

RFP관리번호	2025-반도체·디스플레이-품목공모-04			공모유형	품목공모형	
해당여부	<input checked="" type="checkbox"/> 국가전략기술 <input type="checkbox"/> 탄소중립 <input type="checkbox"/> 글로벌 R&D <input type="checkbox"/> 미래소재 <input type="checkbox"/> 전략연구사업(MPX예정) <input type="checkbox"/> 국방전략기술(예정)					
국책연구기획 평가전문분야1	PM분야	반도체· 디스플레이	RB분야	SI반도체	RB세부분야	뉴로모픽 소자
사업명	원천기술개발사업 - 차세대지능형반도체기술개발(소자)(R&D) - 신소자원천기술개발					
RFP명	공동집적 뉴런-시냅스 소자 기반 뉴로모픽 컴퓨팅 하드웨어 개발					
	(TRL : [시작] 4단계 ~ [종료] 5단계)					
RFP유형코드	사업목적·내용	성과물 특성		지원대상	보안과제 분류	일반
	R	1	-	1		
1. 추진배경						
<div> <input type="checkbox"/> 추진배경 <ul style="list-style-type: none"> ○ 현재 사용되고 있는 인공지능 기술은 데이터를 처리하는 데에 많은 행렬 연산과 메모리 접근이 필요하지만, 기존 소프트웨어 기반의 GPU나 CPU 구조는 효율적인 연산을 수행하는 데에 한계가 있음. 하드웨어 기반 인공지능 기술 구현은 전력 효율과 실시간 데이터 처리 성능 면에서 큰 이점이 있기 때문에, 인공지능이 처리해야 할 데이터가 방대해지고 있는 현 시점에서 매우 필수적인 기술임. ○ Spiking Neural Network (SNN) 인공 신경망 모델은 이벤트 기반의 정보 처리를 수행하며, 기존의 Deep Neural Network (DNN) 형태의 신경망 모델에서 모든 뉴런이 연산을 지속적으로 수행하는 것과는 달리 필요할 때에만 연산을 수행하기 때문에, SNN을 활용한 하드웨어 기반의 인공지능 모델은 기존 소프트웨어 기반의 인공지능 모델보다 에너지 효율적으로 연산을 수행할 수 있음. ○ 기존 SNN을 활용한 뉴로모픽 칩의 대표적인 사례로 Intel의 Loihi와 IBM의 TrueNorth가 있음. 이들은 CMOS 회로로 구현된 뉴런과 SRAM 기반의 시냅스로 구성되어 데이터를 효율적으로 처리할 수 있는 구조를 갖추고 있음. 그러나 회로 기반 뉴런과 시냅스가 차지하는 하드웨어 면적이 크고 전력 소비가 높은 한계가 있음. ○ 에너지 효율과 소자 성능 측면에서 우수한 뉴런과 시냅스 소자에 대한 연구가 다수 이루어지고 있으나, 뉴런과 시냅스를 한 시스템 내에서 구현하기에는 CMOS 회로와의 공정 호환성과 수율 측면에서 어려움이 있어 이를 구현한 결과는 찾기 어려움. ○ 뉴런, 시냅스, 인터페이스 회로가 통합된 초고집적 뉴로모픽 시스템을 구현하기 위해서 동일 웨이퍼 내에서 공동 집적하는 기술 혹은 적층 기술이 필요함. 예를 들면, 뉴런, 시냅스, 인터페이스 회로를 적층 구조로 구현하여, 집적도를 높임과 동시에 소자와 회로 간의 라인 길이를 감소시켜 데이터 전송 속도를 높이고 집적 시 발생하는 비이상성을 최소화 할 수 있음. </div>						

□ 기획 주안점

- 연구 사업단의 최종 목표 및 진행 방향에 대한 본 연구 개발 과제의 수월성을 위해 관련 선행 연구 결과를 기 확보할 필요성이 있음
- 뉴런과 시냅스의 개별 소자에 관한 연구를 넘어, 이들을 하나로 통합하여 인공신경망 모델을 구동할 수 있는 연구가 필요함. 이를 위해서는 뉴런과 시냅스 소자를 통합하는 하드웨어 플랫폼과 이를 지원하는 소프트웨어 프레임워크 연구가 필수적임.
- (하드웨어 플랫폼) 실리콘 동일 웨이퍼 상에서 뉴런, 시냅스, 인터페이스 회로를 공동 집적하기 위해 뉴런과 시냅스 소자의 CMOS 공정 호환성이 보장되어야 함. 특히, 시냅스 소자의 경우 학습과 추론에 적합한 특성이 요구되며, 높은 집적도를 가지기 위해 고집적 시냅스 어레이 개발 방안을 제시해야 함. 또한, 뉴런과 시냅스의 동작을 연결하는 인터페이스 회로 설계 방안을 제시해야 하며 뉴런, 인터페이스 회로, 시냅스 어레이가 동일 chip 내 공동집적 되어야 함.
- (소프트웨어 프레임워크) 개발된 하드웨어 플랫폼을 Pytorch 등의 범용 딥러닝 프레임워크와 결합할 수 있어야 함. 신경망 모델 수행의 핵심 연산을 실시간으로 구동하여 전체 모델의 정확도와 시스템의 에너지 효율성 검증할 수 있는 소프트웨어 프레임워크 개발이 필요함.
- (에너지 효율성) 세계 최고 수준의 에너지 효율성을 가진 뉴로모픽 컴퓨팅 시스템 개발을 목표로 ($\geq 1,000$ TOPS/W), 하드웨어의 각 컴포넌트 및 소프트웨어 프레임워크에 대한 에너지 절감 방안을 제시할 수 있어야 함.

2. 연구개발목표

○ 최종 목표 :

- CMOS 공정 호환성을 갖는 뉴런/인터페이스 회로/시냅스의 동일 chip 내 공동집적을 통한 고집적 · 고에너지 효율의 뉴로모픽 하드웨어 구현
- 통합 시스템의 1,000 TOPS/W 급 에너지 효율적 인공 신경망 모델 구현과 이를 검증할 하드웨어 플랫폼 및 소프트웨어 프레임워크 개발

○ 세부 목표 :

(1) 뉴런 소자 개발

- 주파수 > 1 MHz 처리 가능 (단일 펄스 계산이 아닌 초당 firing 횟수 10^6 회 이상)
- 동작 전압 < 1.5 V
- 뉴런 셀의 단일 integration-and-fire 동작 소모 에너지 < 25 fJ/spike

(2) 시냅스 소자 개발

- 실질적인 뉴로모픽 하드웨어를 구현하기 위해서 신경망 모델의 응용 목적에 맞는 적합한 특성을 지닌 시냅스 소자 1종 이상 제시가 필요함
- Retention > 1 year
- Endurance $> 10^8$ pulses

- 동일 펄스 진폭의 arbitrary pulse scheme (potentiation/hold/depression 동작)에서 가중치 update linearity¹⁾ 및 symmetry > 90% 의 특성을 보이는 시냅스 소자
- 시냅스 읽기 동작 에너지 < 1 fJ/operation
- 시냅스 potentiation/depression 각각의 동작 에너지 < 0.1 fJ/operation

¹⁾ update linearity: 뉴로모픽 소자(특히 시냅스 소자)가 가중치(weight)를 업데이트할 때, 입력 신호(예: 전압 또는 전류)와 이에 따른 가중치 변화(ΔW) 간의 관계가 얼마나 선형적으로 유지되는지를 나타내는 특성임.

(3) 시냅스 어레이 개발

- 어레이 크기: 단층 32×32 혹은 수직 적층 $16 \times 16 \times 3$ 층 이상
- 어레이 내 수율 > 90%

(4) 뉴런-시냅스 간 인터페이스 회로 개발

- 뉴런과 다수의 시냅스 사이의 안정적인 신호 전달을 위한 인터페이스 회로의 개발이 요구됨.
- 임피던스 매칭을 고려한, 프리-뉴런과 시냅스, 그리고 포스트-뉴런과 시냅스 연결 인터페이스 회로 각 1종 이상 제시가 필요함
- 각 인터페이스 회로 동작 속도 > 1 MHz, 에너지 소모 < 10 fJ/operation

(5) 뉴런/인터페이스/시냅스 통합 공동 집적 기술 개발

- 각 유닛을 외부에서 연결하는 방식에 비해 높은 집적도 및 데이터 접근성, 낮은 공정 비용을 위해선 동일 chip에서 공동 집적하는 기술 개발이 필요함
- 통합 뉴로모픽 칩 구축 (예: 동일 chip 내 공동 집적기술, 모노리틱 3D 집적기술)

(6) 하드웨어 플랫폼 및 소프트웨어 프레임워크 개발 및 응용 검증

- 뉴런-시냅스 기반 하드웨어 동작 검증을 위해 실제 어플리케이션 동작 구현이 요구됨. 뉴런, 시냅스의 단일 컴포넌트가 아닌, 뉴런, 시냅스가 결합된 형태의 실시간 통합 시스템 구현 및 검증이 필요함.
- 뉴런 ≥ 32 , 시냅스 어레이 $\geq 32 \times 32$ 동시 동작
(혹은 뉴런 ≥ 16 , 시냅스 어레이 $\geq 16 \times 16 \times 3$ 층 동시 동작)
- 개발된 뉴런-시냅스 소자 기반의 뉴로모픽 컴퓨팅 하드웨어를 범용 딥러닝 소프트웨어 프레임워크와 결합이 요구되며, 32×32 (혹은 $16 \times 16 \times 3$ 층) 이상 시냅스 어레이의 온-디바이스 인공신경망 학습 및 추론 구현이 필요함.
- 하드웨어는 뉴런-시냅스와 동일 Chip에 설계하거나, FPGA와 같은 Host 프로세서와 비동기식 동작이 가능한 독립적으로 온전한 시스템 구현이 필요함.
- 인공신경망 모델 기반의 어플리케이션 검증 > 3종
- 인공신경망 모델 구동 시 실효 전력 효율성 $\geq 1,000$ TOPS/W²⁾

²⁾ TOPS/W: tera operations per second per watt. 1 OP(operation)는 하나의 뉴런의 발화 혹은 뉴런 막전위 역학(Membrane potential dynamics)을 의미함. 또는 이에 준하는 실질적 단위를 정의와 함께 제시해야 함.

[참고] 국내외 관련 기술 주요 성능지표

평가 항목 (주요 성능)	단위	세계 최고수준	연구개발 전 국내 수준	목표설정 근거
		성능수준	성능수준	
1. CMOS 기반 SNN 칩 에너지 효율	TOPS/W	15	-	Intel News, 2024
2. 뉴런-시냅스 소자 기반 SNN 칩 에너지 효율	TOPS/W	4.97	-	X Li, et al., <i>Nature Nanotechnology</i> . 2020
3. 시냅스 소자 소모 에너지	fJ	5	15	Li Tu, et al., <i>RCS Advances</i> . 2018 / G. Noh, et al., <i>Advanced Materials</i> . 2022
4. 뉴런 소자 소모 에너지	pJ/spike	0.14	20	J. Han, et al., <i>IEEE Electron Device Letters</i> . 2018 / D. Lee, et al., <i>Advanced Electronic Materials</i> . 2019

3. 성과목표

- 국내·외 특허 출원 및 등록 자율 제시(국가 R&D 연구비 10억당 출원 4건/등록 2건 내외 수준 고려)
 - 출원 및 등록 특허는 기여율이 50% 이상인 특허에 한하여 성과를 인정함
(최종평가위원회 검토사항)
- SCI(E)급 논문 게재 건수 자율 제시(국가 R&D 연구비 10억당 5편 내외 수준 고려)
 - 게재 논문은 기여율 50% 이상인 논문에 한하여 성과를 인정함
(최종평가위원회 검토사항)
- JCR 상위 10% 이내 논문 게재 건수 자율제시
- 본 과제를 통한 석·박사 배출 예상 인원 자율 제시

4. 지원기간/예산/추진체계

- 연구개발기간: '25.4.~'27.12. (총 33개월 내외)
- 정부지원연구개발비: 총 2,651백만원 내외 ('25년 723백만원 내외)

총 연구기간('25.4.~'27.12. / 33개월)		
1차년도	2차년도	3차년도
'25.4.~'25.12.	'26.1.~'26.12.	'27.1.~'27.12.
723백만원	964백만원	964백만원

※ 연차별 연구비 규모 및 연구기간은 정부예산 사정에 따라 변경 가능

- 선정 과제 수: 1개 과제 (단위과제 - 공동 또는 단독)
- 과제형태: (일반)연구개발과제
- 주관연구개발기관: 대학/출연(연)/기업부설연구소 등
- 기술료 징수여부: 징수

5. 특기사항

- 실제 제출하는 과제명은 연구자의 아이디어가 포함될 수 있는 제목으로 연구계획서 제출
- 복수의 공동연구과제 또는 위탁과제를 통한 역할 분담 가능
- 과제 목표에 부합하도록 소자, 회로구현을 위한 연차별 통합추진체계를 구체적으로 제시
- 회로 및 시뮬레이션 전문가를 포함하여 회로 구현 및 실용화 가능성 점검 필요
- 제안서 상에 제시된 소자에 대한 선행개발 이력 및 결과 제시 필요
- 다음 사항을 연구계획서에 제시 필요
 - 제안된 소자 및 시스템 아키텍처에 대한 구성도(단위 블록별 H/W, S/W 파트 명기)
 - 개발한 최종 결과물에 대한 실험적 검증 내용(방향성 및 범위 제시)
 - 전력효율(Power efficiency)에 대한 지표 및 목표산출과정 제시
 - 최종 결과물을 기반으로 실제 응용 가능한 시스템/어플리케이션 등 활용 방안 1종 이상