

# RZ/V2MA DRP-AI PERFORMANCE REPORT

REV.7.20  
SEPTEMBER 2022

R11AN0629EJ0720

# DRP-AI PROCESSING TIME

No.	Model structure	Processing time [msec]	Note
1	resnet18	15.2	Pre-trained model : <b>Torchvision 0.13.1 ResNet</b> <a href="https://pytorch.org/hub/pytorch_vision_resnet/">https://pytorch.org/hub/pytorch_vision_resnet/</a>
2	resnet34	24.9	References : ResNet <a href="#">Deep residual learning for image recognition</a>
3	resnet50	35.3	
4	MobileNetV1	15.2	
5	mobilenet_v2	19.0	Pre-trained model : <b>Torchvision 0.13.1 MobileNet V2</b> <a href="https://pytorch.org/hub/pytorch_vision_mobilenet_v2/">https://pytorch.org/hub/pytorch_vision_mobilenet_v2/</a>
6	YOLOv3	143.5	Pre-trained model : <b>Darknet YOLOv3/Tiny YOLOv3 (COCO)</b> <a href="https://pjreddie.com/darknet/yolo/">https://pjreddie.com/darknet/yolo/</a>
7	Tiny YOLOv3	22.8	References : YOLOv3 <a href="#">YOLOv3: An Incremental Improvement</a>
8	YOLOv2	72.9	Pre-trained model : <b>Darknet YOLOv2/Tiny YOLOv2 (VOC)</b> <a href="https://pjreddie.com/darknet/yolov2/">https://pjreddie.com/darknet/yolov2/</a>
9	Tiny YOLOv2	26.0	References : YOLOv2 <a href="#">YOLO9000: Better, Faster, Stronger</a>
10	HRNet	59.8	Pre-trained model : <b>MMPose HRNet (COCO)</b> <a href="https://mmpose.readthedocs.io/en/latest/papers/algorithms.html">https://mmpose.readthedocs.io/en/latest/papers/algorithms.html</a> References : HRNet <a href="#">Deep High-Resolution Representation Learning for Human Pose Estimation</a>
11	DeepLabV3	322.7	Pre-trained model: <b>Torchvision 0.13.1 DeepLabV3</b> <a href="https://pytorch.org/hub/pytorch_vision_deeplabv3_resnet101/">https://pytorch.org/hub/pytorch_vision_deeplabv3_resnet101/</a> References : DeepLabV3 <a href="#">Rethinking Atrous Convolution for Semantic Image Segmentation</a>

This processing time includes pre-processing and post-processing by DRP-AI.  
CPU pre-processing and post-processing time is not included.

\* Next page for detail of DRP-AI processing.

Above processing time is measured under following conditions.

- RZ/V2MA Linux Package v1.0.0
- RZ/V2MA DRP-AI Support Package v7.20  
Time measurement: C++ timespec\_get() function
- DRP-AI Translator v1.80

# DETAIL OF DRP-AI PROCESSING (RESNET)

DRP-AI processing includes pre-processing and post-processing as follows.

## resnet18/34/50

Resize

Cast to FP16

Normalize

Inference  
(resnet)

Softmax

Sort

Labeling



<Pre-processing>

**Resize** : Input image size(640x480 BGR) to HWC tensor size(224x224x3).

**Cast to fp16** : Cast to fp16 from int8 for DRP-AI processing.

**Normalize**: Normalized using 'mean' and 'std'.



<Post-processing>

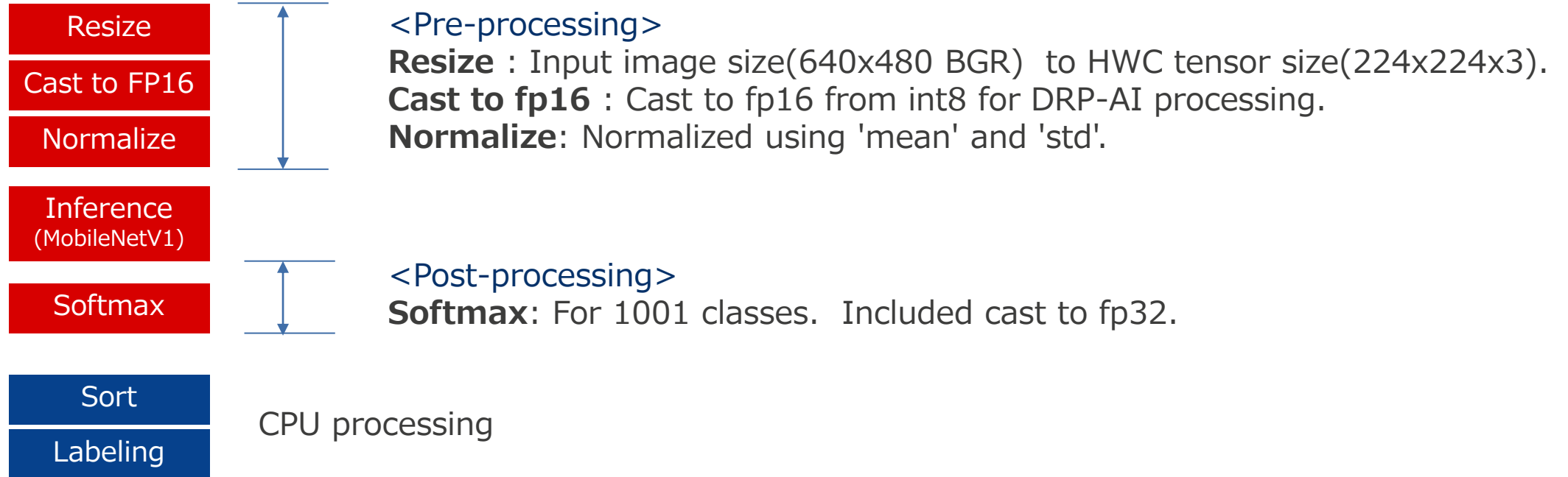
**Softmax**: For 1000 classes. Included cast to fp32.

CPU processing

# DETAIL OF DRP-AI PROCESSING (MOBILENETV1)

DRP-AI processing includes pre-processing and post-processing as follows.

## MobileNetV1



# DETAIL OF DRP-AI PROCESSING (MOBILENETV2)

DRP-AI processing includes pre-processing and post-processing as follows.

## mobilenet\_v2

Resize

Cast to FP16

Normalize

Inference  
(mobilenet\_v2)

Softmax

Sort

Labeling

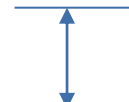


<Pre-processing>

**Resize** : Input image size(640x480 BGR) to HWC tensor size(224x224x3).

**Cast to fp16** : Cast to fp16 from int8 for DRP-AI processing.

**Normalize**: Normalized using 'mean' and 'std'.



<Post-processing>

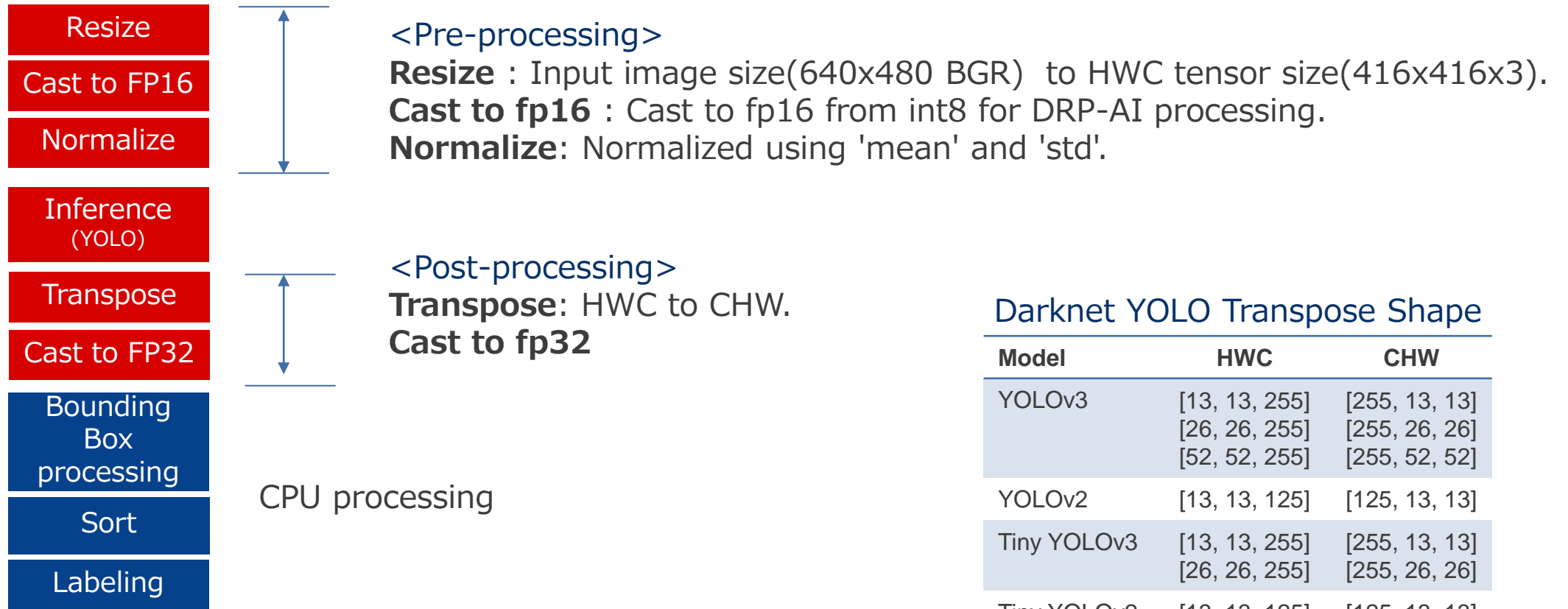
**Softmax**: For 1000 classes. Included cast to fp32.

CPU processing

# DETAIL OF DRP-AI PROCESSING (YOLO)

DRP-AI processing includes pre-processing and post-processing as follows.

## YOLOv3/YOLOv2/Tiny YOLOv3/Tiny YOLOv2



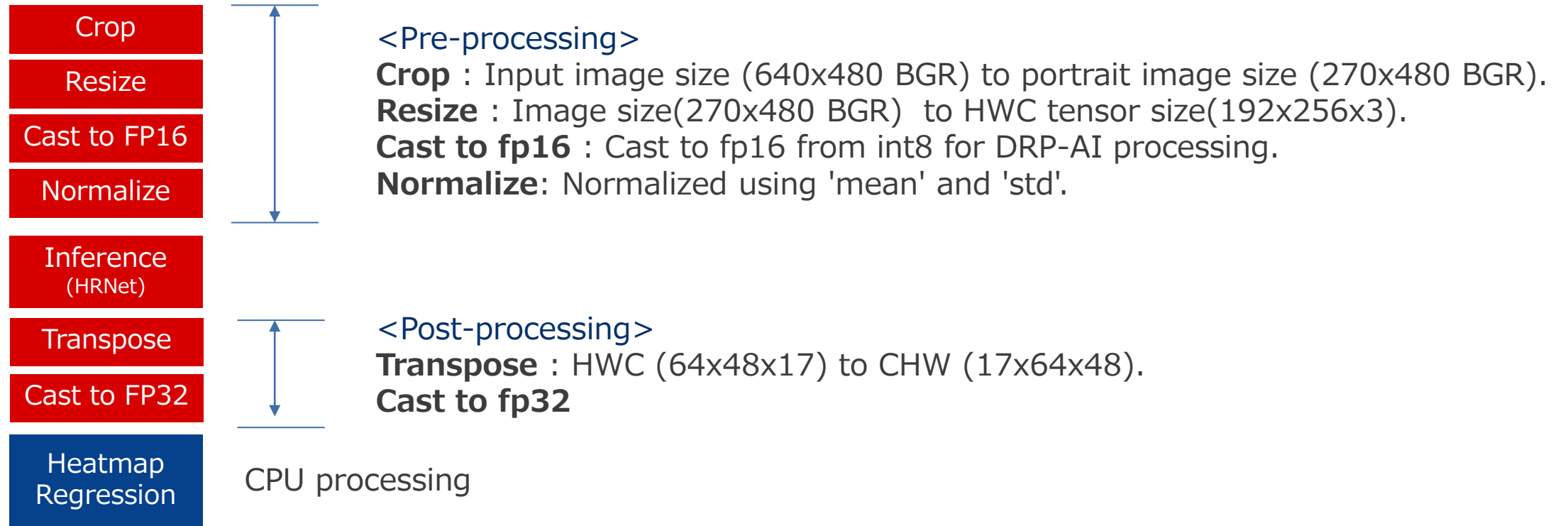
Darknet YOLO Transpose Shape

Model	HWC	CHW
YOLOv3	[13, 13, 255] [26, 26, 255] [52, 52, 255]	[255, 13, 13] [255, 26, 26] [255, 52, 52]
YOLOv2	[13, 13, 125]	[125, 13, 13]
Tiny YOLOv3	[13, 13, 255] [26, 26, 255]	[255, 13, 13] [255, 26, 26]
Tiny YOLOv2	[13, 13, 125]	[125, 13, 13]

# DETAIL OF DRP-AI PROCESSING (HRNET)

DRP-AI processing includes pre-processing and post-processing as follows.

## HRNet



# DETAIL OF DRP-AI PROCESSING (DEEPLABV3)

DRP-AI processing includes pre-processing and post-processing as follows.

## DeepLabV3

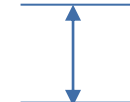
Resize

Cast to FP16

Normalize

Inference  
(DeepLabV3)

Argmax



<Pre-processing>

**Resize** : Input image size(640x480 BGR) to HWC tensor size(234x416x3).

**Cast to fp16** : Cast to fp16 from int8 for DRP-AI processing.

**Normalize**: Normalized using 'mean' and 'std'.

<Post-processing>

**Argmax**: HWC (234x416x21) to HWC (234x416x1). Includes cast to uint8.



---

[Renesas.com](https://www.renesas.com)