어 결과

이러한 조합적인 특성은 다른 언어에도 있어요

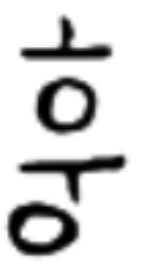
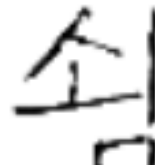
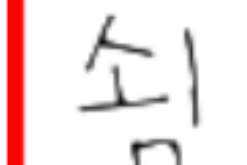
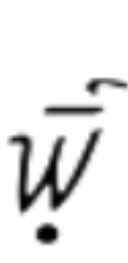
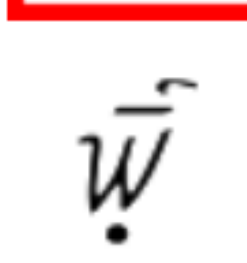
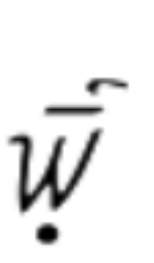
EMD	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒
FUNIT	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒
AGIS-Net	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒
Ours	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒
GT	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒

๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒	๒๒
----	----	----	----	----	----	----	----	----	----	----	----	----	----

태국 글자

기존 방법들도 그럴싸해 보이는데?

사실 학습 때 본 글자 중에 비슷한 애를 가져온 것 뿐

								
								
								
GT	EMD (output)	EMD (NN)	FUNIT (output)	FUNIT (NN)	AGIS-Net (output)	AGIS-Net (NN)	Ours (output)	Ours (NN)

생성한 글자와 가장 비슷한 글자를 (학습 데이터 + 정답 데이터) 에서 찾아보니...

더 적은 글자로, 한국을 넘어 세계로!

LF-Font

Paper: <https://arxiv.org/abs/2009.11042>

Code: <https://github.com/clovaai/lffont>

박송*, 전상혁*, 차준범, 이바도, 심현정

DM-Font 의 한계점

28글자나 필요해?

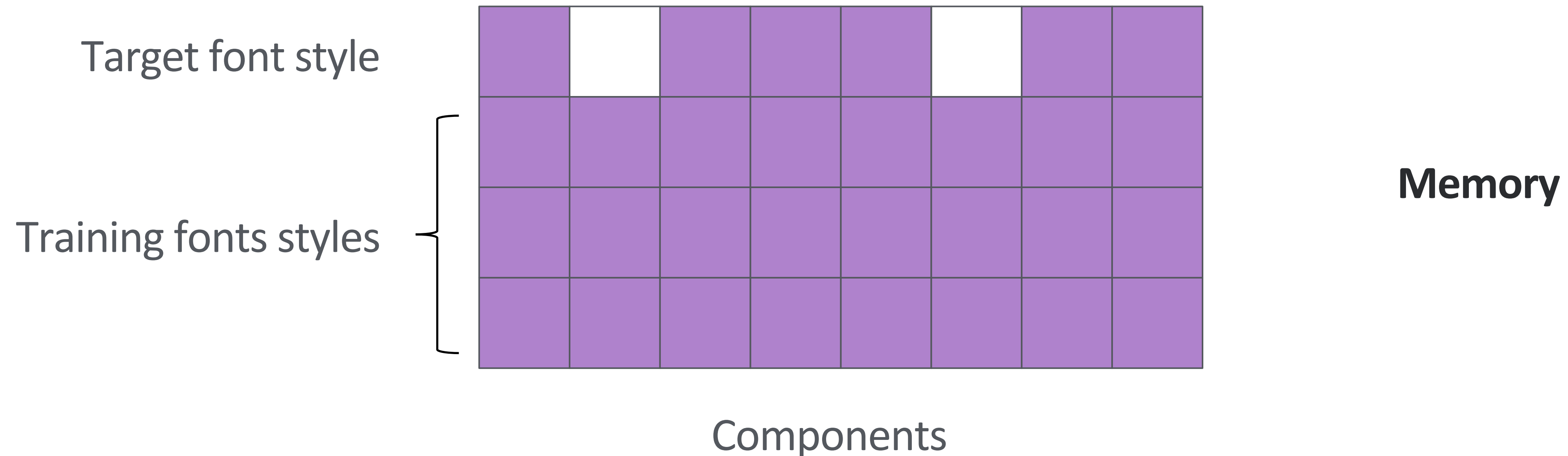
퀄리티 좀 떨어져도 되니까 4글자로 만들고 싶어요

한자는 못 만들어?

한자는 컴포넌트가 너무 많아요 (371개+)

더 적은 글자로

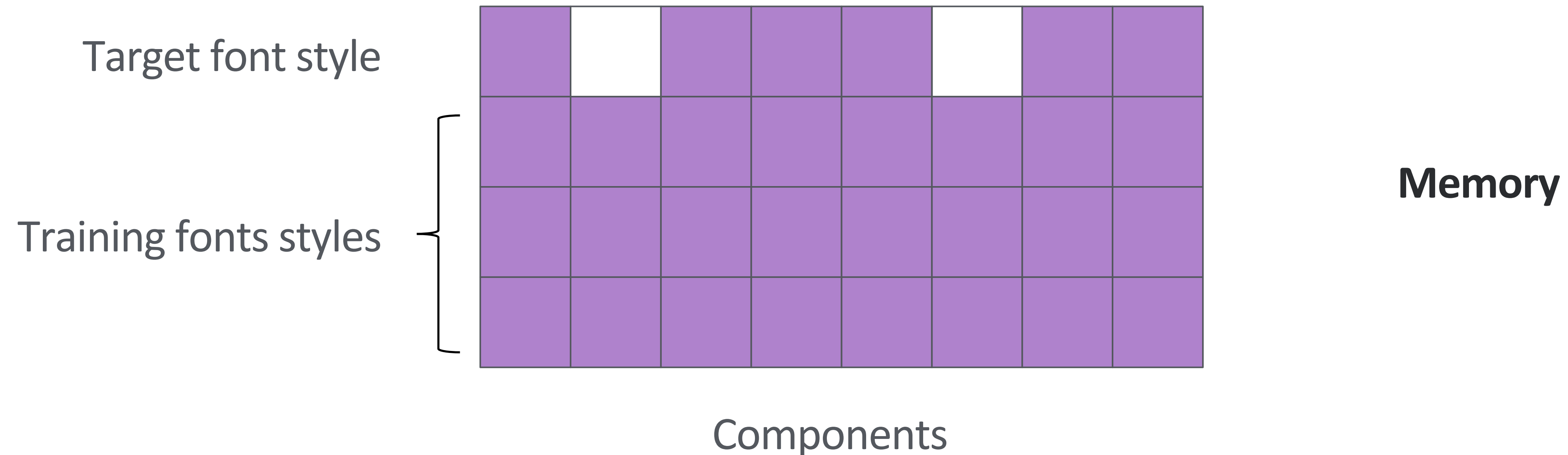
글자수를 줄이면 메모리에 빈 공간이 생김
이 공간을 어떻게 채우지? 🤔



더 적은 글자로

Matrix factorization!

추천 시스템에서 많이 본 것 같은 그림



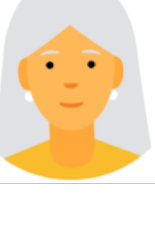
Matrix factorization

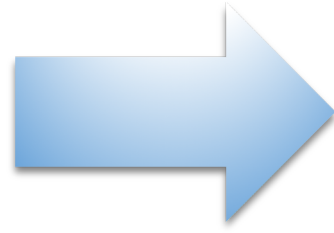
1번 친구와 2번 친구는 슈렉을 좋아할까요?

	Harry Potter	The Triplets of Belleville	Shrek	The Dark Knight Rises	Memento

Matrix factorization

1번 친구와 2번 친구는 슈렉을 좋아할까요?

					
	Harry Potter	The Triplets of Belleville	Shrek	The Dark Knight Rises	Memento
	✓	✗		✓	✗
	✗	✓		✗	✓
	✓	✓	✓	✗	✗
	✗	✗	✗	✓	✓



1	-1		1	-1
-1	1		-1	1
1	1	1	-1	-1
-1	-1	-1	1	1

A

Matrix factorization

$$A \approx UV^T$$

Factorization!

					
	Harry Potter	The Triplets of Belleville	Shrek	The Dark Knight Rises	Memento
	✓	✗		✓	✗
	✗	✓		✗	✓
	✓	✓	✓	✗	✗
	✗	✗	✗	✓	✓

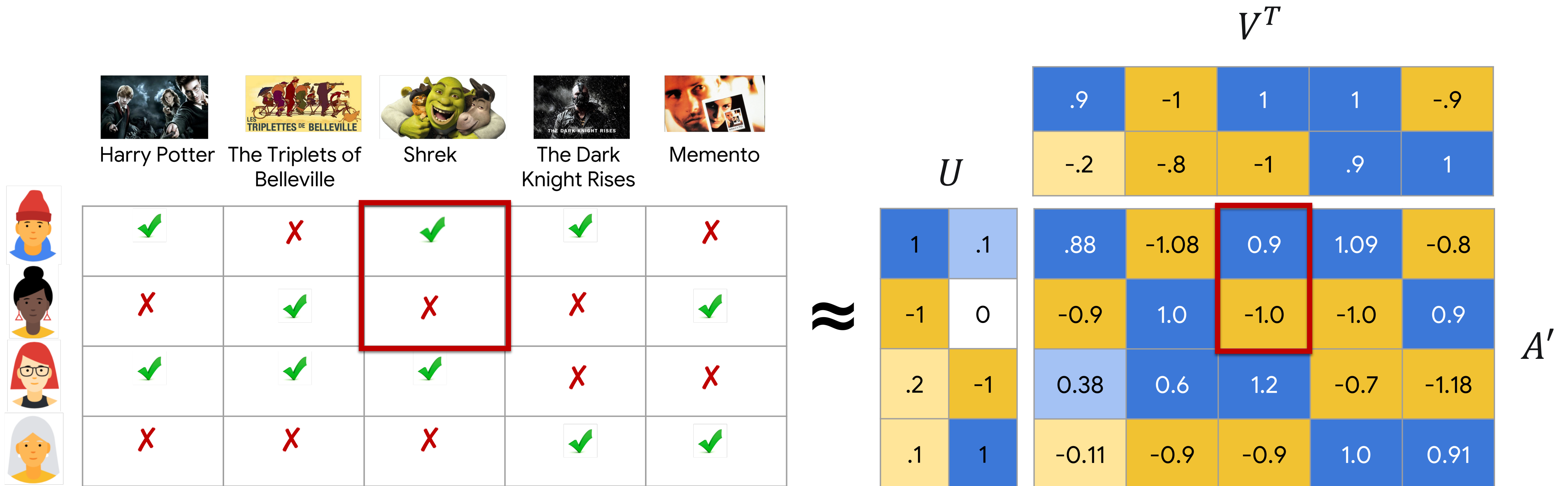


		V^T				
		.9	-1	1	1	-.9
		-.2	-.8	-1	.9	1
U						
1	.1	1	-1		1	-1
-1	0	-1	1		-1	1
.2	-1	1	1	1	-1	-1
.1	1	-1	-1	-1	1	1
		A				

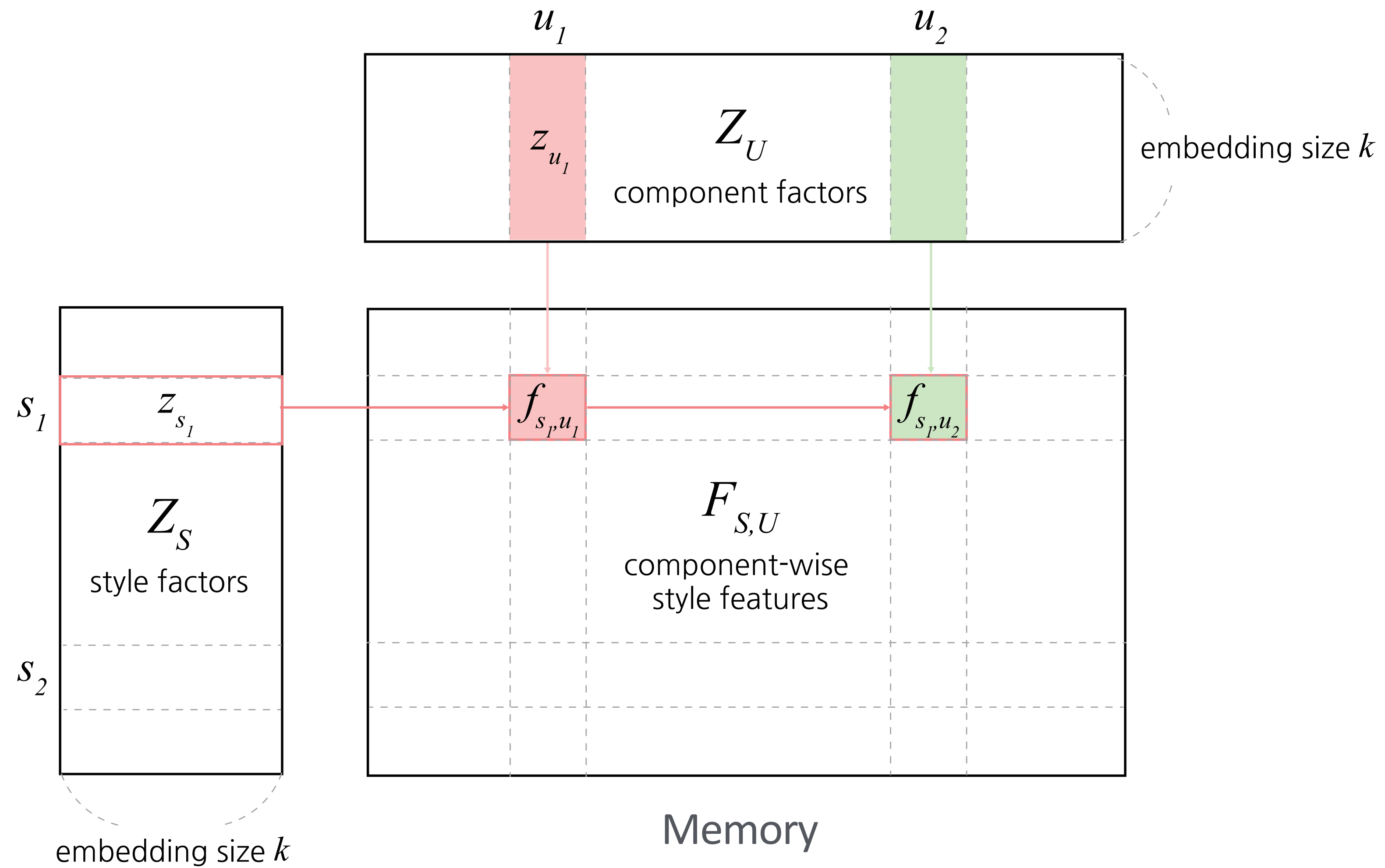
Matrix factorization

$$A \approx UV^T = A'$$

친구들의 성향과 영화의 특성을 통해 예측해 볼 수 있어요



Matrix factorization



End-to-end factorization module

$$Model \rightarrow A \rightarrow UV^T \rightarrow A'$$

Matrix factorization

1. $Model \rightarrow A$
2. $A \rightarrow UV^T$ (최적화 필요)



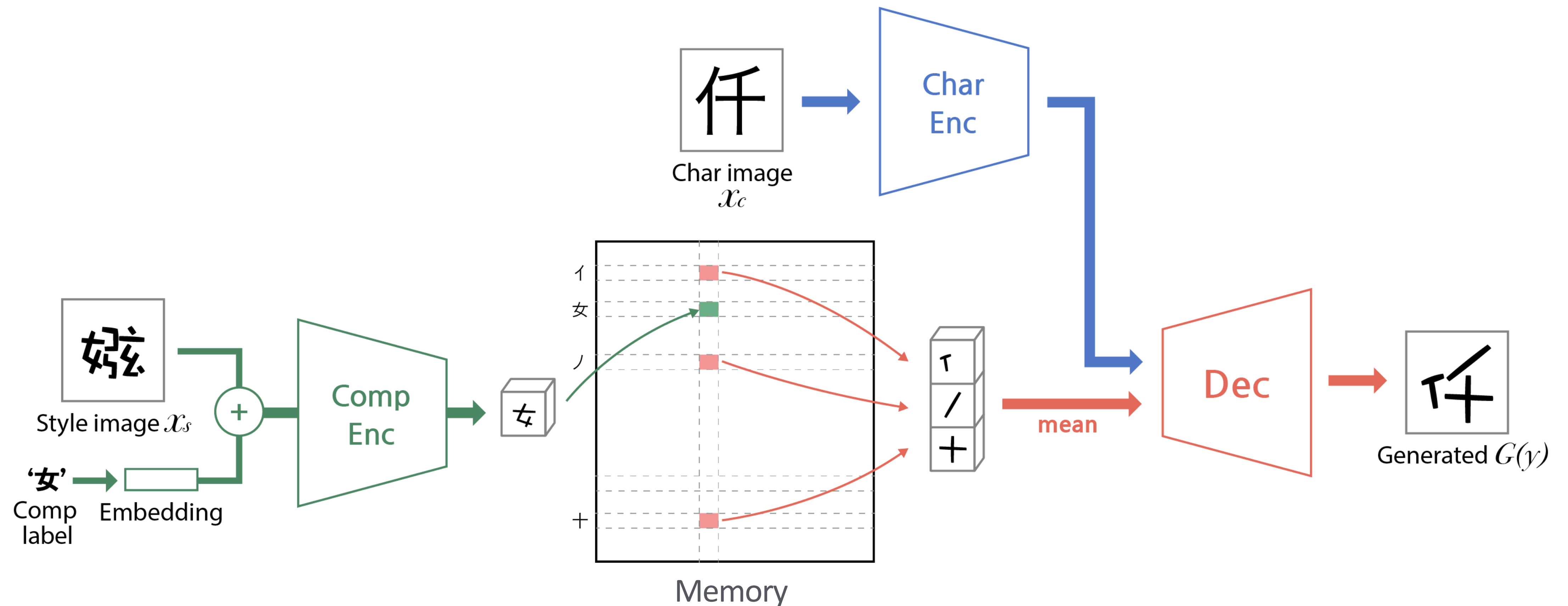
End-to-end factorization module

1. $Model \rightarrow (U, V)$

MF 를 모델에 집어넣어 생성 시 최적화
필요 없도록 개선

생성 모델

Factorization 외에도 달라진 게 좀 있지만...



4글자로 한글 폰트 만들기

이제 4글자로도 한글 폰트를 만들 수 있어요

LF-Font	죄 룩 맥 몽	고 스	르 스	즈 스	보 스
GT	죄 룩 맥 몽	고 스	르 스	즈 스	보 스
LF-Font	섯 윙트 연 객	죄 스	름 스	건 스	갤 스
GT	섯 윙트 연 객	죄 스	름 스	건 스	갤 스

한국을 넘어 세계로

8글자로 한자 만들기

Source	审	弱	和	峰	潮	税	博	城	尊	精	崮	棒	常	瑶	晁	省	城	荷	呼	猗	屋	群
EMD	审	弱	和	峰	潮	税	博	城	尊	精	崮	棒	常	瑶	晁	省	城	荷	呼	猗	屋	群
AGIS-Net	审	弱	和	峰	潮	税	博	城	尊	精	崮	棒	常	瑶	晁	省	城	荷	呼	猗	屋	群
FUNIT	审	弱	和	峰	潮	税	博	城	尊	精	崮	棒	常	瑶	晁	省	城	荷	呼	猗	屋	群
Ours	审	弱	和	峰	潮	税	博	城	尊	精	崮	棒	常	瑶	晁	省	城	荷	呼	猗	屋	群
GT	审	弱	和	峰	潮	税	博	城	尊	精	崮	棒	常	瑶	晁	省	城	荷	呼	猗	屋	群

더 자세한 내용과 더 많은 결과

논문과 코드가 공개되어 있어요

사실 이번 발표에선 디테일들은 다 빠졌어요

DM-Font

Paper: <https://arxiv.org/abs/2005.10510>

Code: <https://github.com/clovaai/dmfont>

LF-Font

Paper: <https://arxiv.org/abs/2009.11042>

Code: <https://github.com/clovaai/lffont>

정리

- 2019년 한글날 손글씨 공모전 256자, 30분
- 사람은 더 잘 하지 않나요? 28자, 20초 (DM-Font)
- 글자를 더 줄이고 싶어요
- 한자도 만들고 싶어요 4자 (LF-Font)

동전의 뒷면

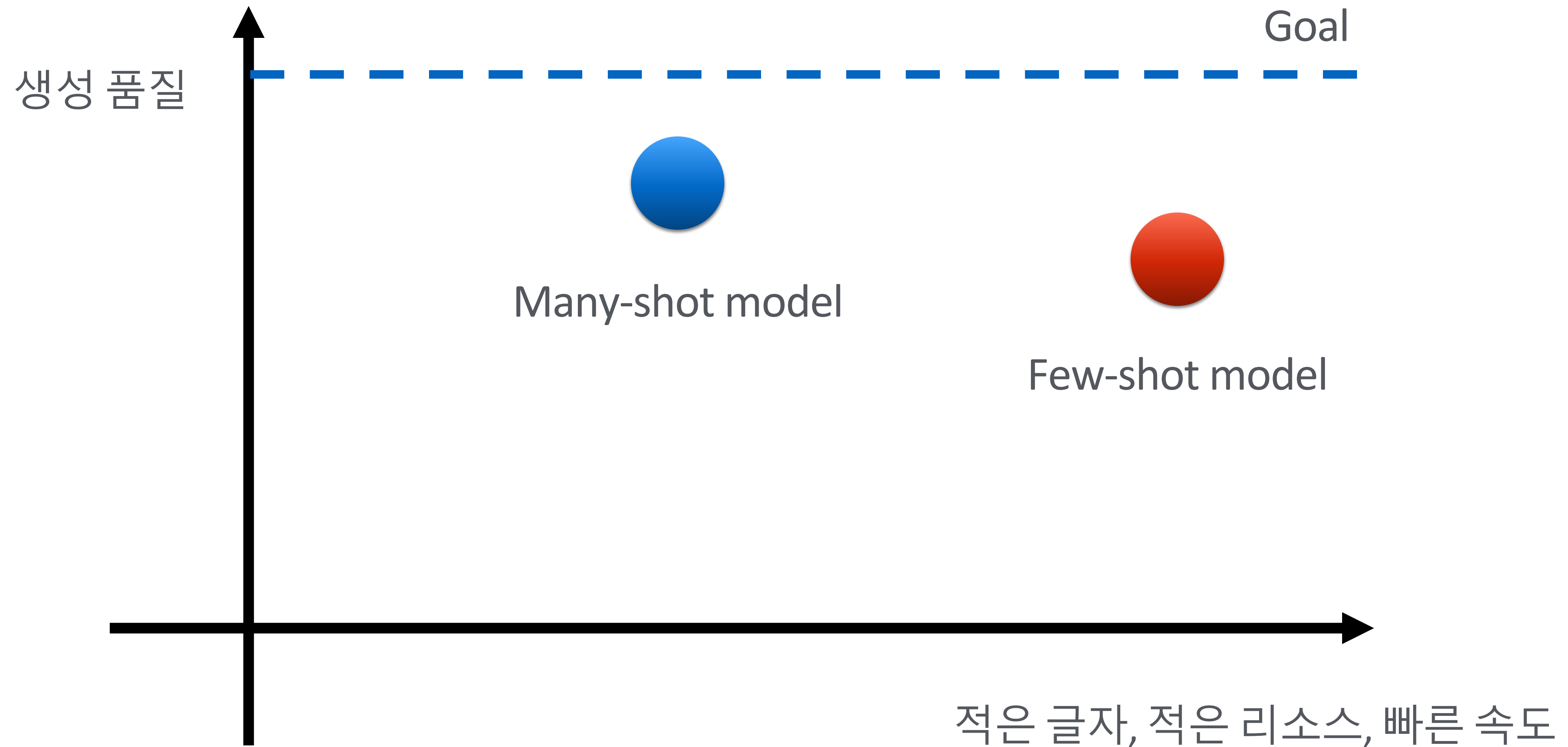
그럼 최신 모델 갖다 쓰면 되겠네요?
4글자면 되고 속도도 빠르다면서요?

아니요 🙄

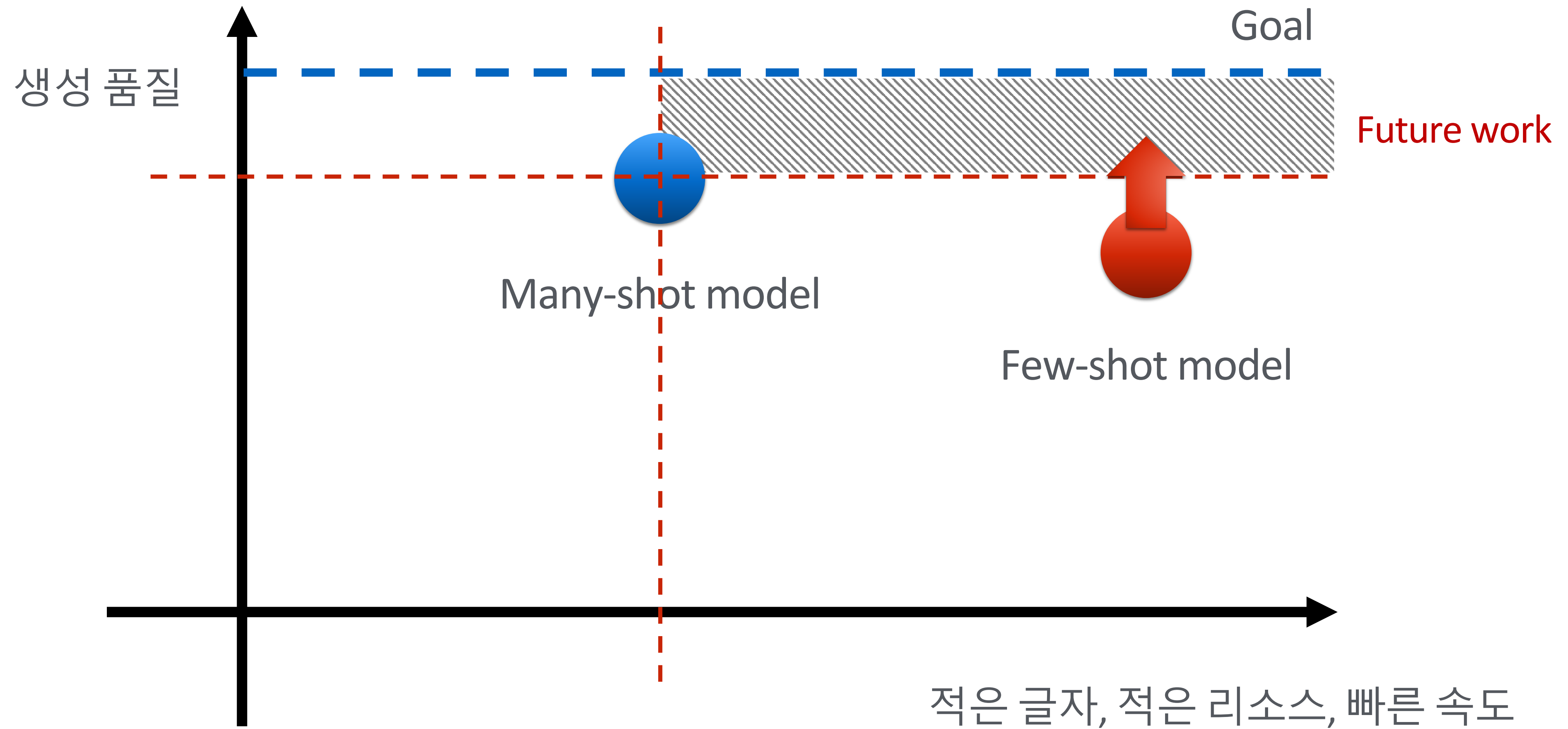
- 글자가 많을수록 더 잘 만들어요
- Finetuning 을 하면 더 잘해요



동전의 뒷면



동전의 뒷면

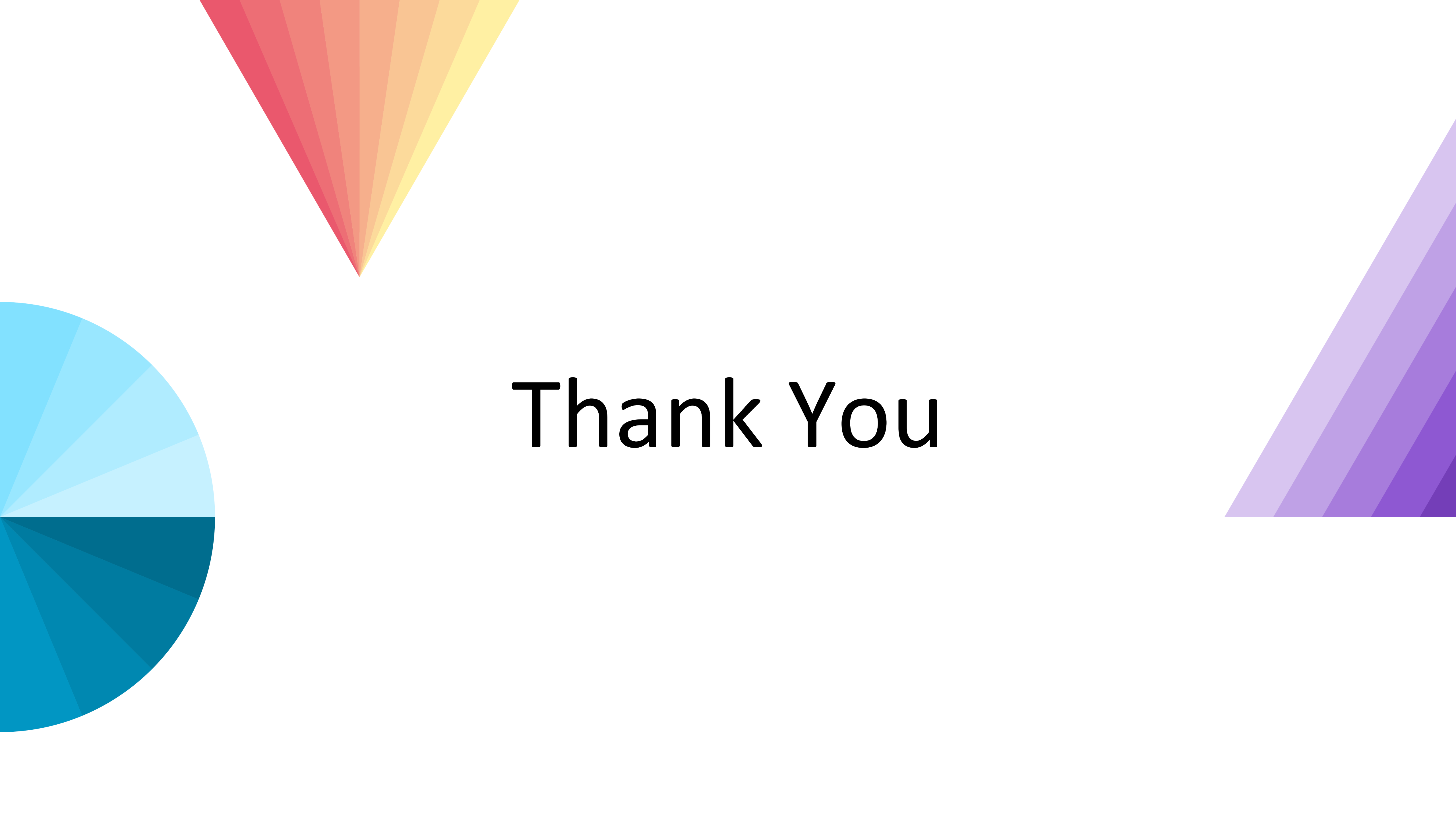


함께 고생하신 분들





Q & A



Thank You