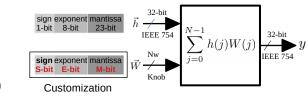
Trans-Precision Neural Network Deployment for Low-Power Embedded Systems

Abstraction levels:

Multiply-accumulate unit Hybrid custom floating-point computation

- Model deployment
- 2. System infrastructure
- 3. Streamlined acceleration



4. Optimized processing