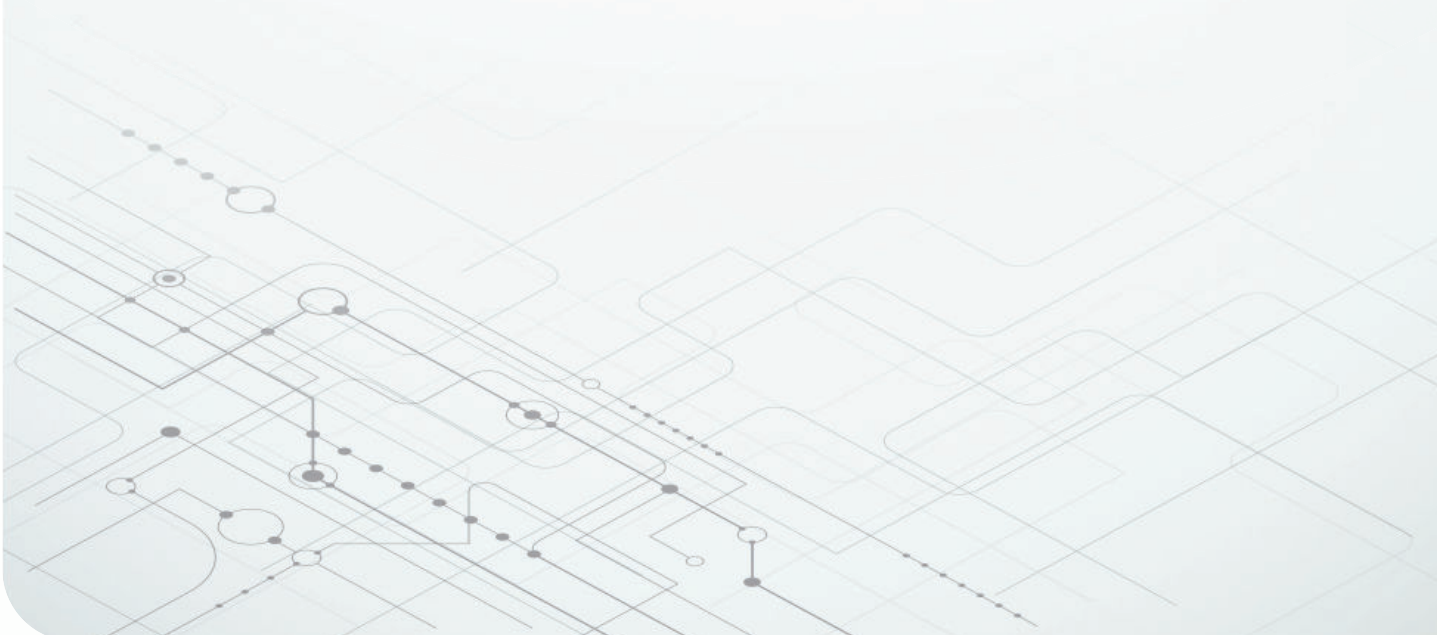


組み込みプロセッサの採用でエッジ側インテリジェンスを強化



Alec May
Systems Manager
Jacinto™ High-Performance Computing Processors



概要



1

はじめに

このホワイトペーパーでは、エッジ人工知能 (AI) の目的と利点を解説するとともに、組み込みプロセッサおよびソフトウェアの進化により、さまざまなアプリケーションにおいて AI の実装がこれまで以上に容易になっていることについて説明します。



2

エッジ AI の利点

エッジ AI の利点と、電子機器においてローカル推論を可能にするその能力について学びましょう。



3

AI はどのようにエッジへと移行しているのか

組み込みハードウェア / ソフトウェア エンジニアが直面してきた課題と、テキサス・インスツルメンツがそれらに対処する方法について解説します。



4

エッジ AI のスケーラビリティ

スケーラビリティと再利用性に関する課題を解決するために、テキサス・インスツルメンツが提供するハードウェアとソフトウェアをご確認ください。

はじめに

広く利用されているクラウドベースの AI ソリューションのアクセシビリティと使いやすさにより、ほとんどの人が AI 向けに設計されたモデルやツールに簡単に取り組むことができます。

しかし、すべての AI イノベーションがクラウドで起こっているわけではありません。組み込みプロセッサ設計の進歩により、AI 機能は現在、ノートパソコンやスマートフォンといった消費マ製品に加え、ビデオドアベル、車載システムにおける画像処理、エネルギー インフラや産業システムにおけるモーターなど、さまざまなバッテリー駆動アプリケーションにも搭載されています。

エッジ AI とは、データの出所に近い場所で AI モデルをローカルで実行する能力であり、これらの製品の応答性、効率、

信頼性、安全性を強化します。クラウドからエッジへの移行を可能にする組み込みプロセッサは、デジタル信号処理 (DSP) 向けの特化型コアなどのコンポーネントを統合したもので、AI をエッジ側に取り込むために必要な時間と専門知識を最小限に抑える、使いやすい GUI ベースのツールがサポートしています。

このホワイトペーパーでは、エッジ AI の進化と利点、およびそれを可能にしているハードウェアとソフトウェアの進歩について説明しています。

AI とは

現在、多くの人々は AI と聞くと、テキスト生成や画像生成を思い浮かべる傾向があります。しかし、最も単純なアルゴリズムでさえ、技術的には文字通りの意味での AI の例です。

AI の広範さとその複数の使用事例により、図 1 に示すように、機械学習やディープラーニングなどのいくつかのサブドメインが生まれています。

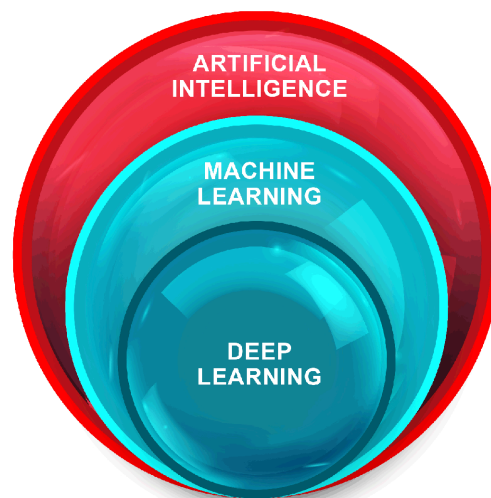


図 1. 異なる AI サブドメイン間の関係。

組み込みアプリケーションで使用される AI の大半は機械学習であり、これは機械やアルゴリズムがデータから問題の解決方法を「学習」する分野です。例えば、車両が画像データの共通パターンを解析して、歩行者と障害物を識別するようなケースが挙げられます。機械学習モデルはトレーニング データから学習しますが、そのデータには、学習すべきパターンをより正確に識別するために正解ラベル (検証済みの正確なデー

タ) が付与されていることがあります。このトレーニングプロセスにより、機械学習モデルはデータ内のパターンを識別し、将来の推論に使用できます。

機械学習の分野では、非常に複雑な問題を正確に解決できる能力があるため、ディープラーニングが非常に一般的な実装の 1 つになりました。ただし、そのような実装には多くの計算リソースが必要です。ディープラーニングは、人間の脳のニューロンに触発されたデータモデルである多層ニューラルネットワークを使用します。ニューラル ネットワークは、パターンが複雑すぎて人が見分けたりルール化したりできない問題を解決するために用いられます。

エッジ AI の利点

AI およびそのサブドメインは通常、クラウド上またはローカルサーバー上で処理を実行できます。従来、クラウドベース AI が一般的に使用されてきましたが、これは強力な AI を実行するために必要なコンピューティング能力は、大規模サーバー以外では容易に達成できなかったからです。しかし、組み込みプロセッサの計算能力と電力効率が向上してきた結果、エッジ AI の普及が進んでいます。

図 2 は、データの受信と処理、およびクラウドベースのリソースとのやり取りに関して、エッジ AI とクラウド AI がどのように異なっているかを示しています。

エッジ AI は多くの場合、開発中のモデルトレーニングにクラウドまたはデスクトップのリソースを使用します。モデルを組み込みデバイスに展開すると、新しいデータについて、独立してローカルでモデルの推論と決定を行うことができます。

最近まで、AI の特に有用な事例では、平均的なコンシューマエレクトロニクスが提供できる処理能力を上回る処理能力を必要としていました。これが、機械学習モデルのトレーニングや実装が、多くの場合クラウドベースのリソースで行われてきた理由です。クラウドベースの実装はハードウェアへの投資を最小限に抑えることで利便性を提供しますが、AI の採用も制限します。クラウドベースの AI 実装は、クラウドへのアクセス、すなわちネットワーク接続がない環境では動作しません。さらに、エッジ AI は、クラウドベース AI と比較して、セキュリティ、安全性、応答性の向上にも貢献します。

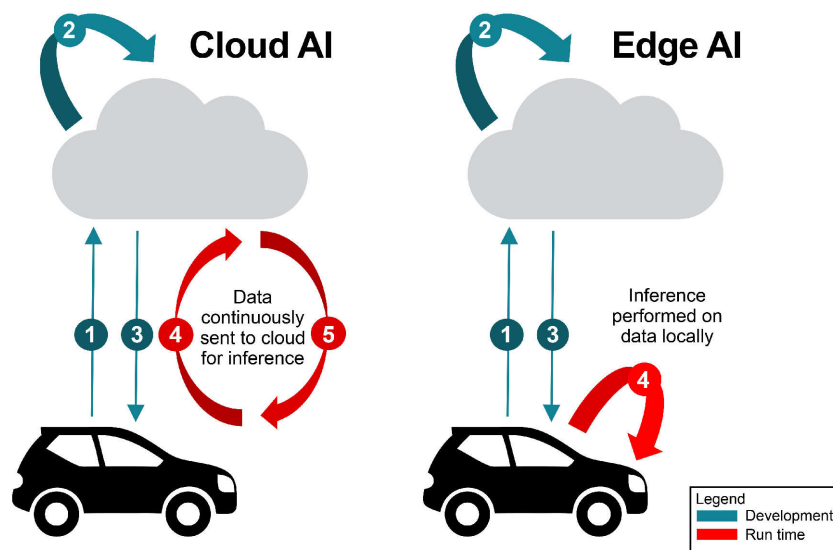


図 2. クラウドベース AI とエッジ AI の比較。

半導体の進歩と AI ツールチェーンの改良に伴い、組み込みプロセッサとマイコン (MCU) に AI ソリューションを直接実装することで、AI をエッジに移行することができるようになりました。AI をエッジで処理するとは、計算や AI による推論をデータを取得するセンサの近くで行うことを意味し、電子機器が収集するセンサ データの増加を考えると重要なことといえます。クラウドとの間での大量のデータの送受信はコストがかかり複雑で単一障害点となる可能性があるため、データ量が増加するとクラウドのみのリソースは実用的ではなくなります。

ネットワーク エッジで AI モデルを実行すると、たとえば衝突検出のため、車内に搭載しているカメラセンサなど、センサデータに基づく推論や意思決定のレイテンシを通常、短縮できます。エッジ AI 機能を活用すると、車両はクラウドからの推論を待たずに、推論をより高速に実行し、外部刺激にリアルタイムで応答できるようになります。このローカル推論は、物理 AI (図 3 参照) を通じて実際の動作へとつながりますが、これは、感知、解釈、応答を行い、工場の現場でロボットが箱を動かしたり、車両が自動的にブレーキをかけたりするなど、現実世界での動作を実行するシステムを指します。



図 3. ヒューマノイド ロボットにおけるエッジ AI と物理 AI の簡単な比較

これらのシステムは、エッジ AI を活用して高度な認識と機械的な動作を組み合わせることで、機械が人間と安全に協働しながら動作できるようにします。

ネットワーク接続への依存度の低減など、エッジ AI にはクラウドベース AI と比較した場合、その他にもいくつかの利点があります。エッジ AI は、クラウドへのアクセスが不可能な環境でも動作し、ネットワーク障害によるダウンタイムの発生を最小限に抑えます。また、クラウドベースの AI にはネットワーク

接続が必要なため、アクセスのサービス料が繰り返し発生する可能性があり、消費者向け製品を設計する際のビジネスモデルとしては困難な場合があります。

AI はどのようにエッジへと移行しているのか

組み込みプロセッサの処理能力や消費電力の制約に加え、高度な社内プログラミング スキルやリソースが求められることが、エッジ AI の広範な普及を制限してきました。AI 計算のパフォーマンス要件を満たすことができる組み込みデバイスは、多くの場合サイズが大きく、多くの電力を消費して大量の熱を発生しました。

エッジ AI に必要な計算処理をより効率的に加速するために、専用ハードウェア ソリューションが登場しましたが、いくつかのトレードオフにより、エッジ アプリケーションでの広範な採用は制限されていました。グラフィックス処理装置 (GPU)、フィールドプログラマブル ゲート アレイ (FPGA)、および特定用途向け集積回路 (ASIC) などの専用ハードウェア ソリューションは、優れた性能を実現しています。しかし、これらのソリューションは通常、高い消費電力 (特に GPU や FPGA の場合) や柔軟性の低さ (ASIC の場合) のいずれかによって制約を受けます。

統合型ニューラル プロセッシング ユニット (NPU) が、AI 機能を組み込みシステムに直接実装するためのソリューションとして登場しました。汎用プロセッサとは異なり、NPU は現代のニューラル ネットワークの中核を成す行列積、畳み込み演算、および活性化関数を実行するために特化して設計されています。計算負荷の高い処理をメイン CPU からオフロードすることで、NPU は推論速度を大幅に向上させるとともに消費電力を削減し、エッジ展開において重要な二つの要件を満たします。

これらのコンポーネントの効果を理解するために、エッジ AI 市場の異なる用途向けに設計されたテキサス インストルメンツの 2 個の NPU、すなわち高性能向けの C7™ NPU と低消費電力でスケーラブルな TinyEngine™ NPU について見ていきます。

C7 NPU

C7 NPU は、高性能で電力効率の優れた AI アクセラレータであり、**TDA54-Q1** および **TDA4VE-Q1** システム オンチップ (SoC) に統合されています。TI のマイクロプロセッサおよび SoC には、ビジョン アプリケーションの計算ニーズに対応するため、複数世代にわたって C7 NPU のバージョンが搭載されています。この NPU は TI の長年にわたる DSP の実績に基づいており、コストや消費電力を犠牲にすることなく、AI ソリューションを効率的に実現できます。

C7 NPU により、TI プロセッサは複数の AI 処理を同時に実行でき、カメラやレーダー、LiDAR など複数のセンサからのデータを同時に処理するシステムにおいて重要な役割を果たします。

TDA54-Q1 は、C7 NPU を活用することで、先進運転支援システム、インフォテインメント、およびロボティクスにおけるエッジ AI を実現します。

TinyEngine NPU

TinyEngine NPU (図 4 の簡略化されたブロック図に示す) は、マイコン向けの専用ハードウェア アクセラレータであり、ディープラーニングの推論動作を最適化することで、バッテリー動作製品を含むリソース制約が厳しいデバイスで AI ワークロードを処理するときにレイテンシと消費電力を削減します。

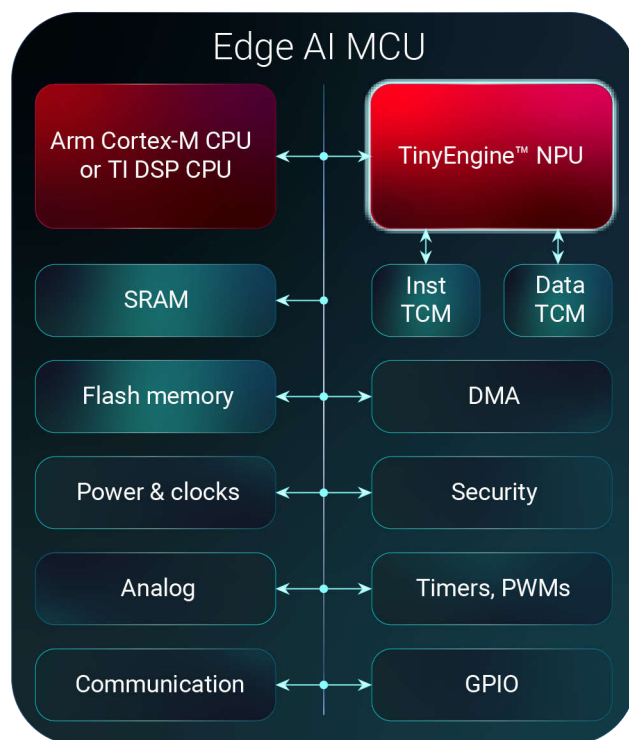


図 4. TinyEngine NPU を内蔵した TI エッジ AI マイコンの簡略ブロック図

この NPU は、アプリケーション コードを実行するメイン CPU と並行して、機械学習アルゴリズムを実行します。TinyEngine NPU 搭載マイコンは、アクセラレータ非搭載の同等のマイコンと比べて、最大でレイテンシを 90 分の 1、推論あたりのエネルギー消費を 120 分の 1 に抑えることができます。

TMS320F28P550SJ C2000™ マイコンは、モーター ベアリングとソーラー アーク故障の検出に TinyEngine NPU を使用し、メイン CPU を離れてリアルタイムのモーター制御を処理することができます。**AM13E23019** は、家電製品、ロボット、産業用システムの適応型制御と予防保守に適した、TinyEngine NPU を先進的なリアルタイム制御アーキテクチャ (最大 4 個のモーター) と組み合わせています。

MSPM0G5187 Arm®Cortex®-M0+ ベースのマイコンは、専用の TinyEngine NPU を用いて、メイン CPU とは独立し

ディープ ニューラル ネットワーク モデルを実行し、ウェアラブル健康モニタ、家電製品、および産業システムにおけるモーターの予知保全などのエッジ AI 機能を実現します。

エッジ AI ソフトウェアの革新

組込みデバイスにおける効率的な AI 計算のためのハードウェアの進化に加えて、オープンソース コミュニティや半導体メーカーも、最小限のプログラミング スキルで AI モデルのテストやデプロイを容易にしています。AI をより使いやすく(場合によっては GUI ベースに) することで、追加のリソースやトレーニングへの投資の必要性を軽減できます。

AI モデルにある程度精通している設計者にとっては、PyTorch や TensorFlow などのオープンソース ツールを用いて独自のデータセットに合わせたモデル アーキテクチャを学習させ、ONNX や LiteRT (旧称 TensorFlow Lite) といった組込み向けフォーマットにエクスポートできます。その後、モデルはデバイス上で対応するオープンソースのランタイム ソフトウェアを用いて実行されます。

これらのオープンソース ツールは、組込みプラットフォームの詳細を抽象化することでエッジ AI 開発を支援し、ハードウェア アクセラレーションのバックエンド (デリゲートとも呼ばれます) にアクセス可能な一貫したインターフェイスを提供します。これらのバックエンドは、ハードウェア アクセラレータへのモデルのデリゲーションに対して設計者がより細かく制御できるよう、追加の設定項目を提供する場合があります。

TI の **CCStudio™ Edge AI Studio** は、リモートの TI ハードウェアと GUI を活用して、TI の組込みデバイス上でのエッジ AI アプリケーション開発を簡素化し、加速する Web ベースのツール群です。このツールには、Model Composer、Model Analyzer、モデル選択ツール、Model Maker が含まれているため、設計者は評価基板に物理的に接続することなく、モデルとその性能を迅速に評価できます。

エッジ AI のスケーラビリティ

組込みマイコンまたはマイクロプロセッサを使用して製品を開発する場合、製品が長期的にどのように進化し、スケーリングするかを考慮することは常に重要です。エンジニアは、1 つのマイクロプロセッサを使用したソリューションの開発に数か月

を費やし、その後、開発した製品をより高性能のプロセッサに更新するとき最初から開発をやり直さなくてはならない、というようなことはしたくありません。

これらの組込みデバイスを開発する半導体メーカーは、機能、性能、コストの面でスケーラビリティを備えた製品ポートフォリオを構築する必要があります。このアプローチにより、AI 向けのさまざまな組み込みプロセッサ間でシームレスな移行戦略を確保でき、開発者が異なるデバイス間で作業内容をできるだけ簡単に再利用できるようになります。

エッジ AI も例外ではありません。たとえば、家庭用ロボットを製作する設計者は、サラウンドビジョン向けに 3 台のカメラを搭載したハイエンド版と、フロントカメラが 1 台のみのエントリーレベル バージョンの両方を製造したいと考える場合があります。エッジ AI アクセラレーションに対応したスケーラブルなデバイス ポートフォリオにより、ハイエンド モデルからエントリーレベル モデルへのソフトウェア移植が可能となり、両製品の開発に必要なリソースを最小限に抑えることができます。スケーラビリティにより、開発者は製品の進化に応じて、研究開発への投資をあるプラットフォームから別のプラットフォームへと移行することもできます。

まとめ

エッジ AI はまだ比較的新しいものですが、私たちの日常生活を再構築する可能性、特に、ほとんどすべての用途でより高い応答性とパフォーマンスを実現する能力が注目されつつあります。低消費電力でコスト効率の優れた組み込みプロセッサや、直観的なソフトウェアとモデルトレーニング ツールの進歩により、どのような経験レベルの設計者にとっても導入障壁はこれまでになく低くなっています。これが、エッジ AI デバイスや、私たちが日々使っており、信頼している電子機器の中で動作を管理したりデータを収集したりする、重要なコンポーネント (たとえばセンサーや電力供給、コネクティビティ向けの半導体など) の各世代で継続することが期待できます。

その他の資料

- [TI のエッジ AI プロセッシングのポートフォリオと設計リソース](#)をご確認ください。
- 以下の技術記事で、TI のエッジ AI アクセラレーション対応のマイコンについてご覧ください
 - [エッジ AI アクセラレーションを採用した Arm® Cortex®-M0+ マイコンはどのようにエレクトロニクスの脳力を向上させるか](#)
 - [産業用オートメーションおよび家電製品の設計において、エッジ AI に対応したモーター制御を実現](#)
- TinyEngine NPU と、エッジ AI アクセラレーションを採用した組み込み設計にもたらす利点に関しては、製品概要『[TI の TinyEngine™ NPU で、より多くの組み込みシステムにおいてエッジ AI アクセラレーションを活用する](#)』をご覧ください。

重要なお知らせ:ここに記載されているテキサス・インスツルメンツ社および子会社の製品およびサービスの購入には、TI の販売に関する標準の使用許諾契約への同意が必要です。お客様には、ご注文の前に、TI 製品とサービスに関する完全な最新情報のご入手をお勧め致します。TI は、アプリケーションに対する援助、お客様のアプリケーションまたは製品の設計、ソフトウェアのパフォーマンス、または特許の侵害に対して一切責任を負いません。ここに記載されている他の会社の製品またはサービスに関する情報は、TI による同意、保証、または承認を意図するものではありません。

すべての商標は、それぞれの所有者に帰属します。

重要なお知らせと免責事項

TI は、技術データと信頼性データ (データシートを含みます)、設計リソース (リファレンス デザインを含みます)、アプリケーションや設計に関する各種アドバイス、Web ツール、安全性情報、その他のリソースを、欠陥が存在する可能性のある「現状のまま」提供しており、商品性および特定目的に対する適合性の黙示保証、第三者の知的財産権の非侵害保証を含むいかなる保証も、明示的または黙示的にかかわらず拒否します。

これらのリソースは、TI 製品を使用する設計の経験を積んだ開発者への提供を意図したものです。(1) お客様のアプリケーションに適した TI 製品の選定、(2) お客様のアプリケーションの設計、検証、試験、(3) お客様のアプリケーションに該当する各種規格や、その他のあらゆる安全性、セキュリティ、規制、または他の要件への確実な適合に関する責任を、お客様のみが単独で負うものとし、

上記の各種リソースは、予告なく変更される可能性があります。これらのリソースは、リソースで説明されている TI 製品を使用するアプリケーションの開発の目的でのみ、TI はその使用をお客様に許諾します。これらのリソースに関して、他の目的で複製することや掲載することは禁止されています。TI や第三者の知的財産権のライセンスが付与されている訳ではありません。お客様は、これらのリソースを自身で使用した結果発生するあらゆる申し立て、損害、費用、損失、責任について、TI およびその代理人を完全に補償するものとし、TI は一切の責任を拒否します。

TI の製品は、[TI の販売条件](#)、[TI の総合的な品質ガイドライン](#)、[ti.com](#) または TI 製品などに関連して提供される他の適用条件に従い提供されます。TI がこれらのリソースを提供することは、適用される TI の保証または他の保証の放棄の拡大や変更を意味するものではありません。TI がカスタム、またはカスタマー仕様として明示的に指定していない限り、TI の製品は標準的なカタログに掲載される汎用機器です。

お客様がいかなる追加条項または代替条項を提案する場合も、TI はそれらに異議を唱え、拒否します。

Copyright © 2026, Texas Instruments Incorporated

最終更新日 : 2025 年 10 月