Specifications (ONNX OPs)

Operators	Parameters		Supported Status	CHIP						Note
	Name	Value		52x	32x	56x	336	530	331	
Abs	NA	NA	supported (Tool)	v	v	v	v	v	v	no other limitation
Acos			not supported	x	x	x	x	x	x	
Acosh			not supported	x	x	x	x	x	x	RT : Square->Sub->Sqrt->Add->Log
Add	NA	NA	supported (Tool)	v	v	v	v	v	v	no other limitation
And			not supported	x	x	x	x	x	x	
ArgMax			not supported	x	x	x	x	x	x	
ArgMin			not supported	x	x	x	x	x	x	
Asin			not supported	x	x	x	x	x	x	
Asinh			not supported	x	x	x	x	x	x	RT : Square->Add->Sqrt->Add->Log
Atan			not supported	x	x	x	x	x	x	
Atanh			not supported	x	x	x	x	x	x	RT : Sub->Add->Div->Log->Div
AveragePool	auto_pad : string (default is NOTSET)	"NOTSET" "SAME_UPPER" "SAME_LOWER" "VALID"	supported (Tool) not supported not supported not supported	v	v	v	v	v	v	1. The input size should satisfy the following conditions : global pooling : width <= 1020 height <= 1020 local pooling : width + pad_left + pad_right - kerl_size_w > 0 height + pad_top + pad_bottom - kerl_size_h > 0 ========== HW native spec : kernel={2x2,3x3} && pad={0x0,1x1} && stride={1x1,2x2} kernel={4x4,5x5} && pad={0x0,2x2} && stride={1x1,2x2} when countIncludePad=0 and pad!=0, the pool will turn to CPU. HW supported spec by splitting: kernel={4x4} && pad={0x0} && stride={4x4} kernel={8x8} && pad={0x0} && stride={8x8} ========== CPU: Except for the hw spec above, the pool will turn to CPU
	ceil_mode : int (default is 0)	0/1	supported (Tool) 0 / 1							
	count_include_pad : int (default is 0)		supported (Tool) 1							
	kernel_shape : list of ints (required)		supported (Tool) 2 ~ 255							
	pads : list of ints		supported (Tool) 1 ~ 255							
	strides : list of ints		supported (Tool) 1 ~ 255							
BatchNormalization	epsilon : float (default is 1e-05)		supported (Tool)	v	v	v	v	v	v	no other limitation
	momentum : float (default is 0.9)		not used when inference							
	training_mode : int (default is 0)		supported (Tool) 0							
Bernoulli			For sampling, not supported	x	x	x	x	x	x	
BitShift			not supported	x	x	x	x	x	x	
Cast			For type conversion, not supported	x	x	x	x	x	x	
CastLike			For type conversion, not supported	x	x	x	x	x	x	
Ceil			not supported	x	x	x	x	x	x	
Celu			not supported	x	x	x	x	x	x	
Clip	NA	NA	supported (Tool)	v	v	v	v	v	v	no other limitation
Compress			not supported	x	x	x	x	x	x	
Concat	axis : int (required)		supported (Tool) 0 ~ 3	v	v	v	v	v	v	not real layer
ConcatFromSequence			not supported	x	x	x	x	x	x	
Constant			For constant generation, not supported	x	x	x	x	x	x	
ConstantOfShape			For constant generation, not supported	x	x	x	x	x	x	
Conv	auto_pad : string (default is NOTSET)	"NOTSET" "SAME_UPPER" "SAME_LOWER" "VALID"	supported (Tool) supported (Tool) supported (Tool) supported (Tool)	v	v	v	v	v	v	1. Weight and bias should be initialized tensor. 2. The input size should satisfy the following conditions : width + pad_left + pad_right - kerl_size_w > 0 height + pad_top + pad_bottom - kerl_size_h > 0 ========== HW (satisfy one of the following conditions) 1、dilations = {1,1} 2、dilations = {2,2} && group=1 && (stride= {1,1} || stride ={2,2}) ========== CPU 1、dilations != {1,1}
	dilations : list of ints		supported (Tool) 1 supported (NT98530 DSP) 2/3/4/12/24/36 @ kernel_shape=3, stride=1							
	group : int (default is 1)		supported (Tool) 1 ~ no limitation							
	kernel_shape : list of ints		supported (Tool) 1 ~ 11							
	pads : list of ints		supported (Tool) 0 ~ no limitation							
	strides : list of ints		supported (Tool) 0 ~ no limitation							
ConvInteger			For type conversion, not supported	x	x	x	x	x	x	
ConvTranspose	auto_pad : string (default is NOTSET)	"NOTSET" "SAME_UPPER" "SAME_LOWER" "VALID"	supported (Tool) not supported not supported not supported	v	v	v	v	v	v	1. Weight and bias should be initialized tensor. 2. Only support 4-dimension input. 3. Kernel and stride should satisfy the following combination k_w*k_h stride 11*11 1, 2, 4, 8 9*9 1, 2, 4, 8 7*7 1, 2, 4, 8 5*5 1, 2, 4, 8 4*4 1, 2, 4, 8 3*3 1, 2, 4, 8 2*2 1, 2, 4, 8 1*1 1, 2, 4, 8 1*7 1, 2, 4, 8 7*1 1, 2, 4, 8 6*1 4 1*5 1, 2, 4, 8 5*1 1, 2, 4, 8 1*3 1, 2, 4, 8 3*1 1, 2, 4, 8
	dilations : list of ints		supported (Tool) 1							
	group : int (default is 1)		supported (Tool) 1 ~ no limitation supported (NT98530 DSP) 1 or group=input channel							
	kernel_shape : list of ints		supported (Tool) 1 ~ 11 supported (NT98530 DSP) 1 ~ 100							
	output_padding : list of ints		only supported on (NT98530 DSP) 1 ~ 2^32-1							
	output_shape : list of ints		only supported on (NT98530 DSP) 1 ~ 65535							
	pads : list of ints		supported (Tool) 0 ~ no limitation							
	strides : list of ints		supported (Tool) 1/2/4/8 supported (NT98530 DSP) 1 ~ 2^32-1							
Cos			not supported	x	x	x	x	x	x	
Cosh			not supported	x	x	x	x	x	x	RT : Exp->Pow->Add->Div
CumSum			not supported	x	x	x	x	x	x	
DepthToSpace			not supported	x	x	x	x	x	x	
DequantizeLinear			For type conversion, not supported	x	x	x	x	x	x	
Det			not supported	x	x	x	x	x	x	
Div	NA	NA	supported (ACL)	v	v	x	v	v	x	1. CPU layer 2. supported (NT98530 DSP)
Dropout			not supported	x	x	x	x	x	x	
DynamicQuantizeLinear			not supported	x	x	x	x	x	x	
Einsum			not supported	x	x	x	x	x	x	
Elu			not supported	x	x	x	x	x	x	
Equal	NA	NA	not supported	x	x	x	x	x	x	
Erf			For error handling, not supported	x	x	x	x	x	x	
Exp	NA	NA	supported (ACL)	v	v	x	v	v	x	1. CPU layer 2. supported (NT98530 DSP)
Expand			not supported	x	x	x	x	x	x	
EyeLike			For sampling, not supported	x	x	x	x	x	x	
Flatten	axis : int (default is 1)		supported (Tool) 1	v	v	v	v	v	v	not real layer
Floor	NA	NA	supported (ACL)	v	v	x	v	v	x	1. CPU layer
Function			For processing, not supported	x	x	x	x	x	x	
Gather			not supported	x	x	x	x	x	x	
GatherElements			not supported	x	x	x	x	x	x	
GatherND			not supported	x	x	x	x	x	x	
Gemm	alpha : float (default is 1.0)		supported (Tool) 1	v	v	v	v	v	v	1. Only support 2-dimensions
	beta : float (default is 1.0)		supported (Tool) 1							
	transA : int (default is 0)		supported (Tool) 0/1							

	transB : int (default is 0)		supported (Tool) 0/1							
GlobalAveragePool	NA	NA	supported (Tool)	v	v	v	v	v	v	1. The input size should satisfy the following conditions : width <= 1020 height <= 1020
GlobalLpPool			not supported	x	x	x	x	x	x	
GlobalMaxPool	NA	NA	supported (Tool)	v	v	v	v	v	v	1. The input size should satisfy the following conditions : width <= 1020 height <= 1020
Greater			not supported	x	x	x	x	x	x	
GreaterOrEqual			not supported	x	x	x	x	x	x	
GridSample			For sampling, not supported	x	x	x	x	x	x	
GRU			not supported	x	x	x	x	x	x	
Hardmax			not supported	x	x	x	x	x	x	
HardSigmoid			not supported	x	x	x	x	x	x	
HardSwish			not supported	x	x	x	x	x	x	
Identity			not supported	x	x	x	x	x	x	
If			not supported	x	x	x	x	x	x	
InstanceNormalization			not supported	x	x	x	x	x	x	
IsInf			not supported	x	x	x	x	x	x	
IsNaN			not supported	x	x	x	x	x	x	
LayerNormalization	axis: int		0~5	x	x	x	x	v	x	1. only supported on (NT98530 DSP) 2. axis supports [width, height, channel, batch, time]
	epsilon: float		-2^31 ~ 2^31-1	x	x	x	x	v	x	
	stash_type: int		not supported	x	x	x	x	x	x	
LeakyRelu	alpha : float (default is 0.01)		supported (Tool) -16 ~ 15.99 (NT98530, NT98331) -8 ~ 7.99 (Others)	v	v	v	v	v	v	no other limitation
Less			not supported	x	x	x	x	x	x	
LessOrEqual			not supported	x	x	x	x	x	x	
Log	NA	NA	supported (ACL)	v	v	x	v	v	x	1. CPU layer
LogSoftmax			not supported	x	x	x	x	x	x	
Loop			For conditional operation, not supported	x	x	x	x	x	x	
LpNormalization			not supported	x	x	x	x	x	x	
LpPool			not supported	x	x	x	x	x	x	
LRN			not supported	x	x	x	x	x	x	
LSTM	activation_alpha : list of floats		not supported	v	v	v	v	v	v	1. The input size should satisfy the following conditions : width * height <= 51200
	activation_beta : list of floats		not supported							
	activations : list of strings		not supported							
	clip : float		not supported							
	direction : string (default is forward)	"forward" "reverse" "bidirectional"	supported (Tool) not supported not supported							
	hidden_size : int		supported (Tool)							
	input_forget : int (default is 0)		supported (Tool) 0							
	layout : int (default is 0)	0/1	supported (Tool) 0							
MatMul	NA	NA	supported (Tool)	v	v	v	v	v	v	1. The input size should satisfy the following conditions : channel : 1 ~ 2048
MatMulInteger			For type conversion, not supported	x	x	x	x	x	x	
Max	NA	NA	supported (Tool)	v	v	v	v	v	v	no other limitation
MaxPool	auto_pad : string (default is NOTSET)	"NOTSET" "SAME_UPPER" "SAME_LOWER" "VALID"	supported (Tool) not supported not supported not supported	v	v	v	v	v	v	1. The input size should satisfy the following conditions : global pooling : width <= 1020 height <= 1020 local pooling : width + pad_left + pad_right - kerl_size_w > 0 height + pad_top + pad_bottom - kerl_size_h > 0 ========== HW native spec : kernel={2x2,3x3} && pad={0x0,1x1} && stride={1x1,2x2} kernel={4x4,5x5} && pad={0x0,2x2} && stride={1x1,2x2} HW supported spec by splitting: kernel={4x4} && pad={0x0} && stride={4x4} kernel={8x8} && pad={0x0} && stride={8x8} kernel={4x4,5x5} && pad={1x1} && stride={1x1,2x2} kernel_h==kernel_w && pad_h==pad_w && pad<kernel/2 && kernel>=6 && stride={1x1,2x2} ========== CPU : Except for the hw spec above, the pool will turn to CPU
	ceil_mode : int (default is 0)	0/1	supported (Tool) 0/1							
	dilations : list of ints		supported (Tool) 1							
	kernel_shape : list of ints (required)		supported (Tool) 1 ~ 255							
	pads : list of ints		supported (Tool) 0 ~ 255							
	storage_order : int (default is 0)	0/1	supported (Tool) 0							
	strides : list of ints		supported (Tool) 1 ~ 255							
MaxRoiPool	pooled_shape : list of ints (required)		supported (Tool) 6/7/14	v	v	x	v	v	v	1. 56x is not support 2. roi_num = 1~64 3. width = 1~128 4. height = 1~128 5. channel = 1~2048 6. pool_h & pool_w should be the same 7. There are two usages of single batch and multi-batch - Single batch: multiple ROIs can be set, only static settings (the number of ROIs are set in prototxt) - Multi-batch: not support ROI [batch idx, x1, y1, x2, y2] can be set arbitrarily for any batch idx, the input feature batch and ROI are in a one-to-one relationship, one batch contains one ROI, so the number of batch of input feature = the number of ROI = batch of output feature 8. the input type of ROI should be UINT16 feature-in.
	spatial_scale : float (default is 1.0)		supported (Tool)							
MaxUnpool			not supported	x	x	x	x	x	x	
Mean			not supported	x	x	x	x	x	x	
MeanVarianceNormalization			not supported	x	x	x	x	x	x	
Min			not supported	x	x	x	x	x	x	
Mod			not supported	x	x	x	x	x	x	
Mul	NA	NA	supported (Tool)	v	v	v	v	v	v	no other limitation
Multinomial			For sampling, not supported	x	x	x	x	x	x	
Neg			not supported	x	x	x	x	x	x	
NegativeLogLikelihoodLoss			For training, not supported	x	x	x	x	x	x	
NonMaxSuppression			not supported	x	x	x	x	x	x	
NonZero			not supported	x	x	x	x	x	x	
Not			not supported	x	x	x	x	x	x	
OneHot			For sampling, not supported	x	x	x	x	x	x	
Optional			For conditional operation, not supported	x	x	x	x	x	x	
OptionalGetElement			For conditional operation, not supported	x	x	x	x	x	x	
OptionalHasElement			For conditional operation, not supported	x	x	x	x	x	x	
Or			not supported	x	x	x	x	x	x	
Pad	mode : string (default is constant)	"constant" "reflect" "edge"	supported (Tool) not supported not supported	v	v	v	v	v	v	no other limitation
Pow	NA	NA	supported (ACL)	v	v	x	v	v	x	1. CPU layer 2. not support negative input when exp is not 1 3. supported (NT98530 DSP)
PRelu	NA	NA	supported (Tool)	v	v	v	v	v	v	no other limitation
QLinearConv			For type conversion, not supported	x	x	x	x	x	x	
QLinearMatMul			For type conversion, not supported	x	x	x	x	x	x	
QuantizeLinear			For type conversion, not supported	x	x	x	x	x	x	
RandomNormal			For sampling, not supported	x	x	x	x	x	x	
RandomNormalLike			For sampling, not supported	x	x	x	x	x	x	
RandomUniform			For sampling, not supported	x	x	x	x	x	x	
RandomUniformLike			For sampling, not supported	x	x	x	x	x	x	
Range			For conditional operation, not supported	x	x	x	x	x	x	
Reciprocal			not supported	x	x	x	x	x	x	
ReduceL1			not supported	x	x	x	x	x	x	RT : Abs->Sum
ReduceL2			not supported	x	x	x	x	x	x	ST : L2Norm
ReduceLogSum			not supported	x	x	x	x	x	x	RT : Sum->Log
ReduceLogSumExp			not supported	x	x	x	x	x	x	RT : Exp->Sum->Log
ReduceMax			not supported	x	x	x	x	x	x	

ReduceMean	axes : list of ints		supported (Tool)							1. input, axes, keepdim need to be the following combination (x means any value) : input axes keepdim xxxx 2+3 0,1 xxxx 2+2 0 xxxx 1+2+3 0,1 xxxx 3 0,1 (should satisfy channel*height<2048) xxx 2 0,1 xxxx 1 0,1 xxx 1 0,1 xx 1 0,1 2. input size should satisfy the following conditions: height <= 1020 width <= 1020
	keepdims : int (default is 1)	0/1	supported (Tool) 0/1	v	v	v	v	v	v	
ReduceMin			not supported	x	x	x	x	x	x	
ReduceProd			not supported	x	x	x	x	x	x	
ReduceSum			not supported	x	x	x	x	x	x	
ReduceSumSquare			not supported	x	x	x	x	x	x	RT : Sqrt->Sum
Relu	NA	NA	supported (Tool)	v	v	v	v	v	v	no other limitation
Reshape	allowzero : int (default is 0)		supported (Tool)	v	v	v	v	v	v	1. do not support input/output dimension size > 4
Resize	antialias : int (default is 0)	If set to 1, "linear" and "cubic" interpolation modes will use an antialiasing filter when downscaling	not supported							no other limitation
	axes : list of ints		not supported							
	coordinate_transformation_mode : string (default is half_pixel)	"half_pixel" "align_corners" "asymmetric" "tf_crop_and_resize"	supported (ACL) supported (ACL) supported (Tool) - asymmetric is only supported under (scale=1/2/4/8 && mode=nearest && nearest_mode=floor) not supported							
	cubic_coeff_a : float (default is -0.75)	coefficient 'a' used in cubic interpolation	not supported							
	exclude_outside : int (default is 0)		not supported							
	extrapolation_value : float (default is 0.0)		not supported	v	v	v	v	v	v	
	keep_aspect_ratio_policy : string (default is stretch)	"stretch" "not_larger" "not_smaller"	not supported not supported not supported							
	mode : string (default is nearest)	"nearest" (default), "linear" "cubic"	supported (ACL) (Tool) supported (ACL) not supported							
	nearest_mode : string (default is round_prefer_floor)	"round_prefer_floor" (default), "round_prefer_ceil" "floor", "ceil"	not supported not supported supported (ACL) (Tool) not supported							
ReverseSequence			not supported	x	x	x	x	x	x	
RNN			not supported	x	x	x	x	x	x	
RoiAlign			not supported	x	x	x	x	x	x	
Round	NA	NA	supported (ACL)	v	v	x	v	v	x	1. CPU layer
Scan			For processing, not supported	x	x	x	x	x	x	
Scatter (deprecated)			Deprecated, not supported	x	x	x	x	x	x	
ScatterElements			not supported	x	x	x	x	x	x	
ScatterND			not supported	x	x	x	x	x	x	
Selu			not supported	x	x	x	x	x	x	
SequenceAt			not supported	x	x	x	x	x	x	
SequenceConstruct			not supported	x	x	x	x	x	x	
SequenceEmpty			not supported	x	x	x	x	x	x	
SequenceErase			not supported	x	x	x	x	x	x	
SequenceInsert			not supported	x	x	x	x	x	x	
SequenceLength			not supported	x	x	x	x	x	x	
SequenceMap			not supported	x	x	x	x	x	x	
Shape			not supported	x	x	x	x	x	x	
Shrink			not supported	x	x	x	x	x	x	
Sigmoid	NA	NA	supported (Tool)	v	v	v	v	v	v	no other limitation
Sign			not supported	x	x	x	x	x	x	
Sin	NA	NA	supported (ACL)	v	v	x	v	v	x	1. CPU layer
Sinh			not supported	x	x	x	x	x	x	RT : Exp->Pow->Sub->Div
Size			not supported	x	x	x	x	x	x	
Slice	NA	NA	supported (Tool)	v	v	v	v	v	v	not real layer
Softmax	axis : int (default is -1)		supported (Tool) axis: -1/1/2/3 supported (ACL) axis: 0~3 supported (NT98530 DSP) axis: 1~3	v	v	v	v	v	v	HW native spec a) axis = 1 (all chip) width = 1 ~ 1023 height = 1 ~1023 channel = 2 ~ 128 width*height <= (1<<19)-1 b) axis = 3 (530/336) width = 2 ~ 2048 height = 1 ~ 1023 channel = 1 ~ 2047 HW supported spec by splitting (using [platform/softmax] = 1 under 530/336 with the following limitation) : 1. permute(0,2,3,1 or 0,1,3,2) will be added before or after softmax to adjust dimensions, height or width of the permutes cannot exceed hw limit. ========== CPU In the cases above: Softmax enabling postprocEn will turn to CPU. When the dimension size of axis exceeds HW spec, it will turn to CPU Except for the hw spec above, it will turn to CPU.
SoftmaxCrossEntropyLoss			For training, not supported	x	x	x	x	x	x	
Softplus			not supported	x	x	x	x	x	x	
Softsign			not supported	x	x	x	x	x	x	
SpaceToDepth			not supported	x	x	x	x	x	x	
Split	axis : int (default is 0)		supported (Tool)	v	v	v	v	v	v	not real layer
SplitToSequence			not supported	x	x	x	x	x	x	
Sqrt	NA	NA	supported (ACL)	v	v	x	v	v	x	1. CPU layer 2. supported (NT98530 DSP)
Squeeze			not supported	x	x	x	x	x	x	
StringNormalizer			For processing, not supported	x	x	x	x	x	x	
Sub	NA	NA	supported (Tool)	v	v	v	v	v	v	no other limitation
Sum			not supported	x	x	x	x	x	x	
Tan			not supported	x	x	x	x	x	x	
Tanh	NA	NA	supported (Tool)	v	v	v	v	v	v	no other limitation
TfidfVectorizer			not supported	x	x	x	x	x	x	
ThresholdedRelu			not supported	x	x	x	x	x	x	
Tile			not supported	x	x	x	x	x	x	
TopK			not supported	x	x	x	x	x	x	ST : Sorting
Transpose	perm : list of ints		supported (Tool)	v	v	v	v	v	v	NCHW → NHWC 1. width = 1~1023 2. height = 1~1023 3. channel = 1~4095 4. width * height <= 524288 when channel=1 or (width==1&&height==1), the permute will be converted to reshape. NCHW → NCWH 1. width = 1~1023 2. height = 1~1023 3. channel = 1~4095 when height==1 or width==1, the permute will be converted to reshape. The order can be transformed by the combination of ①NCHW → NHWC and ②NCHW → NCWH NCHW → NHCW (is equal to ① → ②) NCHW → NWHC (is equal to ② → ①) NCHW → NWCH (is equal to ① → ①) Through the combined order, each permute needs to meet its limitation. ========== Other order (CPU layer)

										1. width = 1 ~ 65535 2. height = 1 ~ 65535 3. channel = 1 ~ 65535 4. batch = 1 ~ 65535
Trilu			not supported	x	x	x	x	x	x	
Unique			not supported	x	x	x	x	x	x	
Unsqueeze			not supported	x	x	x	x	x	x	
Upsample (deprecated)	NA	NA	supported (Tool) 1x, 2x, 4x, 8x (nearest)	v	v	v	v	v	v	no other limitation
Where			not supported	x	x	x	x	x	x	
Xor			not supported	x	x	x	x	x	x	
Customer Layer			not supported	x	x	x	x	x	x	