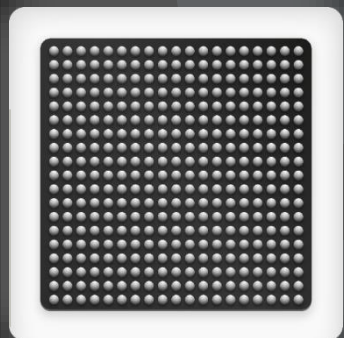




# Hailo AI Software Suite

Version 2024-10



October 2024

# Hailo AI Software Suite October 2024 Release

2024-07

3.28

4.18

2.12

3.29

## AI Software Suite Version 2024-10

- Hailo Dataflow Compiler version 3.29
- HailoRT version 4.19
- Model Zoo version 2.13
- TAPPAS version 3.30





2025-01

3.30

4.20

2.14

3.31

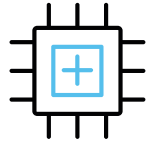
Software features legend:  Release  Preview

# Hailo AI Software Suite Version 2024-10



## Ease of Use

- Improved Transformers' parsing and compilation robustness



## New Capabilities

- Added HailoNet (GStreamer plugin) Windows support



## Enhanced Performance

- Multi-process service significant performance improvements



## Pre-Trained Models

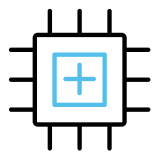
- Multi-view 3D object detection pipeline based on PETR v2
- New models: CAS-ViT, Yolo v10 models, CLIP text encoders



## Application Examples

- Hailo-15 example supports multi-scale detection

## New Capabilities



Added HailoNet Windows support, enabling GStreamer apps development for Windows

## Enhanced Performance



Multi-process service significant performance improvements, improving the throughput of complex multi-stream multi-model multi-device systems



Hailo-15 throughput improvements across many models

## Ease of Use



Improved Transformers' parsing and compilation robustness

# New AI Models in Model Zoo

Advanced Driver-Assistance Systems



New task: Multi-view 3D object detection, enabling ADAS use cases with Hailo-8. This task's pipeline is based on PETR v2, which is a Transformer-based BEV model



New models: CAS-ViT, Yolo v10 base and x-large, CLIP text encoders

Bird Eye's View

## Application Examples



Hailo-15 detection & face landmarks example now supports multi-scale detection, allowing to detect both small and large objects

# Hailo-8 Measured Benchmarks\*

| NN Model              | Input Resolution | FPS  | Power (W) | FPS/W |
|-----------------------|------------------|------|-----------|-------|
| Classification        |                  |      |           |       |
| ResNet-50 v1          | 224×224          | 1371 | 3.7       | 375   |
| MobileNet_v2_1.0      | 224×224          | 2597 | 2.2       | 1157  |
| EfficientNet_M        | 240×240          | 984  | 4.2       | 232   |
| ViT_base              | 224×224          | 139  | 2.7       | 51    |
| Object Detection      |                  |      |           |       |
| SSD_MobileNet_v1      | 300×300          | 1016 | 2.2       | 463   |
| YOLOv5m               | 640×640          | 242  | 5.3       | 45    |
| Semantic Segmentation |                  |      |           |       |
| stdc1                 | 1024×1920        | 58   | 3.1       | 19    |

<https://hailo.ai/products/ai-accelerators/hailo-8-ai-accelerator/#hailo8-benchmarks>

\* Notes:

- 1. Batch size is 8
- 2. Measurements were taken at room temperature through PCIe interface on Hailo-8 evaluation board
- 3. System host: Intel® Core™ i5-9400 CPU @ 2.90GHz; Models compiled with Hailo Dataflow Compiler version 3.29.0 (SW version 2024-10)

# Hailo-8L Measured Benchmarks\*

| NN Model                | Input Resolution | FPS  | Power (W) | FPS/W |
|-------------------------|------------------|------|-----------|-------|
| Classification          |                  |      |           |       |
| ResNet-50 v1            | 224×224          | 500  | 1.9       | 267   |
| MobileNet_v2_1.0        | 224×224          | 1739 | 1.7       | 1053  |
| EfficientNet_M          | 240×240          | 436  | 2.3       | 192   |
| Object Detection        |                  |      |           |       |
| SSD_MobileNet_v1        | 300×300          | 367  | 1.4       | 266   |
| Tiny_YOLOv3             | 416×416          | 899  | 3.1       | 291   |
| Semantic Segmentation   |                  |      |           |       |
| deeplab_v3_mobilenet_v2 | 513×513          | 90   | 2.1       | 43    |

<https://hailo.ai/products/ai-accelerators/hailo-8l-ai-accelerator-for-ai-light-applications/#hailo8l-benchmarks>

\* Notes:  
1. Batch size is 8  
2. Measurements were taken at room temperature through PCIe interface on Hailo-8L evaluation board  
3. System host: Intel® Core™ i5-9400 CPU @ 2.90GHz; Models compiled with Hailo Dataflow Compiler version 3.29.0 (SW version 2024-10)

# Hailo-15H Measured Models\*

| NN Model              | Input Resolution | FPS  |
|-----------------------|------------------|------|
| Classification        |                  |      |
| ResNet-50 v1          | 224×224          | 969  |
| MobileNet_v2_1.0      | 224×224          | 3454 |
| ViT Base              | 224×224          | 202  |
| Object Detection      |                  |      |
| SSD_MobileNet_v1      | 300×300          | 1147 |
| YOLOv5m               | 640×640          | 202  |
| Semantic Segmentation |                  |      |
| stdc1                 | 1024×1920        | 28   |

\* Notes:

- 1. Batch size is 8
- 2. DDR, A53 and NN Core are active, but all other peripherals are on IDLE
- 3. Measurements were taken at room temperature on Hailo-15 evaluation board
- 4. Models compiled with Hailo Dataflow Compiler version 3.29.0 (SW version 2024-10)



# Hailo-15M Measured Models\*

| NN Model              | Input Resolution | FPS |
|-----------------------|------------------|-----|
| Classification        |                  |     |
| ResNet-50 v1          | 224×224          | 616 |
| MobileNet_v2_1.0      | 224×224          | 869 |
| ViT Base              | 224×224          | 142 |
| Object Detection      |                  |     |
| SSD_MobileNet_v1      | 300×300          | 591 |
| YOLOv5m               | 640×640          | 138 |
| Semantic Segmentation |                  |     |
| stdc1                 | 1024×1920        | 25  |

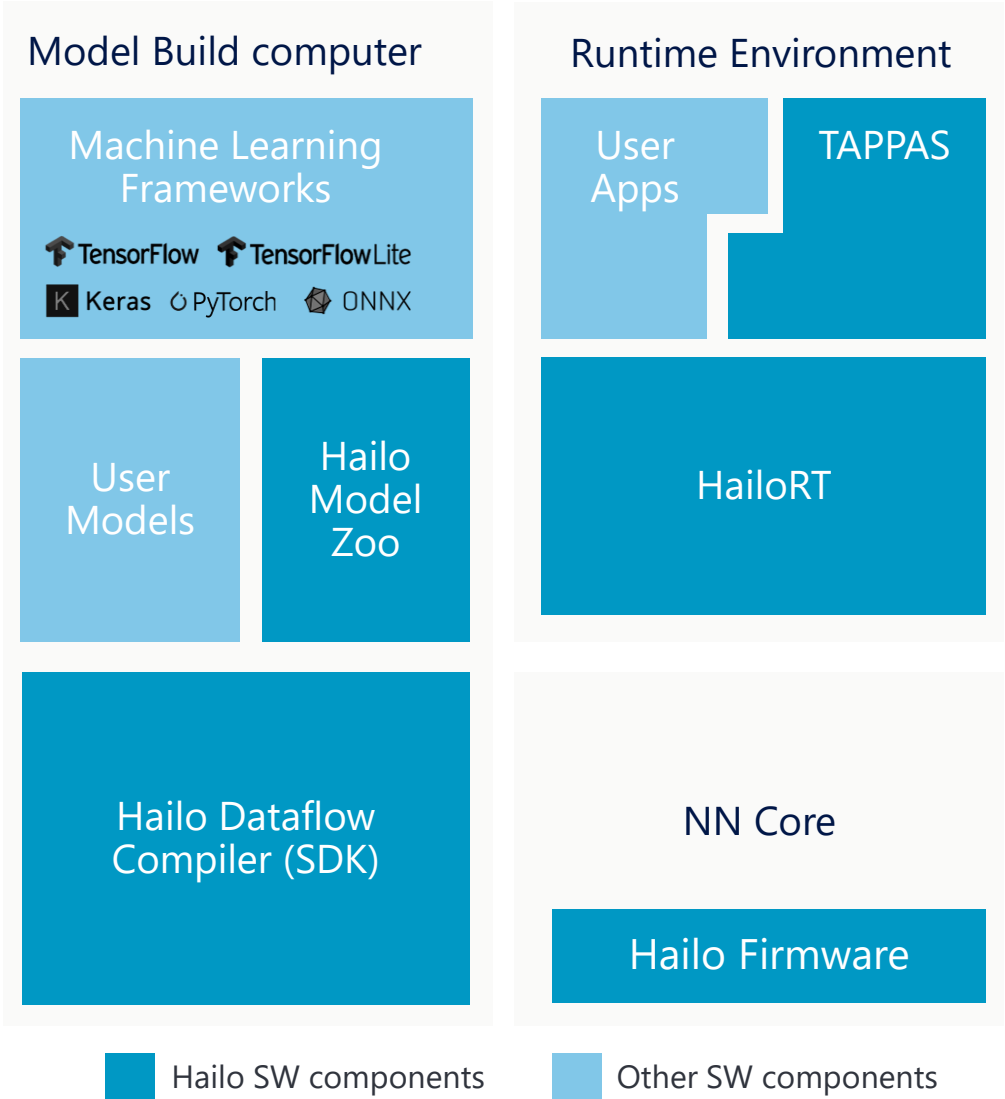
\* Notes:

- 1. Batch size is 8
- 2. DDR, A53 and NN Core are active, but all other peripherals are on IDLE
- 3. Measurements were taken at room temperature on Hailo-15 evaluation board
- 4. Models compiled with Hailo Dataflow Compiler version 3.29.0 (SW version 2024-10)

# Comprehensive AI Software Suite

## Build Environment

- Seamless integration with existing deep learning frameworks
- Large variety of ~100 state-of-the-art and common free models
- Maximizing AI compute performance by efficient utilization of NN core resources

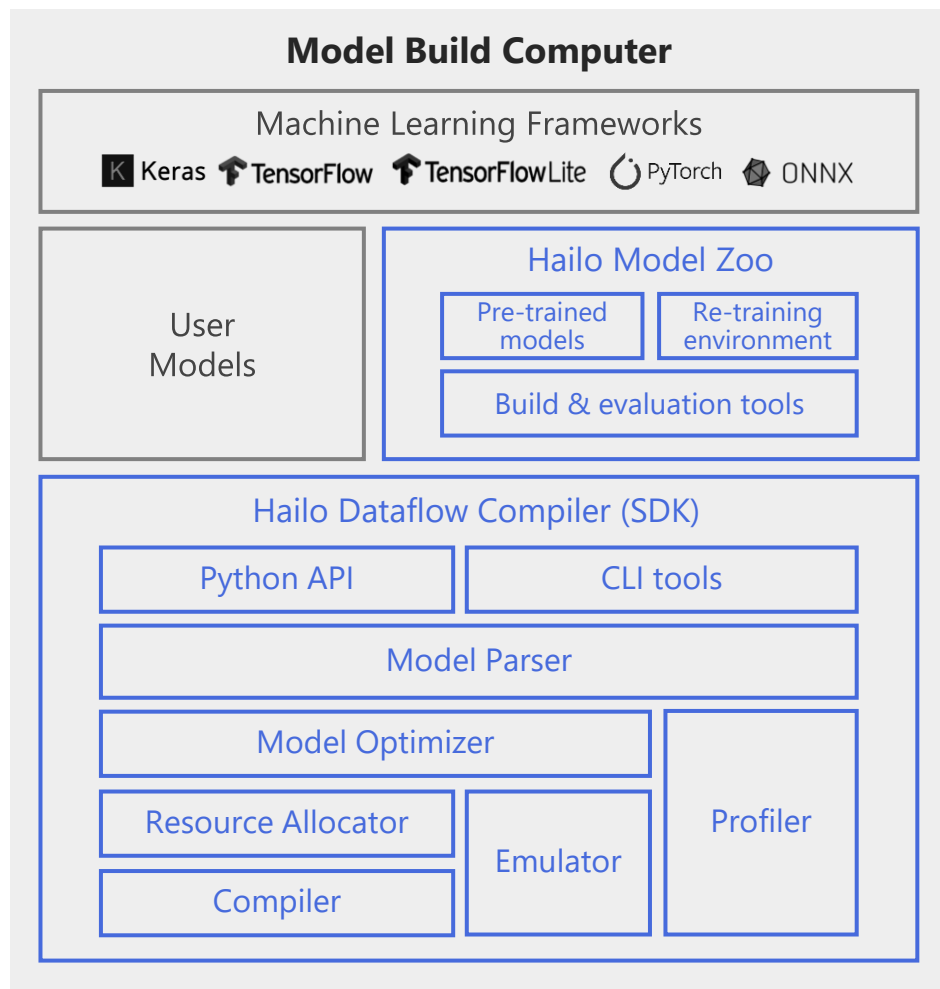


## Runtime Environment

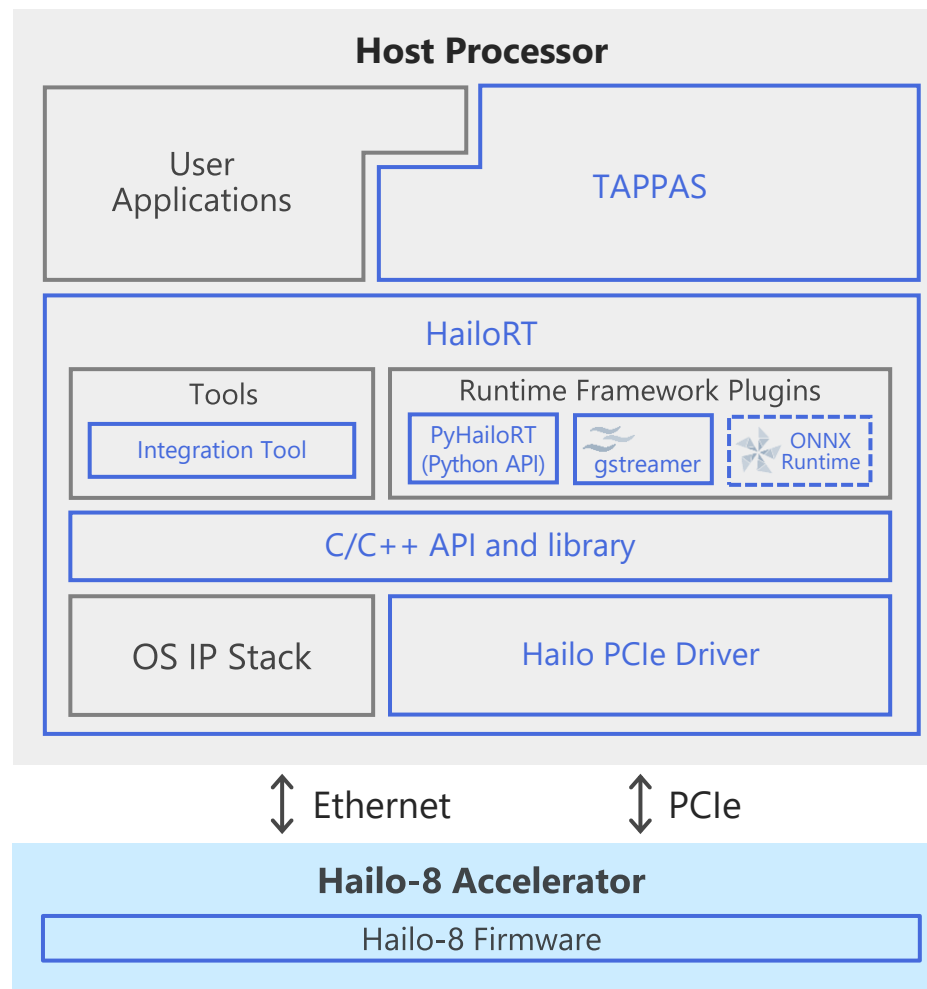
- Application examples, implementing pipeline elements and **pre-trained** AI tasks
- An inference **run-time library** with intuitive **API** for optimized performance (C/C++/Python)
- NN Core is part of Vision Processor or AI Accelerator

# Hailo AI Software Suite for AI Inference Accelerators

## Model Build Environment



## Runtime Environment



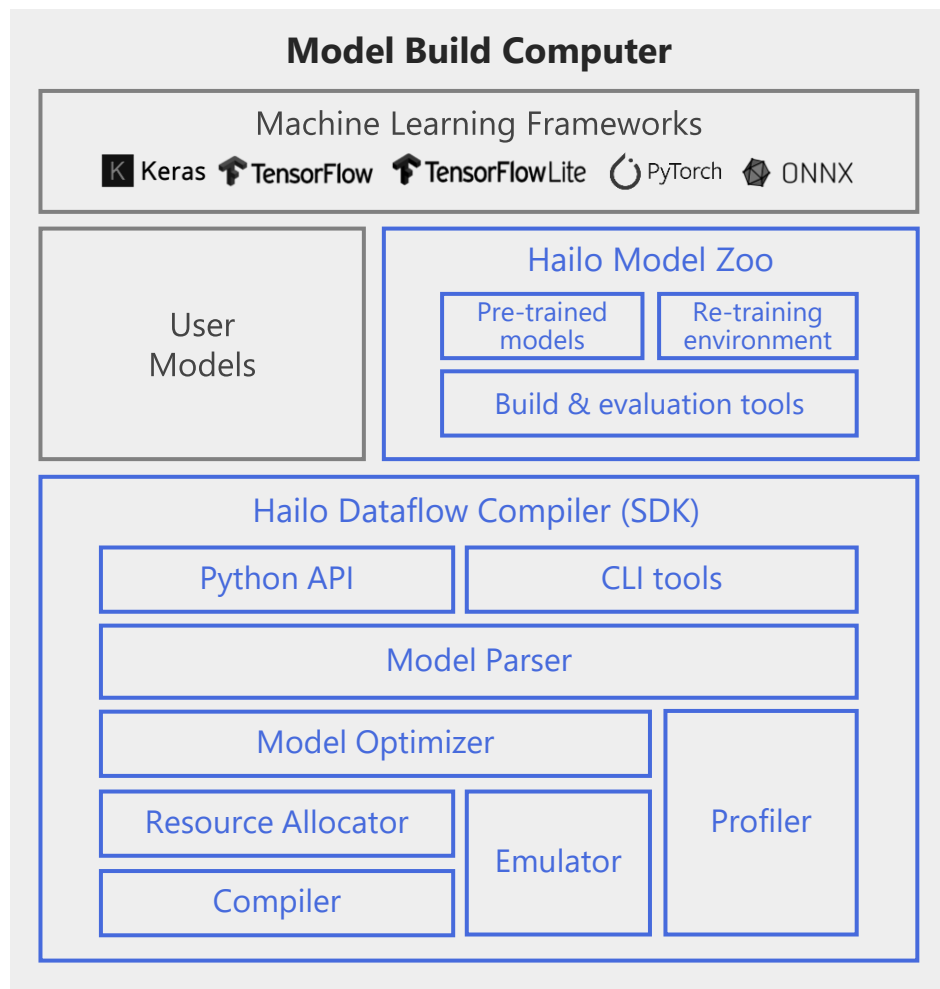
Hailo SW components

In preview

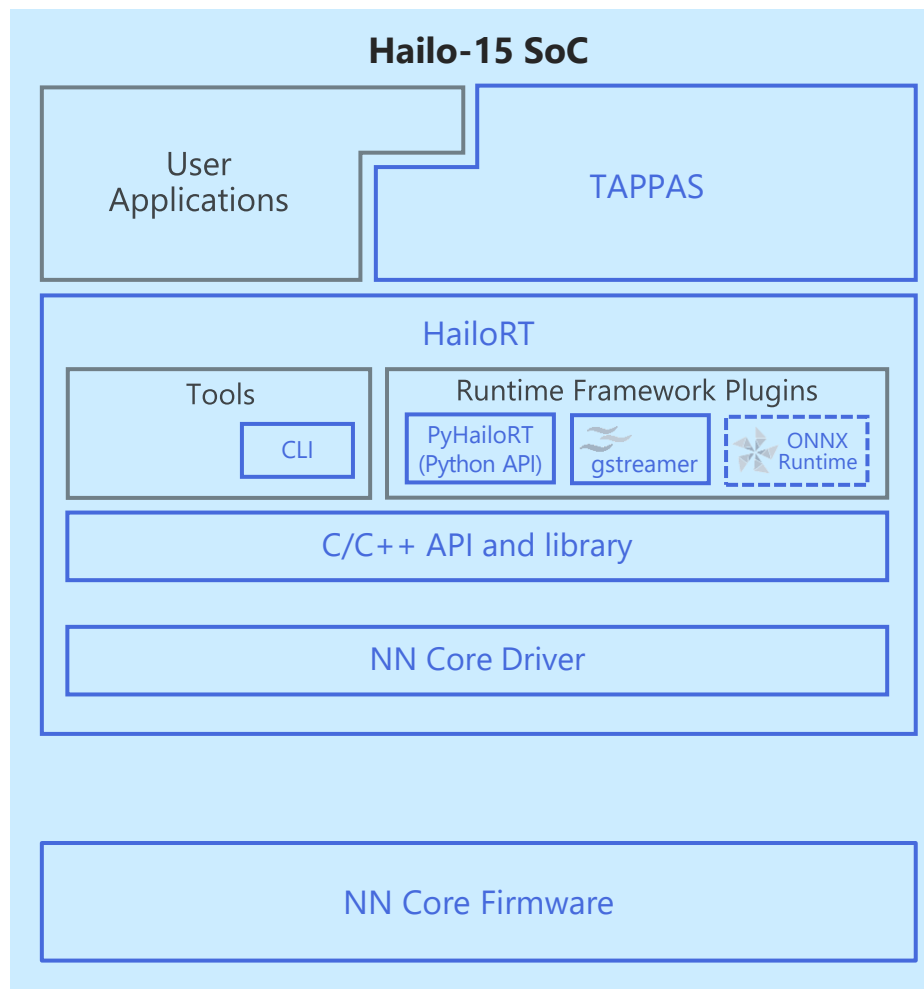
Other SW components

# Hailo AI Software Suite for AI Vision Processors

## Model Build Environment



## Runtime Environment

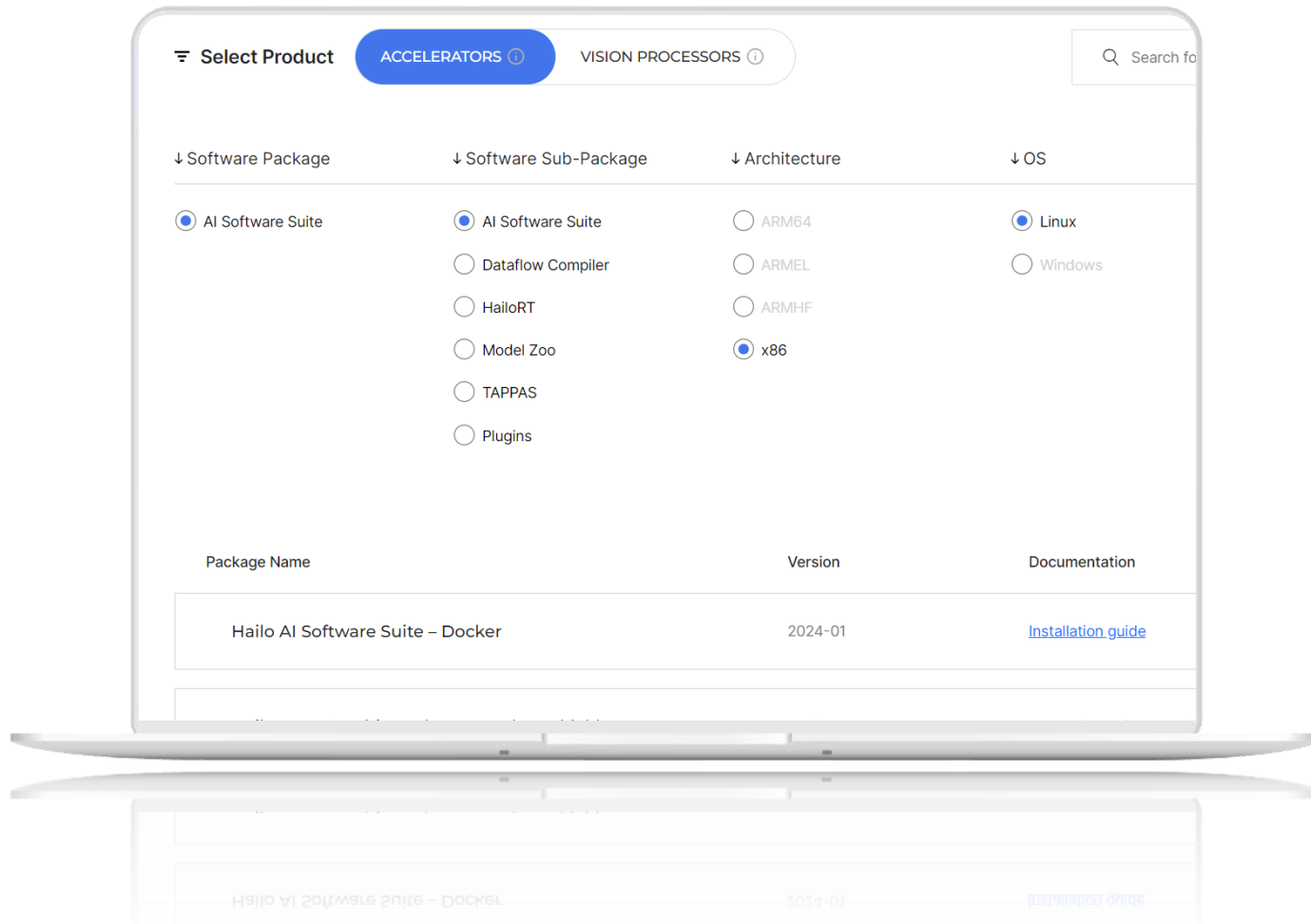


Hailo SW components

In preview

Other SW components

# Software Available in hailo.ai/developer-zone



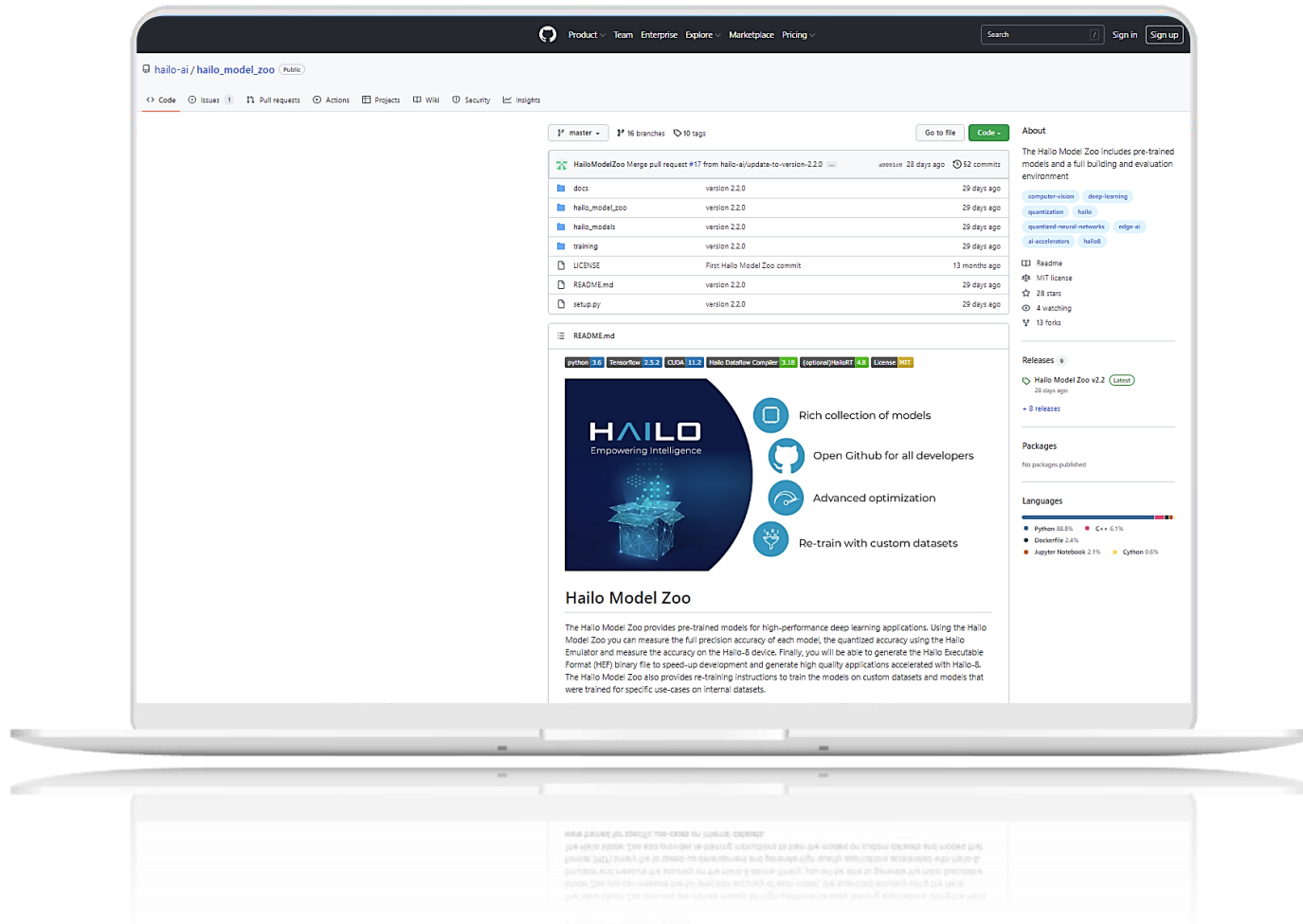
Hailo AI Software Suite and its components are available for download\* at:

<https://hailo.ai/developer-zone/sw-downloads/>

Selection is now available:

- Entire AI Software Suite or selected components
- Architecture
- Operating System
- Python version

# And in [github.com/hailo-ai](https://github.com/hailo-ai)



For example:

[https://github.com/hailo-ai/hailo\\_model\\_zoo](https://github.com/hailo-ai/hailo_model_zoo)

Available in open source:

- hailo\_model\_zoo
- tappas
- hailort-drivers
- hailort



# Thank you.

 Hailo.ai

 [contact@hailo.ai](mailto:contact@hailo.ai)