

# The story begins with AlphaGo



# AI, Machine Learning and Deep Leaning...



### **Artificial Intelligence (AI)**

Intelligence demonstrated by machines



### Machine Learning (ML)

Finding functioning principles through past data and experience, reaching Al ultimately.



### Deep Learning (DL)

The technology that realizes machine learning. (neural network, CNN, etc.)



# Applications of AI



- Driver-assistance system
- Emergency Alert
- Autonomous Vehicles



- Capacity boost
- Error prevention
- Defect inspection



- Language translation
- Natural language processing
- Voice recognition



- Diagnosis in medical imaging
- Drug Discovery
- Treatment queries and suggestions



- Automatic inventory-picking
- Footfall analysis



- Customer retention
- Sales forecasting
- Risk assessment



- Face detection
- Suspicious object detection
- Fire/smoke detection



DNA sequencing



- Investment analysis
- Risk detection
- Credit review

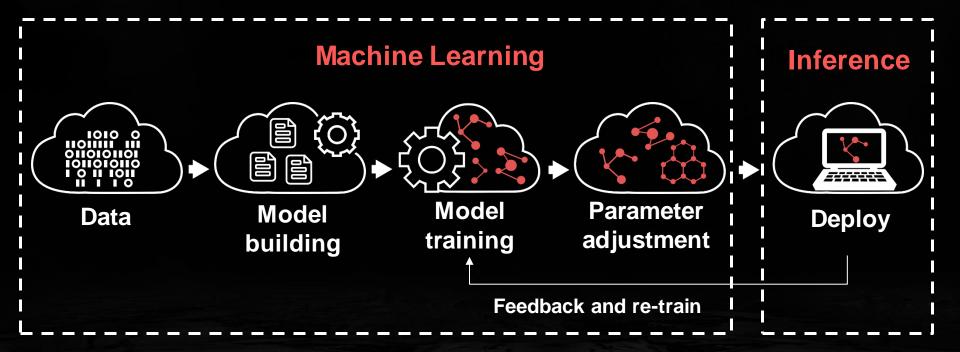


# 3 Critical Factors in Deep Learning...





### **Al Process**





# The thorn in corporations' side...

- Existing equipment with insufficient performance
- Limited budget
- Confidential information not suitable for public cloud training
- Painful environment deployment
- Enormous storage needed to save raw data
- Different management approaches for model training
- Secure and complete disaster recovery solutions required (with backup, data sharing, Internet setting, etc.)





### TS-2888X: QNAP AI NAS

#### **Deep Learning Models**

### Image classification

AlexNet, VGG16, GoogLeNet, ResNet, MobileNet, etc.

#### Face Recognition

MTCNN, DeepFace, Facenet, etc.

#### **Object Detection**

SSD, Yolo v1/v2/v3, R-FCN, RCNN, Faster RCNN, etc.

#### Video Classification

RNN, LSTM, etc.

### Image Segmentation

SegNet, U-Net, FCN, DeepLab v1/v2, etc.

### Voice Recognition

DeepVoice, WaveNet, etc

#### Framework

Caffe
Caffe2
CNTK
MXNet
Neon
PyTorch
TensorFlow

#### Inference

CoreML (iOS)/ OpenVINO/ TensorFlow Lite (Android) / TensorRT (Nvidia)

#### **Container Station**

**NVIDIA** Driver (via App Center)

**Training** 

Intel® MKL/ NVIDIA® CUDA/ OpenCL

**IGD** Driver

QTS 4.3.5

# The all-in-one Al solution for computing and storage



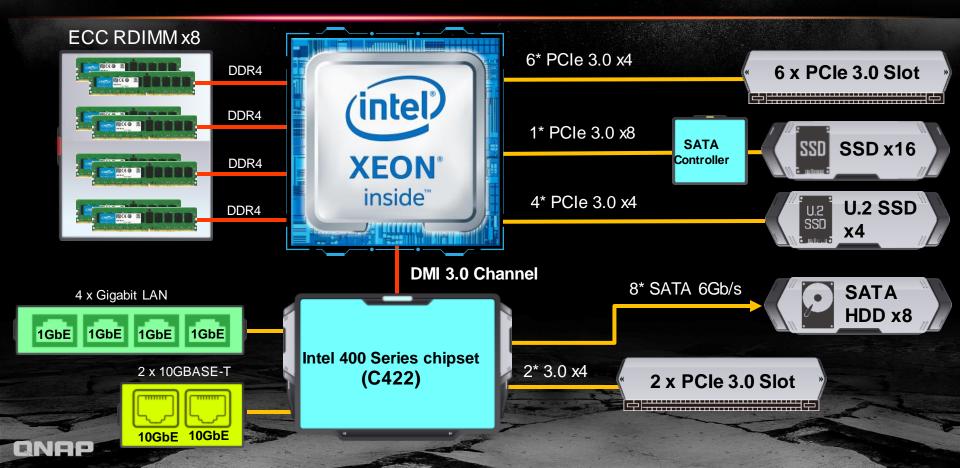


### **Intel Xeon W Processors**

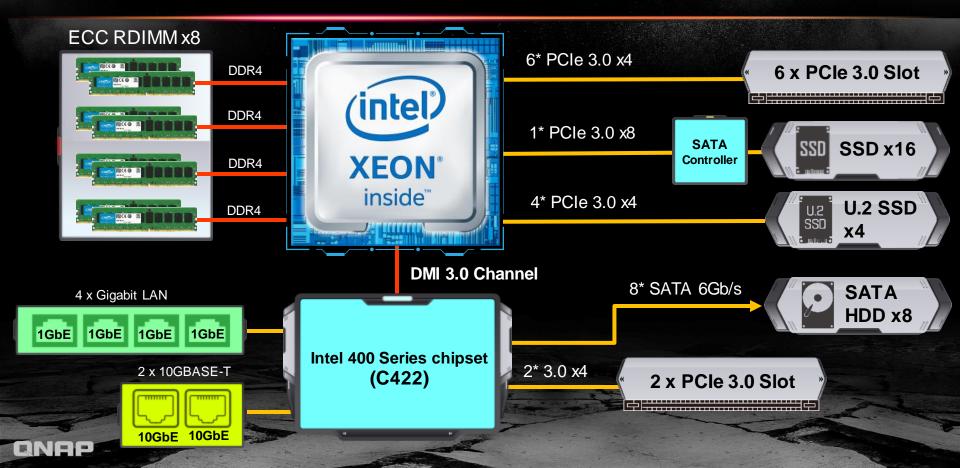
Up to 18 cores and 36 threads with up to 4.5 GHz Intel® Turbo Boost Technology 2.0 frequency, combined with up to 512 GB DDR4 ECC RDIMM 2666 MHz memory support delivering rapid workload loading and processing.



# **Block Diagram**



# **Block Diagram**



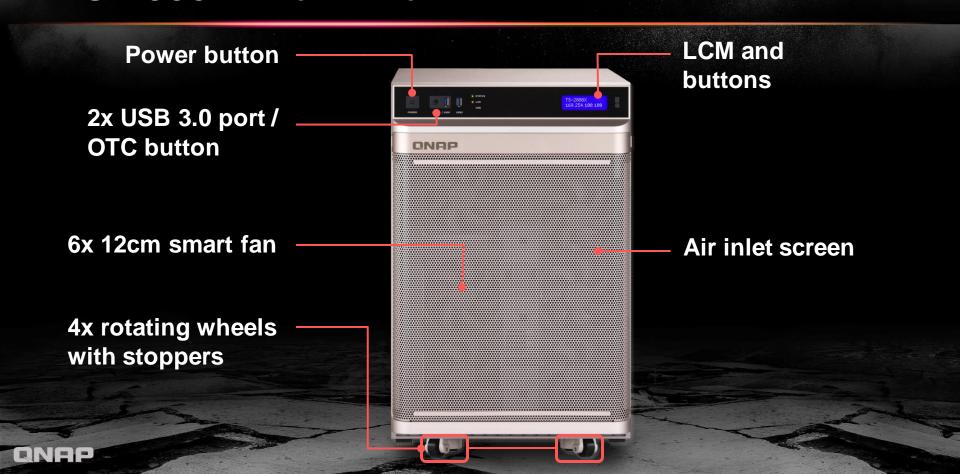
# Hyperconverged Al NAS

### **Computing in your storage**

- Reduce the performance bottleneck due to network
- Provide low latency, high speed data flow.



### **TS-2888X Front View**

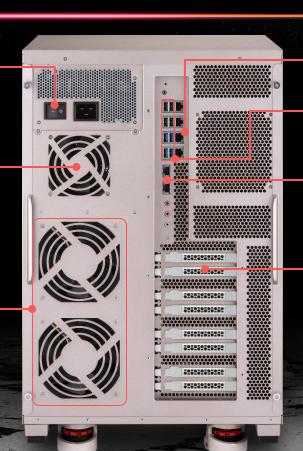


### **Rear View**

2,000 Watt PSU switch and connector

Cooling fan for power

2x 12 cm
Cooling fan for storage zone



4x 1 GbE LAN port

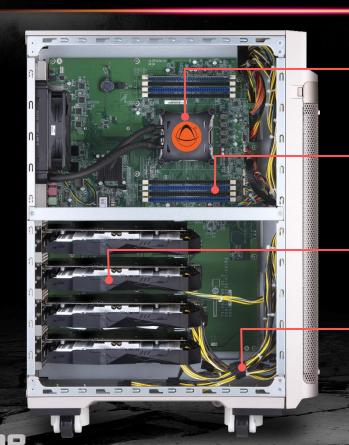
4x USB 3.0 + 4x USB 2.0

2x 10GBASE-T LAN port

8x PCle 3.0 slot



# **Exceptional Computing Capability**



Intel Xeon W Multi-core Processor

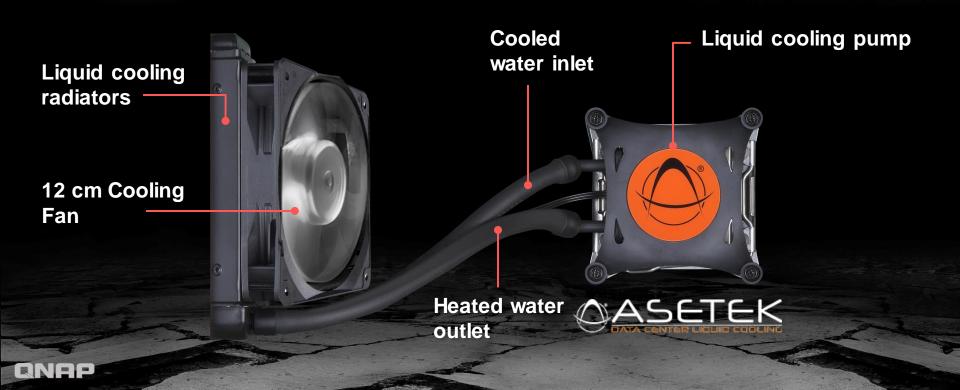
8x DDR4-2666 ECC RDIMM slot , max up to 512GB

8 x PCle Slot, up to 4 GPU adapters

8 x PCle power connector

# **ASETEK CPU Liquid Cooling System**

### **Unparalleled Cooling Performance for 140W Xeon**



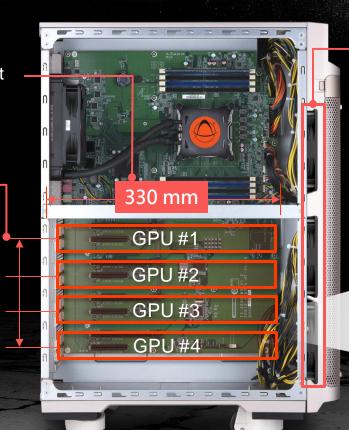
## **Supports up to 4 GPU Cards**

#### Max depth 330mm

- Fit for GPU cards in the market
- Max. 330x140x55mm

# Widen gaps between slots

- Avoid interference
- Better heat dissipation



#### 3x 12 cm cooling fan

- Front-to-back airflow for CPU and GPU cards
- 12cm quiet cooling fans

8x 6+2 pin PCle power connector





# **Supports Mainstream Acceleration Card**



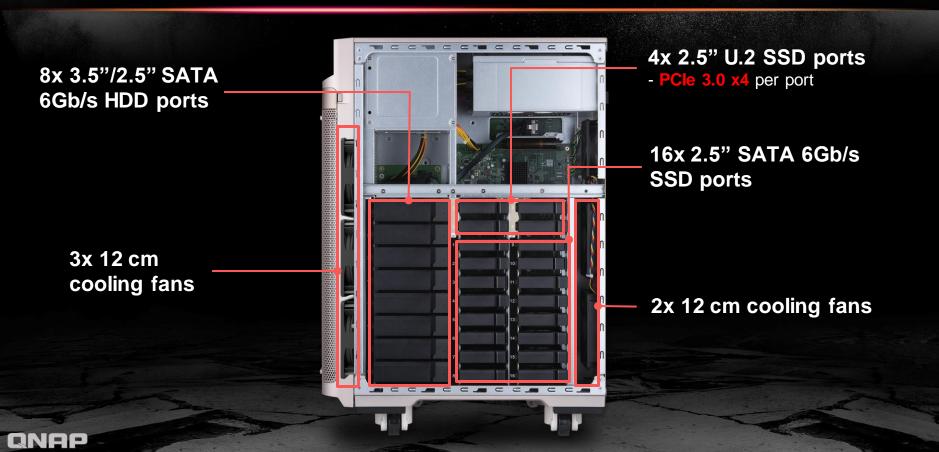




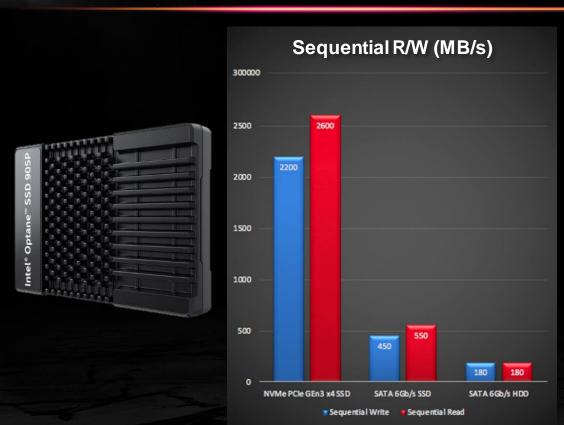


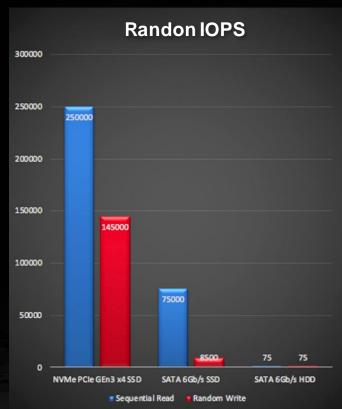
Please visit QNAP website for compatibitily list

# Massive high performance storage



### PCle 3.0 x4 U.2 NVMe SSD



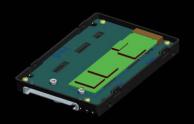




### QDC U.2 to M.2 Drive Converter







#### QDC-U2M2P

#### **Specifications**

Dimension:  $100.2 \times 69.85 \times$ 

15mm (2.5" HDD)

Controller: Asmedia 2812X

RAID: N/A

U.2 interface: PCIe Gen3 x4

M.2 interface: 1 x PCIe Gen3

x4 M.2 port, 2280



#### QDC-U22M2PR

#### **Specifications**

Dimension: 100.2x69.85x

15mm (2.5" HDD)

Controller: Marvell NevoX RAID: Hardware RAID 0,

RAID1

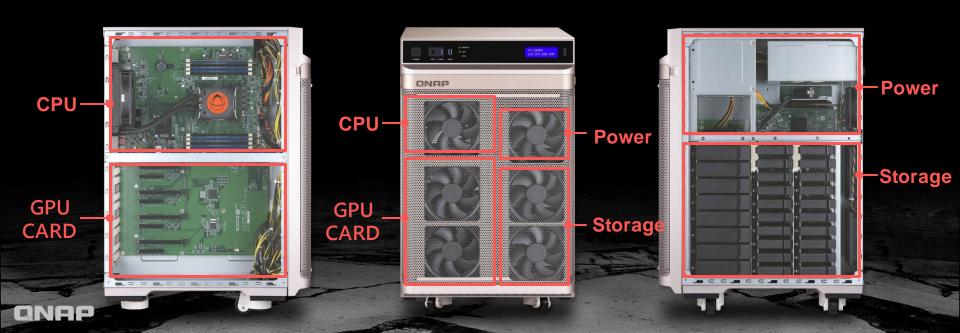
U.2 interface: PCIe Gen3 x4
M.2 interface: 2 x PCIe Gen3

x4 M.2 port, 2280



# **Compartmentalized Smart Cooling**

 Separately detects temperatures in different zones to dynamically control fan speed for quieter operations.



# **Rotating Wheels for Moving Around**



- Shipped with 4 rotating wheels
- Support up to 100 Kg dynamic load

Rubber stopper



Stopper



Note: the wheels are for short distance movements on flat surfaces (e.g. marble and tiled floor) only.

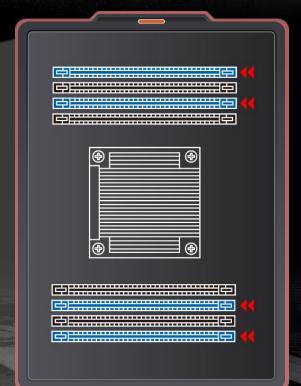
# 4-Channel DDR4 2666 Memory

- Up to 512 GB
- Installing memory in 4 or 8 to take advantage of 4-channel



#### 4 modules

Install in blue slots





# **Memory Accessories**

Ordering P/N	Specification	
RAM-8GDR4ECK0-RD-2666	DDR type: DDR4(288PIN) ECC RDIMM Capacity: 8 GB Manufacturer: Kingston Frequency: DDR4-2666 Manufacturer P/N: KSM26RS8/8MEI	
RAM-16GDR4ECK0-RD-2666	DDR type: DDR4(288PIN) ECC RDIMM Capacity: 16 GB Manufacturer: Kingston Frequency: DDR4-2666 Manufacturer P/N: KSM26RS4/16MEI	
RAM-32GDR4ECK0-RD-2666	DDR type: DDR4(288PIN) ECC RDIMM Capacity: 32 GB Manufacturer: Kingston Frequency: DDR4-2666 Manufacturer P/N: KSM26RD4/32MEI	LRDIMM & RDIMM cannot be used at the same time.
RAM-64GDR4ECS0-LR-2666	DDR type: DDR4(288PIN) ECC LRDIMM Capacity: 64 GB Manufacturer: Kingston Frequency: DDR4-2666 Manufacturer P/N: M386A8K40BM2-CTD	



### **Choose Your AI NAS**

#### TS-2888X-W2123-32G

Xeon® W-2123 4-core/8-thread 3.6GHz processor (Turbo 3.9 GHz) 32 GB DDR4 ECC RDIMM (4 x 8 GB)

#### TS-2888X-W2133-64G

Xeon® W-2133 6-core/12-thread 3.6GHz processor (Turbo 3.9 GHz) 64 GB DDR4 ECC RDIMM (4 x 16 GB)

### TS-2888X-W2145-128G (BTO)

Xeon® W-2145 8-core/16-thread 3.7GHz processor (Turbo 4.5 GHz) 128 GB DDR4 ECC RDIMM (4 x 32 GB)

### TS-2888X-W2195-512G (BTO)

Xeon® W-2195 18-core/36-thread 2.3GHz processor (Turbo 4.3 GHz) 512 GB DDR4 ECC RDIMM (8 x 64 GB)





### **Customized AI NAS**

128 GB, 256 GB and 512 GB for each SKUs

CTO SKUs have longer lead time













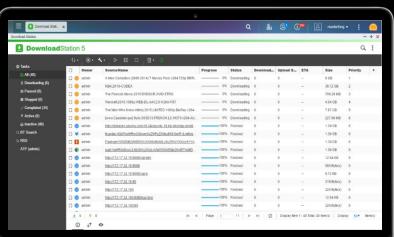
# The ready-to-use Al environment - QuAl

Ready-to-use GPU driver and Container Station enable users to easily and quickly construct containers such as Caffe, TensorFlow, MXNet, and pytorch with GUI in QTS.

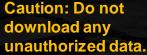


# Your data download companion - Download Station

Download Station helps to download massive training data quickly and easily. For example, more than 9 million images are in OpenImage v4 data sets (20 TB complete sets and 60 GB for competition purposes in 2018), and you can now create download tasks in Download Station to get them all done automatically and on a scheduled basis.



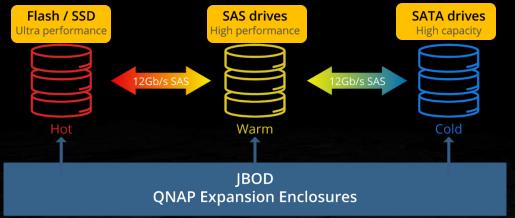






# Strike a Balance Between Performance and Storage – SSD Cache / Qtier

Even with the best GPU card at hand, the bottleneck for model training tends to be determined by IOPS because the training data comprises of a great number of small data sets (usually more than 10 TB). With SSD Cache and Qtier, you can relish the speed of SSD and the capacity of HDD simultaneously.



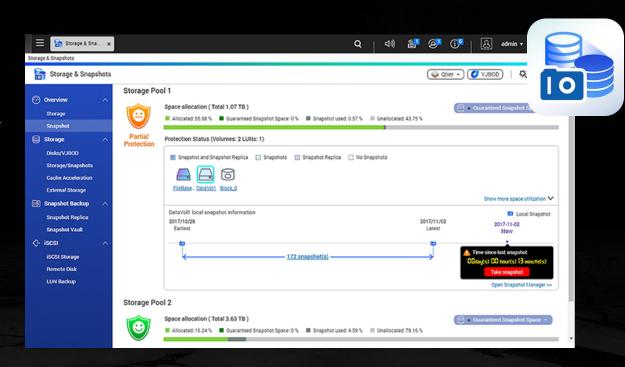




## **QNAP Snapshots**

### Keep the training data/model of every moment, at any time.

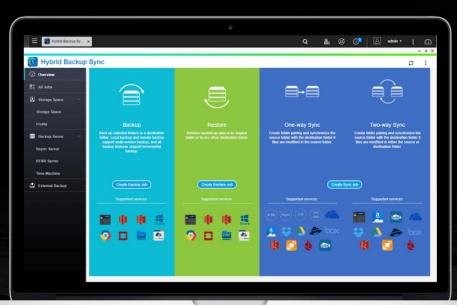
Record the state of your system at any moment with snapshots. You may keep track of the state of your system/data at any time, and mitigate disastrous losses of trained data/model.





# Data Backup/Synchronization— Hybrid Backup Sync

To synchronize trained data in different computers, files, or remote servers, simply arrange the scheduled tasks with Hybrid Backup Sync in one single click!

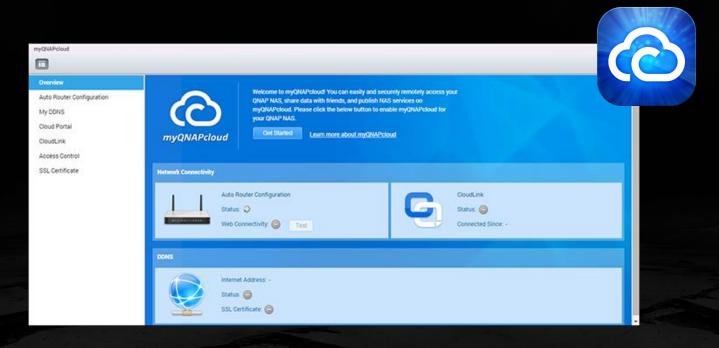




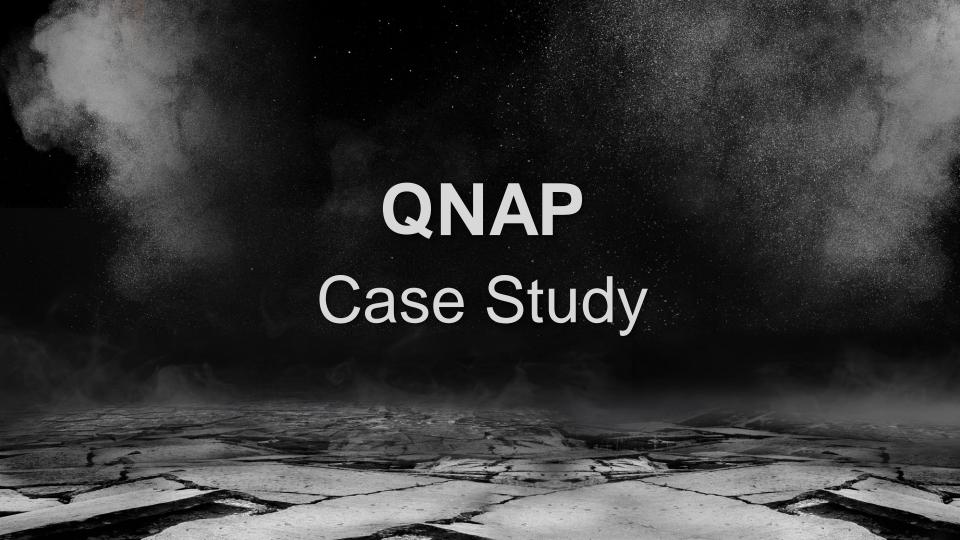


## MyQNAPCIoud - Connect to your AI NAS anytime

Whatever you do, wherever you go, keep in touch with your NAS and attend to the states of your data training/backup through myQNAPcloud.

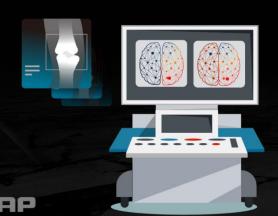






### **Health Care**

- Major Application
   Diagnosis assistance
- Dataset Type
   Medical images (e.g. OCT, MRI, etc.)





### **Process**

01



**OCT Equipment** 

















**Inference Service** 

# Case Study – Choroidal Neovascularization (CNV)

### Are there any lesions?

#### **Prepare the dataset**

- Define the disease categories.
- 2. Prepare 500 to 1,000+ labeled images for each category.
- 3. Start from public dataset.

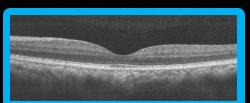
#### Methodology

Using CNN (Convolution Neural Network) topology, for example, VGG16, GoogleNet, ResNet, etc.

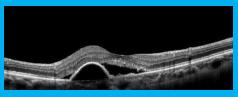
#### **Benchmark**

#### Image Net 1K:

- 1. Top-1 Error Rate: 22.25%
- 2. Top-5 Error Rate: 6.42%



Normal with 99.99% confidence



CNV (Active Wet) with 100.00% confidence

#### **Training platform**

Normally you can have a trained model in few hours or few days when using Transfer Learning with latest GPU Card.

#### Inference platform

- 1. Using Intel CPU with Intel OpenVINO to have a acceptable throughput.
- Consider GPU Card, Al Chip or FPGA card when high throughput is required.

### Case Study- Choroidal Neovascularization (CNV)

### Where are them?

#### **Prepare the dataset**

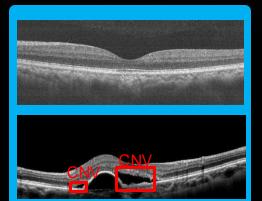
- 1. Define the disease categories.
- 2. Prepare 500 to 1,000+ labeled images for each category.
- 3. Start from public dataset.

#### Methodology

Using object detection topology, for example, SSD, Yolo v1/v2/v3, R-FCN, RCNN, Faster RCNN, etc.

#### **Benchmark**

SSD300: 73.4mAP YOLOv2 544x544: 76.8mAP



#### **Training platform**

Normally you can have a trained model in few hours or few days when using Transfer Learning with latest GPU Card.

#### Inference platform

- 1. Using Intel CPU with Intel OpenVINO to have a acceptable throughput.
- 2. Consider GPU Card, Al Chip or FPGA card when high throughput is required.



## Case Study- Choroidal Neovascularization (CNV)

### Can we mark them?

#### Prepare the dataset

- 1. Define the disease categories.
- 2. At least 500 to 1,000+ well segmented images for each category.
- 3. Start from public dataset.

#### **Training platform**

- 1. Using multi GPU cards is recommended.
- 2. For those real large 3D+ images data, you can consider using Intel Xeon Scalable Processors.

#### Methodology

Using image segmentation topology, for example, SegNet, U-Net, FCN, DeepLab v1/v2, etc.

#### Inference platform

Depend on model size and throughput requirements, the GPU card or Xeon Scalable Processor may required.

#### **Benchmark**

DeepLabv3+: 1.loU Class: 82.1 2.loU Category: 91.8

