ON DEVICE AI

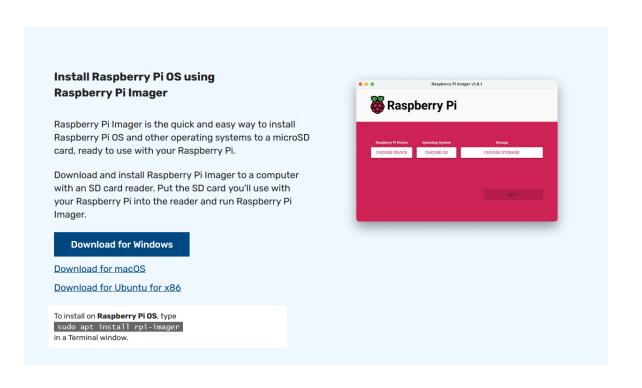
RASPBERRY PI WITH AI

목차

01 환경설정 02 On Device AI? 03 **Use Case** 04 기술적 한계 05 실습

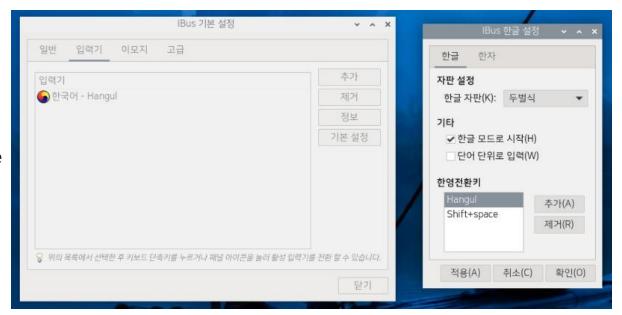
환경 설정

- https://www.raspberrypi.com/software/
- Ubuntu에서 rpi-imager 설치 sudo apt install rpi-imager
- OS를 설치할 SD카드를 PC에 삽입
- \$ rpi-imager (실행)
- 실행 후 64bit OS 선택
- 저장소 선택(SD카드) 후 쓰기 진행



환경 설정

- 한글설정
 - 메뉴에서 Raspberry pi configuration > localization > 'ko(Korean)' & Character Set : UTF-8
- iBus 설치
 - \$ sudo apt update
 - \$ sudo apt upgrade
 - \$ sudo apt install ibus ibus-hangul
- 한글폰트 설치
 - \$ sudo apt install fonts-nanum fonts-unfonts-core
- 한글입력 설정
 - 메뉴 > 기본설정 > iBUS 기본 설정



On Device Al?

- 온디바이스 AI (Edge AI):
 - 클라우드 서버가 아닌 로컬 디바이스(예: 라즈베리파이 5)에서 AI 모델을 직접 실행하는 기술
 - AI 분석과 데이터 수집이 동일한 장소(엣지)에서 일어남
- 필요성 증가 배경
 - 웨어러블 기기, 산업용 IoT, 모니터링 시스템 등 엣지 디바이스 확산
 - AI 기능을 로컬에서 처리해야 할 필요성 증대
- 아키텍처 패러다임 전환
 - 기존: 중앙 집중형 클라우드 AI
 - 변화: 분산형 로컬 인텔리전스 AI
 - 클라우드의 한계를 보완하는 전략적 선택

Use Case

적용 분야

♣ 자율주행 자동차

₩ 산업 자동화

예측 유지보수

◈ 헬스케어 진단

● 증강현실(AR)

이유

실시간 의사결정 필요

네트워크 없는 환경

지연 없이 장비 분석

민감한 개인정보 처리

고속 반응과 저지연 필요

주요 장점

- 클라우드 기반 AI의 한계를 극복(실시간, 미션 크리티컬한 애플리케이션에서 핵심적 역할)
- 개인 정보 보호 및 데이터 보안 강화
- 초저지연 응답 및 실시간 처리 (네트워크를 의존하지 않음 / 오프라인 환경에서도 작동)
- 에너지 효율성과 지속 가능성 (전송 필요 최소화 / 네트워크 대역폭 감소)
- 비용 절감 (인프라 사용량 감소)
- 자동화 및 고급 의사결정 지원 (24시간 무중단 및 사람보다 빠르게 인사이트 도출)

On Device AI의 기술적 한계

- 1. 계산 능력 및 메모리 제약
- 라즈베리파이 5는 강력하지만, 고성능 AI 작업에는 리소스 한계 (예: TensorFlow 전체 실행 어려움)
- 2. 모델 최적화의 어려움
- 온디바이스 개인화는 여전히 복잡
- 고성능 요구 vs 엣지 디바이스 제약
- 개인정보 침해 우려도 존재

- 3. 열 관리 이슈
- 고부하 연산 시 발열 증가
- 열 스로틀링으로 인한 성능 저하 발생
- 적절한 냉각 시스템 필수

On Device AI의 기술적 한계

- 4. 도구 및 개발 생태계 미성숙
- SDK 간 호환성 부족 (예: Core ML, ONNX 등)
- 디버깅·프로파일링 도구 부족
- 플랫폼별 맞춤화 필요

- 5. 대규모 배포의 복잡성
- 다양한 하드웨어에 맞는 배포 자동화 어려움
- 버전 관리, 보안 업데이트, 롤백, 모니터링 필요
- 프라이버시 보장과 운영 효율의 균형 필요

- 6. 윤리적·사회적 문제
- 데이터 편향 → 불공정한 결과 가능성
- AI 결정의 투명성 부족
- 일자리 대체, AI 계층 격차 우려

Raspberry Pi5 SPEC

항목	설명
<u>프로</u> 세서	2.4GHz 쿼드코어 Cortex-A76 탑재 → 이전 모델 대비 성능 대폭 향상
🧠 메모리 (RAM)	4GB / 8GB / 16GB LPDDR4 → 대형 AI 모델 실행 및 병렬 처리 가능
⋒ GPU	VideoCore VII → 이미지 처리 및 경량 ML 작업에 적합
♥ PCIe 2.0 지원	M.2 HAT+ 보드를 통한 SSD 및 AI 가속기 연결 가능
🔅 확장성	GPIO 헤더, 카메라/마이크 모듈 연결 용이 → 다양한 센서 및 입출력 장치 연동
→ 전원 시스템	5V 5A USB-C 전원 공급 필요, 전용 전원 버튼 포함
🚀 AI 프로젝트 적용성	AI 가속기(NPU 등) 연결을 통해 복잡한 AI 추론 작업도 처리 가능
플 플랫폼 진화	기존 CPU/GPU 기반 한계를 극복 → 고성능 엣지 AI 플랫폼으로 전환 가능

실습

