VIDEO SUPER-RESOLUTION

KEEP TRYING AND EXPLORING

Chuckle Lee
Jerry Ho

OUTLINE

1. Recall Last time
2. Improvements
3. Future directions

PART 1

Recall Last time

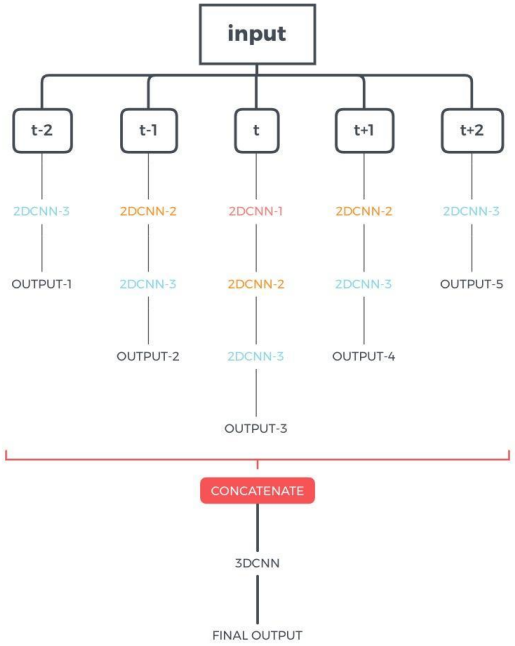
BACKGROUND

- superresolution is a technique that enhance the resolution of imaging system

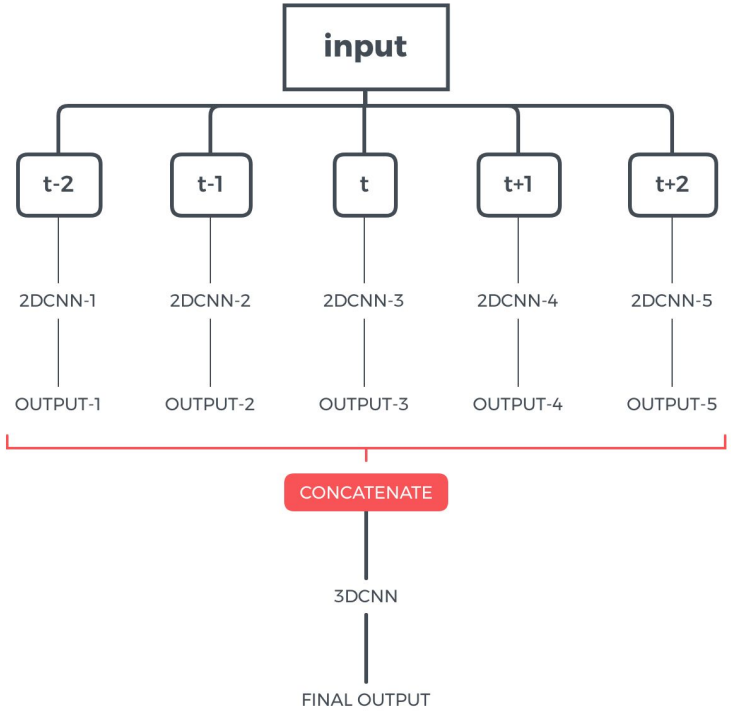
- different methods such as Single-frame deblurring ... has been developed

Two different models

Delta



5-parallel



Measuring instruments

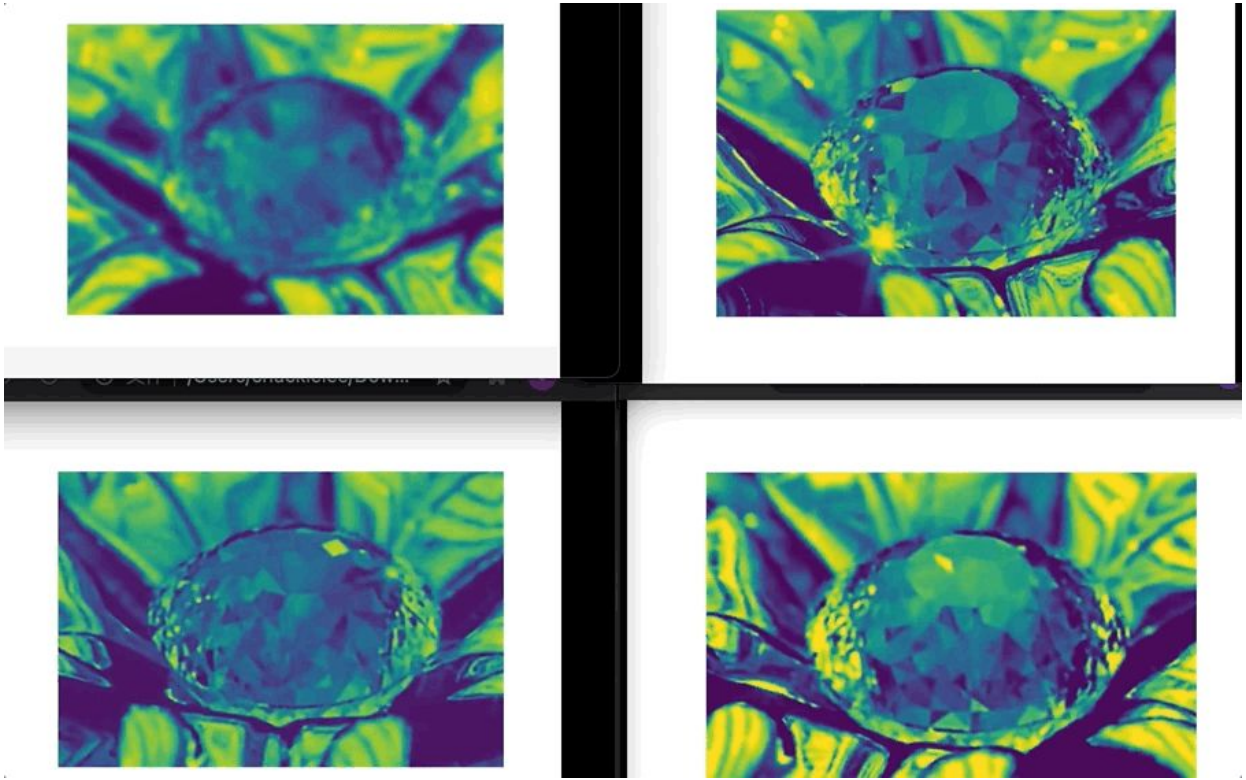
$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right)$$

How close are the values in each pixel similar

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

How close (in terms of strcuture) are two photos similar

Final models last time



PART 2

Improvements

Optimization of Pytorch Code

- Turn off gradient computation during validation (DONE PREVIOUSLY)
- Data parallel (HALF WAY, not yet test on multiple GPU)
- Automatic Mixed Precision (AMP) (DONE)
- Consider using a different learning rate schedule (DONE)

Automatic mixed precision and Grad scale

- autocast operations to datatype float16 or float32 depending on the operations. It can reduce the memory usage and runtime of the model while preserving accuracy

- However, the problem of AMP is that while the gradient is small and the operation datatype is converted to float16, gradient “underflow” occurs. Therefore, grad scaling is used to prevent this problem by scaling the gradient during training.

Statistics and video showing

	MSE	SSIM	PSNR
Raw	9193.6647	0.0392	8.5000
Delta	125.1786	0.9155	27.1675
5-parallel	127.5626	0.9184	27.0820

Video : https://drive.google.com/drive/folders/12zsezJF2uqQ8tmJOY09g9gB7bKL_a0XQ?usp=sharing

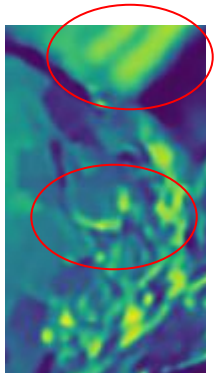
Comparasion of 3 models

	Previous	Delta	5-parallel
Is it a single CNN?	No	Yes	Yes
Does it contain multiple 2D layer stacks?	No	Yes	Yes
Is the recurrent 2D part involved?	No	Yes	Yes
Is the recurrent 3D part involved?	No	Yes	Yes
Is the range of output controlled?	No	No	Yes
Implementation	Tensorflow	Pytorch	Pytorch
Number of parameters	10,577(2D) 63,129(3D)	83,702	259,938

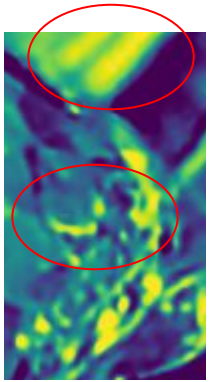
Comparing with last time

- increased the depth of the network
- efficient training after modifying the code allows more epochs

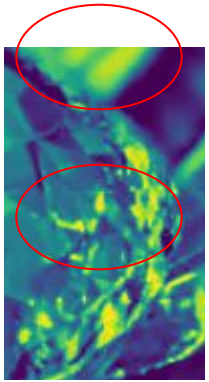
Delta



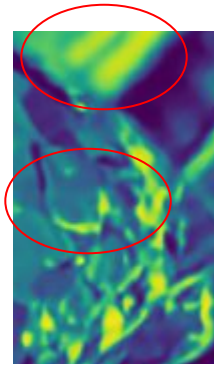
5-parallel (old)



Ground truth

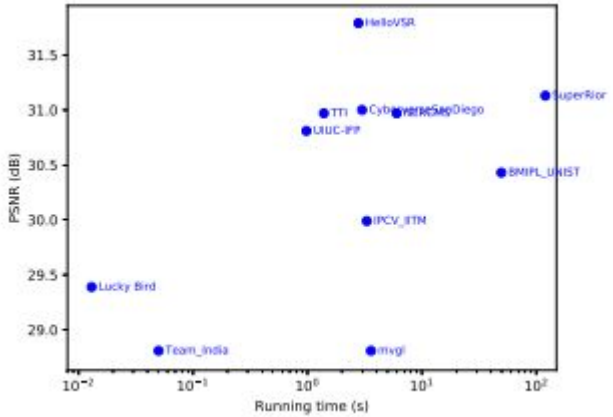


5-parallel (new)



Benchmark dataset

- **the REDS dataset for video super resolution**



Ref from : https://seungjunnah.github.io/Datasets/reds.html

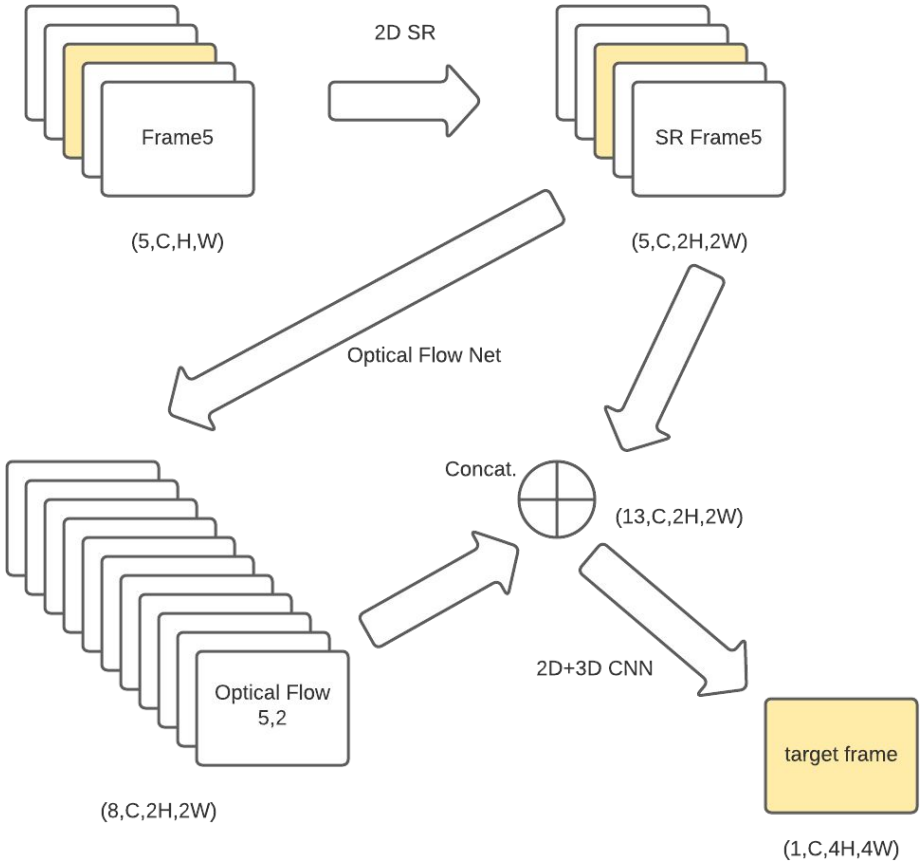
PART 3

WHAT WE ARE GOING TO DO

Optical Flow

commonly used in obtaining temperal information between frames

Plan : 2D + Optical Flow + 2D3D CNN



What to do on REDS

Suppose we have a dataset available for training

Step 1. Use FRSCNN/other 2D CNN to do super resolution for each frame of the whole dataset

Step 2. Use the frame obtained from Step 1 to do training and testing of 5parallel / Delta (Done the python code already)

After training the model , for every incoming video:
Repeat Step 1 and use the model trained from Step 2

THANKS!