cuFLAVR

Aditya Hota, Richard Chen, Kaan Erdogmus
CIS 565 Fall 2021
Milestone 1 Presentation

FLAVR: Flow-Agnostic Video Representations for Fast Frame Interpolation

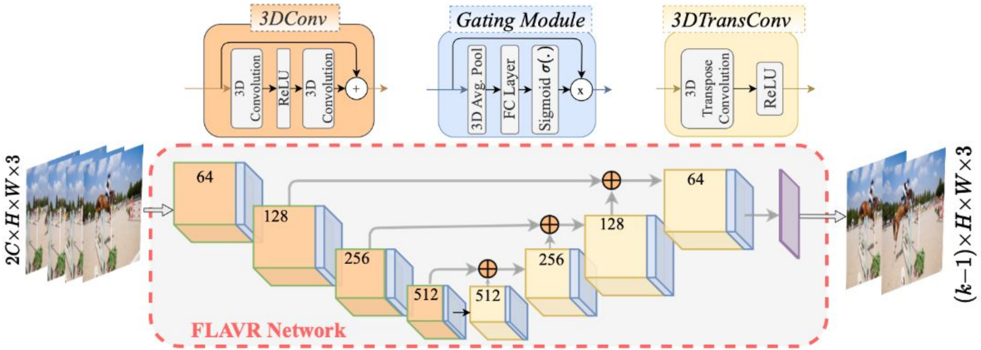
Tarun Kalluri * UCSD

Deepak Pathak CMU

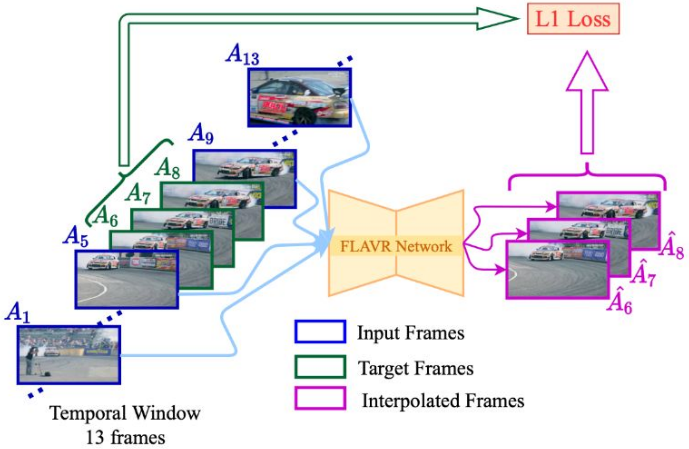
Manmohan Chandraker UCSD

Du Tran Facebook AI

https://tarun005.github.io/FLAVR/



(a) Overview of the proposed architecture



(b) Sampling procedure

Milestones

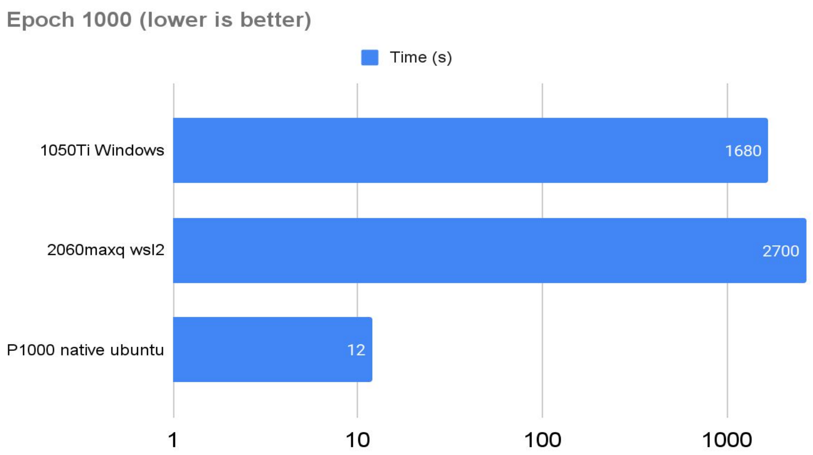
- Milestone 1 (11/17)
  - Setup CUDA neural network code
  - Determine feasibility of implementing special layers in CUDA
  - Become familiar with CUDA NN code and train simple NN for classification
- Milestone 2 (11/29)
  - Implement architecture for FLAVR
  - Training and classification should be working
- Milestone 3 (12/06)
  - Tune all hyperparameters for best results
  - Incorporate camera or object movement into Path Tracer
- Final Deliverable (12/12)
  - Automated pathway for generating interpolated videos
  - Performance analysis (and potential comparison to PyTorch)

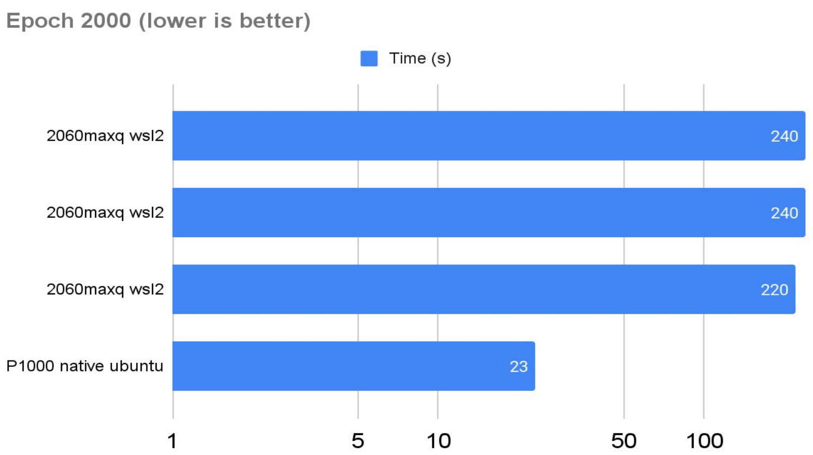
Completed Work

- Milestone 1 (11/17)
  - ✅ Setup CUDA neural network code
  - ⚠️ Determine feasibility of implementing special layers in CUDA
  - ✅ Become familiar with CUDA NN code and train simple NN for classification
- Other progress
  - Set up Linux environment for CUDA NN programming
    - Performance comparison to Windows
  - Reading paper to understand math for spatio-temporal 3D convolution
  - Implementing 2D convolution to understand cuDNN API

CUDA Neural Network

CUDA calls in Windows have much more overhead than in native Linux





Next Milestone

- Milestone 2 (11/29)
    - Finish learning/exploration of cuDNN and cuBLAS
    - Implement forward pass for each conv. and gating layer
        - Test with weights/bias from pre-trained model
    - Implement backwards architecture for FLAVR
    - Training and classification should be working

Concerns

- Difficulties in fully understanding CNN structure of FLAVR
- GPU training time: baseline: we use their hyperparameters for inference
    - they used 8 RTX 2080Tis
- Good cuDNN examples are hard to find
    - Even NVIDIA’s deep learning samples use TF/PyTorch

Q&A

References

Papers:

Kalluri, T., Pathak, D., Chandraker, M., & Tran, D. (2020). Flavr: Flow-agnostic video representations for fast frame interpolation. *arXiv preprint arXiv:2012.08512.*

Tran, D., Wang, H., Torresani, L., Ray, J., LeCun, Y., & Paluri, M. (2018). A closer look at spatiotemporal convolutions for action recognition. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 6450-6459).

3rd Party Code:

CUDA Neural Network Implementation

CUDA Neural Network

CUDA calls in Windows have much more overhead than in native Linux

Epochs	Accuracy	Time	Device
1000	0.89	12min	1050Ti Windows
2000	0.95	4min	2060maxq wsl2
2000	0.94	4min	2060maxq wsl2
1000	0.89	1min45	2060maxq wsl2
2000	0.98	3min40	2060maxq wsl2
1000	0.89	12s	P1000 native ubuntu
2000	0.98	23s	P1000 native ubuntu