Tengine Convert Tool Technical Spec

文档版本 1.1

发布日期 2021-04-02

OPEN AI LAB

变更记录

日期	版本	说明	作者
2020-12-04	1.0	初版	Zhang Bin
2021-04-02	1.1	新增 ONNX HardSwish OP 支持	Tang Qi

目录

1 产品介绍

1.1 背景与目的

Tengine Convert Tool 实现将常见的训练框架(Caffe，ONNX，MXNet，TensorFlow，TensorFlow Lite，Darknet)生成的网络模型转换为 Tengine 专属的网络模型存储文件 tmfile，仅支持 FP32 的网络模型转换。

1.2 产品特点

1) 此工具支持模型转化功能。

2 支持范围

2.1 硬件支持

2.2 操作系统支持

- Ubuntu 18.04 以上

2.3 算子支持

2.3.1 Tengine 算子支持

详见附录 1。

卷积计算方法包括：

- Direct Convolution
- Winograd Convolution
- Gemm Convolution

2.4 FP32 模型支持

2.4.1 Caffe 模型支持

Inception v3	**Inception v4**	**ResNet18**	**ResNet50**	**VGG16**
MobileNet v1	**MobileNet v2**	**MobileNet v3**	**ShuffleNet v2**	**SqueezeNet v1.1**
FasterRCNN	**MobileNetSSD**	**MTCNN**		

2.4.2 ONNX 模型支持

VGG16	**ResNet18**	**ResNet50**	**MobileNet v2**	**ShuffleNet v2**
SqueezeNet v1.1	**YOLO v5s**			

2.4.3 MXNet 模型支持

Inception v3	**VGG16**	**ResNet18**	**ResNet50**	**MobileNet v1**
MobileNet v2	**SqueezeNet v1.1**	**RetinaFaee**	**MobileFaceNets**	

2.4.4 TensorFlow 模型支持

Inception v3	**Inception v4**	**ResNet50**	**Mobilenet v1**	**Mobilenet v2**
SqueezNet v1.1	**DenseNet**			

2.4.5 TensorFlow Lite 模型支持

Inception v3	**Inception v4**	**ResNet v2**	**SqueezeNet v1.1**	**MobileNet v1**
MobileNet v2	**MobileNetSSD**			

2.4.6 Darknet 模型支持

YOLOv3	**YOLOv3 Tiny**	**YOLOv4**	**YOLOv4 Tiny**

附录 1 Tengine Convert Tool 支持算子列表

TENGINE	Caffe		MXNet		TensorFlow	TensorFlow Lite	ONNX	Dark net
ACCURACY	√							
BATCHNORMALIZATION	BatchNorm		BatchNorm		FusedBatchNorm		√	
		ComposedBN						
RESIZE						RESIZE_NEAREST_NEIGHBOR		
CONCAT	√	√	ConcatV2	CONCATENATION	√	route		
CONST								
CONVOLUTION	√	√	Conv2D	CONV_2D	Conv	convolutional		
		DepthwiseConvolution				DepthwiseConv2dNative	DEPTHWISE_CONV_2D	
	ConvolutionDepthwise							
DECONVOLUTION		√		√		Conv2DBackpropInput		
DETECTIONOUTPUT	√							
DROPOUT		√		Copy		√		√	yolo
ELTWISE	√	_minus_scalar	Add	ADD	Add	shortcut		
				_mul_scalar		Sub	SUB	Sub	
		elemwise_add		PROD				
						Rsqrt	RSQRT	
		_div_scalar	RealDiv	DIV	Div			
						Log	LOG	
			Exp	EXP	Exp			
						Pow	POW	
			Sqrt	SQRT				
						Floor	FLOOR	Floor	
			Mul	MUL	Mul			
						Minimum		
			AddN					
FLATTEN		√		√		√		√
FULLYCONNECTED	InnerProduct	√	MatMul	FULLY_CONNECTED	MatMul			
								Gemm
INPUT	Data		FIFOQueueV2					
		Input						
LRN	√		√					
NORMALIZE		√						
PERMUTE	√	transpose						

TENGINE	Caffe	MXNet	TensorFlow	TensorFlow Lite	ONNX	Darknet
POOLING	√	√	AvgPool	AVERAGE_POOL_2D	AveragePool	
			GlobalAveragePool			
			MaxPool	MAX_POOL_2D	MaxPool	maxpool
PRELU	√					
PRIORBOX	√					
REGION	√		region			
RELU	√	Activation	Relu		Relu	
	LeakyReLU		LeakyRelu			
RELU6	√	clip	Relu6			
REORG	√		reorg			
RESHAPE	√	√	√	RESHAPE	√	
ROIPOOLING	√					
RPN	√					
SCALE	√					
SLICE	√				√	
SOFTMAX	√	Activation √	SOFTMAX	√		
	SoftmaxWithLoss					
	SoftmaxOutput					
		SoftmaxActivation				
SPLIT	√	√	√			
DETECTIONPOSTPROCESS				TFLite_Detection_PostProcess		
GEMM						
GENERIC		DecodeWav				
			AudioSpectrogram			
		Mfcc				
LOGISTIC				LOGISTIC		
LSTM		RNN √				
RNN			√			
TANH	TanH	Activation √	√			
SIGMOID	√	Activation	√		√	
SQUEEZE		SQUEEZE √				
PAD			√			
		MirrorPad				
STRIDEDSLICE			√	STRIDED_SLICE		
REDUCTION	√	√	Sum	SUM		
			Mean	MEAN	ReduceMean	
		Asum				
			Sqsum			
		Max				
			Min			
		Prod				

TENGINE	Caffe	MXNet	TensorFlow	TensorFlow Lite	ONNX	Dark net
			L2			
	Logsum					
			Logsumexp			
ARGMAX	√					
ARGMIN			√			
TOPKV2	√					
MAXIMUM			√		Max	
MINIMUM	√					
ADDN		add_n				
SWAPAXIS	√					
GRU		RNN	√			
UPSAMPLE √	UpSampling		upsample			
SHUFFLECHANNEL	√					
RESIZE √	ResizeNearestNeighbor					
			ResizeBilinear			
SPACETOBATCHND	√					
BATCHTOSPACEND			√			
CROP √	√					
PSROIPOOLING		_contrib_PSROIPooling				
ROIALIGN	_contrib_ROIAlign					
EXPANDDIMS			ExpandDims			
UNARY	√					
		abs	Abs			
	neg	Neg				
		ceil	Ceil			
	floor	Floor				
		sin	Sin			
	Asin					
		cos	Cos			
	Acos					
		atan	Atan			
	tan	Tan				
	reciprocal	Reciprocal				
			Square			
	Sqrt					
			Rsqrt			
	Exp					
			Log			
BIAS √						
NOOP						
THRESHOLD √						
HARDSIGMOID						
EMBEDDING √	√	√				
INSTANCENORM		√				

TENGINE	Caffe	MXNet	TensorFlow	TensorFlow Lite	ONNX	Dark net
MVN	√					
ABSVAL	√					
CAST	√					
HARDSWISH					√	
INTERP	√	UpSampling	Upsample			
SELU						
ELU	√	LeakyReLU	ELU	√		
BROADMUL			broadcast_mul			
LOGICAL			LOGICALOR			
				LOGICALAND		
GATHER			GATHER	√		
TRANSPOSE				√	TRANSPOSE	√
COMPARISON		Equal	EQUAL			
				Greater	GREATER	
		GreaterEqual	GREATER_EQUAL			
				Less	LESS	
		LessEqual				
					LESS_GREATER	
SPACETODEPTH			SPACE_TO_DEPTH			
DEPTHTOSPACE					DEPTH_TO_SPACE	
REVERSE		ReverseV2	REVERSE_V2			
SPARSETODENSE				√	SPARSE_TO_DENSE	
CEIL		√	CEIL			
SQUAREDDIFFERENCE				√	SQUARED_DIFFERENCE	
ROUND		√	ROUND			
ZEROSLIKE						
CLIP	Clip			Clip		
POWER	Power					
TILE	Tile					
L2NORMALIZATION					L2_NORMALIZATION	
L2POOL			L2_POOL_2D			
RELU1					RELU_N1_TO_1	
LOGSOFTMAX			LOG_SOFTMAX			
FLOOR				Floor		
REDUCEL2				√		
UNSQUEEZE						√