

2021 年 2 月 5 日

報道関係各位



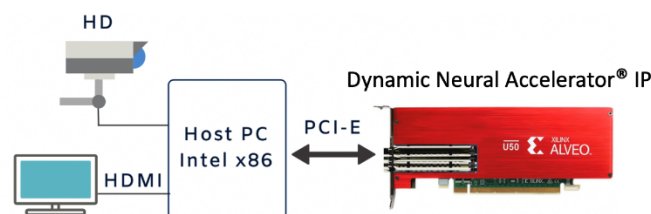
プレスリリース
株式会社 P A L T E K
EdgeCortix 株式会社

P A L T E K、EdgeCortix 株式会社と提携して エッジ AI ハードウェア・アクセラレーション・ソリューションを 市場に投入

株式会社 P A L T E K（本社：横浜市港北区、代表取締役社長：矢吹尚秀、証券コード：7587、以下 P A L T E K）は、エッジデバイス用の人工知能推論プロセッサ開発を手掛ける EdgeCortix 株式会社（本社：東京都品川区、代表取締役：サキャンガ・ダスグプタ、読み方：エッジコーテックス、以下 EdgeCortix）と提携して、エッジ AI 導入における課題解決を実現、また FPGA を活用したことのないソフトウェア製品をソリューションとして販売している会社やシステムインテグレータに対しても、エッジ AI に対する最適なソリューションを提供します。

EdgeCortix は、エッジデバイス上の深層学習推論プロセッサ用の低遅延でエネルギー効率の高いハードウェア IP（知的財産）とシステムソフトウェアを提供する会社です。このパートナーシップにより、P A L T E K は EdgeCortix のハードウェア IP と MERA™コンパイラをザイリンクス社のアクセラレータカード「Alveo™ U50 データセンター アクセラレータカード」※1 に実装・提供することで、ADAS（高度なドライバ支援システム）、自動運転、ロボット、スマートシティ、ドローン、インダストリー4.0 などにおける消費電力の削減が可能となります。

Dynamic Neural Accelerator®による
ザイリンクス社のアクセラレータカードでの
驚異的な高速AI推論と簡単な導入



EdgeCortix IP コアを使用したフル HD 解像度の推論をザイリンクス社のアクセラレータカード

Alveo™ U50 / U50LV 上で実行するためのセットアップ例

■ 今回の協業の背景・概要

近年、AI技術の要素技術であるディープラーニングにより、画像処理や音声処理など、いくつもの分野で劇的な精度向上が実現されてきましたが、現在のサービスはクラウド上のGPUを使って処理することが一般的で、推論時には通信量やネットワーク負荷、レイテンシの問題から、クラウドで効率的に処理できない活用事例が多数存在し、エッジデバイスでの推論処理の実行は重要な技術的課題となっています。

P A L T E Kが販売代理店契約を締結した EdgeCortex は、2019 年に設立され（日本とシンガポール）、エッジデバイス向けに低遅延人工知能推論プロセッサ用のハードウェア IP とソフトウェアを提供する会社です。EdgeCortex のハードウェア IP である Dynamic Neural Accelerator^{®※2} は、P A L T E K が提供するザイリンクス社のアクセラレータカード Alveo™ U50 に搭載でき、応答速度やネットワーク環境などの厳しい制限にも対応できます。「ロボット、ドローン、インダストリー4.0」などの分野では、複雑かつ大容量の演算を高速に処理でき、省スペース化、低消費電力化にも優れている FPGA がエッジ AI 領域では CPU や GPU よりも適しています。

こうした背景の下、エッジデバイスに AI 機能を搭載できる EdgeCortex の高度な技術力と、P A L T E K の量産導入までの提案・販売力および技術サポートを組み合わせることにより、エッジ AI 導入における現場の課題を解決します。また、FPGA を活用したことのないソフトウェア製品をソリューションとして販売している会社やシステムインテグレータに対しても、エッジ AI に対する最適なソリューションを提供することができると考え、このたびの販売代理店契約締結に至りました。

■ 主な取り組み内容・協業施策

本販売代理店契約締結では、両社の強みを組み合わせ、FPGAを活用したエッジAIソリューションの開発・量産導入を加速し、お客様の課題解決に向けた最適なソリューションを両社で提案・実行してまいります。

- EdgeCortex 株式会社 代表取締役 サキャンガ・ダスグプタは、次のように述べています。
「P A L T E K とパートナーシップを結ぶことを大変嬉しく思います。FPGA を含むエッジ AI ハードウェアアクセラレーションビジネスが大幅に成長すると確信します。P A L T E K との協力によって、独自のテクノロジーを活用した完全な統合ソリューションが、エッジ AI 製品の実用化と全体的なカスタマーエクスペリエンスを向上できると強く感じます。P A L T E K は 1982 年の創業以来、FPGA を中心とした半導体製品の販売・技術サポートの豊富な経験を持ち、お客様の製品開発のパートナーとして仕様検討から試作開発、量産までサポートしています。当社はこの経験に支えられて、FPGA を使用したエッジデバイスおよびデータセンターでの非常に高速な推論のための AI プロセッサ IP およびコンパイラテクノロジーが提供できるようになりました。」
- 株式会社 P A L T E K 代表取締役社長 矢吹尚秀は次のように述べています。
「現在、様々な AI サービスが市場に投入されておりますが、その多くはクラウド上の CPU や GPU で処理されています。我々がターゲットとしているロボット、スマートシティ、インダストリー4.0 などの分野では低遅延や低消費電力および長期供給性を求めるお客様が多いため、FPGA によるエッジ AI のニーズが高まっています。EdgeCortex はエッジ AI にフォーカスし、より低遅延、低消費電力に特化したソリューションを提供しており、我々が取り扱うザイリンクス社製品と一緒に提案することで更なる相乗効果が得られることを大変期待しております。」

■ EdgeCortixが提供する Dynamic Neural Accelerator[®] IP DNA-F200について

DNA-F200 は、FPGA 上のディープニューラルネットワーク（DNN）推論アプリケーション向けの EdgeCortix Dynamic Neural Accelerator[®] データフローアーキテクチャベースの IP ファミリーの新製品です。これは、HBM をサポートするザイリンクス社の Alveo[™] U50 / U50LV 適応型アクセラレータカード用に設計しています。DNA-F200（300MHz で 3.7 INT8 TOP/s）およびその前身の DNA-F100（275MHz で 2.3 INT8 TOP/s）は、高性能な畳み込みニューラルネットワーク（CNN）推論用の IP で、超低遅延でエネルギー効率が良く、高スループットなストリーミングデータ用のワークロードに最適です。非常に高速なエッジ AI アプリケーション向けに特別に設計しています。DNA-F200 / F100 は、INT8 ビットバッチサイズ 1 の推論用に高度に最適化されたインストラクションセットで実行され、ResNet、YOLO、SSD、MobileNet、FPN、MonoDepth などのすべての主流な畳み込みニューラルネットワークをサポートしています。また、Vitis^{※3} プラットフォームを備えたすべてのザイリンクス社製ボードまたはカスタムボードもサポートしています。

EdgeCortix は、独自の MERA[™]コンパイラとともに DNA-F200 ビットストリームを提供します。これにより、PyTorch や TensorFlow Lite などの最も一般的なフレームワークで設計されたディープニューラルネットワークを FPGA に最小限の努力でデプロイすることができます。オープンソースの機械学習コンパイラである Apache TVM^{※4} 上で開発された MERA[™]を使用すると、機械学習エンジニアは CPU または GPU 向けに設計されたネットワークを INT8 ビット量子化することでザイリンクス社の Alveo[™]に実装された DNA IP 上で最適に実行できます。コンパイラは、ニューラルネットワークのどの部分をアクセラレータ上で実行できるかを自動的に識別し、その他の演算子をホストプロセッサ上で実行するように判断します。MERA[™]コンパイラには、シミュレータとインタプリタも組み込まれています。コンパイル後、お客様はこれらのツールを使用して、ハードウェアでテストせずにサイクル精度のパフォーマンスシミュレーションを実行したり、INT8 ビットの量子化がネットワークの精度に与える影響を定量化したりできます。

DNA-F200とMERAディープラーニングコンパイラの詳細は<https://www.edgecortix.com/> をご覧ください。

■ 専門用語説明

※1 Alveo[™] U50 データセンター アクセラレータ カード

ディープニューラルネットワーク推論などの高速処理化に優れた性能を発揮するザイリンクス社のアクセラレータカード。詳細は、<https://japan.xilinx.com/products/boards-and-kits/alveo/u50.html> をご確認ください。

※2 Dynamic Neural Accelerator

EdgeCortix が提供する IP で、ハードウェアアーキテクチャと高性能 DNN モデルを組み合わせ、スループット、遅延、電力、サイズなどのメトリックの組み合わせに対して設定された目標を達成し、設計と最適化を同時に行う IP。

※3 Vitis 統合ソフトウェア プラットフォーム

ザイリンクス FPGA、SoC、Versal ACAP でエンベデッド ソフトウェアおよびアクセラレーション アプリケーションを開発するための環境です。詳細は、<https://japan.xilinx.com/products/design-tools/vitis.html> をご確認ください。

※4 Apache TVM

オープンソースでエンドツーエンドの CPU、GPU、アクセラレータ向け機械学習コンパイラフレームワークです。詳細は、<https://tvm.apache.org/> をご確認ください。

※ザイリンクスの名称および、その他本プレスリリースに記載のブランド名は米国およびその他の各国のザイリンクスの登録商標または商標です。その他すべての名称は、それぞれの所有者に帰属します。

※Dynamic Neural Accelerator[®] および MERA[™] は、EdgeCortex株式会社の登録商標です。

■ EdgeCortex株式会社について

EdgeCortex 株式会社は、「低遅延、低コスト、エネルギー効率の高いディープニューラルネットワーク推論のために、クラウドレベルのパフォーマンスを組み込みエッジにもたらす」という企業使命を持って 2019 年に設立しました。

ハードウェアとソフトウェアを統合するコデザイン手法によって設計された独自の人工知能プロセッサ技術が強みです。EdgeCortex は、これまでに日本、シンガポール、アメリカの投資家から 5 億 2500 万円の資金調達に成功しました。電子製造業界の数社との既存のパートナーシップによる実績もあります。

AI ハードウェア IP の作成にソフトウェア中心のアプローチを採用し、Dynamic Neural Accelerator[®] IP コアと MERA[™]コンパイラは、カスタム ASIC や FPGA に簡単にデプロイし動作するように設計されています。

EdgeCortexに関する詳細は、<https://www.edgecortex.com/> をご覧ください。

■ 株式会社 P A L T E Kについて

P A L T E K は、1982年の創業以来、日本のエレクトロニクスメーカーに対して国内外の半導体製品の販売のほか、ハードウェアやソフトウェア等の設計受託サービスも提供し、お客様の製品開発のパートナーとして仕様検討から試作開発、量産までサポートしています。P A L T E K は、「多様な存在との共生」という企業理念に基づき、お客様にとって最適なソリューションを提供することで、お客様の発展に貢献してまいります。

P A L T E Kに関する詳細は、<https://www.paltek.co.jp> をご覧ください。

■この件に関するお問い合わせは下記へお願いします。

1：プレスリリースに関するお問い合わせ

株式会社 P A L T E K

担当者： 広報担当 柴崎、寺田

メールアドレス： pr@paltek.co.jp

所在地： 神奈川県横浜市港北区新横浜 2-3-12 新横浜スクエアビル 6F

電話： 045-477-2016

2：本件に関するお問い合わせ

株式会社 P A L T E K

担当者： FPGA ソリューション事業部

メールアドレス： info_pal@paltek.co.jp

所在地： 神奈川県横浜市港北区新横浜 2-3-12 新横浜スクエアビル 6F

電話： 045-477-2002

3：製品サービスに関するお問合わせ

EdgeCortex 株式会社

担当者： AI ハードウェアアクセラレータチーム

メールアドレス： info@edgecortex.com

所在地： 東京都品川区西五反田 3-15-6

電話： 03-6417-9661