



## Convert and Quantize:

```
MODEL=mcv3.1.1.1-tsa-only.pb
INPUT=self
OUTPUT=Identity
```

```
snpe-tensorflow-to-dlc --input_network $MODEL --input_dim $INPUT "1,256,256,20" --out_node $OUTPUT --output_path test.dlc
--show_unconsumed_nodes
snpe-dlc-quantize --input_dlc test.dlc --input_list=patchlist.txt --output_dlc test_q.dlc --verbose --act_bitwidth=16
--enable_htp
```

## Quantize stage produce an error:

```
/local/mnt/workspace/mlg_user_admin/docker.ci.tmp_htp/build/x86_64-linux-clang/SecondParty/QNN/src/qnn-htp-emu/DSP/HTP/src/
hexagon/src/graph_spillfill.cc:799:ERROR:A single op: q::SlicePad_shape_inplace.tcm (0x1e) requires 0x1400800 bytes of TCM,
which is greater than the TCM size of 0x400000!
```

## Ok, let's try to skip Slice operator:

```
INPUT=functional_3/tf_op_layer_Mul_1/Mul_1
```

```
snpe-tensorflow-to-dlc --input_network $MODEL --input_dim $INPUT "5,256,256,4" --out_node $OUTPUT --output_path test.dlc
--show_unconsumed_nodes
snpe-dlc-quantize --input_dlc test.dlc --input_list=patchlist.txt --output_dlc test_q.dlc --verbose --act_bitwidth=16
--enable_htp
```

## New error:

```
local/mnt/workspace/mlg_user_admin/docker.ci.tmp_htp/build/x86_64-linux-clang/SecondParty/QNN/src/qnn-htp-emu/DSP/HTP/src/h
exagon/src/graph_spillfill.cc:799:ERROR:A single op: q::Transpose_impl (0x11) requires 0x500000 bytes of TCM, which is
greater than the TCM size of 0x400000!
```

## Let's skip Transpose operator:

```
INPUT=functional_3/tf_op_layer_Transpose_1/Transpose_1
```

```
snpe-tensorflow-to-dlc --input_network $MODEL --input_dim $INPUT "256,256,5,4" --out_node $OUTPUT --output_path test.dlc
--show_unconsumed_nodes
snpe-dlc-quantize --input_dlc test.dlc --input_list=patchlist.txt --output_dlc test_q.dlc --verbose --act_bitwidth=16
--enable_htp
```

Successfully quantized

Transpose required 5MB of VTCM.

Tensor size is  $256 \times 256 \times 20 \times \text{sizeof}(\text{uint16})$

Looks like the source and destination tensors both want to be placed into VTCM

But Slice required more than 20MB of VTCM

I have no idea what's the reason for this.