It Takes Two to Tango: Cascading Off-the-shelf Face Detectors

Siqi Yang, Arnold Wilem and Brian C. Lovell
The University of Queensland, Australia

The False Positive Problem

- The existing face detectors, including deep learning based methods, still generate false alarms (false positives).
- Face detection is an initial step of many facial analysis tasks, including facial landmark localisation and face recognition. The unexpected false positives will affect the accuracy and speed of the subsequent tasks.
- Any method aimed to reduce false positives has the potential to improve all existing methods’ performance.

Cascading Off-the-shelf Face Detectors

- The effort to train a new face detection model is enormous.
- Solution: We propose to cascade two pre-trained face detectors in a two-stage framework:

  ![](image)

  - The computational complexity of fusion-based detectors (placed in parallel) increase linearly according to the number of detectors and the overall running time is constrained by the slowest detector.
  - Two questions arise:
    1. Which two detectors should be cascaded?
    2. Which order should they be cascaded?

Correlation and Diversity Metrics

- We define the correlation of overlapping true positives and diversity of overlapping false positives:

\[
C_{T}^{2} = \frac{|T_1|}{|T|}, \quad d_{F}^{2} = 1 - \frac{|F_1|}{|F|}
\]

- Detections from four different face detectors on the FDDB dataset [1]: (a) NPD [2], (b) HeadHunter [3], (c) MTCNN [4] and (d) HR [5]. Green: true positives, red: false positives.

![Images](image)

Proposed Cascade Properties

- A high correlation of TPs, \(C_{T}^{2} \approx 1\).
- A high diversity of FPs, \(d_{F}^{2} \approx 1\).
- Faster detector in the first stage to achieve an overall fast speed.

Conclusions

- We propose three essential cascade properties that guide us in determining the efficacy of the cascaded detector.
- Experimental results show our framework is able to remove a large number of false positives with an insignificant loss of true positive rate.
- We found a pair of face detectors that achieves significantly lower false positive rate with competitive detection rate, which is five times faster than the state-of-the-art detector described in [5].

Experiments

- Runtime analysis on the FDDB dataset [1]:

<table>
<thead>
<tr>
<th>Method</th>
<th>CPU time (SPF*)</th>
<th>TPR (FPPI*–0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VJ [7]</td>
<td>0.575</td>
<td>0.889</td>
</tr>
<tr>
<td>NPD [2]</td>
<td>0.575</td>
<td>0.889</td>
</tr>
<tr>
<td>NPD-HeadHunter</td>
<td>0.575</td>
<td>0.889</td>
</tr>
<tr>
<td>NPD-MTCNN</td>
<td>0.575</td>
<td>0.889</td>
</tr>
<tr>
<td>HeadHunter [3]</td>
<td>0.886</td>
<td>0.886</td>
</tr>
<tr>
<td>HeadHunter-NPD</td>
<td>0.886</td>
<td>0.886</td>
</tr>
<tr>
<td>HeadHunter-MTCNN</td>
<td>0.886</td>
<td>0.886</td>
</tr>
<tr>
<td>MTCNN [4]</td>
<td>0.886</td>
<td>0.886</td>
</tr>
<tr>
<td>MTCNN-NPD</td>
<td>0.886</td>
<td>0.886</td>
</tr>
<tr>
<td>MTCNN-HeadHunter</td>
<td>0.886</td>
<td>0.886</td>
</tr>
<tr>
<td>MTCNN-HR</td>
<td>0.886</td>
<td>0.886</td>
</tr>
<tr>
<td>HR [5]</td>
<td>17.687</td>
<td>17.687</td>
</tr>
<tr>
<td>HR-NPD</td>
<td>17.687</td>
<td>17.687</td>
</tr>
<tr>
<td>HR-HeadHunter</td>
<td>17.687</td>
<td>17.687</td>
</tr>
<tr>
<td>HR-MTCNN</td>
<td>17.687</td>
<td>17.687</td>
</tr>
</tbody>
</table>

- Comparison of our proposed framework and the state-of-the-art face detector:

- Comparisons on FDDB dataset [1] (left) and WIDER FACE validation set [6] (right):

![Images](image)

Acknowledgement: This work has been partly funded by Sullivan Nicolaides Pathology, Australia and the Australian Research Council (ARC) Linkage Projects Grant LP160101787. Arnold Wilem is funded by the Advance Queensland Early-Career Research Fellowship. We thank Dr Shaokang Chen for the useful discussions.