

2025 컴퓨터공학과 캡스톤디자인2 중간 발표

실시간 한영 수어 번역을 위한 어플리케이션 개발

발표자 : 송채린

☐ Team Name : 김송이

☐ Team Members : 이경림, 송채린, 김하현

INDEX

□ 프로젝트 개요 ... 01

□ 시스템 아키텍처 ... 02

□ 핵심 기능 구현 현황 ... 03

□ 기술 시연 ... 04

□ 주요 기술적 성과 및 개선 과정 ... 05

□ 당면 과제 및 해결 전략 ... 06

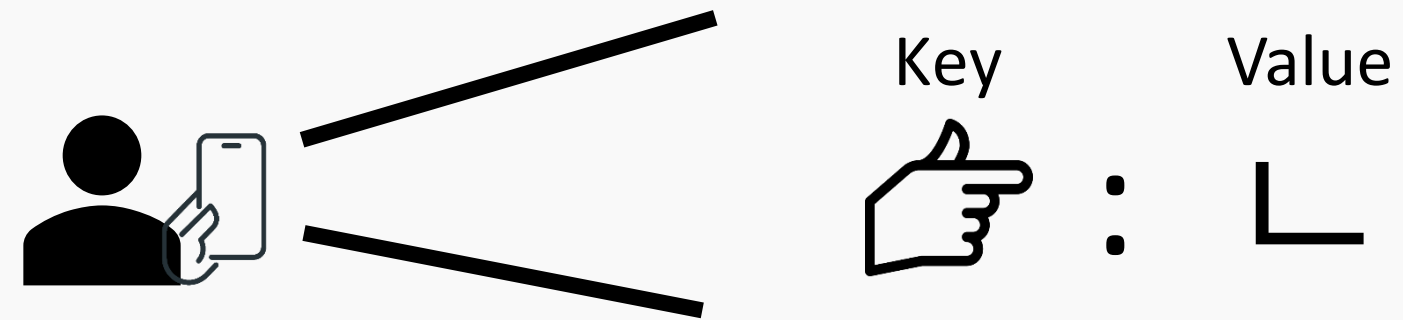
□ 팀원별 역할 및 기여 ... 07

□ 향후 계획 및 리스크 관리 ... 08

01 프로젝트 개요

개발 목표

청각장애인의 의사소통 장벽 해소를 위한 실시간 한영 수어 번역 앱 개발



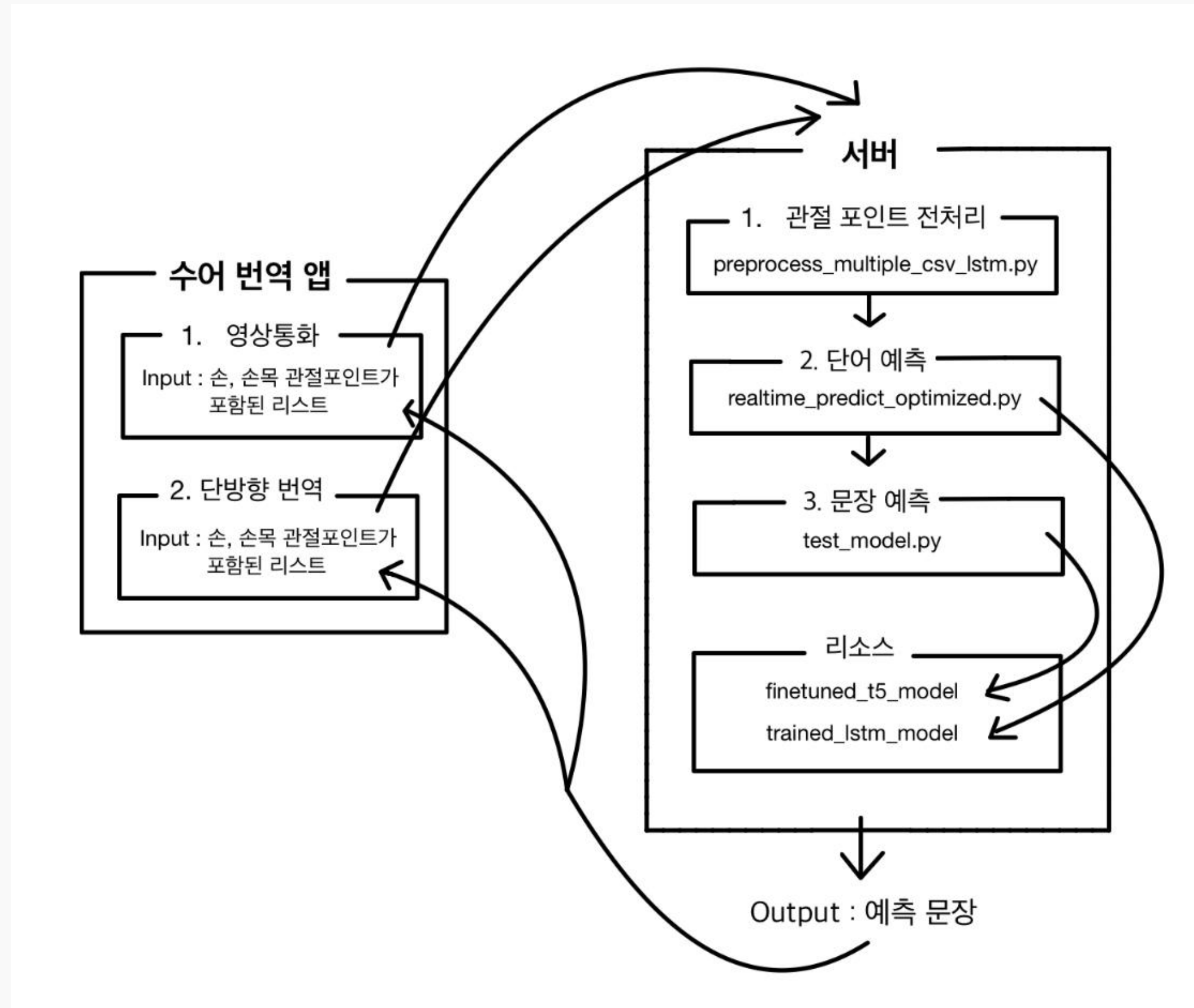
카메라로 인식해 수어 -> 텍스트 변환

핵심기능

□ POINT 01 영상통화 중 실시간 수어 번역

□ POINT 02 카메라를 이용한 단방향 수어 번역

02 시스템 아키텍처



AI 모델파이프라인



STEP 01

MediaPipe



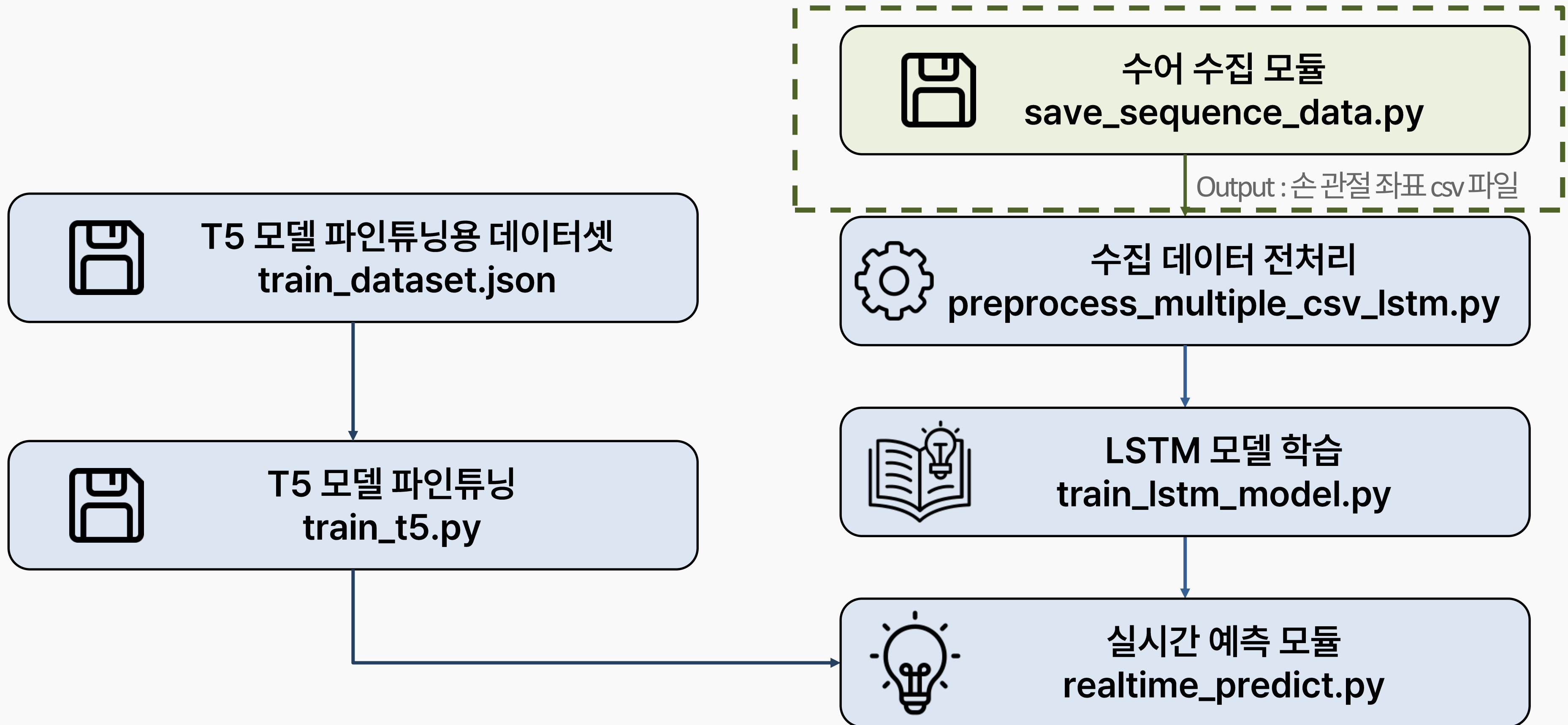
STEP 02

LSTM



STEP 03

T5 Model

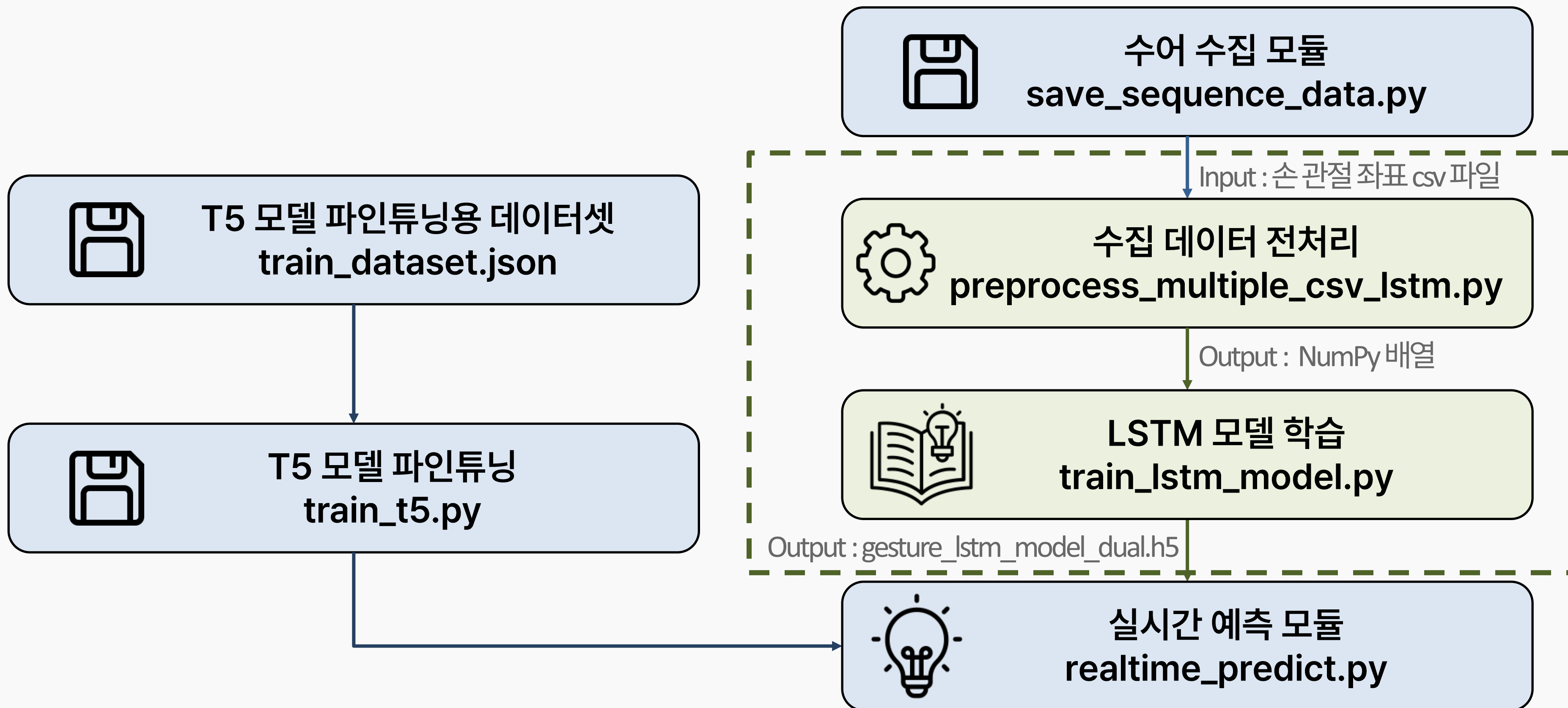


AI 모델파이프라인

 STEP 01 MediaPipe

 STEP 02 LSTM

 STEP 03 T5 Model

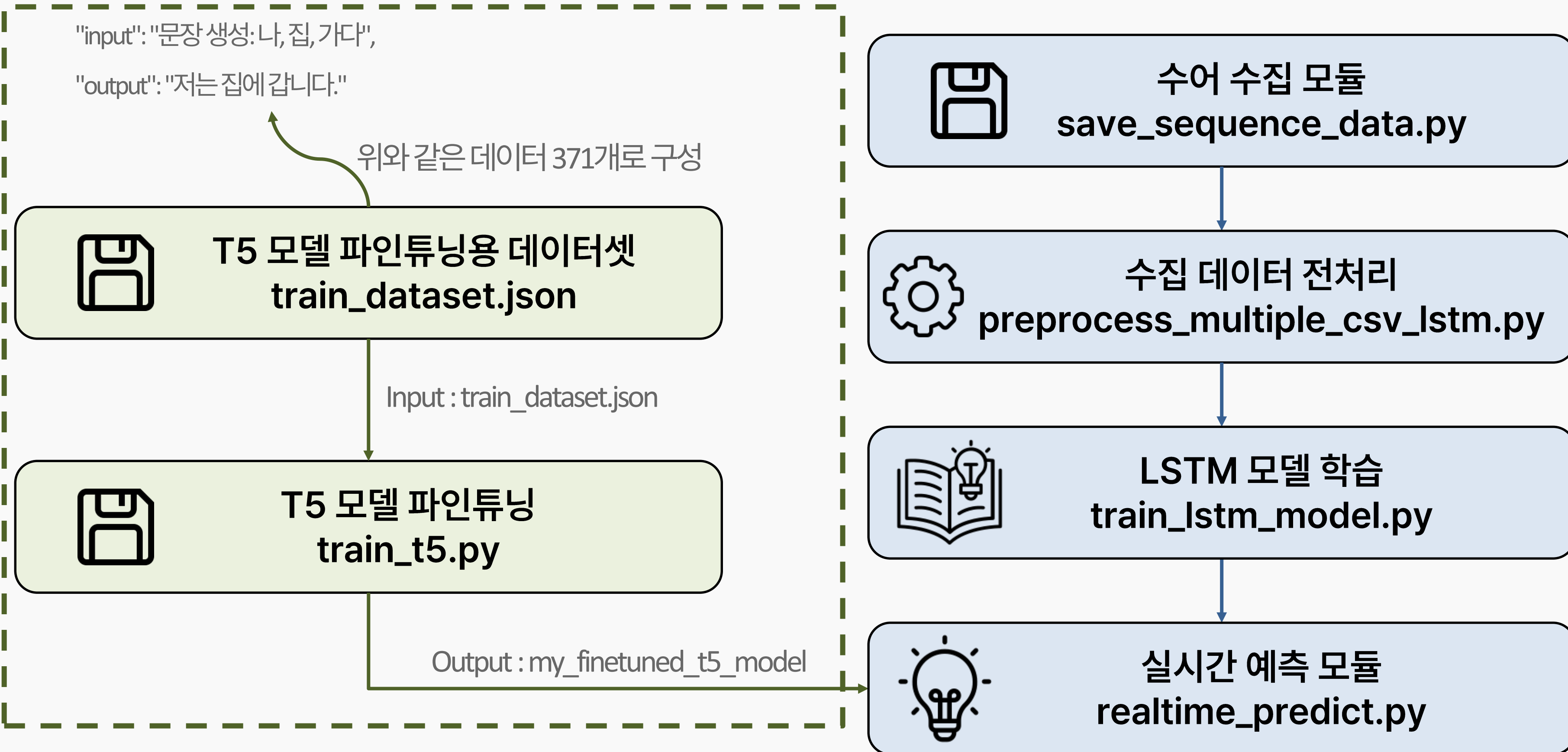


AI 모델파이프라인

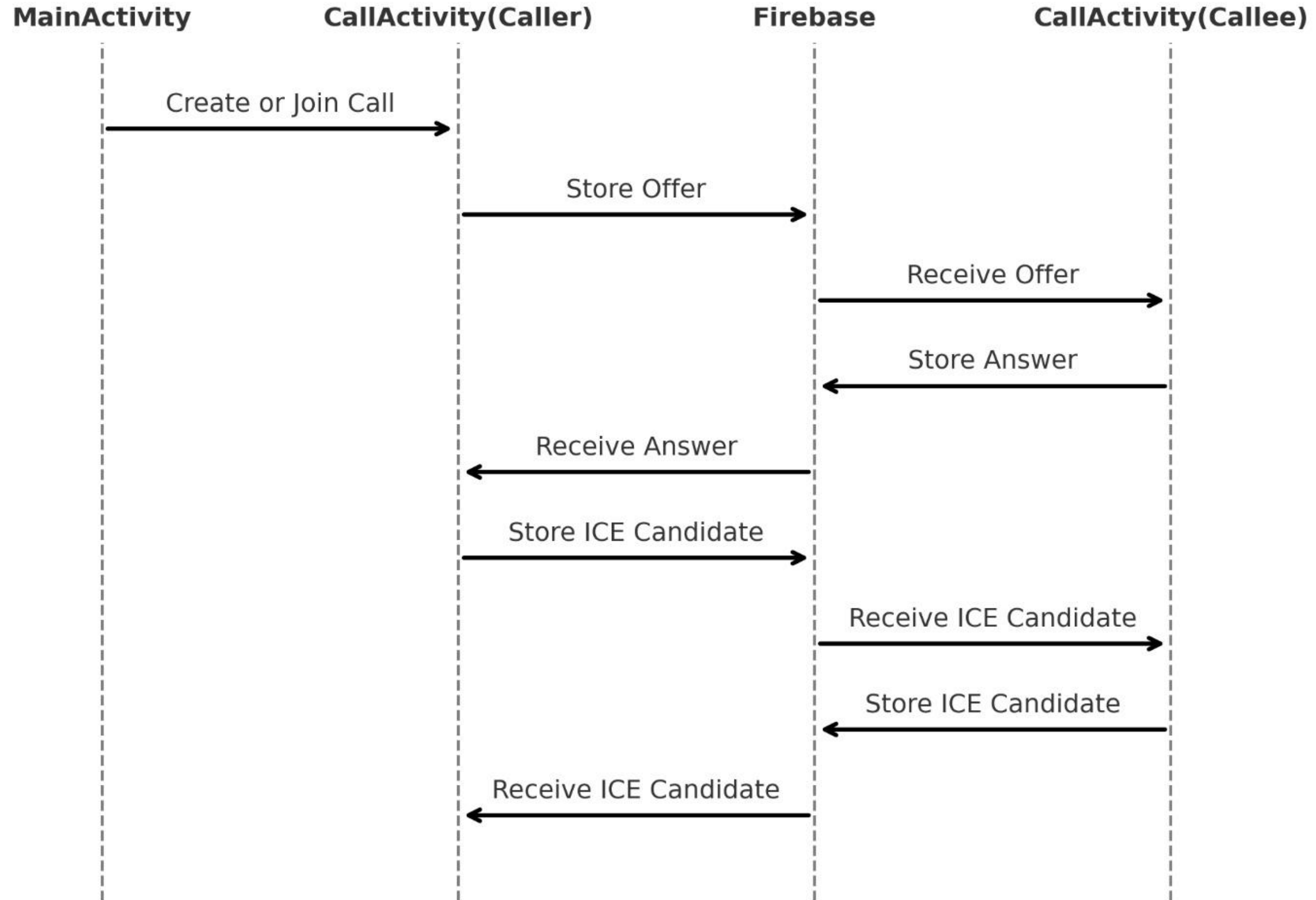
 STEP 01 MediaPipe

 STEP 02 LSTM

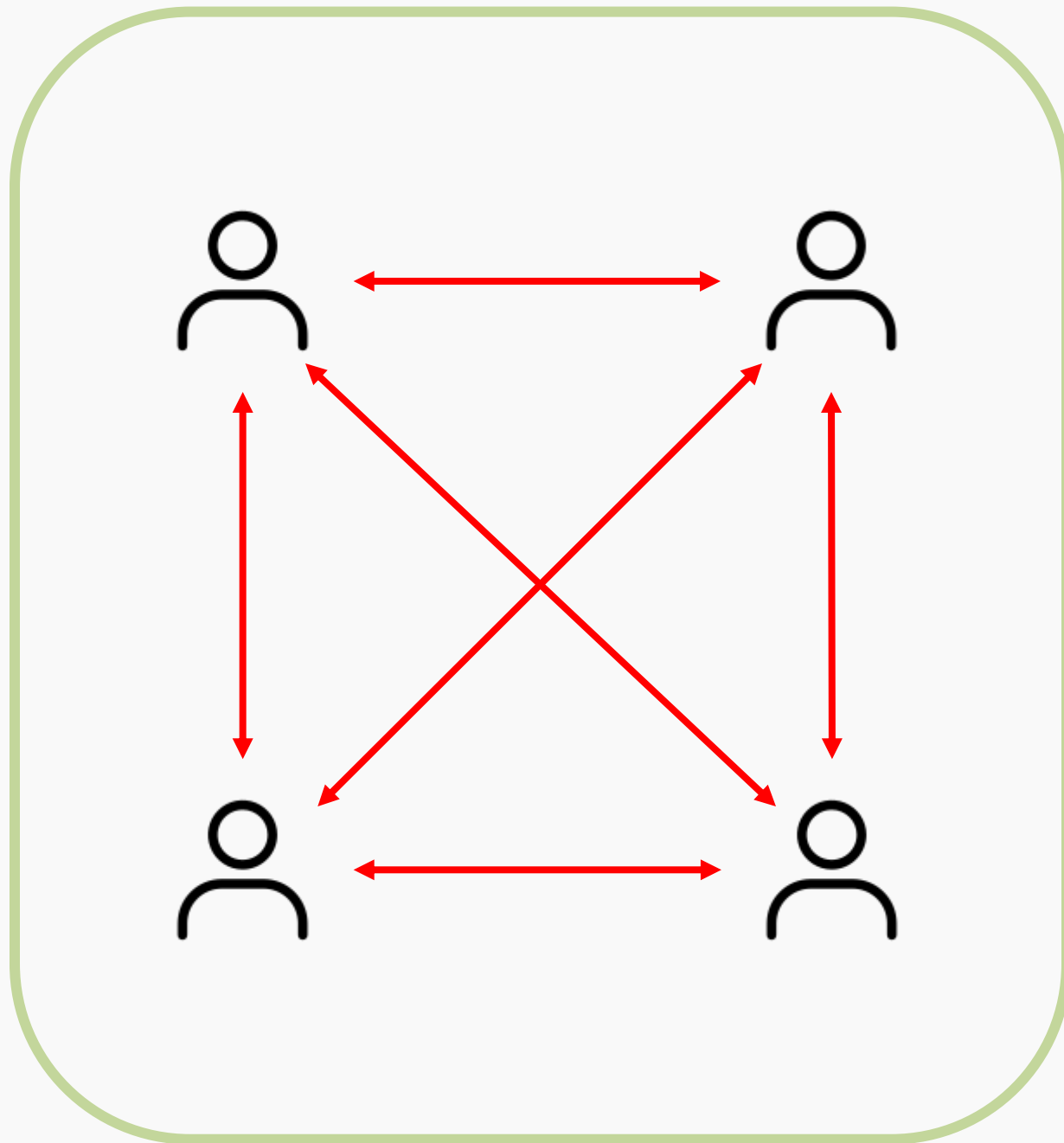
 STEP 03 T5 Model



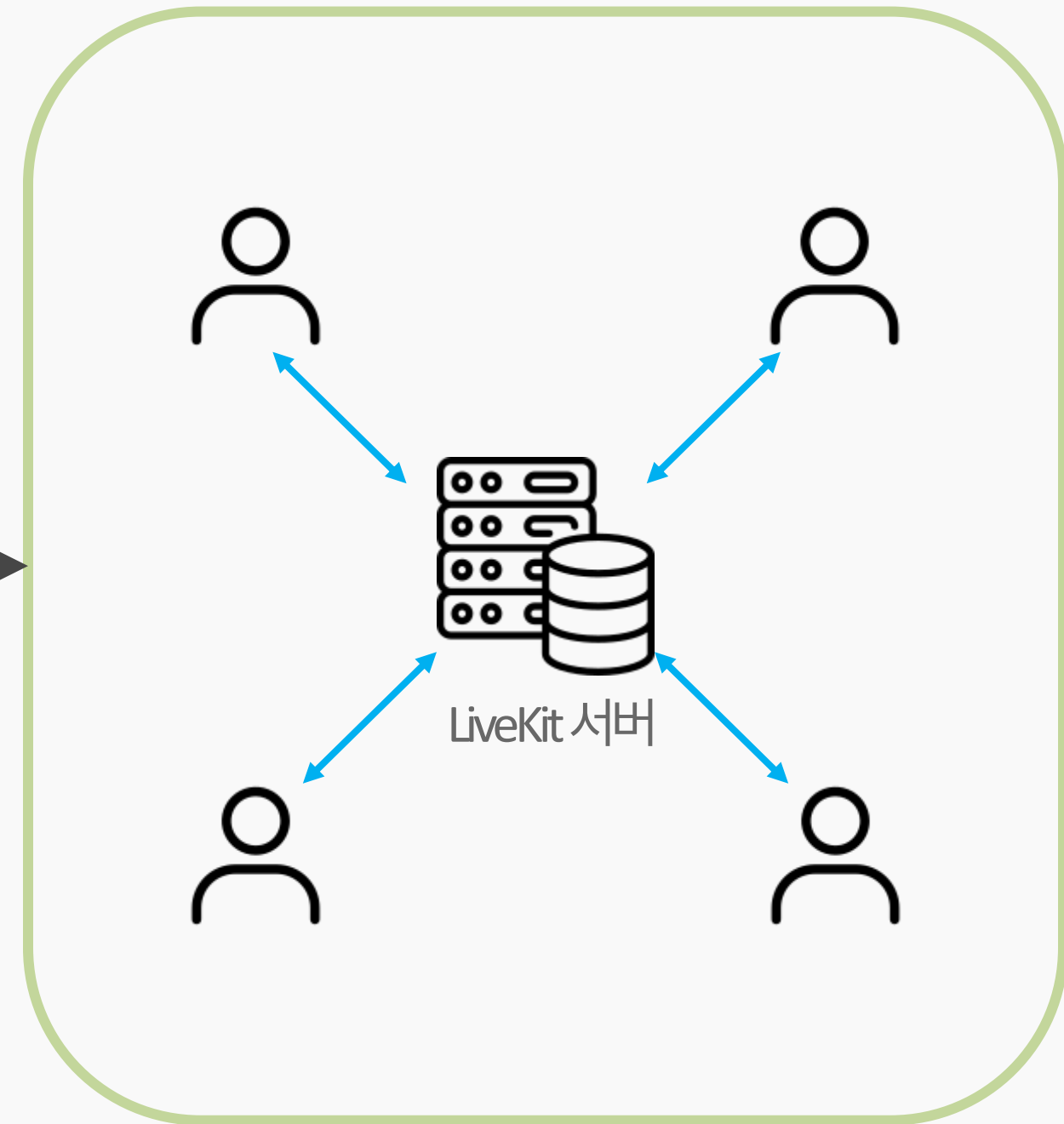
WebRTC + Firebase Signaling Flow



P2P(Mesh) 구조

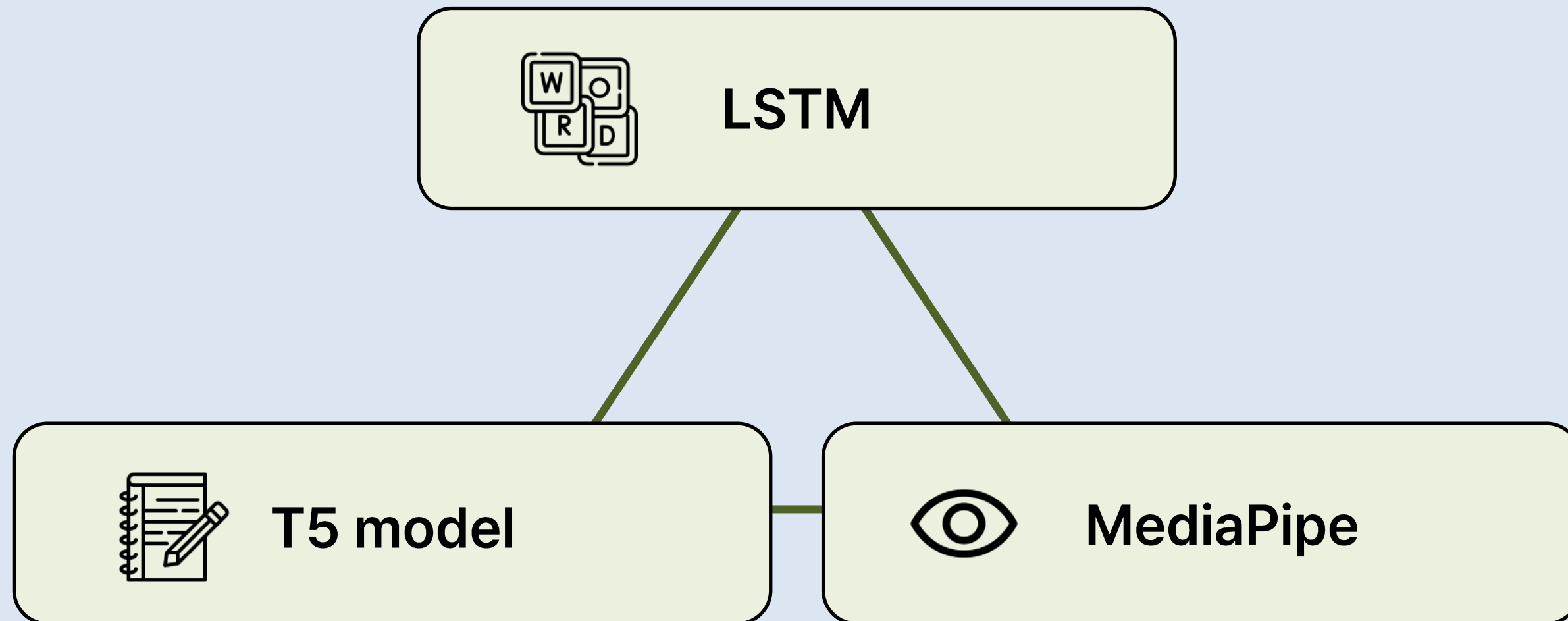


SFU(Star) 구조





Docker, LiveKit 활용한 서버



03 핵심기능구현 현황

기능별 진행 상태

✓ 완료	손 관절 포인트 인식 (Mediapipe)
✓ 완료	수어 단어 예측 (LSTM)
✓ 완료	자연어 문장 생성 (T5)
⚠ 진행중	어플리케이션 UI 및 영상통화 기능
✗ 미구현	서버 구축 및 번역 API 연동

04 기술시연



05 주요 기술적 성과 및 개선 과정

주요 기술적 성과 및 개선 과정



STEP 01

인식 성능 개선



STEP 02

문장 생성 품질 향상



STEP 03

실시간 통신 구조 고도화



동일 동작도 위치에 따라 다르게 인식하는 문제 발생

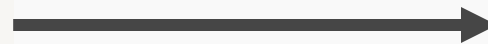
```
# 두 손 데이터 초기화
hand_data = {"Left": [0.0] * 63, "Right": [0.0] * 63}

if results.multi_hand_landmarks and results.multi_handedness:
    for hand_landmarks, hand_handedness in zip(results.multi_hand_landmarks,
results.multi_handedness):
        label_h = hand_handedness.classification[0].label # 'Left' or 'Right'

        coords = []
        for lm in hand_landmarks.landmark:
            coords.extend([lm.x, lm.y, lm.z])
        hand_data[label_h] = coords

        mp_drawing.draw_landmarks(frame, hand_landmarks, mp_hands.HAND_CONNEC
TIONS)

one_frame = hand_data["Left"] + hand_data["Right"]
frame_buffer.append(one_frame)
```



2. 손목 기준 상대 좌표로 정규화합니다.

normalized_left_coords = [0.0] * 63

normalized_right_coords = [0.0] * 63

왼손이 감지되었을 경우

if hand_detected["Left"]:

left_hand_np = np.array(hand_data["Left"]).reshape(21, 3)

left_wrist = left_hand_np[0] # 0번 랜드마크가 손목

relative_left = left_hand_np - left_wrist

normalized_left_coords = relative_left.flatten().tolist()

오른손이 감지되었을 경우

if hand_detected["Right"]:

right_hand_np = np.array(hand_data["Right"]).reshape(21, 3)

right_wrist = right_hand_np[0] # 0번 랜드마크가 손목

relative_right = right_hand_np - right_wrist

normalized_right_coords = relative_right.flatten().tolist()

3. 정규화된 좌표를 하나의 프레임으로 합칩니다.

one_frame = normalized_left_coords + normalized_right_coords



'화면 기준'에서 '손목 기준' 좌표계로 변환하여 인식률을 크게 향상

주요 기술적 성과 및 개선 과정



STEP 01

인식 성능 개선



STEP 02

문장 생성 품질 향상




STEP 03

실시간 통신 구조 고도화


초기 경량 T5 모델이 한국어가 아닌 문장을 생성

KETI-AIR/ke-t5-small

 T5 입력: '문장 생성: 나,집,가다'

✓ T5 생성 문장: 아예原原세상을 여러분이WSJ 바이든litz 김도윤 coalition 물갈이 Islamabad 댄딘 Mattis주길 섬진강 개헌안 블랙리스트 누군가가프로듀스중앙일보 호흡곤란 이승엽 홍성흔 Aleppo블랙리스트 김신회 누군가 삼는 의식한 yen 인플루엔자 정치국 WSJ 봐주국장 laughter위였던 연맹 달러화 정립 모차르트 크라이슬러 물결은 군용기운동 정대세 학생부종합전형 남길 yuangnrاد 별이기도 누구든 위촉 말이오 militants 천송이 희망하는 백남기 출자전환 처우개선KBO관절염

paust/pko-t5-base

 T5 입력: '문장 생성: 나, 집, 가다'
✓ T5 생성 문장: 저는 집에 갑니다.

 다국어 모델(mT5-base)을 기반으로, 한 모델 교체 및 대규모 한국어 데이터 파인튜닝을 통해 자연스러운 문장 생성 성공

주요 기술적 성과 및 개선 과정



STEP 01

인식 성능 개선



STEP 02

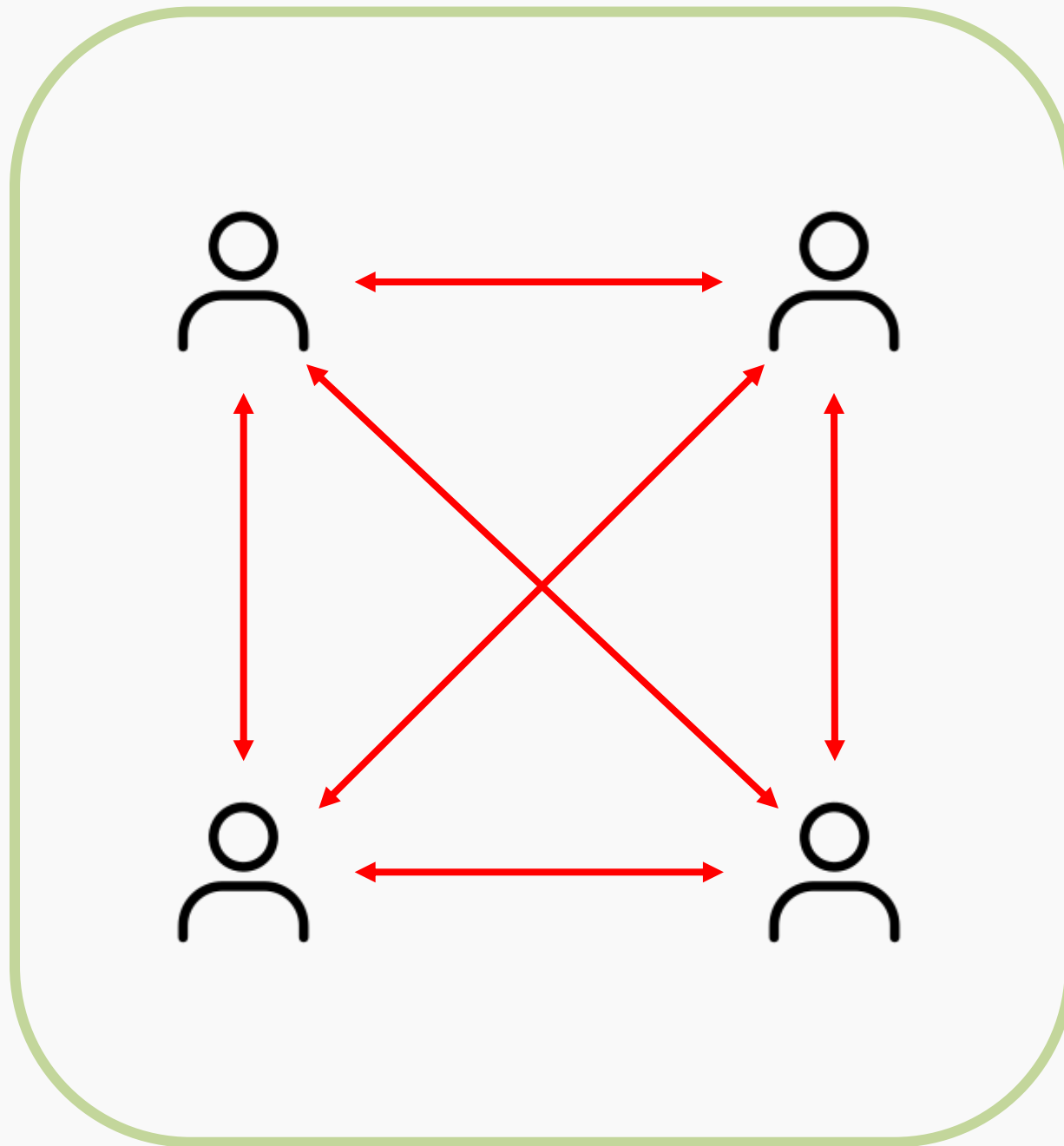
문장 생성 품질 향상



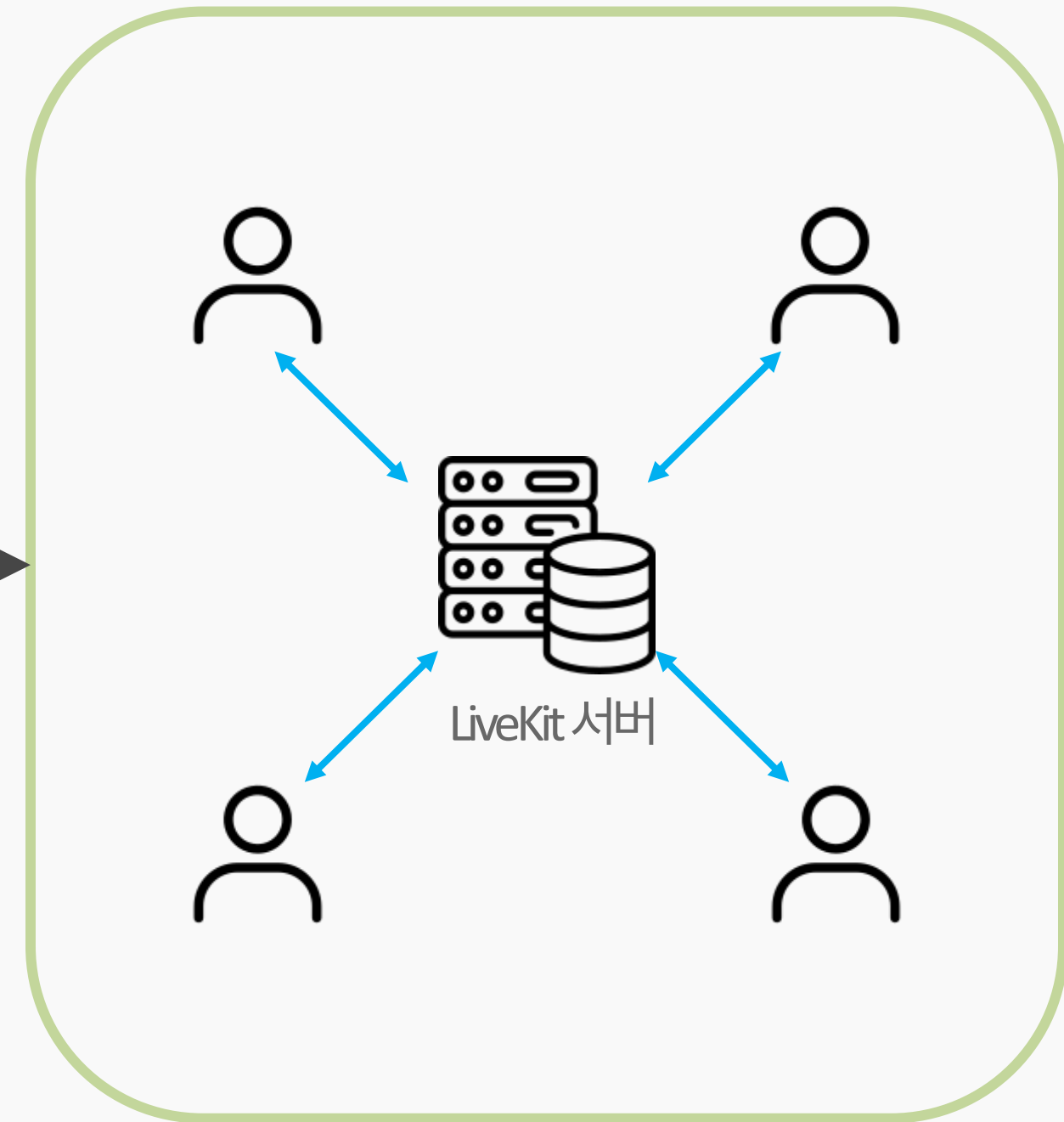
STEP 03

실시간 통신 구조 고도화

P2P(Mesh) 구조



SFU(Star) 구조



06 당면 과제 및 해결 전략

당면 과제 및 해결 전략

 **PROBLEM 01** 모델 크기와 온디바이스의 한계 → 서버-클라이언트 구조로 전환 (서버-모델, 앱-결과값)

 **PROBLEM 02** 모델 학습 시간 과다 소요 → GPU/TPU 가속기 활용으로 학습 시간을 약 1/5로 단축

 **PROBLEM 03** 모델 과적합 문제 → Early Stopping 적용 및 검증 손실 모니터링으로 일반화 성능 확보

07 팀원별 역할 및 기여

팀원별 역할 및 기여



이경림

프로젝트 총괄, GPU 가속 및 데이터셋 확장을 통한 성능 개선



송채린

T5 모델 파인튜닝, 손목 기준 좌표계 변환, AI 기반 시스템 설계 및 구현



김하현

실시간 영상 통신 모듈 구현, WebRTC -> SFU 아키텍처 전환 도입

08 향후 계획 및 리스크 관리

향후 계획

	9월	10월	11월
1단계	번역 API 연동 및 서버 구현		
2단계		앱-서버 통합 및 최종 검증	
3단계			최종 시연 및 보고서 제출

리스크 관리

⚠ 서버 구축 및 연동 작업 지연 가능성

🔑 모듈과 앱 UI 개발을 선행하고,
모듈화된 서버를 단계적으로 통합하여 전체 일정 지연 최소화

Q & A

Thank you
