

Auf dem Markt gibt es FPGA-Familien, die von den Herstellern **speziell mit gehärteten Hardware-Blöcken (Hard IPs) für Tensor- und Matrixmultiplikationen** ausgestattet wurden.

Anstatt Tensormultiplikationen mühsam aus tausenden einzelnen kleinen Logikzellen (LUTs) zusammenzubauen, besitzen diese Chips fertige, dedizierte Rechenwerke, die riesige Datenvektoren in einem einzigen Taktzyklus verarbeiten können. [1]

Die wichtigsten Architekturen und FPGA-Familien teilen sich wie folgt auf:

1. Altera (ehemals Intel) – "AI Tensor Blocks"

Intel/Altera war Vorreiter beim Verbauen dedizierter Tensor-Blöcke direkt in der FPGA-Logikmatrix. Sie ersetzen klassische DSP-Blöcke durch massiv parallele Multiplizierer-Arrays. [1, 2]

-
- **Agilex 5 (insb. D-Series & E-Series):** Die aktuellen, weit verbreiteten Chips für Edge- und Embedded-KI. Sie besitzen variable Präzisions-DSPs mit speziellem **Tensor-Modus**. Ein einziger Block kann pro Taktzyklus **20 INT8-Multiplikationen** gleichzeitig berechnen. Sie beherrschen zudem native AI-Formate wie *Block Floating Point* (Block FP16). [3, 4]
- **Stratix 10 NX:** Das erste reine KI-optimierte High-End-FPGA von Intel. Hier wurden die klassischen DSP-Blöcke komplett durch **AI Tensor Blocks** ersetzt. Ein einziger Block schafft hier sogar **30 INT8-Multiplikationen und 30 Additionen pro Taktzyklus** und unterstützt extrem recheneffiziente Formate wie INT4, INT8 sowie Block-FP12/16. [1, 2, 5, 6]
-

2. AMD (Xilinx) – "AI Engines (AIE-ML)"

AMD verfolgt einen anderen, sehr erfolgreichen Ansatz: Anstatt Tensor-Kerne direkt in das programmierbare Logikgitter (Fabric) einzubetten, integrieren sie ein separates **zweidimensionales Array aus dedizierten Vektorprozessoren** direkt neben der FPGA-Logik. [7, 8]

-
- **Versal Adaptive SoCs (AI Core & AI Edge Serien):** Diese Chips besitzen sogenannte **AI Engines-ML (AIE-ML)**. Dabei handelt es sich um hochentwickelte VLIW/SIMD-Vektorkerne. Ein einziger AIE-ML-Kern kann pro Taktzyklus bis zu **512 INT8- oder INT4-Matrix-Multiplikations-Operationen (MACs)** ausführen. [8, 9, 10]
- **Vorteil:** Sie bieten extrem hohe Durchsatzraten (viele Tera-Operations pro Sekunde, TOPS) für Deep Learning (z.B. CNNs oder Transformatoren) bei sehr geringem Stromverbrauch. [8, 10]
-

3. Achronix – "Machine Learning Processors (MLP)"

Achronix ist ein auf High-End-Datenbeschleunigung spezialisierter Hersteller.

-
- **Speedster7t FPGAs:** Diese Chips verfügen über sogenannte **MLP-Blöcke** (Machine Learning Processors). Ein MLP ist ein massives mathematisches Hard-Makro, das direkt Arrays von Multiplizierern, Addierer-Bäumen, Akkumulatoren und eng gekoppelte lokale Speicher (Register/RAM) enthält. Sie unterstützen standardmäßig Matrix-Vektor- und Matrix-Matrix-Multiplikationen für INT4, INT8, INT16, FP16 und Bfloat16. [1]
-

Zusammenfassung: Welches Konzept für was?

Hersteller / Familie	Spezifischer Baustein	Architektur-Ansatz	Optimiert für
Altera Agilex 5 / Stratix 10 NX	Enhanced DSP / AI Tensor Block	In die FPGA-Matrix eingebettete Tensor-Rechenwerke	Flexibel skalierbare systolische Arrays, direkte Hardware-Pipelines.
AMD Versal (ACAP)	AI Engine-ML (AIE-ML)	Separates Prozessor-Array auf dem Chip neben der FPGA-Logik.	Maximale TOPS/Watt bei Convolutional Neural Networks und großen LLM-Schichten.
Achronix Speedster7t	Machine Learning Processor (MLP)	Große, vorgefertigte Multiplizierer-Blöcke im Fabric.	Extrem schneller Datendurchsatz in Rechenzentren und bei Netzwerk-KI.

Interessieren Sie sich für den Einsatz im **Edge-Bereich (wie autonome Systeme / Bildverarbeitung)** oder für **Rechenzentren / Cloud-Beschleunigung**? Ich kann Ihnen helfen, die passende FPGA-Plattform für Ihr konkretes KI-Modell (z. B. CNN oder Transformer) einzugrenzen. [8, 11, 12]

[1] <https://arxiv.org>

[2] <https://medium.com>

[3] <https://www.altera.com>

[4] <https://www.youtube.com>

[5] <https://jonathan-hui.medium.com>

[6] <https://fpga.org>

[7] <https://adaptivesupport.amd.com>

- [8] <https://www.youtube.com>
- [9] <https://docs.amd.com>
- [10] <https://www.xilinx.com>
- [11] <https://www.amd.com>
- [12] <https://arxiv.org>