



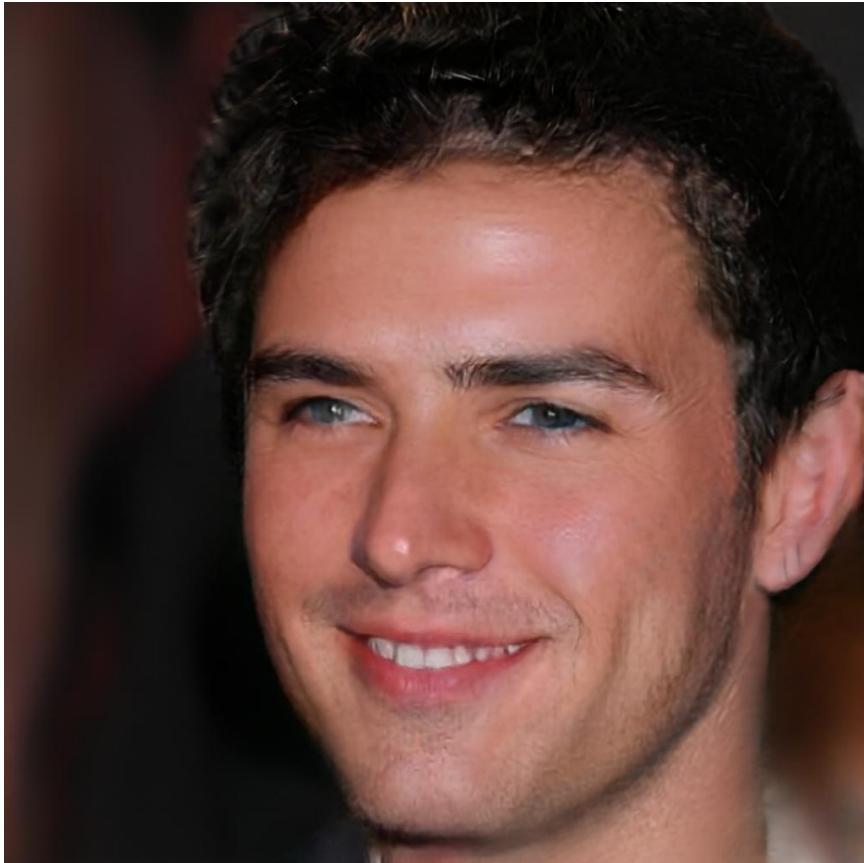
MIXED PRECISION TRAINING FOR PIX2PIXHD

Ting-Chun Wang

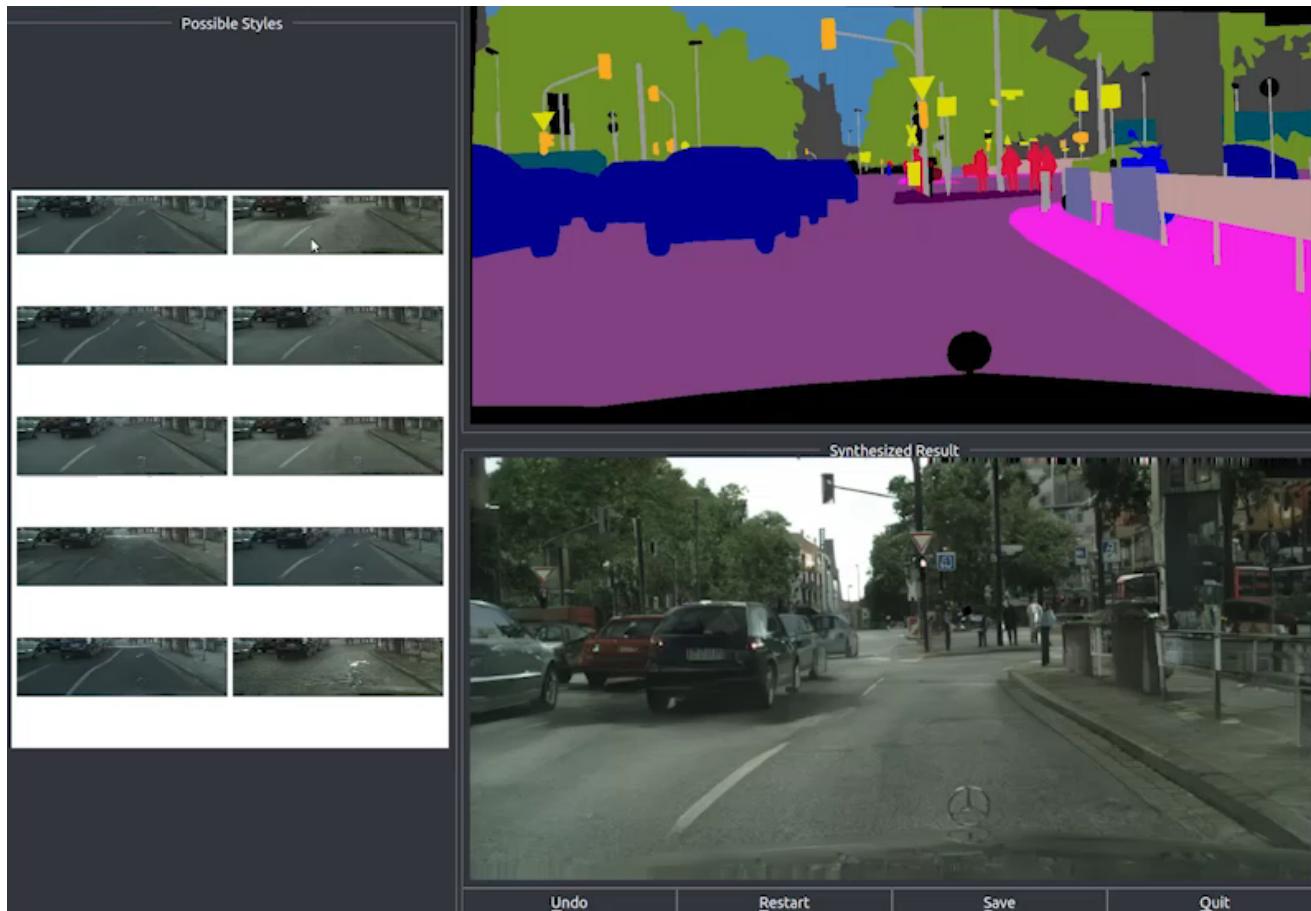
INTRODUCTION



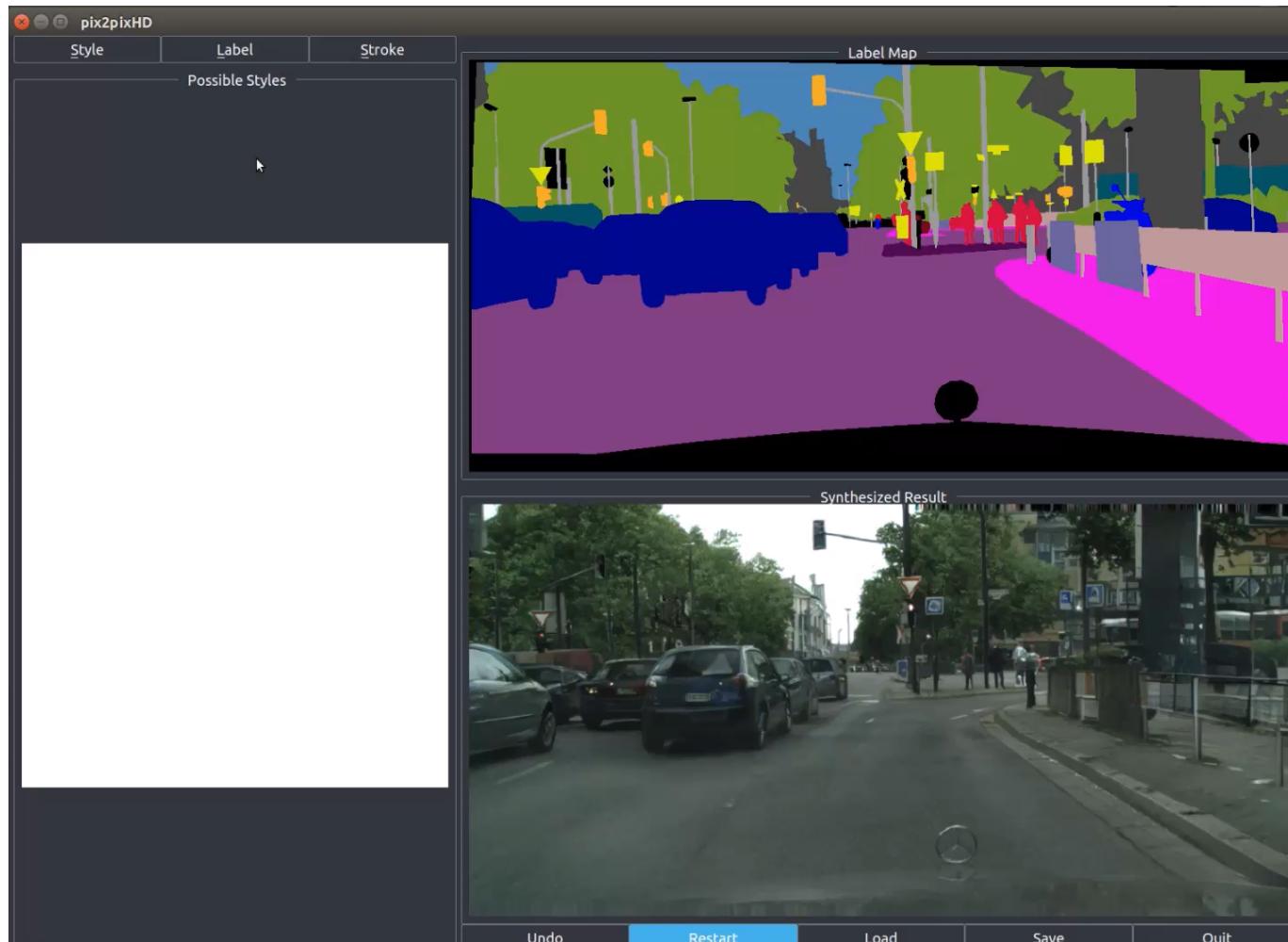
INTRODUCTION



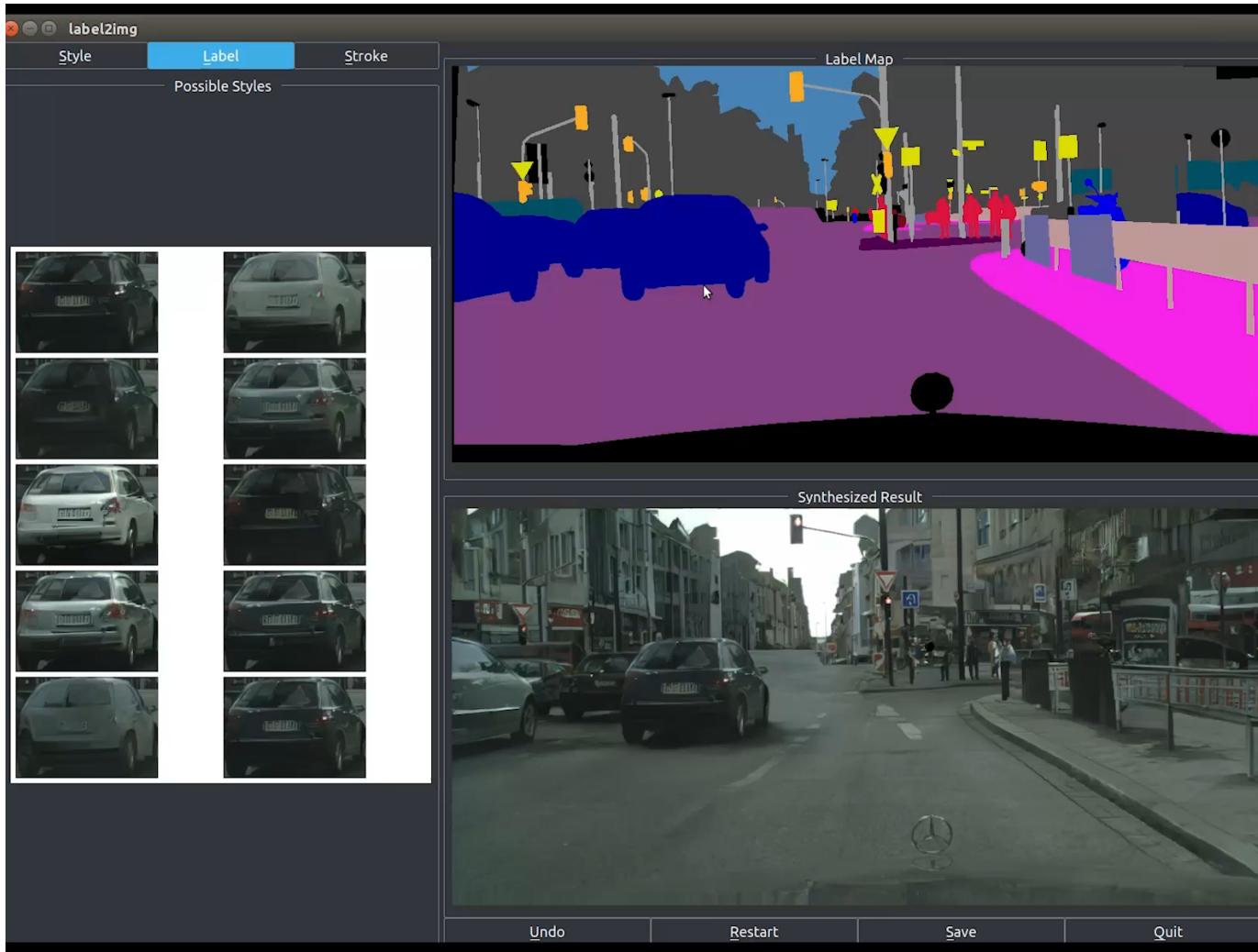
APPLICATIONS: STYLE CHANGING



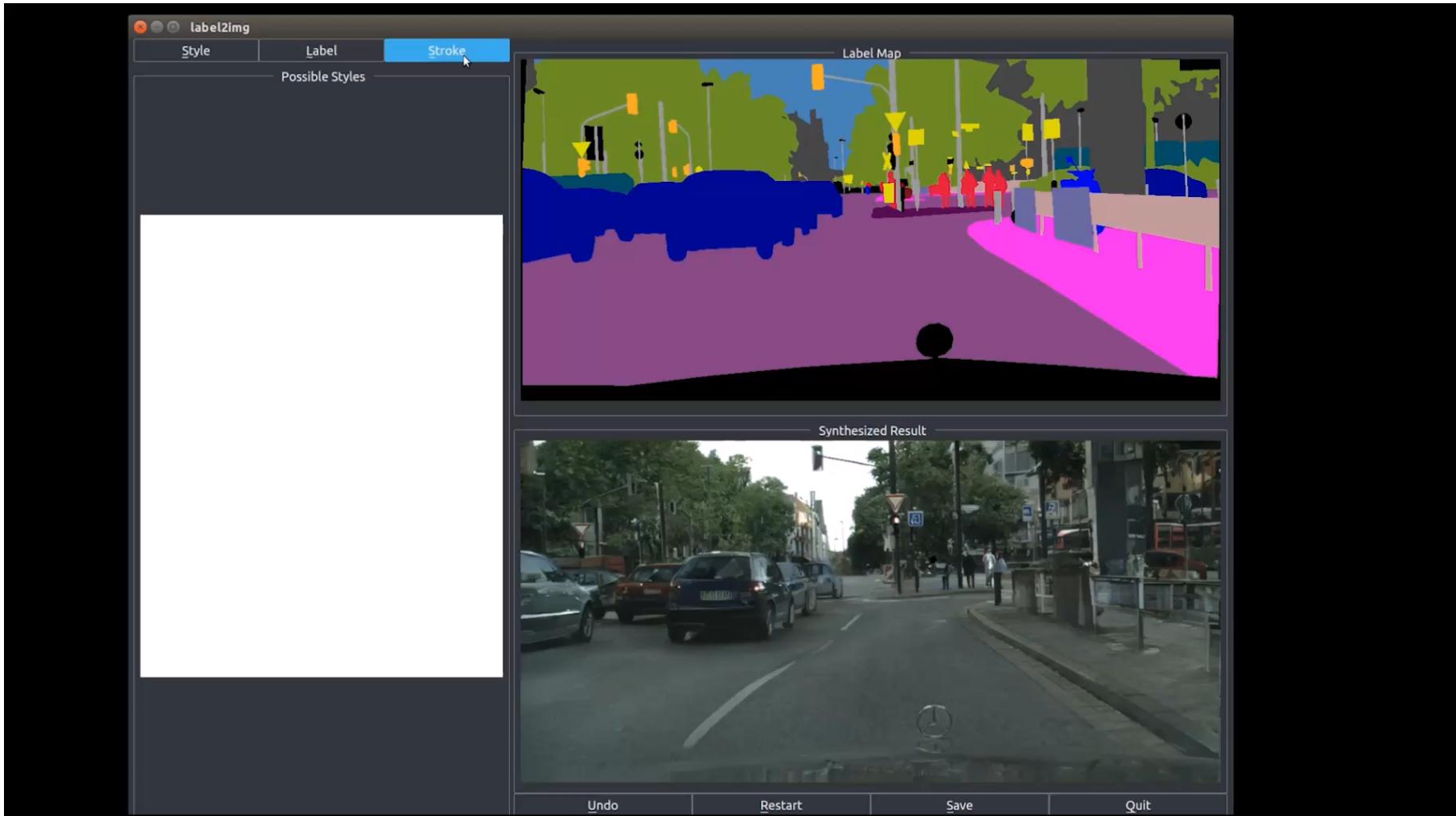
APPLICATIONS: LABEL CHANGING



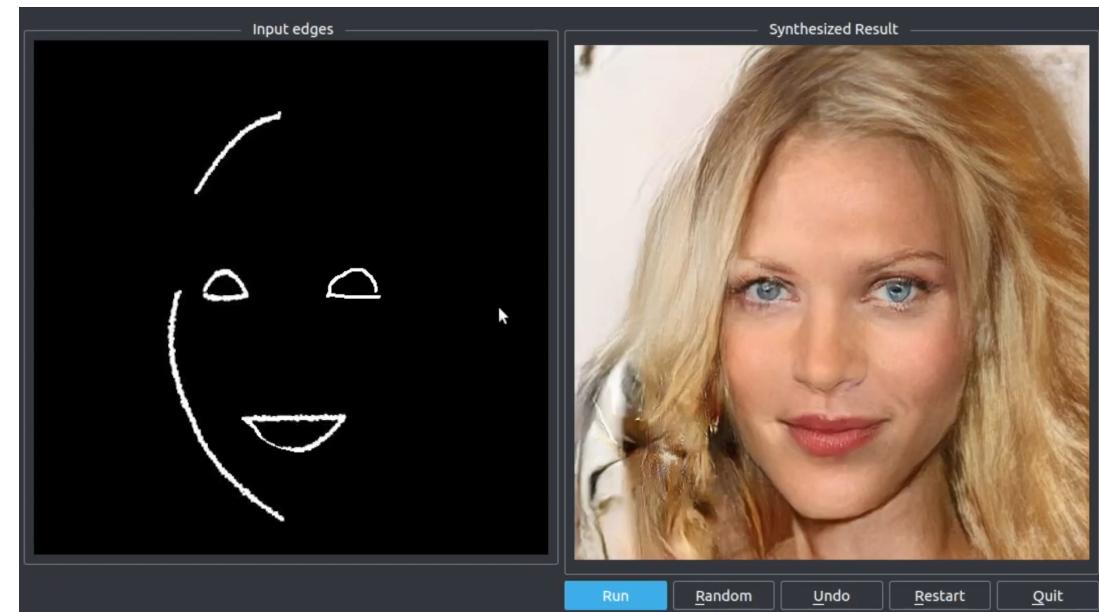
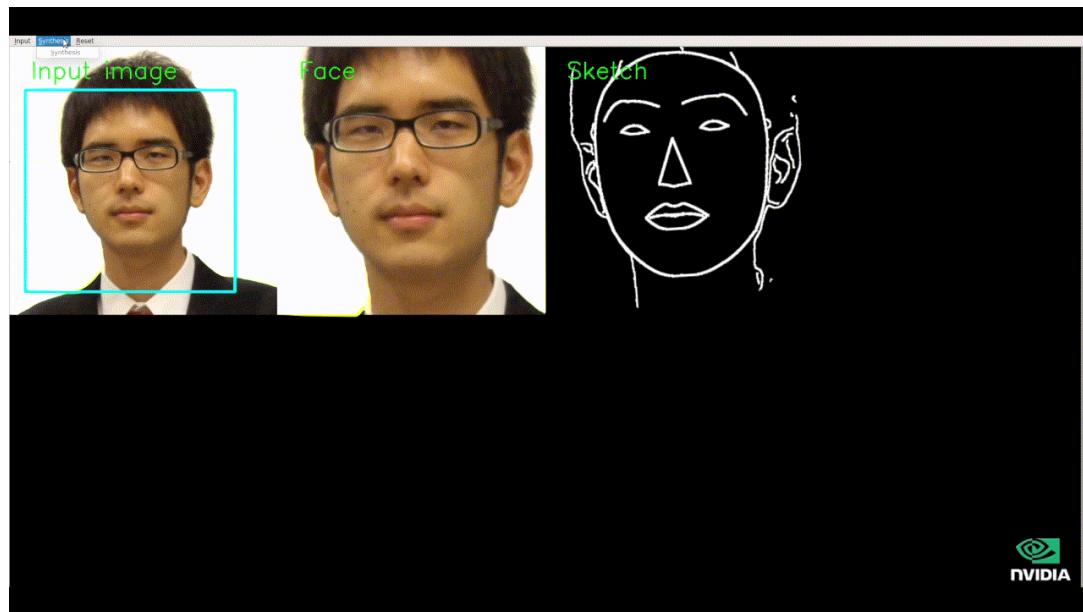
APPLICATIONS: ADDING OBJECTS



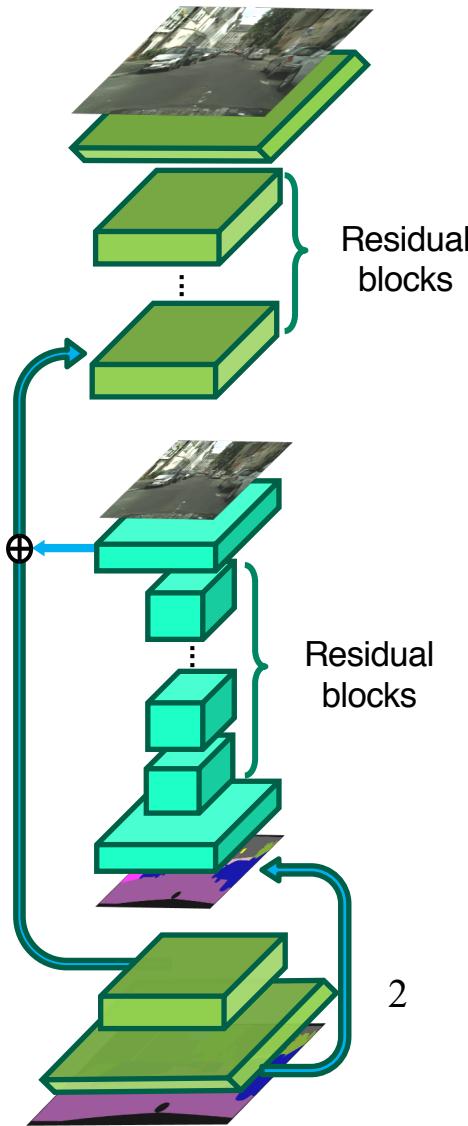
APPLICATIONS: ADDING STROKES



APPLICATIONS: FACE CHANGING



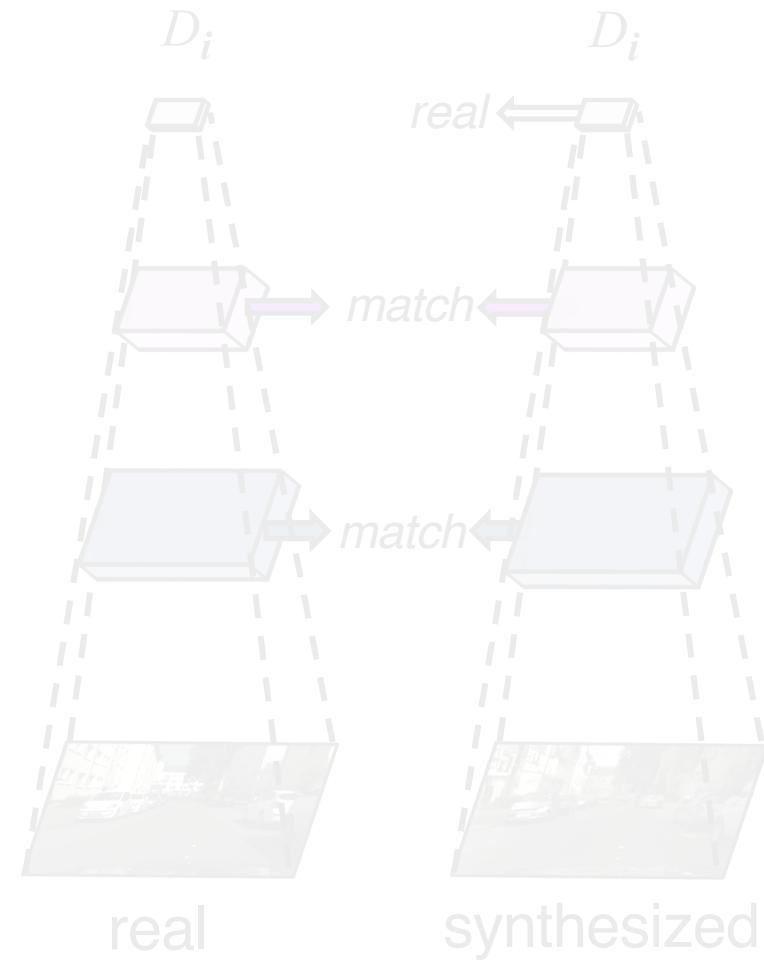
Coarse-to-fine Generator



Multi-scale Discriminators

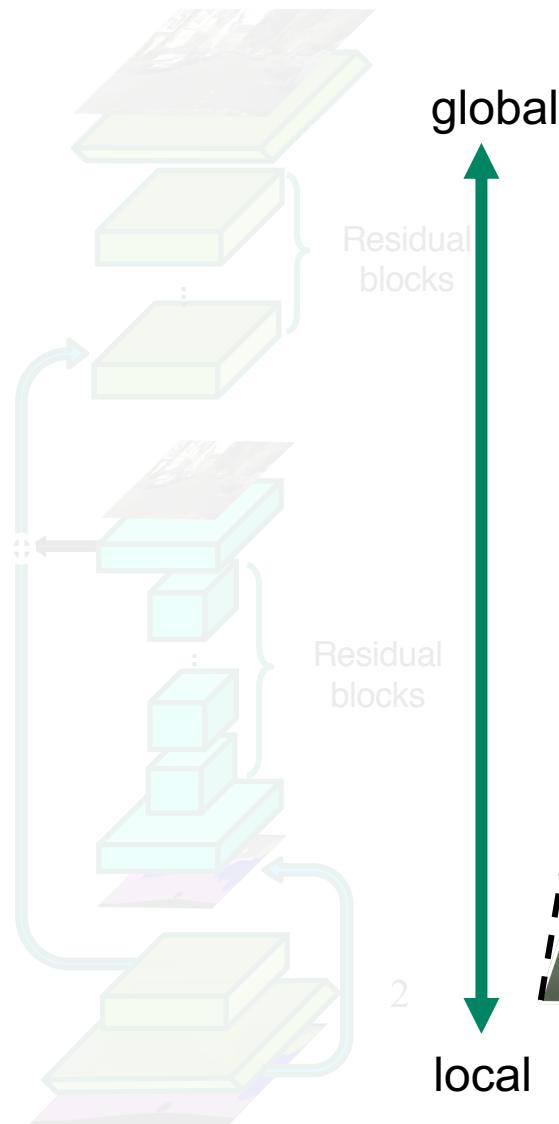


Robust Objective

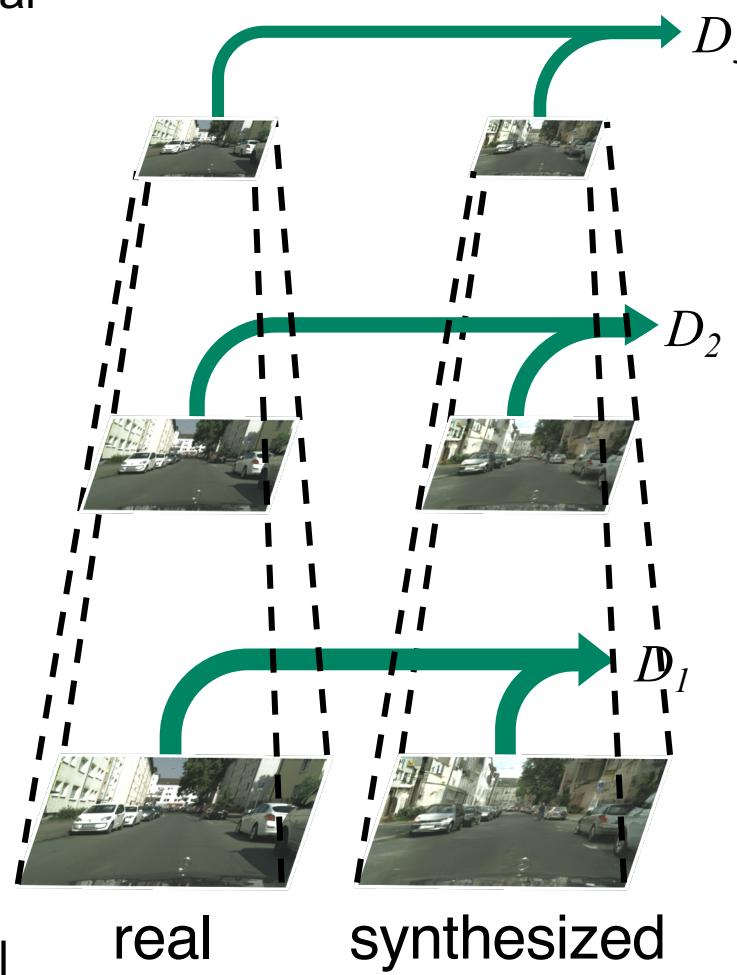


*Similar ideas in Denton et al. 2015, Huang et al. 2017, Chen et al. 2017, Zhang et al. 2017

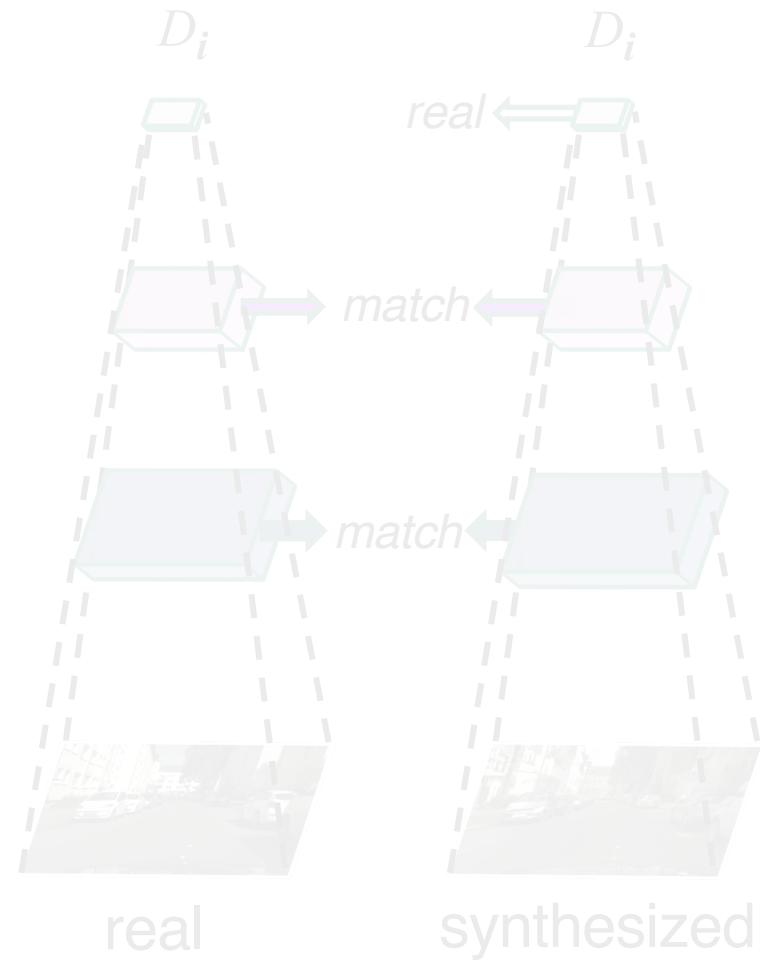
Coarse-to-fine Generator



Multi-scale Discriminators

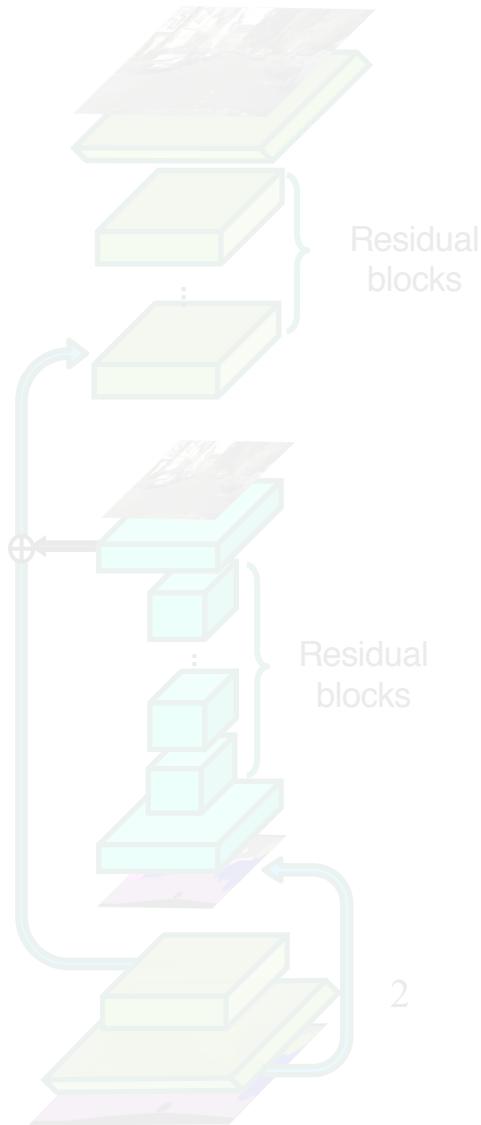


Robust Objective



*Similar ideas in Durugkar et al. 2016, Iizuka et al. 2017, Zhang et al. 2017

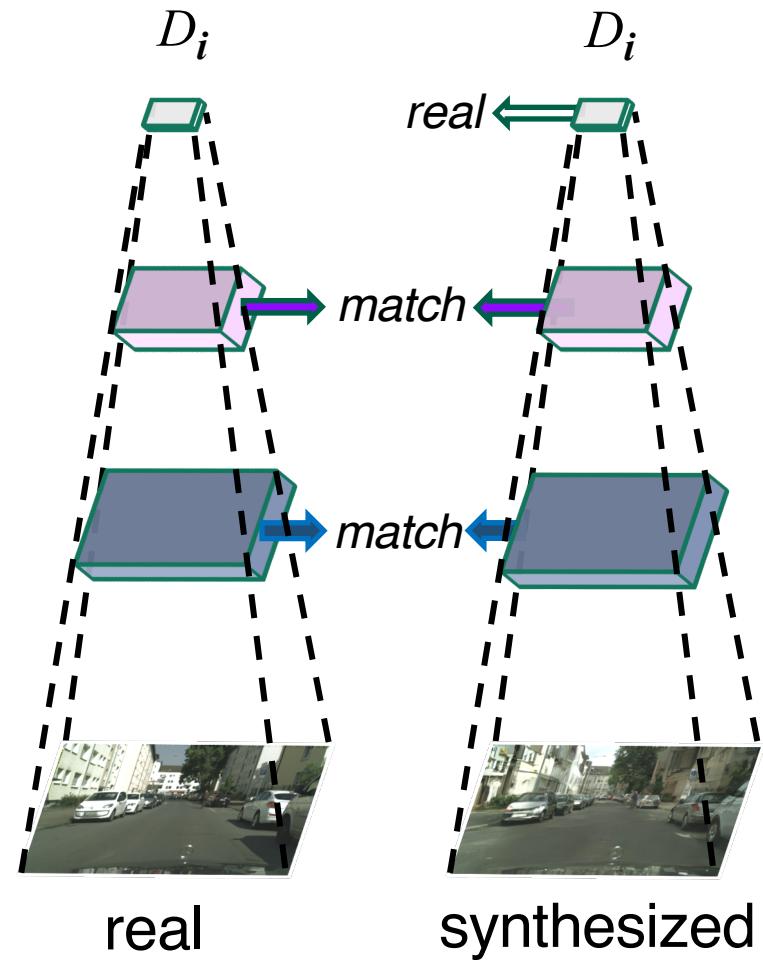
Coarse-to-fine Generator



Multi-scale Discriminators



Robust Objective



*Similar ideas in Larsen et al. 2016

CHANGES REQUIRED

```
from apex import amp
```

```
model, [optimizer_G, optimizer_D] = \  
    amp.initialize(model, [model.optimizer_G, model.optimizer_D],  
                  opt_level=cfg.amp)
```

```
with amp.scale_loss(loss_G, optimizer_G) as scaled_loss:  
    scaled_loss.backward()
```

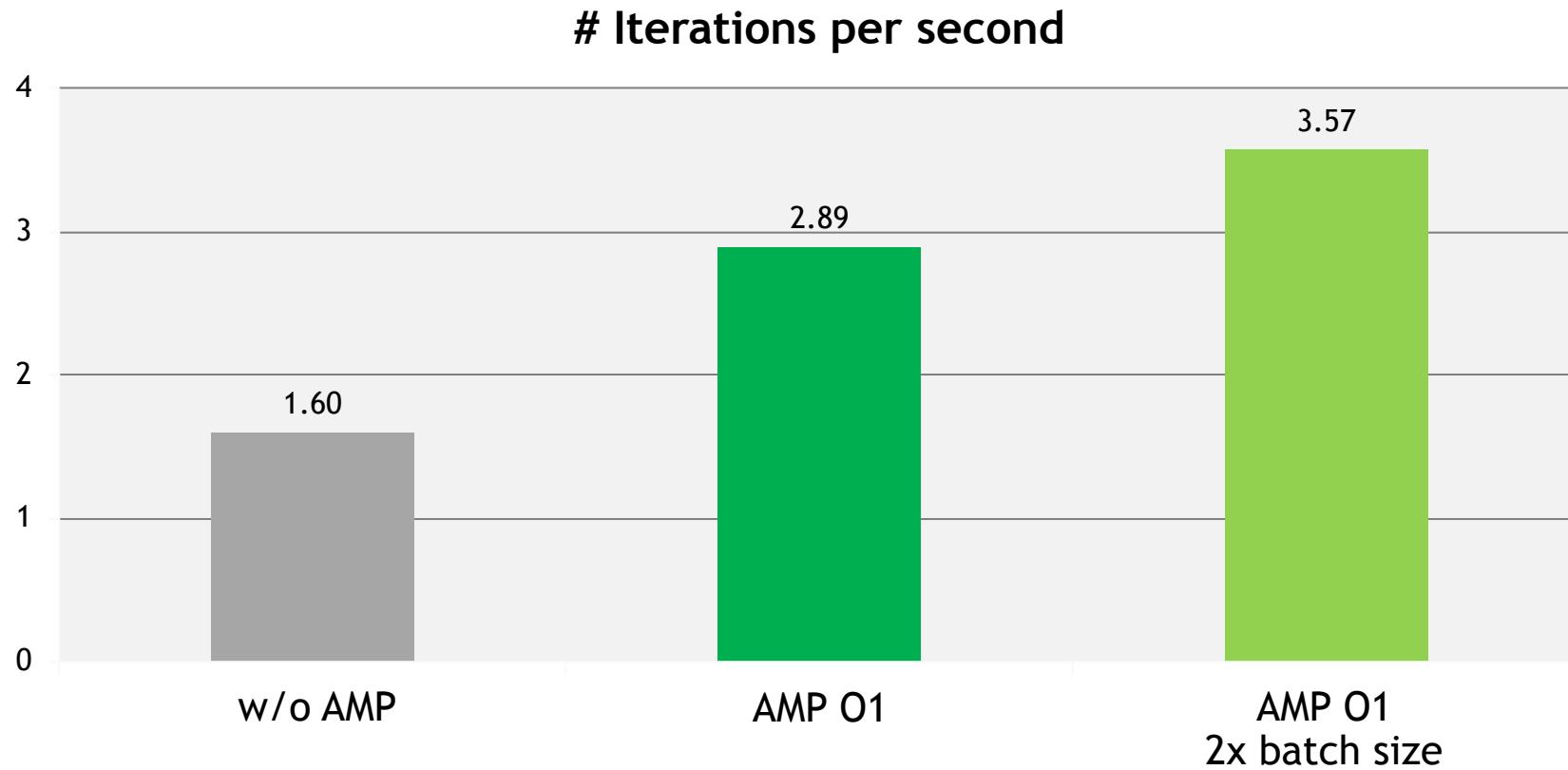
```
with amp.scale_loss(loss_D, optimizer_D) as scaled_loss:  
    scaled_loss.backward()
```

Specs:

Machine: DGX1 V100 16GB
Batch size: 1 image per batch

OBTAINED SPEEDUPS

1.8x faster (2.25x faster with doubled batch size)

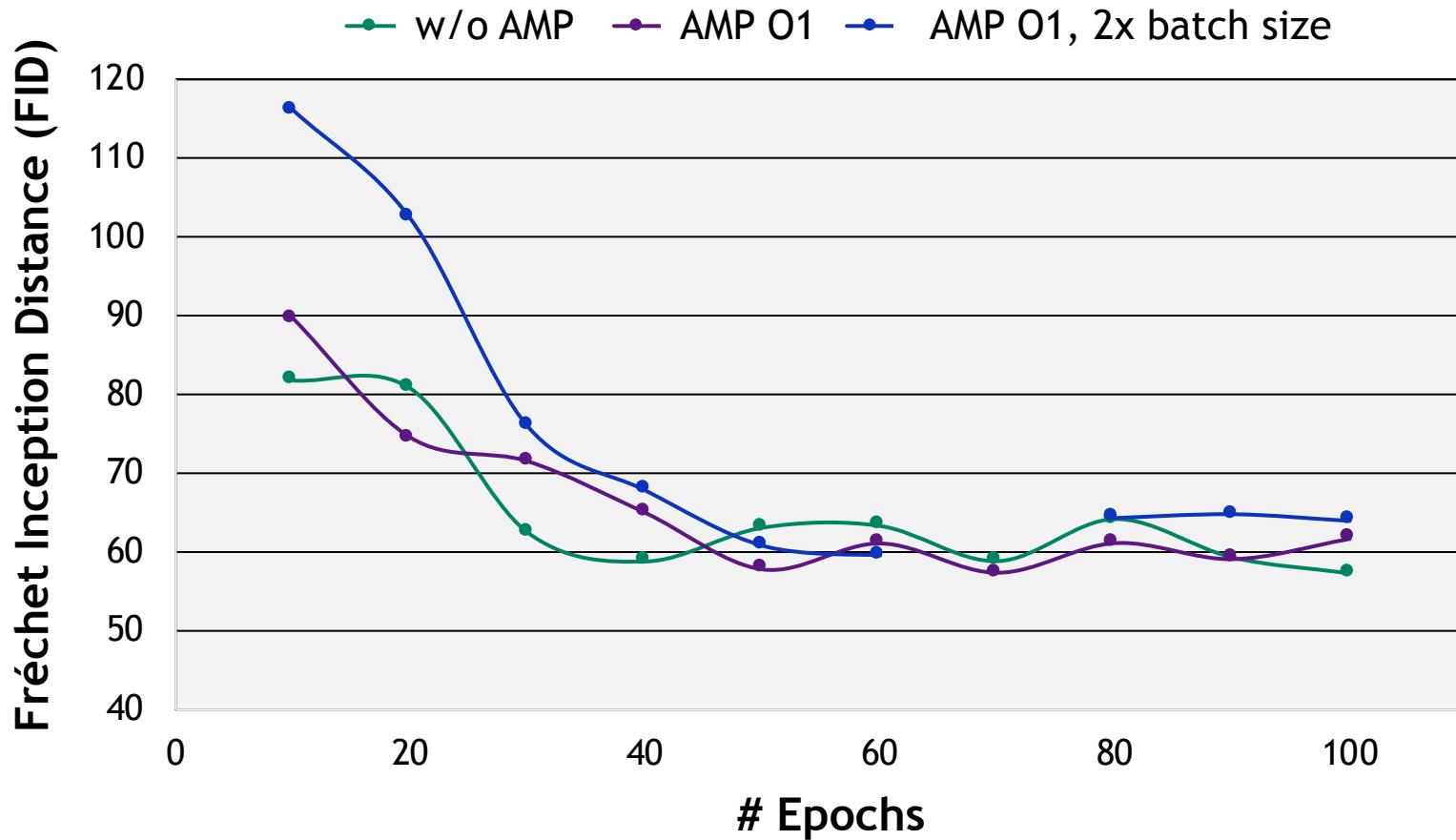


EVALUATION CRITERIA

Frechet Inception Distance

$$\text{FID} = \|\mu_r - \mu_g\|^2 + \text{Tr}(\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{1/2}),$$

EFFECT ON METRICS



- In general, the differences are within the error range
- Usually cannot be differentiated by human eyes

READY TO USE!

<https://github.com/NVIDIA/pix2pixHD>

The screenshot shows the GitHub repository page for 'NVIDIA / pix2pixHD'. The page includes the repository name, a summary of 145 commits, 3,851 stars, and 748 forks. It features tabs for Code, Issues (78), Pull requests (4), Security, Insights, and Settings. Below the tabs, there's a brief description: 'Synthesizing and manipulating 2048x1024 images with conditional GANs' followed by a link. A list of topics is provided: gan, deep-learning, deep-neural-networks, pytorch, pix2pix, image-to-image-translation, generative-adversarial-network, computer-vision, and computer-graphics. The repository stats show 40 commits, 1 branch, 0 releases, 8 contributors, and a view license button. A 'Clone or download' button is also present. The commit history lists several changes, including license changes and AMP support additions. The latest commit was made on Jun 13.

File	Description	Date
data	change license to BSD	3 months ago
datasets/cityscapes	fix inference and multi gpu issues	last year
imgs	pre-release	2 years ago
models	change license to BSD	3 months ago
options	change license to BSD	3 months ago
scripts	add AMP support	4 months ago
util	change license to BSD	3 months ago
.gitignore	first commit	2 years ago
LICENSE.txt	change license to BSD	3 months ago
README.md	update README	4 months ago

Training with Automatic Mixed Precision (AMP) for faster speed

- To train with mixed precision support, please first install apex from: <https://github.com/NVIDIA/apex>
- You can then train the model by adding `--fp16`. For example,

```
#!/scripts/train_512p_fp16.sh
python -m torch.distributed.launch train.py --name label2city_512p --fp16
```

In our test case, it trains about 80% faster with AMP on a Volta machine.

CONCLUSION

Mixed precision training is useful for pix2pixHD training

1. Drop-in replacement when utilizing NVIDIA APEX AMP library

Only need to change 4 lines of code!

2. Training is much faster for the same model and batch size

1.8x speed up

3. Consumes less memory → may train with larger batch size

2.25x speed up using the same model and 2x batch size



NVIDIA®

