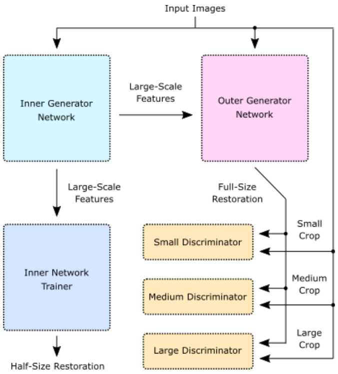
Partial Scan Electron Microscopy with Deep Learning

Jeffrey M. Ede
j.m.ede@warwick.ac.uk

Aim: Reduce electron beam exposure in beam-sensitive applications by training a generative adversarial network to realistically restore electron micrographs from partial scans.

Generative Adversarial Network



Simplified multi-scale generative adversarial network. An inner network generates large-scale features from inputs. These are translated to half-size restorations by a trainer network and recombined with the input to generate full-size restorations by an outer network. Multiple discriminators assess multi-scale crops from full-size real and generated images.

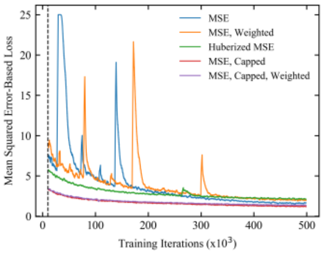
Adversarial roles:

- Micrographs are restored by a two-stage generator that tries to make micrographs that look realistic to multiscale discriminators.
- Multiscale discriminators try to predict whether micrographs are real or generated.

Discriminators only assess whether generated micrographs look realistic; not if they are correct. This can lead to mode collapse: where the generator only produces a subset of outputs. To lift the degeneracy, we add a mean squared error between restored and ground truth images.

Learning Curves

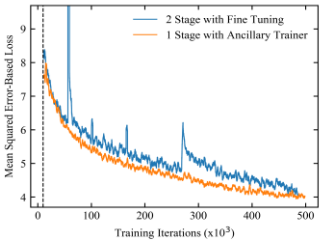
Architecture and learning policy optimizations were selected by comparing learning curves for partial scans with 1/20 coverage. The final architecture is the result of 90+ experiments.



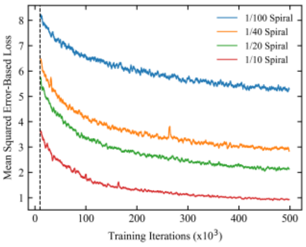
Clipping gradients with an adaptive learning rate stabilizes convergence and results in lower errors than Huberization. Weighting errors by their running means has little effect.

Algorithm 1 Adaptive learning rate clipping (ALRC) of outlier losses. In our experiments, $\mu_1 = 25$, $\mu_2 = 30^2$, $\beta_1 = \beta_2 = 0.999$ and $n = 3$ after initialization.

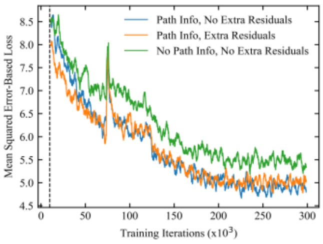
```
Initialize running means, $\mu_1$ and $\mu_2$, with decay rates, $\beta_1$ and $\beta_2$.
Choose number, $n$, of standard deviations to clip to.
while Training is not finished do
    Infer forward-propagation loss, $L$.
    $\sigma \leftarrow (\mu_2 - \mu_1^2)^{1/2}$
    $L_\text{max} \leftarrow \mu_1 + n\sigma$
    if $L > L_\text{max}$ then
        $L_\text{dyn} \leftarrow \text{stop\_gradient}(L_\text{max}/L)L$
    else
        $L_\text{dyn} \leftarrow L$
    end if
    Optimize network by back-propagating $L_\text{dyn}$.
    $\mu_1 \leftarrow \beta_1\mu_1 + (1 - \beta_1)L$
    $\mu_2 \leftarrow \beta_2\mu_2 + (1 - \beta_2)L^2$
end while
```



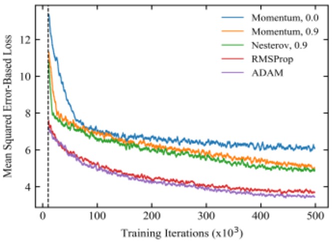
Ancillary inner network guided training is more stable and reaches a lower error than two-stage training with fine-tuning.



Increasing spiral path coverage decreases training errors.



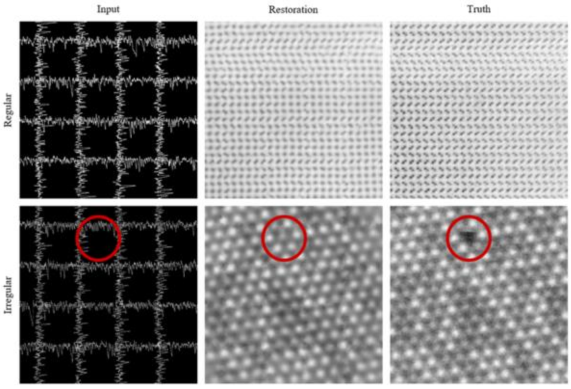
Errors are lower if electron beam path information is concatenated to the network input. Errors are higher for extra residual connections between inner network strided convolutions and mirroring transpositional convolutions.



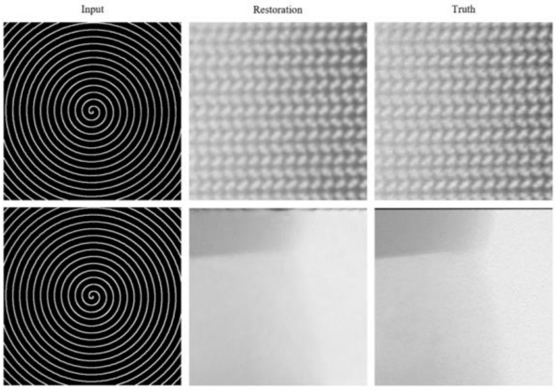
Adaptive momentum-based optimizers, ADAM and RMSProp, outperform non-adaptive momentum-based optimizers, including Nesterov-accelerated momentum. ADAM outperforms RMSProp; however, training hyperparameters and learning protocols were tuned for ADAM.

Example Scans

Electron micrographs can be generated from partial scans, decreasing electron beam damage and acquisition time. Images with regular patterns or predictable structure can be accurately restored with lower coverage. In practice, electron beam exposure can be reduced by more than 20x without noticeable information loss.



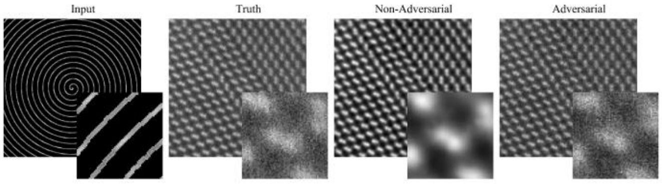
Non-adversarial 512x512 restorations for randomly perturbed grids with 1/20 coverage. Images with regular patterns or predictable structure are accurately restored. Circles accentuate that the network cannot reliably restore irregular images where there is no information.



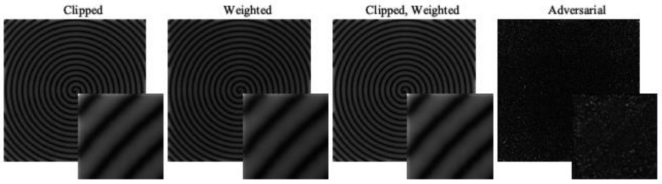
Non-adversarial 512x512 restorations for blurred spiral scans with 1/20 coverage. More uniform coverage allows irregular patterns and unpredictable structure to be restored more accurately.

Performance

Mean squared error training produces blurry images with higher errors away from electron beam paths. In contrast, adversarial training produces images with realistic noise characteristics and less structured errors.



Adversarial and non-adversarial infilling of a blurred test set 512x512 1/20 coverage spiral scan. Adversarial training produces an image with realistic noise characteristics whereas non-adversarial training produces a blurred image. Enlarged 64x64 regions from the top left of each image are inset to ease comparison.



Mean squared errors on 20000 512x512 test set micrographs for training with adaptive learning rate clipping, increased weighting for pixels with high running mean errors, both, and adversarial training. For comparison, errors are linearly transformed to have the same mean and variance. Non-adversarial errors increase away from the electron beam path. In contrast, adversarial errors are less structured. Enlarged 64x64 regions from the top left of each image are inset to ease comparison.

Conclusions

This poster shows that generative adversarial networks can generate realistic micrographs from multiple types of partial scan and for 1/10 to 1/100 coverage. This will enable faster imaging and reduced electron beam damage in low dose applications.

Adversarial training produces more realistic micrographs with lower systematic errors than non-adversarial training. Learning curves demonstrate multiple methods to improve small minibatch training, including adaptive learning rate clipping of outlier losses and the use of an ancillary trainer network.