

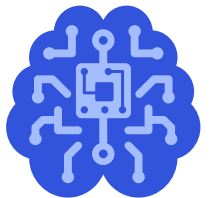
Feb 2020

@qualcomm_tech

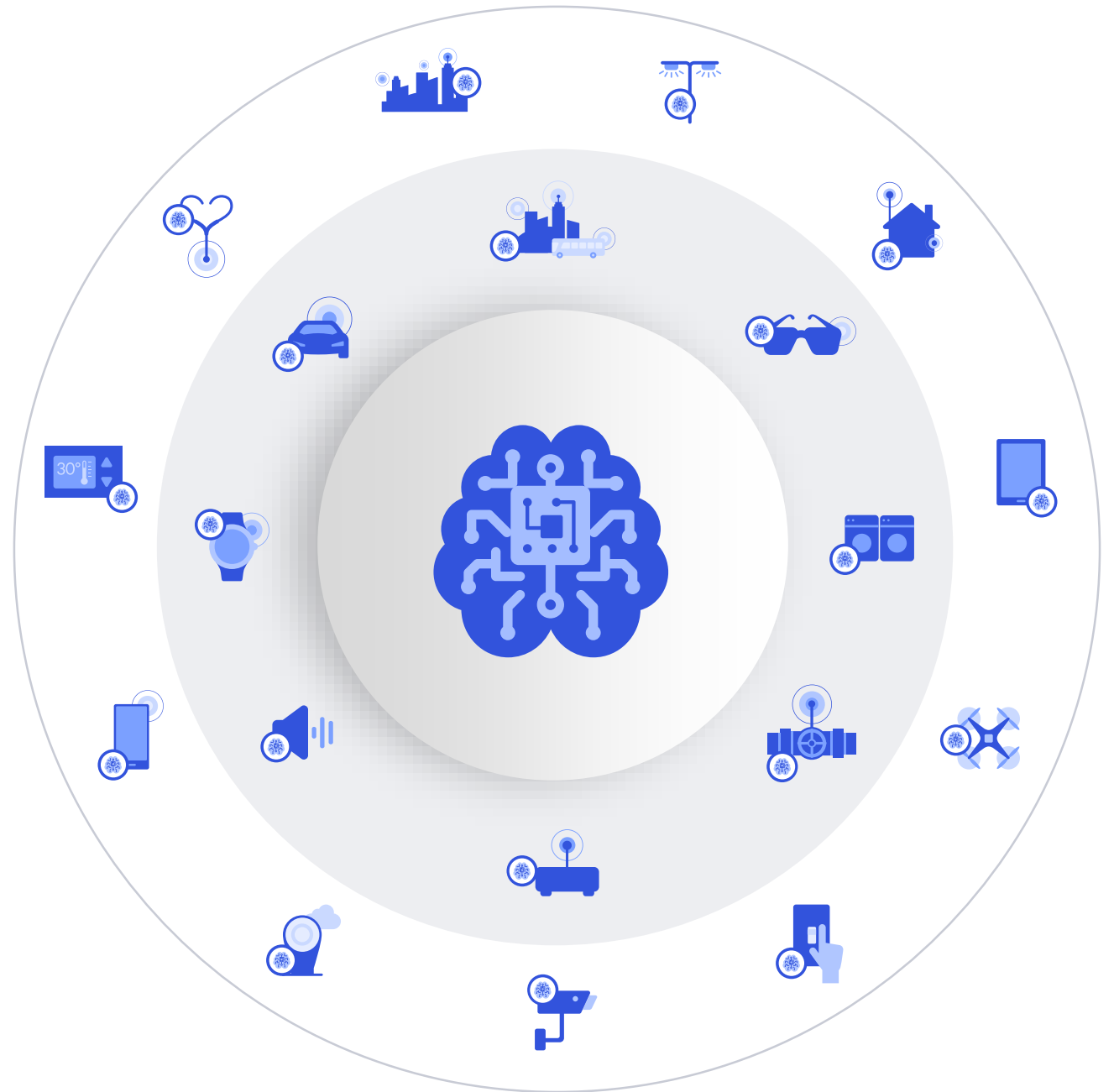
Qualcomm

Making AI ubiquitous

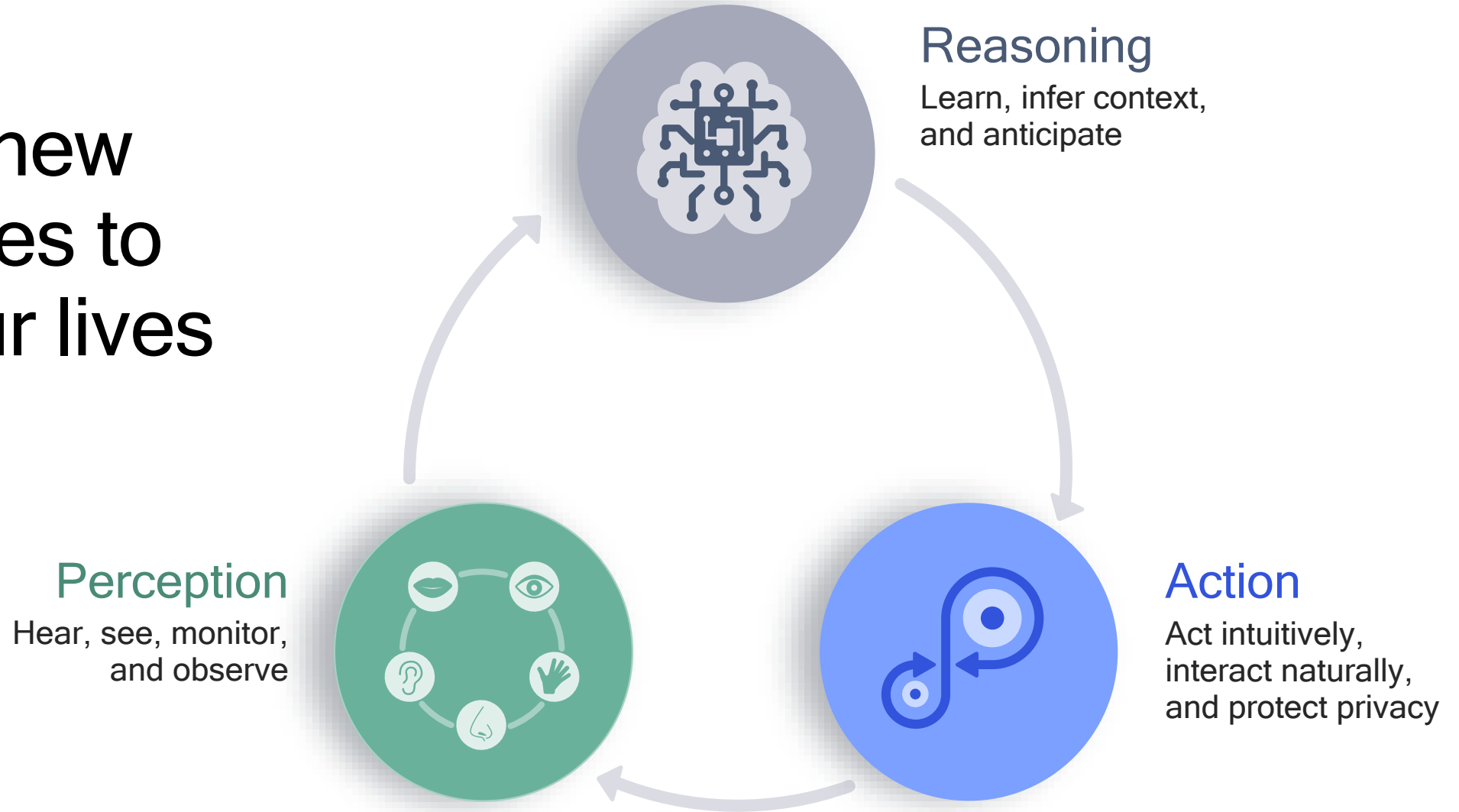
Qualcomm Technologies, Inc.



Devices, machines,
and things are becoming
more intelligent



Offering new capabilities to enrich our lives



A world where virtually
everyone and everything
is intelligently connected





Process data at the source to scale AI and make sense of a digitized world

Past

Cloud-centric AI

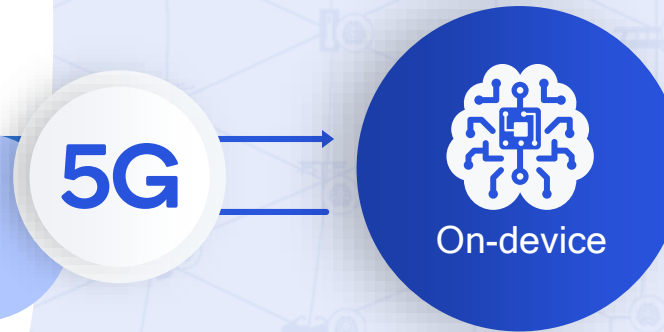
AI training and AI inference
in the central cloud



Today

Partially-distributed AI

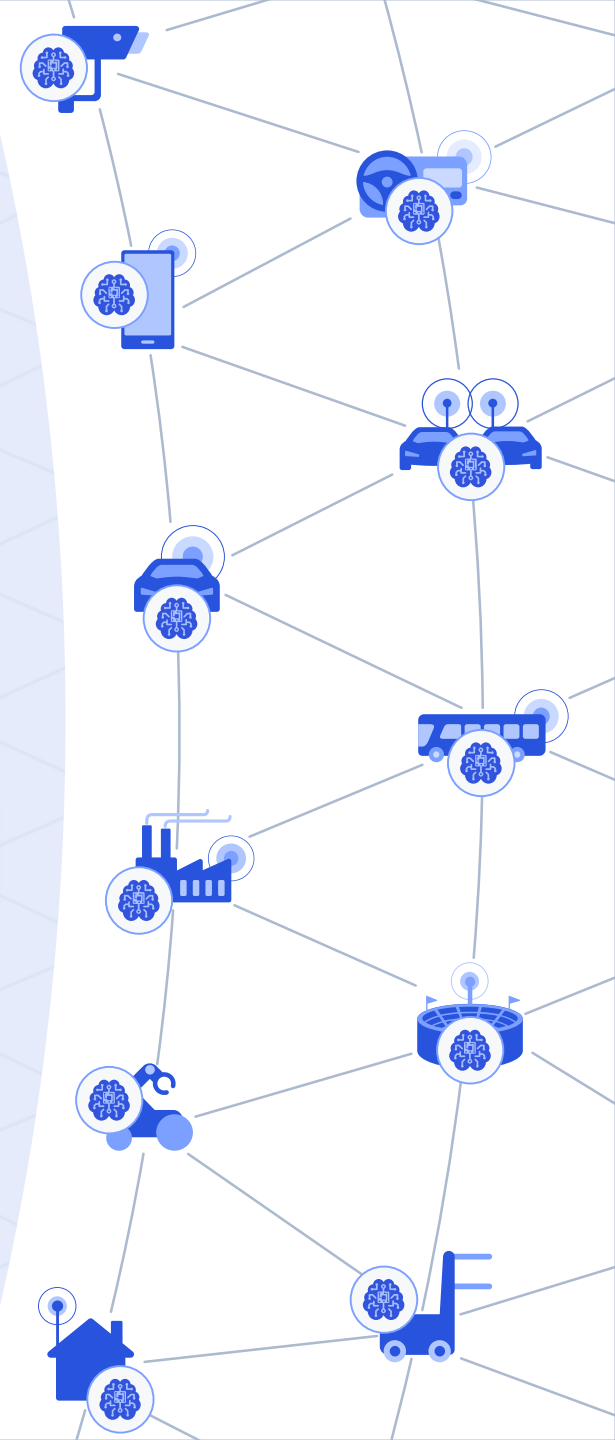
Power-efficient
on-device AI inference



Future

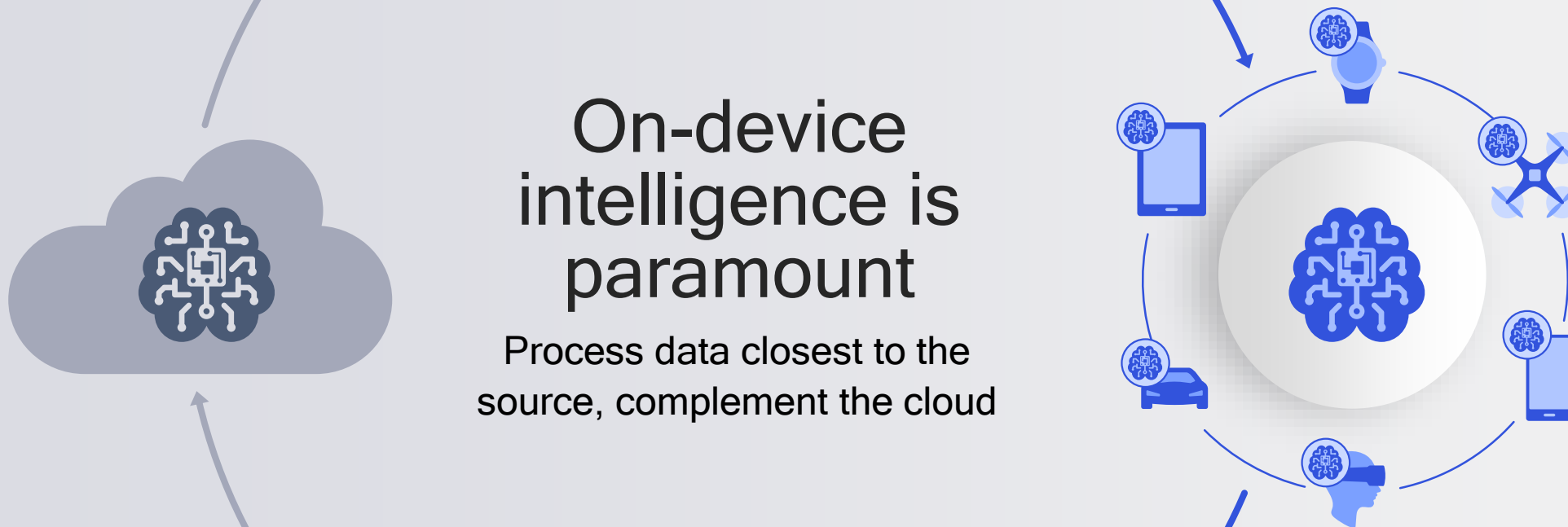
Fully-distributed AI

With lifelong on-device learning



On-device intelligence is paramount

Process data closest to the source, complement the cloud



Privacy

Reliability

Low latency

Efficient use of
network bandwidth

On-device intelligence is quickly gaining momentum

Key segments are expected to see full AI attach rates by 2025

10%

AI attach rate



2018

100%

AI attach rate



2025



Mobile



Automotive



XR



PCs /
Tablets



Smart
speakers

Mobile is
becoming the
pervasive AI
platform

~7.8
Billion

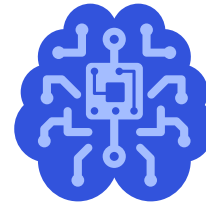
Cumulative smartphone
unit shipments forecast
between 2018-2022

Source: IDC Aug. '18



Mobile scale
changes everything

Bringing AI to the masses



Qualcomm

Mobile computing



Smart homes



Smartphones



Industrial IoT



Wearables



Smart cities



Automotive



Healthcare



Networking



Extended reality

Rapid replacement cycles

Superior scale

Integrated/optimized technologies

AI offers enhanced experiences and new capabilities for smartphones

True personal assistance



Extended battery life



Enhanced connectivity



Qualcomm



Superior photography



Natural user interfaces

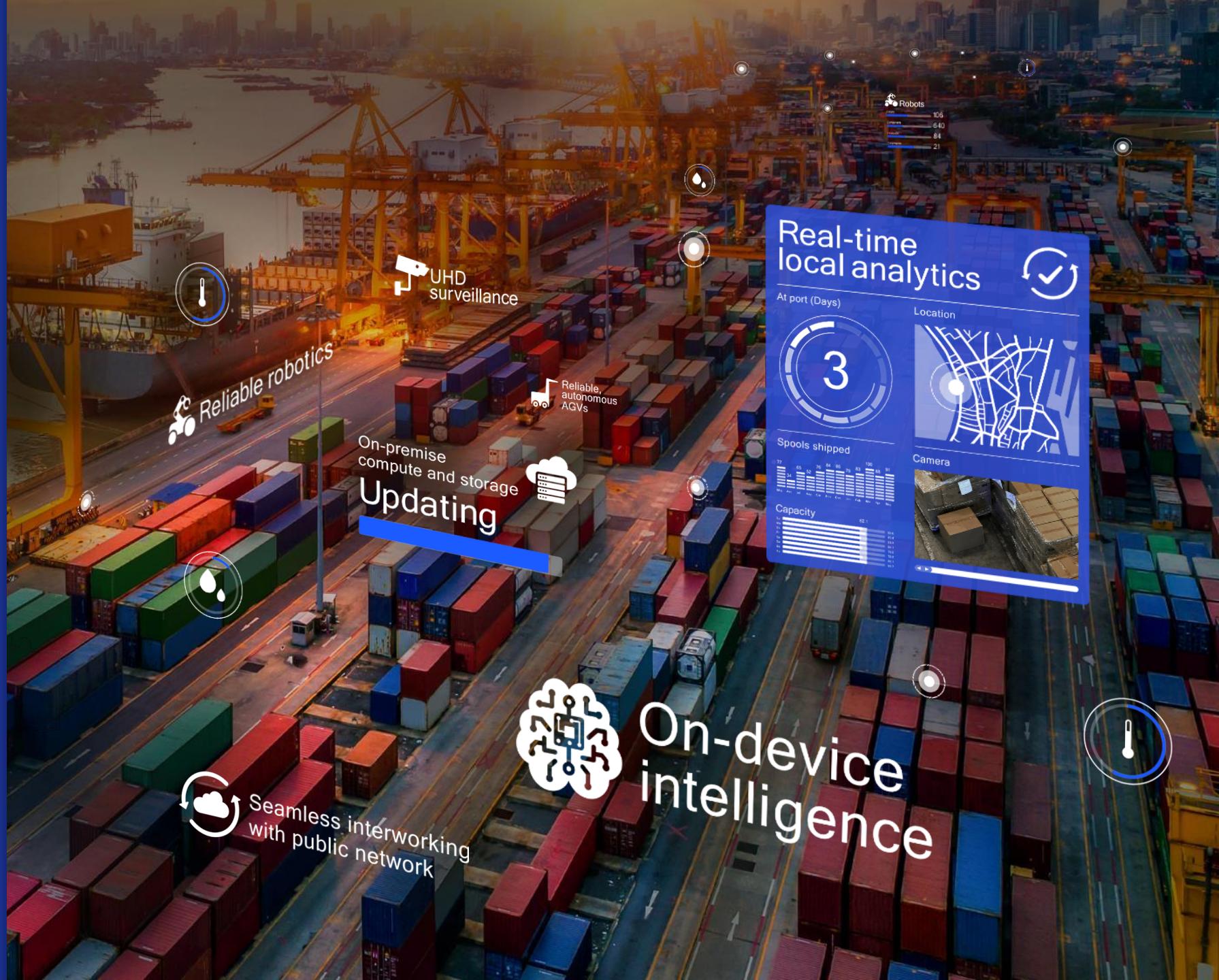


Enhanced security

A new development paradigm
where things repeatedly improve



AI will drive
transformation
across
industries





Boundless mobile XR experiences





Shaping the future of transportation

Personalized driver
settings

Driver awareness
monitoring

Greater autonomous
capabilities





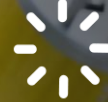
Powering the factory of the future



XR Guided execution



Surveillance



Ultra reliable,
low-latency wireless
connection



Dynamic factory
reconfigurability



5G NR
Private network



Real-time
supply chain
visibility



Predictive maintenance



Autonomous manufacturing and robotics



Smart security for home and enterprise



Smart displays and speakers



Smarter agriculture



More efficient use of energy and utilities



Home hubs and smart appliances



Sustainable cities and infrastructure



Digitized logistics and retail



IoT



AI for IoT across the home, industrial/enterprise, and Smart Cities

Power and thermal efficiency are essential for on-device AI

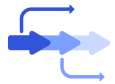
The challenge of AI workloads



Very compute intensive



Large, complicated neural network models



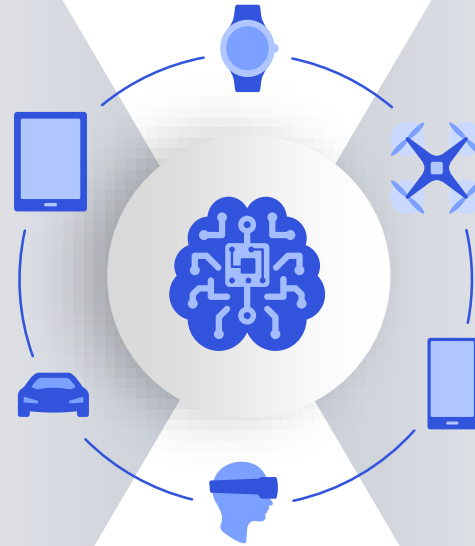
Complex concurrencies



Real-time



Always-on



Constrained mobile environment



Must be thermally efficient for sleek, ultra-light designs



Requires long battery life for all-day use



Storage/Memory bandwidth limitations

Making power efficient AI pervasive

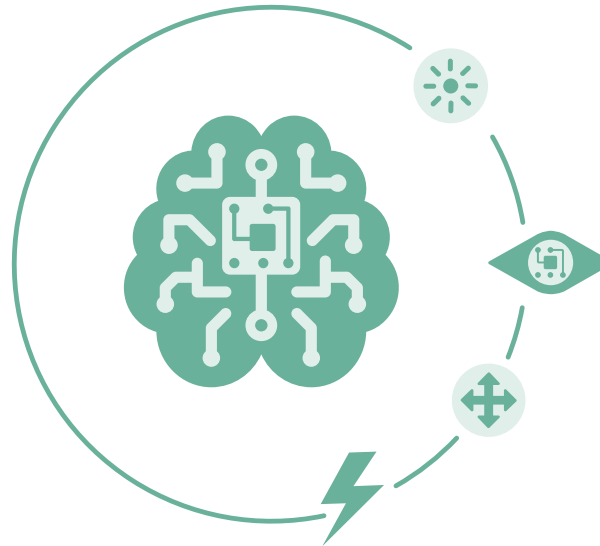
Focusing on high performance HW/SW and optimized network design



Efficient hardware

Developing heterogeneous compute to run demanding neural networks at low power and within thermal limits

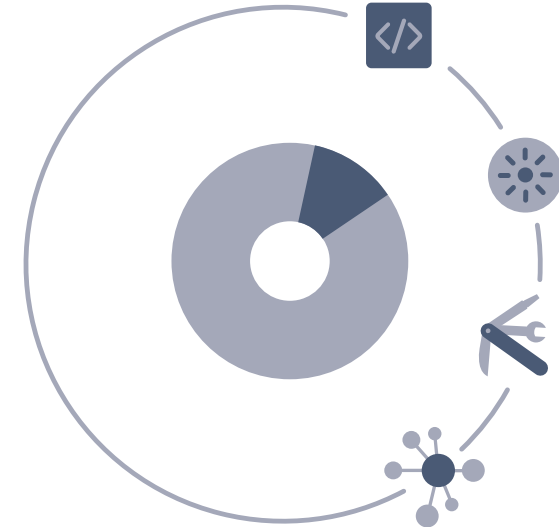
Selecting the right compute block for the right task



Algorithmic advancements

Algorithmic research that benefits from state-of-the-art deep neural networks

Optimization for space and runtime efficiency



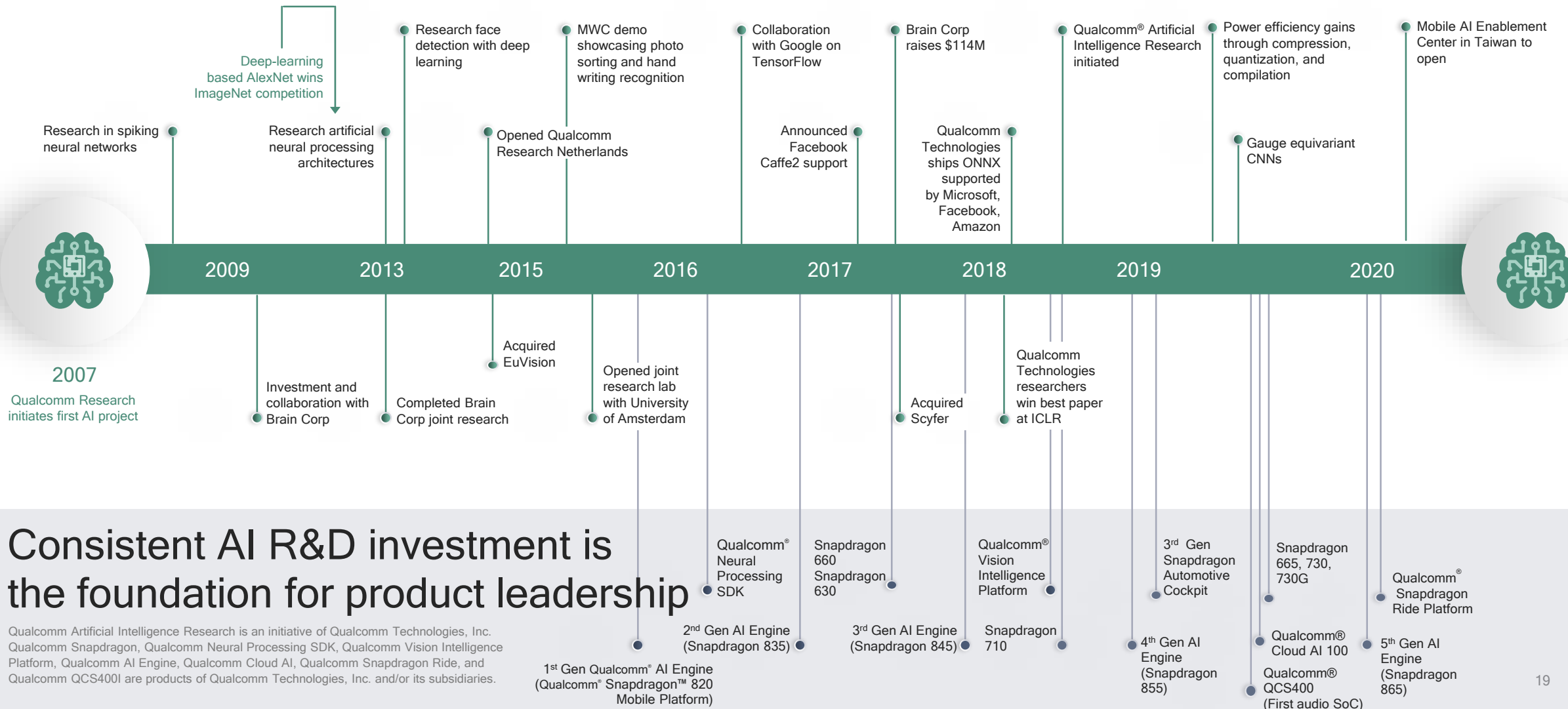
Software tools

Software accelerated run-time for deep learning

SDK/development frameworks

Our AI leadership

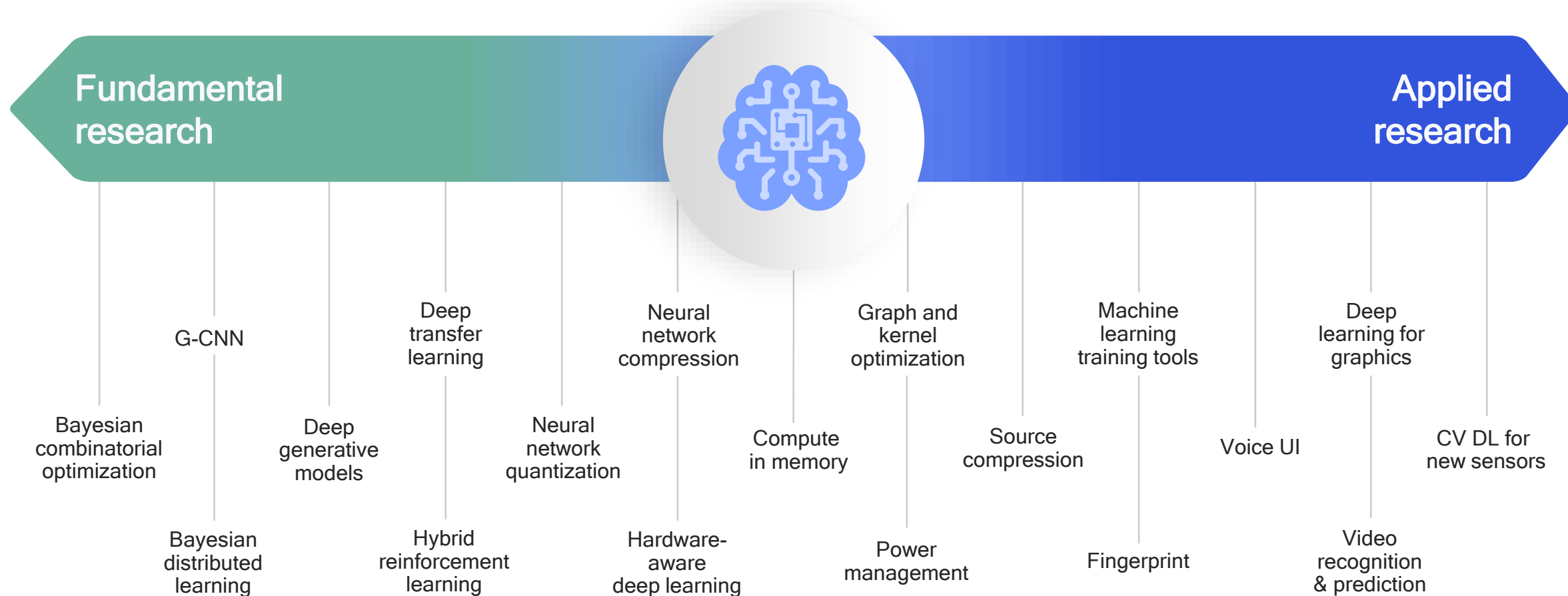
Over a decade of cutting-edge AI R&D, speeding up commercialization and enabling scale

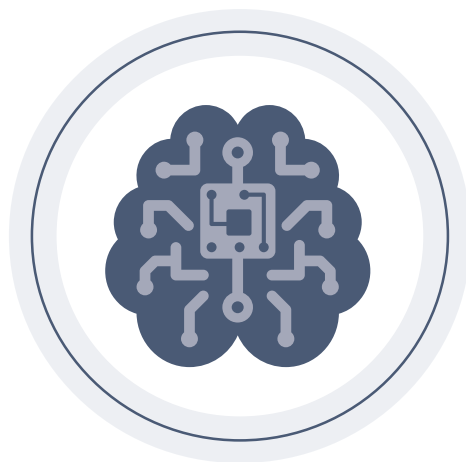


Consistent AI R&D investment is the foundation for product leadership

Qualcomm Artificial Intelligence Research is an initiative of Qualcomm Technologies, Inc. Qualcomm Snapdragon, Qualcomm Neural Processing SDK, Qualcomm Vision Intelligence Platform, Qualcomm AI Engine, Qualcomm Cloud AI, Qualcomm Snapdragon Ride, and Qualcomm QCS400 are products of Qualcomm Technologies, Inc. and/or its subsidiaries.

Leading research and development across the entire spectrum of AI





Can we apply foundational
mathematics of physics, like quantum
field theory, to deep learning?

G-CNN

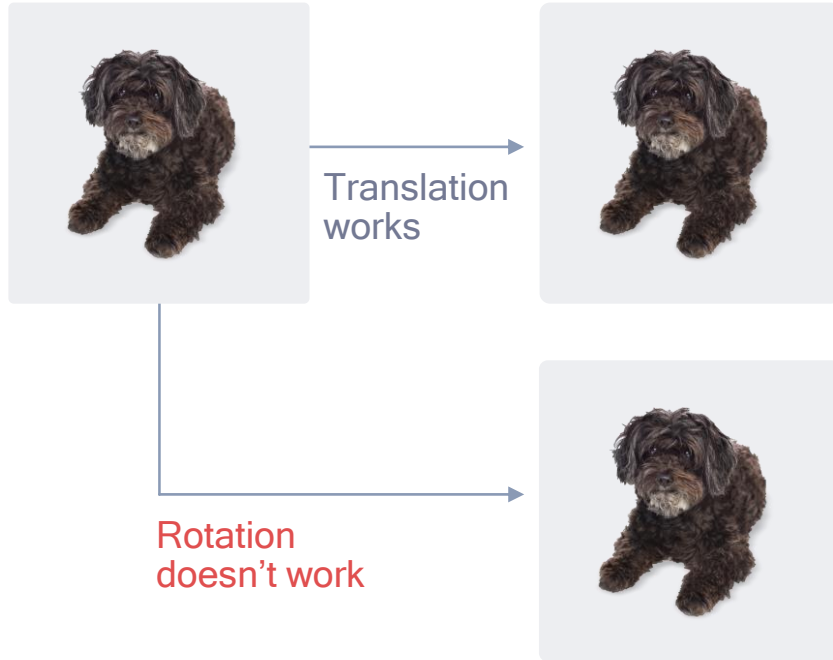
Video



Today's deep learning

Traditional CNNs

Produce state-of-the-art results but...
do not generalize input like rotations



(Convolutional neural networks would need to be retrained with new rotated images to determine new set of parameters—like filter weights)

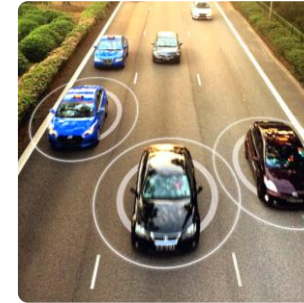
Applying
foundational
mathematics
of physics

Like quantum field theory,
to deep learning

Tomorrow's deep learning

Gauge Equivariant CNNs

No matter how you rotate or move the object,
the generalized model will still identify the object



Rotated objects and
images applicable to
drones, robots, cars,
fisheye-lens cameras,
VR, AR,...



(Generalized CNNs (G-CNN): Gauge equivariant CNN, Group, and Steerable CNN
pioneered by Qualcomm AI Research do not need to be retrained)

Advancing fundamental AI research, such as generalized CNNs

Unifying framework

Gauge equivariant CNN unify special cases like Group CNNs and Steerable CNNs, all pioneered by Qualcomm AI Research

Robust performance, faster training, and fewer training examples required

Equivariance

No matter how you rotate or move the object, it will still be identified

G-CNN can generalize models for different symmetries – traditional CNNs must be retrained

Gauge Equivariant CNN

Broad societal benefits

Use cases like drones, robots, cars, XR, fisheye lenses, 3D gaming, ...

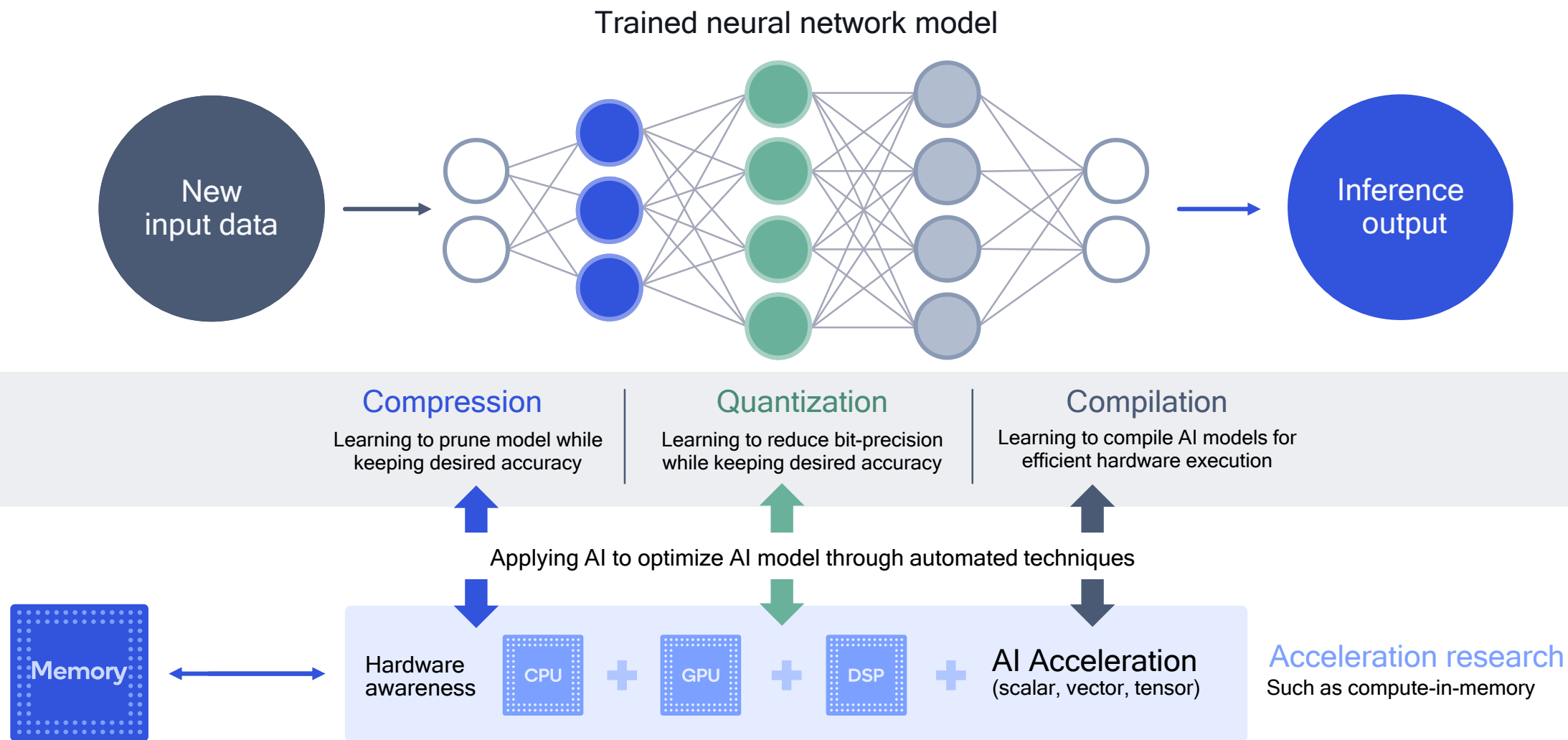
But also areas like state-of-the-art accuracy on climate pattern segmentation

Generalized geometry

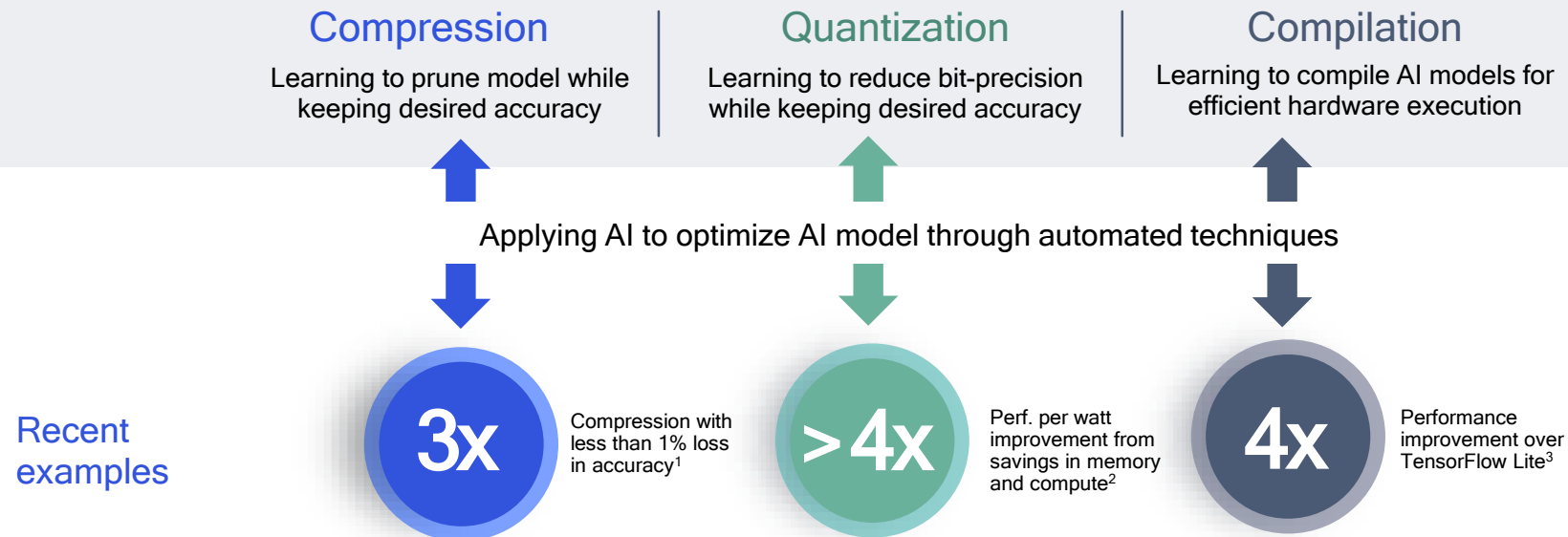
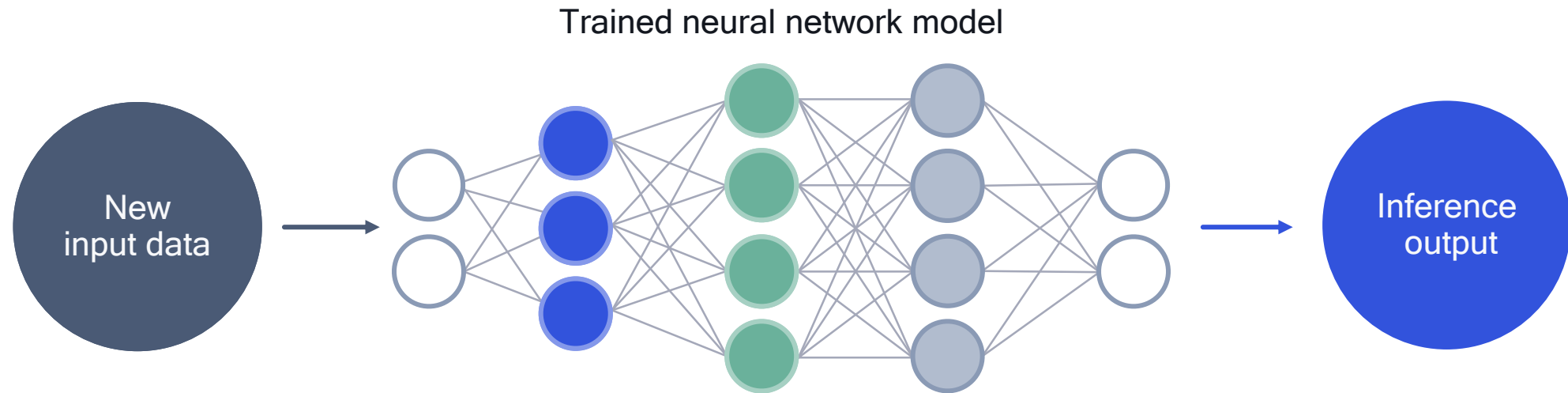
Traditional CNNs work well on narrow field-of-view cameras, but fail on e.g. fish-eye cameras

G-CNN can analyze image data on any curved space, from flat to spherical

Pioneering deep learning research in generalized CNNs



Advancing AI research to increase power efficiency



Advancing AI research to increase power efficiency

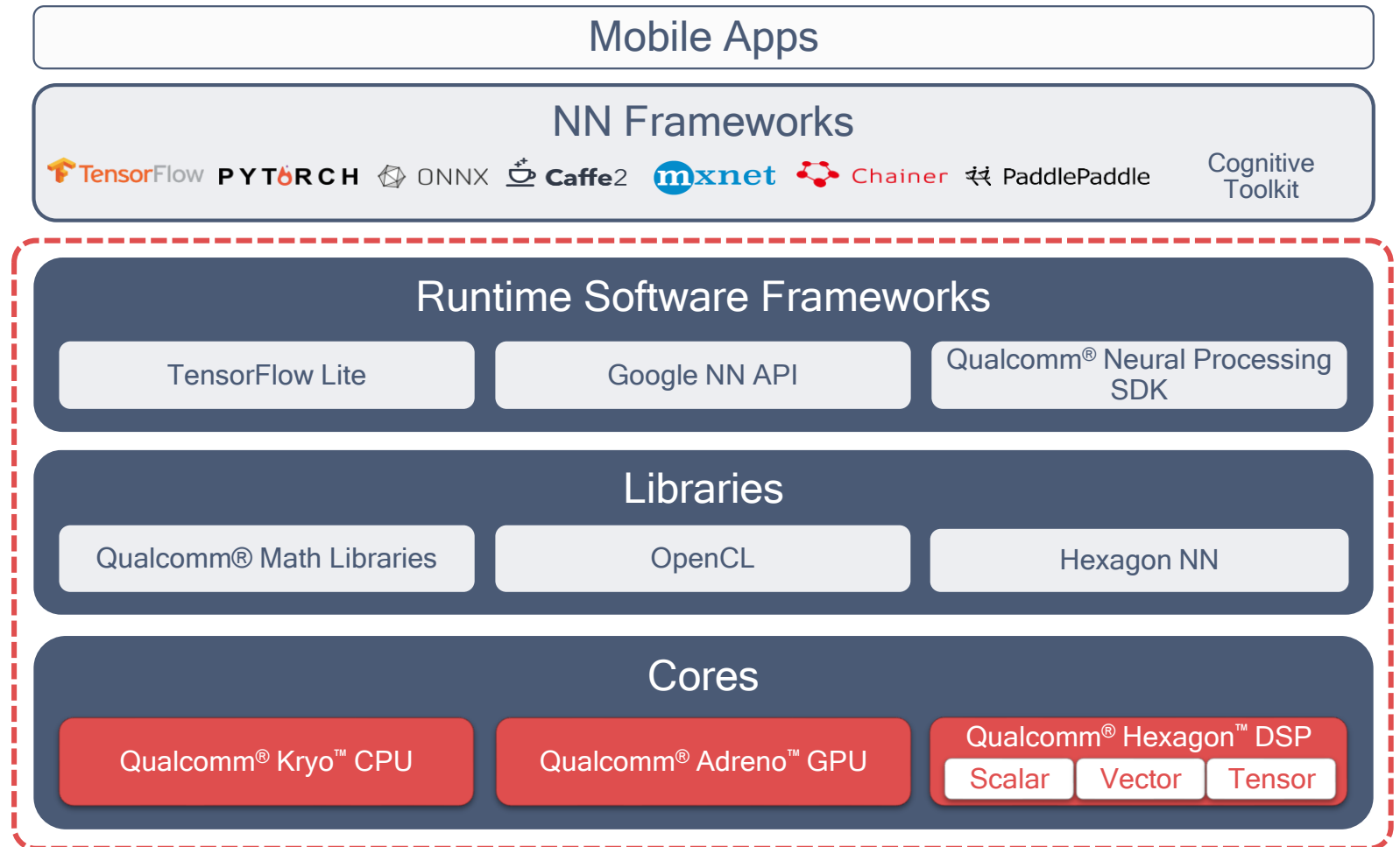
1: With both Bayesian compression and spatial SVD with ResNet18 as baseline. 2: For a quantized INT8 model vs a FP32 model that is not quantized. 3: On average improvement of tested AI models.

Qualcomm® Artificial Intelligence Engine

The hardware and software components for efficient on-device machine learning



5th Gen
AI Engine



Adreno 650

New AI mixed precision instructions

2x higher TOPS
16-bit and 32-bit FP

Hexagon 698

4x higher TOPS
Up to 35% power savings

Deep learning bandwidth compression

LP-DDR5 Memory

30% more bandwidth

5th gen Qualcomm AI engine

15 TOPS

AI highlights



Qualcomm Neural Processing SDK

New features and improvements

NNAPI support

3x number of accelerated operators

Optimized for Google ASR and Google Lens

Hexagon NN Direct

Provide developers direct access to Hexagon

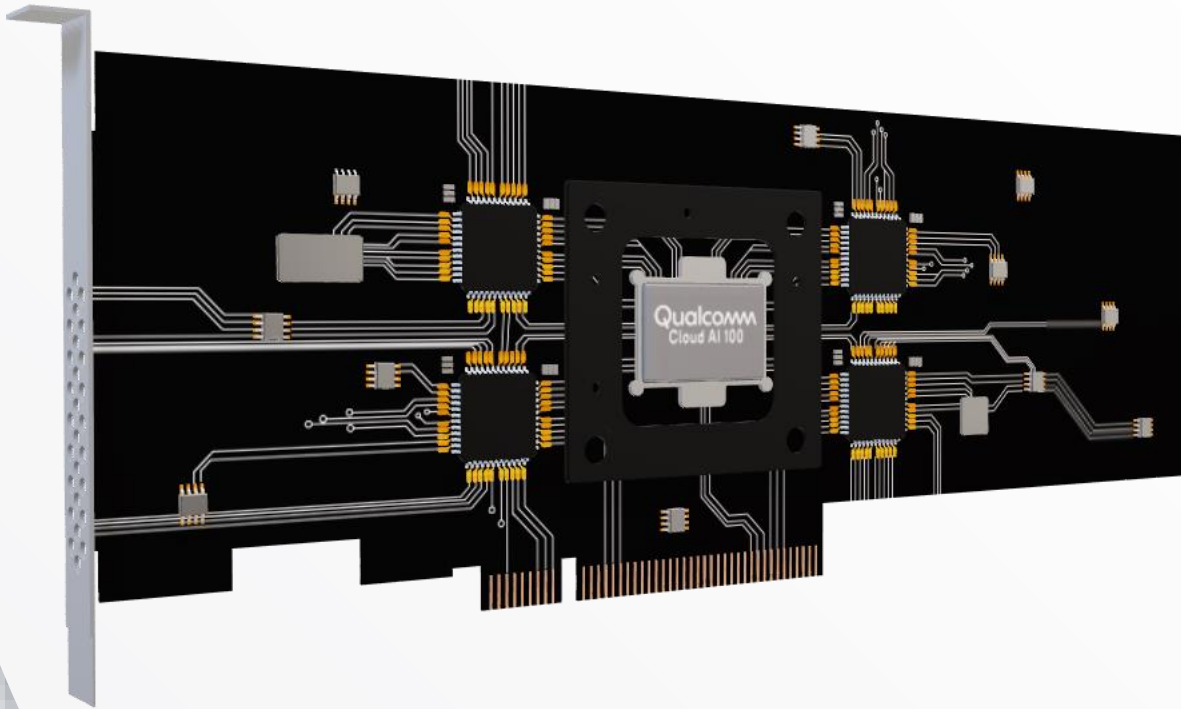
AI model efficiency toolkit

Data free quantization

Quantization aware training

Model compression

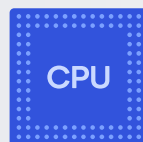
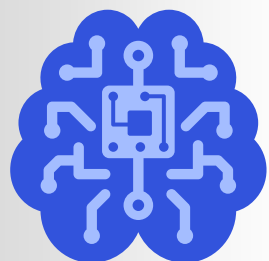
Qualcomm® Cloud AI 100



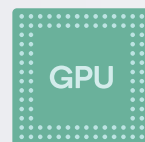
- Built on 7nm
- >350 TOPS Peak AI Performance
- Sampling 2nd half of 2019

Qualcomm® Neural Processing SDK

Software accelerated runtime for the execution of deep neural networks on device



Qualcomm®
Kryo™ CPU



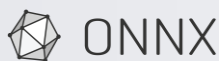
Qualcomm®
Adreno™ GPU



Qualcomm®
Hexagon™ DSP

Efficient execution on Snapdragon

- Takes advantage of Snapdragon heterogeneous computing capabilities
- Runtime and libraries accelerate deep neural net processing on all engines: CPU, GPU, and DSP with HVX and HTA



Model framework/Network support

- Convolutional neural networks and LSTMs
- Support for Caffe/Caffe2, TensorFlow, and user/developer defined layers



Offline
conversion tools



Analyze
performance



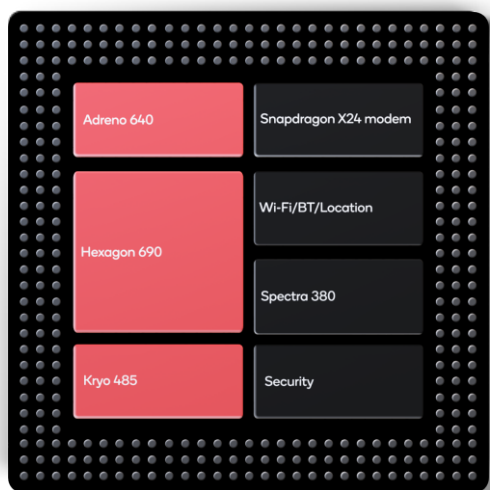
Sample
code



Ease of
integration

Optimization/Debugging tools

- Offline network conversion tools
- Debug and analyze network performance
- API and SDK documentation with sample code
- Ease of integration into customer applications



Qualcomm AI Engine

Frameworks



PYTORCH



Cognitive Toolkit

OS



Ecosystem



Uncanny Vision®



www.163.com

有态度°的门户



大象声科 | 听得见的智能



Features

Noise Suppression

Super Resolution

Night Shot

Face Recognition

Speech Recognition

Object Detection

Video Segmentation

Bokeh

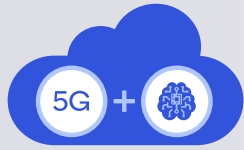
Devices



Qualcomm



Foundational
R&D



5G + AI
technology
leadership

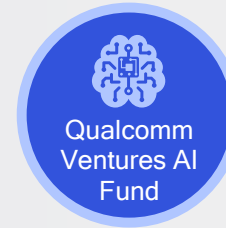


Systems
design
expertise



with AI Engine

Advanced
silicon



Ecosystem
investment



Uniquely positioned to power the
intelligently connected future



Qualcomm

Intelligence is becoming more distributed, with power-efficient on-device AI complementing the cloud

Mobile is democratizing AI and bringing it to new frontiers

Qualcomm Technologies is well positioned to provide superior AI solutions and make AI ubiquitous

Questions?

Connect with Us



www.qualcomm.com/ai



BLOG

www.qualcomm.com/news/onq



[@qualcomm_tech](https://twitter.com/qualcomm_tech)



<http://www.youtube.com/playlist?list=PL8AD95E4F585237C1&feature=plcp>



<http://www.slideshare.net/qualcommwirelessevolution>



Thank you!

Follow us on:   

For more information, visit us at:

www.qualcomm.com & www.qualcomm.com/blog

Nothing in these materials is an offer to sell any of the components or devices referenced herein.

©2018 Qualcomm Technologies, Inc. and/or its affiliated companies. All Rights Reserved.

Qualcomm, Snapdragon, Hexagon, Adreno, and Kryo are trademarks of Qualcomm Incorporated, registered in the United States and other countries. Other products and brand names may be trademarks or registered trademarks of their respective owners.

References in this presentation to “Qualcomm” may mean Qualcomm Incorporated, Qualcomm Technologies, Inc., and/or other subsidiaries or business units within the Qualcomm corporate structure, as applicable. Qualcomm Incorporated includes Qualcomm’s licensing business, QTL, and the vast majority of its patent portfolio. Qualcomm Technologies, Inc., a wholly-owned subsidiary of Qualcomm Incorporated, operates, along with its subsidiaries, substantially all of Qualcomm’s engineering, research and development functions, and substantially all of its product and services businesses, including its semiconductor business, QCT.