Seatbelt Segmentation Using Synthetic Images

Sumit Jha, Isaac Brooks, Soumitry J. Ray, Rajesh Narasimha, Naofal Al-Dhahir, Carlos Busso

UT Dallas MSP Multimodal Signal Processing Laboratory

THE UNIVERSITY OF TEXAS AT DALLAS

Multimodal Signal Processing Lab (MSP)
Erik Jonsson School of Engineering & Computer Science
University of Texas at Dallas, Richardson, Texas - 75080, USA

UTD CRSS

edgetensor

IEEE IV 2023
IEEE INTELLIGENT VEHICLES SYMPOSIUM
Anchorage, Alaska, USA // June 4 – 7, 2023

Motivation

Background:

- Seatbelt segmentation can help enforce correct seatbelt usage
- Computer vision approaches require annotated training data
- Segmentation annotations are expensive
  - How can we train an accurate model with little to no supervision?

Our Approach:

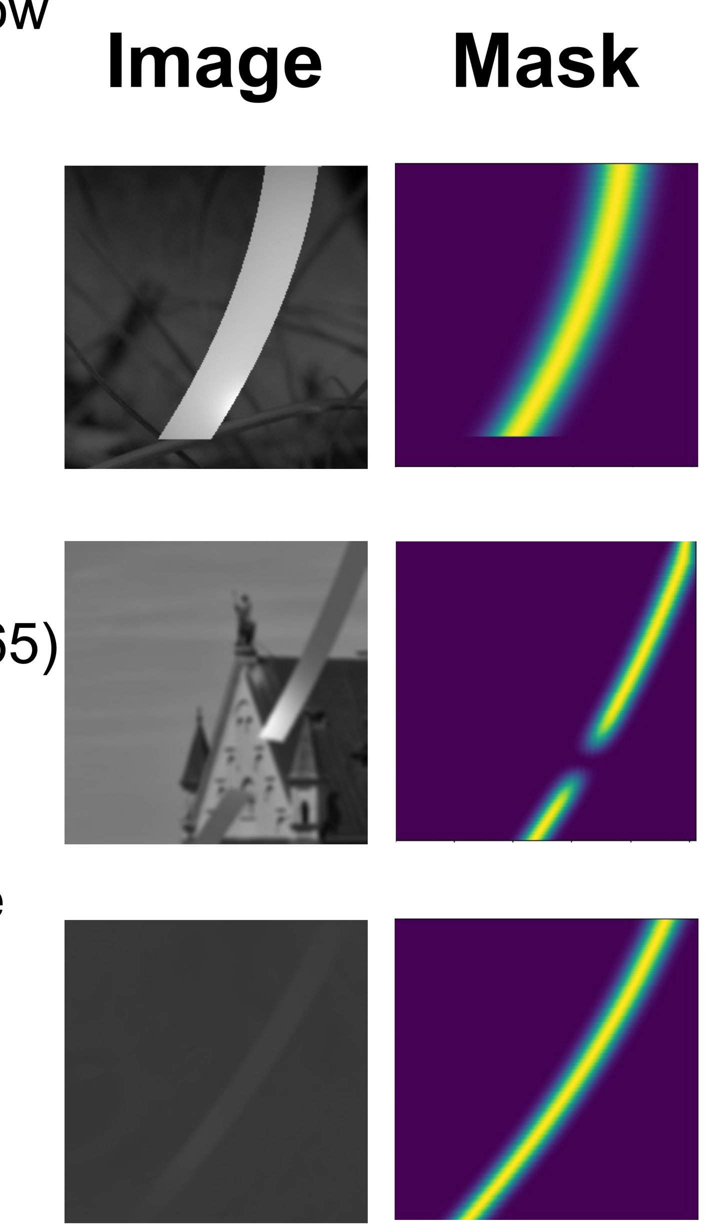
- Generate synthetic seatbelt images for training
- Employ semi-supervised learning techniques to improve performance
- Use a small amount of annotated real images for testing

Synthetic Seatbelt Images

- Observation: Most seatbelt shapes follow a similar pattern

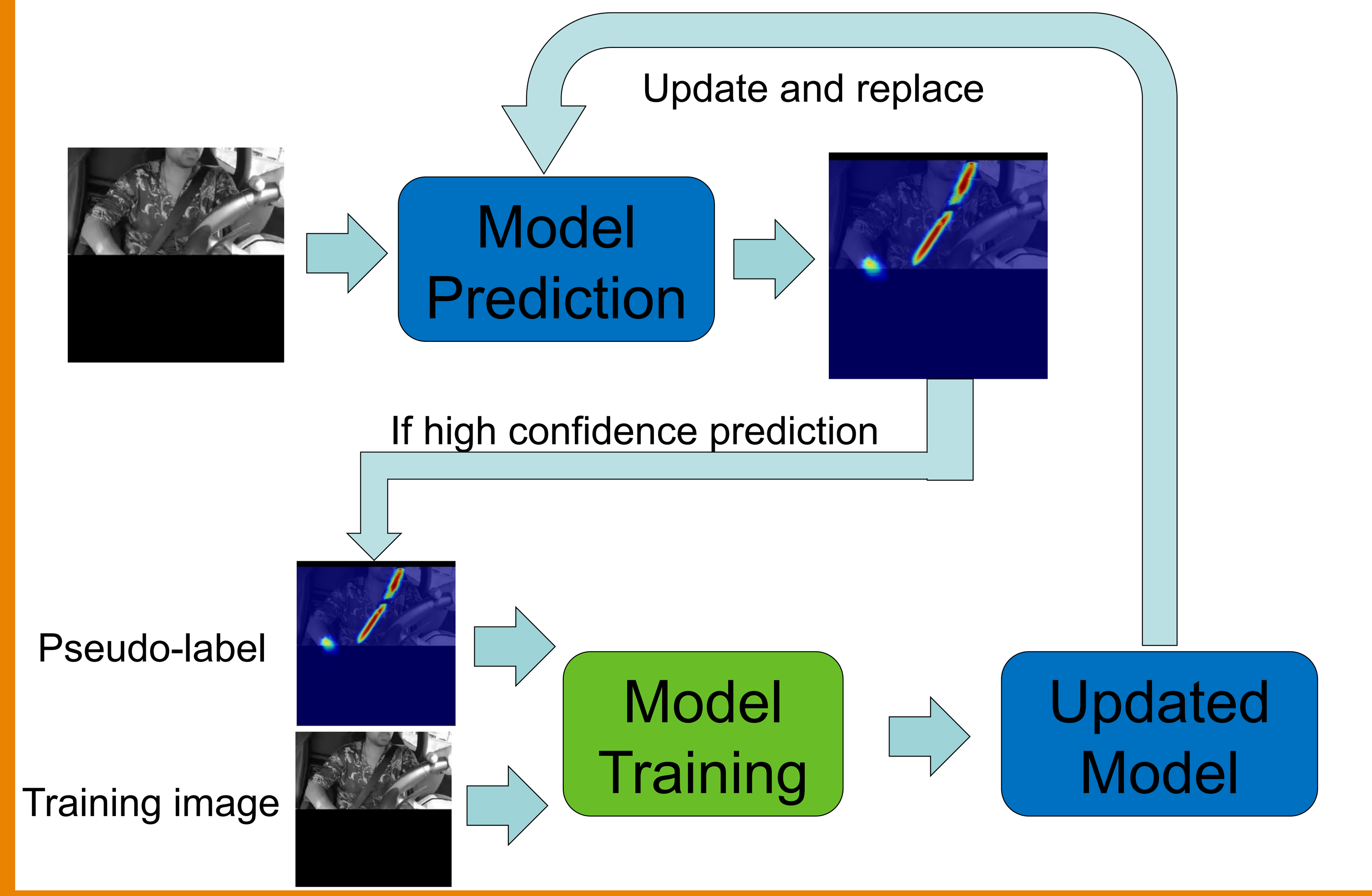
Data Generation

- Interpolate spline between 3 points
  - Two corners, one arbitrary point
- Parameterize thickness, angle, and curvature
- Random background images (Places365)
  - Introduces structured variations in the background
- Apply noise, and occlusions to simulate real-world conditions
- Gaussian blur on the mask to soften prediction boundaries



Pseudo-Labeling

- Step 1: training on only synthetic data
- Step 2: Run predictions on unlabeled naturalistic data
- Step 3: Iteratively update the model using pseudo-labels



Results

Iterative Pseudo-Labeling

- Record performance after N iterations
- Iteration 0 uses only synthetic data

Iterations	Precision	Recall	F1	IoU (%)
0	0.58	0.43	0.50	19.8
1	0.58	0.48	0.53	21.7
2	0.60	0.50	0.54	22.6
3	0.60	0.51	0.55	22.8
4	0.61	0.51	0.55	23.0

- Performance increases steadily, especially in the recall rate
- Four iterations are enough
- These results require no annotated naturalistic samples

Fine-Tuning with Real Data

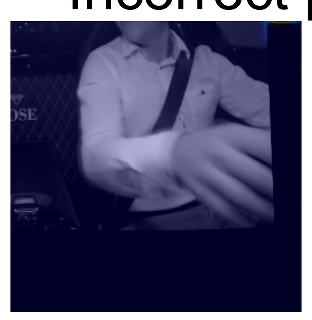
- We fine-tune with a small set of annotated real data
- Compare the performance of Model-Iteration 0 and 4
  - With 200 labeled samples, IoU increases about 10%
  - Performance converges with large amount of annotated data

	Model-Iteration 0				Model-Iteration 4			
Labeled Images	Precision	Recall	F1	IoU (%)	Precision	Recall	F1	IoU (%)
0	0.58	0.43	0.50	19.8	0.61	0.51	0.55	23.0
25	0.63	0.47	0.54	20.3	0.64	0.55	0.59	25.7
50	0.67	0.51	0.58	23.3	0.68	0.57	0.62	27.9
100	0.70	0.57	0.62	27.3	0.71	0.58	0.64	29.9
150	0.71	0.58	0.64	29.4	0.71	0.60	0.65	31.4
200	0.68	0.58	0.62	29.9	0.75	0.60	0.67	32.1
5,828 (all)	0.79	0.71	0.75	41.7	0.79	0.73	0.76	42.7

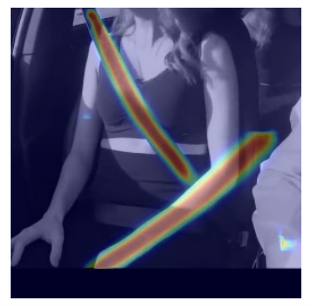
Example Predictions

- Correct predictions

- Incorrect predictions



False negative. The seatbelt is not detected despite the high contrast with the background



False positive. The subject's arm is incorrectly identified as a seatbelt

Conclusions

- Novel seatbelt segmentation training strategy using synthetic training images
- Generate arbitrarily many seatbelt training samples with added augmentations for model robustness
- Using pseudo-labeling, IoU reaches 23.0% with an F1 score of 0.55 without manual annotations
- Adding even small amounts of annotated data (200 images) improves F1 by up to 12% and IoU by 9-10%

Future Work

- Combine pose detection with seatbelt segmentation
  - Leverage the shared context between the two problems

This work was supported by NSF under grant IIP-1950249