

- 우수성과**

모트전이 멤리스터의 열적 동역학 활용 고차원 인공 뉴런 및 시공간적 컴퓨팅 시스템 개발
- 연구자**

김경민 (한국과학기술원 신소재공학과 교수)

발상의 전환으로 차세대 초전력 컴퓨팅 해결책 찾는다

The real importance of memristors is that it is in fact what a brain is built from.

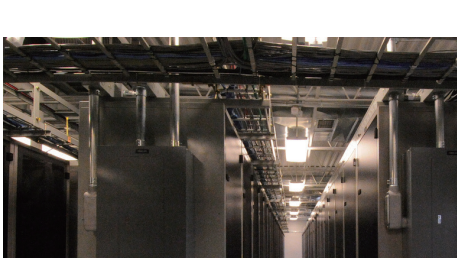
(멤리스터의 진정한 중요성은 사실 뇌가 그것으로 이루어져 있다는 데 있다.)

- 레온 추아(전기공학자, 멤리스터 개념의 창시자)

인공지능은 에너지 효율적일까?

2016년 3월 구글 딥마인드사의 바둑 인공지능 프로그램인 알파고와 프로기사인 이세돌 9단의 대국은 전 세계에 충격을 가져왔다. 아직 바둑은 인간의 영역이라는 많은 이의 예상을 깨고 알파고가 승리를 거두었던 것이다. 이 대국은 인공지능 개발의 역사에 기념비적인 사건이었다. 그로부터 불과 10년이 지나는 사이에 인공지능은 널리 대중화되어 다양한 분야에서 쓰이고 있다.

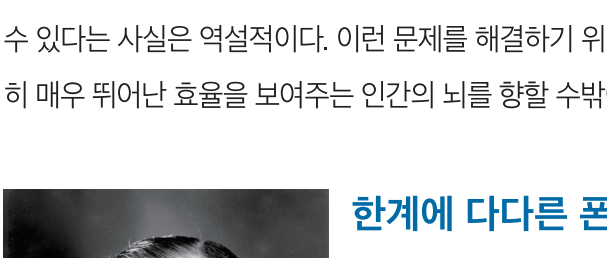
특히 이전까지는 인간의 고유한 영역으로 여겼던 창작 분야에서도 인간 못지않은 활약을 할 수 있다는 사실이 드러나면서 인간의 미래를 걱정하는 목소리도 높아지고 있다. 각국이 인공지능 기술 개발에 사활을 걸고 있는 이유다.



▲ 알파고의 서버, 위키미디어

하지만 아직 인공지능이 인간의 뇌를 따라잡지 못하는 영역이 하나 있다. 바로 뇌의 뛰어난 효율이다. 1000억 개의 뉴런과 100조 개의 시냅스로 이루어진 인간의 뇌가 소비하는 에너지는 고작 20와트에 불과하다. 이세돌과 대결했던 당시 알파고의 소비전력은 약 170kW로 알려져 있다. 자료에 따라서는 그 이상으로 보기도 한다. 170kW라고 해도 사람의 뇌와 비교하면 무려 8500배의 에너지를 소모하는 셈이다. 8500분의 1밖에 안 되는 에너지를 소모하고도 알파고와 대등하게 경기를 치렀다면, 오히려 이세돌의 승리로 볼 수도 있지 않을까?

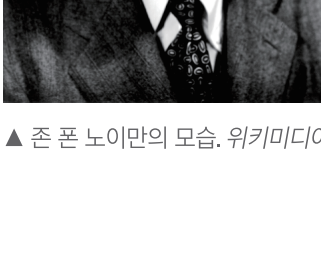
인공지능 기술의 급속한 확산은 에너지 소모량의 증가라는 새로운 문제를 불러일으키고 있다. 2025년 6월 글로벌 컨설팅 기업 액센츄어가 발표한 보고서 '지속가능한 시를 위한 에너지 공급(Powering sustainable AI)'에 따르면, 시를 위한 데이터센터의 전력 사용량은 2030년에 연간 612TWh에 이를 전망이다. 이는 캐나다 전체의 연간 전력 소모량과 맞먹는다.



▲ 데이터센터의 내부 모습, 위키미디어

데이터센터에서 나오는 열을 식히기 위해서는 냉각수가 필요한데, 보고서에서는 30억 입방미터의 물이 필요하게 될 것으로 내다보고 있다. 노르웨이나 스웨덴에서 연간 소비하는 담수의 양과 비슷한 수준이다. 전력 소모에 따른 탄소 배출도 크게 늘어나 시로 인한 탄소 배출량이 전체의 3.4%를 차지한다. 10년 전과 비교하면 11배 늘어난 수치다.

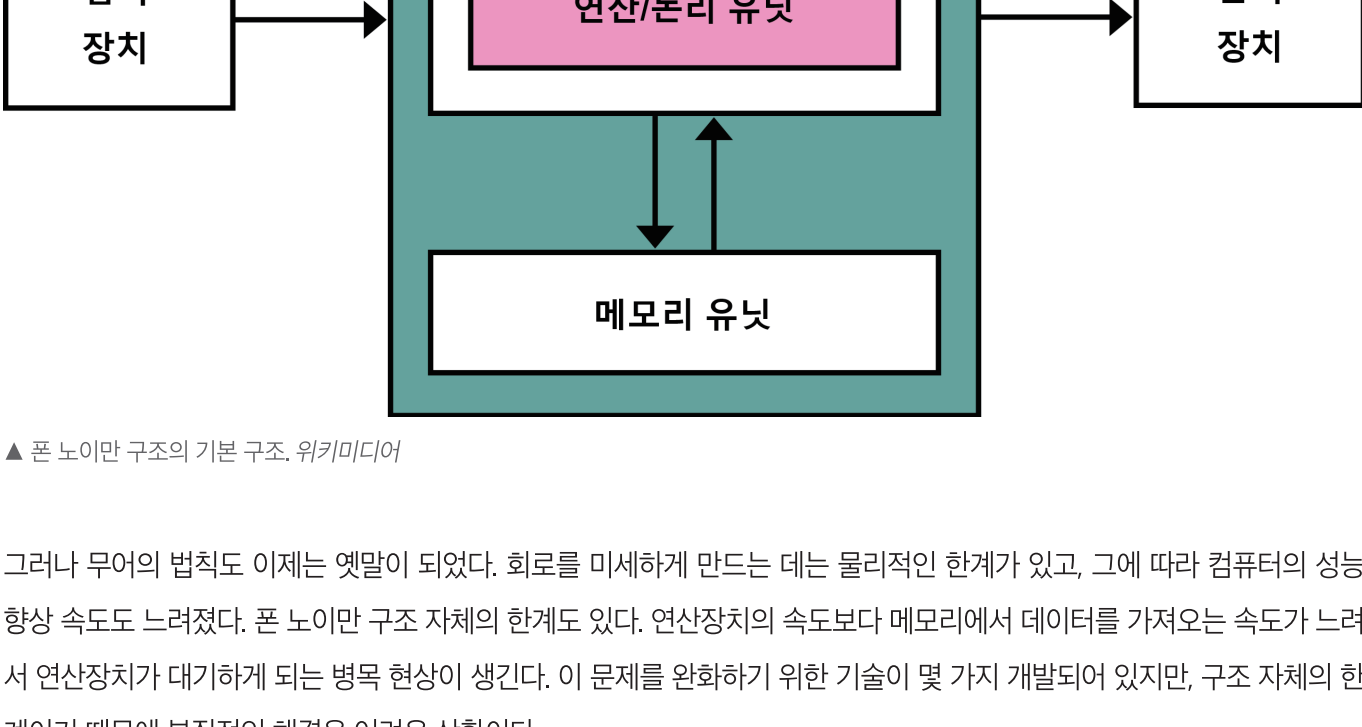
작업의 효율을 높여준다는 이유로 여러 기업에서 채택하고 있는 AI가 에너지 측면에서는 효율적이지 않을뿐더러 기후 위기를 촉진할 수 있다는 사실은 역설적이다. 이런 문제를 해결하기 위해서는 에너지 효율적인 새로운 컴퓨팅 방식이 필요하고, 시선은 자연히 매우 뛰어난 효율을 보여주는 인간의 뇌를 향할 수밖에 없다.



▲ 존 폰 노이만의 모습, 위키미디어

한계에 다다른 폰 노이만 구조

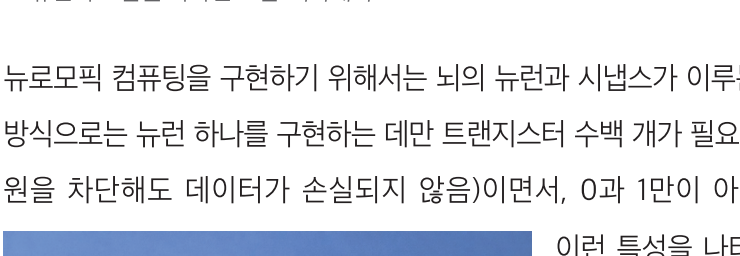
오늘날 우리가 사용하는 컴퓨터는 모두 20세기 중반 존 폰 노이만이 제시한 폰 노이만 구조를 따른다. 폰 노이만 구조에서는 연산장치(CPU)와 메모리, 제어장치가 분리되어 있고, 이들이 서로 신호를 주고받으면서 연산한다. 지금까지 컴퓨팅 성능의 발전은 이 구조 안에서 더 작고 빠른 반도체를 만드는 방식으로 이루어졌다. 같은 면적 안에 더 많은 트랜지스터를 넣음으로써 성능을 높이고 가격을 낮추어 왔다. 이런 추세는 집적 회로의 트랜지스터 수가 약 2년마다 두 배로 증가한다는 '무어의 법칙'을 낳기도 했다.



▲ 폰 노이만 구조의 기본 구조, 위키미디어

그러나 무어의 법칙도 이제는 옛말이 되었다. 회로를 미세하게 만드는 데는 물리적인 한계가 있고, 그에 따라 컴퓨터의 성능 향상 속도도 느려졌다. 폰 노이만 구조 자체의 한계도 있다. 연산장치의 속도보다 메모리에서 데이터를 가져오는 속도가 느려서 연산장치가 대기하게 되는 병목 현상이 생긴다. 이 문제를 완화하기 위한 기술이 몇 가지 개발되어 있지만, 구조 자체의 한계이기 때문에 본질적인 해결은 어려운 상황이다.

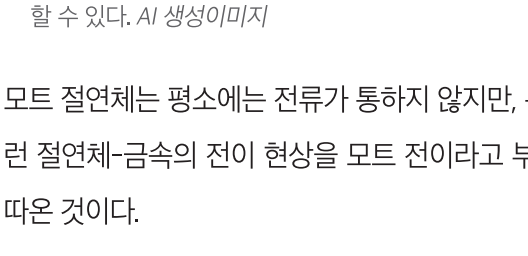
폰 노이만 구조의 한계를 해결하기 위한 시도 중 하나가 바로 인간과 같은 동물의 뇌 구조를 모방한 방식인 뉴로모픽 컴퓨팅이다. 인간 뇌의 뉴런은 사방에 수많은 가지를 뻗어내 서로 연결 구조를 만든다. 뉴런의 가지와 가지가 이어지는 부위가 바로 시냅스다. 새로운 것을 학습하거나 반복적인 행동을 할수록 그와 관련된 시냅스가 강화되거나 새로운 시냅스가 만들어진다.



▲ 뉴런의 모습을 나타낸 그림, 픽사베이

뉴런 하나에 있는 시냅스의 수는 수천 개에 달한다. 즉, 뉴런 하나하나가 다른 수천 개의 뉴런과 연결되어 대단히 복잡한 연결망을 구성한다. 뇌는 폰 노이만 구조와 달리 연산과 저장 기능이 나뉘어 있지 않고 수많은 뉴런이 다발적으로 활성화되거나 억제되면서 엄청난 양의 정보를 동시에 처리한다. 자연이 만든 최고의 병렬 컴퓨팅 시스템인 셈이다. 우리는 이 연결망을 이용해 경험을 기억하거나 문제를 풀어내는 것은 물론 고차원적인 사고까지 해낼 수 있다.

뉴로모픽 컴퓨팅을 구현하기 위해서는 뇌의 뉴런과 시냅스가 이루는 연결을 모방하는 하드웨어가 필요하다. 기존의 디지털 방식으로는 뉴런 하나를 구현하는 데만 트랜지스터 수백 개가 필요하기 때문이다. 시냅스를 모방하기 위해서는 비휘발성(전원을 차단해도 데이터가 손실되지 않음)이면서, 0과 1만이 아닌 여러 단계의 연결 강도를 나타낼 수 있어야 한다.



▲ 멤리스터는 뇌를 모사한 뉴로모픽 컴퓨팅을 가능하게 할 수 있다. AI 생성이미지

이런 특성을 나타내기 위한 소자로 주목받고 있는 것이 바로 '멤리스터(memristor)'다. 멤리스터는 메모리(memory)와 저항(resistor)의 합성어로, 1971년 컴퓨터과학자 레온 추아(Leon Chua)가 제시한 개념이다. 멤리스터는 과거의 전류 흐름에 따라 저항이 달라지며 전력 공급이 끊어져도 가장 최근의 저항을 기억한다. 덕분에 인공 뉴런과 시냅스의 설계가 단순해질 수 있었다.

멤리스터는 한동안 이론에 머물렀지만, 2008년 HP의 연구진이 멤리스터의 특성을 보이는 소자를 개발해 <네이처(Nature)>에 발표하면서 멤리스터 개발의 가능성을 열었다. 지금까지 멤리스터에 적합한 재료로 여러 가지 후보가 등장했는데, 그중에서 모트(Mott) 절연체는 빠른 스위칭과 낮은 전력 소모라는 장점으로 큰 관심을 받고 있다.

모트 절연체는 평소에는 전류가 통하지 않지만, 온도가 올라가면 순간적으로 전류가 통하는 금속성 물질로 변하는 재료다. 이런 절연체-금속의 전이 현상을 모트 전이라고 부른다. 이 현상을 예측한 영국의 물리학자 네빌 모트(Nevil Mott)의 이름에서 따온 것이다.

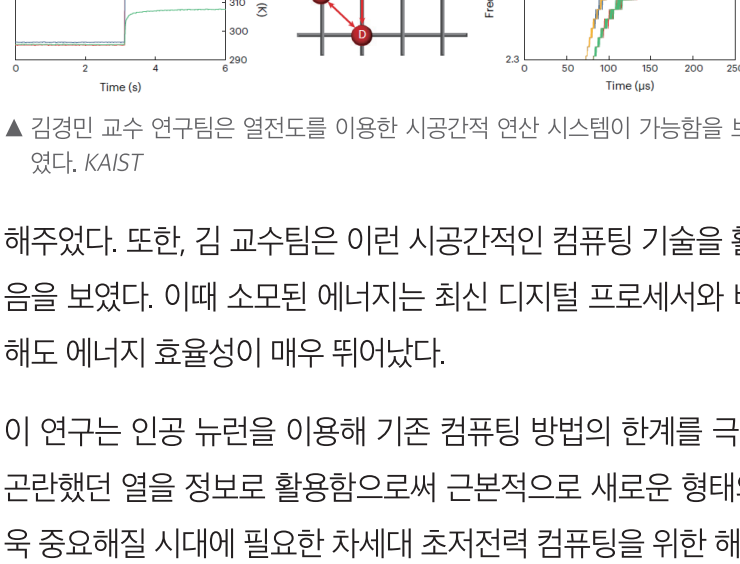
나리오븀 산화물과 같은 모트 절연체를 사용해 금속-절연체-금속 구조를 만들면 모트 전이를 이용해 전류의 흐름을 제어할 수 있다. 과거 전류의 흐름에 따라 저항이 변화하며 정보를 저장하는 멤리스터의 핵심 특성을 보이므로 모트 멤리스터는 뇌의 뉴런을 모사해 신호를 처리할 수 있다.

열 동역학과 멤리스터 이용한 신개념 컴퓨팅

모트 멤리스터를 이용한 뉴로모픽 컴퓨팅 연구가 활발한 가운데 기존 시스템의 한계를 뛰어넘기 위한 새로운 시도 역시 이루어지고 있다. 그중 하나가 한국과학기술원(KAIST) 신소재공학과 김경민 교수팀이 개발한 열 컴퓨팅 기술이다.

전자기기를 사용할 때는 열이 발생하기 마련이고 반도체 역시 마찬가지다. 노트북이나 스마트폰을 오래 사용하면 따뜻해지는 경험을 누구나 해본 적이 있을 것이다. 이렇게 발생하는 열은 에너지 소모를 늘리고 성능 저하를 일으키기 때문에 칩을 설계할 때는 열 발생을 최소화하는 게 관건이다. 앞서 언급했듯이, 인공지능의 급격한 대중화로 인한 데이터 센터의 증가는 발열의 증가라는 새로운 문제도 일으키고 있다.

하지만 김 교수팀은 오히려 열을 적극적으로 활용하는 아이디어를 떠올렸다. 기존의 모트 멤리스터 인공뉴런에서는 내부에서 발생된 열이 쉽게 빠져나가며, 이 경우 열 동역학이 단 하나의 상태 변수로 나타낼 수 있는 수준에 머무른다. 따라서 소자가 표현할 수 있는 동작이 다양하거나 복잡해질 수 없고, 단일 소자가 수행할 수 있는 뉴런의 '기능 밀도(functional density)'에도 한계가 생긴다. 더 높은 기능 밀도를 달성하기 위해서는 다른 소자에 열 정보를 전달할 수 있어야 하는데, 기존의 방식에서는 열을 전기 신호로 바꾸는 과정을 거쳐야 했으며, 그 과정에서 에너지 소모가 크고 일부 정보가 사라질 수도 있어 매우 비효율적이다.



▲ 김경민 교수 연구팀은 열전도를 이용한 시공간적 연산 시스템이 가능함을 보였다. KAIST

김 교수팀은 기능 밀도 문제와 비효율성을 극복하기 위해 유연한 폴리이미드(플라스틱의 일종) 기판에 모트 멤리스터 반도체 소자를 통합했다. 소자 내부의 열 동역학을 활용할 뿐만 아니라 추가적으로 소자와 기판 사이의 열 교환을 적극적으로 사용하기 위해서다.

폴리이미드 기판에 저장된 열은 일정한 시간 동안 유지되며, 이곳 소자로도 퍼져 나간다. 이때 열이 유지되는 시간은 시간적 정보가 되며, 주위로 퍼져 나가는 열은 공간적 정보 역할을 한다. 즉, 열을 전기로 변환하는 과정 없이 열을 통해 직접 다른 소자와 신호를 주고받는 방식을 통해 시공간적인 열 동역학을 컴퓨팅에 사용하는 것이다.

소자와 열을 활용하는 방법은 단순한 소자 구조로도 18가지 유형의 뉴런 작용을 구현할 수 있게 해주었다. 또한, 김 교수팀은 이런 시공간적인 컴퓨팅 기술을 활용해 경로 찾기나 같은 최적화 문제를 해결할 수 있음을 보였다. 이때 소모된 에너지는 최신 디지털 프로세서와 비교해 약 1백만 배 낮은 수준으로, 기존 연구와 비교해도 에너지 효율성이 매우 뛰어났다.

이 연구는 인공 뉴런을 이용해 기존 컴퓨팅 방법의 한계를 극복하고 에너지 소모를 획기적으로 줄이는 한편 처리 곤란했던 열을 정보로 활용함으로써 근본적으로 새로운 형태의 컴퓨팅 방식을 제시했다. 딥러닝과 인공지능이 더욱 중요해질 시대에 필요한 차세대 초저전력 컴퓨팅을 위한 해결책이 될 수 있는 것이다.

《특》 미니 인터뷰

Q 이번 연구성과가 갖는 가장 큰 의미는 무엇인가요?

A 기존의 반도체 기술은 지난 50여 년간 전기 신호를 활용하는 메모리와 CPU 등으로 이루어진 컴퓨터에 초점이 맞춰져 있었습니다. 최근에는 인공지능이 발달하며, GPU를 중심으로 하는 컴퓨팅이 중요해지는 등 기존 컴퓨팅 방식에 대한 도전이 시작되고 있습니다. 앞으로 복잡한 문제 풀이를 위한 새로운 컴퓨팅 기술이 계속 필요해질 텐데, 이번 연구성과가 이런 배경에서 완전히 새로운 해답을 제공할 수 있는 기술입니다.

최근에는 양자컴퓨팅이 큰 관심을 받고 있지만, 양자컴퓨팅은 비현실적으로 복잡하고, 경량화가 어려울 것으로 예상됩니다. 저위 기술은 기존 반도체 기술을 기반으로 하고 있어 양산과 경량화에 큰 장점이 있습니다.

Q 이 주제를 연구하게 되신 계기가 궁금합니다.

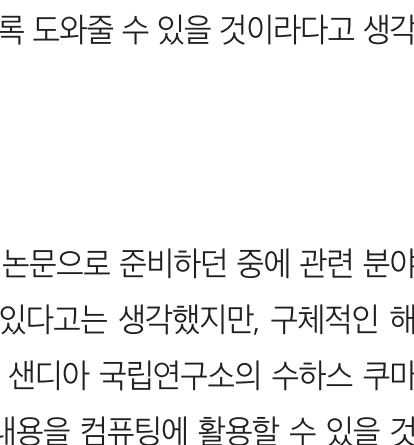
A 전적으로 이 연구를 주도한 김광민 박사의 호기심 때문이었습니다. 제자인 김광민 박사가 석사과정 때 새로운 연구 주제를 탐색하던 중 당시에 조금씩 관심을 받고 있던 모트 전이 멤리스터를 만들어 보았는데, 소자의 특성이 상당히 잘 나왔습니다. 당시에는 소자가 갖는 고차원 동역학에 관한 이해가 없던 상황이었기에 이 소자의 동작을 이해하는 연구부터 시작했지요. 그렇게 5년 정도를 연구하다보니 이 소자가 가진 특성을 깊이 이해할 수 있었고, 이를 바탕으로 이번 연구 성과를 내놓게 되었습니다.

Q 본 연구에 이어서 어떤 연구를 하고 계신가요?

A 앞선 연구에서는 열을 이용하는 컴퓨팅 기술에 관한 아주 기초적인 실험 결과를 보이며 그 가능성을 제시했었습니다. 이제는 저희가 개발한 기술을 활용해서 실제로 컴퓨팅을 수행하는 연구를 하고 있습니다. 이를 위해서는 더 큰 규모의 반도체 소자를 제작해야 하고, 컴퓨팅을 수행하기 위한 알고리즘과 하드웨어를 개발해야 합니다. 궁극적으로 우리 기술은 양자컴퓨팅과 경쟁하는 것을 목표로 하고 있습니다. 열과 전기의 상호 작용을 효과적으로 활용해서 양자컴퓨팅을 능가하는 새로운 컴퓨팅 기술을 개발하고자 합니다.

Q 연구하시는 새로운 컴퓨팅 방식이 상용화된다면, 우리 사회에 어떤 변화를 가져오게 될까요?

A 우리는 매 순간 어떤 결정을 내리며 살아가고 있고, 우리의 삶은 결국 선택의 문제이지 않을까 합니다라고 생각합니다. 우리는 최선의 선택을 위해서 노력하지만, 그 선택이 최선이었을까 하는 점에 대해 항상 고민할 수 밖에 없습니다. 앞으로의 기술은 이러한 선택을 도와주고 우리의 삶에서의 고민을 줄여주는 데 도움을 줄 것입니다. 여기서 말하는 기술은 주어진 환경에서의 다양한 변수들을 고려해서 최선의 선택을 도와줄 수 있는 기술일 것이며 이를 말한 겁니다. 즉, 이는 최적화 문제를 빠르고 에너지 효율적으로 풀아줄 수 있는 컴퓨팅 기술일 것입니다. 이번에 개발한 열 컴퓨팅 기술은 휴대 폰전화, 스마트 워치 등에 탑재될 수 있습니다. 우리 바로 옆에서 복잡한 문제에 대한 답을 주는 역할을 해줄 수 있고, 이는 역할을 하는 거죠. 그렇게 되면 우리들의 최선의 결정을 도와주며내면, 우리 리는 데 도움을 삶을 보다 후회 없는 삶이 되도록 도와줄 수 있을 것이라고 생각합니다. 줄 거라고 생각합니다.



▲ 유연한 폴리이미드 기판을 이용해 웨어러블 반도체 소자로 만들 수 있다. KAIST

Q 이번 연구를 진행하는 과정에서 가장 결정적인 순간을 꼽는다면?

A 저희가 모트 전이 소자에서 고차원 동역학을 발견하고, 2022년에 이를 논문으로 준비하던 중에 관련 분야의 국제 학회에 참석했습니다. 당시 저희 연구가 새로운 기술이 될 수 있다고는 생각했지만, 구체적인 해답을 찾지는 못한 상태였습니다. 그 학회에서 김광민 박사와 당시 미국 샌디아 국립연구소의 수하스 쿠마(Suhas Kumar) 박사가 해당 연구 내용으로 토의했고, 쿠마 박사가 이 내용을 컴퓨팅에 활용할 수 있을 것 같다고 제안해 공동 연구를 하게 되었습니다.

쿠마 박사는 제 이전 직장의 동료였고, 제가 학생들에게 만나게 되면 꼭 이야기를 나누어 보라고 권하곤 합니다. 그런데 마침 제자가 실제로 쿠마 박사와 이야기를 나누는 뒤 공동 연구를 하게 되었지요. 샌디아 연구소는 열 발생을 관찰하는 실험과 컴퓨팅 성능의 검증을 맡아주었고, 이 협력은 결과적으로 저희 기술이 세계적으로 인정받는 데 큰 기여를 했습니다.



▲ 새로운 컴퓨팅 기술을 개발해 낸 김경민 교수 연구팀이 한데 모였다. KAIST

글 고선아(길물빛콘텐츠연구소 대표) 외
기획 한국과학기술기획평가원(김주희 외)

본 원고는 과학기술정보통신부가 선정한 2025 국가연구개발 우수성과 100선 최우수성과에 대한 이해를 돕고자 작성하였습니다.