ONNX Pre-processing WG

Monthly meeting - March 8, 2022

Agenda

- Status
- Next step, working group wrap-up(?)
- Open floor

Status

PRs:

[MERGED] Add Data preprocessing with ONNX: ResNet-50 example
https://github.com/onnx/tutorials/pull/266

[MERGED] Add Resize-19: half_pixel_symmetric
https://github.com/onnx/onnx/pull/4862

[Pending] Using ONNX parser in SequenceMap tutorial
https://github.com/onnx/tutorials/pull/277

Status / Next steps

Goals		
General infrastructure	Utilities to combine models	✅
	Support for heterogeneous batch processing	✅
	Labeling parts of the model	✅
First model published (ResNet)	Resize-18	✅
	CenterCropPad	✅
	Resize-19 (half_pixel_symmetric)	✅ (* Not yet released)
	Publish ResNet preprocessing notebook	✅
	Make Resnet models with preprocessing available in https://github.com/onnx/models/tree/main/vision/classification/resnet and onnx.hub	
	Documentation page about creating models with preprocessing	

Models with pre-processing in onnx.hub

```
{
    "model": "ResNet50-fp32",
    "model_path": "vision/classification/resnet/model/resnet50-v1-12.onnx",
    "onnx_version": "1.7.0",
    "opset_version": 12,
    "metadata": {
        "model_sha": "3f03fdef724b22947eed826f1eef1dc5c34151bb4c37d634f1db89dfa2dd1526",
        "model_bytes": 102576593,
        "tags": [
            "vision",
            "classification",
            "resnet"
        ],
        "io_ports": {
            "inputs": [
                {
                    "name": "data",
                    "shape": [
                        "N",
                        3,
                        224,
                        224
                    ],
                    "type": "tensor(float)"
                }
            ],
            "outputs": [
                {
                    "name": "resnetv17_dense0_fwd",
                    "shape": [
                        "N",
                        1000
                    ],
                    "type": "tensor(float)"
                }
            ]
        },
        "model_with_data_path": "vision/classification/resnet/model/resnet50-v1-12.tar.gz",
        "model_with_data_sha": "9391137cfc8fbec372d7a1e59e272d67550dab72d93cf7c7d6256782262599ea",
        "model_with_data_bytes": 96559469
    }
},
```

Manifest file:

https://raw.githubusercontent.com/onnx/models/main/ONNX_HUB_MANIFEST.json

Access:

```
model = onnx.hub.load("ResNet50-fp32")
```

How do we include preprocessing?

- Separate model "-e2e" (better name?)
- Use "model_with_preprocessing_path"?
- Post also "-preproc" model separately, in case someone want to reuse?

Open floor