25th International Conference
on Pattern Recognition
10 - 15 January, 2021

PA-FlowNet : Pose-Auxiliary Optical Flow Network for Spacecraft Relative Pose Estimation

Zhi-Yu Chen, Po-Heng Chen, Kuan-Wen Chen, Chen-Yu Chan

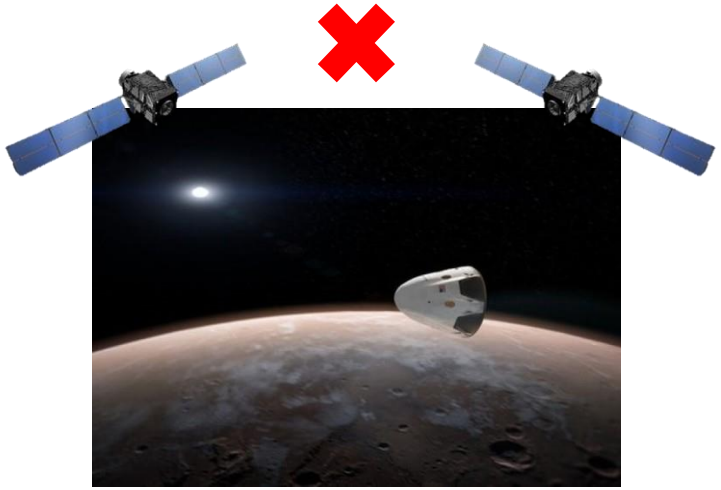
National Chiao Tung University & National Space Organization

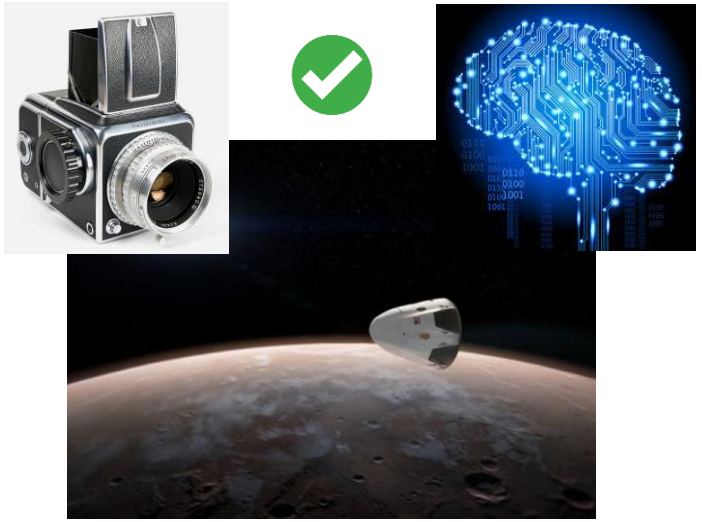
Pre-recorded :
10 December, 2020

Introduction

Backgrounds

- ***GPS*** (***Global Positioning System***) is often used in the global environment.
- It is ***challenging*** to deploy enough satellites on other planets.
- Use ***computer vision technique !***

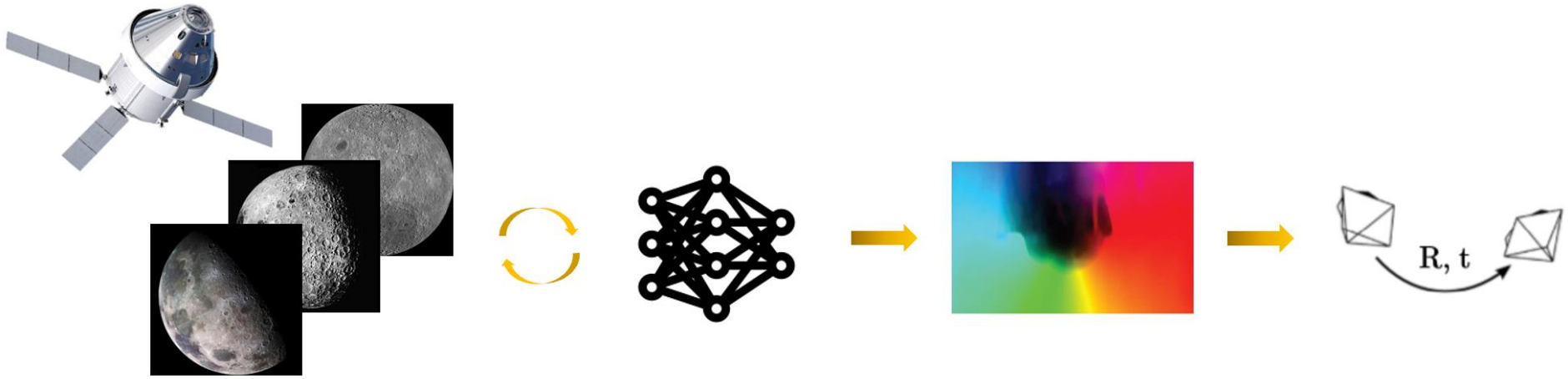




Introduction

Objectives

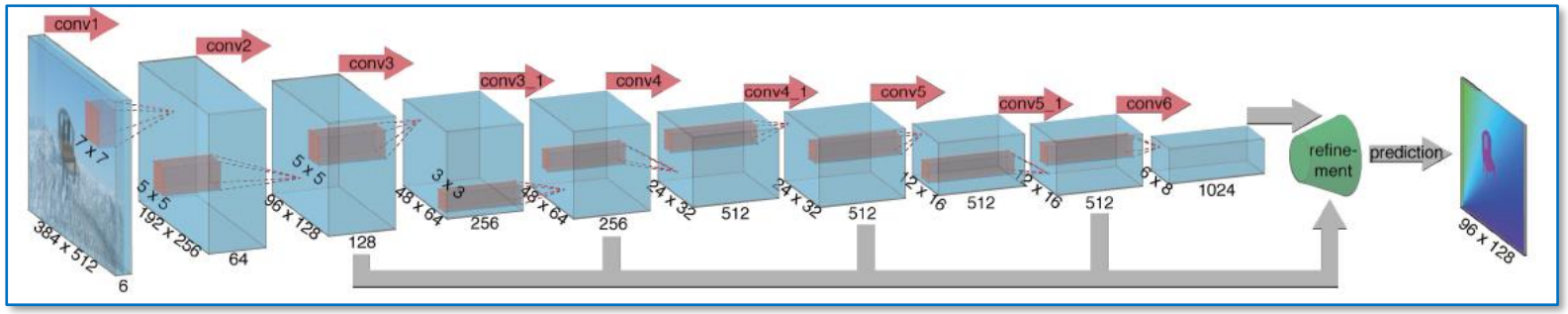
- ***Goal*** : ***To find the appropriate method of computer vision for space localization***.
- ***Complicated foreground*** will not appear in space.
- Use ***optical flow approach !***



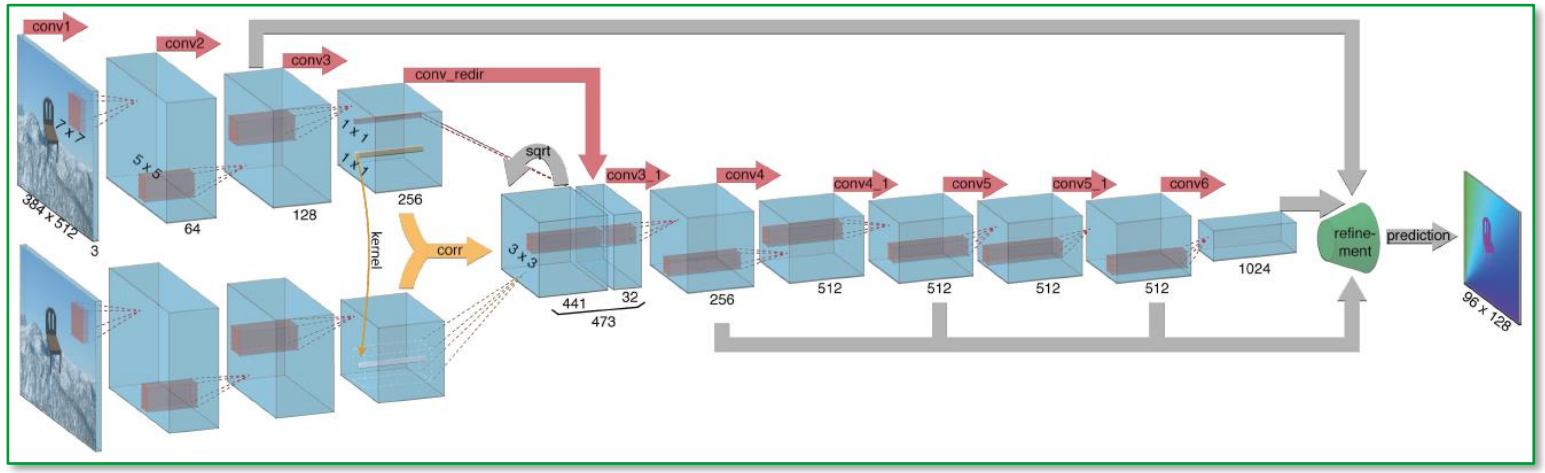
Moon Data → *Neural Network* → *Optical Flow Field* → *Camera Pose*

Previous Work

FlowNet [1]



FlowNetS

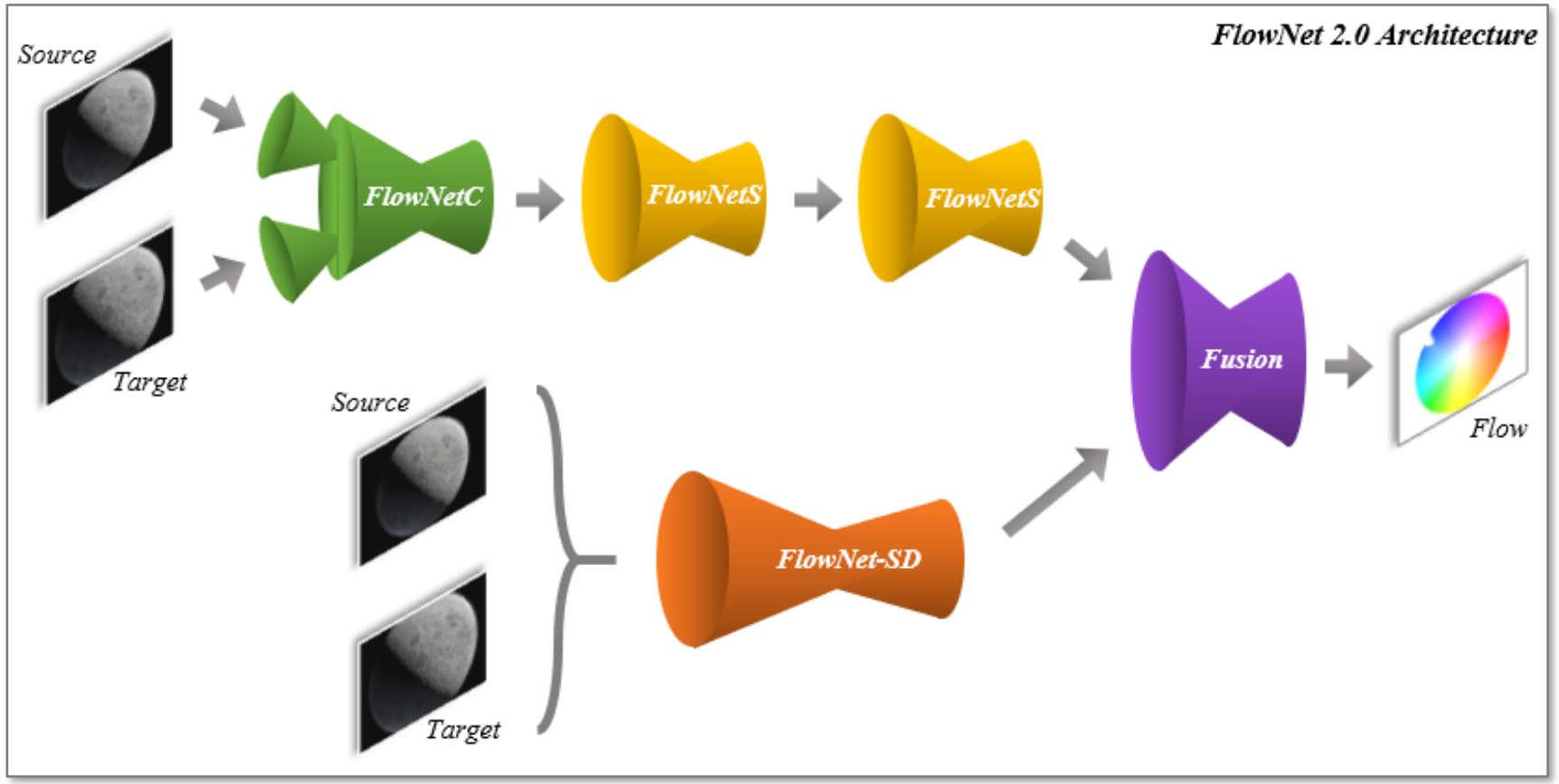


FlowNetC

[1] *A. Dosovitskiy, P. Fischer, E. Ilg, P. Hausser, C. Hazirbas, V. Golkov, P. van der Smagt, D. Cremers, and T. Brox. "FlowNet: Learning optical flow with convolutional networks," in Proc.* ***ICCV, 2015****.*

Previous Work

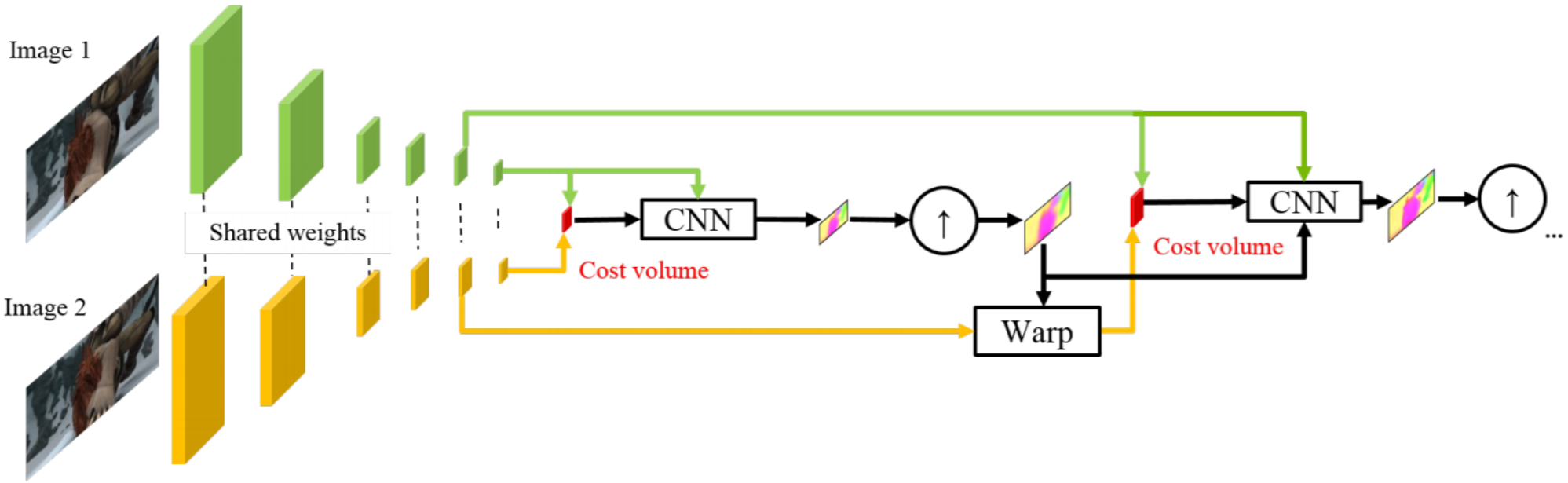
FlowNet2 [2]



[2] *E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, and T. Brox. "FlowNet 2.0: Evolution of optical flow estimation with deep networks," in Proc.* ***CVPR, 2017.***

Previous Work

PWC-Net [3]

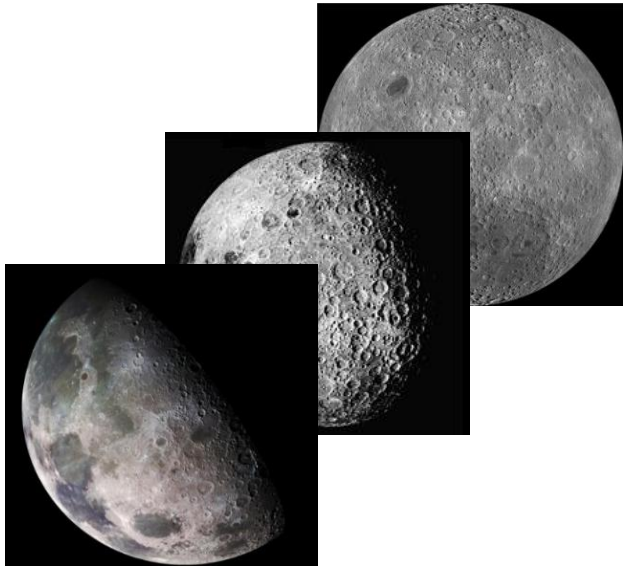


[3] *D. Sun, X. Yang, M.Y. Liu, and J. Kautz.*
"PWC-Net: CNNs for optical flow using pyramid, warping, and cost volume," in Proc. ***CVPR, 2018.***

Basic model for PA-FlowNet

FlowNet2 vs PWC-Net

Moon Datasets



MPI Sintel Dataset

Clean

Final

FlowNet2 > PWC-Net

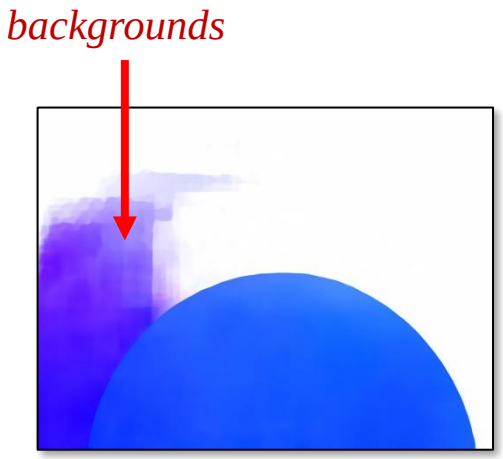
FlowNet2 < PWC-Net

FlowNet2 on Moon Data

- Huge performance gap while the height is ***larger than 3000 kilometers***.
- ***Reason*** : ***the poor flow estimation of the backgrounds***.
- ***Solution*** : Use ***Curriculum Learning*** !

Source

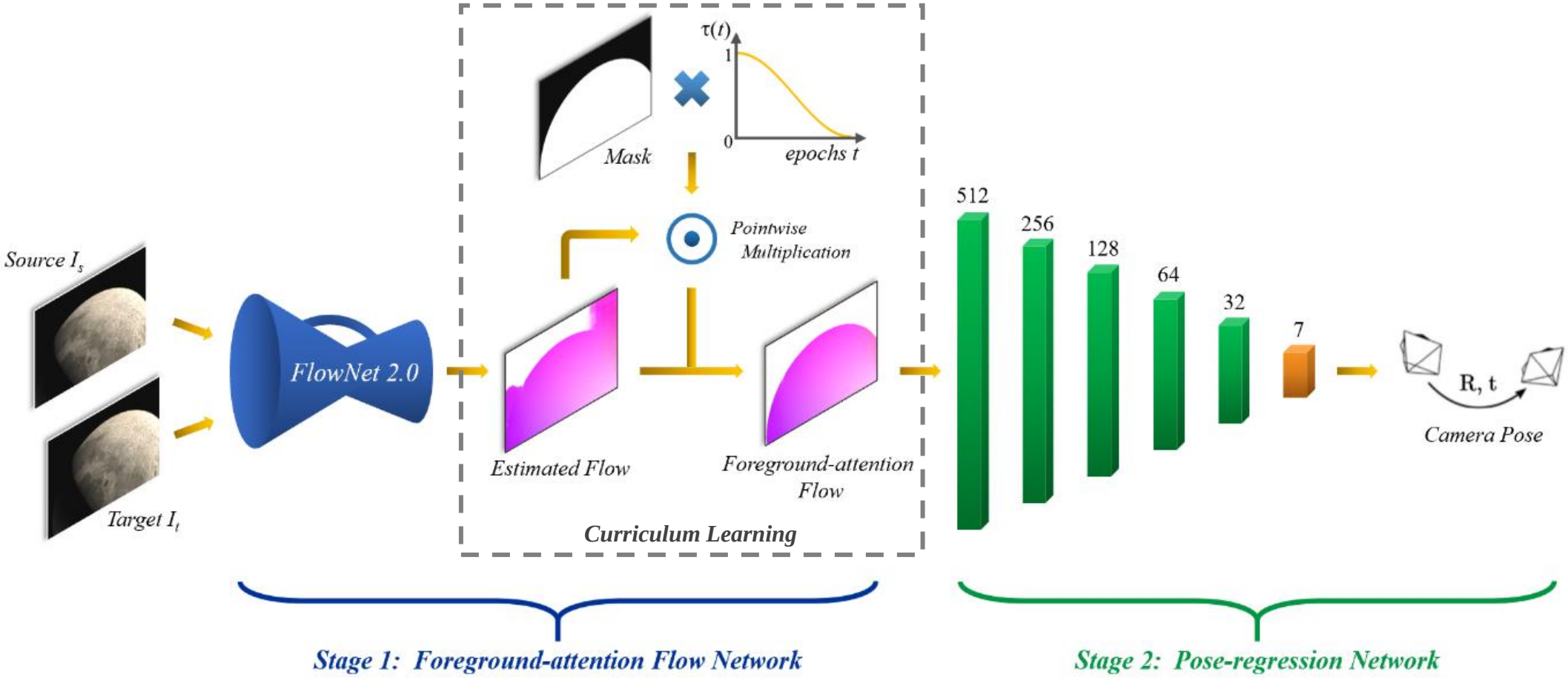
Target



Estimated Flow

Proposed PA-FlowNet

Network Architecture



Proposed PA-FlowNet

Foreground-attention Approach

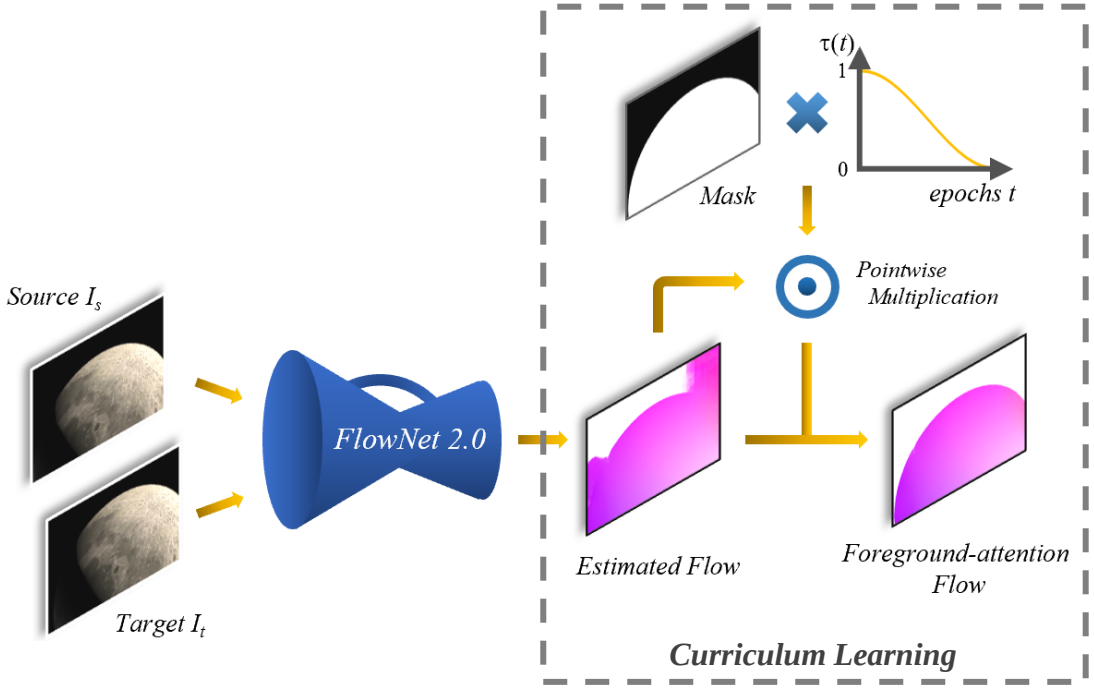
$$f'_{(c,i,j)} = f_{(c,i,j)} + \tau(t)\left[M_{(i,j)} \cdot f_{(c,i,j)}\right]$$

- $f_{(c,i,j)}$: *Estimated Flow*
- $f'_{(c,i,j)}$: *Foreground-attention Flow.*
- $M_{(i,j)}$: *Mask*

$$M_{(i,j)} = \begin{cases} 1, & \text{if } pixel_{(i,j)} \text{ belongs to the foreground.} \\ 0, & \text{if } pixel_{(i,j)} \text{ belongs to the background.} \end{cases}$$

- τ : *attention factor* (*depending on* t)

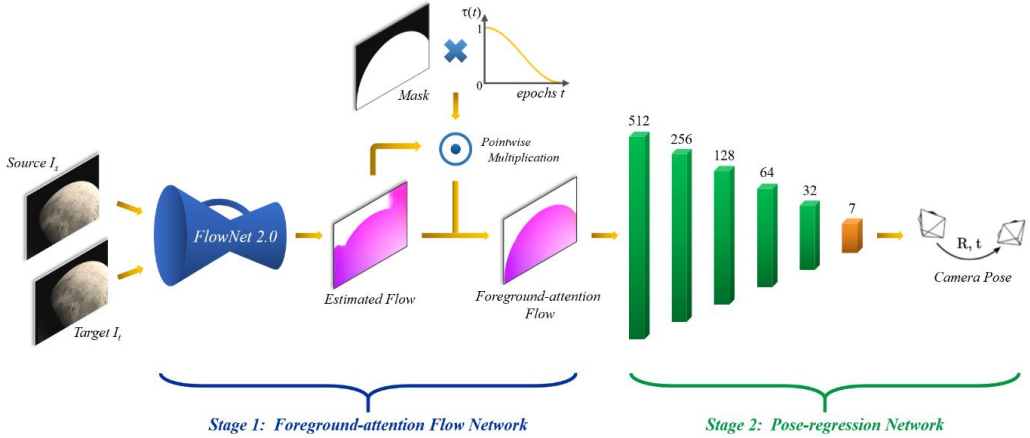
$$\tau(t) = 0.5[1 + cos(\pi t \,/\, total\_epochs)]$$



[4] *M. M. Derakhshani, S. Masoudnia, A. H. Shaker, O. Mersa, M. A. Sadeghi, M. Rastegari, and B. N. Araabi. "Assisted excitation of activations: A learning technique to improve object detectors," in Proc.* ***CVPR, 2019.***

Proposed PA-FlowNet

Training Loss



- ***Multi-scale End-point Error Loss*** *(for* ***Stage 1****)* :

$$\mathcal{L}_{M-epe} = \sum_{l=0}^{N} \alpha_l \sum_x \left\| Flow_{est}^{(l)}(x) - Flow_{gt}^{(l)}(x) \right\|_2 ; \quad \alpha_l = 2^l / (2^N - 1) , \quad N = 4.$$

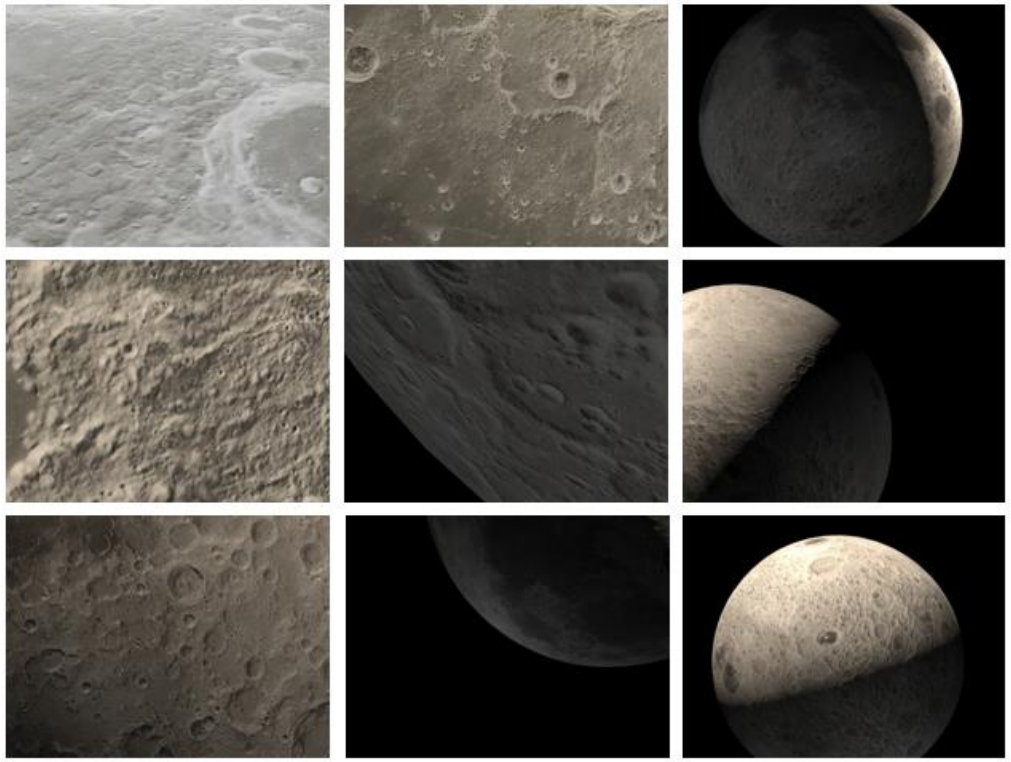
- ***Total Loss*** *(for* ***Stage 2****, End-point Error Loss + Pose Loss) :*

$$\mathcal{L}_{total} = \mathcal{L}_{epe} + \gamma \mathcal{L}_{pose} \; ; \quad \gamma = 10.$$

  - $\mathcal{L}_{epe} = \sum_x \left\| Flow_{est}(x) - Flow_{gt}(x) \right\|_2$.
  - $\mathcal{L}_{pose} = \beta \left\| Q_{gt} - (Q_{est} / |Q_{est}|) \right\|_2 + \left\| t_{gt} - t_{est} \right\|_2 ; \quad \beta = 20.$

Moon64K Dataset

Training & Testing Data



Height (Distance from Moon)	Training Data	Testing Data	Total
50 ~ 100 km	400	100	500
100 ~ 1,000 km	2,000	500	2,500
1,000 ~ 2,000 km	2,000	500	2,500
2,000 ~ 3,000 km	2,000	500	2,500
3,000 ~ 4,000 km	1,600	400	2,000
Aggregate	8,000	2,000	10,000

- **Image Resolution :** **1024 × 768 pixels .**

Moon64K Dataset

Ground-Truth Flow

- ***Source Image Coordinate → Source Camera Coordinate*** :

$$P_s = \lambda \begin{bmatrix} x_s \\ y_s \\ 1 \end{bmatrix} \sim K^{-1} p_s = K^{-1} \begin{bmatrix} u_s \\ v_s \\ 1 \end{bmatrix}$$

- ***Source Camera Coordinate → Target Camera Coordinate*** :

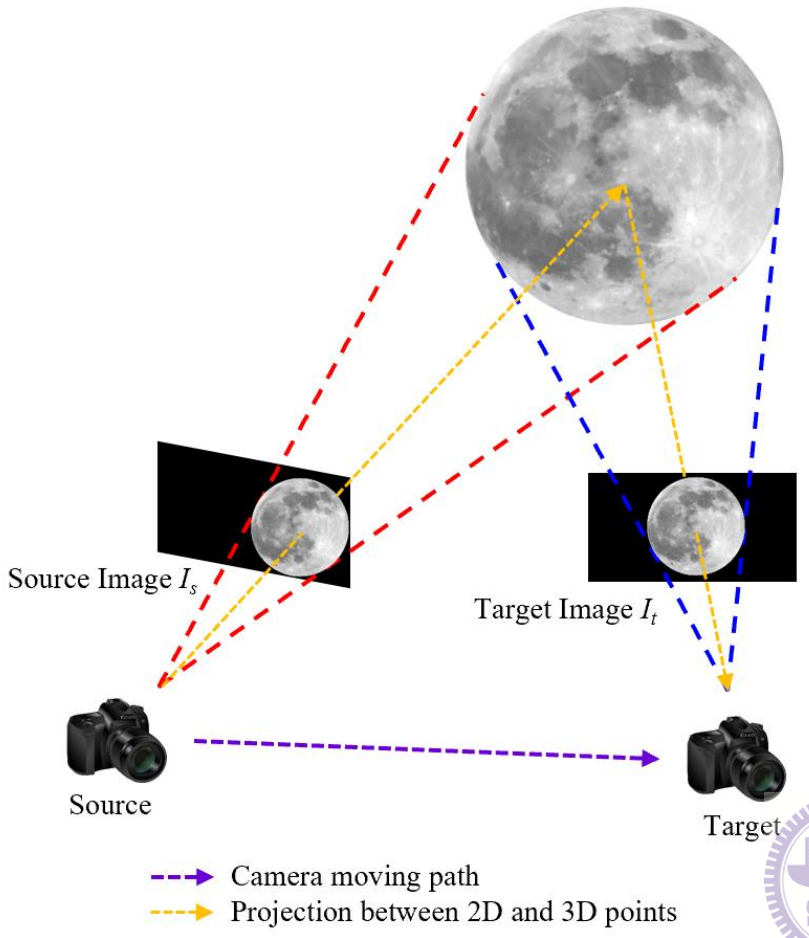
$$P_t = [R \mid t] P_s$$

- ***Target Camera Coordinate → Target Image Coordinate*** :

$$p_t = \begin{bmatrix} u_t \\ v_t \\ 1 \end{bmatrix} \sim K P_t = K \begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix}$$

- ***Solve λ ?*** Use ***spherical equation !***

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r^2$$



Experimental Results

Optical Flow

- Comparison of average end-point error (AEPE) on Moon64K dataset :

Methods	***AEPE***
FlowNetS	*14.967*
FlowNetC	*11.023*
FlowNet2	*6.682*
PWC-Net	*10.059*
Foreground-attention Flow Network *(PA-FlowNet without Pose Regression)*	*3.202*
PA-FlowNet	***2.409***

Experimental Results

Optical Flow

- Comparison of average end-point error (AEPE) with different heights from Moon :

Methods	***Height from Moon (AEPE)***		
	< 1,000 km	*1,000 ~ 3,000 km*	*≥ 3,000 km*
FlowNetS	*13.011*	*15.902*	*15.565*
FlowNetC	*9.020*	*11.839*	*11.987*
FlowNet2	*4.141*	*6.283*	*11.494*
PWC-Net	*8.130*	*10.639*	*10.778*
PA-FlowNet	***2.463***	***2.551***	***1.974***

Experimental Results

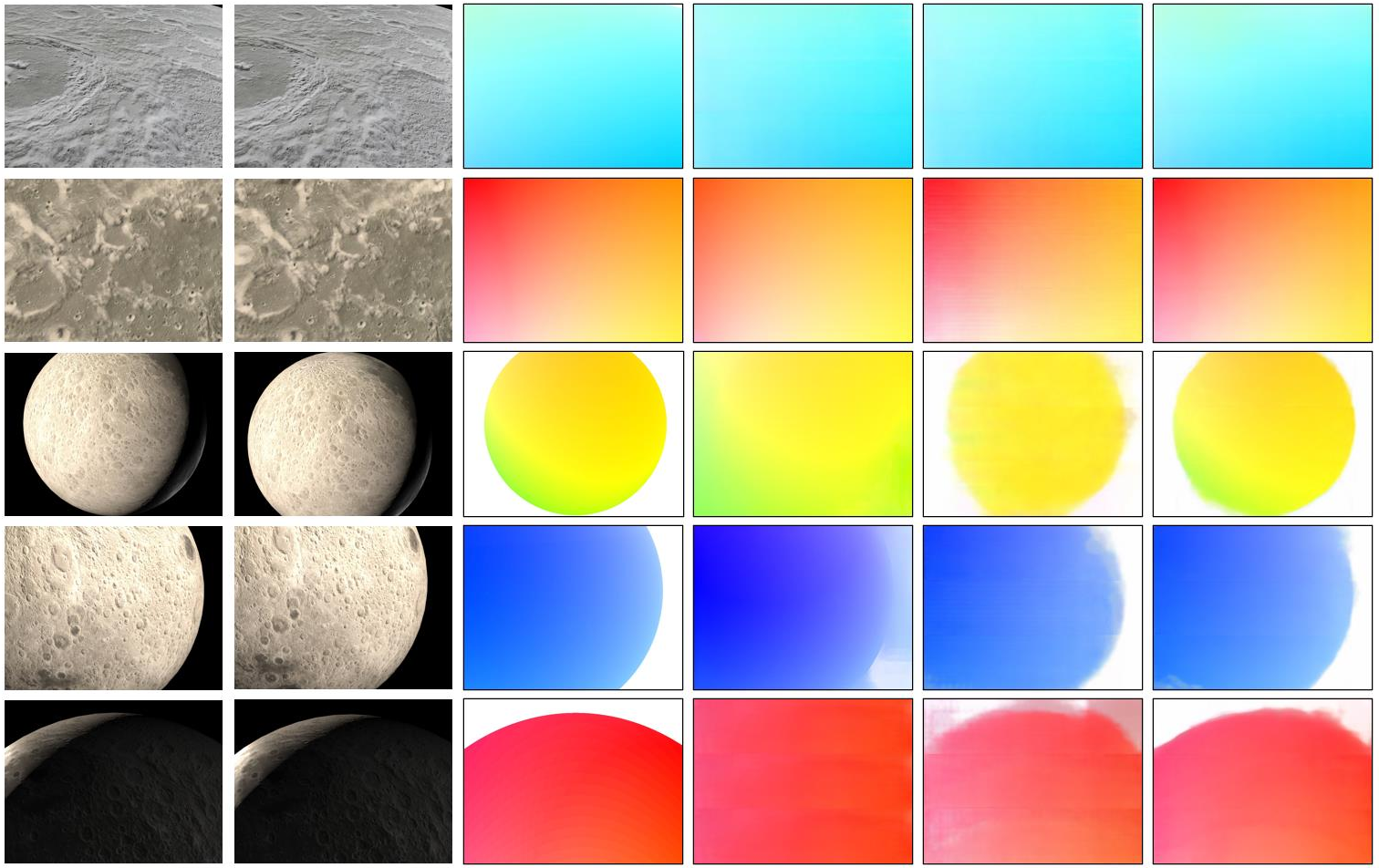
Relative Camera Pose

- Comparison of relative camera pose error on Moon64K dataset.:

Methods	***Relative Camera Pose Error***	
	Orientation Error	*Translation Error*
FlowNet2	*0.987 °*	*0.701*
PWC-Net	*1.025 °*	*0.783*
PA-FlowNet	***0.918 °***	***0.648***

Experimental Results

Optical Flow Results



Source I_s | *Target* I_t | *Ground-Truth* | *PWC-Net* | *FlowNet2* | *PA-FlowNet*

The End

Thanks for Listening!!