CSC413 2021W PA2 Write Up Guolun Li

Part A:

A1.

Table

Description automatically generated

A2.

Graphical user interface, application

Description automatically generated

41.6% accuracy means for each image, on average 40% of the area are colored correctly. The final result looks ok since this is our very first trial, but there’s obviously room for improvement.

A3.

Assuming batch size to be 1, we derive general expressions for input of size (C,W,D, N) for each type of layers first. Here C = # of input channels, W and D are respectively the width and height of each input picture, and N = # of output channels.

We provide justification for # of connections for BatchNorm2d and UpSample layers:

The batchnorm layer computes C different expected values and variances over all possible W\*D pixels. For each channel, W\*D input pixels are fully connected to W\*D output pixels. Thus total # of connection is .

The upsample layer is similar to max pooling layer, but in a reverse way. Each input pixel uniquely connects to 4 output pixels. Thus total # of connection is 4 \* C \* WD.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Conv2d | MaxPool2d | BatchNorm2d | UpSample |
| # of weights |  | 0 |  | 0 |
| # of outputs |  |  |  |  |
| # of connections |  |  |  |  |

With initial image size of 32 x 32:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | # of Inputs | # of Weights | # of Outputs | # of connections |
| Conv1 | (NIC, 32, 32, NF) |  |  |  |
| MaxPool1 | (NF, 32, 32, NF) | 0 |  |  |
| BatchNorm1 | (NF, 16, 16, NF) | 2NF |  |  |
| Conv2 | (NF, 16, 16, 2NF) |  |  |  |
| MaxPool2 | (2NF, 16, 16, 2NF) | 0 |  |  |
| BatchNorm2 | (2NF, 8, 8, 2NF) | 4NF |  |  |
| Conv3 | (2NF, 8, 8, NF) |  |  |  |
| UpSample3 | (NF,8,8,NF) | 0 |  |  |
| BatchNorm3 | (NF,16,16,NF) | 2NF |  |  |
| Conv4 | (NF,16,16,NC) |  |  |  |
| UpSample4 | (NC,16,16,NC) | 0 |  |  |
| BatchNorm4 | (NC,32,32,NC) | 2NC |  |  |
| Conv5 | (NC,32,32,NC) |  |  |  |

Total # of weights =

Total # of outputs =

Total # of connections =

When weight and height are doubled, we can calculate the total #’s in the same way to obtain the following results:

Total # of weights =

Total # of outputs =

Total # of connections =

Part B:

B1:

Graphical user interface, text

Description automatically generated

B2:

Graphical user interface, application

Description automatically generated

B3.

Compared to the previous result, the validation loss decreased from 1.576 to 1.1549(26.7% relative decrement) and the validation accuracy increased from 41.6% to 54.9%(13.3% absolute increment). This is a huge boost in model performance.

Because in a transposed convolution layer, multiple input channels are attached to one output channel, whereas in a upsampling layer, one input channel can only connect to one output channel. As a result, transposed convolutional layer provides more learnable parameters than UpSampling layer for the model to lear to upsample optimally.

B4.

See table below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Layer | parameter | K = 3 | K = 4 | K=5 |
| Conv2d | Padding | P = 1 | P = 1 | P = 2 |
| ConvTrans2d | Padding | P = 1 | P = 2 | P = 2 |
|  | Output padding |  |  |  |

B5.

Smaller the batch size, better the validation loss and output image quality.

Part C:

C1:

Text

Description automatically generated with medium confidence

C2:

Graphical user interface, application

Description automatically generated

C3:

1)The validation loss decreased from 1.1549 to 1.0678(7.5% relative decrement), and the validation accuracy increased from 54.9% to 58.0%(3.1% absolute increment). The model is slightly better than the previous one.

2) Skip connections indeed improve the validation loss and accuracy.

3) Such improvement is not large enough to be considered “qualitative”.

4) There are two reasons why skipnet improve performance:

--Skipnet adds additional trainable parameters to the model. Here we’re using validation loss as a metric, so the model performance literally improves with model complexity.

--Without skipped connections, only the most essential information is retained. By adding skip connections, the model is able to make final decision based on details inside the image.

Part D.1:

D.1.1:

Graphical user interface, text, application, email

Description automatically generated

D.1.2:

Text, application

Description automatically generated

Chart

Description automatically generated

The best mIOU is 0.3420.

D.1.3:

A picture containing grass, colorful

Description automatically generated

Part D.2:

D.2.1:

Text

Description automatically generated

Chart

Description automatically generated

The best mIOU is now 0.3061, which is worse than the previous model.

D.2.2:

A collage of flowers

Description automatically generated with low confidence

A picture containing grass, green, yellow

Description automatically generated