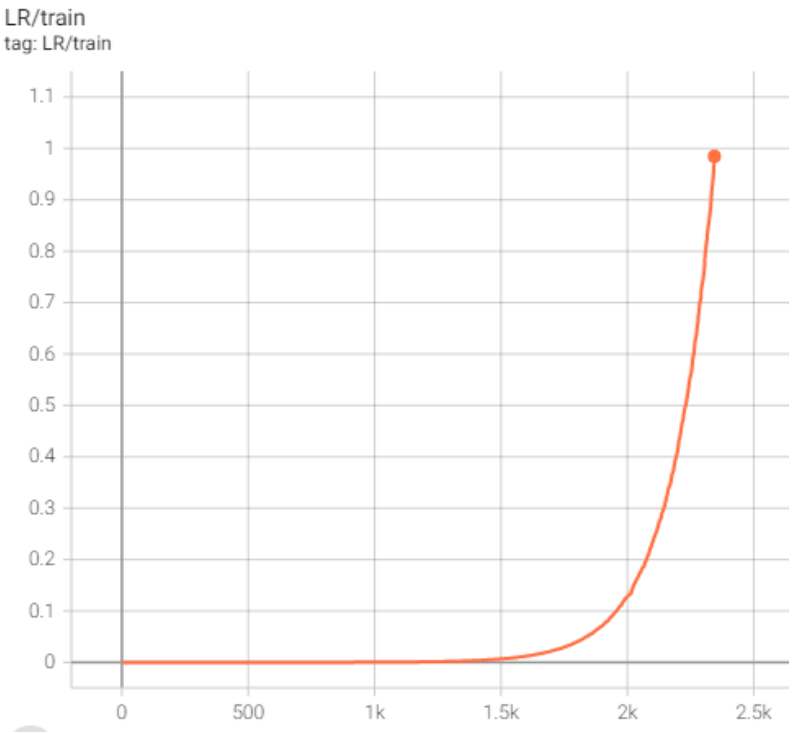
Colossalai and LR Range Test Report

training framework: Colossalai

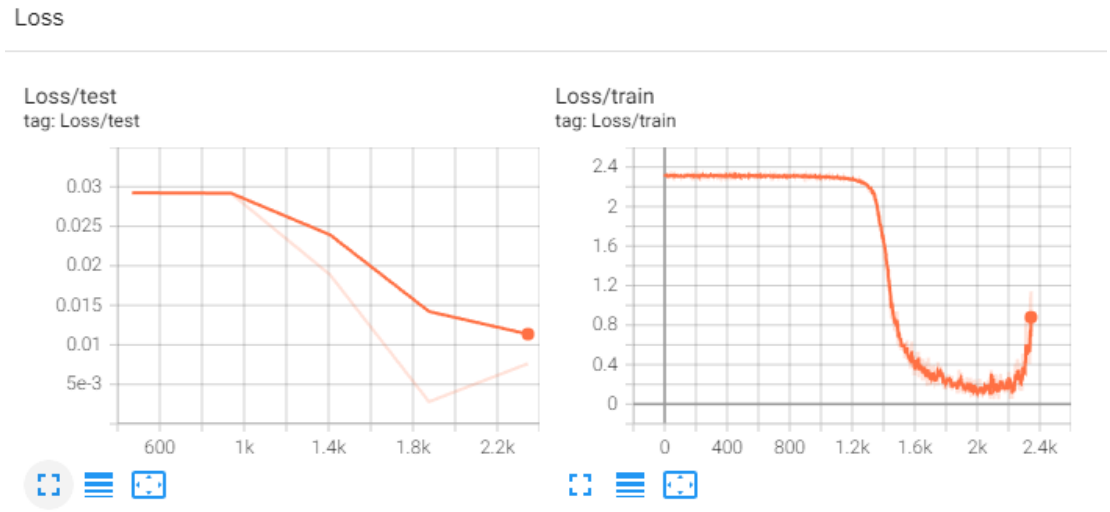
1. Exponentially increase the Learning rate (base)

Learning rate

```
def lrs(batch):
    low = math.log2(1e-5)
    high = math.log2(10)
    return 2**(low+(high-low)*batch/len(train_dataloader)/gpc.config.NUM_EPOCHS)
```



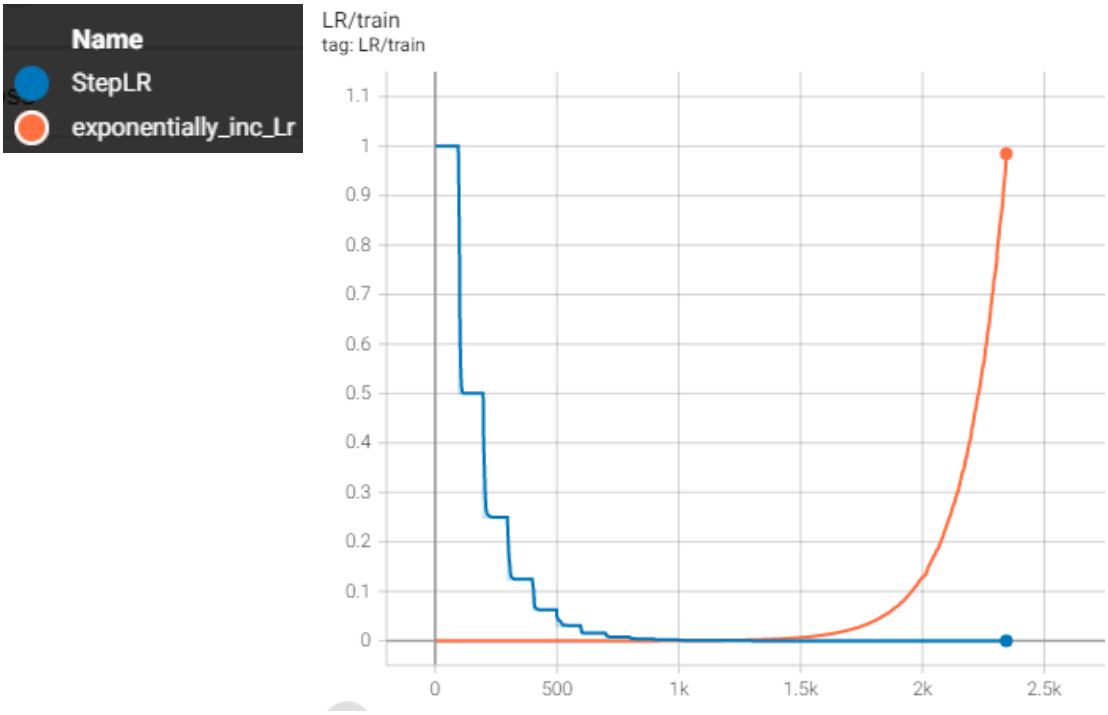
train & valid loss



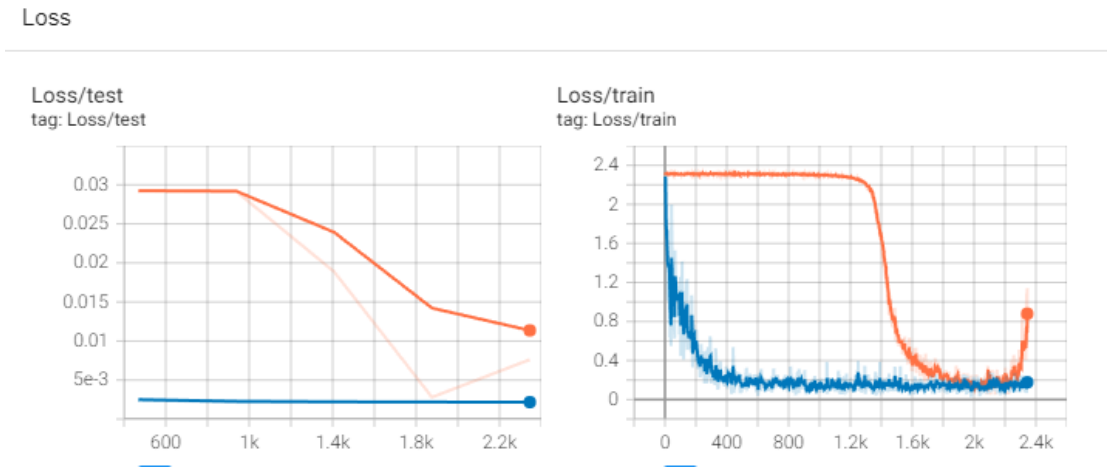
2. Decays the learning rate of each parameter group by gamma=0.5 every 100 epochs.

Learning rate (vs Exponentially decrease Lr)

```
lr_scheduler = torch.optim.lr_scheduler.StepLR(optimizer, step_size=100,
gamma=0.5)
```



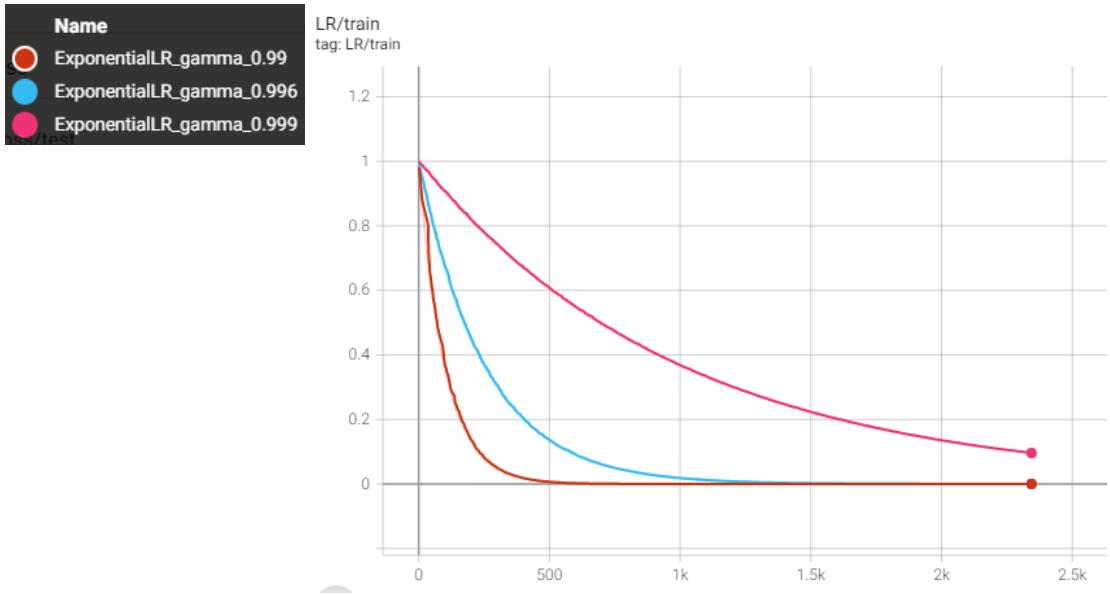
train & valid loss (vs Exponentially decrease Lr)



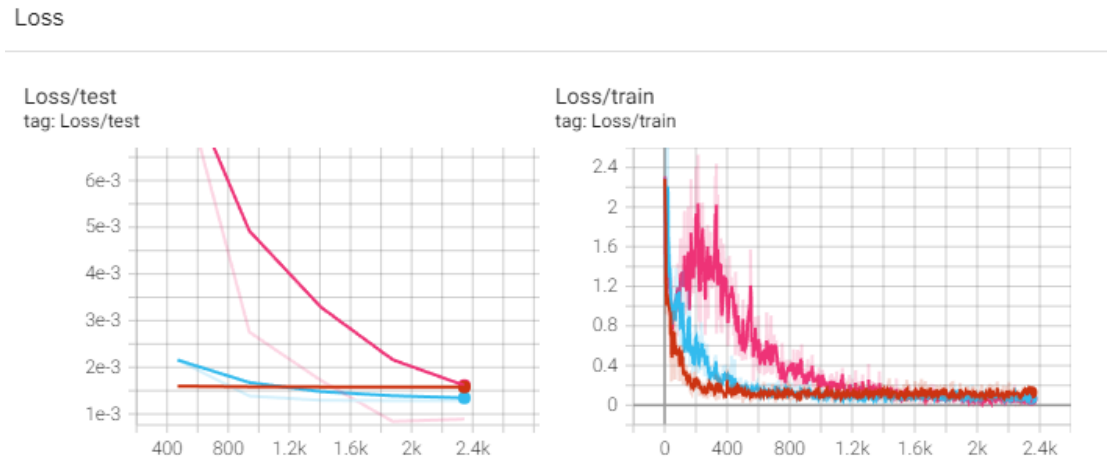
3. Decays the learning rate of each parameter group by gamma every epoch.

Learning rate (comparing different gamma - decaying rate)

```
lr_scheduler = torch.optim.lr_scheduler.ExponentialLR(optimizer, gamma=0.999)
```



train & valid loss (comparing different gamma - decaying rate)

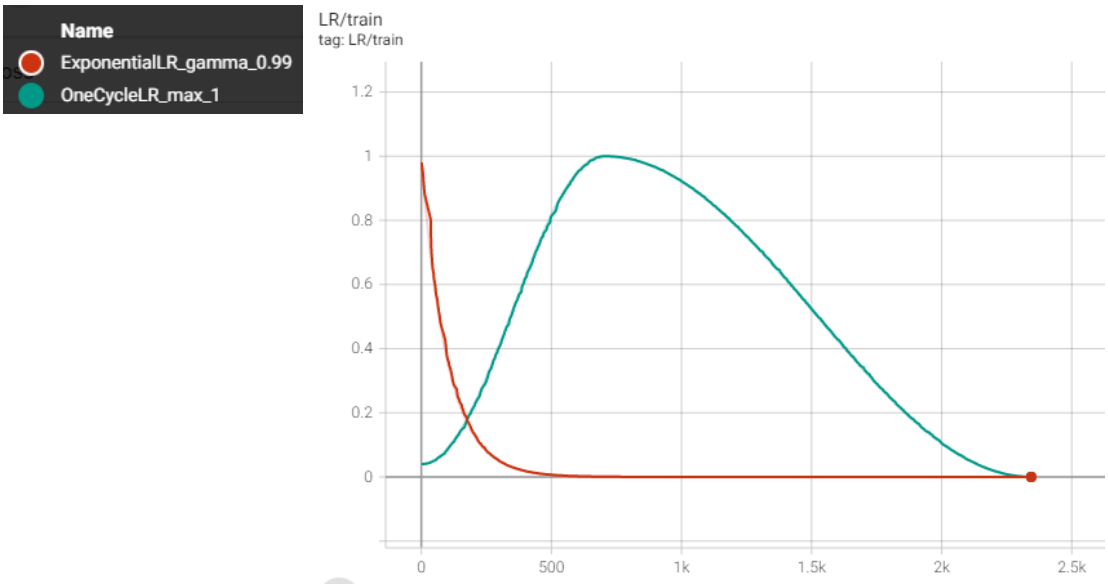


4. Sets the learning rate of each parameter group according to the 1cycle learning rate policy.

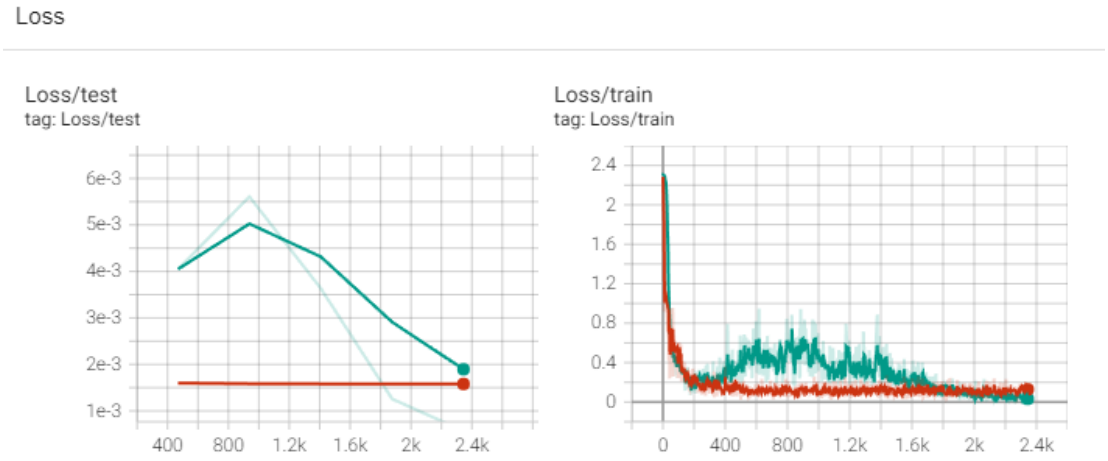
Learning rate

```
lr_scheduler = torch.optim.lr_scheduler.OneCycleLR(optimizer, max_lr=1.1,
steps_per_epoch=len(train_dataloader), epochs=config["NUM_EPOCHS"])
```

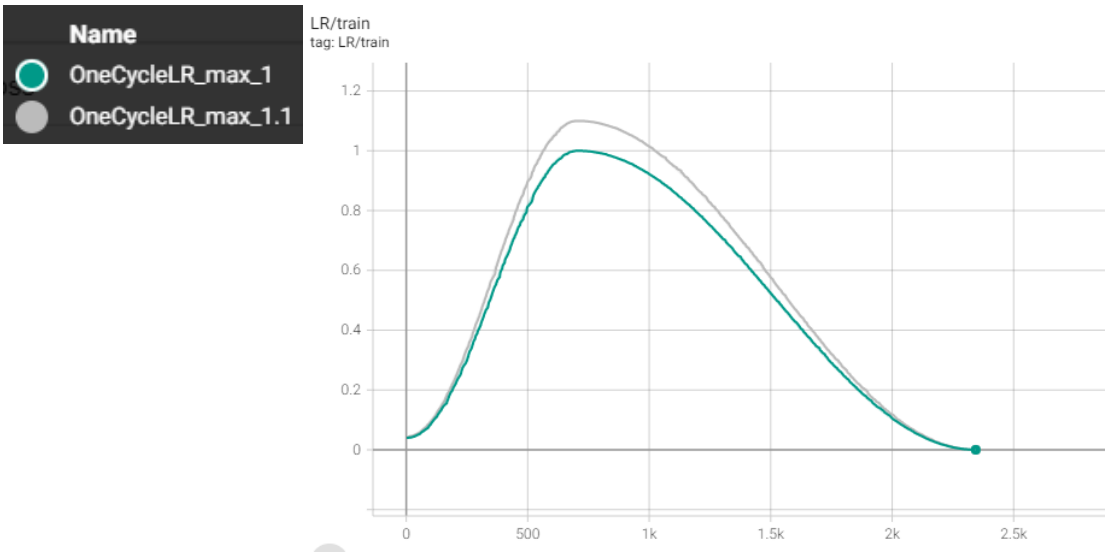
- Compare to Exponentially decrease Lr



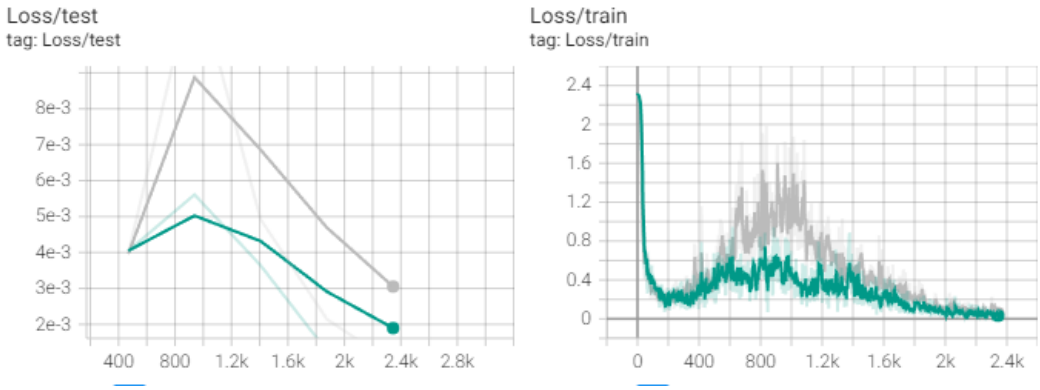
train & valid loss (comparing different gamma - decaying rate)



- Compare different max_lr (1 vs 1.1)



train & valid loss (comparing different gamma - decaying rate)

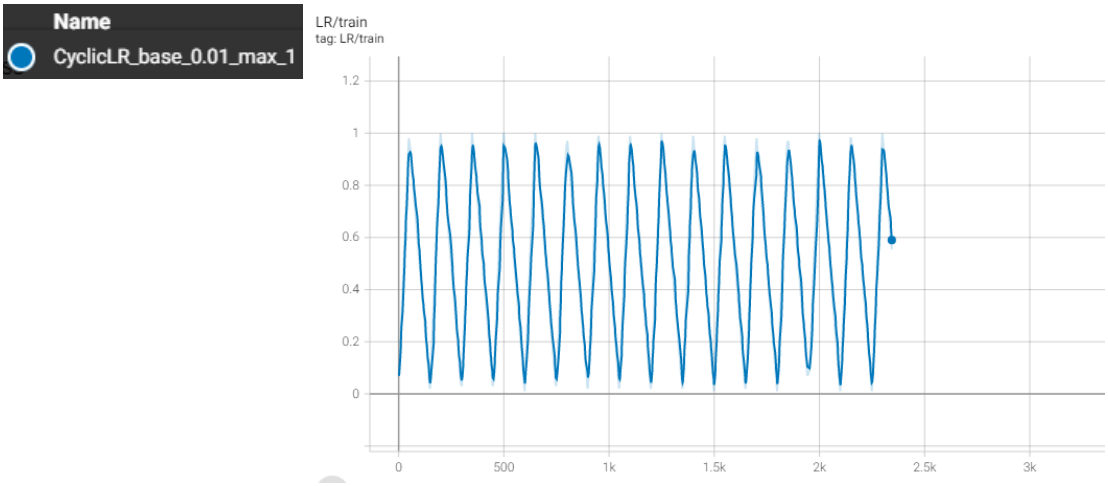


5. Sets the learning rate of each parameter group according to cyclical learning rate policy (CLR).

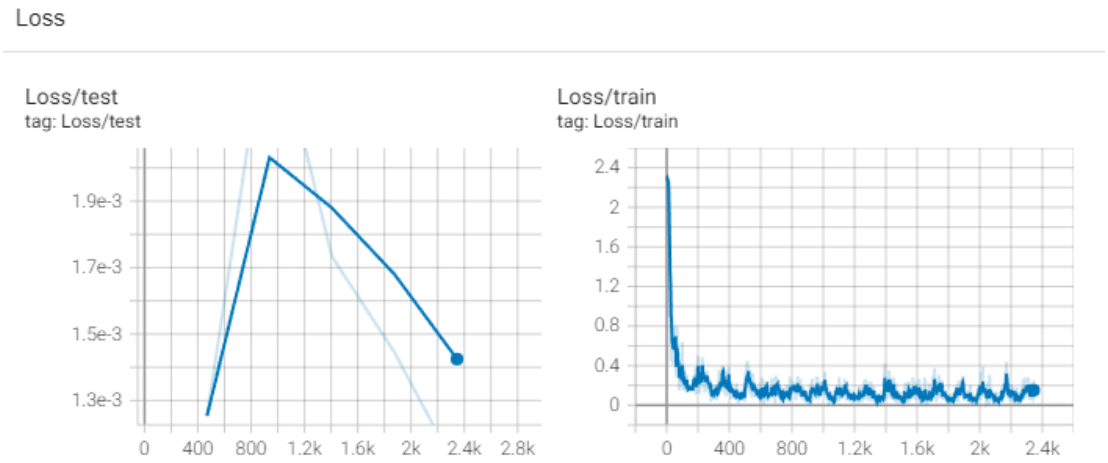
Learning rate

```
lr_scheduler = torch.optim.lr_scheduler.CyclicLR(optimizer, base_lr=0.01,
max_lr=1, step_size_up=50, step_size_down=100)
```

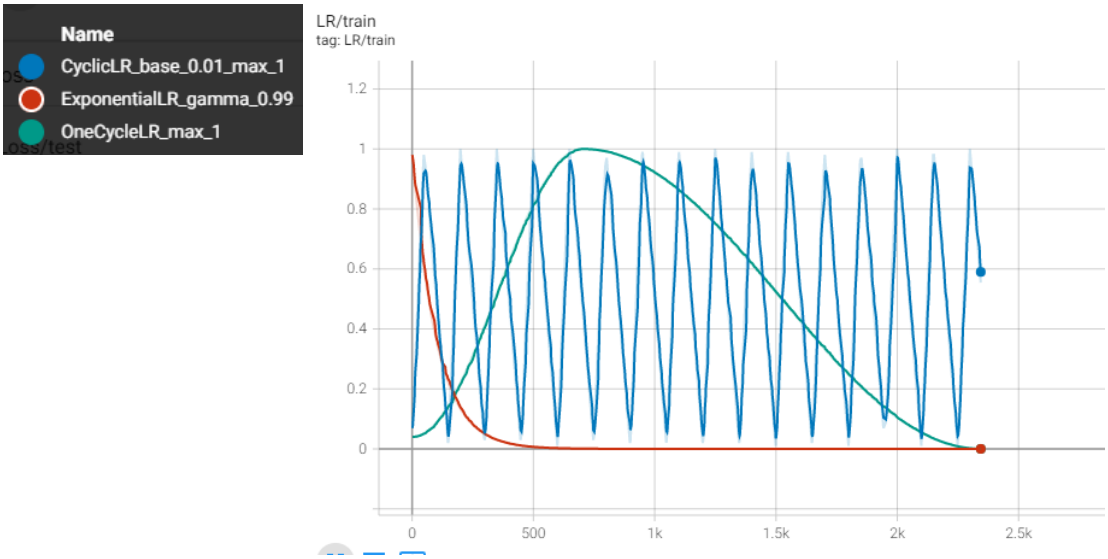
- Compare to Exponentially decrease Lr



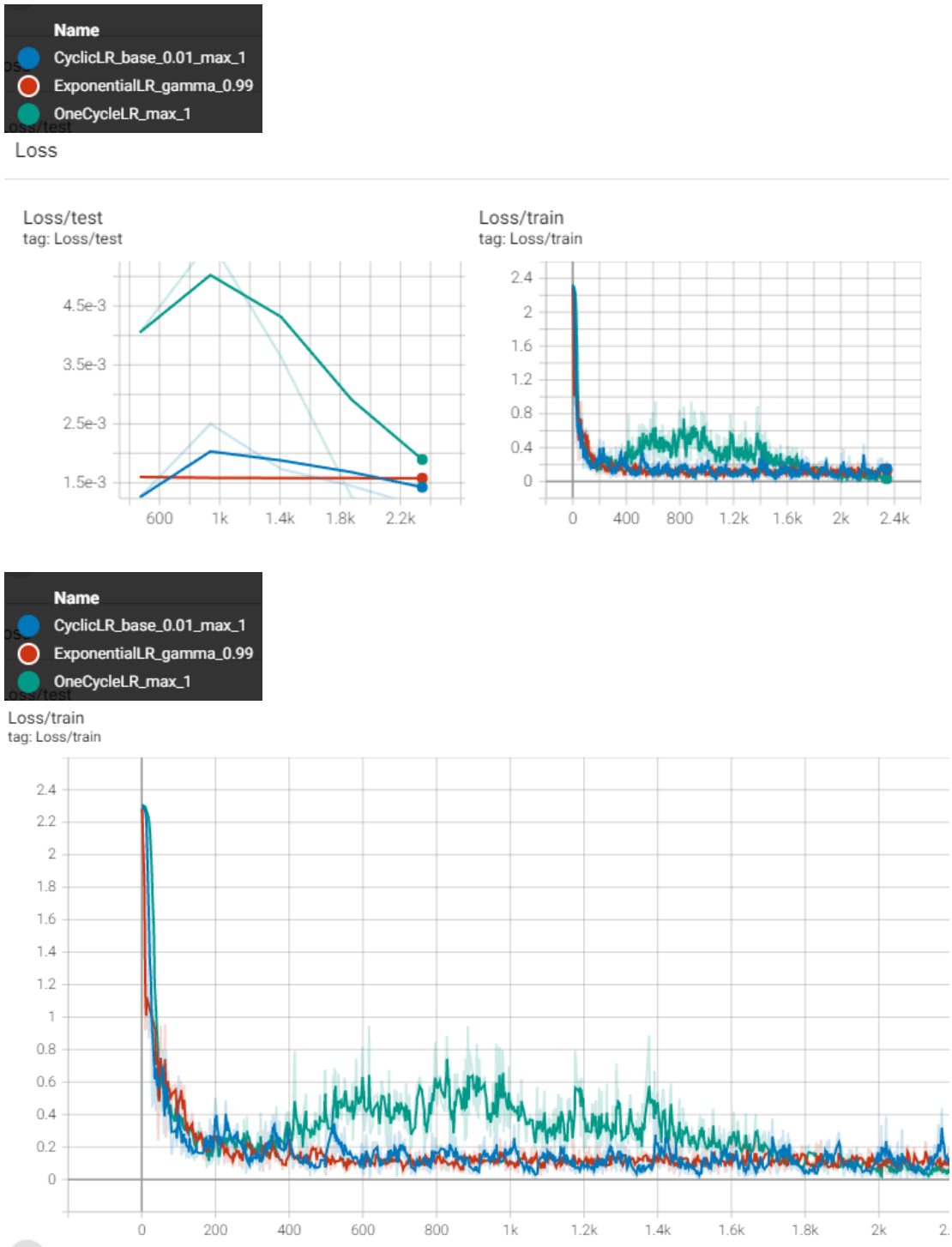
train & valid loss (comparing different gamma - decaying rate)



- Compare different learning schedules



train & valid loss (comparing different gamma - decaying rate)



github link: https://github.com/simasuiyuan/CS5260.git