

tf.nn.quantized_conv2d

```
quantized_conv2d(  
    input,  
    filter,  
    min_input,  
    max_input,  
    min_filter,  
    max_filter,  
    strides,  
    padding,  
    out_type=tf.qint32,  
    name=None  
)
```

Defined in `tensorflow/python/ops/gen_nn_ops.py`.

See the guide: [Neural Network > Candidate Sampling](#)

Computes a 2D convolution given quantized 4D input and filter tensors.

The inputs are quantized tensors where the lowest value represents the real number of the associated minimum, and the highest represents the maximum. This means that you can only interpret the quantized output in the same way, by taking the returned minimum and maximum values into account.

Args:

- `input`: A **Tensor**. Must be one of the following types: `qint8`, `quint8`, `qint16`, `quint16`, `qint32`.
- `filter`: A **Tensor**. Must be one of the following types: `qint8`, `quint8`, `qint16`, `quint16`, `qint32`. filter's input_depth dimension must match input's depth dimensions.
- `min_input`: A **Tensor** of type `float32`. The float value that the lowest quantized input value represents.
- `max_input`: A **Tensor** of type `float32`. The float value that the highest quantized input value represents.
- `min_filter`: A **Tensor** of type `float32`. The float value that the lowest quantized filter value represents.
- `max_filter`: A **Tensor** of type `float32`. The float value that the highest quantized filter value represents.
- `strides`: A list of `ints`. The stride of the sliding window for each dimension of the input tensor.
- `padding`: A `string` from: `"SAME"`, `"VALID"`. The type of padding algorithm to use.
- `out_type`: An optional `tf.DType` from: `tf.qint8`, `tf.quint8`, `tf.qint16`, `tf.quint16`, `tf.qint32`. Defaults to `tf.qint32`.
- `name`: A name for the operation (optional).

Returns:

A tuple of **Tensor** objects (output, min_output, max_output).

- `output`: A **Tensor** of type `out_type`.
- `min_output`: A **Tensor** of type `float32`. The float value that the lowest quantized output value represents.
- `max_output`: A **Tensor** of type `float32`. The float value that the highest quantized output value represents.

Except as otherwise noted, the content of this page is licensed under the [Creative Commons Attribution 3.0 License](#), and code samples are licensed under the [Apache 2.0 License](#). For details, see our [Site Policies](#). Java is a registered trademark of Oracle and/or its affiliates.

Last updated November 2, 2017.

Stay Connected

[Blog](#)

[GitHub](#)

[Twitter](#)

Support

[Issue Tracker](#)

[Release Notes](#)

[Stack Overflow](#)

English

[Terms](#) | [Privacy](#)