Large Scale Asset Extraction for Urban Images

Lama Affara, Liangliang Nan, Bernard Ghanem, Peter Wonka

King Abdullah University of Science and Technology (KAUST), Saudi Arabia

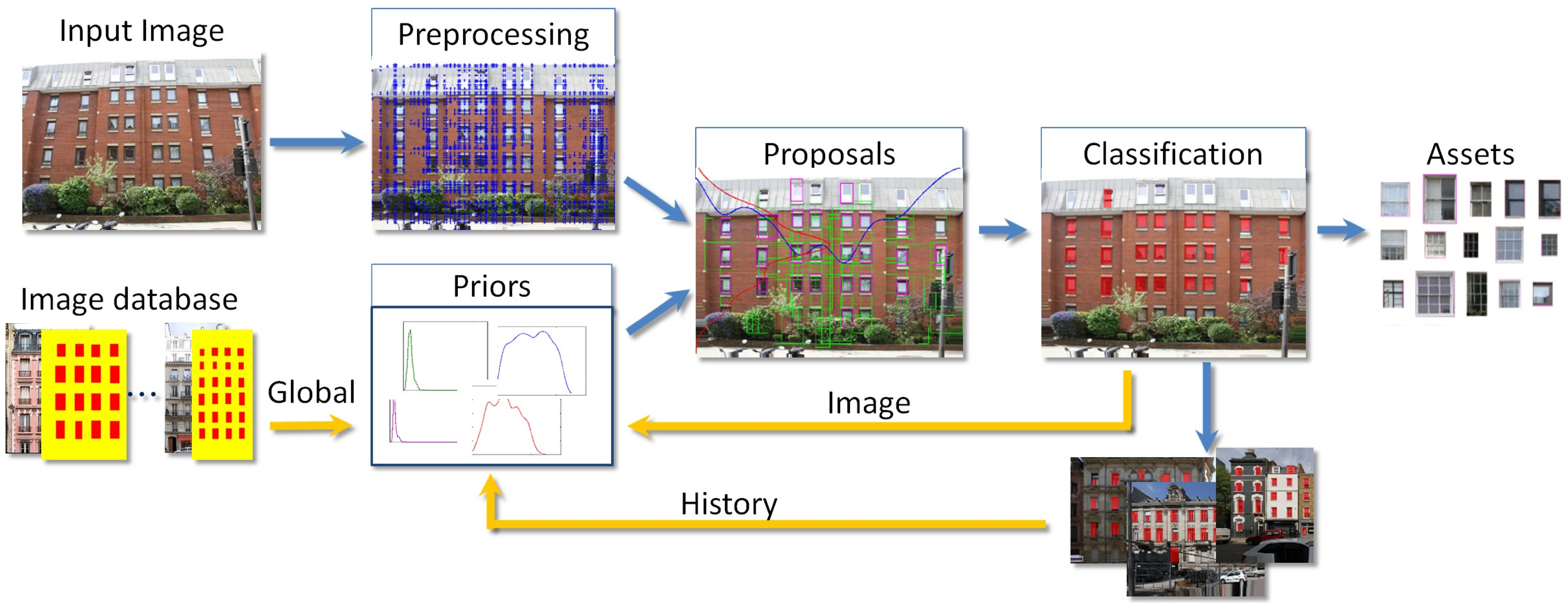
OVERALL PIPELINE

Preprocessing: Given an input image, we first detect line segments and rectify the image based on the detected most dominant facade. We then extract rectangular superpixels to restrict the search space for assets.

Prior Estimation: We employ a particle filter to estimate prior distribution functions (pdf) for the four bounding box parameters. For each parameter, three pdfs are estimated and updated during detection: global prior, history prior, and image prior.

Object Proposals: Object proposals are sampled as the evolving particle states and guided by the search space induced by estimated rectangular superpixels.

Asset Extraction: The proposed objects are classified, and the classifier's output score is used to update the pdfs, and thus guide the sampled object proposals.

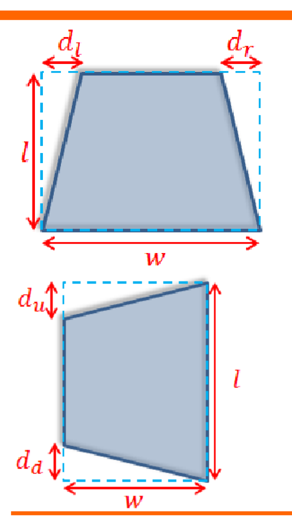


RECTIFICATION

Because assets might be captured from different viewpoints, their corresponding image regions should be rectified.

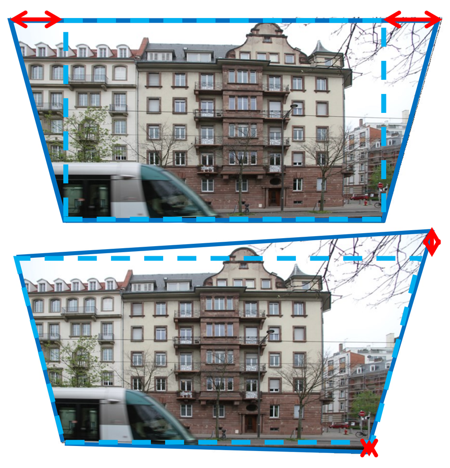
We propose a novel image rectification algorithm that is 18 times faster than previous work. We breakdown the full rectification transformation into a concatenation of two simpler transformations: vertical perspective and horizontal perspective.

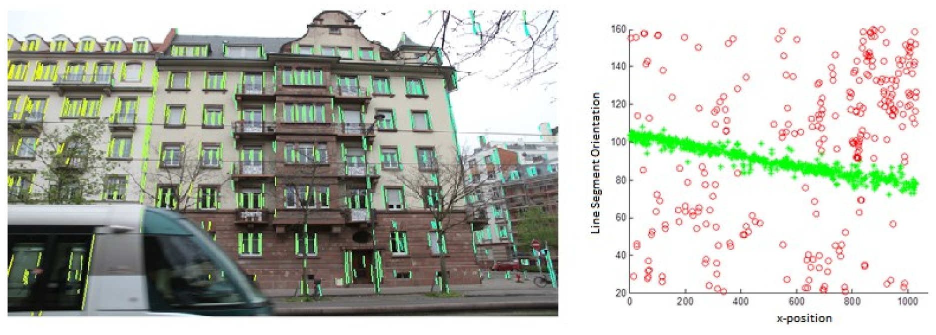
$$H = H_v H_h$$



$$H_v = \begin{bmatrix} 1+\frac{d_r-d_l}{w} & \frac{-d_l}{l} & d_l \\ 0 & 1+\frac{d_r-d_l}{w} & 0 \\ 0 & \frac{d_l-d_r}{wl} & 1 \end{bmatrix}$$

$$H_h = \begin{bmatrix} 1+\frac{d_d-d_u}{l} & 0 & 0 \\ \frac{-d_u}{w} & 1+\frac{d_d-d_u}{l} & d_u \\ \frac{d_d-d_u}{wl} & 0 & 1 \end{bmatrix}$$





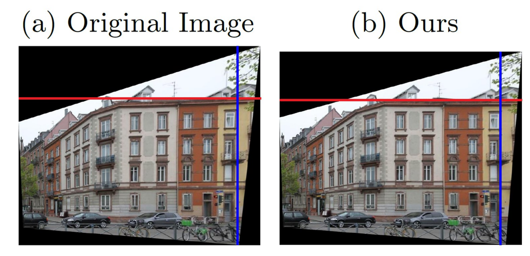
Under perspective transformation, the orientations of the originally vertical segments change linearly with their horizontal position in the image.

$$\underset{d_l, d_r}{\text{minimize}} \sum_{i=1}^{k} \left| \frac{h_1 a_i}{h_3 a_i} - \frac{h_1 b_i}{h_3 b_i} \right|$$

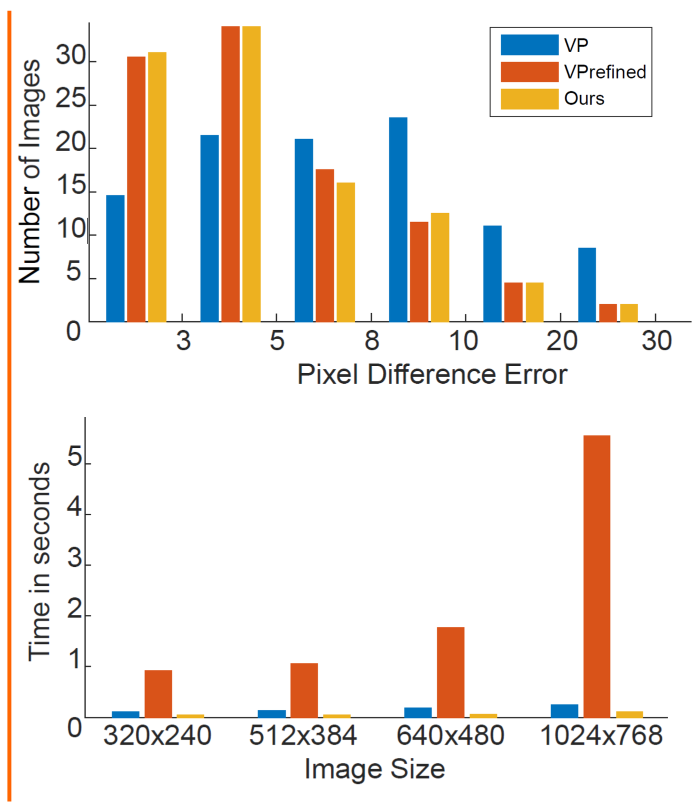
(a_i, b_i): Segment end points

h_k : Homography rows

(a) Original Image (b) Ours



(c) VP (d) VP-refined



GUIDED OBJECT PROPOSALS

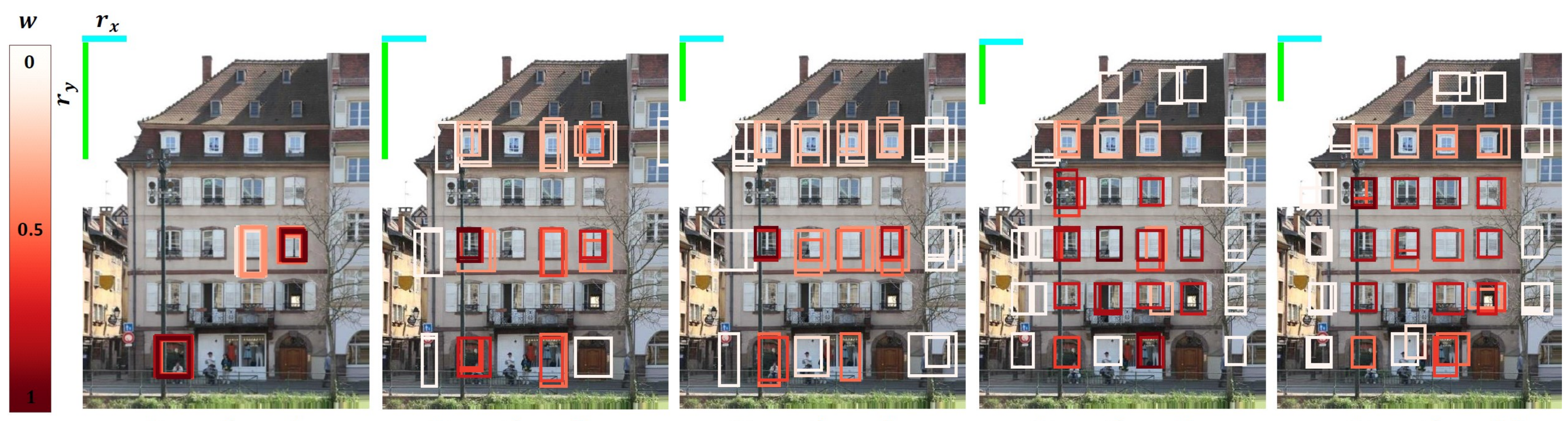
We estimate and update the prior probabilities of bounding box parameters (x-location, y-location, scale, and aspect ratio) . During our detection framework, we keep track of 12 prior distributions: 3 prior distributions for each of the 4 parameters.

Global Prior	• Estimated from training database • Gives general information about the parameters
History Prior	• Updated by the detection history within an image data set • Helps in sampling based on the learned structure
Image Prior	• Updated by the detection scores in the current image • Guides the parameters according to detected windows in the same image

We improve upon state-of-the-art object proposals by using the concept of interleaved proposing and classification.

The priors are updated using a particle filter strategy. The output of the classification at each time stamp is fed back as weights to the particles (bounding boxes), which are sampled based on the updated weights.

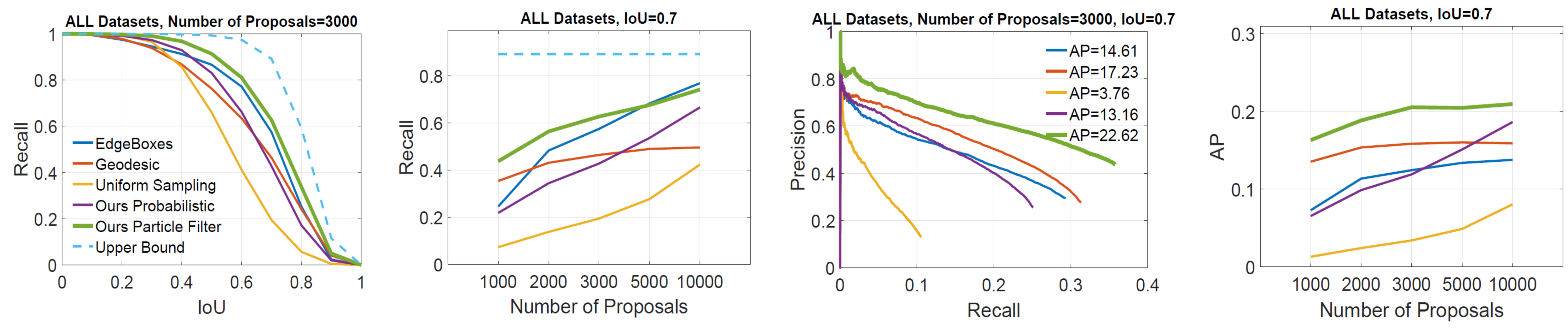
The sampled particles are output as object proposals and they are guided by the image and history updates.

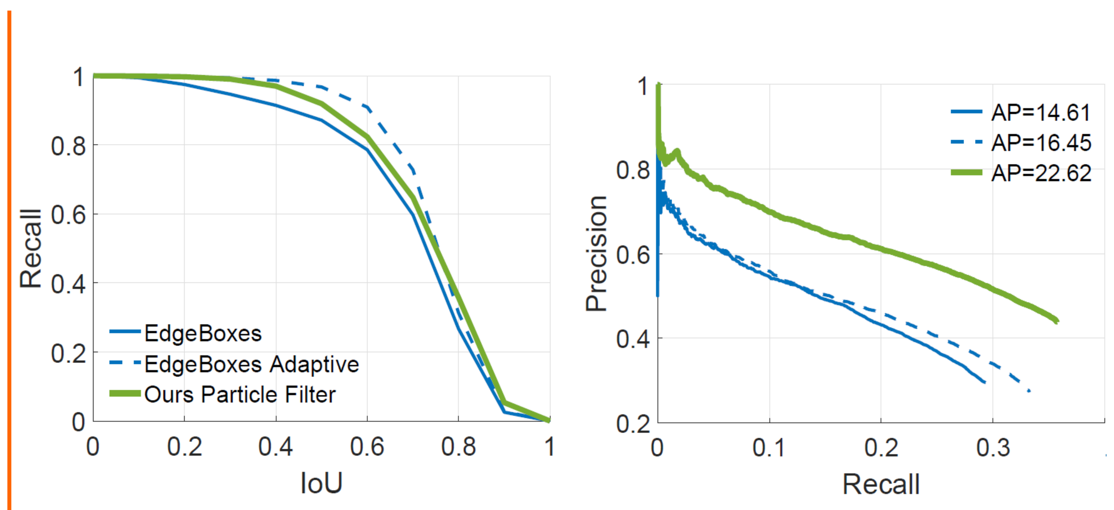


Iteration 1 Iteration 2 Iteration 4 Iteration 7 Iteration 12

The evolution of the particle filter across an image. At the first iteration, particles are sampled around three initial states. The bounding boxes show the sampled particles with the weights represented by the color map on the left. Particles with higher weights are used more often in subsequent iterations.

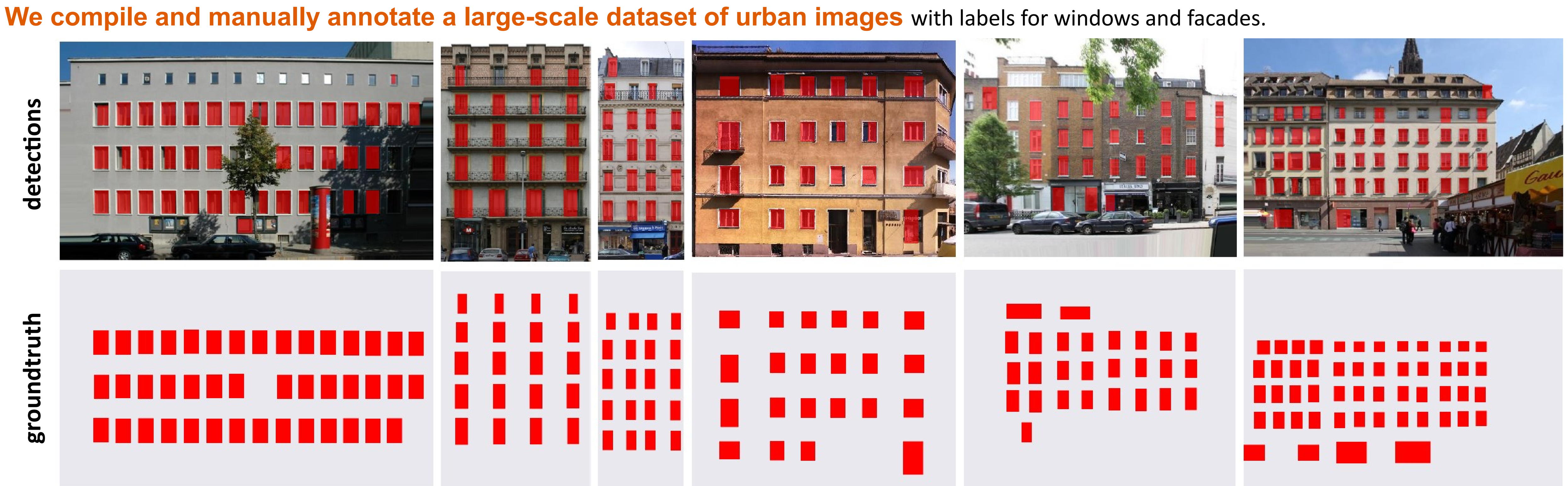
RESULTS





To further show how the priors help get better proposals, we apply the prior weights as a scoring function for the exhaustive space of proposals retrieved by EdgeBoxes. The Figure above shows how adding the priors improves the recall and thus AP of EdgeBoxes.

We compile and manually annotate a large-scale dataset of urban images with labels for windows and facades.



(a) eTrims (b) CMP (c) ECP (d) Graz (e) London (f) Strasburg

Method	Finding Proposals	Asset Classification	Total
Geodesic	0.18	4.32	4.5
EdgeBoxes	0.1	1.26	1.36
Ours Probabilistic	**0.09**	**0.24**	**0.33**
Ours Adaptive	-	-	**1.1**

Running time in seconds for 3000 proposals.