OPTICAL FLOW ESTIMATION WITH DEEP NEURAL NETWORKS

INTELLIGENT ROBOTICS – SEMINAR

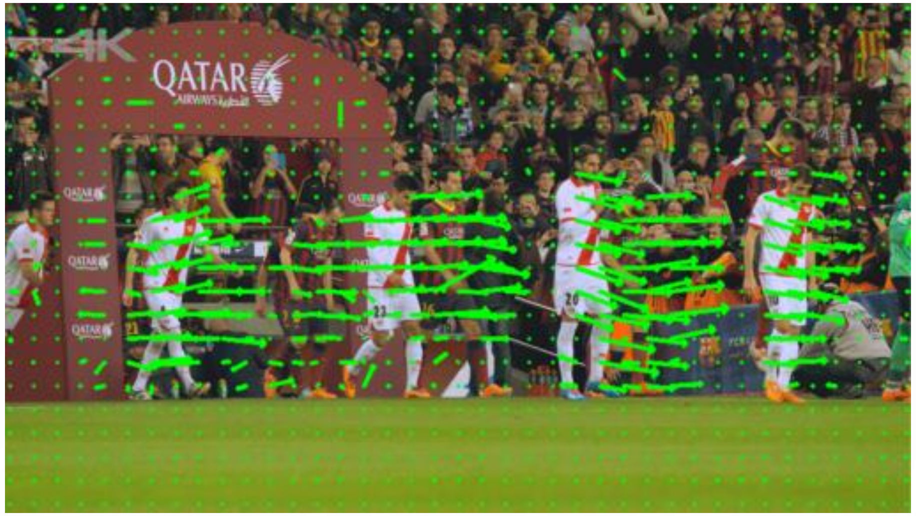
PIA ČUK

25.11.2019

OUTLINE

1. Optical Flow Motivation
2. Neural Networks Basics
3. Optical Flow with Deep Neural Networks
   1. PWC-Net Model
   2. PWC-Net Results
4. Discussion and Outlook

1. OPTICAL FLOW

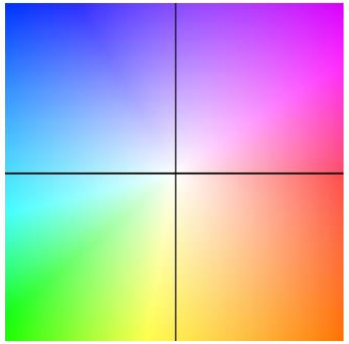


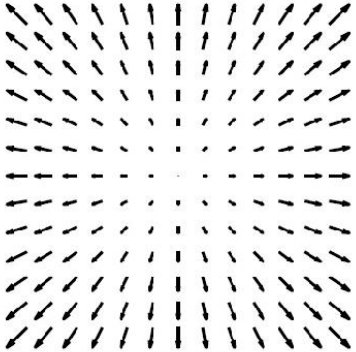
- Motion estimation in video
- “Optical flow is the distribution of apparent velocities of movement of brightness patterns in an image.” [1]
- For subsequent frames, determine displacement vector for each pixel
- https://www.youtube.com/watch?NR=1&v=-F38u9w6YII

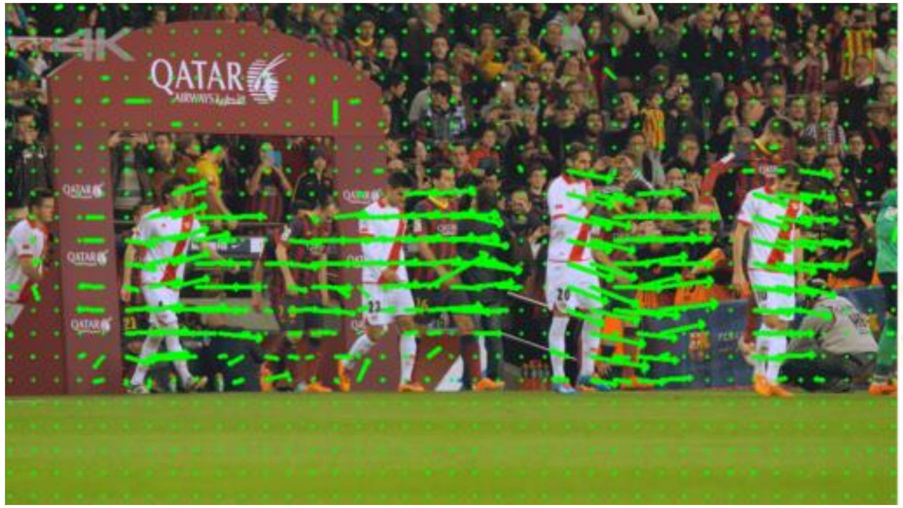
[1] Horn, Berthold KP, And Brian G. Schunck. "Determining Optical Flow." Artificial Intelligence 17.1-3 (1981): 185-203.
https://devblogs.nvidia.com/an-introduction-to-the-nvidia-optical-flow-sdk/, retrieved 18.11.2019

1. OPTICAL FLOW

- Colour code for visualisation:





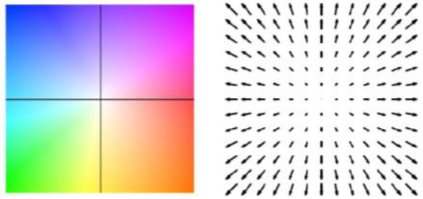


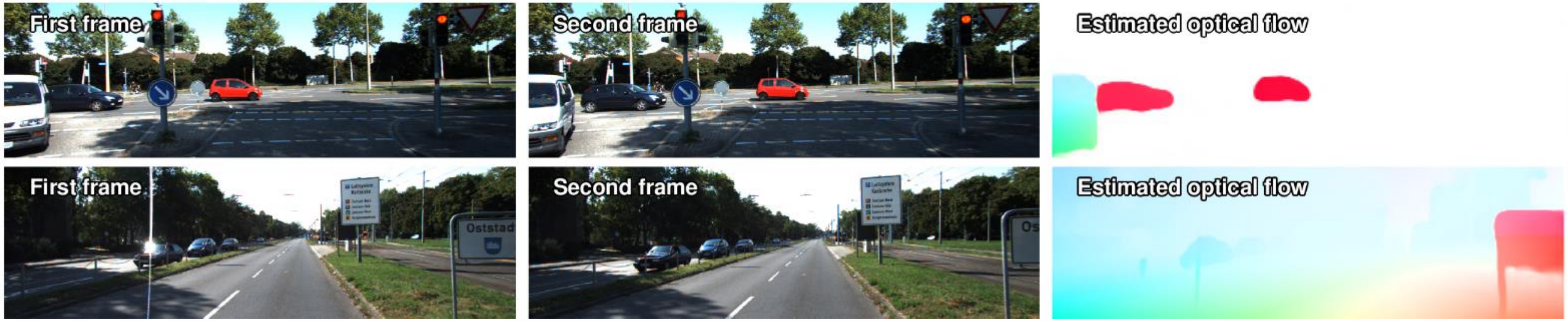
Baghaie, Ahmadreza, Roshan D'Souza, and Zeyun Yu. "Dense descriptors for optical flow estimation: a comparative study." Journal of Imaging 3.1 (2017): 12.
https://devblogs.nvidia.com/an-introduction-to-the-nvidia-optical-flow-sdk/, retrieved 18.11.2019

1. OPTICAL FLOW

- Possible applications: visual odometry, autonomous driving, semantic segmentation…

  →Whenever motion conveys useful information

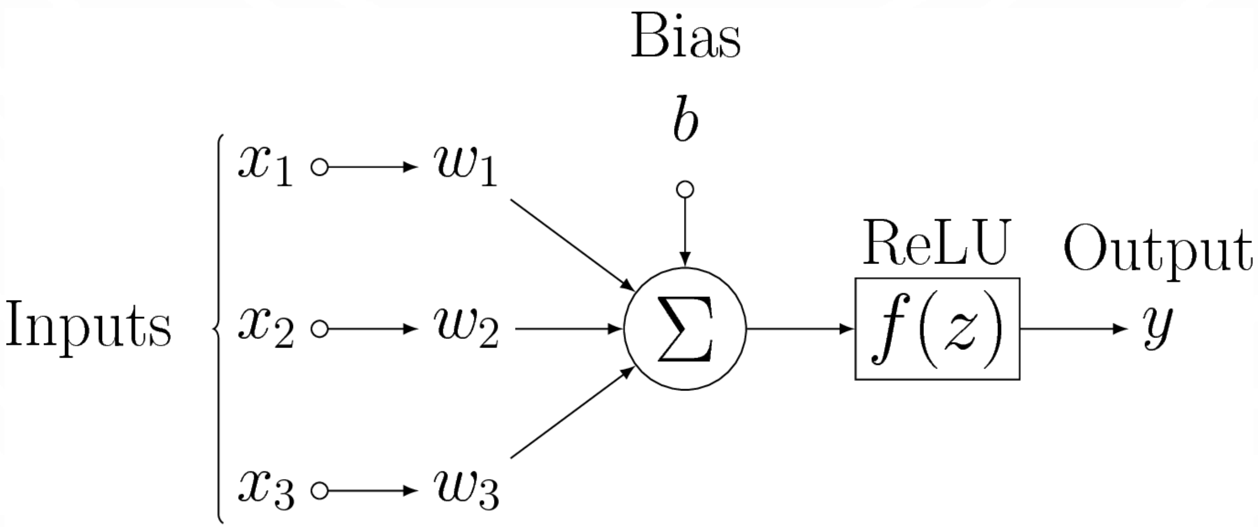




Sun, Deqing, et al. "PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.

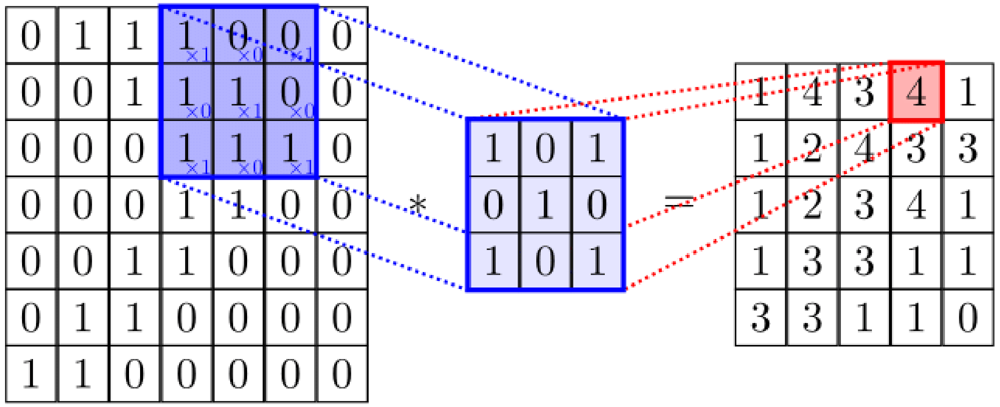
2. NEURAL NETWORKS BASICS

- Inspired by neural networks in the human brain
- Neuron as atomic unit
- Deep neural networks: neurons organised in layers



2.1. CONVOLUTIONAL NEURAL NETWORKS

- Class of deep neural networks well-suited for computer vision
- Use one filter kernel for whole image, “move” it along width, height axes → multiply at every position
- Also called “feature extraction”



3. OPTICAL FLOW WITH DEEP NEURAL NETWORKS

- “Classical” approaches: complex optimization problems, computationally expensive

  →Not suitable for real-time applications

- First DNN approaches: trade-off between accuracy and size of the model

- No end-to-end training

3.1. PWC-NET

Sun, Deqing, et al. "PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.

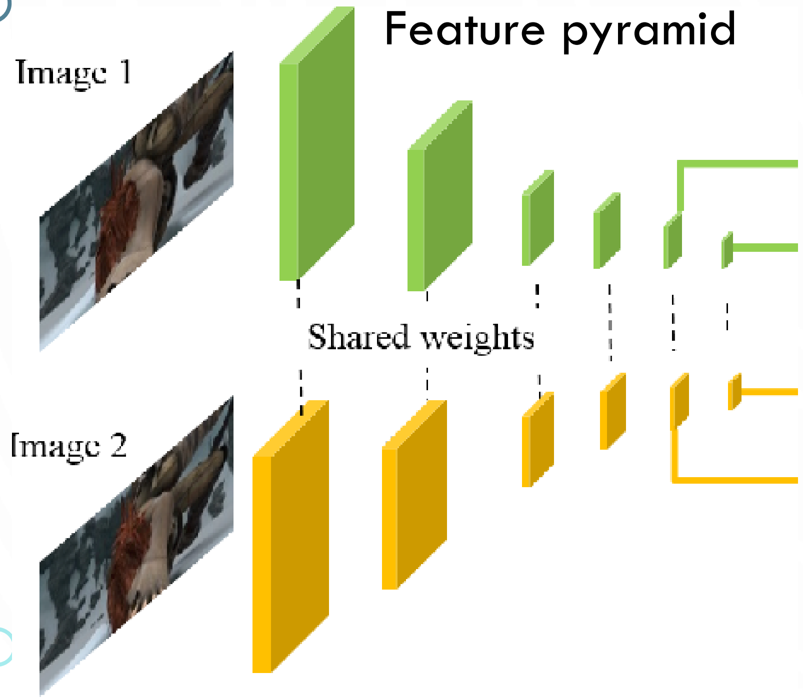
- Uses domain knowledge to reduce complexity
- State-of-the-art accuracy with end-to-end training

3.1. PWC-NET

PWC: Pyramid, Warping, Cost volume

1. Feature extraction from input images with feature pyramid, i.e. convolutional layers
   - Reduction of spatial resolution
2. Optical flow estimation for every level of feature pyramid
   - Start with last convolutional layer, finish on input level
   - Warping and cost volume used in optical flow estimation

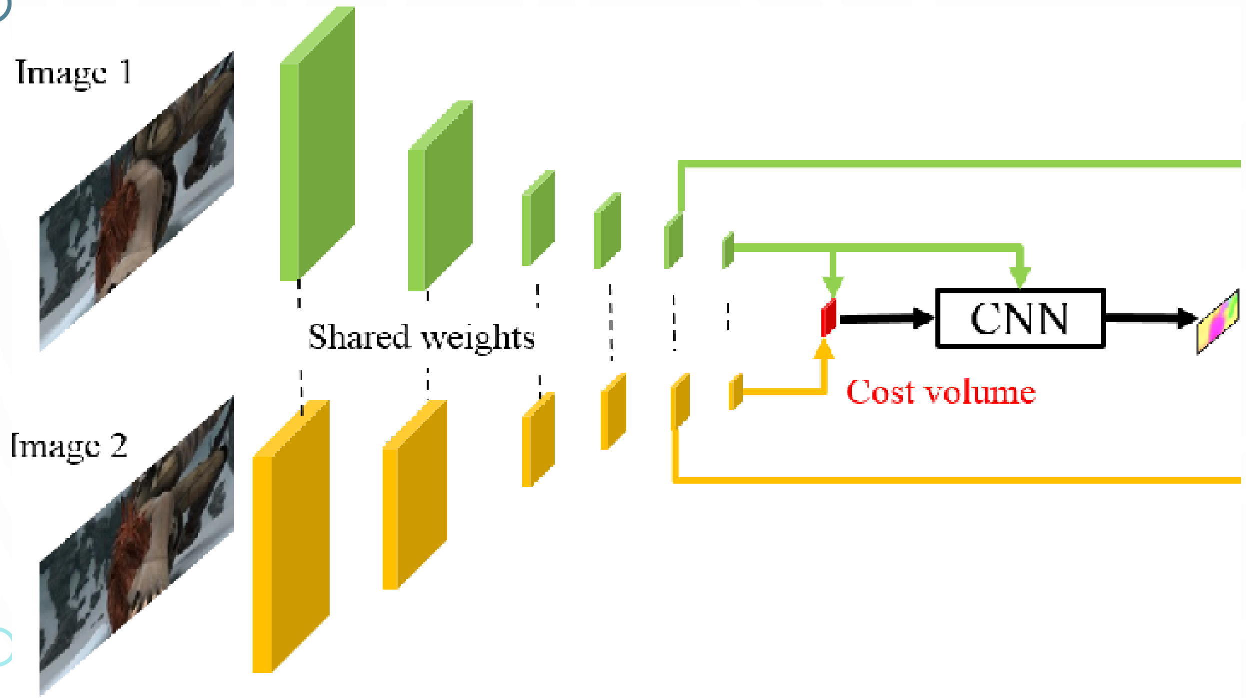
3.1. PWC-NET



1. Compute cost volume: find most similar pixel in features for other image

Sun, Deqing, et al. "Models matter, so does training: An empirical study of cnns for optical flow estimation.“ *arXiv preprint arXiv:1809.05571* (2018).

3.1. PWC-NET



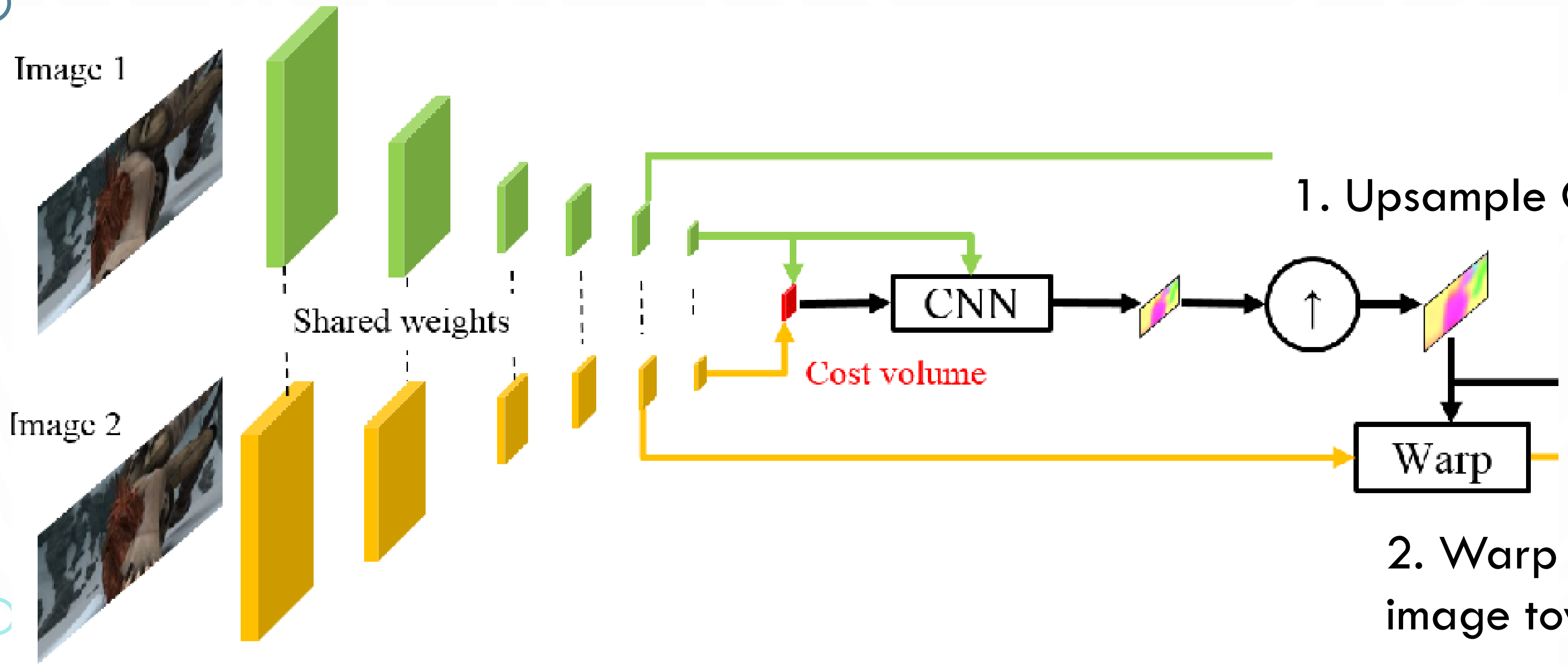
2. Optical flow estimation:

- Cost volume
- Features of first image

→ Output OF for lowest level

Sun, Deqing, et al. "Models matter, so does training: An empirical study of cnns for optical flow estimation.“ *arXiv preprint arXiv:1809.05571* (2018).

3.1. PWC-NET



1. Upsample OF to match spatial dimensions

2. Warp the features of the 2nd image towards the 1st image

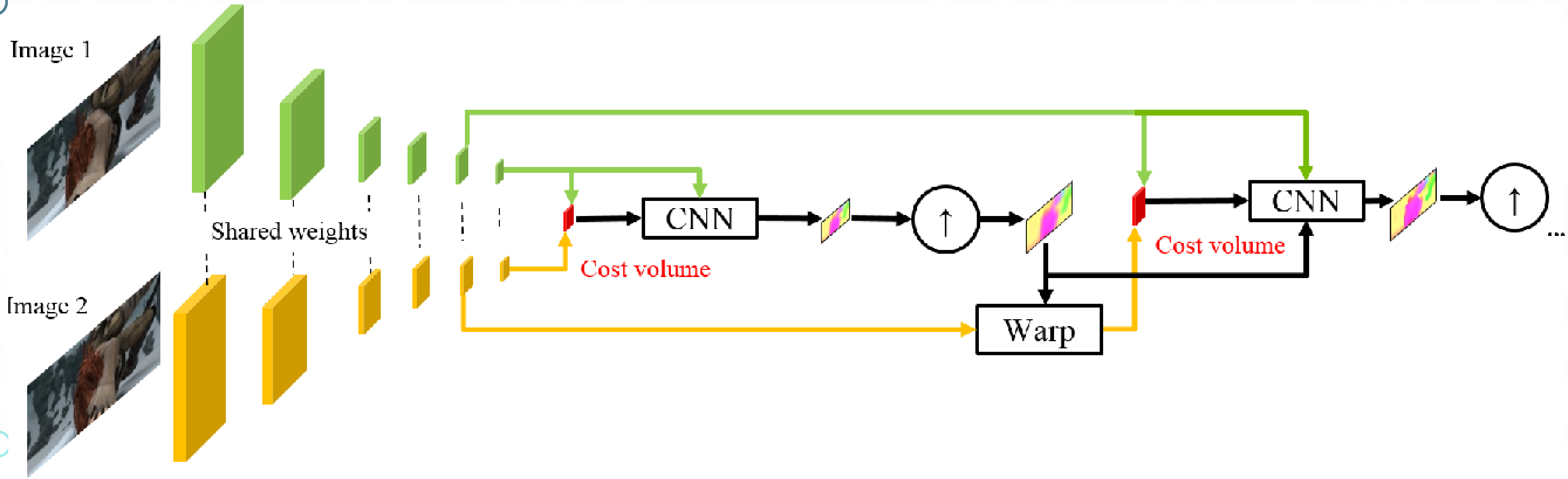
Sun, Deqing, et al. "Models matter, so does training: An empirical study of cnns for optical flow estimation.“ *arXiv preprint arXiv:1809.05571* (2018).

WHY WARPING?

- Second image becomes more similar to first image
- Pixel displacement becomes smaller
- For finding corresponding pixel in cost volume, only need to look at neighbourhood of pixel

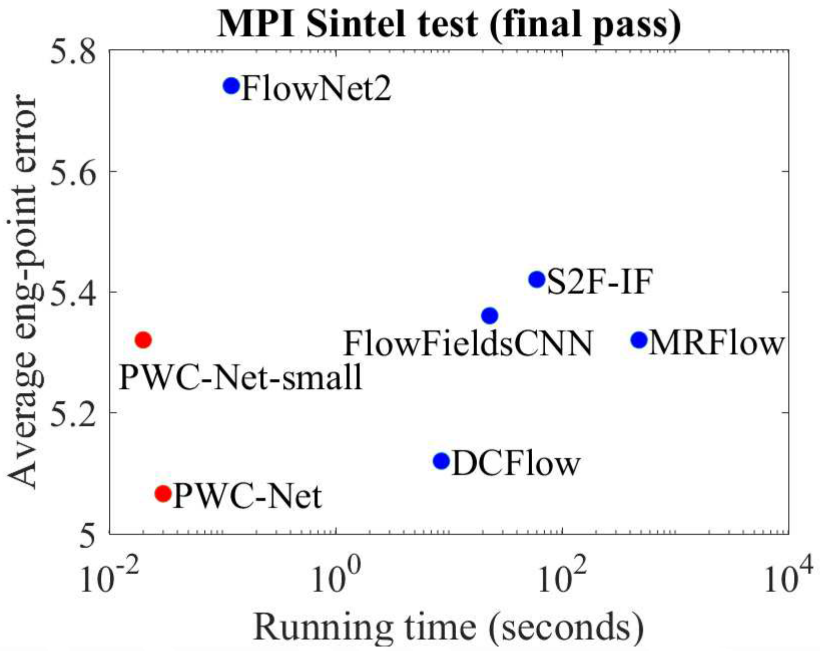
→Computationally much more effective

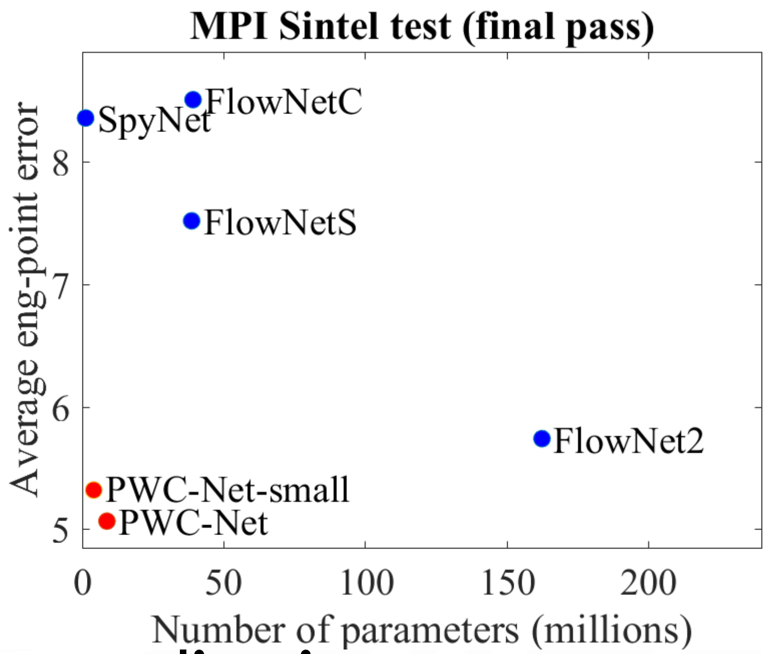
3.1. PWC-NET



Sun, Deqing, et al. "Models matter, so does training: An empirical study of cnns for optical flow estimation.“ *arXiv preprint arXiv:1809.05571* (2018).

3.2. PWC-NET RESULTS





- Inference fast enough for real-time application
- PWC-Net-small for mobile applications

Sun, Deqing, et al. "PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.

3.2. PWC-NET RESULTS

- https://www.youtube.com/watch?v=rCoUcjSz9nQ

Sun, Deqing, et al. "PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.

4. DISCUSSION AND OUTLOOK

- First DNN model to outperform all classical approaches on all popular benchmarks
- Code publicly available: https://github.com/NVlabs/PWC-Net
- Follow-up paper: Sun, Deqing, et al. "Models Matter, So Does Training: An Empirical Study of CNNs for Optical Flow Estimation." *arXiv preprint arXiv:1809.05571* (2018).
- To be improved: occlusion detection, unsupervised training

THANK YOU FOR YOUR ATTENTION!

1. OPTICAL FLOW

- Error metric: Endpoint Error (EPE)
  - Euclidian distance between estimated and ground truth vector for one pixel:

$$\left\|V_{est} - V_{gt}\right\|$$

  - Compute average EPE for all pixels of an image pair: AEPE