

# Video Interpolation by Event-driven Anisotropic Adjustment of Optical Flow

Song Wu<sup>\*1</sup>, Kaichao You<sup>\*2</sup>, Weihua He<sup>\*\*2</sup>, Chen Yang<sup>1</sup>, Yang Tian<sup>2</sup>,  
Yaoyuan Wang<sup>1</sup>, Ziyang Zhang<sup>\*\*1</sup>, and Jianxing Liao<sup>1</sup>

<sup>1</sup> Advanced Computing and Storage Lab, Huawei Technologies Co. Ltd  
{wusong5533, yangchen2017.pku}@gmail.com

{wangyaoyuan1, zhangziyang11, liaojianx}@huawei.com

<sup>2</sup> Tsinghua University

{ykc20, hwh20, tiany20}@mails.tsinghua.edu.cn

**Abstract.** This document is our supporting information of our paper **Video Interpolation by Event-driven Anisotropic Adjustment of Optical Flow**. In this document, we represent our architecture of Event-driven Optical Flow Mask Generation Network and refinement network. Besides, we provide a video demo to show the performance of our proposed methods.

## 1 The Architecture of Proposed Modules

### 1.1 Event-driven Optical Flow Mask Generation Network.

As shown in Fig. 1, our proposed Event-driven Optical Flow Mask Generation Network is a U-Net style network. The input is both event representation and the originals frames. The output is the mask of the bi-directional optical flow in the different direction in a anisotropic way.

### 1.2 The Refinement Network

The architecture of our refinement network in proposed VFI model is shown in Fig. 2. This refinement network is also a U-Net style network which is similar to that in [2].

### 1.3 The Tailored IFNet

As shown in Fig. 3, we tailor the IFNet[1] to better model the motion with event-driven optical mask and event representation.

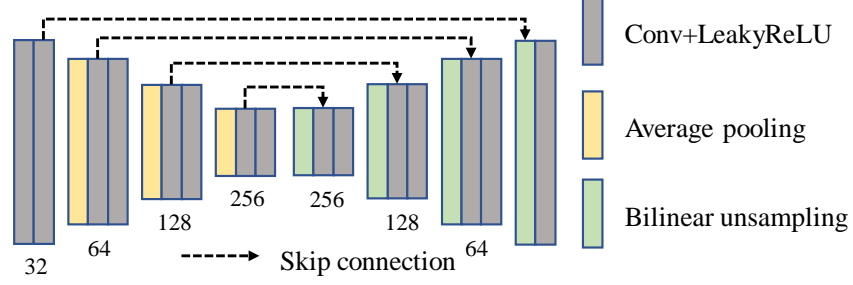
## 2 The Video Demo

Here, we provide a demo to show the performance of our VFI model. The demos are provided in our submitted files.

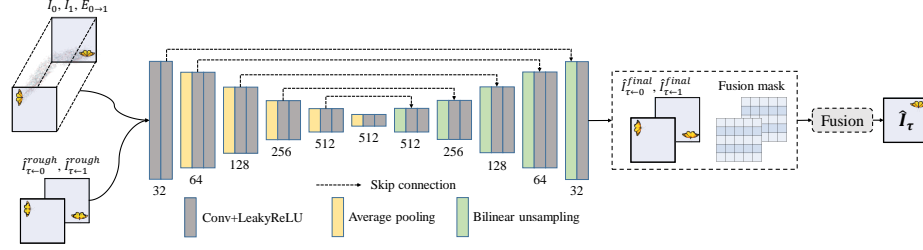
---

\* Song Wu and Kaichao You contribute equally to this paper. Work done while Song Wu, Kaichao You, Yang Tian are interns at Huawei.

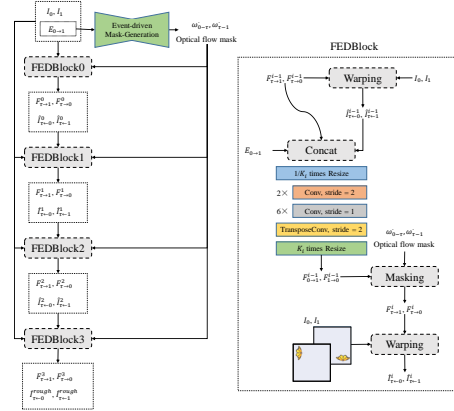
\*\* Weihua He and Ziyang Zhang are corresponding authors.



**Fig. 1.** The architecture our proposed Event-driven Optical Flow Mask Generation Network. The number below each block denotes the channel of corresponding output feature map.



**Fig. 2.** The pipeline of refinement process.



**Fig. 3.** The architecture of tailored IFNet.

## References

1. Huang, Z., Zhang, T., Heng, W., Shi, B., Zhou, S.: RIFE: real-time intermediate flow estimation for video frame interpolation. CoRR **abs/2011.06294** (2020), <https://arxiv.org/abs/2011.06294>
2. Jiang, H., Sun, D., Jampani, V., Yang, M., Learned-Miller, E.G., Kautz, J.: Super slomo: High quality estimation of multiple intermediate frames for video interpolation. In: 2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018. pp. 9000–9008 (2018)