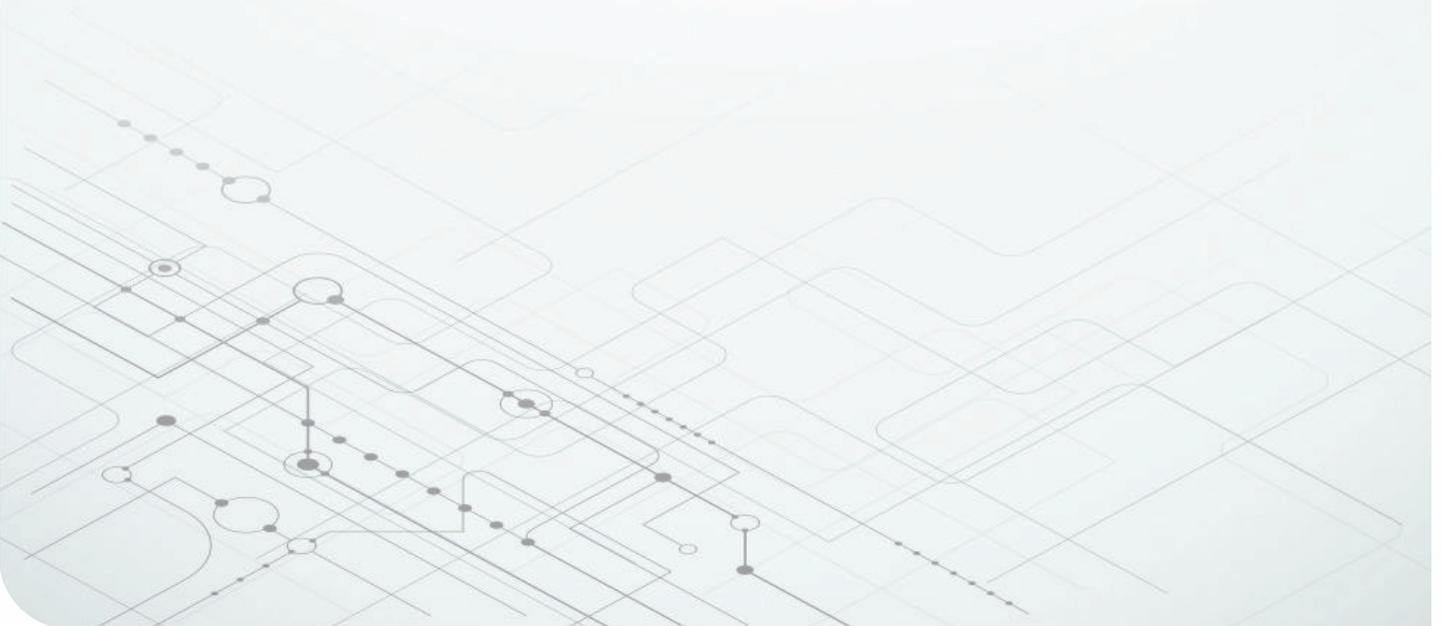


임베디드 프로세서로 더욱 지능적인 에지



Alec May
Systems Manager
Jacinto™ High-Performance Computing Processors



한눈에 보기



소개

이 백서에서는 AI(인공 지능)의 목적과 장점 그리고 임베디드 프로세서와 소프트웨어의 발전으로 다양한 애플리케이션에서 AI를 어떻게 쉽게 구현할 수 있는지 살펴봅니다.

1



에지 AI의 장점

에지 AI의 장점과 전자 장치에서 로컬 추론을 지원하는 기능에 대해 알아보십시오.

2



AI가 에지로 전환하는 방법

임베디드 하드웨어 및 소프트웨어 엔지니어가 직면한 과제와 TI가 이러한 문제를 어떻게 해결하고 있는지 알아보세요.

3



에지 AI의 확장성

TI가 확장성과 재사용성이라는 과제를 해결하기 위해 제공하는 하드웨어 및 소프트웨어를 살펴보세요.

4

소개

널리 사용되는 클라우드 기반 AI 솔루션의 접근성과 사용자 친화성은 거의 모든 사람이 AI를 위해 설계된 모델과 도구에 더 쉽게 참여할 수 있게 해주었습니다.

그러나 모든 AI 혁신이 클라우드에서 일어나는 것은 아닙니다. 임베디드 프로세서 설계의 기술의 발전으로 AI 기능은 이제 노트북 및 휴대폰과 같은 소비자 제품뿐만 아니라, 현관 카메라, 차량용 시스템의 비전 프로세싱, 에너지 인프라 및 산업용 시스템용 모터와 같은 배터리로 구동되는 애플리케이션에서도 사용되고 있습니다.

에지 AI - 데이터 소스 근처에서 로컬로 AI 모델을 실행할 수 있는 기능으로 이러한 제품의 응답성, 효율성, 신뢰성 및 보안을 향상하고 있습니다. 클라우드-에지 변환을 가능하게 하는 임베디드 프로세서는 DSP(디지털 신호 프로세

싱)를 위한 특수 코어와 같은 구성 요소를 통합하고 AI를 에지에 도입하는 데 필요한 시간과 전문 지식을 최소화하는 사용하기 쉬운 GUI 기반 툴의 지원을 받습니다.

이 백서에서는 에지 AI의 발전과 장점에 대해 알아보고 이를 지원하는 하드웨어 및 소프트웨어의 발전에 대해 설명하겠습니다.

AI란 무엇입니까?

오늘날 대부분의 사람들이 AI를 생각하면 텍스트 및 이미지 생성기를 상상하는 경우가 많을 것입니다. 그러나 가장 단순한 알고리즘조차도 기술적으로 문자 그대로의 의미에서 AI의 예시일 뿐입니다.

그림 1에서 볼 수 있듯이 AI의 광범위성과 여러 사용 사례는 머신 러닝과 딥 러닝을 포함한 여러 하위 도메인으로 이어졌습니다.

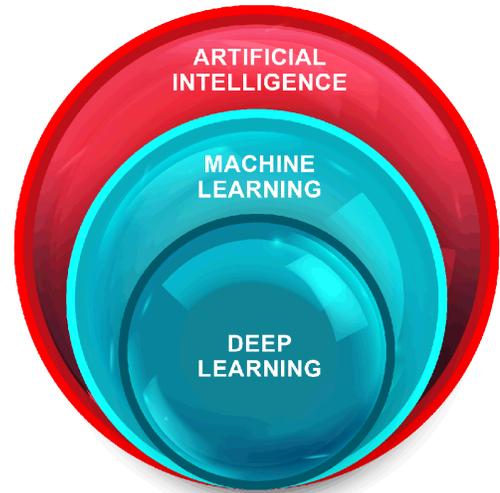


그림 1. 다양한 AI 하위 도메인 간의 관계를 설명합니다.

임베디드 애플리케이션에 사용되는 대부분의 AI는 머신 러닝입니다. 이 하위 도메인은 기계와 알고리즘이 데이터로부터 문제를 해결하는 방법을 "학습"합니다. 일반적인 패턴의 이미지 데이터를 분석하여 보행자와 장애물을 인식하는 차량을 예로 들 수 있습니다. 머신 러닝 모델은 정답 정보(즉, 검증된 정확한 데이터)로 레이블링된 훈련 데이터를 통해 학습하며, 이를 통해 학습할 패턴을 더 잘 식별할 수 있습니다. 이 교육 프로세스를 통해 머신 러닝 모델이 데이터의 패턴을 식별하여 미래의 추론을 수행하는 데 사용할 수 있습니다.

머신 러닝 분야에서 딥 러닝은 매우 복잡한 문제를 정확하게 해결할 수 있는 능력을 갖추고 있어 가장 인기 있는 구현 중 하나가 되었습니다. 하지만 그렇게 하려면 많은 컴퓨팅 리소스가 필요합니다. 딥 러닝은 다층 신경망을 사용하는데, 이는 인간의 두뇌의 뉴런에서 영감을 받은 데이터 모델입니다. 신경망을 통해 개발자는 패턴이 너무 복잡하거나 사용자가 정의 규칙을 작성하기 어려운 문제를 해결할 수 있습니다.

에지 AI의 장점

AI 및 하위 도메인은 일반적으로 클라우드 또는 로컬 서버에서 처리를 수행할 수 있습니다. 클라우드 기반 AI는 역사적으로 영향력 있는 AI 기능을 수행하는 데 필요한 컴퓨팅 성능을 대형 서버 외부에서는 쉽게 달성할 수 없었기 때문에 더욱 일반적인 기능이었습니다. 그러나 임베디드 프로세서의 컴퓨팅 성능과 전력 효율이 높아짐에 따라 에지 AI의 인기가 높아졌습니다.

그림 2에서는 에지 AI 및 클라우드 AI가 데이터를 수신 및 처리하고 클라우드 기반 리소스와 상호 작용하는 방법과 관련하여 어떻게 다른지 볼 수 있습니다.

에지 AI는 개발하는 동안 모델을 교육할 때 에 클라우드 또는 데스크톱 리소스를 사용하는 경우가 많습니다. 모델을 임베디드 장치에 배포한 후에는 독립적으로 로컬에서 새 데이터에 대한 모델 추론 및 결정을 내릴 수 있게 됩니다.

최근까지만 해도 AI의 가장 의미 있는 사례는 일반 소비자 가전제품이 제공할 수 있는 것 이상의 처리 능력을 필요로 했습니다. 이는 머신 러닝 모델이 클라우드 기반 리소스에 대해 훈련되고 구현되는 경우가 많다는 사실을 의미합니다. 클라우드 기반 구현은 하드웨어 투자를 최소화함으로써 편의성을 제공하지만 AI 채택도 제한했습니다. 클라우드 기반 AI는 클라우드 액세스 없이는 어떤 애플리케이션에서도 작동하지 않습니다. 즉, 네트워크 연결이 필요합니다. 또한 에지 AI는 클라우드 기반 AI에 비해 보안, 안전성 및 응답성을 향상시킬 수 있습니다.

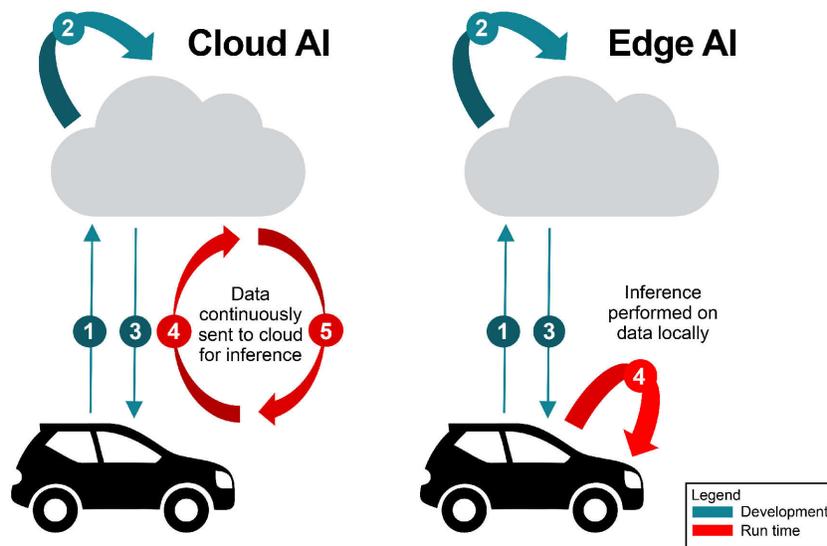


그림 2. 클라우드 기반 AI와 에지 AI의 비교

반도체의 발전과 AI 툴 체인의 개선으로 임베디드 프로세서 및 마이크로컨트롤러(MCU)에 직접 AI 솔루션을 구현하는 것이 AI를 에지로 구현합니다. AI를 에지에 도입한다는 것은 연산 및 AI 추론을 데이터를 제공하는 센서에 더 가깝게 수행한다는 것을 의미하며, 전자 제품에서 수집되는 센서 데이터의 양이 계속 증가하고 있음을 고려하면 이는 중요합니다. 클라우드에서 대량의 데이터를 전송하는 것은 비용이 많이 들고 복잡하며 단일 장애 지점이 될 수 있기 때문에 데이터 볼륨이 증가하면 클라우드 전용 리소스의 실용성이 떨어집니다.

네트워크 에지에서 AI 모델을 실행하면 일반적으로 충돌 감지를 위한 차량의 카메라 센서 등 센서 데이터를 기반으로 한 추론 및 의사 결정의 지연 시간이 줄어듭니다. 에지 AI 기능을 사용하면 차량은 클라우드에서의 추론을 기다리지 않고도 실시간으로 자극에 응답하여 더 빠르게 추론할 수 있습니다. 이를 더 빠르게 만들 수 있습니다. 이러한 로컬 추론은 피지컬 AI(그림 3에 표시됨)를 통해 실행되며, 이는 공장에서 상자를 이동하는 로봇 또는 자동으로 브레이크를 밟는 차량 등 실제 동작을 감지 및 해석하고 대응하는 시스템을 지칭하는 용어입니다.



그림 3. 휴머노이드 로봇의 에지 AI와 피지컬 AI의 간소화된 비교

이러한 시스템은 에지 AI를 사용하여 첨단 인식 기능을 기계적 작동과 결합하여 기계가 인간과 함께 안전하고 협력적으로 작동할 수 있도록 합니다.

에지 AI는 네트워크 연결에 대한 의존도 감소 등 클라우드 기반 AI에 비해 몇 가지 다른 장점이 있습니다. 에지 AI는 클라우드에 액세스할 수 없는 애플리케이션에서 작동하며 네트워크 중단으로 인한 잠재적 다운타임을 최소화합니다. 또한 클라우드 기반 AI에는 네트워크 연결이 필요하기 때문에 액세스를 위한 서비스 요금이 계속해서 발생할 수

있으므로, 소비자 제품을 설계할 때 까다로운 비즈니스 모델이 될 수 있습니다.

AI가 에지로 전환하는 방법

임베디드 프로세서의 처리 및 전력 소비에 대한 제약, 높은 수준의 사내 프로그래밍 전문 지식 및 리소스는 에지 AI의 광범위한 채택에 제한이 되었습니다. AI 계산의 성능 요구 사항을 충족할 수 있는 임베디드 장치는 너무 크고, 너무 많은 전력을 소비하며, 너무 많은 열을 발생시킨 경우가 많습니다.

특수 하드웨어 솔루션이 등장하여 에지 AI를 구현하는 데 필요한 컴퓨팅 작업에 더 나은 가속을 지원하지만 몇 가지 장단점으로 인해 에지 애플리케이션에서의 더 광범위한 채택에 제한이 되었습니다. GPU(그래픽 처리 장치), FPGA(필드 프로그래머블 게이트 어레이) 및 ASIC(애플리케이션별 통합 회로)와 같은 전용 하드웨어 솔루션은 탁월한 성능을 달성했습니다. 그러나 이러한 솔루션은 일반적으로 높은 소비 전력(특히 GPU 및 FPGA의 경우) 또는 제한된 유연성(ASIC의 경우)에 의해 제한됩니다.

통합 NPU(신경망 처리 장치)가 임베디드 시스템에 직접 AI 기능을 구현하기 위한 솔루션으로 떠올랐습니다. 범용 프로세서와 달리, NPU는 현대 신경망의 중추를 형성하는 행렬 곱셈, 합성곱 연산 및 활성화 함수를 실행하도록 특별히 제작되었습니다. NPU는 메인 CPU에서 계산 집약적 작업을 오프로드함으로써 에지 배포에 대한 두 가지 필수 요구 사항인 소비 전력을 줄이면서 추론 속도를 획기적으로 개선합니다.

이러한 구성 요소의 영향을 이해하기 위해 에지 AI 시장의 다양한 부문을 위해 설계된 텍사스 인스트루먼트의 두 가지 NPU, 즉 고성능 애플리케이션용 C7™ NPU와 저전력, 확장형 장치용 TinyEngine™ NPU를 살펴보겠습니다.

C7 NPU

C7 NPU는 TDA54-Q1 및 TDA4VE-Q1 SoC(시스템 온 칩)에 통합된 전력 효율적인 고성능 AI 가속기입니다. TI 마이크로프로세서 및 SoC에는 비전 애플리케이션의 컴퓨팅 요구 사항을 해결하기 위해 여러 세대의 C7 NPU 버전이 포함되었습니다. 이 NPU는 TI의 오랜 DSP 역사에서 비롯되어 비용이나 전력 부담 없이 AI 솔루션에 효율적으로 구동할 수 있게 해줍니다.

또한 C7 NPU를 사용하면 TI 프로세서가 여러 동시 AI 워크로드를 처리할 수 있습니다. 이는 카메라, 레이더, LIDAR 및 기타 센서의 데이터를 동시에 처리해야 하는 시스템에 중요한 기능입니다.

TDA54-Q1은 C7 NPU를 사용하여 첨단 운전자 보조 시스템, 인포테인먼트 및 로봇에서 에지 AI를 지원합니다.

TinyEngine NPU

TinyEngine NPU(그림 4의 간소화된 블록 다이어그램에 표시됨)는 딥 러닝 추론 연산을 최적화하여 배터리 구동 제품을 포함한 리소스 제약적 장치에서 AI 워크로드를 처리할 때 지연 시간과 소비 전력을 줄이는 MCU용 전용 하드웨어 가속기입니다.

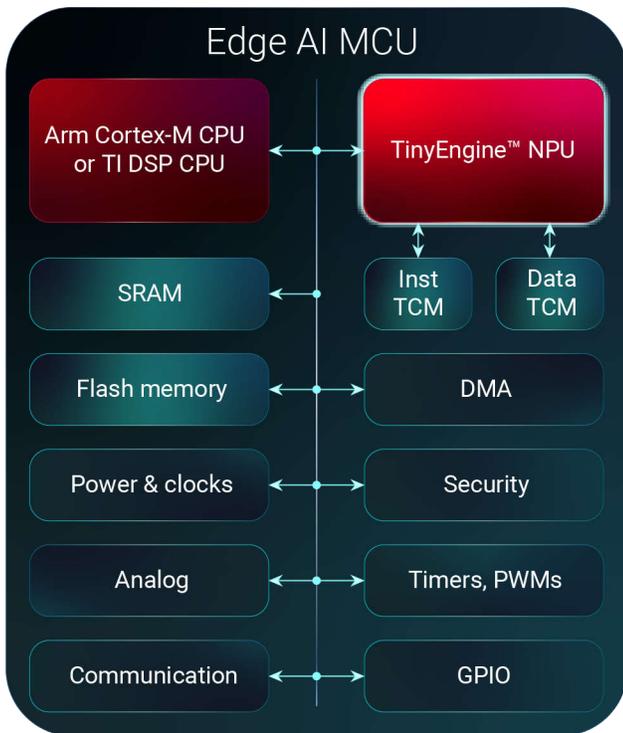


그림 4. TinyEngine NPU가 통합된 TI 에지 AI MCU의 간략한 블록 다이어그램

이 NPU는 애플리케이션 코드를 실행하는 기본 CPU와 병렬로 기계 학습 알고리즘을 실행합니다. TinyEngine NPU가 통합된 MCU는 가속기가 없는 유사 MCU 대비 추론당 지연 시간을 최대 90배 단축하고 에너지 효율을 120배 높여 AI 모델을 실행할 수 있습니다.

TMS320F28P550SJ C2000™ MCU는 모터 베어링 및 태양광 아크 결합 감지를 위해 TinyEngine NPU를 사용하고 메인 CPU가 실시간 모터 제어를 처리하도록 합니다.

AM13E23019 는 가전 제품, 로봇 및 산업용 시스템의 적응형 제어 및 예측형 유지 보수를 위해 TinyEngine NPU와 고급 실시간 제어 아키텍처(최대 4개의 모터)를 결합합니다.

MSPM0G5187 Arm® Cortex®-M0+ 기반 MCU는 전용 TinyEngine NPU를 사용하여 메인 CPU와 독립적으로 심층 신경망 모델을 실행하여 예측형 모터 유지 보수를 위한 웨어러블 건강 모니터, 가전 제품 및 산업용 시스템의 에지 AI 기능을 지원합니다.

에지 AI 소프트웨어 혁신

임베디드 장치의 효율적인 AI 계산을 위한 하드웨어 발전과 함께 오픈 소스 커뮤니티와 반도체 제조업체도 최소한의 프로그래밍 전문 지식을 갖춘 AI 모델을 더 쉽게 테스트 및 배포할 수 있도록 만들고 있습니다. AI를 보다 사용자 친화적으로 만들고 경우에 따라 GUI 기반의 AI를 사용하면 추가 리소스 또는 교육에 투자할 필요성을 줄일 수 있습니다.

AI 모델에 더 친숙한 설계자의 경우, PyTorch 및 TensorFlow와 같은 오픈 소스 툴에서 사용자 지정 데이터 세트의 모델 아키텍처를 훈련할 수 있으며, 이 모델을 ONNX 또는 LiteRT(이전에는 TensorFlow Lite라고 부름)와 같은 임베디드 친화적 형식으로 내보낼 수 있습니다. 그런 다음 이 모델은 장치에서 이에 상응하는 오픈 소스 런타임 소프트웨어로 실행됩니다.

이러한 오픈 소스 툴은 임베디드 플랫폼의 세부 사항을 추상화하여 에지 AI 개발을 지원하며, 하드웨어 가속 백엔드(대리자라고도 함)에 액세스할 수 있는 일관된 인터페이스를 제공합니다. 이러한 백엔드는 추가 설정을 제공하여 설계자가 하드웨어 가속기에 대한 모델 위임을 더 많이 제어할 수 있도록 할 수 있습니다.

TI의 **CCStudio™ Edge AI Studio** 는 원격 TI 하드웨어 및 GUI를 사용하는 TI 임베디드 장치에서 에지 AI 애플리케이션 개발을 간소화하고 가속화하는 웹 기반 툴 모음입니다. 이 툴에는 모델 구성기, 모델 분석기, 모델 선택 툴 및 모델 제작기가 포함되어 있어 설계자가 평가 보드에 물리적으로 연결하지 않고도 모델과 모델의 성능을 빠르게 평가할 수 있습니다.

에지 AI의 확장성

임베디드 마이크로컨트롤러 또는 마이크로프로세서를 사용하여 제품을 개발할 때는 항상 제품이 시간이 지남에 따라 어떻게 발전하고 확장될 수 있는지 고려하는 것이 중요합니다. 엔지니어들은 하나의 마이크로프로세서에 대한 솔루션을 개발하는 데 수개월을 소모하고 제품을 고성능 프로세서로 업데이트할 때 처음부터 시작하고 싶어하지 않습니다.

이러한 임베디드 장치를 만드는 반도체 제조업체는 기능, 성능 및 비용 측면에서 확장성을 갖춘 포트폴리오를 개발해야 합니다. 이러한 접근 방식을 통해 개발자가 여러 장치에서 구현한 작업을 최대한 간단하게 재사용할 수 있도록 다양한 시용 임베디드 프로세서 간에 원활한 마이그레이션 전략이 보장됩니다.

에지 AI도 예외는 아닙니다. 예를 들어, 홈 로봇을 제작하는 설계자는 서라운드 비전용 카메라 3대를 갖춘 고급형 버전과 단일 전방 카메라만 있는 엔트리 레벨 버전을 모두 생산할 수 있습니다. 확장 가능한 에지 AI 가속 장치 포트폴리오를 통해 하이엔드 모델에서 엔트리급 모델로 소프트웨어를 포팅할 수 있어 두 제품을 생산하는 데 필요한 리소스의 양을 최소화할 수 있습니다. 또한 확장성은 개발자가 제품이 진화함에 따라 하나의 플랫폼에서 다음 플랫폼으로 R&D 투자 금액을 이전할 수 있도록 지원합니다.

결론

에지 AI는 여전히 비교적 새로운 개념이지만, 우리 일상을 재구성할 수 있는 잠재력이 가시화되고 있으며, 특히 거의 모든 애플리케이션에 더 많은 응답성과 성능을 제공하는 능력에 중점을 두는 잠재력을 지니고 있습니다. 저전력, 비용 효율적인 임베디드 프로세서와 직관적인 소프트웨어 및 모델 교육 툴의 발전으로 모든 경험 수준의 설계자를 위한 진입 장벽이 그 어느 때보다 낮아졌습니다. TI는 이러한 일이 우리가 상호 작용하고 사용하는 전자 제품 내에서 작동하고 데이터 수집을 관리하는 각 후속 세대의 에지 AI 장치와 중요한 구성 요소(예: 센서, 전력 공급 및 연결용 반도체)로 계속될 것으로 기대합니다.

추가 리소스

- [TI의 에지 AI 프로세싱 포트폴리오 및 설계 리소스](#)를 살펴보세요.
- 다음 기술 문서에서 TI의 에지 AI 가속 MCU에 대해 알아보기
 - [에지 AI 가속 Arm® Cortex®-M0+ MCU가 전자 기기에 더 많은 지능을 제공하는 방법](#)
 - [산업 자동화 및 홈 가전 설계에서 에지 AI 기반 모터 제어 구현](#)
- 제품 개요 문서 [TI의 TinyEngine™ NPU로 더 많은 임베디드 시스템에서 에지 AI 가속 구현](#)에서 TinyEngine NPU와 에지 AI 가속 임베디드 설계에서의 이점에 대해 알아보십시오.

중요 알림: 이 문서에 기술된 텍사스 인스트루먼트의 제품과 서비스는 TI의 판매 표준 약관에 의거하여 판매됩니다. TI 제품과 서비스에 대한 최신 정보를 완전히 숙지하신 후 제품을 주문해 주시기 바랍니다. TI는 애플리케이션 지원, 고객의 애플리케이션 또는 제품 설계, 소프트웨어 성능 또는 특허권 침해에 대해 책임을 지지 않습니다. 다른 모든 회사의 제품 또는 서비스에 관한 정보 공개는 TI가 승인, 보증 또는 동의한 것으로 간주되지 않습니다.

모든 상표는 해당 소유권자의 자산입니다.

IMPORTANT NOTICE AND DISCLAIMER

TI PROVIDES TECHNICAL AND RELIABILITY DATA (INCLUDING DATASHEETS), DESIGN RESOURCES (INCLUDING REFERENCE DESIGNS), APPLICATION OR OTHER DESIGN ADVICE, WEB TOOLS, SAFETY INFORMATION, AND OTHER RESOURCES "AS IS" AND WITH ALL FAULTS, AND DISCLAIMS ALL WARRANTIES, EXPRESS AND IMPLIED, INCLUDING WITHOUT LIMITATION ANY IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT OF THIRD PARTY INTELLECTUAL PROPERTY RIGHTS.

These resources are intended for skilled developers designing with TI products. You are solely responsible for (1) selecting the appropriate TI products for your application, (2) designing, validating and testing your application, and (3) ensuring your application meets applicable standards, and any other safety, security, regulatory or other requirements.

These resources are subject to change without notice. TI grants you permission to use these resources only for development of an application that uses the TI products described in the resource. Other reproduction and display of these resources is prohibited. No license is granted to any other TI intellectual property right or to any third party intellectual property right. TI disclaims responsibility for, and you fully indemnify TI and its representatives against any claims, damages, costs, losses, and liabilities arising out of your use of these resources.

TI's products are provided subject to [TI's Terms of Sale](#), [TI's General Quality Guidelines](#), or other applicable terms available either on [ti.com](#) or provided in conjunction with such TI products. TI's provision of these resources does not expand or otherwise alter TI's applicable warranties or warranty disclaimers for TI products. Unless TI explicitly designates a product as custom or customer-specified, TI products are standard, catalog, general purpose devices.

TI objects to and rejects any additional or different terms you may propose.

Copyright © 2026, Texas Instruments Incorporated

Last updated 10/2025