Tools for Efficient Object Detection

ICCV 2015 Tutorial

Santiago, Chile, December 2015

Organizers:

Rogerio Feris  
IBM Research  

Piotr Dollar  
Facebook AI Research  

Xiaoyu Wang  
Snapchat Research  

Kaiming He  
Microsoft Research  

Ross Girshick  
Facebook AI Research  

Rodrigo Benenson  
Max Planck Institute  

Jan Hosang  
Max Planck Institute
Classification Versus Detection

Classification: WHAT
- Dog
- Bridge

Detection: WHAT and WHERE
- Bridge
- Dog
- Dog
Efficient Object Detection

- Object detection is arguably a harder problem than image classification.

- Usually a large number of image sub-windows need to be scanned in order to localize objects, leading to heavy computational processing.

- Challenge: In many real-world applications, running a fast object detector is as critical as running an accurate object detector.
Applications

MobilEye Forward Collision Warning [Click for video demo]
Applications

Funny Nikon ad: "The Nikon S60 detects up to 12 faces."

Slide credit: Lana Lazebnik
Applications

IBM Intelligent Video Analytics [Click for video demo]
Applications

Body-worn Cameras  [Click for video demo]  (using Fast R-CNN)
Applications

Many more applications require real-time object detection...

- Robotics
- Augmented Reality
- Wildlife Monitoring
- Self-Driving Cars
- Mobile
Tutorial Overview
Goals:

- Cover tools for speeding-up object detection while maintaining high accuracy
- Focus on the state of the art
- Focus on software tools instead of hardware acceleration
- Provide pointers to publicly available source code
How to design a detector running at 100 Hz (CPU only), step by step

(Rodrigo Benenson)

- What makes strong rigid templates
- Integral Channels and Aggregated Features
- Feature Approximation Across Scales
- Cascades
- Geometric Prior

![Diagram showing the process of detecting a rider from different scales.](image)
Region Proposals
(Jan Hosang)

Towards generic object detection: candidate region generation

- Grouping proposal methods
- Window scoring proposal methods
- Metrics and in-depth analysis

Figure credit: Jan Hosang
Regionlets for Generic Object Detection
(Xiaoyu Wang)

- Regionlet representation for handling object deformations
- Classification of region proposals based on boosted detector cascades
- Integration with CNN features

Figure credit: Xiaoyu Wang
Tools for fast CNN-based Detection

Kaiming He (Inference)
Ross Girshick (Training)

“Slow” R-CNN

Fast R-CNN

Faster R-CNN

Figure credit: Kaiming He
<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Presenter</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:00</td>
<td><strong>Introduction</strong></td>
<td>Rogerio Feris</td>
</tr>
<tr>
<td>14:15</td>
<td>Detecting objects at 100 Hz with rigid templates</td>
<td>Rodrigo Benenson</td>
</tr>
<tr>
<td>15:00</td>
<td>Coffee Break</td>
<td></td>
</tr>
<tr>
<td>15:30</td>
<td><strong>Region proposals</strong></td>
<td>Jan Hosang</td>
</tr>
<tr>
<td>16:00</td>
<td><strong>Regionlet Object Detector with Hand-crafted and CNN Features</strong></td>
<td>Xiaoyu Wang</td>
</tr>
<tr>
<td>16:30</td>
<td><strong>Convolutional Feature Maps: Elements of efficient CNN-based object detection</strong></td>
<td>Kaiming He</td>
</tr>
<tr>
<td>17:15</td>
<td>Training R-CNNs of various velocities: Slow, fast, and faster</td>
<td>Ross Girshick</td>
</tr>
<tr>
<td>18:00</td>
<td>Concluding Remarks</td>
<td></td>
</tr>
</tbody>
</table>