

Autoencoders for mSEM artifact detection

mSEM artifacts: general

- mSEM artifacts are a problem for automated downstream analysis -> hinder classification / segmentation / agglomeration
- mSEM artifacts come in various forms and are difficult to detect reliably using humans or hand-designed algorithms
- Variety of artifacts and large data size
 - Problematic situation for classical supervised approaches (e.g. U-net)
- Training a *naïve* U-net or similar (>E6 params) in a supervised fashion would require HUGE amount of labeled training data

mSEM artifacts: data situation

- Huge amount of unlabeled data (even in mag-8-8-1 ~1TB)
- Much less labeled data (> 500 training patches ~ 10MB):
 - Lablog Summary:
<https://mhlablog.net/2020/02/05/training-data-for-artefact-detection-in-msem-data/>
 - Circular Artifact:
<https://webknossos.brain.mpg.de/annotations/Explorational/5e387a00010000501b1a41bd#25238,20546,3546,0,0.513,1>
 - Horse shoe Artifact:
<https://webknossos.brain.mpg.de/annotations/Explorational/5e387a00010000501b1a41bd#19895,15530,3732,0,0.513,12>
 - Scratch Artifact:
<https://webknossos.brain.mpg.de/annotations/Explorational/5e387a00010000501b1a41bd#20088,16049,3925,0,0.513,17>

Supervised CNN vs. unsupervised autoencoder (AE)

CNN

$$f: X \rightarrow Y$$

$$y = f'(x) + \varepsilon$$

$$\text{Min}_p \text{Loss}(f'_p(x) - y)$$

class CNN:

def predict(x):

 y_hat = f(x)

 return y_hat

AE

$$e: X \rightarrow L$$

$$d: L \rightarrow X$$

$$x = d'(e'(x)) + \varepsilon$$

$$\text{Min}_p \text{Loss}(d'_p(e'_p(x)) - x)$$

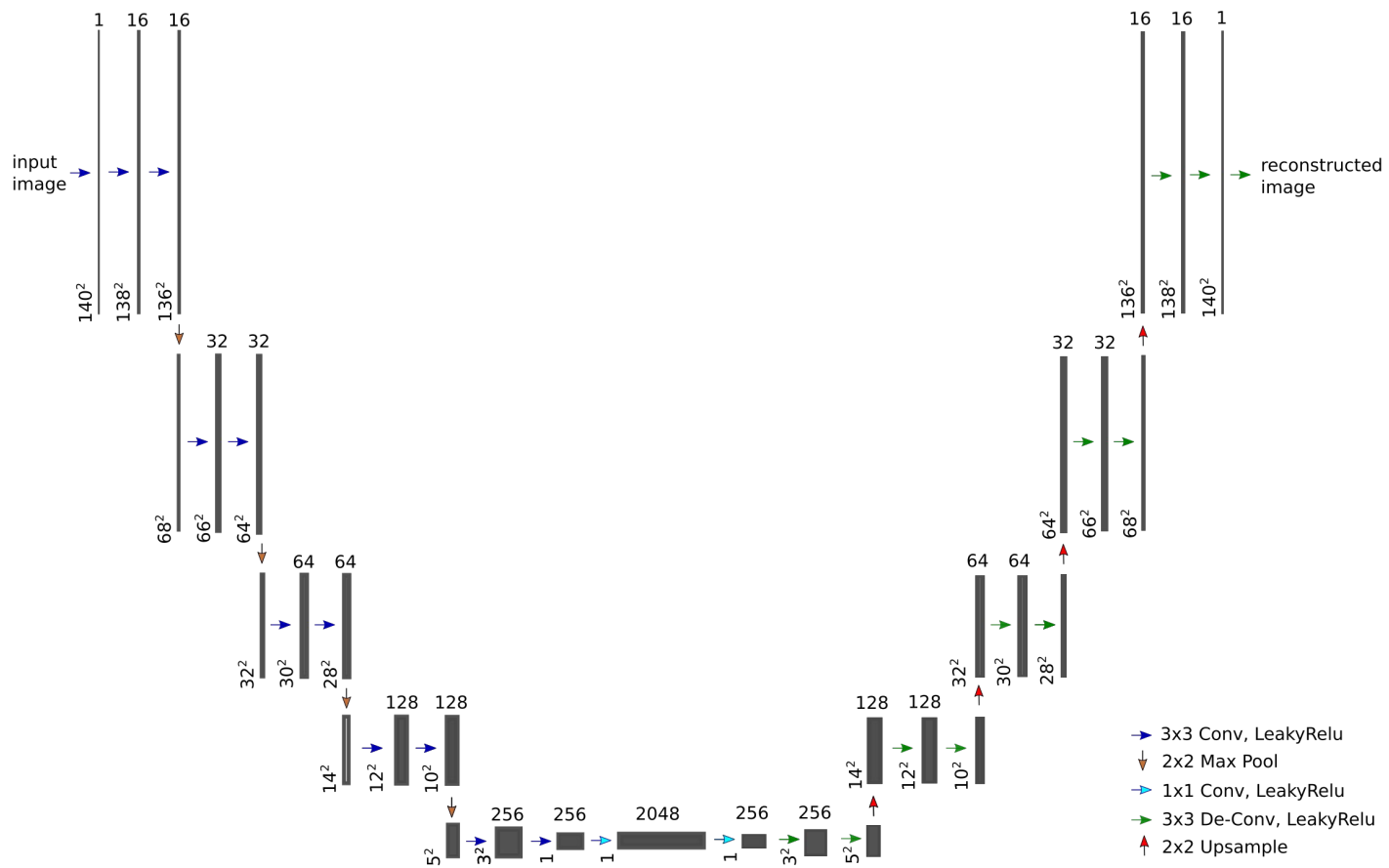
class AE:

def predict(x)

 x_hat = d(e(x))

 return x_hat

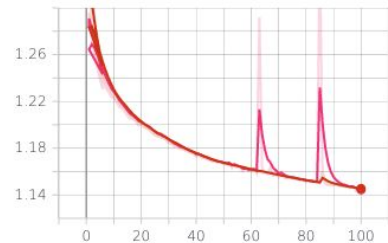
Basic multi-res AE for mSEM data



Basic multi-res AE for mSEM data

epoch_loss

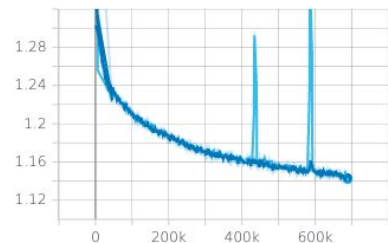
epoch_loss



running_loss_train
epoch_loss_train
running_loss_val
epoch_loss_val

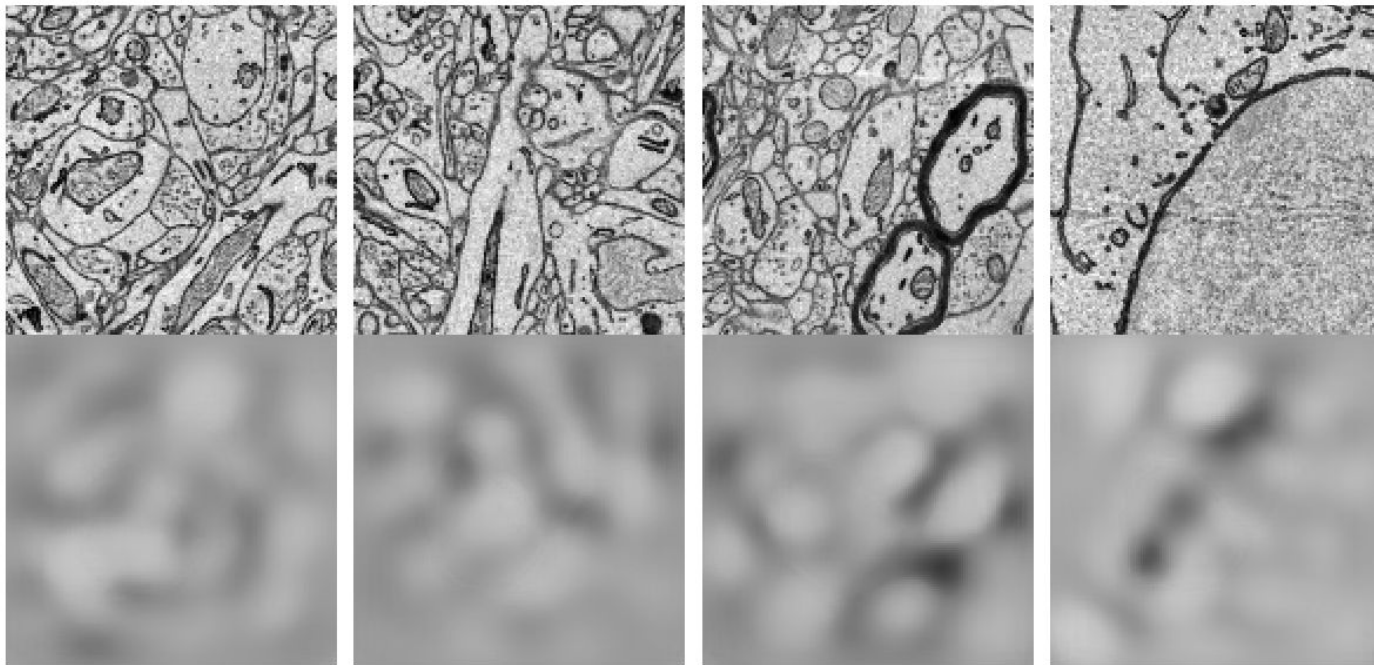
running_loss

running_loss



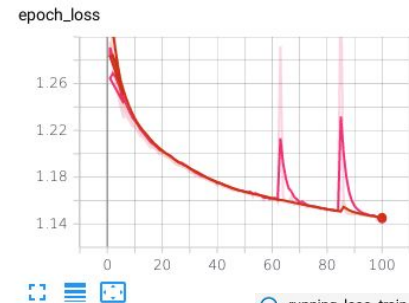
images train
step 3

Sat Jul 11 2020 18:16:28 GMT+0200 (Central European Summer Tim

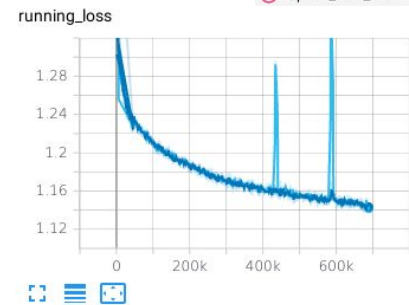


Basic multi-res AE for mSEM data

epoch_loss

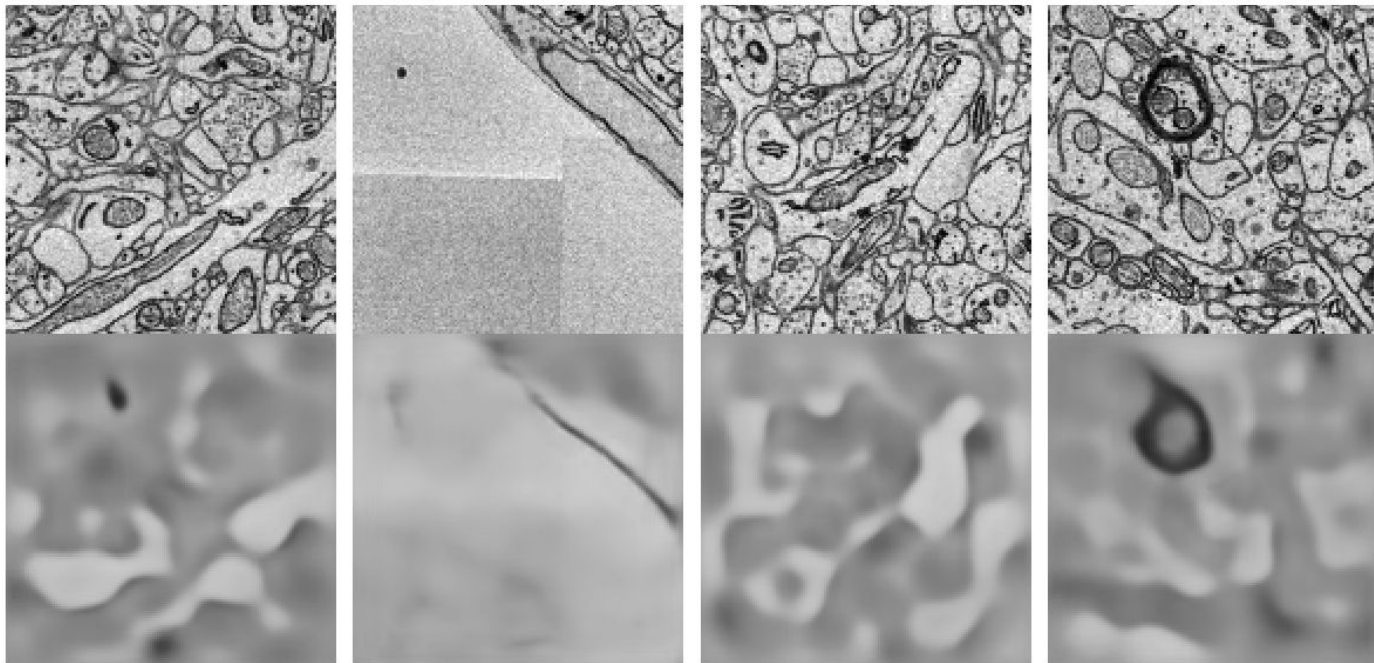


running_loss



images train
step 100

Sun Jul 12 2020 11:35:30 GMT+0200 (Central European Summer Time)



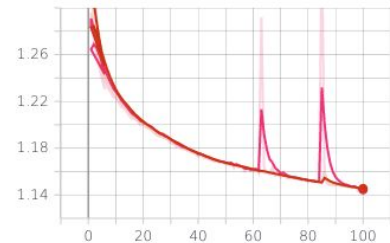
Basic multi-res AE for mSEM data

epoch_loss

images val
step 99

Sun Jul 12 2020 11:26:27 GMT+0200 (Central European Summer Time)

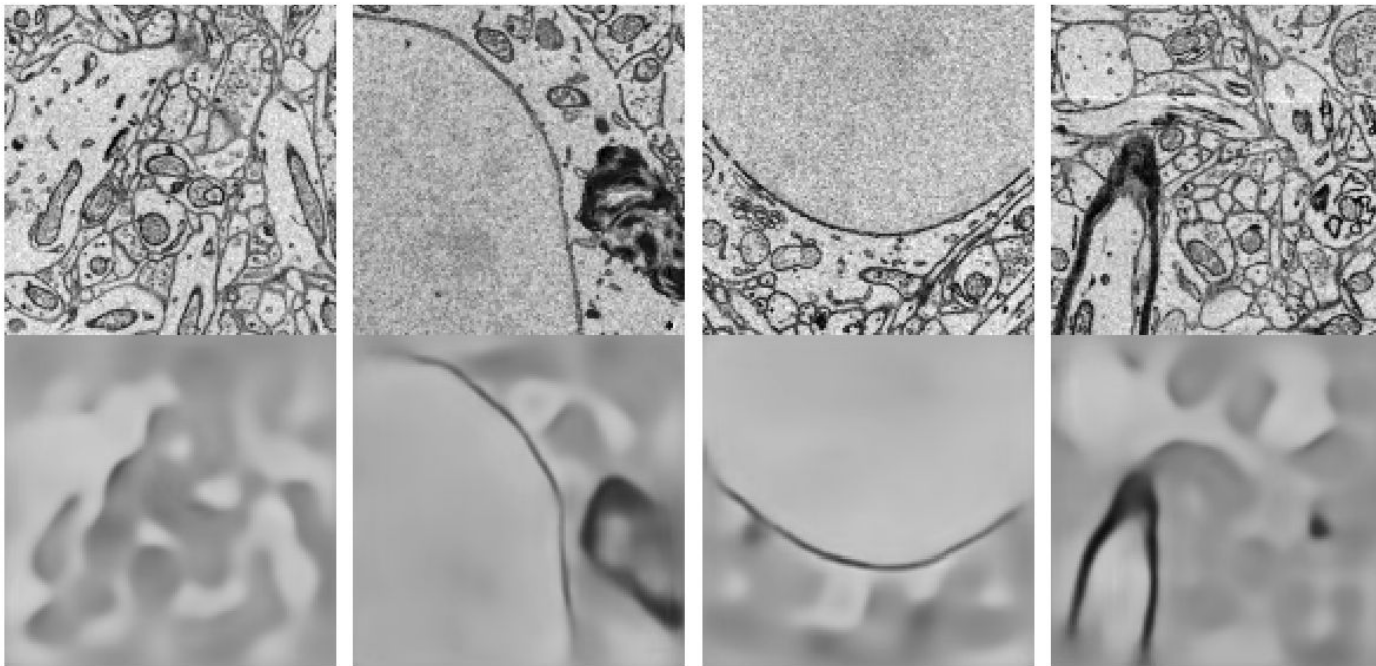
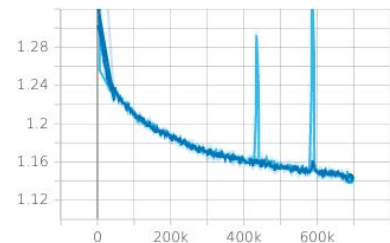
epoch_loss



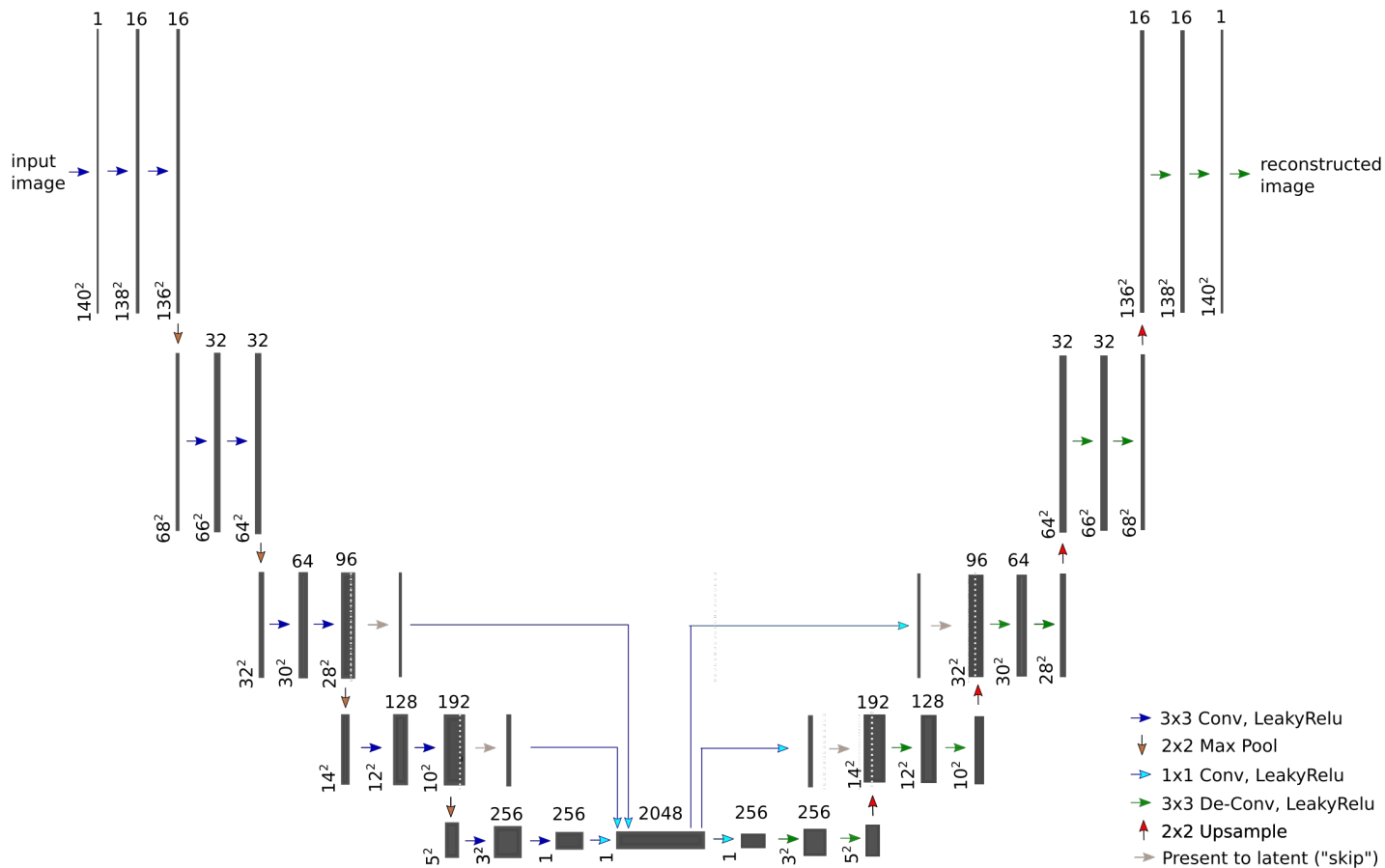
running_loss_train
epoch_loss_train
running_loss_val
epoch_loss_val

running_loss

running_loss



Multi-res “skip” AE for mSEM data



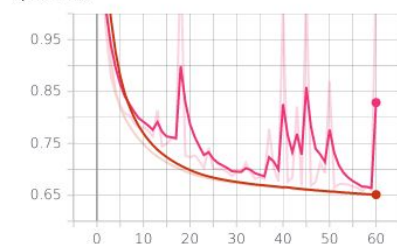
Multi-res “skip” AE for mSEM data

epoch_loss

images train
step 3

Thu Jul 30 2020 09:58:04 GMT+0200 (Central European Summer Time)

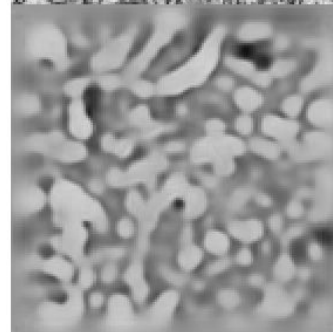
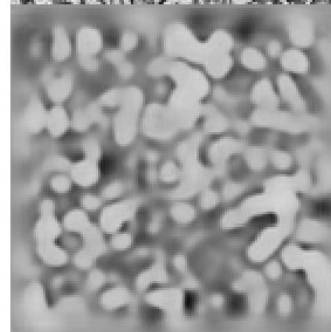
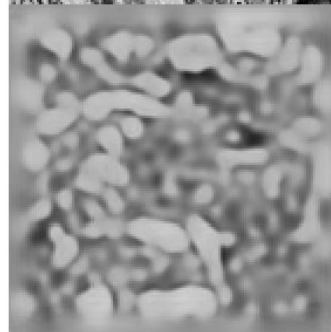
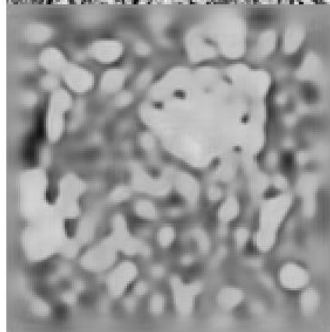
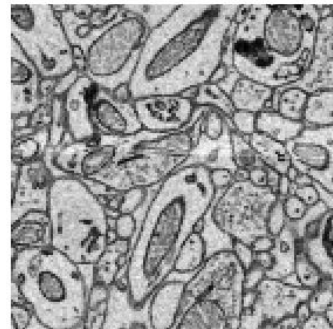
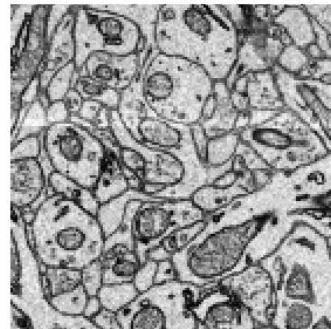
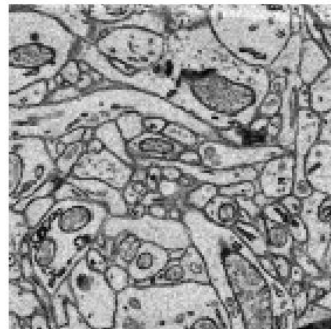
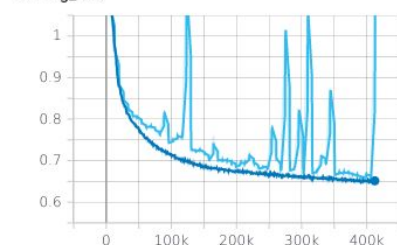
epoch_loss



running_loss_train
epoch_loss_train
running_loss_val
epoch_loss_val

running_loss

running_loss

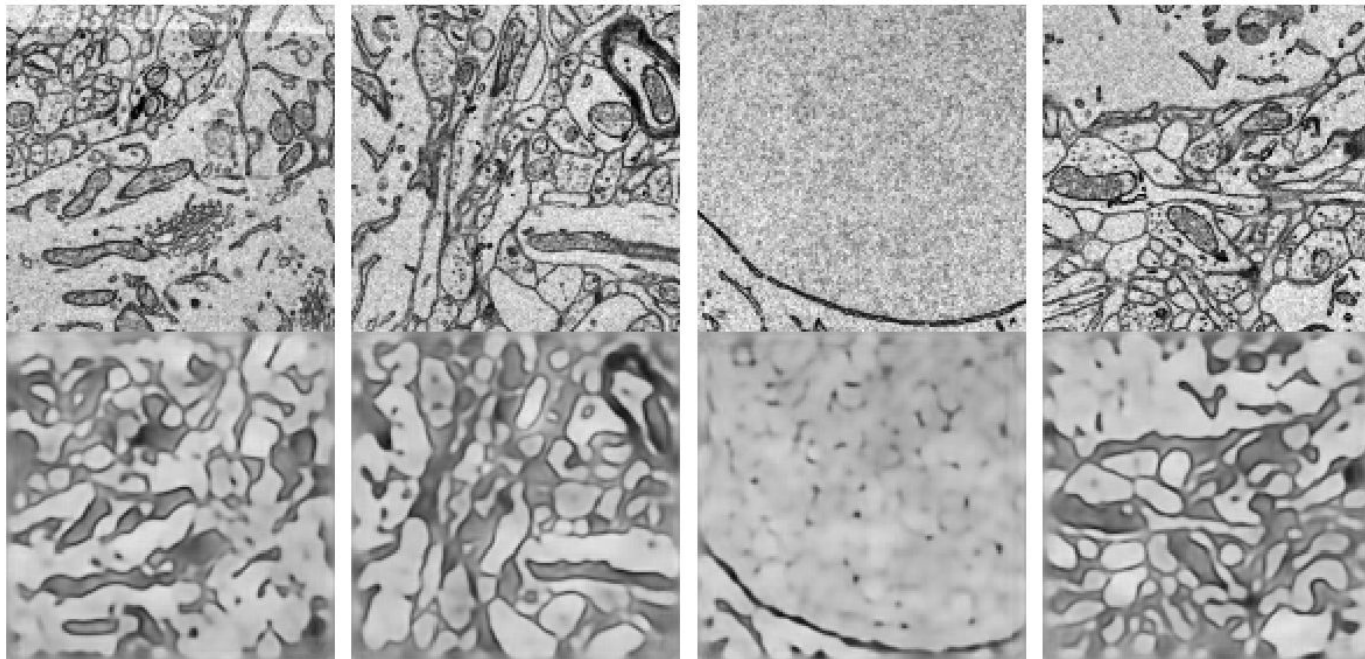
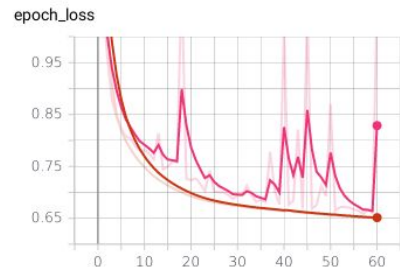


Multi-res “skip” AE for mSEM data

epoch_loss

images train
step 54

Thu Jul 30 2020 20:59:14 GMT+0200 (Central European Summer Time)

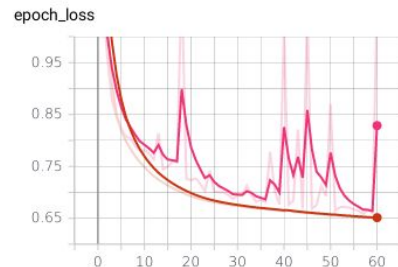


Multi-res “skip” AE for mSEM data

epoch_loss

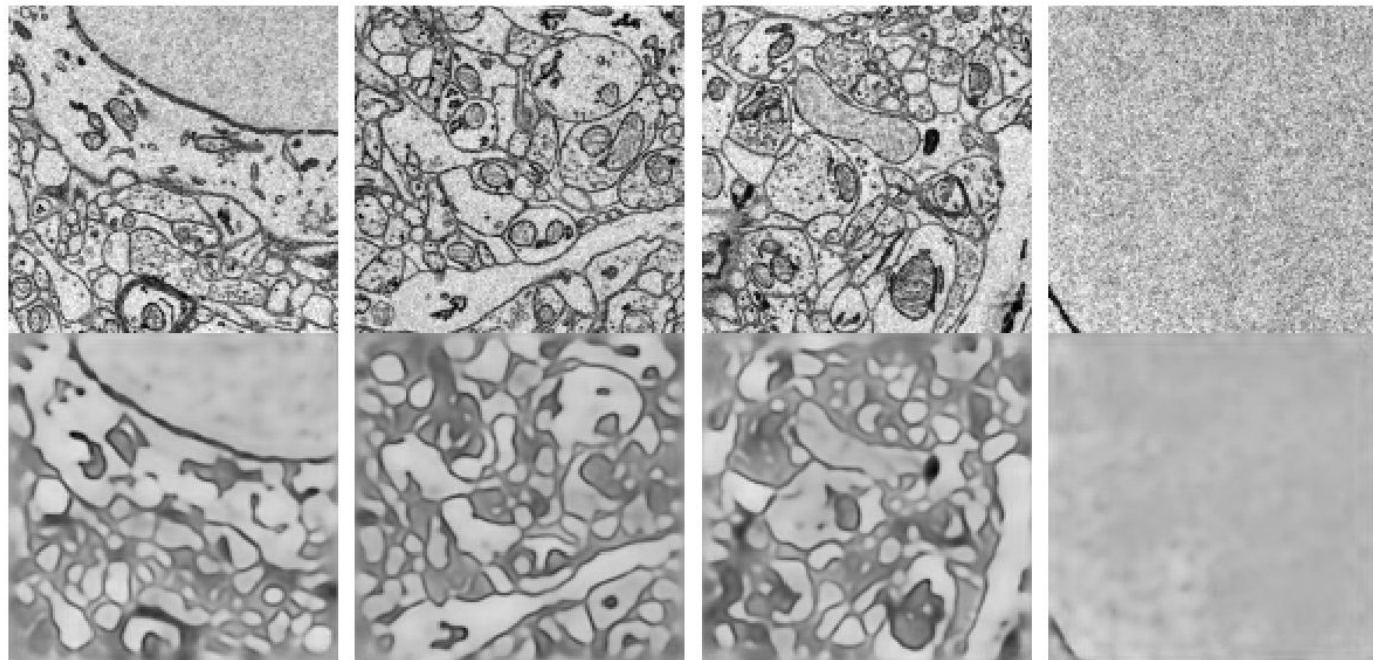
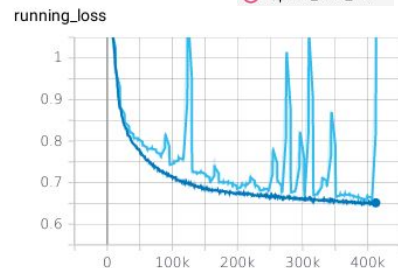
images val
step 59

Thu Jul 30 2020 22:05:10 GMT+0200 (Central European Summer Time)



running_loss_train
epoch_loss_train
running_loss_val
epoch_loss_val

running_loss

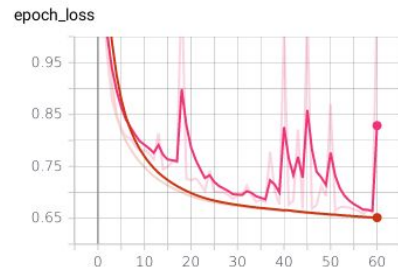


Multi-res “skip” AE for mSEM data

images val
step 59

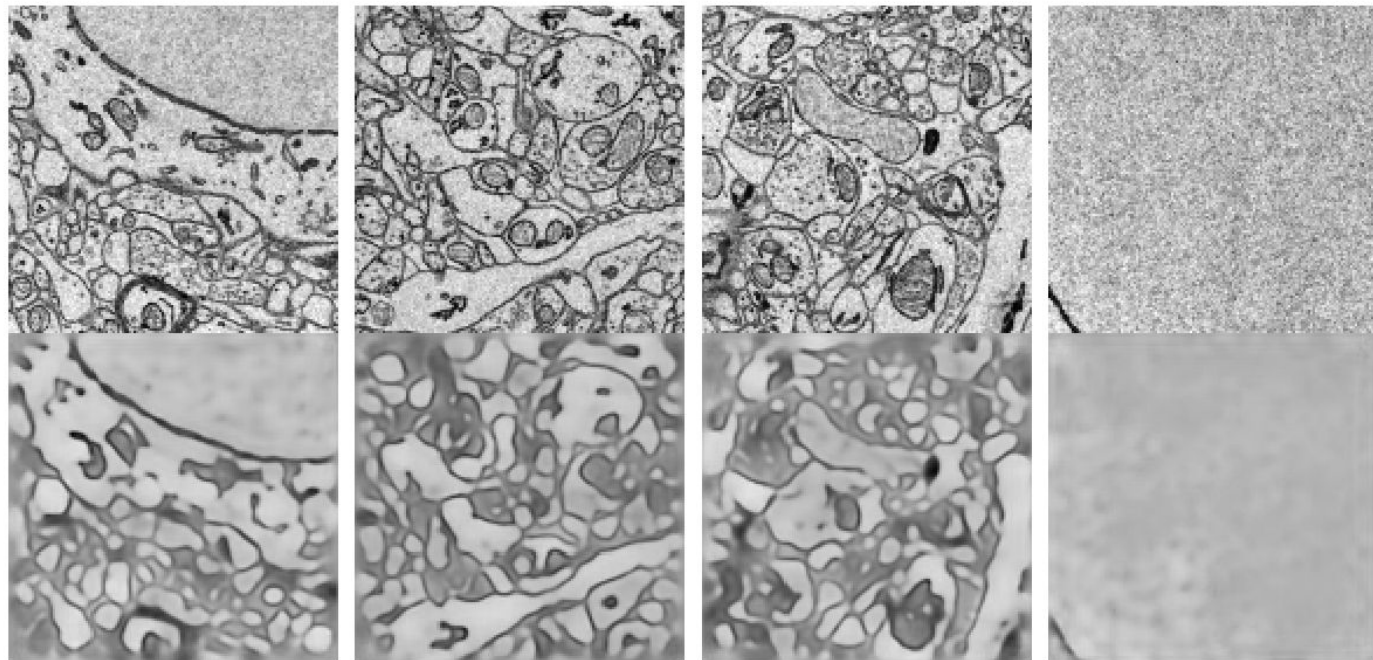
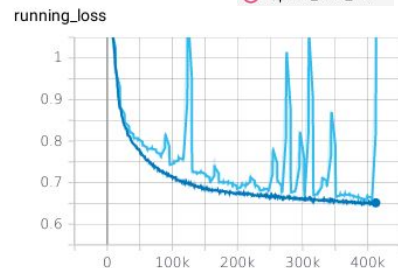
Thu Jul 30 2020 22:05:10 GMT+0200 (Central European Summer Time)

epoch_loss



running_loss_train
epoch_loss_train
running_loss_val
epoch_loss_val

running_loss



Unsupervised AE vs. semi-supervised encoder + classifier

AE

$$e: X \rightarrow L$$

$$d: L \rightarrow X$$

$$x = d'(e'(x)) + \varepsilon$$

$$\text{Min}_p \text{Loss}(d'_p(e'_p(x)) - x)$$

class AE:

```
def predict(x)
    x_hat = d(e(x))
    return x_hat
```

Encoder + classifier

$$e: X \rightarrow L$$

$$c: L \rightarrow Y$$

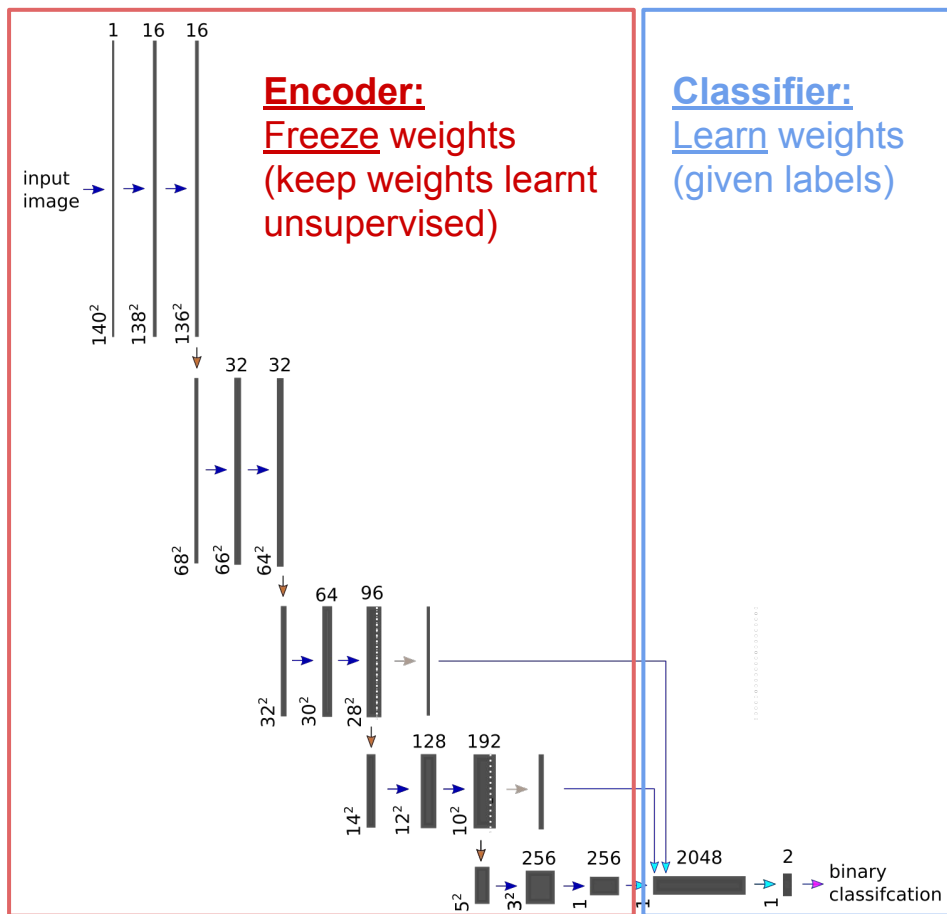
$$y = c'(e'(x)) + \varepsilon$$

$$\text{Min}_p \text{Loss}(c'_p(e'_p(x)) - x)$$

class EncoderClassifier:

```
def predict(x)
    y_hat = c(e(x))
    return y_hat
```

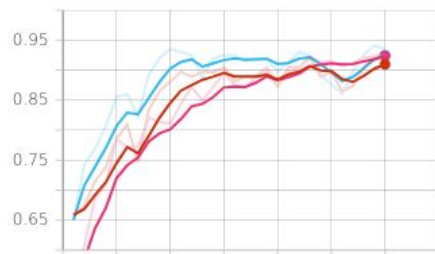
Multi-res “skip” encoder + binary classifier for mSEM data



Multi-res “skip” encoder + binary classifier for mSEM data

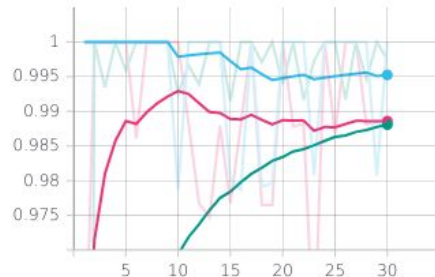
epoch_accuracy

epoch_accuracy



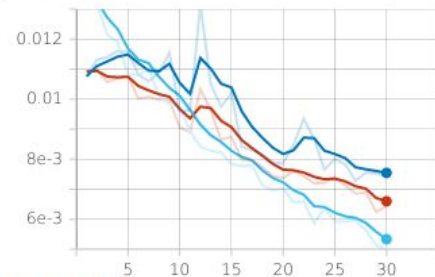
precision

precision



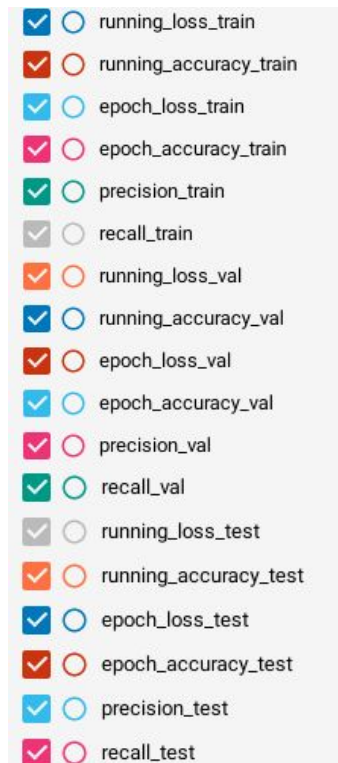
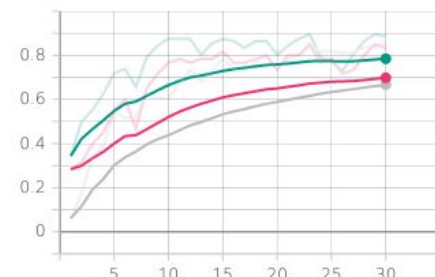
epoch_loss

epoch_loss



recall

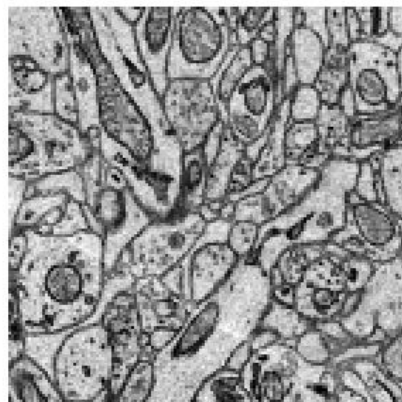
recall



Multi-res “skip” encoder + binary classifier for mSEM data

image_examples_test
step 2

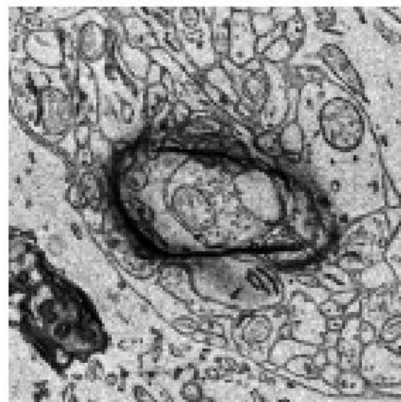
Wed Aug 05 2020 04:32:08 GMT+0200 (Central European Summer Time)



output (class 1): 0.44

prediction class: 0

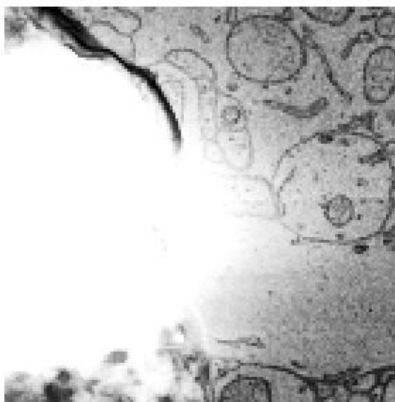
target class: 0



output (class 1): 0.56

prediction class: 1

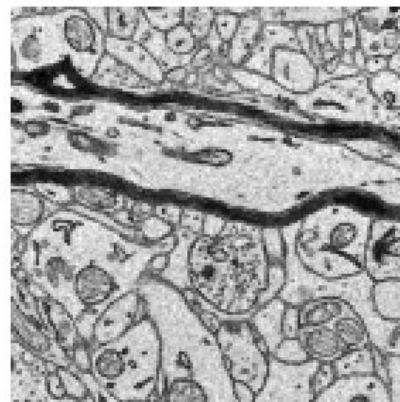
target class: 1



output (class 1): 0.40

prediction class: 0

target class: 1



output (class 1): 0.44

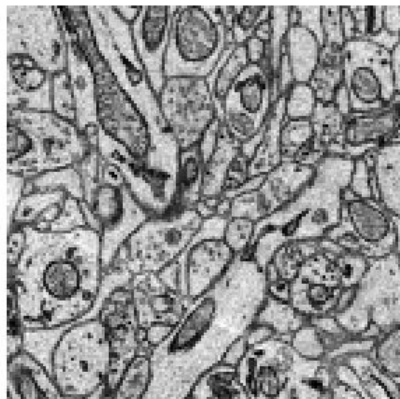
prediction class: 0

target class: 0

Multi-res “skip” encoder + binary classifier for mSEM data

image_examples_test
step 21

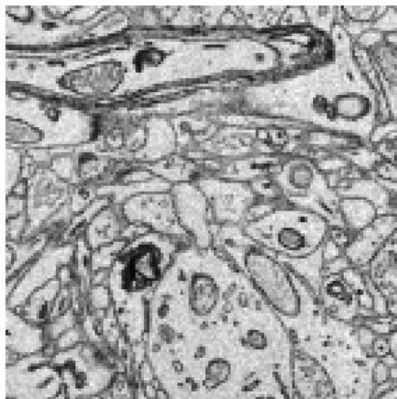
Wed Aug 05 2020 04:55:20 GMT+0200 (Central European Summer Time)



output (class 1): 0.28

prediction class: 0

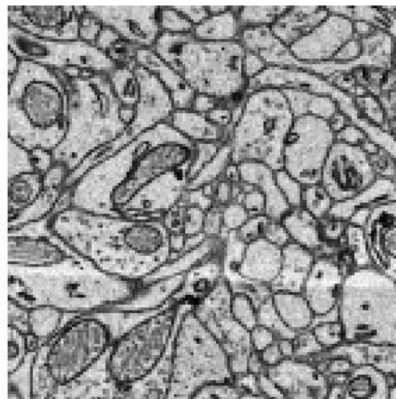
target class: 0



output (class 1): 0.41

prediction class: 0

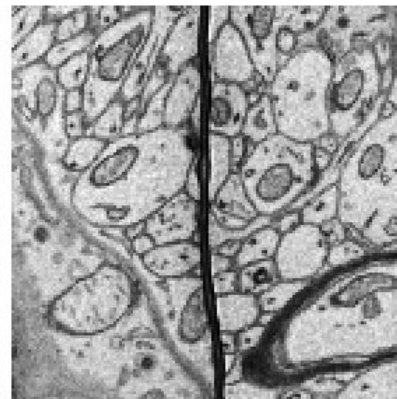
target class: 0



output (class 1): 0.34

prediction class: 0

target class: 0



output (class 1): 0.59

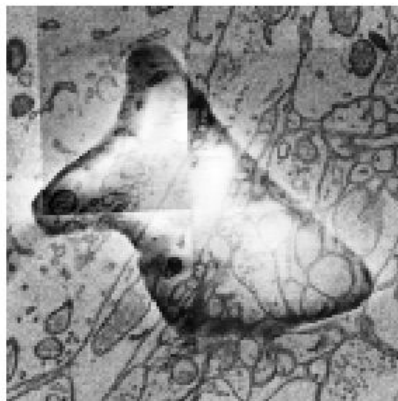
prediction class: 1

target class: 1

Multi-res “skip” encoder + binary classifier for mSEM data

image_examples_test
step 23

Wed Aug 05 2020 04:57:43 GMT+0200 (Central European Summer Time)



output (class 1): 0.58

prediction class: 1

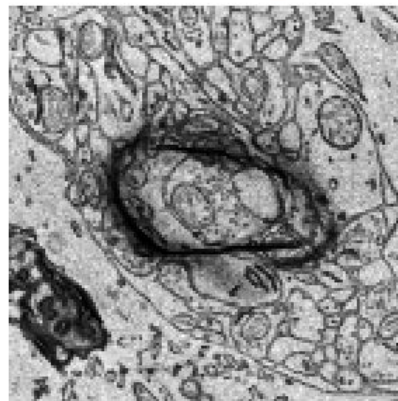
target class: 1



output (class 1): 0.35

prediction class: 0

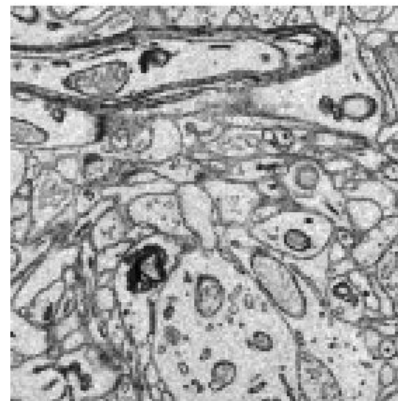
target class: 0



output (class 1): 0.89

prediction class: 1

target class: 1

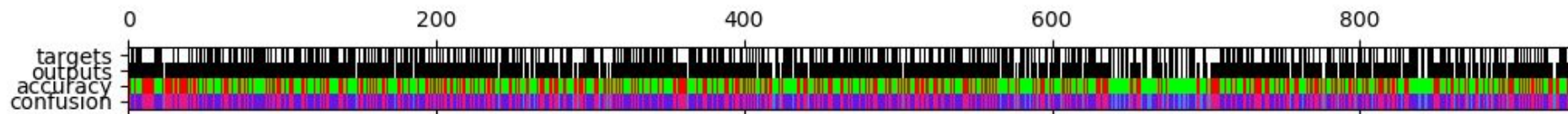


output (class 1): 0.42

prediction class: 0

target class: 0

Multi-res “skip” encoder + binary classifier for mSEM data



target|output

artifact

no artifact

accuracy

frac correct: 569/936=0.61

frac incorrect: 367/936=0.39

confusion

TP: 0.13

TN: 0.87

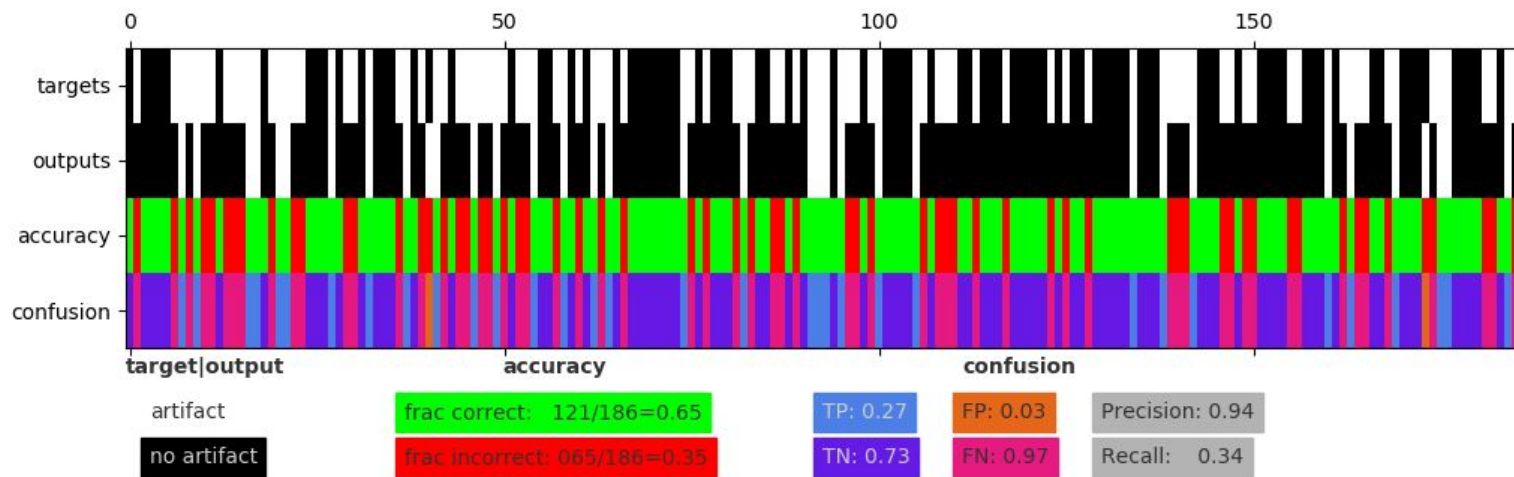
FP: 0.00

FN: 1.00

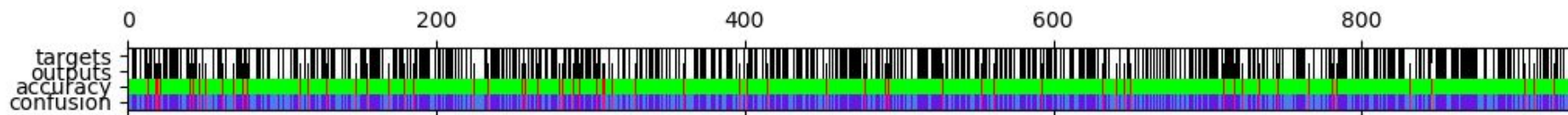
Precision: 1.00

Recall: 0.17

Multi-res “skip” encoder + binary classifier for mSEM data



Multi-res “skip” encoder + binary classifier for mSEM data



target|output

accuracy

confusion

artifact

frac correct: 872/936=0.93

TP: 0.44

FP: 0.02

Precision: 1.00

no artifact

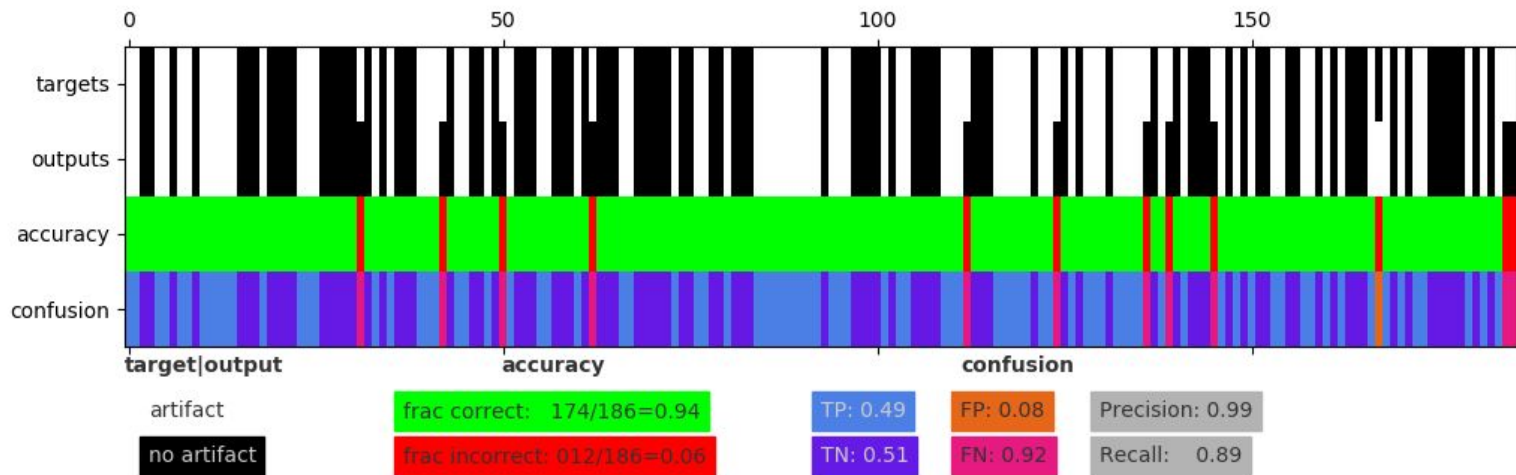
frac incorrect: 064/936=0.07

TN: 0.56

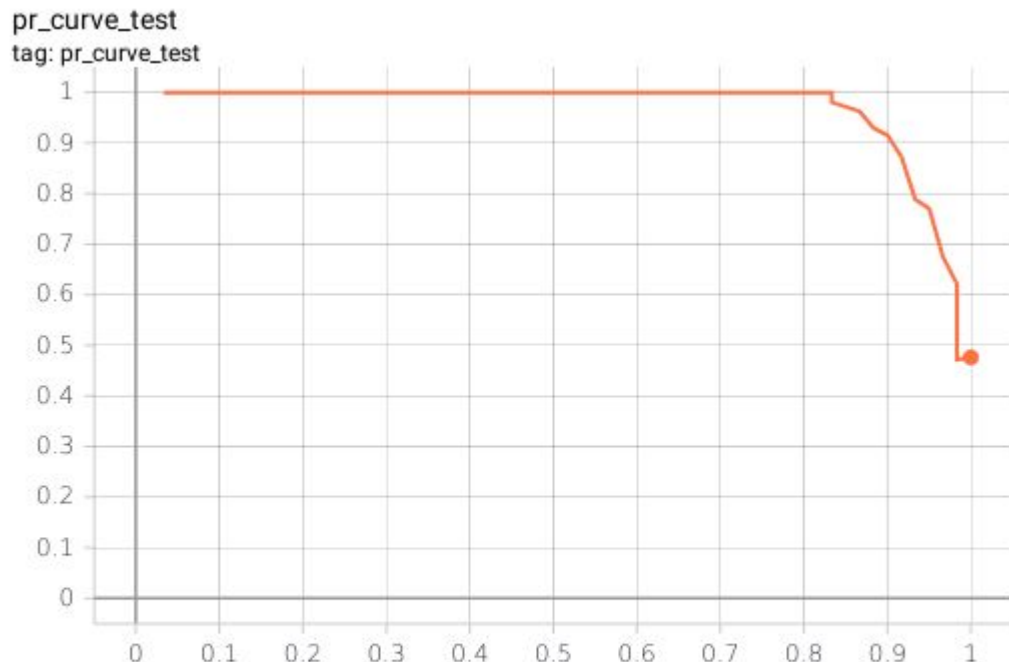
FN: 0.98

Recall: 0.86

Multi-res “skip” encoder + binary classifier for mSEM data



Multi-res “skip” encoder + binary classifier for mSEM data



■ . is at step 600

(Wed Aug 05 2020 05:06:18 GMT+0200 (Central European Summer Time))