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 *naive* 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

$e: X \rightarrow L$

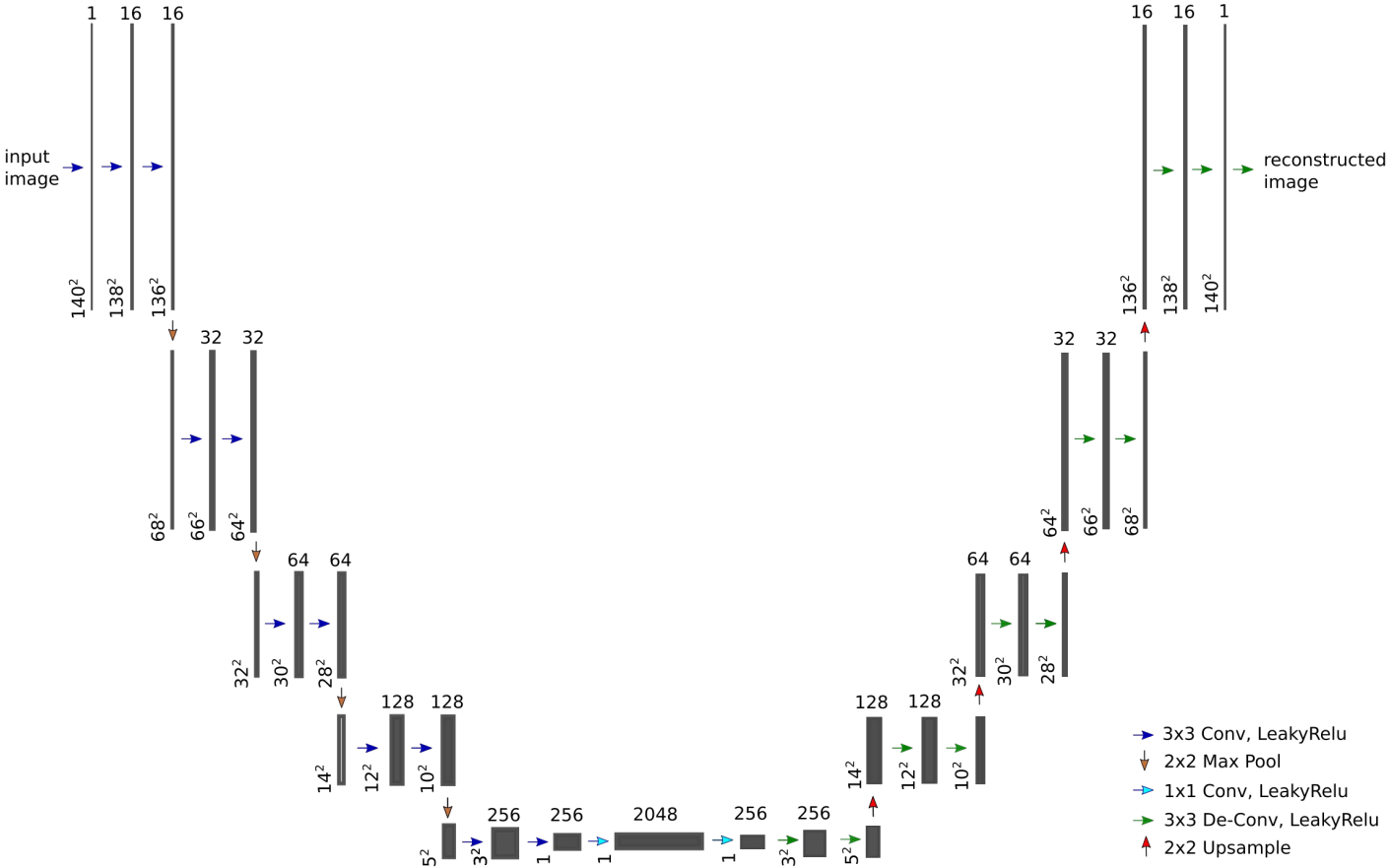
$d: L \rightarrow X$

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

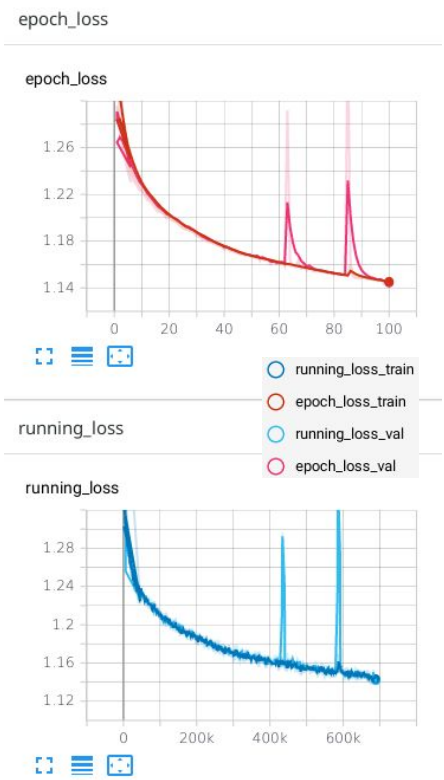
$\text{Min}_p \text{Loss}(d'_p(e'_p(x)) - x)$

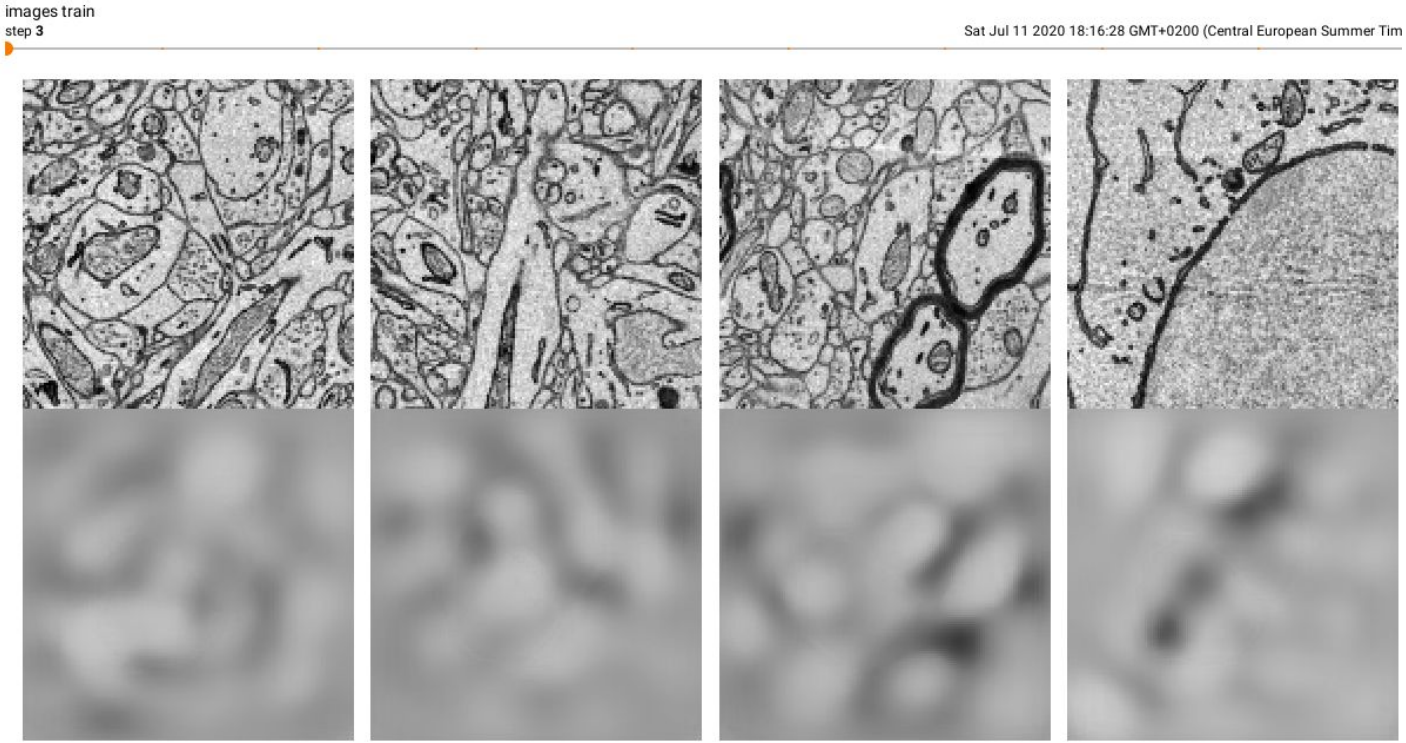
```
class AE:
  def predict(x)
      x_hat = d(e(x))
      return x_hat
```

Basic multi-res AE for mSEM data

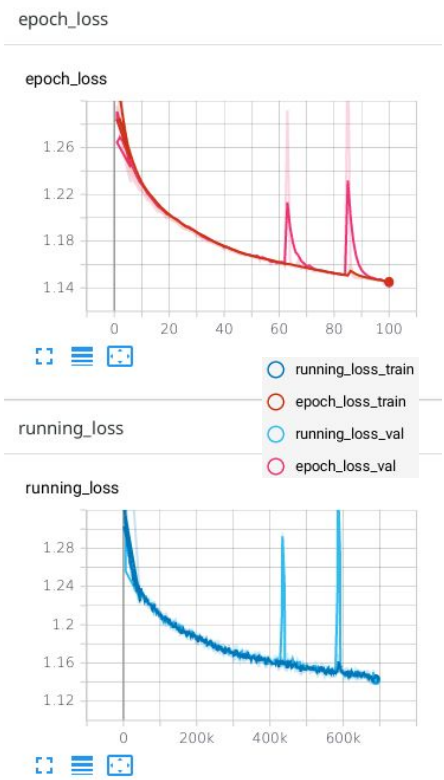


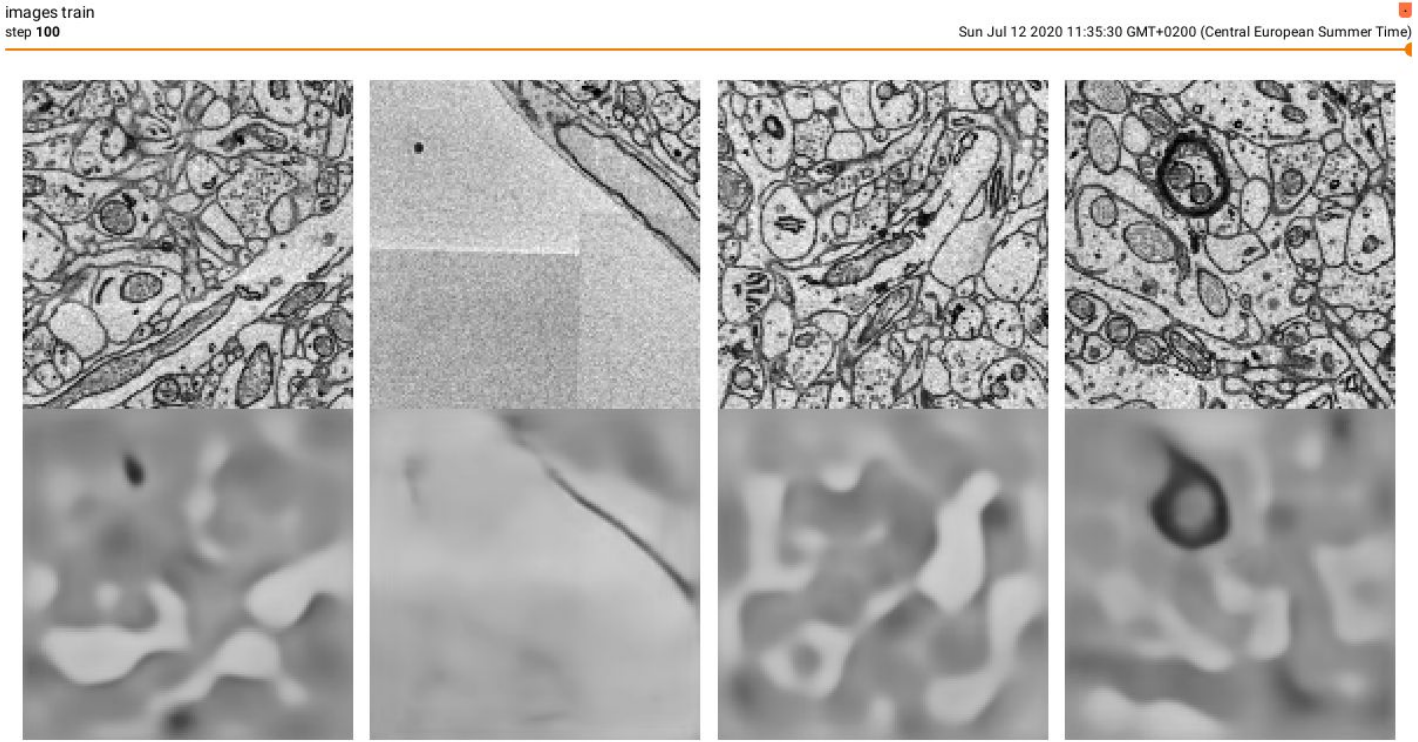
Basic multi-res AE for mSEM data



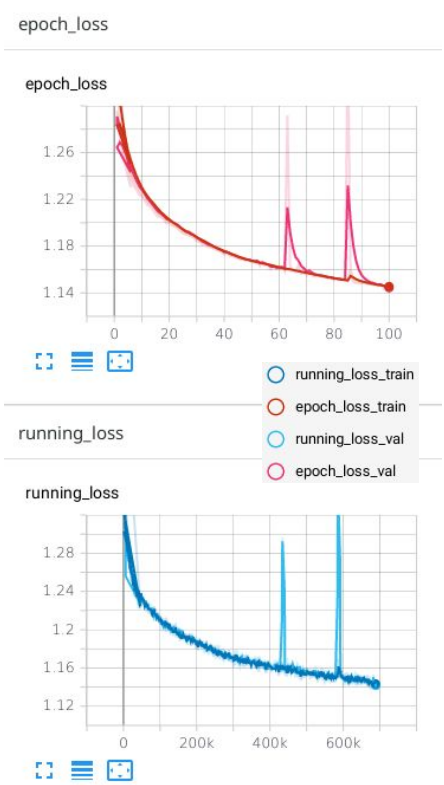


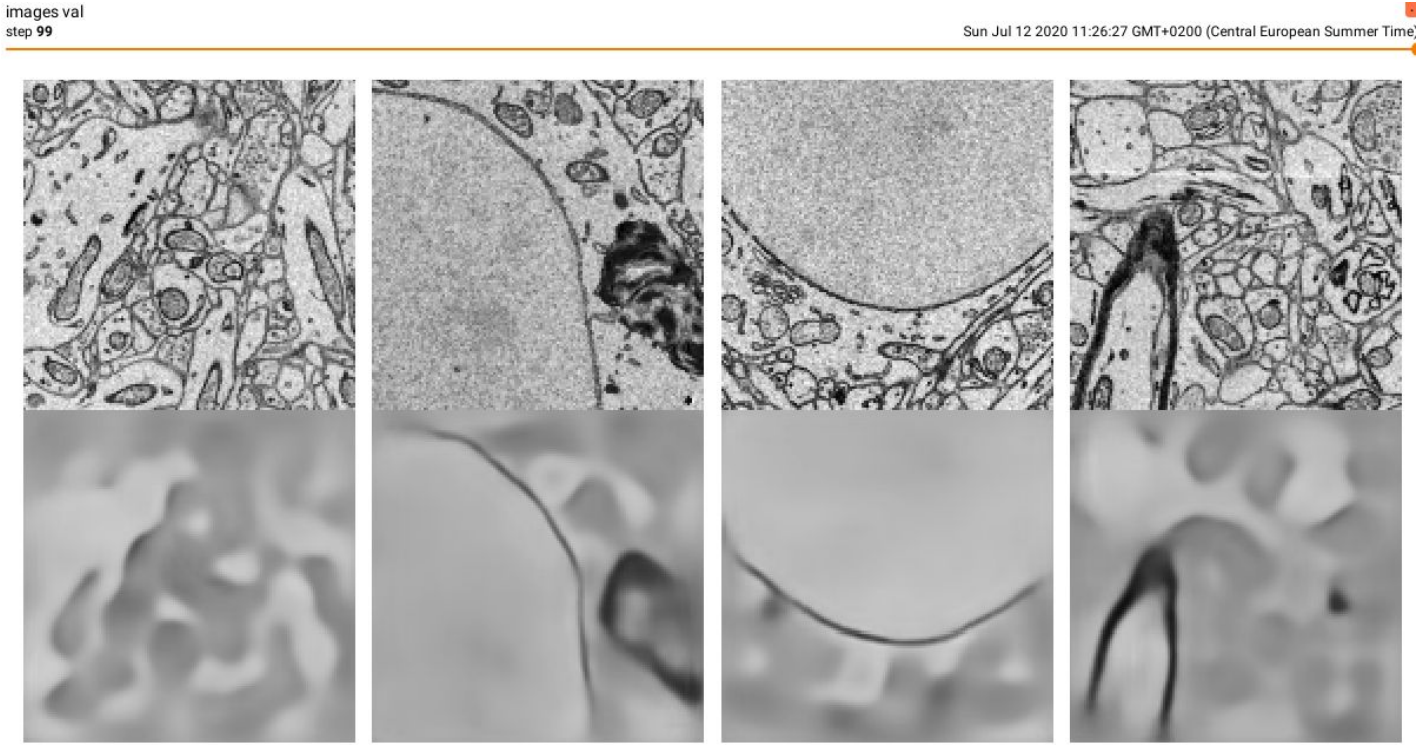
Basic multi-res AE for mSEM data



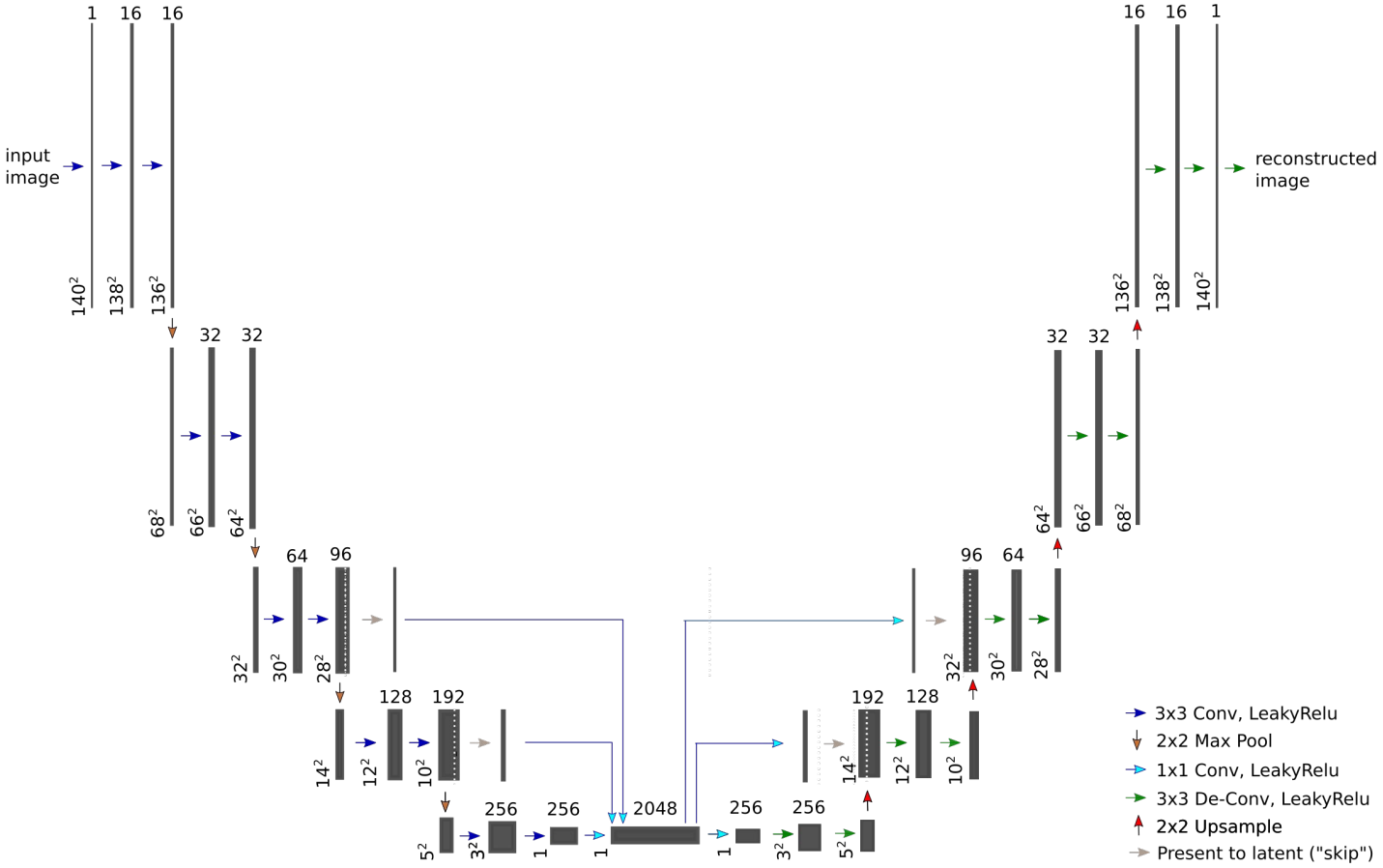


Basic multi-res AE for mSEM data

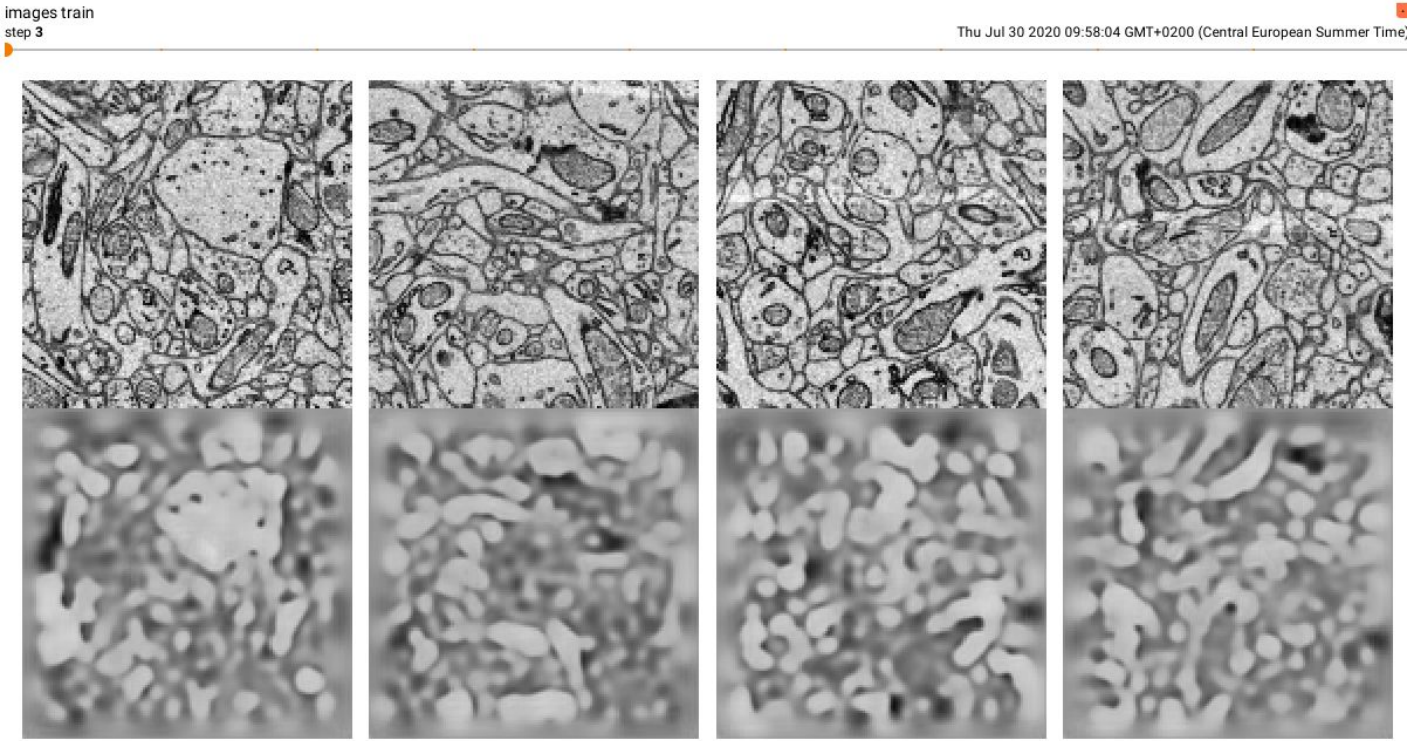




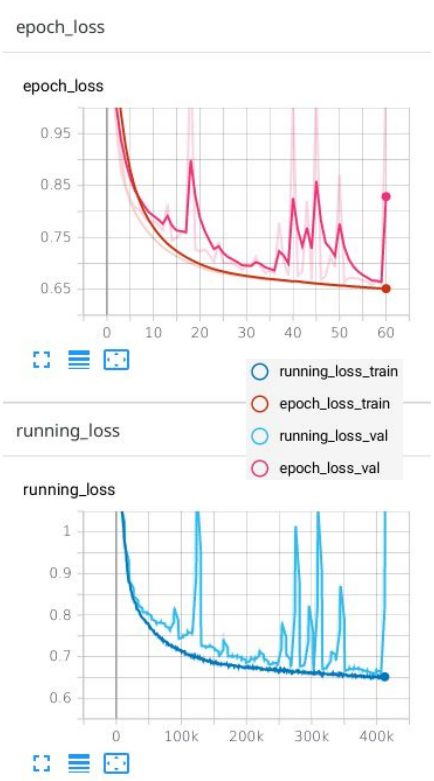
Multi-res “skip” AE for mSEM data

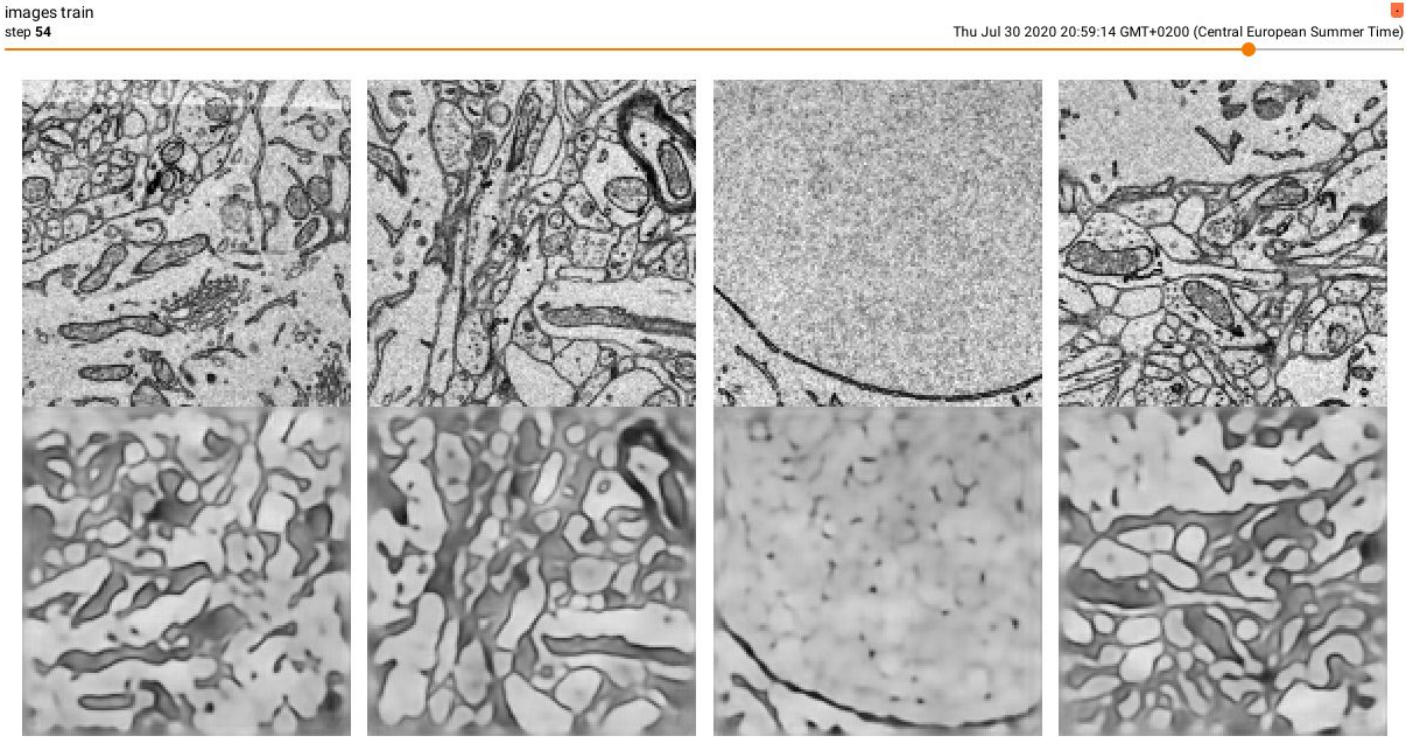


Multi-res “skip” AE for mSEM data

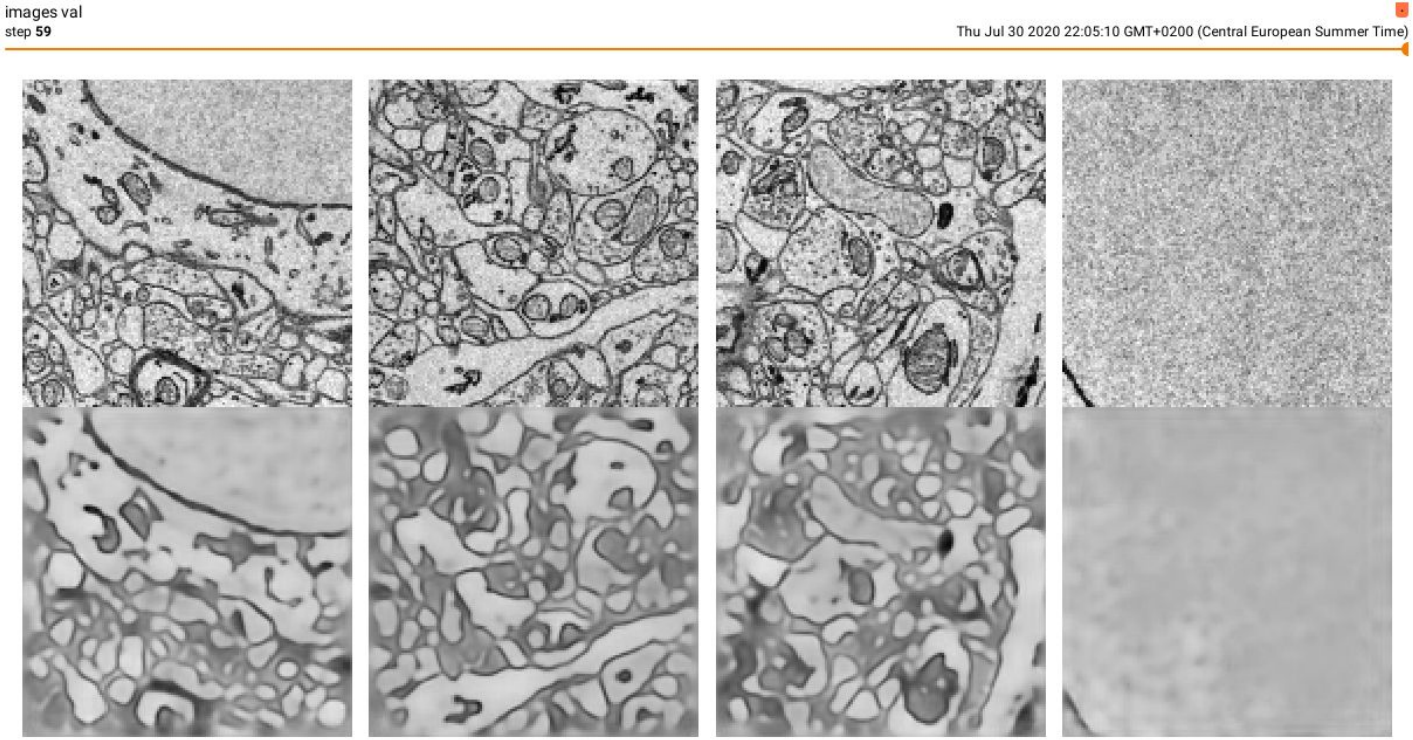


Multi-res “skip” AE for mSEM data

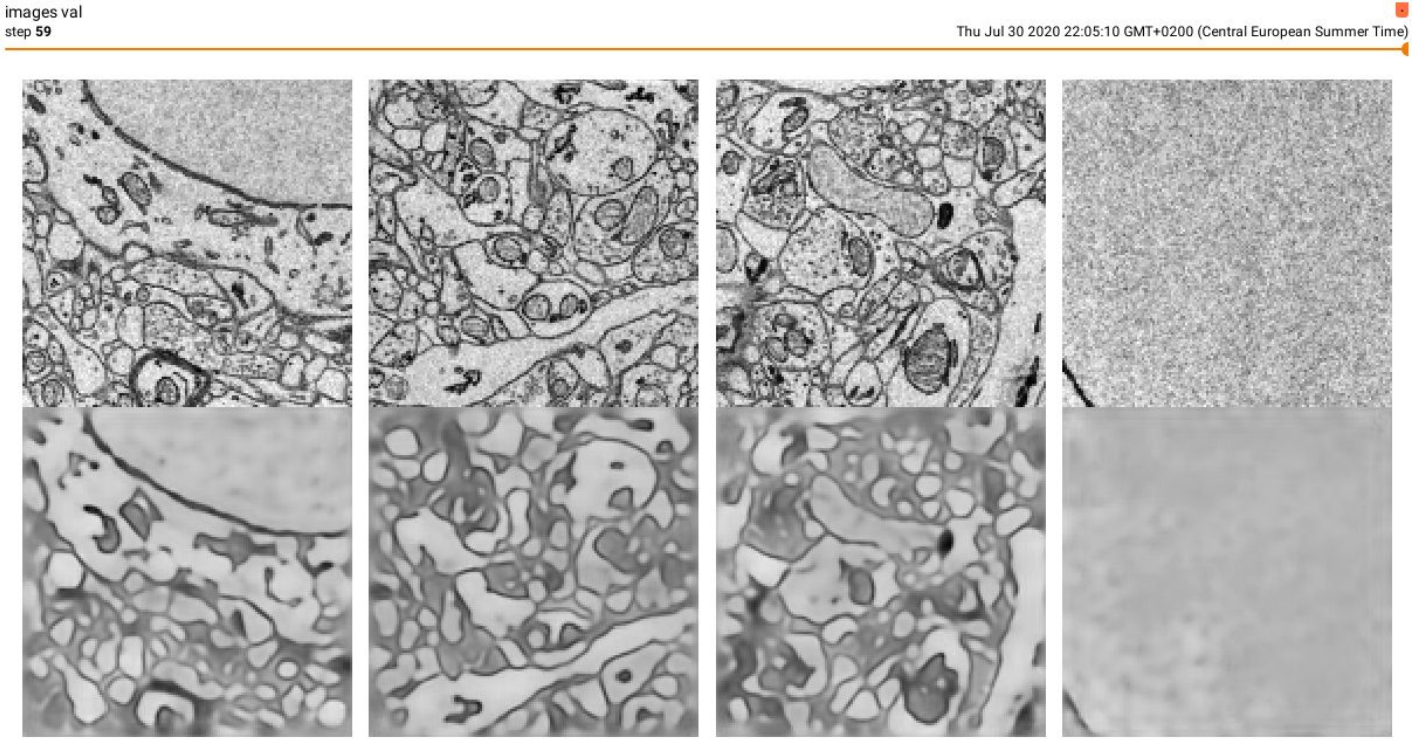




Multi-res “skip” AE for mSEM data



Multi-res “skip” AE for mSEM data



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

$e: X \rightarrow L$

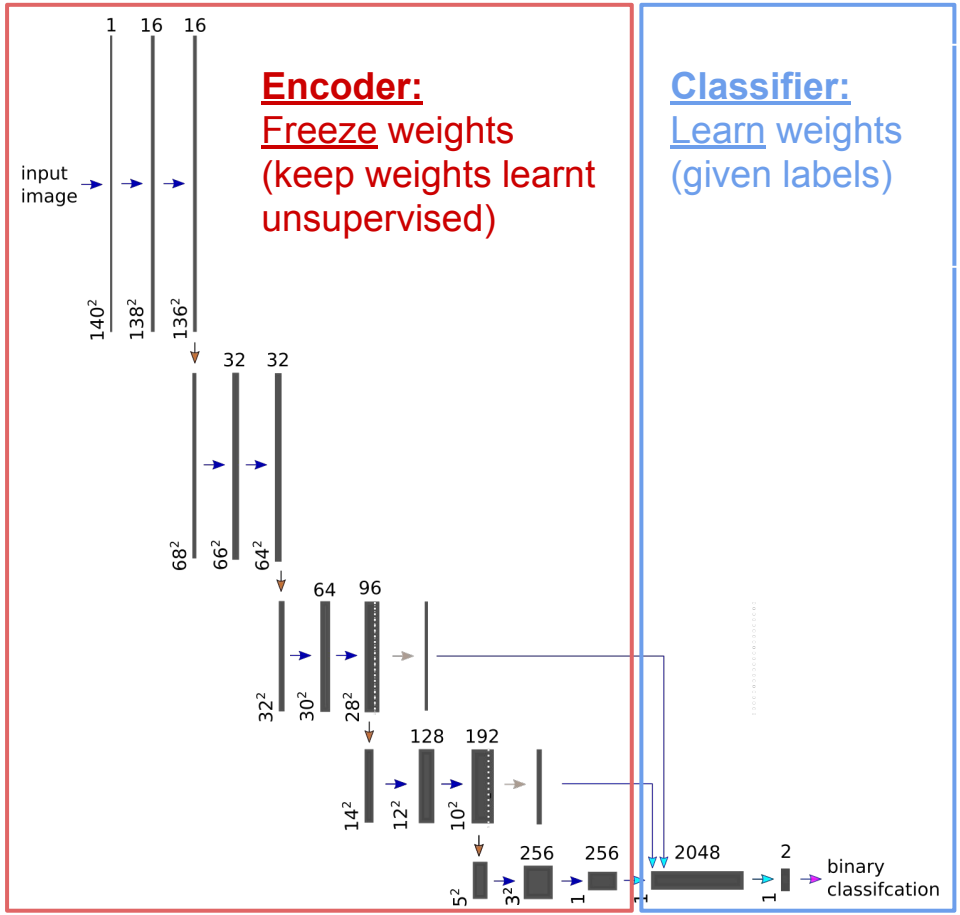
$c: L \rightarrow Y$

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

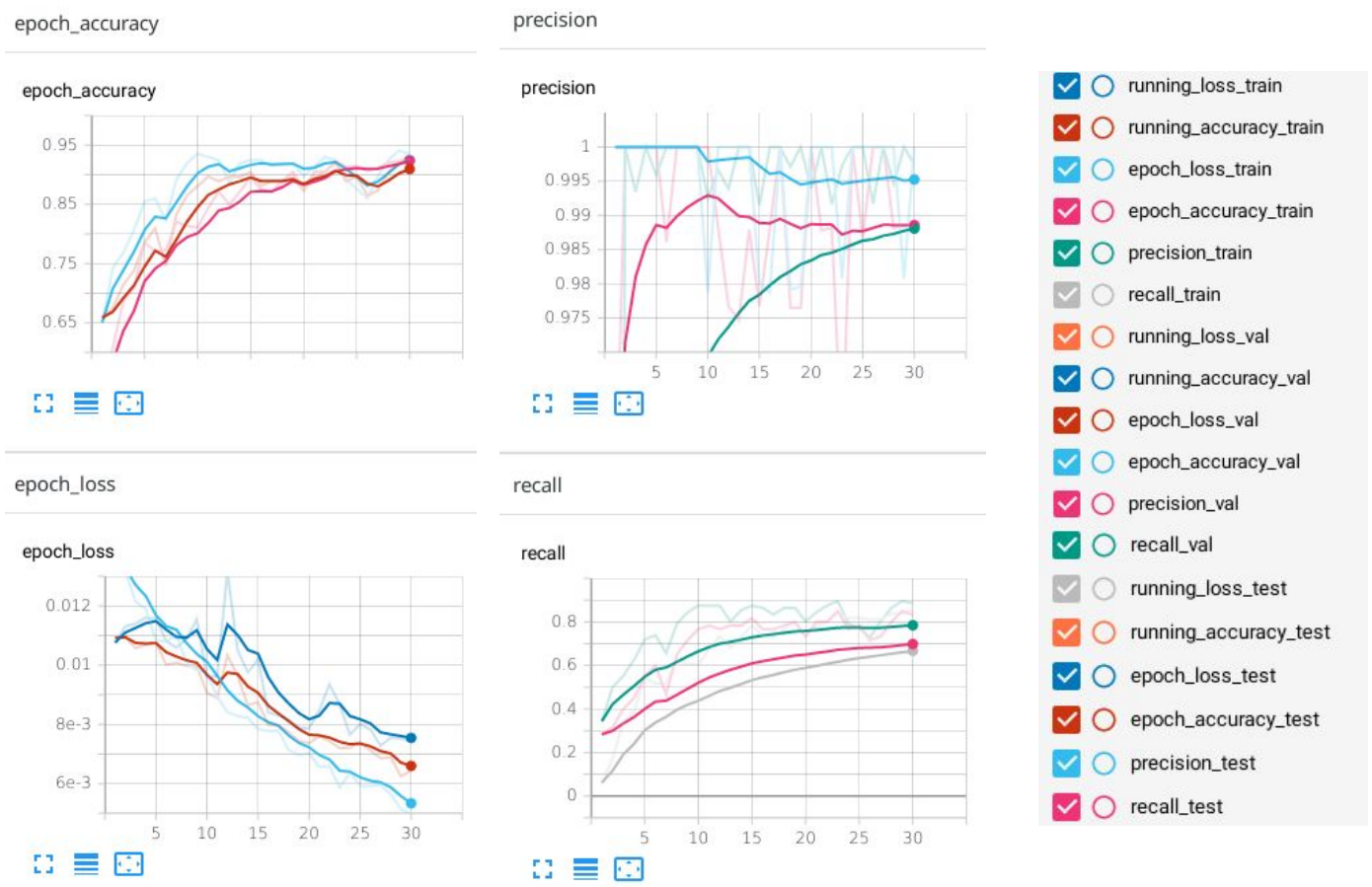
$\text{Min}_p \text{Loss}(c'_p(e'(x)) - x)$

```
class EncoderClassifier:
  def predict(x)
      y_hat = c(e(x))
      return y_hat
```

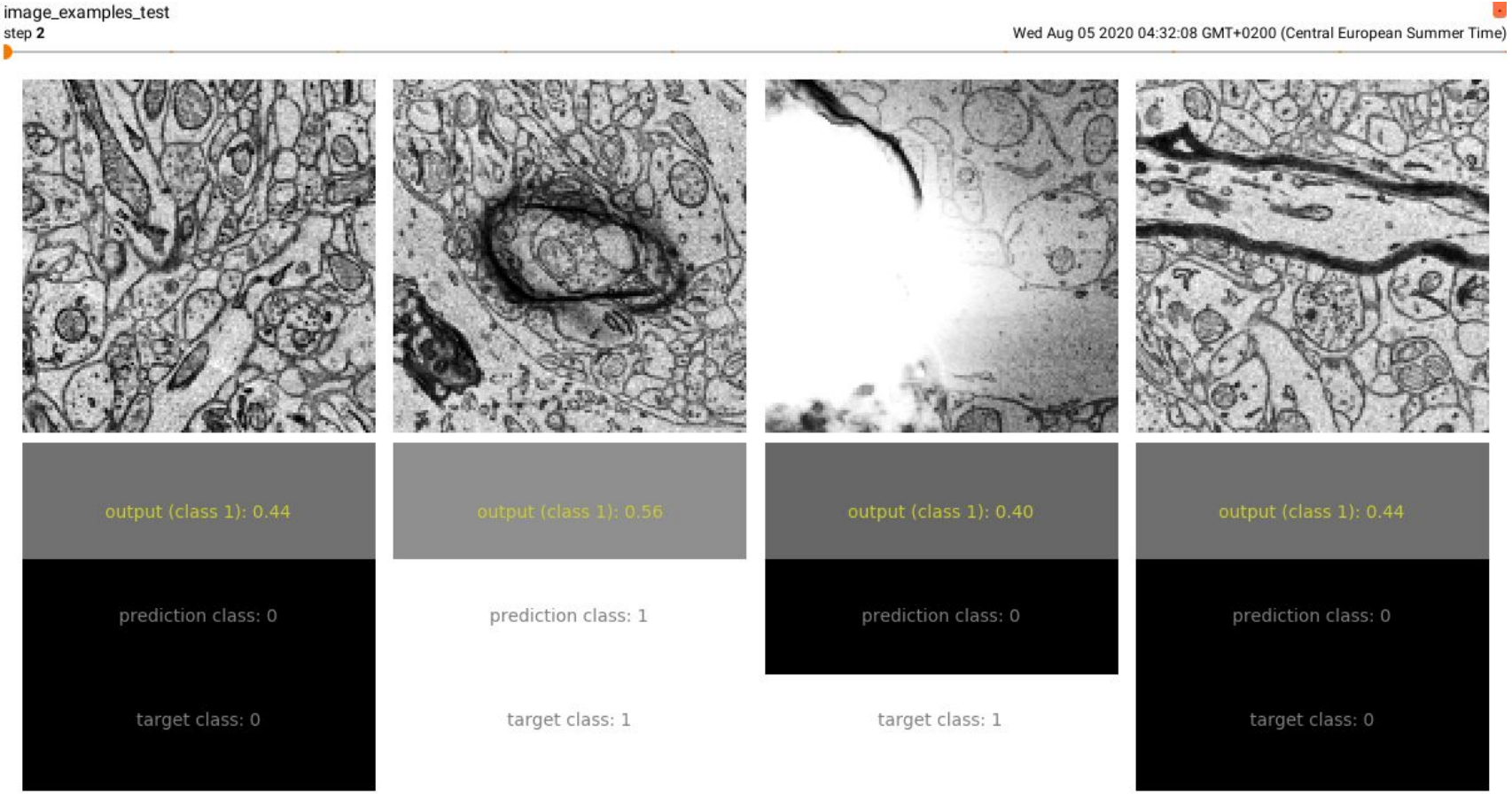
Multi-res “skip” encoder + binary classifier for mSEM data



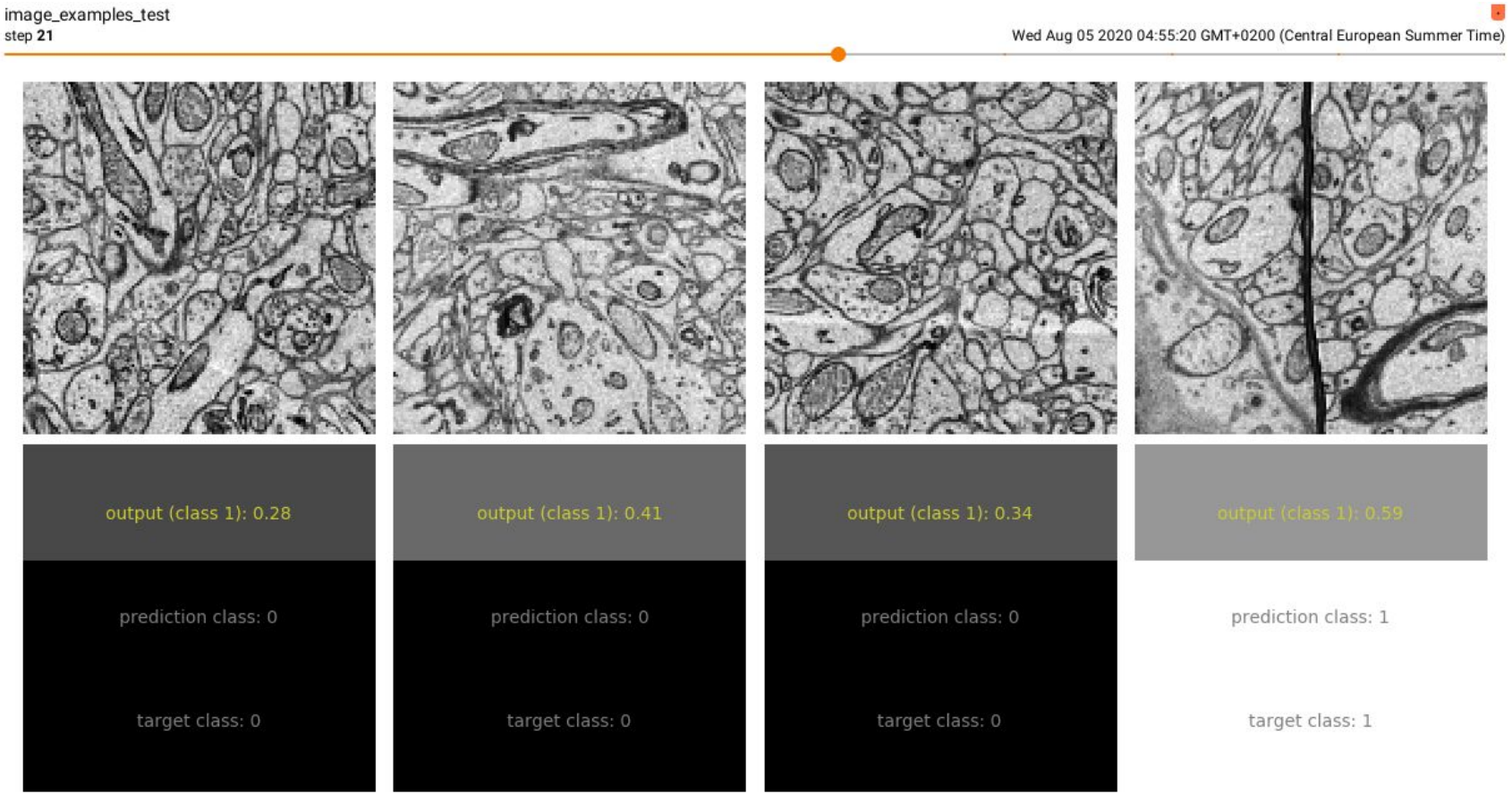
Multi-res “skip” encoder + binary classifier for mSEM data



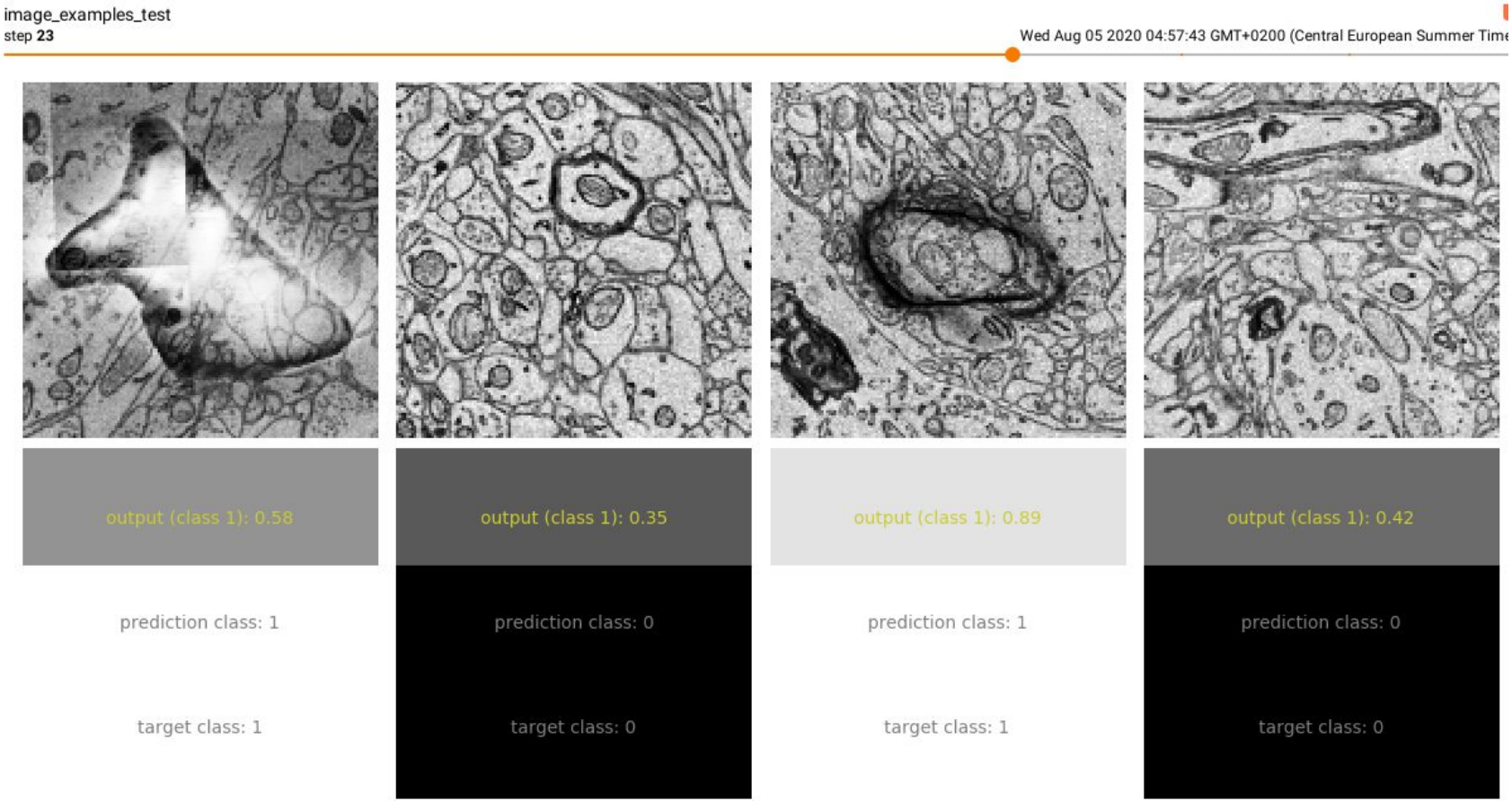
Multi-res “skip” encoder + binary classifier for mSEM data



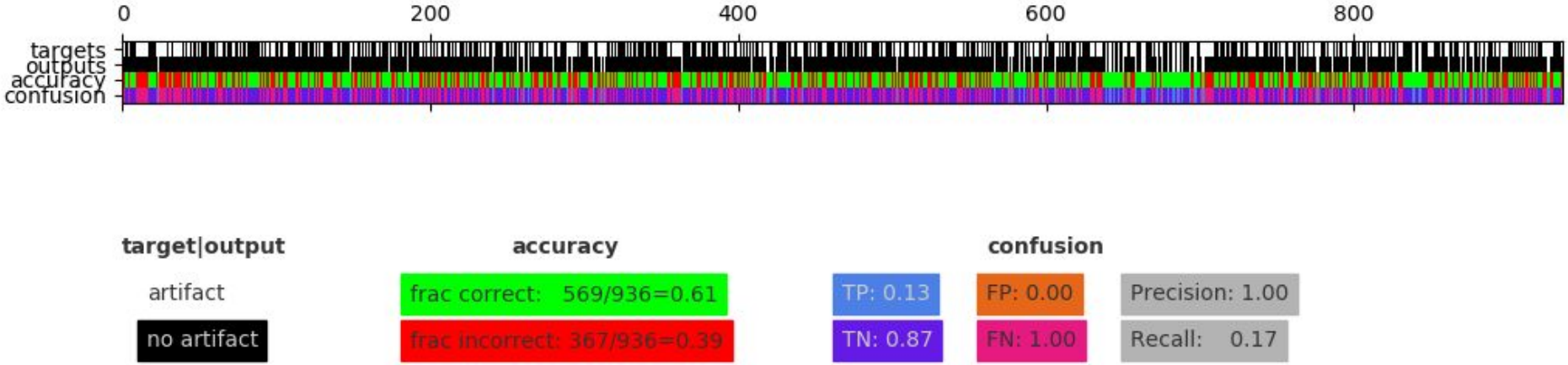
Multi-res “skip” encoder + binary classifier for mSEM data



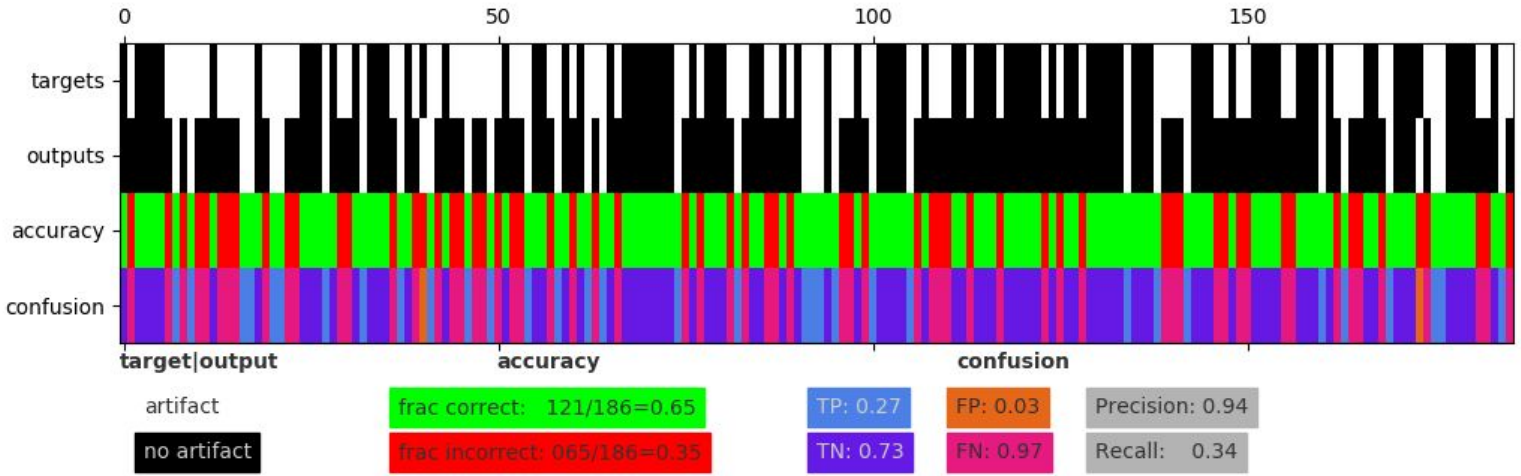
Multi-res “skip” encoder + binary classifier for mSEM data



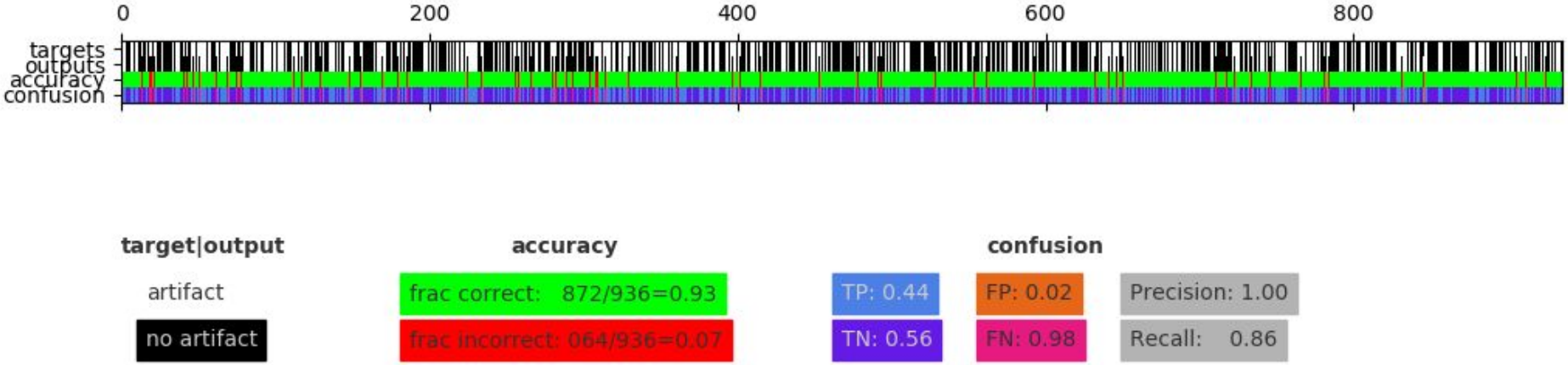
Multi-res “skip” encoder + binary classifier for mSEM data



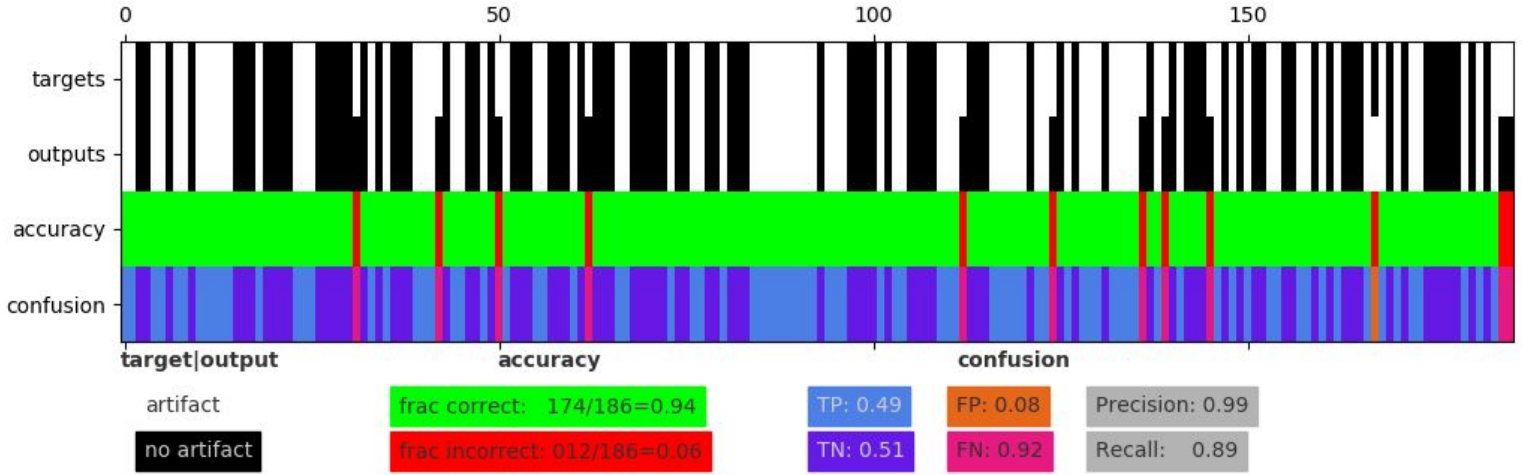
Multi-res “skip” encoder + binary classifier for mSEM data



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



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



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