

# cuFLAVR

---

Aditya Hota, Richard Chen, Kaan Erdogmus  
CIS 565 Fall 2021  
Milestone 2 Presentation

# FLAVR: Flow-Agnostic Video Representations for Fast Frame Interpolation

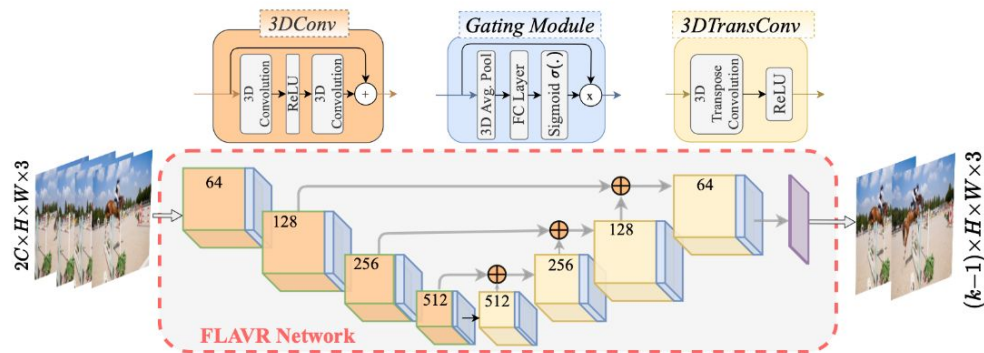
Tarun Kalluri \*  
UCSD

Deepak Pathak  
CMU

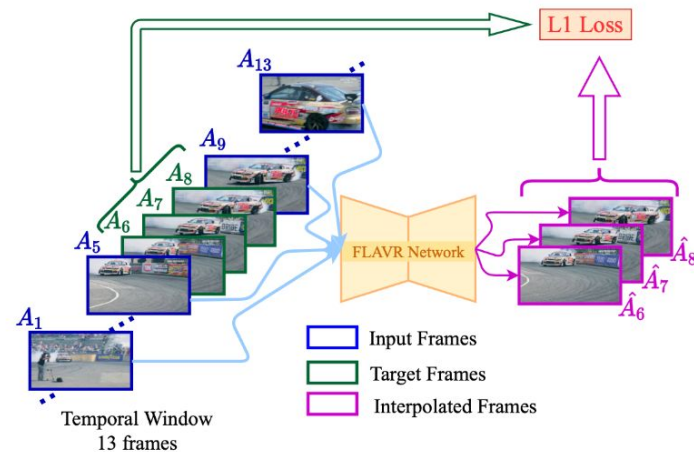
Manmohan Chandraker  
UCSD

Du Tran  
Facebook AI

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



(a) Overview of the proposed architecture









(b) Sampling procedure

# Milestones

- Milestone 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 2 (11/29)
  -  Implement architecture for FLAVR
  -  Training and classification should be working
- Other progress
  -  Implemented 2D convolutions to understand cuDNN API
  -  Set up all latest required dependencies due to new API
    - cuDNN 8.3 with CUDA 11.5 and OpenCV
  -  Extracted data from reference model hyperparameters and weights
  -  Laid out classes and functions for the codebase
  - Read paper to understand math for spatio-temporal 3D convolution
  - Began implementing layers of FLAVR network

# 2D Convolutions using cuDNN

- Challenges
  - API backend changes between cuDNN 7 and 8
    - Needed to incorporate new algorithm finding function
  - Lots of errors during setup process (matching all libraries)

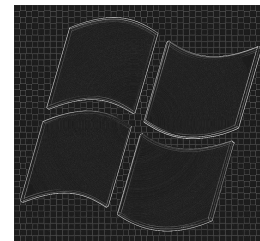
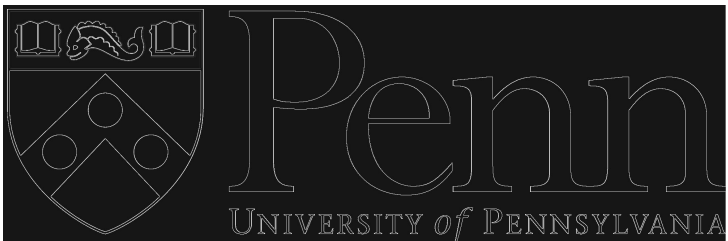


Fig 1. Edge Detection Convolutions

# Reference Model Data Extraction

- Analyzed the schema for the FLAVR reference model
  - Looked at model for 2x Interpolation, 4x and 8x have similar schema and size
- Extracted hyperparameters both for both learning and inference
- Extracted torch.Tensor instances
  - Matched instances with relevant layers and convolutions; 59 tensor instances
- Challenges:
  - Convert pyTorch tensor instances to cuDNN tensor instances
  - Convert overall schema to a format serializable with C++ and simplify the schema



# Next Milestone

- Milestone 3 (12/06)
  - Finish layers for 3D spatio-temporal convolutions and gating
  - Match PyTorch model weights with our cuFLAVR network
  - Convert video into image frames (inference inputs)

# 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\* \(pp. 6450-6459\).](#)

## 2D Convolutions Guide:

Peter Goldsborough: [2D Convolutions using cuDNN](#)



Q&A

---