

01. 문제 정의

02. 솔루션

03. 서비스 로직

04. 아키텍처

05. 워크플로우

06. 데모

07. 유사 서비스와의 차별점

08. 모델 평가

09. Future work

10. 회고

3불!



여러 효과 적용 불가

중간에 적용 불가

중간에 해제 불가









동작 효과 적용

3불! 3가능!



여러 효과 적용 불가 가능

중간에 적용 불가 가능

중간에 해제 불가 가능





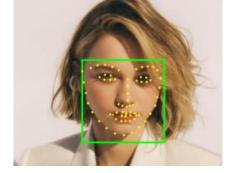
















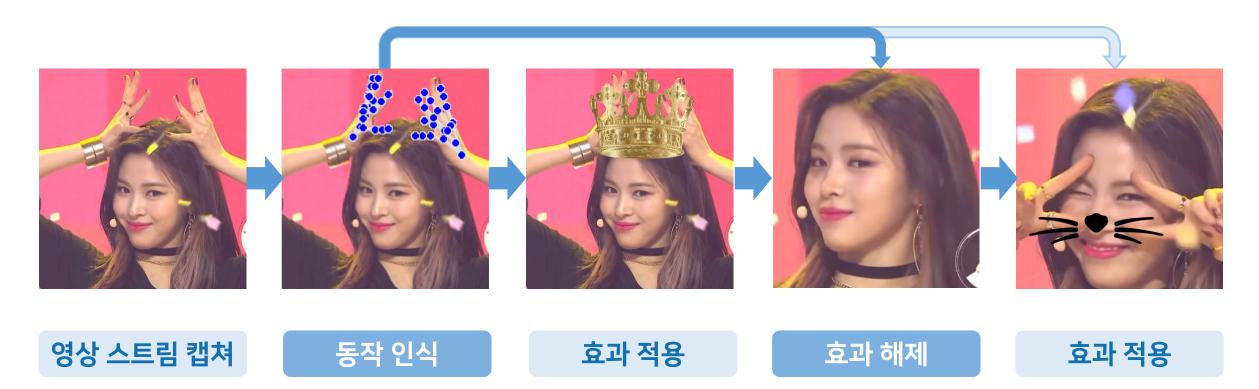
효과 선택

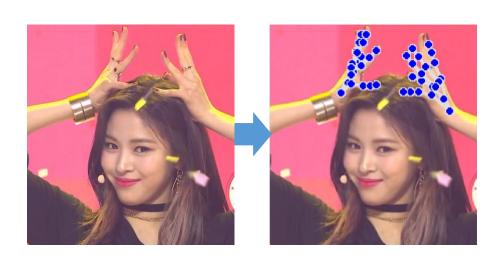
영상 스트림 캡쳐

랜드마크 인식

효과 적용

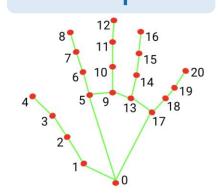
Sorcerer's Hands (Ours)





동작 인식

MediaPipe API



Landmark 추출

- (x, y, z, visibility) * 21 = 84
- 내부 관절 각도(15) = 15
- 한 손 당 99개, 총 198개의 특징

Buffer에 저장 (30회) → 시퀀스

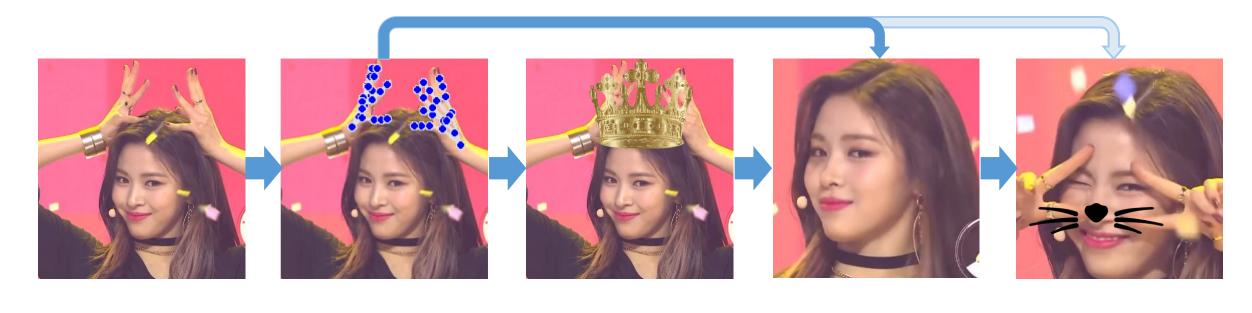
LSTM Model

시퀀스 입력 → 동작 예측

영상 스트림 캡쳐

OpenCV

Real-time Inference & Visualization Layer



영상 스트림 캡쳐

OpenCV

동작 인식

MediaPipe API

LSTM Model

효과 적용

OpenCV

효과 해제

제 효과 적용

- Landmark 기반 위치 및 크기 계산
- 효과 적용 (이미지 합성)

Data Collection & Preprocessing Layer



동작 녹화



Landmark 추출

(x, y, z, visibility) * 21개 Landmark = 84 라벨, 198개의 특징 내부 관절 각도(15) = 15 한 손 당 99개 총 198개의 특징

Feature Engineering

라벨, 198개의 특징 라벨, 198개의 특징

Sequence Buffer

Model Training Layer

LSTM Model

시퀀스 데이터 학습

모델 저장

Data Collection & Preprocessing Layer

동작 녹화

Landmark 추출

Feature Engineering

Sequence Buffer

Model Training Layer

LSTM Model

시퀀스 데이터 학습

모델 저장

Real-time Inference & Visualization Layer

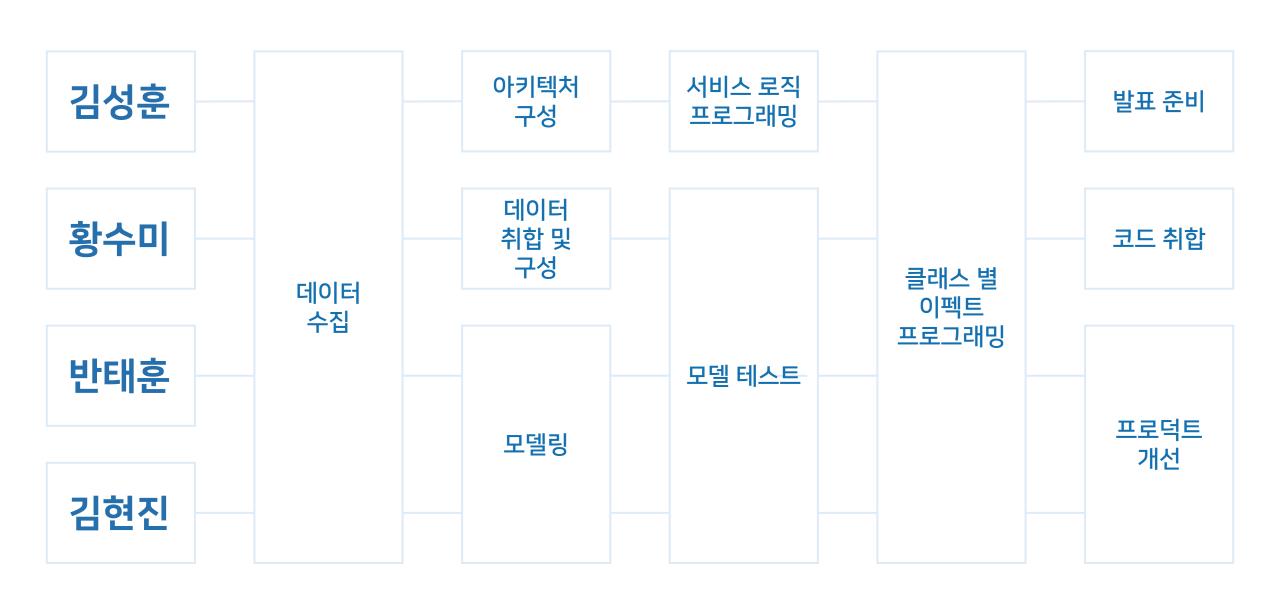
영상 스트림 캡쳐

동작 인식

효과 적용

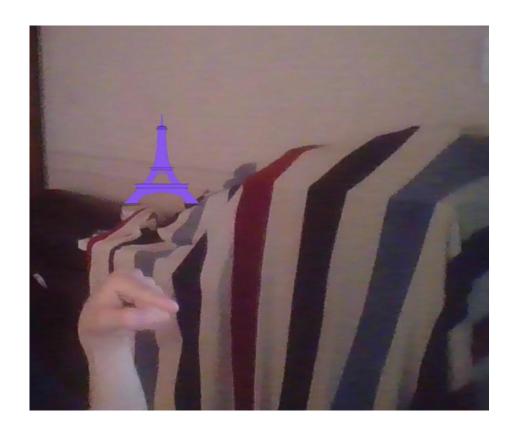
효과 해제

효과 적용

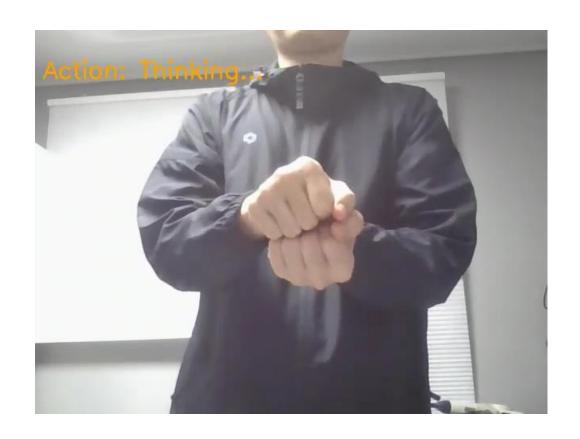














FaceTime APP 손의 모양 (정적 포즈) 인식

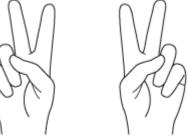




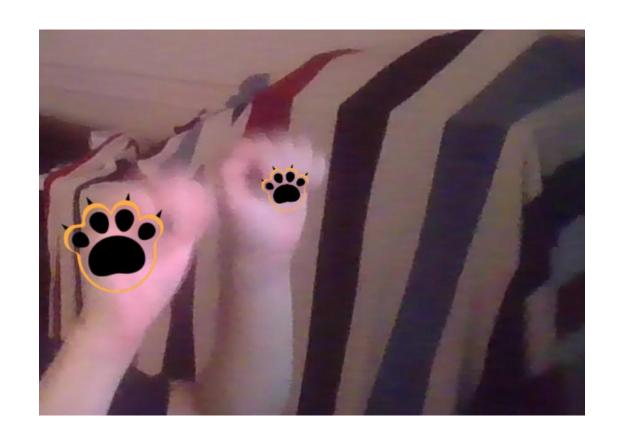


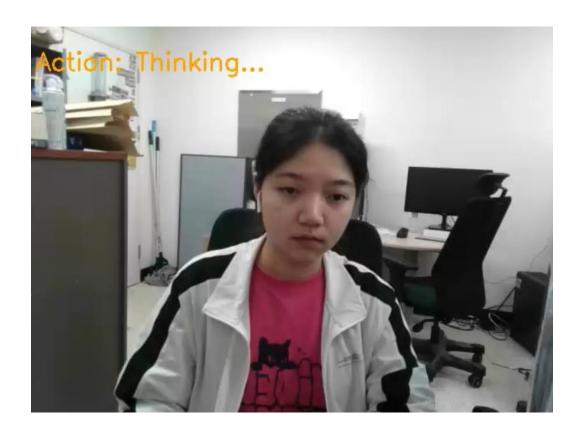






Sorcerer's Hands (Ours) 손의 동작 (동적 포즈) 인식





Sorcerer's Hands (Ours) 손의 동작 (동적 포즈) 인식

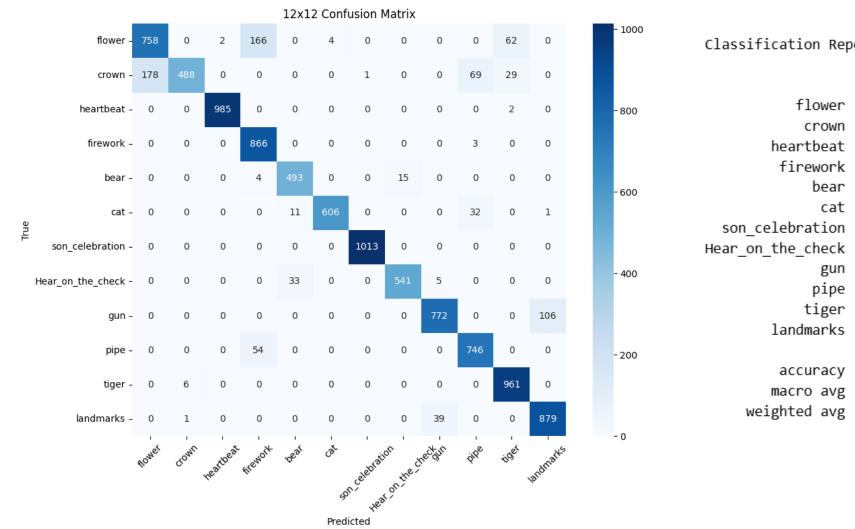




개별 손 모양 순서 인식 방식 vs LSTM을 이용한 동작 인식

12x12 Confusion Matrix																
	flower	914	24	30	13	0	41	0	1	0	3	0	0			Cla
	crown -	- 90	470	0	22	0	0	0	0	0	4	0	16		- 800	CIO
	heart_beat	805	2	159	36	0	1	0	0	0	4	0	0			
	firework ·	- 1	0	0	807	0	5	0	0	0	2	0	3			
	bear -	- 0	0	4	0	540	4	0	5	0	2	0	0		- 600	
	cat ·	- 0	0	0	0	0	665	0	0	0	1	0	0			
	son_celebration	- 1	0	0	0	0	0	740	0	348	0	1	0		- 400	hea
	heart_on_the_cheek	- 0	0	0	0	117	0	0	434	0	0	0	0			
	gun ·	- 0	0	0	0	5	0	549	0	0	11	1	478			
	pipe ·	- 62	0	0	384	13	6	0	49	0	0	378	3		- 200	
	tiger -	- 584	19	0	73	0	57	0	0	0	371	12	23			
	landmarks -	- 0	0	0	0	3	1	514	0	0	0	0	511		0	
nower count hear hear at the profit hear car the profit on the tide mandrants												- 0				
Predicted																

Classification Report:									
	precision	recall	f1-score	support					
			0.50	4005					
flower	0.37	0.89	0.52	1026					
crown	0.91	0.78	0.84	602					
heart_beat	0.82	0.16	0.26	1007					
firework	0.60	0.99	0.75	818					
bear	0.80	0.97	0.88	555					
cat	0.85	1.00	0.92	666					
son_celebration	0.41	0.68	0.51	1090					
heart_on_the_cheek	0.89	0.79	0.83	551					
gun	0.00	0.00	0.00	1044					
pipe	0.00	0.00	0.00	895					
tiger	0.03	0.01	0.02	1139					
landmarks	0.49	0.50	0.50	1029					
accuracy			0.50	10422					
macro avg	0.52	0.56	0.50	10422					
weighted avg	0.46	0.50	0.44	10422					



Classification Report:											
	precision	recall	f1-score	support							
61											
flower	0.81	0.76	0.79	992							
crown	0.99	0.64	0.77	765							
heartbeat	1.00	1.00	1.00	987							
firework	0.79	1.00	0.88	869							
bear	0.92	0.96	0.94	512							
cat	0.99	0.93	0.96	650							
son_celebration	1.00	1.00	1.00	1013							
Hear_on_the_check	0.97	0.93	0.95	579							
gun	0.95	0.88	0.91	878							
pipe	0.88	0.93	0.90	800							
tiger	0.91	0.99	0.95	967							
landmarks	0.89	0.96	0.92	919							
accuracy			0.92	9931							
macro avg	0.92	0.92	0.92	9931							
weighted avg	0.92	0.92	0.92	9931							

Confusion Matrix																				
	flower -	682	0	254	63	0	0	0	0	0	10	17	0		- 800					
	crown -	- 67	94	0	84	0	0	0	0	0	16	0	341		- 700	Classification Rep	ort: precision	recall	f1-score	support
	heartbeat -	136	0	653	52	2	19	4	0	0	81	33	27							
															- 600	flower	0.62	0.66	0.64	1026
	firework -	- 54	0	0	722	7	0	0	0	0	4	30	1			crown	1.00	0.16	0.27	602
	bear -	. 12	0	1	1	494	1	0	36	0	7	3	0			heartbeat	0.70	0.65	0.67	1007
	bear	12	Ü	-	-	151	-	ŭ	50	Ü	,	5	Ü		- 500	firework	0.61	0.88	0.72	818
lec	cat -	- 0	0	0	0	6	645	0	1	0	0	0	14			bear	0.76	0.89	0.82	555
Irue Label															- 400	cat	0.97	0.97	0.97	666
Ĭ	son_celebration -	- 0	0	0	0	0	0	623	0	278	4	0	185			son_celebration	0.40	0.57	0.47	1090
						00		0	477		•					Hear_on_the_check	0.91	0.85	0.88	551
	Hear_on_the_check -	- 0	0	0	0	80	0	0	471	0	0	0	0		- 300	gun	0.19	0.06	0.10	1044
	gun -	- 5	0	0	0	34	0	574	8	67	12	0	344			pipe	0.00	0.00	0.00	895
	3														- 200	tiger	0.12	0.08	0.10	1139
	pipe -	- 34	0	0	200	22	2	0	0	0	0	636	1			landmarks	0.42	0.66	0.51	1029
	tiger -	107	0	27	64	0	0	0	0	0	818	95	28		- 100	accumacy			0.50	10422
																accuracy	0.56	0.54		
	landmarks -	- 0	0	0	0	2	0	342	0	7	0	0	678			macro avg	0.56	0.54	0.51	10422
		o.	20.	, X	, ' \	d	ď	'	,	40	e	05	ا		- 0	weighted avg	0.50	0.50	0.47	10422
		ROWET	Clone	<i>Realtheat</i>	FIENOTH	Ago.	Ş	n celebration eat	on the ch	Recidin	pipe	tiger	landmarks							

Predicted Label











KEEP

- Landmark Detection 관련 배웠던 기술을 좀 더 활용해 볼 수 있어서 좋았다.
- OpenCV 라이브러리 활용에 더 익숙해 질 수 있는 시간이었다. (실시간 처리 등)
- LSTM을 사용하는 이유에 대해서 깊이 고찰해볼 수 있어서 좋았던 거 같아요

Problem

- 시간 부족으로 좀 더 다양한 테스트를 진행해 보지 못한 부분이 아쉽다.
 - 평가 지표를 직접 세우고 테스트해보지 못한 부분 등
- 알고리즘 최적화, 하드웨어 성능 등의 문제로 FPS가 다소 낮아 동작 인식률이 떨어졌다.

Try

• 좀 더 복잡하고 다양한 구분 동작의 연속을 인식하는 모델을 만들어 수어 번역에 도전!