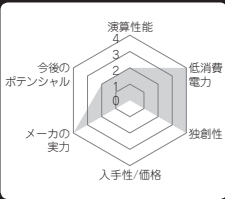


# 世界の名門大学のAIチップ

中森 章



## 16：MITの画像に特化したAIチップ

### Eyeriss (MIT)

#### ● MITの画像ディープ・ラーニングに特化したAIチップ「Eyeriss」

マサチューセッツ工科大学(MIT)で、ディープ・ラーニングにおける推論向けチップEyerissが開発されました。ディープ・ラーニングにGPUを使うのが当たり前の時代、GPUで消費する電力を大幅に削減して、ディープ・ラーニングを組み込みを持ってくるためにEyerissは開発されました。

ディープ・ラーニングの中でもCNN(畳み込みニューラル・ネットワーク)に対応しています。電力削減のために注力したのは、ニューラル・ネットワークの中で最も消費電力が大きいコンボリューション(畳み込み)層です。

#### ● 基本構造

図13にブロックとブロックの中にある演算器PE(プロセッシング・エレメント)を示します。「Filter」

というのが重みに対応し、「Ifmap (feature maps)」というのがデータ(ニューロンの値)に対応します。これらが、 $12 \times 14 (=168)$  個のPEを縦横に駆け巡り積和結果が「Psum (partial Sum)」として出力されます。重みとデータとの容量を比べると、データの方が圧倒的に多いので、外付け(オフ・チップ)DRAMに格納する際は、データはラン・レンジスで圧縮されています。そして、実際に畳み込み(積和)を行う前に伸張して使い、処理が終わると再びラン・レンジスで圧縮されて外付けDRAMに書き戻されます。また、PEではデータの識別子を利用して必要なデータが来たときのみPEを動作させるというデータ・ゲーティングを行っています。こうすることでデータ転送量を減らし、低消費電力に寄与させます。

Eyerissの内蔵バッファ(グローバル・バッファ)は108Kバイトで決して多くはありません。しかし、データ・フローをCNNに最適化することで、このバッ

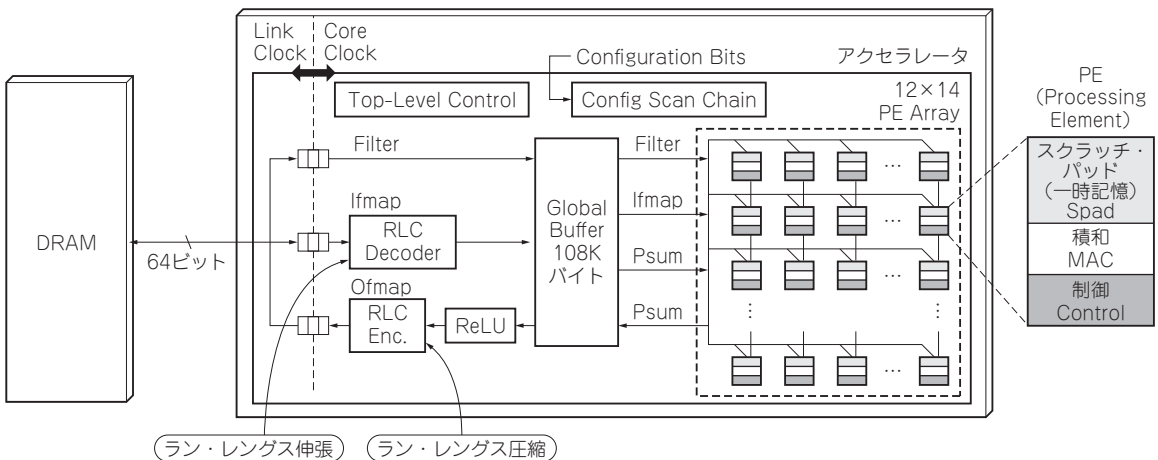


図13 (43) MITの画像ディープ・ラーニングに特化したAIチップ「Eyeriss」