EdgeMask: An Edge-based Privacy Preserving Service for Video Data Sharing

Samira Taghavi and Weisong Shi

Wayne State University
http://thecarlab.org

Outline

- Research problem
- Literature review
- Background
- Contributions
- High level architecture
- Algorithm
- Evaluation
- Future works

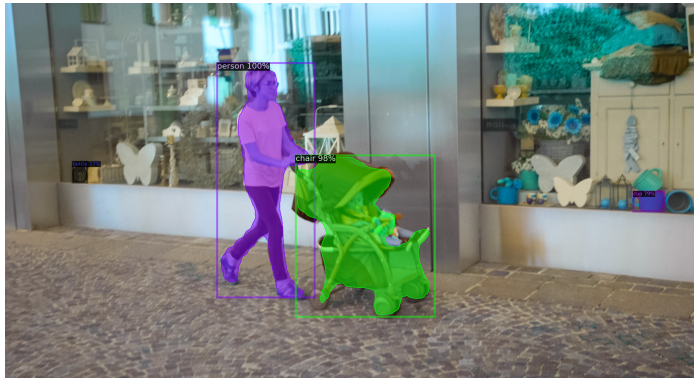
Research problem

- ADAS, SLAM, and smart cities required image or video data from public places

- These data may contain sensitive personal identifying data
- Any using or sharing these raw datasets can have legal implications

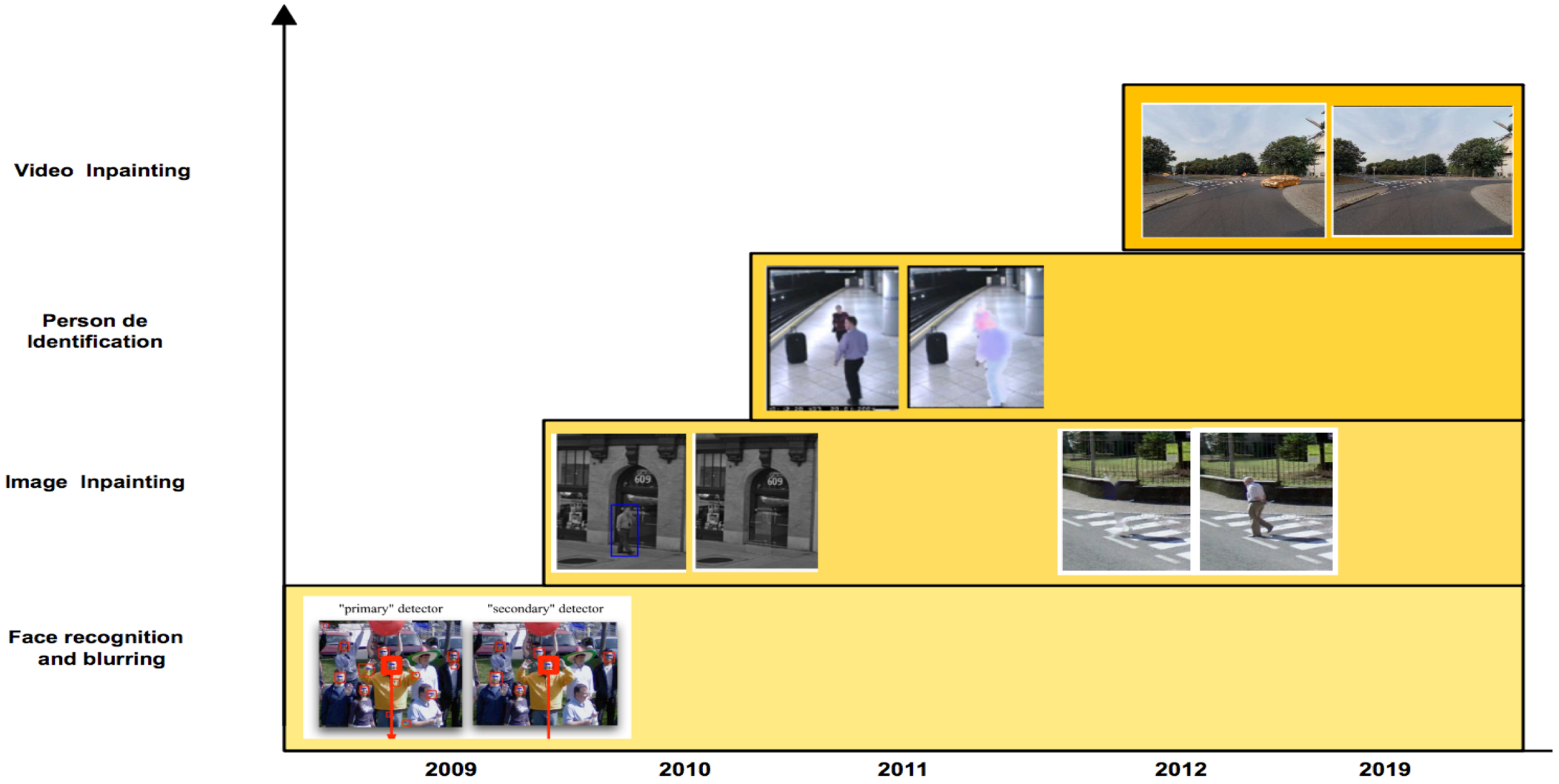
Edge Mask: Goals and objectives

- Privacy protection of individuals in video data
  - Blurring sensitive information



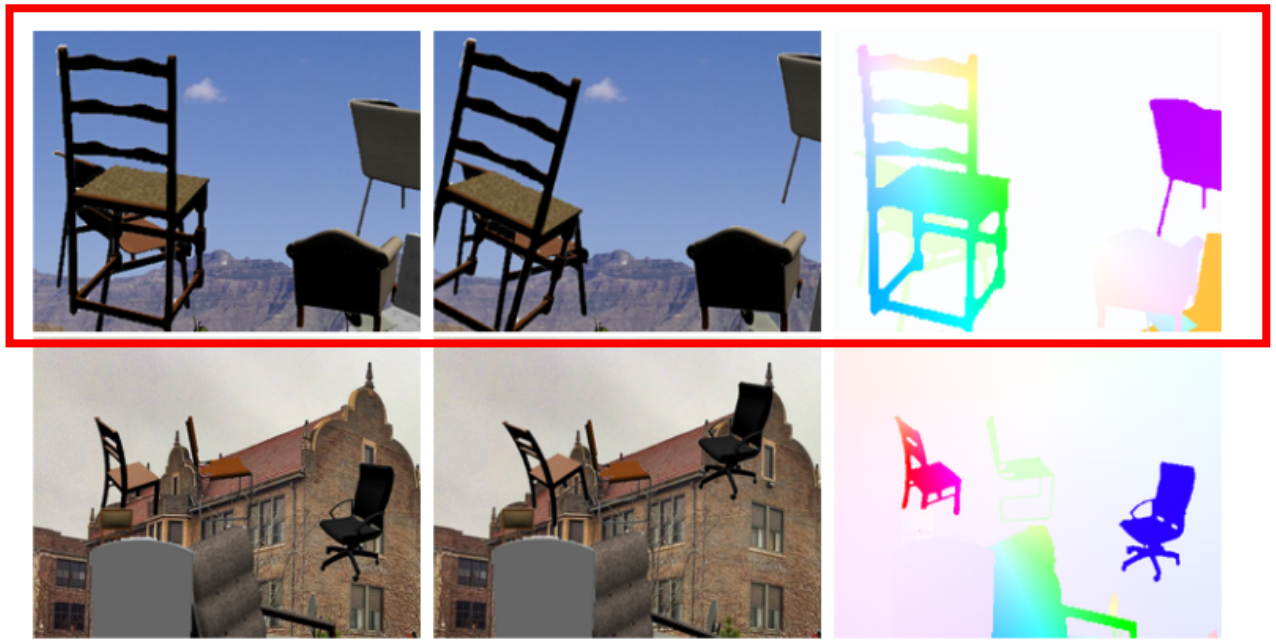
  - Undesired object removal

Related works

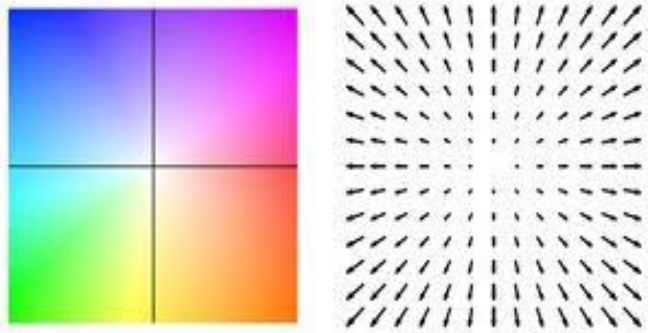


Background

What is optical flow?

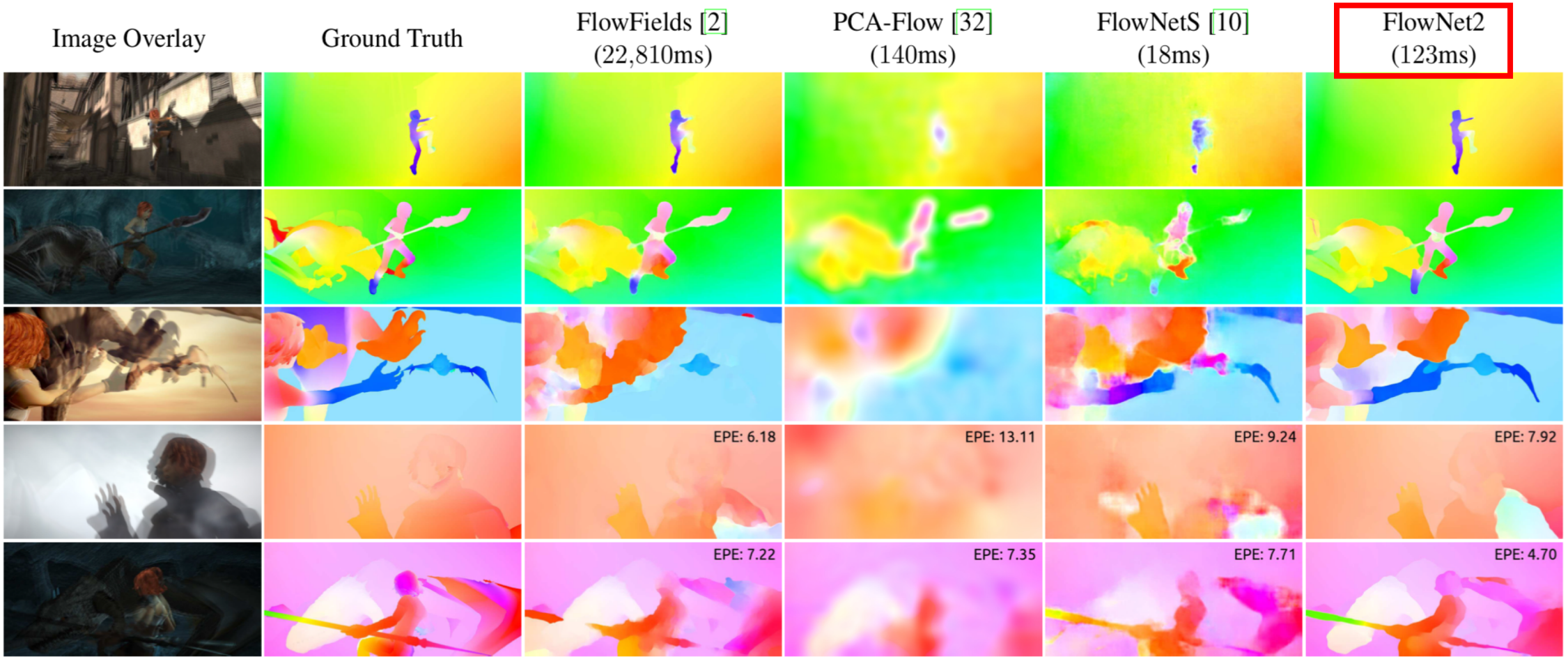


Flying chair dataset



- **2-Dimensional Map**
- **Color : Direction**
- **Saturation: Magnitude**

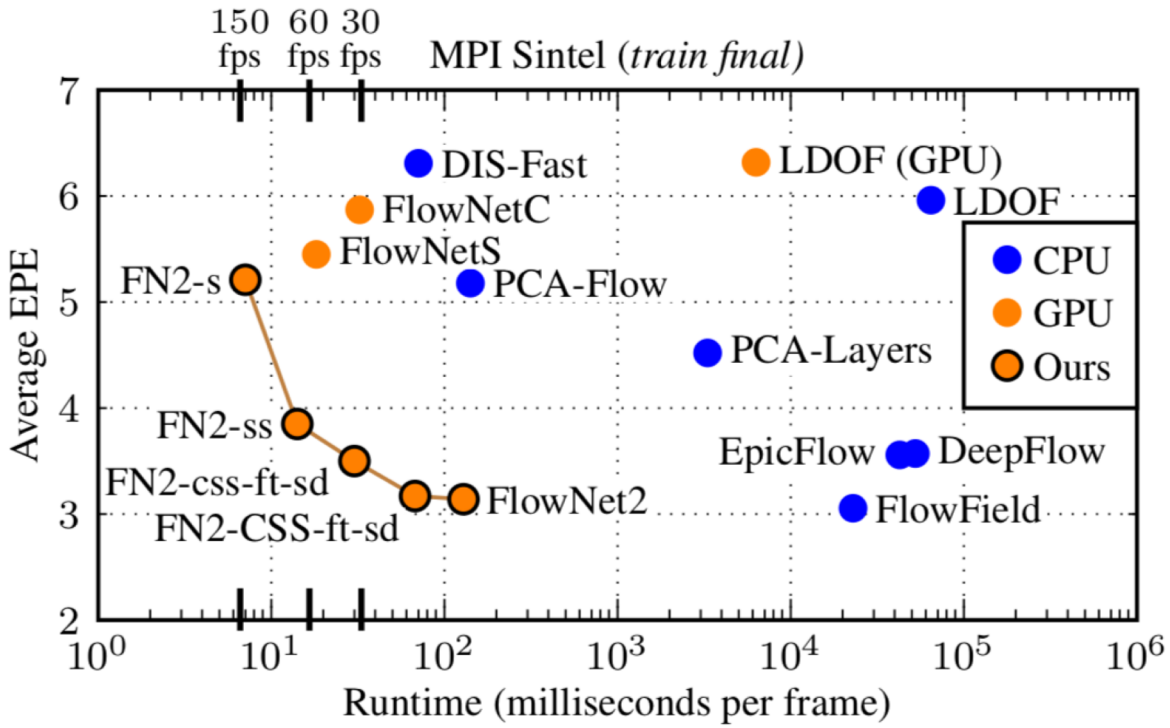
Background



Ilg et al. FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks, CVPR 2018

Background

	Number of Networks			
	1	2	3	4
Architecture	s	ss	sss	
Runtime	**7ms**	14ms	20ms	–
EPE	4.55	3.22	3.12	
Architecture	S	SS		
Runtime	18ms	37ms	–	–
EPE	3.79	2.56		
Architecture	c	cs	css	csss
Runtime	17ms	24ms	31ms	36ms
EPE	3.62	2.65	2.51	2.49
Architecture	C	CS	CSS	
Runtime	33ms	51ms	69ms	–
EPE	3.04	2.20	**2.10**	



Ilg et al. FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks, CVPR 2018

Contributions

❑ System Level

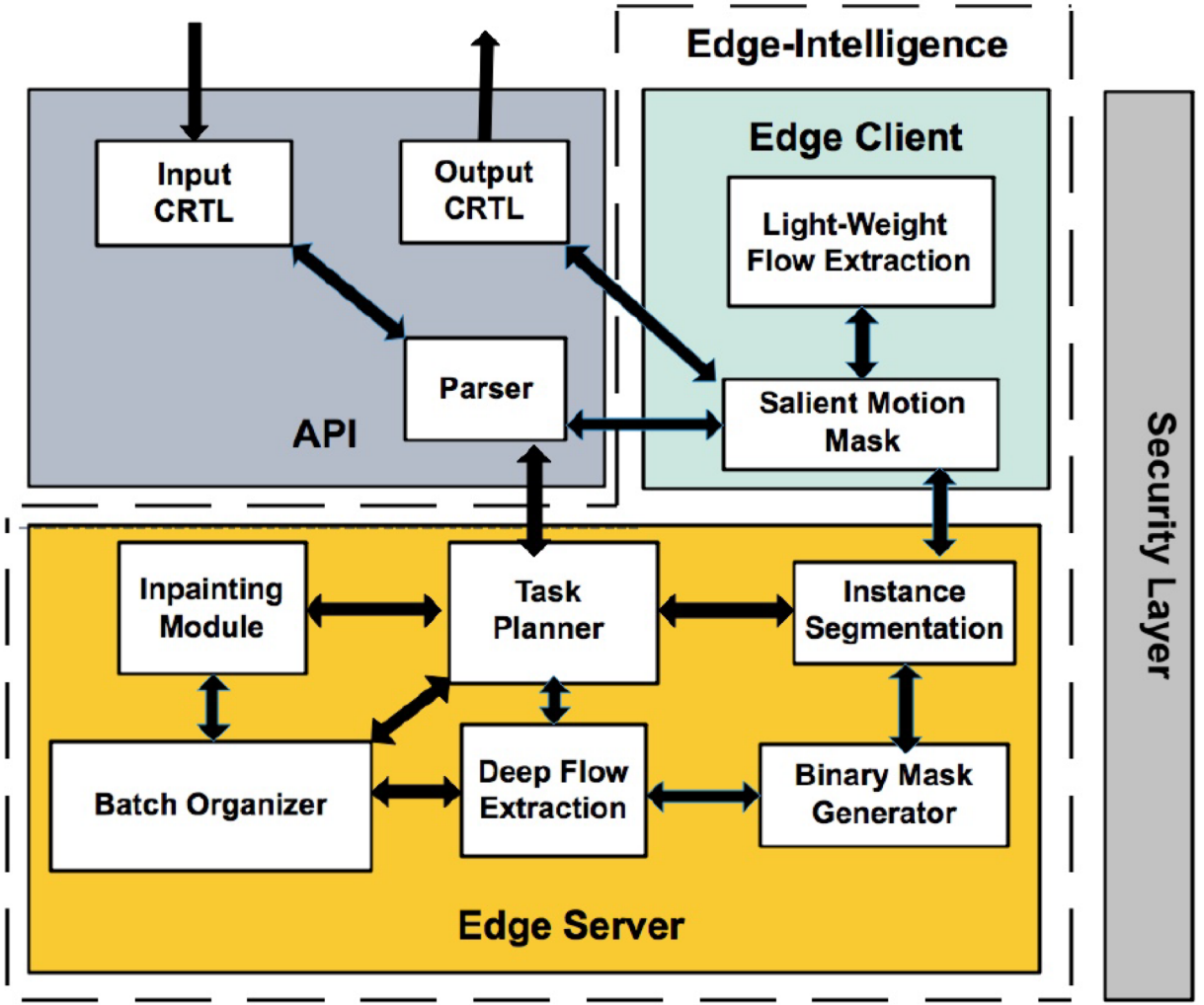
- The first approach that leverages both object blurring and removal technique
- An edge-based real-time privacy preserving framework
- Object segmentation without requiring complex network architectures

❑ Algorithm Level

- Tracking camera velocity and direction based on optical flow for static object segmentation
- Generating object mask for video inpainting based on optical flow and Mask R-CNN

High-Level Architecture

- Edge-Intelligence
- Application Programming Interface
- Security Layer



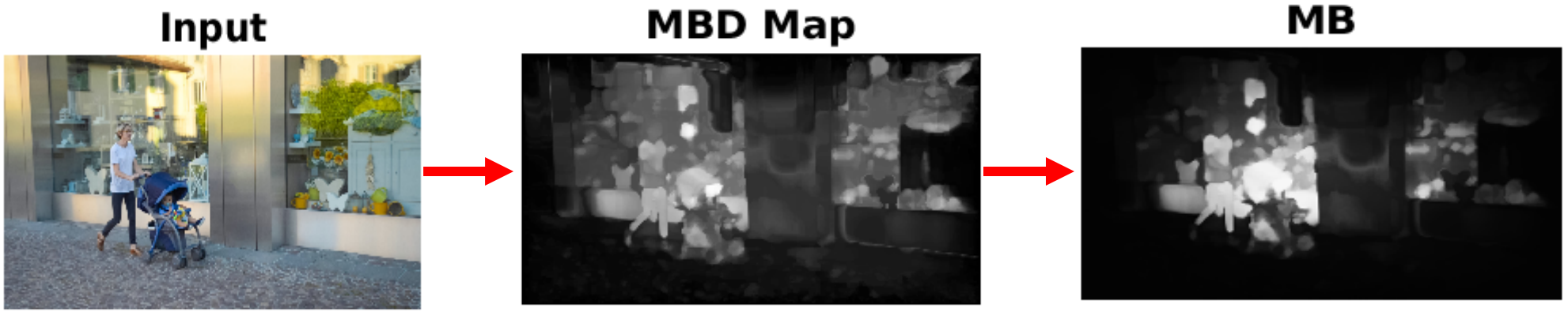
Algorithm Design

Dynamic object segmentation and blurring • FlowNet2.0 • MBD	Static object segmentation and blurring • Mask R-CNN • MBD
Congested scene analysis • FlowNet2.0 • MBD	Video inpainting • Mask R-CNN • MBD • FlowNet2.0 • Deep Flow Guided video inpainting

Dynamic Object Segmentation

"Zhang et al. Minimum Barrier Salient Object Detection at 80 FPS"

- Original image



- Optical flow

Dynamic Object Segmentation

$$S_i^t(F^t) = \sum_{\forall F_j^t \in F_t} d(F_i^t, F_j^t)$$

"d" stands for Minimum Barrier Distance (MBD) Transform

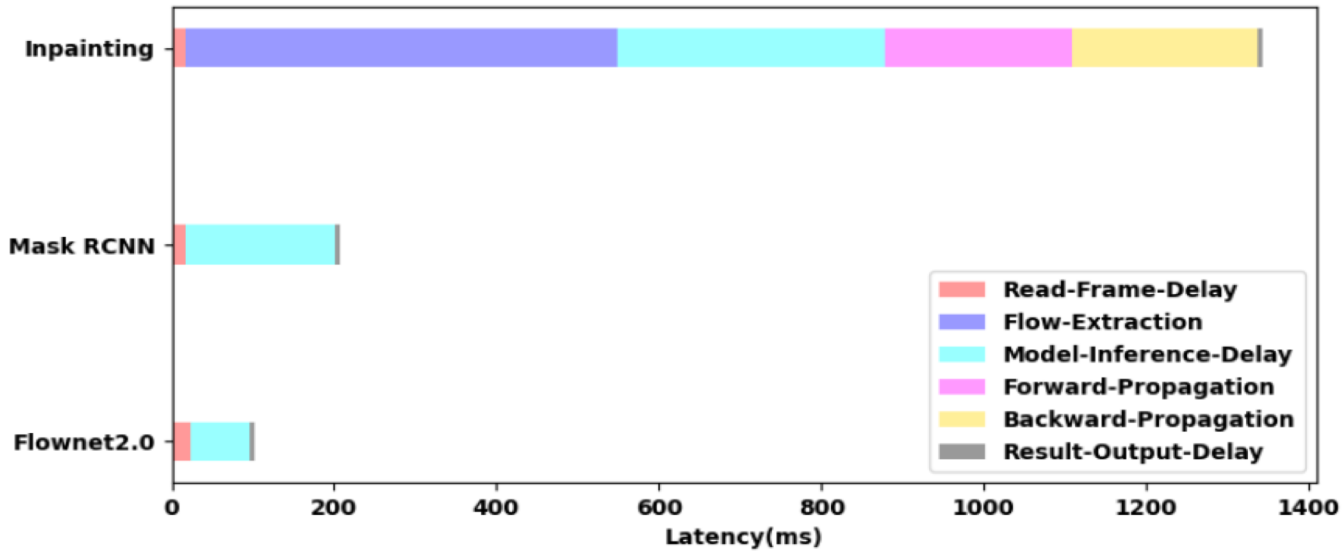
$$\beta_{\mathcal{I}}(\pi) = \max_{i=0}^{k} \mathcal{I}(\pi(i)) - \min_{i=0}^{k} \mathcal{I}(\pi(i)).$$

Static Object Segmentation

1- Find the static background region

$$M_b = \overline{\{M_1, \ldots, M_k\}}$$

$$M_o = \overline{M_b} \ \& \ MBD$$

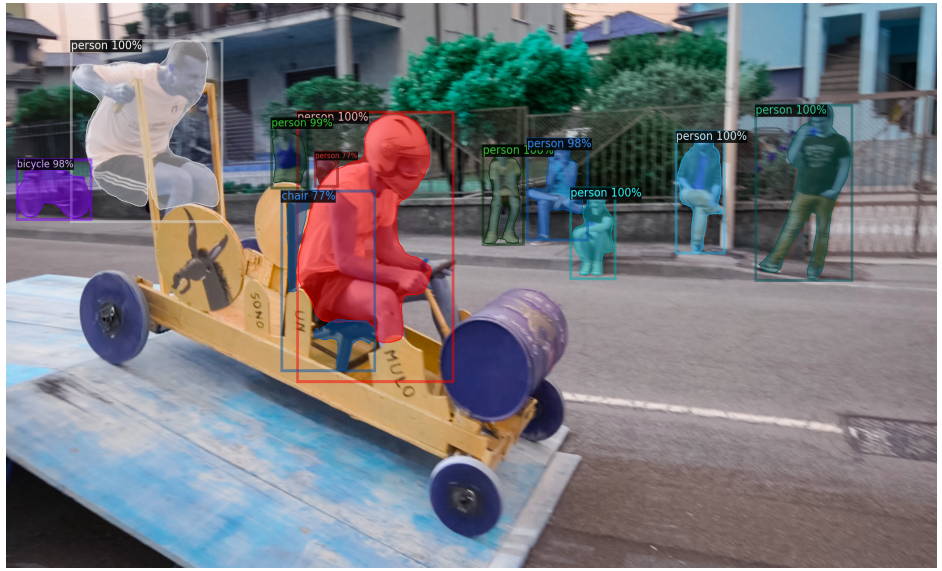


$$\textit{The background region} = \overline{M_o}$$

2- Move the static object masks generated by Mask R-CNN based on the background velocity and direction.

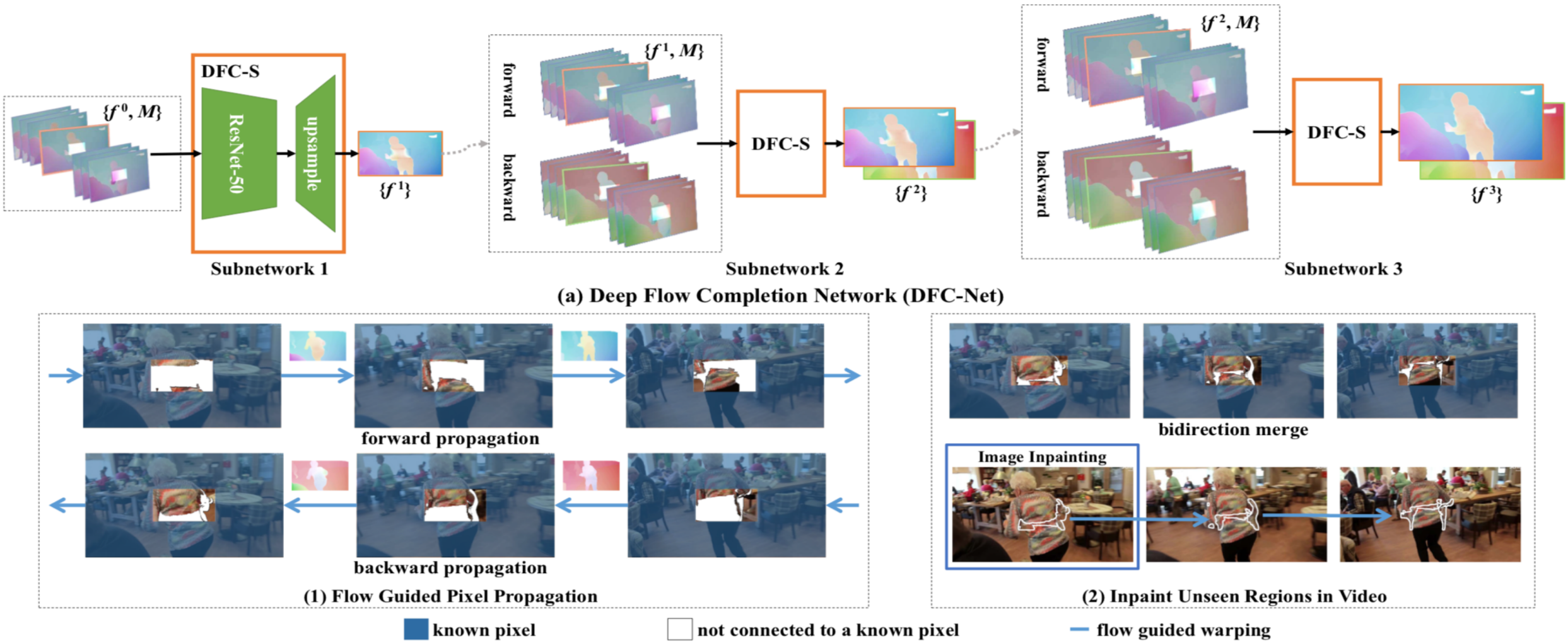
Video Inpainting

- Video inpainting is performed in offline mode
- The required inputs are:
  - Video Frames
  - Object Masks
    - Generated by Mask RCNN and “MBD”
  - Objects’ Optical flows
    - Generated by FlowNet2.0



MBD Map

Video Inpainting

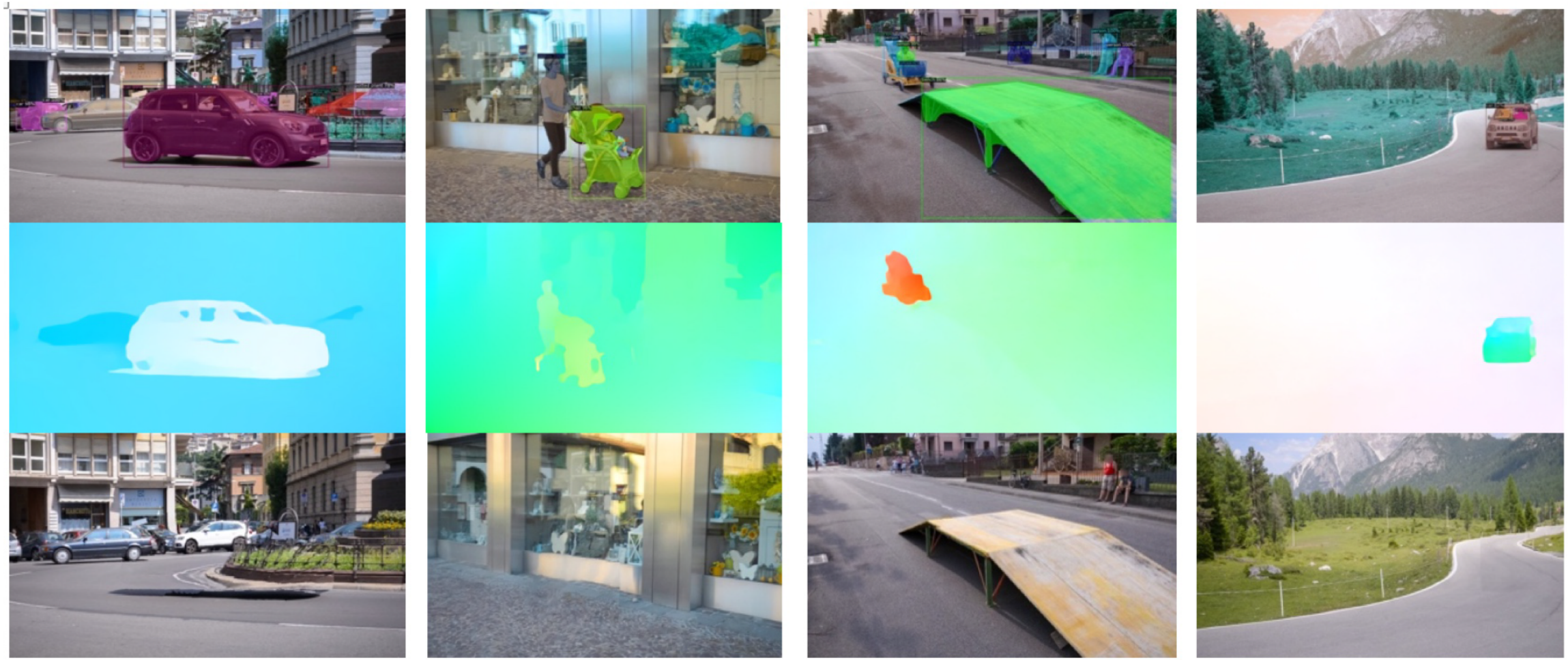


Xu et al. Deep Flow-Guided Video Inpainting, CVPR 2019

Video Inpainting

- Object recognition for parallel video inpainting
  - Object label
  - Color histogram
  - Object direction

Video Inpainting



Experiment Setup

- Dataset

DAVIS 2016

- Hardware setup

Hardware	Model	Memory
GPU	NVIDIA GeForce GTX 1060	3G
VPU	Intel Neural Compute Stick 2	-

- Model setup

Model	Parameter	Value
Mask RCNN	Backbone Pretrained Weights	ResNet101 COCO Dataset
Inpainting	Backbone Flow Extraction Propagation Kernel Enlarged Mask Enlarged	ResNet101 FlowNet2 Bidirectional 50 or 70 YES
Flownet2.0	Backbone Model	ResNet101 FlowNet2.0

Evaluation

- The quality evaluation

Approach	Kernel 70	Kernel 50
Xu et al. [17]	27.838	28.011
EdgeMask	27.851	28.026

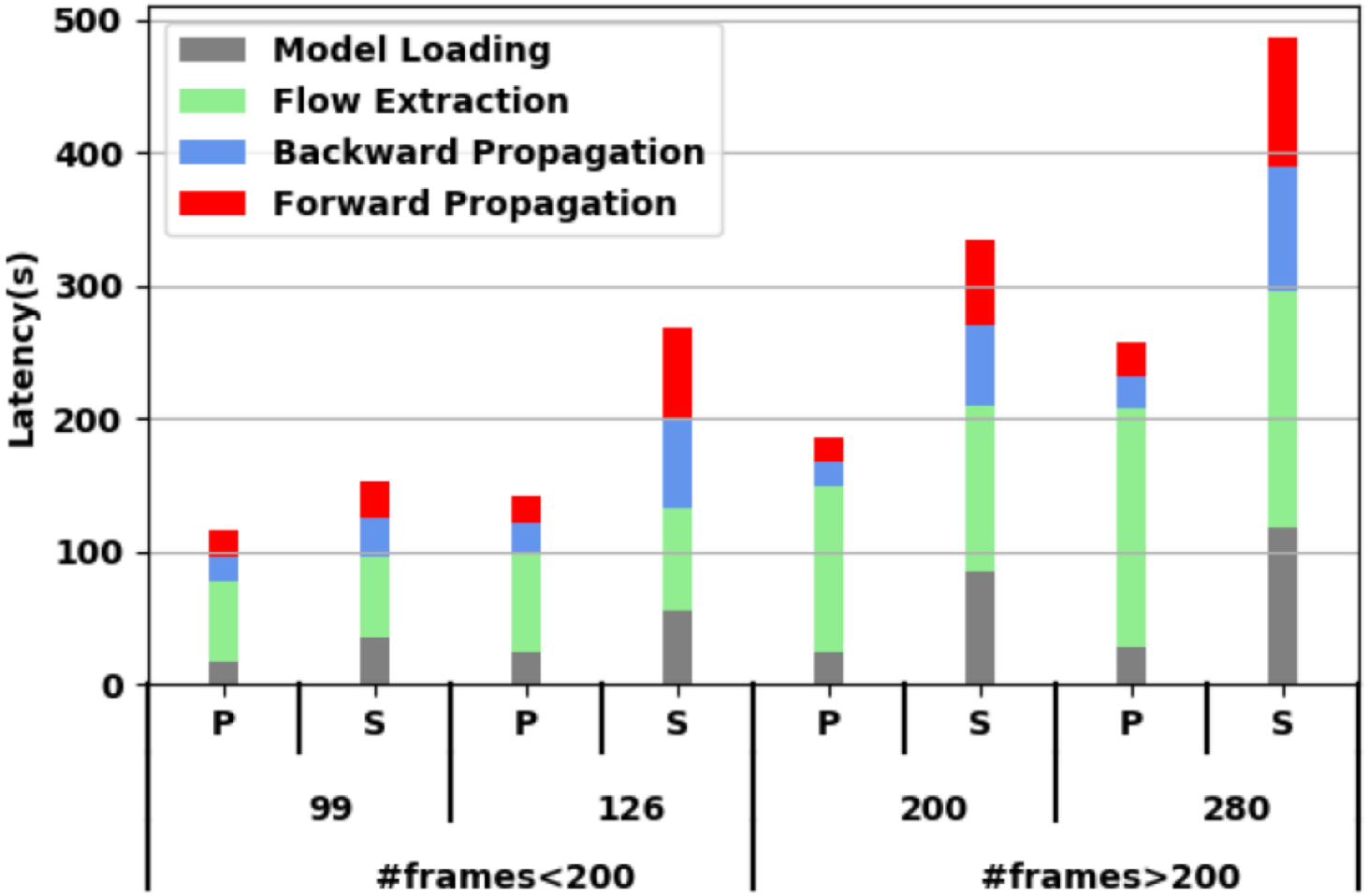
- The computational time based on a single frame

Approach	Object blurring	Object removal
Frome et al. [6]	10.1s	-
Flores et al. [5]	-	31.6s
Agrawal et al. [3]	12s	-
Nodari et al. [12]	-	21.4s
EdgeMask	80ms	1.16s

Evaluation

- Efficiency

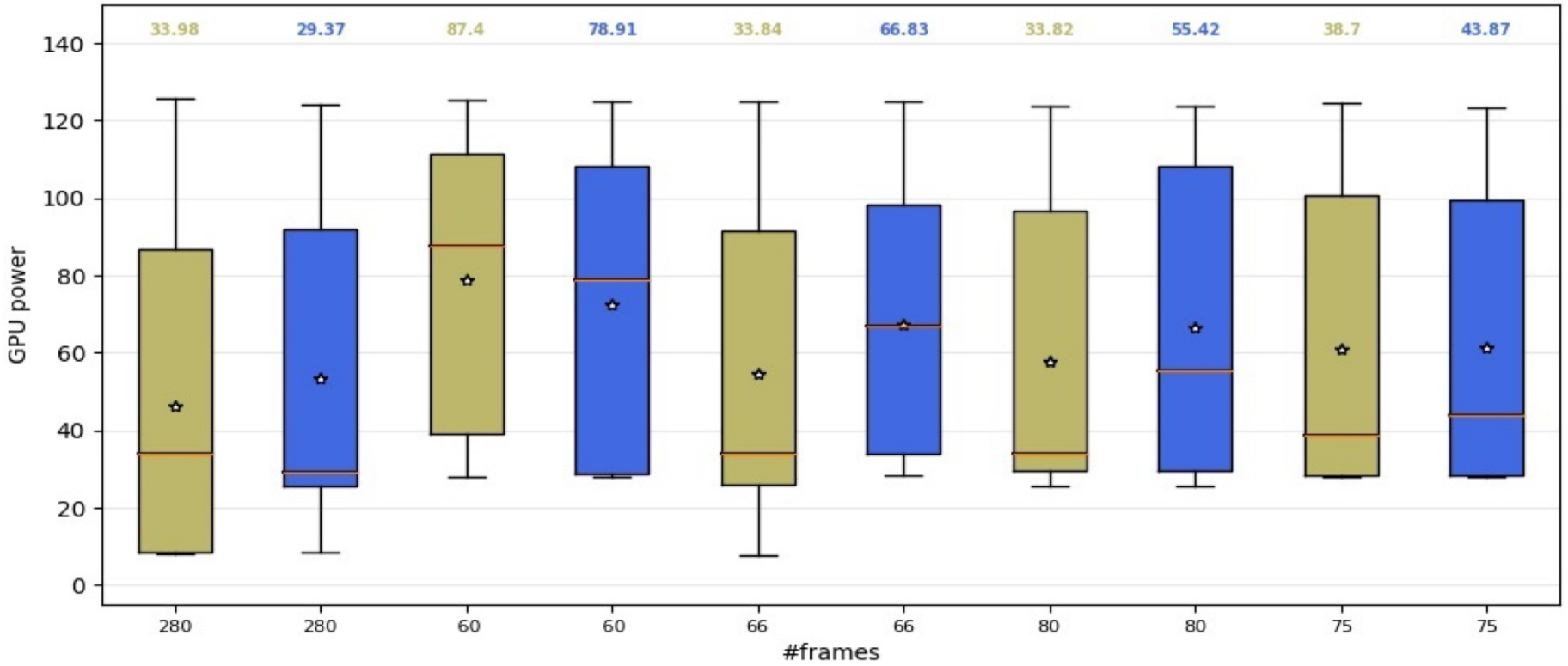
- P stands for parallel
- S stands for serial



Evaluation

- GPU power

The GPU power for parallel and serial computing units (#frames == 280)



Discussion & Future work

- ❑ Neural network pruning for video inpainting algorithm
- ❑ Evaluate in more realistic scenarios
- ❑ Generality of EdgeMask

  EdgeMask workflow can generalize to many other identification tasks

samira.taghavi@wayne.edu