

Computing in Memory AI 가속기 구현 및 On-Sensor AI 응용 적용

고성능 응용에서 저전력 응용으로의 AI 기술 적용 분야 확대



초 고성능 클라우드



고성능 엣지 디바이스



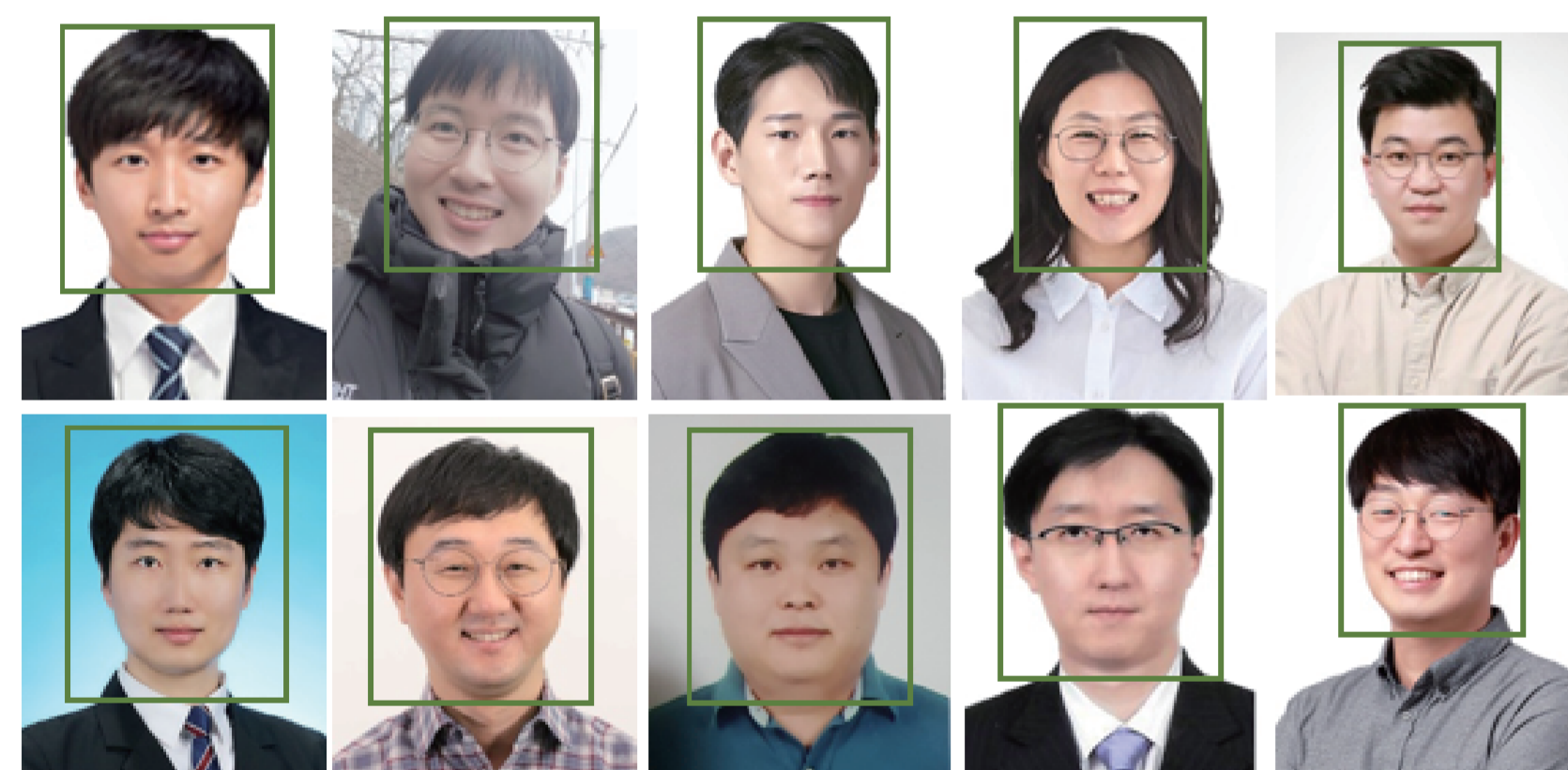
초 저전력 엣지/IoT 디바이스

고성능/고정밀도 AI 기술
 고성능 컴퓨팅 + 고사양 메모리 + 고정밀도 센서
 ★ Peta-/Tera-Operations (초)
 수십 ~ 수백 Watt (Chip)

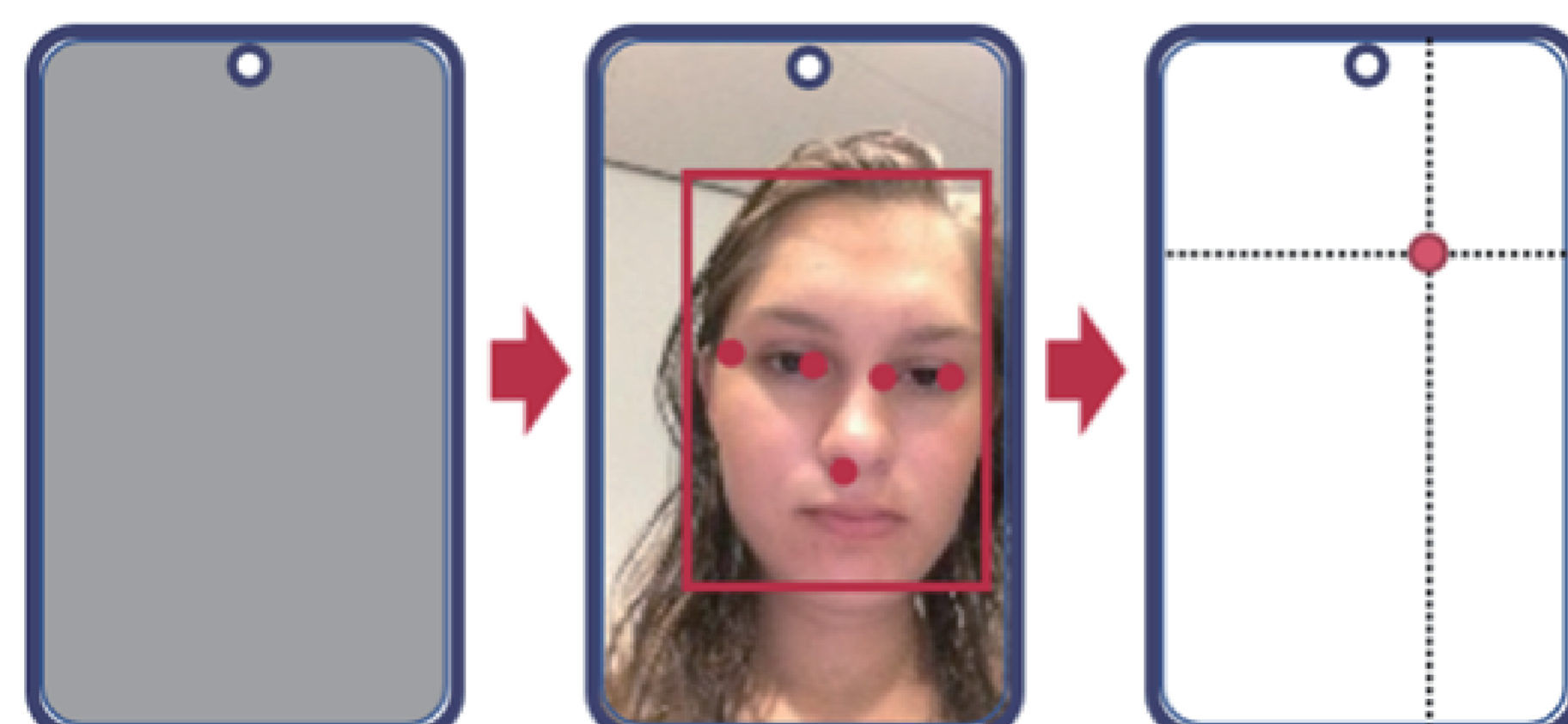
초 저전력 AI 기술
 저전력 컴퓨팅 + 저전력 센서
 Giga/Mega-Operations (초)
 ★ 수 ~ 수십 mW (Chip)

On-Sensor AI 응용 사례

응용 1 Face Detection (on CIS)



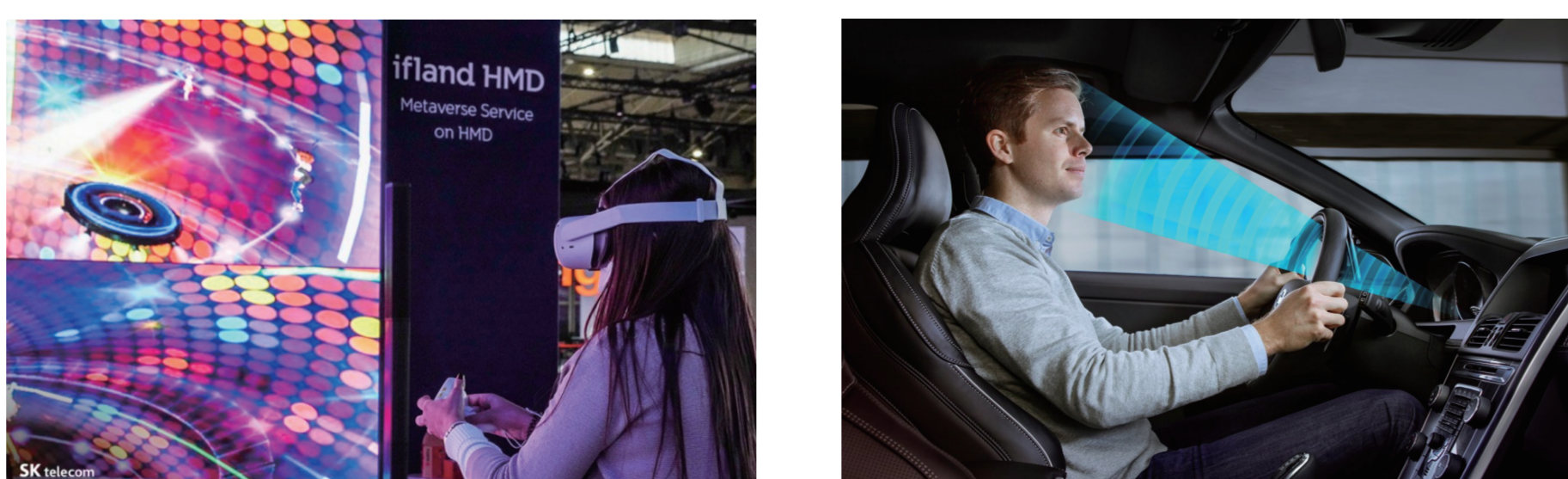
응용 2 Gaze Tracking (on CIS)



응용 3 Hand Detection (on CIS)



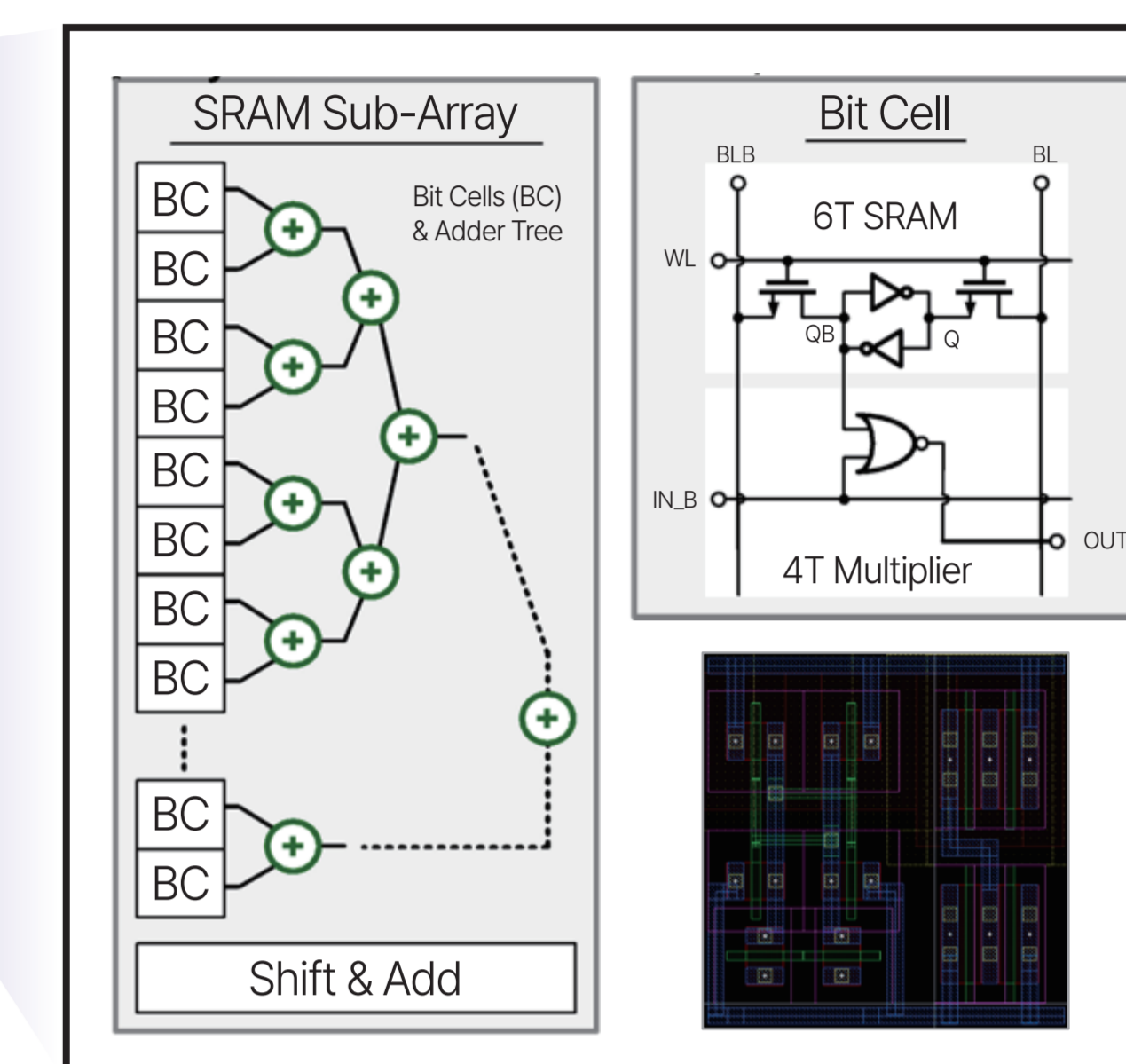
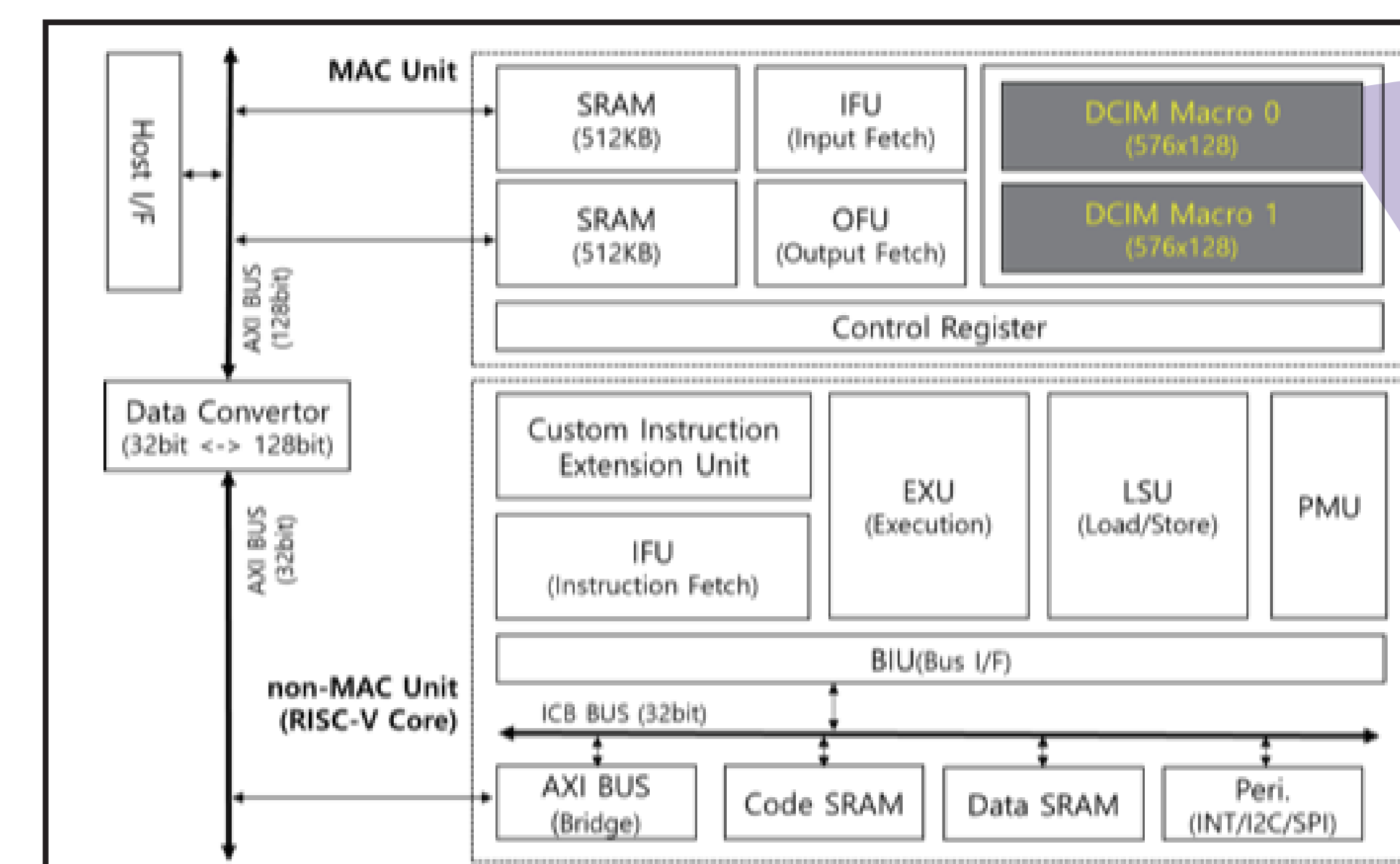
기타 응용 AR/VR HMD, Driver Monitoring 등



On-Sensor AI 요소 기술

- 신경망 최적화: 타깃 디바이스 특화(예: CIS 센서), 응용 특화(예: Face Detection), 신경망 경량화(예: Pruning, 양자화) 등
- 초 저전력 AI 가속기: CiM AI 가속기 구현 기술, Digital CiM용 SRAM Macro 구현 기술, Analog CiM용 ReRAM Macro 구현 기술 등
- Killer 응용 발굴: 지능형 센서 기반 IoT, Surveillance camera 등

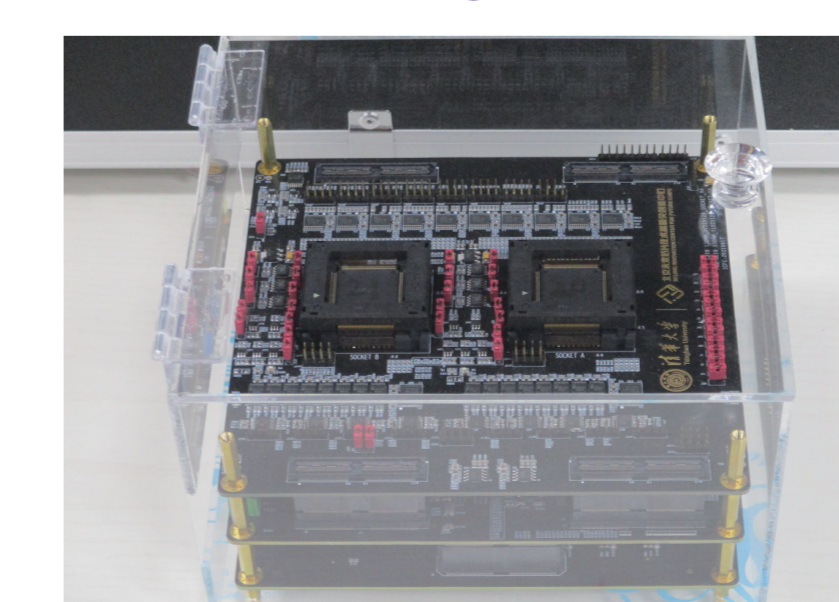
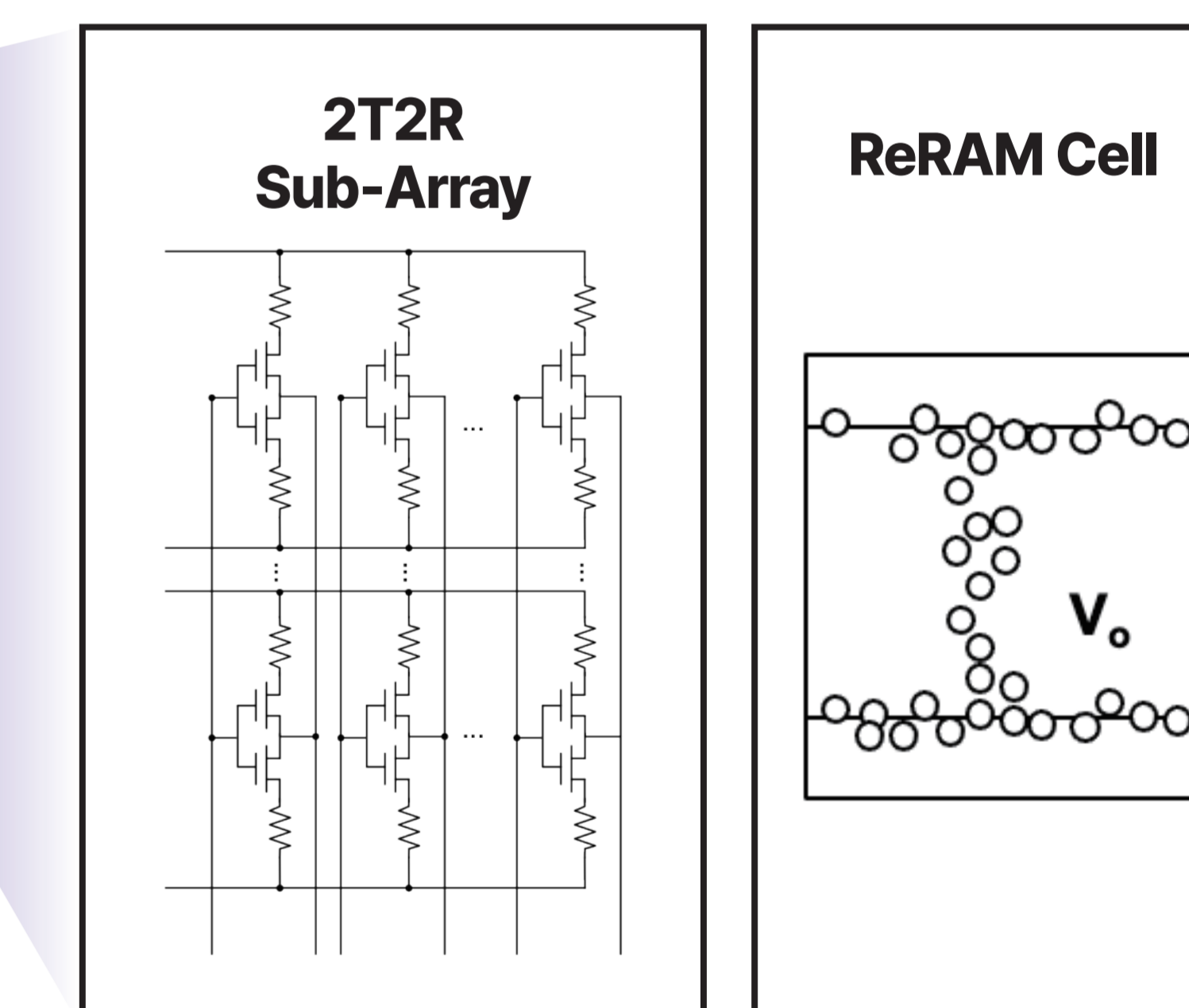
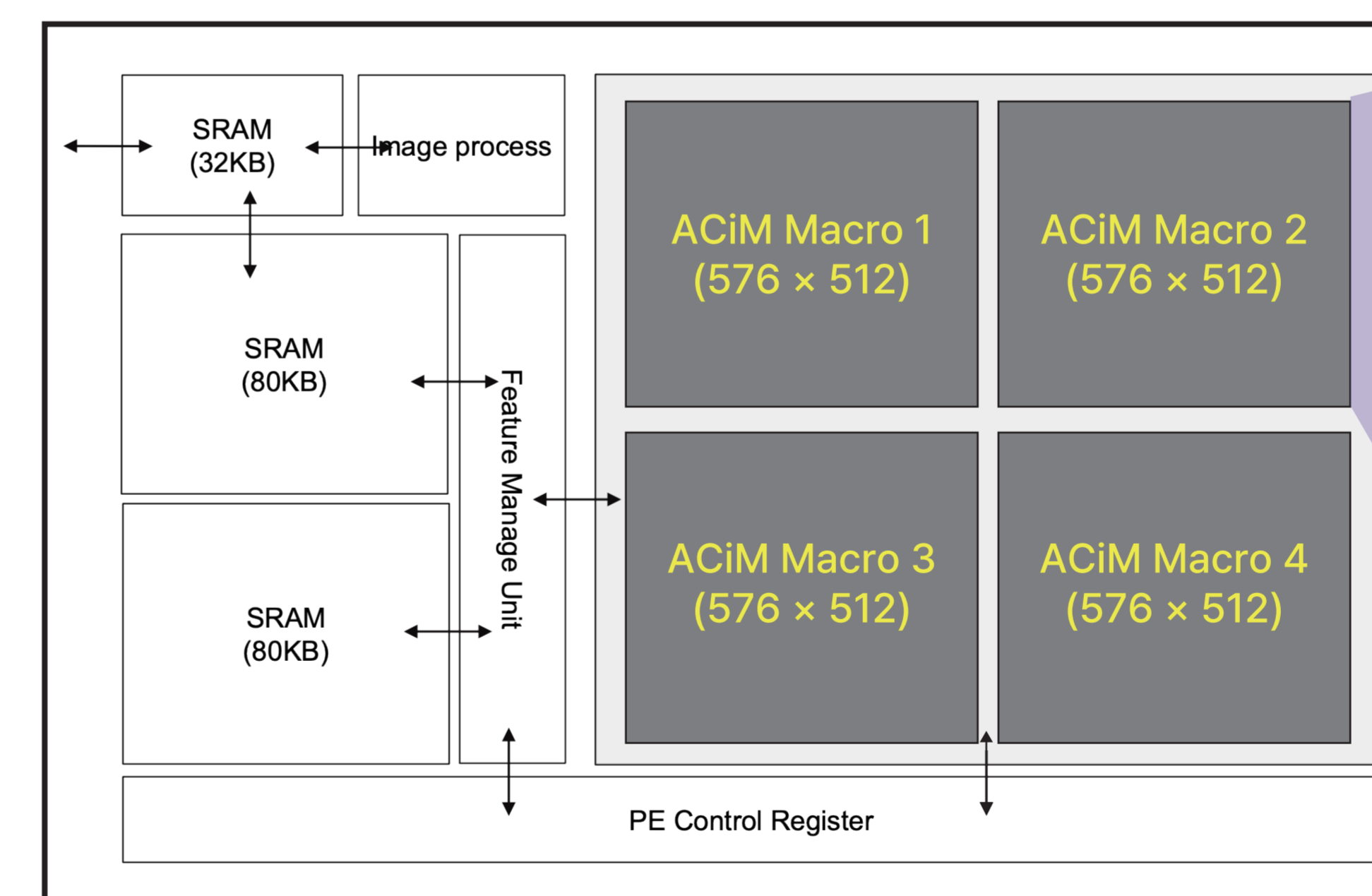
Digital CiM) SRAM based CiM AI Accelerator



	Perf.	Power	Area
@40nm	14.7TOPs	227mW	3mm ²

- TOPs/W: 약 65TOPs/W (DCiM Macro level)
- DCiM Macro: 576×128@INT4

Analog CiM) ReRAM based CiM AI Accelerator



	Perf.	Power	Area
@28nm	1.8TOPs	72mW	10mm ²

- TOPs/W: 약 25TOPs/W
- ACiM Macro: 576×512@INT4