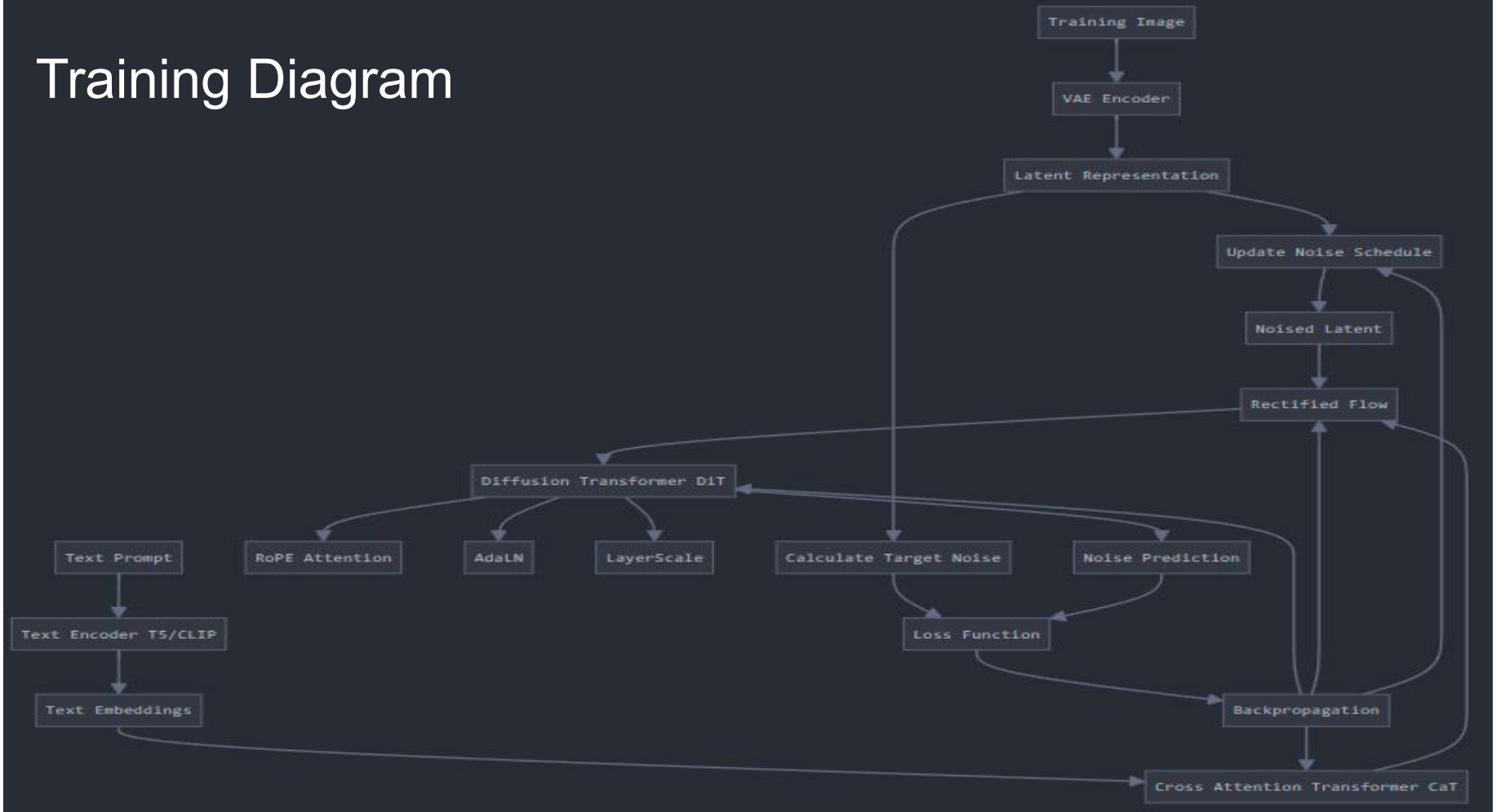
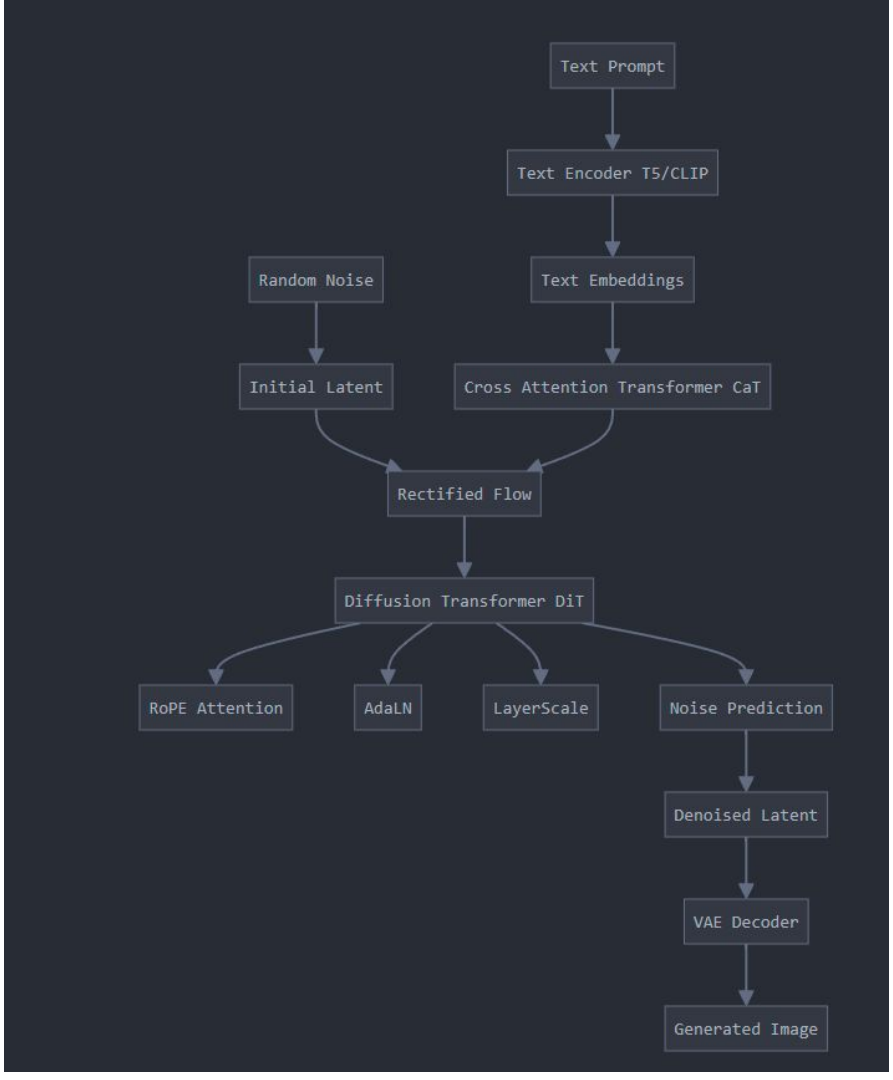
OMI Diffusion Model Pipeline v0.0.0 Foundation Model Report

Dr. Wolfgang Black
10/25/24

Training Diagram



Generation Diagram



Creative Cartoon Creatures Generations (128 - local)

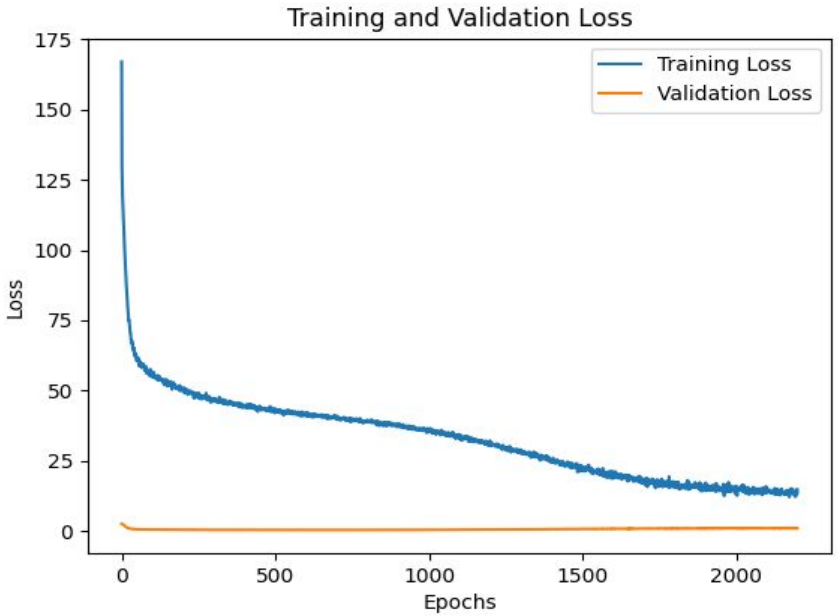
- prompts: ['A green creature with a leaf on its head','a red creature with a fire on its tail', 'a yellow cartoon character with a big smile', 'a cartoon frog character with a crown',]
- resolution: 128x128
- gpu: 1x NVIDIA GeForce RTX 4060 Ti
- gens at epoch 2200, after 578 min
- dataset: 833 rows

Creative Cartoon Creatures Generations (512)

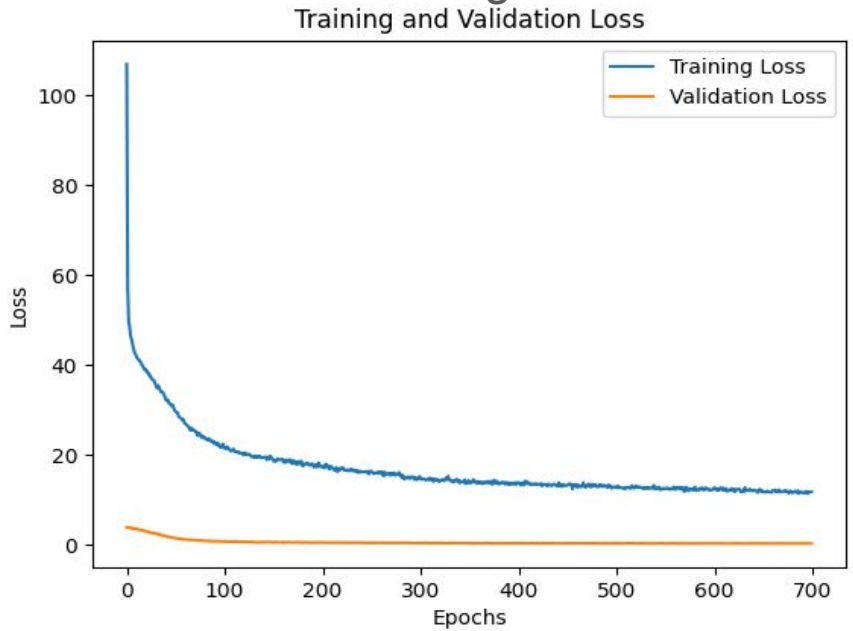
- prompts: ['A green creature with a leaf on its head','a red creature with a fire on its tail', 'a yellow cartoon character with a big smile', 'a cartoon frog character with a crown',]
- resolution: 128x128
- gpu: 3x NVIDIA RTX A4000
- gens at epoch 700, after 627 min
- dataset: 833 rows

Creative Cartoon Creatures Loss plots

Single GPU 128x128 Training



3x GPU 512x512 Training



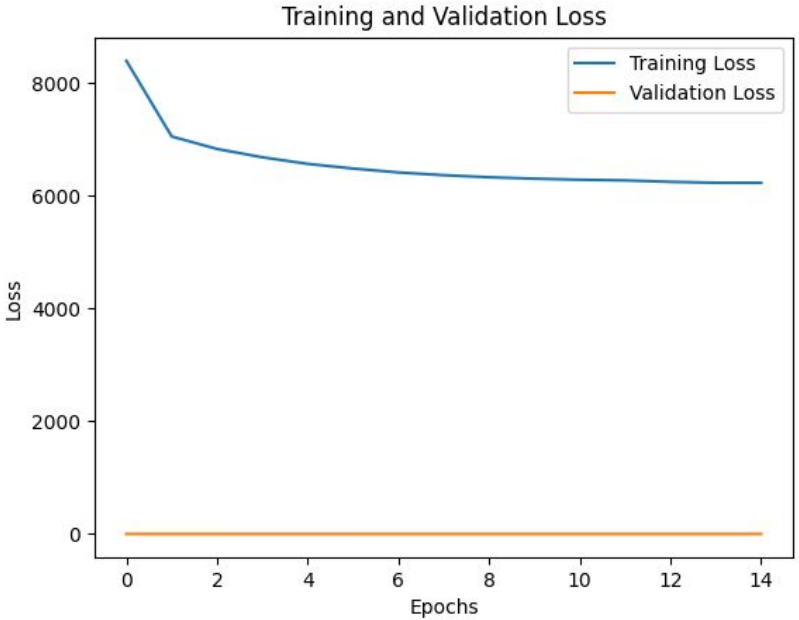
Lshun/Bedroom sample dataset

- prompts: [”]*
    - note, we still have 4 generations during training without any prompts
- resolution: 128x128
- gpu: - gpu: 3x NVIDIA RTX A4000
- gens at epoch 5, after 104 min
- dataset: 288k rows

Image was off of 15 epochs, which took 6 hours

Lshun/Bedroom Loss plots

3x GPU 128x128 Training



Next Steps (10/25/24)

1. Incorporate Model Parallelism to have hybrid parallelism for training
2. Abandon creative cartoon dataset and train on Lshun/Dataset sample with +1bil parameters
    a. Utilize Invoke Resources after testing training loop
3. Compare to a few new papers to see if we want to change any architecture before larger training
4. Consider adding scheduler/cfg/negative prompts