

[← Back to Create Configuration](#)

Configurations

[Create](#)

Configurations Settings

#	Model	Validation Dataset	Target	Precisions	Creation Time	Latency (ms)	Throughput (FPS)	Accuracy (%)	Status	Actions
1	frozen_inference_graph Details	Refined30_coco	Machine: Local Workstation Device: CPU	FP32	16/09/20, 05:37	247.84	3.93	0		

[Compare](#)

Selected Configuration

frozen_inference_graph • Refined30_coco • Local Workstation • CPU

[Profile](#)
[Optimize](#)
[Pack](#)

Select Inference Method

 Single Inference

 Parallel streams(1-4):

 Batch size (1-256):
 Group Inference

[Execute](#)

Inference Tips

Combinations of streams and batch sizes appear in real time on the Inference Results plot below and are characterized with throughput and latency values. Inference is executed asynchronously. To profile multiple combinations of parameters in sequence, use group inference.

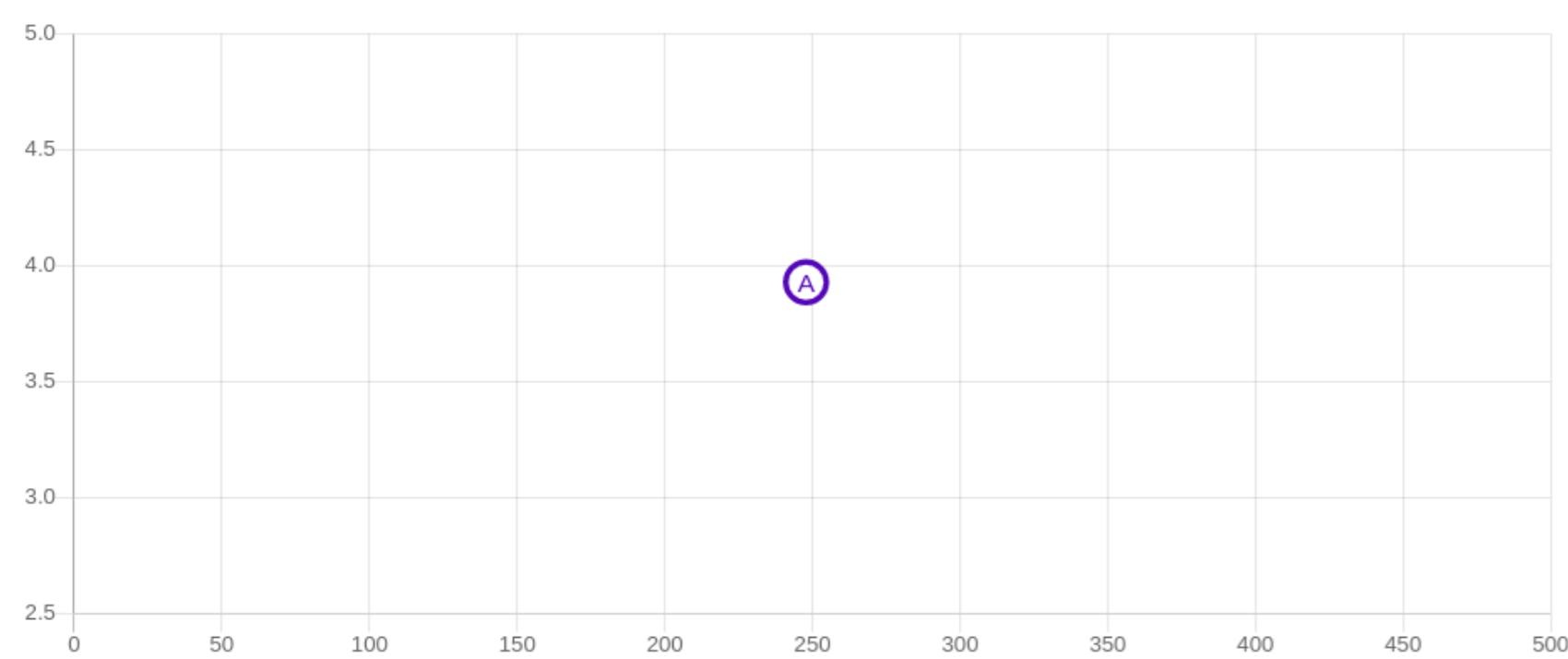
Model Performance Summary

Inference Results

frozen_inference_graph • Refined30_coco • Local Workstation • CPU

Latency Threshold (0-1000)

Throughput, fps



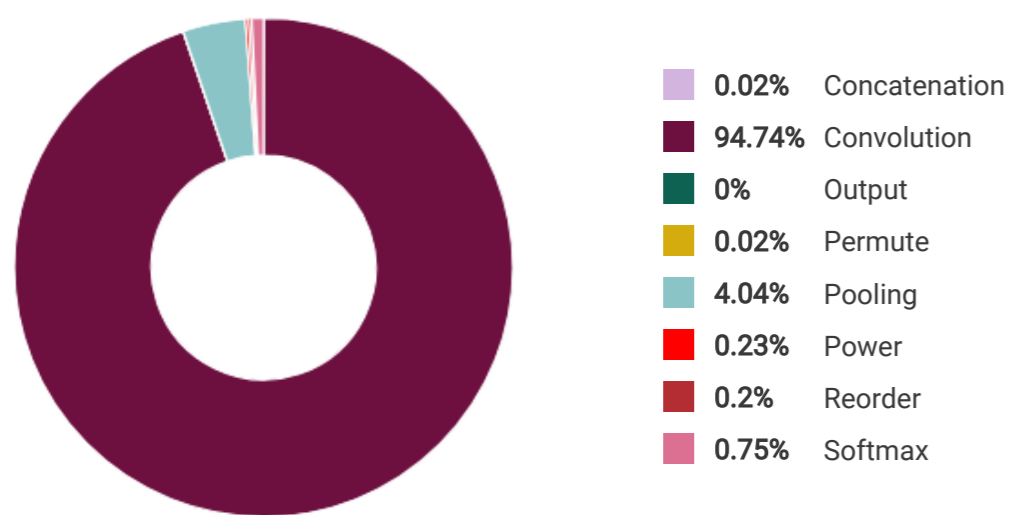
frozen_inference_graph • Refined30_coco • Local Workstation • CPU

 - Selected Point

 - Sweet Spot Point

Execution Time By Layer

frozen_inference_graph • Refined30_coco • Local Workstation • CPU



Layers Table

frozen_inference_graph • Refined30_coco • Local Workstation • CPU

[Visualize original IR](#)
[Visualize runtime graph](#)

Select Column

Select Filter

+ Add new filter

[Apply Filter](#)
[Clear Filter](#)

Execution Order	Layer Name	Layer Type	Execution Time, ms	Precision
0	image_tensor	Input	Not Executed	U8
9	FirstStageFeatureExtractor/InceptionV2/InceptionV2/Mixed_3b/Branch_0/Conv2d_0a_1x1/BatchNorm/FusedBatchNormV3/variance/Fused_Add	Convolution	1.197	FP32
10	FirstStageFeatureExtractor/InceptionV2/InceptionV2/Mixed_3b/Branch_1/Conv2d_0a_1x1/BatchNorm/FusedBatchNormV3/variance/Fused_Add	Convolution	1.237	FP32
11	FirstStageFeatureExtractor/InceptionV2/InceptionV2/Mixed_3b/Branch_1/Conv2d_0b_3x3/BatchNorm/FusedBatchNormV3/variance/Fused_Add	Convolution	3.713	FP32
12	FirstStageFeatureExtractor/InceptionV2/InceptionV2/Mixed_3b/Branch_2/Conv2d_0a_1x1/BatchNorm/FusedBatchNormV3/variance/Fused_Add	Convolution	1.189	FP32
13	FirstStageFeatureExtractor/InceptionV2/InceptionV2/Mixed_3b/Branch_2/Conv2d_0b_3x3/BatchNorm/FusedBatchNormV3/variance/Fused_Add	Convolution	4.808	FP32
14	FirstStageFeatureExtractor/InceptionV2/InceptionV2/Mixed_3b/Branch_2/Conv2d_0c_3x3/BatchNorm/FusedBatchNormV3/variance/Fused_Add	Convolution	7.493	FP32
18	FirstStageFeatureExtractor/InceptionV2/InceptionV2/Mixed_3c/Branch_0/Conv2d_0a_1x1/BatchNorm/FusedBatchNormV3/variance/Fused_Add	Convolution	1.969	FP32
19	FirstStageFeatureExtractor/InceptionV2/InceptionV2/Mixed_3c/Branch_1/Conv2d_0a_1x1/BatchNorm/FusedBatchNormV3/variance/Fused_Add	Convolution	1.781	FP32
20	FirstStageFeatureExtractor/InceptionV2/InceptionV2/Mixed_3c/Branch_1/Conv2d_0b_3x3/BatchNorm/FusedBatchNormV3/variance/Fused_Add	Convolution	5.394	FP32
21	FirstStageFeatureExtractor/InceptionV2/InceptionV2/Mixed_3c/Branch_2/Conv2d_0a_1x1/BatchNorm/FusedBatchNormV3/variance/Fused_Add	Convolution	1.593	FP32
22	FirstStageFeatureExtractor/InceptionV2/InceptionV2/Mixed_3c/Branch_2/Conv2d_0b_3x3/BatchNorm/FusedBatchNormV3/variance/Fused_Add	Convolution	5.017	FP32
23	FirstStageFeatureExtractor/InceptionV2/InceptionV2/Mixed_3c/Branch_2/Conv2d_0c_3x3/BatchNorm/FusedBatchNormV3/variance/Fused_Add	Convolution	8.126	FP32

[Expand Chart](#)
[Expand Table](#)

#	Streams ↑	Batch size ↑	Throughput, fps ↑	Latency, ms ↑	Last status	Filter
A	1	1	3.93	247.84		<input type="checkbox"/>

Latency, ms

Execution Attributes

frozen_inference_graph • Refined30_coco • Local Workstation • CPU

FPS: 3.93

Latency (ms): 247.84

Total Execution Time (ms): 20364.68

Batch size: 1

Streams: 1