

6조
Mini-project 발표 자료

프로그램이름 : 나의 음악 선생님

날짜 : 2024-12-04
조원 : 오태영, 박철우

1. 프로젝트 개요 및 주제:

- 본 프로젝트는 증강현실(AR)과 실시간 음악 인식 기술을 결합하여 피아노 학습을 지원하는 시스템을 개발하였습니다. 이 시스템은 실제 피아노 건반과 가상의 AR 인터페이스를 통합하여 학습자에게 직관적인 피드백을 제공하며, MediaPipe를 활용한 손 동작 인식과 음성 분석을 통해 실시간 학습 가이드를 제공합니다.

2. 문제 정의:

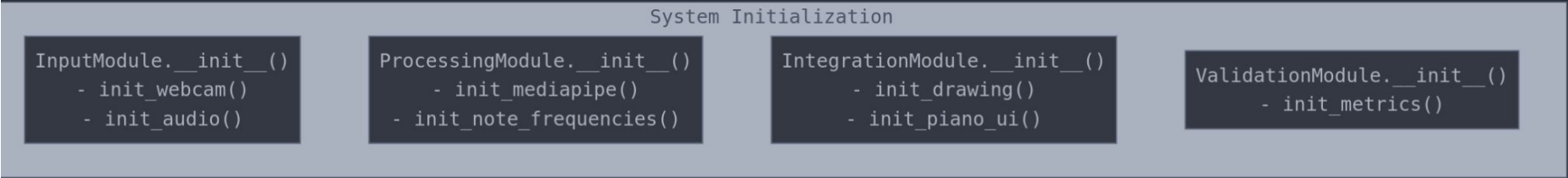
- 악기를 처음 배울때는 악보에 따라 운지법을 맞추어 연주하기가 쉽지않다. 악보, 운지법 모두 익숙하지 않기 때문이다.

따라서, 본 프로그램을 통해 악보없이 운지법에 대한 지시를 받아 연주가 익숙해지도록 연습할 수 있다.

3. 프로젝트 목표:

1. 실시간 상호작용
 - 웹캠을 통한 실시간 손동작 인식
 - 마이크를 통한 실시간 음계 인식
 - 실제 피아노 건반 인식 및 추적
2. 직관적 학습 가이드
 - **AR** 기반 시각적 피드백
 - 실시간 연주 정확도 평가
3. 통합 학습 환경 제공
 - 실제 건반과 **AR** 인터페이스 통합
 - 시각, 청각적 피드백 동기화

4. 시스템 구성도:



1. Input Module

- 웹캠 입력 스트림 처리
- 마이크 오디오 스트림 처리
- 실시간 데이터 캡처 및 전처리

2. Processing Module

- 손동작 인식 및 추적
- 음계 분석 및 매칭
- 실제 건반 인식 및 상태 추적

3. Integration Module

- AR 오버레이 생성
- 시각적 피드백 렌더링
- 사용자 인터페이스 통합

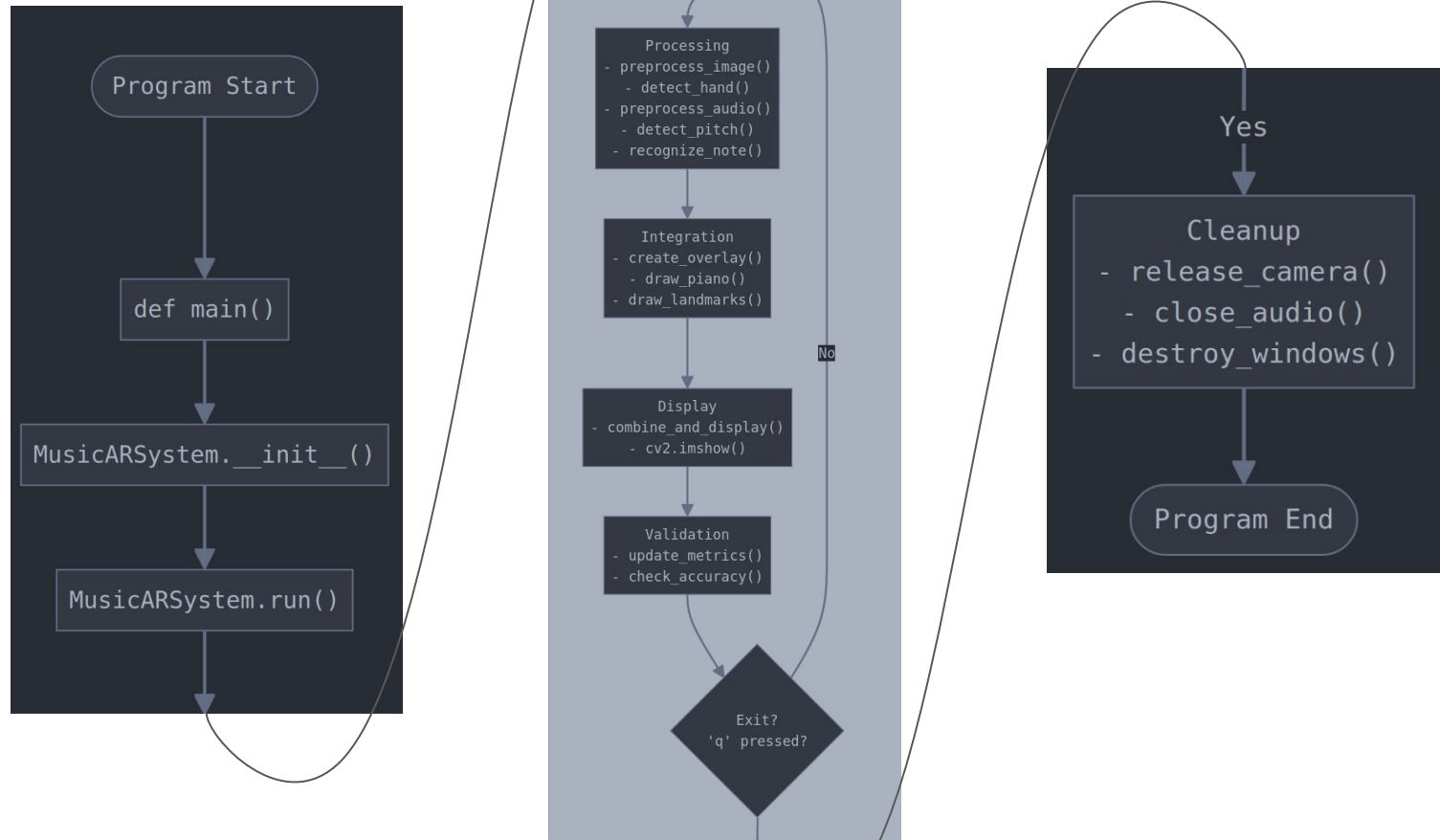
4. Validation Module

- 성능 메트릭스 측정
- 정확도 검증
- 시스템 모니터링

4. 시스템 구성도:



4. 시스템 구성도:



5. 개발 진행:

가능성 확인

원하는 프로그램의 기능을 Chat gpt 및 Claude에 요청하여 기능을 구현

기능정의
명확

↓
구현 결과 예상했던 기능보다 부족하여 기능 정의에 대한 명확성이 중요함을 확인하고 구체적인 기능을 추가적으로 Claude에 요청하여 업그레이드

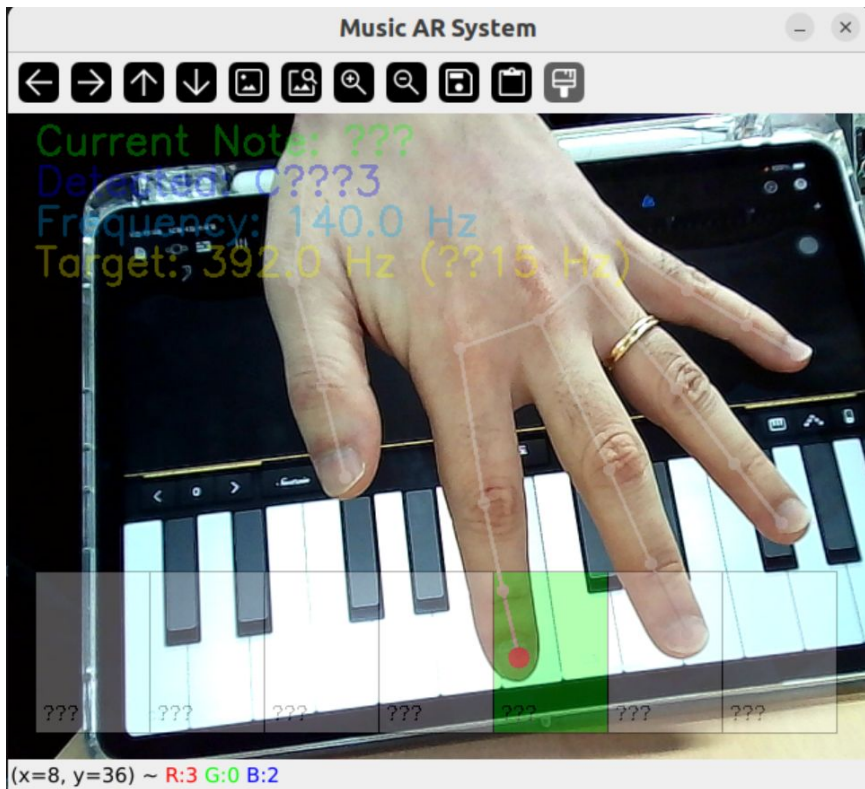
Scope 변경

↓
악기의 종류도 모든 악기에서 피아노건반으로 **scope** 변경 (여러개의 현이나 구멍을 동시에 추적 및 **AR**지시하는 것은 난이도가 높았음)

보완점 개선

↓
추가적인 기능과 화면표시되는 정보를 늘렸으나 추가적인 보완점(노이즈 필터링, 시각적으로 불명확한 부분) 지속 개선중에 있음

6. 시연 및 결과



기능

1. 손동작 인식
 - MediaPipe Hands를 활용한 21개 랜드마크 추적
 - 검지 손가락 끝점 특별 추적
 - 건반 접촉 감지(X)
2. 음계 인식
 - 실시간 오디오 스트림 처리
 - 피치 검출 및 음계 매칭
 - 노이즈 필터링 및 정확도 개선(X)
3. 실제 건반 인식
 - 컬러 기반 건반 검출(X)
 - 건반 상태 추적(X)
 - 손가락-건반 상호작용 감지(X)
4. AR 인터페이스
 - 실시간 오버레이 생성
 - 진행 상태 시각화(X)
 - 정보 디스플레이 통합

7. 고찰:

- 학습과정에서 배운 클래스와 함수구조를 프로그램에 활용하면서 좀 더 복잡하거나 긴 코드에 적응 할 수 있는 기회가 되었다.
- 다양한 기능을 구현하더라도 최초 로우데이터 (이미지, 사운드)에서부터 시작되므로 로우데이터가 어떤 형식으로 저장되고 가공되어 최종 출력되는지 알 수 있었다.
- input -> output 데이터 흐름을 이해할 수 있도록 데이터의 구조를 이해하고 다양한 함수를 활용하여 개발 할 수 있도록 많은 경험을 해야겠다고 생각함
- ML 모델을 더 적극적으로 활용하여 성능을 개선할 수 있는 방향으로 추가적으로 시도해보고 싶음
- 아직 목표로 하는 기능을 완성하진 못하였으나 성능 점검 항목을 통해 구체적인 기능과 수준을

성능 항목 성능 지표*	손 추적 기능 정확도*	손동작+건반 상호작용 감지 성공률*	피치 검출 성능 입력 주파수대비 음표 인식 정확도*	음악 요소 인식 (음표 인식) 연속동작에서 음표 인식 정도*
결과				